



RESEARCH SEMINAR

## Optimization Acceleration for Faster Training Deep Models



**Dr Pan ZHOU**

Senior Research Scientist  
Sea AI Lab  
Singapore

**Date : 10 March 2023 (Fri)**  
**Time : 11:00 am - 12:00 pm**  
**Venue : PQ304**

### Abstract

Training deep networks on large-scale datasets is computationally challenging. In this talk, I provide two solutions for the problem of how to train networks faster: 1) a faster optimizer Adan, and 2) a general acceleration technique, Win, to speed up previous optimizers.

Adan reformulates Nesterov acceleration to estimate the first- and second-order moments of the gradient in Adam for convergence acceleration. Adan with much less training cost can surpass the corresponding SoTA optimizers (e.g. SGD, Adam, AdamW) across many models (e.g., ResNet, ConvNext, ViT, Swin, MAE, LSTM, Transformer-XL, BERT, UNet, GPT). Moreover, Adan shows great tolerance to a large range of minibatch size, e.g., from 1k to 32k.

Win integrates the proximal point method with Nesterov acceleration, and induces an effective plug-in acceleration strategy for current optimizers. Experiments testify to the faster convergence speed and superior performance of Win-accelerated AdamW, Adam, LAMB, and SGD over their non-accelerated counterparts on CV and NLP tasks.

### About the Speaker

Dr Pan ZHOU received Ph.D. Degree from the National University of Singapore in 2019, and obtained a Master's Degree from Peking University in 2016. Now he is a senior research scientist at SEA AI Lab, Singapore. Before, he worked as a research scientist at Salesforce AI Lab from 2019 to 2021. His research interests include optimization, machine learning, and computer vision, and published about 40 papers, including TPAMI, ICML, NeurIPS, CVPR, etc. He was the winner of the Microsoft Research Asia Fellowship 2018.