

# 李嘉昊 (Jarvis)

☎ +(86)15970658312 🗨 JarvisLee0423 ✉ [leejarvis1314@gmail.com](mailto:leejarvis1314@gmail.com) 🏠 [jarvislee0423.github.io](https://github.com/jarvislee0423)

📄 <https://github.com/JarvisLee0423>

AI 工程师 | 计算机视觉算法工程师

## 🎓 教育经历

香港城市大学 CityUHK	2024.09 - 至今
计算机科学博士研究生	导师: 汪建平
研究方向: 自动驾驶语义占用, 3D 场景重建, 深度估计, 双目匹配	
香港中文大学 CUHK	2021.09 - 2022.11
计算机科学理学硕士	导师: 黄健康
研究方向: 2D 目标检测, 图像字幕	
香港浸会大学 HKBU	2017.09 - 2021.07
计算机科学与技术理学学士	导师: 张慧, 李致远, 李树德
研究方向: 深度学习, 量子金融, 图神经网络	

## 📄 论文发表

- [1] Huang, C., Liu, Y., Li, J., Tian, H., Chen, H. (2023). Application of YOLOv5 for mask detection on IoT. *Applied and Computational Engineering*, 29, 1-11.
- [2] Lee, J., Huang, Z., Lin, L., Guo, Y., & Lee, R. (2023). Chaotic Bi-LSTM and attention HLCO predictor-based quantum price level fuzzy logic trading system. *Soft Computing*, 27(18), 13405-13419.

## 🔑 项目经历

第 17 届 F1Tenth 大奖赛 (CPS-IoT 2024)	2024.04 - 2024.05
FSM Speed 团队成员	香港
<ul style="list-style-type: none"><li>取得大奖赛冠军。</li><li>单圈计时赛中, 我们实现了基于地图的算法, 涉及 AMCL 定位模块, RaceLine 生成模块和 MPC 控制模块。</li><li>在 Head-to-Head 竞赛中, 对经典的 Follow-the-Gap 算法进行了优化。下一个时间戳的速度和转向角度将根据距离本车最远点和当前速度进行校准。</li><li>项目主页: <a href="https://xyunaaa.github.io/research/f1tenth/">https://xyunaaa.github.io/research/f1tenth/</a>。</li></ul>	
用于立体匹配的自适应误差感知成本体积 (AEACV-Stereo)	2023.10 - 2024.03
第一作者兼开发者	北京
<ul style="list-style-type: none"><li>提出了基于误差图的动态采样策略, 用于加速双目匹配算法的迭代速度。</li><li>提出了噪声滤波代价体来提升立体匹配算法在 ill-posed 区域的性能。</li><li>代码: <a href="https://github.com/JarvisLee0423/AEACV-Stereo">https://github.com/JarvisLee0423/AEACV-Stereo</a>。</li></ul>	
基于 Transformer 的图像字幕算法	2021.09 - 2022.06
第一作者兼开发者	香港
<ul style="list-style-type: none"><li>提出基于 Full-Transformer 的图像字幕算法。</li><li>提出基于 BERT 的 Encoder 处理文本信息。</li><li>提出基于 Vision-Outlooker 的 Encoder 处理图像信息。</li><li>代码: <a href="https://github.com/JarvisLee0423/Captioning-and-Answering-with-Transformer">https://github.com/JarvisLee0423/Captioning-and-Answering-with-Transformer</a>。</li></ul>	
复杂图在神经网络中的应用	2020.09 - 2020.12
第一作者兼开发者	珠海
<ul style="list-style-type: none"><li>应用复杂图论 (随机图 ER, WS, BA) 在神经网络中, 构建随机网络。</li></ul>	

- 研究网络计算图拓扑结构和网络性能之间的关系。
- 网络性能通过时间复杂度和空间复杂度进行衡量和探索。
- 代码: <https://github.com/JarvisLee0423/RandWiredNN-Model>。

## 长寿与繁荣：鲱鱼和鲭鱼如何影响苏格兰渔业

2020.03 - 2020.04

合作者兼开发者

珠海

- 美国大学生数学建模大赛 (MCM 2020) Meritorious Winner。
- 应用线性回归预测未来 50 年的海平面温度变化，从而探索气候变化对于鲱鱼和鲭鱼迁徙的影响。
- 提出应用高斯分布模拟计算的捕鱼收益和捕鱼地点以及蛋白质腐烂变性之间的经济模型。

## 工作经历

### 字节跳动 - PICO MR 团队

2022.06 - 2024.02

计算机视觉算法工程师 (全职)

北京

- 3DoF PICO-VR 直播:
  - 实现了变形球面相机的立体视差估计系统。
  - 实现了扭曲方向交换系统，将扭曲方向从垂直改变为水平。
  - 修复了 OpenCV 从扭曲的球面视差生成深度以及从深度生成网格的错误。
- 双目视图完整 3D 物体重建算法:
  - 应用一致性和稳健的深度估计方法为两个视频的每一帧生成深度。
  - 应用 SLAM 为两个视频的每一帧生成位姿。
  - 从每个视频中重建一半物体。
  - 应用点云配准算法匹配每个视频的半个物体以生成完整的 3D 物体。
- 基于神经网络的 3D 场景重建算法:
  - 应用单目深度网络估计场景每个视图的深度。
  - 应用基于 RAFT-Stereo 的算法来估计场景每个视图的视差。
- 基于神经网络的深度估计准确度上限的探索:
  - 应用基于 RAFT-Stereo 的大视觉模型探索立体匹配方法的上限。
  - 应用基于多视角立体视觉的大视觉模型探索深度估计方法的上限。
  - 应用基于单目大视觉模型探索深度估计方法的上限。
  - 应用基于 DINO 和 iBOT 的大型视觉模型来提取更好的视觉特征以用于深度估计任务。
  - 使用大模型来提炼小模型以获得更好的性能。
- 深度估计 Codebase 的开发和维护:
  - 基于字节跳动代码引擎 aipack 开发了深度推理代码库。
  - 实现了大量的视觉任务主干，如 ResNet、EfficientNet、MobileNet、Vision Transformer 等。
  - 实现了深度估计任务的所有有用组件，如 Cost Volume、Correlation、FPN、ConvGRU 等。
  - 转移了所有以前的深度代码库，且准确度没有下降。

### 3D 打印实验室

2020.06 - 2020.08

研究助理 (实习)

珠海

- VR 展览系统开发:
  - 使用 Unity 实现 VR 展览系统，包括实时观看和编辑。
  - 实现物体编辑功能，包括拾取物体、缩放、旋转、平移、吸附墙面等。
  - 实现视图编辑功能，包括视图平移、投影、灯光变换。