**河南大学2025届本科生毕业论文（设计、创作）开题报告**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **论文（设计、创作）题目** |  | | | |
| **所在学院** |  | | **专 业** |  |
| **学生姓名** |  | | **学 号** |  |
| **一、本课题研究意义**  随着计算机视觉技术的飞速发展，视频目标检测与跟踪已成为智能监控、自动驾驶、智能安防等领域的关键技术。与静态图像相比，视频数据不仅具有更高的冗余性，还包含丰富的时空信息，如何从连续的视频帧中准确地检测和跟踪多个目标，成为研究者关注的热点问题。  传统的视频目标检测与跟踪方法在复杂场景下往往面临着精度不足和实时性差的问题。深度学习，尤其是卷积神经网络（CNN），在静态图像目标检测领域已取得显著成果，并逐步应用于视频目标检测任务中。深度学习技术能够通过多层网络自动提取图像特征，从而提升目标检测和跟踪的精度。然而，现有的深度学习模型仍存在许多挑战，例如如何提高检测算法的鲁棒性、如何保证视频序列的时空连贯性、如何实现轻量化的检测模型等，这些问题亟待解决。  本研究的意义主要体现在以下几个方面：  **1.** **推动视频目标检测与跟踪技术的发展**  本研究提出的多目标检测与跟踪系统，采用基于深度学习的算法（如YOLOv5和DeepSORT），能够显著提升检测精度与跟踪稳定性。在复杂背景下，系统能够识别和跟踪多个目标，尤其是在目标遮挡、光照变化等动态环境下表现出较强的鲁棒性。这一系统为视频监控、智能安防等应用提供了更为精确和高效的解决方案。  **2.** **推动视频目标检测与跟踪技术的发展**视频目标检测和跟踪不仅要考虑图像中的空间信息，还需要充分利用时序信息。本研究通过深度学习模型的时空特征学习，探索如何在视频序列中提取有用的时空信息，实现目标的准确跟踪。这一思路对于解决传统方法中难以处理的遮挡、重叠等问题具有重要的理论价值和应用潜力。  **3.促进多领域的应用与扩展**本研究的多目标检测与跟踪系统可以广泛应用于自动驾驶、公共安全监控、体育赛事分析等领域。通过结合深度学习与现代Web开发技术，系统能够实现在线展示与实时反馈，这不仅促进了学术研究的发展，也为各行业提供了可行的技术支持。  综上所述，本研究的成果将推动视频目标检测与跟踪技术的研究与应用，提升深度学习技术在实际场景中的应用效果，具有重要的理论价值和广泛的应用前景。 | | | | |
| **二、国内外有关本课题的研究动态**  **1. 目标检测的国内外研究现状**  目标检测是计算机视觉的核心任务，广泛应用于人脸、行人和车辆等领域。它包括两个主要任务：回归目标包围框，确定目标位置和大小；以及分类任务，对已定位的目标进行分类。  在深度学习普及之前，传统目标检测方法依赖滑动窗口技术扫描图像，提取特征并通过分类器进行识别，但这些方法计算量大且鲁棒性差。随着深度学习的应用，目标检测方法被分为两类：两阶段检测器和单阶段检测器。  两阶段检测器，如 R-CNN，首先通过深度学习提取候选框特征，并通过神经网络对感兴趣区域进行特征提取，最后通过分类器分类并调整目标框。尽管提高了检测精度，但存在重复计算等效率问题。He 等人提出使用SPP-net通过空间金字塔池化加速了推理过程，而Fast R-CNN和Faster R-CNN进一步提高了速度和精度。  单阶段检测器直接处理图像中的所有区域，无需生成感兴趣区域，适用于实时检测。YOLO（You Only Look Once）就是典型代表，它直接输出目标框和分类概率，因其速度快，广泛应用于实时视频处理。基于锚框的检测器，如 SSD 和 YOLOv3，后者通过FPN结构增强了小目标检测能力，而YOLOv5改进了CSPBottleneck模块，提升了特征提取能力。尽管基于锚框的检测方法提升了精度，但依然需要预设锚框，并受限于IoU损失的计算。而不基于锚框的检测方法，如CornerNet和CenterNet，采用关键点定位目标，避免了预设锚框的限制，且具有较强的扩展性。  **2. 多目标跟踪的国内外研究现状**  多目标跟踪是目标检测的重要研究方向，随着硬件性能的提升和目标识别技术的进步，跟踪技术得到了迅速发展。多目标跟踪算法分为基于检测的跟踪和无检测的跟踪。基于检测的跟踪首先进行目标检测，得到目标位置后，在时间序列中跟踪目标，最终获取运动轨迹；而无检测的跟踪需要手动标注第一帧目标，后续由跟踪器生成轨迹。两种方法的关键是通过运动或外观特征进行目标关联，生成连续轨迹。  基于检测的跟踪方法中，SORT算法通过IoU和卡尔曼滤波器的预测结果进行匹配，速度快，但存在ID切换问题。继 Wang 等人提出 利用深度学习对目标进行外观特征提取的方法之后，SORT 算法的改进版 DeepSORT方法改进了SORT，加入了Wide ResNet特征提取和余弦距离优化匹配，减少了ID切换，提升了精度。但该方法需要分别推理两个神经网络，计算量大，实时性差。  为了提高速度，联合检测与跟踪（JDE）算法提出了共享网络参数的多任务学习方法，在完成目标检测的同时进行特征提取，提升了效率。Wang 等人在 2020 年提出了 JDE 模型，JDE基于YOLOv3改进，增加了特征提取分支，优化了任务平衡。FairMOT进一步改进JDE，采用CenterNet替代YOLOv3，解决了锚框方法的问题，并通过不确定损失函数平衡任务间的冲突，低维度重识别特征也有助于提升精度和效率。2020 年 Liang 等人提出的 CSTrack[35]中将 JDE 模型中使用的 YOLO v3 替换为 YOLO v5 并为重识别特征提 取分支设计了一个多尺度感知的网络最后提出利用互相关网络来平衡各分支任务 之间的冲突。  针对检测精度问题，MOTDT方法引入区域全卷积神经网络进行目标分类，通过轨迹评分进行后处理，提升了检测准确性。ByteTrack通过二次匹配恢复低可信度目标框，保持轨迹连贯。SimpleTrack进一步通过余弦距离提升匹配精度。  这些方法逐步解决了多目标跟踪中的精度、实时性和效率问题，推动了该领域的技术进展。 | | | | |
| **三、本课题研究的基本内容**  本项目的主要研究方向涵盖了视频目标检测和跟踪算法以及相关数据集的整理和管理，旨在构建一个综合的视频目标检测和跟踪应用系统。这个系统具有以下关键组件：   1. **数据集管理组件：** 该组件负责整理现有的视频目标检测和跟踪数据集，并将它们纳入平台。关键信息如数据集名称、作者、应用场景等都会被记录和维护。 2. **视频目标检测和跟踪算法组件：** 视频目标检测和跟踪是计算机视觉领域的重要研究方向，它们在无人驾驶、视频监控和物联网等领域有着广泛的应用。视频目标检测的任务是在每一帧中识别出目标的类别和位置，而视频目标跟踪的任务是根据已知目标的大小和位置，预测后续帧中的目标位置。与静态图像不同，视频中的目标是动态变化的，这增加了检测和跟踪的复杂性。本项目将研究现有的视频目标检测和跟踪算法，对其性能进行评估，并优化这些算法以适应不同的应用场景。 3. **视频目标检测跟踪应用场景：** 这一组件旨在构建具体的应用场景，通过在线方式展示多目标检测和跟踪的视频。 4. **不同识别算法的效果对比展示组件：**这一组件是为了对比各目标检测算法对于视频的检测效果进行对比。 | | | | |
| **四、本课题拟解决的主要问题及对解决复杂工程问题能力的要求**  本课题拟解决的主要问题是基于深度学习的视频中多目标检测跟踪系统的设计与实现，包括三个层次，一是对现有视频目标检测与跟踪算法的性能评估和优化算法适应不同的应用场景的记录；二是对现有视频检测跟踪数据集的收集并录入平台进行维护；三构建视频多目标检测跟踪应用场景，在线展示不同场景下多目标检测与跟踪视频。四实现不同识别算法的效果对比展示。  本系统需要进行前端页面实现，后台服务实现和可视化展示多目标跟踪监测的任务，熟练运用Web开发、HTML、CSS、VUE开发、JavaScript、JQuery、Spring Boot、Docker、Kafka、NoSQL、YOLOv5、并行处理等技术构建基于深度学习的视频中多目标检测跟踪系统。 | | | | |
| **五、研究方法**  **1. 数据收集与预处理**  **本研究将使用公开的数据集（如MOT、KITTI、COCO等）进行训练和测试。这些数据集包含了视频中的多目标检测和跟踪数据。数据预处理包括图像标注、目标框提取以及视频帧分割，确保输入数据的质量，并通过数据增强技术（如旋转、裁剪、翻转等）提升模型的泛化能力。**  **2. 多目标检测算法的选择与优化**  **本研究将使用基于深度学习的目标检测算法，如 YOLOv5，它在实时目标检测中表现优异。YOLOv5作为单阶段检测算法，其速度快且精度高，适用于视频流实时处理。为提高检测精度，研究还将优化YOLOv5的网络结构，尤其是在小目标和复杂场景中的表现。**  **3. 目标跟踪算法设计与改进**  **目标跟踪将采用 DeepSORT 算法。DeepSORT通过结合外观特征和运动信息，解决了目标丢失和ID切换问题。通过引入基于 Wide ResNet 的特征提取模块，DeepSORT在提高精度的同时，解决了目标跟踪中的身份混淆问题。为提高系统的实时性，本研究将优化算法以减少计算时间并提高系统处理速度。**  **4. 系统架构设计与实现**  **本研究的多目标检测与跟踪系统将采用 Spring Boot 框架作为后端，提供高效的数据处理和算法实现。前端部分将采用 Vue.js 和 JavaScript 实现目标检测和跟踪结果的可视化展示。为提升系统的可扩展性和性能，研究将使用 Docker 容器技术部署系统，并通过 NoSQL 数据库存储检测和跟踪的结果。** | | | | |
| **六、主要创新点（选填）**     1. **系统架构创新**：采用Spring Boot后端与Vue.js前端的高效架构，结合Docker容器化部署，实现系统高可扩展性和高性能，确保多目标检测与跟踪的实时性。 2. **应用场景拓展**：构建多目标检测跟踪应用场景，在线展示不同场景下的实时视频，为自动驾驶、公共安全监控等多领域提供可视化解决方案，推动技术应用落地。 | | | | |
| **七、主要参考文献**  [1]肖禹超.基于视频的多目标跟踪与行为检测[D].北京邮电大学,2024.DOI:10.26969/d.cnki.gbydu.2024.000465.  [2]佟知航.复杂视频场景多目标跟踪及行为检测[D].北京邮电大学,2023.DOI:10.26969/d.cnki.gbydu.2023.003203.  [3]蒋贤皓.基于视频流的多目标检测与跟踪技术研究[D].电子科技大学,2023.DOI:10.27005/d.cnki.gdzku.2023.003322.  [4] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]. IEEE Conference on Computer Vision and Pattern Recognition, Columbus, Ohio, 2014: 580-587.  [5]Glenn Jocher, et al. YOLOv5: A state-of-the-art object detection model for real-time applications[C]. Ultralytics, 2020.  [6]佟知航.复杂视频场景多目标跟踪及行为检测[D].北京邮电大学,2023.DOI:10.26969/d.cnki.gbydu.2023.003203.  [7]韩瑞博.基于改进YOLOv5的矿下视频目标检测技术研究[D].哈尔滨工程大学,2024.  [8]张春平.基于多目标跟踪的监控视频中异常事件检测[D].北京邮电大学,2022.DOI:10.26969/d.cnki.gbydu.2022.003459.  [9]胡尊政.多目标检测及追踪算法研究与应用[D].杭州电子科技大学,2022.DOI:10.27075/d.cnki.ghzdc.2022.001240.  [10]陈波.视频中目标检测与追踪方法的研究与应用[D].电子科技大学,2011. | | | | |
| **具体时间及写作进度安排（表格不够可追加表格）** | | | | |
| **起止日期** | | **主要工作内容** | | |
|  | |  | | |
|  | |  | | |
|  | |  | | |
| **指导教师对开题报告的意见**  \*\*\*\*\*\*\*\*。是否推荐开题。  **指导教师签名： 2025年 1月3日** | | | | |
| **系（室）对本课题开题的意见**  \*\*\*\*\*\*\*\*。是否同意开题。  **负责人签名： 2025年 1月7日** | | | | |