

盛胡丹筠

hudanyun.sheng@outlook.com | <https://www.linkedin.com/in/hudanyunsheng/>

教育背景

佛罗里达大学	美国, 佛罗里达
电子计算机工程硕士	毕业年份: 12/2019
硕士论文: Switchgrass Genotype Classification using Hyperspectral Imagery	
佛罗里达大学	美国, 佛罗里达
工业系统工程硕士	毕业年份: 12/2017
同济大学	中国, 上海
物理学光信息科学与技术学士	毕业年份: 06/2015

工作经历

德克萨斯大学西南医学中心 Quantitative Biomedical Research Center	美国, 德克萨斯州, 达拉斯
数据科学家	09/2021 至今

- 实现基于 PyTorch 框架的 MaskRCNN, 从 H&E 染色的癌症病理图像中实现细胞核检测、分割并分类
- 系统性优化 MaskRCNN 基础结构, 使其具备处理不同精度图片的能力
- 通过重新定义 MaskRCNN 的损失函数, 实现对特定细胞类别的选择性灵活训练适应数据特殊性的模型
- 研发基于 YOLO v5 的分层分类算法; 并赋予 YOLO v5 基础模型图像分割功能
- 完善并包装基于 Python 的 CyTOF 图片处理分析工具包

Donald Danforth 植物科学研究院	美国, 密苏里州, 圣路易斯
数据科学研究员	02/2020 - 09/2021

- 建立并完善包括预处理、处理、后期处理、统计分析及可视化等在内的自动化处理 RGB 图像、热成像图像、以及高光谱图像的算法
- 利用 MaskRCNN 预训练模型对植物叶子进行实例分割; 通过开发叶子追踪算法来帮助理解植物的生命周期
- 参与 PlantCV (一个基于 Python 的开源的用于植物表型研究的图像分析软件包) 的开发: 优化及添加新的图像分析、图像分割、目标检测、特征提取等功能及算法; 通过 GitHub 实现版本控制
- 为整个科研团队提供数据可视化以及统计结果分析

佛罗里达大学医疗中心 Precision and Intelligent Systems in Medicine Partnership Lab	美国, 佛罗里达州, 盖恩斯维尔
数据科学实习生	05/2019-08/2019

- 预处理原始数据, 包括数据清洗、异常检测、批次合并、可视化等, 并确定研究队列
- 从病人的主要生命体征数据中提取时间序列特征
- 文献调研, 并应用包括插值网络在内的算法来处理时序特征的不规律性
- 通过对比分析聚类结果的临床数据来对比时序数据聚类结果

学术研究经历

佛罗里达大学 机器学习及感知实验室(Machine Learning and Sensing Lab)	美国, 佛罗里达州, 盖恩斯维尔
研究助理	05/2017-12/2019

- 研发用于从植物微根管图片中自动检测植物根的机器学习算法
- 搭建并完善植物高光谱图片及热成像图片的自动处理架构
- 研发并应用基于纯正端元提取以及高光谱逆混合原理的自动检测算法, 从高光谱图像中自动检测目标植物像素
- 研发基于孪生神经网络的分类友好型降维算法, 实现相同植物品种不同基因型的分类

出版文章

Yu, Guohao, Alina Zare, Hudanyun Sheng, Roser Matamala, Joel Reyes-Cabrera, Felix B. Fritsch, and Thomas E. Juenger. "Root identification in minirhizotron imagery with multiple instance learning." Machine Vision and Applications 31, no. 6 (2020): 1-13.

专业技能

编程及相关工具: Python (NumPy, Pandas, Scikit-learn, Scipy, ggplot2, Seaborn, matplotlib, PyTorch, Tensorflow, keras, OpenCV), MATLAB, Git & Github (for version control and corporation), Jupyter Notebook, R, SQL, LaTeX
机器学习算法: 类聚算法, 回归分析, 决策树, 随机森林, 邻近算法, 贝叶斯, 主成分分析, 支持向量机, 奇异值分析, 时序数据分析, 神经网络, 图像分析, 计算机视觉