

# 赖彦涛

电话: 13033081901(微信同号) | 邮箱: yantaolai@nuaa.edu.cn

个人主页: yantaolai.github.io

意向岗位: 算法工程师实习



## 教育经历

### 南京航空航天大学

2023 年 9 月-2026 年 4 月

计算机技术 硕士 计算机科学与技术学院 模式分析与机器智能 (工信部) 重点实验室

### 安徽大学

2019 年 9 月-2023 年 6 月

智能科学与技术专业 本科 互联网学院

绩点: 3.65/5, 排名 14/118

## 科研经历

### Pathformer3D: A 3D Scanpath Transformer for 360° Images (与导师共一)

ECCV 2024

• 该论文旨在研究全景图像中人类视线/注意力的转移情况。针对全景图像的扭曲问题,我们采用了一种为全景图像定制的三维卷积网络 SphereNet 来提取全景图像信息;然后将其通过 Transofrmer 编码器来上下文化图像信息;再将其和历史视点信息一起输入 Transformer 解码器中,并采用自回归的方式来生成新视点的隐藏状态;最后该隐藏状态通过我们所设计的一种新颖的三维混合密度网络,这样操作可以将视点的隐藏状态转化为相应的视点分布信息;最后我们通过采样的方式来生成新的视点。该模型在 4 个数据集以及 6 个评价指标上都达到了 SOTA 的水平。

### CLIPGaze: Zero-Shot Goal-Directed Scanpath Prediction Using CLIP (第一作者)

审稿中

• 该论文旨在研究目标导向的人类视线/注意力的转移情况。针对语义对齐问题,我们采用 CLIP 来提取目标的语义信息和多层次场景图像信息;然后我们采用一种新颖的层级式视觉-语义融合模块来对器进行融合;再将融合后的信息和视点的查询信息一起输入并行的 Transformer 的解码器 (类似 DETR) 中来视点的隐藏状态信息;最后采用了多个 MLP 来将其转化为视点的分布信息。我们还采用了图像分割作为代理任务来进行预训练,这显著提升了扫视路径预测效果。实验结果证明,我们的模型在零样本、传统目标存在、传统目标缺失等三种情况下都取得了 SOTA 的水平。

### Speech-directed Human Attention Prediction using Multimodal LLMs and Mamba

进行中

• 本文旨在研究扫视路径预测领域中的增量目标指代任务,即在一条完整的语句中每一个单词对人类注意力的引导情况。针对语义对齐问题,我们拟采用 llava/BLIP-3/PixelLLM 等多模态大语言模型来对图像信息和文本信息进行提取。我们也将尝试采用新颖的 Mamba 结构来对新视点生成问题进行建模。我们也将尝试改进 Pathformer3D 中的混合密度网络,提出一种参数更少,速度更快的新颖混合密度网络。

### 一种面向全景图像的人眼扫视轨迹预测方法 (导师一作, 本人二作) 专利号: 202410512599.0

专利已公开

### 一种目标引导情况下的扫视路径预测方法 (导师一作, 本人二作)

专利已提交

## 实习经历

### 合肥科大讯飞教育发展有限公司

2023 年 3 月-2023 年 6 月

数据资源开发工程师

合肥

• 实习内容: 参与讯飞星火认知大模型的 AI 数据标注, 数据生成等工作

## 项目经历:

### C++ Web 高性能服务器 (Tiny-WebServer)

- 项目描述: 实现一种基于 C++ 的轻量级 Web 服务器
- 相关技术: 线性池, 非阻塞 socket, epoll

### 基于计算机视觉的小麦苗情检测系统设计

- 项目描述: 利用常见的计算机视觉方法, 对小麦的高光谱图像进行苗情检测。
- 相关技术: 主成分分析 (PCA), 支持向量机 (SVM), 随机森林 (RF), 极端梯度提升 (XGBoost)

## 技术细节 & 荣誉

- 研究领域: 计算机视觉, 深度学习, 多模态学习, 少样本学习
- 主要课程: 计算机视觉, 机器学习, 模式识别, 图像处理与分析, 智能控制理论, 智能软件开发与测试技术, 数据库, 自然语言处理
- 编程语言 & 框架: Python, C++, C, Pytorch, Latex, Git, Linux Shell, OpenCV, Java(了解), Tensorflow(了解), Matlab(了解)
- 本硕期间荣誉: 同竣奖学金 (5000 元), 校一/二等奖学金, 校三好学生/优秀共青团员, 研究生一等奖学金 (硕士)
- 其他: CET-4/6