

IMI RESEARCH SEMINAR

Tuesday 24 September 2019 11:00 am – 12:45 pm
Institute for Media Innovation Seminar Room

Research Techno Plaza, XFrontiers Block, Level 03-01, 50 Nanyang Drive, Singapore 637553

* Attendance is on first-come first-served basis due to limited seating

+ Lunch will be served

11:00am – 11:30am

25 mins presentation, 5 mins Q&A



Dr Li Hanhui

Research Fellow

Institute for Media Innovation

Multi-Task Distillation for Indoor Semantic Segmentation

The goal of indoor semantic segmentation is to predict the semantic label of each pixel, and its related techniques could be applied into various applications like scene understanding, robot navigation and human-computer interaction. Current methods on this topic mainly adopt the Encoder-Decoder network architecture, in which the encoder extracts high-level feature from image while the decoder turns feature into semantic prediction. More recently, multi-task distillation methods, which incorporate extra decoders to perform segmentation and other visual tasks simultaneously, such as depth estimation and surface normal estimation, have boosted the performance of baselines significantly. Therefore, in this presentation, we will overview several state-of-the-art multi-task distillation models, and discuss the possible improvements on them.

About the Speaker

Hanhui Li received the Ph.D. degree in Computer Software and Theory from Sun Yat-sen University, Guangzhou, China, in 2018, where he also received the B.S. degree in Computer Science and Technology in 2012. Currently, he is working as a research fellow in Institute for Media Innovation, Nanyang Technological University, Singapore. His research interests include deep learning, image processing and computer vision.

11:30am – 11:50am

15 mins presentation, 5 mins Q&A



Ajay Vishwanath

Research Associate

Institute for Media Innovation

Multimodal Affect Recognition from Audio, Video and Language Cues

Multimodal affect recognition is a challenging problem in affective computing. The aim is to combine unimodal recognition results from audio, video and language cues, to achieve the best prediction of human emotion. In addition to deciding the data set and modelling technique for the unimodal data, the crucial design consideration in multimodal fusion is deciding when to combine the unimodal features. In this work, we find the best combination of dataset, algorithm and fusion techniques. Finally, we present some preliminary results and compare the various design decisions.

About the Speaker

Ajay Vishwanath is a research associate at Institute for Media Innovation (IMI) / BeingTogether Centre (BTC) in Nanyang Technological University, Singapore. He completed his B.E. at Manipal University in Electronics and Communication Engineering, in 2013. Between 2013 and 2016, Ajay worked in software engineering and information technology for various corporations. In 2018, he completed his Masters in Computing from National University of Singapore with a specialization in machine learning. Currently, he is working on the Virtual Assistant platform with a focus on human-machine interaction, affective computing and virtual humans.

11:50am – 12:10pm

15 mins presentation, 5 mins Q&A



Chen Yuedong Donald

Research Associate

Institute for Media Innovation

Facial Motion Prior Networks for Facial Expression Recognition

Deep learning based facial expression recognition (FER) has received a lot of attention in the past few years. Most of the existing deep learning based FER methods do not consider domain knowledge well, which thereby fail to extract representative features. In our work, we propose a novel FER framework, named Facial Motion Prior Networks (FMPN). Particularly, we introduce an addition branch to generate a facial mask so as to focus on facial muscle moving regions. To guide the facial mask learning, we propose to incorporate prior domain knowledge by using the average differences between neutral faces and the corresponding expressive faces as the training guidance.

About the Speaker

Yuedong is now a Research Associate from the Institute for Media Innovation (IMI) at Nanyang Technological University, Singapore. He received his B.Eng and M.Eng from Sun Yat-sen University (SYSU) in 2016 and 2018, respectively. And his current research interests lie in the fields of Face Analysis and Deep Learning.

12:10pm – 12:30pm

15 mins presentation, 5 mins Q&A



Dao Duy Son

Project Officer

Institute for Media Innovation

Object Completion in Clutter

Reasoning occluded parts from the visibility of objects and inferring object order in a clutter room are significant tasks for scene understanding and object manipulation. In this work, we propose a unified framework that simultaneously optimizes three tasks: predict the shape of full object by a modal segmentation, completion the invisible parts of objects by generative network and scene composition by depth information. To evaluate our approach, we present a synthetic dataset which contains multiple room objects in occlusion. We also conduct experiments on this dataset to make comparison between our framework and some state-of-the-art methods in tasks of a modal segmentation and recovery of objects and room scenes.

About the Speaker

Dao Duy Son completed his Bachelor's Degree in Information Technology from Hanoi University of Science and Technology. He is currently working as a Project Officer at Institute for Media Innovation at Nanyang Technological University, Singapore.

12:30pm – 12:45pm

Discussion & Closing Remarks

12:45pm

End of Research Seminar