

# Zeqiang Lai

Email: laizeqiang@outlook.com

GitHub: zeqiang-lai.github.io

---

## RESEARCH INTERESTS

Generative Model, Image Processing, Multi-Modal, Optimization

## EDUCATION

**Beijing Institute of Technology** Beijing, China  
Master in Computer Science and Technology Sep 2020 – July 2023  
Mentors: Professors Ying, Fu.

**Beijing Institute of Technology** Beijing, China  
Bachelor in Computer Science and Technology Sep 2016 – July 2020

## PUBLICATIONS PREPRINTS

**$\nabla$ -Prox: Differentiable Proximal Algorithm Modeling for Large-Scale Optimization**

Zeqiang Lai\*, Kaixuan Wei\*, Ying Fu, Philipp Härtel, and Felix Heide.  
[SIGGRAPH'23] *ACM Transactions on Graphics*, 2023.

**Hybrid Spectral Denoising Transformer with Guided Attention**

Zeqiang Lai, Chenggang Yan, and Ying Fu.  
[ICCV'23] *International Conference on Computer Vision*, 2023.

**Hyperspectral Image Super-Resolution with Real Unaligned RGB Guidance**

Zeqiang Lai, Ying Fu, and Jun Zhang.  
[TNNLS] *IEEE Transactions on Neural Networks and Learning Systems* (Minor Revision).

**Reference-based Hyperspectral Super-Resolution with Spatial Spectral Matching**

Yingkai Zhang, Zeqiang Lai, and Ying Fu.  
[IJCV] *International Journal of Computer Vision* (Under Review).

**Deep Plug-and-Play Prior for Hyperspectral Image Restoration**

Zeqiang Lai, Kaixuan Wei, and Ying Fu.  
[NEUCOM'21] *Elsevier Neurocomputing*, 2021.

**Denoising Diffusion Semantic Segmentation with Mask Prior Modeling**

Zeqiang Lai, Yuchen duan, Jifeng Dai, Ziheng Li, Ying Fu, Hongsheng Li, Yu Qiao, and Wenhai Wang.  
[arXiv'23] *arXiv preprint*, 2023.

**Mixed Attention Network for Hyperspectral Image Denoising**

Zeqiang Lai, and Ying Fu.  
[arXiv'23] *arXiv preprint*, 2023.

**InternGPT: Solving Vision-Centric Tasks by Interacting with ChatGPT Beyond Language**

Zhaoyang Liu\*, Yanan He\*, Wenhai Wang\*, Weiyun Wang\*, Yi Wang\*, Shoufa Chen\*, Qinglong Zhang\*, Zeqiang Lai\*, Yang Yang, Qingyun Li, Jiashuo Yu, Kunchang Li, Zhe Chen, Xue Yang, Xizhou Zhu, Yali Wang, Limin Wang, Ping Luo, Jifeng Dai, and Yu Qiao  
*[arXiv'23] arXiv preprint, 2023.*

## RESEARCH EXPERIENCE

### **Proximal Algorithm Modeling** [SIGGRAPH'23] [NEUCOM'21]

Mentor: Felix Heide (Princeton University) Mar 2022 – Present  
– We introduce  $\nabla$ -Prox as a domain-specific language and compiler for differentiable proximal algorithms that solve large-scale optimization problems.  
– With only a few lines of code,  $\nabla$ -Prox generates performant solvers for diverse problems from computational optics to integrated energy system planning – each with state-of-the-art performance by bi-level optimization with algorithm unrolling, equilibrium learning, and reinforcement learning.

### **Generative Mask Prior Modeling with Diffusion Model** [arXiv'23]

Mentor: Wenhai Wang (Shanghai AI laboratory) May 2022 – Present  
– We investigate the generated mask prior modeling for image segmentation with discrete diffusion models and introduce two novel designs – noise on first prediction and free re-noising, for the training and inference processes.  
– We demonstrate that our mask prior modeling can be integrated with state-of-the-art segmentors for boosting performance in terms of quantitative measurements and visual coherence.

### **Large Language Model for Controlling Tools** [arXiv'23]

Mentor: Wenhai Wang (Shanghai AI laboratory) May 2023 – Present  
– We introduce InternGPT, a system that utilizes LLM as controller to bind tools across vision-language domains, with a new pointing instruction type.  
– We propose InternGPT v2 with a novel LLM controller that enables adding new tools without any retraining and explorations of novel tool calling path, and upgraded architectural server-client designs.

### **Hyperspectral Image Denoising** [ICCV'23] [arXiv'23]

Mentor: Ying Fu (Beijing Institute of Technology) Oct 2020 – Present  
– We design two novel image denoising models, in particular, a hybrid 3D transformer. Exploiting domain knowledges, we demonstrate it outperforms existing state-of-the-art by over 1.5 dB on PSNR and with 12% parameters, 42% running time and 20x fast training convergence.

### **Dual Camera Hyperspectral Imaging System** [TNNLS] [IJCV]

Mentor: Ying Fu (Beijing Institute of Technology) Aug 2021 – Present  
– We build the first dual camera hyperspectral imaging system with unaligned RGB image as guidance. Together with the first real-world paired unaligned dataset, we propose two post-processing super-resolution models that help achieving better image quality than previous single camera system.

## INDUSTRY EXPERIENCE

**Shanghai AI Laboratory**, OpenGVLab Shanghai, China  
Research Intern May 2022 – July 2023  
– Cutting-edge algorithm exploration of diffusion models on perception tasks.

- InternImage: Regular maintenance; Integrated DeepSpeed for low-cost training of huge model; Optimized DCNv3 CUDA operator, etc.
- InternGPT: Regular maintenance; Integrated tools of Stable Diffusion inpainting, Imagebind generation, and DragGAN editing.
- Multi-Modal data enhancement with large language model.
- Exploration of large language model for tools.

**Kuaishou Technology, Y-Tech** Beijing, China  
 Algorithm Intern July 2019 – Mar 2020  
 - Built a virtual anchor by integrating the systems of prosody prediction, sound synthesis, mouth prediction, etc.  
 - Developed two learning-based models for mouth animation prediction from the spectrum of audio, and prosody prediction.

**OPNESOURCE PROJECTS**

**DragGAN (4.8k Stars)** Creator  
 Unofficial implementation of "Drag Your GAN: Interactive Point-based Manipulation on the Generative Image Manifold" at SIGGRAPH 2023

**InternGPT (2.8k Stars)** Major Contributor  
 InternGPT (iGPT) is an open source demo platform where you can easily showcase your AI models.

**InternImage (1.8k Stars)** Contributor  
 Official Implmentaiton of "InternImage: Exploring Large-Scale Vision Foundation Models with Deformable Convolutions" (CVPR 2023 Highlight)

**Anything to Image (147 Stars)** Creator  
 Integration fo ImageBind and Stable Diffusion to achieve anything (audio, text, image) to image generation.

**TFPnP (81 Stars)** Major Contributor  
 Official implementation of "Tuning-free Plug-and-Play Proximal Algorithm for Inverse Imaging Problems" (ICML 2020 Award Paper & JMLR 2022)

**Yan Programming Language** Creator  
 An educational programming language compiler and interpreter, along with an educational framework and evaluation website.

**HONORS AND SCHOLARSHIPS**

National Scholarship (Ministry of Education in China)	2022
Academic Scholarship (Beijing Institute of Technology)	2020,2022
First Prized of Mathematical Contest in Modeling	2018
Scholarship for Outstanding Students	2017, 2018

**ACADEMIC SERVICE**

**Conference and Journal Reviewer**  
 TIP, TGRS, TNNLS  
 ICCV 2021, PBDL 2021, ACM MM 2021, CVPR 2022, CVPR 2023, ICCV 2023

**SKILLS**

**Programming:** Python, C++, Java, HTML, Javascript, Matlab, LLVM  
**Languages:** Chinese (native), English (fluent)

# 赖泽强

18811031367 / laizeqiang@outlook.com / Zeqiang-Lai.github.io

## 教育经历

---

### 北京理工大学

- 硕士学位 | 计算机科学与技术专业 2020.09 – 2023.07
- 学士学位 | 计算机科学与技术专业 2016.09 – 2020.07

## 学术文章

---

- [1] Differentiable Proximal Algorithm Modeling for Large-Scale Optimization SIGGRAPH TOG 2023  
[Zeqiang Lai\\*](#), Kaixuan Wei\*, Ying Fu, Philipp Hartel, Felix Heide
- [2] Hybrid Spectral Denoising Transformer with Guided Attention ICCV 2023  
[Zeqiang Lai](#), Chenggang Yan, Ying Fu
- [3] Hyperspectral Image Super-Resolution with Real Unaligned RGB Guidance TNNLS (Minor)  
[Zeqiang Lai](#), Ying Fu, Jun Zhang
- [4] Reference-based Hyperspectral Super-Resolution with Spatial-Spectral Matching IJCV (Under Review)  
Yingkai Zhang, [Zeqiang Lai](#), Ying Fu
- [5] Deep Plug-and-Play Prior for Hyperspectral Image Restoration Neurocomputing 2022  
[Zeqiang Lai](#), Kaixuan Wei, Ying Fu
- [6] Denoising Diffusion Semantic Segmentation with Mask Prior Modeling Preprint 2023  
[Zeqiang Lai\\*](#), Yuchen Duan\*, Jifeng Dai, Ziheng Li, Ying Fu, Hongsheng Li, Yu Qiao, Wenhai Wang
- [7] Mixed Attention Network for Hyperspectral Image Denoising Preprint 2023  
[Zeqiang Lai](#), Ying Fu
- [8] InternGPT: Solving Vision-Centric Tasks by Interacting with ChatGPT Beyond Language Preprint 2023  
Zhaoyang Liu\*, Yinan He\*, Wenhai Wang\*, Weiyun Wang\*, Yi Wang\*, Shoufa Chen\*, Qinglong Zhang\*, [Zeqiang Lai\\*](#), and more

## 研究经历

---

### 近端优化算法建模 [1][5]

- [1] 提出了  $\nabla$ -Prox, 一个用于可微近端算法建模的领域建模语言和编译器。仅需要数行代码,  $\nabla$ -Prox 即可智能地针对指定优化问题生成集成了迭代展开, 隐微分, 强化学习的可微分求解器。特别地,  $\nabla$ -Prox 支撑了在计算光学任务上一个全新设计空间的探索发现, 并且图像恢复, 医学图像, 线性规划等领域均取得了显著性能。
- [5] 提出了一种即插即用 ADMM 高光谱图像恢复算法。在图像去噪、超分、压缩感知等多个任务实现了最先进的性能。

### 基于扩散模型的生成式分割图先验建模 [6]

- 探索了基于扩散模型的生成式分割图先验建模。基于一个预训练的判别式语义分割方法, 我们提出了两种新的设计用于训练一个生成式的扩散模型对分割图进行约束, 使其在形状, 边缘, 以及类别关系上更加符合分割图先验分布。
- 该方法能够灵活地集成到任意已有的分割方法用于提升指标性能 (mIoU 提升最高可达 3.5) 以及视觉效果。

### 基于大语言模型的工具控制器 [8]

- [4] 构建了 InternGPT, 一个利用 LLM 作为控制器, 调用视觉-语言等各个领域工具的系统, 并引入了一种新的基于鼠标点击的 Instruction 交互类型, 这种交互的重要性在后续火热的 DragGAN 得到进一步体现。

- 正在探索 InternGPT v2, 一种新颖的 LLM 控制器, 可以在无需重新训练的情况下添加新工具, 并能发现现有工作无法找到的工具调用路径, 同时在架构上升级为更为灵活的前后端分离设计。

### 高光谱图像去噪 [2][7]

- [4] 利用光谱图像领域知识设计了一个基于可分离卷积以及谱间注意力的混合 Transformer 网络。实现了仅以 12% 的参数量, 42% 的运行时间实现了高达 1.5 dB 的 PSNR 提升。除此之外, 该模型还实现了 20 倍的收敛速度提升。
- [5] 探索了基于 RNN 的 Transformer Attention 设计, 用于解决自注意力的显存占用问题。

### 高光谱双相机成像系统 [3][4]

- [6] 构建了第一个以不对齐的 RGB 图像作为引导的双相机高光谱成像系统。收集了第一个不对齐真实配对数据集。提出了一个基于光流对齐, 注意力融合的后处理超分辨率模型, 并实现相比单摄像头系统显著更好的图像质量。
- [3] 进一步提出了一种基于全局特征匹配, 可变形卷积, 交叉注意力机制的特征融合方法, 进一步提升了后处理算法性能。

## 实习经历

---

上海人工智能实验室 | 见习研究员 @ OpenGVLab 2022.05 – 2023.07

- 去噪扩散模型在高层视觉任务的研究。
- InternImage 相关维护, 包括 Deepspeed 集成, DCNv3 CUDA 算子优化等。
- 基于语言模型的多模态数据增强研究。
- 基于大语言模型的工具调用控制器研究。

快手科技 | 算法实习生 @ Y-Tech 2019.06 – 2020.02

- 整合韵律分析, 声音合成, 口型预测等模型, 构建了一个虚拟主播系统。
- 开发了两种深度学习模型用于韵律预测以及基于语音的口型预测。

## 项目经历

---

**DragGAN** (4.8k stars) | 一天复现 DragGAN Creator

**InternGPT** (2.8k stars) | 一个集成超多工具的多模态问答系统, 使用 LLM 作为控制器。 Major Contributor

**InternImage** (1.8 stars) | 刷了超多榜的开源卷积视觉大模型 Contributor

**Anything to Image** (147 stars) | 基于 ImageBind 和 Stable Diffusion, 从任意模态到图像的生成工具 Creator

**TFPnP** (81 stars) | ICML 2020 Award Paper 的官方实现的搭建与开源 Joint Creator

**Yan** | 一个用于教学的编程语言, 包含编译器, 解释器, 评测服务器, 评测网站等 Creator

## 荣誉奖项

---

国家奖学金 2022

北京理工大学学业奖学金 2022, 2020

全国大学生数学建模北京赛区一等奖 2018

北京理工大学学业奖学金 2017, 2018