Self-supervised deep learning in imaging science

Hui Ji

National University of Singapore

ABSTRACT

In last few years, deep learning has rapidly become a prominent tool for solving many challenging problems in imaging science. While most existing methods are supervised over a dataset of many degraded/truth image pairs to learn how to predict a truth image from the collected measurement, there is an increasing interest on studying powerful deep learning methods for image recovery that is dataset-free. In this talk, we will introduce a general self-supervised deep learning method for imaging science that does not have any prerequisite on training samples. The key parts include random-sampling-based data augmentation technique and Bayesian-deep-network-based approximate minimum mean square error (MMSE) estimator. Extensive experiments showed that our dataset-free deep learning method can compete well against existing supervised-learning-based solutions to many tasks in imaging science, e.g., image denoising, image deblurring, and compressed sensing, and phase retrieval.

When the rubber hits the road: data science at Grab

Hannes Martin Kruppa

Grab, Singapore

ABSTRACT

In this presentation I will be highlighting a number of Data Science key lessons learned at Grab.

How to reconcile benefits of large models with actual production requirements?

How does time-to-market influence model design?

Where are black-box models appropriate?

How to build for robustness to shifts in user behavior?

Two ways to win: finding the sweet spot in academic/industrial collaboration.

Neuralizing Algorithms

Wee Sun Lee

National University of Singapore

ABSTRACT

Most AI problems that we would like to solve are computationally intractable in the worst case. However, we are mostly interested in solving a small fraction of possible problems, those that occur in the real world. Approximation algorithms work well on some of these problems but are not tuned for the target problem distribution. Machine learning provides the tools to develop approximation algorithms that work well on problems encountered in practice, by training the algorithms on problems sampled from real world distributions. We modify known approximation algorithms and augment them with learnable components to allow learning of more powerful approximation algorithms that are then tuned to work on the target problem distribution. We demonstrate how to do this by developing Factor Graph Neural Network, a high order graph neural network based on loopy belief propagation on factor graphs and Particle Filter Recurrent Neural Network, a recurrent neural network based on particle filters. We also examine the advantages of decomposing a large problem into algorithmic components to simplify the design of a large system, while doing end-to-end learning to ensure that the entire system works well on the target problem distribution.

Machine learning and dynamical systems

Qianxiao Li

National University of Singapore

ABSTRACT

In this talk, we discuss some recent work on the connections between machine learning and dynamical systems. These come broadly in three categories, namely machine learning via, for and of dynamical systems. In the first direction, we introduce a dynamical approach to deep learning theory with particular emphasis on its connections with approximation and control theory. In the second direction, we discuss the approximation and optimization theory of learning input-output temporal relationships using recurrent neural networks and variants, with the goal of highlighting key new phenomena that arise in learning in dynamic settings. In the last direction, we discuss some principled methods that learns stable and interpretable dynamical model from data arising in scientific applications.

Data analytics - from research to real-world impact

Xiaoli Li

*A*STAR, Singapore*

ABSTRACT

Data analytics aims to extract hidden knowledge and insights from large amount of data. With the recent development of technology, more and more data have been accumulated and data analytics can play critical role in this digital world. In this talk, he will share some data analytics research that he and his colleague have done, as well as some real-world applications and impact.

Continual learning, overview and potential directions

Min Lin

Sea Ltd, Singapore

ABSTRACT

Machine learning is often considered to consist of two phases, training and inference. In the training phase, an offline collected data set is iterated for multiple epochs to optimize the network parameters. During the inference phase, the optimized parameters are kept fixed, and the fixed neural network is only used for making predictions. This offline learning scheme has successfully solved many problems, especially, the success stories of deep learning largely follow this scheme. Nevertheless, the ability to "continual learning" is still a desired or even an indispensable property under many scenarios. In this presentation, I will give an introduction to continual learning, including the problem, the existing approaches and potential directions.

Deep Approximation via Deep Learning

Zuowei Shen

National University of Singapore

ABSTRACT

The primary task of many applications is approximating/estimating a function through samples drawn from a probability distribution on the input space. The deep approximation is to approximate a function by compositions of many layers of simple functions, that can be viewed as a series of nested feature extractors. The key idea of deep learning network is to convert layers of compositions to layers of tuneable parameters that can be adjusted through a learning process, so that it achieves a good approximation with respect to the input data. In this talk, we shall discuss mathematical theory behind this new approach and approximation rate of deep network; we will also how this new approach differs from the classic approximation theory, and how this new theory can be used to understand and design deep learning network.

Can algorithms collaborate? The replica exchange method and its spectral gap

XIN TONG

National University of Singapore

ABSTRACT

Gradient descent (GD) is known to converge quickly for convex objective functions, but it can be trapped at local minima. On the other hand, Langevin dynamics (LD) can explore the state space and find global minima, but in order to give accurate estimates, LD needs to run with a small discretization step size and weak stochastic force, which in general slows down its convergence. This paper shows that these two algorithms can "collaborate" through a simple exchange mechanism, in which they swap their current positions if LD yields a lower objective function. This idea can be seen as the singular limit of the replica-exchange technique from the sampling literature. We show that this new algorithm converges to the global minimum linearly with high probability, assuming the objective function is strongly convex in a neighborhood of the unique global minimum. By replacing gradients with stochastic gradients, and adding a proper threshold to the exchange mechanism, our algorithm can also be used in online settings. We further verify our theoretical results through some numerical experiments, and observe superior performance of the proposed algorithm over running GD or LD alone. We will further explain how does replica exchange method improve LD's spectral gap using Poincare type of inequality.

Overview of recent advances in deep reinforcement learning

ZHONGWEN XU

Sea Ltd, Singapore

ABSTRACT

Rapid progress has been observed in the field of Deep Reinforcement Learning (Deep RL). In this talk, we'll first introduce fundamental concepts of reinforcement learning. Then we'll review the representative achievements of deep reinforcement learning, including the first breakthrough of Deep RL – Deep Q-learning Network, which learned from pixel inputs and achieved human performance in Atari games, and the research work of AlphaGo series, which beat the renowned human champions of Go and demonstrated the great power of AI. Finally we'll discuss potential applications of Deep RL and future research directions.

Leveraging synthetic data and self-supervision for 3D hand pose estimation

Yao Yingjie Angela

National University of Singapore

ABSTRACT

We consider vision-based 3D hand pose estimation – a structured prediction problem at the core of many HCI and AR/VR applications. Hand pose estimation has been studied for over two decades in computer vision but remains a highly challenging task. A main reason is that it is incredibly difficult to obtain sufficiently accurate ground truth labels required for standard supervised learning algorithms. In this talk, I will outline a series of works in which we leverage synthetic data and self-supervision to learn models for 3D hand pose and hand shape estimation. Central to our approaches is an efficient yet differentiable neural rendering model that unifies model-based fitting with conventional neural network training. Our experimental results are competitive with fully supervised approaches, highlighting the possibility for learning systems under label-deficient settings.

Learning with limited resources

Joey Tianyi Zhou

*A*STAR, Singapore*

ABSTRACT

Machine learning has experienced a strong growth in recent years, due to increased dataset sizes and computational power, and to advances in algorithms that has made great achievements. However, the state-of-art machine learning methods especially deep models, which are trained over a large number of high-quality data with known labels from standard benchmarks, tend to be generic. Unfortunately, in many tasks, it is often expensive to annotate or even difficult to collect sufficient training data to build machine learning-based information systems. Another challenge is coming from the limited hardware resource. Currently machine learning methods especially deep learning based models are computationally-hungry. During the last decades, we have seen how data is migrating, first from on-premises to cloud data centers and now, from cloud to the "edge" points closer to the source, where it's being generated. Edge AI devices operate with tight resource budgets such as memory, power and computing horsepower. AI technology with high-end GPUs for training and running large neural networks are not suitable for edge AI. The objective of the proposed research is to tackle the low-resource challenges such as limited data or constrained computing resources in real-world applications.