Minxing Luo

mxluocv@gmail.com — +86 159 9695 8078 — Github — Google Scholar

EDUCATION

Nankai University, Tianjin, China — Sep 2023 — Jun 2026
Master of Science in Computer Science
Advisor: Prof. Jian Yang (Link), Visual Computing and Intellectual Perception Lab, College of Computer Science, Nankai University

China University of Mining and Technology, Xuzhou, China — Sep 2019 — Jun 2023
Bachelor of Science: Information Security

PUBLICATIONS

Luo, M., Fan, L., Wang, Q., Wu, G., Luo, Y., Yu, Y., Chen, J., Wang, Y., Fan, Q., & Yang, J. (2026). *Restore text first, enhance image later: Two-stage scene text image super-resolution with glyph structure guidance*. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. (Link)

- **Developed TIGER** (**T**ext–**I**mage **G**uided sup**E**r-**R**esolution), a novel two-stage framework that resolves the trade-off between image quality and text readability by decoupling glyph restoration from image enhancement.
- **Constructed** the UZ-ST dataset—the first Chinese scene text dataset with extreme zoom—and **achieved** state-of-the-art performance in both readability and visual fidelity.

Luo, M., Xia, Z., Chen, L., Li, Z., Zeng, W., Wang, J., Cheng, W., Wang, Y., Zhou, Y., & Yang, J. (2025). *Beyond flat text: Dual self-inherited guidance for visual text generation.* In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops* (pp. 1937-1946). (Link)

- **Introduced** STGen, a training-free framework that eliminates text distortion and background mismatches in complex scenarios (e.g., slanted or curved text).
- **Demonstrated** that integrating semantic rectification with glyph-based structural cues significantly improves text accuracy and visual harmony.

Cheng, W., & **Luo, M.** (2024). MVP: One-shot object pose estimation by matching with visible points. *IEEE Signal Processing Letters*, 31, 2760–2764. (Link)

- **Introduced** a novel, detector-free one-shot object pose estimation method that leverages an explicit image-point relational graph from SfM to prune erroneous and rarely visible 3D points.
- **Integrated** this lightweight graph-based pruning into stacked feature transformers to secure early, high-quality 2D-3D correspondences, outperforming state-of-the-art methods in both accuracy and speed.

Luo, M., Cheng, W., & Yang, J. (2024). *RotationDrag: Point-based image editing with rotated diffusion features* [Preprint]. arXiv. (Link)

- **Developed** RotationDrag, a novel point-based image editing approach that overcomes feature instability during in-plane rotation by tracking handle points via rotated feature maps.
- **Introduced** RotateBench, the first benchmark dedicated to evaluating in-plane rotation, and demonstrated superior optimization and high image fidelity against existing diffusion-based baselines.

ACADEMIC EXPERIENCE

Ming Team, Ant Group — Hangzhou, China
Research Intern — Jan 2026 – Present

- Exploring unified generation-understanding models and diffusion conditioning mechanisms.
- Working on diffusion architecture methodology, with one related work being prepared for a top-tier venue (2nd author).

Imaging Algorithm Center, VIVO — Hangzhou, China
Research Intern — Advisor: Dr. Qingnan Fan (Link) — Mar 2025 – Dec 2025

- **Lead developer** of a diffusion-based image super-resolution framework specializing in **text region enhancement**, addressing critical limitations in existing methods that neglect textual fidelity – a key factor in human visual perception.
- Designed and implemented a **text-aware image super resolution pipeline (*Accepted by CVPR 2026*)** by:
  - Curating high-quality *real-world degraded* text image pairs to overcome dataset limitations;

  - Integrating textual prior into diffusion sampling.
- Designed and implemented an **automated image pairing system** based on the ViVO X200U, addressing the scarcity of high-quality low/high-resolution image pairs in real scenarios.
- Constructed a **real-world dataset** to provide robust data foundations for model training and evaluation.

Visual Computing and Intellectual Perception Lab, Nankai University Tianjin, China
Master Student Jun 2023 — Present

- Supervised by Prof. Jian Yang (Link), Dr. Wentao Cheng (Link), and Dr. Yaxing Wang (Link) during my master's program.
- Initially focused on Visual Localization in Structure from Motion, I later developed a strong interest in generative models and shifted my research focus accordingly.
- Conducted research in Visual Text Generation with Diffusion Models, resulting in three published/submitted papers.
- Served as a teaching assistant for the Digital Logic course, assisting in lab sessions and grading assignments.

China University of Mining and Technology ACM/TINA Team Xuzhou, China
Team Member Jun 2020 — Jun 2021

- Completed rigorous training in foundational programming and classical algorithms with a score of 100 in *Program Practice*.
- Represented the team in national ACM competitions, earning an Honorable Mention at the 2020 ICPC Asia Nanjing Regional Contest and securing Second Prize in the C/C++ University Group A at the 12th Lanqiao Cup (Jiangsu Division).

AWARDS

China National Scholarship in both 2019-2020 and 2020-2021 academic years Xuzhou, China
A highly competitive award granted to the **top 0.2%** of students nationwide for exceptional academic achievements and research potential. Nov 2020 and Dec 2021

Interdisciplinary Contest in Modeling (ICM) Honorable Mention Xuzhou, China
Global mathematical modeling competition organized by COMAP, awarded to top-performing teams for excellence in interdisciplinary analysis and problem-solving. May 2022

Second Prize of the 10th Asia-Pacific University Student Mathematical Modeling Competition Xuzhou, China
Developed and implemented innovative mathematical models to solve complex real-world problems within a 4-day timeframe, demonstrating strong analytical and problem-solving abilities. March 2021

SKILLS

- **Language:** English C1 (**CET-6: 617**, **PTE-A: 88**)
- **Programming:** C++, Python, LaTeX, PHP, SQL, XML/XSL
- **Databases:** MySQL, SQLite
- **Web Dev:** HTML, CSS, JavaScript, Apache Web Server.
- **Soft Skills:** Passionate about swimming, which cultivates resilience, discipline, and focus.

OTHER EXPERIENCES

College debate team leader Xuzhou, China
leader Jun 2020 — Dec 2020

- Teach freshmen basic techniques and rules for debate.
- Organize practice and contact other teams for contests.
- Prepare for University-wise contest.