**北 京 邮 电 大 学**

**2025届本科毕业设计（论文）中期进展情况检查表**

|  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 学院 | | | 电子工程学院 | | | 专业 | | 电子科学与技术 | |
| 学生姓名 | | | 张书凡 | 学号 | | 2021210646 | | 班级 | 2021211206 |
| 指导教师姓名 | | | 金磊 | 所在单位 | | 电子工程学院 | | 职称 | 副研究员 |
| 设计（论文）题目 | | | （中文）面向拥挤场景的高性能人体姿态估计算法研究 | | | | | | |
| （英文）Research on High-Performance Human Pose Estimation for Crowded Scenes | | | | | | |
| 目前已完成任务 | 主要内容： （毕业设计（论文）进展情况，字数一般不少于1000字）  1. 文献调研  在研究的初始阶段，首先对人体姿态估计算法进行了系统性调研，重点关注了传统方法与近年来的深度学习方法。其中，传统方法主要包括基于特征点检测的解析模型和概率图模型，而深度学习方法则分为Top-Down 和 Bottom-Up 两类。Top-Down 方法先通过目标检测器检测人体框，再在每个框内进行姿态估计，代表方法包括 Mask R-CNN、HRNet 等；Bottom-Up 方法则直接检测所有关键点并通过关联策略进行匹配，如 OpenPose、DEKR 及本文复现的 DecenterNet。  阅读 DecenterNet 相关论文后，理解了该方法采用 Decentralized Pose Representation（去中心化姿态表示），用所有可见关键点作为根点，使其在拥挤场景下更具鲁棒性。此外，该方法引入 Decoupled Pose Assessment（解耦姿态评估），通过位置图（Location Map）自适应选择最佳姿态，提高姿态估计的稳定性。  通过对比分析不同方法的优缺点，明确了当前技术在遮挡与交错（Entanglement） 场景下的挑战。例如，在多人交错的复杂场景中，传统方法容易出现关键点识别错误或关联错误，导致姿态估计失败，而 DecenterNet 通过 Limb Disentanglement Learning（肢体解耦学习） 进一步增强了对复杂场景的适应能力。  2. 环境配置与算法复现  在实验环境方面，成功搭建了深度学习框架，包括 PyTorch、CUDA、cuDNN 及相关依赖库，并在 COCO 数据集上完成了 DecenterNet 的复现。具体步骤如下：  配置 Python 运行环境，安装 PyTorch、Torchvision 及必要的依赖包。  下载并准备 COCO 数据集，完成数据格式转换及预处理。  复现 DecenterNet 训练和测试流程，调整超参数，使实验结果与原论文结果保持一致。  评估 DecenterNet 的基础性能，记录 AP（Average Precision） 指标，并与其他方法对比。  实验表明，复现的 DecenterNet 在 COCO test-dev 2017 数据集上达到了 71.2 AP，与论文结果基本一致，验证了复现的有效性。  3. 数据集分析  研究了 COCO、CrowdPose、SkatingPose 数据集的特点，分析了它们在拥挤场景下的适用性：  COCO 数据集是当前人体姿态估计的标准基准，但其数据大部分为普通日常场景，拥挤程度较低。  CrowdPose 主要用于拥挤场景的姿态估计，具有更高的交错度，但缺少可见性标签（Visibility Flags）。  SkatingPose 是一个新的数据集，专门用于评估交错场景下的姿态估计能力，包含可见性标签，并引入 Entanglement Index（交错度指标） 衡量拥挤程度。  进一步分析了 Entanglement Index 的作用，发现该指标能够更精确地衡量 交错程度对姿态估计的影响，并能够用于优化模型训练。  4. 实验环境优化  在训练过程中，优化了超参数设置和数据预处理流程，以提高模型在特定数据集上的收敛速度和准确率，包括：  采用混合精度训练（Mixed Precision Training），减少显存占用，提高训练速度。  进行数据增强（Data Augmentation），包括随机翻转、仿射变换等，提高模型的泛化能力。  使用 Adam 优化器，并调整学习率衰减策略，加快模型收敛。  采用多尺度测试（Multi-Scale Testing），提升模型在不同尺度下的适应性。 | | | | | | | | |
| 是否符合任务书要求进度 是 □ 否 □ | | | | | | | | |
| 尚需完成的任务 | 1. 匈牙利算法应用  在 DecenterNet 的基础上，引入 匈牙利算法（Hungarian Algorithm） 进行 实例分割，以自动匹配检测到的关键点，优化多人姿态估计的精度。主要思路如下：  使用 DecenterNet 预测所有关键点。  计算关键点之间的欧几里得距离，构建 匹配代价矩阵。  采用匈牙利算法 进行 最优匹配，确保每个关键点唯一归属某个人体实例。  2. 算法改进  关键点关联策略优化：结合 图神经网络（Graph Neural Network, GNN） 或 基于 Transformer 的特征提取，增强关键点关联的准确性。  Transformer 结构优化：将 自注意力机制（Self-Attention） 引入 DecenterNet 以提升全局信息捕捉能力，提高模型的泛化能力。  SkatingPose 数据集优化训练策略：针对 SkatingPose 的高交错场景，调整 Loss 权重，使模型更专注于解决 交错区域的关键点匹配 问题。  3. 实验与评估  在 COCO、CrowdPose、SkatingPose 三个数据集上测试改进算法，计算 AP 指标，与原始 DecenterNet 进行对比。  采用消融实验（Ablation Study），分别评估 DPR、DPA、LDL 三个模块的贡献，验证算法改进的有效性。  4. 论文撰写  整理实验结果，完成论文撰写，包括以下部分：  绪论：介绍研究背景、问题定义及研究目标。  相关工作：综述人体姿态估计的现有方法，并分析 DecenterNet 的优势与不足。  方法：详细介绍 DecenterNet 及优化方案。  实验：展示实验结果，并进行可视化分析。  结论：总结研究贡献，并展望未来改进方向。 | | | | | | | | |
| 是否可以按期完成设计（论文） □是 □否 | | | | | | | | |
| 存在问题和解决办法 | 存  在  问  题 | 1. 高度拥挤场景下的识别精度仍有待提高  问题：当前算法在 交错区域的关键点匹配误差较高，容易导致多人姿态估计失败。  2. 计算资源消耗较大，训练速度较慢  问题：实验训练过程较慢，占用大量显存。  3. 论文撰写进度较慢，缺乏完整实验数据支持  问题：实验数据尚未完全收集，影响论文撰写进度。 | | | | | | | |
| 拟  采  取  的  办  法 | 1. 高度拥挤场景下的识别精度仍有待提高  解决方案：  采用匈牙利算法进行关键点匹配优化，提高实例分割的准确性。  引入可见性标签（Visibility Flags），增强模型对遮挡区域的识别能力。  使用注意力机制（Attention Mechanism），重点关注交错区域的特征提取。  2. 计算资源消耗较大，训练速度较慢  解决方案：  使用混合精度训练（Mixed Precision Training），减少计算开销。  调整批次大小（Batch Size） 以提高 GPU 利用率。  采用剪枝（Pruning）和量化（Quantization）技术，优化模型计算量。  3. 论文撰写进度较慢，缺乏完整实验数据支持  解决方案：  尽快完成实验，整理数据，提高论文撰写效率。  制定写作计划，分阶段完成论文撰写。 | | | | | | | |
| 指导教师签字 | |  | | | 日期 | | 年 月 日 | | |
| 检查小组评分及意见 | | 评分： （总分： ）  组长签字： 年 月 日 | | | | | | | |

注：此表仅供参考，各学院应围绕毕设目标达成度，结合人才培养目标、专业认证要求等进行个性化完善。