

朱惠兴 个人简历

☎ 189-3350-4063 ✉ 22530052@zju.edu.cn 📄 GitHub 📄 Personal Page

🎓 教育背景

浙江大学 电子信息（硕士） 光电科学与工程学院 2025.9 – 2028.6
浙江大学 光电信息科学与工程（本科） 光电科学与工程学院 2021.9 – 2025.6
荣誉与奖学金：浙江省政府奖学金、浙江大学优秀毕业生、浙江大学校奖学金、大恒光电创新创业奖学金等。

🔬 科研与项目经历

注意力引导的因果稀疏视频理解框架 2025.11 – 2026.6

项目描述：针对视觉语言模型（VLM）处理长视频或高分辨率图像时 token 激增导致的算力瓶颈，受人类视觉系统选择性注意机制启发，设计了一套因果在线的稀疏视频采集-感知框架，在高 token 压缩比下保持视频理解精度，并结合真实稀疏采集系统进一步降低感知开销。

- **Coarse-to-Fine 两阶段决策模块**：按视觉编码器 tubelet size 将视频划分为时间步，每步内对首帧局部降采样，经 patch embedding 后由轻量决策模块融合运动特征估计各空间 patch 的语义重要性并生成采集掩膜。
- **注意力分数分布对齐**：提取视觉编码器特定 transformer 层中各 token 的注意力分数，训练决策模块的输出分布与之对齐，使其具备高层视觉语义的判别能力。
- **稀疏视频适应微调**：视频动作识别任务上，基于 VideoMAE 预训练模型在稀疏视频上微调，token 丢弃 70% 时在 SSv2 测试集上达完整视频 97% 精度；VideoQA 任务上，对 Qwen2.5-VL 3B 在 LLaVA-Video-178K 数据集上做 LoRA 微调，token 丢弃 80% 时在 VideoMME 上维持全分辨率 98% 以上精度。
- **Baseline 对比与系统收益**：动作识别任务上超越或持平 ToMe、RLT 等方法，VideoQA 任务上接近或持平 VisionZip、VidCom² 等 post-encoder 压缩方法。对 Qwen2.5-VL 3B 推理实现 $1.54\times$ TTFT 加速，节省 41.7% 显存；实际采集时可降低 50%–70% 带宽开销。
- **相关成果**：会议论文一篇（在投），发明专利《一种新型因果稀疏感知系统》（在审）。

大语言模型构建与对齐：From Scratch to RLHF 2026.4 – 2026.6

项目描述：基于 Stanford CS336 课程，从零搭建大语言模型，涵盖 BPE 分词器训练、Transformer 架构搭建与预训练、监督微调及强化学习对齐，涉及 Qwen2.5-Math-1.5B 与 Llama-3.1-8B 两个模型家族。

- **BPE 分词器与预训练**：在 TinyStories（10K 词表）与 OpenWebText（32K 词表）上训练 Byte-level BPE 分词器，并从零搭建含 RMSNorm、RoPE、Causal Multi-Head Self-Attention、SwiGLU FFN 等组件的 Pre-Norm Transformer Decoder；在 TinyStories 数据集上以 4 层 Transformer block 和 512 维 d-model 配置训练，验证集 loss 降至 1.45 以下。
- **监督微调（SFT）**：基于 GPT-OSS-120B 推理轨迹数据对 Qwen2.5-Math-1.5B 进行 SFT，MATH 基准准确率由 2.6%（zero-shot）提升至 47.7%，观察到去掉训练集中错误的推理轨迹能提升精度。对 Llama-3.1-8B 在 safety-augmented UltraChat 200K 单轮指令数据集上进行指令微调，AlpacaEval 长度控制胜率达 31.8%（以 Llama-3.3-70B-Instruct 评判，对比 GPT-4-turbo 参考回答，base 模型仅 2.1%）。
- **强化学习对齐（RLHF）**：实现 GRPO 算法对 Qwen2.5-Math-1.5B 进行推理能力优化，完成学习率、baseline 消融、clip 机制与 off-policy 策略的对比实验，最优配置在 MATH 上达 32.5%；实现 DPO 算法对 Llama-3.1-8B 进行偏好对齐，基于 HH-RLHF 数据集训练，安全性评分由 base 模型的 4% 提升至 36%（经 Llama-3.3-70B 评判），其中 SFT 阶段贡献主要增益（33%），DPO 进一步提至 36%。

基于事件相机的轻量级眼动追踪算法设计 2024.10 – 2025.5

项目描述：针对传统帧相机在眼动追踪中的固有限制，基于事件相机设计完整瞳孔提取与角膜反射光斑定位算法，实现自适应实时眼动追踪（与舜宇集团合作项目）。

- **瞳孔检测**：将事件流预处理为事件帧，基于 U-Net 搭建轻量瞳孔分割网络，引入通道与空间注意力机制提升掩膜精度，拟合掩膜估计瞳孔中心坐标。

- **角膜反射光斑检测**：基于 CenterNet 目标检测框架，对比 ResNet 与 MobileNetV2 等骨干网络的速度与精度，提取各光斑序号与中心坐标。
- **帧相机插帧验证**：使用 TimeLens 对同步灰度视频插帧以对齐事件流的高时间分辨率，对照帧图像与事件帧的提取坐标计算误差，验证事件模态精度。
- **实际效果**：自建数据集上瞳孔中心误差约 3–4 像素，光斑中心误差约 1–2 像素；TensorRT 部署后单帧推理低于 5ms，满足实时眼动追踪需求。
- **相关成果**：发明专利《一种基于事件相机的眼动追踪方法》。

☰ 获奖情况

- 第二十三届全国大学生机器人大赛 RoboMaster 机甲大师超级对抗赛（RMUC）一等奖（全国殿军，雷达站技术组组长）2024.5 – 2024.8
- 第十四届全国大学生数学竞赛省一等奖2022.8
- 浙江省大学生物理创新（理论）竞赛一等奖2022.5
- CET-4 625 CET-6 6302021–2022