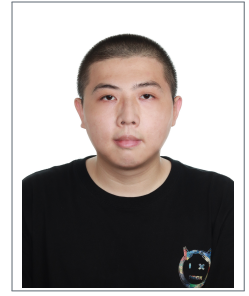


WENYANG LIU

Computer Vision · Multimodal AI · Foundation Models · Content Understanding

Singapore (EP-eligible) · +65 8387 8619 · wenyang001@e.ntu.edu.sg · Homepage · LinkedIn · GitHub · Scholar

Computer Vision & Multimodal AI researcher (Ph.D., NTU Singapore) with 5+ years developing foundation-model-based methods for visual content understanding and real-world vision applications. First author of four CCF-A papers at leading venues, including CVPR, ACM MM, and IEEE TMM ($\times 2$), with 20+ publications and 340+ citations. Experience includes few-/zero-shot learning, open-vocabulary perception, and image/video restoration, with methods validated on public benchmarks and real industrial data and vision systems demonstrated to industry partners.



EDUCATION

Nanyang Technological University, Singapore

Jan 2022 – Nov 2025

Ph.D. in Electrical and Electronic Engineering | Advisors: Prof. Lap-Pui Chau and Assoc. Prof. Kim-Hui Yap

Thesis: *From File Fragments to Images: Vision Image Restoration and Super-Resolution Using Deep Learning*

Chongqing University, China

M.Sc. in Computer Science | Advisor: Assoc. Prof. Weichen Liu

Sep 2017 – Jul 2020

B.Sc. in Internet of Things

Sep 2013 – Jul 2017

TECHNICAL SKILLS

Foundation & Multimodal Models: CLIP, DINOv3, SigLIP, ImageBind, SAM, Grounding DINO, LLaVA, Qwen

Vision & Content Understanding: Object detection, instance segmentation, object tracking, anomaly detection, open-vocabulary perception, image/video understanding

Learning Methods: Few-/zero-shot, multimodal, self-supervised, prompt learning, distillation, PEFT

Programming & Tools: Python, C/C++, PyTorch, TensorFlow, Hugging Face, NumPy, OpenCV, Git

RESEARCH EXPERIENCE

Hyundai-NTU-A*STAR Corporate Lab, Nanyang Technological University

Singapore

Research Fellow

Nov 2025 – Present

Foundation Model Adaptation for Few- & Zero-Shot Anomaly Detection [P1, P9]

- Developed *DriftAD*, a visually guided **CLIP** adaptation framework that adapts text representations to local visual context for few-shot anomaly detection. Achieved state-of-the-art results across 1-, 2-, and 4-shot settings, with up to **1.5-point** gains in 1-shot AUROC and PRO. Published as first author at **ACM MM 2026** [P1].
- Extended the work to zero-shot anomaly detection with *TRACE*, combining **CLIP** semantic evidence with **DINOv3** structural features for cross-model anomaly detection without target-domain examples [P9].

Open-Vocabulary Detection & Segmentation for Robotic Grasping [Patent]

- Built an open-vocabulary perception and grasping pipeline that takes text or reference-image queries and predicts detections, instance masks, and grasp poses for object categories unseen during training.
- Designed the pipeline for clutter, occlusion, reflective surfaces, and packaging interference. Validated it across **77 categories** of real industrial parts and delivered an on-site robotic grasping demo to the industry partner.

School of Electrical and Electronic Engineering, Nanyang Technological University

Singapore

Research Associate

Jan 2021 – Oct 2025

Image & Video Restoration for Degraded Content & Watermark Removal [P2, P5, P11]

- Developed a two-stage framework for bitstream-corrupted JPEG restoration, reaching **38.92 dB PSNR** on 2K images (+**5.52 dB** over EPDN). Published as first author at **CVPR 2023** [P2]. Extended this research through a collaborative **NeurIPS 2023** study on corrupted-video recovery and benchmark development [P11].
- Developed *SSH-Net*, a self-supervised model for watermark removal without paired clean targets. Outperformed PSLNet on unseen noise and watermark settings and published the work in **JVCIR 2025** [P5].

Prompt-Based Lightweight Image Super-Resolution [P3, P10]

- Designed *PromptSR*, a **0.78M-parameter** lightweight super-resolution model using cascade prompting and cross-scale features for efficient image enhancement.
- Improved lightweight super-resolution baselines by up to **0.48 dB**. Extended the work to adaptive superpixel token aggregation in *AsTaSR* [P3, P10].

Multimodal Byte-Visual Fusion for File-Fragment Classification [P4, P6]

- Developed multimodal classifiers for raw file fragments without file extensions, metadata, or valid decoding by combining byte-sequence and visual representations.
- Built *Byte2Image* to encode bit-level patterns as images and *ByteNet* to fuse byte-sequence and visual features. Improved accuracy by up to **12.2%** across two benchmarks and **14 evaluation settings** [P4, P6].

EARLIER RESEARCH EXPERIENCE

College of Computer Science, Chongqing University

Graduate Researcher

Chongqing, China

Sep 2017 – Jul 2020

Efficient Video Processing with Parallelization & DVFS [P7]

- Parallelized H.264 decoding at the task level and applied dynamic voltage and frequency scaling. Improved decoding speed by **27–32%** while reducing energy consumption by **23–25%** across 720p–2160p video [P7].

Energy-Efficient Machine Learning Acceleration [P13, P8]

- Developed *HolyLight*, a nanophotonic accelerator for data-center CNN workloads, achieving **13× higher throughput per watt** than a ReRAM baseline with negligible accuracy loss [P13].
- Co-first-authored *MindReading*, extending the photonic architecture to real-time EEG-based intention recognition. Reduced power by **62.7%** and achieved **1.68×** the throughput per watt of HolyLight-A [P8].

RESEARCH & TECHNICAL LEADERSHIP

- Contributed to the technical development of three research proposals totaling approximately S\$10 million, covering problem definition, methodology, and evaluation planning.
- Developed and validated computer vision methods and system prototypes with Hyundai and Schaeffler for real industrial use cases, including algorithm implementation, system integration, and live demonstrations to partner teams.
- Mentored 20+ final-year and M.Sc. students in computer vision research, guiding implementation, experimental design, evaluation, result analysis, and technical writing.

SELECTED PUBLICATIONS

Selected First- and Co-First-Author Publications

- P1. **W. Liu**, D. Zhang, T. Liu, K. Wu, A. W.-K. Kong. “DriftAD: Visually-Guided Text Drift for Few-Shot Industrial Anomaly Detection.” Accepted at *ACM International Conference on Multimedia (ACM MM)*, 2026. [CCF-A](#) · [Multimodal Content Understanding](#)
- P2. **W. Liu**, Y. Wang, K.-H. Yap, L.-P. Chau. “Bitstream-Corrupted JPEG Images Are Restorable: Two-Stage Compensation and Alignment Framework for Image Restoration.” *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 2023. [CCF-A](#) · [Visual Enhancement & Generation](#)
- P3. **W. Liu**, C. Cai, J. Gao, K. Wu, Y. Wang, K.-H. Yap, L.-P. Chau. “PromptSR: Cascade Prompting for Lightweight Image Super-Resolution.” *IEEE Transactions on Multimedia (TMM)*, 2026. [CCF-A](#) · [JCR Q1](#) · [Visual Enhancement & Generation](#)
- P4. **W. Liu**, K. Wu, T. Liu, Y. Wang, K.-H. Yap, L.-P. Chau. “ByteNet: Rethinking Multimedia File Fragment Classification through Visual Perspectives.” *IEEE Transactions on Multimedia (TMM)*, 2024. [CCF-A](#) · [JCR Q1](#) · [Multimodal Content Understanding](#)
- P5. **W. Liu**, J. Gao, K.-H. Yap. “SSH-Net: A Self-Supervised and Hybrid Network for Noisy Image Watermark Removal.” *Journal of Visual Communication and Image Representation (JVCIR)*, 2025. [CCF-C](#) · [JCR Q2](#) · [Visual Enhancement & Generation](#)
- P6. **W. Liu**, Y. Wang, K. Wu, K.-H. Yap, L.-P. Chau. “A Byte Sequence Is Worth an Image: CNN for File Fragment Classification Using Bit Shift and n-Gram Embeddings.” *IEEE International Conference on Artificial Intelligence Circuits and Systems (AICAS)*, 2023. [Multimodal Content Understanding](#)
- P7. **Wenyang Liu**, Weichen Liu, M. Li, P. Chen, L. Yang, C. Xiao, Y. Ye. “Fine-Grained Task-Level Parallel and Low-Power H.264 Decoding in Multi-Core Systems.” *IEEE International Conference on Parallel and Distributed Systems (ICPADS)*, 2018. [CCF-C](#) · [Efficient AI Systems](#)
- P8. Q. Lou, **Wenyang Liu** (co-first author), Weichen Liu, F. Guo, L. Jiang. “MindReading: An Ultra-Low-Power Photonic Accelerator for EEG-Based Human Intention Recognition.” *Asia and South Pacific Design Automation Conference (ASP-DAC)*, 2020. [CCF-C](#) · [Efficient AI Systems](#)

Submitted Manuscripts

- P9. **W. Liu**, Y. Zhang, T. Liu, K. Wu, A. W.-K. Kong. “TRACE: Text-Residual Anchoring for Cross-Model Evidence Decoding in Zero-Shot Anomaly Detection.” Submitted to AAAI (CCF-A), 2027. [Multimodal Content Understanding](#)
- P10. **W. Liu**, T. Liu, C. Cai, J. Gao, K. Wu, K.-H. Yap, C.-W. Lin. “AsTaSR: Adaptive Superpixel Token Aggregation for Lightweight Image Super-Resolution.” Submitted to *IEEE TMM* (CCF-A, JCR Q1), 2026. [Visual Enhancement & Generation](#)

Selected Collaborative Publications

- P11. T. Liu, K. Wu, Y. Wang, **W. Liu**, K.-H. Yap, L.-P. Chau. “Bitstream-Corrupted Video Recovery: A Novel Benchmark Dataset and Method.” *Advances in Neural Information Processing Systems (NeurIPS)*, 2023. [CCF-A](#) · [Visual Enhancement & Generation](#)
- P12. C. Cai, T. Liu, J. Gao, **W. Liu**, K. Wu, R. Wang, Y. Wang, S. C. Liew. “From Semantics, Scene to Instance-awareness: Distilling Foundation Model for Open-vocabulary Grounded Situation Recognition.” *ACM International Conference on Multimedia (ACM MM)*, 2025. [CCF-A](#) · [Multimodal Content Understanding](#)
- P13. Weichen Liu, **Wenyang Liu**, Y. Ye, Q. Lou, Y. Xie, L. Jiang. “HolyLight: A Nanophotonic Accelerator for Deep Learning in Data Centers.” *Design, Automation & Test in Europe Conference (DATE)*, 2019. [CCF-B](#) · [Efficient AI Systems](#)
- P14. R. Wang, S. Wang, R. Gong, C. Cai, J. Gao, W. Wang, **W. Liu**, K.-H. Yap. “MotionAnimate: Animate Human Images with Pose Motion for Vivid and Temporally Consistent Video Generation.” *Information Fusion*, 2026. [JCR Q1](#) · [Visual Enhancement & Generation](#)
- P15. J. Gao, C. Cai, R. Wang, **W. Liu**, K.-H. Yap, K. Garg, B. S. Han. “CL-HOI: Cross-Level Human–Object Interaction Distillation from Multimodal Large Language Models.” *Knowledge-Based Systems*, 2025. [JCR Q1](#) · [Multimodal Content Understanding](#)

RECOGNITION & PATENTS

Patent: “Grasp Pose Estimation for Wrapped Parts in Unstructured Industrial Environments” (2026)

Chunhui Cup Innovation and Entrepreneurship Competition, Winner (2023)

Academic Scholarships: A-Class (2017–2020), C-Class (2016)