Yuanhao Xiong

yhxiong@cs.ucla.edu | +1 424-407-6112 | `https://xyh97.github.io`

EDUCATION

University of California, Los Angeles   **Los Angeles, USA**
Ph.D. student in Computer Science   *Sep. 2019 - Present*

Zhejiang University   **Hangzhou, China**
B.Eng. in Information Engineering   *Aug. 2015 - Jun. 2019*

RESEARCH INTERESTS

My research interests lie in developing machine learning algorithms in various related fields including **self-supervised learning**, **adversarial learning**, **learning to learn**, and **optimization**.

RESEARCH EXPERIENCE

Amazon   **Remote**
Applied Scientist Intern, Advisor: Wei-Cheng Chang   *June. 2021 - Present*

Self-Supervised Learning for Extreme Multi-label Classification
- Proposed a new and general setting for XMC, extreme zero-shot XMC (EZXMC), where no supervision is needed and merely raw texts of instances and labels are accessible.
- Developed an effective self-supervised method called SSXMC, and demonstrated that SSXMC could outperform its unsupervised counterparts significantly as well as the state-of-the-art supervised extreme classifiers when fine-tuned on a limited number of positive pairs.

University of California, Los Angeles   **Los Angeles, CA, U.S.**
Research Assistant, Advisor: Prof. Cho-Jui Hsieh   *Sep. 2019 - Present*

Graph Network-based Scheduler
- Designed a novel Graph Network-based Scheduler (GNS) by constructing a directed graph for the underlying neural network of the target problem, to improve the optimization performance.
- Evaluated GNS on benchmarking datasets, Fashion-MNIST and CIFAR10 for image classification, and GLUE for language understanding.
- Demonstrated that GNS was both informative to to encode rich information from intermediate layers and generalizable to different architectures.

Optimizer Benchmarking
- Proposed a new benchmarking protocol to evaluate both end-to-end efficiency (training a model from scratch without knowing the best hyperparameter) and data-addition training efficiency (the previously selected hyperparameters are used for periodically re-training the model with newly collected data).
- Evaluated end-to-end efficiency considering hyperparamter tuning with Hyperband.
- Designed a new protocol for assessing the hyperparameter sensitivity to data shift.
- Tested our framework on 7 different optimizers with popular tasks.

Learning to Learn for Adversarial Training
- Leveraged a learning-to-learn (L2L) framework to train an optimizer with recurrent neural networks (RNN), providing update directions and steps adaptively for the inner problem.
- Improved robust accuracy of PGD-based adversarial training and TRADES throught co-training optimizer's parameters and model's weights.

Learning to Learn by Zeroth-Order Oracle

- Extended the learning to learn (L2L) framework from first-order to zeroth-order (ZO) optimization.
- Proposed a novel RNN optimizer architecture which learns both parameter update rule and Gaussian sampling rule for ZO optimization.

AI Labs at Didichuxing Inc.   **Beijing, China**
Research Intern, Advisor: Prof. Yan Liu   *Jul. 2018 - Oct. 2018*

Improved Adaptive Optimization Algorithm

- Demonstrated that extreme learning rates in adaptive methods could lead to poor performance.
- Provided new variants of Adam, employing dynamic bounds on learning rates to achieve a gradual transition from Adam to SGD and gave a theoretical proof of convergence.

PUBLICATIONS

(* indicates equal contribution.)

[1] **Yuanhao Xiong**, Wei-Cheng Chang, Cho-Jui Hsieh, Hsiang-Fu Yu, and Inderjit Dhillon. Extreme zero-shot learning for extreme text classification. *arXiv preprint: 2112.08652*, 2022.

[2] **Yuanhao Xiong**, Li-Cheng Lan, Xiangning Chen, Ruochen Wang, and Cho-Jui Hsieh. Learning to schedule learning rate with graph neural networks. *Proceedings of the 10th International Conference on Learning Representations (**ICLR**)*, 2022.

[3] **Yuanhao Xiong**, Xuanqing Liu, Li-Cheng Lan, Yang You, Si Si, and Cho-Jui Hsieh. How much progress have we made in neural network training? A new evaluation protocol for benchmarking optimizers (Under Review). *arXiv preprint: 2010.09889*, 2020.

[4] **Yuanhao Xiong** and Cho-Jui Hsieh. Improved adversarial training via learned optimizer. In *Proceedings of the European Conference on Computer Vision (**ECCV**)*, 2020.

[5] Yangjun Ruan, **Yuanhao Xiong**, Sashank Reddi, Sanjiv Kumar, and Cho-Jui Hsieh. Learning to learn by zeroth-order oracle. In *Proceedings of the 8th International Conference on Learning Representations (**ICLR**)*, 2020.

[6] Chacha Chen, Hua Wei, Nan Xu, Guanjie Zheng, Ming Yang, **Yuanhao Xiong**, Kai Xu, and Zhenhui Li. Toward a thousand lights: Decentralized deep reinforcement learning for large-scale traffic signal control. In *Proceedings of the 34th AAAI Conference on Artificial Intelligence (**AAAI**)*, 2020.

[7] **Yuanhao Xiong**, Guanjie Zheng, Kai Xu, and Zhenhui Li. Learning traffic signal control from demonstrations. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management (**CIKM**)*, 2019.

[8] Guanjie Zheng, **Yuanhao Xiong**, Xinshi Zang, Jie Feng, Hua Wei, Huichu Zhang, Yong Li, Kai Xu, and Zhenhui Li. Learning phase competition for traffic signal control. In *Proceedings of the 28th ACM International Conference on Information and Knowledge Management, (**CIKM**)*, 2019.

[9] **Yuanhao Xiong***, Liangchen Luo*, Yan Liu, and Xu Sun. Adaptive gradient methods with dynamic bound of learning rate. In *Proceedings of the 7th International Conference on Learning Representations (**ICLR**)*, 2019.

[10] Mengyuan Lee, **Yuanhao Xiong**, Guanding Yu, and Geoffrey Ye Li. Deep neural networks for linear sum assignment problems. *IEEE Wireless Communications Letters*, 7(6):962–965, 2018.