

# WU CHENGYUE

Hong Kong

☎ +86-15019472631 ✉ [wcy010805@gmail.com](mailto:wcy010805@gmail.com)

## EDUCATION

---

**The University of Hong Kong**

*Ph.D. in Computer Science*

**Sept.2023 — June.2027 (expected)**

*Hong Kong, China*

**Harbin Institute of Technology, Honors School(Top 5% in HIT)**

*B.E. in Computer Science - CGPA : 95.6/100 - Rank :1/144(0.7%)*

**Aug.2019 — June.2023**

*Harbin, China*

**Harbin Institute of Technology(Shenzhen)**

*Exchange student in Computer Science*

**Aug.2021 – Aug.2022**

*Shenzhen, China*

## HONORS & AWARDS

---

**National Scholarship**

2019 — 2020

**Hong Kong PhD Fellowship**

2023

**Bytedance Scholarship** (200,000 RMB, 20 PhD students in China and Singapore)

2025

**HKU Presidential PhD Scholarship**

2023

**Outstanding Graduate of Heilongjiang Province**

2023

**Samsung Scholarship**

2020 — 2021

**Tencent Scholarship**

2021 — 2022

**Best Student Paper Award in ICDIS 2022**

2022

**Runner-up in NTIRE Challenge CVPR 2022 Image Inpainting**

2022

**First Prize of School Scholarship**

2019 — 2021

**First Prize in Chinese National Mathematics Modelling Contest for College Students**

Oct. 2021

**Honorable Award in American Mathematics Modelling Contest for College Students**

Mar. 2021

## PUBLICATIONS

---

- **Wu C**, Zhang H, Xue S, et al. Fast-dLLM v2: Efficient Block-Diffusion LLM[J]. ICLR 2026
- **Wu C**, Zhang H, Xue S, et al. Fast-dLLM: Training-free Acceleration of Diffusion LLM by Enabling KV Cache and Parallel Decoding[J]. ICLR 2026
- **Wu C**, Chen X, Wu Z, et al. Janus: Decoupling visual encoding for unified multimodal understanding and generation[J]. CVPR 2025.
- **Wu C**, Gan Y, Ge Y, et al. Llama pro: Progressive llama with block expansion[J]. ACL 2024.
- **Wu C**, Wang T, Ge Y, et al.  $\pi$ -Tuning: Transferring Multimodal Foundation Models with Optimal Multi-task Interpolation[C]//International Conference on Machine Learning. PMLR, 2023: 37713-37727.
- **Wu C**, Ge Y, Guo Q, et al. Plot2Code: A Comprehensive Benchmark for Evaluating Multi-modal Large Language Models in Code Generation from Scientific Plots[J]. NAACL, 2025.
- Lu Z, **Wu C**, Chen X, et al. Hierarchical diffusion autoencoders and disentangled image manipulation[C]//Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision. 2024: 5374-5383.
- Wang C, Chen X, **Wu C**, et al. Automatic time series forecasting model design based on pruning[J]. Applied Soft Computing, 2024: 111804.
- Wang Z, Duan S, **Wu C**, et al. Generative data augmentation for non-iid problem in decentralized clinical machine learning[C]//2022 4th International Conference on Data Intelligence and Security (ICDIS). IEEE, 2022: 336-343.
- Lu, Z., Huang, D., Bai, L., Qu, J., **Wu, C.**, Liu, X., & Ouyang, W. (2024). Seeing is not always believing: benchmarking human and model perception of AI-generated images. Advances in Neural Information Processing Systems, 36.

- Lu, Z., Wang, Z., Huang, D., **Wu, C.**, Liu, X., Ouyang, W., & Bai, L. (2024). Fit: Flexible vision transformer for diffusion model. arXiv preprint arXiv:2402.12376.
- Wang, J., Shao, W., Chen, M., **Wu, C.**, Liu, Y., Wu, T., ... & Luo, P. (2024). Adapting llama decoder to vision transformer. arXiv preprint arXiv:2404.06773.
- Ma, Y., Liu, X., Chen, X., Liu, W., **Wu, C.**, Wu, Z., ... & Ruan, C. (2024). JanusFlow: Harmonizing Autoregression and Rectified Flow for Unified Multimodal Understanding and Generation. arXiv preprint arXiv:2411.07975.
- Xiong, J., Liu, G., Huang, L., **Wu, C.**, Wu, T., Mu, Y., ... & Wong, N. (2024). Autoregressive Models in Vision: A Survey. arXiv preprint arXiv:2411.05902.

## PROJECTS

---

### NTIRE Challenge CVPR 2022 Image Inpainting [↗](#)

Feb.2022 — Apr.2022

- Propose to explore the types of mask used in the training process. At the same time, using our mask generation strategy can effectively improve the results of the model.
- Use joint spatial and frequency loss in spatial domain and frequency domain with a regular term to reconstruct the image.
- The proposed GLaMa was ranked first in terms of PSNR, LPIPS and SSIM in the NTIRE 2022 Image Inpainting Challenge Track 1 Unsupervised.

## INTERNSHIP

---

### Nvidia [↗](#)

Feb. 2025 – Present

#### *Research Intern*

Shenzhen, China

- Acceleration on Diffusion Large Language Models.

### DeepSeek [↗](#)

May 2024 – Nov.2024

#### *AGI Research Intern*

Beijing, China

- Research on Multi-Modal Large Language Models' understanding and generation application.
- Research on the data centric training for pretrained multimodal models.

### Tencent [↗](#)

Sept.2022 – May 2024

#### *Research Intern*

Shenzhen, China

- Research on the Large Language Models and Vision Language Models.
- Research on the data centric training for pretrained multimodal models.

### The University of Hong Kong. [↗](#)

Sept.2022 – Jun.2023

#### *Research Assistant*

Hong Kong, China

- Research on the similarity between intra- and inter-modal tasks.
- Research on the parameter-efficient transfer learning of vision-language pretrained models.

### TikTok [↗](#)

Jan.2022 – Apr.2022

#### *Backend Developer Intern*

Beijing, China

- Aimed to design and implement microservices with load balancing and high availability, based on the fact of providing TikTok's international live broadcast with the global distributed IDC architecture.
- Develop unified API interfaces to microservices for heterogeneous downstream data sources. Support configurable API, thus preventing manual rewrite and reducing human resources by up to 50%.

## TECHNICAL SKILLS

---

**Program Languages:** Python, Java, C, C++, SQL, Go

**Certification:** Certified Software Professional (CSP) Top 5.57%

**Language:** CET-6: 582, IELTS: 7.5, Mandarin (mother tongue)

**Developer Tools:** Pycharm, VS Code, Goland, IntelliJ Idea Ultimate

**Technologies/Frameworks:**Pytorch, Scikit-Learn, Linux, GitHub, KiteX