重庆师范大学

专业学位硕士研究生论文开题报告

及实施计划

所在学院

专业学位类型

专业领域

研究生

指导教师

入学年月

计算机与信息科学学院

工程硕士

计算机技术

刘 晨

杨德刚

2020年9月

重庆师范大学研究生院

2022年 7 月

**说 明**

一、开题报告作为学位论文撰写和学位申请的重要材料，应及时完成，不得随意更改。确需更改开题报告，应由学生提出申请，导师签字同意，院学位评定分委员会审查盖章后交由校学位评定委员会办公室审核同意。

二、本表一式二份，A4纸张双面打印，研究生院、学院各一份。

1. **论文开题报告**

|  |  |
| --- | --- |
| **拟定论文题目: 雾霾环境下车辆检测与前向车辆测距研究与应用** | |
| **选题分类（√）** | **1. 基础研究 2. 应用研究√ 3. 综合研究 4. 其它：** |
| **选题来源 （√）** | **1. 国家基金项目 2. 教育部人文社科项目**  **3. 省部级项目 4. 其它√** |
| **开题报告内容（选题价值、文献综述、研究或开发设计、可行性分析等）**  **一、选题价值**  在当今时代，世界工业的发展已经达到非常成熟的阶段，伴随其而来的则是日益严重的环境污染。工业废气、机动车尾气的大量排放，非法燃烧垃圾，绿化不足等等现象导致了在秋冬季节空气颗粒浓度上升，室外能见度下降。这种雾霾天气不仅给人们的日常生活以及交通出行带来了很大困扰，在监控等场景下也使摄像机的成像质量下降。 在户外安防监控场景下，监控视频也容易受到雾天天气的影响。传统监控摄像头成像本身已很难保证监控目标的角度、大小合适，成像清晰，又由于监控任务在时间线上的不可逆转重现性，雾霾的影响会大大的降低监控视频或图像的利用价值，给道路、社区等户外公共场所的监控质量和安全带来很大的影响。  随着我国社会经济的不断发展，人们的生活水平逐渐提升，自驾出行已成主流出行方式，机动车数量也在持续增加，据统计，截至2019年底，我国汽车总量相比2018年增加2122万辆，达到2.6亿同比增长8.8%[1]。在智能交通场景下，浓雾天气会导致机动车驾驶员视野变差，安全性降低，易发生交通事故致使交通堵塞。为解决交通拥堵和避免交通事故，并拥有快速、及时的交通调度能力，智能交通技术（Intelligent Transportation System, ITS）应运而生[2]。ITS主要是将计算机视觉技术应用到交通行业当中，对道路信息进行立体化获取，并赋予计算机自主判断和决策能力，结合交通数据，进行车辆检测与跟踪，通过车流量分析路况，合理规划红绿灯时间等多方面处理，从而形成智能化的交通检测系统。  智能交通系统由多个部分组成[3]，各个部分在处理交通问题上负责的领域不同，系统之间的相互联系使得智能交通系统拥有强大的信息获取能力，车辆检测与跟踪作为智能交通技术的关键技术之一，使用摄像头对道路中行驶的车辆进行拍摄，通过计算机图像处理技术，将车辆目标分离出来，并对其进行标注与定位，从而识别拍摄到的车辆；车辆跟踪是在摄像头监控中，从连续的视频序列识别车辆，在视频中匹配并判断是否为同一辆车，对识别出的车辆进行跟踪与计数，快速获得道路车流量信息。智能交通系统对这些信息进行后续处理，判断当前道路是否堵塞，并将决策信息通过信息服务系统在网络平台上告知交通管理部门以及驾驶员，交通管理部分根据车流量信息对道路资源进行合理分配，自主调整交通信号灯，缓解交通拥堵压力，并引导行驶车辆避开拥堵道路；驾驶人员也可根据道路信息选择畅通道理，减少驾驶时间，在一定程度上减少交通事故的发生。因此智能交通下的车辆检测和跟踪不仅对提高道路交通管理具有很大帮助，而且对建立智慧城市、平安城市有重大的社会意义。  计算机视觉利用计算机和感知设备模仿人类的视觉系统，将人眼所能感知世界的视觉能力赋予计算机，进而完成一些特定的任务。目标检测一直是计算机视觉研究中的重点问题，也是很多计算机视觉系统应用的基础，目标检测的任务是给定的图像或视频中检测到每个独立的个体。 | |
| 传统基于图像的车辆检测与识别方法，通常需要标注者通过先验知识对目标定位，手工提取特征，每当出现一类新的车型，又需要重新设计特征，在实际应用场景中，鲁棒性差，泛化能力弱，所以检测准确率与检测速度均比较低，无法满足当前智能交通系统发展的要求。近年来，随着数据量呈指数型的增长，以及硬件设备的数据计算能力快速发展，深度学习发展迅速，并应用于目标检测的研究领域。基于深度卷积神经网络(Conventional Neural Network，CNN)[4]的多种目标检测框架相继被提出，其大致可以分为两类：基于候选区域的深度学习目标检测算法和基于回归方法的深度学习目标检测算法，也称为两阶段（two-stage）和单阶段（one-stage）检测方法。基于候选区域的检测算法首先在图像上采用区域候选网络（Region Proposal Network，RPN）[5]产生候选区域，然后在候选区域内使用卷积神经网络对目标进行分类和定位，代表算法有 R-CNN[6]、Fast R-CNN[7]、SPP-Net[8]等，此类方法精度高，但实时性差。基于回归思想的检测算法不需要候选区域，直接在特征图中检测目标位置，代表的算法有YOLO[9]、SSD[10]、RetinaNet[11]等，此类方法速度快，但精度较低。基于CNN的目标检测算法在检测时需要的人为参与量较低，特别是单阶段目标检测框架，只需要为网络提供大量的带标注的图像进行训练，在达到某种要求后便可以实现对图像目标进行端到端的检测识别，在检测正确率上高于传统检测方法，并且具有再学习能力。  然而，在实际的交通道路应用中，通常存在目标尺度差异大、背景复杂和密集目标等情况，由于交通摄像头监控范围较大，拍摄距离远，或因为视角等问题，因此交通对象均为小目标物体。在复杂交通场景中存在许多小目标，由于背景较为复杂，现有的目标检测和跟踪算法对小目标的检测与跟踪效果较差，甚至出现漏检等问题，无法满足智能交通系统对交通目标的高精度、高速度检测与追踪的要求，车辆检测方法在识别准确率以及检测速度上仍需要进一步提升。因此，如何能够提高车辆检测效率和准确性，并进一步提升车辆跟踪算法的性能，克服雾天场景对于智能交通的影响，这对解决城市交通问题、预防或减少交通事故以及道路规划问题具有重要的现实意义。  **二、文献综述**  **1、去雾算法**  图像去雾分成两个不同的研究方向，一是基于图像复原的去雾方法，根据大气退化模型，进行响应去雾处理；二是基于图像增强的去雾方法，直接在图像上去除噪声，提高对比度。  基于图像复原的去雾方法，主要运用雾天大气退化模型[12]。通过大量的图像数据分析假设其衰减和光线等各因素和参数的影响，建立雾天条件下图像成像模型，转而求解图像降质的逆过程。Tan[13]等人基于图像对比度高区域及大气光的变化与景深的关系，建立了马科夫模型的损失函数方程，使用局部对比度进行优化。该方法适用于灰度图和彩色图像，恢复图像的质量较好，但在对比度变化较大的区域容易使色彩失真，造成处理过度。Kartz[14]等提出目标的反射与景深相互独立，使用幂函数重尾分布先验建立反射率梯度模型。此方法在处理图像时具有良好的边缘保留效果，而处理彩色图像时易造成色彩饱和。Ding[15]等针对 He 暗通道先验提出了使用四叉树分割法对图像进行分割，再使用软抠图优化透射率分布。Fang[16]等提出了使用细化透射率分布代替软抠图的算法。此外，还有众多学者通过使用雷达等测距仪器来获取实际的景深，再代入数据通过大气退化模型来复原图像。  近年来，Retinex 图像增强方法出现在去雾研究中，Retinex 理论模型是在人类感官视觉模型的基础上发展起来的，具有色彩恒常性和动态领域压缩的特点，Retinex 算法不断发展应用于图像去雾中并取得了良好的效果。自Retinex理论提出以来，它经历了三个发展阶段。1971年，Edwin Land[17]提出该原理用于路径算法，该算法从当前像素的邻域中随机选用下一个像素，然而该算法的缺陷是计算量过大。1997 年，Jobson 等人[18]提出在对数领域中计算得到光照值的方法，在该领域中计算光照反射量从而降低光照的影响。2003 年，Kimmel[19]提出将多种 Retinex 算法统一为变分形式的方法。在图像去雾方面，学者们结合雾天图像特点提出基于 Retinex 去雾算法。张洪坤等人[20]提出基于非线性分散均值漂移平滑的 Retinex 去雾算法，其算法对图像清晰化处理后产生的光晕、颜色失真等问题都有一定的补偿作用。李菊霞等人[21]提出了基于改良的多尺度Retinex 去雾算法，其算法不仅能够使图像更合乎人眼感官系统，同时还提高运算速度。E Salari 等人[22]提出针对图像细节复原的去雾算法，该方法结合多尺度Retinex算法，算法的复杂度小，速度快。  **2、目标检测算法**  目标检测算法是计算机视觉领域的重要分支，对图像数据中的物体类别进行检测，并能实现对该物体的定位。在目标检测发展的20多年来，目标检测技术取得了长足的进步，以2014年左右为分界线，分为两个阶段，一个是使用传统算法的“传统目标检测算法”阶段，另一个是基于深度学习的“深度学习目标检测算法”阶段。以下将从传统目标检测算法与深度学习目标检测算法两个阶段阐述该目标检测技术的研究进程。  传统的目标检测算法通常分为四个步骤：（1）选取候选区域：在给定的图像上使用多尺度的滑动窗口进行遍历；（2）对候选区域提取特征：通过特征提取器来手工提取目标特征，如方向梯度直方图特征(Histogram of gradient, HOG)[23]、尺度不变特征变换算法（scale invariant feature transform，SIFT）[24]、加速稳健特征（Speeded Up Robust Features，SURF）[25]等；（3）使用分类器分类：分类器对提取到的特征进行识别，如 SVM (Support Vector Machine）[26]、AdaBoost[27]等。  2001年，Viola[28]提出了用于人脸检测的V-J检测器，在相同的检测精度下，检测速度超越当时其他算法数十倍；N.Dalal和B.Triggs在2005年提出了方向梯度直方图(HOG)特征检测器，该检测器提取区域内各个方向的梯度特征，在行人检测中检测效果良好。由于人工提取目标特征过程是十分困难的，时间复杂度高且鲁棒性差，泛化能力弱，不适用于复杂场景。  随着数据爆炸式增长和计算机硬件计算能力的大幅提升，基于深度学习的目标检测算法逐渐成为主流，可大致分为两类，一种是基于候选区域的目标检测算法，称为两阶段（two-stage）检测方法，如R-CNN、Fast R-CNN、SPPNet等；另一种是基于回归方法的目标检测算法，称为一阶段（one-stage）检测方法，如YOLO、SSD等。  2014年，Ross Girshick 等人提出了首个两阶段目标检测算法R-CNN（Region with CNN features），利用选择性搜索[29]（Selective Search）的方法提取候选框，利用CNN网络进行特征提取，使用SVM算法对特征进行分类。之后，R-CNN算法的作者Ross Girshick借鉴SPP-Net的思路，提出了Fast R-CNN，使用原理和功能相似的感兴趣池化层（ROI pooling layer）替代SPP层，将算法的检测精度提高到了70%左右。  一阶检测算法省略了候选区生成计算部分，直接构建端到端的神经网络对目标进行检测分类，大幅缩短检测时间。2016年，Joseph Redmon针对检测速度创造性的提出了YOLO（you only look once）算法, 不需要选取候选框作为检测区域，将整张图片作为网络的输入，将目标检测作为回归问题来解决。同年，刘伟等人提出了 SSD（Single Shot multi-box Detector）检测算法，SSD在检测时使用不同的尺度对输出的特征图进行分类预测，提升了小目标检测的性能。此后，Joseph Redmon在YOLO算法的基础相继提出了YOLOv2[30]，yolov3[31]，检测精度均有所提高。2020年Alexey Bochkovskiy等人在yolov3算法的基础上提出了YOLOv4[32]，将CSPDarknet53作为骨干网络，使用 PANet[33]（Path Aggregation Network）进行参数聚合，结合了目标检测领域大量的 tricks，如跨阶段部分连接( CSP)、Mish 激活函数、Mosaic 数据增强、DropBlock 正则化等，实现了检测速度和精度的最佳权衡。  **3、目标跟踪算法**  视觉目标跟踪算法的主要任务是在一组连续的图像序列中寻找目标。将视频文件以帧为单位，从初始帧检测到目标位置，并预测后续每帧中该目标所在位置坐标，使得之后的图像序列能够进行正确的目标跟踪。  近年来，随着深度学习研究不断发展，以卷积神经网络（Convolutional Neural Network, CNN）为代表的深度神经网络被广泛地被应用在目标检测、目标跟踪等领域，并取得显著提升。2015年。上海交通大学的马超等人提出分层卷积特征的跟踪算法（Hierarchical Convolutional Features for Visual Tracking，HCF）[34]，利用VGGNet[35]中三种不同的卷积层（Conv3-4,Conv4-4,Conv5-4）作为特征提取层，从三个特征提取层分别输入到三个相关滤波器中，最后得到的三个权重图进行加权融合，从而得到目标的精确位置。Danelljan等人将CNN与传统的SRDCF算法相结合，使用CNN提取特征替换原来的特征，提出了DeepSRDCF[36]跟踪算法，取得了较好的效果。2016 年，Nam 等人[37]提出了多域神经网络MDNet，MDNet预训练一个包含共享层和多个独立域分支的特定域层组成，每一个域对应一个训练序列，在训练过程中共享层学习通用的目标表示。held等人首次将孪生网络引入目标跟踪领域，提出了基于回归网络的通用目标跟踪算法（Generic Object Tracking Using Regression Networks，GOTURN）GOTURN[38]使用大量带标签的图像、视频进行离线训练，将前一帧与当前帧的图像剪裁后送入卷积神经网络进行特征提取，将这些特征图级联并输入全连接层，回归得出预测位置信息。Bertinetto 等人提出的全卷积孪生（SiameseFully-Convolutional, SiamFC）[39]利用全卷积网络来提取模板和搜索区域的深度特征，根据模板与搜索区域得到相似性度量函数，通过相似度寻找最大概率目标位置。Valmadre等人对 SiamFC 改进得到 CFNet[40]跟踪算法，CFNet 推导了相关滤波器的前向和反向传递函数，并添加到模板分支中来更新模型。  **三、项目的研究内容、研究目标,以及拟解决的关键科学问题**  **1、研究内容**  本研究是基于YOLO目标跟踪算法及其应用，研究内容主要分为如下几个方面：  1.1、数据收集与预处理  车辆检测的数据集对车辆识别模型训练非常重要，因为道路上的车辆类型多种多样且环境比较复杂，自然光照的强弱都会影响车辆检测模型的精准度和帧速率。因此构建一个适合车辆实时检测的数据集是非常重要的。目前常用的有 PASCAL VOC2012、KITTI 和 COCO 数据集。这些数据集虽然在车辆检测中具有良好的表现，但是想要实现车辆实时检测的功能，现有的数据集难以达到很好的训练效果，因此还需要制作一个适合车辆实时检测的数据集。  本课题拟通过摄像头采集道路上车辆，模拟交通监控摄像头拍摄视频，使用LabelImg标注工具制作数据集。Voc2012数据集中存在车辆照片，可通过LabelImg标注工具重新制作数据集，与自制的数据集相结合，保证数据集体量足够训练与验证。  1.2、目标检测算法改进  YOLOv5目标检测网络结构由主干网络 (Backbone) 、颈部网络(Neck)和输出网络(Prediction)组成。在主干网络的层面上，YOLOv5使用了CSPDarknet模块，其大致原理是利用CSPDarknet提取特征信息。使用CSPDarknet在一定程度上可以减少梯度信息的重复。将输入分为两部分，一部分进行Bottleneck分支特征提取，另一部分通过一个卷积层过渡，最后在深度上进行堆叠。CSPDarknet结构的主要目的是使特征学习网络实现更丰富的梯度组合，增强CNN的学习能力，同时减少计算瓶颈、减少将近一半的显存占用。在 Backbone 结束时使用了空间金字塔池化层（The Spatial Pyramid Pooling Layer，SPP Layer），其优势在于不用固定图像输入的尺寸即可产生相同大小的输出特征，同时融合了多尺度池化信息。在颈部网络结构的层面上，YOLOv5使用了PANET来对特征进行融合，即在FPN特征金字塔结构之后加入了一个自底至顶的路径。深层特征更加关注目标的整体信息，浅层特征更多关注目标的纹理和位置信息，PANet增加的这一路径将低层信息又传到高层，同时减少了浅层纹理位置信息传到深层的信息流通所需跨越的卷积层数，使得多层次、多尺度之间的信息流通更加顺畅。颈部结构网络的作用是用来生成一个特征金字塔结构，以此实现检测不同大小和尺度的目标。在输出网络结构的层面上，使用Head网络结构来预测物体的位置，并生成相应地预测框。  本文拟使用轻量化的特征提取网络Enhanced ShuffleNet作为主干网络，可大幅度减少网络参数和计算量，加快检测速度；由于交通道路背景较为复杂，在网络结构中的特定位置增加深度可分离卷积层(SeparableConv)、或者是轻量化的模块，例如GhostBottleneck模块等。  深度可分离卷积（depthwise separable convolution），由depthwise(DW)和 pointwise(PW)两个部分结合起来，用来提取特征。而pointwise convolution其实就是普通的卷积，只不过其采用的卷积核。对于depthwise separable convolution，其首先是采用 depthwise convolution 对不同输入通道分别进行卷积，然后采用pointwise convolution将上面的输出再进行结合。宽度因子，对每一个模块要使用的卷积核数量进行Reduction；分辨率因子，对输入数据以及由此在每一个模块产生的特征图进行Reduction。  结合的卷积层计算量为：  普通卷积模块就是直接普通的进行非线性，而 Ghost 模块是先进行一次普通的卷积（即卷积-批归一化-非线性激活），也就是说经过 Ghost 模块的卷积变化之后，网络结构输出的特征图数量急剧减少，也就达到了减少计算量的目的，最后在特征图减少的基础上进行相关的线性变化。  其中理论加速比(speed-up ratio)计算公式为：  理论压缩比(compression ratio)计算公式为  Ghost Bottleneck 模块是一个可复用模块可以拿到现有的网络中替换掉 bottleneck模块，从而减少计算了，降低模型体积。  网络输入维度为，即原始图片预处理之后的长宽都为640的3通道RGB彩色图片，通过网络后，大中小3个尺度的检测层输出维度为，其中为划分的网格个数，为每个尺度对应预设先验框的个数，为需要预测的类别数。以大尺度为例，该维度网络结构的检测层输出维度为。网络模型主要预测参数有边界框相关参数，边界框置信度和第类别的置信度。这些参数需要解码公式得到最终的预测框。    蓝色框是标签边界框,分别为标签边界中心点坐标和宽高尺寸，是该标签边界框所占据网格与左上角的网格距离；红色的是先验框，是先验框的宽高。  在目标检测中，由于所处的位置，拍摄角度等原因，会得到一些相互遮挡的图像。假设有两个相互遮挡的目标，若前面的目标A的置信度较高，为 0.92，后边的目标B 置信度较低，为 0.74，两个目标的交并比为 0.56。在这个情况下，YOLO 的阈值默认为 0.5，在阈值 0.5 以上的情况下，原网络计算的方式就是保留目标A，将后边的目标B置信度重置为0，因此会得到前边的物体被检测出来，后边的物体被舍弃的情况，结果只有目标A 被检测出来，没有目标B。  本文拟设计一种自适应 NMS 算法，使得在交并比高于阈值时，自动调整阈值，减少自动忽略后边被遮挡物体的情况。  1.3、注意力机制  注意力机制自提出以来就被广泛应用于深度学习的各个领域，各领域的蓬勃发展为注意力机制的进步做出了积极贡献，各领域也对注意力机制进行了专业化改造，视觉领域设计了通道注意力、空间注意力等针对图像卷积的注意力模型，对注意力模型进行了扩展。虽然现有的注意力模型能够满足现有深度学习模式下的大多数需要，但是注意力新原型模型的提出和建立仍然必要，新的注意力原型可以衍生出更多的注意力模型，这些更多的注意力模型将为深度学习的进步提供强大的动力。  在通道上的特征融合时，一般的卷积计算是将所有的通道进行合并，相较于一般的卷积操作，SENet网络结构则是没有直接将各个通道上的特征进行合并，而是利用网络模型自动学习特征，结构如图。    原特征图C×W×H进行全局平均池化，然后得到了一个 1×1×C大小的特征图，这个特征图具有全局感受野；输出的1×1×C特征图，再经过两个全连接神经网络，最后用一个类似于循环神经网络中门的机制。通过参数来为每个特征通道生成权重，其中参数被学习用来显式地建模特征通道间的相关性；使用Excitation 得到的结果作为权重，然后通过乘法逐通道加权到U的C个通道上，完成在通道维度上对原始特征的重标定，并作为下一级的输入数据。  本文拟提出一种空间注意力机制，融合通道注意力机制，使目标检测网络在复杂背景下获得更好的特征，提升对小目标检测精度。  1.4、目标跟踪算法改进  在跟踪算法中，对首帧图片中初始化目标，为目标分配跟踪器；在处理下一帧之前，采用关联算法将当前帧检测结果与上一帧中目标进行匹配，找出最适配组合即为同一目标在不同帧中的运动形态，在关联过程中，对目标消失、突然加速、被遮挡等特殊情况进行判断，并进行轨迹修正，与此同时当前帧检测目标成为新的轨迹预测对象，开始下一阶段匹配，直至视频结束。    轨迹预测主要任务是根据检测器得到的该目标的最优中心点坐标输入到滤波器中，进行下一帧可能会到达的中心点的预测。核心是根据已知坐标求出下一帧坐标最大概率值，现有许多成熟的滤波算法已实现通过系统的输入值，观测系统状态，同时消除输入数据中的背景噪音干扰，输出最优估计值。考虑到本文的跟踪对象是处于连续变化的环境中，且均平面上线性移动，故拟选择在线性变换中预测效果优异的卡尔曼滤波。将当前 T 帧的检测器输出目标框最优中心点作为卡尔曼滤波器输入，对 T+1 帧时该目标位置进行预测，一般分为预测量和更新量。  YOLO\_V5 检测器的目标框尺寸参数为(，常规认为目标在短时间内宽高不变，且保持匀速运动，则在T 帧时，可将检测器中的目标 a 的测量结果表示为，其中表示目标在 x 水平，y 竖直方向的速度分量。将 帧时的检测值代入下式则可得预测量。  其中为预测点的状态转移矩阵，为预测过程噪声矩阵，为状态协方差矩阵，表示外部干扰。更新最优估计点是通过上一时刻最优估计点的预测值加上外部噪声干扰修正所得，同时新的不确定性因素由上一时刻不确定性结合外部干扰预测而得。  在求出预测值后，利用此时刻的检测网络所测得目标中心位置对此预测值进行修正，设为帧的检测所得中心坐标值，则可求出检测值与预测值残差：  其中为检测矩阵，从而可得最终的最优预测值：  其中为卡尔曼滤波增益值：  其中为测量过程中的噪声矩阵。  在实际跟踪场景中，同一目标在前后两帧中的目标框坐标变化程度较小，利用匈牙利算法将该目标的上帧检测目标框与本帧目标框进行对比，构成其代价矩阵，同时进行 id 匹配。当两者边框面积的 IOU 值，即两帧中目标边框重合度越高，说明两帧中为同一个目标的概率越大。  本文拟利用DeepSORT算法对车辆进行跟踪，使用改进后的YOLO v5目标检测算法作为 算法中的检测器。DeepSORT算法利用卡尔曼滤波(Kalman Filter)和匈牙利算法(Hungarian algorithm)的有效结合，将检测算法得到的车辆边界框信息和卡尔曼滤波预测的边界框信息进行正确的关联。此时利用匈牙利匹配算法将两者的边界框信息进行匹配。匈牙利算法是在多项式时间内，对任务分配问题求解的组合优化算法，常为指派问题提供解决方案。利用两帧的边框面积以及 IOU 值构成匈牙利算法的代价矩阵，然后利用匈牙利算法找出最优匹配结果，通过调节 IOU 值对应权重，有效减少由于相机抖动造成 IOU 误差。具体见公式如下：  示中，分别表示在当前帧与上一帧目标的边框面积，表示匈牙利算法代价矩阵中对应元素的值，表示当前帧与本帧目标框对匈牙利算法代价矩阵中的贡献比例，表示当前帧与本帧目标框面积关系对匈牙利算法代价矩阵中的贡献比例。  在实验的过程中，遍历当前帧时的所有目标框与帧时刻该目标框的交并比，选出最大交并比值，将与固定阈值作比较;当大于等于 时，证明帧产生此交并比值的目标框就是帧待匹配的目标，更新id值、目标框信息以及卡尔曼滤波器中的状态协方差矩阵。  在实际应用场景中，视频中除了目标物体外还存在大量背景物，在目标物运动时被背景物遮挡而短暂消失，极易导致跟踪偏差或失败等问题，且遮挡按照运动时间可分为部分遮挡、严重遮挡、完全遮挡，本节提出针对性方案解决了遮挡情况而导致的目标关联错误问题，以对一连串的错误轨迹做修正。传统目标跟踪算法在被遮挡时，无法正确的跟踪目标主要是两方面原因引起：（1）缺少针对性判别机制，一旦运动环境复杂，目标物极易被遮挡，如跟踪器中缺少判别机制则无法确定此帧目标消失是由于被遮挡还是离开环境，同时也无法对跟踪误差进行修正。（2）被遮挡目标的消失会导致数据关联失败，轨迹预测发生偏差，跟踪器中缺少误差处理模块。  针对第一个问题，本课题拟提出一种遮挡判别机制，根据前后时间对应不同检测结果置信度进行区分遮挡状态。首先跟踪过程中目标k与环境内每个剩余目标间欧几里得距离值为d，此距离为两框体中心位置距离，但当目标过大或宽高比过大时，d不能代表相邻目标框之间的最近距离值，故结果检测器中检测各目标的真实边框信息。对于第二个问题，通常情况下，目标在环境中运动时，目标形状、大小不会发生突变，可利用这一特性修正轨迹，拟提出一种轨迹修正策略，根据当前帧目标宽度和高度相比于前一帧宽度和高度的变化量，判断被遮挡部分，并对目标该方向的目标值进行修正。  1.5、去雾算法的改进  通过对雾天图像观察经验，结合道路目标的景深位置，本课题拟对图像进行局部高斯滤波处理；再通过双边滤波器卷积处理作入射图像的照度估计，最后使用非线性函数增强反射图像。  **2、研究目标**  针对研究现状和研究所存在的问题，本文提出基于YOLO的目标检测算法，与目标跟踪框架相结合，构建实时目标跟踪算法，并将算法应用到车辆检测与跟踪系统。  （1）本课题旨在基于YOLO的目标跟踪方法，经过改进使得该算法更适合于交通车辆小目标的检测。本课题将从数据集的搜集与处理出发，从处理好数据配对训练开始，构建轻量化的特征提取网络，探索基于卷积的注意力网络，构造损失函数和聚类算法，建立具有较高精度的实时目标跟踪网络结构。  （2）将目标跟踪算法应用与车辆检测与跟踪系统，对交通车辆进行车辆检测、车辆跟踪车流量的统计及车辆行为判断。  **3、拟解决的关键问题**  （1）交通车辆数据集中车辆目标较小，背景复杂，雾天车辆目标不明显，现有的YOLOV5对小目标不敏感，存在小目标检测精度较低的问题。  （2）车辆行驶时存在被背景物遮挡而短暂消失的情况，极易导致跟踪偏差或失败，在复杂交通场景下进行目标跟踪是本项目的一个挑战。  **四、拟采取的研究方案及可行性分析**  **1、技术路线**  本课题研究的技术路线由数据集制作与处理、模型改进、模型训练和系统开发总共4个部分组成。此4部分内容贯穿整个目标跟踪检测过程，以下对此4部分的进行简单的技术路线介绍。  本课题拟通过摄像头在天桥采集道路上车辆，模拟交通监控摄像头拍摄视频，使用LabelImg标注工具制作数据集。预处理Voc数据集，与自制的数据集相结合，完善数据量，对模型训练有正向作用。  本课题模型改进部分主要针对检测算法优化、引入注意力机制和跟踪算法优化。检测网络优化主要依托更换轻量化特征提取网络，减少计算参数，引入自适应 NMS 算法，减少自动忽略后边被遮挡物体的情况；引入注意力机制选用空间注意力与通道注意力相结合的方式实现；跟踪算法优化通过将改进的目标检测算法作为检测器，输入到跟踪器，提高跟踪精度和速度，提出新的遮挡判别机制和轨迹修正策略，改善遮挡目标的检测与跟踪精度。  本课题通过模型训练实现车辆检测与跟踪，并根据训练结果改进或调整模型参数，以达到模型最好检测与跟踪效果。  本课题通过需求分析，针对交通车辆实际需求做出总体设计，采用基于.NET 平台的 3 层分布式系统架构设计，C#是主要的开发语言，并配合微软公司推出的 Visual Studio 2010 集成开发环境。    **2、研究方法**  2.1、文献综述法  大量收集国内外目标检测算法和目标跟踪算法的期刊和会议论文，了解所使用的，技术路线，以及每个技术的优缺点，并以此进行相关理论分析。最终能够从理论上寻找可行性技术路线。  2.2、实验法  大量阅读论文后，筛选一些与本研究相关的论文，并对其代码进行测试了解检测与跟踪效果，尝试发现其中的不足点，并且寻找改进方案。同时，可以尝试加入一些自己的创新想法或者加入其它目标检测思想来达到改进的目的。算法使用轻量化的思想，轻量化网络可以大幅度降低目标检测与目标跟踪算法的计算量，同时提高算法的检测速度，实现实时的检测与跟踪算法，但同时也会牺牲一部分精度，因此需做对比实验，筛选合适的轻量化网络，以平衡检测与跟踪算法的精度和速度。  2.3、对比分析法  首先利用目前已有的技术进行主客观评价指标的对比，发现各自的优缺点以及合适的应用场景，针对各自的情况，选其优点，并改进其缺点，最终能够实现在车辆检测与跟踪模块上比已有的方法更好的效果。  **3.实验方案**  根据本课题的研究内容、研究方法和技术路线，对各研究部分拟定以下研究方案：  3.1、目标检测评价指标  (1)mAP：目标检测中较为常用的评价指标。是评价目标检测模型精度高低的指标，其大致原理是计算所有类别的精确率(Precision)和召回率(Recall)，并计算所有类别 AP 的平均值，AP 是每个类别自身的平均精度。其中精确率又叫查准率，召回率又叫查全率，公式如下：  True Positive(TP)：（一般取0.5）的检测框数量。  False Positive(FP):目标检测框的数量，也就是说检测到相同值得其余的检测框的数量。  False Negative (FN): 没有检测到的 GT 的数量。  True Negative (TN): 在 mAP 评价指标中不会使用到。  一般来说，AP 这一评价指标是指计算 P-R 曲线下的所有面积，而 mAP 指的是是计算所有类别 AP 的平均值。  其中，表示第q个类别的平均精度，Q表示类别数目。  (2)推理速度：指的是车辆检测模型在检测完固定的图像帧后，所耗费的时间。  (3)模型参数量：模型参数量一般衡量模型的大小和复杂度。  3.2、目标跟踪问题评价指标  （1）MOTA（Multiple Object Tracking Accuracy）: MOTA 主要考虑的是跟踪中所有对象匹配错误的情况，主要是 FP、FN、IDS，可以非常直观表现跟踪过程中保持轨迹的性能, 计算公式如下:  其中，FN 为假阴性，即误报，整个视频误报数量之和；FP 为假阳性，即漏报，整个视频漏报数量之和；IDSW 为 ID Switch，即 ID 切换总数，整个视频误配数量之和；GT（Ground Truth）为实际物体的数量，整个视频 GT 数量之和。MOTA 越接近 1 表示跟踪其性能越好，由于有跳变数的存在，当看到 MOTA 可能存在小于 0 的情况。  （2）MOTP（Multiple Object Tracking Precision）：多目标跟踪框的精度，计算公式如下:  其中，指的是时刻跟踪结果与为其匹配的真实标注轨迹信息边界框之间的交叠程度，为时刻匹配成功的边界框个数。MOTP 越大，表示跟踪框与真实标注框越接近。  3.3、系统实现三个对比实验  （1）使用测试数据集，进行数据预处理之后将改进的YOLOV5目标检测算法与原有的YOLOV5目标检测算法进行对比实验，比较mAP、推理速度和模型参数量。  （2）将本章算法与现有目标跟踪算法中性能较为优越的 4 种算法同时应用在测试数据集中，分别为：基于多通道颜色空间特征进行目标跟踪的 CN 算法，基于 HOG 特征的相关滤波算法 KCF，采用多跟踪器的快速跟踪算法 MEEM，以及与本章算法一样都采用深层卷积神经网络特征的 CCOT 算法，比较4种算法与本课题MOTA和MOTP指标。  （3）选取包含多次遮挡，且遮挡时间较长的视频序列，适用于对本课题提出的遮挡解决模块做针对性测试。在前述4种算法和本课题方法中测试，并比较MOTA和MOTP指标。  **4、关键技术**  本课题研究主要使用了以下关键技术：  4.1、轻量化特征网络  目标检测算法使用分类网络作为特征提取网络，本课题采用更为轻量化的特征提取网络，减少计算参数，提高检测速度。  4.2、自适应 NMS 算法  交通目标较为密集，存在遮挡情况，引入自适应 NMS 算法，在交并比高于阈值时，自动调整阈值，减少自动忽略后边被遮挡物体的情况。  4.3、混合注意力机制  传统的空间注意力机制计算复杂，每个通道的注意力权重相同，形式固定，注意力单一，使用混合注意力模型可结合空间注意力与通道注意力机制，使得网络的空间注意力丰富，有利于网络的特征提取，对网络的训练有着正向促进作用。  4.4、YOLOv5与跟踪框架相结合  使用改进后的YOLO v5目标检测算法作为跟踪算法中的检测器，提出新的遮挡判别机制和轨迹修正策略，改善遮挡目标的检测与跟踪精度。  **5、可行性分析**  5.1、研究问题的可行性  本课题所研究的目标跟踪算法是国内外的研究热点，并且国内外对目标跟踪算法研究具有一定的基础，本课题基于这些基础对模型进行改进和创新，并对模型的基本理论作以解释，所以本课题所研究的问题具有可行性。  5.2、研究方案的可行性  本课题的研究方案是对研究内容进行推导，一步一步的构成整个研究，在层次和结构上循序渐进，不存在跳跃研究，并且每个研究步骤都可达，不存在不能完成的情况，所以本课题的研究方案具有可行性。  5.3、研究技术的可行性  本课题使用深度学习作为研究框架，使用YOLOv5算法作为基本模型对目标检测与跟踪算法进行研究，研究的主要技术都是整个领域的基础，研究的创新和改进部分也是基于这些基础在思想上进行创新，在技术上进行改进，所以本课题的研究技术具有可行性。  5.4、研究环境的可行性  本课题研究者所在实验室拥有一台GPU计算设备，保障了本课题实验研究的进行，所以本课题具有研究环境的可行性。   1. **预期研究成果**   1、将轻量化特征提取网络应用到目标检测当中，大幅度减少计算参数，提高检测速度；引入混合注意力机制，提高对小目标的检测精度；提出一种NMS, 减少自动忽略后边被遮挡物体的情况。  2、将改进的目标检测算法与目标跟踪框架相结合，提高跟踪精度和速度；针对跟踪中出现的遮挡问题，提出一种遮挡判别机制和轨迹修正策略，建立健壮的目标跟踪算法。  3、将目标跟踪算法应用在车辆检测与跟踪系统。实现对交通车辆检测，并根据需求构建车速检测模块和前车安全距离检测等，构建前端界面，实时显示当前数据，将检测和跟踪的结果保存到数据库中，实现历史车辆统计信息的查询，构建完整车辆检测与测距系统。  **参考文献**  [1]马社强,丁立民,刘东,赵丹.我国道路交通安全状况及挑战[J].中国人民公安大学学报(自然科学版),2020,26(04):35-41.  [2]陆化普,李瑞敏.城市智能交通系统的发展现状与趋势[J].工程研究-跨学科视野中的工程,2014,6(01):6-19.  [3] 臧利林,贾磊,秦伟刚,张立东.基于环形线圈车辆检测系统的研究与设计[J].仪器仪表学报,2004(S1):329-331.  [4] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." Advances in neural information processing systems 25 (2012).  [5]Liu G, Wang C. A novel multi-scale feature fusion method for region proposal network in fast object detection[J]. International Journal of Data Warehousing and Mining (IJDWM), 2020, 16(3): 132-145.  [6] Girshick R, Donahue J, Darrell T, et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation[C]. Computer Vision and Pattern Recognition, 2014:  580-587.  [7] Girshick R. Fast R-CNN[C]. International Conference on Computer Vision, 2015: 1440-1448.  [8] He K, Zhang X, Ren S, et al. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition[J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2014,  37(9):1904-16.  [9] Redmon J, Divvala S K, Girshick R, et al. You Only Look Once: Unified, Real-Time Object Detection[C]. Computer Vision and Pattern Recognition, 2016: 779-788.  [10] Liu W, Anguelov D, Erhan D, et al. SSD: Single Shot MultiBox Detector[C]. European Conference on Computer Vision, 2016: 21-37.  [11] Lin T Y, Goyal P, Girshick R, et al. Focal Loss for Dense Object Detection[C]//Proceedings of the IEEE International Conference on Computer Vision. 2017: 2980-2988.  [12] 张瑞平. 雾天条件下图像的恢复研究[D]. 天津:天津大学,2007.  [13] Tan R T. Visibility in bad weather from a single image[C]// 2008 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR 2008), 24-26 June 2008, Anchorage, Alaska, USA. IEEE, 2008.  [14] Kratz L, Nishino K. Factorizing scene albedo and depth from a single foggy image[C]//Proceedings of the 12th IEEE International Conference on Computer Vision(CVPR)，2009．Kyoto：IEEE，2009．1701-1708．  [15] Ding M，Tong R F．Efficient dark channel based image dehazing using quadtrees[J]．Science  China Information Sciences，2012，56(9)：1-9．  [16] Fang F，Li F，Yang X M，et a1．Single image dehazing and denoising with variational method[c]//Proceedings of the 2010 International Conference on Image Analysis and Signal Processing(IASP)，Xiamen，China：IEEE，201 0：219-222．  [17] Mc C E H L A J J. Lightness and retinex theory[J]. Journal of the Optical Society of America,  1971, (1): 1-11.  [18] DJ J, Z R, GA. W. Properties and performance of a center/surround retinex[J]. IEEE Transactions on Image Process, 1997: 451-462.  [19] Kimmel R, Elad M, Shaked D, et al. A Variational Framework for Retinex[J]. International Journal of Computer Vision, 2003, 52(1):7-23.  [20] 张洪坤, 薛模根, 周浦城. 基于非线性扩散均值漂移的 Retinex 雾天图像清晰化算法[J].图学学报, 2013, 34(2):47-51.  [21] 李菊霞, 余雪丽. 雾天条件下的多尺度 Retinex 图像增强算法[J]. 计算机科学, 2013, 40(3): 299-301.  [22]Salari E, Li M, Ouyang D. Foggy image enhancement based on Principal Component Analysis[A]. 2013 IEEE 56th International Midwest Symposium on Circuits and Systems (MWSCAS)[C]. Columbus, OH: IEEE, 2013: 1259-1262.  [23] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]. 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05). IEEE, 2005, 1: 886-893  [24]Lindeberg, Tony. "Scale invariant feature transform." (2012): 10491.  [25] Bay H, Tuytelaars T, Van Gool L. Surf: Speeded up Robust Features[C]//European Conference on Computer Vision. Springer, Berlin, Heidelberg, 2006: 404-417.  [26] Suykens J A K, Vandewalle J. Least squares support vector machine classifiers[J]. Neural processing letters, 1999, 9(3): 293-300.  [27] Zhu J, Zou H, Rosset S, et al. Multi-class Ada Boost[J]. Statistics & Its Interface, 2006, 2(3):349-360.  [28] Paul Viola, Michael J. Jones. Robust Real-Time Face Detection[J]. International Journal of Computer Vision, 2004, 57(2):137-154.  [29] Uijlings J R R,Van De Sande K E A,Gevers T,et al.Selective search for object recognition[J]. International Journal of Computer Vision,2015,115(3):211-252.  [30] Redmon J, Farhadi A. YOLO9000: Better, Faster, Stronger[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition(CVPR), Honolulu, HI, USA, 2017 : 6517-6525.  [31] Farhadi A, Redmon J. Yolov3: An incremental improvement[J]. Computer Vision and Pattern Recognition, cite as, 2018.  [32] Wu D, Lv S, Jiang M, et al. Using channel pruning-based YOLO v4 deeplearning algorithm for the real-time and accurate detection of apple flowers in natural environments[J]. Computers and Electronics in Agriculture, 2020, 178:105742.  [33] Liu Shu ， Qi Lu ， Qin Haifang ， Shi Jianping,et al. Path aggregation network for instance segmentation[J]. Computer Vision and Pattern Recognition, 2018:8759-8768.  [34] MA C, HUANG J-B, YANG X, et al. Hierarchical convolutional features for visual tracking[C]. IEEE International Conference on Computer Vision (ICCV), 2015:3074-3082.  [35] K. Simonyan, A. Zisserman. Very deep convolutional networks for large-scale image recognition [J]. arXiv preprint arXiv:1409.1556, 2014.  [36] M. Danelljan, G. Häger, F. S. Khan, et al. Convolutional features for correlation filter based  visual tracking [C]. IEEE International Conference on Computer Vision Workshops, 2015: 58-66.  [37]NAM H, HAN B. Learning multi-domain convolutional neural networks for visual tracking[C]. Internaltional Conference on Computer Vision and Pattern Recogintion (CVPR), 2016:4293-4302.  [38] Held D, Thrun S, Savarese S. Learning to Track at 100 FPS with Deep Regression Networks[C].European Conference on Computer Vision, Springer, Cham, 2016.  [39] BERTINETTO L, VALMADRE J, HENRIQUES J F, et al. Fully-convolutional siamese networks for object tracking[C]. European Conference on Computer Vision (ECCV), 2016:850-865.  [40] VALMADRE J, BERTINETTO L, HENRIQUES J, et al. End­to­end representation learn­ing for correlation filter based tracking[C] // IEEE Conference on Computer Vision andPattern Recognition (CVPR). 2017 : 5000 – 5008. | |

二、**论文工作实施计划**

|  |  |  |
| --- | --- | --- |
| **论文工作的进度与安排** | | |
| **起止时间** | **进度安排** | **备注** |
| **2022.3-2022.4** | **确定论文的方向和资料收集** |  |
| **2022.5-2022.6** | **完成开题报告和论文提纲** |  |
| **2022.7-2023.1** | **完成论文初稿** |  |
| **2023.1-2023.3** | **论文修改** |  |
| **2023.3-2023.4** | **完成论文** |  |
| **2023.4-2023.5** | **论文送审并准备论文答辩** |  |
|  |  |  |
|  |  |  |

1. **指导教师综合意见**

|  |
| --- |
| **对报告人论文的选题价值、文献综述、研究或开发设计、可行性等的综合意见**  该生对选题与老师进行了充分讨论，符合计算机技术专业毕业论文要求，参考了许多文献，具有一定的实用价值。本选题是学生所学专业知识的延续，符合学生专业发展方向，对于提高学生的基本知识和技能，对于提高学生的研究能力有益。研究方法和研究计划基本合理，难度合适，学生能够在预定时间内完成该论文的设计。  指导教师：  年 月 日 |

**四、开题报告论证记录**

|  |
| --- |
| **包括时间、地点、参加人、建议、提问及回答要点等**  时间：2022年7月7日  地点：学院楼409会议室  参加人：陈勇（组长）、牛彦敏、王艳霞  问题1：当前研究进度如何，做了哪方面的工作？  回答：本学期开始进行了毕业论文准备，阅读相关文献，并做了部分实验，实验结果表明效果有提升。  问题2：你的数据集从哪里来？  回答：通过摄像机拍摄交通车辆采集自制数据集，并使用部分公开数据集与自制数据集混合，保证数据量满足训练要求。  建议1：部分参考文献年代久远，建议参考近几年的文献。  建议2：该应用背景已有较多人做了相关研究，建议优化应用背景，落实到具体场景上去。  记录（签名）：  年 月 日 |

|  |
| --- |
| **对报告人确定的论题的选题价值、可行性、是否同意撰写论文等做出结论**  主持人（签名）：  年 月 日 |
| **院学位评定分委员会意见**  主席（签名）：  年 月 日 |
| **校学位评定委员会办公室审核意见**  盖章：  年 月 日 |

**五、报告会结论**

**六、计划变动情况（无变动此页不填）**

|  |
| --- |
| **计划变动原因：**  计划变动申请人（签字）：    年 月 日 |
| **导师意见**    导师（签字）：  年 月 日 |
| **院学位评定分委员会意见**    主席（签字）：  年 月 日 |
| **校学位评定委员会办公室意见**    盖章：  年 月 日 |