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

- Palmprint recognition methods are sensitive to image noise and need an effective denoising algorithm.
- First attempt at end-to-end denoising of low-resolution palmprint images by neural networks.
- Design a generative adversarial network (GAN)-based model to address multiple types of noise in palmprint image and reserve more orientation information with using Gabor loss function in training.
- Collect Data from PolyU palmprint database and IITD database to build train/test dataset and generate noisy image by adding different types of noise.
- Outperforms existing state-of-the-art methods in both image denoising quality and palmprint recognition accuracy in test dataset with different types of noise. Average EER (equal error rate) of palmprint recognition decreased from 10.841% to 1.532% after denoising.

Visual Token Transformer for Image Restoration May 2020 - Jun 2021

- First attempt to use visual token-based transformer in image restoration.
- Neural network learn to divide images into different groups and map them to visual tokens without manual rules.
- Design transformer block based on visual token to extract the non-local/multi-scale self-similarity of image.
- Token-based transformer reduces computation cost from $O(n^2)$ to $O(n)$ compared to vanilla transformer with comparable image restoration performance.
- Included in 'NTIRE 2021 Challenge on Image Deblurring' (**CVPR Workshop 2021**). (In Top 10 methods)
- Project report (Applied in various low-level vision tasks): Visual Token Transformer for Image Restoration.pdf.

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 propose OSDiffST, a novel style transfer method based on a 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.
- Use two additional loss functions to align color distribution and improve structural similarity for enhancing visual quality and maintaining multi-view consistency across images from different viewpoints.
- Research paper is accepted by the AI for Visual Arts Workshop and Challenges in ECCV 2024.

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 outperforms CondInst by 8.4 mask AP on OCHuman dataset (occluded and crowded scenario).

Microsoft Research Asia, Research intern Jul 2021 - Jul 2022

- **Rotated object detection (multi-directional table detection)**
  - 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 Intern Program.

Work Experience

The Hong Kong Polytechnic University (PolyU), Research Assistant Apr 2022 - Now

- **Artificial Intelligence and Signal Processing Laboratory**
  - Accelerated Diffusion in Image Processing Task (e.g., Style Transfer, Image Translation)
  - Efficient Video Super-Resolution
  - Old movie restoration and enhancement
  - Novel-view synthesis with 3D Gaussian Splatting
  - Image Processing and Diffusion in vision-language model safety and defense

Microsoft, Applied Scientist in Bing Aug 2022 - Mar 2024

- **Bing News - Recommendation system**
  - Explainable AI
    - Use SHAP to calculate feature contribution to ranking score for a better explanation of model's output.
    - Show users why he/she sees this recommended content based on the recall path with a mapping method.
    - Currently applied to all Bing News channels (e.g. Edge homepage), while collecting user's feedback to modify the mapping method.
  - Dynamic quota allocation
    - Train a classification model to determine whether a news recommendation request is triggered by user or by prerender/other backend tasks. (Result: AUC > 0.8 in test dataset built on Bing News Recommendation database)
    - Based on the classification result, reduce the quota of each recall path in Ranker for requests predicted to be 'Not User-trigger' 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 Related
    - Use LLM (GPT-3.5) to label challenging samples from web search results and obtain 1.3M new training samples for answer triggering model.
    - Recall in test dataset improved from 0.54 to 0.73 after training with new training set with LLM labeling.
    - Product performance: 3% increase in answer trigger rate (answer triggers Bing real estate application) in Bing search, and 4.1K gain in DAU (Daily Active Users) of Bing real estate application.

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.
- Marcos V. Conde, Zhijun Lei, Wen Li, Christos Bampis, Ioannis Katsavounidis, Radu Timofte, **Yushen Zuo** et al. "AIM 2024 Challenge on Efficient Video Super-Resolution for AV1 Compressed Content", ECCV 2024 Workshop.

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
AIM 2024 Challenge on Efficient Video Super-Resolution for AV1 Compressed Content - 2nd place	Aug 2024