Shun Hing Institute of Advanced Engineering
信興高等工程研究所

Report and Research Highlights
2020 - 2024

August 2024

香　港　中　文　大　學
The Chinese University of Hong Kong

Printed in August 2024, this report contains information known as at July 2024.

Every effort has been made to ensure that the information contained in this report is correct at the time of printing. Shun Hing Institute of Advanced Engineering (SHIAE) – CUHK, reserves the right to make appropriate changes to the materials presented in this report without prior notice.

For further information, please visit our website: http://www.shiae.cuhk.edu.hk

Contents

Introduction of SHIAE

Mission of SHIAE

The MISSION of the Institute is to spearhead, conduct, promote and co-ordinate research in advanced engineering. There is no end to the list of areas to be explored and the plan is to give priority to research topics that are both exciting and innovative. The Institute also aspires to transferring its research results to industry for practical application and to put across to the community at large the role of engineering as a driving force for human development through educational activities.

As a pioneering institute exploring the forefront of the engineering science, The Shun Hing Institute of Advanced Engineering will

- spearhead state-of-the-art advanced engineering research
- create and sustain synergy with world-class researchers
- develop with and transfer to industries cutting edge technologies
- promote appreciation of engineering in society through educational programmes

The Shun Hing Education and Charity Fund was founded the late by Dr. William Mong Man Wai with the aim of enhancing educational opportunities for the younger generations. The Fund has already sponsored numerous educational and research programmes in Hong Kong, the Mainland, and overseas educational institutions. Himself an engineer and a firm believer in advancing the quality of life through the development of science and technology, Dr. Mong had been there to support the establishment and growth of this Institute from the beginning.

Centre of Excellence at CUHK

The Chinese University of Hong Kong (CUHK) is an internationally renowned institution of higher learning devoted to quality teaching and both academic and applied research. The University has established 29 research institutes and a number of research centres with a view to pursuing up-front research endeavours with focused goals and objectives. The Shun Hing Institute of Advanced Engineering was established in 2004 with the generous support of the late Dr. Willian Mong. It plays a crucial part in the research infrastructure of CUHK, particularly for the Faculty of Engineering, which is committed to the development of the state-of-the-art technologies in various advanced engineering areas. The Institute is now progressing to the

third decade of its milestone, and we are extremely grateful to have received continual staunch support and guidance from Mr. David Mong Tak Yeung, Chairman and CEO of the Shun Hing Group and the Shun Hing Education and Charity Fund.

As a strategic centre of excellence at CUHK, the Institute supports greater regional and international research collaborations, and strives to attract talent from the world over to achieve greater internationalization, a vision strongly advocated by every member of the University.

Commitment of the Faculty of Engineering

The Faculty of Engineering was founded in 1991 and was built upon existing strengths with added talent from all over the world. The Faculty has been able to attract some of the best minds. Many received their training in leading universities in North America, Great Britain and Australia. Most of them have extensive experience in industry and many are leaders in their fields. This team of top-notch talent is gathered to nurture local talent through educational programmes, and break new frontiers in research through innovative and exciting research endeavours.

The positioning of The Shun Hing Institute of Advanced Engineering in the William M.W. Mong Engineering Building is deliberate as a key nucleating point to integrate research endeavours in the Engineering Faculty and its neighbours. Our members join hands with their counterparts from the Faculties of Science and Medicine in many interesting research collaborations. It is the ambitious goal of the Faculty of Engineering that the Institute should become a lighthouse for the local technology landscape to herald the migration towards high value-added technology and an information economy.

The mission of the Institute is to spearhead, conduct, promote and co-ordinate research in advanced engineering. There is no end to the list of areas to be explored and the plan is to give priority to research topics that are both exciting and innovative. The Institute also aspires to transferring its research results to industry for practical application and to put across to the community at large the role of engineering as a driving force for human development through educational activities.

Building on Strength and The Way Ahead

Many of the Institute's research projects are built upon areas in which the Faculty has already achieved outstanding performance. These are areas that have great potential for further technological advancement and in line with industrial development in Hong Kong. The Institute provides a vibrant R&D environment to spur new discoveries and speed up their translation into applications. Since 2012, we have expanded our scope to cover new frontiers in Renewable Energy striving to answer tomorrow's energy challenges. In year 2017, we further expand the research scope in Multimedia Technologies to include Artificial Intelligence, Big Data Analytics and Deep Learning as well. In the past few years, the Engineering Faculty has recruited many young and talented researchers, and the Institute has given priority to provide them with the needed research support as far as possible.

Technology Transfer

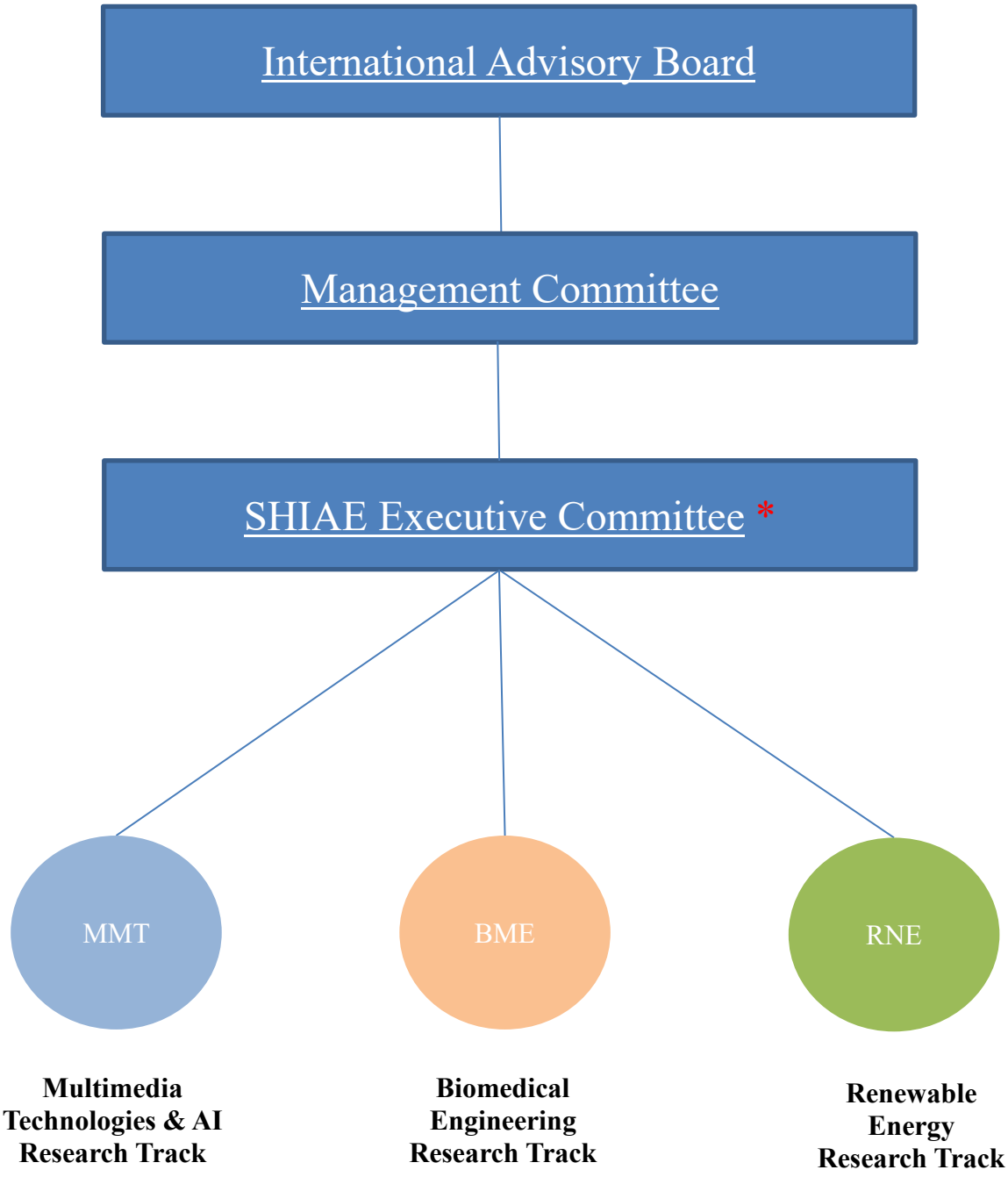
Synergy with industry is the ultimate goal of research and development in Hong Kong. External experts have been brought in to the Institute to lead research projects that could benefit the industrial sector.

The technology transfer arm of the Faculty of Engineering plays an important role in the traffic between the Institute and industry. The Institute houses an array of top-notch research and development activities encompassing contract research, spin-off companies, and consultancies.

Contribution to Society

The Institute has been making contributions to the progress of Hong Kong through a wide range of educational activities like training courses, seminars, symposiums which disseminate the latest technologies to promote appreciation of engineering in society and arouse interest of the younger generations in engineering.

Organization



We also provide support and sponsorship to the Faculty of Engineering in organizing prestigious academic conferences in Hong Kong so as to raise our international profile.

* In compliance with CUHK's guidelines in strengthening the governance of research units, an Executive Committee was formed to oversee the daily operation of the Institute in April 2020, headed by the Director, while the Dean of Engineering served as the Chairman of the Management Committee.

Composition of International Advisory Board

Chairman:

Dr. David T.Y. MONG 蒙德揚先生

Chairman & Group CEO
Shun Hing Electronic Holdings Limited
Hong Kong

Members:

Professor Victor ZUE

Delta Electronics Professor of Electrical Engineering and Computer Science Massachusetts Institute of Technology
U.S.A

Dr. Harry SHUM 沈向洋博士

Director
International Digital Economy Academy
U.S.A.

Professor Yongmin KIM

Professor
Department of Convergence IT Engineering
PoHang University of Science and Technology
Korea

Professor C.C. Jay KUO

Professor of Electrical Engineering and Computer Science
Viterbi School of Engineering University of Southern California
U.S.A.

Professor Wai-yee CHAN 陳偉儀教授

Pro-Vice-Chancellor
The Chinese University of Hong Kong
Hong Kong

Professor Wing-shing WONG 黃永成教授

Emeritus Professor
Department of Information Engineering
The Chinese University of Hong Kong
Hong Kong

Professor TSANG, Hon Ki 曾漢奇教授

Dean of Engineering
The Chinese University of Hong Kong
Hong Kong

Professor Pak Chung CHING 程伯中教授

Director of Shun Hing Institute of Advanced Engineering
Research Professor of Electronic Engineering
The Chinese University of Hong Kong
Hong Kong

Composition of Management Committee

Chairman: **Professor TSANG, Hon Ki**
Dean of Engineering

Deputy Chairman: **Professor Pak Chung CHING**
Director, Shun Hing Institute of Advanced Engineering (ex-officio)

Members:

Mr. Gary NG
Managing Director of Shun Hing Technology Co., Limited

Professor Tan LEE
Department of Electronic Engineering

Professor Wei-Hsin LIAO
Department of Mechanical and Automation Engineering

Professor Soung-chang LIEW
Department of Information Engineering

Professor Anthony Man-cho SO
Department of Systems Engineering and Engineering Management

Professor Raymond Kai-yu TONG
Department of Biomedical Engineering

Professor Benny C.Y. ZEE
Director, Office of Research and Knowledge Transfer Services

Member and Secretary: **Professor John C.S. LUI**
Department of Computer Science and Engineering

Composition of Executive Committee

(*with effect from May 1, 2024*)

Chairman: **Director, Shun Hing Institute of Advanced Engineering**
(ex officio)
Professor Pak Chung CHING

Members: **Professor Jonathan Chung Hang CHOI**
Department of Biomedical Engineering

Professor John C.S. LUI
Department of Computer Science and Engineering

Professor Li ZHANG
Department of Mechanical and Automation Engineering

Secretary: **Miss Kaia LI**
Shun Hing Institute of Advanced Engineering

Shun Hing Fellows and Research Associate

The Institute has launched a Shun Hing Distinguished Scholar Program with an aim to attract distinguished scholars to pursue research collaboration with our faculty and to strengthen our research profile.

Dr. Qian Qi 2024-2025
National University of Singapore, Singapore

Dr. Tan Yanchao 2024-2025
Zhejiang University, China

Dr. LIU Haoran 2024
Xi'an Jiaotong University, China

Dr. Lyu Wen 2024
Tsinghua University, China

Dr. Qian Junning 2023-2025
The Chinese University of Hong Kong, Hong Kong

Dr. Rui XIE 2023-2025
Tsinghua University, China

Dr. LIU Haoming 2023-2024
Peking University, China

Dr. SUN Xiangguo 2023-2024
The Chinese University of Hong Kong, Hong Kong

Dr. Zhang Rui 2023-2024
The Chinese University of Hong Kong, Hong Kong

Mr. Wang Bo 2023-2024
Beijing University of Posts and Telecommunications

Mr. WANG Yinqiao 2022-2023
Shandong University, China

Dr. Gao Shiming 2021-2023
The Chinese University of Hong Kong, Hong Kong

Dr. M. Erfan Kazemi 2020-2022
Hong Kong University of Science and Technology, Hong Kong

Financial Status of SHIAE

INCOME AND EXPENDITURE STATEMENT 2023-2024

(Fiscal Year: April 1, 2023 – March 31, 2024)

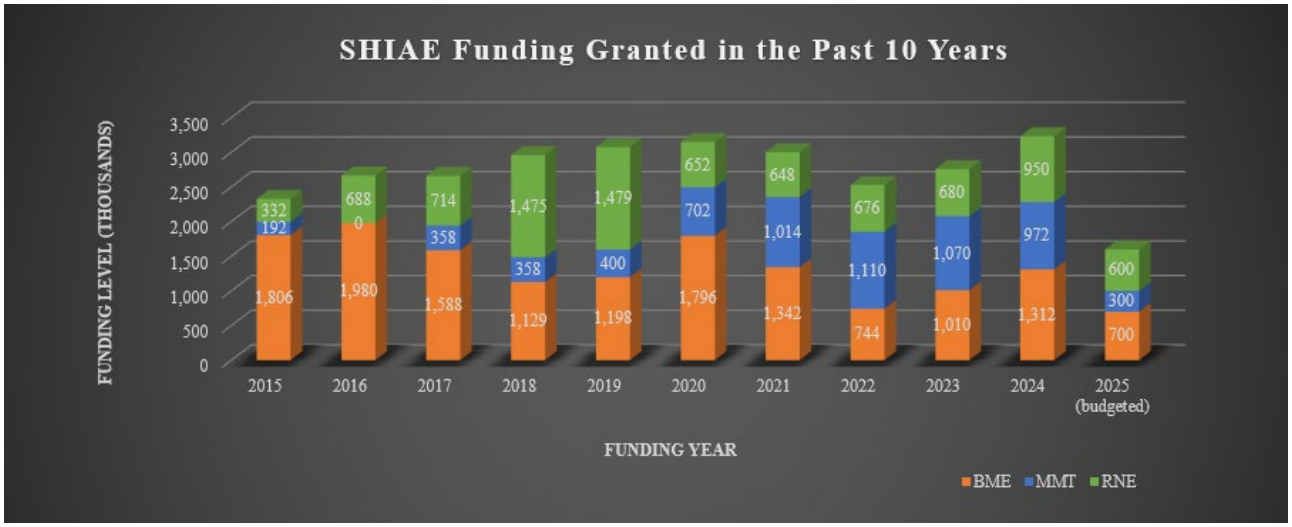
		Notes	As at 31 March 2024	As at March 31, 2023
INCOME				
Funding Source				54,500,000
SHECF new donation (1st biannual installment on April 1, 2023)			6,500,000	-
Accumulated fund brought forward			5,278,605	-
Interest and investment income			519,307	7,978,231
	Sub-total:		12,297,912	62,478,231
EXPENDITURE				
Research Funding		(1)	2,760,000	55,372,200
Remaining fund from completed projects			(238,906)	(4,736,185)
Operating cost			186,699	6,563,611
	Sub-total:		2,707,793	57,199,626
BALANCE as at 31 March 2024			**9,590,119**	**5,278,605**

APPROVED BUDGET 2024-2025

(Fiscal Year: April 1, 2024 – March 31, 2025)

		Notes	
INCOME			
Accumulated fund brought forward			9,590,119
Projected interest and investment income			200,000
	Sub-total:		9,790,119
EXPENDITURE			
Research Funding			
On-going projects (Year 2023 batch)			1,634,000
Newly funded projects (Year 2024 batch)			1,500,000
Operating cost			
Staff and Admin. Cost			70,000
Office Expenses			10,000
Distinguished lectures			10,000
Activities Sponsorship			100,000
	Sub-total:		3,324,000
Projected Balance in March 2022			**6,466,119**

Note (1) Annualized Research Funding to each research areas granted in the past ten years.



This figure shows the distribution of the SHIAE funding granted to each track of research projects, namely Biomedical Engineering (BME), Multimedia Technology & AI (MMT) and Renewable Energy (RNE) annually.

Note (2) Total funding for each batch of projects (in HK$ '000)

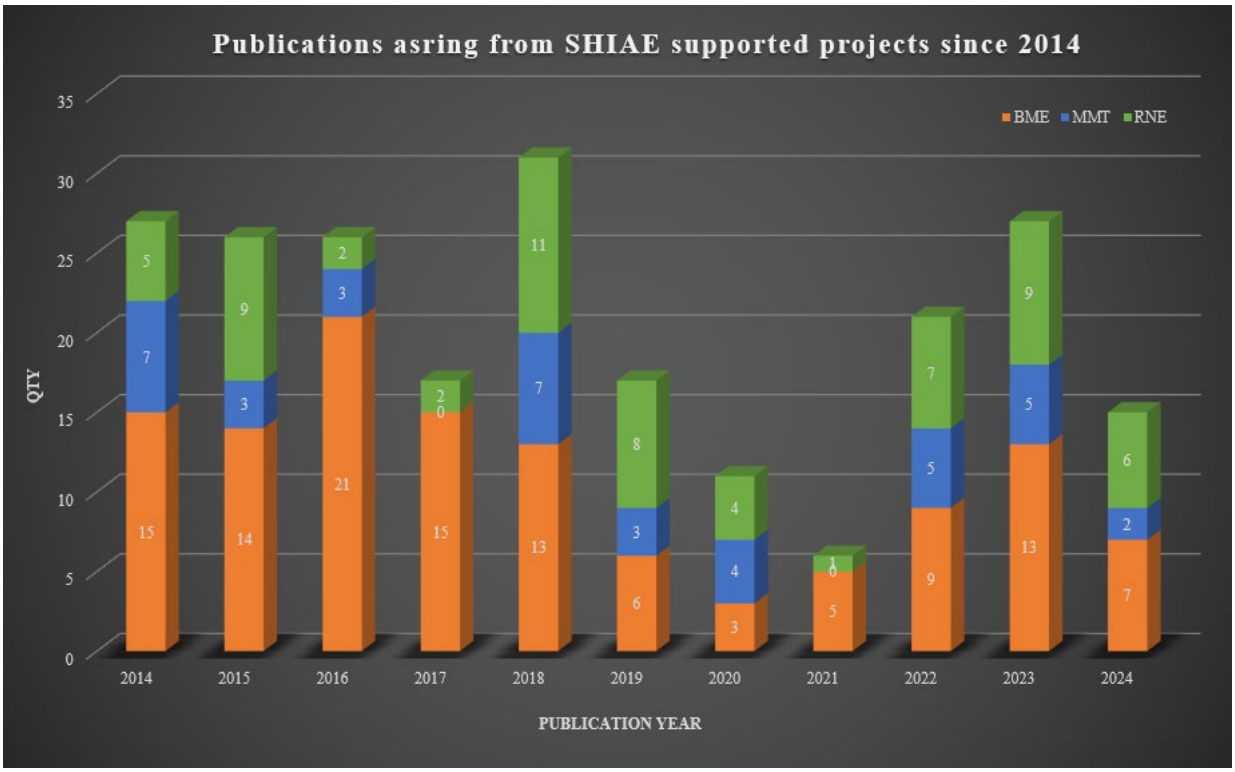
Funding Year / (No. of projects)	2025 (committed)	2024	2023	2022	2021	2005 - 2020
Year 2005 / (6)	-	-	-	-	-	6,108
Year 2006 / (5)	-	-	-	-	-	3,175
Year 2007 / (7)	-	-	-	-	-	4,146
Year 2008 / (4)	-	-	-	-	-	3,976
Year 2009 / (5)	-	-	-	-	-	3,306
Year 2010 / (5)	-	-	-	-	-	2,789.2
Year 2011 / (4)	-	-	-	-	-	2,476
Year 2012 / (5)	-	-	-	-	-	3,040
Year 2013 / (4)	-	-	-	-	-	2,948
Year 2014 / (3)	-	-	-	-	-	2,004
Year 2015 / (4)	-	-	-	-	-	2,656
Year 2016 / (4)	-	-	-	-	-	1,340
Year 2017 / (4)	-	-	-	-	-	2,660
Year 2018 / (4)	-	-	-	-	-	2,962
Year 2019 / (4)	-	-	-	-	-	3,077
Year 2020 / (5)	-	-	-	-	1,600	3,150
Year 2021 / (5)	-	-	-	1,404	1,404	
Year 2022 / (3)	-	-	1,126	1,126	-	
Year 2023 / (5)	-	1,634	1,634	-	-	
Year 2024 / (5)	1,600	1,600	-	-	-	
WOSP2007	-	-	-	-	-	25
	1,600	3,234	2,760	2,530	3,004	49,838.2

Accumulated Total: **HK$62,966.200**

Research Reports and Highlights

Academic Publications

So far **76 projects** have been successfully completed and **513 articles** arising from the results of these research projects have been published in international conference proceedings and journals. Many of their academic outcomes have received Best Paper Awards from prestigious journals and top-tier conference with international recognition. The other **8 on-going projects** are also progressing well with encouraging results produced. All publications generated by each individual project are kept in the archive of SHIAE office. The chart below shows the number of academic publications produced from 2014 onward.



Renewable Energy Track

Research Reports (2020-2024) In Renewable Energy

Newly Funded Projects (2024-2026)	* Segmented High-entropy Thermoelectric Materials for Geothermal Heat Harvesting
	* Development of carbon-14 detection system at part-per-quadrillion level based on doubly resonant photoacoustic spectroscopy
Continuing Projects (2023-2025)	* A Hierarchical Carbon-Centric Management System for Energy Storage-Assisted Data Center
(2022-2024)	* Metasurface based all-optical CNN for real-time and power - efficient machine vision

Completed Projects	
(2021-2023)	* Development of high-efficient Triply Periodic Minimal Surface (TPMS) heat exchangers for solar thermal energy applications using additive manufacturing
	* Printable aerogels for on-chip thermal management and thermoelectric power generation
(2020-2022)	* Modeling of environmental effects on performance degradation of offshore wind turbine blades made by carbon fiber reinforced plastics (CFRPs)
(2019-2021)	* Energy Management System for Large-scale Electric Vehicle Charging with Renewable Generation and Energy Storage
(2018-2020)	* Development of a Novel Cooling Tower with Free Daytime Radiative Cooling for Reducing Energy Consumption in Buildings
	* Megahertz Current Sensor for Megahertz Renewable Energy Converter
	* Development of High-Performance Triboelectric Nanogenerators for Renewable Blue Energy Harvesting

(Funded Year)

SEGMENTED HIGH-ENTROPY THERMOELECTRIC MATERIALS FOR GEOTHERMAL HEAT HARVESTING

Principal Investigator: Professor Ady SUWARDI
Department of Electronic Engineering CUHK

Project Start Date: 1 July 2024

ABSTRACT

Thermoelectric material is one of the most efficient technologies to harvest electricity from low grade (<200°C) heat sources, such as geothermal and manufacturing plants. To achieve high efficiency, high performance materials are desirable, especially between room temperature to 200°C. Specifically, with an average figure of merit (zT) of 1.5, efficiency > 10% can be realized. While the current stateof-the-art room temperature (25°C) zT is around 1.5, it drastically decreases as temperature approaches 200°C. On the other hand, while there are materials with zT > 1.5 at 200°C, they have low zT at room temperature. One intuitive approach is to combine materials with high performance at the respective temperatures into segmented legs. However, the fundamental transport properties are far from straightforward. To overcome this, we propose to study the fundamental electronic and thermal transports of segmented thermoelectrics. We intend to approach this via high-entropy alloying using Bi2Te3 and GeTe-based alloys as the candidate materials due to their intrinsically high performance at room temperature and also up to 200°C. We hypothesize that such segmentation approach will potentially enable average zT of 1.5 and beyond. The insights derived from this study will bring the community a step closer towards realizing efficient thermoelectric devices.

INNOVATION AND PRACTICAL SIGNIFICANCE:

Engineering an efficient thermoelectric device starts with innovative material. In this project, we intend to study and develop a "winning" material by combining materials segmentation approach together with high-entropy strategy. This strategy is expected to not only improve the properties of a single material, but also result in higher overall efficiency of the combined segments. In terms of practical applications, the use of thermoelectrics for low-grade (up to 200 °C) heat harvesting can help to increase energy efficiency and alleviate the high carbon emissions. In addition, the nature of the proposed project is highly relevant to the abundance of geothermal resources across greater China and the rest of the regions. Therefore, more research efforts should be invested to fully utilize them. In terms of technology adoption, the development of optimized high-entropy material compositions and segmentation conditions does not involve significant change in processing conditions. Therefore, the strategy can easily be adopted by thermoelectric material manufacturers. On a more immediate term, the findings from this project will serve as a seed for a more downstream project for prototype development, for example via innovation and technology fund.

PROJECT OBJECTIVES:

1. Synthesize and investigate the thermal and electrical properties of high-entropy thermoelectric materials, with target figure-of-merit zT of >1.5.
2. Investigate the fundamental electronic and phonon transport properties of segmented thermoelectrics to maximize the combined performance (zT).
3. Using segmentation strategy to achieve single-leg thermoelectric with device conversion efficiency of > 10.

DEVELOPMENT OF CARBON-14 DETECTION SYSTEM AT PART-PER-QUADRILLION LEVEL BASED ON DOUBLY RESONANT PHOTOACOUSTIC SPECTROSCOPY

Principal Investigator: Professor WANG Zhen
Department of Mechanical and Automation Engineering
CUHK

Co-investigator(s): Professor REN Wei [(1)]

[(1)] Dept. of Mechanical and Automation Engineering

Project Start Date: 1 July 2024

ABSTRACT

As many governments have announced the carbon neutrality goal, carbon sources apportion requires precision instruments to detect radiocarbon, which is the marker to differentiate the fossil fuel and renewable energy source emissions. However, the present commercial instruments based on accelerator mass spectrometry and cavity ring-down spectroscopy have issues of long measurement time, extremely high price and large footprint. To address this issue and achieve on-site and on-line radiocarbon detection, we propose to develop a carbon-14 detection instrument based on the novel doubly resonant photoacoustic spectroscopy. It is the one and only photoacoustic technique which has enough sensitivity for atmospheric radiocarbon detection, and inherits the features of low cost and compact sensing system. With the completion of the project, we expect to deliver a new type of radiocarbon instrument which can fulfill the in-situ distributed measurement requirement of carbon sources apportion, and provide technical support for Hong Kong to become an international city with accurate carbon traceability and trading.

INNOVATION AND PRACTICAL SIGNIFICANCE:

Carbon-14 (14C) is a very rare and important carbon isotope with a natural abundance about 1 part-per-trillion (ppt, 0.0000000001%). It is radioactive, decaying with a half-life of about 5730 years which makes it a significant radiolabel in archaeology, oceanography, climate sciences, nuclear plant monitoring and medical science [1-5].
Hong Kong's Climate Action Plan 2050 outlines the strategies and targets for combating climate change and achieving carbon neutrality [6]. Carbon sources apportion is the premise of carbon neutrality and future carbon trading market construction in Greater Bay Area. Quantifying 14CO2 has been proven to be a promising method to measure the CO2 from fossil fuels [7]. The basic principle is that the fossil fuel is totally depleted of 14C because of the much longer existence time. It is a marker to differentiate the fossil fuel emissions from other renewable power sources. A widespread use of biofuels is a key recommendation of the Paris Agreement to tackle global warming issues induced by the greenhouse-gas emissions from fossil fuels. Considering the low concentration of 14CO2 (1.2 ppt) in pure CO2, we need instruments with ultra-high sensitivity to trace and elucidate carbon cycle. To construct a regional carbon emission network, the instrument should have small footprint and affordable price for distributed monitoring.
Additionally, for the purpose of carbon reduction, 10 nuclear power plants in total will be in Guangdong province with distances about 50-400 km away from Hong Kong in the future [8]. 14CO2 artificially generated in nuclear reactors may pose a risk to living organisms if not properly confined. The precise and real-time

monitoring of 14CO2 in radioactive waste streams generated during nuclear decommissioning is important for establishing the best-suited nuclear waste management. However, the concentration of 14CO2 may change by orders of magnitude at different locations. For example, the biodegradation of radioactive waste leads to 14CO2 emissions with an activity concentration of 10 ppb to 1 ppm. In this scenario, we need instruments with a wide dynamic range and in-situ measurement.
The radiocarbon dating is a well-established technique in archaeology. In specialized corporations of Hong Kong, radiocarbon dating is one of the key scientific methods for authenticating antiques [9]. The basic principle is the amount of 14C in organisms keeps an equilibrium with living environment and subsequently decreases as a result of radioactive decay after death. The date of death can be determined by measuring the remaining amount of the isotope with extremely low concentration. This radiocarbon dating technique can determine the age of carbon-containing samples up to about 50,000 years old. However, due to the extremely high price and room size of the professional instrument, called Accelerator Mass Spectrometry (AMS), the dating work needs to be only conducted in a few professional labs outside Hong Kong.
The project aims to develop a 14CO2 sensing instrument which can provide ultra-high sensitivity and wide dynamic range with portable size and low cost. The project will provide the third type of cutting-edge technique world-wide for 14CO2 detection and make breakthroughs in resolving bottlenecks.

[1] Bronk Ramsey, Christopher. "Radiocarbon dating: revolutions in understanding." Archaeometry 50.2 (2008): 249-275.
[2] Shepard, F. P., and J. R. Curray. "Carbon-14 determination of sea level changes in stable areas." Progress in oceanography 4 (1965): 283-291.
[3] Heaton, Timothy J., et al. "Radiocarbon: A key tracer for studying Earth's dynamo, climate system, carbon cycle, and Sun." Science 374.6568 (2021): eabd7096.
[4] Povinec, P. P., et al. "Forty years of atmospheric radiocarbon monitoring around Bohunice nuclear power plant, Slovakia." Journal of Environmental Radioactivity 100.2 (2009): 125-130.
[5] Kratochwil, Nicole A., et al. "Nanotracing and cavity-ring down spectroscopy: A new ultrasensitive approach in large molecule drug disposition studies." PloS one 13.10 (2018): e0205435.
[6] https://www.info.gov.hk/gia/general/202110/08/P2021100800588.htm?fontSize=1
[7] Basu, Sourish, et al. "Estimating US fossil fuel CO2 emissions from measurements of 14C in atmospheric CO2." Proceedings of the National Academy of Sciences 117.24 (2020): 13300-13307.
[8] https://www.rfa.org/cantonese/news/nuke-09152022093318.html
[9] http://www.antiqueauthentication.com/Radiocarbon/

PROJECT OBJECTIVES:

1. The acoustic resonator designed for silicon-based cantilever is lacking in existing photoacoustic sensors, which limits the sensitivity. We will design a high Q-factor acoustic resonator to amplify the acoustic waves by standing wave effect. The resonator with the cantilever will enhance the sensitivity of photoacoustic gas sensors by at least one order of magnitude.

2. The photoacoustic signal is proportional with laser power. The sensitivity of gas sensors using mid-infrared lasers is limited by lack of mechanism which can enhance the laser power. We plan to develop a high efficiency locking technique between mid-infrared laser and high finesse optical resonator. The intracavity constructive interference will enhance laser power, thus the photoacoustic signal by 2-3 orders of magnitude.

3. The first generation of 14CO2 instrument based on photoacoustic spectroscopy will be produced by combining the acoustic resonator and optical resonator. The instrument should have parts-per-quadrillion sensitivity and compact size. To validate the performance and feasibility, we will design experiments to compare with commercial instruments.

WORK IN PROGRESS

A HIERARCHICAL CARBON-CENTRIC MANAGEMENT SYSTEM FOR ENERGY STORAGE-ASSISTED DATA CENTERS

Principal Investigator: Professor CHEN Yue
Department of Mechanical & Automation Engineering
CUHK

Research Team Members:
Rui XIE, SHIAE Research Associate (1)
Dongxiang YAN, Postdoctoral Fellow (1)
Wenyi ZHANG, Postdoctoral Fellow (1)
Shihan HUANG, PhD student (1)
Tao TAN, PhD student (1)

(1) Dept. of Mechanical and Automation Engineering

Reporting Period: 01 July 2023 – 1 May 2024
(to be completed in June 2025)

INNOVATION AND PRACTICAL SIGNIFICANCE:

With the explosion of data-driven workloads, the energy demand and carbon emissions of cloud/edge data centers grow sharply, threatening environmental sustainability. In the current practice, data centers balance their carbon emissions with carbon offsets such as signing power purchase agreements with large-scale renewable energy projects. This practice, however, has some potential problems: 1) Large-scale renewable generation sites are often located far from data centers. Huge transmission losses occur when their generated electricity is used. These losses need to be compensated by other carbon-intensive energy sources. 2) Even if data centers' net carbon emissions are offset to zero, they still emit massive amounts of carbon into the environment. In fact, reducing absolute carbon emissions is the only way to eventually achieve sustainability. This requires emphasizing carbon efficiency in data center operations. This project aims to develop a hierarchical carbon-centric data center management system. By adopting a bottom-up approach, we provide carbon-efficient solutions to individual, regional, and overall data center systems, respectively. If successful, it can benefit: 1) end-users, by allowing them to gain revenue through participation in carbon emission reduction; 2) data centers, by lowering their carbon emissions and operational costs while promoting the use of their co-located renewable generations; 3) the power grid, by alleviating the impact of volatile renewable generations on system reliability. The outcome of this project will be turned into a set of software packages. The PI will then reach out to our industry partners to address critical issues in engineering practices, such as compatibility with the existing management system. This may bring about follow-up funding and opportunities for implementation and commercialization.

ABSTRACT

The energy demand of data centers increases dramatically with the explosion of data-driven workloads. This positions data centers among the main contributors to global carbon emissions. In fact, the increasing carbon footprint is a more serious problem than merely the rising energy demand. Noticing that energy efficiency and carbon efficiency are not necessarily correlated, this project aims to develop a carbon-centric data center management system. Unlike previous research that focused on energy costs, it elevates carbon to a first-priority metric in the design. The main difficulties come from the diversity of computing tasks, the volatility of carbon-free energy sources, and the invisibility of the carbon footprint of grid power. This project adopts a bottom-up approach with the following tasks: (1) Enhancing carbon efficiency within an edge data center by co-optimizing computing task assignment and energy storage utilization considering diverse user requirements;

(2) Enhancing carbon efficiency regionally across edge data centers considering the tradeoff between reduced emission and lower latency; and (3) Enhancing carbon efficiency of the overall system by coordinating the operations of cloud and edge data centers through well-designed prices. A simulation platform will be built for demonstration and validation. This project will contribute to carbon emission reduction and sustainability.

1. OBJECTIVES AND SIGNIFICANCE

(1) For each edge data center, develop a computing task and energy storage co-optimization approach to enhance the use of co-located renewable generation.
(2) For regional edge data center clusters, develop a computing task and energy storage co-sharing approach to take advantage of the complementarity between their renewable generations.
(3) For the overall data center system, develop a carbon-integrated electricity price-based coordination mechanism to reduce the use of carbon-intensive grid power.
(4) Construct a simulation platform for demonstration and validation.

2. RESEARCH METHODOLOGY

First, we will focus on improving the carbon efficiency of an edge data center. Edge data centers typically have low power density, allowing them to be self-powered by their co-located renewable energy sources without relying on grid power. Hence, it is possible to have a zero-carbon footprint as long as the volatility of renewable energy sources is well-tackled. We will create appropriate models to characterize the diverse features of computing tasks from end-users. With these models, an online computing task assignment approach will be developed to match the energy consumption of the edge data center with the co-located renewable generation. Furthermore, we will co-optimize energy storage utilization with computing task assignments to promote the use of clean renewable energy.

Second, we will focus on improving the carbon efficiency of edge data center clusters based on the fact that the aggregate renewable power supply across a region is smoother and more predictable. Thus, shifting computing tasks across data centers would help when an individual edge data center struggles with its co-located volatile renewable generation. However, shifting computing tasks across a larger region will trade off low latency of computation for decreased aggregate renewable power volatility. We will characterize the latency and the associated disutility for end-users when their computing tasks are shifted. A computing task sharing approach considering the tradeoff above will be developed. Furthermore, as some computing tasks are non-shiftable, edge data centers may still need carbon-intensive grid power. A computing task and energy storage co-sharing approach will be proposed to further reduce reliance on grid power by lowering the energy supply-demand mismatch.

Third, we will focus on improving the carbon efficiency of the overall data center system via proper carbon-integrated electricity prices. The response of edge data centers to carbon-integrated electricity prices will be modeled through sensitivity analysis. Based on this, a price-based coordination mechanism will be developed to regulate the service demand of edge data centers. Furthermore, we will design incentives to encourage edge data center clusters to support low-carbon operation of cloud data centers via their individual/shared energy storage. With the proposed technologies, a simulation platform will be built for demonstration and validation.

3. RESULTS ACHIEVED SO FAR

We have developed a distributed online algorithm for combined computing workload and energy coordination of data centers. The proposed algorithm is prediction-free and easy to implement. It well addresses the two major challenges of data center operation: (1) The high uncertainty due to the unpredictable computing workload and renewable generation, and (2) the need for a distributed implementation framework to avoid the high computational burden and privacy leakage. As shown in Fig. 1, the proposed algorithm can achieve the lowest accumulated operation cost compared to other existing online algorithms. Moreover, the proposed accelerated alternating direction method of multipliers (ADMM) algorithm can achieve an operation cost close to the centralized benchmark in a distributed manner, as illustrated in Fig. 2. The small gap is the efficiency

loss due to the truncation for acceleration purposes. The proposed algorithm can improve the renewable power utilization of data centers, thereby improving their carbon efficiency.

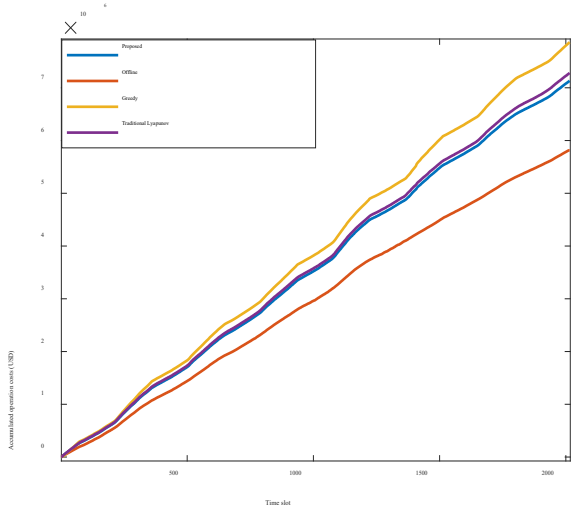


Fig. 1 Accumulated operation costs under the offline benchmark and three online algorithms

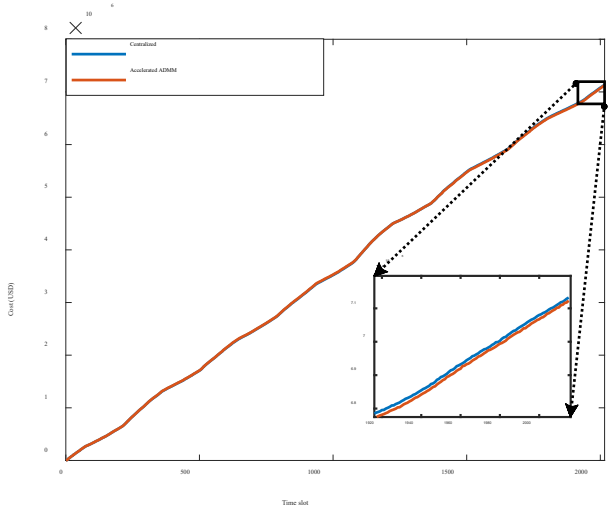


Fig. 2 Cost traces of the proposed algorithm and its centralized counterpart

We have developed an Aumann-Shapley price-based method to allocate carbon responsibility within a power network. Electricity consumers should be responsible for at least part of the carbon responsibility since they are the cause of electricity generation. In order to encourage the adoption of low-carbon practices in data centers, it is crucial to properly allocate and price the carbon responsibility of data centers at various buses within the power network. The proposed Aumann-Shapley price-based allocation method possesses several desirable properties, such as cost sharing, scale invariance, monotonicity, additivity, and consistency. These properties contribute to a fair and effective allocation of carbon responsibility. Based on this, we have developed a carbon-integrated electricity pricing method. The prices for a modified IEEE 30-bus system are shown in Fig. 3, which effectively captures the contributions of electricity demand from various locations to the total system emissions. We can observe from Table 1 that with the proposed pricing method and energy storage, we can achieve the lowest total emission and renewable power curtailment.

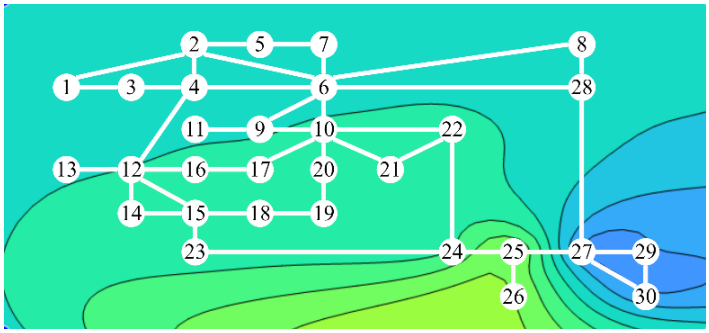


Fig. 3 Carbon-integrated electricity prices for a modified IEEE 30-bus system

Table 1 Results with/without energy storage and carbon responsibility allocation

Case	Proposed	A1	A2	A3
Energy storage	√	√	×	×
Carbon responsibility allocation	√	×	√	×
Total generation cost ($/h)	3387	3121	3443	3173
Total emission (kgCO_2/h)	30546	53701	31063	54457
Renewable curtailment rate	1.84%	1.84%	3.25%	3.25%

We have developed a prediction improvement approach to forecast renewable power and electricity demand more accurately by aggregating predictions from the system operator and distributed agents. Fig. 4 shows that with the improved predictor (Best Linear Predictor), the variation ranges of prediction errors are greatly narrowed. In the figure, the green area represents the variation range of the original prediction error, while the blue area represents the variation range of the improved prediction error. Furthermore, a robust optimization method has been proposed to leverage the improved prediction for better decision-making. The proposed method can lower operation costs and enhance carbon efficiency by reducing the uncertainty associated with renewable generation.

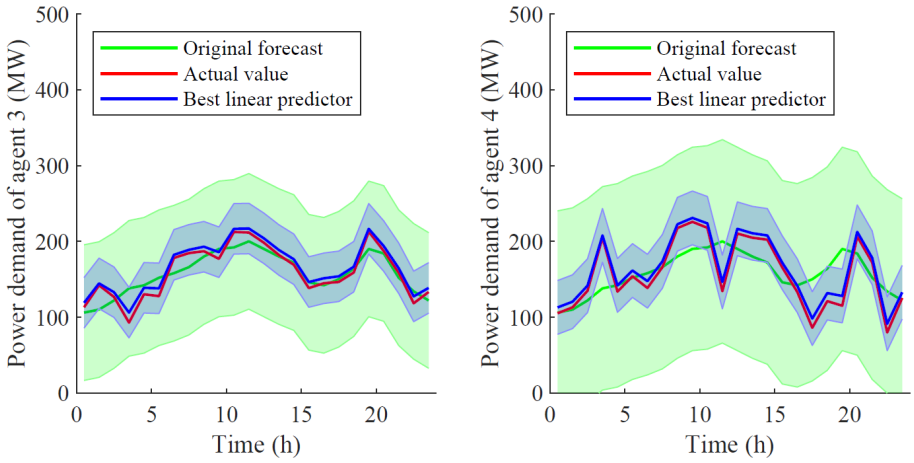


Fig. 4 Original and improved uncertainty sets of a wind farm (left) and an electricity demand (right)

4. PUBLICATION AND AWARDS

4.1. Publications

J[1] S. Huang, D. Yan, and Y. Chen, “An Online Algorithm for Combined Computing Workload and Energy Coordination within a Regional Data Center Cluster,” International Journal of Electrical Power & Energy Systems, vol. 158, pp. 109971, 2024.
J[2] R. Xie, and Y. Chen, “Real-time Bidding Strategy of Energy Storage in an Energy Market with Carbon Emission Allocation Based on Aumann-Shapley Prices,” IEEE Transactions on Energy Markets, Policy and Regulation, 2024, early access.
J[3] R. Xie, P. Pinson, Y. Xu, and Y. Chen, “Robust Generation Dispatch with Purchase of Renewable Power and Load Predictions,” IEEE Transactions on Sustainable Energy, 2024, early access.
J[4] D. Yan, S. Huang, and Y. Chen, “Real-time Feedback Based Online Aggregate EV Power Flexibility Characterization,” IEEE Transactions on Sustainable Energy, vol. 15, no. 1, pp. 658 – 673, 2024.
J[5] Y. Zhang, Y. Su, Y. Chen, and F. Liu, “Asynchronous Distributed Charging Protocol for Plug-in Electric Vehicles,” Journal of Economy and Technology, vol. 1, pp. 29-47, 2023.
C[1] R. Xie, and Y. Chen, “Privacy-Preserving Aggregated Load Forecasting Based on Vertical Federated Learning”, Nexus Forum, Hong Kong, China, pp. 1-6, May 9-10, 2024.
C[2] M. Yang, and Y. Chen, “Robust Operation of Distribution Systems with Uncertain Renewable Generation via Energy Sharing”, The 3rd Conference on Fully Actuated System Theory and Applications, Shenzhen, China, pp. 1-6, May 10-12, 2024.

4.2. Awards

[1] 2023 Best Paper Award of Journal of Economy and Technology

WORK IN PROGRESS

METASURFACE BASED ALL-OPTICAL CNN FOR REAL-TIME AND POWER -EFFICIENT MACHINE VISION

Principal Investigator: Professor HUANG Chaoran
Department of Electronic Engineering
CUHK

Research Team Members:
Dr. Lyu Wen, Postdoctoral Fellow [(1)],
Bo Wang, Research Assistant [(1)],
LUO Zeyu, Junior Research Assistant[(1)]

[(1)] Dept. of Electronic Engineering

Reporting Period: 01 July 2022– 31 May 2023
(to be completed in August 2024)

INNOVATION AND PRACTICAL SIGNIFICANCE:

Device-level innovation*: This is the first time that the photonic metasurface is proposed for CNNs. Photonic metasurface was originally used for more conventional applications such as optical imaging [10]. Recently, there has been a surge of research on using photonic metasurface for computing [11]. However, those studies have only demonstrated simple arithmetic operations (e.g., differentiators [12], [13]). The photonic metasurface, according to our preliminary results, can perform convolution operations to the optical images encoded in visible light, and extract multiple image features from the entire image at one time. The proposed all-optical CNN is completely passive, thus can bring unprecedented low latency and power efficiency as compared to electronic accelerators.*

System-level innovation: *The optical convolutional accelerators that have been demonstrated so far need to manipulate how the image pixels enter the photonic devices, which as a result, need to preprocess the image at the digital domain first and then convert the processed signal back to the optical domain[5], [9]. The proposed all-optical CNN, in contrast, can process the whole image encoded on light directly at the photonic domain, without power-hungry domain crossings (e.g., optical-to-electrical-optical and analog-to-digital-to-analog conversions). As such, our all-optical CNN can significantly reduce the system overhead in terms of power consumption, latency, device footprint, and cost.*

Application-level innovation*: The proposed all-optical CNN can be used as a general-purpose AI accelerator, since it can compute the convolutional layer in only sub-picosecond. This means that the proposed all-optical CNN can achieve extremely large computing throughput as many AI accelerators target. However, we want to set a more ambitious goal from the application level – we want to fully exploit the unprecedented low latency and power efficiency that our all-optical CNN uniquely has, in order to benefit those extremely latency-, power-sensitive, and computation-intensive applications [3].*

In the following section, we will discuss some of the targeted applications and how our all-optical CNN can bring innovative and practical significance to them.

ABSTRACT

Convolutional neural networks (CNNs) have become the dominant artificial intelligence (AI) algorithm in various computer vision tasks. However, the trend towards ever-larger CNN models has posed significant stress on hardware resources. Most AI accelerator hardware, such as GPUs, focus on having high computing throughput but at the expense of huge power consumption. Therefore, the same hardware cannot be used in

edge applications where computing power is scarce. Meanwhile, many edge applications, for example, autonomous vehicles, require hardware to make decisions within a very short response time (i.e., low latency). Unfortunately, the state-of-the-art digital hardware cannot offer the matched response time. This project proposes an all-optical CNN accelerator to overcome the two major bottlenecks in digital electronic accelerators–power consumption and latency–from a new perspective. The proposed all-optical CNN is based on a passive metasurface consisting of a large array of collectively designed subwavelength “pillars”. By manipulating the light encoded with image information, the metasurface can perform convolution operations per second and generate various features of the entire image 'on-the-fly' with sub-picosecond delay and nearly zero power consumption. The proposed all-optical CNN will bring critical performance improvement to the edge devices for machine vision which is desperate for real-time and low-power information processing.

1. OBJECTIVES AND SIGNIFICANCE

1.1. Objectives

(1) To design optimized all-optical cnn hardware that can provide ultra-low power consumption and latency.

(2) To experimentally validate the design of all-optical cnn in performing machine vision applications and benchmark its performances against the state-of-the-art edge computing devices

(3) To explore and demonstrate power- and latency-sensitive machine vision applications that all-optical cnn hardware can bring critical performance improvement.

1.2. Significance of this project

The CNN accelerator to be developed in the project will overcome the two major bottlenecks in digital electronic accelerators in power consumption and latency. It can perform convolutional operations to the optical image 'on-the-fly' with sub-picosecond delay and nearly zero power consumption, thus promising to bring critical performance improvement to the edge devices for machine vision which is desperate for real-time and low-power information processing.

2. RESEARCH METHODOLOGY

Theoretical modeling, device design, and fabrication: We carry out the theoretical modeling and the design of the metasurface, in which sub-wavelength Si cubes are arranged with different in-plane rotation angles on the Si2O3 substrate. Based on the 3D finite-difference time-domain (FDTD) method, the transmission and phase accumulation of a single Si cube with various sizes and the focusing effect of the element structure of the metasurface are investigated. The Si cube arrays will be collectively designed using the particle swarm optimization method. The devices are fabricated in the commercial silicon photonic foundry.

Experimental characterizations: We experimentally characterize the fabricated metasurface in performing large-scale CNN and directly in the photonic domain. We use spatial light modulators (SLMs) to generate input optical images. The optical images will be first filtered by a wavelength-selective filter with a central wavelength that is matched with the operation wavelength of our metasurface. The filtered optical images are then illuminated onto the metasurface. The output of the metasurface is the extracted features (which are still optical images) of the input optical images. Finally. a high-speed camera will be installed at the metasurface output to detect the extracted information, which are further processed by the subsequential digital processor. We will perform a comprehensive link-to-link analysis to the proposed all-optical metasurface processor and benchmark against state-of-the-art edge devices (for example, edge TPU). We will evaluate the classification accuracy, total power consumption, processing latency, throughput of the entire machine vision link incorporating the all-optical metasurface processor and the digital processor.

3. RESULTS ACHIEVED SO FAR

We have successfully completed the first two objectives of this project, including chip design, fabrication, and experimentally validate the design of all-optical metasurface in performing machine vision applications. We developed two different schemes.

In the first scheme, we develop an optical metasurface that acts as a convolutional layer in a CNN. The optical metasurface can implement convolution operations to optical images, and produce multiple feature maps in parallel. Moreover, the optical metasurface can be compactly assembled in front of the commercial image sensor with only a 10 μm spacing and without any additional components in between. When the optical image incidents on the metasurface, its complex field is modified by the metasurface, leading to an elementwise multiplication between every pixel of the optical image and kernels. After that, the image sensor detects the processed optical image, and the subsequent digital processor sums the detected light in each sub-block to produce the corresponding feature maps.

We proposed a new metasurface design, which overcomes crosstalk between adjacent sub-blocks of the metasurface caused by light diffraction. This restricts the ultimate accuracy of the designed kernel. In our design, we draw inspiration from the design of metalenses, which can converge light to a single point at the focus plane. However, typical metalenses cannot modify light transmission to achieve a continuous range of weight. Therefore, to achieve a continuous range of weight, we further introduce Mie resonance into the design, which can alter the light transmission by trapping the light in its resonance. By combining metalens and Mie resonance, our design reduces the crosstalk-induced error from 7.3x10-2 to 1.8x10-5 and realizes the computing accuracy of 6.5 bit. For comparison, the computing accuracy of IBM TrueNorth, an analog neuromorphic computer, is 5 bit, and that of Google TPU, a digital AI accelerator, is 8 bit.

We performed an experimental evaluation of the performance of our metasurface as the convolutional layer in a CNN for MNIST digital recognition. The CNN is comprised of a convolutional layer for feature extraction, an averaging pooling layer for reducing data size, and a fully connected layer for producing recognition results. The CNN is first trained on a digital computer. Using the trained weights, we design and fabricate a metasurface to realize the convolutional layer that has 4 kernels with a size of 3 × 3. The following layers are realized in a digital computer. A photomask with the same pixel size as the metasurface block is used to produce an optical image on the metasurface. We use 133 photomasks with various MNIST digits to test the accuracy of the CNN experimentally. The confusion matrix is plotted in Figure 1 , showing an accuracy of 91.7 %,
The results are presented at the Conference on Lasers and Electro-Optics 2023, San Jose, United States, the flagship optoelectronics conference. The presentation is selected as the highlighted talk (< 3% of the accepted papers). A journal article is under preparation.

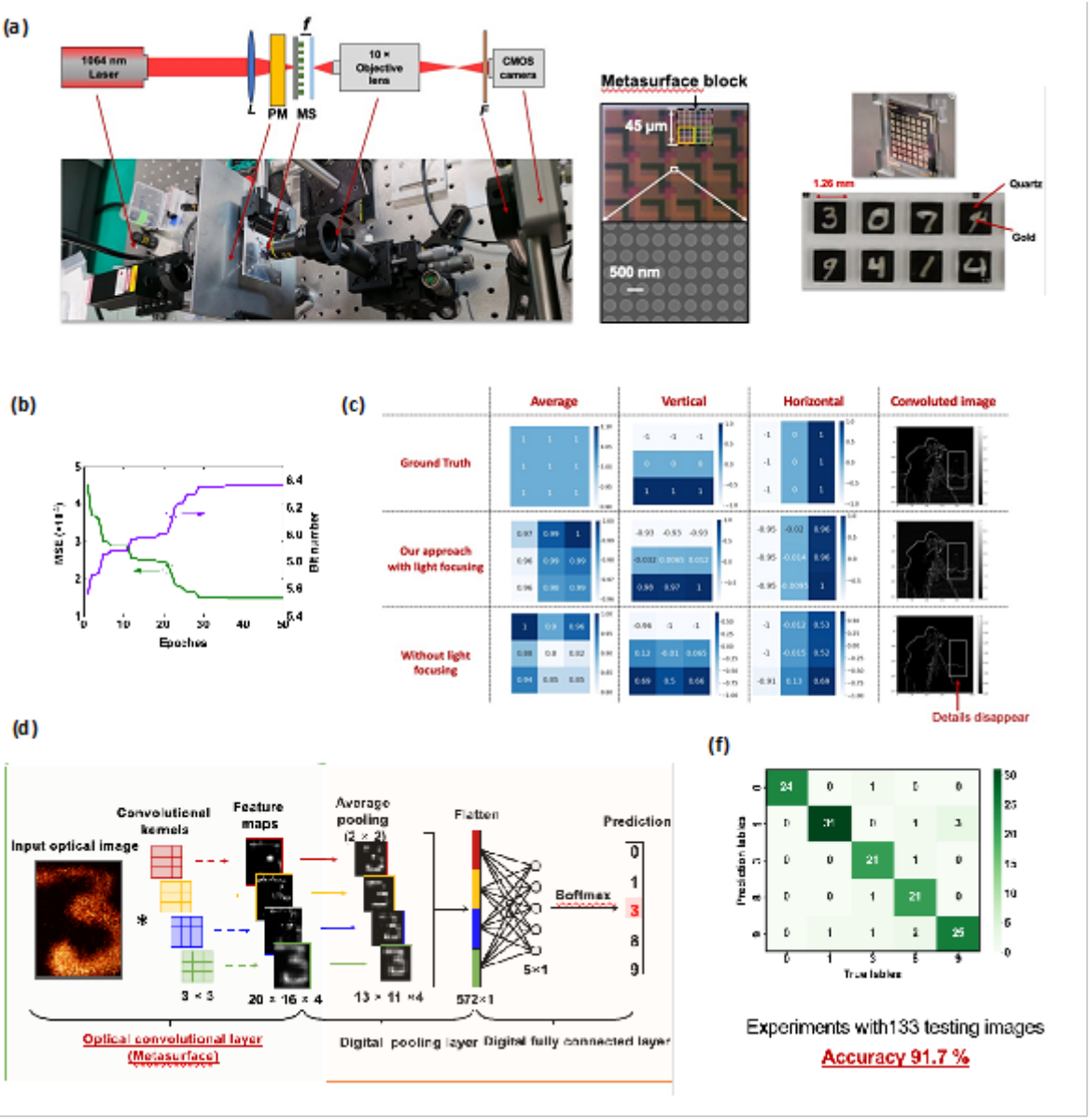


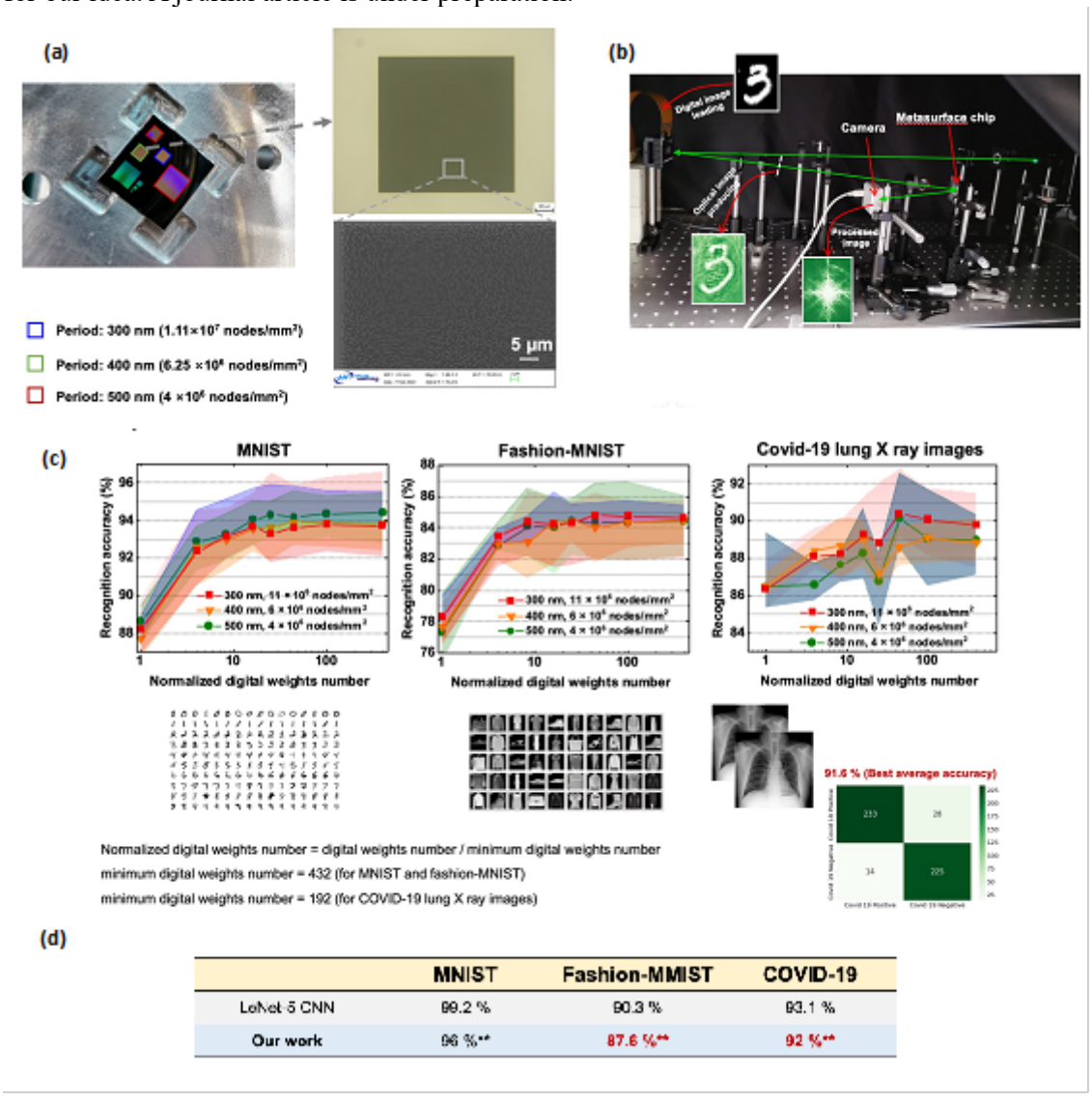
Figure 1. (a) Experimental setup (left), optical image of metasurface (middle), and optical photomask for MNIST dataset (right); (b) Optimization of metasurface design using particle swarm method; (c) Accuracy comparison of our approach with conventional design. (d) metasurface-based CNN architecture (e) experimentally tested confusion matrix for MNIST digits classification.

In the second scheme, we proposed an all-optical reservoir computing using metasurface. Reservoir computing is a type of recurrent neural network that involves training only the readout layer, while leaving the internal "reservoir" of the network fixed. The reservoir consists of a randomly connected network of nodes, and its dynamics are driven by the input data. The readout layer, which is a linear or nonlinear classifier, is trained to map the reservoir's state to the desired output. The advantage of reservoir computing is that the reservoir's fixed topology and random connections allow it to efficiently represent complex input-output mappings without the need for fine-tuning of the internal weights.

A metasurface is a surface composed of millions of subwavelength structures. These structures can manipulate the properties of light, such as its transmission, phase, polarization, and amplitude, with high spatial resolution. This allows for precise control over the behavior of light waves, enabling functionalities that are not readily available with natural materials. The millions of subwavelength structures form a large-scale network with millions of nodes, projecting the incident optical image to high-dimensional reservoir states. The reservoir's complex and high-dimensional transient response to the input signal is capable of universal computation.

In our design, the radii of the pillars are randomly sampled from a Gaussian distribution. The incident optical images, upon passing through the metasurface, are captured by the CMOS sensor array. The optical detection process measures the square of the input light's amplitude, which acts as the nonlinearity activation function within the reservoir. Finally, the readout layer is performed digitally using logical regression. We conducted experimental tests on our chip using three distinct machine vision benchmark tasks: the MNIST dataset, Fashion-MNIST dataset, and Covid-19 X-ray lung images. To evaluate its performance, we compared our chip with the LeNet-5 convolutional neural network (CNN) model, which is a widely recognized benchmark for these tasks. The passive metasurface device has comparable accuracy in all these tasks compared to LeNet-5 (Figure 2).

We are exploring the real-world biomedical image applications of our chip and system. After that, we will file a patent for our idea. A journal article is under preparation.



	MNIST	Fashion-MNIST	COVID-19
LeNet-5 CNN	99.2 %	90.3 %	93.1 %
Our work	96 %**	87.6 %**	92 %**

Figure 2. (1) metasurface chip, (2) Experimental setup; (c) accuracy vs node number; (d) performance comparison of our approach with LeNet-5 CNN.

4. PUBLICATION AND AWARDS

M. Luo, S. Xiao, T, Xu, H. Tsang, C. Shu, and C. Huang, Ultra-compact optical convolutional accelerators based on polarization-independent metasurfaces. In Conference of Laser and Electro-optics, SM3J-1, Optica Publishing Group, 2023. (Highlighted presentation, < 3%).

COMPLETED

DEVELOPMENT OF HIGH-EFFICIENT TRIPLY PERIODIC MINIMAL SURFACE (TPMS) HEAT EXCHANGERS FOR SOLAR THERMAL ENERGY APPLICATIONS USING ADDITIVE MANUFACTURING

Principal Investigator: Professor Xu SONG
Department of Mechanical & Automation Engineering, CUHK

Research Team Members:
Shiming Gao, Postdoc – Shun Hing Fellow [(1)];

[(1)] Dept. of Mechanical and Automation Engineering, CUHK

Project Start Date: 1st July 2021
Completion Date: 30th June 2023

INNOVATION AND PRACTICAL SIGNIFICANCE:

The key innovations come from both the design and the manufacturing: the design of conformal TPMS-based HEs with external freeform geometry and μSLM processing of fully-dense pure copper. The conformality of smooth and continuous TPMS with large surface-area-to-volume ratio enables a high heat transfer efficiency with acceptable pressure drop. The high volume energy density μSLM processing of pure copper, which hasn't been reported before, will significantly improve the density of copper parts with smooth surface finishing and increase the heat transfer coefficient of HEs further, thanks to the high material thermal conductivity. With continuous process improvement of μSLM, printing high-performance copper components with high flexibility and complexity in geometric designs becomes possible.

Due to the better performance of HEs, higher solar energy conversion rate and less traditional energy consumption can be achieved, which provides more incentives to encourage people to install the solar thermal system and reduces the consumption of fossil fuel. Beyond aforementioned application, this proposed design method also benefits other type of HE designs used in a variety of thermal control systems, such as refrigeration cycles, automotive industry and electronic equipment, etc. Hence, there is good potential and wide economic and environmental impact derived from this project.

ABSTRACT

Heat exchangers (HEs) are a key component in the solar thermal system, which is claimed to be one of the most efficient systems to harvest solar energy as the renewable energy source. However, traditional heat exchangers have always suffered from the problems of leakage, low heat exchange effectiveness, and bulky volume. The emerging additive manufacturing (AM) technology offers the product design more freedom, which enables the fabrication of high-efficiency heat exchangers to become possible. A series of heat exchangers with different materials including stainless steel, Inconel 718, Ti-6AL-4V, and AlSi10Mg fabricated, but their designs are still based on improvements to some traditional heat exchanger structures and the materials used are still low thermal conductivity materials. The Triply Periodic Minimal Surface (TPMS) has properties of a large surface-area-to-volume ratio and smooth staggered fluid channels, but the conformal filling of TPMS structures in HEs is still difficult. Pure copper is widely known for its high thermal conductivity, but the AM process still faces many challenges. Combining the advantages of the pure copper and TPMS porous structure to design HEs should yield amazing results, which provide a new way to fabricate a high-performance HE. In this project, a new HE design method with conformal Triply Periodic Minimal Surface (TPMS) core structures is proposed, a novel micro SLM (μSLM) processing technique for pure copper fabrication is explored, and a series of prototype pure copper heat exchangers with different cell sizes are fabricated and their performances are evaluated. The newly proposed shape-function-based transformation

method successfully achieves the smooth and continuous core structure in heat exchangers, which makes the complex external geometry of HE designs become possible. The optimal parameter window for micro SLM processing pure copper is found. A relative density (RD) of 99.6% and an electrical conductivity of 76.1% are observed in as-built pure copper components and the top/side/bottom surface roughness achieves a relatively low value of 3.3/14.8/19.2 μm, respectively. The pure copper TPMS heat exchanger is successfully fabricated and its performance is evaluated compared to stainless steel TPMS HE and traditional plate HE. It is found that the heat exchange effectiveness of copper TPMS HE is about 10% ~ 30% higher than that of SS316L TPMS HE and about 60% ~ 90% higher than that of SS316L plate HE. The convective heat transfer coefficients of copper TPMS HE is about 1.2 ~1.4 times that of SS316L TPMS HE and about 1.5 ~1.9 times that of SS316L plate HE. The volumetric heat transfer coefficient of copper TPMS HE is about 18 ~25 times that of polymer TPMS HE and outperforms similar plate HEs. Furthermore, the effect of TPMS cell size (cell size=4,5,6) and its gradient (cell size=4-6) on thermo-hydraulic characteristics of additive-manufactured pure copper HEs are investigated. The results demonstrate that the effectiveness of cell size=4 outperforms that of others. As the flow rate increases, the overall heat transfer coefficient and pressure drop increase, while the total thermal resistance decreases. The smaller the cell size, the higher the overall heat transfer coefficient and pressure drop, and the lower the total thermal resistance. The cell size=6 has the advantage in PEC due to the lowest pressure drop. The cell size=4 has a greater volume-based power density but also consumes a larger pressure drop per unit length. The present works have demonstrated the feasibility to fabricate ultra-compact HE with ultra-high performance.

1. OBJECTIVES AND SIGNIFICANCE

The objective of this project is to develop an ultra-compact heat exchanger with ultra-high performance. The proposed HE aims to: (1) realize higher heat transfer effectiveness than the current commercial heat exchangers; (2) explore the influence of cell size on thermo-hydraulic characteristics of AMed HE; (3) achieve the heat exchange core structure conformal filling in the complex geometry so that it can better adapt to the application environment; (4) eliminate the leakage risk and reduce the total volume; (5) reduce post-maintenance costs; (6) satisfy special decoration or customization requirement. The project results are expected to provide some guidance for further heat exchanger design.

Since the proposed HE has high heat transfer effectiveness, a higher solar energy conversion rate will be achieved. Problems that previously could not meet demand due to low conversion efficiency may be resolved. Given the high energy prices of oil and gas, this high efficiency HE will provide more incentives to encourage people to install the solar thermal system and reduce the consumption of fossil fuels. Furthermore, the proposed design method will also benefit other types of HE designs used in a variety of thermal control systems, such as refrigeration cycles, automotive industry, and electronic equipment, etc. Hence, there is good potential and wide economic and environmental significance derived from this project.

2. RESEARCH METHODOLOGY

The research methodology is implemented in four steps as described by follows:

2.1 Conformal TPMS core design based on shape function transformation and HE geometry selection

The method for generating complex geometry HEs filled with conformal TPMS structures is proposed by the PI as follows. Step 1: Select/create the macroscopic geometry design for HE (such as tube, plate, spiral, and helix) and keep them of the same weight or volume for performance evaluation; Step 2: Obtain the above-mentioned primary geometry that TPMS needs to infill in 3D solid format, which will be meshed by hexahedral elements; Step 3: obtain the mesh file with node locations and element connection information as the input for the transformation and assembly step; Step 4: Choose the regular TPMS structure type, generate it and rotate it to the preferred orientation based on the design requirement using Euler angles; Step 5: Conduct the iso-parametric transformation of coordinates for each cell/element such that the conformality in the mesh can be passed onto TPMS structures and assemble these cells/elements into the target geometry based on the element connection information. Step 6: Enclose the TPMS core structures with the HE macroscopic geometry using Boolean operation. Repeat Steps 1-6 until the best combination of heat transfer coefficient, compactness, and

pressure drop can be obtained.

2.2 μSLM process optimization for pure copper

Spherical fine pure copper powder (<25μm) was employed for μSLM fabrication to ensure good bonding between particles in the fine beam size μSLM setup. Printing of a single track will be first conducted to investigate the melting and solidification behavior under different process parameters, including laser power and scanning speed. The printing quality can be evaluated based on the morphology of a single track, including width, height, stability, and continuity, which is used to narrow down the processing window by identifying the most uniform and consolidated scan track. Then, the single wall will be fabricated to study the interlayer bonding to minimize the tendency of distortions, cracking, and pore formation, which is mainly affected by penetration depth under different parameters. Subsequently, the effect of hatch distance and scan strategies on distortion, surface roughness, dimensional accuracy, and density will be investigated by printing cubes and curvilinear geometry to obtain a fully dense part with good dimensional accuracy. In addition, the support structure re-design also has to be conducted in terms of pillar diameter, density, and angle to minimize the part distortion. Furthermore, the electrical and thermal conductivity of pure copper components will be measured, which also helps to determine the optimal process parameter set. This set will be employed to print the pure copper TPMS HEs. Stainless steel TPMS HEs will also be fabricated separately for benchmarking studies.

2.3 Numerical analysis to explore thermo-hydraulic characteristics

To better understand the temperature and velocity field distribution inside the heat exchanger during heat transfer, CFD analysis was conducted using Ansys Fluent 21.0 software. The liquid medium will be assumed to be incompressible and the fluid flow in HE will be supposed to be steady-state turbulent flow with Shear Stress Transport (SST) $k-\omega$ model. The whole heat transfer process will be governed by the continuity equation, energy equation, and momentum equations. For the FEM model setting, the velocity inlet and pressure outlet will be chosen as boundary conditions to ensure that they are consistent with the experimental conditions. 10 boundary layers in the fluid region near solid walls will be generated to guarantee calculation accuracy. Grid independence will be checked before simulation implementation. Furthermore, the temperature and velocity distributions as a function of volumetric flow rate and cell size will be investigated. The heat transfer mechanism will be studied.

2.4 Heat exchanger experimental apparatus construction and performance evaluation

Each fabricated HE will be mounted on a test loop and insulated to minimize the heat loss to the ambiance. T-type thermocouples will be employed to measure the inlet and outlet temperatures for each side, and the data will be recorded using a four-channel acquisition system. The pressure drop will be measured using two differential pressure transmitters. Fluids on both sides are supplied by two pumps and water tanks, and volumetric flow rates are measured by two flow meters. SS316L heat exchangers with preliminarily selected geometries will be printed to compare their performance in terms of heat transfer rate, pressure drop, and mass flow rate. The best shape among them will be selected and fabricated using pure copper. The final Cu TPMS HE will be benchmarked against the traditional HE of the same weight or volume to demonstrate its superior performance. Furthermore, pure copper TPMS HEs with different cell sizes (cell size=4,5,6,4-6) will be selected to explore their influence on thermo-hydraulic performance. The heat exchange effectiveness, overall heat transfer coefficient, pressure drop, thermal resistance, etc. will be calculated and the changing tendency will be illuminated.

3. RESULTS ACHIEVED

3.1 Heat exchanger with conformal TPMS core structure

The method for generating complex geometry HEs filled with conformal TPMS structures is proposed by our group and a variety of heat exchangers with different shapes are designed as shown in Fig. 1. The heat exchanger integrates the water inlet and outlet connectors into the shell and core structure, which effectively eliminate the risk of water leakage. The TPMS structure conformal filling along with the HX shape will reduce the pressure drop, saving pump power. Furthermore, we attempt to share the wall in the design of the heat exchanger so that the heat transfer effectiveness can be further improved.

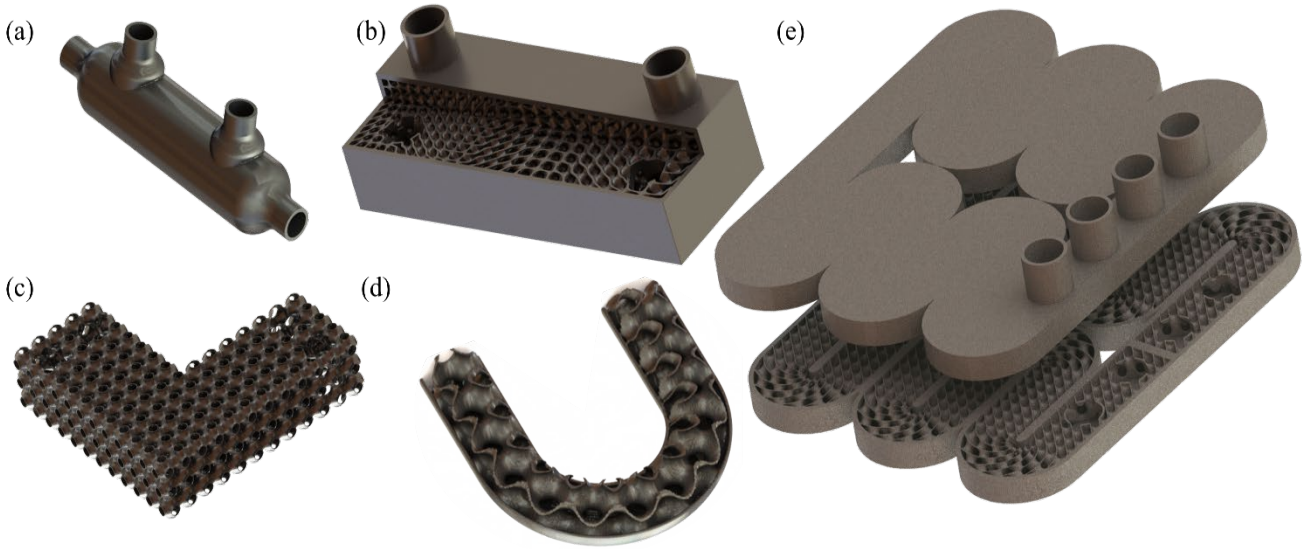


Fig. 1 The HEs with different external geometries: (a) tube design; (b) cube design; (c) L design; (d) U shape design; (e) conformal shape design with shared wall

3.2 µSLM process optimization for pure copper fabrication

The process windows for pure copper fabrication are explored in this project. The thin walls printed by the single tracks with different parameters are first investigated and the dimensions and roughness were measured as shown in Fig. 2. It is found that track densification can be realized when the laser power is in the range of 160-200W and the scanning speed is in the range of 350-800 mm/s. When the scanning speed is 650 mm/s, a top surface Ra of 3µm and a side surface Ra of 5µm are achieved. Bulk specimens are then fabricated to further explore copper formability in the µSLM process. The relative density of 99.6% and top/side surface roughness Ra of 4.2/7.8µm are achieved in bulk samples when the laser power is 200W, the scanning speed is 600 mm/s, and the hatch distance is 0.05 mm. The hatch distance (HD) study indicates that 0.05mm HD is the optimal parameter for the high-precision copper fabrication, both low HD and high HD result in the lack of fusion (LOF) phenomenon. Finally, the TPMS pure copper parts are fabricated and their minimum top-wall thickness reaches about 100 µm and minimum side-wall thickness reaches about 149 µm, respectively. The found optimal parameter window provides a guarantee for the manufacture of ultra-compact pure copper heat exchangers.

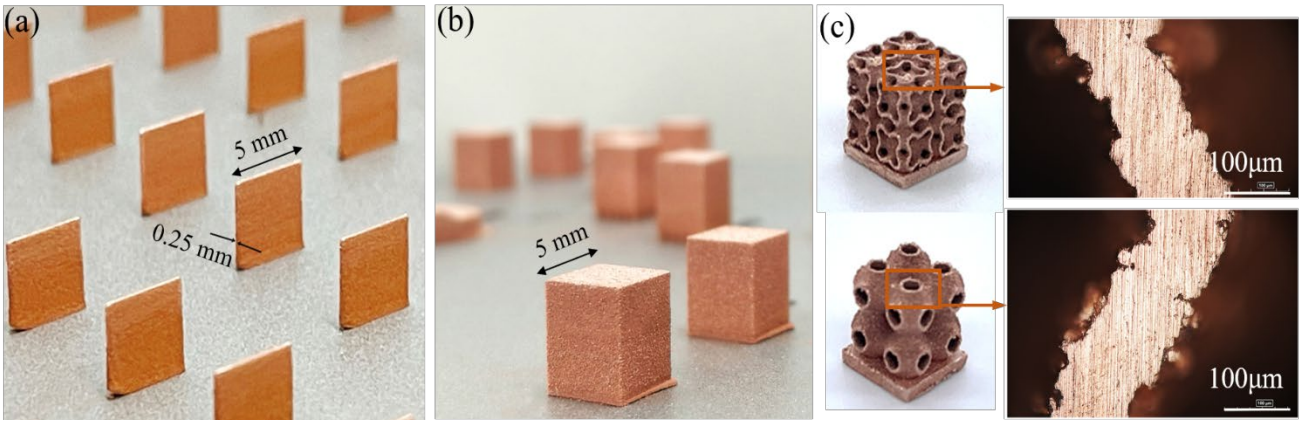


Fig. 2 The pure copper parts fabricated by our µSLM machine: (a) single walls; (b) cubes; (c) TPMS structures.

3.3 Thermo-hydraulic characteristics for heat transfer

The thermo-hydraulic characteristics of TPMS and plate heat exchangers are first investigated by simulation method. Both types of HEs exist in some flow-dead zones. However, compared to the plate HE, the designed

TPMS HE exhibits a more uniform velocity distribution in the edge regions on both sides. An eccentric-like velocity distribution is also observed in the fluid channels, and this situation is pronounced in the center of the cross-section. The interleaved channels in TPMS HE induces more complex flow patterns such as secondary flows and eccentric helical motion flows than Plate HE. These complex flow patterns combined with the constant changes in flow directions will intensify fluid mixing to some extent. The fluid boundary layer on the TPMS wall is probably squeezed. A higher convective heat-transfer coefficient is then probably obtained, thereby inducing a more uniform temperature distribution in TPMS HE than Plate HE.

The influence of cell size and its gradient on thermo-hydraulic characteristics of TPMS HE made of copper material is further studied. Four prototypes with different cell sizes are fabricated, and their heat transfer and fluid flow properties are explored and compared by numerical methods. The velocity distribution shows that the eccentric flow pattern goes throughout the heat exchanger and presents a rotation tendency as shown in Fig. 3. The eccentric flow patterns on the right area show a left-right distribution, while the flow patterns on the other area mainly present an up-down distribution. This satiation indicates the flow mechanism in HE is highly related to its TPMS morphology, inlet/outlet port distribution, and adopted flow patterns. Apparently, the smaller the cell size, the denser the eccentric flow in HE, and the more uniform the temperature distribution. The explored mechanism can allow people to better understand the heat transfer process and provide a reference for the design of novel TPMS HEs.

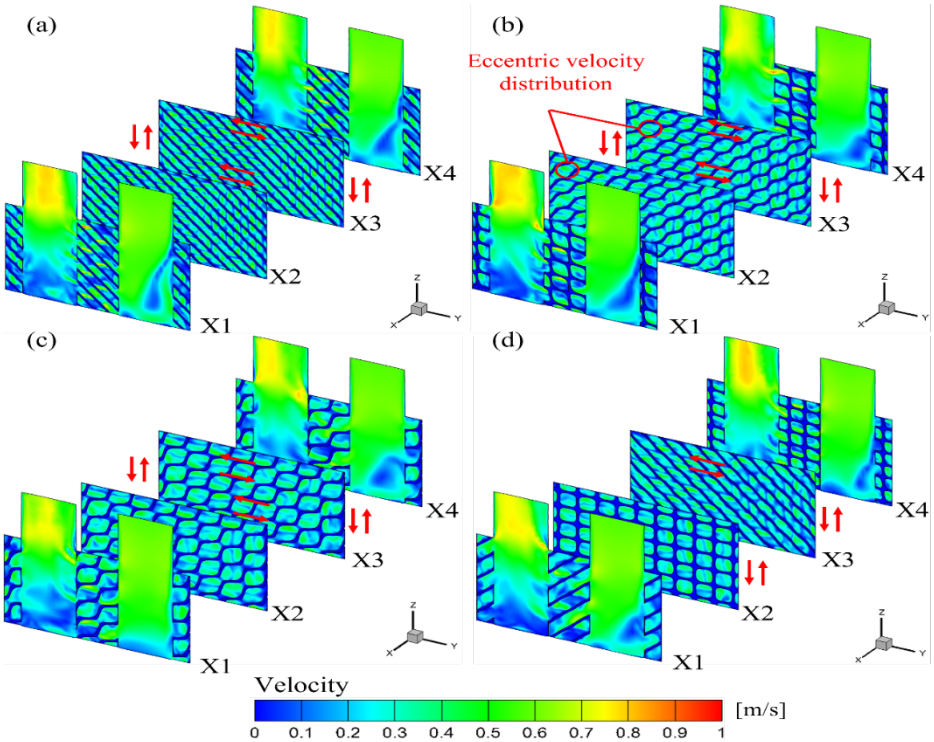


Fig. 3 The velocity distribution in a series of X planes for different cell sizes HE: (a) cell size=4; (b) cell size=5; (c) cell size=6; (d) cell size= 4-6.

3.4 Pure copper heat exchangers and performance evaluation

To evaluate the performance of our proposed pure copper TPMS heat exchangers, we also fabricated another two typical heat exchangers (SS316L TPMS HE and SS316L plate HE) as the references. Three HEs were then systematically studied and compared through ourself-developed experimental setup as shown in Fig. 4. The experimental results showed that the heat exchange effectiveness of copper TPMS HE is about 1.1 ~1.3 times that of SS316L TPMS HE and about 1.6 ~ 1.9 times that of SS316L plate HE. The convective heat transfer coefficient of copper TPMS HE is about 20 %~ 40% higher than that of SS316L TPMS HE and about 50% ~ 90% higher than that of SS316L plate HE. The calculated volumetric heat transfer coefficient of copper TPMS HE is about 18 ~25 times that of polymer TPMS HE and outperforms similar plate HEs.

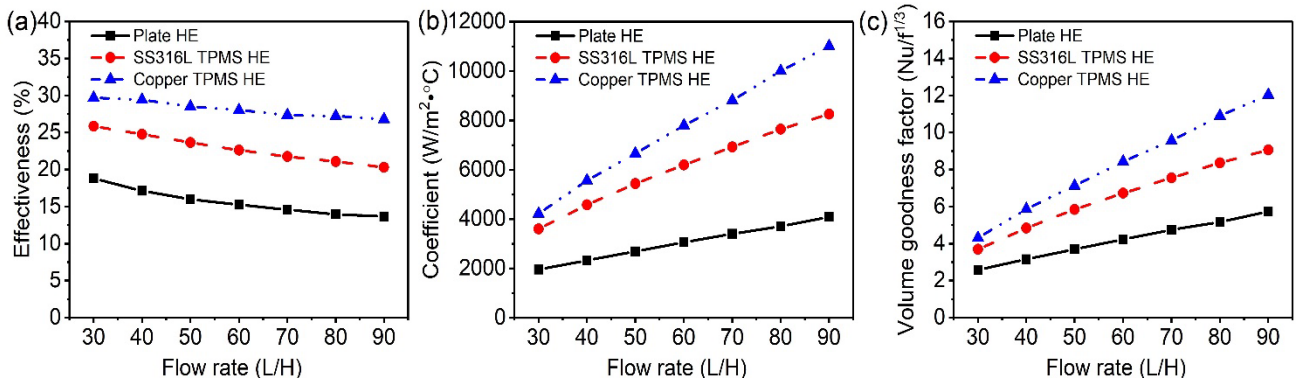


Fig. 4 The performance of proposed pure copper TPMS HE: (a) heat exchange effectiveness; (b) convective heat transfer coefficient; (c) volume goodness factor.

The influence of cell size and its gradient on the heat transfer performance of copper TPMS HEs are further studied by experimental methods. The corresponding results are shown in Fig. 5. It is found that the heat transfer rate increases with the increasing flow rate, while the effectiveness decreases as the flow rate increases. The difference in effectiveness between different heat exchangers does not vary significantly with flow rate, and the effectiveness of cell size=4 outperforms that of others. As the flow rate increases, the overall heat transfer coefficient U and pressure drop ΔP increase, while the total thermal resistance 1/UA decreases. The smaller the cell size, the higher the U and ΔP, and the lower the 1/UA. The PEC increases with the increase of flow rate. The cell size=6 has the advantage in PEC due to the lowest pressure drop. The cell size=4 has a greater volume-based power density but also consumes a larger pressure drop per unit length.

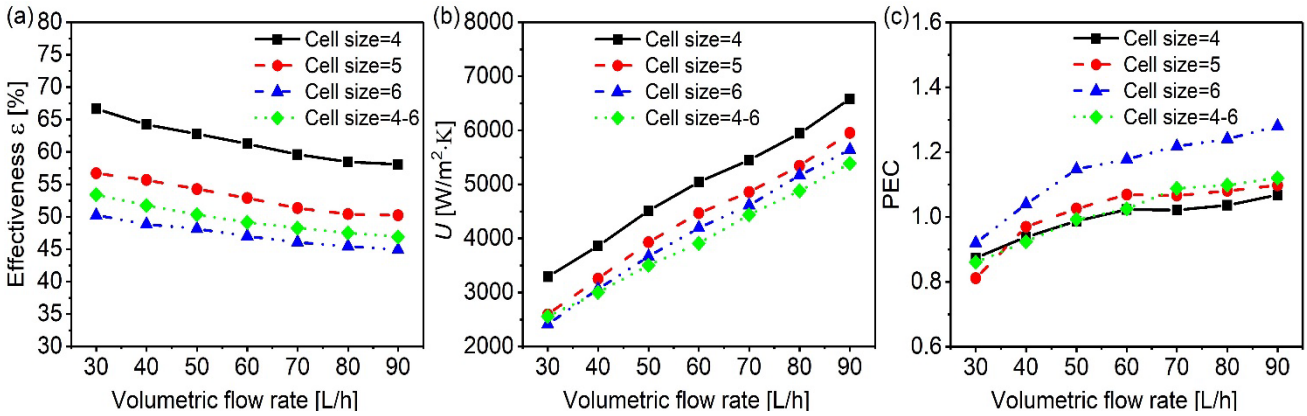


Fig. 5 The heat transfer properties of heat exchangers with different cell size: (a) heat transfer effectiveness ε; (b) overall heat transfer coefficient (U); (c) performance evaluation coefficient (PEC).

4. PUBLICATION AND AWARDS

Total ten (10) high quality journal papers have been published [1-10]. Moreover, the uniquely designed ultra-compact heat exchanger is awarded the "Red Dot Award: Design Concept 2022" [11].

1. S Gao, J Ding, S Qu, H Liu, X Song, Numerical and experimental investigation of additively manufactured shell-lattice copper heat exchanger, *International Communications in Heat and Mass Transfer* (2023) 147: 106976.
2. S Gao, S Qu, J Ding, H Liu, X Song, Influence of cell size and its gradient on thermo-hydraulic characteristics of triply periodic minimal surface heat exchangers, *Applied Thermal Engineering* (2023) 232: 121098.
3. Z Hu, S Gao, J Tai, S Qu, J Ding, X Song, Z Fan, Columnar grain width control for SS316L via hatch spacing manipulation in laser powder bed fusion, *Materials Research Letters* (2023) 11: 231-238
4. J Fu, J Ding, L Zhang, S Qu, X Song, MW Fu, Development of conformal shell lattices via laser powder bed fusion and unraveling their mechanical responses via modelling and experiments, *Additive*

Manufacturing, (2023) 62: 103406
5. X Guo, J Ding, X Li, S Qu, JYH Fuh, WF Lu, X Song, W Zhai, Interpenetrating phase composites with 3D printed triply periodic minimal surface (TPMS) lattice structures, *Composites Part B: Engineering* (2023) 248: 110351
6. J Fu, J Ding, S Qu, L Zhang, MY Wang, MW Fu, X Song, Improved light-weighting potential of SS316L triply periodic minimal surface shell lattices by micro laser powder bed fusion, *Materials & Design* (2022) 222: 111018
7. S Qu, J Ding, J Fu, M Fu, X Song, Anisotropic material properties of pure copper with fine-grained microstructure fabricated by laser powder bed fusion process, *Additive Manufacturing* (2022) 59: 103082
8. J Ding, S Qu, L Zhang, MY Wang, X Song, Geometric deviation and compensation for thin-walled shell lattice structures fabricated by high precision laser powder bed fusion, *Additive Manufacturing* (2022) 58: 103061
9. J. Fu, H. Li, X. Song, MW. Fu, Multi-scale defects in powder-based additively manufactured metals and alloys, *Journal of Materials Science & Technology* (2022) 122: 165–199 **(ESI highly cited paper)**
10. S Qu, J Ding, J Fu, M Fu, B Zhang, X Song, High-precision laser powder bed fusion processing of pure copper, *Additive Manufacturing* (2021) 48 (Part A): 102417

Red Dot Design Award: Design Concept (2022) Title: Ultra-compact heat exchanger fabricated by 3D printing
https://www.red-dot.org/project/ultra-compact-heat-exchanger-fabricated-by-3d-printing-60272

Ultra-Compact Heat Exchanger Fabricated By 3D Printing

← Back

Download

This ultra-compact heat exchanger, created using additive manufacturing, maximises heat transfer in a confined space. It features a mathematically perfect Triply Periodic Minimal Surface (TPMS) that is intrinsically smooth and continuous, allowing it to be used as the core structure for the compact heat exchanger.

This TPMS design achieves a much larger heat transfer capability (more than ten times improvement) due to its unique properties and high surface- area-to-volume ratio.

The incoming hot and cold fluids will flow through these channels to complete the heat exchange. The conformal filling of these TPMS structures in the heat exchanger significantly improves heat exchange efficiency and reduces fluid pressure drop. Because of its unique internal TPMS structure and integrated one- piece design, this heat exchanger is very durable and highly efficient, lowering maintenance costs.

Red Dot Award: Design Concept | Ready to Launch | 3D Printed

reddot award 2022
winner

Credits

Institution:
The Chinese University of Hong Kong, Hong Kong

Faculty Advisor:
Prof. Song Xu

Design:
Ding Junhao, Dr. Gao Shiming, Liu Hui, Mo Haoming, Qu Shuo

COMPLETED

PRINTABLE AEROGELS FOR ON-CHIP THERMAL MANAGEMENT AND THERMOELECTRIC POWER GENERATION

Principal Investigator: Professor Guohua HU
Department of Electronic Engineering, CUHK

Research Team Members:
LIU Yang[(1)], and WEN Yingyi[(1)]

[(1)] Dept. of Electronic Engineering, CUHK

Project Start Date: 1st July 2021
Completion Date: 31st July 2023

INNOVATION AND PRACTICAL SIGNIFICANCE:

In this project, we propose additive printing of aerogels for on-chip thermal management and thermoelectric power generation. Though aerogels are promising for thermal management, their brittleness presents challenges in subtractive processing for any practical applications. We propose the addition of low-dimensional materials may allow additive printing and manufacturing of the aerogels. A method to achieve this remains elusive and, if implemented, will bring tremendous innovation and impact to the forefront of research on aerogels. We envisage that the printable aerogels can bring an exciting solution for the thermal management of the circuit boards. The circuit electronics such as transistors, diodes, capacitors and resistors can be significantly affected by the ambient temperature. Reliable thermal management is essential for their stable, high-performance operation. The printing strategy can be readily adapted to industrial-scale high-speed press for real-world aerogel manufacturing, bringing huge impact and economic benefits. Besides, we suggest the printable aerogels may enable high-performance on-chip thermoelectric power conversion from the ambient thermal waste, arising from the ultralow thermal conductivity of the aerogels and the good thermoelectric property of the low-dimensional materials. Though the thermoelectric power may not be valid for many applications, we suggest it can be exploited as a power source for low-power electronics, for instance, IoT sensors. Currently, one of the major limitations of the ubiquitous deployment of IoT sensors is the power supply. We believe the printed aerogels with the on-chip thermoelectric power generation capability will be of great significance to IoT development and the fourth-generation industrial revolution.

ABSTRACT

As ultralight solids with 99% air, aerogels present an ultralow thermal conductivity, making them promising for thermal management. However, their brittleness renders subtractive processing and manufacturing challenging. Recent studies show printing of aerogels is possible with the addition of nanomaterials. The ultralow thermal conductivity and the printability present the potential for up-scaled, additive manufacturing of aerogels for practical thermal management applications. In this project, we aim to develop printable aerogels by adding low-dimensional materials, and on this basis, on-chip aerogel printing. We propose using the printed aerogels in thermal insulation of circuit electronics from the ambient heat dispensation and temperature variations. Operation at a stable temperature ensures the high performance of the electronics. The low-dimensional materials may also allow energy conversion from the ambient thermal waste through a thermoelectric effect. We expect the printed aerogels, with the ultralow thermal conductivity of the aerogels and the thermoelectric effect of the low-dimensional materials, can enable efficient on-chip thermoelectric power generation. This may be exploited as a power source for widespread low-power electronics raining from IoT sensors and flexible/wearable electronics to bioelectronics and biomedical implants.

1. OBJECTIVES AND SIGNIFICANCE

Objectives:

1) Developing aerogels with the addition of low-dimensional materials (LDMs). Aerogels, with their brittleness, present challenges for subtractive processing. Mixing LDMs into aerogels may allow additive manufacturing, e.g., by printing.
2) Realising printable aerogels and on-chip printing. Based on *Objective 1*, the rheology of the aerogel and LDM precursor blends will be tuned to attain non-Newtonian shear-thinning fluids for printing. Then, aerogel printing on circuit boards will be investigated.
3) Validating on-chip thermal management and thermoelectric power generation. The thermal insulation and thermoelectric power generation performance of the printed aerogels will be assessed. The materials and printing will be engineered to achieve optimal performance.

Significance:

1) Research impact: Additive printing of aerogels is an emerging technology to enable up-scaled, high-performance thermal management. However, achieving printing of aerogels has thus-far remained challenging. We believe our method of incorporating LDMs will be a viable strategy, and if implemented, will be impactful for the forefront of research on aerogels. Successful aerogel printing will bring the following impacts at the commercial level.
2) Thermal management: Thermal management is dispensable for electronics, optics, sensors, bioelectronics, *etc*. For instance, the operating temperature of electronics can significantly affect their performance. Effective thermal management is hence critical. We believe the proposed printed aerogels can provide effective on-chip thermal insulation of the electronics from the ambient heat dissipation and temperature variations. This method can also be readily adapted to the existing industrial-scale high-speed press to deliver real-world additive aerogel manufacturing for thermal management, bringing the application and economic benefits.
3) Thermoelectric power generation: LDMs, *e.g.* carbon nanotubes (CNTs) and graphene, present promising thermoelectric properties. They are capable of recycling the ambient thermal waste and converting it into electrical power. The conversion efficiency depends on the *ZT* factor, defined by the *Seebeck* coefficient, the electrical and thermal conductivities, and the device operating temperature. Owing to the thermoelectric effect of LDMs and the thermal insulation of aerogels, the printed aerogels may enable efficient on-chip thermoelectric power generation that may be exploited as a power source for low-power electronics, for instance, IoT sensors, flexible/wearable electronics, bioelectronics and biomedical implants.

2. RESEARCH METHODOLOGY

The proposed methodology is set out as (1) realising aerogel incorporation with LDMs, (2) realising printable aerogels and additive printing of aerogels, and (3) investigating and validating the thermal management and thermoelectric power generation capabilities of the printed aerogels.

(1) Realising aerogel incorporation with LDMs: 1. LDMs, including graphene and CNTs, will be produced via liquid-phase exfoliation. Through sonication, centrifugation, vacuum filtration and drying steps, LDMs will be obtained. Besides, other LDMs, e.g. MoS2 in the metallic phase, will also be produced and investigated; 2. The LDMs will then be mixed with silica aerogel precursors for aerogel development. A typical method: LDMs, ethyl silicate, 1-pentanol and water will be mixed and blended, and then stirred vigorously to obtain a stable silica-LDM precursor blend; 3. The produced silica-LDM precursor blend will be placed in a closed box. The box will be filled with ammonia gas. Ideally, under this situation, the blend will be solidified to form the proposed LDM aerogels.

(2) Realising printable aerogels and additive printing of aerogels: Once we have developed a working protocol to produce the silica-LDM aerogels, we will then investigate the development of printable silica-LDM inks. In a typical method, the composition of the silica-LDM precursor blend (including the concentrations of the LDMs and the aerogel precursors as well as the solvent selection) will be precisely

controlled to tune the surface tension and viscosity of the blend. A suitable rheology of the silica-LDM precursor blend will be non-Newtonian shear-thinning. For inkjet printing, a viscosity of 1-10 mPa·s may be appropriate; for the PCB board printer, a viscosity of 100-1,000 mPa·s may be appropriate. To further tune the rheology, we will also consider the addition of polymers such as cellulose. After we have developed inks with the suitable rheology and viscosities, we will then proceed to print with the inkjet printer and the PCB board printer. After the structures are printed, the structures will then be solidified with ammonia, as described in Objective 1. On this basis, we will investigate printing on circuit boards.

(3) Investigating and validating the thermal management and thermoelectric power generation capabilities of the printed aerogels: After we have demonstrated on-chip aerogel printing, we will then assess the performance of the printed aerogels for thermal management and thermoelectric power generation. For thermal management, we will apply high and low temperatures to the circuit boards for a certain amount of time and then measure the resultant temperature of the circuit electronics. Ideally, the circuit electronics shall remain at a stable temperature. Otherwise, the materials and printing processes will be engineered to achieve high thermal insulation performance. For thermoelectric power generation, we will again apply high and low temperatures to the circuit boards and then measure the generated electric voltage from the printed aerogels. Again, ideally, the printed aerogels can effectively generate thermoelectric power. Otherwise, the materials and the printing processes will be engineered to achieve a high performance.

3. RESULTS ACHIEVED

We have carried out the following experiments in line with the above research plan:

For ***Obj 1***:

- We have achieved the fabrication of silica aerogels from the solution precursors. The silica aerogels are of a light weight due to its porous structure (Fig. 1a), and can achieve a good thermal insulation in high and low temperatures (Fig. 1b and c).
- Based on state-of-the-art research on thermoelectric, we have identified two-dimensional bismuth telluride (Bi_2Te_3) as the potential material for thermoelectric power generation. We have therefore conducted chemical exfoliation production of Bi_2Te_3 (Fig. 1d) and demonstrated thermoelectric power generation from the printed Bi_2Te_3 thin-film (Fig. 1e and f). Besides, to modulate the carrier concentration for enhancements in the thermoelectric performance, we have identified Stark electrostatic modulation as a potential engineering method. Based on this, we have thus far demonstrated the effectiveness of the Stark electrostatic modulation for carrier concentration modulation in two-dimensional materials (please see the attached manuscript). However, further work is needed (and also arranged) to modulate the carrier concentration in the Bi_2Te_3 for an optimized thermoelectric power generation performance and the development of high-performance thermoelectric electronics. Note that the thermoelectric voltage generation (~0.2 mV/°C) though not optimized yet may be exploited in the thermal sensing and low-power electronics applications.

For ***Obj 2***:

- We have investigated slot-die printing (attached video 1) and 3D printing (attached video 2) of the silica aerogels. Printing requires well-refined ink compositions to attain well-suited rheology. We have tuned the composition and the physical property of the silica aerogel precursors such that the silica aerogels are printable. However, the slot-die and the 3D printing processes still have to be further optimized to deliver scalable, high-resolution on-chip printing of silica aerogels.

For ***Obj 3***:

- As mentioned above, we have conducted thermal management investigation of the produced silica aerogels (Fig. 1b and c). The silica aerogels show a good thermal insulation in heating and cooling conditions.

- However, as we have not yet optimized the slot-die and the 3D printing of aerogels, we have yet to demonstrate on-chip printed aerogels and investigate the thermal management and thermoelectric power generation from the on-chip printed aerogels.

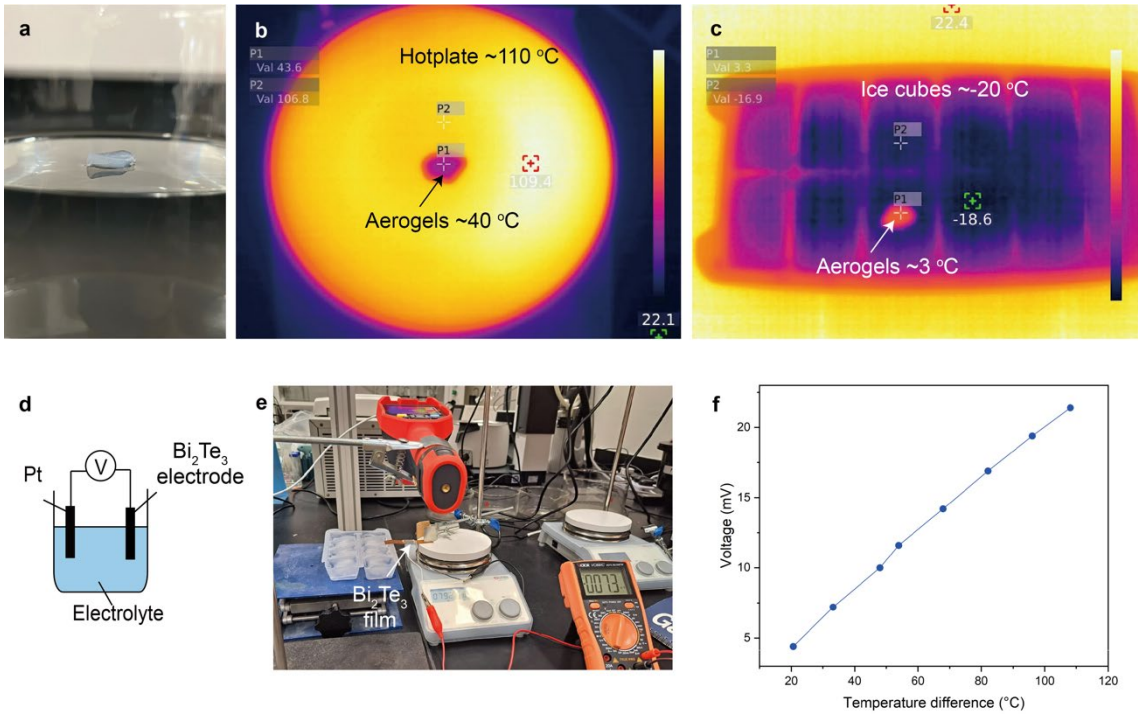


Figure 1. Printable aerogels: (a) A piece of silica aerogel floating on water. (b) Thermal insulation of the aerogel in the heating condition, showing an effective insulation of high temperature. (c) Thermal insulation of the aerogel in the cooling condition, showing an effective insulation of high temperature. (d) Schematic figure showing chemical exfoliation of Bi_2Te_3. (e) The set-up showing thermoelectric power generation from a printed Bi_2Te_3 thin-film. (f) Measured thermoelectric voltage generation with respect to the temperature difference, giving ~0.2 mV/°C.

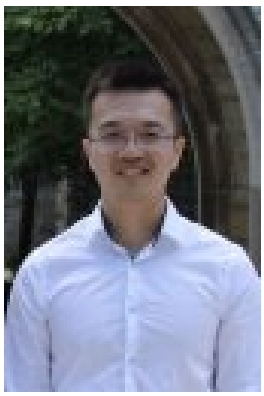
4. PUBLICATION AND AWARDS

C[1] Y. Wen, Y. Liu, and G. Hu, "Printable thermoelectric aerogel for thermal management and energy recycling of power ICs," MRS Fall Meetings, Boston, USA, 2022.
J[1] S. Liu, X. Fan, Y. Wen, P. Liu, Y. Liu, J. Pei, W. Yang, L. Song, D. Pan, T. Ma, Y. Lin, G. Wang*, and G. Hu*, "Conduction modulation of solution-processed two-dimensional materials," submitted.
J[2] S. Liu, X. Fan, Y. Wen, P. Liu, Y. Liu, J. Pei, W. Yang, L. Song, D. Pan, T. Ma, Y. Lin, G. Wang*, and G. Hu*, "Reservoir activation with high-order nonlinear solution-processed MoS_2 devices," submitted.
J[3] L. Song, P. Liu, J. Pei, F. Bai, Y. Liu, S. Liu, Y. Wen, L. W. T. Ng, K.-P. Pun, S. Gao, M. Q.-H. Meng, T. Hasan, and G. Hu, "Spiking Neurons with Neural Dynamics Implemented Using Stochastic Memristors," *Advanced Electronic Materials*, 2300564, 2023.

COMPLETED

MODELING OF ENVIRONMENTAL EFFECTS ON PERFORMANCE DEGRADATION OF OFFSHORE WIND TURBINE BLADES MADE BY CARBON FIBER REINFORCED PLASTICS (CFRPS)

Principal Investigator: Professor ZHANG, Weizhao
Department of Mechanical & Automation Engineering, CUHK



Research Team Members:
M.Erfan Kazemi, PosDoc researcher [(1)]
Wanrui Zhang, PhD student [(1)]

[(1)] Department of Mechanical and Automation Engineering

Project Start Date: 1st July 2020
Completion Date: 30th June 2022

INNOVATION AND PRACTICAL SIGNIFICANCE:

The proposed research work will lead to the first-ever multiscale modeling method to analyze and predict performance degradation of CFRPs under extreme offshore working conditions, i.e., seawater immersion and high temperature. At microscopic and mesoscopic levels, the method will provide (1) innovative mechanical-chemical-physical analysis to investigate the degradation mechanisms that are difficult to be captured by pure experiments; and (2) cost-efficient numerical characterization of CFRPs constitutive law with environmental effects considered. The obtained constitutive law is then to be implemented into the state-of-the-art macroscopic model for offshore wind turbine blade design optimization and damage monitor. Upon completion, this method can achieve efficient design and accurate failure prediction for blades during their lifecycles. Hence, this method can greatly reduce manufacturing, usage and maintenance cost of offshore wind turbines, expanding application of renewable wind energy significantly. Furthermore, this modeling method can be expanded to other applications under extreme environment, such as marine vehicles and buildings that can enlarge space for human activities and alleviate rising issues resulted from limited resources.

ABSTRACT

The pressing need to reduce greenhouse gas emissions has led to increasing installation of offshore wind turbines, which usually have blades made with carbon fiber reinforced plastics (CFRPs) that have high performance-to-weight ratios. Modern methods utilize finite element analysis (FEA) to design CFRP blades at high speed and low cost, but none of these FEA models consider degradation of CFRPs under extreme offshore conditions, which have been experimentally proved to significantly impair long term performance of blades. To efficiently design CFRP blades with long lifecycle and avoid high safety factors, this project aims to develop a cutting-edge MD (molecular dynamics)-FEA integrated modeling method to simulate performance degradation of CFRP wind turbine blades with offshore environmental effects included. Upon completion, a state-of-the-art numerical tool will be established to aid material selection and new material design for CFRPs that are to be applied under marine or other harsh conditions. The underlying principles about electrochemistry erosion at CFRP-metal interfaces can also be obtained to guide construction of CFRP part joints and selection of component materials for water-involved applications. Furthermore, a flexible numerical tool that can customize blade design and guide blade damage monitor can also be developed.

1. OBJECTIVES AND SIGNIFICANCE

1. To numerically investigate performance degradation of carbon fiber reinforced plastics (CFRPs) under salty-wet-hot conditions that are common under extreme offshore environment but not yet systematically researched. A state-of-the-art numerical tool will be established to aid material selection and new material design for CFRPs that are to be applied under marine or other harsh conditions.

2. To understand electrochemistry erosion of metal components assembled with CFRP parts because such erosion will significantly impair performance of machines under seawater immersion. Underlying principles about this electrochemistry erosion will be obtained to guide construction of CFRP part joints and selection of component materials for water-involved applications.

3. To create a flexible numerical tool that can customize blade design. Blades need to be built differently as their working conditions change, but currently there is no one-for-all approach for blade design. By applying the proposed modeling method, optimal blades with highest performance-to-cost ratios can be obtained, which will boost utilization of offshore wind turbines to generate clean renewable electricity. In addition, this modeling method can also guide blade damage monitor to improve inspection and maintenance efficiency.

2. RESEARCH METHODOLOGY

In order to fill in the current gaps, an innovative multiscale method was established, and the development roadmap of this method is illustrated in Figure 1. In this method, the smallest length scale is micro level and the modeling is achieved via molecular dynamic (MD) simulation that can capture movement of material molecules, so salty-wet-hot effects on properties of CFRPs can be investigated. MD simulation can also assist polymer research since detailed physical-chemical-mechanical mechanisms that are difficult to be observed by experiments can be studied with high fidelity. Furthermore, in the project, MD results are integrated with the follow-up models with larger scales, namely, the mesoscopic unit cell model and the macroscopic part model.

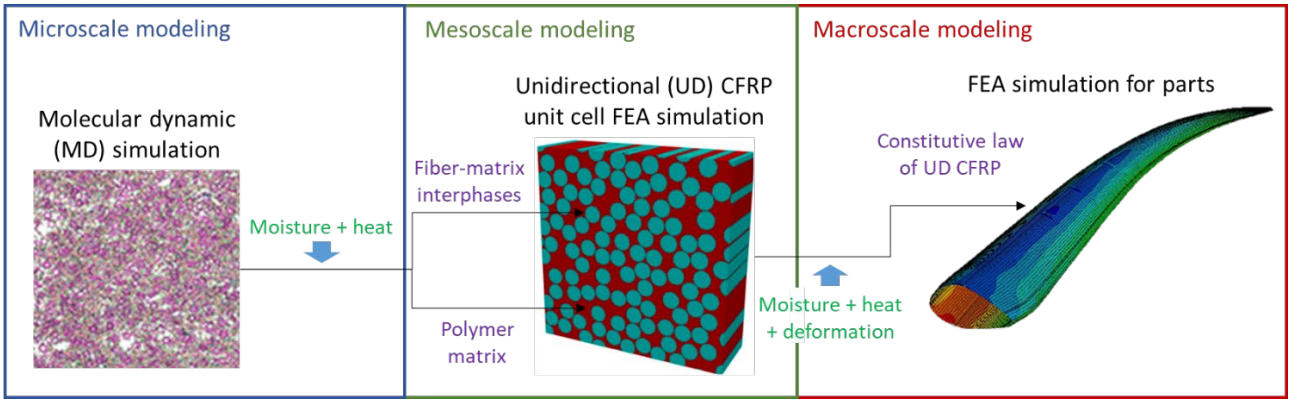


Figure 1. Roadmap of multiscale modeling method for CFRP parts under offshore environment.

The mesoscopic unit cell model is realized via FEA for its accuracy. The unit cell is designed to have unidirectional (UD) fiber alignment, which is the most popular choice for current wind turbine application. In the mesoscopic model, environmental effects on mechanical properties of various constituents of the CFRPs are characterized via MD simulation combined with experiments. Previous research has revealed that influence of salt on pure CFRPs is negligible, so only moisture and temperature are considered in the characterization. After mesoscopic material properties are determined, moisture, temperature and mechanical deformation are applied to the unit cell model to (1) study CFRP degradation dynamically; and (2) provide homogeneous constitutive law of CFRPs as a function of environmental effects for macroscopic part analysis. The homogenization is achieved via an analytical approach for efficiency.

For macroscopic FEA modeling of CFRP parts, material constitutive law is characterized by the unit cell model. The remaining challenge is to capture the non-uniform fiber directions that are caused by part surface curvature and dominate CFRP anisotropy. Previous benchmark tests have validated that preforming models developed by the PI can predict fiber directions in parts with large curvature variation. These models are utilized here with proper modification to establish direct connection between virtual and real world for design, failure

perdition and damage monitor of the parts.

3. RESULTS ACHIEVED

Project achievement:
1. Developed a four-stage theory based on experimental analysis and numerical modeling to describe hydrothermally ageing of CFRPs based on fundamental mechanisms instead of phenomenology.

2. Predicted performance degradation of CFRPs under seawater conditions by establishing an innovative numerical tool with high-fidelity molecular dynamic and finite element models

3. Established a virtual characterization tool via representative volume element analysis of CFRPs to efficiently identify interface strength between fibers and resin under various ageing conditions with less than 3.3% error.

Commercialization activity:
The PI's team is currently collaborating with BWSea Technology (Shanghai) Co. Limited. to further develop the project outcomes, aiming to extend it to virtual and efficient design and long-term maintenance of next-generation marine engineering equipment made of lightweight CFRPs.

4. PUBLICATION AND AWARDS

J[1] Han Z, Xiong Y, Feng Y, Zhang W, Zhang W. Manufacturing woven carbon fabric electrodes with high areal capacity via high-pressure spray of active lithium-ion particles. Manufacturing Letters. 2023 Aug 1;35:152-9.

J[2] Feng Y, Zheng W, Zhao Z, Liang B, Ye H, Ge Q, Zhang X, Zhang W. Characterization and Finite Element Modeling for Thermoset Resin of Carbon Fiber Prepregs During Curing. Journal of Manufacturing Science and Engineering. 2022 Aug 1;144(8):081007.

J[3] Feng Y, Han Z, Li R, Zhang W. Numerical modeling for curing of unidirectional carbon fiber reinforced polymer based on micromechanics in Laplace domain. Composites Science and Technology. 2022 Sep 29;228:109637.

J[4] Zhang W, Kazemi M, Mokarizadeh A, Sahraei A, Arab B, Zhang W. Experimental testing and molecular dynamics (MD) simulation for ageing of carbon fibre-reinforced polymers (CFRPs) under hygro-thermal-salt conditions. Composites Science and Technology. (*under review*)

J[6] Zhang W, Liu M, Han Z, Zou J, Liang B, Zhang W. A constitutive modelling method to predict the mechanical behaviour of hydrothermally aged epoxy resin and carbon fiber composites using finite element analysis (FEA). (*in preparation*)

J[7] Zhang W, Liu M, Han Z, Zou J, Liang B, Zhang W. Integrated numerical-experimental study on hydrothermal ageing mechanism of carbon fiber reinforced polymers (CFRPs). (*in preparation*)

COMPLETED

ENERGY MANAGEMENT SYSTEM FOR LARGE-SCALE ELECTRIC VEHICLE CHARGING WITH RENEWABLE GENERATION AND ENERGY STORAGE

Principal Investigator: Professor Yunjian XU
Department of Mechanical & Automation Engineering, CUHK

Research Team Members:
Dr. Jin Jiangliang [1]

[1] Dept. of Mechanical and Automation Engineering, CUHK

Project Start Date: 1st July 2019
Completion Date: 30th June 2021

INNOVATION AND PRACTICAL SIGNIFICANCE:

Hong Kong has more than 11,000 plug-in EVs in late July 2017. Many major countries, including the U.S. and China, have witnessed fast-growing adoption of plug-in EVs and intermittent renewable generation. The intermittency and stochasticity of renewable generation (from solar and wind) impose significant challenges on the real-time supply-demand balance of electric power system operation. The key innovation of the proposed project is two-fold: i) a stochastic optimal control based computational approach that optimally schedules the charging of a large number of EVs by explicitly taking into account the stochasticity in future renewable generation, EV arrivals, and electricity prices, and ii) a hardware-in-the loop (HIL) simulator that demonstrates the advantages of the developed energy management system (over existing technical approaches) with actual hardware components simulating real-world power system conditions. The developed HIL demonstrator will facilitate the technology transfer and follow-up funding applications for the potential commercialization of our research results on the cost-minimizing coordinated charging of a large number of EVs for a power distribution system with significant renewable generation.

ABSTRACT

This project aims to develop an energy management system for electric vehicle (EV) charging stations equipped with an energy storage system (e.g., reused EV batteries) and distributed renewable generation (e.g., rooftop solar and small-scale wind generation). The quickly growing adoption of EVs and intermittent renewable generation will impose significant challenges on the secure and efficient operation of electric power systems. We develop a novel approach (that combines the advantages of stochastic optimal control techniques and data-driven approaches) to harness the inherent flexibility in (deferrable) EV charging load for renewable generation integration and operational cost reduction. Success of the proposed research would maximize the economic and environmental benefits of EV adoption for Hong Kong.

This project seeks to develop i) a software package that optimally coordinates the charging of a large number of (up to 1000) EVs and the operation of an energy storage system to minimize the long-term expected system cost, under random renewable generation, EV arrivals, and electricity prices, and ii) a hardware-in-the loop (HIL) demonstrator that implements and tests the developed computational approaches and power electronic controllers in a realistic hardware environment simulating real-world EV charging station and power distribution system conditions.

1. OBJECTIVES AND SIGNIFICANCE

1. Develop scalable algorithmic approaches to compute the optimal scheduling policies for up to 1000 EV chargers and an energy storage system with intermittent renewable generation.

2. Develop a software module that trains and updates probabilistic models describing future renewable generation and EV arrivals with real-world data.

3. Develop a software package for real-time decision making on the charging of (up to 1000) EVs, based on the updated probabilistic information about future renewable generation and EV arrivals.

4. Implement the developed software package in a hardware-in-the-loop (HIL) simulation platform that verifies the performance of the developed energy management system with actual hardware components in real-world power distribution system environment.

Hong Kong has more than 11,000 plug-in EVs in late July 2017. Many major countries, including the U.S. and China, have witnessed fast-growing adoption of plug-in EVs and intermittent renewable generation. The intermittency and stochasticity of renewable generation (from solar and wind) impose significant challenges on the real-time supply-demand balance of electric power system operation. The key innovation of the proposed project is two-fold: i) a stochastic optimal control based computational approach that optimally schedules the charging of a large number of EVs by explicitly taking into account the stochasticity in future renewable generation, EV arrivals, and electricity prices, and ii) a hardware-in-the loop (HIL) simulator that demonstrates the advantages of the developed energy management system (over existing technical approaches) with actual hardware components simulating real-world power system conditions. The developed HIL demonstrator will facilitate the technology transfer and follow-up funding applications for the potential commercialization of our research results on the cost-minimizing coordinated charging of a large number of EVs for a power distribution system with significant renewable generation.

2. RESEARCH METHODOLOGY

Through a stochastic optimal control formulation, we seek to fully characterize a policy that optimally coordinate the charging of a large number of EVs and energy storage operation so as to minimize the (daily or weekly) expected operational costs of EV charging stations (including electricity cost and penalty cost for not fulfilling EVs' charging demands). Based on the optimal policy characterizations, we will develop and implement scalable computational approaches in a software package making real-time optimal decisions under random EV arrivals, renewable generation, and electricity prices. We will build a hardware-in-the-loop (HIL) demonstrator that simulates the performance of the developed energy management system in the actual hardware environment of an EV charging station equipped with solar PV generation and energy storage systems.

3. RESULTS ACHIEVED

This project focuses on the joint scheduling of EV charging and energy storage devices under random renewable generation and electricity prices. In the past 9 months, we have established structural characterizations of optimal policies, and developed efficient algorithms to compute the optimal scheduling policies with real-world data inputs, for two different application scenarios with constant and adjustable EV charging power.

1) For the case with constant (non-adjustable) charging power, the decision on EV's charging is binary. We formulate the scheduling problem as a restless multi-armed bandit (RMAB) problem. Relaxing the scheduling problem into multiple independent single-arm scheduling problems, we define the Lagrangian priority value as the greatest tax under which it is optimal to activate the arm. We propose a Lagrangian priority policy which processes EVs in the order of their Lagrangian priority values, and establish its asymptotic optimality as the system scales. Numerical results on real-world data show that the proposed Lagrangian priority policy achieves

22%-49% higher average reward than the classical Whittle index policy.

These results will be published at a flagship journal in the field of automatic control [J1].

2) For the case with constant (non-adjustable) charging power, we have further considered the joint scheduling of EV charging load and energy storage devices under random renewable generation and electricity prices [J3]. We establish an index based priority rule that is shown to be optimal under arbitrary system dynamics: tasks with less laxity should be processed first, and for two tasks with the same laxity, the task with a later deadline should have the priority. Based on the established optimal control policy characterizations, we propose to apply deep reinforcement learning (RL) methods to compute the total charging power (to all EVs). Numerical results on real-world data show that the proposed approach significantly outperforms existing RL methods combined with the earliest deadline first (EDF) priority rule (by reducing 46-53% of system cost).
These results have been published at a flagship journal in the field of automatic control [J3].

3) For the case with adjustable charging power, the decision on EV's charging is in a continuous action space. We formulate the cost-minimizing scheduling problem faced by an EV charging station operator as a dynamic program. When the number of EVs is large, the formulated dynamic program cannot be exactly solved by brute-force methods due to the curse of dimensionality. We show that given the total amount of energy charged into all EVs, the optimal energy allocation (among EVs) follows a threshold policy, which attempts to charge/discharge all EVs' battery to a target State-of-Charge (SoC), which enables the development of scalable computational approaches. The proposed approach achieves close-to-optimal performance in numerical experiments with real-world electricity pricing and solar generation data.

These results have been published at a flagship journal in the field of automatic control [J4].

4) Based on the established theoretic and algorithmic results, we develop a hardware-in-the-loop (HIL) simulation platform that implements the developed software package to verify its performance with actual hardware components in real-world power distribution system environment. Experimental results on a scaled-down 2.5kW EV charging system with OPAL-RT OP4510 simulator show that the proposed EV charging scheduling system reduces 20.1% and 9.6% of the total cost resulting from uncontrolled charging and a popular optimization based EV charging scheduling algorithm, respectively.

These results are submitted to a flagship journal in the field of power electronics [J5].

To sum up, we have achieved the four objectives listed in the proposal (see Section 1 of this report). In particular, we have developed a software module that trains and updates probabilistic models describing future renewable generation and EV arrivals with real-world data via deep neural networks. With the developed software module, we have developed scalable algorithms that can compute the optimal scheduling policies for 1000 EV chargers and an energy storage system with random renewable generation and electricity prices, for both cases with constant and adjustable EV charging power. Compared with the state-of-the-art (online optimization and reinforcement learning) approaches, the developed algorithmic approach has demonstrated significant improvement in both learning speed and system cost reduction.

Benchmarking the performance of the developed energy management system with actual hardware components in real-world power distribution system conditions, the developed HIL demonstrator will facilitate the technology transfer and follow-up funding applications for the potential commercialization of our research results.

4. PUBLICATION AND AWARDS

Please list out and number all the publications and/or awards produced under the funded project. All these publications must be directly acknowledged the SHIAE funding support and stated the affiliation with the Institute. The list can be numbered in alphabetic order. When referring to them for the submission in CD, name

the file with corresponding reference number in square brackets as “81150xx-J[1].pdf”.

J[1] L. Hao, Y. Xu*, and L. Tong, “Asymptotically optimal index policies for deadline scheduling with processing rate limits,” accepted to *IEEE Trans. on Automatic Control*, IEEE, 2021.
J[2] J. Jin and Y. Xu*, “Segregated linear decision rules for distributionally robust control with linear dynamics and quadratic cost,” *IEEE Systems Journal*, 15(1), pp. 355-364, IEEE, 2020.
J[3] J. Jin, L. Hao, Y. Xu*, J. Wu, and Q.-S. Jia, “Joint scheduling of deferrable demand and storage with random supply and processing rate limits”, 66(11), pp. 5506-5513, *IEEE Trans. on Automatic Control*, IEEE, 2021.
J[4] J. Jin, Y. Xu, and Z. Yang, “Optimal deadline scheduling for electric vehicle charging with energy storage and random supply,” *Automatica*, 119, pp. 109096, Elsevier, 2020.
J[5] Y. Zhao and Y. Xu, “Model Predictive Control based Cost-Minimizing Energy Management System for PV-Integrated Electric Vehicle Charging Stations,” submitted to *IEEE Trans. on Power Electronics*.
C[1] L. Hao and Y. Xu*, “Index Policies for Stochastic Deadline Scheduling with Time-varying Processing Rate Limits,” *American Control Conference*, Denver, CO, USA, July 1-3, 2020.

COMPLETED

DEVELOPMENT OF A NOVEL COOLING TOWER WITH FREE DAYTIME RADIATIVE COOLING FOR REDUCING ENERGY CONSUMPTION IN BUILDINGS

Principal Investigator: Professor Chun CHEN
Department of Mechanical & Automation Engineering, CUHK

Research Team Members:
Xinxian Yu, Ph.D. Student [(1)], Haiqiang Zhang, Ph.D. Student [(1)]

[(1)] Dept. of Mechanical and Automation Engineering, CUHK

Project Start Date: 1st July 2018
Completion Date: 30th June 2020

INNOVATION AND PRACTICAL SIGNIFICANCE:

Conventional cooling towers drag the outdoor air to cool the cooling water for rejecting heat to the atmosphere. Lower cooling water temperature results in a higher COP of chillers. Therefore, it is worthwhile to enhance the heat rejection in cooling towers without consuming additional energy. The innovation of this work is to develop a cooling tower with free daytime radiative cooling. The proposed cooling tower utilizes free and renewable cooling to lower the cooling water temperature. Consequently, the COP of chillers is expected to increase by 10 to 20%. A prototype will be fabricated and tested in this project. With the collaboration with the heating, ventilation, and air-conditioning (HVAC) industry, we will actively see further development of the prototype and potential technology transfer. If successful, the novel cooling towers can be potentially used in commercial and residential buildings to reduce the energy consumption and the associated carbon dioxide emissions.

ABSTRACT

The air-conditioning systems in buildings consume about 30% of the total electricity in Hong Kong. In a typical heating, ventilation, and air-conditioning (HVAC) system, the conventional cooling tower drags the outdoor air to cool the cooling water for rejecting heat to the atmosphere. To reduce the energy consumption, this project proposed to develop a novel cooling tower with renewable sky radiative cooling. A basin coated with a film of radiative cooling metamaterial, as a sky radiative cooler, will be implemented into the cooling tower.

First, this project developed a method to compare the cooling performance of different daytime radiative cooling materials. This simulation-based method consists of the basic radiation theory, the standard solar spectrum, and six standard model atmospheres, for comparing the cooling performance of different daytime radiative cooling materials. The proposed simulation-based comparison method was then used to compare four representative daytime radiative cooling materials selected from the literature. The results demonstrated the effectiveness of the proposed method in comparing the cooling performance of different radiative cooling materials tested under different environmental conditions. With the radiative cooling material, the cooling water temperature will decrease without consuming additional energy, so that the coefficient of performance (COP) of chillers will increase. The design of the proposed cooling tower will be supported by thermodynamic modeling. The proposed project will offer a novel cooling tower that can utilize renewable cooling and reduce the energy consumption in buildings.

1. OBJECTIVES AND SIGNIFICANCE

1.1. Objectives

The first objective is to propose a simulation-based method for comparing the cooling performance of different daytime radiative cooling materials. With the proposed comparison method, the engineers can choose the most suitable daytime radiative cooling material for a specific environment. Thus, the system can yield a better energy performance with the best radiative cooling material.

The second objective is to propose a novel cooling tower with renewable sky radiative cooling. The proposed cooling tower will utilize sky radiative cooling to reduce the condenser temperature, so that the COP of the chillers can be increased. Thus, the new system can be potentially applied in buildings to reduce the energy consumption.

The third objective is to develop a numerical model for predicting the cooling performance of radiative cooling materials under different climate. The development of this model can evaluate the radiative cooling materials' cooling performance under different locations, which would guide the future design of the cooling system in buildings.

1.2. Significance of this project

The increasing commercial potential of radiative cooling has led to much greater interest in selecting the optimal materials over others under specific environmental conditions. Therefore, it is worthwhile to compare the cooling performance of different daytime radiative cooling materials tested under different environmental conditions, in order to guide the suitable selection of these materials.

Conventional HVAC cooling towers drag the outdoor air to cool the cooling water for rejecting heat to the atmosphere. Lower cooling water temperature results in a higher COP of chillers. Therefore, it is worthwhile to enhance the heat rejection in cooling towers without consuming additional energy. The innovation of this work is to develop a cooling tower with renewable sky radiative cooling. The proposed cooling tower utilizes free and renewable sky radiative cooling to lower the cooling water temperature. Consequently, the COP of chillers is expected to increase by 5 to 10%. With the collaboration with the HVAC industry, we will actively see further development of the prototype and potential technology transfer. If successful, the novel cooling tower with sky radiative cooling can be potentially used in commercial and residential buildings to reduce the energy consumption and the associated carbon dioxide emissions.

2. RESEARCH METHODOLOGY

2.1 Proposed method for comparing the cooling performance of radiative cooling materials

Here, we present a simulation-based method, consisting of the basic radiation theory, the standard solar spectrum, and six standard model atmospheres, for comparing the cooling performance of different daytime radiative cooling materials. The proposed simulation-based comparison method was then used to compare four representative daytime radiative cooling materials selected from the literature. The details can be found in the published paper J[1]. The properties of the four materials are summarized in Table 1.

Table 1. Properties of the four daytime radiative cooling materials selected from the literature.

Reference	Materials and structural designs	Location	Daytime cooling power (W/m2)	Temperature reduction (oC)
Raman et al.	Multilayer structures	Stanford, CA, USA	40.1	4.9
Zhai et al.	Spheres distributed in a matrix material	Cave Creek, AZ, USA	93	N/A
Mandel et al.	Polymer coating	Phoenix, AZ, USA	96	6
		New York, NY, USA	83	5
		Chattogram, Bangladesh	26	3
Zhao et al.	Glass-polymer hybrid metamaterial	Boulder, CO, USA	45	10.6

2.2 Cooling tower with sky radiative cooling

This study proposed to implement a passive sky radiative cooler between the cooling tower and the condenser. Figure 1 shows a schematic of the system setup. This investigation focused on a counterflow cooling tower without the fan speed control. The high-temperature cooling water from the condenser entered into the cooling tower, and the spray water droplets had heat and mass transfer with the included air. The water leaved the cooling tower with a lowest possible temperature of the outdoor wet-bulb temperature. The cool temperature then entered into a radiative cooler, which consisted of a basin covered by a metamaterial film with radiative materials. Through radiative heat transfer to the outer space, the cooling water was further cooled before flowing back to the condenser. Through this process, the condenser temperature was decreased, so that the COP of chiller could be improved.

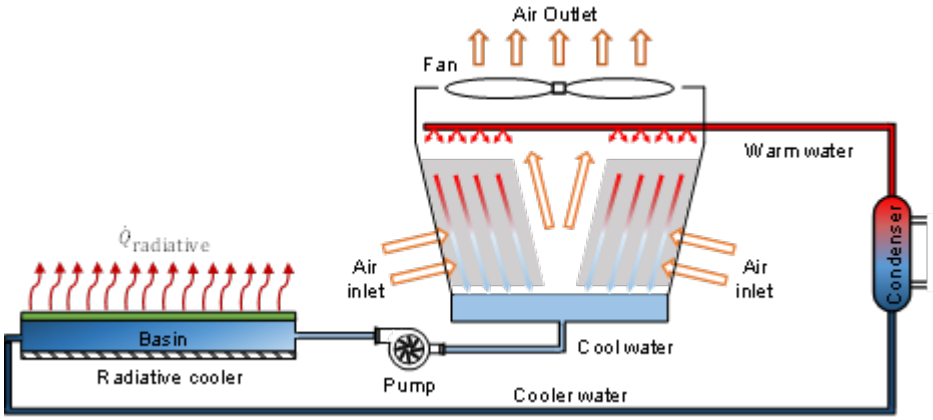


Figure 1. Schematic of the proposed cooling tower with renewable passive radiative cooling.

Figure 2 shows the schematic of the metamaterial film with radiative materials. The upper layer can reject heat through the infrared irradiance to the cold sink of outer space. The radiative material has strong emission between 8 and 13 μm, the atmospheric transmission window. The lower layer can reflect the solar irradiance so that the radiative cooler can still work in the daytime. With the use of the radiative cooler, the system may achieve a cooling water temperature lower than outdoor wet-bulb temperature. However, it is still unknown if the proposed cooling tower with passive sky radiative cooling could actually improve the COP of chiller. Therefore, this study will develop a numerical model for the analysis in the following section.

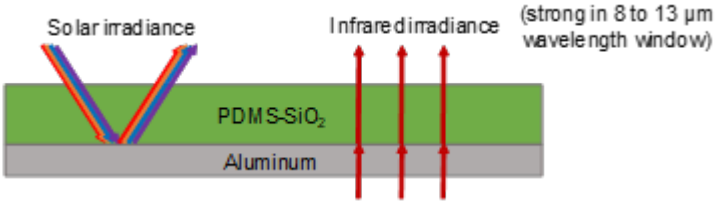


Figure 2. Schematic of the metamaterial film that can reflect the solar irradiance and reject heat through the infrared irradiance to the cold sink of outer space.

2.3 Mathematical model

Considering the steady-state condition, the radiative emission power of the cooler is equal to the net incident power absorbed by the cooler:

$$2\pi A_{rc,a}\int_0^{\pi/2}\sin\theta\cos\theta\int_0^{\infty} I_{BB}(T_{rc},\lambda)\varepsilon_{rc}(\lambda,\theta)d\lambda d\theta = A_{rc,a}\int_0^{\infty} I_{AM1.5}\varepsilon_{rc}(\lambda)d\lambda$$
$$+2\pi A_{rc,a}\int_0^{\pi/2}\sin\theta\cos\theta\int_0^{\infty} I_{BB}(T_{atm},\lambda)\varepsilon_{rc}(\lambda,\theta)\varepsilon_{atm}(\lambda,\theta)d\lambda d\theta + A_{rc,a}h_{rc,a}(T_{atm}-T_{rc}) + m_w c_{p,w}(T_{w2}-T_{w1}) \quad (1)$$

where A_{rc} (m^2) is the cooling surface area exposed to the sky, θ (sr) is the radiation angle, I_{BB} ($W/m^2 \cdot sr \cdot m$) is the spectral radiance of a blackbody, T_{rc} (K) is the temperature of the radiative cooler, ε_{rc} (unitless) is the

emissivity of the radiative cooling material, $I_{AIM1.5}$ (W/m²·m) is the solar illumination, T_{atm} (K) is the temperature of the ambient air, $\delta_{w,atm}$ (atm-cm) is the absolute vertical water vapor column in the atmosphere, ε_{atm} (unitless) is the emissivity of the atmosphere, $h_{rc,a}$ (W/m²·K) is the convective heat transfer coefficient between the cooler and the surrounding air, T_{w1} (K) is the temperature of the water entering the radiative cooler (or leaving the cooling tower), and T_{w2} (K) is the temperature of the water leaving the radiative cooler (or entering the condenser). The spectral radiance of a blackbody can be calculated by

$$I_{BB}(T,\lambda)=\frac{2hc^2}{\lambda^5}\frac{1}{\exp(\frac{hc}{\lambda k_B T})-1} \tag{2}$$

where h is the Planck's constant (6.626×10^{-34} J·s), c is the speed of light (2.998×10^{8} m/s), λ (m) is the wavelength, and k_B is the Boltzmann constant (1.381×10^{-23} m²·kg/s²·K). The emissivity of the atmosphere, which can be calculated based on the "box model" proposed by Granqvist and Hjortsberg:

$$\varepsilon_{atm}(\lambda,\theta)=\begin{cases}1 & \lambda<8\ \mu m\\ 1-t(\lambda,\delta_{w,atm})^{1/\cos\theta} & 8\ \mu m\le\lambda\le13\ \mu m\\ 1 & \lambda>13\ \mu m\end{cases} \tag{3}$$

where t (unitless) is the atmospheric transmittance, which can be calculated using the software MODTRAN.

This study focused on counterflow cooling tower with a constant fan speed. Based on the cooling tower theory proposed by Merkel, the heat and mass transfer process occurs between water, interfacial film, and air. The interfacial film was assumed to be of saturated air. Under steady-state condition, the water heat loss is equal to the air heat gain:

$$m_w c_{p,w} dT_w = m_a dh_a = K\alpha(h_{film}-h_a)dV \Rightarrow \frac{K\alpha V}{m_w}=\int_{T_{w1}}^{T_{w3}}\frac{c_{p,w}}{h_{film}-h_a}dT_w \tag{4}$$

where m_w (kg/s) is the inlet water mass flow rate, $c_{p,w}$ (kJ/kg·K) is the specific heat of water, T_w (K) is the water temperature, m_a (kg/s) is the air mass flow rate, h_a (kJ/kg) the enthalpy of bulk air, K (kg/s·m²) is the unit conductance of mass transfer from the water-air interface to main airstream, α (m²/m³) is the ratio of water-air interface area to the cooling tower volume, h_{film} (kJ/kg) is the enthalpy of saturated air at the bulk water temperature, V (m³) is the cooling volume of the tower, and T_{w3} (K) is the temperature of water leaving the cooling tower (or entering the radiative cooler).

The COP of the chiller can be calculated by:

$$COP=\frac{Q_{evap}}{m_w c_{p,w}(T_{w,3}-T_{w,2})-Q_{evap}}=\frac{T_{evap}}{\beta_1\cdot T_{w,3}+\beta_2-T_{evap}}\eta \tag{5}$$

where Q_{evap} (W) is the heat absorbed by the evaporator, T_{evap} (K) is the average temperature of evaporator, T_{cond} (K) is the average temperature in condenser, β_1 and β_2 are constants, and η (unitless) is the internal efficiency of the chiller. Solving the equations above, the temperatures, including T_{rc}, T_{w1}, T_{w2}, T_{w3}, and T_{cond}, can be obtained.

3. RESULTS ACHIEVED

Figure 3 depicts the calculated net cooling power, which takes into consideration both the absorbed solar energy and the emissive power. Note that the daytime radiative cooling materials can be compared

meaningfully only when the environmental conditions are fixed. This example demonstrates that the direct comparison of the measured cooling powers reported in the literature is not reasonable, while the proposed simulation-based method can be used to fairly compare the cooling performance of different daytime radiative cooling materials.

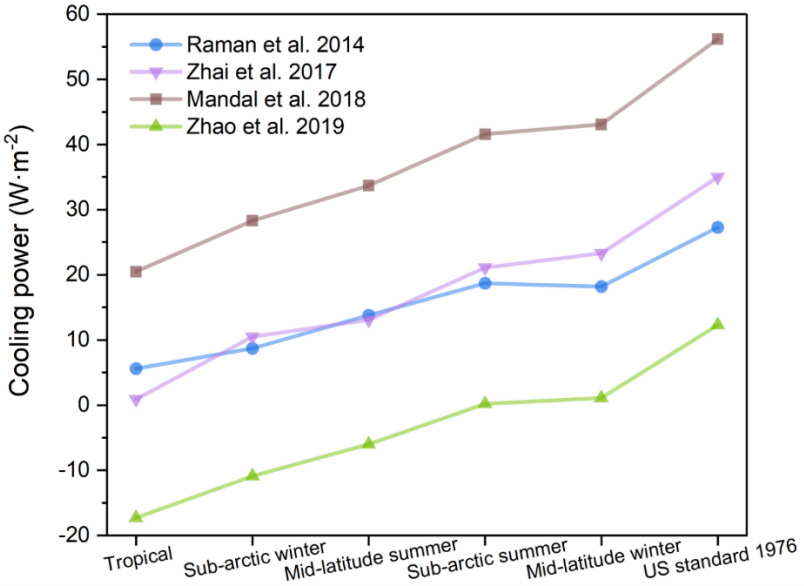


Figure 3. Calculated cooling power of the four daytime radiative cooling materials using the six standard model atmospheres.

As shown in Figure 4, when the ambient temperature was higher than 20 °C and the relative humidity was lower than 20%, the potential cooling power was greater than 110 W/m^2. Namely, a hot, dry climate was preferable for maximizing the potential cooling power. Meanwhile, when the ambient temperature was above 20 °C and the relative humidity was higher than 80%, the potential cooling power was lower than 68 W/m^2. That is to say, a hot, humid climate significantly limited the maximum possible cooling power of a daytime radiative cooler. Furthermore, when the ambient temperature was lower than 10 °C, the potential cooling power ranged from 68 to 112 W/m^2; thus, it was relatively insensitive to the relative humidity.

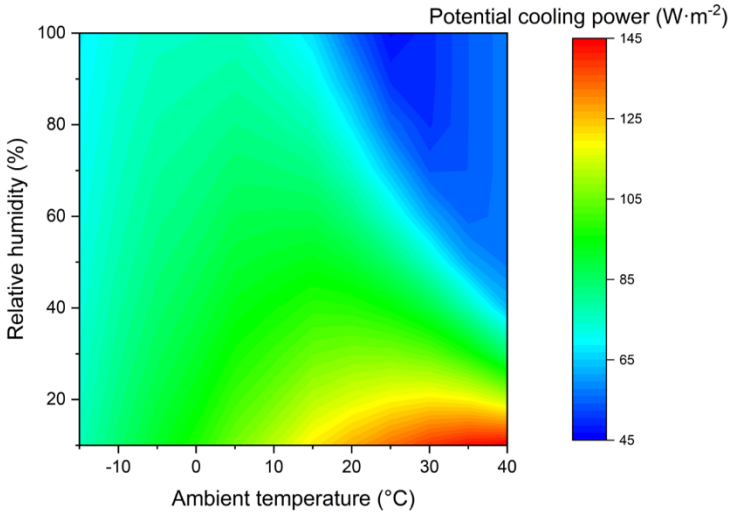


Figure 4. Influence of ambient air temperature and relative humidity on the potential cooling power for ideal broadband radiative coolers.

With the best material identified above, this study used the developed numerical model to predict the improvement in COP of chiller and reduction in chiller electricity use for several buildings. The outdoor air temperature and relative humidity were set the monthly average values in Hong Kong for a whole year. The cooling load per unit floor area was set at 60 W/m^2. The rooftop area available for the sky radiative cooler was

assumed to be 80%. The cooling tower volume and water flow rate were set at the average value from the data collected from the literature. The K·α was set at 0.445 according to the U.S. Department of Energy (DOE) report. The radiative cooler material was selected to be of high emissivity in the wavelength range of 8 to 13 μm, while very low emissivity in other wavelength ranges. Figure 5 plots the percentage of reduction in chiller electricity use and increase in chiller COP for the buildings with different numbers of floor. When the number of floor increased, the corresponding cooling load increased, so that the cooling water flow rate tended to increase. In that case, with the same sky radiative cooling area, the reduction in cooling water temperature decreased. Thus, the cooling tower with sky radiative cooling performed better for lower floor buildings.

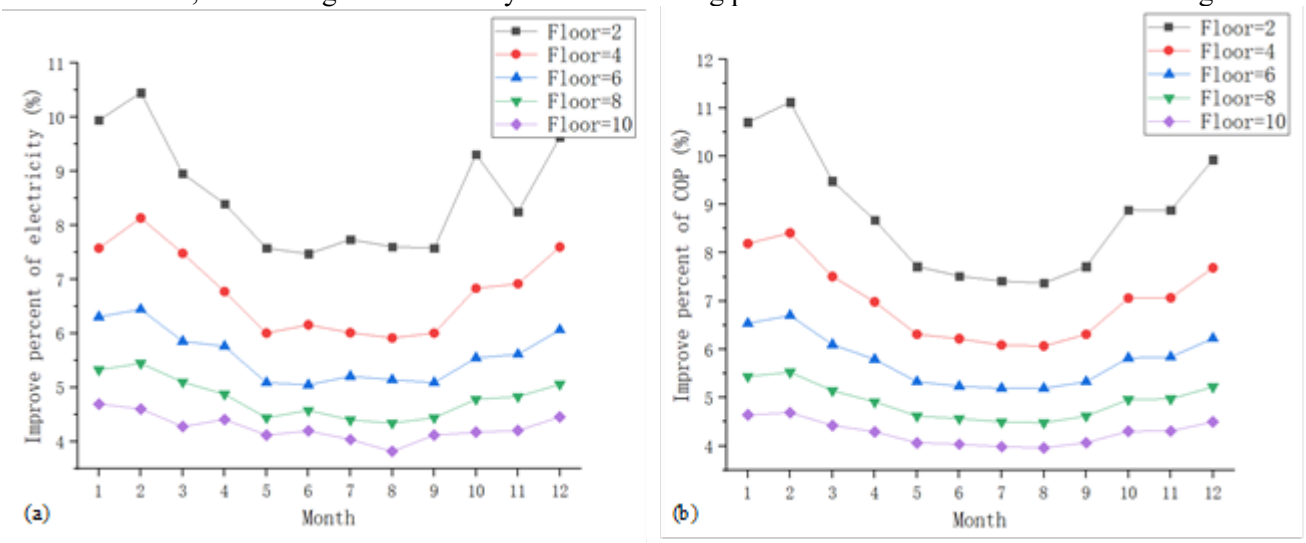


Figure 5. (a) Percentage of reduction in chiller electricity use, and (b) percentage of increase in chiller COP.

To better understand the free cooling potential from using the radiative coolers, this study calculated the potential cooling powers for different cities in China. Figure 6 presents distribution maps of potential cooling power for ideal broadband radiative coolers in January and July. A comparison shows that the potential cooling power was greater in winter than in summer. Furthermore, the national distribution was more uniform in winter than in summer. In summer, the potential cooling performance in the southeast was worse than that in other regions. The distribution map can be a quick reference for engineers or designers in determining whether radiative cooling has good potential to save energy in particular regions.

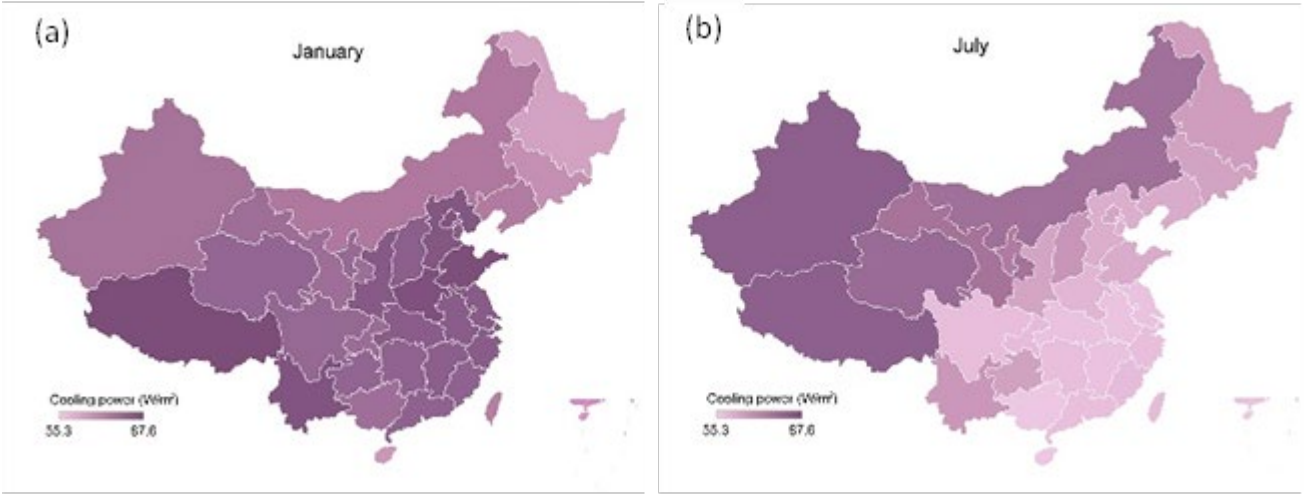


Figure 6. Potential cooling power for ideal broadband radiative coolers in (a) January and (b) July.

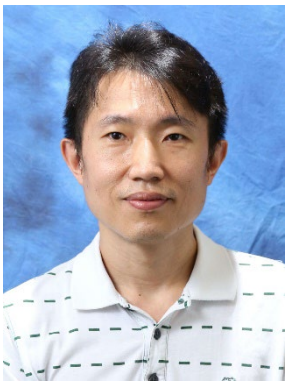
4. PUBLICATION AND AWARDS

J[1] X. Yu and C. Chen, “A simulation study for comparing the cooling performance of different daytime radiative cooling materials”, Solar Energy Materials and Solar Cells, Elsevier, Netherlands, 110459, 1 June 2020.
C[1] X. Yu and C. Chen, “A simulation study for comparing the cooling performance and cooling potential of daytime radiative cooling”, Proceedings of the 16th Conference of the International Society of Indoor Air Quality & Climate (Indoor Air 2020), South Korea, Paper ABS-0804, 1-4 November 2020.

COMPLETED

MEGAHERTZ CURRENT SENSOR FOR MEGAHERTZ RENEWABLE ENERGY CONVERTER

Principal Investigator: Professor Loh Poh Chiang Andrew
Department of Electronic Engineering, CUHK



Research Team Members:
XIN Zhen, Professor [2]
DING Wenlong, Dr. [3]
MING Lei, Mr. [1]

[1] Dept. of Electronic Engineering, CUHK
[2] Department of Electrical Engineering, Hebei University of Technology
[3] Hong Kong Applied Science and Technology Research Institute

Project Start Date: 1st July 2018
Completion Date: 31st August 2020

INNOVATION AND PRACTICAL SIGNIFICANCE:

The developed current sensor will have a frequency bandwidth up to 10 megahertz. It is thus around 40 times wider than that of a commonly used Hall current sensor. Such wide bandwidth permits the realization of fast control and reliable protection in the newest generation of power converters built using WBG devices. The resulting converters can be sized for renewable energy generation or miniaturized for routinely used laptop adapter, EV charger and phone charger to name only a few. There will be a revolutionary reduction of size and weight, enabled by the high-frequency operation of WBG devices. In addition, the developed sensor is of great commercial competitiveness due to its small size and low cost, made possible by its integration into an inexpensive printed circuit board. This merit will promote the coming era of WBG devices with extremely high efficiency, but has its commercialization hindered many years due to its high price. Moreover, the substantial market volume and the all-important cost factor have led to the development of a wide range of alternative current sensing methods. The market demand for isolated current sensors will, in fact, hit 100 million in 2020. This is roughly equivalent to the amount of iPhones produced in 2012 [1]. Therefore, the developed current sensor is revolutionary, yet with great potential and competitiveness to be converted into a commercial product.

ABSTRACT

Power converter is an essential interface for tying a renewable source (e.g. solar and wind) to the power system. It usually consists of three main components, known respectively as the microprocessor, power semiconductor devices and sensors. The microprocessor can analogously be viewed as the “brain” of the power converter, while the sensors and power devices can be viewed as its “sense organs” and “limbs”. Recently, technological progress has pushed operational speed of its “brain” to the gigahertz range, while development of wide bandgap (WBG) devices has allowed its “limbs” to reach the megahertz range with high efficiency and power density (power per unit volume). Undoubtedly, these are significant achievements motivated by the booming renewable energy industry, which next will propel rapid development of the “sense organs”. “Sense organs”, mostly current sensors, are presently limited below hundreds of kilohertz. The megahertz responses of the “limbs” are thus not sensed properly, which in practice, can cause control, protection and monitoring of the power converter to deteriorate. These are nontrivial issues, judging from the amount of power generated by a typical renewable source. There is thus an intense need to develop an advanced current sensing technique that can sense in the megahertz range with a high noise immunity and a low manufacturing cost. Presently, no existing commercial sensors can meet the targets, which is why it will be the next trend of rapid development.

1. OBJECTIVES AND SIGNIFICANCE

Improving renewable energy generation has constantly demanded better performances from the power converters, especially their efficiencies and power densities [2]. This, to some extent, contributes to the invention of WBG devices as the next-generation promising building blocks for the power converters. Building a full WBG power converter is however not straightforward with many barriers yet to be resolved [3]. Some barriers are typical like costs and reliabilities, while others are less obvious but technically important to both manufacturers and users. One prominent one from the latter is the requirement for high-performance current sensors [4], which in this project, is met by systematically finding answers to fulfill the following objectives. The ultimate challenging sensor developed is compact, ultrafast, nonintrusive, low cost and highly suitable for controlling and protecting next-generation power converters.

1.1. Ultrafast current sensor

The developed current sensor must have a frequency bandwidth up to 10 MHz to meet requirements imposed by fast protection of next-generation power converters. These converters with their WBG devices can switch at extremely high frequencies up to several megahertz, in order to realize high efficiencies and high power densities demanded by modern renewable energy generation. They however face tougher short-circuit protection, since fast switching currents through WBG devices are significantly more challenging to measure. They are thus not satisfactorily protected by widely available Hall current sensors, whose highest bandwidth is only up to 250 kHz. Reliable protection of WBG power converters is therefore lacking at present, which certainly is haphazard during unforeseen abnormal operating conditions. It is thus essential to develop a new current sensor that can measure up to 10 MHz, in order to better convince manufacturers to adopt WBG devices for building power converters.

1.2. Nonintrusive current sensor

The developed current sensor must have a very small insertion inductance, usually smaller than 1nH. This is essential, since high-speed WBG power devices are highly sensitive to parasitic inductances, which if significant, can degrade performance of a WBG device towards that of a conventional silicon device. That then causes an optimally designed converter to become unstable or malfunction unintentionally. Low insertion parasitic components, especially inductance in the nH range, is therefore a strict requirement to be met by current sensors designed for next-generation power converters.

1.3. Compact current sensor

The developed current sensor must be extremely compact, so as not to compromise desired high power density of a power converter. High power density, in general, is desired, since it equates to better utilization of energy resources at a low operating cost. It is even more essential given that operating cost has roughly doubled every 10 years since 1970. It is thus not an option to use existing bulky current sensors. To better visualize that, the recent winning photovoltaic inverter from the Google Little Box Challenge has achieved its 2-kW power rating with only a volume of about two iPhone 5s. Its employed industrial current sensor (typical among PV inverters) is however as large as an Apple Watch, which unquestionably, is not acceptable. In other words, research focus must now shift from classical power semiconductors, topologies and control to the relatively new area of designing compact current sensors, in order to preserve the past trend of doubling power density per decade.

1.4. Low-cost current sensor

The developed current sensor must be inexpensive when manufactured in large volume, like all new product developments, including transistors and the entire topological design. Such cost saving however cannot be met by a typical Hall current sensor, which despite its mature technology, is still expensive. A Hall sensor will hence undesirably raise price of a future WBG power converter. Such rise will however not occur with the developed new current sensor, whose structure can conveniently be embedded in a print-circuit board. In other

words, new opportunities for lowering cost open up, which certainly is the reason for exploring new current sensor in this proposed research.

2. RESEARCH METHODOLOGY

The project consists of a number of tasks, investigated by a postdoctoral fellow and a postgraduate Ph.D. student. Their core task is to enormously reduce the volume of a conventional Rogowski current sensor for implementation as a new printed-circuit-board (PCB) Rogowski current sensor. While doing so, important characteristics of the sensor, like its wide-bandwidth and ability to work in a hostile transient and interfering environment, have either been retained or improved. Such retentions or improvements are possible because of strict execution of a well-organized research plan and methodology to be described next.

2.1. Structural layout of sensing coil as sensor head

Performance of a Rogowski current sensor is closely related to its number of coil turns. On one hand, high number of turns is desirable, since it increases accuracy of measurement with its large mutual inductance. On the other hand, with more turns, its self-inductance increases faster than its mutual inductance. The outcome is a drop of resonance frequency and hence bandwidth of the coil. Such compromise becomes more impactful with a tiny Rogowski current sensor occupying a small volume. It is thus meaningful to find the best tradeoff between dynamics and accuracy of the sensing coil for implementing the developed sensor head.

2.2. Development of techniques for nullifying effects from interfering electrostatic fields

Power converters, such as micro photovoltaic inverters, are shrinking in size, but retaining their high power ratings. The result is an increase of field strength within each converter, which to the internally installed current sensors, equates to a hostile operating environment. Such hostility becomes more intense upon incorporating much faster switching speeds of all WBG devices, as compared to conventional power devices. It is thus essential for the developed PCB Rogowski current sensor to have an excellent common-mode immunity, so as to reject all interfering fields that can deteriorate accurate extraction of current information in the power converter. One common solution is to fit a shield around the measuring coil, but that will substantially reduce bandwidth of the overall sensor. It is thus not an appropriate method for usage within a modern fast-switching power converter. Instead, alternative cost-effective solutions, capable of rejecting interfering fields, are proposed in the project.

2.3. Investigation of parametric design rules

Accurate design rules for tuning parameters of the developed current sensor are important and are usually applied before commercialization and mass production. Their purpose is to optimize effects from physical and other design parameters on the measuring performances of the sensor. That purpose additionally requires an equivalent model of the sensing coil, which has hence been derived to complete the set of design tools for the proposed sensor.

2.4. Incorporation of dc current measurement to developed Rogowski current sensor

One disadvantage of the Rogowski coil is its inability to detect a DC current. It is thus not recommended, if DC current information is essential for both control and protection. Instead, a recently proposed hybrid sensor, combining Rogowski and anisotropic magneto-resistive (AMR) technologies, can be considered. The hybrid sensor utilizes an AMR element for sensing DC and low-frequency components of the current and a Rogowski coil for measuring its high-frequency components. The sensor however suffers from high cost and significant design difficulties, which together, render it unsuitable for a highly dense power converter. Some peers have since opted for an alternative approach, whose DC measuring principle is to compute averages of switching current pulses. This naturally suits a power converter, whose high-frequency switching creates lots of narrow current pulses for sensing by a Rogowski coil. The sensed pulses can then be averaged by a simple averaging circuitry. The approach is thus inexpensive and hence further investigated in the project.

2.5. Experimental test-bed and performance evaluation

Performances of the developed current sensor, in terms of its rejection of external interfering fields, DC current measurement and other characteristics, must be tested, which for the project, are done with a double-pulse test-bed in the laboratory. Multiple of the proven sensors are then fabricated for installation within a highly dense power converter, such as the highly compact 18-switches matrix converter. The target is to test overall coordinated short-circuit protection and current control of the converter. Complementing it, a second similar matrix converter has also been built, but installed with multiple existing Hall (or other types of) sensors. Results obtained from both converters can then be normalized to offer a fair comparison.

3. RESULTS ACHIEVED

The developed current sensor has a proven wide frequency bandwidth of up to 10 megahertz. It is thus almost 40 times wider than that of an existing commonly found Hall current sensor. The proven wide bandwidth then permits newest generation of power converters, built with WBG devices, to be reliably protected and accurately controlled. These technical expectations have been proven in experiments, whose details can be found from peer-reviewed papers published and listed below. The proven merits will then promote a rapid era of WBG usage, which although is highly efficient, has presently been hindered by high (but rapidly falling) prices, complicated protection and control issues.

Physical wise, the developed sensor is commercially revolutionary due to its small size, lightness and ready implementation on an inexpensive printed circuit board. A photograph showing compactness of the sensor has been provided together with this report. The sensor is thus competitively attractive with potential to capture a huge market share. Such market share is eye-catching, since in 2020, the total market volume of isolated current sensors is about 100 million or, in context, roughly equivalent to the amount of iPhones produced in 2012 [1]. The commercial potential of the developed sensor is thus sizable, in addition to its technical enabling of WBG power converters for populated renewable energy generation (and / or miniaturized electronic adapters and EV chargers to name only a few).

[1] J. Krapp, “No Need for External Sensors in New Inverters,” Semikron Report, May. 2016.

4. PUBLICATION AND AWARDS

J[1] L. Ming, Z. Xin, W. Liu, and P.C. Loh, “Structure and Modelling of Four-Layer Screen-Returned PCB Rogowski Coil with Very Few Turns for High-Bandwidth SiC Current Measurement,” *IET Power Electronics*, IET, UK, pp. 765-775, 2020.

J[2] L. Ming, W. Ding, C. Yin, Z. Xin, and P.C. Loh, “A Direct Carried-Based PWM Scheme with Reduced Switching Harmonics and Common-Mode Voltage for Current Source Converter,” *IEEE Transactions on Power Electronics*, IEEE, USA, early access, 2020.

C[1] L. Ming, Z. Xin, C. Yin, P.C. Loh, and Y. Liu, “Screen-Returned PCB Rogowski Coil for the Switch Current Measurement of SiC Devices,” *IEEE Applied Power Electronics Conference and Exposition*, IEEE, Anaheim, CA, USA, pp. 958-964, 17-21 March 2019.

C[2] L. Ming, Z. Xin, C. Yin, M. Chen, and P.C. Loh, “Integrator Design of the Rogowski Current Sensor for Detecting Fast Switch Current of SiC Devices,” *IEEE Energy Conversion Congress and Exposition*, IEEE, Baltimore, MD, USA, pp. 4551-4557, 29 September-3 October 2019.

C[3] L. Ming, W. Ding, P.C. Loh, and Z. Xin, “An FPGA-Based Fault-Tolerant Method for Reliable Current Commutation of Direct Matrix Converter,” *IEEE International Power Electronics and Motion Control Conference*, IEEE, Nanjing, Jiangsu, China, 29 November-2 December 2020 (received ***Best Paper Award***).

COMPLETED

DEVELOPMENT OF HIGH-PERFORMANCE TRIBOELECTRIC NANOGENERATORS FOR RENEWABLE BLUE ENERGY HARVESTING

Principal Investigator: Professor Yunlong ZI
Department of Mechanical & Automation Engineering, CUHK

Research Team Members:
Ms. Jingjing Fu[(1)], Dr. Yubiao Sun[(1)], Dr. Xiaoyi Li[(1)],
Dr. Jiaqi Wang[(1)], Mr. Haoyu Wang[(1)]

[(1)] Dept. of Mechanical and Automation Engineering, CUHK

Project Start Date: 1st July 2018
Completion Date: 30th June 2020

INNOVATION AND PRACTICAL SIGNIFICANCE:

Innovation: *The innovation of the proposed research lies in the development of high-performance TENGs for blue energy harvesting. Traditional water energy harvesting uses electromagnetic generators (EMG), which are usually huge, heavy, expensive, and technically difficult for construction in deep water. Further, studies have demonstrated that EMG are extremely inefficient in harvesting the low-frequency mechanical energy that is generated by ocean waves. TENG possess several advantages over EMG technology as they are lightweight and able to float, cost-effective and easy to produce, and efficient in harvesting low-frequency energy. Our research will focus on maximizing TENG output through addressing two crucial fundamental issues: the air-breakdown effect inside TENG and the parasitic capacitances brought by seawater, which have never been systematically studied before. For the first time, we propose to develop TENGs with novel designs to address these issues. It is expected that the output performance of TENG could increase 10-100 times through our proposed research.*

Practical significance: *Considering the challenges of energy security and environmental protection, developing renewable energy sources is of critical importance for Hong Kong. With Hong Kong's extensive coastline and water area percentage of 59.8%, Hong Kong is strategically placed to take advantage of a convenient, clean and renewable power source. Especially for areas with complex coastlines that are not suitable for water shipping, electrical generation is the best option to utilize them. Additionally, the ocean currents and tropical storms that are common in Hong Kong, provide an abundant amount of mechanical energy that could potentially be converted to electrical energy. Lastly, developing the blue energy harvesting technology to replace the fossil fuels will also decrease the emissions of pollutants and greenhouse gases, which is critical for environmental protection. Therefore, developing blue energy harvesting through TENG is beneficial to Honk Kong on many levels including the mediating the energy crisis, promoting environmental protection, and advancing both economic and social development in Hong Kong.*

ABSTRACT

Electricity is the world's fastest growing form of end-use energy consumption. Between 2015 and 2040, world net electricity generation will increase by 45%. Non-renewable fossil fuels still account for >60% of electricity generation. However, 70% of the earth's surface is covered by ocean, which represents a huge untapped clean and renewable energy source. Estimated to provide power of over 75 TW, ocean energy could satisfy energy demands around the world. To effectively harvest this "blue energy" especially the low-frequency mechanical energy generated by ocean waves, three-dimensional networks of triboelectric nanogenerators (TENG) have been proposed. To test this concept, however, TENG units first need to be refined to optimize their output performance. To date, factors that limit the performance include the challenge that achievable charge density is limited by the phenomenon of air breakdown; additionally, the parasitic capacitances brought by the

conductive seawater may suppress the performance of TENG. We propose experimental and theoretical studies that will focus on mitigating these limiting factors by developing novel structural and material designs, greatly enhancing the output performance of TENG. The proposed research will lay the cornerstone for further technologic advancement in large-scale harvesting of kinetic water energy using TENG units.

1. OBJECTIVES AND SIGNIFICANCE

1.1 Objectives:

1. To simulate and experimentally demonstrate a TENG design that will minimize the air-breakdown effect using controlled high pressure and inert gas environments; (Completed)
2. To simulate and experimentally develop optimized structural/material designs to minimize the influences of the parasitic capacitances brought by the conductive seawater; (Completed)
3. To determine the optimized designs and produce a TENG that provides maximal available output, and to compare with other technologies used in blue energy harvesting. (Completed)

1.2 Significance: Considering the challenges of energy security and environmental protection, developing renewable energy sources is of critical importance for Hong Kong. With Hong Kong's extensive coastline and water area percentage of 59.8%, Hong Kong is strategically placed to take advantage of a convenient, clean and renewable power source. Especially for areas with complex coastlines that are not suitable for water shipping, electrical generation is the best option to utilize them. Additionally, the ocean currents and tropical storms that are common in Hong Kong, provide an abundant amount of mechanical energy that could potentially be converted to electrical energy. Lastly, developing the blue energy harvesting technology to replace the fossil fuels will also decrease the emissions of pollutants and greenhouse gases, which is critical for environmental protection. Therefore, developing blue energy harvesting through TENG is beneficial to Honk Kong on many levels including the mediating the energy crisis, promoting environmental protection, and advancing both economic and social development in Hong Kong.

2. RESEARCH METHODOLOGY

2.1 Theoretical modeling and simulation of water-wave TENG: The basic model of the water-wave TENG is built and simulated by COSMOL Multiphysics, a cross-platform finite element analysis simulation software. It can provide the potential and electric field distribution which can be used to determine the breakdown status and the voltage output. It can also simulate the balanced charge status so that the available charge output can be given.

2.2 Experimental studies on the breakdown and parasitic-capacitance effects: In the measurement circuit as shown in **Fig. 2a**, the charging TENG is used to provide the high voltage required to approach the breakdown conditions in the test TENG. The charge transfer was measured by the electrometer with mark of "Q", and the voltage was calculated by the measured current through the electrometer with mark of "A" multiplying the shunt resistance R. The voltage at the turning points is recorded as the breakdown voltage, and then it was mapped into the V-Q diagram of TENG to outline the breakdown region. A high-speed camera is used to record the sparks during breakdown. In the non-breakdown part of the V-Q plot, the capacitance can be estimated by the slope.

2.2 Fabrication and electrical measurement of water-wave TENG output: A hollow plastic ball was used as the shell. Two copper films were attached to the inner side of the shell. A silicone rubber solid ball was put into the plastic ball to act as the moving part. Water-wave energy was simulated by a linear motor to supply a periodical external force with specified frequency. A pressure sensor was simultaneously placed under the TENG to calibrate the applied force. And the electrical properties of TENGs including output voltage, power, reliability and consistency were investigated by using a digital oscilloscope with a 100 MΩ probe, or a Keithley 6514 electrometer.

2.3 Fabrication of the optimized water-wave TENG design: To fill in the high-pressure CO_2, we designed and fabricated a ball-shell TENG with a screw cap, in which the high-pressure CO_2 gas can be filled and packaged inside (**Fig. 3a**). Silicone sealing layer is applied after tightening the cap to maintain the high-pressure environment inside. To reduce the parasitic capacitance from the water, we fabricated an additional layer made from packaging foams, with thickness of ~0.5 cm.

2.4 Development of the pressure vessel for quantitative analysis of the environmental factors: A chemical reaction chamber (from Gongyi Yuhua) was modified as a pressure vessel for testing TENG in controlled environments. A low-speed bidirectional motor was installed to provide the driven motion for TENG, and a framework was built inside to support the TENG. The electrodes were connected to the electrometers outside of the vessel through copper wires.

3. RESULTS

3.1 The modeling and simulations of the water-wave TENG

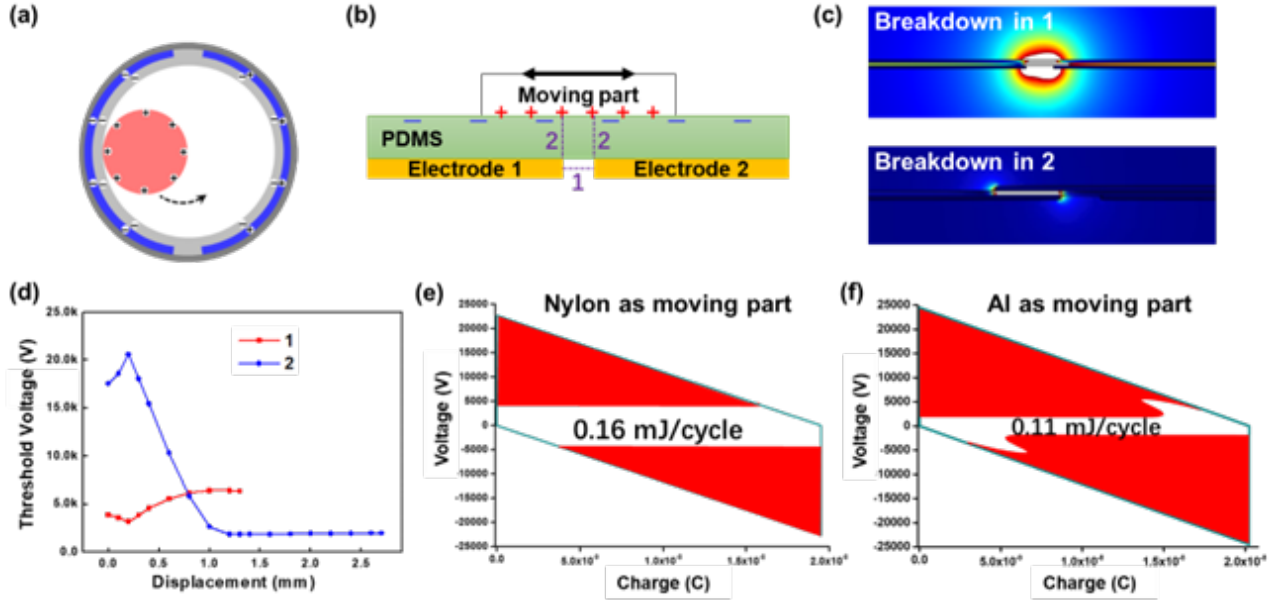


Figure 1: The modeling and simulation of water-wave TENG.

The original ball-shell structured TENG to harvest wave energy used in this project is schematically shown in **Fig. 1a**. Due to the triboelectrification, the electrons will be transferred from the electrodes to the ball. With the oscillations provided by the water waves, the inner moving ball can rotate from one electrode to the other one, which can electrostatically induce the potential difference between two electrodes, as the driven force of the current in the external circuits. To understand the breakdown effect in this water-wave TENG, the sliding-freestanding (SF) mode can be used to model this TENG. The schematic diagram of the SF-mode TENG is shown as **Fig. 1b**. There are two types of potential breakdown simulated, including horizontal air breakdown (type 1), and vertical dielectric breakdown (type 2). The model and simulations on the parasitic-capacitance effect have already been demonstrated and shown in the proposal previously.

The device breakdown criterion is set as below: if the minimum electric field along any line between electrodes or triboelectric surfaces is larger than the air-breakdown threshold electric field (3×10^6 V/m), then the breakdown happens, and there will be electron transfer between electrodes or triboelectric surfaces which can greatly impact the performance. The simulation of the electric field distribution is conducted in COMSOL Multiphysics software package, as shown in **Fig. 1c**. The simulated threshold voltages corresponding to different types of breakdown are plotted versus the displacement, as shown in **Fig. 1d**. And then we plotted the V-Q diagrams showing the breakdown regions in red, as shown in **Fig. 1e-f**. We notice that even though Al-PDMS friction can provide a little higher charge density, the TENG with insulating nylon as the moving layer can allow much more effective energy output (0.16 mJ/cycle) than that with conductive Al as the moving layer (0.11 mJ/cycle), due to the suppressed type-2 breakdown. Therefore, to avoid type-2 breakdown, we can choose insulating materials for the moving ball.

3.2 Experimental studies on the breakdown and parasitic-capacitance effects

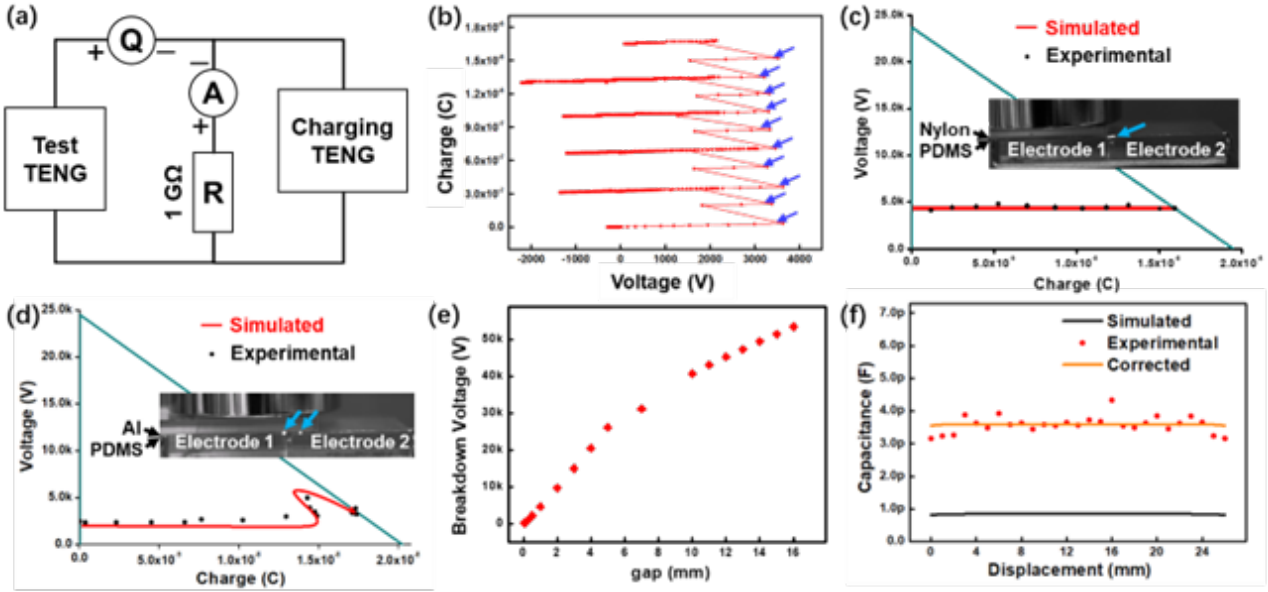


Figure 2: Experimental studies on the breakdown and parasitic-capacitance effects.

To directly observe the air-breakdown effect, we developed a measurement circuit to evaluate the breakdown effect, as shown in **Fig. 2a**. The hypothesis is, if there is breakdown happening, the measured V-Q plot will show clear turning points. The typically measured Q-V plots are shown in **Fig. 2b**, with the arrows indicating the breakdown points. As plotted in the V-Q diagram of the TENG, the experimental measured breakdown points are very consistent with the simulated results (**Fig. 2c-d**). The breakdown sparks are clearly observed through a high-speed camera, as indicated by arrows in the **insets of Fig. 2c-d**. We also experimentally demonstrated that the type-1 breakdown voltage can increase with the gap distance between electrodes (**Fig. 2e**), as a method to decrease the breakdown effect. This effect can be suppressed further within high-pressure inert environment.

We also revealed the parasitic-capacitance effect. Due to this effect, the measured capacitance values are much larger than the simulated results, as shown in **Fig. 2f**. After several repeated test, we can consider the average value of the difference between the simulated and the experimental capacitances as the parasitic capacitance. And then the corrected theoretical capacitance is calculated and plotted as the orange line, which is consistent with experimental results.

3.3 Optimized design of the water-wave TENG

In order to suppress the breakdown effect through the high pressure and inert gas environment, we modified the TENG structure design, as shown in **Fig. 3a inset**. To compare the output electric performances under different environment, the 1 atm air and 1.5 atm CO_2 were filled inside, respectively, and the results are shown in **Fig. 3a-b**. From these plots, we can conclude that the 1.5 atm CO_2 can greatly enhance the open-circuit voltage and short-circuit charge by both around 1.6 times, by suppressing the air-breakdown effect.

When the TENG is operated in water, the tested voltage and charge outputs are reduced to ~80 V and ~22 nC, respectively, as shown in **Fig. 3c-d**, due to the parasitic-capacitance effect from the water. To reduce that, COMSOL simulations were conducted and confirmed that thickness increase of ~1 cm in water TENG may achieve the optimized output energy density. Hence, we fabricated a device with relative thick shell ~ 1 cm, as shown in **Fig. 3c inset**. With that, the voltage and charge outputs are recovered to be ~190 V and ~40 nC, respectively (**Fig. 3c-d**). Therefore, compared to that without the ~1 cm thick shell, the energy output increases by ~4.3 times. The overall capacitance is reduced from ~0.28 nF to be ~0.21 nF. These results are well consistent with our theoretical simulations. In the meanwhile, the total volume of the TENG with the thick shell increased from 1.13×10^{-4} m^3 to 2.68×10^{-4} m^3, which is ~2.37 times increase. Consequently, the increase of the output energy density is only about 1.8 times, which is quite limited in enhancing the output performance.

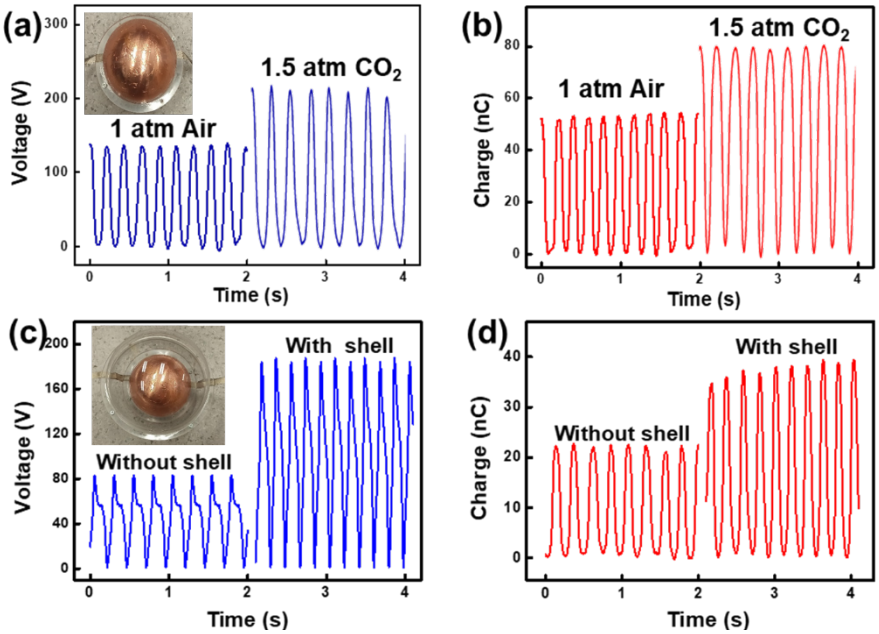


Figure 3: Optimized design of water-wave TENG and the high-pressure vessel.

3.4 Quantitative analysis of impacts from the pressure in high-pressure vessels

Even though we have demonstrated the optimized design, we still cannot evaluate and optimize the performance quantitatively, since it is impossible to control and maintain the internal environments very precisely. To address this issue and to test the device performance under even higher pressures, we have a modified steel-made chemical reactor to provide the desired environment for the output performance evaluation of TENG. (**Fig. 4a inset**) This reactor can provide quantitative controls on environmental factors, including pressure, gas concentration, temperature, *etc*. A micro motor is used to provide the mechanical motions to mimic the water waves with various frequency and magnitude. When TENG is placed inside the vessel, the electrical measurement can be conducted under precisely-controlled environmental parameters.

The open-circuit voltages and short-circuit charge transfers were measured in different pressures, for both contact-separation (CS) mode and sliding freestanding (SFT) mode operations. As shown in **Fig. 4 a and b**, the the open-circuit voltage (as limited by the breakdown effect) has almost linear relationship with the pressure for both CS and SFT modes TENGs, which is consistent with the theoretical calculations. **Fig. 4 c and d** show the increase of the short-circuit charge transfers with the pressure in both modes, which reach relative stable values after ~ 4-5 atm, possibly due the limitation of the triboelectric performance of materials. The increase of this charge transfer is lower than the calculated values, showing the potential of furtherly increasing the charge density through other methods (e.g. charge injection). The overall increase in output energy density reaches over 25 times in CS mode in 10 atm and almost 5 times in SFT mode in 6.5 atm, respectively, as shown in **Fig. 4 e and f**. As shown in **inset of Fig. 4f**, the LED bulbs lighted by the SFT-TENG have obviously increased brightness when the pressure increase from 1 atm to 7.68 atm, demonstrating the effectiveness of the high pressure environment for promoting the output performance.

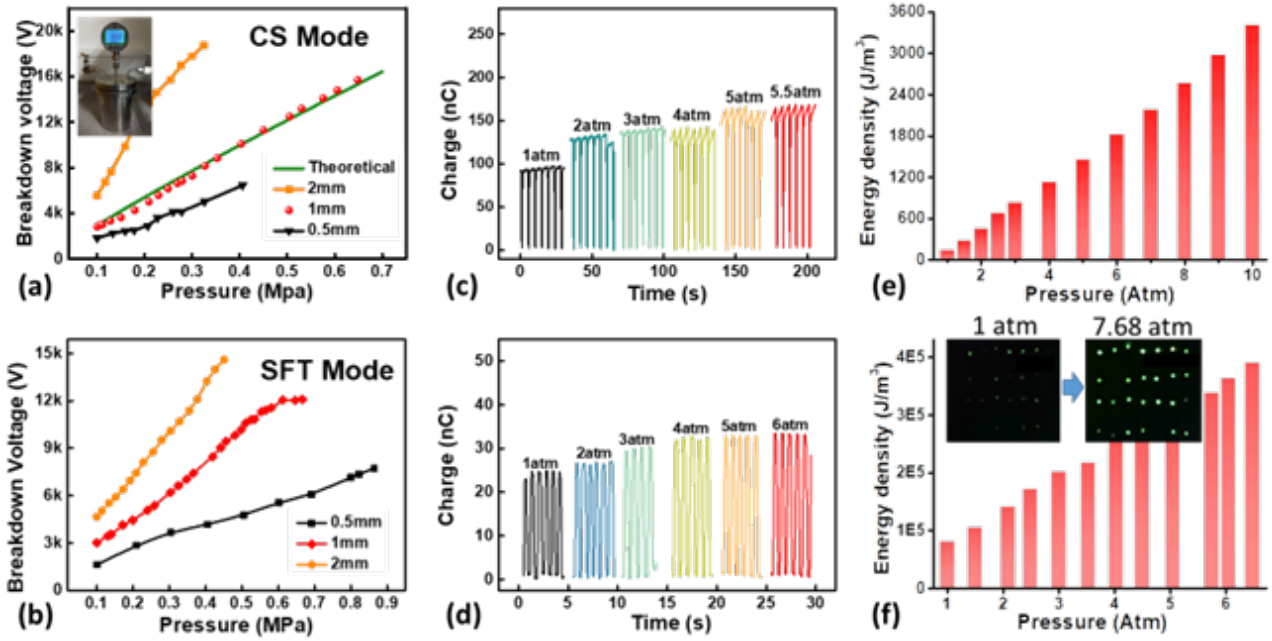


Figure 4. The enhancement of the output performance in high-pressure gas environments.

3.5 Demonstration of optimized water-TENG for harvesting simulated and real ocean waves

By summarizing previous studies about optimization of the water-TENGs, we fabricated the optimized device with the thick-shell structure (thickness ~ 1 cm), which was filled in ~1.5 atm CO_2 gas. Unfortunately due to limited technologies in the lab for packaging, we cannot achieve even higher inner pressure. We purchased a commercial removable water pool setup in the lab, filled in water, and used a wave machine to generate water waves (**Fig. 5a**). When the water-TENG was placed inside, it can naturally float on the water surface (**Fig. 5b**). When the wave machine is on, the water-TENG can shake with the waves, and generate power to light LEDs to make a sign to indicate directions (**Video S1**). We also demonstrated the power generation capability of water-TENGs in real ocean environment, located in the coast of Victoria Harbor in Hong Kong, (**Video S2**) which demonstrates the potential application of the developed devices serving as floating self-charging platform on the ocean.

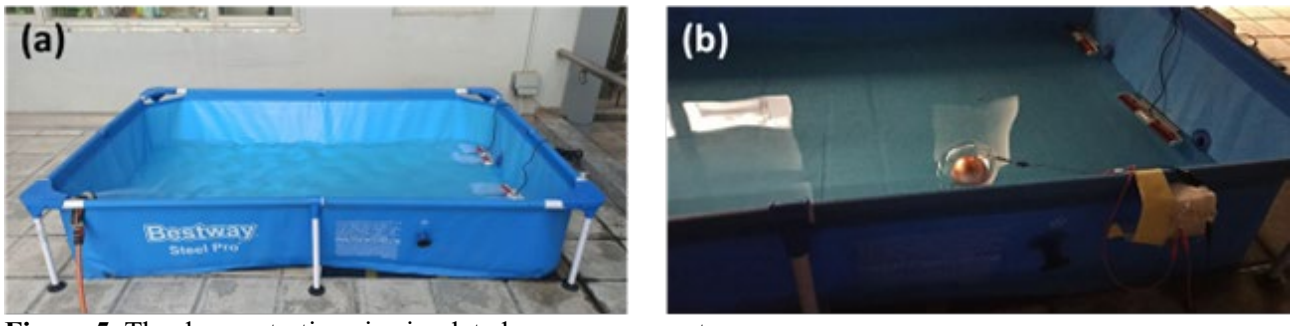


Figure 5. The demonstrations in simulated ocean wave systems.

3.6 Conclusions and perspectives

In summarizing this project, the discharge effect and parasitic-capacitance effect in water-TENG have been systematically studied, and optimized device was fabricated and demonstrated for ocean wave energy harvesting. All objectives have been completed with deliverables achieved. In these results, it has been demonstrated that high-pressure gases with relative high breakdown threshold, such as CO_2, is quite effective to enhance the output energy density of the device by 5-25 times. The thick-shell strategy can increase the output performance by suppressing the parasitic-capacitance effect. However, it also increase the volume, making the output energy density increase less than twice. The improved output performance of TENG was demonstrated in both simulated and real ocean waves. The future studies can focus on the packaging

technology aspect, which can package the high-pressure gas environment inside the TENG, for applications in the future large-scale energy harvesting from the ocean.

Video S1: Power generation from a water-TENG in simulated water waves.
Video S2: Power generation from water-TENGs in real water waves in Victoria Harbor, Hong Kong.

4. PUBLICATION AND AWARDS

J[1] J. Fu, G. Xu, C. Li, X. Xia, D. Guan, J. Li, Z. Huang, **Y. Zi**, "Achieving Ultrahigh Output Energy Density of Triboelectric Nanogenerator in High-Pressure Gas Environment", *Advanced Science*, Wiley, Germany, DOI: 10.1002/advs.202001757 (2020)
J[2] X. Xia, H. Wang, H. Guo, C. Xu, **Y. Zi**, "On the Material-Dependent Charge Transfer Mechanism of the Contact Electrification", *Nano Energy*, Elsevier, Netherlands, 78, 105343 (2020)
J[3] H. Wang, J. Wang, X. Xia, D. Guan, **Y. Zi**, "Multifunctional Self-Powered Switch toward Delay-Characteristic Sensors", *ACS Applied Materials & Interfaces*, American Chemistry Society, United States, 12 (20), 22873-22880 (2020)
J[4] L. Wang, Y. Liu, Q. Liu, Y. Zhu, H. Wang, Z. Xie, X. Yu, **Y. Zi**, "A Metal-Electrode-Free, Fully-Integrated, Soft Triboelectric Sensor Array for Self-Powered Tactile Sensing", *Microsystems & Nanoengineering*, Nature Publishing Group, United Kingdom, 6 (1), 1-9 (2020)
J[5] J. Wang, S. Li, X. Chen, **Y. Zi**, "High-Voltage Applications of the Triboelectric Nanogenerator - Opportunities Brought by the Unique Energy Technology", *MRS Energy & Sustainability*, Cambridge University Press, United Kingdom, 6, E17 (2020)
J[6] J. Wang, C. Meng, Q. Gu, M.C. Tseng, S.T. Tang, H.S. Kwok, J. Cheng, **Y. Zi**, "Normally Transparent Tribo-Induced Smart Window", *ACS Nano*, American Chemistry Society, United States, 14 (3), 3630-3639 (2020)
J[7] G. Xu, D. Guan, X. Yin, J. Fu, J. Wang, **Y. Zi**, "A Coplanar-Electrode Direct-Current Triboelectric Nanogenerator with Facile Fabrication and Stable Output", *EcoMat*, Wiley, Germany, DOI: 10.1002/eom2.12037 (2020)
J[8] Y. Liu, L. Wang, L. Zhao, X. Yu, **Y. Zi**, "Recent Progress on Flexible Nanogenerators toward Self-Powered Systems", *InfoMat*, Wiley, Germany, 2 (2), 318-340 (2020)
J[9] X. Xia, H. Wang, P. Basset, Y. Zhu, **Y. Zi**, "Inductor-Free Output Multiplier for Power Promotion and Management of Triboelectric Nanogenerators toward Self-Powered Systems", *ACS Applied Materials & Interfaces*, American Chemistry Society, United States, 12, 5, 5892-5900 (2020)
J[10] J. Wang, H. Wang, X. Li, **Y. Zi**, "Self-Powered Electrowetting Optical Switch Driven by a Triboelectric Nanogenerator for Wireless Sensing", *Nano Energy*, Elsevier, Netherlands, 66, 104140 (2019)
J[11] J. Fu, X. Xia, G. Xu, X. Li, **Y. Zi**, "On the Maximal Output Energy Density of Nanogenerators", *ACS Nano*, American Chemistry Society, United States, 13 (11), 13257-13263 (2019)
J[12] X. Xia, J. Fu, **Y. Zi**, "A universal standardized method for output capability assessment of nanogenerators", *Nature Communications*, Nature Publishing Group, United Kingdom, 10:4428 (2019)
J[13] G. Xu, X. Li, X. Xia, J. Fu, W. Ding, **Y. Zi**, "On the force and energy conversion in triboelectric nanogenerators", *Nano Energy*, Elsevier, Netherlands, 59, 154-161 (2019).
J[14] H. Wang, Q. Zhu, Z. Ding, Z. Li, H. Zheng, J. Fu, C. Diao, X. Zhang, J. Tian, **Y. Zi**, "A fully-packaged ship-shaped hybrid nanogenerator for blue energy harvesting toward seawater self-desalination and self-powered positioning", *Nano Energy*, Elsevier, Netherlands, 57, 616-624 (2019).
J[15] F. Chen, Y. Wu, Z. Ding, X. Xia, S. Li, H. Zheng, C. Diao, G. Yue, **Y. Zi**, "A Novel Triboelectric Nanogenerator Based on Electrospun Polyvinylidene Fluoride Nanofibers for Effective Acoustic Energy Harvesting and Self-powered Multifunctional Sensing", *Nano Energy*, Elsevier, Netherlands, 56, 241-251 (2019).
J[16] X. Li, G. Xu, X. Xia, J. Fu, L. Huang, **Y. Zi**, "Standardization of Triboelectric Nanogenerators: Progress and Perspectives", *Nano Energy*, Elsevier, Netherlands, 56, 40-55 (2019).
J[17] Z. Wen, J. Fu, L. Han, Y. Liu, M. Peng, L. Zheng, Y. Zhu, X. Sun, **Y. Zi**, "Toward Self-powered Photodetection Enabled by Triboelectric Nanogenerators", *Journal of Materials Chemistry C*, The Royal Society of Chemistry, United Kingdoms, 6, 11893-11902 (2018).

Research Reports (2020-2024) In Biomedical Engineering

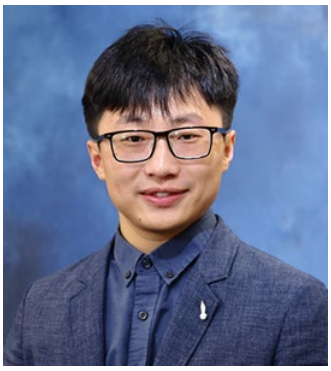
Newly Funded Projects (2024-2026)	* Cost-efficient highly potent antimicrobial peptide discovery for livestock farming antibiotic alternatives with protein language model-powered AI methods
	* Development of Mitochondria-targeting, Single-atom Nanozyme for Accelerated Bone Regeneration
Continuing Projects (2023-2025)	* Design, Optimization, and Experimental Validation of a Handheld Variable-Curvature Hybrid-Structure Robotic Instrument (HVHRI) for Maxillary Sinus Surgery
	* Smart bandage with integrated organic electronic sensor and iontronic drug delivery platform for advanced chronic wound care
(2022-2024)	* Coupling MoS2 Field-Effect Biosensors with Hybridization Chain Reaction Self-Assembly Amplification for Highly Sensitive and Label-free Nucleic Acid Detection

Completed Projects (2021-2023)	* Endoluminal Image-based Motion Optimization and Motion Understanding (EIMOMU) Towards Semi-Supervised Stiffness Tunable Flexible Robotic Procedures
(2020-2022)	* Development of deep-learning assisted ultrahigh-resolution endoscopic OCT for visualization and quantification of volumetric small airway microstructures in vivo * A dexterous robot for endoscopic submucosal dissection in the colon
(2019-2021)	* Development of a folded bilayer scaffold for intestinal tissue engineering * Optogenetic regulation of hormone production for glucose homeostasis maintenance
(2018-2020)	* Development of Highly Sensitive Quantitative Phase Microscopy forLabel-free Imaging of Neuronal Network Activities

(Funded Year)

COST-EFFICIENT HIGHLY POTENT ANTIMICROBIAL PEPTIDE DISCOVERY FOR LIVESTOCK FARMING ANTIBIOTIC ALTERNATIVES WITH PROTEIN LANGUAGE MODEL-POWERED AI METHODS

Principal Investigator: Professor LI Yu
Department of Computer Science & Engineering, CUHK



Co-investigator(s):
Dr. Irwin King [(1)],

Research Team Members:
Dr. Christopher K C LAI [(2)],
Dr. Lei Dai [(3),]

[(1)] Dept. of Computer Science & Engineering, Faculty of Engineering
[(2)] Dept. of Microbiology, Faculty of Medicine
[(3)] SIAT, Shenzhen, China

Project Start Date: 1 July 2024

ABSTRACT

From 2020, the addition of antibiotics has been banned in China's feed (Figure 1), aiming to reduce the harm caused by abusing antibiotics. To maintain normal development of animal agriculture, it is essential to find antibiotic alternatives, which are needed for animal normal growth. Antimicrobial peptides (AMPs), as natural antibacterial drugs, are less vulnerable to drug resistance because of their unique mechanism of action and are ideal substitutes for antibiotics. However, only 7 antimicrobial peptides are currently approved by FDA, so it is important to discover more useful antimicrobial peptides. Previous AMP mining methods omitted the tertiary structure of proteins, failed to efficiently find AMPs with strong bactericidal activity, and were not tested on animal pathogen strains. This project uses evolutionary and tertiary structural information based on protein language models to discover new potent AMPs from animal data. Our pipeline is very cost-efficient and 84% of the candidates identified by our method are potent AMPs. Preliminary results have shown 8 experimentally validated highly potent novel AMPs from the organism and their microbes' genome data, even more potent than medicinal AMP polymyxin B and swine-gut AMPs like PR-39, Cecropin P1, and Porcine beta-defensin 2 (Figure 2). These novel AMPs are promising to be applied in animal husbandry. The preliminary work has been invited to submit to Nature Biomedical Engineering after presubmission enquiry (Figure 3). We will refine the pipeline and discover more potent AMPs with a larger dataset in this proposed project to create larger impact and for transformational purposes.

INNOVATION AND PRACTICAL SIGNIFICANCE:

Innovation:

1. We develop the first method to using protein language model for AMP discovery. Our preliminary results show that it's very efficient and sensitive and can discover AMPs omitted by the previous method because of its awareness of structural information.

2. We are the first team to discover AMPs from both the bacterial genomes and host genome, dedicated to animal husbandry. The interaction between host and microbiome makes us easier to discover novel AMPs.

3. Based on our preliminary study, we have experimentally validated 62 candidates from our method. Impressively, 84% of our candidates are potent AMPs, better than all previous methods, and 8 of them even surpass medicinal AMP polymyxin B and swine-gut AMPs like PR-39, Cecropin P1, and Porcine beta-defensin 2. We will discover more AMPs on other animal genomes and further refine the comprehensive evaluation in this project.

Practical significance:

1. We propose a paradigm for cost-efficient and effective ways to discover and develop peptide drugs. Such a platform can give rise to a new drug discovery start-up.

2. We discovered novel highly potent AMPs and will discover more. They can be potentially used in the actual animal husbandry industry.

PROJECT OBJECTIVES AND LONG-TERM IMPACT:

1. To develop an AMP prediction model based on protein language models. The model should be fast and accurate so that we can apply it onto the metagenomic data and discover novel AMPs.
2. To discover highly potent AMPs for animal husbandry with the new pipeline from the large-scale animal metagenomic data. Based on our preliminary results, we will extend it to other animals than swine.
3. To biosynthesize and validate the above discovered AMP candidates to see if these AMPs are suitable to be used as feed additives in animal husbandry. We will perform bacterial level test, toxicity test, and animal test.

Our preliminary results have shown eight of our identified candidates even surpass medicinal AMP polymyxin B and swine-gut AMPs like PR-39, Cecropin P1, and Porcine beta-defensin.
Upon finishing the animal test, we will push the transformation to patent the AMPs and seek for the opportunities to apply them into the animal husbandry industry and create real-life impact.

DEVELOPMENT OF MITOCHONDRIA-TARGETING, SINGLE-ATOM NANOZYME FOR ACCELERATED BONE REGENERATION

Principal Investigator: Professor LI Zhong Alan
Department of Biomedical Engineering, CUHK

Co-investigator(s):
Mr. XIONG Tiandi [(1)],

[(1)] Dept. of Biomedical Engineering

Project Start Date: 1 July 2024

ABSTRACT

Critical-sized bone defects (CSBD) do not heal spontaneously and pose significant challenges in healthcare. Osteogenic differentiation of bone-resident skeletal stem/progenitor cells (SSPCs) is a highly energy-consuming process that plays an essential role in bone regeneration. Therefore, biomaterials capable of regulating SSPC energy metabolism and promoting its osteogenic differentiation are highly desirable in CSBD management. Herein, we propose to develop a novel nanozyme that targets mitochondria (the powerhouse of cell) and boosts cellular energy production by mimicking the function of cytochrome c oxidase, a natural enzyme present in mitochondria. We will introduce single Cu and Fe atoms to dendritic silica nanoparticles characterized by a large specific surface area and modify them with mitochondria-targeting triphenylphosphonium. We will comprehensively examine the chemical and structural properties and catalytic activities of the nanozyme and evaluate their ability to enhance the osteogenic differentiation of SSPCs in vitro. Finally, we will use a mouse model to assess the nanozyme's ability to facilitate CSBD regeneration in vivo and explore the associated molecular mechanisms. Our innovative mitochondria-targeting, single-atom nanozyme holds promising potential in CSBD treatment and is expected to greatly relieve the heavy socioeconomic burden of CSBD and other bone diseases such as fracture and osteoporosis.

INNOVATION AND PRACTICAL SIGNIFICANCE:

The major innovation of the current work lies in the design of a biomaterial based on cell-derived ECM, which potentially enables the delivery and long-term stabilization of cell-derived therapeutic signaling factors. We have already filed a provisional patent application in May 2019 (Anna Blocki and Marisa Assuncao, Chinese University of Hong Kong (2019) "Process and material for tissue healing" US provisional patent application: 62/848,971), protecting this material design.

By utilizing this biomaterial, we are able to address major limitations of cell-based therapies, such as limited engraftment and survival of transplanted cells, limited therapeutic efficacy of conditioned media, immunological concerns, when allogeneic cell sources are utilized (ECM is highly conserved and thus not evoke an immune response), etc. At the same time, the engineered biomaterial exhibits the necessary complex bioactivity to guide complex tissue healing processes, in contrast to selected biologics or simple scaffolds

The unique techniques utilized, enable the synthesis of larger amounts of biomaterial, thereby ensuring a stable/reproducible bioactivity and are necessary for future scale-up production.
The bioactive material can be stored and thus utilized off-the-shelf. It can be processed and incorporated in all types of materials including tissue scaffolds, implants, wound dressings and (injectable) hydrogels. Hence, just by itself or incorporated into other materials, it can be applied to tissue areas with chronically inflamed and

dysregulated microenvironments, thereby modulating and turning the diseased environment into a pre-healing one. This will advance the healing and regeneration process in non-healing and non-regenerative tissues such as osteoarthritis and beyond.

PROJECT OBJECTIVES AND LONG-TERM IMPORT:

Specific objectives
The specific objectives of this project include:
a) Synthesizing a CoC-mimicking nanozyme with dispersed Cu and Fe single atoms on dendritic mesoporous silica nanoparticles (DMSN);
b) Optimizing the biocompatibility and osteogenic potential of the nanozymes using in vitro SSPC cultures and investigating the underlying molecular mechanisms; and
c) Evaluating the nanozyme's ability to promote CSBD repair in vivo using a mouse model.

Expected outcomes
The expected outcomes of this project include:
a) A versatile DMSN-based platform that can be used to prepare various nanozymes and drug delivery systems targeting mitochondria;
b) A CcO-mimicking nanozyme capable of generating increased energy to promote osteogenic differentiation of SSPCs; and
c) A 3D, nanozyme-enhanced therapeutic construct for accelerated CSBD regeneration.

Significance and long-term impact
The CcO-mimicking nanozyme invented in this study has a low manufacturing cost, high stability, and sustained catalytic activities across a wide pH and temperature range, making it a promising commercial product as a substitute for natural CcO enzyme. Additionally, this bionic technology can offer a new platform for biomaterial researchers, chemists and chemical engineers to study the kinetics of enzyme- catalyzed reactions, thus extending the application of the developed nanozyme beyond biomedical applications.
With the ability to enhance the osteogenic differentiation of SSPCs, the nanozyme can broaden the application scenarios of stem cell therapies and tissue engineering strategies for treating various bone diseases, including CSBD, bone fracture, and osteoporosis, as well as a wide range of other conditions involving aberrant cellular energy metabolism, ultimately improving patient outcomes and reducing the associated burden on the healthcare system.

WORK IN PROGRESS

DESIGN, OPTIMIZATION, AND EXPERIMENTAL VALIDATION OF A HANDHELD VARIABLE-CURVATURE HYBRID-STRUCTURE ROBOTIC INSTRUMENT (HVHRI) FOR MAXILLARY SINUS SURGERY

Principal Investigator: Professor MA Xin[1,4]
Department of Mechanical and Automation Engineering, CUHK

Co-Investigator:
Weibin Li[2], Xianfeng Xia[3], Samuel Au[1,4], Zheng Li[3,4]

Research Team Members: Yi Yang[4],, Xuchen Wang[1,4], Puchen Zhu[1,4]

(1) Dept. of Mechanical and Automation Engineering, The Chinese University of Hong Kong.
(2) School of Computer Science, Sun Yatsen University.
(3) Dept. of Surgery, The Chinese University of Hong Kong.
(4) Multi-Scale Medical Robotics Center, Hong Kong.

Reporting Period: 1st July 2023 – 31st May 2024

INNOVATION AND PRACTICAL SIGNIFICANCE:

To include a paragraph to highlight specifically the innovation and practical significance of your work. Both VC and the donor would like to see more research endeavors be directed to innovation and technology transfer for the betterment of mankind.

We will develop a dexterous, compact HVHRI with high bending and torsional stiffness for sinus surgery. Compared with existing robotic flexible instruments, the HVHRI has larger reachability and dexterity, which will reduce the trauma to the patients in surgery and the possibility of complications after surgery.

ABSTRACT

Existing robotic flexible medical tools for maxillary sinus surgery are still low in dexterity and big in diameter, which results in big damage to the patients. We will develop a novel 4-DOF handheld hybrid-structure, variable-curvature robotic instrument (HVHRI) (including a 1-DOF variable-curvature flexible bending section and a 3-DOF distal grasper), which is compact in size (diameter is 3.5 mm; actuation system weighs < 800g) and can provide sufficient bending (14 $N{\cdot}mm$) and torsional stiffness (0.5°/$N{\cdot}mm$). To our knowledge, it is the thinnest flexible instrument for sinus surgery. And it is capable of sharply bending at the distal end (bending radius is 1.5 mm). To further enlarge the reachable space and dexterity of the HVHRI inside the maxillary sinus, a novel structure parameter optimization framework (maximizing the reachable space and dexterity of the HVHRI) will be studied. The dexterity and reachability of the existing two-segment flexible tool, the hybrid-structure flexible tool, and the HVHRI will be compared by simulations and experiments. Besides, we will propose two intuitive control methods for the HVHRI (handheld and integrated with robot arm). Several phantom and cadaver experiments will be conducted to validate the feasibility, dexterity, reachability, bending stiffness and torsional stiffness of the HVHRI.

1. OBJECTIVES AND SIGNIFICANCE

1. Existing robotic medical tools for maxillary sinus surgery are still straight, low in dexterity and big in diameter (> 4 mm), which results in big damage to the patients (a big incision is often needed in the face). We will develop a novel 4-DOF HVHRI that has a small diameter (3.5 mm), high reachability and dexterity. With the HVHRI, surgeons can dramatically reduce the trauma caused to the patients in surgery.
2. Most of the existing medical instruments are operated manually. However, the enhanced dexterity adds to the difficulty of manual operation, which in turn imposes upon surgeons longer learning curve and more

workload. To facilitate sterilization, reduce surgeons' learning time and avoid surgeons' fatigue, we will develop a compact actuation system for the HVHRI (with weight < 800g) which can be easily handled by hand or integrated with robot arm. This actuation system enables convenient sterilization and easy change of instruments for the surgeons.

3. A novel structure optimization framework will be developed for maximizing dexterity and reachability of HVHRI inside the maxillary sinus area.

4. To validate the feasibility, operability, and workload of the HVHRI, we will conduct several simulations and experiments. Simulations will be conducted to compare the dexterity and reachability between existing medical instruments and the HVHRI. Several phantom and cadaver experiments will be conducted to validate the feasibility, dexterity, reachability, bending stiffness and torsional stiffness of the HVHRI. Results of the simulations and experiments will provide surgeons with detailed data when they choose instruments.

2. RESEARCH METHODOLOGY

2.1. Development of a novel variable-curvature, hybrid-structure manipulator (VHM) that has a small diameter (3.5 mm) and high reachability and dexterity in the maxillary sinus area.

As is shown in Fig. 1(a), we will propose a novel VHM. (1) In this VHM, we will develop a novel 3-DOF gripper (Fig. 1(b)), which can be driven by four cables. The gripper (the diameter is 3.5mm; the length is 9 mm) is more compact than existing ones and can bend 90° in two directions. (2) Besides, we will develop a 1-DOF variable-curvature flexible bending section (Fig. 1(c)). The joint is produced on metal tubes by laser cutting. And guide rings can be easily made by sheet metal forming technology to constrain the 6 actuating cables. Different maximum bending angles (α_i) will be set at each joint of the flexible bending section. In operation, due to the different joint limit of each joint (α_i), the curvatures of the flexible bending section can be manipulated to increase the reachability of the VHM. (3) Besides, we will design a novel metal-woven mesh to further increase the overall torsional stiffness of the manipulator. And the FEA simulation results (Fig. 1(d)) show that the torsional stiffness of the metal-woven mesh is twice better than the existing nitinol flexible bending section. All in all, the VHM contains a 3-DOF gripper, a 1-DOF flexible bending section, and a novel metal-woven mesh.

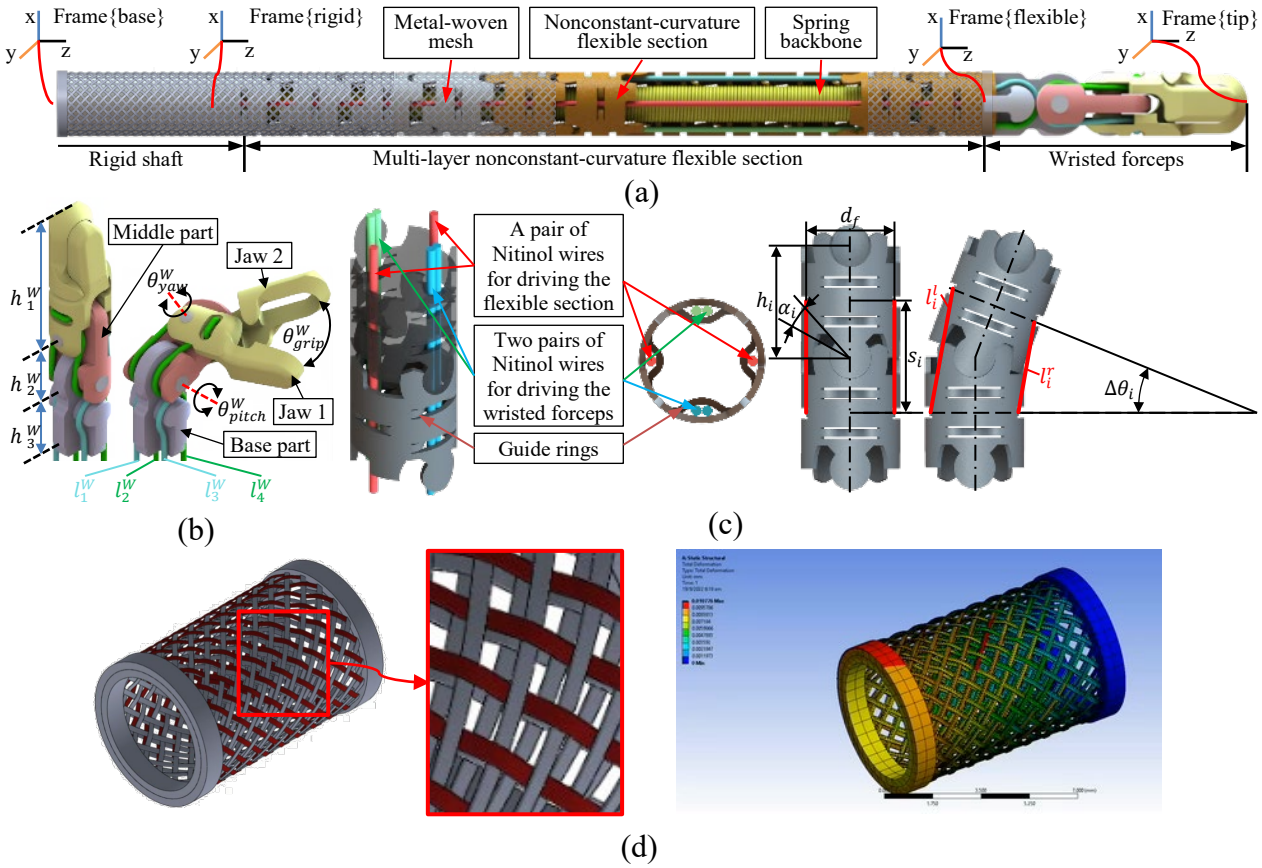
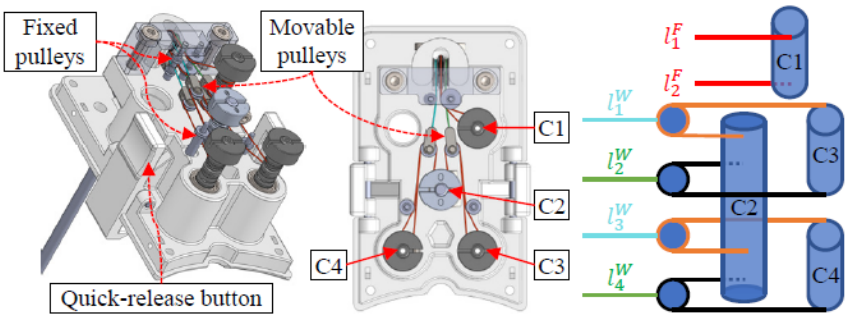


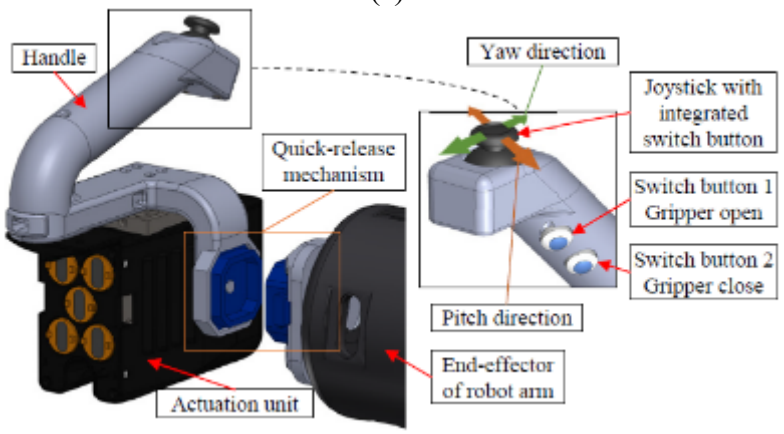
Fig. 1. (a) Overview of the variable-curvature, hybrid-structure manipulator. (b) The structure of the 3-DOF gripper. (c) The structure of a single joint of the 1-DOF variable-curvature flexible bending section. (d) The metal-woven mesh for further increasing the overall torsional stiffness of the manipulator and its FEA simulation.

2.2. Development of a compact actuation system (with weight < 800g) that can be easily handled by hand or integrated with robot arm.

First, we will develop a backend with a novel decoupling mechanism (see Fig. 2(a)) to connect HVM and the actuation system. With this backend, we can use three motors to separately control the motion of the 3-DOF gripper (pitch, yaw and open/close) and use one motor to control the motion of the flexible bending section so that the coupling motions of the gripper and the flexible section can be reduced. In addition, as shown in Fig. 2(b), we will develop a novel interface to operate the HVHRI. The interface will contain a joystick and two buttons. The joystick is used to control the manipulator and the two buttons are used to switch the gripper between “open” and “close”. The actuation system contains four motors. The manipulator can be detached via a quick-release structure, which is convenient for sterilization.



(a)



(b)

Fig. 2. (a) Backend with decoupling mechanism. (b) Interface for operating the HVHRI.

2.3. Structure optimization for the 4-DOF HVHRI by maximizing the dexterity and reachability inside the maxillary sinus area.

As is shown in Fig. 3, we will propose a novel structure optimization framework for the VHM. In this framework, we will further consider both the dexterity and reachability of the instruments in the maxillary sinus area. To my knowledge, this is the first time to optimize the structure of surgical instruments considering both the reachability and dexterity for the sinus surgery. Besides, more structure parameters will be optimized (including the different maximum bending angles for the joints of the flexible bending section) in the framework.

Algorithm 1 Structure Parameters Optimization Framework

Input: Positions of environment cloud points (P_{envir}), positions of target space's cloud points (P_{target}), and positions of waypoints (P_{way}).

Output: Number of sections (N_{seg}), length of each section (L_i, $i = 1, 2, \cdots, N_{seg}$), joint limit of each section ($\theta_{limit-i}, i = 1, 2, \cdots, N_{seg}$).

```
for N_seg = 1 : N_seg^max do
    for i = 1 : N_seg do
        for L_i = L_i^min : L_i^max do
            for θ_limit-i = θ_limit-i^min : θ_limit-i^max do
                p_i = [L_i, θ_limit-i]
            end for
        end for
    end for
    Dex_global^max(p) = MaxOpt(Dex_global(p), P_envir, P_target, P_way)
end for
[L_1, θ_limit-1, ..., L_N_seg, θ_limit-N_seg]_optimal = Argmax(Dex_global^max(p))
```

Fig. 3. Structure optimization framework for HVHRI

2.4. Experimental validation for the HVHRI.

To evaluate the performance of the HVHRI, we will conduct several experiments in six aspects: 1) tracking the trajectories of the HVHRI's motions to verify the proposed kinematics and decoupling mechanisms; 2) a bending stiffness test on the HVHRI to show how much force the hybrid structure endures; 3) a torsional stiffness test on the HVHRI to show how many twists it overcomes; 4) comparing the reachability and dexterity of the HVHRI and other flexible instruments; 5) a feasibility study in a 3D maxillary sinus phantom to show the reachability and dexterity of the HVHRI; and 6) cadaver experiments for validating the feasibility of the HVHRI.

3. RESULTS ACHIEVED SO FAR

3.1. Development of a novel variable-curvature, hybrid-structure manipulator (VHM) that has a small diameter (3.5 mm) and high reachability and dexterity in the maxillary sinus area.

In this project, we prototyped HVHRI (see Fig. 4(a)) comprises three major components: a detachable instrument unit, an actuation unit, and a handheld user interface. **The detachable instrument unit** consists of a VHM and a backend transmission mechanism (BTM). The VHM includes a pair of 3-DOF wristed forceps and a 1-DOF multi-layer variable-curvature flexible section. **The 3-DOF wristed forceps:** In order to address the challenges associated with reducing the diameter of surgical instruments for MSS, a specialized design of the 3-DOF wristed forceps with a 3.5 mm diameter was developed. There will be challenges to directly scale down the diameter of a conventional Φ8 mm instrument to Φ3.5 mm: (1) accommodating six driving cables within limited space, (2) ensuring high stiffness of thinner driving cables (Φ0.5 mm to Φ0.2 mm), and (3) avoiding sharp bending angles

for driving cables with small pulley diameters (Φ3.8 mm to Φ1.6 mm). To overcome these challenges, a novel approach is employed in the design. The 3-DOF wristed forceps incorporate a four-cable wrist mechanism, utilizing symmetrically fixed cables with high friction paths on the jaws to control pitch ($\boldsymbol{\theta}^{\boldsymbol{W}}_{\boldsymbol{pitch}}$), yaw ($\boldsymbol{\theta}^{\boldsymbol{W}}_{\boldsymbol{yaw}}$), and grip ($\boldsymbol{\theta}^{\boldsymbol{W}}_{\boldsymbol{grip}}$) motions. By using only four driving cables, more space is made available to accommodate thicker cables, thus effectively enhancing stiffness. Nitinol wires with a diameter of 0.3 mm are selected as the

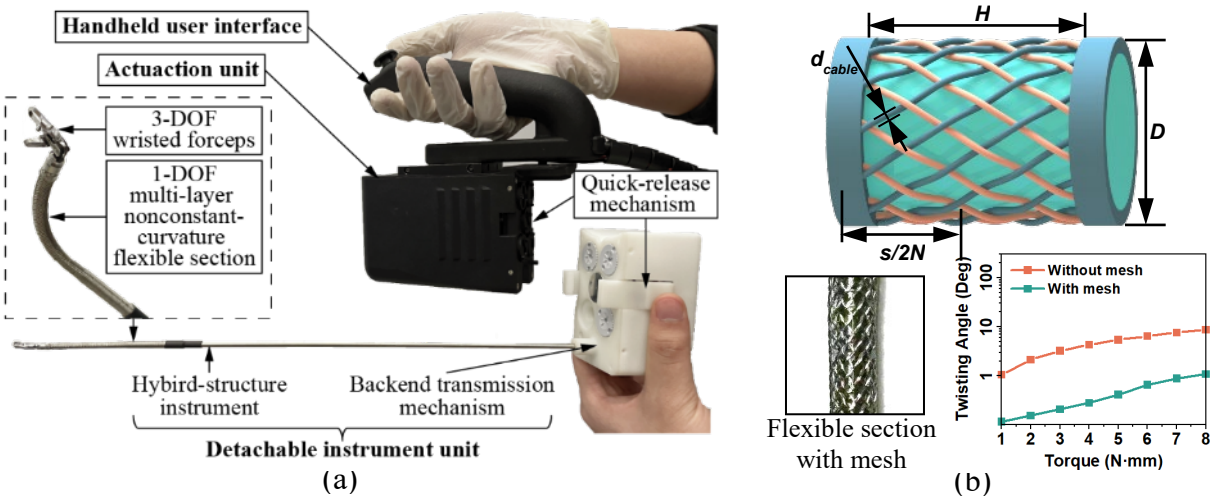


Fig. 4. (a) Overview of the HVHRI prototype. (b) Model, prototype, and torsional stiffness test of the metal-woven mesh.

driving cable, which provides up to 50 N tension. To avoid bending the driving cables with sharp angles, the four driving cables are designed to be in contact with the smooth convex sliding surfaces of the parts to eliminate the need of pulleys with smaller diameters. As well, the streamlined design of the wristed forceps assembly comprised only six parts, facilitating both manufacturability and assemble ability. **The 1-DOF multi-layer variable-curvature flexible section with high stiffness:** Flexible medical tools face several challenges that hinder their effectiveness in various procedures. Firstly, the use of constant curvature limits flexible instruments' dexterity and reachability, particularly in complex anatomical structures such as the nasal and maxillary sinus cavities. Secondly, existing robotic flexible instruments are often low in bending and torsional stiffness, which impacts their stability during tissue manipulations. This compromises the surgeon's ability to perform delicate maneuvers accurately. Lastly, the diameter of current robotic flexible instruments for sinus surgery typically ranges from 4 to 5 mm, which may pose challenges when operating in tight space and may limit their access to certain areas within the maxillary sinus. To overcome these challenges, we propose a multi-layer variable-curvature flexible section, allowing for improved dexterity, reachability, and enhanced stiffness for the HVHRI. The flexible section consists of three layers, including 1) a metal-woven mesh, 2) a variable-curvature flexible section, and 3) a spring backbone. ***First layer:*** The metal-woven mesh is woven with 24 braided steel cables of 0.1 mm diameter (see Fig. 4(b)). The mesh is tightly welded to the flexible section to enhance the torsional stiffness of the flexible section and ensure smooth insertion through the nasal cavity during the surgery. Based on the torsional stiffness test results, the metal-woven mesh decreased an average of 89.41% twisting angle of the original flexible section. ***Second layer:*** The flexible section is manufactured by laser-cutting with features of variable curvature, small diameter, and complex decoupling guide rings for driving cables. The variable curvatures are achieved by setting different joint limitations on the laser-cutting processed bendable portion. Note that the variable curvatures are optimized with the structure parameters optimization framework to improve the reachability and dexterity of the HVHRI in the maxillary sinus cavity. Based on the optimization results in Section IV-B, the flexible section has two segments with the same single joint length. The numbers of single joints in two segments are N_{F1}=5 and N_{F2}=7. And the joint limitations of the two segments are α_1=5° and α_2=30°. The flexible section is manufactured from a stainless-steel tube with a small outer diameter of 3.4 mm. And guide rings are manufactured by stamping on the laser-cutting processed tube. In addition, since the wristed forceps are attached at the distal end of the flexible section, the driving cables of the forceps are inevitably influenced by the bending motion of the flexible section. To avoid this coupling effect, two series of guide rings are positioned along the center axis of each side of the flexible section to pass through the two pairs of driving cables of the wristed forceps. This decoupling structure ensures that the wristed forceps are not significantly influenced by the bending motion of the flexible section, thus improving maneuverability during surgical

procedures. ***Third layer:*** The bending stiffness of the single flexible section is low, due to the small diameter of the flexible section. Therefore, a stiff spring (Φ 2 mm) is installed at the center of the flexible section, serving as a backbone to provide curvatures and enhance the bending stiffness for the flexible section. Besides, we use Nitinol wires as the driving cables of the flexible section to further improve its bending stiffness and make it easy to assemble.

3.2. Development of a compact actuation system (with weight < 800g) that can be easily handled by hand or integrated with robot arm.

In this project, we developed the actuation system of the HVHRI to be as compact and light as possible, as it is a handheld surgical device. To meet design criteria, as shown in Fig. 5(a), the BTM of the actuation system includes a pulley set decoupling mechanism that utilizes only four sets of capstans to control all 4-DOF motions of the hybrid-structure instrument of the HVHRI. This mechanism efficiently transfers the coupled translation motions of the six driving cables to the rotation motions of the four sets of capstans, which greatly improves the controllability of the HVHRI. Specifically, one set of capstans (C1) controls the bending of the flexible section, while the other three sets of capstans (C2, C3, and C4) control the pitch motion, the yaw motion of Jaw 1, and the yaw motion of Jaw 2 of the wristed forceps. With this pulley set decoupling mechanism, only four motors are used in the system (Fig. 5(b)). By contrast, without this pulley set decoupling mechanism, at least five motors would be required to control the wristed forceps and the flexible section, adding at least 20% unnecessary weight to the system. In addition, the quick-release mechanism on the BTM includes an adapter and two spring-return switches. During operation, the BTM is secured to the adapter by the switches. Pressing the switches allows easy detachment of the BTM from the actuation unit. The detached instrument unit only contains mechanical components, which can be sterilized individually through methods such as high temperature and liquid immersion.

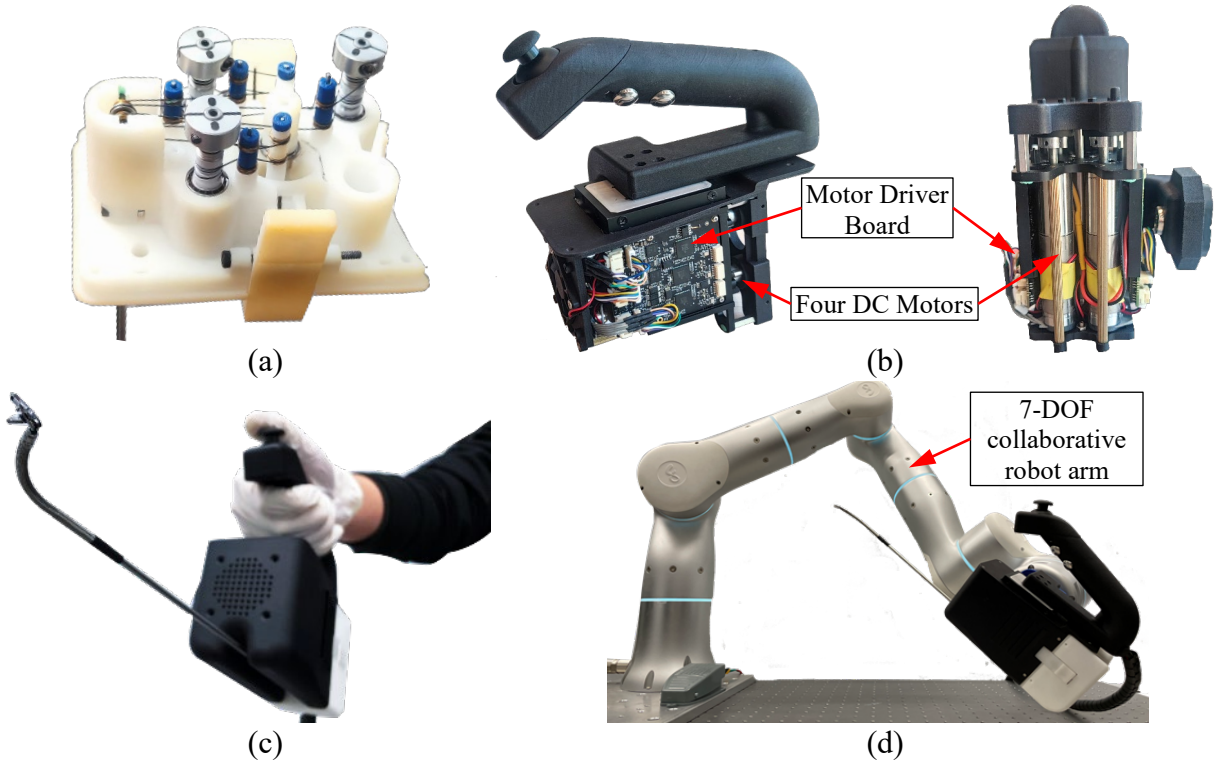


Fig. 5. (a) Prototype of the BTM. (b) Four-motor actuation unit and user interface of the HVHRI. (c) HVHRI works as a standalone handheld surgical instrument. (d) HVHRI works as a robot-arm-assisted surgical instrument.

The handheld user interface of the HVHRI is designed for easy to control during the MSS. As shown in Fig. 5(b), the interface includes a handle and a joystick with an integrated button for controlling the pitch and yaw motions of the wristed forceps, as well as the bending in the pitch direction of the flexible section. The integrated button serves as a switch to change the control mode. Two additional buttons are integrated into the interface to control the opening and closing of the forceps. In addition, a quick-release mechanism is designed between the

handle frame and the end-effector of the robot arm, enabling easy switching between robot-arm-assisted and standalone device as needed during the surgery (see Figs. 5(c)(d)).

3.3. Structure optimization for the 4-DOF HVHRI by maximizing the dexterity and reachability inside the maxillary sinus area.

We developed a structure optimization software for maximizing the dexterity and reachability of the HVHRI. The inputs of this software are position information sets of the environment cloud points ($\mathcal{P}_{envir}$), target space cloud points ($\mathcal{P}_{target}$), and waypoints ($\mathcal{P}_{way}$). In the development of the optimal flexible section curvature, we generated the anatomical structures of the nasal cavity and maxillary sinus cavity based on CT scan data obtained from a real phantom (Patient "Meyer", PHACON GmbH., Germany [15]) (see Figs. 6(a) and (b)). This process allows us to obtain the inputs including 1) a set of environment cloud points encompassing both the nasal cavity and the maxillary sinus cavity, 2) a set of target space cloud points representing the inner surface of the maxillary sinus cavity, and 3) a set of waypoints including the nasal entrance and the surgical incision opened between the nasal cavity and the maxillary sinus cavity.

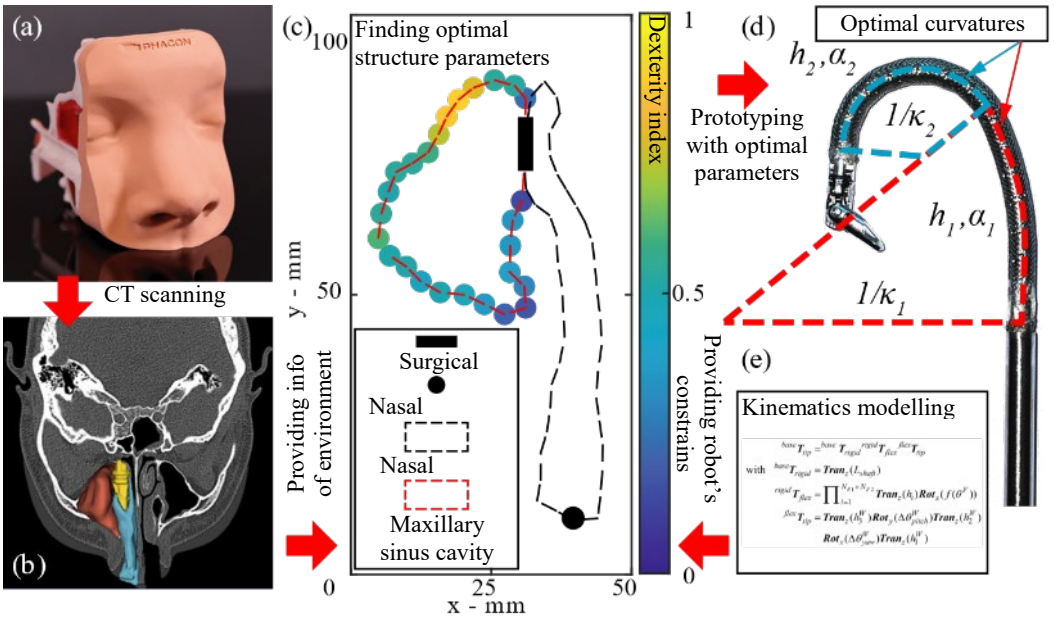


Fig. 6. Optimization procedure of the framework. (a) Sinus phantom constructed based on actual patient scans. (b) CT scan data based on the phantom. (c) Dexterity indices distribution based on the optimal flexible section curvatures. (d) Structure parameters with optimal flexible section curvatures. (e) Kinematics modeling of the flexible section serving as constrains.

As shown in Fig. 6(d), the output of this optimization framework is a parameter vector representing the optimal curvature of the flexible section ($\boldsymbol{s} = [N_{seg}, N_{F,1}, h_1, \alpha_1, \dots, N_{F,}, h_{N_{seg}}, \alpha_{N_{seg}}]^T$), which include the number of segments (N_{seg}), the number of joints in each segment ($N_{F,i},\ i = [1, N_{seg}]$), the joint length in each segment (h_i), and the joint limitation in each segment (α_i). Then, the joint in the i-th segment has a curvature of $\kappa_i \in [0, 2\tan(0.5\alpha_i)/h_i]$.

The optimization process of this framework is mainly maximizing the reachability and dexterity of the HVHRI (see Fig. 6(c)) under configuration limits and constrains of kinematics (see Fig. 6(e)). The reachability index (R) is defined as a ratio of the reachable surface area (A_{reach}) to the entire surface area (A_{all}) of the target space, which is $R = A_{reach}/A_{all}$. A commonly used rapidly exploring random tree algorithm is introduced to ensure the accurate deployment of the HVHRI without encountering collisions, thereby assessing its reachability effectively. The dexterity service spheres are used for every discrete point on the inner surface of the maxillary sinus cavity to show all orientations achieved by the tip of the HVHRI. The dexterity index ($D(\mathrm{P})$) of the point P is defined as a ratio of the area of the service region $\left(A^{\mathrm{P}}_{service}\right)$ to the total area of the void region without any obstacle $\left(A^{\mathrm{P}}_{void}\right)$ at the point P, which is $D(\mathrm{P}) = A^{\mathrm{P}}_{service}/A^{\mathrm{P}}_{void} \in (0, 1]$. Then, we get an objective

function by weighting and adding the reachability index and mean of the dexterity indices of the target space. Therefore, the flexible section curvature optimization framework is formulated as:

$$\max \quad \alpha_R R + \alpha_D \frac{1}{J}\sum_{j=1}^{J} D(\boldsymbol{p}_{target}^{j})$$

$$\text{s.t.} \begin{cases} \| \boldsymbol{p}_{target}^{j} - \boldsymbol{f}(\boldsymbol{T}_w, \boldsymbol{q}, \boldsymbol{s}) \| = 0, j \in [1, J] \\ \min(\| \boldsymbol{p}_{way}^{k} - \boldsymbol{g}(\boldsymbol{T}_w, \boldsymbol{q}, \boldsymbol{s}) \|) \le L_k, k \in [1, K] \\ q^m \in [q_{min}^m, q_{max}^m], m \in [1, M] \\ s^n \in [s_{min}^n, s_{max}^n], n \in [1, N] \\ inShape(alphaShape(\mathsf{P}_{envir}), \boldsymbol{g}(\boldsymbol{x}_0, \boldsymbol{q}, \boldsymbol{s})) = 1 \\ \boldsymbol{p}_w \in \mathsf{S}_{Limitation} \end{cases}$$

where α_R and α_D are two constants serving as weights of the reachability and dexterity indices, respectively. $\boldsymbol{p}_{target}^{j} \in \mathcal{P}_{target}$ is the position information of j-th point of the target space cloud points. $\boldsymbol{p}_{way}^{k} \in \mathcal{P}_{way}$ is the position information of k-th waypoint. $\boldsymbol{T}_w$ is the transformation matrix between the world coordinate system and the HVHRI's base coordinate system, which can be left multiplied by ${}^{base}\boldsymbol{T}_{tip}$, the transformation matrix between the Frame {base} to the Frame {tip} of the HVHRI in Fig. 6(e), to get the HVHRI's tip transformation matrix in the world coordinate system. $\boldsymbol{q} = [q^1, \dots, q^M]^T$ is the kinematic configuration vector, which is $\boldsymbol{q} = [\theta^F, \theta_{pitch}^W, \theta_{yaw}^W]^T$ in the HVHRI. $\boldsymbol{f}(\boldsymbol{T}_w, \boldsymbol{q}, \boldsymbol{s})$ is the position vector of the HVHRI's tip point based on the flexible section curvatures and the kinematics in Fig. 6(e). $\boldsymbol{g}(\boldsymbol{T}_w, \boldsymbol{q}, \boldsymbol{s})$ is the position vector of HVHRI's body points based on the flexible section curvatures and the kinematics in Fig. 6(e). L_k is the distance tolerance between the k-th waypoint and the HVHRI's body. $alphaShape(\cdot)$ is a MATLAB function that constructs a bounding area or volume around a given set of 3D points. $inShape(\cdot)$ is a MATLAB function that determines whether a given point is inside an $alphaShape$. $\boldsymbol{p}_w$ is the position vector of the HVHRI's base in the world coordinate system constrained in space outside the human head $\mathsf{S}_{Limitation}$.

Based on the mechanical design of the HVHRI, we specified several constraints as:

$$\begin{cases} N_{seg} \le 2 \\ h_i \ge 4 \text{ mm}, \alpha_i \le 30^\circ, i = [1, N_{seg}] \\ L_1 = 5 \text{ mm}, L_2 = 5 \text{ mm} \\ \theta^F \in [-\sum_{i=1}^{N_{seg}} N_{Fi}\alpha_i, \sum_{i=1}^{N_{seg}} N_{Fi}\alpha_i] \\ \theta_{pitch}^W \in [-90^\circ, 90^\circ], \theta_{yaw}^W \in [-90^\circ, 90^\circ] \end{cases}$$

where h_i and α_i are constrained by the displacement of six guide rings in every single joint. L_1 is determined by the radius of the nasal entrance since the first waypoint locates at the center of the entrance. L_2 is determined by the radius of the surgical incision between the nasal cavity and the maxillary sinus cavity since the second waypoint locates at the center of the incision. Then, the entire framework was implemented in MATLAB, employing the Interior-Point algorithm to obtain a parameter vector representing the optimal flexible section curvatures for the HVHRI, which was $\boldsymbol{s}_{optimal} = [N_{seg} = 2, N_{F,1} = 5, h_1 = 4 \text{ mm}, \alpha_1 = 5°, N_{F,2} = 7, h_2 = 4 \text{ mm}, \alpha_2 = 30°]^T$ (see Fig. 6(d)). Utilizing these parameters, the dexterity index and mean dexterity indices of the proposed robot were calculated as 100% and 48%, respectively, based on sectional CT scan data of the maxillary sinus cavity (see Fig. 6(a)). For comparative purposes, we also computed the dexterity indices for a constant-curvature hybrid-structure forceps (CHF) with identical maximum bending angle (235°) which had a parameter vector, $\boldsymbol{s}_{CHF} = [N_{seg} = 1, N_{F,1} = 12, h_1 = 4 \text{ mm}, \alpha_1 = 19.58°]^T$. As shown in Fig. 7(b), the CHF exhibited a low reachability index of 35% and a low mean of the dexterity indices of 17%. In addition, the dexterity indices on a target point for both instruments are shown in Fig. 7. Notably, the HVHRI demonstrated a 122% higher dexterity index at the target point compared with the CHF.

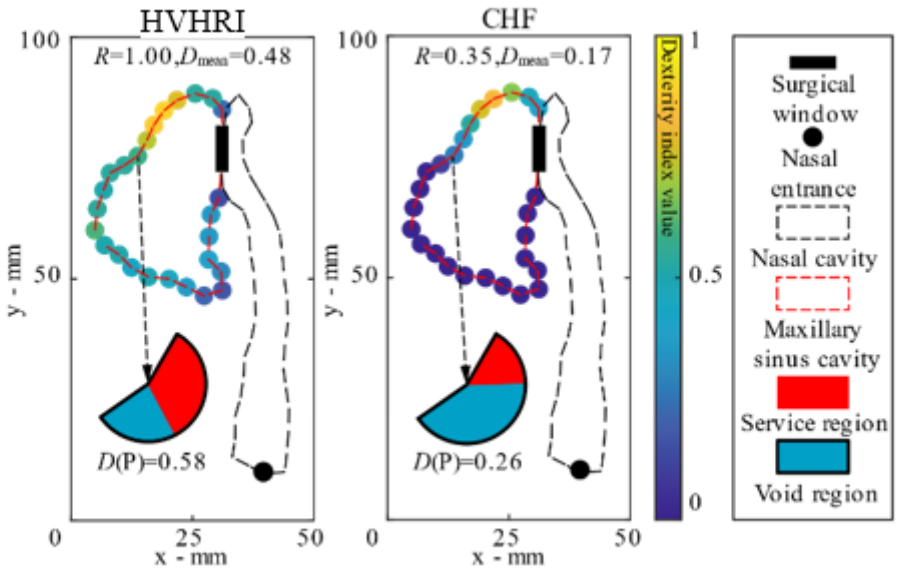


Fig. 7. Dexterity indices distributed in the maxillary sinus cavity of the HVHRI and the CHF

3.4. Experimental validation for the HVHRI.

3.4.1. TORSIONAL AND BENDING STIFFNESS TESTING

To assess the HVHRI's ability to maintain its shape during tissue manipulation, a bending stiffness test was conducted. As shown in Fig. 8(a), the actuation unit, along with the detachable instrument, was secured on a test bench. A 10 g weight hanger was suspended from the wristed forceps. The test involved gradually adding 20 g weights to the hanger until reaching a total weight of 70 g. The same stereo vision system used in the kinematics verification experiment was employed to measure the displacement of the base coordinate and the distal end of the wristed forceps. The test was carried out with the flexible section bent at 0°, 45°, and 90° to assess the bending stiffness under different configurations. The experimental results, as shown in Fig. 8(b), quantified the bending stiffness by representing the ratio of deflection to the total length under a 70 g (≈0.7 N) load. It was determined that the HVHRI exhibited a bending stiffness of 5.1%-11.1%. This finding indicates that the HVHRI fulfills the design requirements and is at least 197% stiffer than the instrument developed by Hong et al., which has a bending stiffness of 33% with a larger diameter of 4 mm.

The torsional stiffness of the HVHRI was tested across various bending configurations of the flexible section. The testing setup, as shown in Fig. 9(a), comprised a rotation stage and a torque sensor (Nano 17-E, ATI Industrial Automation, USA). Each test involved bending the flexible section to angles of 0°, 45°, and 90°, while the rotation stage was utilized to twist the HVHRI to specific angles ranging from 0° to 10°. The torque sensor was used to measure the twisting torques exerted on the robot. The results, as shown in Fig. 9(b), revealed that when a torque load of 12 N·mm was applied, the twisting angles of the HVHRI in the three bending configurations were 10°, 9°, and 8°, respectively. For easier comparison with other instruments, the torsional

stiffness is represented as a twisting angle per flexible section length under 6 N·mm load. Then, we can obtain that the torsional stiffness of the HVHRI is 0.1 deg/mm, which fulfills the design requirement and is 150% stiffer than the flexible tool developed by Zhang et al., which has a 0.25 deg/mm torsional stiffness.

3.4.2. KINEMATICS TESTING

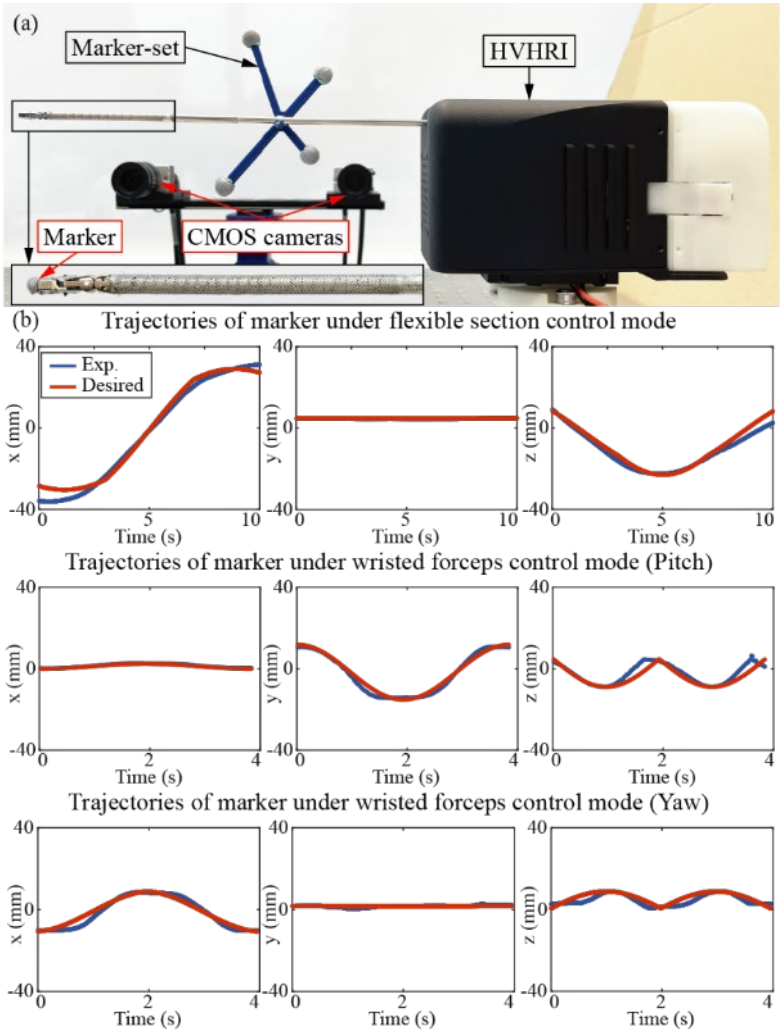


Fig. 10. Kinematics verification. (a) Experiment setup. (b) Measured and desired trajectories of the marker under two control modes.

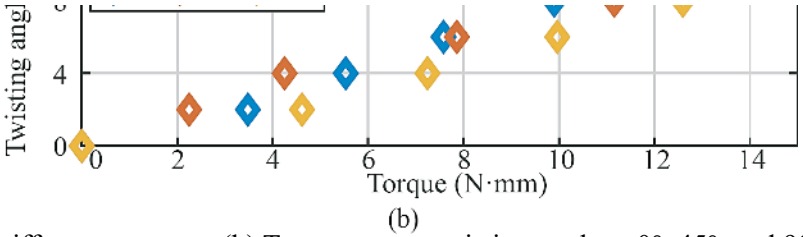


Fig. 9. (a) Torsional stiffness test setup. (b) Torque versus twisting angle at 0°, 45°, and 90° bending angles of the flexible section, which are marked in blue, red, and yellow, respectively.

In this experiment, we verified the kinematics of the HVHRI with the proposed two-mode control method. As shown in Fig. 10(a), the actuation unit with the detachable instrument of the HVHRI was fixed on the test bench. A stereo vision system including two CMOS cameras (DALSA Nano-m2420, Teledyne, Canada) (accuracy < 0.1 mm) was used to measure the 3D positions of various markers attached to the marker set, and the distal end of the wristed forceps. Four markers of the marker set formed a rigid body, which was fixed on the rigid shaft of the HVHRI to obtain the system's base coordinate.

As shown in Fig. 10(b), the position errors between the measured and desired trajectories of the marker were calculated for evaluating the kinematics model with the two-mode control method. Under the flexible section control mode, the 1-DOF flexible section was bent from -150° to 150°. The minimum error, the mean error, and the maximum error were 0.35 mm, 3.35 mm, and 7.26 mm, respectively. Then, under the wristed forceps control mode, the pitch and yaw joints of the wristed forceps were rotated from -90° to 90°, separately. The minimum

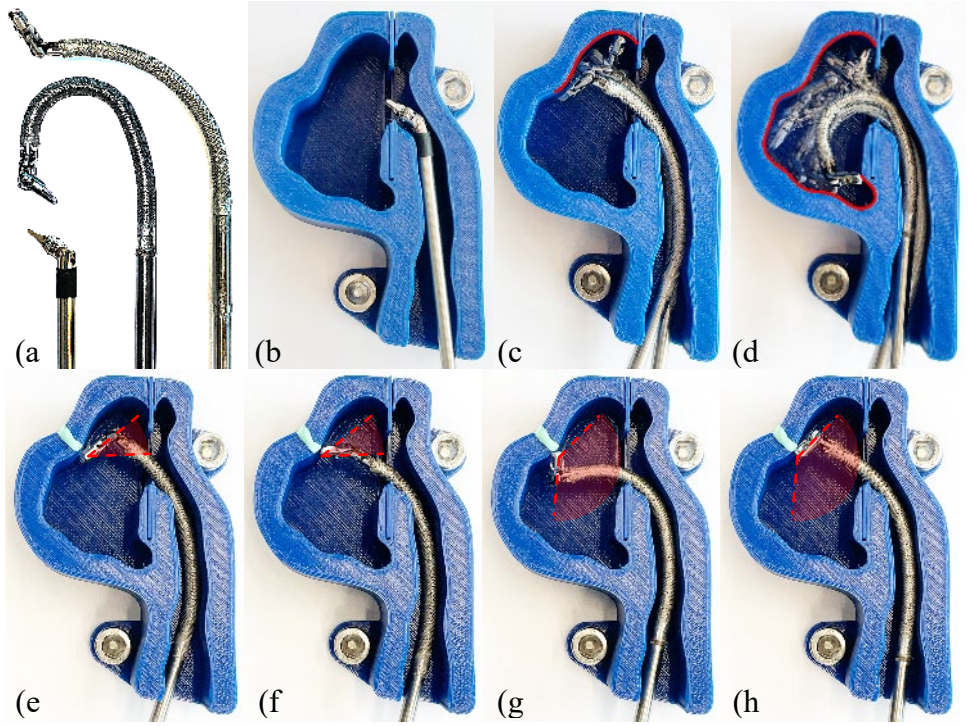


Fig. 11. (a) Prototypes of the WF (left), CHF (right), and HVHRI (middle). (b), (c), and (d) are the reachability test results of the WF, CHF, and HVHRI, respectively. CHF's reachable edge is marked yellow. HVHRI's reachable edge is marked red. (e) and (f) are the dexterity test results of the CHF, which has a yellow service region. (g) and (h) are the dexterity test results of the HVHRI, which has a red service region.

errors of the pitch and yaw motions were 0.25 mm and 0.11 mm. The mean errors of the pitch and yaw motions were 2.26 mm and 2.24 mm, respectively. And the maximum errors of the pitch and yaw motions were 7.07 mm and 5.34 mm, respectively.

3.4.3. REACHABILITY AND DEXTERITY COMPARISON

In this experiment, we compared the reachability and dexterity of the HVHRI with a pair of 3-DOF wristed forceps (WF) and a pair of 4-DOF CHF (see Fig. 11(a)) in a 3D-printed maxillary sinus and nasal cavity phantom, which was created based on a sectional CT scan data obtained in Section IV. Notably, all three instruments shared identical wristed forceps. Besides, the CHF featured a constant-curvature flexible section capable of achieving a maximum bending angle of 235°, which matched the capabilities of the HVHRI.

In the reachability test, an incision was made with a width of 10 mm between the nasal cavity and the maxillary sinus. All three instruments were operated to insert from the nasal entrance, pass through the incision, and try to touch the edge of the phantom's inner surface. In this 2D case, the reachability (R_{2D}) can be calculated as a ratio between the length of the reachable edge (L_{reach}) and the length of the entire edge (L_{all}), namely $R_{2D} = L_{reach}/L_{all}$. The test results are shown in Figs. 11(b)-(d). With a rigid shaft, the WF could not even reach the maxillary sinus, and therefore its reachability was 0. As the CHF was stuck in the nasal cavity to keep a constant curvature, its reachability was only 31%. The HVHRI had the largest reachability of 88%.

In the dexterity test, as the WF could not even reach the maxillary sinus cavity, only the CHF and the HVHRI were operated to insert from the nasal entrance, pass through the incision, and try to touch a point on the edge of the phantom's inner surface with different approaching angles. As shown in Figs. 11(e)-(f), the result showed that the CHF had a dexterity of 18% while the HVHRI had a dexterity of 69%.

3.4.4. USER STUDY FOR COMPARING THE WORKLOAD FOR DIFFERENT INTERFACES

As shown in Fig. 12, we designed three different handheld interfaces to conduct user study for comparing the workload. There were 7 male volunteers from the Multiscale Medical Robotics Centre, all right-handed, with normal or corrected to normal vision, and ages ranging from 24 to 31 years, participated the study. One of them had a medical background. The average collision number of style II (0.43 times) was fewer than the other two styles (style I: 1.71 times, style III: 1.57 times). The results of the NASA-TLX were shown in Fig. 13. The physical demand,
temporal demand, effort, and frustration of gripping style I were significantly lower than style II and III ($p < 0.05$). And the performance of style I was significantly better than style II and III.
The results indicated the gripping style II was the most suitable in these three styles.

3.4.5. FEASIBILITY TEST IN PHANTOM

To evaluate the feasibility of the HVHRI in the MSS, a realistic 3D maxillary sinus phantom (SN-as, PHACON GmbH., Germany) was employed to simulate polyps removal. As shown in Fig. 14 (c), two simulated polyps were placed in the anterior and posterior areas of the maxillary sinus cavity of the phantom. To visualize and guide the operation of the HVHRI inside the phantom, a hybrid-structure handheld robotic endoscope (HHRE) with robot-arm assistance was employed (see Fig. 14(a)). A surgical background researcher, who had taken a 15-minute training session on the operation of the HVHRI and the HHRE, performed the simulated polyps removal task. During the task, the researcher inserted the HVHRI and the HHRE together through the nasal cavity. Once the simulated polyps were shown in the endoscope's view, the position of the HHRE was fixed using the robot arm. Subsequently, the researcher focused on operating the HVHRI to grasp the simulated polyps. Ultimately, both the HVHRI and the HHRE were operated to extract the polyps from the phantom. The entire experiment was carried out three times. And it took an average of 130 s to locate, grasp, and retrieve the first simulated polyp and an average of 180 s for the second simulated polyp. As shown in Fig. 14(b), the results demonstrated that the HVHRI was effectively navigated through the 10 mm diameter surgical incision and successfully removed the two polyps from the maxillary sinus cavity.

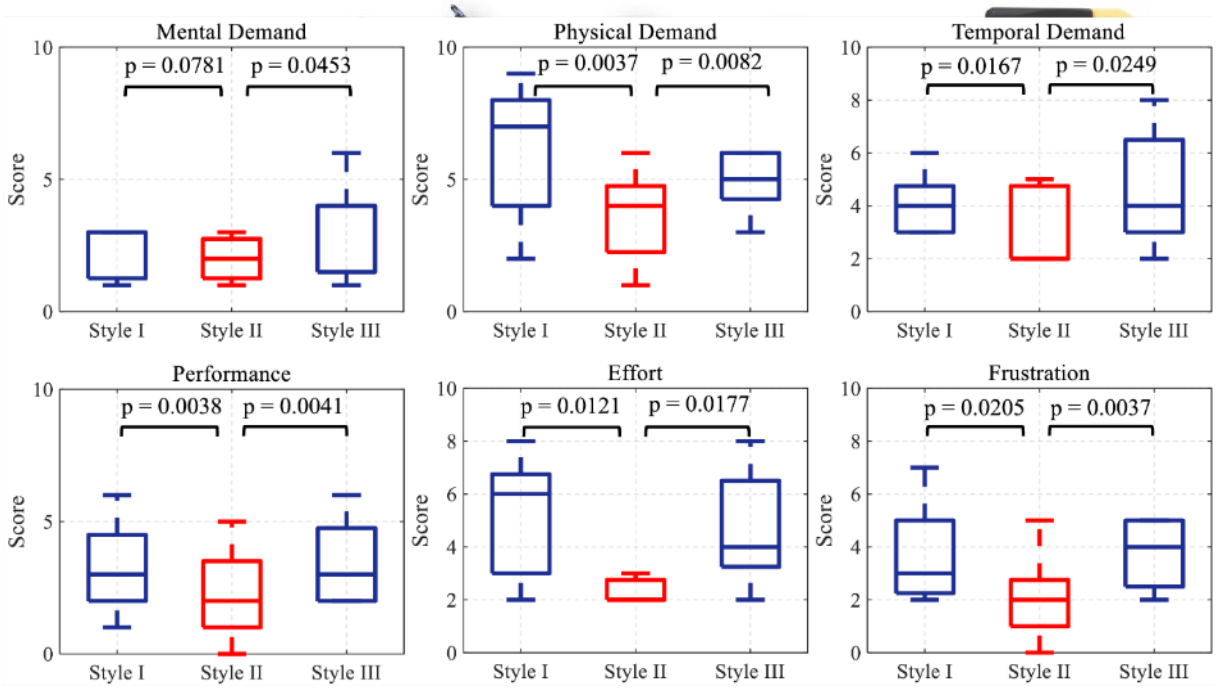


Fig. 13. The results of the NASA-TLX. The statistical analysis was conducted and the significant level was set at 0.05.

4. PUBLICATION AND AWARDS

J[1] Yi Yang, Puchen Zhu, Weibing Li, Richard M. Voyles, **Xin Ma*, "**A Fractional-Order Gradient Neural Solution to Time-Variant Quadratic Programming With Application to Robot Motion Planning", *IEEE Transactions on Industrial Electronics*, accepted, April, 2024.

J[2] Mei Liu, Kun Liu, Puchen Zhu, Guoqian Zhang, **Xin Ma***, and Mingsheng Shang*, "Data-Driven Remote Center of Cyclic Motion (RC2M) Control for Redundant Robots with Rod-Shaped End-Effector", *IEEE Transactions on Industrial Informatics*, accepted, February, 2024.

J[3] Xuchen Wang, **Xin Ma***, Puchen Zhu, Wee Shen Ng, Huayu Zhang, Xianfeng Xia, Russell H. Taylor, and Kwok Wai Samuel Au, "Design, Optimization, and Experimental Validation of a Handheld Nonconstant-Curvature Hybrid-Structure Robotic Instrument for Maxillary Sinus Surgery", *IEEE-ASME Transactions on Mechatronics,* accepted, 2024.

C[1] Yi Yang, Weibing Li*, Jianshu Zhou, Junda Huang, Jinfei Hu, Richard M. Voyles, **Xin Ma***, "PTC-FOZNN: A Strictly Predefined-Time Convergent Fractional-Order Recurrent Neural Network for Solving Time-Variant Quadratic Programming", *IEEE ICCA*, accepted, April, 2024.

C[2] Huayu Zhang, Jiajun An, Tianle Pan, Upinder Kaur, Zhijian Wang, Qiguang He*, **Xin Ma***, "Miniature Reconfigurable M odu lar Soft Robot s Using Liquid Crystal Elastomer Actuation", 6th International Conference on Reconfigurable Mechanisms and Robots, accepted, 2024.

C[3]Yi Yang, Puchen Zhu, Weibing Li, Richard M. Voyles, **Xin Ma***, "A Fractional-Order Recurrent Neural Network Model for Time-Variant Quadratic Programming in Robot Motion Planning", IEEE/ASME International Conference on Advanced Intelligent Mechatronics, accepted, 2024.

A[1] Internet + National Competition Gold Medal," 多層疊剛柔混合結構機械人系統", Instructor: **Xin MA** and Kwok Wai Samuel Au, December, 2023.

A[2] Third Prize, Professor Charles Kao Student Creativity Award, "應用於耳鼻喉手術中狹窄空間的掌上型混合結構的手術機器人系統", Instructor: **Xin MA** and Kwok Wai Samuel Au, July, 2023.

P[1] **Ma Xin**, Zhang Zihao, Wang Xuchen, Kwok Wai Samuel Au, "Cam-based Backend Transmission Mechanism for Driving Four-cable Wristed Instrument", 2023-8-7, US Patent Pending, US 63/531,314.

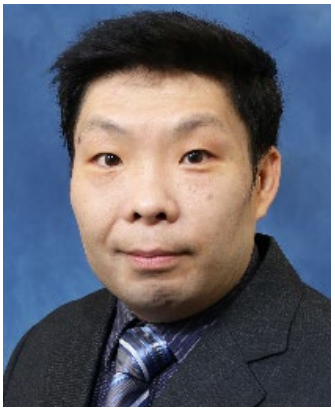
WORK IN PROGRESS

SMART BANDAGE WITH INTEGRATED ORGANIC ELECTRONIC SENSOR AND IONTRONIC DRUG DELIVERY PLATFORM FOR ADVANCED CHRONIC WOUND CARE

Principal Investigator: Professor MAK Wing Cheung
Department of Biomedical Engineering, CUHK

Research Team Members:
Junning QIAN, PostDoc[(1)]

[(1)] Department of Biomedical Engineering, CUHK



Reporting Period: 1 July 2023 – 31 May 2024

INNOVATION AND PRACTICAL SIGNIFICANCE:

"Theranostics" are emerging biomedical strategies that integrate therapeutic and diagnostic/biosensing components for the monitoring and management of chronic clinical conditions. From the innovative technology aspect, our project aims to develop a novel wearable theragnostic smart bandage platform with integrated sensors and drug delivery patches to monitor and actively react to modulate wound healing progress revolutionizing existing chronic wound addressing technologies with improved patient healthcare outcomes. From the material science engineering aspect, our project explores the unique properties of conductive polymers with mixed electronic and ionic conductivities as organic electronic platforms to perform both the electrochemical biosensing and ionic drug release feature, with additional advantages of being flexible, organic in nature, processable for large-scale fabrication and biocompatible, which has not been demonstrated in theragnostic smart bandage platform. The success of this project would open new prospects for the next generation of smart bandages with sensing and therapeutic features for the betterment of wound management for pressure ulcers (disabled person/neuromuscular diseases/stroke patients with long-term beds), diabetic foot ulcers (diabetes and elderly) ischemic ulcers (elderly) and venous ulcers (elderly). In fact, Asia is the hotspot for diabetes and facing huge challenge in ageing population. Therefore, we expect the new theranostics smart bandage developed in this project could offer new healthcare tools for the betterment of mankind in Hong Kong, Greater Bay Area, China and beyond. Moreover, as the skin is the largest organ in the body, the developed smart bandage platform may potentially as new wearable biomedical tools for monitoring and treatment of other skin diseases.

ABSTRACT

Chronic wound management is a major healthcare challenge that affects millions of patients (1.5-2.0% of the world population) suffering from slow healing processes, susceptibility to infection and risk of amputation, together with financial burdens to central healthcare system. Incidents are expected to rise with ageing population and increase in diabetic cases (i.e. diabetic foot ulcers). Emerging smart bandages with integrated sensors have demonstrated to support monitoring of wound healing progress with improved healthcare outcomes, yet there is limited development in integrating both sensor and on-demand drug delivery for wound modulation. Our project aims to develop a novel closed-loop smart bandage platform with integrated electronic sensor to track the wound conditions and iontronic on-demand drug delivery patch for modulation of infection incidents at early stage. We plan to develop wearable organic electronics with conducting polymers feature with unique mixed electronic-and-ionic conductivity properties for fast-response and reversible pH monitoring (indicator for infection), and iontronic-controlled release of antibiotics via doping/de-doping mechanism to modulate infections. The success of this project would open new prospects for next generation of smart bandages with sensing and therapeutic features for better wound management,

reducing the frequency of dressing replacement and visits to medical facilities with improved patient compliance.

1. OBJECTIVES AND SIGNIFICANCE

Objectives:
1. To develop a conducting polymer-based wearable platform for selective pH sensing and optimize the pH response with a wide pH window covering pH 4-10. (Completed on schedule)
2. To achieve high performance in fast pH response, reversible pH monitoring and long-term stability based on the smart bandage. (Completed on schedule)
3. To develop and optimize conducting polymer iontronic delivery patches and demonstrate the electrochemical triggered release. (On Track, Continue in 2nd year)
4. Integrate both pH monitoring and drug delivery (On Track, Continue in 2nd year)

Significance:
The outcome of the project will immediately impact the field of smart bandages with innovative sensing and therapeutic features for the advancement of chronic wound management which remains a global challenge affecting millions of patients. Moreover, chronic wounds are often associated with the ageing population and diabetes communities with Asia as the hotspot (including Hong Kong), and the developed smart bandages could potentially benefit and support the long-term healthcare system. Furthermore, as the skin is the largest body organ, the developed smart bandage platform may further innovate as new wearable biomedical tools for monitoring and treatment of other skin diseases.

For the first step, measuring the pH of wound exudate is the most important indicator for assessing bacterial infection and monitoring the wound healing progress. In the 1st phase of the project, we have successfully developed a PANI-based wearable pH sensor on a wearable and breathable fabric substrate tailored for the smart bandage application. The performance of the pH sensor delivered good sensitivity, wide pH sensing range (4- 10), good reversibility for pH cycling, fast response time (~1sec) and good stability over 12 hours of pH measurement. The details of the achievements are presented in Section 3.

2. RESEARCH METHODOLOGY

2.1 Materials and instruments

Screen-printed electrodes (SPEs, DropSens 110) were purchased directly from Metrohm. Polyimide film (PI Film: thickness of 75 μm) was purchased from ubisoft tape Co., Ltd. Polyimide fabric (PI Fabric: density of 60 g·cm-2, W-60) was purchased from Jiangsu xiannuo new materials technology Co., Ltd. Aniline and hydrochloric acid (HCl) were purchased from Macklin and Sigma-Aldrich, respectively. Solid-state silver/silver chloride (Ag/AgCl) was purchased from Dupont. All chemicals were of analytical grade and used without any further purification. Computer-controlled CO2 laser platform (HL 40-5g, 10.6 μm, Full Spectrum Laser LLC) was used to produce laser-induced graphene (LIG), which was controlled by the software Lightburn. Electrochemical workstation (PalmSens4, Europe) was used for electrochemical tests. The microstructure of electrodes was observed on a optical microscope (Axioscope, Zeiss, Germany). A commercial pH meter (FiveEasy Plus pH meter FP20-Std-Kit, Mettler Toledo, Switzerland) was used as a standard for pH measurement.

2.2 Fabrication of LIG:

LIG was prepared in different patterns using a commercial 10.6 μm CO2 laser cutter system (Gweike Cloud RF, China) on both PI Film and PI Fabric under ambient conditions. The resulting LIG on two substrates were noted as PI-Film-LIG and PI-Fabric-LIG, respectively. The laser parameters were optimized to be 8.5% of the full laser power (full power is 60 W), with a scan speed of 170 mm·s-1 and a defocus distance of 15 mm.

2.3 Electrodeposition of polyaniline (PANI) on LIG:

Firstly, a three-electrode system was used for PANI preparation. Pt wire and Ag/AgCl electrode were used as counter electrode and reference electrode, accordingly. SPEs, PI-Film-LIG, and PI-Fabric-LIG were used as

the working electrodes, respectively. The electrodeposition of PANI was carried out at room temperature. This three-electrode system was immersed in the 1 M HCl solution containing 0.1 M aniline. The electrochemical polymerization of aniline was then carried out using the cyclic voltammetry (CV) method in the potential range from -0.2 to 1 V vs. Ag/AgCl with a scanning rate of 50 mV·s-1. The number of CV cycles was five. After the electrochemical polymerization was completed, the prepared working electrodes were washed with deionized water and alcohol, and then dried under 60 °C for further use.

2.4 Fabrication and integration of a wearable pH sensor on bandage:
The two-electrode system was employed as the pH sensor based on the PI-Fabric-LIG. The working electrode was the as-prepared PANI decorated PI-Fabric-LIG, and the reference electrode was a solid-state Ag/AgCl electrode. This reference electrode was stencil printed with Ag/AgCl ink on the surface of PI-Fabric-LIG and dried under 60 °C for 1 hour. Then 5 µL of PVB polymer containing KCl fine powder (30% w/v) was dropped cast on top of the Ag/AgCl using a mask, followed by importing the reference electrode in 3 M KCl for 24 hours. After that, the surface of the reference electrode was coated with 5 µL of PVB polymer to further improve the stability of the fabricated pH sensor. Finally, the wearable pH sensor was assembled on the bandage by fixing the pH sensor layer between the lower dressing and the upper adhesive layers.
2.5 Electrochemical tests:
The response of the pH sensors was recorded and characterized by measuring the electromotive force (EMF) between the working electrode and the Ag/AgCl reference electrode in 0.1 M PBS with varied pH of 4-10.

3. RESULTS ACHIEVED SO FAR

We have successfully completed the content on building a conducting polymer-based wearable platform in objective 1. As shown in Figure 1, the conventional substrates such as screen-printed electrodes (SPE) are thick, rigid, non-porous and non-breathable, while the flexible PI films are flexible, however, the compacted PI film is a non-porous and non-breathable substrate which is not an ideal substrate material for bandage application. In contrast, we developed a PANI-based pH sensor on PI fabric that is flexible, thin, comfortable, porous and breathable suitable for wearable bandage devices. Additionally, the electrodeposition technique employed for synthesizing the conducting polymer-PANI pH sensing layer meticulously retains the porous morphology and integrity of the PI fabric substrates and their inherent architecture. As noted, the morphological characteristics of various substrates remain predominantly unaltered after the electrodeposition of PANI.

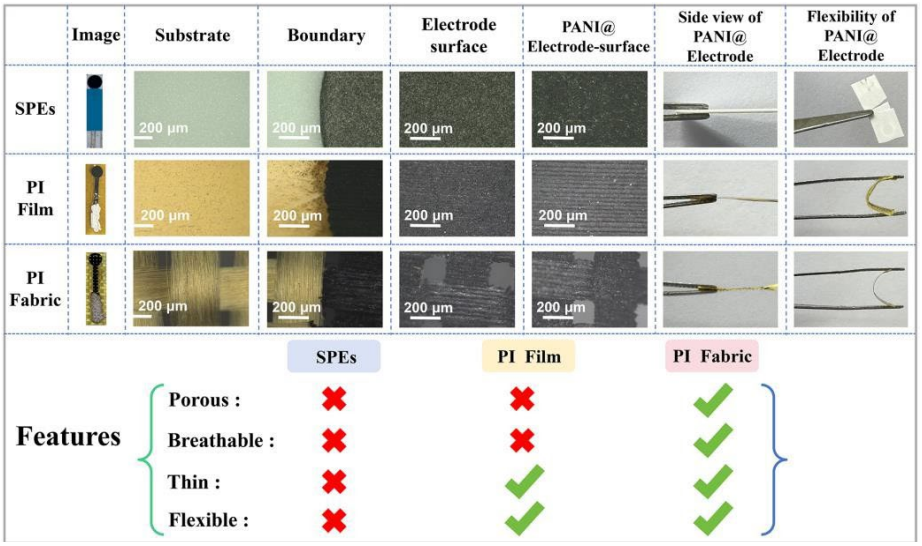


Figure 1. Microstructure, optical images and structural features of the SPEs, PI-film, PI-fabric based electrodes.

Next, we completed the optimization of the pH response with a wide pH window covering pH 4-10 in objective1.1.1. By analyzing the electrochemical characteristics and pH sensing performance of

different electrode platforms (Figure 2a-b), the platform of PI fabric decorated with PANI exhibits the best sensitivity of 70.592 mV/pH with a linear pH range covering pH 4-10 which is outperformed compared with other conventional platforms (Figure 2c-d).

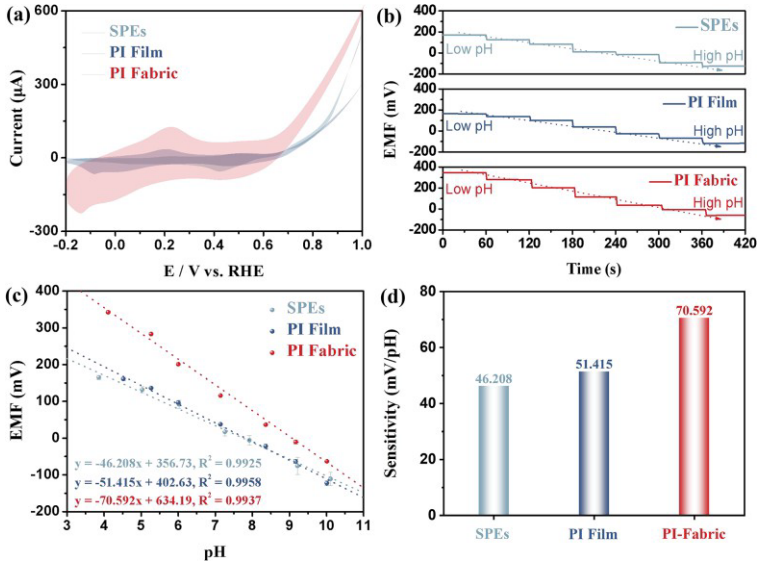


Figure 2. (a) CV curves of PANI decorated SPEs, PI Film and PI Fabric; (b) Potentiometric curves of EMF response stepwise increasing with pH 4-10; Corresponding calibration plots (c) and sensitivity values (d) of EMF versus pH.

We have achieved to develop a high-performance pH sensor of PI fabric with fast pH response, reversible pH monitoring and long-term stability for smart bandage in objective 1.1.2. The PANI-fabric pH sensor is porous, breathable, thin (0.06 mm), flexible and lightweight (6.38 mg·cm^{-2}) (Figure 3a-f). The integrated smart bandage comprises three components: dressing material, PANI-fabric pH sensor, and adhesive unit (Figure 3g), showcasing comfortable wearable features on the skin (Figure 3h). The PANI-fabric pH sensor showed a fast response to pH change within one second (Figure 3i). The pH sensitivity of the PANI-fabric pH sensor reaches 65.279 mV/pH with a linear pH range covering pH 4-10 (Figure 3i-j). The high sensitivity and wide pH sensing window of the PANI-fabric pH sensor surpasses most literature-reported pH sensors. [1-3] After completing the cycle of pH sensing from 4-10-4, it can retain 98.6% of its initial EMF value (Figure 3k). During a 12-hour stability test, the PANI-fabric pH sensor maintains the original signal of 95.1%, 93.7% and 96.9% compared with the starting points in pH values of 4, 7 and 10, accordingly (Figure 3l). These results demonstrated the excellent stability of the PANI-fabric pH sensor compared to most literature-reported pH sensors only showing a relatively short stability duration, typically within 1 hour. [4-6] Furthermore, after a pH value mutation from 10 to 4, the PANI-fabric pH sensor maintains 99.3% of the initial EMF value (Figure 3m) showing a high repeatability in pH sensing.

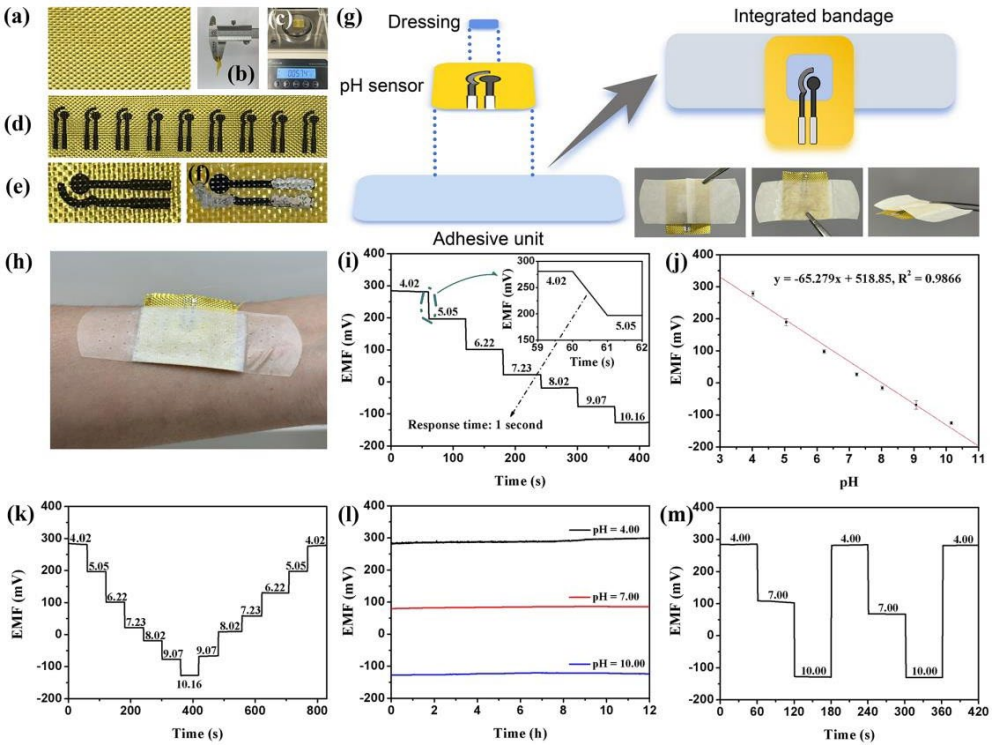


Figure 3. (a) Image, (b) thickness and (c) weight of PI Fabric; (d) Scalable synthesis of PI-Fabric-LIG; (e) enlarged area in Figure 3d; (f) Image of integration of PI-Fabric-LIG with solid-state Ag/AgCl, PANI and silver ink; (g) Schematic of craft for smart bandage and physical images from different perspectives; (h) wearability of smart bandage; (i) Potentiometric curve of EMF response stepwise increasing with pH 4.02-10.16, insert: enlarged area from pH values of 4.02-5.05; (j) Corresponding calibration plot of EMF versus pH; (k) Reversibility of smart bandage to stepwise increase pH value from 4 to 10 followed by stepwise decreasing from 10 to 4; (l) Long-term stability of smart bandage at pH values of 4, 7 and 10; (m) Repeatability of smart bandage at pH values of 4, 7 and 10.

Summary:

We have completed the development of the pH sensor on a breathable fabric substrate for smart bandage with good sensitivity, wide pH sensing range (4-10), good reversibility for pH cycling, fast response time (~1sec) and good stability over 12 hours pH measurement (completed on schedule). We will continue in the 2nd phase of the project to develop the conducting polymer iontronic delivery patches and finally integrate the smart bandage with the pH sensor and drug delivery patch (to be continued).

References

[1] X.Q. Cui, et. al, A wearable electrochemical sensor based on β-CD functionalized graphene for pH and potassium ion analysis in sweat, Talanta, 2022, 245, 123481.

[2] J.H.Yoon, et. al., Highly self-healable and flexible cable-type pH sensors for real-time monitoring of human fluids, Biosensors and Bioelectronics, 2020, 150, 111946.

[3] C.H. Zhu, et. al., A dual-functional polyaniline film-based flexible electrochemical sensor for the detection of pH and lactate in sweat of the human body, Talanta, 2022, 242, 123289.

[4] K. Chawang, et. al., Porous polypropylene membrane based pH sensing for skin monitoring," IEEE Access, 2022, 10, 111675-111687.

[5] K. Singh, et. al., Impact of yttrium concentration on structural characteristics and pH sensing properties of sol-gel derived Y2O3 based electrolyte-insulator-semiconductor sensor," Materials Science in Semiconductor Processing, 2020, 105, 104.

[6] L.R. Wang, et. al., Skin-like hydrogel-elastomer based electrochemical device for comfortable wearable biofluid monitoring,
Chemical Engineering Journal, 2023, 455, 140609.

WORK IN PROGRESS

COUPLING MOS2 FIELD-EFFECT BIOSENSORS WITH HYBRIDIZATION CHAIN REACTION SELF-ASSEMBLY AMPLIFICATION FOR HIGHLY SENSITIVE AND LABEL-FREE NUCLEIC ACID DETECTION

Principal Investigator: Professor Zhaoli GAO
Department of Biomedical Engineering, CUHK

Research Team Members:
Dong Wook JANG [(1)]
Jiawen YOU [(1)]

[(1)] Dept. of Biomedical Engineering, the Chinese University of Hong Kong

Reporting Period: 1st July 2022 – 31st May 2023
(to be completed in August 2024)

INNOVATION AND PRACTICAL SIGNIFICANCE:

The objective of this project is to develop a highly sensitive and label-free platform for nucleic acid detection, based on MoS2 transistor arrays coupled with an HCR-based signal amplification scheme. The proposed project is novel and highly notable for multiple reasons: 1) The application of large-scale MoS2 transistor arrays in biosensing, which has not yet been reported, could result in a substantial enhancement in the limit of detection, compared to conventional 2D devices without a band gap, e.g., graphene-based devices; 2) Scalable fabrication of biosensor arrays based on MoS2 has yet to be achieved; and 3) Involving signal amplification in MoS2-based biosensors is a novel methodology for nucleic acid detection. Project success would bring accuracy to nucleic acid-based diagnosis, which is critically needed for the management of various diseases, such as early-stage cancers, and viral infectious diseases. We expect the all-electronic biosensors developed through this project would offer a diagnostic platform that would be widely applicable in the healthcare system of Hong Kong and that of the world over.

ABSTRACT

Nucleic acid testing plays a crucial role in clinical diagnoses. Existing approaches, such as polymerase chain reaction (PCR), are sensitive to inhibitors due to their enzyme application and fluorescent labeling. The past decade has witnessed the development of 2D bioelectronics for highly sensitive testing of biomolecules, such as nucleic acid associated with various diseases. 2D nanomaterials hold tremendous promise for low-level and label-free detection of nucleic acids, offering the prospects of simple and multiplexed testing, with high accuracy and specificity. However, the zero-bandgap nature of traditional 2D nanomaterials, such as graphene, and the binding-affinity-dependent limit of nucleic acid detection inhibit further advancement in the label-free detection of ultralow-level nucleic acid. The objective of our proposal is to develop a highly sensitive and label-free biosensing system for nucleic acid detection, through the combined use of gapped monolayer MoS_2 and hybridization chain reaction-based signal amplification scheme. The project success would open a new prospect for next generation ultrasensitive MoS2-field-effect-transistor (MoS2-FET) biosensors that could potentially exceed the detection limit of graphene-FET.

1. OBJECTIVES AND SIGNIFICANCE

1) To synthesize monolayer MoS_2 with controlled uniformity on a large scale, through chemical vapor deposition
2) To develop scalable fabrication of MoS_2-FET sensor arrays functionalized with hairpin probe DNA
3) To develop a target recycling and hybridization chain reaction (TRHCR) scheme that can work on MoS_2-FET devices, for sensitive detection of nucleic acid at aM levels

4) To test and optimize the sensing protocols with MoS_2-FET biosensors for maximum sensitivity, reproducibility, and reliability in physiological conditions

The proposed project is novel and highly notable for multiple reasons: 1) The application of large- scale MoS_2 transistor arrays in biosensing, which has not yet been reported, could result in a substantial enhancement in the limit of detection, compared to conventional 2D devices without a band gap, e.g., graphene-based devices; 2) Scalable fabrication of biosensor arrays based on MoS_2 has yet to be achieved; and 3) Involving signal amplification in MoS_2-based biosensors is a novel methodology for nucleic acid detection. Project success would bring accuracy to nucleic acid-based diagnosis, which is critically needed for the management of various diseases, such as early-stage cancers, and viral infectious diseases. We expect the all-electronic biosensors developed through this project would offer a diagnostic platform that would be widely applicable in the healthcare system of Hong Kong and that of the world over.

1. RESEARCH METHODOLOGY

Task 1: Controlled large-scale growth of MoS_2

We intend to develop a feasible CVD method to grow uniform MoS_2 thin films with precise control over their layer number and crystal orientations. A Si/SiO_2 growth substrate will be coated with a layer of NaCl (1% aqueous solution), followed by spin coating of saturated ammonia heptamolybdate (AHM) solution as the Mo source and being heated to ~ 800 °C. A sulfur source will then be placed upstream of the furnace, and a flow of 500 sccm Ar will carry the sulfur vapor into the chamber for a 30-60 min growth period. After growth, the sample will be rapidly cooled to room temperature.

Our hypothesis is that CVD growth could lead to the controlled synthesis of a uniform MoS_2 monolayer across the entire material. The feasibility of our proposal was verified by trial growths of triangular MoS_2 crystals on SiO_2 substrates, and we plan to achieve more precise control, with further exploration of the governing mechanism of CVD conditions on MoS_2 growth.

Task 2: Scalable fabrication and functionalization of MoS_2 sensor arrays

We will carry out photolithography to fabricate the MoS_2 biosensor arrays in a scalable way. A sub-monolayer of gold will be deposited onto the MoS_2 by physical vapor evaporation to a nominal thickness of 2 Å. The MoS_2-Au heterostructure will be transferred onto a SiO_2 substrate with prefabricated Cr/Au electrodes. Fabrication of the MoS_2 transistor arrays will be implemented with the use of photolithography to pattern the transferred MoS_2, followed by plasma etching. The sensor array will be annealed at 225 °C to allow the formation of monodispersed AuNPs. An Al_2O_3 passivation layer will be deposited *via* atomic layer deposition to prevent leakage current from the source to drain. Finally, the arrays of MoS_2-AuNP sensors will be functionalized with thiolated hairpin probe DNA through Au−S bonds.

Task 3: MoS_2-FET performance testing and characterization

We will design, develop, and validate a TRHCR scheme, to quantify DNA using the MoS_2 sensing platform with accuracy and precision. We will verify the improved LOD and enhanced specificity of TRHCR with commercially available DNA (ThermoFisher Scientific), using four different concentrations (in the clinically relevant range of 10 fM, 1fM, 100 aM, and 10 aM), each in 1× SSC buffer solutions. We will also examine whether the biosensor responses reflect specific binding of the target DNA, by testing the MoS_2 biosensor against various negative controls, i.e., target DNA with a single-base mismatch at the 3′ or 5′ end.

2. RESULTS ACHIEVED SO FAR

Result 1: CVD growth of high-density 2H-phase monolayer MoS_2

We have investigated the epitaxial CVD growth of MoS_2 through the assistance of NaCl halide salt as the promotor species. Previous studies have shown that NaCl can react with molybdenum precursor to form molten Na-Mo-O ($Na_2Mo_2O_7$) and volatile oxyhalides (MoO_2Cl_2). This reaction has been found to foster the epitaxial growth of MoS_2 in several ways, such as by increasing the vapor pressure of the molybdenum feedstock through the high diffusivity of molybdenum oxyhalide species and by reducing the activation energy needed for MoS_2 formation [1]. As demonstrated in our results (see **Fig. 1**) and consistent with previously reported reports [2-4], incorporating NaCl into the reaction system enhances the overall monolayer flake size of MoS_2 and inhibits its layered growth.

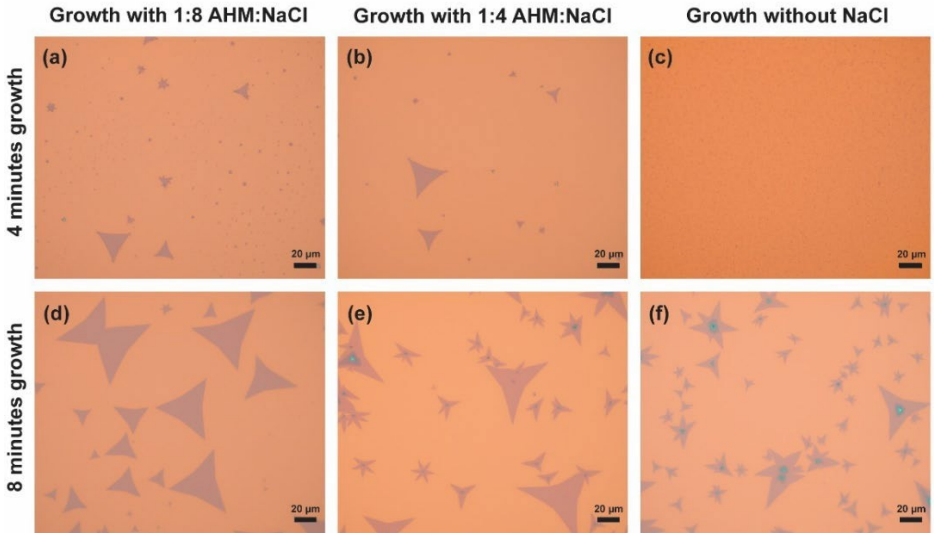


Figure 1. Efficacy of halide salt on the epitaxial growth of monolayer MoS_2 assisted by NaCl in the CVD process. The diagram illustrates the impact of halide salt on the growth of monolayer MoS_2. AHM is an abbreviation for ammonium heptamolybdate, which is the molybdenum precursor used in our CVD growth process. The figures display the growth results when different ratios of AHM to NaCl are used for growth periods of (a-c) 4 minutes and (d-f) 8 minutes.

By further fine tunning the recipe, we have now been able to grow dense monolayer flakes with average flake size of approximately 40 µm (**Fig. 2**). Large and dense (preferably continuous) monolayer MoS_2 flakes are crucial to produce FET devices with high yield and performance.

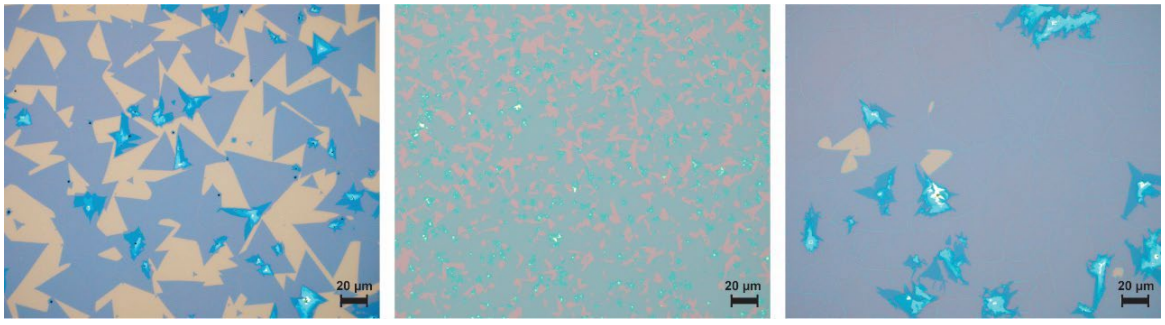


Figure 2. Halide salt assisted CVD growth of dense monolayer MoS_2. Figure above shows the dense growth of monolayer MoS_2 flake.

To further verify the quality of the as-grown material, we used Raman spectroscopy and atomic force microscopy (AFM) as the primary characterization methods. The Raman results, shown in **Fig. 3a**, reveal the peak positions of E_{2g} and A_{1g} (activation modes at 384 cm^{-1} and 403 cm^{-1}, respectively), with a peak difference of 19 cm^{-1}. This confirms the characteristic peak of monolayer MoS_2. Furthermore, the absence of the 1T-phase characteristic

peaks at the J1, J2, J3 acoustic modes, as well as a large full-width-half-maximum (FWHM) of the E2g peak, indicate the semiconducting 2H crystalline phase nature of the monolayer MoS2. Consistent with the Raman result, the AFM data in **Fig. 3b** further confirms the monolayer nature of the as-grown MoS2, revealing a thickness profile of roughly 0.76 nm.

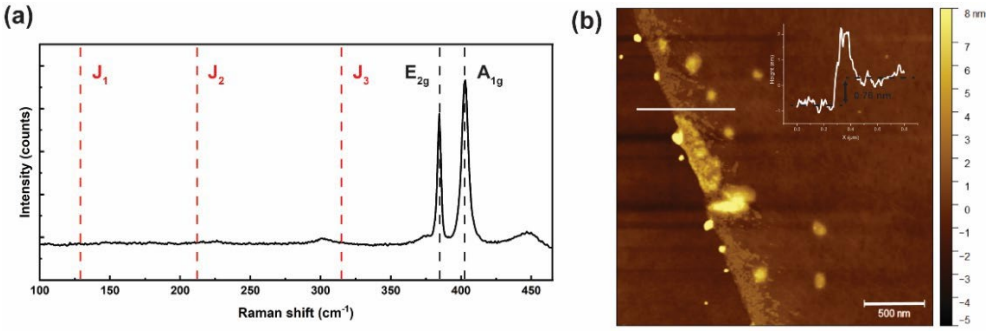


Figure 3. (a) Raman and (b) AFM characterization of as-grown monolayer MoS2 by CVD Result 2:

Fabrication of MoS2-FET and analysis of transfer curve measurements

We have fabricated FET devices based on our CVD-grown MoS2. **Fig. 4a** illustrates the design of the FET electrode, with the channel size tailored to the typical flake size of the monolayer MoS2. **Fig. 4b** shows the transfer curve measurement results for the MoS2-FET. The device revealed an on-off ratio of approximately 6300 and a transconductance of 4.32 x 10^{-10} S under 0.1V biased, with a calculated mobility of μ=0.826 cm^2/Vs.

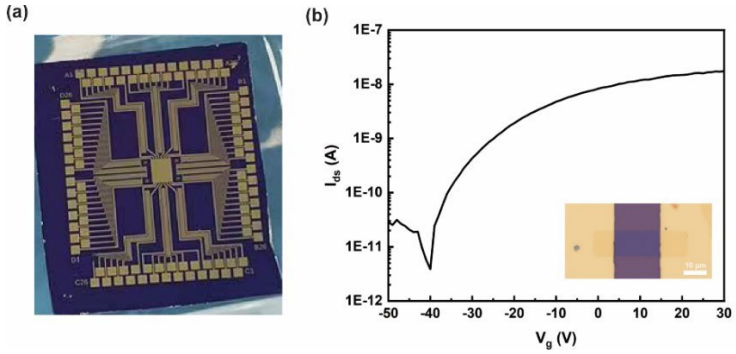


Figure 4. MoS2-FET array and transfer curve measurement. The figure above displays (a) the design of our MoS2-FET electrode, tailored to fit the typical flake size of the monolayer MoS2, and
(b) the transfer curve measurement obtained from an as-fabricated FET chip following the design outlined in panel a.

3. REFERENCES

1. Lei, J., et al., *Salt-Assisted MoS2 Growth: Molecular Mechanisms from the First Principles.* Journal of the American Chemical Society, 2022. **144**(16): p. 7497-7503.
2. Singh, A., et al., *NaCl-assisted substrate dependent 2D planar nucleated growth of MoS2.* Applied Surface Science, 2021. **538**: p. 148201.
3. Chen, L., et al., *Study on the catalyst effect of NaCl on MoS2 growth in a chemical vapor deposition process.* CrystEngComm, 2021. **23**(31): p. 5337-5344.
4. Zhang, L., et al., *Damage-free and rapid transfer of CVD-grown two-dimensional transition metal dichalcogenides by dissolving sacrificial water-soluble layers.* Nanoscale, 2017. **9**(48): p. 19124-19130.

COMPLETED

ENDOLUMINAL IMAGE-BASED MOTION OPTIMIZATION AND MOTION UNDERSTANDING (EIMOMU) TOWARDS SEMI-SUPERVISED STIFFNESS TUNABLE FLEXIBLE ROBOTIC PROCEDURES

Principal Investigator: Professor REN Hongliang
Department of Electronic Engineering, CUHK

Research Team Members:
Mengya Xu [(1)],
An Wang [(1)],

[(1)] Dept. of Electronic Engineering, CUHK

Project Start Date: 1st July 2020
Completion Date: 6 October 2023

INNOVATION AND PRACTICAL SIGNIFICANCE:

The novel aspects of this project include:

A new multi-curvature high-DOF (degrees of freedom) robotic endoscopy system (>5 DOFs, fit in existing sub-centimeter endoscopic instrument channels) that simultaneously incorporates decoupled motion optimization (Aim 1) whilst varying appropriate stiffness (Aim 2) within a confined surgical endoluminal space. The new portable system will enable dexterous and modular interchangeable endoscopic manipulations without changing the current surgical workflows.

A first image-based end-to-end deep neural network of motion understanding (Aim 3) for endoluminal surgical procedures. This technology is a first of its kind for annotating robot-assisted endoscopic surgery scenes involving stiff-tunable flexible instruments, enabling expert-level real-time surgical monitoring and guidance to safeguard the critical and labor-intensive surgical procedures. In addition to intraoperative supervision, it can be directly deployed for document entry tasks, surgical training, and post-operative analysis.

A new image-based data-driven visual guidance module simultaneously estimates the relative poses (position and orientation) of the flexible end-effectors and involuntary moving surgical targets. Incorporating the semantic motion understanding (Aim 3), image-based visual guiding algorithm will drive the flexible manipulator towards dynamic targets along the optimally planned trajectories employing optimally tuned motion kinematics (Aim 1) real-time, which can ensure easy manipulation and best endoscopic view-angle configuration.

In the long run, the investigated technologies will potentially motivate and benefit many other surgical procedures through natural orifices or other transluminal techniques. The extension of this proposed Endoluminal Image-based Motion Optimization and Motion Understanding (EIMOMU) framework to access the confined cavity via natural orifices, e.g., mouth, anus or vagina, could affect hundreds of thousands of people annually. Hence, the proposed research topics have broad social, clinical and biomedical industrial impacts.

ABSTRACT

We develop the ambient clinical intelligence system, aiming at surgical motion understanding in robot-assisted Minimally Invasive Surgery (MIS). We contribute to intraoperative surgical decision-making and automate the creation of reports during patient treatment. Current deep learning models degrade the performance when applied to different domains with domain shifts and new classes. Thus, we also design the class incremental domain adaptation to overcome this limitation.

1. OBJECTIVES AND SIGNIFICANCE

Objective:

To develop the ambient clinical intelligence system, an Artificial Intelligence (AI)-powered, Natural Language Processing - enabled solution aimed at surgical motion understanding. It can automatically recognize the surgical interaction, generate surgical instructions, and document patient encounters accurately and efficiently during the surgery.

Significance:

The ambient clinical intelligence solution, aimed at surgical motion understanding in robot-assisted Minimally Invasive Surgery (MIS), is critical for context-aware decision support, surgical path planning, documenting patient treatment, evaluating surgical quality, tracking patient care, and surgical training. Complete and accurate clinical documentation enables treatment plans to be exchanged throughout care teams, ensuring continuity of care and a transparent and efficient reimbursement procedure.

2. RESEARCH METHODOLOGY

2.1 Surgical Interaction Recognition

A fine-grained Surgical Interaction Recognition Network (SIRNet) is proposed with the goal of predicting surgical interaction triplets. The projected SIRNet includes a multi-head self-attention mechanism that learns the relationships between surgical interaction triplets without defining them first, whereas a multi-head cross-attention mechanism learns the relationships between endoscopic pictures and each triplet. The bipartite matching loss, which takes into consideration the permutation and combination of tools, verbs, and targets, is utilized to make proper learning and prediction for each component in the surgical interaction triplet. Furthermore, a weight attention module is created to weigh the relevance of each predicted surgical interaction triplet and each component in the triplet when predicting final valid surgical interaction triplets. Many state-of-the-art classification models, including ViT and Swin Transformer, are compared to the proposed SIRNet.

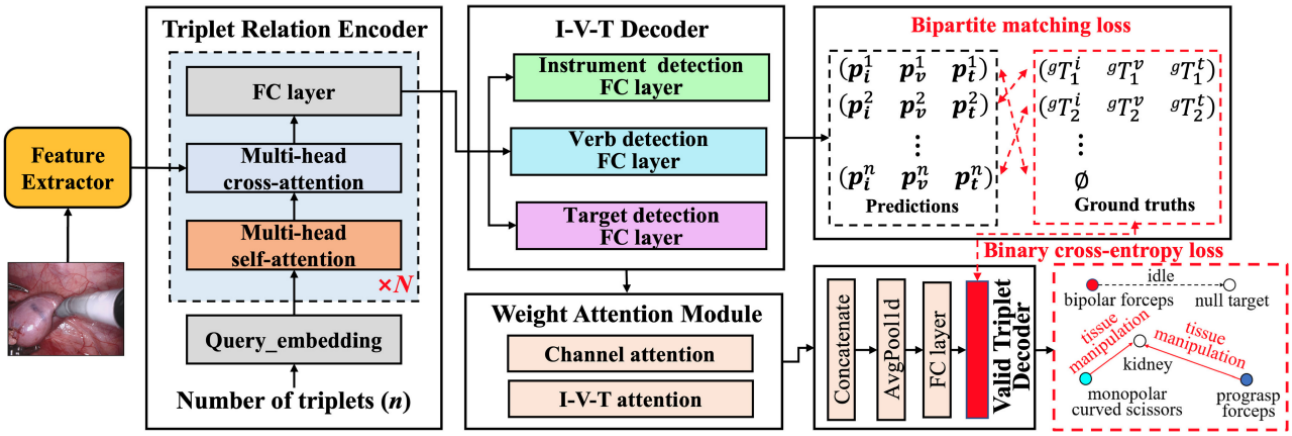


Fig. 1. The architecture of SIRNet, from J[1]. A triplet relation encoder uses multi-head self-attention to learn the relationships between each surgical interaction triplet and multi-head cross-attention to learn the relationships between picture characteristics and each triplet. An I-V-T decoder is presented to recognize instruments, verbs, and targets simultaneously without sorting triplet labels, and bipartite matching loss is used to oversee set prediction outcomes. The weight attention module calculates the relative relevance of each queried triplet and its components. The final surgical interaction triplets are output using a valid triplet decoder.

2.2 Documenting Surgical Procedures

Medical documentation requirements have increased, and this load has been identified as one of the key contributors to physician burnout. In robot-assisted surgery, creating surgical reports targeted at understanding the surgical environment can help with documenting entry duties and post-operative analyses. Current deep learning

models have compromised performance when applied to diverse domains with domain changes. In addition, new equipment and surgical tissue variations have developed in robotic surgery. We suggest using a multi-layer transformer-based approach to tackle new classes and domain shifts in the target domain to provide surgical reports during robotic surgery.

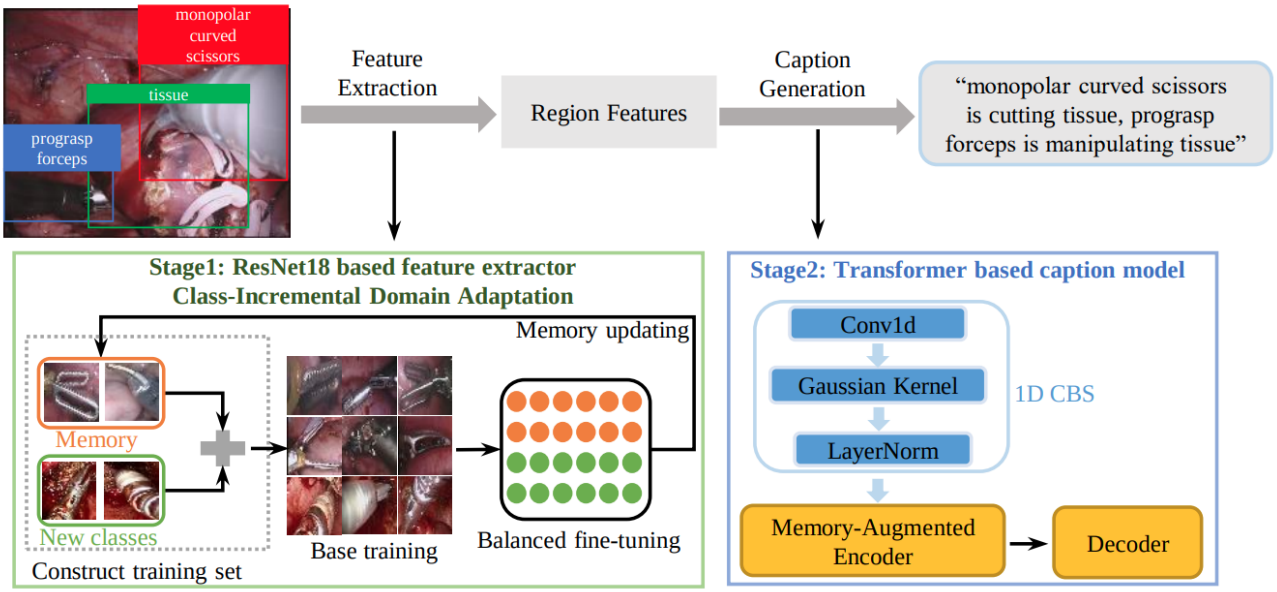


Fig. 2. Overall workflow, from C[1]. The input image is sent into the ResNet18-based feature extractor augmented with CIDA, and output region features. Inside the Transformer-based caption model with 1D CBS, the encoder takes in the region features and understands the relationship between regions. The decoder receives the encoder's output and generates the medical report.

A class-incremental (CI) learning method with supervised contrastive (SupCon) loss is combined with a feature extractor to adapt incremental classes and extract domain invariant features. Curriculum by one-dimensional Gaussian Smoothing (CBS) is combined with a multi-layer transformer-based caption prediction model to create the caption from the extracted feature. CBS uses anti-aliasing to smooth the feature embedding and aids the model in learning domain invariant features. With both the feature extractor and the captioning model, we use label smoothing (LS) to calibrate prediction probability and gain better feature representation. The proposed methodologies are tested using datasets from two surgical domains: nephrectomy operations and transoral robotic surgery.

2.3 Improving Documenting Surgical Procedures with Faster Inference Speed

Surgical captioning plays an important role in surgical instruction prediction and report generation. However, the majority of captioning models still rely on the heavy computational object detector or feature extractor to extract regional features. In addition, the detection model requires additional bounding box annotation which is costly and needs skilled annotators. These lead to inference delays and limit the captioning model to deploy in real-time robotic surgery. For this purpose, we design an end-to-end detector and feature extractor-free captioning model by utilizing the patch-based shifted window technique. We propose a Shifted Window-Based Multi-Layer Perceptrons Transformer Captioning model (SwinMLP-TranCAP) with faster inference speed and less computation.

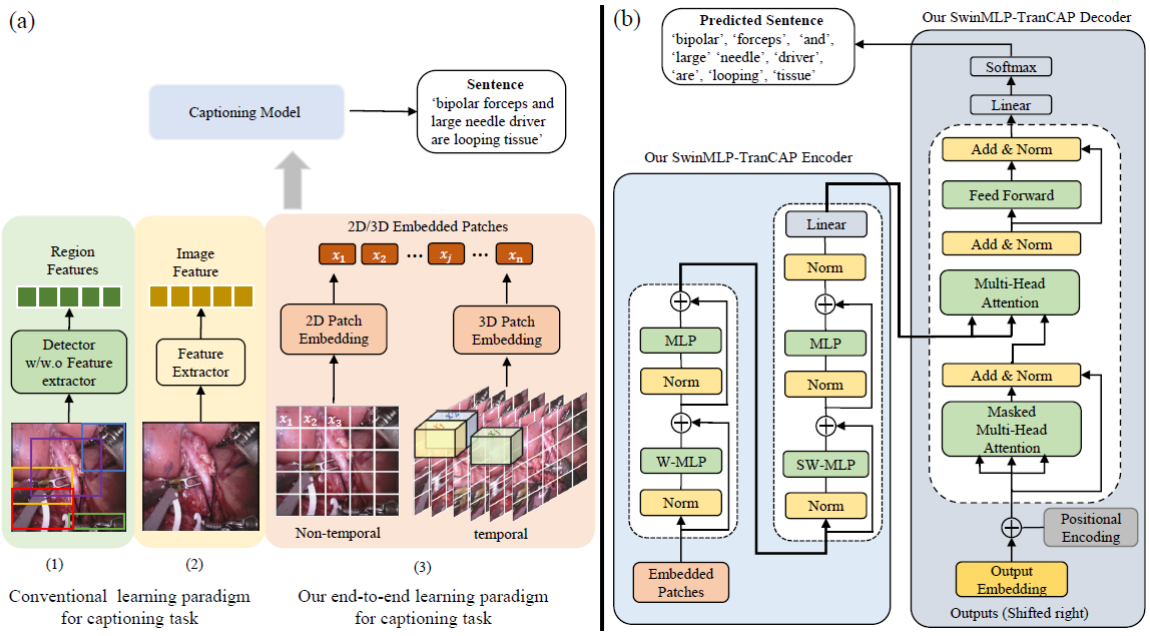


Fig. 3. (a) Comparisons of conventional learning paradigm and our learning paradigm for the captioning task and (b) our proposed end-to-end SwinMLP-TranCAP model, from [4]. (1) Captioning models based on an object detector w/w.o feature extractor to extract region features. (2) To eliminate the detector, the feature extractor can be applied as a compromise to the output image feature. (c) To eliminate the detector and feature extractor, the captioning models can be designed to take the patches as the input representation directly.

2.4 Gradient-Based Localized Feature Extraction for Surgical Downstream Tasks

Several approaches have been introduced to understand surgical scenes through downstream tasks like captioning and surgical scene graph generation. However, most of them heavily rely on an independent object detector and region-based feature extractor. Encompassing computationally expensive detection and feature extraction models, these multi-stage methods suffer from slow inference speed, making them less suitable for real-time surgical applications. The performance of the downstream tasks also degrades from inheriting errors of the earlier modules of the pipeline. This work develops a detector-free gradient-based localized feature extraction approach that enables end-to-end model training for downstream surgical tasks such as report generation and tool-tissue interaction graph prediction. We eliminate the need for object detection or region proposal and feature extraction networks by extracting the features of interest from the discriminative regions in the feature map of the classification models.

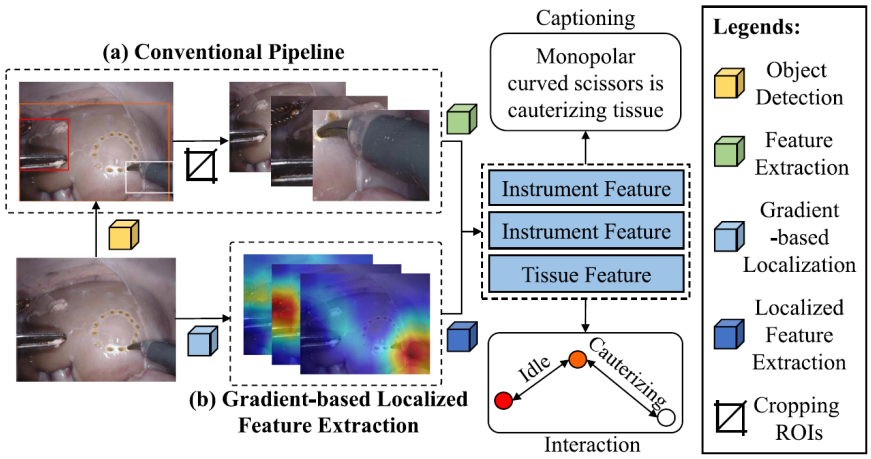
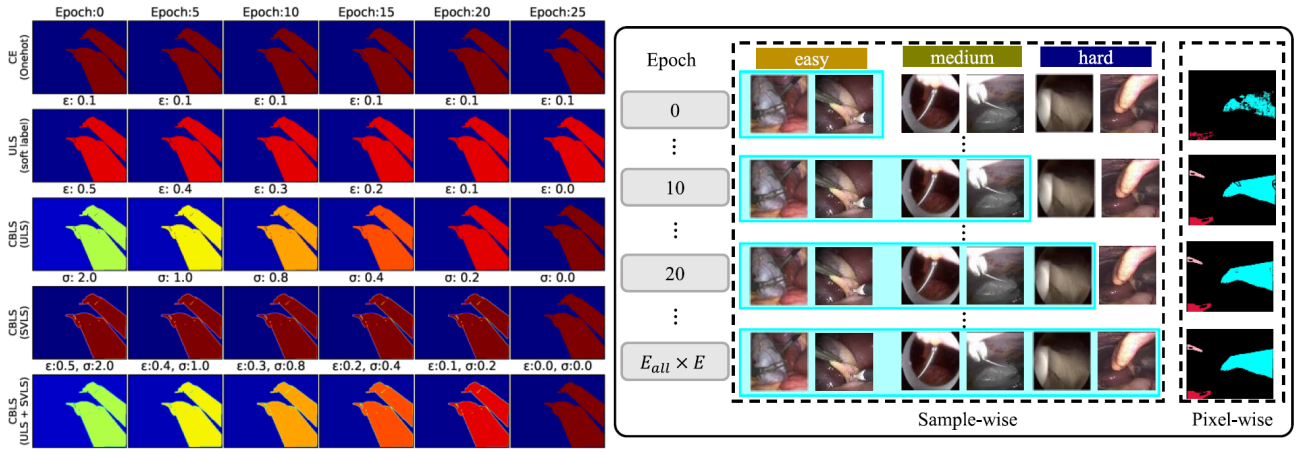


Fig. 4. An overview of our proposed gradient-based localized feature extraction method (b), against the conventional approach of deep learning pipelines (a) for downstream tasks, from [7].

7

2.4 Improving Surgical Scene Understanding with Curriculum Learning

Curriculum learning and self-paced learning are the training strategies that gradually feed the samples from easy to more complex. They have captivated increasing attention due to their excellent performance in robotic vision. We design a paced curriculum by label smoothing (P-CBLS) using paced learning with label smoothing (LS) for documenting surgical procedures in a curriculum manner.



(a) Curriculum Learning by Label Smoothing (b) Confidence-aware Paced Learning

Fig. 5. (a) Curriculum Learning by Label Smoothing. (b) Confidence-aware Paced Learning, including sample-wise paced learning, and pixel-wise paced learning, from [3].

2.5 Surgical-VQA: Visual Question Answering in Surgical Scenes

Expert surgeons are scarce and are often overloaded with clinical and academic workloads. This overload often limits their time answering questionnaires from patients, medical students, or junior residents related to surgical procedures. At times, students and junior residents also refrain from asking too many questions during classes to reduce disruption. While computer-aided simulators and recordings of past surgical procedures have been made available for them to observe and improve their skills, they still hugely rely on medical experts to answer their questions. Having a Surgical-VQA system as a reliable 'second opinion' could act as a backup and ease the load on the medical experts in answering these questions. The lack of annotated medical data and the presence of domain-specific terms has limited the exploration of VQA for surgical procedures. In this work, we design a Surgical-VQA task that answers questionnaires on surgical procedures based on the surgical scene.

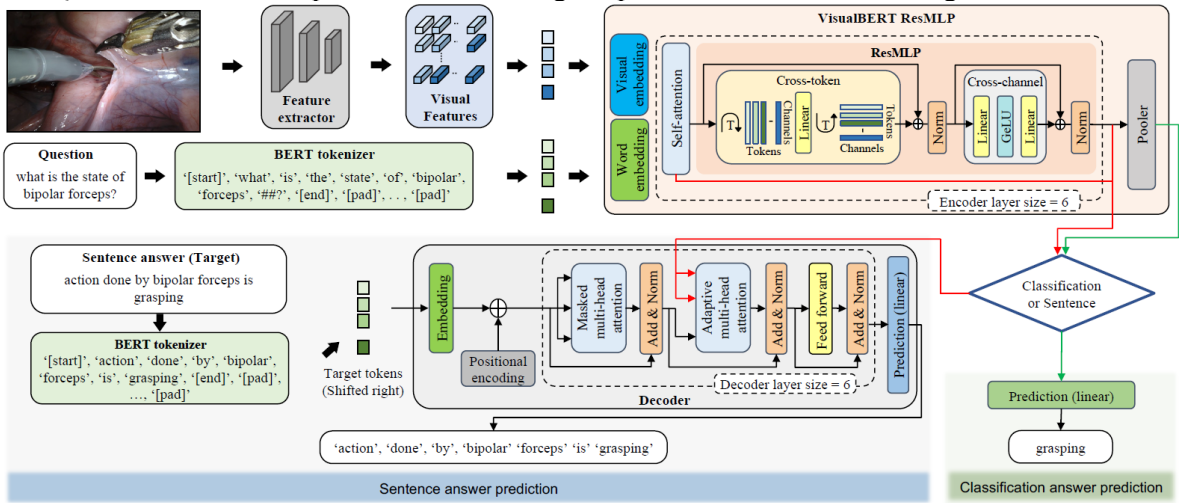


Fig. 6. Architecture: Given an input surgical scene and questions, its text and visual features are propagated through the vision-text encoder (VisualBERT ResMLP), from [5]. (i) Classification-based answer: The encoder output is propagated through a prediction layer for answer classification. (ii) Sentence-based answer: The encoder is combined with a transformer decoder to predict the answer sentence word-by-word (regressively).

2.6 Compliant and Flexible Robotic System with Parallel Continuum Mechanism for Transoral Surgery

As one of the minimally invasive surgeries (MIS), transoral robotic surgery (TORS) contributes to excellent

oncological and functional outcomes. This paper introduces a compliant and flexible robotic system for transoral surgery, consisting of an execution part with flexible parallel mechanisms and a positioning part with a continuum structure. A pilot cadaveric study that mimics the procedure of the TORS using an intact cadaveric human head was conducted to evaluate the feasibility and efficiency of this robotic system. Both the initial setup time and the time cost by the robot to safely access the deep surgical area in the upper aerodigestive tract are shortened due to the enlarged workspace, compact structure, and increased flexibility. The proposed surgical robotic system is preliminarily demonstrated to be feasible for TORS, especially for the in-depth surgical sites in the upper aerodigestive tract.

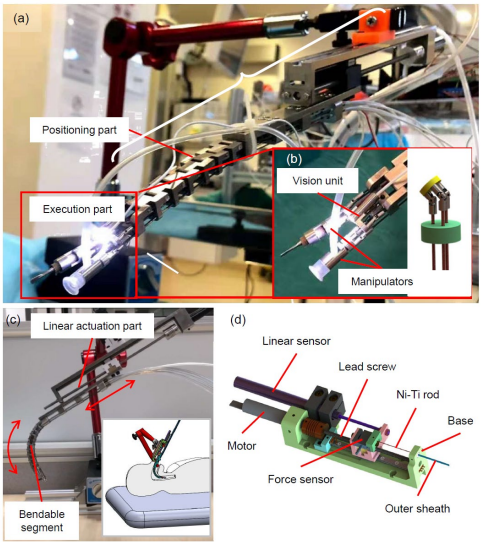


Fig. 7. Overview of the surgical robot. (a) Prototype. (b) Execution part. (c) Positioning part. (d) Part of the actuator to drive one Ni-Ti rod.

3. RESULTS ACHIEVED

3.1 Project Achievement:

(1) Automatically recognizing surgical interaction triplets (Instrument-Verb-Target), as shown in Table 1.
(2) Automatically documenting surgical procedures, as shown in Fig. 8.
(3) Achieve domain adaption by dealing with the domain shift, and novel instruments appear in the target domain (see Table 2).
(4) Achieve faster inference speed and less computation by eliminating the need for the detection model (see Table 3, 4, and 5, and Fig. 9)
(5) Improving task performance and robustness by incorporating the curriculum learning (see Table 5).
(6) Achieve visual question answering in the surgical scene. (see Table 6 and Fig. 10)
(7) Design a compliant and flexible surgical robotic system with a parallel continuum mechanism (see Fig. 11 and 12).

Table 1. The comparison of proposed SIRNet with other SOTA models, from J[1].

Method	mAP (%)↑	CP (%)↑	CR (%)↑	CF1 (%)↑	OP (%)↑	OR (%)↑	OF1 (%)↑	HL↓
Our SIRNet*	**42.63**	**26.93**	16.56	13.91	45.56	**27.82**	34.55	0.0856
Tripnet* [12]	40.55	9.47	6.15	6.95	**70.59**	9.30	16.44	**0.0789**
ML-GCN# [24]	36.72	17.33	12.17	11.98	50.82	26.61	**34.93**	0.0811
RMA# [25]	27.84	0.00	0.00	0.00	0.00	0.00	0.00	0.1124
ADD-GCN# [26]	37.46	16.95	**18.84**	**16.14**	40.08	27.30	32.48	0.0932
Mlp_mixer# [19]	37.02	14.31	9.43	7.74	41.90	15.16	22.26	0.0859
Swin transformer# [27]	34.92	24.26	7.82	9.15	47.62	21.53	29.66	0.0829
ViT# [17]	35.58	24.64	11.88	13.20	48.24	27.13	34.73	0.0828
DETR# [18]	28.02	0.00	0.00	0.00	0.00	0.00	0.00	0.0812

Note: The mAP, CP, CR, CF1, OP, OR, OF1, and HL are short for mean average precision, average per-class precision, average per-class recall,average per-class F1, average overall precision, average overall recall, average overall F1, and hamming loss. Please note that mAP is the most important metric in fine-grained SIR. ∗ denotes that the method treats a triplet as three separate components: instrument, verb, and target. # denotes that the method treats a triplet as an object.

3.2 Potential for Commercialization and Technology Transfer

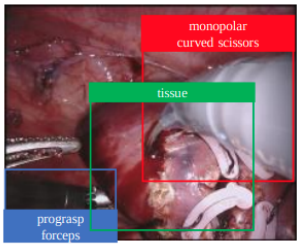
The documentation of patient care is the responsibility of physicians. Traditional clinical recording procedures have resulted in a poor patient-provider experience, reduced time spent dealing with patients, and a loss of work-life balance. Physicians spend substantial time in front of the computer performing administrative activities. As a result, patients are less content with their whole experience, and doctors, who have spent years studying medicine, cannot perform at the highest level of their licensure and are burnt out. Outside of office hours, physicians spend an additional 1 to 2 hours every night doing additional computer and clerical work.

Our ambient clinical intelligence solution can be commercialized from the following aspects:

(1) Release doctors from heavy documenting work.
(2) Provide surgical training for non-skilled surgeons.
(3) Generate surgical instruction to provide decision support during surgery.
(4) Make healthcare experiences more personalized.
(5) Design the visual question-answering system to ease the load on the medical experts in answering questions.

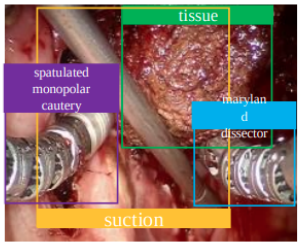
Table 2 Evaluation metrics of the proposed models in SD and TD, from C[1].

Domain	Stage 1 Feature Extractor	Stage 2 Caption Model	Metric						
			BLEU-1↑	BLEU-2↑	BLEU-3↑	BLEU-4↑	METEOR↑	ROUGE↑	CIDEr↑
SD	Res [10]	X-LAN [19]	0.5733	0.5053	0.4413	0.3885	0.3484	0.5642	2.0599
	Res [10]	M2T [6]	0.5703	0.5097	0.4572	0.4156	0.3817	0.599	2.5385
	Res [10]	M2T [6]+DANN [8]	0.5995	0.5318	0.4748	0.4301	**0.5995**	0.5994	2.4672
	Res [10]	Xu et al [24]	0.5875	0.5190	0.4599	0.4123	0.3621	0.5982	2.5930
	Res+**CI**	M2T	0.5571	0.4947	0.4395	0.3932	0.3609	0.5791	2.319
	Res+**CICL**	M2T+**CL**	0.6204	0.5498	0.4923	0.4452	0.3532	0.6017	2.6524
	Res+**CISC**	M2T+**CL**	**0.6246**	**0.5624**	**0.5117**	**0.472**	0.38	**0.6294**	**2.8548**
TD one shot	Res [10]	M2T [6](direct)	0.2408	0.098	0.0319	0.	0.1051	0.2407	0.1348
	Res [10]	M2T [6](fine-tune)	**0.5678**	0.4534	0.3891	0.3305	0.2759	0.5006	1.936
	Res+**CI**	M2T	0.5439	0.4568	0.403	0.352	**0.2886**	**0.5279**	2.2741
	Res+**CICL**	M2T+**CL**	0.5088	0.4277	0.3758	0.3286	0.2687	0.5071	2.3617
	Res+**CISC**	M2T+**CL**	**0.5626**	**0.472**	**0.417**	**0.3648**	0.2857	0.5147	**2.4641**
TD few shot	Res [10]	M2T [6](direct)	0.5331	0.4567	0.4114	0.3712	0.2738	0.5348	2.7496
	Res [10]	M2T [6](fine-tune)	0.5677	0.4807	0.4285	0.3836	0.289	0.5669	2.9209
	Res [10]	M2T [6]+DANN [8]	0.6338	0.5367	0.4819	0.4321	0.3173	0.5794	3.0407
	Res [10]	Xu et al [24]	0.6286	0.5422	0.4919	**0.4457**	0.3235	0.5921	3.3620
	Res+**CI**	M2T	0.6156	0.534	0.4859	0.4400	0.3189	0.5975	3.2223
	Res+**CICL**	M2T+**CL**	0.6314	0.5434	0.4912	0.4444	0.3262	0.6003	**3.3930**
	Res+**CISC**	M2T+**CL**	**0.6455**	**0.5518**	**0.4935**	0.4387	**0.328**	**0.6021**	3.3913



GT: monopolar curved scissors is cutting tissue, prograsp forceps is manipulating tissue
Base: bipolar forceps and prograsp forceps and monopolar curved scissors are idle
Ours: monopolar curved scissors is cutting tissue, bipolar forceps is retracting tissue

Source domain



GT: suction is suctioning blood, spatulated monopolar cautery and maryland dissector are manipulating tissue
Base: maryland dissector is grasping tissue
Ours: spatulated monopolar cautery is cauterizing tissue, suction is suctioning blood

Target domain

Fig. 8. The predicted caption of our approaches for source and target domain, from C[1].

Table 3 Evaluation metrics of the proposed efficient SwinMLP-Tran model, from [4].

Model			DAISI Dataset				EndoVis18 Dataset			
Det	FE	Captioning Model	B4	MET	SPI	CID	B4	MET	SPI	CID
FasterRCNN [12]	RN18 [7]	Tran [5]	✗				0.363	0.323	0.512	2.017
		Self-Seq [13]					0.295	0.283	0.496	1.801
		AOA [8]					0.377	0.371	0.58	1.811
YLv5x [9]	RN18 [7]	Tran [5]	✗				0.427	0.328	0.577	**3.022**
✗	RN18 [7]	Self-Seq [13]	0.296	0.207	0.330	2.827	0.446	0.353	0.531	2.674
		AOA [8]	0.349	0.246	0.403	3.373	0.427	0.322	0.533	2.903
		Tran [20]	0.454	0.308	**0.479**	**4.283**	0.426	0.335	0.524	2.826
✗	3DRN18	Tran [5]	✗				0.406	0.345	0.586	2.757
Ours		Swin-TranCAP	0.346	0.237	0.378	3.280	**0.459**	0.336	0.571	3.002
		SwinMLP-TranCAP	**0.459**	**0.308**	0.478	4.272	0.403	0.313	0.547	2.504
		V-SwinMLP-TranCAP	✗				0.423	**0.378**	**0.619**	2.663

Table 4 Proof of less computation cost of our SwinMLP-Tran captioning model, from [4].

Model			Proof of less-computation cost		
Det	FE	Captioning Model	FPS	N_Parameters(M)	GFLOPs
FasterRCNN [12]	RN18 [7]	Tran [5]	8.418	28.32+46.67	251.84+25.88
YLv5x [9]	RN18 [7]	Tran [5]	9.368	97.88+46.67	1412.8+25.88
✗	RN18 [7]	Tran [20]	11.083	11.69+46.67	1.82+25.88
Ours		Swin-TranCAP	10.604	165.51	19.59
		SwinMLP-TranCAP	**12.107**	**99.11**	**14.15**

Table 5 Performance of our gradient-based localized feature extraction method, from [7].

Method	**Model**		**FPS**	**No. of Params**	**Detection**	**Captioning** [11]			**Interaction** [8]	
	Det. Model	**FE**			**mAP@0.5**	**BLEU-1**	**BLEU-4**	**CIDEr**	**Recall**	**mRecall**
Conventional approach	Faster RCNN [2]	RN50	18.88	41.12 M	0.5538	0.4822	0.3182	2.04	0.2815	0.1530
	YOLOv5x [36]		25.05	109.73 M	0.5640	0.5767	0.4122	2.90	0.2940	0.1623
	YOLOv5s [36]		34.15	30.54 M	0.5450	0.5932	0.4443	2.79	0.2822	0.1413
	YOLOv7-E6E [37]		19.89	188.49 M	0.5670	0.6180	0.4566	2.85	0.3412	0.1386
	YOLOv7 [37]		35.78	60.03 M	0.5290	0.6129	0.4470	2.84	0.3055	0.1555
Ours (LN-FE)	Grad-CAM [19]	RN50	36.26	47.04 M	**0.6396**	0.6243	0.4587	3.41	0.2850	**0.1885**
Ours (L-FE)		✗	65.79	**23.51 M**		0.5984	0.4425	2.87	0.3272	0.1794
Ours (SL-FE)			**67.59**		✗	0.5900	0.4368	3.22	0.3523	0.1716
Ours (LN-FE)	Grad-CAM++ [20]	RN50	23.58	47.04 M	0.6237	**0.6532**	**0.5059**	**4.02**	0.3322	0.1444
Ours (L-FE)		✗	33.55	**23.51 M**		0.5731	0.4096	2.94	0.3358	0.1623
Ours (SL-FE)			35.10		✗	0.5530	0.3859	2.84	**0.3529**	0.1717

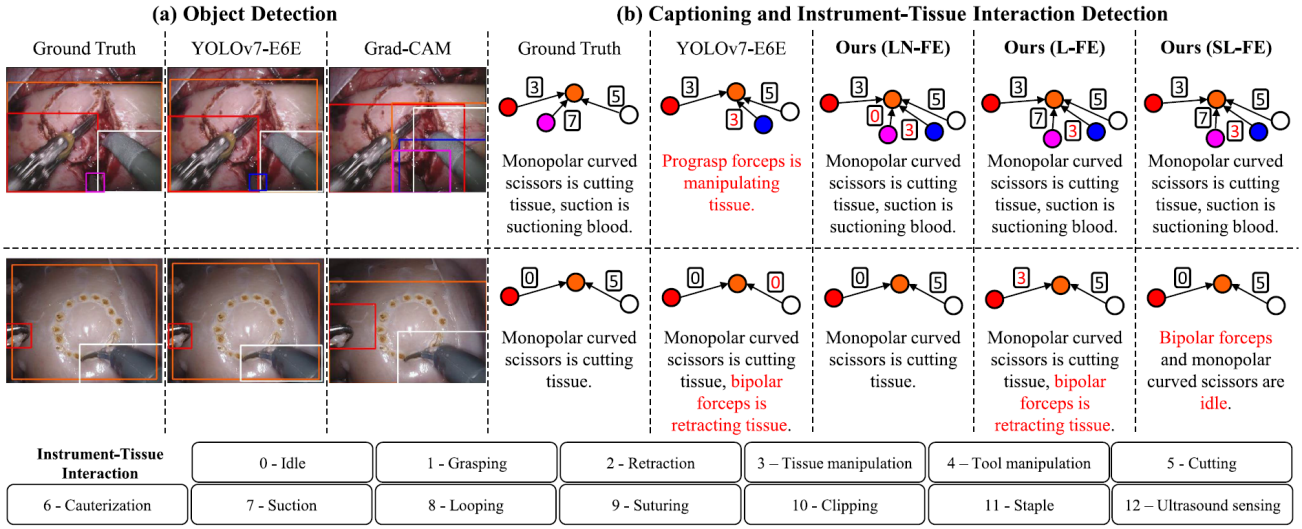


Fig. 9. Visualization of (a) Object detection with the conventional approach (YOLOv7-E6E) and the gradient-based localization (Grad-CAM). (b) Comparison of the scene captioning and interaction detection performance of our proposed feature extraction variants and the conventional object detection approach (YOLOv7- E6E)

Table 5 Evaluation metrics of the model augmented by curriculum learning technique (P-CBLS), from [3].

Model		BLEU-4	METEOR	ROUGE	CIDEr
Baseline	X-LAN [60]	38.85	34.84	56.42	205.99
	Xu et al [46]	44.52	35.32	60.17	265.24
	M2T [55]	42.29	35.67	58.95	274.78
	LS (M2T)	46.21	35.45	60.53	296.04
Ours	CBLS	45.07	**36.52**	60.25	309.65
	P-CBLS	**47.46**	35.67	**62.63**	**340.67**

Table 6 Evaluation metrics Comparison of our model and other models for sentence-based answering, from [5]

Model	EndoVis-18-VQA (S)				Cholec80-VQA (S)			
	BLEU-3	BLEU-4	CIDEr	METEOR	BLEU-3	BLEU-4	CIDEr	METEOR
MedFuse [20]	0.212	0.165	0.752	0.148	0.378	0.333	1.250	0.222
VisualBert [15] + TD	**0.727**	**0.694**	5.153	**0.544**	**0.963**	**0.956**	**8.802**	**0.719**
VisualBert ResMLP + TD	0.722	0.691	**5.262**	0.543	0.960	0.952	8.759	0.711

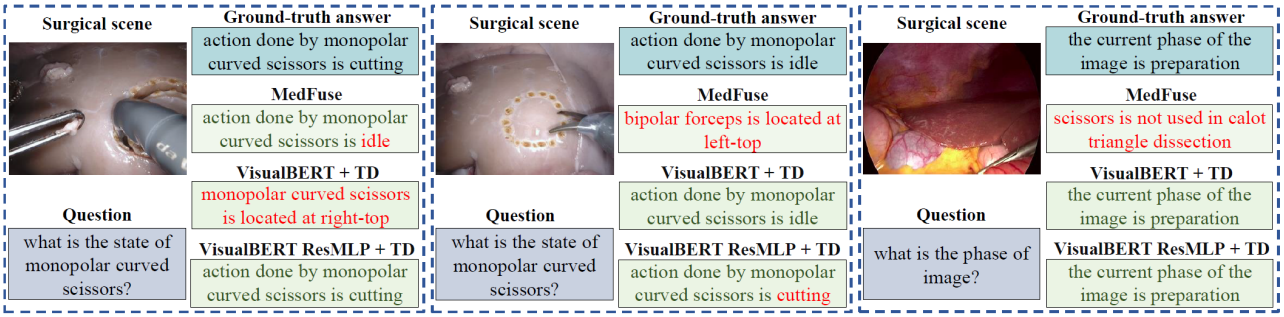


Fig. 10. Predicted answers generated by our model, from [5].

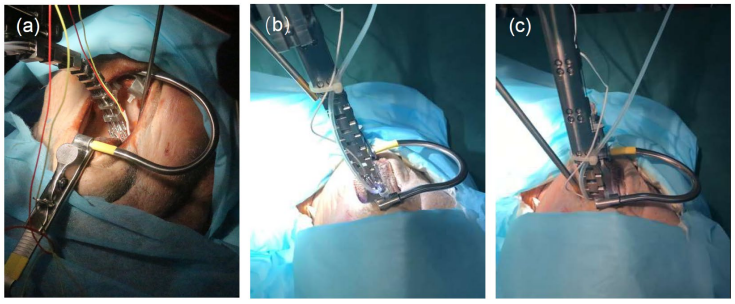


Fig. 11 Experimental results of the cadaveric simulations. (a) Initial setup. (b) Initial state. (c) Final state, from [6].

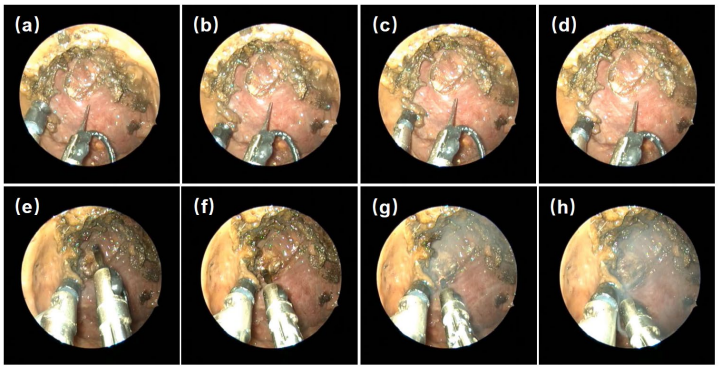


Fig. 12. Endoscopic view of manipulators during the cadaver trial. (a–d) The left manipulator is equipped with vacuum suction to grasp the target tissue. (e–h) The process of the right manipulator equipped with the monopolar electrotome to cut the target tissue, from [6].

4. PUBLICATION AND AWARDS (*CORRESPONDING AUTHOR)

[1] Xu, Mengya, Mobarakol Islam, Chwee Ming Lim, and Hongliang Ren*. "Class-Incremental Domain Adaptation with Smoothing and Calibration for Surgical Report Generation." In International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 269-278. Springer, Cham, 2021.

[2] Li, Ling, Xiaojian Li*, Shuai Ding, Zhao Fang, Mengya Xu, Hongliang Ren*, and Shanlin Yang. "SIRNet: Fine-Grained Surgical Interaction Recognition." IEEE Robotics and Automation Letters 7, no. 2 (2022): 4212-4219.

[3] Xu, Mengya, Mobarakol Islam, Ben Glocker, and Hongliang Ren*. "Confidence-Aware Paced-Curriculum Learning by Label Smoothing for Surgical Scene Understanding." IEEE Transactions on Automation Science and Engineering (2023).

[4] Xu, Mengya, Mobarakol Islam, and Hongliang Ren. "Rethinking surgical captioning: End-to-end window-based mlp transformer using patches." In International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 376-386. Cham: Springer Nature Switzerland, 2022.

[5] Seenivasan, Lalithkumar, Mobarakol Islam, Adithya K. Krishna, and Hongliang Ren*. "Surgical-vqa: Visual question answering in surgical scenes using transformer." In International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 33-43. Cham: Springer Nature Switzerland, 2022.

[6] Li, Changsheng, Xiaoyi Gu, Xiao Xiao, Chwee Ming Lim, and Hongliang Ren*. "Compliant and flexible robotic system with parallel continuum mechanism for transoral surgery: A pilot cadaveric study." Robotics 11, no. 6 (2022): 135.

[7] Pang, Winnie, Mobarakol Islam, Sai Mitheran, Lalithkumar Seenivasan, Mengya Xu, and Hongliang Ren*. "Rethinking Feature Extraction: Gradient-Based Localized Feature Extraction for End-To-End Surgical Downstream Tasks." IEEE Robotics and Automation Letters 7, no. 4 (2022): 12623-12630.

[8] Wang, An, Mobarakol Islam, Mengya Xu, Yang Zhang, and Hongliang Ren. "Sam meets robotic surgery: an empirical study on generalization, robustness and adaptation." In International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 234-244. Cham: Springer Nature Switzerland, 2023.

[9] Wang, An, Mengya Xu, Yang Zhang, Mobarakol Islam, and Hongliang Ren. "S 2 ME: Spatial-Spectral Mutual Teaching and Ensemble Learning for Scribble-Supervised Polyp Segmentation." In International Conference on Medical Image Computing and Computer-Assisted Intervention, pp. 35-45. Cham: Springer Nature Switzerland, 2023.

[10] Wang, Guankun, Long Bai, Yanan Wu, Tong Chen, and Hongliang Ren. "Rethinking exemplars for continual semantic segmentation in endoscopy scenes: Entropy-based mini-batch pseudo-replay." Computers in Biology and

Medicine 165 (2023): 107412.
[11] Islam, Mobarakol, Lalithkumar Seenivasan, S. P. Sharan, V. K. Viekash, Bhavesh Gupta, Ben Glocker, and Hongliang Ren. "Paced-curriculum distillation with prediction and label uncertainty for image segmentation." International Journal of Computer Assisted Radiology and Surgery 18, no. 10 (2023): 1875-1883.
[12] Wang, An, Mobarakol Islam, Mengya Xu, and Hongliang Ren. "Generalizing Surgical Instruments Segmentation to Unseen Domains with One-to-Many Synthesis." In 2023 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 4608-4614. IEEE, 2023.
[13] Seenivasan, Lalithkumar, Mobarakol Islam, Mengya Xu, Chwee Ming Lim, and Hongliang Ren. "Task-aware asynchronous multi-task model with class incremental contrastive learning for surgical scene understanding." International Journal of Computer Assisted Radiology and Surgery 18, no. 5 (2023): 921-928.
[14] Wang, An, Mobarakol Islam, Mengya Xu, and Hongliang Ren. "Curriculum-based augmented fourier domain adaptation for robust medical image segmentation." IEEE Transactions on Automation Science and Engineering (2023).
[15] Mengya Xu, and Hongliang Ren. "DAUA-Plane: Depth Anything with Uncertainty Map and Attention-based Plane Reconstruction of Surgical Scenes." International Conference on Biomimetic Intelligence and Robotics & Medical Robotics Forum (2024). YuanhuaTech Best paper award

COMPLETED

DEVELOPMENT OF DEEP-LEARNING ASSISTED ULTRAHIGH-RESOLUTION ENDOSCOPIC OCT FOR VISUALIZATION AND QUANTIFICATION OF VOLUMETRIC SMALL AIRWAY MICROSTRUCTURES IN VIVO

Principal Investigator: Professor Wu YUAN
Department of Biomedical Engineering, CUHK

Research Team Members:
Yuhan JU, Research Associate[(1)] ,
Tariq ABAXI, Research Associate[(1)] ,
Shaodi ZHU, Research Associate[(1)] ,
Chao XU, PhD Student[(1)]

[(1)]Department of Biomedical Engineering, CUHK

Project Start Date: 1st July 2020
Completion Date: 31st July 2022

INNOVATION AND PRACTICAL SIGNIFICANCE:

There are several novel aspects to this project:

First, the new endoscopic OCT system, including a novel high-speed 800-nm OCT engine (with an A-line speed up to 250 kHz), a fast rotary joint (for imaging speed more than 20 frames/sec), and a miniaturized and achromatic imaging endoscope (of a diameter less than 2 mm), is able to offer a superior imaging resolution (versus 1300-nm OCT) for accurate delineation of small airway microstructures, such as the epithelium, lamina propria, airway smooth muscle, glands, and blood vessels, in vivo and in video rate.

Second, with a powerful deep-learning neural network, we will be able to cope with the increasing volume of OCT images and the shortage of available human expertise, and achieve an automatic segment, quantify, and visualize the important tissue compartments in small airways in 3D. Such a tool would potentially provide a critical step to perform objective pathology assessment in vivo and in real time.

Third, the portable configuration of our endoscopic OCT system will help the future clinical translation. The small form factor of the imaging endoscope enables its direct deployment through the working channel of bronchoscope in clinic for minimal-invasively imaging the small airways in patients.

In summary, the capability to directly visualize and quantify airway smooth muscle and other critical tissue compartments in small airways in vivo, which are the critical pathology sites related to asthma and COPD, offers an unique opportunity for longitudinally studying the pathogenesis, determining the severity, phenotyping, guiding and monitoring the response to treatment in related pulmonary diseases.

ABSTRACT

In vivo visualization and quantification of the microstructures of small airways (of a lumen diameter <2 mm) allows a better understanding and management of asthma and chronic obstructive pulmonary disease (COPD). At present, the resolution and contrast of the clinically available imaging technologies (such as micro-CT and MRI) and conventional optical coherence tomography (OCT) operating at 1300 nm are insufficient to resolve the fine microstructures of small airways in vivo. In addition, the high-speed OCT generates a large amount of volumetric imaging data (up to tens of GBs), making manual image reading extremely laborious if not impossible.

1. OBJECTIVES AND SIGNIFICANCE

To overcome these obstacles, we aim to develop a deep-learning aided high-speed, ultrahigh-resolution endoscopic OCT technology for automatic volumetric measurements of small airway pathology in vivo.

Aim 1. Develop a portable, high-speed OCT engine and a fast rotary joint operating at 800 nm to enable the ultrahigh resolution and video-rate imaging of the small airways of patients in clinic.

Significance: The OCT system will operate at 800 nm with a broad spectral bandwidth, offering cellular resolution (≤2 μm) to delineate fine microstructures in small airway walls. The new OCT engine will acquire the image with an A-line speed up to 250 kHz. The high-speed rotary joint (not commercially available at 800 nm) will perform circumferential scan of the endoscope at speed more than 20 frames/sec to enable a real-time volumetric imaging of small airways.

Aim 2. Design and fabricate achromatic endoscopes of compact size (<2 mm) to work with bronchoscope for accurately assessing fine microstructures in the peripheral small airways in vivo.

Significance: Our OCT endoscope will afford high-resolution 3D in vivo "optical biopsy" of internal luminal organs, such as airways, enabling disease diagnosis and screening where traditional biopsy suffers from sampling error or risk of complications.

Aim 3. Develop a deep-learning neural network algorithm for automatic segmentation of small airway microstructures with the accuracy comparable to human expert and incorporate the deep-learning algorithm into the portable OCT system for potential clinical use.

Significance: The deep-learning assisted real-time tissue assessment on 3D OCT images in vivo will enhance the clinical viability of our endoscopic OCT imaging technique to perform objective pathology assessment in vivo and in real time.

2. RESEARCH METHODOLOGY

Aim 1: The new 800-nm endoscopic OCT engine will adopt a spectral-domain design, consisting of a broadband laser, a broadband fiber coupler and high-speed spectrometer. As for the high-speed rotatory joint, we will use a monolithic design and high-precision CNC fabrication to achieve a compact size and a rotation speed of more than 20 revolutions/sec.

Aim 2: We will use a small GRIN lens (of 1 mm or 0.7 mm diameter) to engineer a micro endoscope of a <2 mm diameter, so it can go through the working channel of bronchoscope for the clinic use.

Aim 3.1: The deep-learning algorithm will consist of an augmentation method specially designed for OCT images and a sequential model to capture spatial and temporal features for 3D segmentation. We will validate the performance of this neural network and compare the results with those of OCT expert reviewers.

Aim 3.2: The hardware-accelerated inference framework will be built with Nvidia TensorRT platform and Nvidia GPUs, to realize real-time processing capability.

3. RESULTS

3.1 Major Activities:

During the project, we have focused on Aims 1, 2 and 3.1 of the proposal: (1) development and characterization of 800-nm portable OCT system; (2) development the miniature endoscope and testing its performance for in vivo imaging; and (3) development and evaluation of deep-learning algorithm for automated airway segmentation.

By the end of the project, we have completed all research tasks of Aim 1, Aim 2 and Aim 3.1.

3.2 Significant Results:

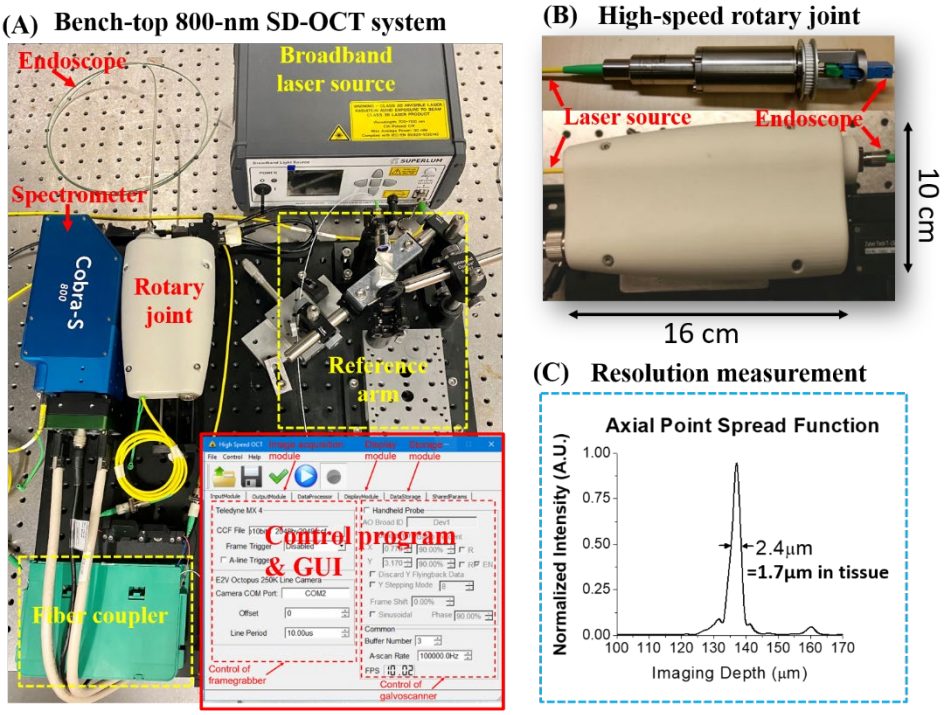


Figure 1. (**A**) Photo of the bench-top endoscopic spectral-domain OCT system operating at 800 nm, inset: control program and its GUI. (**B**) Customized broadband high-speed rotary joint. (**C**) The measurement of resolution is about 1.7 μm in tissue (2.4 μm in air).

Aim 1: We first finished the **bench-top version of our ultrahigh resolution OCT system at 800 nm**. In addition to customizing a portable, broad spectral bandwidth light source (Superluminescent diodes, Superlum) and a fast spectrometer of speed of 250 kHz (Cobra-S, Wasatch), we have customized a compact, broadband, high-speed, fiber-optic rotary joint in order to perform high-speed imaging for a proximal scanning endoscope (see Aim 2 below). This new rotary joint is able to correct the chromatic aberration and achieve one-way throughput >70%, optical coupling variation <±5% during continuous rotation, and back-reflection <-60 dB. With this rotary joint the imaging speed can be improved to up to 20 fps. **Figure 1** shows photos of the bench-top endoscopic OCT system operating at 800 nm with its control software and GUI (**Figure 1A**), the compact rotary joint (**Figure 1B**), and the resulting resolution of our OCT system (**Figure 1C**, measured with endoscope).

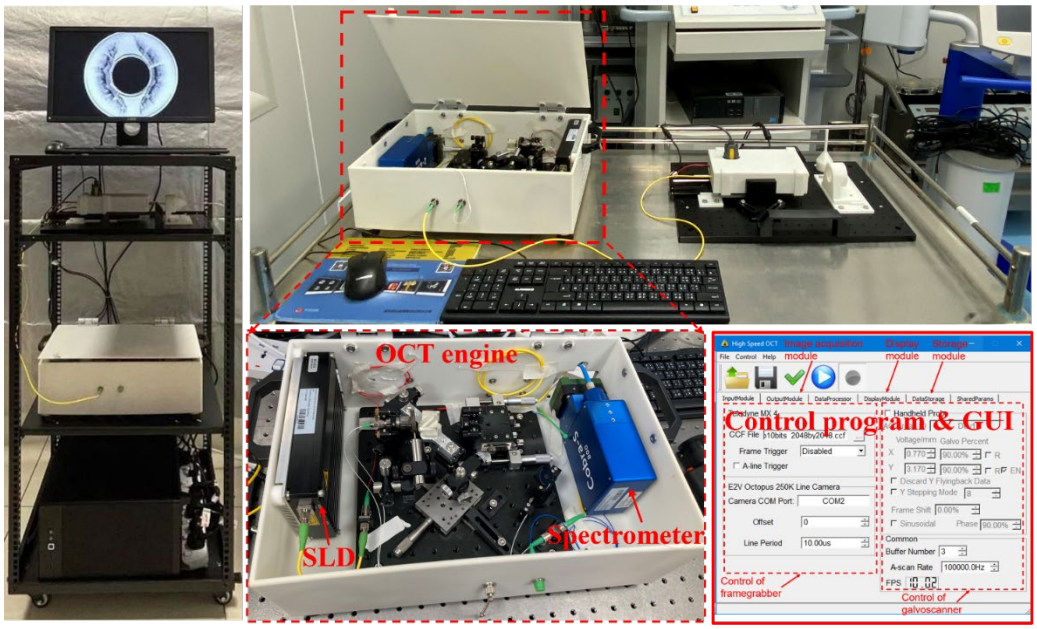


Figure 2. Photos of portable OCT system operating at 800 nm, the OCT engine, the control program and its GUI.

Based on the first benchtop system, we have further developed **a portable 800-nm ultrahigh resolution OCT** system based on spectral-domain OCT (SD-OCT) configuration. With this rotary joint the imaging speed can be

improved to up to 20 frames per second. **Figure 2** shows photos of the portable OCT system with its imaging engine and control software and GUI.
Aim 2: We have successfully developed a new endoscope (**Figures 3A and B**) that operated at 800 nm wavelength range with unprecedentedly high imaging resolution and speed while maintaining a highly compact size. The new endoscope is based on a 1-mm GRIN lens and incorporates a customized 1 mm diffractive compound lens, capable of rotating at speeds greater than 20 revolutions per second (compared to the previous 10 frames per second imaging speed). The probe has an overall outer diameter of 1.3 mm including the protective metal guard. The unique design of the endoscope enables achromatic beam focusing so all wavelength within the OCT light source spectrum can be focused to a single spot for achieving ultrahigh axial resolution. The small size enables the endoscope to pass through the accessory port of a standard clinical bronchoscope (**Figure 3C**) and reach terminal bronchioles with a diameter <2mm.
We further validate the portable system for in vivo lung imaging in animal models. The in vivo pig lung imaging protocol was approved by the Animal Experimentation Ethics Committee of CUHK. The pig was anesthetized according to the approved protocol and OCT airway imaging was performed under the guidance of bronchoscopy and HRCT (**Figure 4A and B**). Briefly, the imaging endoscope was deployed to the small airway through the 2.2-mm working channel of a bronchoscope (Ambu aScope 4 Broncho). The endoscope was placed to the preselected section of lung by first using the bronchoscopic view (**Figure 4C**) and then real-time imaging of HRCT (**Figure 4D**). We tested our system in 2 pigs. A representative cross-sectional OCT image clearly illustrates the microstructures of small airways, including the epithelium and alveoli (**Figure 4E**).

Figure 4. (**A**) The validation study on pig model, the μ-EBOCT airway imaging is guided with bronchoscope and high-resolution CT (HRCT). (**B**) HRCT C-arm is used to guide the deployment of endoscope in pig lung. (**C**) and (D) the OCT endoscope under the bronchoscopic and HRCT views to precisely guide the deployment of endoscope in lungs. (E) The representative OCT images in deep lung showing the healthy alveoli of pig.
Aim 3.1: To facilitate the volumetric visualization and quantification of small airway microstructures, we further trained and tested a ResNet18 neural network. The performance of our segmentation network was first evaluated using the test dataset. The automated segmentation results demonstrate a high similarity to the ones that were manually labelled (ground truth) by one of the experienced OCT reviewers, who segmented the OCT images by referring to corresponding histology (**Figures 5a-c**). An IoU of more than 0.8 was achieved for automated segmenting each tissue compartment in airway wall and airway lumen with the average IoU of about 0.92 (**Fig. 5d**). The trained and tested neural network was then applied to the series of OCT images of the small airways. and the segmented microstructures were then reconstructed along the longitudinal lumen direction in a 3D fashion (**Figure 6)**. This permitted quantitative evaluation of the architecture and volume of each tissue compartment.
We have parameterized the unwrapped en face views of each tissue compartment (**Figures 7a-c**) by encoding tissue thickness in color for an 18-mm long sheep small airway segment. One can clearly appreciate the longitudinally organized collagenous structures of basement membrane anchoring the similarly oriented epithelium (**Figures 7a and b**). Running perpendicular to the longitudinally oriented basement membrane, is the circumferentially oriented airway smooth muscle (**Figure 7c**). We further assessed the cross-sectional area of each tissue compartment longitudinally. We observed that the areas of the epithelium, basement membrane, airway smooth muscle, adventitia and cartilage increase gradually along the catheter pull-back direction (distal to the proximal direction) (**Figures 7a-d)**, leading to an increased cross-sectional airway wall area (**Figure 7e)**. Linear fitting was performed in one representative airway to indicate the trend of changes of cross-sectional areas for each tissue compartment and the entire airway wall over the length of the airway **(Figure 7)**

3.3 Key Outcomes:

The key outcomes of this project include:
(1) successful development of a bench-top and portable ultrahigh-resolution endoscopic OCT system operating at 800 nm for future use in a clinical setting.
(2) successful development of a compact rotary joint to enable high-speed endoscopic OCT imaging at 800 nm, a critical feature for in vivo imaging studies.
(3) successful development of novel OCT endoscopes at 800nm achieving ultrahigh axial resolution to enable direct visualization of smooth muscle in small airways in vivo.
(4) successfully validate the portable imaging system in animal models.

(5) successfully development of the deep-learning segmentation method for quantification and visualization of 3D microstructures of small airways with an accuracy comparable to the experienced OCT reviewers.

3.4. Other Achievements:

(1) Based on the preliminary results generated in the current project, we have successfully secured an ITSP project from ITC in 2022 (**ITS/240/21**, 2022-08-01 to 2024-01-31) to explore the clinical potential of the portable ultrahigh-resolution endoscopic imaging system and its commercialization potential.

4. PUBLICATIONS (* CORRESPONDING AUTHOR)

J[1] Jiarui Wang, Chao Xu, Shaodi Zhu, Defu Chen, Haixia Qiu, Alexander K. N. Lam, Christopher K. S. Leung, **Wu Yuan***, "A Generic and Effective System Dispersion Compensation Method: Development and Validation in Visible-Light OCT," ***Photonics***, 10, 892 (2023).

J[2] Shaodi Zhu, Zhenming Xie, Yuzhi Chen, Shiyue Liu, Yiu-Wa Kwan, Shuwen Zeng, **Wu Yuan***, Ho-Pui Ho*, "Real-Time Detection of Circulating Tumor Cells in Bloodstream Using Plasmonic Fiber Sensors", ***Biosensors***, 12, 968 (2022).

J[3] V. K. Viekash, Lalithkumar Seenivasan, S. P. Sharan, Mengya Xu, Zhen Li, Defu Chen, Haixia Qiu, **Wu Yuan ***, Hongliang Ren*, ***Pattern Recognition***, under review.

COMPLETED

A DEXTEROUS ROBOT FOR ENDOSCOPIC SUBMUCOSAL DISSECTION IN THE COLON

Principal Investigator: Professor CHENG Shing Shin
Department of Mechanical and Automation Engineering, CUHK

Co-investigator(s): TEOH, Yuen Bun Anthony[2]

Research Team Members:
Dezhao GUO, MPhil student [1]
Wenhui ZENG, Postdoctoral fellow [1]
Jibiao CHEN, PhD student [1]
Junyan YAN, PhD student [1]

[1] Dept. of Mechanical and Automation Engineering
[2] Dept. of Surgery

Project Start Date: 1st August 2020
Completion Date: 31st July 2022

INNOVATION AND PRACTICAL SIGNIFICANCE:

The proposed research focuses on the development of a dexterous endoscopic robot for performing endoscopic submucosal dissection (ESD) in the colon. Currently, ESD is performed manually using an instrument such as an ITknife 2 that passes through the 3.2mm-diameter instrument channel of a colonoscope. The limited degree-of-freedom (DoF) of only translation of the instrument requires the distal part of the colonoscope to be maneuvered together with the instrument, leading to unstable visualization that hinders effective execution of ESD. The procedure also requires coordination between an endoscopist and an assistant nurse who are responsible for manipulating the colonoscope and the instrument, respectively, which is a challenging task. The proposed research addresses these challenges by developing a two-segment dexterous endoscopic robot with five DoFs, namely distal segment bending in pitch and yaw, proximal segment bending in pitch, robot translation, and self-axis robot rotation. The ability to steer the ESD instrument dexterously allows the colonoscope to act as a stable visualization platform. The teleoperation control of the multi-DoF robot also allows the entire ESD procedure to be handled by one endoscopist, potentially leading to higher performance and productivity. While dual-arm flexible endoscopic robotic systems have been developed previously for endoscopic intervention, many issues have been identified that hinder their clinical translation, including high system complexity, coupled motion between segments in the arms, and lack of robust haptic-based controller. The proposed dexterous endoscopic robot, which is already a significantly less complex system, also addresses these concerns. Its two segments will be decoupled through purposeful design based on a novel mechanics model to increase motion accuracy. The integration of hysteresis compensation and the Fiber Bragg Grating (FBG)-based force sensor as feedback in its controller ensures more accurate and safe control. The robot, miniaturized to fit inside the instrument channel of a regular colonoscope, will directly replace the existing manually-operated ITknife 2 and thus create minimal disruption to the standard procedure workflow, leading to an increase in its potential for clinical translation.

ABSTRACT

Endoscopic submucosal dissection (ESD) is an effective minimally invasive treatment for early colorectal cancer. However, it is challenging to perform ESD with existing instruments due to the lack of distal dexterity and stable

visualization. While dual-channel flexible endoscopic systems have been developed for general transluminal surgery in the recent decade, their performance has been limited by the lack of operation stability, miniaturization issue, system complexity, and lack of haptic feedback. Therefore, the proposed research aims at developing a dexterous/steerable endoscopic robot for single-channel colonoscope to perform ESD. The 4- DoF robot, based on the notched tube design, has been miniaturized to 2.5 mm diameter (among the world's smallest steerable endoscopic instruments for ESD) to fit the 2.8 mm diameter instrument channel of a standard endoscope (gastroscope or colonoscope). It has been designed to be handheld so that a single operator can handle the endoscope and the steerable endoscopic robot simultaneously while staying by the patient's bedside to take care of the patient's intraoperative needs. An accurate kinematic model that takes into account the hysteresis behavior due to the long flexible actuation has been developed and taken into account in the control. A feedforward position controller based on the hysteresis-compensated kinematic model has also been implemented to provide improved position tracking capability. The robot was finally evaluated for its performance through benchtop performance and tested in a colon phantom.

1. OBJECTIVES AND SIGNIFICANCE

1.1 Objectives

- To develop a novel multi-degree of freedom (DoF) dexterous/steerable endoscopic robot for endoscopic submucosal dissection (ESD) in the colon.
- To develop an accurate kinematic model and characterize hysteresis to improve control accuracy.
- To develop an error-compensated controller for the robot.
- To experimentally evaluate the performance of the dexterous endoscopic robot in benchtop experiments and colon phantom.

1.2 Significance of the research

- More accurate and dexterous control of the instrument during ESD
- Stable endoscopic visualization
- ESD by only one endoscopist, improving performance and surgical outcome
- Safer interaction between the instrument and the inner colon wall due to improved position control
- Direct adoption in a standard single-channel colonoscope with minimal disruption to current procedure

2. RESEARCH METHODOLOGY

Task 1A: Design and fabrication of a steerable end effector of the endoscopic robot
Task 1B: Development of the flexible robot body and the actuation system of the endoscopic robot
Task 2A: Kinematic modeling of the steerable end effector of the endoscopic robot
Task 2B: Hysteresis modeling of the cable behavior in the flexible body of the endoscopic robot
Task 3A: Implementation of a model-based hysteresis-compensated feedforward position controller for the endoscopic robot
Task 4A: Experimental evaluation of the endoscopic robot performance in benchtop experiments
Task 4B: Experimental evaluation of the endoscopic robot performance in colon phantom

1. RESULTS ACHIEVED

1.1: Develop a novel multi-degree of freedom (DoF) dexterous endoscopic robot for endoscopic submucosal dissection (ESD) in the colon.

A novel hand-held steerable endoscopic robot (HSER) has been developed for submucosal dissection (ESD), consisting of three parts: 1) a steerable manipulator, 2) a detachable hand-held actuator with customized interface and 3) control unit and master PC. It has a total of four degrees-of-freedom (DOFs): two for end effector bending, one for instrument insertion and one for electrical surgical knife operation. As shown in Fig. 1, the steerable manipulator can be passed through the working channel of a standard endoscope (gastroscope or colonoscope) to be used as the instrument for ESD.

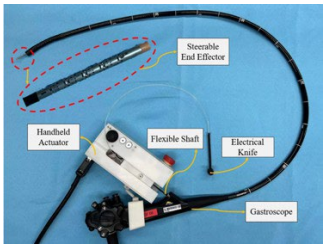


Fig. 1. Overview of the HSER attached with the gastroscope.

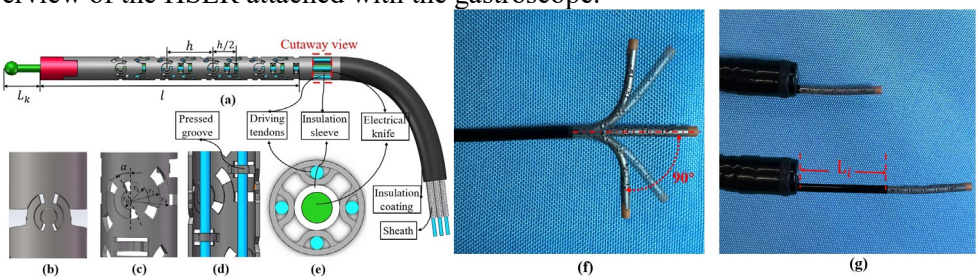


Fig. 2. The prototype of the steerable manipulator. (a) the overview of steerable manipulator. (b) Symmetric interlocking circular pattern. (c) Structure of optimized interlocking circular pattern. (d)-(e) Cross-sectional view of the steerable end effector. (f) Bending performance of end effector. (g) Insertion performance of proposed HSER.

As shown in Fig. 2(a), the 2.5 mm diameter steerable manipulator among the world's smallest steerable endoscopic instrument capable of performing ESD. It is composed of a long flexible shaft, four tendon- sheaths (steel tendon and wire coil), an electrical surgical knife, and a steerable end effector. The electrical surgical knife is obtained directly from the commercial product, Dual Knife-J (KD-655L, Olympus, Japan). One of the main innovations in the project is the steerable end effector with circular patterns fabricated out of a stainless-steel tube by laser micromachining. It is designed to be contact-aided compliant mechanism. A series of compliant joint patterns, numbered as N, are cut isometrically along the tube central axis in two orthogonal planes, allowing the end effector to bend towards four directions. As an asymmetrical topology structure, the circular topology pattern forms an interlocking mechanism and increases the contact-aided area, which contribute to the improved stiffness at maximal bending angle. Besides, a pressed groove sits underneath each circular pattern in each joint, allowing the passage of actuation tendons along the wall of the steerable end effector and welded at its distal end.

The structure of the hand-held actuator is shown in Fig. 3. It can be divided into three parts: motor unit, transmission unit and external unit. In the motor unit, two DC motors (RE13, Maxon, Switzerland) and one linear motor (LA16, INSPIRE-ROBOTS, China) are installed to output the driving motions. The steerable manipulator is attached with the transmission unit. As shown in Fig. 3(c), the transmission unit contains three linear motion modules, including two lead screws for the end effector's two DOF bending and one chute-type linear guide for the linear translation of the electrical knife. The tendons that control the left-right bending are fixed with wire connector I, while tendons that control the up-down bending are fixed with wire connector II. In order to create tendon pretension, all the tendons are looped around pins and then fixed by set screws. Therefore, every bending DOF could be controlled only by an individual motor. In the electrical knife operation, the linear motion of C-link will guide the electrical knife to move linearly, and the water sink furnishes the robot with the irrigation function. The transmission unit and the motor unit can be quick- connected and then driven along the endoscope instrument channel, guided by the lead screw in the external unit. All the control interface and electrical boards are also installed on the external unit.

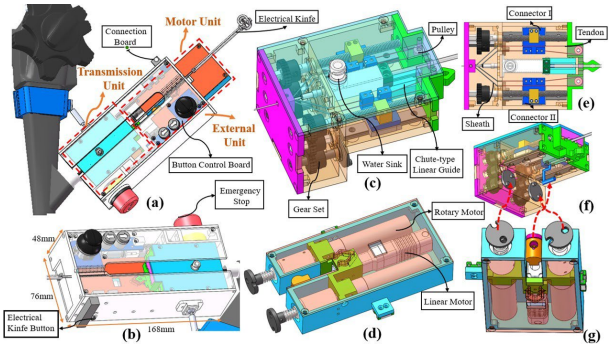


Fig. 3. Mechanical design of the hand-held actuator. (a) Overall structure of hand-held actuator. (b) Side view of hand-held actuator. (c,e,f) Multi-perspective views of transmission unit. (d,g) Multi-perspective views of motor unit.

1.2 : To develop an accurate kinematic model and hysteresis model to improve control accuracy.

Under the assumption of piecewise constant curvature (PCC), the frames are established as shown in Fig. 4. The base frame B is located at the center of instrument channel of the gastroscope, while the tip frame D is attached to the distal end of the end effector. The length of the centerline is l, and its curvature is denoted as k. φ is the deflection angle that the bending plane forms with the x axis. The initial distance between base frame B and frame C is defined as $d0$. Since the robot would move distance Li along z-axis, the actual distance is $d0 + Li$. As a result, the forward kinematics (FK) is derived and homogeneous transformation matrix from the base frame B to the distal tip frame E is given by:

$$
\begin{aligned}
{}^{B}_{D}T &= {}^{B}_{C}T \cdot R_{\varphi} \cdot T_{p} \cdot R_{\theta} \\
&= \begin{bmatrix} \cos\varphi\cos\theta & -\sin\varphi & \cos\varphi\sin\theta & \frac{\cos\varphi(1-\cos\theta)}{k} \\ \sin\varphi\cos\theta & \cos\varphi & \sin\varphi\sin\theta & \frac{\sin\varphi(1-\cos\theta)}{k} \\ -\sin\theta & 0 & \cos\theta & \frac{\sin\theta}{k} + d_0 + L_i \\ 0 & 0 & 0 & 1 \end{bmatrix}
\end{aligned} \tag{1}
$$

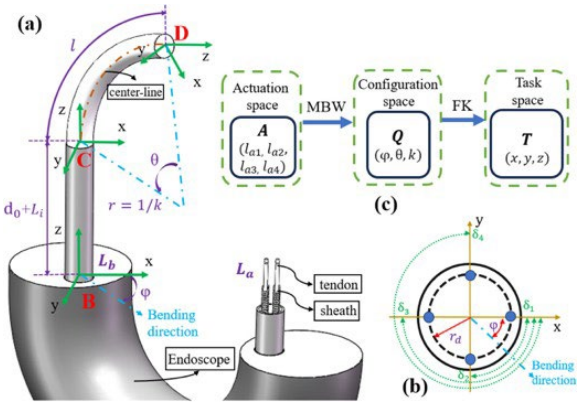


Fig. 4. (a) Geometrical model of steerable end effector. (b) Cross-sectional view of the tendon location. (c) Kinematics transformation among actuation, configuration and task space.

For the steerable end effector, the arc parameters Q = (φ, θ, k) can be directly determined by the cable displacements ΔLb at the base frame. The mapping between tendon space Lb = ($lb1$, $lb2$, $lb3$, $lb3$) and configuration space Q can be expressed by:

$$\Delta l_{bi} = l - l_{bi} = r_d\theta\cos(\sigma_i - \varphi) \tag{2}$$

where $\sigma i = (i-1)\,\pi$, $i = 1, 2, 3, 4$

$$
\begin{cases} \varphi = arctan(\dfrac{\Delta l_{b2}}{\Delta l_{b1}}) \\ \theta = \dfrac{\sqrt{\Delta l_{b1}^2 + \Delta l_{b2}^2}}{r_d} \\ k = \dfrac{\theta}{l} \end{cases} \tag{3}
$$

However, owing to the existence of hysteresis orginated from the nonlinear friction and the backlash phenomenon, the cable displacements at actuator ΔLa are not equal to the displacements ΔLb at the end effctor's base. In other words, the relationship between arc parameters Q and actuator lengths La = ($la1$, $la2$, $la3$, $la3$) cannot be obtained by Eq. (3). As a result, we proposed a new modified Bouc-Wen (MBW) model to more accurately represent the relationship between actuation space and configuration space. The function of MBW is presented as follow:

$$\begin{cases} \epsilon(\dot{x}(t)) = \lambda_1 + \dfrac{\lambda_2}{1+e^{-\dot{x}(t)}} \\ \Theta(x,t) = \epsilon(\dot{x}(t)) \cdot x(t) + \lambda_3 \cdot z(t) \end{cases} \tag{4}$$

$$\dot{z}(t) = |\dot{x}(t)|[\text{sgn}(\dot{x}(t)) - \beta|z(t)|^{n-1}z(t) + \gamma] \tag{5}$$

where x(t) is the cable displacement, Θ(x,t) is the output bending angle, and z(t) is the hysteresis variable. The smooth function $\varepsilon(\dot{x}(t))$ employs the relative velocity $\dot{x}(t)$ to capture the hysteresis behavior.

$H(\lambda 1, \lambda 2, \lambda 3, n, \beta, \gamma)$ are the parameters that control the shapes of the hysteresis loops in both loading and unloading phases. In order to match the model output with the experimental bending angles, the genetic algorithm (GA) optimization is utilized to minimize the value of the fitting function given by:

$$Min\ f(x) = \frac{1}{N}\sum^{N}(\Theta_m(i) - \Theta_p(i))^2 \tag{6}$$

where Θm represents the measured bending angle, Θp denotes the predicted output generated by the proposed model. Under relevant constraints and a generation of 600, the parameters of MBW would be identified.

Then, the cable displacements $\Delta La1$ and corresponding bending angle $\theta 1$ are collected when the robot only bends in the XZ plane. We then performed the same test on YZ plane and obtained $\Delta La2$ and $\theta 2$. After combining the two sets of data into Eq. (4), Eq. (6) would allow us to determine the corresponding hysteresis parameters $H1$, $H2$ as well as bending relationship $\Theta 1(\Delta La1)$, $\Theta 1(\Delta La2)$. As a result, the end effector's curvature along the x-axis (kx) and y-axis (ky) can be formulated with the actuator displacements as:

$$\begin{cases} k_x = \dfrac{\theta_x}{l} = \dfrac{\Theta_1(\Delta l_{a1})}{l} \\ k_y = \dfrac{\theta_y}{l} = \dfrac{\Theta_2(\Delta l_{a2})}{l} \end{cases} \tag{7}$$

The overall bending curvature k, deflection angle φ, and overall bending angle θ can be expressed by:

$$\begin{cases} k = \sqrt{k_v^2 + k_h^2} = \dfrac{\sqrt{\Theta_v^2 + \Theta_h^2}}{l} \\ \varphi = arctan(\dfrac{k_v}{k_h}) = arctan(\dfrac{\Theta_v}{\Theta_h}) \\ \theta = lk = \sqrt{\Theta_v^2 + \Theta_h^2} \end{cases} \tag{8}$$

3.3: To develop a position controller for the robot.

A feedforward control technique based on the proposed hysteresis-incorporated kinematics method is implemented to improve the end effector's position tracking performance. After inputting the experimental data into proposed MBW model, we would get the empirical hysteresis parameters $H@$ as well as predicted hysteresis displacement $\tilde{z}$. Assuming that the input cable displacements to the robotic system is $xs(t)$, the command bending angle θd equals to the robotic system output bending angle θout as follow:

$$\theta_{out} = \tilde{\epsilon}(\dot{x}_s(t)) \cdot x_s(t) + \tilde{\lambda_3} \cdot z(t) = \theta_d \tag{9}$$

$$\tilde{\epsilon}(\dot{x}_s(t)) = \tilde{\lambda_1} + \frac{\tilde{\lambda_2}}{1+e^{-\dot{x}_s(t)}} \tag{10}$$

Substitute z(t) with the estimated hysteresis displacement $\tilde{z}$, we obtain the expression for $xs(t)$ as:

$$x_s(t) = \frac{1}{\tilde{\epsilon}(\dot{x}_s(t))}(\theta_d - \tilde{\lambda_3} \cdot \tilde{z}(t)) \tag{11}$$

$$\dot{\tilde{z}}(t) = |\dot{x}_s(t)|[\text{sgn}(\dot{x}_s(t)) - \tilde{\beta}|\tilde{z}(t)|^{\tilde{n}-1}\tilde{z}(t) + \tilde{\gamma}] \tag{12}$$

Hence, if the hysteresis displacement is well estimated by predicted value, which means $\tilde{z} = z(t)$, the output bending angle θ will follow the input command value θd.

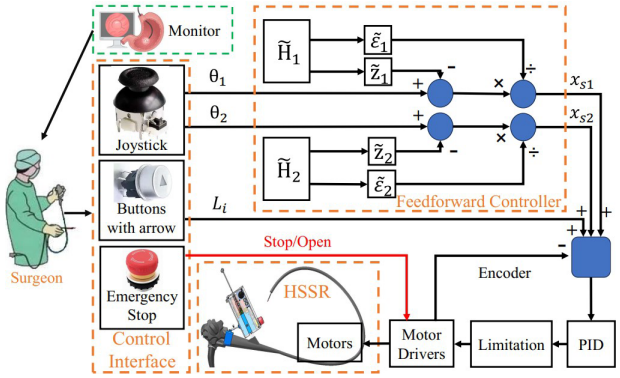


Fig. 5. Control diagram of hand-held position control

The overall control framework of the hand-held control is shown in Fig. 5. The control strategy is composed of two parts: the low-level PID control and the hysteresis-based feedforward controller. The user controls the bending angle of the steerable end effector by pushing the joystick. The pushing angles are defined as $[\theta pitch, \theta yaw]$. By multiplying a coefficient radio Ks with the pushing angles, we can get the mapping relationship between the command robot bending angle and the joystick pushing angle as: $[\theta 1, \theta 2] =$

$Ks[\theta pitch, \theta yaw]$. Then the command bending angles $\theta 1$, $\theta 2$ in XZ and YZ plane are input into the feedforward controller. The end effector's final bending angle is illustrated as:

$$\theta = K_s \sqrt{\theta_{pitch}^2 + \theta_{yaw}^2} \tag{13}$$

The robot insertion is controlled by the buttons with forward arrow and backward arrow (labeled as $B1$ and $B2$ respectively). Once they are pressed, the steerable manipulator would move forward or backward at a constant speed vi. Here we represent the status of button with function $sgn(B)$:

$$sgn(B) = \begin{cases} 1, & \text{the button is pressed} \\ 0, & \text{the button is not pressed} \end{cases} \tag{14}$$

3.4 : To experimentally evaluate the performance of the endoscopic robot in benchtop experiments and colon phantom.

Several benchtop experiments were performed to evaluate the performance of the steerable endoscopic robot. The experimental setup in Fig. 6(a) was designed to validate the hysteresis-incorporated kinematic model of the proposed robot in three-dimension space. We firstly tested the bending movements of the robot in XY- plane and YZ-plane, which were only actuated by connector I or connector II modules solely. The end effector bent from straight posture (0°) to two opposite directions (-90° and 90°), and the data of cable displacements and corresponding bending angles were recorded. Then, training these data using the proposed MBW model allowed us to obtain their corresponding hysteresis parameters in the two planes. After implementing the feedforward controller, the end effector was commanded to move in two planes that incline 45° with the YZ- plane. The measured four trajectories and the desired counterparts were presented in Fig. 6(b). The maximum error between the experiment and the kinematic model for trajectories ①-④ are 1.46 mm, 1.98 mm, 3.32 mm,2.27 mm, while root mean square errors (RMSE) are 0.119 mm, 0.248 mm, 0.130 mm and 0.446 mm. It is clearly shown that the kinematics model could precisely follow the experimental bending behavior regardless of the bending directions.

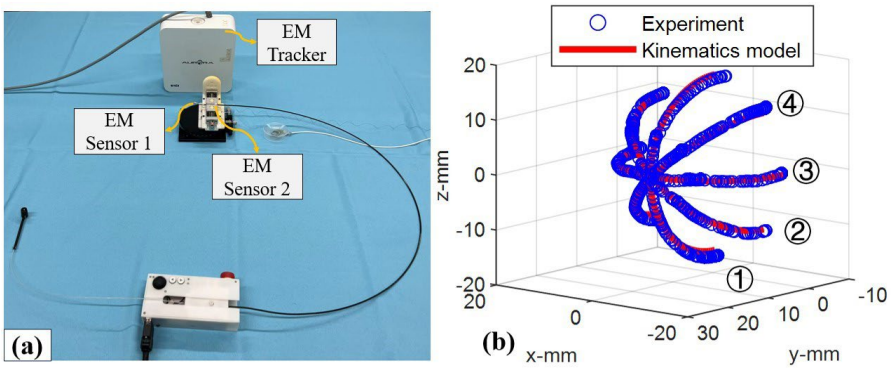


Fig. 6. (a) Experiment set up of kinematics verification. (b) The desired and measured trajectories in XY plane, YZ plane and 45° planes.

As shown in Fig. 7, we have also performed an experiment in the colon phantom by manipulating the steerable endoscopic robot in different directions after inserting it through the gastroscope into the colon phantom. It is shown that we were able to smoothly and flexibly manipulate the instrument inside the confined space of the colon through the working channel of a gastroscope.

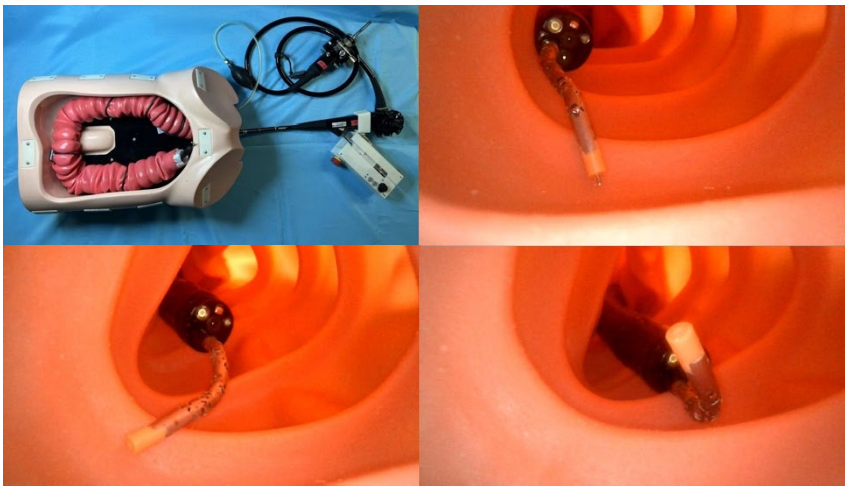


Fig. 7. Demonstration of manipulating the hand-held steerable endoscopic robot in a phantom colon.

4. PUBLICATIONS AND AWARDS

J[1] Wenhui Zeng, Junyan Yan, Yilun Hong, and **Shing Shin Cheng***. Numerical analysis of large deflection of the cantilever beam subjected to a force pointing at a fixed point. *Applied Mathematical Modelling*, 92, 719-730, 2020.

J[2] Wenhui Zeng, Junyan Yan, Kim Yan, Xu Huang, Xuefeng Wang, and **Shing Shin Cheng***. Modeling a symmetrically-notched continuum neurosurgical robot with non-constant curvature and superelastic property. IEEE Robotics and Automation Letters 6, no. 4, 6489-6496, 2021.

J[2] Dezhao Guo, Ai Dong, Junyan Yan, Anthony Yuen Bun Teoh, and Shing Shin Cheng. Design, Modeling and Control of a hand-held Steerable Endoscopic Robot for Endoscopic Submucosal Dissection. IEEE Robotics and Automation Letters, 2023. (Under review)

C[1] Wenhui Zeng, Junyan Yan, Xu Huang, and **Shing Shin Cheng***. Motion Coupling Analysis for the Decoupled Design of a Two-Segment Notched Continuum Robot. In 2021 IEEE International Conference on Robotics and Automation (ICRA), pp. 7665-7671, 2021.

COMPLETED

DEVELOPMENT OF A FOLDED BILAYER SCAFFOLD FOR INTESTINAL TISSUE ENGINEERING

Principal Investigator: Professor CHAN Hon Fai
Department of Biomedical Engineering, CUHK
Institute for Tissue Engineering and Regenerative Medicine

Research Team Members:
Xiaoyu Zhao, PhD student [1,2,3];
Yanlun Zhu, PhD student [1,2]; Hu Meng [1,2,3]

(1) Institute for Tissue Engineering and Regenerative Medicine, The Chinese University of Hong Kong, Shatin, Hong Kong SAR, People's Republic of China
(2) Key Laboratory for Regenerative Medicine of the Ministry of Education of China, School of Biomedical Sciences, Faculty of Medicine, The Chinese University of Hong Kong, Shatin, Hong Kong SAR, People's Republic of China
(3) Shun Hing Institute of Advanced Engineering, The Chinese University of Hong Kong, Shatin, Hong Kong SAR, People's Republic of China

Project Start Date: 1st July 2019
Completion Date: 30th June 2021

INNOVATION AND PRACTICAL SIGNIFICANCE:

The major innovation of the current work lies in the design of the folded bilayer hydrogel scaffold. All scaffold materials reported previously for the production of TESI consist of single-phase materials such as polylactic-co-glycolic acid (PLGA), poly-ε-caprolactone, chitosan etc, without taking into consideration the distinct properties of different layers of the intestinal wall. Decellularized tissue is an exception but its application is limited by human tissue source. The bilayer hydrogel scaffold is expected to outperform current designs by better recapitulating the multi-layered structure of small intestine. In addition, we are the first group demonstrating the fabrication of a folded scaffold by mechanics-guided folding, which represents a new paradigm in tissue engineering. The technique is novel but at the same time simple for scale-up production. We will file a patent application on the scaffold materials and fabrication technique after they are optimized, and actively pursue preclinical and clinical trial to bring the potential therapy to patients.

ABSTRACT

Short bowel syndrome can be caused by birth defect or surgical removal of part of the intestine resulting from a number of diseases. It is associated with high morbidity and mortality. Current therapies are mostly ineffective, ultimately requiring an intestinal transplant but complications arise. Construction of tissue-engineered small intestine (TESI) in vitro, in which patient cells are cultured on a biomaterial scaffold, offers a therapeutic alternative. However, the current scaffold designs do not take into consideration the distinct properties of different layers of the native intestine wall. In addition, the dense mucosal folds, which are present along the inner surface of intestine and assume an important functional role, have not been reproduced. Recently, we have engineered a folded bilayer hydrogel scaffold with tunable mechanical properties via a theory- and simulation-guided approach for the first time. A stretchable hydrogel is stretched before another hydrogel is layered on top. Subsequent relaxation induces controlled folding of the top layer, and the process is reversible. To further develop the strategy for clinical translation, we propose to 1) optimize the properties of the bilayer hydrogel scaffold in terms of the biomaterial used; 2) direct the differentiation of primary intestinal organoid in the bilayer scaffold. The successful completion of the project will improve the current design of TESI and advance the field of intestinal tissue engineering.

1. OBJECTIVES AND SIGNIFICANCE

Short bowel syndrome (SBS) affects 3-4 individuals per million people[1]. A number of conditions can lead to SBS, such as congenital anomalies and intestinal resection following Crohn's disease. Since SBS can cause malabsorption and malnutrition, it is associated with high morbidity and mortality, especially in children in which the estimated 5-year mortality approaches 40%[2]. Current therapies aiming to increase absorption or restore intestinal length are ineffective. Intestinal transplantation is the last resort but is limited by tissue supply and complications such as graft rejection[3].

The development of a tissue-engineered small intestine (TESI) by incorporating patient's own intestinal cells (e.g. organoids derived from intestine tissue) into a scaffold represents an attractive option for autologous transplantation[3]. However, none of the reported TESI has been applied in clinic so far. Possible reasons include a lack of control of organoid differentiation and complications arising from multiple surgeries required for the in vivo maturation of TESI[4]. Generating TESI resembling native intestine in vitro would be advantageous. Nevertheless, the current scaffold fabrication approaches fail to recapitulate mucosal fold, which increases surface area for absorption and endows the intestine with the flexibility to contract and expand, and mimic the properties of distinct layers of the intestine wall, consisting mainly of mucosa, submucosa, muscle layers and serosa. Our lab has recently demonstrated the fabrication of a folded bilayer hydrogel scaffold, mimicking the architecture of intestinal fold, based on the concept of surface instability[5].

To further develop the strategy for clinical translation, we proposed to achieve the following aims.
Aim 1: Optimize the properties of the bilayer hydrogel scaffold in terms of the biomaterial used;
Aim 2: Direct the differentiation of primary intestinal organoid in the bilayer scaffold by applying mechanical and biochemical cues.

We expect to identify an optimal scaffold formulation for organoid culture, and apply cues to direct organoid to differentiate into different intestinal cell types.

2. RESEARCH METHODOLOGY

2.1. Aim 1: Optimize the properties of the bilayer hydrogel scaffold in terms of the biomaterial used
Previously, we fabricated the bilayer scaffold using gelatin as the top layer but it tended to detach from the base hydrogel easily[5]. In this study, we fabricated a folded bilayer scaffold using polyacrylamide and alginate (base hydrogel) and alginate (top hydrogel) based on a modified protocol from our previous work[5]. We first developed a soft and stretchable base hydrogel using polyacrylamide and alginate. While the base hydrogel was under tension (e.g. 100%), the top hydrogel (alginate and Matrigel) was added before tension was released to generate the folds under compression. To improve the adhesion between two layers, we adopted alginate as the top layer which can form an interpenetrating network with alginate present in the base hydrogel. We performed bright-field microscopic and micro-CT imaging to analyze the folding structure and bonding between two layers. To determine the microstructure, we examined the pore of the scaffold using scanning electron microscopy (SEM). To assess the biocompatibility of the scaffold, we cultured two cell lines, mouse aortic endothelial cells (MAEC) and mouse fibroblast cells (L929), with the scaffold and assessed their viability by MTT assay over time. Finally, we measured the water content of the bilayer scaffold by quantifying the swelling ratio (i.e. wet weight/dry weight) of the lyophilized scaffold immersed in PBS buffer over time.

2.2. Aim 2: Direct the differentiation of primary intestinal organoid in the bilayer scaffold by applying mechanical and biochemical cues
We isolated the intestinal crypt from mouse intestine following an established protocol[6, 7]. The percentage of Lgr5+ positive cells, which represent the intestinal stem cells, was quantified using flow cytometry. The isolated

cells were then seeded onto a Matrigel scaffold and cultured in cell culture medium containing R-spondin for organoid formation. Immunostaining was then performed to visualize the presence of Lgr5+ stem cells. Since folding induced compression in the top hydrogel layer, we investigated whether mechanical compression modulated the growth and differentiation of organoid into various epithelial cell types such as goblet cells, enterocytes, Paneth cells. A weight was placed atop the Matrigel containing the organoids to simulate the mechanical effect. Polymerase chain reaction (PCR) and immunostaining were performed to analyze expression of markers specific for the differentiated cell types under compression and in the bilayer scaffold.

3. RESULTS ACHIEVED

3.1. Development of bilayer folded scaffold

Based on the concept of mechanical/surface instability (Fig. 1A), we fabricated the folded bilayer scaffold by first stretching polyacrylamide/alginate hydrogel before coating the surface with a layer of alginate and Matrigel hydrogel. After relaxation of the bottom layer, folding pattern was produced which was reversible upon stretching and relaxation (Fig. 1B). Micro-CT imaging showed a regular folding pattern that resembled the structure of porcine intestine (Fig. 1C). Bright-field microscopic imaging revealed that the top hydrogel layer in our new scaffold design adhered to the base layer tightly (Fig. 1D). SEM imaging showed that pores with an average size of ~20μm were observed in the alginate hydrogel (Fig. 1E), which facilitate cell migration in the scaffold. To determine the biocompatibility of the bilayer scaffold, we cultured MAEC and L929 cells with the scaffold and measured their viability over 6 days (Fig. 1F). No significant decrease in cell viability was observed, suggesting no detectable cytotoxicity. Finally, the bilayer scaffold was lyophilized before immersed into PBS at room temperature. The swelling ratio, calculated as the wet weight/dry weight and presented in Fig. 1G, was ~30 at the equilibrium state.

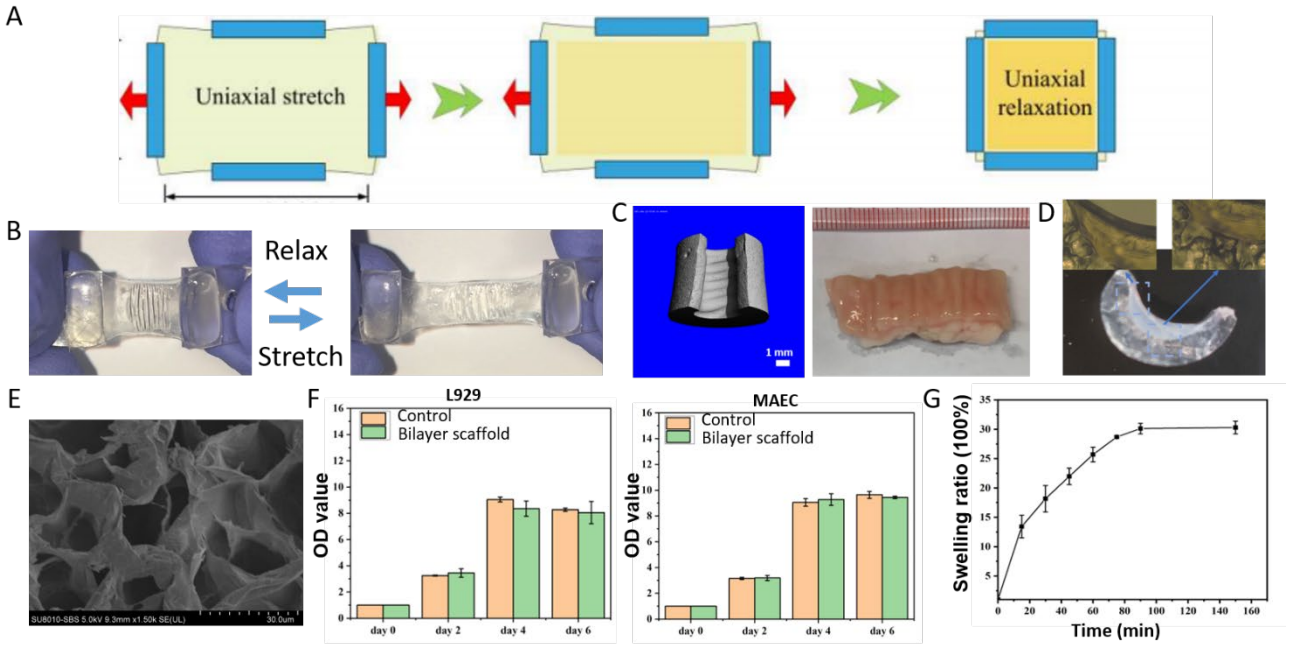


Figure 1. Fabrication and characterization of folded bilayer hydrogel scaffold. (A) Schematic diagram showing folded scaffold formation by mechanical/surface instability. (B) Reversible folding-unfolding triggered by stretching and relaxing the scaffold. (C) Micro-CT imaging of the scaffold and an image showing a section of the porcine intestine. (D) Bright-field imaging of the bilayer scaffold. (E) SEM image of the alginate hydrogel (top layer). (F) Viability of L929 and MAEC cultured with the bilayer scaffold measured by MTT assay. (G) Swelling ratio of the bilayer scaffold immersed in PBS over time.

3.2 Differentiation of intestinal organoid by applying mechanical and biochemical cues

Following established protocols[6,7], we isolated the intestinal crypt from mouse intestine. Flow cytometry suggested ~10% of cells were Lgr5+ (Fig. 2A), which is similar to the percentage of stem cell population found in the native intestine[8]. Culturing the isolated crypt cells in Matrigel with the addition of biochemical cues (R-spondin) led to the formation of 3D intestinal organoid in 3 days. Immunostaining showed that these organoids were derived from Lgr5+ stem cells (Fig. 2B). Given that folding was induced by compression of the top hydrogel layer, we simulated the mechanical effect by culturing the organoids under compression. Formation of crypt-like structure, a sign of intestinal organoid differentiation, was observed when organoids were cultured under compression for 5 days in expansion medium containing R-spondin (Fig. 2C). Our PCR results showed that the expression of genes related to differentiation, including Chr-A (marker of enteroendocrine cells), Lyz (marker of Paneth cells), Atoh1 (marker of intestinal secretory cells), and Hes1 (marker of secretory lineage differentiation), was significantly upregulated under compression (Fig. 2D). On the other hand, the expression of mTert, which is related to cell division, was downregulated. This suggested compression could promote organoid differentiation. To verify the findings, we cultured intestinal organoids in the bilayer folded scaffold and examined the expression of certain marker. We detected increased expression of Lyz in the folded scaffold compared with unfolded one by immunostaining (Fig. 2E), which is consistent with the results obtained when organoids were cultured under compression. As the next step, we will further characterize the differentiation of organoids in the bilayer scaffold, and implant the scaffolds in vivo to assess whether the bilayer folded scaffold containing organoids can be used for intestinal regeneration.

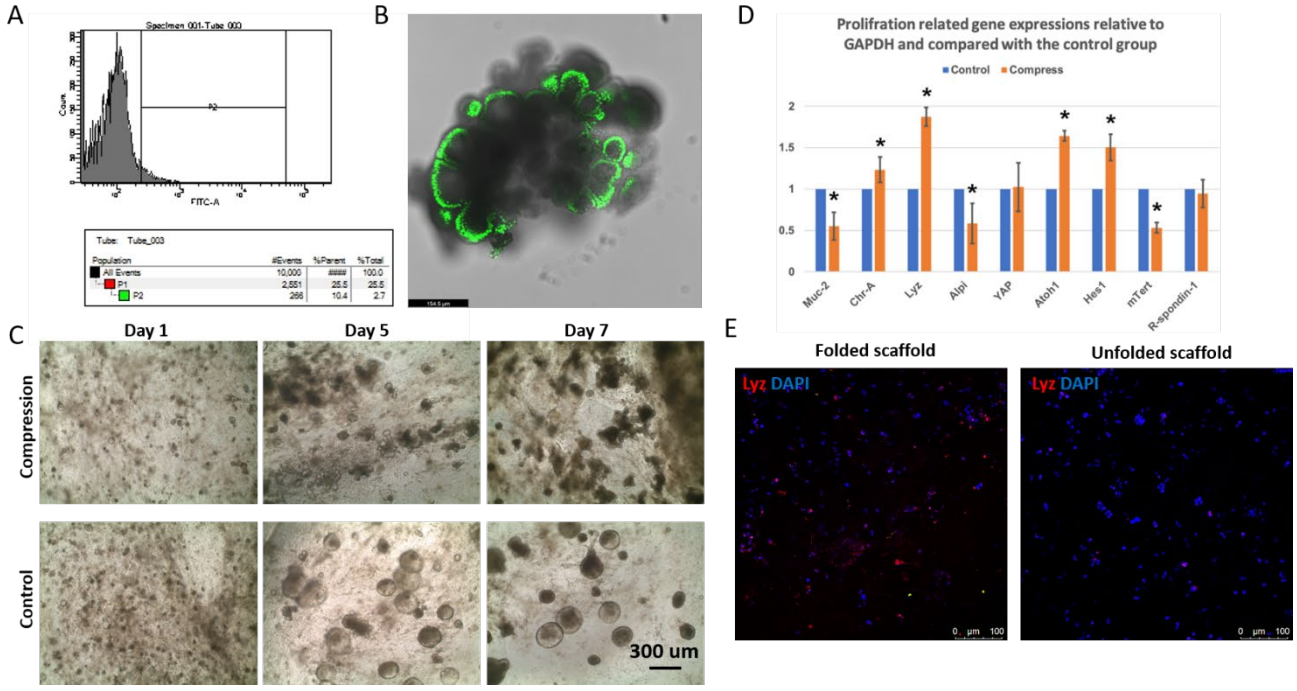


Figure 2. Culture and differentiation of intestinal organoids under biochemical and mechanical cues. (A) Flow cytometry results of cells isolated from intestinal crypt expressing Lgr5. (B) Immunostaining of Lgr5 of intestinal organoids cultured in Matrigel for 7 days. (C) Morphologies of intestinal organoids cultured under compression or in control for 7 days. (D) Expression of various genes related to intestinal differentiation and proliferation measured by PCR. (E) Immunostaining of Lyz in intestinal organoids cultured in bilayer folded and unfolded scaffold for 7 days.

3.3 Next step and commercialization plan

We plan to submit a patent application on the design of the bilayer scaffold after finalizing the protocol of fabrication. We will actively seek financial support from government agencies (e.g., Innovation and Technology Commission)

or private investor(s) to pursue downstream R&D to further characterize the engineered tissue, improve the culture configuration, optimize and scale-up production, and perform quality control of the process. The long-term goal will be to generate functional intestinal tissues for treating SBS. We will demonstrate the engraftment and maturation of the engineered intestinal tissue in small animal models (e.g. mouse), followed by testing in large animals (e.g., pigs). We have established a collaboration with Dr. Guanggai Xia from the Shanghai Jiao Tong University Affiliated Sixth People's Hospital to pursue preclinical and clinical studies.

4. PUBLICATION AND AWARDS

4.1. Publication

J[1] S. Deng, Y. Zhu, X. Zhao, J. Chen, R.S. Tuan, H.F. Chan, "Efficient fabrication of monodisperse hepatocyte spheroids and encapsulation in hybrid hydrogel with controllable extracellular matrix effect" *Biofabrication*, IOP Publishing, 14:015002, 2022

J[2] Y. Zhu, S. Deng, X. Zhao, G. Xia, R. Zhao, H.F. Chan, "Deciphering and engineering tissue folding: A mechanical perspective" *Acta Biomaterialia*, Elsevier, 134 (15): 32-42, 2021

4.2. Conference presentation

C[1] H.F. Chan, R. Zhao, G. Parada, H. Meng, K.W. Leong, L. Griffith, X. Zhao, "Development-inspired engineering of folded mucosa guided by mechanics model" 9th WACBE World Congress in Bioengineering, WACBE, Taipei, 2019

4.3. Award

H.F. Chan, "Young Scholar Award", World Association for Chinese Biomedical Engineers (2019)

5. REFERENCES

[1] P. Seetharam, G. Rodrigues, Short bowel syndrome: a review of management options, Saudi journal of gastroenterology : official journal of the Saudi Gastroenterology Association, 17 (2011) 229-235.

[2] L.Y. Martin, M.R. Ladd, A. Werts, C.P. Sodhi, J.C. March, D.J. Hackam, Tissue engineering for the treatment of short bowel syndrome in children, Pediatric research, 83 (2018) 249-257.

[3] R.G. Spurrier, T.C. Grikscheit, Tissue engineering the small intestine, Clinical gastroenterology and hepatology : the official clinical practice journal of the American Gastroenterological Association, 11 (2013) 354-358.

[4] M.R. Ladd, L.Y. Martin, A. Werts, C. Costello, C.P. Sodhi, W.B. Fulton, J.C. March, D.J. Hackam, The Development of Newborn Porcine Models for Evaluation of Tissue-Engineered Small Intestine, Tissue engineering. Part C, Methods, 24 (2018) 331-345.

[5] H.F. Chan, R. Zhao, G.A. Parada, H. Meng, K.W. Leong, L.G. Griffith, X. Zhao, Folding artificial mucosa with cell-laden hydrogels guided by mechanics models, Proceedings of the National Academy of Sciences of the United States of America, 115 (2018) 7503-7508.

[6] C.N. Grant, S.G. Mojica, F.G. Sala, J.R. Hill, H. Shimada, N.C. Zachos, T.C. Grikscheit, Human and mouse tissue-engineered small intestine both demonstrate digestive and absorptive function, Am J Physiol Gastrointest Liver Physiol, 308 (2015) G664-G677.

[7] E.R. Barthel, A.L. Speer, D.E. Levin, F.G. Sala, X. Hou, Y. Torashima, C.M. Wigfall, T.C. Grikscheit, Tissue engineering of the intestine in a murine model, Journal of visualized experiments : JoVE, (2012) e4279.

[8] X. Zhang, T. Yu, W. Zhong, J. Li, Z. Xia, Y. Yuan, Z. Yu, Q. Chen, Lgr5 positive stem cells sorted from small intestines of diabetic mice differentiate into higher proportion of absorptive cells and Paneth cells in vitro, Development, Growth & Differentiation, 57 (2015), 453-465

COMPLETED

OPTOGENETIC REGULATION OF HORMONE PRODUCTION FOR GLUCOSE HOMEOSTASIS MAINTENANCE

Principal Investigator: Professor Liting DUAN
Department of Biomedical Engineering, CUHK

Research Team Members:
Peiyuan Huang, Ph.D. Student[(1)];
Qiuli Qi [(1)]; Ning Zhang [(1)]; Kandy Yeung [(1)];
Pengfan Zheng [(1)]; Yanqing Xu [(1)]

[(1)]Department of Biomedical Engineering, CUHK

Project Start Date: 1st July 2019
Completion Date: 31st August 2021

INNOVATION AND PRACTICAL SIGNIFICANCE:

Diabetes is a progressive and complex disease featured by chronically deregulated blood glucose levels affecting more than 422 million people in the world. In Hong Kong, one in ten people suffer from diabetes, and the prevalence will continue to increase due to the rapid growth of the aging population. Aberrant insulin production is a hallmark of both type 1 and type 2 diabetes. Since no cure is available yet, current treatment strategies include strict food control along with lifelong regular injections of insulin. Many research efforts are devoted to developing alternative methods to increase insulin concentration for diabetic patients. Novel methods that enable precisely controlled and rapid production of hormones regulating blood glucose levels hold great promise to provide effective treatments. Here we propose to design and construct optogenetic systems that use light to control the production of blood glucose regulating hormones. We expect light-inducible production of insulin or glucagon-like peptide can offer a promising approach to treat diabetes. Besides, we plan to establish an orthogonal optogenetic system to express different hormones under the light stimulation of different wavelengths. We expect the proposed orthogonal system will be able to balance blood glucose levels bidirectionally, thus presenting a great tool to study the glucose homeostasis and the mechanisms underlying diabetes.

1. OBJECTIVES AND SIGNIFICANCE

1.1. Objectives

1. To develop optical control of insulin production in living cells as a potential treatment for type 1 diabetes
2. To develop optical control of glucagon-like peptide (GLP) production in living cells as a potential treatment for type 2 diabetes
3. To develop an orthogonal optogenetic system to independently produce insulin (or GLP) and glucagon for bidirectional maintenance of blood glucose level

1.2. Significance:

The generation of cells producing insulin or other related hormones would provide an unprecedented cell source for drug discovery and cell transplantation therapy targeting diabetes. To this end, presented here is a previously unavailable strategy that uses LIGHT to precisely and rapidly control the production of one or two hormones in engineered cells for glucose homeostasis. The novel optical strategy offers many key advantageous over the currently available cell-based methods to balance the blood glucose level (e.g., transplantation of healthy islet cells

or stem cell-derived pancreatic cells), including (1) Precise control over the production of the related hormones with high spatial and temporal resolution; (2) Fast on-and-off switch of the hormone production; (3) The capacity to bidirectional influence glucose homeostasis by orthogonally inducing the production of two hormones with opposite effects on the blood glucose concentration; (4) Good scalability.

2. RESEARCH METHODOLOGY

Many optical approaches for regulating gene expression have been established based on photo-mediated protein-protein interactions (Figure 1A). Protein A and B will bind to each other upon light illumination. Such light-gated hetero-dimerization allows the optical reconstitution of split transcription factors, where A is fused to DNA binding domain (DBD) and protein B is coupled with transcriptional activation domain (AD). Light stimulation dimerizes DBD and AD, which reconstitute the transcription factors and drive the expression of the gene of interest located after the promoter region.

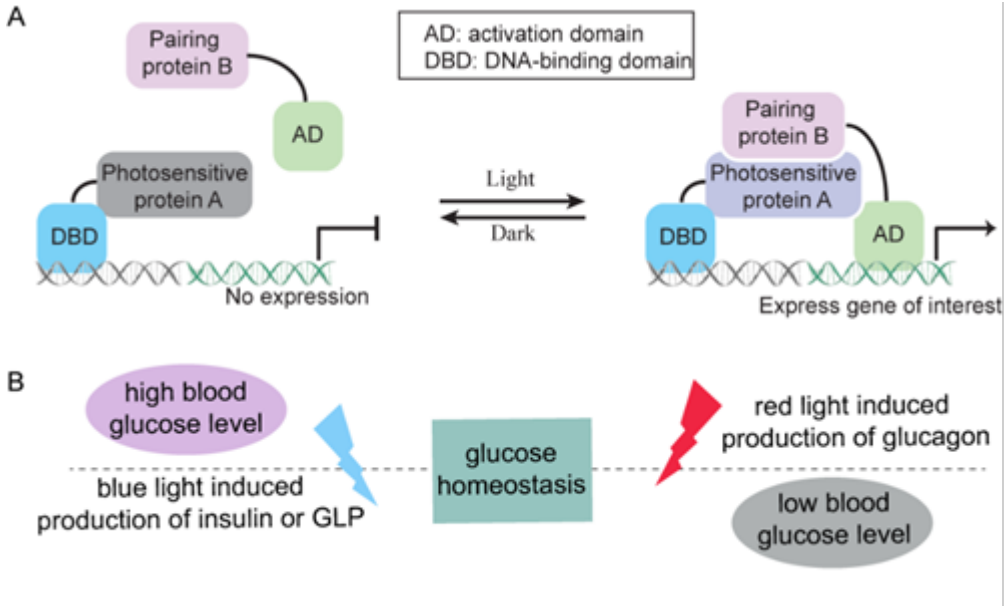


Figure 1. (A) The design of light-induced gene expression system. (B)The design of the orthogonal system to produce insulin and glucagon for directional maintenance of glucose homeostasis.

Task 1: To develop optical control of insulin production in living cells

In this task, we will construct optogenetically controlled expression and production of insulin in engineered mammalian cells as a potential cell transplantation therapy to treat type 1 diabetes. We will adopt an optical strategy that uses blue light to induce the production of the target protein. The sequence of insulin will be incorporated into the system. HEK293T cells will be used to express the designed optogenetic system. The cell culture medium will be collected after transfected cells undergo 1-day blue light stimulation. The insulin level in the medium will be assayed by ELISA.

Task 2: To develop optical control of glucagon-like peptide (GLP) production inside cells

In this task, we will develop an optogenetic system to optically induce the production of glucagon-like peptide (GLP) in engineered mammalian cells as a potential therapy to treat type 2 diabetes. Here we propose to use light to induce the production of GLP in cells to maintain the glucose homeostasis. Similar to the strategy described in Task 1, the sequence encoding a short variant of human GLP-1 (GLP-1) will be inserted into the light-inducible gene expression system. Genetically engineered HEK293T cells will be generated that are transfected with this

optogenetic system. The level of secreted GLP-1 will be assayed after cells are subject to 24-h blue light stimulation.

Task 3: To develop an orthogonal system to produce insulin and glucagon for bidirectional regulation of blood glucose level

In task 3, we plan to develop an orthogonal system in which the secretion of insulin (or GLP) and glucagon can be separately controlled by light of different wavelengths. We will utilize a red light-induced gene expression system to trigger the expression of glucagon. HEK293T cells will express both the optical-controlled insulin (or GLP) production system (described in Task 1 and 2) and the optical-regulated glucagon secretion system (Figure 1B). Blue light required for the expression of insulin (or GLP) (450 nm) is very well separated from the red light (660 nm) indispensable for light-triggered production of glucagon, which allows independent optical control over the production of different hormones.

3. RESULTS

3.1. We have achieved optogenetic control of insulin expression in living cells

3.1.1 Comparison of EL222 and LightOn gene expression system

There are two widely used blue light-mediated gene expression systems available, LightOn and EL222 systems. In the LightOn system, a trans-activator GAVPO contains the VVD domain responsive to blue light, the truncated version of DNA-binding domain (DBD) Gal4 (1-65 aa), and the p65 transcription activation domain (AD). Various target genes can be inserted into the reporter vector after the upstream activating sequence of Gal proteins (UASG). GAVPO rapidly forms dimers upon blue light activation and binds to UASG areas driving the expression of target genes. The other system, VP-EL222, consists of a light-responsive bacterial transcription factor (EL222) fused with a VP16 transcriptional activation domain that will dimerize and bind to DNA to initiate gene expression upon blue light illumination. The reporter gene is placed under the control of five copies of the EL-222 binding clone 1-20 bp (C120) sequence.

First, we compared which system can provide better expression efficiency under our experimental conditions. We constructed reporter plasmids containing the gene encoding mCherry (mCh) fluorescent protein under the control of EL222-binding C120 sequence, or the GAVPO-binding UASG sequence, respectively. HEK293T cells were transfected by each set of plasmids and subjected to the same blue light irradiation conditions (450 nm blue light, 1 mW/cm^2, 1 sec on, 29 sec off). Figure 2A shows the representative images of HEK293T cells illuminated for 24 hours or left in the dark as a control. We quantified the mean of relative cell fluorescence in each group. The LightOn system strongly enhanced mCh expression which is more than 20 times higher than its dark control group. Contrastingly, EL222 systems only yielded a slight increase in mCh expression in the group under the same light illumination (Figure 2B). Thus we chose to use the LightOn system in all the following experiments as the platform to control gene expression by blue light due to its high efficacy.

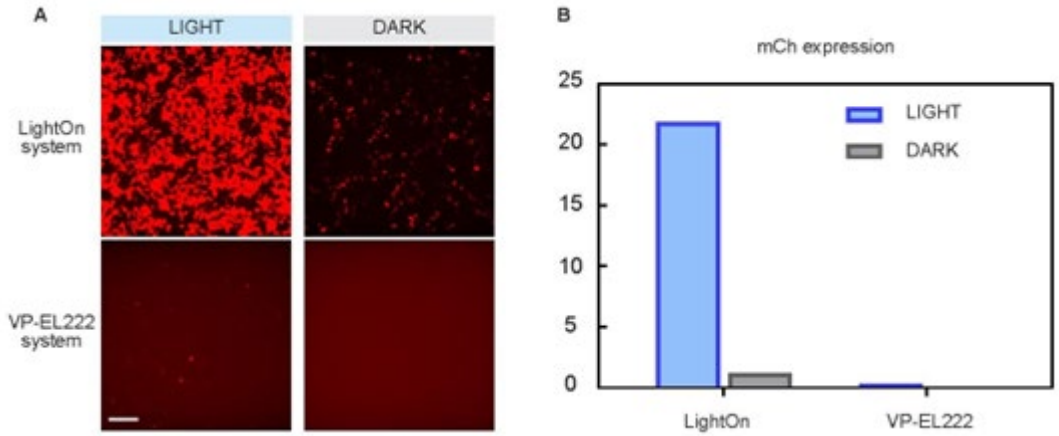


Figure 2. Comparison of two different blue light-inducible gene expression systems. (A) HEK293T cells were transfected by plasmids for the LightOn or EL222 expression system and subjected to blue light irradiation conditions (450 nm blue light, 1 mW/cm^2, 1 sec on, 29 sec off) or kept in dark as a control. The LightOn system led to significantly increased expression of mCh by light, while the EL222 system only resulted in a slight increase. (B) The LightOn system can induce >20 fold change of expression upon light stimulation, quantified by mean fluorescence intensity. Scale bars, 200 μm.

3.1.2 Optogenetic control of IL-10 expression and secretion

To validate the ability of engineered cells to secret target protein molecules, we examined light-controlled secretion of IL-10, an anti-inflammatory cytokine molecule that plays pivotal roles in regulating inflammation and immune response. We also use this light-regulated IL-10 expression to optimize the experimental conditions, because quantitative measurement of IL-10 is relatively easy. The human IL-10 gene was incorporated into the pUASG5 reporter vector, forming the plasmid pUASG5_IL-10. HEK293T cells were co-transfected by GAVPO and the reporter plasmid pUASG5_IL-10 and subjected to blue light stimulation (450 nm blue light, 1 mW/cm^2, 1 sec on, 29 sec off) for 24 hours. Immunostaining results showed that blue light illumination was efficient to upregulate the expression levels of IL-10 cytokine proteins in HEK293T cells. This system is also applicable in other cell lines such as NIH 3T3 fibroblast cells and retinal ganglion cells RGC-5, though the IL-10 expression levels were lower than that in HEK293T cells (Figure 3A). To quantitatively decide the expression fold change of IL-10 molecules, we used the enzyme-linked immunosorbent assay (ELISA) to measure IL-10 concentrations. Upon 12 hours of blue light illumination, HEK293T cells secreted nearly threefold IL-10 compared to the cells that were kept in the dark. Secreted IL-10 concentration was further increased by 4.3-fold after being stimulated by blue light for 24 hours (Figure 3B). Therefore, for all the following experiments, we will exert 24 hr light stimulation for best expression efficiency.

3.1.3 Optogenetic control of insulin expression

After optimizing the conditions for light-induced expression of mCherry and IL-10, we demonstrated the optogenetic control of insulin expression and secretion in HEK293T cells. We constructed the reporter plasmid by inserting the mouse insulin gene downstream of UASG elements (pUASG5_mInsulin). HEK293T cells were co-transfected with GAVPO and pUASG5_mInsulin. Blue light stimulation successfully induced enhanced levels of insulin expression in HEK293T cells, compared with cells kept in dark.

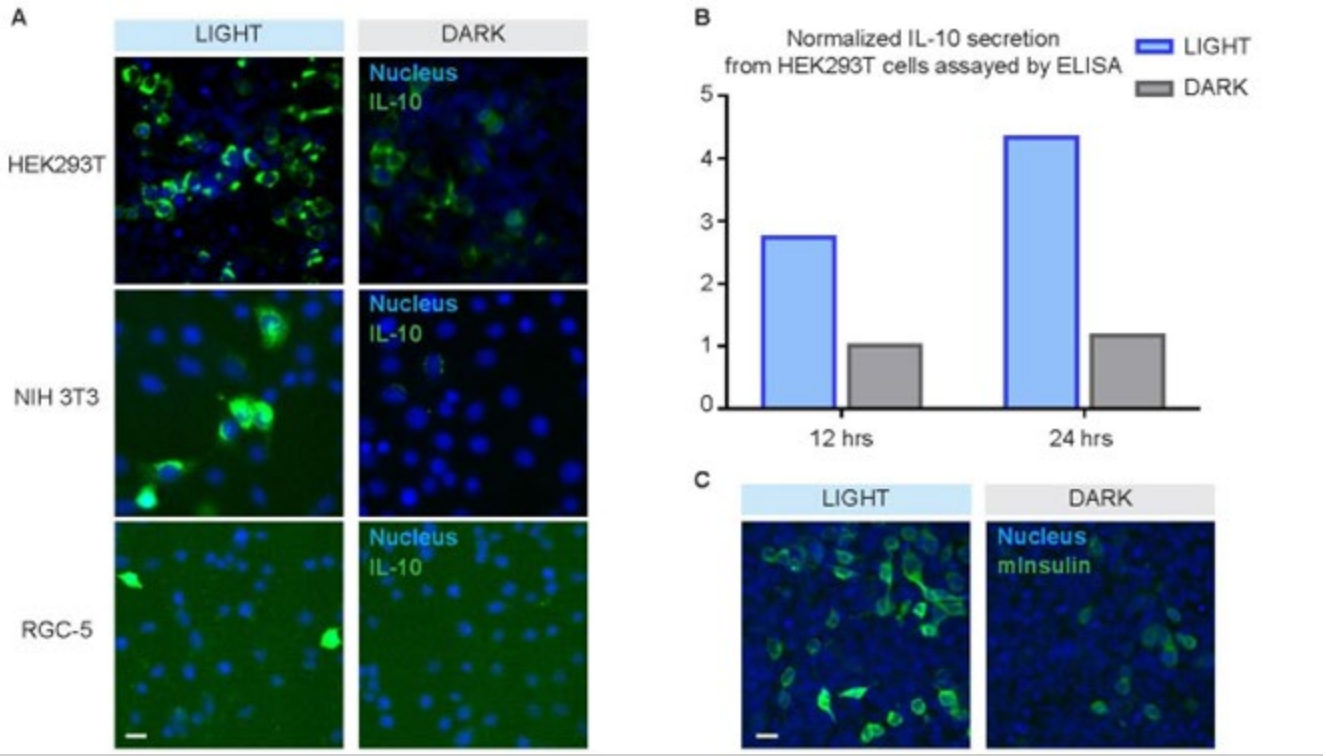


Figure 3. Optical control of target protein expression and secretion. (A) Light-controlled IL-10 expression from LightOn system in different cell lines. (B) ELISA assay of IL-10 secretion from transfected HEK293T cells after 12 and 24 hours blue light activation. (C) Light-controlled mInsulin expression from transfected HEK293T cells. Scale bars, 20 μm.

3.2. We have achieved optogenetic control of GLP production inside cells

In this task, we have achieved optical regulation of a short variant of human glucagon-like peptide 1 (shGLP-1) expression in engineered mammalian cells using the LightOn system. Gene encoding shGLP-1 was inserted into the LightOn system. To monitor the expression levels of GLP-1, we constructed the reporter plasmid by attaching a P2A self-cleavable peptide and green fluorescent protein (GFP) at the C-terminus of shGLP-1, forming pUASG5_shGLP-1-P2A-GFP. Production of peptides will be further cleaved into shGLP-1 peptide and a GFP protein in 1:1 ratio so that the GFP signal can indicate the expression of shGLP-1. HEK293T cells were co-transfected with GAVPO and pUASG5_shGLP-P2A-GFP. One set of cells was subject to 24 hours blue light illumination (450 nm, 1 mW/cm^2, 1 sec on, 29 sec off), while the other set was kept in the dark. The shGLP-1 expression was increased by 20.6-fold compared to the control group kept in the dark as indicated by the GFP signal (Figure 3). The results show that the expression of GLP can be robustly modulated by light via optogenetic methods.

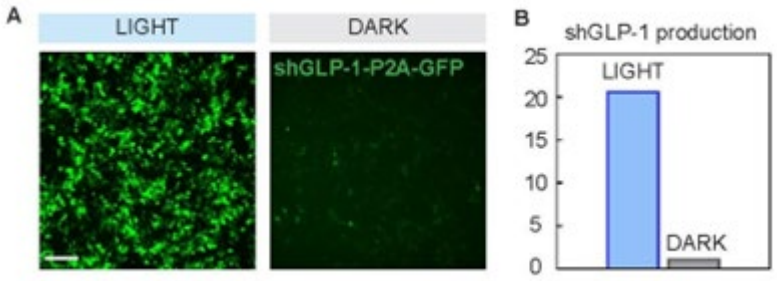


Figure 4. Optical control of shGLP-1 expression. HEK293T cells were co-transfected with GAVPO and pUASG5_shGLP-1-P2A-GFP, the two components of LgithOn system to express shGLP-1. (a) Illuminated HEK293T cells showed a drastic increase of GFP signal compared with the cells kept in the dark, which indicates the greatly enhanced expression of shGLP-1 expression induced by blue light illumination. (b) Normalized shGLP-1 production fold change in engineered HEK293T cells was measured by the average intensity of GFP in HEK293T cells. Scale bar, 200 μm.

3.3. We have achieved orthogonal control of insulin and glucagon production where expression of insulin can be modulated by red light and expression of glucagon can be regulated by blue light

To develop an orthogonal gene expression system, we chose to use a gene expression system that is modulated by near-infrared (NIR) light. The well-separated spectrum between NIR and blue light will permit independent control of expressing two different target proteins. First, we validated a near-infrared light-triggered gene expression system harnessing the NIR light-responsive module BphP1 from bacteria. The system is composed of BphP1 fused to a VP16 transcriptional activation domain and a Gal4-Q-PAS1 that localizes in the cell nucleus and binds to reporter plasmid DNA upon NIR light stimulation. We co-transfected HEK293T cells with the optimized BphP1-Q-PAS1 system and a reporter plasmid pUASG5_mCherry. Fluorescence data revealed that expression levels of mCherry protein were increased under 740 nm pulsed light (10 sec on, 10 sec off, 1 mW/cm^2) after 24 hours (Figure 4A). The result indicated that this BphP1-based method is effective in controlling gene expression by red light.

Next, we investigated the efficiency of red light-induced insulin expression. HEK293T cells co-transfected with BphP1-Q-PAS1 and pUASG5_mInsulin-P2A-GFP were subjected to the same red light stimulation or kept in the dark for 24 hours. Figure 3B shows the insulin expression was increased by 2.7-fold from cells illuminated by 740 nm red light.

To build a bi-directional control of insulin or glucagon expression, we incorporated the red light gene expression system with the blue light gene expression system in the same cell cultures. We inserted the gene encoding the glucagon (1-36 aa) tagged by mCherry fluorescent protein at the N-terminus into pUASG5 reporter, forming pUASG5_mCh-glucagon. One group of HEK293T cells were transfected with the blue light-activatable glucagon expression system (GAVPO and pUASG5_mCh-glucagon), while another group of cells was transfected with the red light-inducible insulin expression system (BphP-Q-PAS1 and pUASG5_mInsulin-P2A-GFP). Cells were harvested and counted 10 hours after transfection. Equal amounts of live cells from the two groups were mixed and plated into two cell culture plates that were placed under blue light or red light stimulation, respectively. After 24 hours, cell samples were fixed and the expression levels of two products were examined by fluorescence microscopy. As shown in Figure 4, the mCh-glucagon levels were increased in the group of cells subjected to blue light illumination, while the red light-stimulated group of cells showed increased signals from insulin expression. These data demonstrated effective multiplexing of the blue and red light optogenetic tools in controlling gene expression. Moreover, by achieving these results, we have demonstrated the orthogonal control of insulin and glucagon by light of different wavelengths.

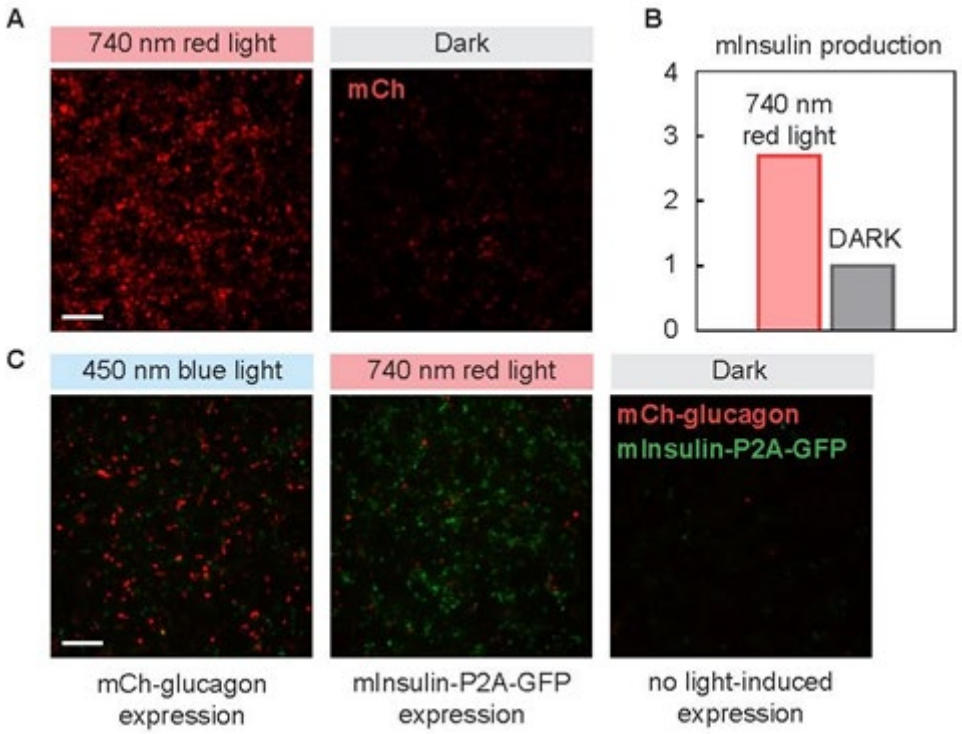


Figure 5. Orthogonal control of glucagon and insulin production. (A) HEK293T cells transfected by BphP1-Q-PAS1 and pUASG5_mCh showed increased mCh expression under 740 nm red light illumination. (B) Normalized mInsulin expression levels of cells under 740 nm red light illumination or kept in the dark. The insulin expression was increased by 2.7-fold from cells illuminated by 740 nm red light, compared with cells kept in dark. (C) Orthogonal control of glucagon and insulin expression under different wavelengths. HEK293T cells transfected by the blue light-activatable glucagon expression system and the red light-activatable insulin expression system were subjected to light of different wavelengths or kept in the dark. Blue light stimulation induces the expression of glucagon, and red light triggers the production of insulin. Scale bars, 200 μm.

4. PUBLICATIONS

J[1] P. Huang, A. Liu, Y. Song, J.M. Hope, B. Cui and L. Duan*, "Optical activation of TrkB signaling", *Journal of Molecular Biology*, 432(13), pp.3761-3770. [2020] link
J[2] Y. Song, P. Huang, X. Liu, Z. Zhao, Y. Wang, B. Cio and L. Duan*, "Light-inducible deformation of mitochondria in live cells", *Cell Chemical Biology*, online. [2021] link
J[3] P. Huang, Z. Zhao and L. Duan*, "Optogenetic activation of intracellular signaling based on light-inducible protein-protein homo-interactions", *Neural Regeneration Research*, 17(1), p.25. [2022] link

COMPLETED

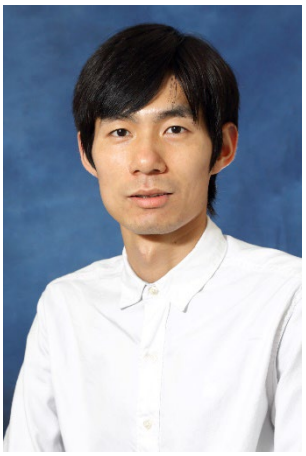
DEVELOPMENT OF HIGHLY SENSITIVE QUANTITATIVE PHASE MICROSCOPY FORLABEL-FREE IMAGING OF NEURONAL NETWORK ACTIVITIES

Principal Investigator: Professor Renjie ZHOU
Department of Biomedical Engineering, CUHK

Co-Investigator (if any): Prof. Vincent Chi Kwan Cheung [2]

Research Team Members:
Yujie Nie, PhD student [1], Mengxuan Niu, PhD student [1],
Yijin Wang, PhD student [1], Kam Pang So, Junior Research Assistant [2]

[1] Department of Biomedical Engineering, CUHK
[2] Dept. of Biomedical Science, the Chinese University of Hong Kong



Project Start Date: 1st July 2018
Completion Date: 30th September 2020

INNOVATION AND PRACTICAL SIGNIFICANCE:

The development of optical recording techniques for mapping action potentials is important as it offers many key advantageous over the traditional electrophysiology techniques (e.g., patch clamping), such as:

a. *Noninvasiveness without physical contacts,*
b. *Spatial and temporal resolvability,*
c. *High-throughput measurement capability.*

However, label-free optical imaging techniques (i.e., without using fluorescent tags), that do not suffer from photobleaching and slow kinetics of fluorescent proteins, still have not been developed for long term and high speed recording of action potential signals on excitable mammalian cells. Among possible label-free techniques, interferometric microscopy, particularly the quantitative phase microscopy (QPM), is promising in satisfying the sensitivity and speed requirements needed for imaging action potentials. The PI, Prof. Renjie Zhou who is an expert on QPM, has recently implemented MEMS-based mirrors and highly sensitive cameras into QPM systems. He recently co-developed a theory for understanding the phase noise limit in such systems, which has led to achieve 10-4 temporal sensitivity with 10 ms temporal resolution in a QPM system. In order to image neuronal action potentials as proposed in this project, we need to further achieve 10-5 temporal sensitivity and 1 ms temporal resolution. Therefore, we will develop a QPM system that integrates the following technical innovations:

a. *A high stable interferometric microscopy design,*
b. *A better usage of the dynamic range of a high well-depth camera,*
c. *Capability of operating in the reflection-mode QPM system to inheritably provide 10x better sensitivity.*

Our system will enable us to image, for the first time, the electrical activities of cultured neurons, such as those from induced pluripotent stem cells, without fluorescence labeling. This work will promote strong collaborations with the School of Biomedical Sciences at CUHK through the Co-I Prof. Vincent Cheung who is a neurobiologist. By mapping the neural network activities of multicellular organisms, e.g., Caenorhabditis elegans, it will

establish our technique as an important neural imaging tool for revealing functional maps of complex nervous systems in the future.

ABSTRACT

With the median life expectancy in Hong Kong and many developed countries rapidly increasing, aging related neurological disorders, such as stroke, Alzheimer's disease, and dementia, are of great social and economic relevance. The challenge in neuroscience lies in the ability to monitor neuronal network activities to study mechanisms of increasingly complex behaviors under normal and disease conditions. These demands of neuroscience can only be met by developing novel neural imaging technologies at single neuron level. Neuronal network activities are characterized by electrical impulses called the action potentials. Fluorescence-based imaging techniques using bright voltage-sensitive dyes can directly map the action potential signals, but they suffer from photobleaching of the fluorescent proteins, preventing them from long-term neuronal network functional studies. Therefore, the development of a label-free optical imaging technique (i.e., without using fluorescent tags) is critical in solving this issue. During this project, we proposed to develop a high sensitivity interference microscopic technique for imaging the action potential signals in neuronal networks. We envision our work will enable many important discoveries in neuroscience.

1. OBJECTIVES AND SIGNIFICANCE

Objectives

We aim to develop label-free optical recording system for monitoring neuronal network activities. To realize this system, we will achieve the following objectives:

1. Design and build a versatile common-path quantitative phase microscopy (QPM) system.
2. Implement a high full-well-capacity camera and develop a filtering scheme.
3. Apply the super-high sensitivity QPM system for mapping the action potential signals of neurons.
4. Extend the QPM system for reflection-mode measurements

Significance

Long-term monitoring of the neuronal activities in target neurons in the living brain is essential for unraveling the organization and functions of the nervous system. However, this is a very challenging task that can only be met by developing novel in vivo neural imaging technologies at single neuron levels. Neuronal activities are characterized by electrical impulses called the action potentials. Traditional electrophysiology based on patch clamping does not offer spatial resolution and it's invasive. Fluorescence-based imaging techniques using bright voltage-sensitive dyes can directly map the action potential signals, but they suffer from photobleaching and phototoxicity of the fluorescent proteins, preventing them from long-term neuronal function studies, and this method cannot be applied to in vivo imaging in human. Therefore, the development of a label-free and noninvasive in vivo optical imaging technique (i.e., without using fluorescent labels) is critical in solving this issue. With unmatched advantages including noninvasiveness, microscopic resolution, and the high throughput measurement capability, we envision this new imaging method will foster many new discoveries in neuroscience and provide great promise for human clinical applications.

2. RESEARCH METHODOLOGY

For each objective, we outline the related methodology as follows:

1. Development of highly sensitive common-path quantitative phase microscopy system

As a foundation step, we will achieve a QPM system with a mechanically stable design. Diffraction phase microscopy (DPM) system is the candidate design for stability. To realize high temporal sensitivity, we will first explore how the mechanical vibration from environment affects the phase sensitivity. To quantify temporal phase sensitivity and spatial phase sensitivity of both QPM systems, we will measure sample-free interferograms under different environmental conditions, including (1) during day & air conditioner on; (2) day & air conditioner off; (3) night & air conditioner on; and (4) night & air conditioner off. Under each condition, multiple interferogram stacks will be recorded for further analysis. Using this analysis, we will determine how does the environmental disturbance

(mainly the mechanical vibrations of the building structure) affect the sensitivity. We will also determine the best achievable sensitivity.

2. Further improvement of sensitivity with high well capacity camera and filtering
In an earlier work done by the PI, a key finding is that the detection sensitivity in QPM is determined by the photon shot noise. Due to the limited full-well-capacity, ~ 30,000 e-, for most cameras, the phase sensitivity noise is worse than 5 milliradiances. Recently, cameras with more 2 million electrons of full-well-capacity has been manufactured. If implemented to our system, 10^{-4} phase sensitivity, i.e., measuring < 10 nanometer displacement on the neuron membrane, could be achieved. However, this is still not enough to observe the action potential signals, thus, a spatiotemporal filtering method will be further developed to select the best detection frequency band. With this filtering method, we expect to achieve 10^{-5} detection sensitivity for observing membrane displacement due to action potentials.

3. Deployment of the system for imaging neuronal networks and dynamics
During the system deployment, we will first focus on the optimization our imaging instrument for neural imaging. We plan to first validate our technology by recording action potentials of neuron-like cells, such as modified HEK 293T cells. For validation, a fluorescent imaging part will be added to the system for detecting the response from the spontaneous spiking. After that, we will attempt imaging cultured mammalian neurons, including those derived from induced pluripotent stem cells, so that our imaging technology can become an efficient and easy-to-use method for characterizing the membrane properties of different types of cultured neurons. After that, we will extend the field-of-view of our system, which will be crucial for us to later map the neuron network of a whole model organism such as that of the C. elegans.

4. Development of reflection-mode QPM for in vivo neural imaging
Due to larger refractive index contrast, reflection-mode QPM design can potentially give 10x better temporal sensitivity than transmission-mode. In our earlier publication (R. Zhou *et al*., Optics Express 25, 130-143 (2017)), we have developed an imaging model that can be used to simulate the depth-sectioning effect in a reflection-mode QPM system. Together with our collaborators at M.I.T., we have recently developed an inverse scattering model that can be used for 3D object reconstruction in the reflection-mode QPM system. As a highly challenging work, it will open new avenues for in vivo optical imaging of the brain without the need of fluorescent labeling.

3. RESULTS ACHIEVED

3.1 Development of highly sensitive common-path quantitative phase microscopy system

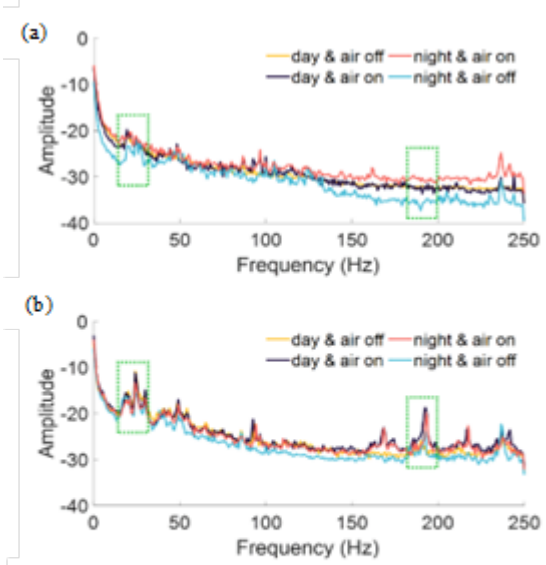


Fig. 2. Averaged frequency spectra of the phase maps measured with (a) common-path and (b) non-common-path QPM systems, respectively. Y-axis is in log scale with unit dB. Selected bandwidth: 15-30 Hz; 185-200 Hz.

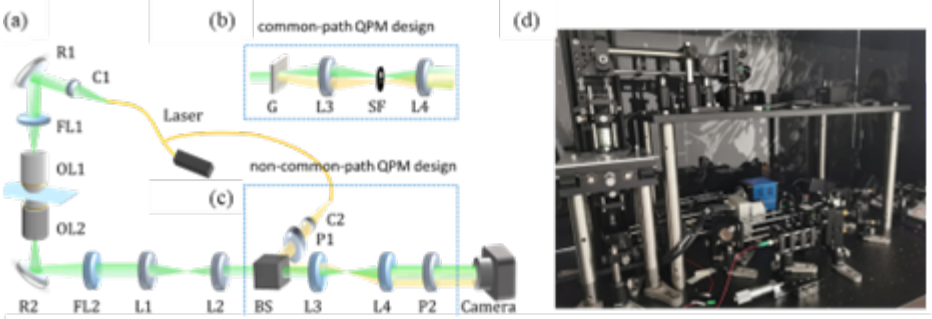
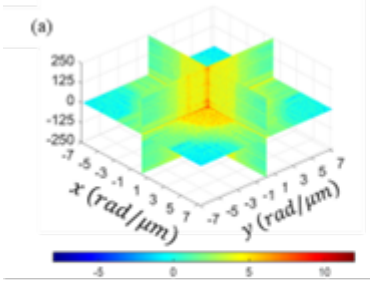


Fig. 1. (a)-(c) Schematic design of the microscopy system, common-path interferometer, and non-common-path interferometer, respectively. (d) Physical setup.

We first constructed two different QPM systems, as schematically shown in Fig. 1 (a)-(c), for exploring sensitivity enhancement. To quantify temporal phase sensitivity and spatial phase sensitivity of both QPM systems, we measured sample-free interferograms under four different environmental conditions: (1) during day & air conditioner on; (2) day & air conditioner off; (3) night & air conditioner on; (4) night & air conditioner off. Under each condition, 25 sample-free interferogram stacks were recorded at 500 frames per second (fps) with an exposure time of 1337 μs. After analysis, we found that common-path QPM system can achieve high sensitivity by isolating most of the environmental disturbance. At condition 4 (night & air conditioner off) when environmental disturbance is minimized, the lowest phase noise occurred, and the best temporal phase sensitivity was around 0.08 nm. We found the environmental disturbance (mainly the mechanical vibrations of the building structure) is prominent at the 15-30 Hz frequency band. In both the low frequency region and the high frequency region, spectrum peaks are observed in the non-common-path system. From the spectrum analysis shown in Fig. 2, we concluded that common-path design achieved better phase sensitivity through isolating both high frequency and low frequency noise induced by mechanical vibrations from the environment.

3.2 Further improvement of sensitivity with high well-capacity camera and a filtering scheme



(b)

Temporal BW (Hz)	Spatial BW (rad/μm)	OPL_t (nm)
15-55	2π	0.0056
	π	0.0044
	π/2	0.0034
180-220	2π	0.0043
	π	0.0031
	π/2	0.0019

Fig. 3. (a) Spatiotemporal spectrum along three different planes in 3d frequency domain. Colormap is in log scale. (b) Bandpass filtering over the selected spatiotemporal bands. BW: bandwidth.

3.2.1 Implement a super-high well capacity camera

Due to the limited electron well depth, ~ 30,000 e^- for most cameras, the temporal phase sensitivity is worse than $2\pi/\sqrt{30000}$ 0.5 nm. By employing a super-high well depth camera (~2 million electron well depth; Q-2A750/CXP, Adimec) in the common-path QPM systems, we were able to achieve a sensitivity of better than 0.05 nm. Further improvement can be achieved by applying frame summing method. The principle is to effectively increase the electron well depth by summing multiple raw interferograms together to create a new summed interferogram for phase retrieval. The temporal phase sensitivity can be further improved to around 0.008 nm, which is a 10x improvement from the original.

3.2.2 Spatiotemporal filtering method for sensitivity enhancement

To further beat the temporal sensitivity limit, we applied a temporal and spatial bandwidth filtering method (spatiotemporal filtering method in short) to diminish the vibrational noise. We draw a 3D spatiotemporal map by taking the Fourier transform along x,y ant t directions of the selected image stack (Fig. 3(a)). The resulted temporal phase sensitivities by selecting different selected temporal and spatial bandwidth are shown in Fig. 3 (b). Through spatiotemporal filtering, the temporal phase sensitivity is pushed down to 2 picometers.

3.3 Deployment of the QPM system for imaging neuronal networks and dynamics

3.3.1 Fluorescence imaging of neuron-like HEK 293T cells

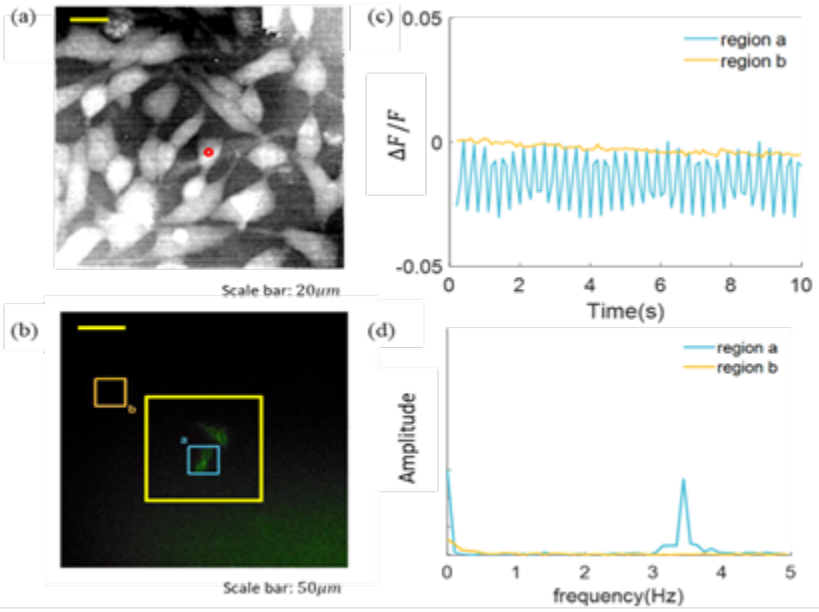


Fig. 4. (a) Phase imaging of HEK 293T cells. (b) Fluorescence imaging of spontaneous spiking HEK 293T cells. (c) Fluorescence responses in the selected regions in (b). (d) The frequency spectrum of the selected regions in (b).

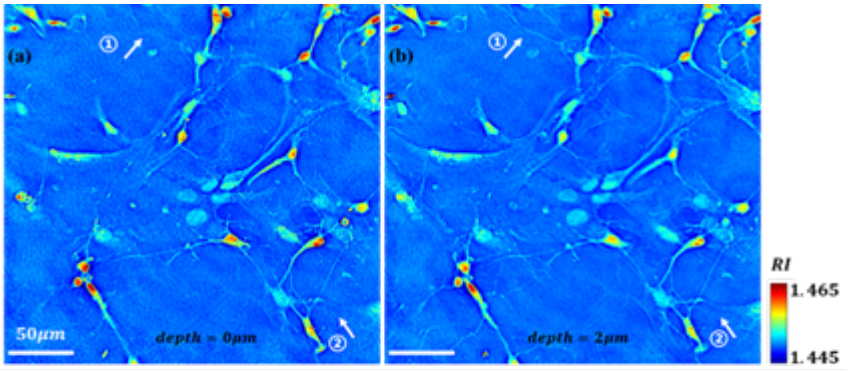


Fig. 5. Imaging of fixed C57 mouse embryonic primary neuron cells and their network. (a) Refractive index map at the focus plane. (b) Refractive index map at the plane above the focal plane. As indicated by the white arrows, connections between two cells show up at different depth.

HEK 293T cell originates from the same precursor line as neurons which is commonly used for neuroscience studies. To image spiking HEK 293T cell (a modified HKE293 cell line), we built an epi-fluorescence imaging system and integrated with our QPM system. For fluorescence imaging, we used a light-emitting diode (LED) with wavelength of 470 nm (M470L4, Thorlabs) as the illumination light source. 470-nm illumination light was filtered at the exit of LED with band-pass filter centered at 475 nm (87-788, Edmund). The results in Fig. 4 has confirmed the spontaneous spiking signal from HEK 293T cells. We are still actively working on matching the fluorescence spiking signals with the recorded phase maps.

3.3.2 Imaging of neuronal network

For imaging multi-cellular organisms and their neuronal networks, we extended the field-of-view of our system to about 300 x 300 μm^2, while maintaining the spatial resolution. We were able to map the neuron network and characterize their properties (Fig. 5).

3.4 Development of the reflection-mode QPM system for in vivo neural imaging

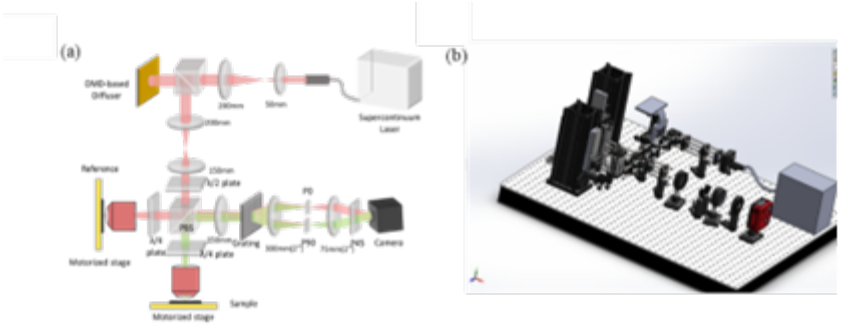


Fig. 6. (a) Schematic design of the reflection-mode QPM system. (b) The 3D layout of the system.

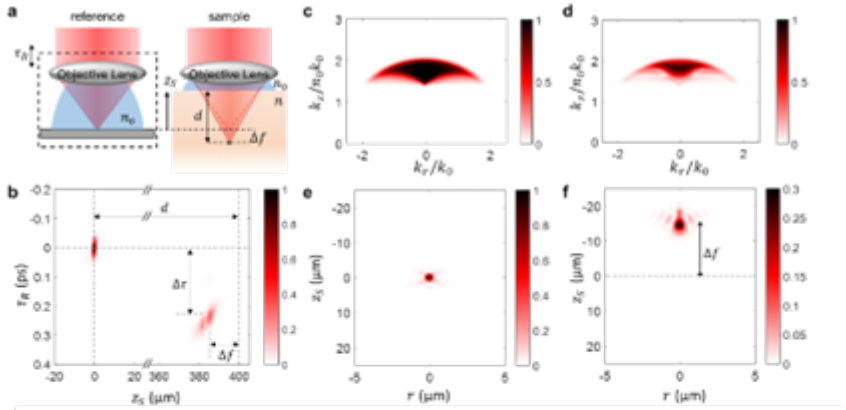


Fig. 7. Inverse scattering model development for the reflection-mode QPM system. (a) Measurement scheme used in the numerical simulation. (b) Point spread function inside the medium, assuming $\boldsymbol{d = (0,\ 400)\ \mu m}$. (c) and (d) Transfer functions of the system in k-space for $\boldsymbol{d = (0, 400)\ \mu m}$ with $\boldsymbol{\tau_R = (0, 233)\ fs}$, respectively. (e) and (f) Point spread functions in the real space, obtained by taking the inverse Fourier transform of transfer functions in c-d, respectively.

According the proposed design, we built a reflection-mode QPM system (Fig. 6). A super-continuum laser (NKT EXR-15) is used as the illumination. A digital micromirror device (DMD) is used as a diffuser to generate dynamic speckle patterns at a high-speed rate. For the illumination source, we used a short-wave infrared light that can achieve better penetration depth up to 10 mm. A NIR camera (model ARTCAM-031TNIR) is placed in the image plane, in which the object is in focus. Another polarizer is introduced in front of the camera to achieve interference between the orthogonally polarized sample and reference beams. We have performed preliminary tests on this system.

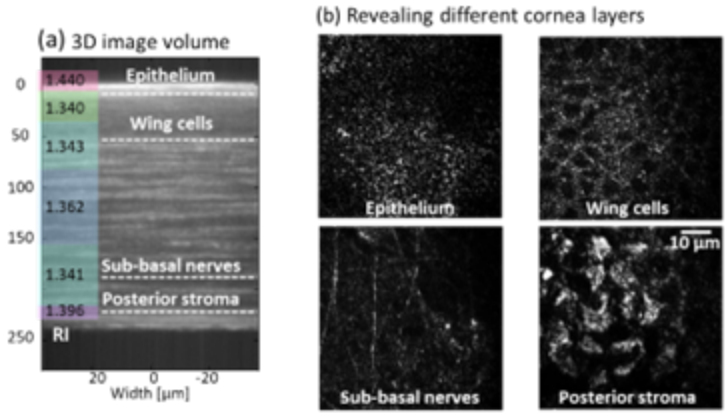


Figure 8. Preliminary in vivo imaging results from a rat cornea. (a) Imaging at different layers in the cornea. (b) En-face images of different depths, where different cell structures are clearly revealed.

At the same time, we developed an inverse scattering model for the system. In the following, we briefly describe the inverse scattering model development. To model light propagation in thick multi-layered structures, we solved the wave equation under a modified Born approximation model. To relate the sample information with the measurement under broadband speckle-field illumination, we derived the transfer function (TF) for achieving rTPM. To illustrate the behavior of TF and point spread function (PSF), we performed a numerical simulation by considering a point scatterer within a single-layered medium and using a water immersion objective with NA=1 for imaging (Fig. 7).

For proof-of-concept demonstration, we collaborated with M.I.T and Prof. Mårten Brelén at CUHK to measure rat cornea. Cornea z-stack intensity images and reconstructed RI in different layers are shown in Fig. 8 (a). En-face images at different depths in the cornea are shown in Fig. 8(b), where different cell structures are clearly revealed. This is a very promising preliminary result that demonstrates the possibility of developing a reflection-mode TPM system for in vivo neural imaging. In the future, through collaborations we will further develop our system for neural imaging.

4. PUBLICATION AND AWARDS

Journal Publications:

[J1] Y. Nie and R. Zhou, "Beating Temporal Phase Sensitivity Limit in Off-axis Interferometry based Quantitative Phase Microscopy," APL Photonics, under revision.

[J2] X. Shu, S. Sansare, D.Jin, X. Zeng, K. Y. Tong, R. Pandey, and R. Zhou, "Artificial Intelligence Enabled Reagent-free Imaging Hematology Analyzer," under review.

[J3] C. Zheng, D. Jin, Y. He, H. Lin, J. J. Hu, Z. Yaqoob, P. T. C. So, and R. Zhou, "High Spatial and Temporal Resolution Synthetic Aperture Phase Microscopy," Advanced Photonics 2, 065002 (2020).

[J4] M. Niu, G. Luo, X. Shu, F. Qu, S. Zhou, Y. P. Ho, N. Zhao, and R. Zhou, "Portable Quantitative Phase Microscope for Material Metrology and Biological Imaging," Photonics Research 7, 1253-1259 (2020).

[J5] R. Pandey, R. Zhou, R. Bordett, C. Hunter, K. Glunde, I. Barman, T. Valdez, and C. Fincke, "Integration of Diffraction Phase Microscopy and Raman Imaging for Label-free Morpho-molecular Assessment of Live Cells," Journal of Biophotonics 12, e201800291 (2019).

Conference Publications:

[C1] Y. Nie, Y. Wang, R. Zhou, "Development of high sensitivity quantitative phase microscopy for label-free imaging of nanoscale dynamics," Proc. SPIE 11249, Quantitative Phase Imaging VI, 112491W (14 February 2020).

Research Reports (2020-2024) In Multimedia Technologies and AI

Newly Funded Projects (2024-2026)	* Building Personalized Multi-modal Large Language Models
Continuing Projects (2023-2024)	* Robustifying Decentralized Training of Deep Learning Models with Fully Stochastic Optimization Algorithms
	* Intelligent Analysis of User Psychological Traits via Online Multi-Modal Social Fingerprint
(2022-2024)	* Intelligent Mobile Robot-Augmented Self-Supervised Wi-Fi Sensing
Completed Project (2021 - 2023)	* A Single-Stage Deep Framework for Fast and Accurate 3D Object Detection and Tracking from Point Cloud
(2020 - 2022)	* Improving Workflow Recognition of Robot-assisted Surgery via Exploration into Future Scenes for Surgical Video Analysis
(2019 - 2021)	* Deep Learning Based Audio-visual Recognition of Cantonese Disordered Speech

(Funded Year)

BUILDING PERSONALIZED MULTI-MODAL LARGE LANGUAGE MODELS

Principal Investigator: Professor YUE Xiangyu
Department of Information Engineering
CUHK

Project Start Date: 1 July 2024

ABSTRACT

Large Language Models (LLMs), e.g. ChatGPT, have become increasingly popular and play a crucial role in Artificial Intelligence. These models excel at understanding and acting upon instructions conveyed in natural language. However, current LLMs can only understand a very limited number of data types (image, text), and cannot provide personalized responses tailored to individual users (e.g. all users use the same ChatGPT model).

To address this, our proposal focuses on developing personalized LLMs in three key areas. Firstly, we aim to create a Multi-modal LLM capable of understanding various data types such as images, text, audio, video, etc. This involves designing a system that can adapt and process multiple data types efficiently. Secondly, we want to improve how these models respond to user queries by incorporating personal knowledge and context into their answers. This will involve retrieval-augmented generation and aligning features across different types of data. Finally, we aim to ensure that these models perform consistently across different domains by developing a method to adapt to changes in data distribution.

Overall, our goal is to make LLMs understand a wide range of data modalities and generate personalized responses, ultimately enhancing their utility in various real-world applications.

INNOVATION AND PRACTICAL SIGNIFICANCE:

The proposed research aims to enhance Large Language Models (LLMs) for personalized applications across various domains. Currently, LLMs e.g. ChatGPT have limitations in understanding multiple data modalities, hindering their effectiveness in real-world scenarios where information exists in diverse formats such as images, text, audio, and more. This project seeks to address this challenge by developing a Multi-modal LLM capable of processing and comprehending different data types. By designing a unified data encoder and implementing a mixture of experts approach, the scalability and adaptability of the model will be significantly improved. Furthermore, the project focuses on enhancing the context-awareness and relevance of responses generated by Multi-modal LLMs through cross-modal retrieval-based approaches and cross-modal contrastive learning methods. Additionally, a domain-adaptive learning approach will be developed to ensure the model's robustness to data distribution shifts. The innovations proposed in this research will enable LLMs to understand and process more data modalities (text, image, video, audio, infrared, MRI, IMU, etc.), and generate personalized responses tailored to individuals. This will have significant practical implications, with applications spanning various sectors including healthcare, education, finance, and beyond.

PROJECT OBJECTIVES:

1. Develop a Personalized Multi-modal LLM: This model will cater to diverse data modalities and user contexts, setting a foundation for AI technologies that are finely tuned to individual preferences and needs, ultimately transforming user interactions across sectors such as education, healthcare, and public services.

2. Create a General-purpose Multi-modal LLM: By supporting a large array of data modalities with computational efficiency, this model will expand the application of AI in creating personalized services and products, significantly impacting technology adoption and the user experience in both personal and professional settings.

3. Incorporate User-Specific Knowledge Efficiently: With this objective, the project aims to make AI interactions more relevant and insightful, enriching user experience and fostering advancements in AI applications that are sensitive to the nuances of individual user histories and preferences.

4. Make the Multi-modal LLM Domain-Adaptive: By considering individual data distribution shifts, the project will result in AI models that are robust to changes in user environments and preferences, ensuring long-term reliability and effectiveness of AI systems in dynamic real-world scenarios.

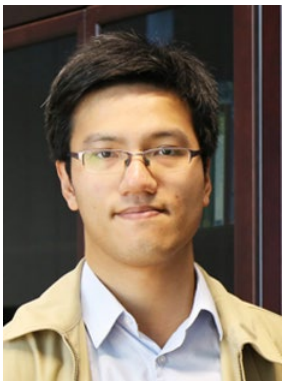
WORK IN PROGRESS

ROBUSTIFYING DECENTRALIZED TRAINING OF DEEP LEARNING MODELS WITH FULLY STOCHASTIC OPTIMIZATION ALGORITHMS

Principal Investigator: Professor Hoi-To Wai
Department of Systems Engineering and Engineering Management, CUHK

Research Team Members:
Chung Yiu Yau, PhD student [(1)],
Haoming Liu, Research Associate [(1)]

[(1)] Dept. of Systems Engineering and Engineering Management, CUHK



Reporting Period: 1 July 2023 – 31 May 2024

INNOVATION AND PRACTICAL SIGNIFICANCE:

This project entails major innovations on algorithm design principles in decentralized DNN training. This is both practical and timely with the recent advocation of federated learning by Google and other leading tech companies. The proposed innovative design tackles a challenging problem of adapting decentralized optimization algorithms to practical network environments via a general and flexible design principle. We anticipate that the results can greatly benefit downstream machine learning (ML) applications on emerging scenarios such as IoT devices, distributed computing, etc. Together with the strong technical component on theoretical analysis of the algorithms, which will be a significant contribution on its own right, the project is anticipated to help advance the state-of-the-art in ML.

ABSTRACT

The popularization of Internet-of-Things (IoT) networked devices has rendered our world more connected than ever. Simultaneously, it raised the challenge of how to leverage the collective intelligences of networked devices in a decentralized manner, e.g., can smart phones cooperate to classify images with their private dataset? can a fleet of autonomous cars coordinate among themselves to reach a certain destination? These challenges have prompted the development of decentralized optimization algorithms which combine training (stochastic gradient) with peer-to-peer communication. Meanwhile, the widespread use of deep neural networks (DNNs) has deployed large models with impressive performance. Training such large models in a distributive manner could waste a lot of bandwidth since these algorithms require devices to communicate the entire model frequently.

This project proposes a fully stochastic optimization framework that adapts to completely random communication graph and compression. Leveraging on the PI's recent works, this project aims to (A) robustify the decentralized training process of DNNs against links failure and/or long latency, (B) advance theories of decentralized stochastic optimization, especially for algorithm designs with compression, and (C) conduct numerical experiments on realistic time varying network to provide empirical evidence that supports the algorithmic/theoretical findings.

1. OBJECTIVES AND SIGNIFICANCE

1. To robustify decentralized training algorithms of deep neural networks (DNNs) against dynamical network environment, we address a key open issue of existing algorithms in training DNNs as

network overheads slow down the training process. We develop a stochastic primal dual framework for decentralized optimization that adapt to dynamical network environments.

2. To advance theories of stochastic decentralized optimization. While partial results have been reported, a *fully stochastic* algorithm as Objective 1 that simultaneously adapt to compression and random graphs with link failures has not been studied. Our framework also offers new mechanisms for analyzing similar stochastic algorithms.

3. To deploy and test the proposed framework on realistic network environment to support the theories in Objective 2. Training of DNNs such as ResNet-50, VGGNet-16 will be simulated on a testbed accounting for network latency, etc.

2. RESEARCH METHODOLOGY

This project aims at improving the robustness of decentralized training of DNN via advances in **algorithm design**, **theories**, and **applications**. We inquire the following research questions:

> *Can we suggest a general design principle for **robustified** stochastic decentralized optimization with compression on dynamical graphs? What are the general convergence guarantees? How do the algorithms work in real network environments?*

To formulate the decentralized DNN training problem, we treat each device as a node and consider n nodes by $\mathcal{V} = \{1, ..., n\}$, linked through a connected graph denoted by $\mathcal{G} = (\mathcal{V}, \mathcal{E})$:

$$\min_{\boldsymbol{\theta}_i \in \mathbb{R}^d, i=1,...,n} \; F(\boldsymbol{\theta}_1, ..., \boldsymbol{\theta}_n) := \tfrac{1}{n} \textstyle\sum_{i=1}^n f_i(\boldsymbol{\theta}_i) \quad \text{s.t.} \quad \boldsymbol{\theta}_i = \boldsymbol{\theta}_j, \; \forall \, (i,j) \in \mathcal{E}, \tag{1}$$

where $\boldsymbol{\theta}_i$ is the DNN's weights, $f_i(\boldsymbol{\theta}_i)$ is a local function such as the cross-entropy loss for the local data, both held by node i. Note that $f_i(\boldsymbol{\theta}_i)$ is non-convex in general and is difficult to compute as it encapsulates a lot of training samples.

To tackle (1), notice that the main challenge is the consensus constraint, which can be represented as a linear equality: set $\boldsymbol{\theta} = (\boldsymbol{\theta}_1; \cdots; \boldsymbol{\theta}_n) \in \mathbb{R}^{nd}$ and $\boldsymbol{B} \in \mathbb{R}^{|\mathcal{E}| \times n}$ be an incidence matrix whose rows correspond to the edges of $\mathcal{G}$. The requirement $\boldsymbol{\theta}_i = \boldsymbol{\theta}_j, \; \forall \, i, j \in \mathcal{E}$ is equivalent to $(\boldsymbol{B} \otimes \boldsymbol{I}_d)\boldsymbol{\theta} = \boldsymbol{0}$. The **key idea** is that $\boldsymbol{B} \otimes \boldsymbol{I}_d$ can be replaced by its randomization $\boldsymbol{B}(\xi) \otimes \boldsymbol{I}(\xi)$:

$$(\boldsymbol{B} \otimes \boldsymbol{I}_d)\, \boldsymbol{\theta} = \boldsymbol{0} \iff \mathbb{E}[\; \underbrace{\boldsymbol{B}(\xi)}_{\text{random graph}} \otimes \underbrace{\boldsymbol{I}(\xi)}_{\text{random compression}} \;]\, \boldsymbol{\theta} = \boldsymbol{0}. \tag{2}$$

The linear constraint can thus be interpreted as a *stochastic linear equality* encompassing random graph and compression. With $\boldsymbol{B}(\zeta)$ taken as a random graph, taking $\boldsymbol{I}(\zeta) = \mathrm{Diag}(\boldsymbol{i}(\zeta))$ as $\boldsymbol{i}(\zeta) \sim \mathrm{Ber}(p)^d$ models the case of random sparsification that communicates $pd \ll d$ numbers.

Leveraging on the stochastic equality (2), we consider the stochastic augmented Lagrangian function for (1): with $\boldsymbol{\Theta} = (\boldsymbol{\theta}_1^\top, ..., \boldsymbol{\theta}_n^\top)^\top \in \mathbb{R}^{nd}$

$$\mathcal{L}(\boldsymbol{\Theta}, \boldsymbol{\lambda}) := \mathbb{E}_\xi \left[\frac{1}{n} \sum_{i=1}^n f_i(\boldsymbol{\theta}_i; \xi) + \eta \langle \boldsymbol{\lambda}, \boldsymbol{B}(\xi) \otimes \boldsymbol{I}(\xi) \rangle + \frac{\gamma}{2} \| \boldsymbol{B}(\xi) \otimes \boldsymbol{I}(\xi)\, \boldsymbol{\Theta} \|^2 \right]. \tag{3}$$

Using (3), the *fully stochastic proximal primal-dual algorithm* (FSPPD) algorithm can be built as a primal-dual stochastic gradient method: let $\widetilde{\mathbf{L}}(\xi^{t+1}) = \boldsymbol{B}(\xi^{t+1})^\top \boldsymbol{B} \otimes \boldsymbol{I}(\xi^{t+1})$, for any $t \geq 0$,

$$\begin{cases} \boldsymbol{\Theta}^{t+1} = \boldsymbol{\Theta}^t - \gamma \widetilde{\mathbf{L}}(\xi^{t+1}) \boldsymbol{\Theta}^t - \eta \widehat{\boldsymbol{\lambda}}^t - \alpha \nabla \mathbf{f}(\boldsymbol{\Theta}^t; \xi^{t+1}), \\ \widehat{\boldsymbol{\lambda}}^{t+1} = \widehat{\boldsymbol{\lambda}}^t + \beta \widetilde{\mathbf{L}}(\xi^{t+1}) \boldsymbol{\Theta}^t, \end{cases} \tag{4}$$

where $\alpha, \beta > 0$ are additional step size parameters. Notice that (4) is adapted to random graph with coordinate-wise sparsification and is a special case of the *stochastic forward-backward* algorithm in [1]. Meanwhile, in the special case of (4) with static graph, the former can be used to recover classical algorithms such as ProxPDA [2], DIGing [3]. We anticipate FSPPD to achieve fast convergence as these existing algorithms.

To achieve our objectives and long term goals, the project will be divided into the following work packages to be achieved in order:

1. **Design principle for FSPPD**: In addition to studying the random sparisfication idea as in (2) which constitute a "linear compression" method, we will also study "nonlinear compression" method such as incorporating randomized quantization in the FSPPD framework.

2. **Convergence properties of FSPPD**: We will study the expected convergence of FSPPD using both empirical and theoretical analysis. Our main analysis technique is to investigate the algorithm using a Lyapunov function approach.
3. **Training large DNNs with FSPPD**: The FSPPD algorithm will also be tested on realistic communication network environment.

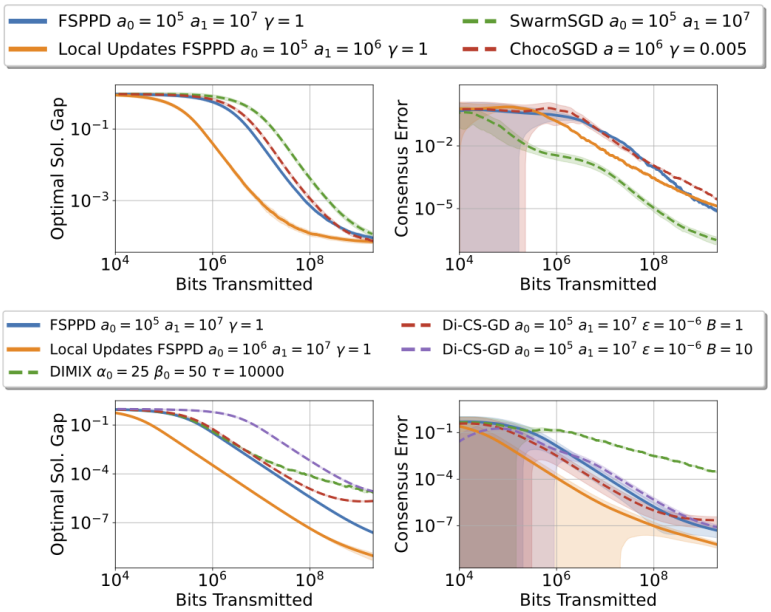
References

[1] P. Bianchi, W. Hachem, and A.Salim. A fully stochastic primal-dual algorithm. Optimization Letters, 15(2):701–710, 2021.
[2] D. Hajinezhad and M. Hong. Perturbed proximal primal–dual algorithm for nonconvex nonsmooth optimization. Mathematical Programming, 176(1-2):207–245, 2019.
[3] A. Nedic, A. Olshevsky, and W. Shi. Achieving geometric convergence for distributed optimization over time-varying graphs. SIAM Journal on Optimization, 2017.

3. RESULTS ACHIEVED SO FAR

The project team is composed of 1 PhD student (Mr. Yau Chung Yiu) and 1 Research Associate (Dr. Liu Haoming). Furthermore, an incoming PhD student (Mr. Rongxin Du from Zhejiang University) will join the team in summer 2024. Over the past year, we have made progresses in algorithm development and theoretical analysis. We list the achievements of this project so far:

1. The paper [C1] has been accepted and presented at the IEEE CDC 2023, Singapore in December 2023. In the paper, we proposed the FSPPD framework using the stochastic linear equality formulation and provided preliminary convergence analysis for the case of strongly convex objective function. Here are numerical experiments presented:

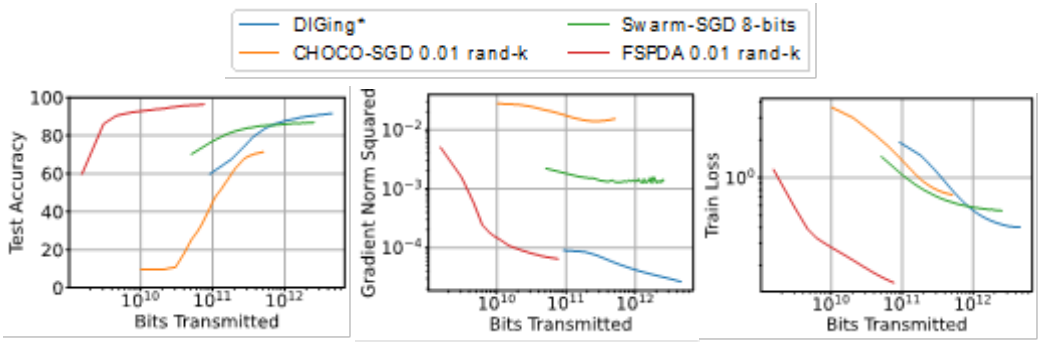


We have tested the algorithms on a simple linear regression problem and showed that it achieves a better communication efficiency compared to SOTA algorithms.

2. We have completed the convergence analysis of FSPPD in the **non-convex optimization** setting. Through a unique stochastic Lyapunov function design, we obtained the following result: under suitable and mild assumptions, it holds

$$\frac{1}{T}\sum_{t=0}^{T-1}\mathbb{E}\left\|\frac{1}{n}\sum_{i=1}^{n}\nabla f_i(\bar{\boldsymbol{\theta}}^t)\right\|^2 = \mathcal{O}\left(\frac{L\sum_{i=1}^{n}\sigma_i^2}{\sqrt{nT}}\right)$$

where the omitted constant and step size criterion depends on the stochasticity of the random graph and sparsification properties. Currently the analysis works only for uniform sparsification and edge selection. We have also extended the numerical experiments for DNN training and achieved the following results:



As a preliminary example, the above experiments are performed on training a 2-layer feedforward neural network in classifying the handwritten digits in MNIST. Observe that the proposed approach outperforms several SOTA algorithms in terms of the communication complexity.

Observe that the proposed FSPPD algorithm has significantly outperformed the competing algorithms in terms of communication efficiency. It has thus partially achieved the planned **Objectives 1, 2, and 3**.

3. We are preparing a journal paper, to be submitted to the IEEE Transactions on Automatic Control, during the summer of 2024. In addition, our results will be presented at an invited special session at the Asilomar Conference, Asilomar, CA in October, 2024.

4. The above result has been further extended to incorporate nonlinear compression schemes (e.g., quantization), thus achieving the second component in **Objective 1.** The idea of the extension is to develop the algorithm through a majorization-minimization approach and integrating with nonlinear gossiping algorithm.

5. Currently, we are working on the convergence analysis of the extended algorithms and conducting numerical experiments, thus achieving the second component in **Objectives 2 & 3**. We plan to complete and submit a journal paper by the end of October in 2024.

4. PUBLICATION AND AWARDS

[C1] C.-Y. Yau, H.-T. Wai, “Fully Stochastic Distributed Convex Optimization on Time-Varying Graph with Compression”, in Proc. of IEEE CDC, 2023.

WORK IN PROGRESS

INTELLIGENT ANALYSIS OF USER PSYCHOLOGICAL TRAITS VIA ONLINE MULTI-MODAL SOCIAL FINGERPRINTS

Principal Investigator: Professor CHENG Hong
Department of Systems Engineering and Engineering Management, CUHK

Research Team Members:
SUN Xiangguo, PhD student [(1)],
FENG Zijin, PhD student [(1)]

[(1)] Dept. of Systems Engineering and Engineering Management, CUHK

Reporting Period: 1 July 2023 – 31 May 2024

INNOVATION AND PRACTICAL SIGNIFICANCE:

Innovation
Psychological analysis in online social networks is very challenging but definitely worthy of our investigation. Although traditional psychological studies provided many insightful observations to the inner personality of people, they heavily rely on users' offline surveys or questionnaires and the analyzed results may vary a lot by different experts. It can hardly meet the current demand because nowadays people transfer more of their activities online and they need faster, more accurate, and more timely estimation of their real-time mental status to help them for better decision. As an alternative, online social networks record users' real-time data with more personalized behavior traces and support highly intelligent psychological analysis in a self-service manner. However, online social networks are far more complicated than traditional questionnaires because they usually have no clear indicator of one's personality traits. How to uncover the underlying psychological motivations from various behavior data, and how to deal with the consistency and conflict of different types of data during the psychological study have become two intractable challenges. To this end, we propose to learn the private features of each kind of data, and the shared features across different data modalities. In this way, we can seamlessly integrate the fact that different people usually have different sensibilities to different expression forms, and further improve the reliability of our psychological analysis with the help of multi-modal fusion. To further study the impact of social environments, we introduce the hypergraph learning (a hypergraph allows one edge to connect multiple nodes, which is perfect to model various social environments) to study the interaction between user's inner traits and their online environments.

Significance
Although there are some existing studies trying to detect one's psychological status from textual or visual data, most of them ignore more profound influence among different data modalities, and users'

various social environments. By contrast, this project considers how to deal with the consistency and conflict of multiple data modalities (including but not limited to text, images, network, etc.), and further study the social impact from the hypergraph level. These technical innovations will guarantee our project more feasible to the real-world applications and obtain more insightful findings from user's psychological analysis. The outcome will benefit the whole society from different perspectives as illustrated in the following. For example, the developed system can be easily integrated in online interviewing so that the interviewer and interviewee can both recognize their matching degree between specific jobs and user's inner traits. The online studying system can analyze students' behaviors to customize more personalized teaching plans for different students. Universities can leverage this system to care

about their students' psychological states and provide timely counselling. Online social platforms can also integrate this system to find psychological abnormality to build a healthier cyberspace.

ABSTRACT

Caused by the long-term impact of COIVD-19, there are an increasing number of people suffering from serious mental diseases and students discouraged by low-effective remote study. Thus an important issue is how to improve people's mental healthiness in the post-pandemic era. Under this background, effective psychological analysis would be very crucial and beneficial as it can identify risks of mental diseases, help people understand themselves better, and improve their performance by more personalized suggestions. Unfortunately, traditional psychotherapy services are usually too specialized, expensive and with limited availability, which is far from flexible and efficient. There is a pressing need to develop more intelligent technologies which can be easily applied to various applications. By capitalizing on our previous research on online social data mining and interdisciplinary study, this project will develop an AI-driven system to analyze users' online social behavior for their psychological traits and prompt suggestions if they present the high risk of mental diseases. The outcome of this project can support large-scale deployment and usage for online interviewing, online education, remote mental intervention, and so on, and can be easily embedded in many online platforms.

1. OBJECTIVES AND SIGNIFICANCE

Objectives:

1. Learning the underlying psychological patterns from user's online behaviors.
2. Handling the consistency and conflicts of different data modalities to improve the reliability of psychological analysis.
3. Studying the social interaction influence with user's psychological traits.
4. Developing friendly systems/plug-ins for psychological self-evaluation and more applications.

Significance:

This project studies how to deal with the consistency and conflict of multiple data modalities (including but not limited to text, images, network, etc.), and further study the social impact from the hypergraph level. These technical innovations will make our project more feasible to the real-world applications and obtain more insightful findings from user's psychological analysis. The outcome will benefit the whole society from different perspectives. Online social platforms can integrate this system to find psychological abnormality to build a healthier cyberspace.

2. RESEARCH METHODOLOGY

2.1. Learning Psychological Patterns from Different Types of Records

To learn the underlying psychological patterns, we use deep learning-based models to obtain text, voice, and visual representations from the networks. We then design a prompting component for each channel to reconcile these AI models' prior knowledge and the specific psychological analysis task. In this way, the learning process can be conducted more efficiently with less training burden. This is particularly helpful considering that psychological signals are sparsely labeled in most online social networks.

2.2. Improving the Psychological Analysis by Multi-modal Fusion

We enhance psychological analysis reliability by merging different modal data, focusing on consistency, and addressing conflicts between modalities. We identify shared features across data by finding the largest overlapping subspace in their latent spaces and learning private features by identifying informative components, allowing for an integrated multi-modal data approach.

2.3. Modeling Social Impact and Online Environment for User Psychological Traits

We use hypergraphs to model social interactions and environments online, which can connect multiple nodes with one edge, mirroring real-world social groups. This approach enables us to study psychological transitions and evolving environments, crucial for the early detection of depressive disorders and for aiding users in adjusting their self-perception for better decision-making in real-life social settings.

2.4. An Intelligent System for Self-service Psychological Analysis

We develop a self-service psychological analysis system, which can function independently or as a plugin for services like online education or psychotherapy. This system analyzes users' public information, generating psychological reports. Users can interact with the system to receive personalized advice, promoting immediate and improved psychological well-being.

3. RESULTS ACHIEVED SO FAR

3.1. Hypergraph Clustering

We designed a novel random hypergraph model, HEM, which simplifies the existing random hypergraph model while preserving the essential features of real hypergraphs. The simplification paves the way to an efficient modularity computation for quality clustering. On top of it, we proposed a new hypergraph clustering algorithm PIC. Our method substantially outperformed existing methods in terms of clustering quality, and achieved up to five orders of magnitude faster clustering time. It provides a useful tool for analyzing complex social structures for the analysis of social influence in personality analysis. This work has been published in Proceedings of the ACM on Management of Data 2023 (J[1]).

3.2. Graph Neural Network Prompting

We designed novel multi-task prompting techniques for graph neural networks to bridge the gap between general graph knowledge and specific application needs. By integrating NLP-inspired prompting mechanisms, we can enhance the adaptability of graph models to various tasks without the need for extensive re-training. Specifically, we first unify the format of graph prompts and language prompts with the prompt token, token structure, and inserting pattern. Then, to further narrow the gap between various graph tasks and state-of-the-art pre-training strategies, we study the task space of various graph applications and reformulate downstream problems to the graph-level task. Afterward, we introduce meta-learning to efficiently learn a better initialization for the multi-task prompt of graphs so that our prompting framework can be more reliable and general for different tasks. This approach is instrumental in setting new standards for graph analysis, particularly in how models are pre-trained and fine-tuned for diverse applications. The prompting techniques can be used to abstract important psychological knowledge from complex social relations in our project. This work has been published at KDD2023 Conference (C[1]) and received the Best Paper Award (Research Track).

3.3. Adversarial Reinforcement Learning

Our pioneering work in adversarial reinforcement learning for scoring systems introduced robust methodologies for evaluating and enhancing the scoring mechanisms. We proposed a “counter-empirical attacking” mechanism that can generate “attacking” behavior traces and try to break the empirical rules of the scoring system. Then an adversarial “enhancer” is applied to evaluate the scoring system and find the improvement strategy. By training the adversarial learning problem, a proper scoring function can be learned to be robust to the attacking activity traces that are trying to violate the empirical criteria. The “counter-empirical attacking” mechanism allows for the dynamic testing and improvement of scoring systems, making them more resilient against potential manipulations and better suited for real-world applications. Our approach has been proved to be effective across different platforms such as the financial and resource management systems, demonstrating its versatility and impact. Based on this technique, we can score the mental status for an online user and give a quantitative analysis. This work has been accepted by IEEE Transactions on Knowledge and Data Engineering (TKDE) (J[2]) and will appear soon.

3.4 Survey on Using LLMs in Graph Learning Tasks

Recently, Large Language Models (LLMs) have been leveraged in graph learning tasks to surpass traditional Graph Neural Networks (GNNs) based methods and yield state-of-the-art performance. In this survey, we presented a

comprehensive review and analysis of existing methods that integrate LLMs with graphs. First of all, we proposed a new taxonomy, which organizes existing methods into three categories based on the role (i.e., enhancer, predictor, and alignment component) played by LLMs in graph-related tasks. Then we systematically surveyed the representative methods along the three categories of the taxonomy. Finally, we discussed the remaining limitations of existing studies and highlight promising avenues for future research. This work has been accepted by IJCAI 2024 Conference (C[2]).

4. PUBLICATIONS AND AWARDS

4.1. Publications

J[1] Zijin Feng, Miao Qiao, Hong Cheng. "Modularity-based Hypergraph Clustering: Random Hypergraph Model, Hyperedge-cluster Relation, and Computation," Proceedings of the ACM on Management of Data, 1(3), 1-25, 2023.

C[1] Xiangguo Sun, Hong Cheng, Jia Li, Bo Liu, Jihong Guan. "All in One: Multi-task Prompting for Graph Neural Networks," KDD, ACM, Long Beach, pp. 2120-2131, 2023.

J[2]. Xiangguo Sun, Hong Cheng, Hang Dong, Bo Qiao, Si Qin, Qingwei Lin. "Counter-Empirical Attacking based on Adversarial Reinforcement Learning for Time-Relevant Scoring System," Accepted by IEEE Transactions on Knowledge and Data Engineering, 2024.

C[2] Yuhan Li, Zhixun Li, Peisong Wang, Jia Li, Xiangguo Sun, Hong Cheng, Jeffrey Xu Yu. "A Survey of Graph Meets Large Language Model: Progress and Future Directions," IJCAI, Jeju, 2024.

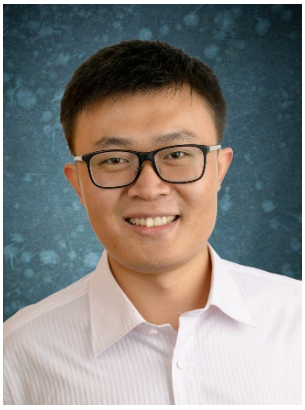
4.2. Awards

C[1] received the **Best Paper Award (Research Track)** at KDD 2023 Conference. This is the first time in Hong Kong and Mainland China to receive this award (**News Release**).

WORK IN PROGRESS

INTELLIGENT MOBILE ROBOT-AUGMENTED SELF-SUPERVISED WI-FI SENSIN

Principal Investigator: Professor He CHEN
Department of Information Engineering, CUHK



Co-Investigator (if any):
Dr. WU Chenshu, Assistant Professor [(1)]

Research Team Members:
Dr. ZHANG Rui, Postdoctoral Fellow [(2)],
Ms. HE Jing, Research Assistant [(2)]
Ms. WANG Qijia, Research Assistant [(2)],
Ms. HUANG Bingyuan, Research Assistant [(2)]

[(1)] Dept. of Computer Science, HKU
[(2)] Dept. of information Engineering, CUHK

Reporting Period: 1st June 2022 – 31st May 2023

INNOVATION AND PRACTICAL SIGNIFICANCE:

Intelligent robots are entering our daily lives from laboratories and playing increasing roles in housework, home security monitoring, and indoor entertainment. These robots are typically equipped with Wi-Fi modules for connecting them to the Internet. Meanwhile, Wi-Fi sensing technologies are becoming more mature with the fast developments in the past ten years. The proposed marriage between intelligent mobile robots and Wi-Fi sensing has great potential to accelerate the broad applications of both technologies in homes for boosting people's safety and security. On the one hand, Wi-Fi sensing-augmented mobile robots could become more favorable thanks to the added value brought by the existing Wi-Fi module. On the other hand, the mobile capability of intelligent robots serves as an effective means to address Wi-Fi sensing's hurdle of complicated environment-dependent algorithm optimization for different homes, essential to make Wi-Fi sensing ubiquitous. Nevertheless, such a marriage is not technically straightforward. Instead, the marriage presents new technical challenges and calls for the design and validation of new signal processing and sensing algorithms.

To realize such a promising marriage between intelligent mobile robots and Wi-Fi sensing, we will develop a clean slate mobile robot-based self-supervised Wi-Fi sensing framework in this project. The proposed framework will leverage intelligent robot's onboard sensors and mobile capability to double-check the events detected by Wi- Fi sensing algorithms. The proposed algorithms can perform continuous and automated adaptation in a self-supervised manner towards higher and higher detection accuracy in homes with different environments. Furthermore, we will tackle a brand new and challenging problem unique to mobile robot-based Wi-Fi sensing– – sensing while moving. Specifically, the robot will be enabled to remove the impact of its self-movement to Wi- Fi signals before detecting other targeted movements.

Our preliminary research shows that the off-the-shelf Wi-Fi devices can detect indoor motions, and the optimization of Wi-Fi sensing algorithms for different environments can be self-supervised by fusing visual data. The proposed research, if successfully executed, will push the practicality of Wi-Fi sensing technologies to new heights, and accelerate the wide adaptation of intelligent mobile robots in homes, thereby boosting the safety and security of people whole live alone, particularly older adults during the pandemic.

ABSTRACT

This project aims to develop and validate a clean slate self-supervised Wi-Fi sensing framework through building prototypes. For the first time, our framework proposes to employ the increasingly popular in-home intelligent mobile robots for addressing a pain point of state-of-the-art Wi-Fi sensing technology–Sensing algorithms need to be manually calibrated for different environments (e.g., homes) to guarantee the detection accuracy. Wi-Fi sensing technologies, which leverages advanced machine learning (ML) algorithms to analyze and interpret Wi- Fi signals for the realization of human activity recognition (e.g., falls of older adults), health monitoring, and object detection, have advanced substantially in the past ten years and have been commercialized recently. However, today's Wi-Fi sensing systems, due to their environment-dependent nature, need to be carefully tuned for different homes to ensure their detection accuracy. Such tuning can be complicated and require frequent human interventions, hindering the broad penetration of Wi-Fi sensing. To circumvent the pain point, this project will devise new algorithms to leverage the onboard smart sensors (e.g., camera) and mobile capabilities of in-home intelligent robots to realize self-supervised optimization of Wi-Fi sensing algorithms with almost no human intervention. For example, a robot, which detects a human presence (i.e., a potential intrusion) in an empty home from its Wi-Fi sensing module, will move to the suspicious area to confirm it through other onboard sensors before raising the alarm. This project extends a seed project funded by the Worldwide Universities Network.

1. OBJECTIVES AND SIGNIFICANCE

Intelligent robots are entering our daily lives from laboratories and playing increasing roles in housework, home security monitoring, and indoor entertainment. These robots are typically equipped with Wi-Fi modules for connecting them to the Internet. Meanwhile, Wi-Fi sensing technologies are becoming more mature with the fast developments in the past ten years. The proposed marriage between intelligent mobile robots and Wi- Fi sensing has great potential to accelerate the broad applications of both technologies in homes for boosting people's safety and security. On the one hand, Wi-Fi sensing-augmented mobile robots could become more favorable thanks to the added value brought by the existing Wi-Fi module. On the other hand, the mobile capability of intelligent robots serves as an effective means to address Wi-Fi sensing's hurdle of complicated environment-dependent algorithm optimization for different homes, essential to make Wi-Fi sensing ubiquitous. Nevertheless, such a marriage is not technically straightforward. Instead, the marriage presents new technical challenges and calls for the design and validation of new signal processing and sensing algorithms.

To realize such a promising marriage between intelligent mobile robots and Wi-Fi sensing, we will design and prototype an intelligent mobile robot-based self-supervised Wi-Fi sensing system. More specifically, the project has the following three interconnected objectives:

1. To develop intelligent mobile robot-based self-supervised Wi-Fi sensing algorithms, which enable the robot 1to detect and localize indoor motions via analyzing Wi-Fi signals, confirm detected anomalies through other onboard sensors like cameras, and automatically calibrate the underlying Wi-Fi sensing model for different environments in a self-supervised manner.
2. To devise more robust Wi-Fi sensing algorithms that can detect human motions even when the robot is moving.
3. To design and prototype a self-supervised Wi-Fi sensing system based on the TurtleBot2 platform available in PI's lab to validate and demonstrate the effectiveness and robustness of the proposed algorithms.

This project will demonstrate a novel union between intelligent in-home mobile robots and Wi-Fi sensing. To the best of our knowledge, the proposed union will be first-of-its-kind and thus has great potential to open a new research line in the field. Furthermore, we believe this is a "win-win" union that will accelerate the widespread penetration of both technologies into every household, as Wi-Fi networks have achieved so far. The proposed research, if successfully executed, will push the practicality of Wi-Fi sensing technologies to new heights, and accelerate the wide adaptation of intelligent mobile robots in homes, thereby boosting the safety and security of people whole live alone, particularly older adults during the pandemic.

2. RESEARCH METHODOLOGY

To circumvent the identified problems in the state-of-the-art Wi-Fi sensing systems, we will develop an

interdisciplinary approach to automate the anomaly confirmation process in this project. Specifically, increasingly popular intelligent in-home robots will be employed to execute anomaly confirmation through leveraging their mobility and onboard sensors such as cameras. Our approach will enable a self-supervised calibration of wireless sensing algorithms for different environments with almost no human intervention. We will adopt the following research plan and methodology to accomplish the project if funded:

To conduct research for delivering Objective 1: The first objective is to develop intelligent mobile robot-based self-supervised Wi-Fi sensing algorithms, which facilitate the following functionalities of the robot: (1) detect and localize indoor motions, (2) confirm detected anomalies with onboard sensors, and (3) automatically calibrate the Wi-Fi sensing model for different environments in a self-supervised manner. We propose a two- step signal processing scheme to be computationally efficient for realizing the first functionality. In the first step, we estimate the target speed by evaluating the autocorrelation function of the channel state information (CSI) extracted from the Wi-Fi module. We then combine the speed values extracted from each subcarrier and receiving antenna with the maximum ratio combining (MRC) strategy. After that, we will apply a recurrent neural network model (RNN) with long short-term memory (LSTM) blocks to exploit the temporal relationship of target speed and thus detect the suspicious anomalies. Once an anomaly is detected, in the second step, we will execute the multiple signal classification (MUSIC) algorithm to estimate the angle-of- arrival (AoA) of incoming signals incurred by the suspicious motion. We will integrate open-source visual processing algorithms for realizing the second functionality in our framework for executing anomaly confirmations. When it comes to the third functionality, after each anomaly confirmation, the corresponding temporal speed sequence and anomaly confirmation result will self-supervise the training of RNN. Built on the initial training of the RNN that can be done based on standard templates generated by band-relaxed segmental local normalized dynamic time warping (SLN-DTW) and DTW barycenter averaging (DBA) algorithms, the proposed anomaly confirmation result can automatically calibrate the RNN- based Wi-Fi sensing model efficiently.

To conduct research for delivering Objective 2: Our framework raises a brand-new technical problem for Wi- Fi sensing—sensing while moving. This distinguishes our work from previous designs of Wi-Fi sensing systems, where the Wi-Fi signal transceivers are considered to stay in a fixed location. Any motion of the transceiver will make the information extraction much more challenging. This is because Wi-Fi sensing relies on extracting information from the signals reflected by the target (e.g., human), and reflections are weak. Nevertheless, in our project, the in-home robot can be a platform for multiple tasks, thus cannot be static for most time. Furthermore, when a suspicious motion is detected, and the robot is on the way to confirm it, the sensing functionality should be maintained in case another abnormal event occurs. We thus need to develop new sensing algorithms that enable the continuous monitoring capability for the intelligent robot. To that end, inspired by the successive interference cancellation (SIC) technique, which is used in wireless communications and allows a receiver to decode two or more data packets that arrived simultaneously, we will develop a SIC-alike Wi-Fi sensing algorithm that allows an intelligent robot to detect a potential human motion masked by its self-movement. More specifically, by leveraging the onboard sensor like IMU, the robot will be able to estimate its current velocity, carefully characterize the effects of its self-movement to the received Wi-Fi signals, and then remove the effects before detecting other targeted movements. We will design and compare both model-based and model-free approaches to characterize the impacts of self- movement.

To conduct research for delivering Objective 3: Built on the mobile robotic platform established in the PI's lab, the project team will design and implement a mobile robot-based self-supervised Wi-Fi sensing system prototype to evaluate and showcase the effectiveness and robustness of the proposed algorithms in Objectives 1-2 in real-world environments. To that end, the team will substantially modify the Linux kernel of the mini- PC that serves as the "brain" of the mobile robot, so that it will output the CSI of each received Wi-Fi packet to the application layer. After that, we will implement the proposed algorithms on the application layer and build a tunnel between our application with the Robot Operating System (ROS) so that the outputs of our algorithms can be used to control the robots' movements.

3. RESULTS ACHIEVED

At the time of writing this report, we have made significant progress in detecting human motion on mobile robots using Wi-Fi signals. The system we built allows a moving robot to detect the presence of an individual behind a corner and estimate their movement directions, effectively preventing potential collisions. This effort represents the first passive Wi-Fi detection system specifically designed for mobile robots, which does not require direct line-of-sight observation of the target. This technology can be applied to various applications, such as fall detection. By identifying anomalies beyond its field of view, the robot can respond to a diverse range of human actions in complex indoor environments and proactively approach suspicious areas when further investigation is necessary. A photo of the developed experimental platform is shown to the right.

The mobile robot detects a person behind the corner by estimating AoA of received Wi-Fi signals using the MUSIC algorithm. The presence of a person introduces a new value to the AoA estimation result compared to a scenario without a person. However, the AoA estimation results are influenced by both human and robot movements, making it challenging to determine the human's movement direction. The challenge of predicting potential collisions as the robot moves towards the corner lies in isolating human motion data from the received Wi-Fi signals. We employ principal component analysis (PCA) on the CSI extracted from the Wi-Fi module to identify the most stable amplitude path, designating it as the reference path consisting only of the robot's motion. Drawing inspiration from the SIC technique, we develop a robot movement removal algorithm that filters the robot's motion from the entire CSI dataset, resulting in CSI data containing solely human motion information. We then use the mean of the correlation coefficients between all adjacent subcarriers in the CSI dataset to revel the person movement directions. Specifically, the correlation coefficients reflect the fluctuation of the path through which signals propagate from the transmitter to the receiver. It increases as the propagation length shortens and decrease as the path length increases.[1]

Fig. 1 displays the correlation coefficients of raw CSI data and the data with robot self-motion removal in three scenarios: no person behind the corner, a person moving towards the corner, and a person moving away from the corner. The mobile robot moves towards the corner from another direction in all three scenarios and thus has no line-of-sight view of the person. Fig. 1 reveals that human movements are obscured by the robot's motion in the raw data, causing the correlation coefficient to increase in all three scenarios. After implementing our proposed algorithm, the correlation coefficient remains constant when no person is present, increases as the person approaches the corner, and decreases when the person moves away, confirming the robot's passive sensing capabilities while in motion.

[1] Y. Hu, M. Z. Ozturk, B. Wang, C. Wu, F. Zhang and K. J. R. Liu, "Robust Passive Proximity Detection Using Wi-Fi," in IEEE Internet of Things Journal, vol. 10, no. 7, pp. 6221-6234, 1 April 2023, doi: 10.1109/JIOT.2022.3224701.

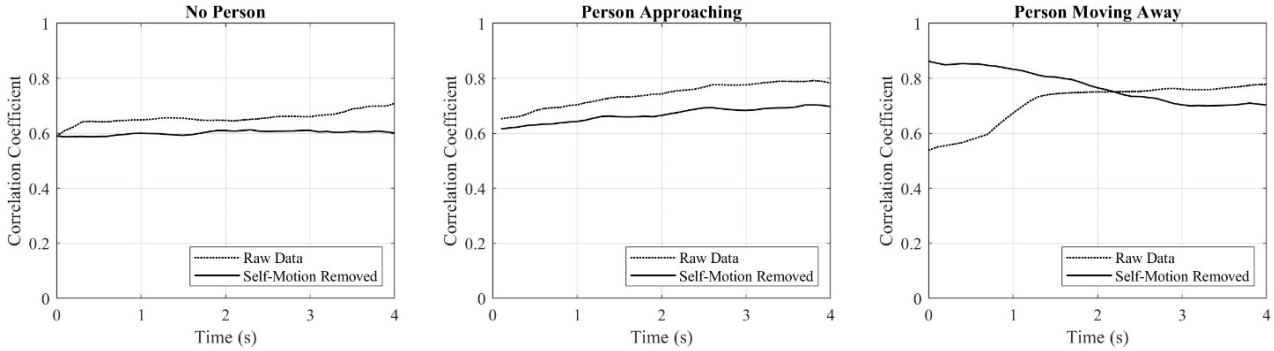


Fig. 1 The subcarrier correlation coefficients of three considered scenarios.

In conclusion, we have successfully developed a robust Wi-Fi sensing algorithm capable of detecting human motion even when the robot is moving. This accomplishment brings us closer to fulfilling the second project objective and more than half of the third project objective, and lays the groundwork for the remaining tasks. Specifically, with this efficient event-triggering mechanism, the robot can more actively and effectively verify occurrences using other onboard sensors, while simultaneously calibrating its underlying Wi-Fi sensing model autonomously through a self-supervised approach. We are drafting a paper based on the proposed algorithm and plan to submit it to a top conference of the field by August 2023.

In addition to passive Wi-Fi sensing on mobile robots, we are tackling the project objectives from two other angles. Firstly, we have developed a novel algorithm to estimate multipath delays, facilitating more refined Wi-Fi sensing applications. We observed that due to the limited bandwidth of practical wireless systems, a single multipath component may appear as a discrete pulse comprising multiple taps in the digital delay domain. This phenomenon, known as channel leakage, complicates multipath delay estimation and has been largely overlooked in existing Wi-Fi sensing research. To address this issue, we exploit the limited number of paths in physical environments and frame the estimation problem as a sparse recovery challenge. We propose using a sparse Bayesian learning (SBL) method to estimate the sparse vector, effectively determining both the number of physical paths and their associated delay parameters. Our simulation results indicate that our algorithm accurately identifies path numbers and achieves superior precision in path delay estimation and channel reconstruction compared to two benchmarking schemes. The work has been submitted to a signal processing conference for possible publication.

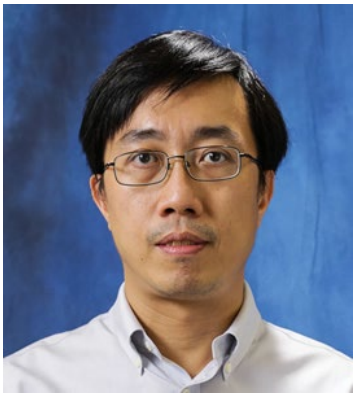
Secondly, besides WiFi sensing for robots, our team has also been investigating new visual sensing techniques for cooperative mobile robots, which loosely connects to the anomaly confirmation part of this project. In this line of research, we propose a new optical ISAC (OISAC) scheme for cooperative mobile robots by integrating camera sensing and screen-camera communication (SCC). As a case study, we consider the leader- follower formation control problem, an essential part of cooperative mobile robotics. The proposed OISAC scheme enables the follower robot to simultaneously acquire the information shared by the leader and sense the relative pose to the leader using only RGB images captured by its onboard camera. We design and conduct real-world experiments involving uniform and non-uniform motions to evaluate the proposed system and demonstrate the advantages of applying OISAC over a benchmark approach that uses extended Kalman filtering (EKF) to estimate the leader's states. Our results show that the proposed OISAC-augmented leader- follower formation system achieves better performance in terms of accuracy, stability, and robustness. The work has been submitted to a wireless networking conference for possible publication.

The PI has presented the core concept of this proposal and demonstrated our preliminary results to Computime Group Limited, a Hong Kong Stock Exchange-listed smart home/building company. They expressed appreciation for our technology, which seamlessly integrates in-home intelligent robots and Wi-Fi sensing, and demonstrated interest in sponsoring us for larger grant applications, such as the Innovation and Technology Fund (ITF). This support would enable our team to file patents and advance the prototype closer to becoming a commercial product.

COMPLETED

A SINGLE-STAGE DEEP FRAMEWORK FOR FAST AND ACCURATE 3D OBJECT DETECTION AND TRACKING FROM POINT CLOUD

Principal Investigator: Professor Chi Wing FU
Department of Computer Science and Engineering, CUHK



Research Team Members:
WANG Tianyu[(1)]; WANG Yinqiao[(1)];
ZHENG Wu[(1)]; HONG Mingxuan[(1)]; WAN Yuxuan[(1)]

[(1)]Department of Computer Science and Engineering, CUHK

Project Start Date: 1st July 2021
Completion Date: 30th June 2023

INNOVATION AND PRACTICAL SIGNIFICANCE:

In this project, we aim to develop an efficient and effective single-stage deep framework for detecting and tracking 3D objects. Specifically, we will focus on the autonomous driving scenario, which is a challenging but highly valuable domain with a huge market and impact. Compared with RGB images, 3D point clouds provide more precise and robust 3D information for handling situations such as fog and night time, and point-cloud-based 3D object detectors have been shown to achieve higher precisions than those based on RGB images. So, in this project, we will focus on designing and developing a single-stage deep framework for the 3D detection and tracking tasks on 3D point clouds. To overcome the limitations of existing point-cloud-based methods, we will aim particularly for both high precision and real-time speed by formulating our framework with novel innovations in four areas: (i) to investigate the neural network architecture; (ii) to investigate new constraint formulation; (iii) to investigate new data augmentation models; and (iv) to investigate new models for robust 3D tracking. These are cutting-edge directions in the research of point-cloud-based detectors and trackers.

As high precision and high speed are essential for practical usage and reliability in 3D perception applications, this research work will promote robustness and smoothness in 3D perception, particularly for enhancing the perception reliability and improving the safety. In the course of the research, we will also carefully perform various experiments with real 3D data to evaluate and demonstrate the precision and efficiency, as well as the practicality, of our approach.

ABSTRACT

3D object detection and tracking are enabling components in 3D perception on real-scanned point clouds and have great values for supporting many practical applications, e.g., autonomous driving, robotics, and logistic planning. While the problem has attracted lots of attention in both the academics and industries, it is particularly challenging, since one has to aim not only for high precision but also high efficiency. Existing methods are either too slow, due to extra overhead in the deep neural network for refining the predictions, or not sufficiently accurate, due to simple network structures for aiming high efficiency. This project offers a new solution to the problem, in which we will develop a knowledge distillation approach to build a single-stage deep framework to aim for both high precision and real-time speed. Our approach has four parts: (i) a lightweight deep framework for extracting robust features without excessive computation overhead; (ii) new

and effective constraints to guide the training and enhance the regression precision; (iii) new data augmentation models to effectively enlarge the sample diversity and to improve the generalizability of the deep framework; and (iv) a self-supervised model for training the framework to track 3D objects in time-varying point clouds with temporal coherence.

1. OBJECTIVES AND SIGNIFICANCE

This project aims to achieve efficient and accurate 3D object detection and tracking by taking a knowledge distillation approach to design and develop a single-stage deep framework. Technically, we have the following four objectives:

1. To articulate a lightweight deep framework with knowledge distillation, through which we train and produce a lightweight backbone to extract robust features without excessive computation overhead;
2. To formulate new and effective constraints to guide the training of the detector and improve the overall regression precision of the predicted bounding boxes;
3. To develop new data augmentation models for point-based 3D scenes, such that we can effectively enlarge the sample diversity to alleviate the overfitting by producing more meaningful augmented samples; and
4. To further extend our deep framework to support efficient and high-quality 3D object tracking in time-varying point clouds.

Our innovation will overcome the limitations of existing approaches and produce the next-generation deep framework that demonstrates both high precision and high speed.

2. RESEARCH METHODOLOGY

This project mainly has four tasks. The detailed research plan and methodology for each task are as follows.

Task-1: Articulate a Lightweight Deep Framework with Knowledge Distillation.
First, we will explore three subtasks to build an effective and lightweight backbone to boost both precision and speed:

i. We will explore knowledge distillation models, upon which we will formulate a lightweight deep framework and supervise it with both hard and soft targets. Compared with manually-annotated hard targets, soft targets often have higher entropy but larger quantity, so we can distill more information from them into the network. Particularly, we will try to distill knowledge from a training network and ensemble it to produce soft targets to further supervise the final network.

ii. Also, we will explore joint tasks training to improve feature representation in the framework. Fundamentally, point-cloud-based 3D object detection, 3D scene segmentation, and 3D object tracking are relevant subtasks for scene understanding. Hence, we will explore different ways of sharing deep features and formulate multiple heads in our framework for joint optimization.

iii. Another area is to cascade a weak detector and a strong detector for easy and hard scenes. Having said that, we will use weak (simple but fast) detectors to process easy scenes that constitute the main part of all the cases, while for the hard ones, we plan to transfer them to subsequent strong (complex but more precise) detectors. We plan to approach the problem by disentangling the task and designing an assigner network to discriminate the input samples. In doing so, we can optimize both the weak and strong detectors to obtain better performance by harnessing weak detectors for high efficiency and strong detectors for high precisions.

Task-2: Formulate Efficient Constraints. Constraints (losses) play important roles in training a deep neural network. To handle issues in 3D point clouds, e.g., distorted object shapes, severe occlusions, viewpoint variations, conventional constraints such as Smooth-L1 may not work well and it is often hard to acquire sufficient information from sparse 3D points for accurately predicting all the box dimensions. These motivate us to focus our effort in the second task on loss design and modeling:

i. First, we will explore a new formulation of IoU loss to improve the overall regression precision of the predicted bounding boxes. Here, we will consider the object's geometric characteristics with a specific formulation to prioritize the box orientation and center alignment in the IoU loss.
ii. Second, we aim to balance the constraints to alleviate overfitting in network training caused by the unbalanced training samples. One of our ideas is to consider the IoU as a weight factor, since often, only a few predicted boxes are lowly-overlapped with the ground truths. Besides, we will consider the box center distance from viewpoint as another weight factor, etc.

The consistency constraint between the soft targets and predictions is another crucial factor in knowledge distillation. To minimize the misalignment, we will formulate a new dynamic pairing strategy as a module in our framework and formulate a new consistency loss for the supervision.

Task-3: Learn to Augment 3D Objects in Scenes. Data augmentation has been heavily employed in tasks related to 2D images, but very little work has been done for augmenting 3D point sets. Founded upon the PI's recent work in CVPR 2020 (which is a new augmentation model for object-level point clouds), we aim to extend the model for augmenting point-based 3D scenes to enlarge the data space and to produce more meaningful augmented data samples to enhance the network's generalizability:

i. First, we will generalize the model in our previous work and formulate an automated strategy to augment objects in point-based scenes by taking into account the object classes and shapes in the scene. Compared with conventional augmentation models, our model can work with point-based 3D scenes and learn to augment scenes instead of simply following pre-defined heuristics.
ii. Second, we will extend the model to be dual scale (scene & object levels). Here, we will first pre-segment each input scene into parts based on the distance to camera and the object orientation, then learn to assemble parts from different scenes to produce new 3D scenes. In object level, we will formulate a relation module in the augmentor network to learn to compose object parts into new objects and also to mimick how 3D point clouds are affected by occlusions, changes in distance, and diverse object shapes.
iii. Third, we will consider learning a probabilistic latent space to enhance the diversity and distinctives of the sampled features for few-shot object detection. In scene level, we will learn geometric prototypes with a variational autoencoder (VAE). In object level, we will cluster the scene-level features into object-level features and learn class-specific prototypes with another VAE. In the end, we use multiple feed-forward network to generate instance-level 3D box positions and dimensions as the detection results.

Task-4: Learn to Account for Temporal Coherency. Tracking aims to predict not only the target object location but also its 3D velocity vector by considering the temporal coherency when processing a temporal-varying 3D point cloud. Compared with static perception in detection, tracking provides a dynamic perception and helps the detector to infer the object movement, benefiting a better analysis and movement prediction. In this task, we will first build a baseline for tracking based on our detection framework by correlating features extracted independently over time, associating the related 3D objects, and learning temporal features from the extracted point features. After that, we will extend the baseline using a self-supervision model. Particularly, one of our insights is that moving 3D objects forward in time and then backward in time should essentially uphold the 3D location of the objects, we thus can formulate new losses, e.g., forward-backward and backward-forward losses, to self-train the framework. This strategy also helps to leverage a vital challenge for processing 4D point clouds, i.e., sample points at a time moment may not perfectly match points in subsequent time moments due to sampling variation and occlusion. Further, we will explore different forms of feature extraction, similarity metrics, and associating algorithm in building our tracking framework.

In this project, we will mainly work on autonomous driving data, which includes not only vehicles but also cyclists, pedestrians, etc. Particularly, we will explore various driving scenarios and aim to deliver a practical system, while aiming for high-quality publications in top international venues.

3. RESULTS ACHIEVED SO FAR

We have completed the method development for all the four tasks and validated the performance of our new methods on various 3D point cloud datasets for object detection, including KITTI, Waymo, and nuScenes; we also validated our methods for 3D hand-mesh reconstruction. Details are as follows:

1) Simulate Multimodality by Knowledge Distillation to Boost the 3D Object Detection Performance. In this project, we developed a new approach to boost a single-modality (LiDAR) 3D object detector by teaching it to simulate features and responses that follow a pre-trained multi-modality (LiDAR-image) detector. Our approach needs LiDAR-image data only when training the single-modality detector, and once the detector is well-trained, it only needs LiDAR data as input during the inference. Our technical contributions in this work are summarized as follows. First, we designed a novel framework called S2M2-SSD (see the illustration of the framework in Figure 1 below) with the following components: the response distillation to focus on the crucial response samples and avoid most background samples; sparse-voxel distillation to learn voxel semantics and relations from the estimated crucial voxels; a fine-grained voxel-to-point distillation to better attend to features of small and distant objects; and instance distillation to further enhance the deep-feature consistency. Our novel approach perfectly meets the sheer practical need of autonomous driving and efficiently boosts single-modality 3D object detection for (i) high efficiency, since our approach needs to process only LiDAR data at inference; (ii) high precision, since our network outperforms the SOTA LiDAR-only detectors; and (iii) high robustness, since our approach is capable of simulating LiDAR-image features for detecting objects in varying lighting conditions, even at night time. Table 1 below shows the experimental results of our method on the large-scale nuScenes dataset. Our approach outperforms all SOTA LiDAR-only 3D detectors and even surpasses the baseline LiDAR-image detector on the key NDS metric, filling ~72% mAP gap between the single- and multi-modality detectors. This work was published in the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2022 as an oral paper.[1]

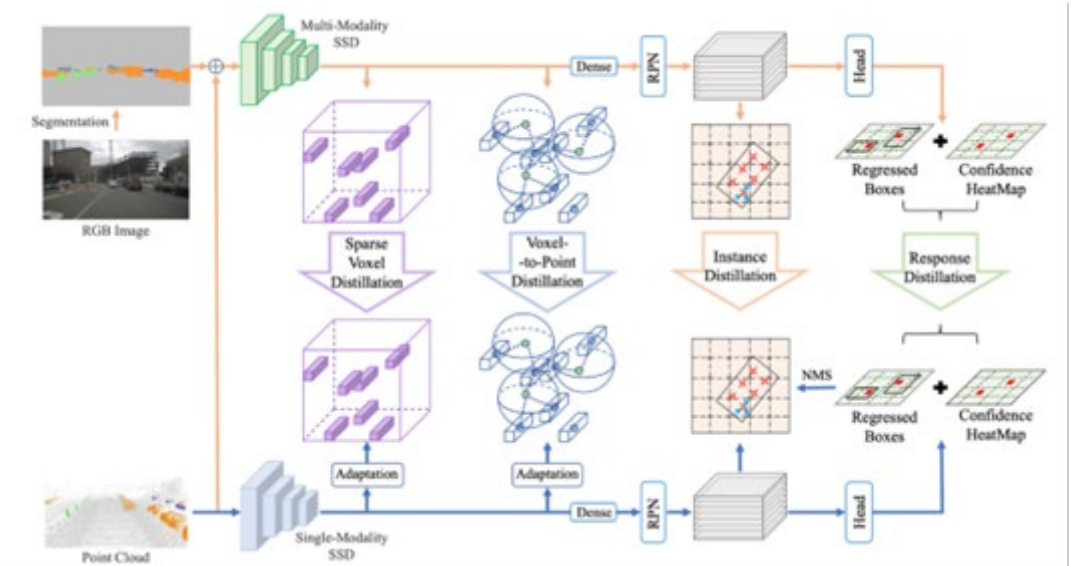


Figure 1. Our S2M2-SSD framework. The single-modality SSD (bottom) takes only a point cloud as input, whereas the multi-modality SSD (top) further takes a segmented image. The training has two phases: (i) pre-train the multi-modality SSD (orange arrows); and (ii) train the single-modality SSD (orange & blue arrows) to learn to produce features and responses comparable with the pre-trained multi-modality SSD by designing four levels of knowledge distillation: response distillation, sparse-voxel distillation, voxel-to-point distillation, and instance distillation. At the testing (only the blue arrows), we only need a point cloud as the input of the well-trained single-modality SSD for the object detection.

Table 1. Comparison with SOTA LiDAR-only detectors on nuScenes test set. Our S2M2-SSD attains the highest NDS and mAP. The percentages in () mean the proportions of S2M2-SSD's gains on the single-modality SSD relative to the single- and multi-modality gap. * means the SSD is built on our improved version of CenterPoint.

Method		Modality	NDS	mAP
WYSIWYG	CVPR 2020	LiDAR	41.9	35.0
PointPillars	CVPR 2019	LiDAR	45.3	30.5
PointPainting	CVPR 2020	LiDAR	58.1	46.4
CVCNet	NeurIPS 2020	LiDAR	64.4	55.3
PMPNet	CVPR 2020	LiDAR	53.1	45.4
SSN	ECCV 2020	LiDAR	58.1	46.4
CBGS	NeurIPS 2019	LiDAR	63.3	52.8
CenterPoint	CVPR 2021	LiDAR	65.6	58.0
Multi-modality SSD*	-	LiDAR+RGB	69.1	64.0
Single-modality SSD*	-	LiDAR	67.3	60.1
S2M2-SSD (Ours)	CVPR 2022	LiDAR	69.3	62.9
Improvement		-	+2.0 (111%)	+2.8(72%)

2) Semi-supervised Learning on Point Cloud. We proposed a new approach to alleviate the label-hungry problem in a semi-supervised setting for 3D point cloud instance segmentation. To leverage the unlabeled data to boost the model performance, we present a novel Two-Way Inter-label Self-Training framework named TWIST; see Figure 2 below for an illustration of our framework. It exploits inherent correlations between semantic understanding and instance information of a scene. Specifically, we consider two kinds of pseudo labels for semantic- and instance-level supervision. Our key design is to provide object-level information for denoising pseudo labels and make use of their correlation for two-way mutual enhancement, thereby iteratively promoting the pseudo-label qualities. TWIST attains leading performance on both ScanNet and S3DIS, compared to recent 3D pre-training approaches, and can cooperate with them to further enhance performance, e.g., +4.4% AP 50 on 1% label ScanNet data-efficient benchmark. This work was published in the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2022.

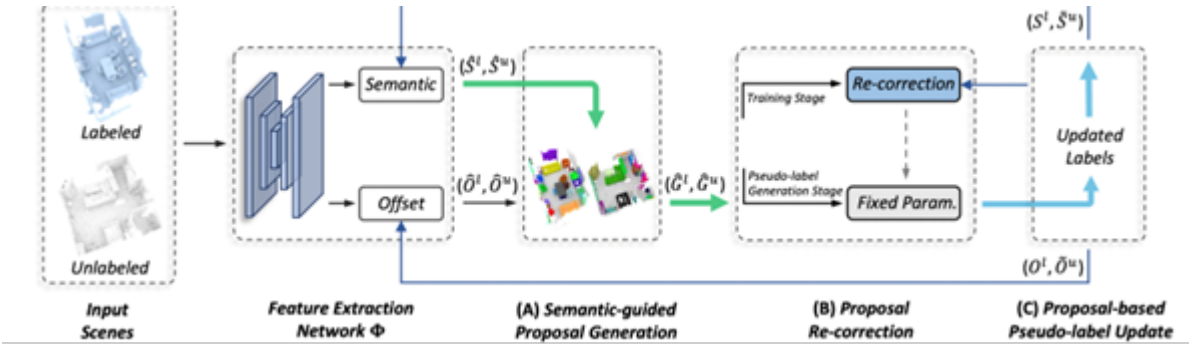


Figure 2. Our TWIST framework. Given a pair of labeled and unlabeled point clouds (P^l, P^u), the feature extraction network Φ first predicts per-point semantic classes (S^l, S^u) and per-point offsets (O^l, O^u). At the training stage of each self-training round, $\hat{S}^l$ and $\hat{O}^l$ are forwarded by module A to generate instance proposals $\hat{G}^l$, which is then fed into module B for evaluation/rectification. ($\hat{S}^l, \hat{O}^l$) is supervised by ground-truth labels (S^l, O^l) and ($\hat{S}^u, \hat{O}^u$) is supervised by pseudo labels ($\tilde{S}^u, \tilde{O}^u$). At the pseudo-label generation stage, only the unlabeled point cloud is processed. We pass $\hat{S}^u$ and $\hat{G}^u$ through modules A, B, and C for object-level denoising and finally update pseudo labels ($\tilde{S}^u, \tilde{O}^u$). The thick arrows in green and blue represent the two-way mutual enhancement between pseudo labels $\tilde{S}^u$ and $\tilde{O}^u$.

.

3) Simulating Multi-Frame Point Clouds to Boost 3D Object Detection Performance. To boost a detector for single-frame 3D object detection, we present a new approach to train it to simulate features and responses following a detector trained on multi-frame point clouds. Our approach needs multi-frame point clouds only when training the single-frame detector, and once trained, it can detect objects with only single-frame point clouds as inputs during the inference. We design a novel Simulated Multi-Frame Single-Stage object Detector (SMF-SSD) framework to realize the approach (see Figure 3 below for an illustration of the framework). In our framework, we first design the multi-view dense object fusion strategy to densify ground-truth objects to generate a multi-frame point cloud; then, we formulate the self-attention voxel distillation to facilitate one-to-many knowledge transfer from multi- to single-frame voxels; multi-scale BEV feature distillation to transfer knowledge in low-level spatial and high-level semantic BEV features; and adaptive response distillation to activate single-frame responses of high confidence and accurate localization. Experimental results on the large-scale Waymo test set show that our SMF-SSD consistently outperforms all state-of-the-art single-frame 3D object detectors for all object classes of difficulty levels 1 and 2 in terms of both mAP and mAPH. This work was published in ACM Multimedia 2022.

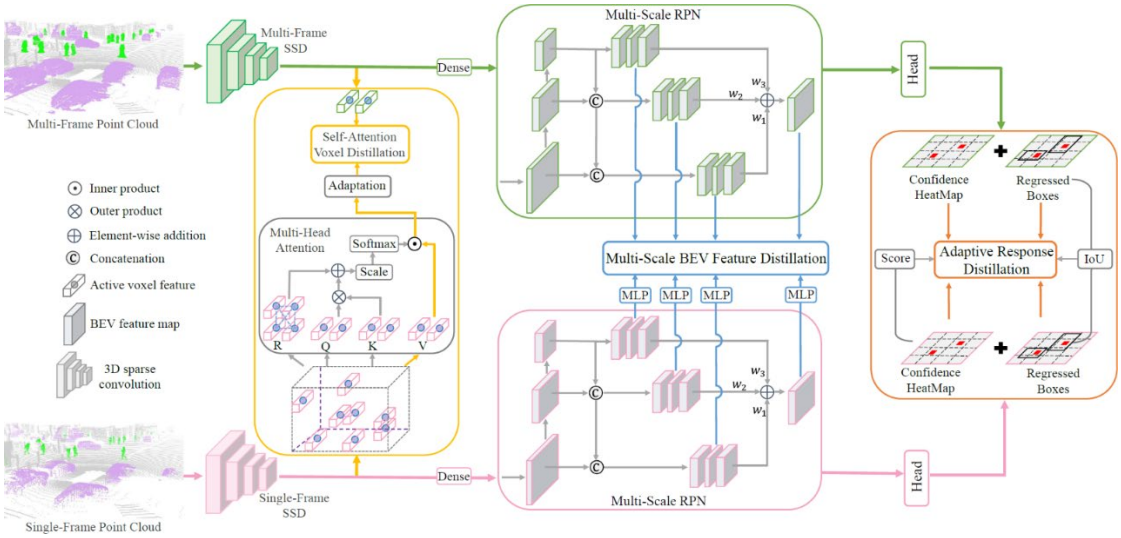


Figure 3. Our SMF-SSD framework. The single-frame SSD (bottom) takes a single-frame point cloud as input, whereas the multi-frame SSD (top) adopts a multi-frame one, in which we only densify ground-truth objects in the multi-frame point cloud by our designed multi-view dense object fusion. At testing (only pink arrows on the bottom), object detection can be done using the single-frame SSD and we only need a single-frame point cloud as input.

4) Learning to Augment Sparse Point-based Objects to Boost 3D Object Detection Performance. We design Sparse2Dense, a novel framework (see the illustration of our framework shown in Figure 4 below) to efficiently boost 3D detection performance by learning to densify input point clouds in latent space, addressing the issues that existing detectors often miss small, distant, and incomplete objects with sparse or few points. Specifically, we first train a dense point 3D detector (DDet) with dense point clouds as input and design a sparse point 3D detector (SDet) with raw point clouds as input. Importantly, we formulate the lightweight S2D module and the point compensation module in SDet to densify 3D features and train SDet to produce 3D features that follow the dense 3D features in DDet. So, in testing, SDet simulates dense 3D features from raw (sparse) point cloud input without requiring dense inputs. We evaluate our method on the Waymo Open and Domain Adaptation dataset, showing its high performance and efficiency over state-of-the-art 3D object detectors. This work was published in NeurIPS 2022.

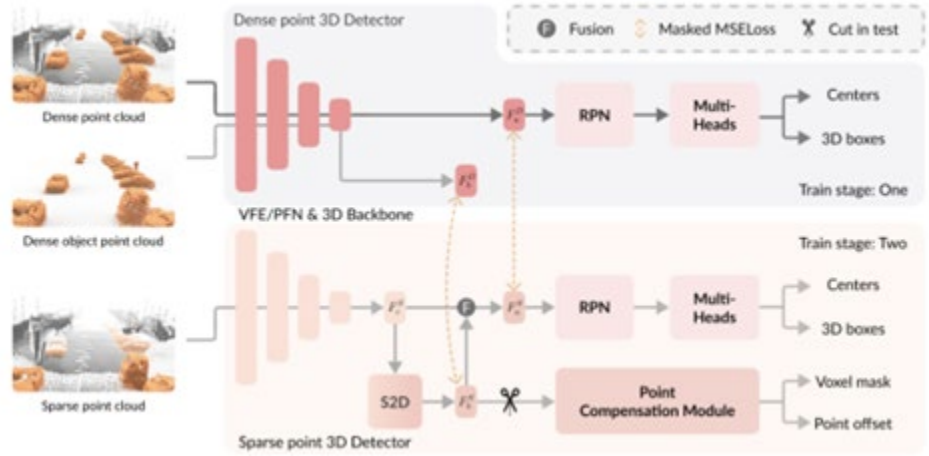


Figure 4. Our Sparse2Dense. First, we train the Dense point 3D Detector (DDet) with a dense point cloud as input, then adopt DDet to initialize the Sparse point 3D Detector (SDet). Then, we fix DDet when training SDet. During the training, we take both dense point cloud and dense object point cloud as the inputs. Further, we design the point compensation module as an auxiliary task to recover a dense point cloud from the densified 3D features. At testing, we only need SDet for 3D object detection on raw point clouds without using DDet and the point compensation module.

5) Leveraging Temporal Features to Resolve 3D Occlusion Problem

Real-time 3D hand mesh reconstruction is challenging, especially when the hand is holding some object. Beyond the previous methods, we design H2ONet to fully exploit non-occluded information from multiple frames to boost the reconstruction quality. First, we decouple hand mesh reconstruction into two branches, one to exploit finger-level nonoccluded information and the other to exploit global hand orientation, with lightweight structures to promote real-time inference. Second, we propose finger-level occlusion-aware feature fusion, leveraging predicted finger-level occlusion information as guidance to fuse finger-level information across time frames. Further, we design hand-level occlusion-aware feature fusion to fetch non-occluded information from nearby time frames. We conduct experiments on the Dex-YCB and HO3D-v2 datasets with challenging hand-object occlusion cases, manifesting that H2ONet is able to run in real-time and achieves state-of-the-art performance on both the hand mesh and pose precision. This work was published in the IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2023.

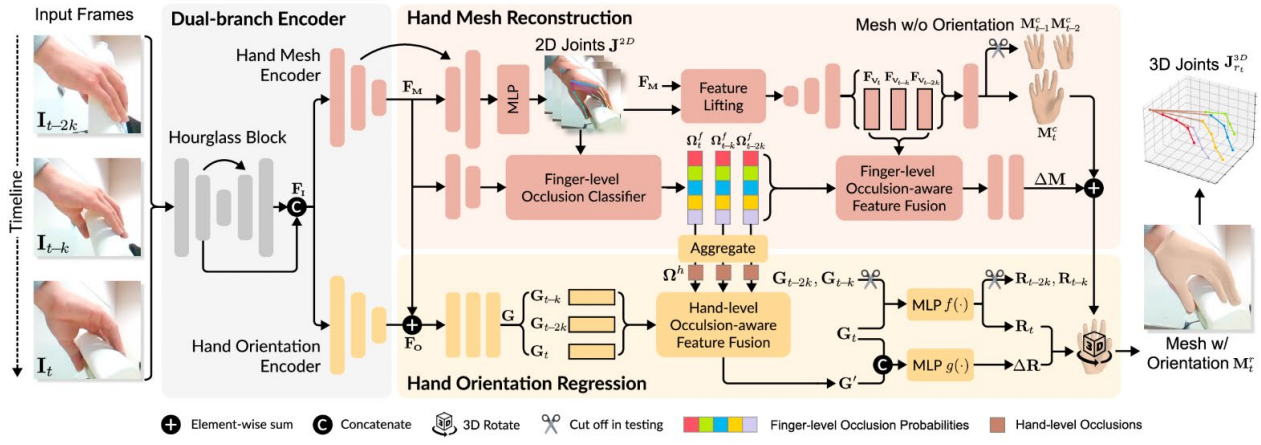


Figure 6. The H2ONet architecture: (i) the dual-branch encoder extracts general and task-specific features; (ii) the hand mesh reconstruction module focuses on constructing hand meshes at canonical poses by predicting finger-level occlusions and taking these information as guidance to fuse the multi-frame information; and (iii) the hand orientation regression module predicts the global hand orientation using the predicted hand-level visibility (equivalently, the occlusion) across multiple frames for support the estimation of the hand orientation.

6) Prototypical Variational Autoencoder for Few-shot 3D Point Cloud Object Detection

Few-Shot 3D Point Cloud Object Detection (FS3D) is a challenging task, aiming to detect 3D objects of novel classes using only limited annotated samples for training. Considering that the detection performance highly relies on the quality of the latent features, we design a VAE-based prototype learning scheme, named prototypical VAE (P-VAE), to learn a probabilistic latent space for enhancing the diversity and distinctiveness of the sampled features. The network encodes a multi-center GMM-like posterior, in which each distribution centers at a prototype. For regularization, P-VAE incorporates a reconstruction task to preserve geometric information. To adopt P-VAE for the 3D object detection framework, we formulate Geometric-informative Prototypical VAE (GP-VAE) to handle varying geometric components and Class-specific Prototypical VAE (CP-VAE) to handle varying object categories. In the first stage, we harness GP-VAE to aid feature extraction from the input scene. In the second stage, we cluster the geometric-informative features into per-instance features and use CP-VAE to refine each instance feature with category-level guidance. Experimental results show the top performance of our approach over the state of the arts on two FS3D benchmarks. Quantitative ablations and qualitative prototype analysis further demonstrate that our probabilistic modeling can significantly boost prototype learning for FS3D. Very recently in early October, this work was accepted for publication in NeurIPS 2023.

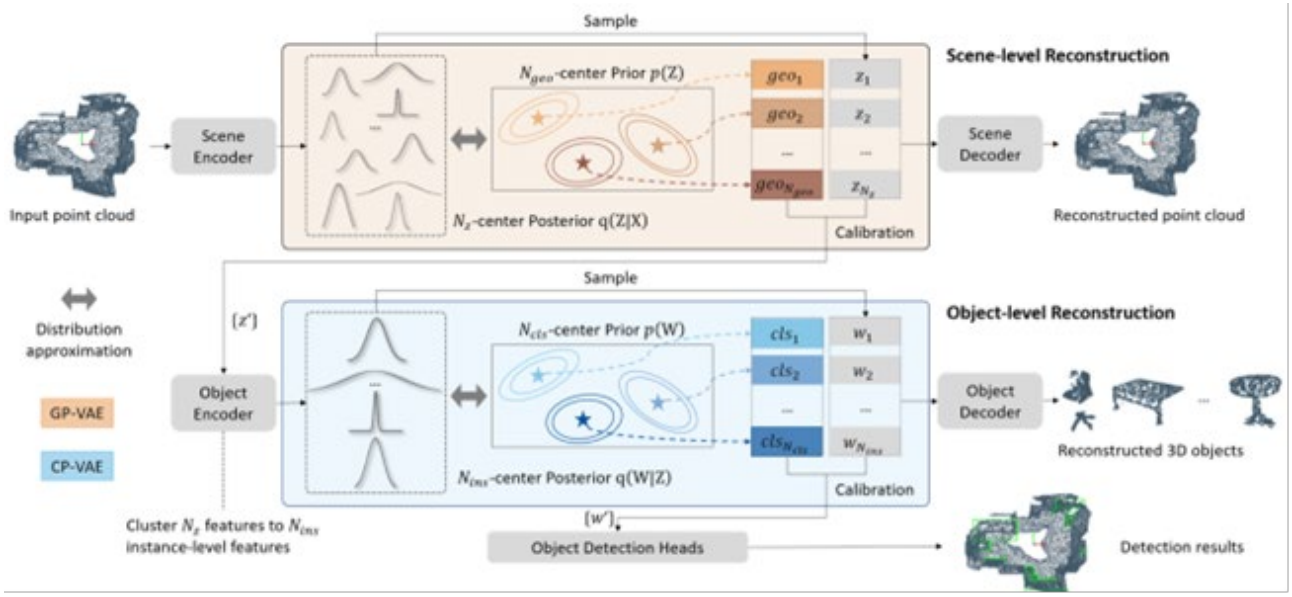


Figure 5. Illustration of our pipeline. We incorporate two crucial tasks to facilitate the framework to learn representative prototypes: a scene-level reconstruction task with GP-VAE and an object-level reconstruction task with CP-VAE.

7) Simulating Multi-View Features to Boost 3D Reconstruction Performance

Estimating 3D hand mesh from RGB images is a long-standing track, in which occlusion is one of the most challenging problems. Existing attempts towards this task often fail when the occlusion dominates the image space. In this paper, we propose SiMA-Hand, aiming to boost the mesh reconstruction performance by Single-to-Multi-view Adaptation. First, we design a multi-view hand reconstructor to fuse information across multiple views by holistically adopting feature fusion at image, joint, and vertex levels. Then, we introduce a single-view hand reconstructor equipped with SiMA. Though taking only one view as input at inference, the shape and orientation features in the single-view reconstructor can be enriched by learning non-occluded knowledge from the extra views at training, enhancing the reconstruction precision on the occluded regions. We conduct experiments on the Dex-YCB and HanCo benchmarks with challenging object- and self-caused occlusion

cases, manifesting that SiMA-Hand consistently achieves superior performance over the state of the arts. This work was submitted to AAAI recently in September and just passed the Stage 1 review in AAAI.

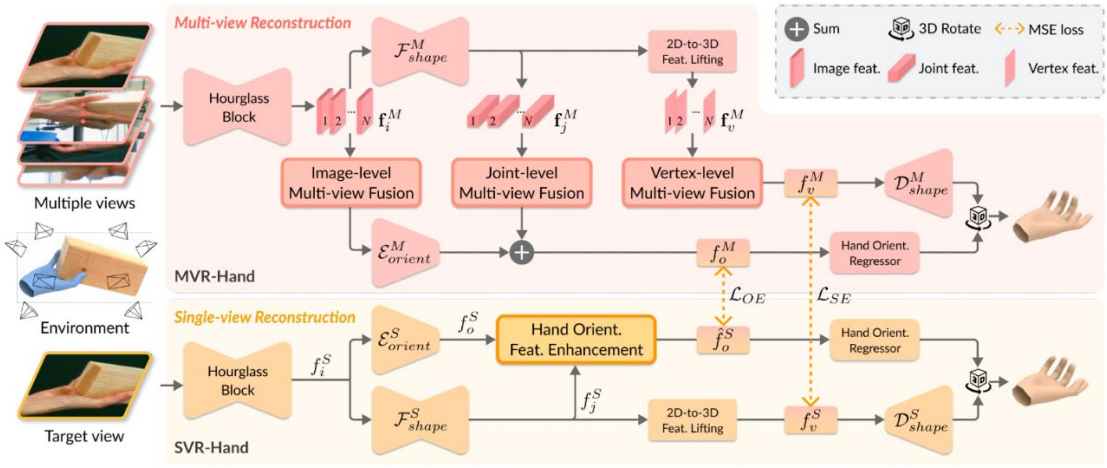


Figure 7. The architecture of SiMA-Hand: (i) the MVR-Hand takes multiple views of the hand as input for 3D hand-mesh reconstruction by fusing multi-view features at image, joint, and vertex levels; and (ii) the SVR-Hand takes only one view as input and learns to output a high-quality 3D hand mesh, even under a severely-occluded situation (see the target view at lower left), with both shape and orientation feature enhancement from the MVR-Hand.

4. PUBLICATIONS

[1] Zheng, W., Hong, M., Jiang, L., and Fu, C.-W., "Boosting 3D Object Detection by Simulating Multimodality on Point Clouds", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022. (Oral)

[2] Chu, R., Ye, X., Liu, Z., Tan, X., Qi, X., Fu, C.-W., and Jia, J., "TWIST: Two-Way Inter-label Self-Training for Semi-supervised 3D Instance Segmentation", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2022.

[3] Zheng, W., Jiang, L., Lu, F., Ye, Y., and Fu, C.-W., "Boosting Single-Frame 3D Object Detection by Simulating Multi-Frame Point Clouds", *ACM International Conference on Multimedia (ACM MM)*, 2022.

[4] Wang, T., Hu., X and Fu, C.-W., "Sparse2Dense: Learn to Densify 3D Features to Boost 3D Object Detection", *Conference on Neural Information Processing Systems (NeurIPS)*, 2022.

[5] Wang, T., Hu., X., Heng, P.-A., and Fu, C.-W., "Instance Shadow Detection with A Single-Stage Detector", *IEEE Transactions on Pattern Analysis and Machine Intelligence (T-PAMI)*, 2022.

[6] Xu, H., Wang, T., Tang, X., and Fu, C.-W. "H2ONet: Hand-Occlusion-and-Orientation-Aware Network for Real-Time 3D Hand Mesh Reconstruction", *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023.

[7] Tang, W., Yang, B., Li, X., Heng, P.-A., Liu, Y., and Fu, C.-W. "Prototypical Variational Autoencoder for Few-shot 3D Point Cloud Object Detection", *Conference on Neural Information Processing Systems (NeurIPS)*, 2023.

Under review:

[8] Wang, Y., Xu, H., Heng, P.-A., and Fu, C.-W. "SiMA-Hand: Boosting 3D Hand-Mesh Reconstruction by Single-to-Multi-view Adaptation", Submitted to AAAI 2024, now under review.

IMPROVING WORKFLOW RECOGNITION OF ROBOT-ASSISTED SURGERY VIA EXPLORATION INTO FUTURE SCENES FOR SURGICAL VIDEO ANALYSIS

Principal Investigator: Professor Qi DOU
Department of Computer Science and Engineering, CUHK

Research Team Members:
Yonghao[(1)]; WEI, Wang[(1)];
GAO, Xiaojie[(1)]; MA, Wenao[(1)]

[(1)] Department of Computer Science and Engineering, CUHK

Project Start Date: 1st July 2020
Completion Date: 30th June 2022

INNOVATION AND PRACTICAL SIGNIFICANCE:

In this project, we will address the problem of surgical workflow recognition for endoscopic video analysis, which serves as the fundamental role for enhancing context-awareness in robot-assisted surgery. As mentioned, existing approaches only make predictions based on previous video frames of the current state. Such blindness to future information consideration would lead to insufficient contextual clues, inconsistent segment-level smoothness, unreliable prediction at complex scenes, and suboptimal understanding of the on-going surgical activity. This project will address the above limitations with novel learning methods proposed from three hierarchical aspects: (1) we will first establish a reinforcement learning framework with a tree search algorithm, which is fundamentally novel from existing supervised convolutional networks based frameworks, in order to allow the future frames being explored through the value network. (2) we will innovate a video prediction approach for synthesizing future frames for the important use of online mode when there is be no future frame available in real-time. (3) we will incorporate the stochasticity for longer-term future scene predictions, which will be able to take into account future uncertainty in human-like manner for abnormality alert and risk assessment. In the last, we will demonstrate the practical significance of the proposed techniques on two robotics surgery scenarios: i.e., da Vinci robot-assisted suturing which is the most common process, and the cholecystectomy procedure which is more complex. It is promising that the surgical workflow recognition with human-like decision intelligence will ultimately facilitate surgeon performance, improve intervention outcome and benefit patient care.

ABSTRACT

Artificial intelligence for promoting surgical context-awareness based on endoscopic video analysis serves as a core enabler for reliable cognitive assistance in the next generation of robot-assisted surgery. Towards this mission, automated workflow recognition, i.e., being aware of current state of the surgical procedure being performed, is the fundamental task, which is the prerequisite for almost any form of intelligent assistance for automated surgery. Existing methods have formulated this problem into frame-wise or video clip-wise classification tasks and adopted computer vision approaches in supervised learning. However, they are limited to only analyzing previous frames to current state, without exploring ahead to future frames, which is far from how humans make decisions. This project addresses the limitation and hence improves workflow recognition accuracy, by exploring how to predict future information in video data with pixel-wise future scene synthesis

technologies. To our knowledge, this is the first attempt to systematically study how to incorporate the crucial future information for robot-assisted surgery context-awareness. The outcomes of this research not only advance methods and knowledge in deep learning and surgical robotics perception, but also promote AI technique transfer to the highly challenging yet important interdisciplinary area of medical interventions.

1. OBJECTIVES AND SIGNIFICANCE

This project addresses the problem of surgical workflow recognition for endoscopic video analysis, which serves as the fundamental role in robotic surgery. As mentioned, current methods only make predictions based on previous video frames of the current state. Such blindness to future information consideration would lead to suboptimal understanding of the on-going surgical activity. This project solves such limitation with novel learning methods proposed from three hierarchical aspects: (1) we first establish a reinforcement learning framework which is fundamentally novel compared to existing supervised CNN based frameworks, in order to allow the future frames being explored. (2) we innovate a video prediction approach for synthesizing future frames for the important use of online mode when there is no future frame available in real-time. (3) we incorporate the stochasticity for longer-term future scene predictions, which can take into account future uncertainty in human-like manner for abnormality alert and risk assessment. In the last, we demonstrate the practical significance of the proposed techniques on both robotic systems and real videos.

2. RESEARCH METHODOLOGY

This project mainly has four tasks. The detailed research plan and methodology for each task are as follows.

Task-1: A Reinforcement Learning Framework for Incorporating the Future Frames
Existing deep learning frameworks for endoscopic surgical video analysis are designed for frame-wise or video clip-wise problem formulation, which mainly rely on supervised learning and cannot support dynamic explorations of information inherent in future frames. Our previous ICRA 2020 paper proposed a new method to preliminarily explore the effectiveness of an AlphaGo style reinforcement learning framework, endowing the new capability of looking ahead into future frames for the learning system. Based on this promising related

work, we have made several improvements to our reinforcement learning agent. More specifically, we have integrated more advanced feature extraction networks and developed new reward functions aimed at enhancing performance in surgical scenarios.
Task-2: Enabling Strict Online Mode through Deterministic Future Frame Prediction
To further extend the online usability of the RL framework developed in Task-1, we solve the challenges for enabling strict online mode prediction, in which the workflow recognition is conducted in real-time surgery where no future frames are available. Specifically, we address the prediction of future frames in surgical video sequences, which is the first time to try to predict video frames in medical field. Surgical videos have its special characteristics in which the motion is more subtle and volatile than natural videos. For instance, the tool of grasper in cholecystectomy procedures has motion with its main body when its pointed-end interacting with tissues. Although existing works have been proposed for general video prediction, the complex evolution of detailed pixels in surgical video is still challenging and unexplored. The key technical challenges to be conquered lie in how to keep the tendency of video frames while guaranteeing video quality in resolution. We have proposed novel methods to resolve the above challenges.
Task-3: Modeling Future Prediction Uncertainty by Incorporating Stochasticity
Building the predictive model by learning from the collected expert videos reveals great potential to enable the AI system forecast what will happen in near or long future and make better decisions by considering future possible transformations. With the obtained richer knowledge by benefiting through deterministic future video prediction, it is natural and interesting to further explore various possibilities of future evolution for the given video clip, especially with the fact that longer future always accompanies more uncertainties. Modeling the distribution of such future uncertainties by incorporating prediction stochasticity helps make the system more

realistic and flexible. We have developed a more comprehensive representation of the available videos with variational autoencoder, where the distributions of latent random variables are learned by deep neural networks in a generative manner.

Task-4: Evaluation and Demonstration with Robot-assisted Surgery Applications

Last, we have evaluated and demonstrated the practical significance of the above three aspects of methodologies on two robotics surgery scenarios, i.e., da Vinci robot-assisted procedure which is the most common task in robotic surgery; and cholecystectomy procedure which is a more complicated and clinically important surgical workflow.

3. RESULTS ACHIEVED

During the support of this project, we have published a total of 18 papers in top-tier conferences and journals. These works are related to surgical data science, medical image analysis and AI-powered surgical robotics. Our contributions have been well recognized by the research community, with our publications winning 3 prestigious Best Paper Awards. The research objectives in this project are well attained. We highlight representative results as follows.

1) A Novel Reinforcement Learning Framework

Based on preliminary ICRA 2020 work, we further developed a new SV-RCNet-PPO model, which is a more advanced reinforcement learning framework for workflow recognition in surgical videos. Our SV-RCNet-PPO model first adopts a deep SV-RCNet to extract representative visual features while maintaining the temporal information within the feature simultaneously. Based on the extracted temporal visual features, we then design a novel reward function leveraging the new proximal policy optimization (PPO) to effectively optimize model parameters in a reinforcement learning framework. Figure 1 shows an overview of our SV-RCNet-PPO model.

On the widely adopted Cholec80 benchmark, our SV-RCNet-PPO achieves 88.2% accuracy, 83.1% precision, and 84.6% recall for online workflow recognition, which outperforms the SOTA SV-RCNet model (without our reinforcement learning framework) by 2.9%, 2.4%, and 1.1% respectively. These results demonstrate the effectiveness of our proposed novel reinforcement learning framework.

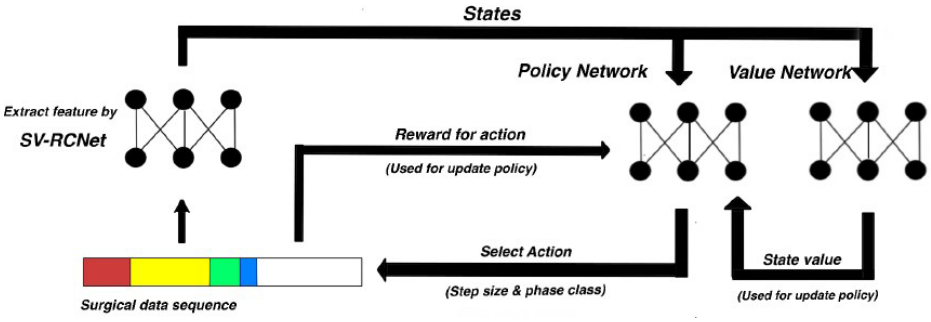


Figure 1. Overview of our SV-RCNet-PPO reinforcement learning framework for workflow

2) Future Frame Prediction for Surgical Videos

We propose a ternary prior guided variational autoencoder (TPG-VAE) model for robot-assisted surgical video prediction, which generates stable and realistic future frames [1]. Besides widely used content distribution, our TPG-VAE model proposes to learn the motion distribution, which is novel to handle the small movements of surgical tools. Furthermore, we add the invariant prior information from the gesture class into the generation process to constrain the latent space of our model. By considering unique characteristics of surgical scenarios

relative to general robotic videos, our proposed TPG-VAE model achieves the state-of-the- art performance for future frame prediction on the JIGSAWS dataset. The Figure 2 presents qualitative results compared with other methods. Generally, leveraging capturing the movements of the two robot arms, TPG- VAE model produces the closest predictions towards the ground truth. Additionally, in [2], we study the challenging one-shot instrument segmentation for robotic surgical videos, in which only the first frame mask of each video is provided at test time, such that the pre-trained model (learned from easily accessible source) can adapt to the target instruments. In [3], we design an adaptive instrument segmentation via meta learning, which is more practical for application as it can solve the knowledge gap between source domain (data from source hospital) and target domain (data from other hospital).

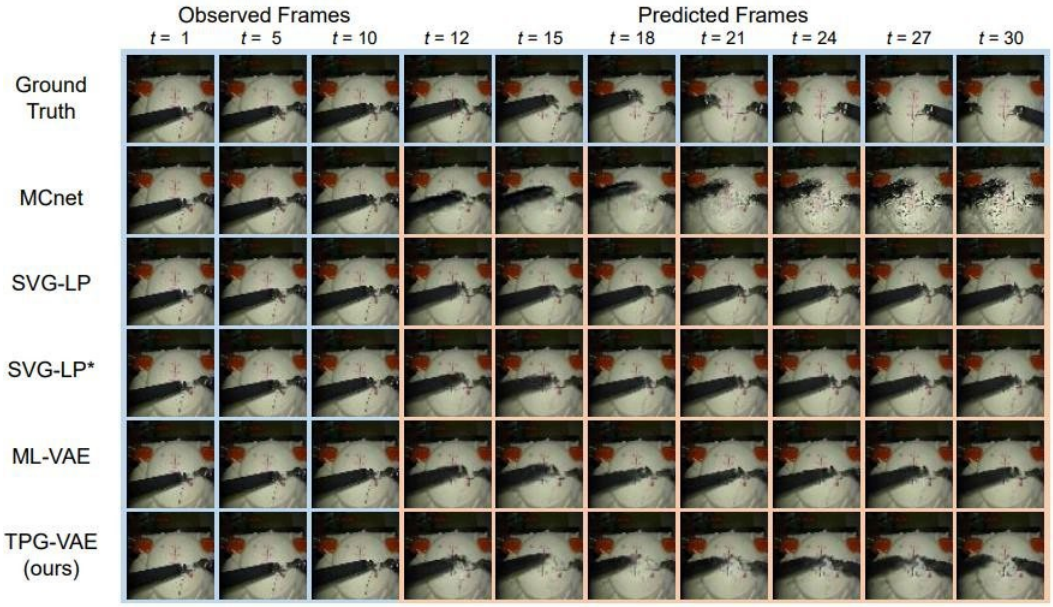


Figure 2. Future frame prediction results for the gesture of "Positioning Needle" from different

3) Modeling Temporal Dynamics for Future Prediction Uncertainty

We study the potential of leveraging temporal information / constrains in surgical sceneries. Three works have been conducted on surgical scene understanding. First, we apply the motion flow from robotic surgical video to semi-supervised tool segmentation [4]. Unlike most previous methods using unlabeled frames individually, we propose a dual motion-based method to learn motion flows for segmentation enhancement by leveraging temporal dynamics. Second, we propose a temporal memory relation network (TMRNet) [5] for relating long-range and multi-scale temporal patterns to augment the present features. We establish a long-range memory bank to serve as a memory cell storing the rich supportive information. Our results indicate that the adopt of temporal information significantly improves the performance surgical workflow recognition. Last, we introduce Transformer in surgical workflow analysis, to reconsider complementary effects of spatial and temporal representations. We propose a hybrid embedding aggregation Transformer, named Trans-SVNet [6], to effectively interact with the designed spatial and temporal embeddings. The Figure 3 shows recognition results via color-coded ribbon under four settings from all the three datasets. We can see that aggregating embedding from ResNet and TeCNO wisely, our Trans-SVNet trained with single recognition task already contributes to more consistent and robust predictions.

Intuitively, reliable and unreliable predictions exhibit different patterns in their corresponding temporal features (e.g., consistent motion flow and inconsistent motion flow). We leverage insights from our previous works to inject temporal information into our framework, effectively modeling the uncertainty of future predictions. By integrating this module of future frame uncertainty estimation into our workflow recognition framework, the enhanced system can more sophisticatedly consider future information by incorporating stochasticity.

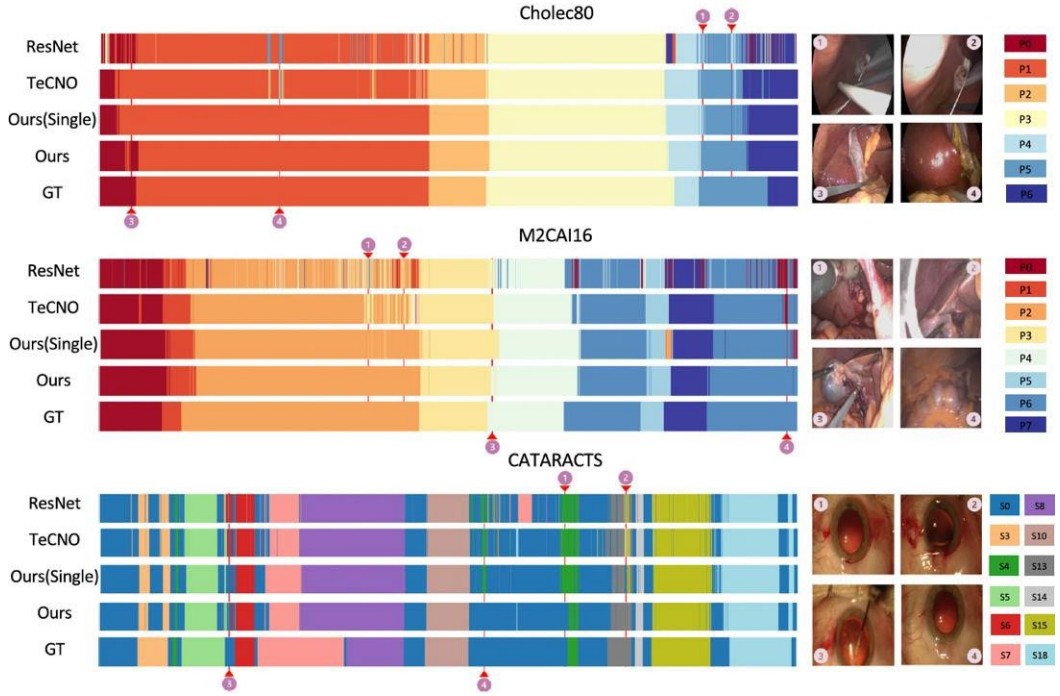


Figure 3. Color-coded illustration for three complete surgical videos.

4) Evaluation and Demonstration with Robot-assisted Surgery Applications

We have evaluated and demonstrated the so far developed techniques in da Vinci Research Kit (dVRK) robot platform at T Stone Robotics Institute, as well as real surgical video data. Figure 4 presents the adopted dVRK data in which we recorded video of peg transfer task, which is commonest task studying robotic surgery. The surgical workflow recognition accuracy has reached as high as 96% on metric of Edit Score [7].

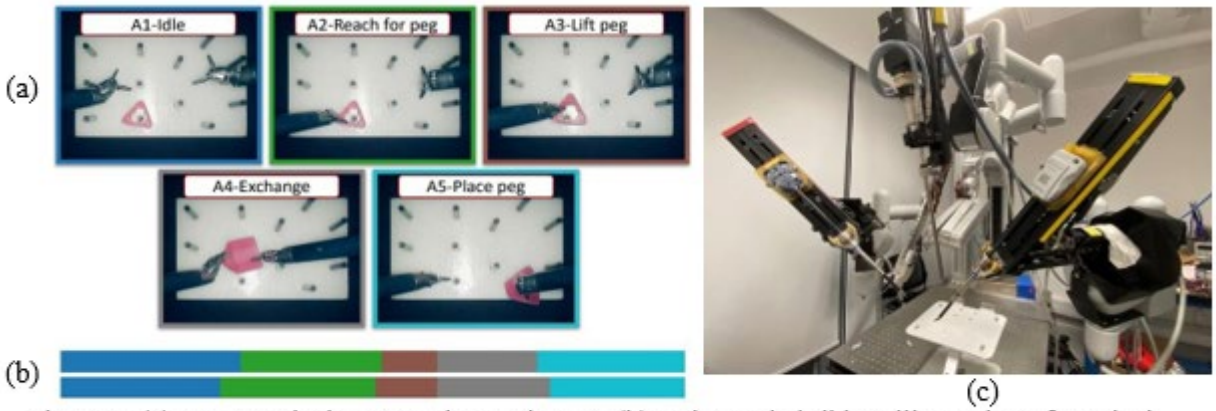


Figure 4. (a) Some typical gestures in our dataset. (b) Color-coded ribbon illustration of surgical sequences. (c) da Vinci Research Kit platform.

In addition, we applied our reinforcement learning framework on real surgical video data to recognize surgical workflow. Three metrics are used for quantitative evaluation, including the accuracy (AC), the precision (PR), and the recall (RE). In our model, there are two sets of hyper-parameters for adjusting. One is the long and short step sizes $\{ks, kl\}$, which is related to the action of reinforcement learning agent. The other is the weight α, which denotes the weight on the encouragement of a large action step and is related to the reward function. For parameter α, considering its simplicity as unary parameter, we first tested the cases with different values for it. Value of α is set as {0.01, 0.02, 0.05, 0.1, 0.2, 0.5, 1.0, 2.0} respectively, and then the corresponding models are trained and evaluated. The models with α value as 0.05, 0.1 and 0.2 obtained the best metrics in evaluation. Therefore, we then narrowed the value difference within the interval between 0.05 and 0.2 and evaluated the models with those α values. Based on the test results, we initially found the most optimal value for α as 0.1. Based on the optimal α value, we then adjusted the values of step sizes $\{ks, kl\}$ for models and evaluated the performance of them. Table 1 gives the experimental results with different parameters.

Table1. Quantitative evaluation results with different reinforcement learning hyper-parameters on video.

Parameters ($\{k_S, k_l\}, \alpha$)	Accuracy (AC)	Precision (PR)	Recall (RE)
({14, 80}, 0.1)	86.4%	78.2%	77.9%
({16, 60}, 0.1)	86.7%	78.7%	77.8%
({14, 40}, 0.1)	86.1%	80.0%	78.4%
({8, 60}, 0.1)	87.1%	74.8%	73.8%

5) Enhancing Robotic Surgery through Augmented Reality in Surgical Video Analysis

Further, surgical phase recognition techniques are integrated with augmented reality (AR) to enhance robotic surgery. We utilize advanced AR methods to improve surgical procedures. In [8], we propose a novel online reconstruction pipeline called E-DSSR which can reconstruct the surgical scene with only stereo endoscopic videos as input, and handle the cases of tissue deformation, tool occlusion and the camera movement simultaneously. In [9], we endeavor to reconstruct highly deformable surgical scenes captured from single-viewpoint stereo endoscopes. We embark on adapting the emerging neural rendering framework to the regime of deformable surgical scene reconstruction. In [10], we present the first completely Augmented Reality-based remote mentoring system with a head-mounted display for robotic surgical education. Experiments and the achieved promising results have demonstrated its potential as a new reliable, flexible and low-cost solution for next-generation remote mentoring.

4. PUBLICATION AND AWARDS

Awards:

- MICCAI 2020 1st Place in Surgical Workflow Recognition Grand Challenge
- ICRA 2021 Best Paper Award in Medical Robotics (1 out of 1690 papers, reported by CUHK News)
- IJCARS-MICCAI Best Paper Award in 2021
- MICCAI Young Scientist Publication Award in 2022 (a.k.a. Best Paper Award)

FYP ESTR4999 (2020/21): Topic: Reinforcement Learning for Workflow Recognition in Surgical Videos

Publications:

[1] X. Gao, Y Jin, Z. Zhao, Q. Dou and P. A. Heng, "Future frame prediction for robot-assisted surgery," Information Processing in Medical Imaging (IPMI), pp. 533-544, 2021. (Oral)

[2] Z. Zhao, Y. Jin, J. Chen, B. Lu, C. Ng, Y. Liu, Q. Dou, and P. A. Heng, "Anchor-guided online meta adaptation for fast one-shot instrument segmentation from robotic surgical videos," Medical Image Analysis, vol. 74, pp. 102240, 2021.

[3] Z. Zhao, Y. Jin, B. Lu, C. Ng, Q. Dou, Y. Liu, and P. A. Heng, "One to many: adaptive instrument segmentation via meta learning and dynamic online adaptation in robotic surgical video", IEEE International Conference on Robotics and Automation (ICRA), pp. 13553-13559, 2021.

[4] Zhao, Z., Jin, Y.*, Gao, X., Dou, Q., and Heng, P. A. "Learning Motion Flows for Semi-supervised Instrument Segmentation from Robotic Surgical Video," International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, pp. 679-689, 2020.

[5] Y. Jin, Y. Long, C. Chen, Z. Zhao, Q. Dou and P. A. Heng, "Temporal memory relation network for workflow recognition from surgical video", IEEE Transactions on Medical Imaging, vol. 40, no. 7, pp. 1911-1923, 2021.

[6] Y. Jin, Y. Long, X. Gao, D. Stoyanov, Q. Dou, and P. A. Heng, "Trans-svnet: Hybrid embedding aggregation transformer for surgical workflow analysis," International Journal of Computer Assisted Radiology and Surgery, vol. 17, no. 12, pp. 2193–2202, 2022. (IJCARS-MICCAI Best Paper Award)

[7] Y. Long, J. Wu, B. Lu, Y. Jin, M. Unberath, Y. Liu, P. A. Heng, and Q. Dou, "Relational graph learning on visual and kinematics embeddings for accurate gesture recognition in robotic surgery," IEEE International Conference on Robotics and Automation (ICRA), pp. 13346-13353, 2021. (IEEE ICRA Best Paper Award in Medical Robotics)

[8] Y. Long, Z. Li, C. Yee, C. Ng, R. Taylor, M. Unberath and Q. Dou, "E-DSSR: Efficient dynamic surgical scene reconstruction with transformer-based stereoscopic depth perception," Medical Image Computing and Computer Assisted Intervention (MICCAI), vol. 12904, pp. 415-425, 2021.

[9] Y. Wang, Y. Long, Y. Fan, S. Fan, and Q. Dou, "Neural rendering for stereo 3D reconstruction of deformable tissues in robotic surgery," Medical Image Computing and Computer Assisted Intervention (MICCAI), vol. 13437, pp. 431-441, 2022. (MICCAI Young Scientist Publication Award)

[10] Y. Long, C. Li, and Q. Dou, "Robotic surgery remote mentoring via AR with 3D scene streaming and hand interaction", Computer Methods in Biomechanics and Biomedical Engineering: Imaging & Visualization, vol. 11, no. 4, pp. 1027–1032, 2022.

[11] H. Yang, C. Chen, M. Jiang, Q. Liu, J. Cao, P. A. Heng, and Q. Dou, "DLTTA: Dynamic learning rate for test-time adaptation on cross-domain medical images," IEEE Transactions on Medical Imaging, vol. 41, no. 12, pp. 3575-3586, 2022.

[12] L. Yang, J. Jiang, X. Gao, Q. Wang, Q. Dou, and L. Zhang, "Autonomous environment-adaptive microrobot swarm navigation enabled by deep learning-based real-time distribution planning," Nature Machine Intelligence, vol 4, no. 5, pp. 480-493, 2022.

[13] W. Ma, C. Chen, S. Zheng, J. Qin, H. Zhang, and Q. Dou, "Test-time adaptation with calibration of medical image classification nets for label distribution shift", Medical Image Computing and Computer Assisted Intervention (MICCAI), vol. 13433. pp. 313-323, 2022.

[14] M. Jiang, H. Yang, X. Li, Q. Liu, P. A. Heng, and Q. Dou, "Dynamic bank learning for semi-supervised federated image diagnosis with class imbalance," Medical Image Computing and Computer Assisted Intervention (MICCAI), vol. 13433, pp. 196-206, 2022.

[15] W. Ma, S. Zheng, L. Zhang, H. Zhang, Q. Dou, “Rapid model transfer for medical image segmentation via iterative human-in-the-loop update: from labelled public to unlabeled clinical datasets for multi-organ segmentation in CT,” IEEE International Symposium on Biomedical Imaging (ISBI), pp. 1-5, 2022.
[16] Q. Dou, T. Y. So, M. Jiang, Q. Liu, V. Vardhanabhuti, G. Kaissis, Z. Li, et al, “Federated deep learning for detecting COVID-19 lung abnormalities in CT: a privacy-preserving multinational validation study,” Nature Partner Journal Digital Medicine, vol. 4, no. 1, 2021.
[17] J. Xu, B. Li, B. Lu, Y.H. Liu, Q. Dou, and P. A. Heng, "Surrol: An open-source reinforcement learning centered and dvrk compatible platform for surgical robot learning,” IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), pp. 1821-1828, 2021.
[18] Z. Wang, Q. Liu and Q. Dou, “Contrastive cross-site learning with redesigned net for COVID-19 CT classification,” IEEE Journal of Biomedical and Health Informatics, vol. 24, no. 10, pp. 2806-2813, 2020.

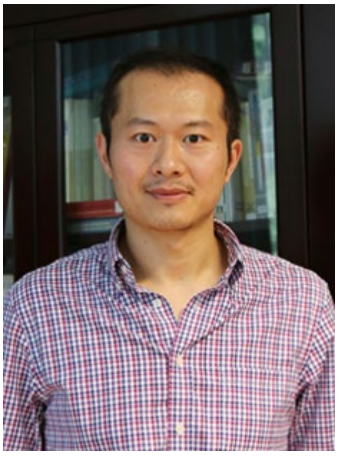
DEEP LEARNING BASED AUDIO-VISUAL RECOGNITION OF CANTONESE DISORDERED SPEECH

Principal Investigator: Professor Xunying LIU
Department of Systems Engineering and Engineering Management, CUHK

Co-Investigator (if any): Helen Meng[(1)]

Research Team Members:
Shansong Liu[(1)] , Shoukang Hu[(1)] , Jianwei Yu[(1)] , Mengzhe Geng[(1)] , Zi Ye[(1)] , Zengrui Jin[(1)]

[(1)]Department of Systems Engineering and Engineering Management, CUHK



Project Start Date: 1st July 2019
Completion Date: 31st July 2021

INNOVATION AND PRACTICAL SIGNIFICANCE:

Automatic speech recognition (ASR) for disordered speech is a challenging task. Speech disorders such as dysarthria lead to severe degradation of speech quality, highly variable voice characteristics and large mismatch against normal speech. State-of-the-art speech recognition systems designed for normal speech often produce poor recognition accuracy when applied to disordered speech. Human speech production is a bimodal process based on audio-visual representation. The visual information is invariant to acoustic signal corruption and can provide complementary information to the speech recognizer. This motivates the use of visual information to improve disordered speech recognition. However, among people with speech disorders, their underlying medical conditions such as Parkinson disease and co-occurring disabilities increase the difficulty to record high quality audio-visual data in large amounts that are necessary for audio-visual speech recognition (AVSR) system development. For example, in addition to the degradation of voice quality, head movements and different angles facing the camera are often found in recoding.

To the best of our knowledge, this project is the first attempt among the international multimedia, AI and speech technology research communities to use deep learning based AVSR approaches for disordered speech recognition. It is also the first one dedicated for Cantonese disordered speech. The outcome of this research will allow easier and more natural communication for Cantonese speaking people suffering from speech disorders with the outside world, improve their social inclusion and quality of life, and to support research and development efforts to create speech based assistive technology applications for such people. These can increase the cost effectiveness and quality of care and health service for them. This research will form a strong basis for future work on disordered speech recognition for Mandarin and other Chinese dialects to help a much larger number of similarly affected people in China.

ABSTRACT

Speech disorders such as dysarthria are commonly found among the elderly population which disrupt their verbal communication with the outside world. Speech disorders not only introduce a negative impact on their quality of life but also increase the cost of care. As Hong Kong is ageing rapidly, the number of people with

speech disorders will further increase. Due to the large mismatch between normal and disordered speech, state-of-the-art speech recognition systems designed for normal speech often produce very low recognition accuracy when applied to disordered speech. Hence, there is a pressing need to develop new technologies for building disordered speech recognition systems of high accuracy performance. By capitalizing on our previous research which produced the best accuracy on disordered English speech recognition, and also motivating from the bimodal nature of human speech production, this project will develop novel deep learning based audio-visual speech recognition (AVSR) techniques for Cantonese disordered speech. Exemplar ASVR systems and recipes for publicly available Cantonese disordered speech corpora will be developed to provide insights for designing state-of-the-art AVSR systems on such data. The outcome of this project allows easier and more natural communication for Cantonese speaking people in Hong Kong and Great Bay Area who suffer from speech disorders with the outside world. This project will also form a strong basis for future research on disordered speech recognition for Mandarin and other Chinese dialects to help a much larger number of similarly affected people in China.

1. OBJECTIVES AND SIGNIFICANCE

Automatic speech recognition (ASR) for disordered speech is a challenging task. Speech disorders such as dysarthria lead to severe degradation of speech quality, highly variable voice characteristics and large mismatch against normal speech. State-of-the-art speech recognition systems designed for normal speech often produce poor recognition accuracy when applied to disordered speech. Human speech production is a bimodal process based on audio-visual representation. The visual information is invariant to acoustic signal corruption and can provide complementary information to the speech recognizer. This motivates the use of visual information to improve disordered speech recognition. However, among people with speech disorders, their underlying medical conditions such as Parkinson disease and co-occurring disabilities increase the difficulty to record high quality audio-visual data in large amounts that are necessary for audio-visual speech recognition (AVSR) system development. For example, in addition to the degradation of voice quality, head movements and different angles facing the camera are often found in recoding.

To the best of our knowledge, this project is the first attempt among the international multimedia, AI and speech technology research communities to use deep learning based AVSR approaches for disordered speech recognition. It is also the first one dedicated for Cantonese disordered speech. The outcome of this research will allow easier and more natural communication for Cantonese speaking people suffering from speech disorders with the outside world, improve their social inclusion and quality of life, and to support research and development efforts to create speech based assistive technology applications for such people. These can increase the cost effectiveness and quality of care and health service for them. This research will form a strong basis for future work on disordered speech recognition for Mandarin and other Chinese dialects to help a much larger number of similarly affected people in China.

The project objectives and long-term impact are summarized below:

1. Using visual information to improve dysarthric speech recognition performance

2. Deriving novel deep AVSR methods to robustly model limited amounts of audio-visual disordered speech data

3. Deriving fast AVSR model adaptation methods to handle speaker dependent impairment characteristics

4. Developing exemplar AVSR systems and recipes to provide insights for designing state-of-the-art dysarthric speech recognition systems

5. Allow easier and more natural speech based communication for people suffering from speech disorders with the outside world, improve their social inclusion and quality of life

2. RESEARCH METHODOLOGY

1) Development of baseline ASR and AVSR systems for Cantonese disordered speech
First, baseline deep neural network based ASR and ASR systems will be constructed using CUDYS, the largest publicly available CUHK Cantonese disordered speech corpus. When developing these systems, state-of-the-art deep learning based ASR model architectures that have been successfully adopted in our CUHK English dysarthric speech recognition system will be used to build the baseline Cantonese ASR system for disordered speech. Comparable deep neural network based Cantonese AVSR systems using a fusion of acoustic and visual features will also be constructed to serve as the baseline AVSR approach.

2) Gaussian Process neural network based AVSR systems for modelling limited disordered speech
When given limited amounts of highly variable dysarthric speech data, it is a challenging problem to construct accurate ASR and AVSR systems. In order to address this issue, Bayesian neural networks and Gaussian process neural networks (GPNNs) will be explored to model limited dysarthric audio-visual speech data. Compared with Bayesian neural networks that only consider parameter uncertainty, GPNNs can model the uncertainty associated with both the choice of hidden activation functions and their parameters. With such flexibility they can produce highly compact deep model structures for modelling limited data and therefore reduce the risk of over-fitting. This was confirmed by our pilot study on normal speech recognition tasks. The GPNN acoustic models consistently outperformed the standard DNN baselines of fixed hidden activations with comparable numbers of parameters. As expected, the largest performance gap between the baseline DNN and GPNN systems were found when the DNNs systems are either under-fitting or over-fitting.

In the proposed research, feed forward architecture based Bayesian DNN and GPNN ASR and AVSR systems will first be investigated for disordered Cantonese speech. In order to model longer temporal contexts, novel forms of recurrent deep model architectures with Bayesian and Gaussian Process activations in place of standard fixed parameter LSTM cells and gated recurrent units will be investigated. New GP activations based on more compact forms of basis kernels will also explored to further improve their performance. Finally, performance analysis over different Bayesian DNNs and GPNN models will be conducted to produce their ranking order in terms of the trade-off between their model compactness and recognition accuracy performance. This can provide insights for optimal ASR and AVSR system design on limited audio-visual Cantonese disordered speech data.

3) Development of Bayesian Gated deep neural network based AVSR systems
A commonly used approach in state-of-the-art DNN based AVSR systems is to concatenate the acoustic and visual features at the input layer. This approach works well when the visual features being used are robust and contain sufficient discriminative information for classifying speech. However, in addition to the well-known degradation of voice quality, there are several new challenges when developing AVSR systems for people with speech disorders. First, their underlying medical conditions such as cerebral palsy and Parkinson disease combined with co-occurring disabilities increase the difficulty to record high quality visual data. For example, head movements and different angles facing the camera are often found. These make the accurate detection of lip regions very difficult, and the subsequent extracted visual features unreliable to use. Second, in common with the audio data, the diverse causes leading to speech disorders and the resulting symptoms create a large variability among individual impaired speakers.

In order to address these issues, we will use a novel Bayesian gated neural network (BGNN) based AVSR architecture in this paper. This is realized by positioning an additional multiplicative gating layer between the

input and first hidden layer. This layer's outputs are used to dynamically weight the contributions from visual features before they are further concatenated with acoustic features. This allows a more flexible fusion of acoustic and visual features that can learn to suppress non-discriminant visual data. Speaker dependent BGNN models are constructed to handle speaker level variability. In order to address the data sparsity issue, a posterior distribution over the gating layer weight and bias parameters is used to model their uncertainty given limited and variable data. Efficient variational inference will be also used in BGNN system training.

4) Fast deep AVSR model adaptation to diverse speaker impairment characteristics
The highly diverse impairment characteristics at speaker level creates large variation in audio-visual disordered speech data. Speaker adaptation techniques play a key role in reducing the mismatch between AVSR systems and target users. Model based DNN adaptation techniques often require a significant amount of data to robustly learn speaker dependent adaptation parameters. For example, in the commonly used learning hidden unit contributions (LHUC) based DNN adaptation, speaker dependent high dimensional hidden layer output scaling vectors are used. When limited adaptation data from individual impaired speakers are available, the standard LHUC method is prone to over-fitting and poor generalization. To address the issue, Bayesian learning of hidden unit contributions (BLHUC) will be investigated in our research. In contrast to current adaptation methods requiring large amounts of disordered speech data, this technique is expected to allow deep learning based AVSR systems to be rapidly adaptable to individual impaired speakers using very little data, for example, a few seconds of audio-visual disordered speech.

3. RESULTS ACHIEVED

Deep neural networks (DNNs) based automatic speech recognition (ASR) systems are often designed using expert knowledge and empirical evaluation. In our recent research, a range of neural architecture search (NAS) techniques are used to automatically learn two types of hyper-parameters of state-of-the-art factored time delay neural networks (TDNNs): i) the left and right splicing context offsets; and ii) the dimensionality of the bottleneck linear projection at each hidden layer. These include the DARTS method integrating architecture selection with lattice-free MMI (LF-MMI) TDNN training; Gumbel-Softmax and pipelined DARTS reducing the confusion over candidate architectures and improving the generalization of architecture selection; and Penalized DARTS incorporating resource constraints to adjust the trade-off between performance and system complexity. Parameter sharing among candidate architectures allows efficient search over up to 7^28 different TDNN systems. Experiments conducted on the 300-hour Switchboard corpus suggest the auto-configured systems consistently outperform the baseline LF-MMI TDNN systems using manual network design or random architecture search after LHUC speaker adaptation and RNNLM rescoring. Absolute word error rate (WER) reductions up to 1.0% and relative model size reduction of 28% were obtained. Consistent performance improvements were also obtained on a UASpeech disordered speech recognition task using the proposed NAS approaches (Hu et al. IEEE ICASSP21)

Automatic speech recognition (ASR) for disordered speech is a challenging task. People with speech disorders such as dysarthria often have physical disabilities, leading to severe degradation of speech quality, highly variable voice characteristics and large mismatch against normal speech. It is also difficult to record large amounts of high quality audio-visual data for developing audio-visual speech recognition (AVSR) systems. To address these issues, a novel Bayesian gated neural network (BGNN) based AVSR approach is proposed in our recent research. Speaker level Bayesian gated control of contributions from visual features allows a more robust fusion of audio and video modal- ity. A posterior distribution over the gating parameters is used to model their uncertainty given limited and variable disordered speech data. Experiments conducted on the UASpeech dysarthric speech corpus suggest the proposed BGNN AVSR system consistently outperforms state-of-the-art deep neural network (DNN) baseline ASR and AVSR systems by 4.5% and 4.7% absolute (14.9% and 15.5% relative) in word error rate (Liu et al. ISCA Interspeech2019)

Audio-visual speech recognition (AVSR) technologies have been successfully applied to a wide range of tasks. When developing AVSR systems for disordered speech characterized by severe degradation of voice quality and large mismatch against normal, it is difficult to record large amounts of high quality audio-visual data. In order to address this issue, a cross-domain visual feature generation approach is proposed in our recent research. Audio-visual inversion DNN systems constructed using widely available out-of-domain audio-visual data were used to generate visual features for disordered speakers for whom video data is either very limited or unavailable. Experiments conducted on the UASpeech corpus suggest that the proposed cross-domain visual feature generation based AVSR system consistently outperformed the baseline ASR system and AVSR system using original visual features. An overall word error rate reduction of 3.6% absolute (14% relative) was obtained over the previously published best system on the 8 UASpeech dysarthric speakers with audio-visual data of the same task (Liu et al. Interspeech2020).

Disordered speech recognition is a highly challenging task. The underlying neuro-motor conditions of people with speech disorders, often compounded with co-occurring physical disabilities, lead to the difficulty in collecting large quantities of speech required for system development. Our recent research investigated a set of data augmentation techniques for disordered speech recognition, including vocal tract length perturbation (VTLP), tempo perturbation and speed perturbation. Both normal and disordered speech were exploited in the augmentation process. Variability among impaired speakers in both the original and augmented data was modelled using learning hidden unit contributions (LHUC) based speaker adaptive training. The final speaker adapted system constructed using the UASpeech corpus and the best augmentation approach based on speed perturbation produced up to 2.92% absolute (9.3% relative) word error rate (WER) reduction over the baseline system without data augmentation, and gave an overall WER of 26.37% on the test set containing 16 dysarthric speakers (Geng et al. ISCA Interspeech2020).

Development of high performance ASR systems for medical and healthcare applications, for example, disordered speech, requires sufficient training data. Neuro-motor conditions leading to speech impairments, often co-occur with physical disabilities, and lead to the difficulty in collecting large quantities of impaired speech required for system development. To this end, data augmentation techniques play a key role in current ASR systems. In contrast to existing data augmentation techniques only modifying the speaking rate or overall spectral shape, fine grained spectra-temporal differences between impaired and normal speech are modelled using generative adversarial networks in our recent research. The resulting data augmentation process transforms normal speech spectra into those closer to disordered speech. Experiments on the UASpeech task suggests adversarial data augmentation approaches consistently outperformed the baseline augmentation using tempo or speed perturbation. An overall WER reduction of 3.05% (9.7% relative) was obtained over the baseline system using no data augmentation. The LHUC speaker adapted system using the best adversarial data augmentation approach gave a state-of-the-art WER of 25.89% on the UASpeech test set [80] (Jin et al. ISCA Interspeech 2021).

Automatic recognition of disordered speech remains a highly challenging task to date. Sources of variability commonly found in normal speech including accent, age or gender, when further compounded with the underlying causes of speech impairment and varying severity levels, creates large diversity among speakers. To this end, speaker adaptation techniques play a vital role in current ASR systems. Motivated by the spectral-temporal level differences between disordered and normal speech that manifest in articulatory imprecision, decreased volume and clarity, slower speaking rates and increased dysfluencies, novel deep spectra-temporal subspace embedding features derived by SVD decomposition of disordered speech are used in our recent research to facilitate both accurate speech intelligibility assessment and input feature based speaker adaption of state-of-the-art hybrid DNN and CTC end-to-end disordered speech recognition systems. Experiments conducted on the largest benchmark UASpeech dysarthric speech corpus suggests the spectra-temporal deep features adaptation approach consistently outperformed baseline i-vector adaptation by up to 2.63% absolute (8.6% relative) word error rate (WER) reduction after data augmentation and LHUC speaker adaptation are applied. The final speaker adapted system using the spectral basis embedding features gave an overall WER of 25.6% on the UASpeech test set of 16 dysarthric speakers (Geng et al. ISCA Interspeech2021).

Despite the rapid progress of automatic speech recognition (ASR) technologies in the past few decades, recognition of disordered speech remains a highly challenging task to date. Disordered speech presents a wide spectrum of challenges to current data intensive deep neural networks (DNNs) based ASR technologies that predominantly target normal speech. Our recently published IEEE/ACM TALSP journal paper presents recent research efforts at the Chinese University of Hong Kong (CUHK) to improve the performance of disordered speech recognition systems on the largest publicly available UASpeech dysarthric speech corpus. A set of novel modelling techniques including neural architectural search, data augmentation using spectra-temporal perturbation, model based speaker adaptation and cross-domain generation of visual features within an audio-visual speech recognition (AVSR) system framework were employed to address the above challenges. The combination of these techniques produced the lowest published word error rate (WER) of 25.21% on the UASpeech test set 16 dysarthric speakers, and an overall WER reduction of 5.4% absolute (17.6% relative) over the CUHK 2018 dysarthric speech recognition system featuring a 6-way DNN sys- tem combination and cross adaptation of out-of-domain normal speech data trained systems. Bayesian model adaptation further allows rapid adaptation to individual dysarthric speakers to be performed using as little as 3.06 seconds of speech. The efficacy of these techniques were further demonstrated on a CUDYS Cantonese dysarthric speech recognition task (Liu et al. IEEE/ACM TASLP2021)

An example video demo system for three speakers of mild, low and very low intelligibility levels from the UASpeech data set can be accessed via the following link:

https://speechsystemdemo.github.io/demo/M14B2M3.html

4. PUBLICATION AND AWARDS

J[1] Shansong Liu, Mengzhe Geng, Shoukang Hu, Xurong Xie, Mingyu Cui, Jianwei Yu, Xunying Liu and Helen Meng. Recent Progress in the CUHK Dysarthric Speech Recognition System, IEEE/ACM Transactions on Audio, Speech and Language Processing, Volume 29, Pages 2267-2281, 2021. [DOI]
J[2] Rongfeng Su, Xunying Liu, Lan Wang and Jingzhou Yang. Cross-Domain Deep Visual Feature Generation for Mandarin Audio-Visual Speech Recognition, IEEE/ACM Transactions on Audio, Speech and Language Processing, Volume 28, Issue 1, December 2020, Pages 185-197. [DOI]
C[1] Liu, S., Hu, S., Wang, Y., Yu, J., Su, R., Liu, X., Meng, H. (2019) Exploiting Visual Features Using Bayesian Gated Neural Networks for Disordered Speech Recognition. Proc. Interspeech 2019, 4120-4124, doi: 10.21437/Interspeech.2019-1536, **ISCA Student Paper Award Nomination**
C[2] Liu, S., Hu, S., Liu, X., Meng, H. (2019) On the Use of Pitch Features for Disordered Speech Recognition. Proc. Interspeech 2019, 4130-4134, doi: 10.21437/Interspeech.2019-2609
C[3] Hu, S., Liu, S., Chang, H.F., Geng, M., Chen, J., Chung, L.W., Hei, T.K., Yu, J., Wong, K.H., Liu, X., Meng, H. (2019) The CUHK Dysarthric Speech Recognition Systems for English and Cantonese. Proc. Interspeech 2019, 3669-3670

C[4] Shoukang Hu, Sirui Xie, Hehui Zheng, Chunxiao Liu, Jianping Shi, Xunying Liu, Dahua Lin; Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2020, pp. 12084-12092 [DOI]

C[5] Geng, M., Xie, X., Liu, S., Yu, J., Hu, S., Liu, X., Meng, H. (2020) Investigation of Data Augmentation Techniques for Disordered Speech Recognition. Proc. Interspeech 2020, 696-700, doi: 10.21437/Interspeech.2020-1161

C[6] Liu, S., Xie, X., Yu, J., Hu, S., Geng, M., Su, R., Zhang, S.-X., Liu, X., Meng, H. (2020) Exploiting Cross-Domain Visual Feature Generation for Disordered Speech Recognition. Proc. Interspeech 2020, 711-715, doi: 10.21437/Interspeech.2020-2282

C[7] Jianwei Yu, Shixiong Zhang, Jian Wu, Shahram Ghorbani, Bo Wu, Shiyin Kang, Shansong Liu, Xunying Liu, Helen Meng, Dong Yu. AUDIO-VISUAL RECOGNITION OF OVERLAPPED SPEECH FOR THE LRS2 DATASET, **IEEE Signal Processing Society Travel Grant Winner**, IEEE ICASSP2020, Barcelona, Spain.

C[8] Shoukang Hu, Xurong Xie, Shansong Liu, Mingyu Cui, Mengzhe Geng, Xunying Liu, Helen Meng. Neural Architecture Search For LF-MMI Trained Time Delay Neural Networks, IEEE ICASSP2021, Toronto, Canada.

C[9] Geng, M., Liu, S., Yu, J., Xie, X., Hu, S., Ye, Z., Jin, Z., Liu, X., Meng, H. (2021) Spectro-Temporal Deep Features for Disordered Speech Assessment and Recognition. Proc. Interspeech 2021, 4793-4797, doi: 10.21437/Interspeech.2021-60

C[10] Jin, Z., Geng, M., Xie, X., Yu, J., Liu, S., Liu, X., Meng, H. (2021) Adversarial Data Augmentation for Disordered Speech Recognition. Proc. Interspeech 2021, 4803-4807, doi: 10.21437/Interspeech.2021-168

Awards and Accomplishments

1. Professor Yue Chen received 2023 Best Paper Award of Journal of Economy and Technology.

   Y. Zhang, Y. Su, Y. Chen, and F. Liu, "Asynchronous Distributed Charging Protocol for Plug-in Electric Vehicles," Journal of Economy and Technology, vol. 1, pp. 29-47, 2023.

2. Professor Xu Song received Red Dot Design Award: Design Concept (2022).

   Ultra-compact heat exchanger fabricated by 3D printing https://www.red-dot.org/project/ultra-compact-heat-exchanger-fabricated-by-3d-printing-60272

3. Professor Andrew Loh received Best Paper Award of IEEE International Power Electronics and Motion Control Conference in 2022.

   L. Ming, W. Ding, P.C. Loh, and Z. Xin, "An FPGA-Based Fault-Tolerant Method for Reliable Current Commutation of Direct Matrix Converter," IEEE International Power Electronics and Motion Control Conference, IEEE, Nanjing, Jiangsu, China, 29 November-2 December 2020

4. Professor CHAN Hon Fai received Young Scholar Award, World Association for Chinese Biomedical Engineers in 2019.

5. Professor Qi DOU received ICRA 2021 IEEE ICRA Best Paper Award in Medical Robotics.

   Y. Long, J. Wu, B. Lu, Y. Jin, M. Unberath, Y. Liu, P. A. Heng, and Q. Dou, "Relational graph learning on visual and kinematics embeddings for accurate gesture recognition in robotic surgery," IEEE International Conference on Robotics and Automation (ICRA), pp. 13346-13353, 2021.

6. Professor Qi DOU received IJCARS-MICCAI Best Paper Award in 2022.

   Y. Jin, Y. Long, X. Gao, D. Stoyanov, Q. Dou, and P. A. Heng, "Trans-svnet: Hybrid embedding aggregation transformer for surgical workflow analysis," International Journal of Computer Assisted Radiology and Surgery, vol. 17, no. 12, pp. 2193–2202, 2022.

7. Professor Qi DOU received MICCAI Young Scientist Publication Award in 2022.

   Y. Wang, Y. Long, Y. Fan, S. Fan, and Q. Dou, "Neural rendering for stereo 3D reconstruction of deformable tissues in robotic surgery," Medical Image Computing and Computer Assisted Intervention (MICCAI), vol. 13437, pp. 431-441, 2022.

Commercialization Endeavors

1. Endoluminal Image-based Motion Optimization and Motion Understanding (EIMOMU) Towards Semi-Supervised Stiffness Tunable Flexible Robotic Procedures
by **Professor REN Hongliang** *(page 103)*

Professor Hongliang Ren is planning to deploy the ambient clinical intelligence solution developed in his project to the following aspects:

(1) Release doctors from heavy documenting work.
(2) Provide surgical training for non-skilled surgeons.
(3) Generate surgical instruction to provide decision support during surgery.
(4) Make healthcare experiences more personalized.
(5) Design the visual question-answering system to ease the load on the medical experts in answering questions.

2. Development of a folded bilayer scaffold for intestinal tissue engineering
by **Professor CHAN Hon Fai** *(page 123-124)*

Professor CHAN Hon Fai is going to submit a patent application on the design of the bilayer scaffold after finalizing the protocol of fabrication and he will actively seek financial support from government agencies (e.g., Innovation and Technology Commission) or private investor(s) to pursue downstream R&D to further characterize the engineered tissue, improve the culture configuration, optimize and scale-up production, and perform quality control of the process. The long-term goal will be to generate functional intestinal tissues for treating SBS.

3. Development of a novel flexible surgical robot with haptic sensation
by **Professor Zhang Li**

It is worth to mention that a SHIAE funded project pioneered by Prof. Zhang Li has recently come up with very successful commercialization result and it has been widely reported by local newspapers.

中大新技術 人體深層組織採樣免開刀

【本報訊】香港中文大學昨日發布一項「破裂誘導的功率放大機制」的新研發，可用以開發為進入人體深層組織採樣的臨床醫學用微型機械人，免卻病人開刀之苦。中大研究團隊表示，該技術的原理就如射箭，其結構不算複雜，成本相對便宜，除了作為醫療器械，還能應用在太空探索、農業等方面，不過，目前開發的裝置只能使用一次。

製光驅動水凝膠彈射器

這項技術是受到噴瓜種子的傳播模式啟發，以此為基礎開創了破裂誘導的功率放大運動機制，並藉此研發了光驅動水凝膠彈射器。採用這種新技術，一個直徑僅7毫米、厚3毫米，如鈕扣型電池般大小的水凝膠彈射器，起步速度可達時速約27公里，垂直發射距離可接近兩米。

中大工程學院機械與自動化工程學系教授張立表示，技術對物料要求不高，成本相宜。但現時運用這技術時，彈射器底部會破裂，故每個彈射器只能用一次。研究團隊正嘗試改用可「自癒」的水凝膠物料製作，使其可重複使用。對於新技術的原理，張立用射箭作類比，以射箭為例，當手指拉開弓弦時就會逐漸累積能量，一鬆手便瞬間釋放大量能量。

張立還表示，新技術有望取代現有微型機械人內的驅動部件，大幅增加其功率輸出，使微型機械人在人體內執行支架運載、在腸道等深層組織內採樣及切除組織等任務，讓更多病人受惠。在農業範疇，光驅動水凝膠發射器讓機械人有足夠的驅動和輸出能力，攜帶種子等彈射至遠處。此外，因應新技術的運動能力，它也可能應用於無重力或微重力的環境，在月球或外太空探測中驅動機械人進行任務。

■團隊展示大小僅幾毫米的新型光驅動水凝膠彈射器。

Shun Hing Institute of Advanced Engineering (SHIAE)
The Chinese University of Hong Kong
Shatin, N.T., Hong Kong

Office : Room 702, William M.W. Mong Engineering Building
Tel : (852) 3943 4351
Fax : (852) 3943 4354
Email : applications@shiae.cuhk.edu.hk
Website : http://www.shiae.cuhk.edu.hk