



陈胤达

☎ (+86)13058626611

✉ cyd0806@mail.ustc.edu.cn

🌐 <https://ydchen0806.github.io/>

🌐 <https://scholar.google.com/citations?user=hCv1j5cAAAAJ&hl=en&oi=ao>



## 🎓 教育背景

中国科学技术大学 & 上海 AI Lab 📍 合肥 & 上海

信息与通信工程, 博士研究生

2024.7 ~ 2027.6 (预计)

- 研究方向: 机器学习理论, 自监督预训练, 多模态大模型, 图像编码、压缩
- 导师: 吴枫, 熊志伟, 汤晓鸥
- 核心课程: 算法设计与分析, 统计学习, 深度学习, 强化学习
- 荣誉奖励: 国家自然科学基金博士生项目负责人 (2024 年)

中国科学技术大学 📍 合肥

计算机技术, 硕士

2022.9 ~ 2024.7

- 荣誉奖励: 研究生国家奖学金 (2022 年)

厦门大学 📍 厦门

遥感 & 经济学双学位, 本科

2018.9 ~ 2022.7

- 综合排名: 1/31
- 荣誉奖励: 厦门大学学术之星 (2021 年), CDA 一级认证 (2022 年), Kaggle Expert
- 导师: 张原野

## 🔬 科研项目

国家自然科学基金博士生项目

2025.1 ~ 2027.12

国自然项目 负责人 (经费 30 万, 2024 年安徽省信息口唯一入选)

[1] 神经科学大模型的构建与训练

- 面向脑神经科学研究的百亿级基础大模型开发
- 融合多模态神经影像数据的预训练方法设计

大规模自监督预训练

2022.5 ~ 2023.12

Conference & Journal 一作 & 共同一作

[1] **[LONG ORAL]** Self-supervised neuron segmentation with multi-agent reinforcement learning, IJCAI **[CCF-A]**

23

- 基于强化学习方法改进 MAE 掩码策略, 自动选取掩码率和掩码方案。
- 核心技术: 多智能体强化学习框架、自适应掩码策略优化、自监督特征学习。
- 引入多智能体框架实现更高效的自监督特征学习, 分割精度提升 12%。

[2] MaskTwins: Dual-form Complementary Masking for Domain-Adaptive Image Segmentation, ICML 25 **[CCF-A]**

- 从稀疏信号重构角度提出互补掩码新理论, 严格证明对偶形式互补掩码在提取领域不变特征上的理论优势。
- 核心技术: 对偶形式互补掩码理论、互补掩码一致性学习、无监督域适应框架。
- 在自然图像分割上提升 2.7% mIoU, 生物图像分割提升 3.2% IoU。

[3] TokenUnify: Scalable Autoregressive Visual Pre-training with Mixture Token Prediction, ICCV 25 **[CCF-A]**

- 提出图像自回归预训练方式与 Mamba 框架相结合, 体现长序列和低计算量的优势。
- 核心技术: 自回归视觉预训练、Mamba 架构集成、混合 token 预测策略。
- 展示良好的 scaling law, 并给出相应的理论证明, 模型性能与参数呈对数线性关系。

[4] EMPOWER: Evolutionary Medical Prompt Optimization With Reinforcement Learning, JBHI **[SCI 一区]**

- 提出首个针对医疗领域的进化 prompt 优化框架, 结合领域知识和强化学习。
- 核心技术: 医学术语注意力机制、多维度质量评估 (clarity/specificity/relevance/accuracy)、保结构进化算法、语义验证模块确保临床指南一致性。
- 事实错误降低 24.7%, 领域特异性提升 19.6%, 在盲测中获得 15.3% 更高的临床医生偏好。

[5] Learning multiscale consistency for self-supervised electron microscopy instance segmentation, ICASSP **[CCF-B]**

24

- 基于多尺度特征对比学习和特征重构, 实现高性能预训练策略。
- 核心技术: 多尺度特征对比学习、特征一致性损失函数、自监督表征学习。
- 提出特征一致性损失函数, 克服尺度变化带来的表征差异, 准确率提升 9%。

[6] **[在投]** Generative Text-Guided 3D Vision-Language Pretraining for Unified Medical Image Segmentation, Submit to PR

- 基于大语言模型生成图像描述, 进行多模态图文对比学习预训练。
- 核心技术: 文本引导的 3D 视觉预训练、vision-language 对比学习、生成式文本引导机制。
- 创新性地引入生成式文本引导机制, 解决医学图像标注稀缺问题, 实现零样本分割。

**[1] MaskFactory: Towards High-quality Synthetic Data Generation For Dichotomous Image Segmentation, [CCF-A] NeurIPS 24**

- 通过刚性和非刚性形变编辑掩码，再利用 ControlNet 生成对应的 mask-image pair。
- 核心技术：形变驱动的掩码编辑、ControlNet 引导的图像生成、合成数据质量评估。
- 合成数据在下游分割任务中表现接近真实数据，仅有 2% 性能差距，大幅降低标注成本。

**[2] Vid-TTA: Low-Cost Test-Time Adaptation for Robust Video Editing, WACV 26 [CCF-C]**

- 首次将测试时适应 (TTA) 应用于视频编辑，在推理过程中动态微调 UNet backbone。
- 核心技术：motion-aware frame reconstruction 识别关键运动区域、prompt perturbation 增强模型鲁棒性、meta-learning 驱动的动态 loss 平衡机制。
- 显著改善视频时序一致性并缓解 prompt overfitting，计算开销低，可作为现有模型的即插即用增强。

**[3] [在投] Joint Semantic and Coded Generation for Conditional Latent Coding, Submit to TCSVT**

- 提出联合语义和编码生成框架，结合文本描述与极少量编码特征 (0.01 bpp) 引导可控整流流模型生成高保真参考图像。
- 核心技术：ControlNet for DiT 架构、迭代参考对齐 (IRA)、动态特征融合 (DFS)、迭代概率精炼 (IPR)、理论鲁棒性证明。
- 相比 VVC 压缩性能提升 15.8% BD-rate，成功统一图像生成与压缩，填补语义表示与编码表示间的关键空白。

## 深度学习图像编码与压缩

2023.8 ~ 至今

**[1] [ORAL] Conditional Latent Coding with Learnable Synthesized Reference for Deep Image Compression, [CCF-A] AAAI 25**

- 构建图像相似度字典检索相似图像，用于改进熵模型的概率估计。
- 核心技术：条件隐变量编码、可学习的合成参考框架、自适应熵建模、Transformer 编解码架构。
- 提出可学习的合成参考框架，在同等比特率下 PSNR 提升 0.6dB，压缩性能领先。

**[2] BIMCV-R: A Landmark Dataset for 3D CT Text-Image Retrieval, MICCAI 24 [CCF-B]**

- 构建首个开源的 3D CT 图文对数据集，包含 10,000 对高质量医学图像与描述。
- 核心技术：3D CT 图文对齐、多模态检索框架、关键词索引系统。
- 实现高效的图文信息检索和关键词搜索，召回率较传统方法提升 25%。

**[3] [在投] Learned Image Coding with Generative Reference of Conditional Latents, Submit to TPAMI**

- AAAI 25 oral 文章的拓展工作，进一步探究参考图像对图像编码的助益。
- 核心技术：生成式参考图像合成、多源参考融合 (本地字典 + 网络检索 + 图像生成)、条件隐变量生成框架、鲁棒性理论分析。
- 通过三种方式获取参考图像，压缩性能提升 15%，提出条件隐变量生成框架，有效解决参考图像不可用情况下的性能退化问题，并给出扰动鲁棒性理论证明。

**[4] [在投] UniCompress: Enhancing Multi-Data Medical Image Compression with Knowledge Distillation, Submit to TCSVT**

- 通过多模态知识先验实现隐式神经网络压缩多个数据，压缩率提升 40%。
- 核心技术：知识蒸馏、多模态融合、隐式神经表示、医学影像共性特征提取。
- 基于知识蒸馏提取多种医学影像共性特征，减少 20% 存储空间同时保持诊断质量。

## 大模型工程

2023.9 ~ 至今

**[1] 图像编码，帧内预测大模型**

- 主导设计 10 亿参数级别编码架构，比传统编码标准提升 30% 压缩率。

**[2] 医学图像分割，神经元分割大模型**

- 主要负责团队中的预训练部分，具有 64 卡 A40 大规模集群预训练经验。
- 掌握 DDP、DeepSpeed 等大模型框架，实现 300 亿参数模型的高效训练与优化。

## 🧑 实习经历

## 腾讯 (IEG) 📍 上海

## 青云计划实习生

2025.8 ~ 至今

- 专注游戏视频场景理解技术研发，构建智能视频内容分析系统。
- 基于多模态大模型设计游戏场景自动识别与分类算法，准确率达到 95% 以上。
- 开发游戏视频关键帧提取和场景转换检测模型，为游戏内容创作提供技术支持。

## 中国人民解放军总医院 (301 医院) 📍 北京

## 数据压缩小组，研究实习生

2023.9 ~ 2024.2

- 协同戴琼海院士团队进行高效数据压缩的研究。
- 设计了医学影像特定的压缩算法，针对 CT、MRI 等模态优化，压缩效率提升 35%。

## 帝国理工学院 📍 伦敦 (remote)

## Data Science Institute，研究实习生

2022.11 ~ 2023.8

- 协同 Rossella Arcucci 副教授进行多模态预训练的研究，并投稿期刊论文一篇。
- 开发了图像-文本对比学习框架，在医学诊断任务上准确率达到 93.5%。

- 负责数据挖掘类课程的设计与讨论, 主讲聚类和 Transformer 小节。
- 指导 20 名本科生完成机器学习项目, 组织 2 次校内竞赛活动。

- 协助朱炯副教授完成国土经济统计, 进行宅基地信息的视觉特征提取。
- 开发卫星影像分析工具, 自动识别土地利用变化, 准确率达 85%。

🏆 荣誉奖项

• 国家自然科学基金博士生项目	2024.12
- 负责人, 安徽省信息类唯一	
• 研究生国家奖学金	2022.12
- 获奖率 1%	
• 厦门大学学术之星	2021.12
- 本科生唯一获奖者	
• “景润杯”数学竞赛专业组	2021.09
- 厦门大学第一名	
• “互联网+”大赛	2021.08
- 福建省金奖	
• 全国大学生数学竞赛非专业组决赛	2021.05
- 全国二等奖	
• “挑战杯”全国大学生课外学术科技作品竞赛	2021.05
- 福建省一等奖	
• 全国大学生数学竞赛非专业组	2020.11
- 福建省第一名	

🔧 专业技能

- 编程能力: Python, MATLAB,  $\LaTeX$ , C, C++, Java
- 深度学习: PyTorch, TensorFlow, DeepSpeed, DDP
- 英语能力: TOEFL(110), GRE(328)
- 专业工具: Git, Docker, CUDA, HPC

👨‍🔬 学术服务

- 期刊 & 会议审稿人: CVPR 2025, NeurIPS 2024, ICML 2025, ICLR 2024, MICCAI 2025, ACM MM 2024, AISTATS 2024, IJCV, TIP

📰 最新动态

• 一篇论文被 ICCV 2025 接收	2025.06.26
• 一篇论文被 ACL 2025 findings 接收	2025.05.15
• 一篇论文被 ICML 2025 接收	2025.05.01
• 一篇论文被 AAAI 2025 选为口头报告	2025.01.18
• 两篇论文被 AAAI 2025 接收	2024.12.11
• 成功入选博士生自然科学基金项目负责人	2024.12.06
• 一篇论文被 NeurIPS 2024 接收	2024.10.10
• 两篇论文被 MICCAI 2024 接收	2024.06.18