Chuang Lin

Curriculum Vitae

(+61) 468820052
chuang.lin@monash.edu
https://clin1223.github.io/

Biography

I'm currently a third-year Ph.D. student at Monash University, supervised by Prof.Jianfei Cai and Prof.Gholamreza (Reza) Haffari. Previously, I received my M.Eng. degree (2019) from Harbin Institute of Technology and B.Eng. degree (2017) from Jilin University.
I'm broadly interested in computer vision and deep learning. My current research focuses on vision-and-language, domain adaptation.

Research Experience

03/2020-05/2022 **Research Intern**, *ByteDance, Beijing, China*.
Mentor: Dr. Changhu Wang and Dr. Zehuan Yuan

07/2019-02/2020 **Research Intern**, *National University of Singapore(NUS)*, School of Computing.
NUS-Tsinghua Extreme Search Center (NExT)
Advisor: Prof. CHUA Tat Seng and Dr. Meng Lei

05/2018-09/2018 **Research Intern**, *DiDi, Beijing, China*.
Mentor: Dr. Pengfei Xu

Publications

2022 **Learning Object-Language Alignments for Open-Vocabulary Object Detection**.
- **Chuang Lin**, Peize Sun, Yi Jiang, Ping Luo, Lizhen Qu, Gholamreza Haffari, Zehuan Yuan, Jianfei Cai.
- *IEEE Conference on International Conference on Learning Representations (ICLR)*, 2023.
- Formulate object-language alignment as a set matching problem between a set of image region features and a set of word embeddings. It enables us to directly train an open-vocabulary object detector from image-text pairs in a much simple and effective way.

Multimodal Transformer with Variable-length Memory for Vision-and-Language Navigation.
- **Chuang Lin**, Yi Jiang, Jianfei Cai, Lizhen Qu, Gholamreza Haffari, Zehuan Yuan.
- *IEEE Conference on European Conference on Computer Vision (ECCV)*, 2022.
- Propose Multimodal Transformer with Variable-length Memory (MTVM) for visually-grounded natural language navigation by modeling the temporal context explicitly. MTVM enables the agent to keep track of the navigation trajectory by directly storing activations in the previous step in a memory bank.

2021 **Domain-Invariant Disentangled Network for Generalizable Object Detection**.
- **Chuang Lin**, Zehuan Yuan, Sicheng Zhao, Changhu Wang, Jianfei Cai.
- *IEEE Conference on Computer Vision and Pattern Recognition (ICCV)*, 2021.
- Address the problem of domain generalizable object detection, which aims to learn a robust detector from multiple "seen" domains so that it can generalize well to other "unseen" domains.

2020 **Multi-source Domain Adaptation for Visual Sentiment Classification**.
- **Chuang Lin**, Sicheng Zhao, Lei Meng, CHUA Tat Seng.
- *In AAAI Conference on Artificial Intelligence*, 2020.
- Propose a novel multi-source domain adaptation method for visual sentiment classification via adversarial learning. It learns to find a unified sentiment latent space where data from both the source and target domains share a similar distribution.

2019 **CycleEmotionGAN: Emotional Semantic Consistency Preserved CycleGAN for Adapting Image Emotions**, *(Oral)*.

- Sicheng Zhao, **Chuang Lin**, Pengfei Xu, Sendong Zhao, Ravi Krishna, Guiguang Ding, Kurt Keutzer.
- In *AAAI Conference on Artificial Intelligence*, 2019.
- Propose a novel cycle-consistent adversarial model for unsupervised domain adaptation problem in image emotion classification. It can adapt source domain images to have similar distributions to the target domain without the use of aligned image pairs.

2018 **Cycle-consistency Based Hierarchical Dense Semantic Correspondence**, *(Oral)*.

- **Chuang Lin**, Hongxun Yao, Wei Yu, Xiaoshuai Sun.
- *IEEE International Conference on Image Processing (ICIP)*, 2018.
- Introduced the cycle-consistency guidance from higher-layer flow field to improve the lower-layer correspondence estimation. Our model outperforms the state-of-the-art methods on public benchmark Caltech101 and PASCAL in 2017 and early 2018, e.g. FCSS and Object-aware in CVPR 2018.

2017 **Multi-level Semantic Representation for Flower Classification**, *(Oral)*.

- **Chuang Lin**, Hongxun Yao, Wei Yu, Wenbo Tang.
- *Pacific-Rim Conference on Multimedia (PCM)*, 2017.
- Captured the subtle differences between sub-categories, with the complementary strengths of object-level models and part-level models. (Note that this work was done when I was an undergraduate)

Education Background

12/2020-Present **Ph.D.**, Faculty of Information Technology.
Monash University, Australia
Advisor: Prof. Jianfei Cai and Prof. Gholamreza (Reza) Haffari

09/2017-07/2019 **M.Eng.**, *School of Computer Science and Technology*.
Harbin Institute of Technology, China
Advisor: Prof. Hongxun Yao

09/2013-07/2017 **B.Eng.**, *School of Computer Science and Technology*.
Jilin University, China

Projects

2018.05–09 **DiDi**, *Beijing*, **Research Intern**, Mentor: Dr. Pengfei Xu.

Data Augmentation for Traffic Sign Detection .
Generated traffic sign images for the categories with few samples to tackle the data imbalance challenge using GANs.

- Preprocessed the categories with few samples by simulating with large-sample classes through traffic sign mask.
- Learned the traffic signs enhanced by simulation using CycleGAN to deal with the problem of data imbalance collected by driving recorders.
- Improved the detection accuracy 25% - 40% per small-sample class with new generated traffic sign images. The final model is applied to the DiDi Map Update Service.

Skills

Programming Languages: Python *(proficent)*, Matlab, C

Deep Learning Framework: PyTorch, caffe *(Linux)*