

Zinuo Li

Personal Email: 1273180880lzn@gmail.com
Student Email: zlii0362@student.monash.edu
Phone: +86-138-2427-6641
[Research Gate](#)
[Google Scholar](#)
[Personal Homepage](#)



PERSONAL INFO

I am a highly motivated individual with a strong passion for academic pursuits. Due to some personal and familial factors, my initial GPA did not reflect this level of dedication in the first two years. However, starting from the third year, my GPA has been maintained at a high level. I got the highest score in the professional course of my school in the Embedded Systems course (99.00). At the same time, I was selected by exceptional research institutions to serve as a research assistant and have published high-quality papers.

EDUCATION

- | | |
|---|-------------------|
| University of Macau , Research Assistant | 2021.09 - Current |
| • Focusing on Computer Vision. | |
| Chinese Academy of Sciences , Research Assistant | 2022.10 - Current |
| • Shenzhen Institute of Advanced Technology (SIAT), Medical Imaging. | |
| Huizhou University , Computer Science, <i>B.Eng</i> | 2019.09 - 2023.06 |
| • GPA 83.34/100, with average GPA 92.95/100 in last two years (top 3%). Graduation thesis 90.61 (top 1%). | |
| • Teaching Assistant of Linux System in Huizhou University. | |
| • English Level: IELTS band 7.0. | |

RESEARCHES

- **Current Research:** Low-level Computer Vision.
- **Interested Research:** Computer Vision, Multimodal, Medical Imaging, etc.
- **Preparing Researches:** One research about **real large-scale film style simulation** and one research about **visual prompt tuning** are being conducted, targeting to **CVPR**.

SELECTED PUBLICATIONS (*INDICATES EQUAL CONTRIBUTION)

High-resolution Document Shadow Removal via A Large-scale Real-world Dataset and A Frequency-aware Shadow Erasing Net

*Zinuo Li**, Xuhan Chen*, Chi-Man Pun and Xiaodong Cun
ICCV23 accepted [C], Core A*

- We acquire over 7k couples of high-resolution (2462×3699) images of real-world documents with various samples under different lighting circumstances, which is 10 times larger than existing datasets.
- We decouple the high-resolution images in the frequency domain, where we can learn the low-frequency details and high-frequency boundary individually via the carefully designed network structure.
- The proposed method shows a clearly better performance than previous methods in terms of visual quality and numerical results.
- For dataset and code, the website and repo are being constructed, please refer to [Github](#) once it is available.

A Large-scale Film Style Dataset for Learning Multi-frequency Driven Film Enhancement

*Zinuo Li**, Xuhan Chen*, Chi-Man Pun and Shuqiang Wang
IJCAI23 published [C], Core A*

- We are the first to construct a large-scale high-quality dataset with 3 groups of different film style and a total of 5,285 high-quality images, called FilmSet.
- To learn the features in FilmSet properly, we present FilmNet, a novel multi-frequency framework based on Laplacian Pyramid for simulating film styles and subsequently retouching normal photos.
- We demonstrate our model is superior to the SOTA methods via extensive experiments on our dataset and other publicly accessible benchmark datasets.
- For dataset and code please refer to [Github](#).

UWFormer: Underwater Image Enhancement via a Semi-Supervised Multi-Scale Transformer

Zinuo Li, Xuhan Chen, Shenghong Luo, Chi-Man Pun and Shuqiang Wang
AAAI24 in submission [C], Core A*

- We introduce a new loss function Subaqueous Perceptual Loss via a Mean-Teacher based Semi-supervised Transformer network for enhancing images accross multiple frequencies.
- We propose a Nonlinear Frequency-aware Attention mechanism and a Multi-scale Fusion Feed-forward network in Transformer blocks, which have been demonstrated to gain better performance.

- Experiments using full-reference and non-reference underwater benchmarks demonstrate that our method outperforms SOTA.

MedPrompt: Cross-Modal Prompting for Multi-Task Medical Image Translation

Xuhang Chen*, Zinuo Li*, Pingping Xu, Huisi Wu, Chi-Man Pun and Shuqiang Wang
AAAI24 in submission [C], Core A*

- We present MedPrompt, a multi-task framework that efficiently translates different modalities.
- Specifically, we propose the Self-adaptive Prompt Block, which dynamically guides the translation network towards distinct modalities. Within this framework, we introduce the Prompt Extraction Block and the Prompt Fusion Block to efficiently encode the cross-modal prompt.
- Extensive experimental results involving five datasets and four pairs of modalities demonstrate that our proposed model achieves state-of-the-art visual quality and exhibits excellent generalization capability using only one-time training process.

WaveEnhancer: Unifying Wavelet and Transformer for Image Enhancement

Zinuo Li*, Xuhang Chen*, Chi-Man Pun and Shuqiang Wang
CGI23 accepted [C], Core B

- We present WaveEnhancer, a novel wavelet domain transformer-based enhancement framework for enhancing normal photos
- We demonstrate our model is superior to the state-of-the-art methods via extensive experiments on our dataset and other publicly accessible benchmark datasets.

An Improved Matting-SfM Algorithm for 3D Reconstruction of Self-Rotating Objects

Zinuo Li*, Zhen Zhang*, Shenghong Luo, et al.
Mathematics published [J], JCR-Q1

- When an object is in self-rotating like the Earth, the traditional SfM algorithm cannot reconstruct it well, or even collapse. We reveal the reason why conventional SfM cannot reconstruct self-rotating objects.
- We propose a new algorithm called Matting-SfM. It was proven that Matting-SfM algorithm possessed more accurate results and reached SOTA, which solved the problem that the self-rotating objects could not be reconstructed.

SKILLS AND EXPERIENCES

Experiences and Awards:

- **National First Prize of China College Student Computer Designing Competition:** As the leader, led the team to win the first prize at the national level in the artificial intelligence track (4.5% of 6516 projects).
- **Two National University Student Innovation Projects:** One as the leader focusing on 3D reconstruction and one as a member concentrating on geometeorological research.
- **Project leader of Guangdong Climbing Innovation Program:** Concentrating on the 3D reconstruction of the foot arch and the determination of the health level.

Skills: Strong English skills, able to read various AI papers and understand their principles, as well as strong writing skills.

Coding: Python, Pytorch, OpenCV, Latex, etc.