PAPER LIST

November 6: Training Deep Neural Networks
Peri-LN: Revisiting Normalization Layer in the Transformer Architecture
Jeonghoon Kim, Byeongchan Lee, Cheonbok Park, Yeontaek Oh, Beomjun Kim, Taehwan Yoo, Seongjin Shin, Dongyoon Han, Jinwoo Shin, Kang Min Yoo. ICML 2025.

November 13 Convolutional Neural Networks
ShiftwiseConv: Small Convolutional Kernel with Large Kernel Effect
Dachong Li, Li Li, Zhuangzhuang Chen, Jianqiang Li. CVPR 2025.
https://github.com/lidc54/shift-wiseConv

November 20: Understanding and Visualizing CNNs
Direct Ascent Synthesis: Revealing Hidden Generative Capabilities in Discriminative Models
Stanislav Fort, Jonathan Whitaker. ArXiv Preprint ArXiv:2502.07753v1, February 2025.
https://github.com/stanislavfort/Direct_Ascent_Synthesis

November 27: Recurrent Neural Networks
Vision-LSTM: xLSTM as Generic Vision Backbone
Benedikt Alkin, Maximilian Beck, Korbinian Pöppel, Sepp Hochreiter, Johannes Brandstetter. ICLR 2025.
https://github.com/nx-ai/vision-lstm

December 4: Attention and Transformers
An Image is Worth More Than 16x16 Patches: Exploring Transformers on Individual Pixels
Duy Kien Nguyen, Mido Assran, Unnat Jain, Martin R. Oswald, Cees G. M. Snoek, Xinlei Chen. ICLR 2025.
https://github.com/insitro/ChannelViT

December 11: Graph Neural Networks
Improving the Effective Receptive Field of Message-Passing Neural Networks
Shahaf E. Finder, Ron Shapira Weber, Moshe Eliasof, Oren Freifeld, Eran Treister. ICML 2025.
https://github.com/BGU-CS-VIL/IM-MPNN

December 18: *Project Progress Presentations*

December 25: Language Model Pretraining
BERTs are Generative In-Context Learners
David Samuel. NeurIPS 2024.
https://github.com/ltgoslo/bert-in-context

January 1: *No Classes - New Year's Day*

January 8: Large Language Models
SFT Memorizes, RL Generalizes: A Comparative Study of Foundation Model Post-training
Tianzhe Chu, Yuexiang Zhai, Jihan Yang, Shengbang Tong, Saining Xie, Dale Schuurmans, Quoc V Le, Sergey Levine, Yi Ma. ICML 2025.
https://github.com/LeslieTrue/SFTvsRL