Yushen Zuo

Tel: (+86) 13630049521 · Email: zuoyushen12@gmail.com · Page: Personal Page

Education

Xidian University, Electronic Engineering, *Bachelor* Aug 2015 - Jun 2019

- GPA: 3.90 / 4.0 (Top 5%), Outstanding Graduates

Tsinghua University, Department of Automation, *Master* Sept 2019 - Jun 2022

- GPA: 3.58 / 4.0
- Research interests: low-level vision; image translation and generation; object detection and segmentation; vision-language model safety

Research Experience

Low Resolution Palmprint Image Denoising - Neurocomputing 2019 Jan 2019 - Jun 2019

- Design a generative adversarial network (GAN)-based model to address multiple types of noise in palmprint image and reserve more orientation information with Gabor loss in training.
- Achieve state-of-the-art performance in both image denoising and palmprint recognition in test dataset.

Visual Token Transformer for Image Restoration May 2020 - Jun 2021

- Design transformer block based on visual token to extract the non-local/multi-scale self-similarity of image.
- Reduce computation cost of Transformer from $O(n^2)$ to $O(n)$ with comparable image restoration performance.
- Included in the paper of NTIRE 2021 Challenge on Image Deblurring in **CVPR 2021**. (In Top 10 methods)

Multi-View Consistent Style Transfer with One-Step Diffusion Jun 2024 - Aug 2024

- Focus on the stylization of multi-view images in 3D scenes and proposed OSDiffST, a novel style transfer method based on one-step diffusion model.
- Incorporate LoRA adapters to rapidly adapt the pre-trained diffusion model for style transfer. Propose a vision condition module for efficient style information extraction and injection.
- Research paper is accepted by the AI for Visual Arts Workshop and Challenges (AI4VA) in **ECCV 2024**.

Safeguarding Vision-Language Models from Gaussian Noise - ICCV 2025 Submission Jul 2024 - Jan 2025

- The first to provide a systematic vulnerability analysis, revealing that mainstream VLMs lack inherent robustness to Gaussian noise visual perturbations.
- Propose Robust-VLGuard, a dataset with novel image-text misalignment cases and Gaussian noise augmentation to improve VLM robustness without sacrificing helpfulness.
- Extend the defense scope to visual adversarial attacks and propose DiffPure-VLM, a diffusion-based defense framework that converts adversarial noise into Gaussian-like noise, enabling effective defense via Gaussian noise-augmented fine-tuned VLMs.

4KAgent - Research Intern Project in TACO group, TAMU Jan 2025 - Now

- Leveraging agentic systems to address complex image restoration tasks and upscale images to 4K resolution.
- Design multi-agent system for image analysis and restoration execution. Propose Q-MoE policy for better quality.
- Design profile module for users to customize the system to meet diverse restoration requirements.

Test Time Scaling in Advanced Text-to-Image Framework Feb 2025 - Now

- Focus on the path search Test Time Scaling in ODE / Flow matching based Text-to-Image frameworks. (e.g., Flux, Stable Diffusion 3 / 3.5)
- Collaborator: Hunyuan-DiT team, Tencent.

Intern Experience

Youtu Lab, Tencent, Research intern Oct 2020 - May 2021

- **UniInst: Detection free and NMS free instance segmentation - CN114332457A [P]**
  - Instance-aware One-to-one Assignment: Use Hungarian matching to assign the best matching feature point to the target as positive point according to the classification score and segmentation mask accuracy.

- MaskIOU Branch: During training, learn to predict the IOU of the generated Mask. During inference, multiply it's IOU prediction for generated masks with the classification score as the final confidence.
- Achieve state-of-the-art mask AP on COCO test-dev 2017 dataset and OCHuman dataset.

Microsoft Research Asia, Research intern Jul 2021 - Jul 2022

- **Rotated object detection (multi-directional table detection in PDF image)**
  - Design an anchor-free two-stage detector for rotated object detection.
  - Design sequence-invariant loss and relative-offset for rotated object detector training.
  - Stable performance under different image rotation angles in production dataset (F-score fluctuation ≤ 0.02).
  - Achieve state-of-the-art performance in production dataset and contribute to Azure OCR API.
  - 'Stars-of-tomorrow' award of Microsoft Research Asia Internship Program.

Work Experience

Microsoft, Applied Scientist in Bing Aug 2022 - Mar 2024

- **Bing News - Recommendation system**
  - Dynamic quota allocation
    - Train a classification model to determine whether a recommendation request is triggered by user.
    - Adjust the quota of each recall path in Ranker based on classification result to reduce computational cost.
    - Product performance: Reduce $\sim$20% computing resources usage without losing performance.
- **Bing Whole Page - Large Language Model Application**
  - Answer triggering in Bing Search - Real Estate Vertical
    - Use GPT-3.5 to label challenging samples from web search results and obtain 1.3M new training samples.
    - Train answer triggering model based on new training set augmented with samples by LLM labeling.
    - Product performance: 4.1K gain in DAU (Daily Active Users) of Bing real estate vertical.

The Hong Kong Polytechnic University (PolyU), Research Assistant Apr 2024 - Now

- **Artificial Intelligence and Signal Processing Laboratory**
  - Accelerated Diffusion in Image Processing Task (e.g., Style Transfer (AI4VA@ECCV2024))
  - Image and Video Super-Resolution
    - **2nd place** in the AIM 2024 Challenge on Efficient Video Super-Resolution for AV1 Compressed Content in **ECCV 2024**.
    - **1st place** in NTIRE 2025 Challenge on Short-form UGC Image Super-Resolution ($4\times$) in **CVPR 2025**.
  - Novel View Synthesis under sparse view with 3D Gaussian Splatting
    - Enhance 3D Gaussian Splatting under sparse view based on local depth and semantic regularization.
    - Our research paper is accepted by **ICASSP 2025**.

Publications

- Shengjie Chen, Shuo Chen, Zhenhua Guo, **Yushen Zuo**. "Low-resolution palmprint image denoising by generative adversarial networks", Neurocomputing, 2019, 358: 275-284.
- Seungjun Nah, Sanghyun Son, Suyoung Lee, Radu Timofte, Kyoung Mu Lee, **Yushen Zuo** et al. "NTIRE 2021 Challenge on Image Deblurring", 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). IEEE, 2021: 149-165.
- **Yushen Zuo**, Jun Xiao, Kin-Chung Chan, Rongkang Dong, Cuixin Yang, Zongqi He, Hao Xie, Kin-Man Lam. "Towards Multi-View Consistent Style Transfer with One-Step Diffusion via Vision Conditioning", ECCV 2024 Workshop.
- Zongqi He, Zhe Xiao, Kin-Chung Chan, **Yushen Zuo**, Jun Xiao, Kin-Man Lam. "See In Detail: Enhancing Sparse-view 3D Gaussian Splatting with Local Depth and Semantic Regularization", ICASSP 2025.
- Jiawei Wang*, Yushen Zuo*, Yuanjun Chai, Zhendong Liu, Yicheng Fu, Yichun Feng, Kin-Man Lam. "Safeguarding Vision-Language Models: Mitigating Vulnerabilities to Gaussian Noise in Perturbation-based Attacks", ICCV 2025 Submission.

Honors & Awards

First-class scholarship, outstanding student in 2016, 2017, 2018

The first prize (Shaanxi Division) of the National College Student Mathematics Competition	Aug 2017
Meritorious winner in Interdisciplinary Contest in Modeling (ICM)	May 2018
Outstanding Graduates	Jun 2019
Champion of the 1st Ocean Target Detection International Challenge (1 / 151)	Dec 2020
Kaggle NFL 1st and Future - Impact Detection, Silver medal (23 / 459)	Jan 2021
CVPR 2021 NTIRE Image Deblurring Challenge - Track1. Low Resolution (10 / 60)	Mar 2021
'Stars-of-tomorrow' award of Microsoft Research Asia Internship Program	May 2022
AIM 2024 Challenge on Efficient Video Super-Resolution for AV1 Compressed Content - 2nd place	Aug 2024
CVPR 2025 NTIRE Challenge on Short-form UGC Image Super-Resolution (4×) - 1st place	Mar 2025

Skills

Programming	Python (PyTorch, NumPy, Scikit learn. etc.), C/C++, HTML/CSS, SQL.
Miscellaneous	Linux, Shell (Bash/Zsh), LATEX(Overleaf/Markdown), Microsoft Office, Git.