

杨梓瀚

13631036545 | crazyyoung@foxmail.com

教育背景

北京航空航天大学 (985, 211) | 硕士 | 电子信息 (人工智能方向) (gpa:3.61/4.0) 2024.09 - 2027.06

修读课程: 人工智能、计算机视觉、算法设计与分析、互联网络技术、数理统计、矩阵理论与应用

北京邮电大学 (211, 双一流) | 本科 | 计算机科学与技术 2019.09 - 2024.06

主修: 数据结构与算法、计算机网络、操作系统、数据库; 辅修: 机器学习、Python 程序设计、c++程序设计

科研经历

ACM MM 2026 (CCF A, 第一作者)

Towards Accurate Small Object Detection via Dynamic Frequency Suppression with Label Disambiguation

针对小目标视觉信息不足、频域噪声干扰及标签歧义导致的定位不准确问题, 提出 DyFrFPN 动态频率抑制网络与 LDM 标签消歧模块, 通过自适应过滤频域冗余与噪声、抑制歧义样本提升检测精度。

ECCV 2026 (CV 三大顶会, 第二作者)

Adaptive Spectrum-Aware Feature Disentangled Network for Small Object Detection

提出 ASD 全频谱特征解耦框架与 CPD 类别原型蒸馏策略, 实现全频谱噪声抑制与类内一致性增强。

实习经历

多模态算法实习生 | 短剧多模态理解与场景级语义检索

patch-X 科技 2026.04 - 2026.07

- 在一部 60 集短剧上搭建多模态内容理解链路: 结合镜头变化与均匀覆盖抽取代表帧, 对齐字幕时间轴, 使用 YuNet/SFace 完成人脸检测与剧内人物聚类, 并通过带时间戳的 Contact Sheet 驱动商用 VLM 生成事件、角色关系及镜头语义等结构化标签。
- 优化 VLM 视觉证据组织方式, 在相同 60 集数据及 Schema 校验口径下, 将模型请求数由 304 降至 68、图像输入数由 1,223 降至 68; 同时将其中的 YuNet/SFace 人物分析模块复用与团队已有的 20 部短剧数据, 在无需人工角色标注的情况下生成剧内匿名人物簇与跨场景轨迹伪标签。
- 基于该链路生成的约 10K 场景-文本对构建 CLIP 双塔检索模型, 在视觉塔顶部 4 个 Transformer Block 注入 LoRA, 通过 Temporal Attention、时间对齐字幕融合与角色轨迹感知聚合学习场景表征; 轨迹编号仅用于剧内视觉证据分组, 不跨剧复用且不输入文本塔。
- 采用 InfoNCE、In-batch Negative 及基于匿名人物簇构造的“同人物、不同事件”Hard Negative; 按剧集划分训练与测试, 在约 10K 场景全库检索下, $R@1/R@5/R@10$ 达 27%/55%/69%, 相较 Frozen CLIP + Frame Mean Pooling, $R@10$ 由 44% 提升至 69%。

项目经历

全模态语义 ID 驱动的生成式广告召回系统 | 腾讯广告推荐个人项目

2026.01 - 2026.03

- 基于腾讯广告推荐公共数据构建生成式召回链路, 覆盖百万级用户行为序列与 66 万候选广告池; 用 decoder-only causal Transformer 建模用户曝光/点击序列, 融合广告稀疏属性、行为类型、时间间隔与多模态 embedding 输出用户兴趣向量。
- 基于广告多模态 embedding 训练 4-level RQ-KMeans semantic ID (codebook=1024), 与 item ID、广告主、类目 embedding 融合, 提升长尾广告语义共享能力 (长尾 HR@10 相对提升约 34%)。
- 采用 InfoNCE 对比学习对齐用户与广告向量, 结合 in-batch negatives 与大负样本库, 并围绕负样本规模、模型深宽、semantic ID 层数完成 scaling/ablation。
- 推理阶段离线预计算候选广告向量并用 Faiss HNSW 做毫秒级 Top-K 检索 (约 4.3ms); 最终 HR@10=0.103, NDCG@10=0.061, 较 item-ID-only Transformer 提升约 24%/19%。

技术技能

- 深度学习: 熟练使用 PyTorch, 具备模型训练、调试与优化实战经验; 熟悉 Transformer 等模型架构, 了解 PPO、DPO、GRPO 等强化学习对齐方法, 熟悉 SFT、RLHF 流程;
- 熟悉推荐系统召回、排序与重排全链路, 掌握协同过滤、矩阵分解、双塔模型、多路召回及 ANN 向量检索; 熟悉 Wide & Deep、DeepFM、DCN-v2、DIN/SIM 等排序模型与 MMoE/PLE、ESMM 多任务建模; 掌握负采样、特征交叉、样本回流与特征快照、AUC/GAUC 离线评估、冷启动与 MMR/DPP 多样性重排;
- 工程能力: 熟练使用 Python/C++, 熟悉 JAVA; 熟悉全栈开发全链路; 熟悉 Docker、Git、Linux 开发环境; 熟悉 vLLM 推理服务化部署与 SSE 流式接口。

荣誉奖项

国家发明专利 (已授权) 1 项, 发明专利 (审查中) 1 项; CCF-CSP 计算机软件能力认证 310 分 (全国累计排名前 2.84%); 全国大学生数学竞赛 三等奖; CET-6; 学业奖学金 2 项