

Shuai Wang

Ph.D. Student, Department of Computer Science and Technology, Nanjing University

wangshuai.mcg@gmail.nju.edu.cn · GitHub: WANGSSSSSS

Profile

Optimistic, responsible, and self-motivated researcher with a strong interest in new ideas and a clear understanding of personal strengths and limitations. My research interests include scalable training and fast sampling for generative models, especially diffusion models, as well as visual perception and representation learning.

Education

Nanjing University , Department of Computer Science and Technology Direct-track Ph.D. Student in Computer Science and Technology	2022.09 - Present
Northwestern Polytechnical University , School of Computer Science B.Eng. in Computer Science and Technology	2018.09 - 2022.06

Awards and Honors

- 2018-2019 **National Scholarship**
- 2019 FIRA RoboWorld Cup RoboSot Race, Development Group, **Runner-up**
- 2019-2020 National College Student Mathematics Competition, Non-Mathematics Group, Shaanxi Division, **First Prize**
- 2020 ICRA RoboMaster AI Challenge, **International Third Prize**
- 2022-2023 First-Class Academic Scholarship
- 2022-2023 **Tencent Scholarship** (Master's level)
- 2024-2025 **National Scholarship** (Ph.D. level)

Publications

Deep Equilibrium Object Detection ICCV 2023

Shuai Wang, Yao Teng, Limin Wang

- Identified parameter redundancy and insufficient performance saturation in query-based object detectors.
- Converted the FFN detector decoder into an RNN-style formulation to reduce redundancy, and further explored the upper performance bound with a deep equilibrium model.

Deep Equilibrium Object Detection and Segmentation TPAMI

Shuai Wang, Yao Teng, Limin Wang

- Journal extension of DEQDet, adding instance segmentation applications and more extensive detection and segmentation experiments.

FlowDCN: Exploring DCN-like Architectures for Fast Image Generation with Arbitrary Resolution NeurIPS 2024

Shuai Wang, Zexian Li, Tianhui Song, Xubin Li, Tiezheng Ge, Bo Zheng, Limin Wang

- Proposed Groupwise-MultiScale-DCN as a basic building block to explore DCN-style architectures for image generation.
- Compared with DiT/SiT, reduced parameters by 8%, reduced FLOPs by 20%, and required only one-fifth of the total training images.
- Achieved FID 2.00 and state-of-the-art sFID 4.35 on ImageNet 256 x 256.
- Achieved FID 2.44 and state-of-the-art sFID 4.53 on ImageNet 512 x 512.

Differentiable Solver Search for Fast Diffusion Sampling ICML 2025

Shuai Wang, Zexian Li, Tianhui Song, Xubin Li, Tiezheng Ge, Bo Zheng, Limin Wang

- Searched for generalizable solver parameters with a very small number of trainable parameters.
- For flow-matching models, improved 10-step FID to 2.40 for SiT and 2.35 for FlowDCN, significantly outperforming existing solvers.
- For DDPM models, improved 10-step FID to 2.33 for DiT, significantly outperforming DPM-Solver++ and UniPC.

DDT: Decoupled Diffusion Transformer ICLR 2026

Shuai Wang, Zhi Tian, Weilin Huang, Limin Wang

- Decoupled high-frequency decoding and low-frequency encoding in generative model pre-training.
- Achieved a leading FID of 1.26 on ImageNet 256 x 256, outperforming other methods under fair comparison.
- Achieved a leading FID of 1.28 on ImageNet 512 x 512, substantially outperforming competing methods.

PixNerd: Pixel Neural Field Diffusion CVPR 2026

Shuai Wang, Ziteng Gao, Chenhui Zhu, Weilin Huang, Limin Wang

- Used neural fields to assist large-patch pixel decoding.
- The class-to-image model achieved FID 1.78 on ImageNet 256 x 256; under the same compute budget, it approached latent-model performance and substantially outperformed other pixel-space generative models.
- The text-to-image model achieved GenEval 0.73 and DPG 80.9, approaching several advanced text-to-image models and clearly outperforming pixel-space generative baselines.

UniDDT: Unifying Multimodal Understanding and Generation with Decoupled Diffusion Transformer ECCV 2026

Shuai Wang, et al.

- Unified multimodal understanding and generation with a decoupled diffusion transformer framework.
- Validated the scalability of DDT-style architectures across different visual representation spaces.

Project and Research Experience

NPU Soccer Robot Base - FireFly 2019.10 - 2021.01

- Responsible for the real-time object detection module in the RoboMaster AI Challenge.
- Responsible for full-field localization and object detection based on sentry cameras in the RoboMaster AI Challenge.

NPU 3D Vision Laboratory 2020.09 - 2021.01

- Participated in the Rocket Force Artificial Intelligence Challenge; achieved the highest accuracy and twice the speed of the second-place team.
- Contributed to a research project on stereo disparity estimation based on RAFT-style optical-flow methods.

SenseTime 2021.12 - 2022.05

- Participated in autonomous driving perception tasks for SenseAuto.

Alibaba Alimama 2024.04 - 2024.12

- Worked on model acceleration for intelligent creative generation, collaborating with full-time engineers on open-source models such as Flux-Turbo, which exceeded 10,000 downloads.

ByteDance 2024.12 - 2026.12

- Contributed to Seedream 5.0 Pro, focusing on large-scale rendered data and text editing.
- Seedream 5.0 Pro is a multimodal image generation model designed for professional content production, supporting high-fidelity visual synthesis, instruction-following image editing, and reasoning-driven creation.