University of Chicago, Computer Science Department
Machine Learning Seminar
PRESENTS:

Yi Fang, Asst. Professor
NYU Abu Dhabi and NYU Tandon

Title: Deep Learning for 3D Shape Recognition and Registration

Abstract: With the availability of both large 3D datasets and unprecedented computational power, researchers have been shifting their focus on applying deep learning to address the challenges in specific tasks such as 3D classification, registration, recognition, and correspondence. The deep learning model often takes input as a grid structured input to effectively exploit discrete convolutions on the data as its fundamental building blocks. However, the irregular Non-Euclidean 3D data representation poses a great challenge for directly applying standard convolutional neural networks to 3D applications such as object recognition from 3D point clouds, 3D shape registration and matching, 3D localization and mapping and so on. In this talk, to address the challenges of learning with irregular 3D data representation, I will discuss our lab's recent efforts in the development of an encoder-free design of deep neural network architecture, which we apply to 3D deep learning for shape recognition and registration. The mainstream 3D deep learning efforts require explicitly designed encoders to extract deep shape features and/or spatial-temporal correlation features from irregular 3D data representations. By contrast, we acknowledge the challenges in designing an explicit form of an encoder to extract deep features from unstructured 3D data. As a result, we propose our novel approaches to work around this issue by implicitly extracting the deep features towards various 3D tasks. Our key novelty is that we present a novel unified concept of task-specific latent code (TSLC). The TSLC can take on different forms depending on the nature of the task. It represents a 3D shape descriptor for shape recognition, a 3D shape spatial correlation tensor for shape alignment, or a spatial-temporal descriptor for 3D group registration. The TSLC captures the geometric information from unstructured 3D data essential to each task and is used as input to a task-specific decoder to produce the desired output. Particularly, our approach starts with a randomly initialized TSLC. Next, at training time, we jointly optimize the latent shape code and update the neural network decoder's weights towards the minimization of a task-specific loss, while at inferencing time we hold the decoder's weights fixed and only optimize the TSLC. Our novel encoder-free approach brings forth two unique advantages: 1) it avoids the inclusion of an explicit 3D feature encoder for irregular 3D data representation. 2) It enhances the flexibility of feature learning for unseen data. The new design centers around the combination of optimization and learning, enabling further fine-tuning on the test data for better generalization abilities. By contrast, the conventional neural network does not have the flexibility in fine-tuning at the testing phase. We conducted experiments on a variety of tasks, including the unsupervised learning of 3D registration, 3D correspondence, and 3D recognition. Qualitative and quantitative comparisons on these experiments demonstrate that our proposed method achieves superior performance over existing methods.

Bio: Yi Fang is an Assistant Professor of Electrical and Computer Engineering at the NYU Abu Dhabi and NYU Tandon. He directs the NYU Multimedia and Visual Computing Lab (http://wp.nyu.edu/mmvc). He received his Ph.D. degree from Purdue University, West Lafayette in 2011. Upon one year of industry experience in Siemens in Princeton, New Jersey and in Riverain Technologies in Dayton, Ohio, and a half-year academic experience as a senior staff scientist at Department of Electrical Engineering and Computer science, Vanderbilt University, Nashville, he joined New York University Abu Dhabi as an Assistant Professor of Electrical and Computer Engineering. Professor Fang's research focuses on 3D Computer Vision and Machine Learning with applications to robotics and autonomous driving. He is currently working on the development of 3D deep learning technologies in large-scale visual computing, cross-domain and cross-modality models, and their various industrial applications. Professor Fang has published more than 50 papers in international journals and conferences including IEEE TPAMI, TIP, TCybernetics, TNNLS, CVPR, ICCV, ECCV, NeurIPS, AAAI, IJCAI, ACMMM, IROS, etc. Professor Fang has served as the publicity chair for the 27th ACM International Conference on Multimedia (ACMMM) in Nice, France. He has served as co-chair for the IROS 2019 session "Human-centered Robotics" and will serve as the program chair for the Asian Conference on Machine Learning (ACML 2021) and the area chair of Conference on Computer Vision and Pattern Recognition (CVPR 2021).

Host: Yuxin Chen **Zoom link for the seminar: https://uchicago.zoom.us/j/814358865**

Tuesday, April 14 10:30 am