Chen Xu

Email: chenxu24568@gmail.com   **Date of Birth**: 1999.4.8   **Phone**: (+86)18852005156

RESEARCH
INTERESTS

Computer Vision: Generative Models, Vision-Language Models
Machine Learning: Deep-learning, Self-supervised learning

EDUCATION

Nanjing University   Nanjing, China
Master of Science in MCG Group   August 2021 – June 2024
Supervisor: Prof. Limin Wang (NJU MCG)
Nanjing University   Nanjing, China
Bachelor in National Basic Subject Talent Training Plan (Computer   August 2017 – June 2021
Science)   GPA: 4.38/5

PUBLICATIONS

• **Chen Xu***, Tianhui Song*, Weixin Feng, Xubin Li, Tiezheng Ge, Bo Zheng, Limin Wang. Accelerating Image Generation with Sub-path Linear Approximation Model. *European conference on computer vision (**ECCV'24**), **Oral**,* 2024.

• **Chen Xu**, Yuhan Zhu, Haocheng Shen, Boheng Chen, Yixuan Liao, Xiaoxin Chen, Limin Wang. Progressive Visual Prompt Learning with Contrastive Feature Re-formation. *International Journal of Computer Vision (**IJCV'24**),* 2024.

• **Chen Xu***, Yuhan Zhu*, Guozhen Zhang, Haocheng Shen, Yixuan Liao, Xiaoxin Chen, Gangshan Wu, Limin Wang. DPL: Decoupled Prompt Learning for Vision-Language Models. *In submission.*

• Yuhan Zhu, Guozhen Zhang, **Chen Xu**, Haocheng Shen, Xiaoxin Chen, Gangshan Wu, Limin Wang. Efficient Test-Time Prompt Tuning for Vision-Language Models. *In submission.*

HONORS &
SCHOLARSHIPS

• **Second Prize** in Postgraduate Academic Scholarship   2022&2023
• **First Prize** in Postgraduate Academic Scholarship (20%)   2021
• **Second Prize** in the people's scholarship in China   2019
• **Silver Award** in ACM-ICPC National Invitational Tournament (Informally participated)   2018
• **Third Prize** in Academic Outstanding Scholarship for Elite Class   2017

RESEARCH
EXPERIENCE

SPLAM: Accelerating Image Generation with Sub-path Linear Approximation Model
Advisor: Prof. Limin Wang   Nov 2023 - Mar 2024
• As the first author, proposed a novel approach for accelerating diffusion models for image generation. Motivated by *consistency models*, this work treats the separated PF-ODE trajectory as a series of PF-ODE sub-paths divided by sampled points. It introduces Sub-path linear (SL) ODEs to form a progressive and continuous error estimation along each PF-ODE sub-path, enabling the construction of denoising mappings with reduced cumulative error. An efficient distillation method is additionally provided for incorporating pre-trained models such as Stable Diffusion models.
• Extensive experiments on COCO 2014, and COCO 2017 datasets demonstrate SPLAM's remarkable training efficiency and state-of-the-art performance on image generation benchmarks with significantly fewer computational resources. The code is available at https://github.com/MCG-NJU/SPLAM.

DPL: Decoupled Prompt Learning for Vision-Language Models

Advisor: Prof. Limin Wang   April 2023 - Sep 2023

• As the first author, presented a novel prompt learning architecture for vison-language models, namely Decoupled Prompt Learning (DPL). This work decouples the attention operation, which includes prompt and instance inputs, into four sub-processes. DPL then strategically re-combines such sub-processes to emphasize the generalization effect while maintaining the effective adaptation ability.

• DPL achieves state-of-the-art results on base-to-new generalization, cross-dataset transfer, and domain generalization benchmarks with significantly fewer parameters (48x less), demonstrating its strong generalization capability. The paper is available at https://arxiv.org/abs/2308.10061.

ProVP-Ref: Progressive Visual Prompt Learning with Contrastive Feature Re-formation

Advisor: Prof. Limin Wang   Aug 2022 - Nov 2022

• We explored the potential of visual prompt learning in vision-language models. As the first author, I proposed a progressive visual prompt learning (ProVP) architecture, which strengthens cross-layer prompt interactions when deeply propagating image embeddings. A contrastive feature re-formation technique is further proposed to maintain alignment with the original CLIP visual feature distribution to prevent the deterioration of model generalization ability.

• ProVP-Ref is evaluated on three representative benchmarks: few-shot learning, base-to-new generalization, and cross-dataset transfer, and surpasses all previous methods both in adaptation and generalization capabilities. The code is available at https://github.com/MCG-NJU/ProVP.

INTERNSHIP   **Image Generation Research Internship**   Hangzhou, China

Alibaba Group   Nov 2023 – Mar 2024

Focusing on **generative models:**

• Build an efficient and effective framework for accelerating latent diffusion models based on Stable Diffusion models.

• The framework has shortened the inference step for text-to-image models to 2 or 4 steps to generate one image, thereby enabling fast image generation with low latency and high image quality.

Computer Vision Research Internship   Shenzhen, China

Vivo AI Lab   Aug 2022 – Feb 2023

Focusing on **low shot learning and transfer learning of vision-language models:**

• Build a progressive prompt learning framework based on CLIP, which stabilizes the prompt training and enables efficient transfer learning.

• Introduce the learning strategy into practical applications, help to decrease the demand of training data to 30% when transferring CLIP to adapt to vulgar image recognition tasks.

SKILLS   • Programming: Python, PyTorch, LaTeX, C++, C
              • Languages: Mandarin, English