

The 11th International Doctoral Forum

Technical Program Book

Venue: International Lecture Hall, Building CII
Graduate School at Shenzhen, Tsinghua University
地点：清华大学深圳研究生院 CII 一楼多功能厅

December 6-8, 2018
Shenzhen, China



WELCOME MESSAGES

Dear PhD Candidates,

As the Chairs of the 11th International Doctoral Forum, we are delighted to invite you to attend the forum to be held in December 6-8, 2018, Shenzhen, China. This forum is designed to encourage communication between PhD candidates who are working in the field of Audio, Speech, Natural Language Processing, Computer Vision, Multimedia, AI, Networking, Big Data and Web Mining.

The forum consists of invited speeches, panel discussion and technical presentations. Besides technical activities, it also provides opportunities to network with like-minded participants through social events.

We would like to express our gratitude to the organizing committee, the technical program committee, the steering committee, authors, participants and many other contributors. Without the support of these people, it would be impossible for us to have such an attractive forum. We also would like to express heartfelt gratitude to the host of this forum, Graduate School at Shenzhen, Tsinghua University, for its elaborate preparations.

It's our great pleasure to welcome you to Shenzhen and the 11th International Doctoral Forum. We hope this forum will be productive and satisfying, we wish your stay in Shenzhen would be both valuable and enjoyable.

Sincerely Yours,

Prof. Juanzi Li (Tsinghua University)

Prof. Helen Meng (The Chinese University of Hong Kong)

Prof. Yanning Zhang (Northwestern Polytechnical University)

General chairs

HOST

Graduate School at Shenzhen, Tsinghua University

The Graduate School at Shenzhen, Tsinghua University, was jointly founded by Tsinghua University and the Shenzhen Municipal Government for cultivating top level professionals and carrying out scientific and technological innovations.

Directly affiliated with Tsinghua University in Beijing, the Graduate School plays an important role in Tsinghua's commitment to achieve international prominence. The new campus is in the University Town of Shenzhen, set amidst the natural beauty of Xili Lake and the Tanglang Mountain.

Following the principle of 'one university, one brand', the two campuses share the same culture and style, as well as the same goal of excellence. The primary goal of the Graduate School is to cultivate professionals with leadership capability, international vision, enterprising spirit and a wide range of knowledge.

The Graduate School serves the regional economy and society by developing top quality education, initiating original researches, and facilitating the transfer of high technology to industry.

With help from Shenzhen, the Graduate School is building up its research competitiveness and is developing at a fast pace by strengthening key disciplines, recruiting top talent, and establishing well equipped laboratories.



ABOUT Shenzhen

Shenzhen was a fishing village prior to 1978. It was chosen to become the first of China's Special Economic Zones under Deng Xiaoping. It has rapidly grown into a massive metropolis within decades, and becomes the manufacturing heart of the global tech industry. It is now home to some of China's most high-profile companies and has been regarded as the birthplace of China's economic miracle.

Window of the World

Window of the World is a theme park located in the western part of the city of Shenzhen. It has about 130 reproductions of some of the most famous tourist attractions in the world squeezed into 48 hectares. The 108-meter-tall Eiffel Tower dominates the skyline and the sight of the Pyramids and the Taj Mahal all in proximity to each other are all part of the appeal of this theme park.



Dapeng Fortress

Dapeng Fortress is located in the middle of Dapeng Peninsula in east Shenzhen. First built in the 27th year of Hongwu reign of the Ming dynasty (1394), it has a history of over 600 years. When it was set up in the Ming dynasty to resist Japanese pirates, its full name was the Dapeng Fortress of 1,000 Households, and it had played an important role in China's fight against Japanese pirates and British colonial invaders in the Ming and Qing dynasties. Standing 6 meters tall and extending 1,200 meters, the Fortress covers an area of about 100,000 sq meters, and its walls are of granite and bluestone bricks. Dapeng Fortress is so important that Shenzhen is now also known as "Pengcheng (literally the city of roc)".

Shenzhen Bay Park

Shenzhen Bay Park is the longest park along the coast in Shenzhen. It extends from the Futian Mangrove Ecological Park in the east to the Shenzhen Bay Checkpoint in the west. The 1.08km² park is home to 13 regional theme parks for sports, entertainment, sightseeing and leisure. The entire length of the park is motor vehicle-free. Cycle paths and picnic spaces are featured along the park. The park offers a panoramic view covering the Shenzhen skyline and New Territories, Hong Kong.



Xichong Beach

Xichong is a scenic spot in the Dapeng New District of Shenzhen City. It is located in the southern part of the Dapeng Peninsula, facing the South China Sea with Dapeng Bay to the west and Daya Bay to the east.

Xichong is located within Dapeng National Park, a recreational and scenic area. In 2006 Chinese National Geographic named Dapeng Peninsula one of China's eight most beautiful coastlines. To its north is the Qiniangshan, the second-highest mountain in the city. Xichong and its sister beach Dongchong are popular with backpackers, hikers and surfers.



FORUM TECHNICAL PROGRAM

TECHNICAL PROGRAM AT A GLANCE

Date	Time	Event	Speaker	Venue	Chair
Dec. 5	Check-in & Welcome Reception				
	17:30-19:00	Welcome Reception		Canteen	
	20:00	Hotel Check-in		Aii Life Residence	
Dec. 6	Opening Ceremony, Keynote Speech, Invited Talk, Science Park Visit				
	8:30-9:00	Registration		International Lecture Hall, Building CII	
	9:00-9:30	Opening Ceremony			Runnan Li
	9:30-9:45	Group Photo & Tea Break			
	9:45-10:45	Invited Talk 1 <i>Knowledge-rich Speech Processing: Beyond Current Deep Learning</i>	Chin-Hui Lee		Xuya Jia
	10:45-11:15	Invited Talk 2 <i>Knowledge-Guided Natural Language Processing</i>	Zhiyuan Liu		
	11:15-11:45	Invited Talk 3 <i>Edge AI for Data-Intensive Internet of Things</i>	Guoliang Xing		Yuewen Cao
	11:45-12:15	Invited Talk 4 <i>Mental Health Computing via Harvesting Social Media Data</i>	Jia Jia		
	12:15-13:00	Lunch		Canteen	
	13:00-17:30	Science Park Visit <i>UBTech Robotics & Tencent Headquarter at Binhai Mansion</i>		Science Park	
	18:00-19:30	Dinner		Canteen	
	20:00	Hotel		Aii Life Residence	
Dec. 7	Invited Talk, Sessions, Banquet and Closing Ceremony				
	8:30-9:00	Invited Talk 5 <i>Meeting the New Challenges in Speech Processing: Some NPU-ASLP Approaches</i>	Lei Xie	International Lecture Hall, Building CII	Qing Wang
	9:00-9:30	Invited Talk 6 <i>Deep Visual Scene Understanding</i>	Bolei Zhou		
	9:30-10:00	Invited Talk 7 <i>Event Level Video Captioning based on Attentional RNN</i>	Chun Yuan		
	10:00-10:30	Tea Break		International Lecture Hall, Building CII	
	10:30-12:00	Special Sessions			
		Oral Sessions (A, B, C)		Classrooms, Building CIII	
	12:00-13:30	Lunch		Canteen	
	13:30-15:00	Oral Sessions (D, E, F, G)		Classrooms, Building CIII	
	15:00-15:30	Tea Break			
	15:30-17:00	Oral Sessions (H, I, J)			
	17:30-19:00	Banquet, Closing Ceremony & Award Presentation		Canteen	Xixin Wu
20:00	Hotel		Aii Life Residence		
Dec. 8	Check-out & Social Event				
	Morning	Hotel Check-out		Aii Life Residence	
	All day	Social Event (自由活动, 不统一安排)			

INVITED TALKS

Speaker: Chin-Hui Lee, Professor, Georgia Institute of Technology

Title: Knowledge-rich Speech Processing: beyond Current Deep Learning

Time: 09:45 – 10:45, December 6, 2018

Venue: International Lecture Hall

Biography:



Chin-Hui Lee is a professor at School of Electrical and Computer Engineering, Georgia Institute of Technology. Before joining academia in 2001, he had accumulated 20 years of industrial experience ending in Bell Laboratories, Murray Hill, as a Distinguished Member of Technical Staff and Director of the Dialogue Systems Research Department. Dr. Lee is a Fellow of the IEEE and a Fellow of ISCA. He has published over 500 papers and 30 patents, with more than 42,000 citations and an h-index of 80 on Google Scholar. He received numerous awards, including the Bell Labs President's Gold Award in 1998. He won the SPS's 2006 Technical Achievement Award for "Exceptional Contributions to the Field of Automatic Speech Recognition". In 2012 he gave an ICASSP plenary talk on the future of automatic speech recognition. In the same year he was awarded the ISCA Medal in scientific achievement for "pioneering and seminal contributions to the principles and practice of automatic speech and speaker recognition".

Abstract:

Deep neural networks (DNNs) are becoming ubiquitous in designing speech processing algorithms. However, the robustness issues that have hindered a wide-spread deployment of speech technologies for decades still have not been fully resolved. In this talk, we first discuss capabilities and limitations of deep learning technologies. Next, we illustrate three knowledge-rich techniques, namely: (1) automatic speech attribute transcription (ASAT) integrating acoustic phonetic knowledge into speech processing and computer assisted pronunciation training (CAPT), (2) Bayesian DNNs leveraging upon speaker information for adaptation and system combination, and (3) DNN-based speech pre-processing, demonstrating better acoustics leads to more accurate speech recognition. Finally, we argue that domain knowledge in speech, language and acoustics is heavily needed beyond current blackbox deep learning in order to formulate sustainable whitebox solutions to further advance speech technologies.

Speaker: Zhiyuan Liu, Associate Professor, Tsinghua University

Title: Knowledge-Guided Natural Language Processing

Time: 10:45 – 11:15, December 6, 2018

Venue: International Lecture Hall

Biography:



Zhiyuan Liu is an associate professor at the Department of Computer Science and Technology, Tsinghua University. He received his Ph.D. degree in Computer Science from Tsinghua in 2011. His research interests include representation learning, knowledge graphs and social computation, and has published more than 60 papers in top-tier conferences and journals of AI and NLP including ACL, IJCAI and AAAI, cited by more than 3500 according to Google Scholar.

Abstract:

Recent years have witnessed the advances of deep learning techniques in various areas of NLP. However, as a typical data-driven approach, deep learning suffers from the issue of poor interpretability. A potential solution is to incorporate large-scale symbol-based knowledge graphs into deep learning. In this talk, I will present recent works on knowledge-guided deep learning methods for NLP.

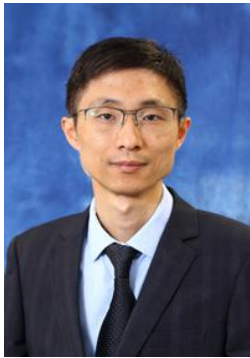
Speaker: Guoliang Xing, Professor, The Chinese University of Hong Kong

Title: Edge AI for Data-Intensive Internet of Things

Time: 11:15 – 11:45, December 6, 2018

Venue: International Lecture Hall

Biography:



Guoliang Xing is currently a Professor in the Department of Information Engineering, the Chinese University of Hong Kong. Previously, he was a faculty member at Michigan State University, U.S. His research interests include Embedded AI, Edge/Fog Computing, Cyber-Physical Systems, Internet of Things (IoT), security, and wireless networking. He received the B.S. and M.S. degrees from Xi'an Jiao Tong University, China, in 1998 and 2001, the D.Sc. degree from Washington University in St. Louis, in 2006. He is an NSF CAREER Award recipient in 2010. He received two Best Paper Awards and five

Best Paper Nominations at several first-tier conferences including ICNP and IPSN. Several mobile health technologies developed in his lab won Best App Awards at the MobiCom conference and were successfully transferred to the industry. He received the Withrow Distinguished Faculty Award from Michigan State University in 2014. He serves as the General Chair for IPSN 2016 and TPC Co-Chair for IPSN 2017.

Abstract:

Internet of Things (IoT) represent a broad class of systems which interact with the physical world by tightly integrating sensing, communication, and compute with physical objects. Many IoT applications are data-intensive and mission-critical in nature, which generate significant amount of data that must be processed within stringent time constraints. It's estimated that 0.75 GB of data can be produced by an autonomous vehicle each second. The existing Cloud computing paradigm is inadequate for such applications due to significant or unpredictable delay and concerns on data privacy.

In this talk, I will present our recent work on Edge AI, which aims to address the challenges of data-intensive IoT by intelligently distributing compute, storage, control and networking along the continuum from Cloud to Things. First, I will present ORBIT, a system for programming Edge systems and partitioning compute tasks among network tiers to minimize the system power consumption while meeting application deadlines. ORBIT has been employed in several systems for seismic sensing, vision-based tracking, and multi-camera 3D reconstruction. Second, I will briefly describe several systems we developed for mobile health, smart cities, volcano and aquatic monitoring, which integrate domain-specific physical models with AI algorithms. We have conducted several large-scale field deployments for these systems, including installing a seismic sensor network at two live volcanoes in Ecuador and Chile.

Speaker: Jia Jia, Associate Professor, Tsinghua University

Title: Mental Health Computing via Harvesting Social Media Data

Time: 11:45 – 12:15, December 6, 2018

Venue: International Lecture Hall

Biography:



Dr. Jia Jia is a tenured associate professor in Department of Computer Science and Technology, Tsinghua University. Her main research interest is affective computing and human computer speech interaction. She has been awarded ACM Multimedia Grand Challenge Prize (2012), Scientific Progress Prizes from the National Ministry of Education as the First Person-in-charge (2016), IJCAI Early Career Spotlight (2018), ACM Multimedia

Best Demo Award (2018) and ACM SIGMM Emerging Leaders (2018). She has authored about 70 papers in leading conferences and journals including T-KDE, T-MM, T-MC, T-ASLP, T-AC, ACM Multimedia, AAAI, IJCAI, WWW etc. She also has wide research collaborations with Tencent, SOGOU, Huawei, Siemens, MSRA, Bosch, etc.

Abstract:

Psychological stress and depression are threatening people's health. It is non-trivial to detect stress or depression timely for proactive care. With the popularity of social media, people are used to sharing their daily activities and interacting with friends on social media platforms, making it feasible to leverage online social media data for stress and depression detection. In this talk, we will systematically introduce our work on stress and depression detection employing large-scale benchmark datasets from real-world social media platforms, including 1) stress-related and depression-related textual, visual and social attributes from various aspects, 2) novel hybrid models for binary stress detection, stress event and subject detection, and cross-domain depression detection, and finally 3) several intriguing phenomena indicating the special online behaviors of stressed as well as depressed people. We would also like to demonstrate our developed mental health care applications at the end of this talk.

Speaker: Lei Xie, Professor, Northwestern Polytechnical University

Title: Meeting the New Challenges in Speech Processing: Some NPU-ASLP Approaches

Time: 08:30 – 09:00, December 7, 2018

Venue: International Lecture Hall

Biography:



Lei Xie is currently a Professor in the School of Computer Science, Northwestern Polytechnical University, Xian, China. From 2001 to 2002, he was with the Department of Electronics and Information Processing, Vrije Universiteit Brussel (VUB), Brussels, Belgium, as a Visiting Scientist. From 2004 to 2006, he worked in the Center for Media Technology (RCMT), City University of Hong Kong. From 2006 to 2007, he worked in the Human-Computer Communications Laboratory (HCCL), The Chinese University of Hong Kong. His current research interests include audio, speech and language processing, multimedia and human-computer interaction. He is currently an associate editor of IEEE/ACM Trans. on Audio, Speech and Language Processing. He has published more than 140 papers in major journals and proceedings, such as IEEE TASLP, IEEE TMM, Signal Processing, Pattern Recognition, ACM Multimedia, ACL, INTERSPEECH and ICASSP.

Abstract:

Speech has become a popular human-machine interface due to fast development of deep learning, big data and super-computing. We can see many applications in smartphones, TVs, robots and smart speakers. However, for further wide deployments of speech interfaces, there are still many challenges we have to face, such as noise interferences, inter- and intra-speaker variations, speaking styles and low-resource scenarios. In this talk, I will introduce several approaches, recently developed in the Audio, Speech and Language Processing Group, Northwestern Polytechnical University (NPU-ASLP) team, to meet these challenges in speech recognition, speech enhancement and speech synthesis.

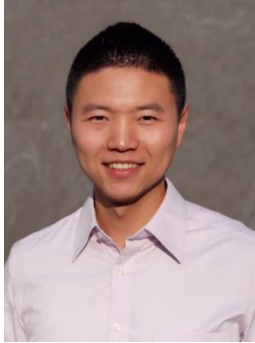
Speaker: Bolei Zhou, Professor, The Chinese University of Hong Kong

Title: Deep Visual Scene Understanding

Time: 09:00 – 09:30, December 7, 2018

Venue: International Lecture Hall

Biography:



Bolei Zhou is an Assistant Professor with the Information Engineering Department at the Chinese University of Hong Kong. He received his PhD in computer science at Massachusetts Institute of Technology (MIT). His research is in computer vision and machine learning, focusing on visual scene understanding and interpretable deep learning. He received the Facebook Fellowship, Microsoft Research Fellowship, MIT Greater China Fellowship, and his research was featured in media outlets such as

TechCrunch, Quartz, and MIT News.

Abstract:

Deep learning has made great progress in computer vision, achieving human-level object recognition. However, visual scene understanding, which aims at interpreting objects and their spatial relations in complex scene context, remains challenging. In this talk I will first introduce the recent progress of deep learning for visual scene understanding. From the 10-million image dataset Places to the pixel-level annotated dataset ADE20K, I will show the power of data and its synergy with interpretable deep neural networks for better scene recognition and parsing. Then I will talk about the trend of visual recognition from supervised learning towards more active learning scenario. Applications including city-scale perception and spatial navigation will be discussed.

Speaker: Chun Yuan, Associate Professor, Tsinghua University

Title: Event Level Video Captioning based on Attentional RNN

Time: 09:30 – 10:00, December 7, 2018

Venue: International Lecture Hall

Biography:



Chun Yuan is currently an Associate Professor with the Division of Information Science and Technology at Graduate school at Shenzhen, Tsinghua University. He received his M.S. and Ph.D. degrees from the Department of Computer Science and Technology, Tsinghua University, Beijing, China, in 1999 and 2002, respectively. He once worked at the INRIA-Rocquencourt, Paris, France, as a Post-doc research fellow from 2003 to 2004. In 2002, he worked at Microsoft Research Asia, Beijing, China, as an intern. His research interests include computer vision, machine learning and multimedia technologies. He is now the executive vice director of “Tsinghua-CUHK Joint Research Center for Media Sciences, Technologies and Systems”

Abstract:

Video understanding is a hotspot and challenge subject featured by jointly knowledge of natural language processing (NLP) and computer vision. More and more commercial application of online multimedia content requires better automatic understanding of video events. Unlike image captioning, video captioning faces more obstacles. First, video is complex data form to get and utilize feature, comparing to image. The temporal change makes sufficient information and different methods have their own shortages in mining temporal information. Second, in the task of captioning, the generation of sentence is required to extract dynamic information from videos. While some methods deal well with short and monotone actions, mining with longer and more complex actions is next goal. Third, some new tasks like captioning multiple events, call for new algorithm to get event-level processing. When generating sentences, correctly generate words like “continue” or “another” is one manifestation of good exploit context information.

SPECIAL SESSION TALKS

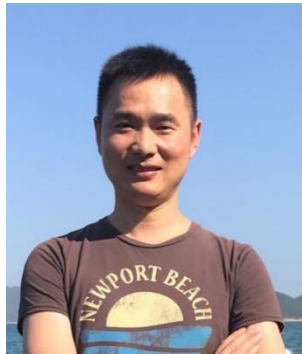
Speaker: Pengfei Liu, CTO of SpeechX

Title: Developing a Personalized Emotional Conversational Agent for Learning Spoken English

Time: 10:30 – 12:00, December 7, 2018

Venue: International Lecture Hall

Biography:



Dr. Pengfei Liu received his B.E. and M.E. degrees from The East China Normal University and the Ph.D. degree from The Chinese University of Hong Kong. His research areas are natural language processing and deep learning, particularly on sentiment analysis and dialog systems. He developed the SEEMGO system which ranked 5th in the task of aspect-based sentiment analysis at SemEval-2014, and received the Technology Progress Award in JD Dialog Challenge in 2018. Dr. Liu previously worked at SAP Labs China in Shanghai, The Chinese University of Hong Kong, and Wisers AI lab in Hong Kong, where he led a team to conduct research on deep learning-based sentiment analysis. He is currently the CTO of SpeechX.

Abstract:

The spoken English skill is critical but challenging for non-native learners in China due to lack of enough practice, while improving spoken English is in large demand among learners of different ages. This talk presents our ongoing project at SpeechX on developing a personalized emotional conversational agent which aims to provide a virtual partner for language learners to practice their spoken English. Such an agent is personalized based on each learner's English level and interests, and meanwhile gives appropriate responses according to the learner's emotions. Developing the agent involves a lot of research challenges such as consistency and personalization in dialog systems, multimodal emotion recognition, expressive speech synthesis and so on. In this talk, we will briefly introduce our work responding to these challenges, present a preliminary proof-of-concept prototype and discuss future research perspectives.

About company:

SpeechX was founded in Hong Kong and Shenzhen in 2016. The founders are mainly from the Human-Computer Communications Laboratory (HCCL) in The Chinese University of Hong Kong. The mission of SpeechX is to empower language learning with AI to be more efficient, productive and enjoyable.

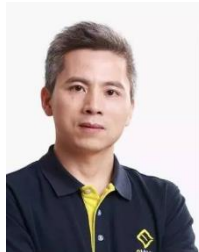
Speaker: 陈文明，壹秘科技创始人

Title: 专业商务智能语音的应用及挑战

Time: 10:30 – 12:00, December 7, 2018

Venue: International Lecture Hall

Biography:



陈文明，深圳壹秘科技有限公司创始人，中欧国际工商学院 EMBA。在音视频、智能语音、智能家居、物联网领域工作 18 年；曾于 TCL 就职 10 余年，历任研发总经理、产品总经理、电声事业部总经理、创新事业部总经理；2016 年 8 月创立深圳壹秘科技有限公司。

Abstract:

报告将以壹秘产品及服务的应用场景及市场潜力作为切入点，分享深圳壹秘科技有限公司争做智能语音单项技术应用冠军的心路历程，进而从前端语音处理的技术瓶颈、后端语言处理技术的挑战机遇两方面阐述专业商务智能语音的应用和挑战。

About company:

深圳壹秘科技有限公司成立于 2016 年，专注移动办公产品创新及智能服务。研发的人工智能会议服务系统，基于智能语音前端阵列算法技术、自然语言处理技术、网络通讯技术，服务于全球移动办公及智能会议市场。公司现有研发人员 25 人，研发具有自主知识产权的语音阵列算法、语音通话及 AI 服务产品，拥有多项软件著作权和发明专利。在全球渠道、创新算法研发、新型产品及智能服务开发等方面有较强的竞争力。

Speaker: 杨嵩，好未来 AI LAB 语音技术负责人

Title: AI 在教育领域落地的探索

Time: 10:30 – 12:00, December 7, 2018

Venue: International Lecture Hall

Biography:



杨嵩，历任思必驰高级语音工程师、苏州驰声研发主管、好未来 AI LAB 语音技术负责人。研究方向为语音识别、语音评测。一直致力于中高考英语口语机器评分，在线教育课堂质量自动化评估等方面工作，在该领域拥有多项专利。2014 年获中国人工智能学会颁发的“吴文俊人工智能科学技术奖进步奖”。

Abstract:

好未来教育集团以“科技推动教育进步”作为自己的使命，深入发掘 AI 技术和教育场景的结合点。针对教学资源不均衡，优质师资不足的现状，发展各个场景的“AI 助手”辅助教学；针对学生能力发展不平衡，推广个性化教学。此外为教育的各个环节引入不同的 AI 评测技术；在线下课堂教学中提供智慧教室的解决方案，让教室拥有眼睛（摄像头），耳朵（麦克风），大脑（云）及其他器官（答题器，ipad），引入音视频量化教学过程，评价课堂的教学质量；在线上课堂通过识别和分析课堂内容，评价师生间的交互状况，抽取相关特征对师生进行匹配，提高教学效率。好未来以 AI 技术为引擎，持续探索未来教育的新模式。

About company:

北京世纪好未来教育科技有限公司（NYSE: TAL）是一个以智慧教育和开放平台为主体，以素质教育和课外辅导为载体，在全球范围内服务公办教育，助力民办教育，探索未来教育新模式的科技教育公司。好未来全面布局教育产业，构建智慧教育、教育云、内容及未来教育、K12 及综合能力和国际及终身教育五大事业群，旗下共有学而思、学而思网校等十多个业务品牌。连续三年入选“最具价值中国品牌 100 强”。

PAPER PRESENTATIONS

Oral Session A: Network Architectures and Algorithms & Web Mining	
Time: 10:30 – 12:00	Venue: CIII-201
Chair: Yao Ge, CUHK	
Affiliation: CUHK Speaker: Yao Ge Title: Robust Secrecy Design for MIMO SWIPT with Artificial Noise and Full-Duplex Receiver Jamming	
Affiliation: CUHK Speaker: Yiding Yu Title: Deep-Reinforcement Learning Multiple Access for Heterogeneous Wireless Networks	
Affiliation: Tsinghua Speaker: Yuyu Luo Title: DeepEye: Towards Automatic Data Visualization	
Affiliation: CUHK Speaker: Xin Li Title: Aspect Term Extraction with History Attention and Selective Transformation	
Affiliation: Tsinghua Speaker: Hongyu Lu Title: Between Clicks and Satisfaction: Study on Multi-Phase User Preferences and Satisfaction for Online News Reading	
Oral Session B: Computer Vision and Image Processing & Multimedia	
Time: 10:30 – 12:00	Venue: CIII-202
Chair: Linkun Lv, Tsinghua	
Affiliation: Tsinghua Speaker: Zhen Zheng Title: VersaPipe: A Versatile Programming Framework for Pipelined Computing on GPU	
Affiliation: The Polytechnic University of Hong Kong Speaker: Dongmei Mo Title: Generalized Robust Regression for Jointly Sparse Subspace Learning	
Affiliation: Tsinghua Speaker: Guangxing Han Title: Semi-Supervised DFF: Decoupling Detection and Feature Flow for Video Object	
Affiliation: CUHK Speaker: Hang Zhou Title: Talking Face Generation by Adversarially Disentangled Audio-Visual Representation	
Affiliation: Tsinghua Speaker: Minhao Tang Title: A Universal Optical Flow Based Real Time Low Latency Omnidirectional Stereo Video System	

Oral Session C: Networking and Big Data

Time: 10:30 – 12:00

Venue: CIII-203

Chair: Menghao Zhang, Tsinghua

Affiliation: Tsinghua

Speaker: Kun Qian

Title: Awakening Power of Physical Layer: High Precision Time Synchronization for Industrial Ethernet

Affiliation: Tsinghua

Speaker: Xue Li

Title: ReGraph: A Graph Processing Framework that Alternately Shrinks and Repartitions the Graph

Affiliation: Tsinghua

Speaker: Menghao Zhang

Title: Control Plane Reflection Attacks in SDNs: New Attacks and Countermeasures

Affiliation: Tsinghua

Speaker: Nan Geng

Title: Flow-level Traffic Engineering in Conventional Networks with Hop-by-Hop Routing

Affiliation: Tsinghua

Speaker: Dongbiao He

Title: Joint Rate and FoV adaptation in immersive video streaming

Oral Session D: Networking and Big Data

Time: 13:30 – 15:00

Venue: CIII-201

Chair: Minghua Ma, Tsinghua

Affiliation: Tsinghua

Speaker: Xuya Jia

Title: Link Fault Protection and Traffic Engineering in Hybrid SDN Networks

Affiliation: Tsinghua

Speaker: Minghua Ma

Title: Robust and Rapid Adaption for Concept Drift in Software System Anomaly Detection

Affiliation: Tsinghua

Speaker: Xiaoli Zhang

Title: Generic and Agile Service Function Chain Verification on Cloud

Affiliation: Tsinghua

Speaker: Zhan Yulong

Title: A More Scalable Scheduling Algorithm for FPGA-based Time-Triggered Network

Affiliation: Tsinghua

Speaker: Wenkai Cui

Title: Real-time Total Focusing Method for Ultrasonic Imaging of Multilayered Object

Oral Session E: Audio, Speech and Natural Language Processing

Time: 13:30 – 15:00

Venue: CIII-202

Chair: Peng Gao, CUHK

Affiliation: CUHK

Speaker: Peng Gao

Title: Question-Guided Hybrid Convolution for Visual Question Answering

Affiliation: Northwestern Polytechnical University

Speaker: Xiaochun An

Title: A Kullback-Leibler Divergence Based Recurrent Mixture Density Network for Acoustic Modeling in Emotional Statistical Parametric Speech Synthesis

Affiliation: Northwestern Polytechnical University

Speaker: Ke Wang

Title: A Pitch-Aware Approach to Single-Channel Speech Separation

Affiliation: CUHK

Speaker: Siyuan Feng

Title: Exploiting Speaker and Phonetic Diversity of Mismatched Language Resources for Unsupervised Subword Modeling

Oral Session F: Audio, Speech and Natural Language Processing

Time: 13:30 – 15:00

Venue: CIII-203

Chair: Liumeng Xue, NWPU

Affiliation: Northwestern Polytechnical University

Speaker: Qing Wang

Title: Unsupervised Domain Adaptation via Domain Adversarial Training for Speaker Recognition

Affiliation: CUHK

Speaker: Xixin Wu

Title: Feature Based Adaptation for Speaking Style Synthesis

Affiliation: Northwestern Polytechnical University

Speaker: Liumeng Xue

Title: A Comparison of Expressive Speech Synthesis Approaches based on Neural Network

Affiliation: CUHK

Speaker: Songxiang Liu

Title: Voice Conversion Across Arbitrary Speakers based on a Single Target-Speaker Utterance

Oral Session G: Multimedia	
Time: 13:30 – 15:00	Venue: CIII-204
Chair: Yuanpeng Xiong, Tsinghua	
Affiliation: Tsinghua Speaker: Weinan Shi Title: TOAST: Ten-Finger Eyes-Free Typing on Touchable Surfaces	
Affiliation: Tsinghua Speaker: Xiyu Yan Title: Online Visual Tracking with High-order Pooling	
Affiliation: Tsinghua Speaker: Yan Yukang Title: VirtualGrasp: Leveraging Experience of Interacting with Physical Objects to Facilitate Digital Object Retrieval	
Affiliation: Tsinghua Speaker: Suping Zhou Title: Emotion Inferring from Large-scale Internet Voice Data: A Multimodal Deep Learning Approach	
Oral Session H: Audio, Speech and Natural Language Processing	
Time: 15:30 – 17:00	Venue: CIII-201
Chair: Xiaochun An, NWPU	
Affiliation: Northwestern Polytechnical University Speaker: Shan Yang Title: Enhancing Hybrid Self-Attention Structure with Relative-Position-Aware Bias for Speech Synthesis	
Affiliation: CUHK Speaker: Jianwei Yu Title: Development of the CUHK Dysarthric Speech Recognition System for the UA Speech Corpus	
Affiliation: Northwestern Polytechnical University Speaker: Yan Li Title: Integrating Deep and Shallow Models for Multi-Modal Depression Analysis — Hybrid Architectures	
Affiliation: CUHK Speaker: Xu Li Title: Unsupervised Discovery of Non-native Phonetic Patterns in L2 English Speech for Mispronunciation Detection and Diagnosis	

Oral Session I: Audio, Speech and Natural Language Processing

Time: 15:30 – 17:00

Venue: CIII-202

Chair: Shoukang Hu, CUHK

Affiliation: CUHK

Speaker: Shoukang Hu

Title: Gaussian Process Neural Networks for Speech Recognition

Affiliation: Northwestern Polytechnical University

Speaker: Haifeng Chen

Title: Bipolar Disorder Recognition with Histogram Features of Arousal and Body Gestures

Affiliation: Tsinghua

Speaker: Yequan Wang

Title: Sentiment Analysis by Capsules

Affiliation: Northwestern Polytechnical University

Speaker: Xiang Hao

Title: An Attention-based Neural Network Approach for Single Channel Speech

Affiliation: CUHK

Speaker: Yuewen Cao

Title: Rapid Style Adaptation Using Residual Error Embedding for Expressive Speech Synthesis

Oral Session J: Artificial Intelligence

Time: 15:30 – 17:00

Venue: CIII-203

Chair: Yuanpeng Xiong, Tsinghua

Affiliation: Tsinghua

Speaker: Shuang Hu

Title: Model Error Correction in Data Assimilation by Integrating Neural Networks

Affiliation: Tsinghua

Speaker: Guorun Yang

Title: SegStereo: Exploiting Semantic Information for Disparity Estimation

Affiliation: Tsinghua

Speaker: Zhuobin Zheng

Title: Self-Adaptive Double Bootstrapped DDPG

Affiliation: Tsinghua

Speaker: Kun Kuang

Title: Stable Prediction across Unknown Environments

THE INTERNATIONAL LECTURE HALL, BUILDING CII, GRADUATE SCHOOL AT SHENZHEN, TSINGHUA UNIVERSITY MAP



- ★ CII一楼国际会议中心（注册&主会场）
Building CII, International Lecture Hall
(Registration & Conference Venue)
- ▲ 快乐食间餐厅（午餐&晚餐）
Happy Meal Time (Lunch & Dinner)



- ★ CII一楼国际会议中心（注册&主会场）
Building CII, International Lecture Hall
(Registration&Conference Venue)
- ▲ 快乐食间餐厅（午餐&晚餐）
Happy Meal Time (Lunch&Dinner)

ACKNOWLEDGMENTS

Organizing Chairs:

Juanzi Li, Tsinghua University

Helen Meng, The Chinese University of Hong Kong

Yanning Zhang, Northwestern Polytechnical University

Steering Committee Chairs:

Zhiyong Wu, Tsinghua University

Jia Jia, Tsinghua University

Lei Xie, Northwestern Polytechnical University

Zhiyuan Liu, Tsinghua University

Hong Cheng, The Chinese University of Hong Kong

Minghua Chen, The Chinese University of Hong Kong

Organizing Committee Members:

Runnan Li, THU

Shuai Yang, THU

Shaoguang Mao, THU

Mu Wang, THU

Yao Du, THU

Xixin Wu, CUHK

Xin Yi, THU

Jingheng Xu, THU

Xiao Li, THU

Xiaochun An, NPU

Dongyang Dai, THU

Kun Zhang, THU

Liangqi Liu, THU

Hui Lu, THU

Yulan Chen, THU

Jiakun Hu, THU

Huirong Huang, THU

Bin Su, THU

Yingmei guo, THU

Xiangyu Liang, THU

Chao Wang, THU

Michael Lao Ban Teng, THU

Xiong Cai, THU

Liyan Dai, THU

Legend:

THU	Tsinghua University
CUHK	The Chinese University of Hong Kong
NPU	Northwestern Polytechnical University

SPONSORS

Tsinghua University

The Chinese University of Hong Kong

Northwestern Polytechnical University

Dolby Laboratories International Services (Beijing) Ltd

TAL AI Lab

eMeet

Tencent AI Lab

Shenzhen SpeechX Ltd



[illegible]
