Shen Zhuoran

cmsflash99@gmail.com | +1 425-428-3693 | cmsflash.github.io | github.com/cmsflash

Work Experience

Pony.ai, *Software Engineer, Prediction Department* *Nov. 2021 – Present*

- Responsible for the development of the next-generation, end-to-end, general-purpose vehicle trajectory prediction model for autonomous driving.

Google, *AI Resident, Google Brain, Google Research* *Oct. 2019 – Aug. 2021*

- Designed global self-attention networks (GSA-Nets), a novel family of neural networks for computer vision that use efficient attention mechanisms to fully replace convolution. Demonstrated superior trade-offs for accuracy vs. parameters, computation, and latency over CNNs.
- Worked on vision Transformer for open-world localization (OWL-ViT), a simple zero/few-shot object detection framework. Proposed the simplest recipe to transfer image-text pretraining models to detection. Achieved a new state-of-the-art for one-shot detection by a wide margin and competitive zero-shot results. To publish a paper at ECCV 2022.
- Developed an on-device age detector using cross-domain knowledge distillation. Deployed the model to user devices to support privacy-preserving data filtering for a confidential project.

Tencent, *Research Intern, Applied Research Center, Platform and Content Group* *Jul. 2019 – Sep. 2019*

- Proposed a novel efficient method for information propagation through an arbitrarily long video with constant complexity w.r.t. video length and linear complexity w.r.t. resolution. Presented a first-author paper at ECCV 2020.
- Developed the first real-time video object segmenter with state-of-the-art accuracy to support an on-device smart video editing feature of WeSee (a TikTok-like app in China with 40M+ DAUs).

SenseTime, *Research Intern, Intelligent Perception and Services Team, Smart City Group* *Jun. 2017 – Jun. 2019*

- Proposed a novel efficient attention mechanism, which reduces the memory and computational complexities of the attention mechanism from quadratic to linear. Demonstrated significant improvement in performance-cost trade-offs on a variety of tasks including object detection, instance segmentation, stereo depth estimation, and temporal action localization. Presented a first-author paper at WACV 2021.
- Integrated the efficient attention mechanism to the BG-wide object detection framework that supports research teams of 100+ people.

Education

The University of Hong Kong, *BEng Computer Science, GPA: 3.85/4.30, Standing: 1/111* *Sep. 2015 – Jun. 2019*

Awards

- **First Runner-up**, *ACM-HK Programming Contest 2017*

Publications and Preprint

- M. Minderer, A. Gritsenko, A. Stone, M. Neumann, D. Weissenborn, A. Dosovitskiy, A. Mahendran, A. Arnab, M. Dehghani, **Shen Z.**, X. Wang, X. Zhai, T. Kipf, N. Houlsby. (2022) *Simple Open-Vocabulary Object Detection with Vision Transformers.* ECCV 2022.
- **Shen Z.**, Zhang M., Zhao H., Yi S., Li H. (2021). *Efficient Attention: Attention with Linear Complexities*. WACV 2021.
- **Shen Z.**, I. Bello, R. Vemulapalli, Jia X., Chen C.-H. (2020). *Global Self-Attention Networks for Image Recognition*. arXiv: 2010.03019.
- Li Y.*, **Shen Z.***, Shan Y. (2020). *Fast Video Object Segmentation using the Global Context Module*. ECCV 2020. **Equal contribution.*

Skills

- **Languages**: Python, C++, Shell script, Markdown, LaTeX
- **Technologies**: TensorFlow, Keras, PyTorch, Git, Slurm, Vim, NumPy, OpenCV, Bazel, Django
- **Skills**: Deep learning, machine learning, computer vision, neural networks, motion prediction