分类号：TP391　　学校代码：10363
密　级：公开　　学　　号：2220320160

安徽工程大学
Anhui Polytechnic University

硕士学位论文

题目　交通道路场景下的目标检测方法研究

论 文 作 者　郑泊文
校 内 导 师　陆华才
校 外 导 师　汪泽银
专业学位类别　电子信息
研究方向（领域）　模式识别与智能系统

论文提交日期：2025 年 5 月 10 日

分类号：TP391　　单位代码：10363

密　级：公 开　　学　　号：2220320160

题 目　交通道路场景下的目标检测方法研究

题 目　Investigation of Object Detection Techniques in Traffic Road Scenarios

学生姓名：郑泊文

校内导师：陆华才

校外导师：汪泽银

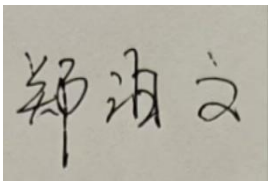
专　　业：电子信息

研究方向：模式识别与智能系统

论文答辩日期：2025 年 5 月 10 日

安徽工程大学学位论文原创性声明

本人郑重声明：我恪守学术道德，崇尚严谨学风。所呈交的学位论文，是本人在导师的指导下，独立进行研究工作所取得的成果。除文中已明确注明和引用的内容外，本论文不包含任何其他个人或集体已经发表或撰写过的作品及成果的内容。论文为本人亲自撰写，我对所写的内容负责，并完全意识到本声明的法律后果由本人承担。

学位论文作者签名： 

日期：2025 年 5 月 10 日

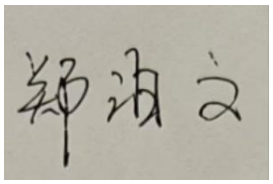
安徽工程大学学位论文版权使用授权书

学位论文作者完全了解学校有关保留、使用学位论文的规定，同意学校保留并向国家有关部门或机构送交论文的复印件和电子版，允许论文被查阅或借阅。本人授权安徽工程大学可以将本学位论文的全部或部分内容编入有关数据库进行检索，可以采用影印、缩印或扫描等复制手段保存和汇编本学位论文。

保密 □，在 __ 年解密后适用本版权书。

本学位论文属于

不保密 √。

学位论文作者签名：  指导教师签名： 

日期：2025 年 5 月 10 日　　日期：2025 年 5 月 10 日

交通道路场景下的目标检测方法研究

摘 要

计算机视觉已经成为社会发展和科学进步的一个重要组成部分，近年来由于卷积神经网络的快速发展，使得计算机视觉开始融入我们的日常生活，并对我们的生活产生了极为深远的影响。在交通道路场景下对目标的检测已然成为了热点研究方向之一，同时也是促进城市化快速发展的重要环节。

针对交通道路场景中目标的检测，被检测目标和环境的不同都会对检测结果产生影响。本文主要针对交通道路场景下目标检测与定位方法问题进行研究，通过不同方法和思路对目标检测算法进行改进，从而提升检测的准确性。本文主要的研究工作如下：

（1）对交通道路场景下目标检测的发展历史进行系统性的概述和总结。从理论部分对交通道路场景的目标检测的现状及所面临的挑战进行分析概述。

（2）对交通道路场景中目标识别进行理论分析。详细介绍了卷积神经网络的组成部分，并对现如今常见的 One-stage 目标检测算法和 Two-stage 目标检测算法进行对比分析，接着介绍了目标检测算法的评价指标，并对本人所使用的数据集与一些常见的交通道路数据集进行对比，并阐述了本文选用 YOLOv7 目标检测算法的原因及优势。

（3）针对交通道路场景中的常见目标进行检测，提出了 CTD-YOLOv7 算法（Common Target Detection-YOLOv7）。首先介绍了本文所使用的 YOLOv7 目标检测算法的模型，接着使用 SENet 模块实现特征图通道信息的关联，使得模型对各个通道的特征具有更强的辨别能力。然后采用NAM注意力机制，利用权重的贡献因子来改善注意力机制的性能。最后介绍了本文所使用的实验设备及实验环境。

（4）针对夜间环境中的目标进行检测，在 CTD-YOLOv7 的基础上，提出了 NTD-YOLOv7 算法（Night Target Detection-YOLOv7）。为了使算法在夜间具有较好的检测效果，本文首先对图像进行预处理，使用基于拉普拉斯算子的锐化方法对图像进行处理，通过对比局部灰度来判断暗色背景区域，对局部的暗色背景区域进行直方均衡化和 Gamma 变换处理，从而提高图像的清晰度。在已

改进的基础上引入 GSConv 模块来减少训练过程中的参数量和计算量，接着选择 Shufflenetv2×1.5 作为特征提取网络，使网络整体结构在减少参数量的同时能具有较高的跟踪精度，再采用 Hard-Swish 激活函数来大幅度降低延迟成本，然后用 SIoU 函数替换 EIoU 函数，从而优化损失函数，提高鲁棒性和泛化能力。

（5）针对雨雪雾等恶劣环境中的目标进行检测，在 NTD-YOLOv7 的基础上，提出了 AWTD-YOLOv7 算法（Adverse Weather Target Detection-YOLOv7）。通过使用暗通道先验算法和基于引导滤波的操作实现去雾和去雨雪的目的。然后，在已经实现的去雨雪雾的基础上，通过使用 DIP 模块和 CNN-PP 模块的结合提高在雨雪雾环境下目标检测准确度。

最后的实验结果表明，本文改进的 YOLOv7 算法对不同环境下的交通道路场景均具有较好的目标检测效果。

关键词：交通道路；目标检测；神经网络；交通环境；YOLOv7

Investigation of Object Detection Techniques in Traffic Road Scenarios

Abstract

Computer vision has emerged as a critical component of social development and scientific advancement. In recent years, the rapid evolution of convolutional neural networks has facilitated the integration of computer vision into daily life, profoundly impacting various aspects of human existence. Target detection in traffic road scenarios has become a prominent research focus, playing a pivotal role in advancing urbanization.

For target detection in traffic road scenarios, variations in targets and environmental conditions can significantly influence detection outcomes. This paper investigates target detection and localization methods in traffic road scenarios, proposing enhancements to the target detection algorithm through diverse approaches and methodologies to improve detection accuracy. The primary research contributions are outlined as follows:

(1) A systematic review and summary of the historical development of target detection in traffic road scenarios is provided. From a theoretical perspective, the current state, and challenges of target detection in traffic road scenarios are analyzed and synthesized.

(2) A theoretical analysis of target recognition in traffic road scenarios is conducted. The components of convolutional neural networks are elaborated upon, and common One-stage and Two-stage object detection algorithms are compared and analyzed. Subsequently, evaluation metrics for object detection algorithms are introduced, and the dataset utilized in this study is contrasted with other prevalent traffic road datasets. The rationale and advantages of selecting the YOLOv7 object detection algorithm are expounded.

(3) Common target detection in traffic road scenarios is addressed and proposed the CTD-YOLOv7(Common Target Detection-YOLOv7). The YOLOv7 object

detection model employed in this paper is introduced. The SENet module is utilized to capture channel-wise correlations in feature maps, enhancing the model's discriminative capabilities for each channel's features. Furthermore, the NAM attention mechanism is adopted to improve attention performance by incorporating weight contribution factors. Finally, the experimental equipment and environment utilized in this study are detailed.

(4) Detect targets in nighttime environments and proposed the NTD-YOLOv7(Night Target Detection-YOLOv7) on the basic of CTD-YOLOv7. To enhance the detection performance of the algorithm at night, this paper proposes a preprocessing method that includes image sharpening using the Laplace operator, determination of dark background areas by comparing local gray levels, and square equalization as well as Gamma transformation for local dark regions to improve image clarity. Based on these improvements, the GSConv module is integrated into the architecture to reduce the number of parameters and computational cost during training. Shufflenetv2×1.5 is selected as the feature extraction network to ensure high tracking accuracy while minimizing the parameter count. The Hard-Swish activation function is employed to significantly reduce latency costs. Furthermore, the EIoU loss function is replaced with the SIoU function to optimize the loss function and enhance robustness and generalization capabilities.

(5) Target detection under adverse weather conditions, such as rain, snow, and fog, is investigated and proposed the AWTD-YOLOv7(Adverse Weather Target Detection-YOLOv7) on the basic of NTD-YOLOv7. By employing the dark channel prior algorithm and guided filter-based operations, rain, snow, and fog removal is achieved. Subsequently, the combination of the DIP module and CNN-PP module improves the accuracy of target detection in rainy, snowy, and foggy environments following the removal of precipitation and fog.

The final test results demonstrate that the enhanced YOLOv7 algorithm presented in this paper achieves excellent target detection performance across traffic road scenes in various environments.

Keywords：Traffic route; Target detection; Neural network; Traffic environment; YOLOv7

目 录

第1章 绪 论

1.1 研究背景及意义

随着科学技术的不断发展，人们的生活质量也在不断提高，出行方式也在多样化发展。根据中国汽车工业协会在 2023 年发布的数据显示，我国汽车产业的销量有着小幅增长，但是随着出行车辆数量的增加，交通事故的发生率也在增高。当今由于司机的各种不当行为，包括酗酒、超速、过度疲惫、缺乏安全意识以及不当的操作，导致全球范围内发生的交通事件数量正在逐年增高。据统计，全球每年的交通事故有 60%都发生在晚上，夜间事故的死亡数更是占交通事故死亡总数的 50%，这也使得在夜间进行驾驶存在更大的危险性，光照的不足会使得司机们的能见范围受到限制，加大了司机们的驾车难度，而且司机们在夜间驾驶车辆通常没有很好的状态，这也会使司机们的疲惫状况更加严重，进而会增大驾驶车辆的危险性。根据公安部道路交通事故统计报告，我国每年大约有 10%左右的交通事故直接与雨雪雾等恶劣天气有关。研究表明，当雾天能见度降低到 500 米以下时，会对公路交通产生影响，当雾天能见度降低到 200 米以下时，会对公路交通产生显著影响，当雾天能见度降低到 50 米以下时，会对公路交通产生十分严重影响。而在雨雪环境条件下，路面摩擦系数、驾驶员可视距离等指标明显降低，且雨雪天气下的路面摩擦系数最低不足 0.2，加上不良天气容易造成驾驶员反应时间延长、心理紧张，发生事故的概率则会相应增加。据我国在 2023 年国民经济和社会发展的统计公报显示，在道路交通事故中万车死亡人数为 1.38 人，因此降低交通道路事故的发生率和死亡率十分重要，为了解决这一问题，我国除了提高驾驶员的素质要求和道路安全准则外，还需要在出行方式上解决问题，其中就包括智能驾驶系统的出现，智能驾驶系统通过融合车辆识别[1-5]、行人识别、交通标志识别、车牌识别和驾驶员识别等技术，可以很好的提高汽车安全性能，进而减少交通事故的发生。

具有智能驾驶系统的汽车逐渐成为全球各大汽车厂商的重要研究方向，其中就包括提高驾驶安全性、降低驾驶复杂度和提升驾驶效率等，这就需要汽车对交通道路的周围环境有很强的感知能力。环境感知技术不仅是汽车智能化的基础，而且也是汽车实现高精度自动驾驶的关键，因此基于交通道路环境感知

技术的检测与跟踪对提高汽车安全行驶具有重要作用。

随着科技的发展，利用视觉感知技术来检测物体的位置已成为当今汽车制造商的一项关键挑战，这种技术的应用范围从图片到高精度的数字化建模，再到最新的人工智能，都涉及到了对物体的检测、识别、跟踪、定位等多个维度，从而有效地帮助汽车完成操作。

传统目标检测算法对图像的特征提取能力有限，因此很难保证实时性的实际要求，而且检测准确率也不高。但是随着计算机技术的不断进步，基于深度学习的目标识别技术取得了进步，它具有精度高、操作灵活、反应时间短、效果显著稳定等特征。目前基于交通道路的目标检测技术正在不断的快速发展，这也使得其具有很重要的研究价值和实际意义。

1.2 国内外研究现状

目标检测通过对输入的视频或图片找出其感兴趣的区域，并在这些感兴趣区域中预测目标的类别和位置，最终输出目标的边界框。传统目标检测算法与深度学习目标检测算法的区别在于其特征提取方式。传统目标检测算法常常使用手工设计的特征提取器来提取图像中的特征，通过从图像中提取出目标的形状、边缘、颜色等信息，并将其转换成能够用于分类器的向量形式。相较于其他目标检测技术，深度学习目标检测技术使用深层卷积神经网络[6]（Convolutional Neural Network, CNN）来捕捉和处理输入数据，从而实现从目标信息到模型的转换。

提取出来的特征对原始图像特征描述是否准确直接决定了后续分类器的判断。1996 年，Ojala 等[7]提出了局部二值法（Local Binary Pattern, LBP），这一技术可以用来捕捉和分析图像的细节，它利用灰度值的变化把图像转换成二维的表示形式，从而可以更好地捕捉和表达图片的细节。1999 年，David 等[8]提出了尺寸不变特征（Scale Invariant Feature Transform, SIFT），它的核心理念是通过识别图片上的重要元素，来确保它们的位置、大小、亮暗程度以及外界环境条件都保持稳定。2005 年，Dalal 等[9]提出了一种新的技术梯度直方图特征（Histogram of Oriented Gradients, HOG），它能够准确地捕捉到图片的细节，从而更加清晰地展示其内在的结构和功能。2006 年，Bay 等[10]提出了加速鲁棒特性（Speeded Up Robust Feature, SURF），通过计算图像的 Hessian 矩阵来检测关

键点，并对关键点周围的区域进行快速积分，进而来描述这些关键点。Cortes等[11]提出了支持向量机（Support Vector Machine, SVM），通过利用高维空间的特征将两个类分开，以便更好地理解和分析，通过使用多种核函数将数据从低维空间转换到高维空间，从而实现对两个物体之间的准确划分，其中包括一条直线和一个弯曲的形状，而这两个形状之间的差距也必须尽量小，从而实现对复杂的非线性系统的优化。Jin 等[12]针对 SVM 分类器在正负样本比例失调场景下的分类性能衰减问题，创新性的引入欠采样技术和 Easy Ensemble 集成学习方法，该方法通过构建平衡化的训练数据子集，有效地缓解了因行人目标与背景区域数量级差异导致的模型偏倚现象。Freund 等[13]提出的 AdaBoost 算法通过迭代调整样本权重与分类器组合策略，为集成学习领域树立了重要示范。受到 Freund 的启发，文学志等[14]在车辆识别任务中实现了双重创新，一方面构建了基于 Haar 特征的多尺度纹理表征体系，另一方面对 AdaBoost 框架进行特征选择机制的优化。2013 年，Felzenszwalb 等[15]提出了一种具有可变性的目标检测算法（Deformable Parts Model, DPM），它的核心理念在于把被观察的对象视为一系列的有机结构，通过构建一个具有相应特征的概率图来描述它们之间的关系。为了使模型能更好的适应不同类型的目标，DPM 的每个部件都可以独立旋转缩放，采用滑动窗口的方式来实现窗口在图像中的移动，通过计算各个窗口与模型之间的匹配程度，进而得到目标物体在图像各个位置中存在的概率，这也使得 DPM 可以通过自主学习达到目标识别的目的。

在 2007 年到 2009 年，虽然目标识别算法刚开始发展，但 Felzenszwalb 仍凭借着 DPM 优越的性能连续三次获得了 PASCAL VOC 挑战赛的冠军。虽然 DPM 作为传统目标检测算法的开山作之一，其也是存在很多的缺点，例如其相对简单的特征表示方法会导致其无法很好的扫描具有复杂纹理和形状信息的目标。除此之外，DPM 也很容易受到各种外界因素的影响，计算量巨大也是其检测速度慢的原因之一。

1.3 交通道路技术的挑战与前景

随着目标检测技术的不断发展，交通道路的目标检测方法取得了显著的进步，但是要实现更高精度的交通道路目标检测，还需要更多的努力和投入。目标检测技术将会在未来的时间里受到更多的关注，并朝着更高的水平发展。目

标检测方法未来研究与发展的重点和趋势如下：

（1）随着科技的进步，基于视觉的目标检测方法正在被更先进的技术所取代，这些技术不仅可以节省计算资源，还可以提高识别速度。为了满足这一需求，越来越多的轻量级车载应用正在开发出来，它们可以在移动设备上安装，从而更便捷地实现目标检测。

（2）随着现代技术的不断改善，许多新型的目标检测方法正在朝着将 3 种或更多种传感器有机地整合在一起的方向前进。这样不仅可以收集到大量的目标信息，而且还能够精细地捕捉到目标的各种特征，从而提供更加精确和可靠的检测结果。

（3）近年来越来越多的人开始关注于探索不同类型的目标，如远程、微型、大型、单目标和多目标等。尤其是在处理数据量有限的情况下，探索这些类型的目标变得更加困难。因此未来的技术将会更加重视通过改善采集与输入的数据，并结合各种传感技术来收集更多的数据。

（4）目前研究的交通道路检测技术已经不仅限于路况良好的城市道路，而且还可以应用于复杂的环境，如多传感器融合，以解决传感器在复杂环境下可能出现的故障和失效问题。

1.4 本文研究内容及结构安排

本文针对交通道路场景下的目标检测技术进行研究分析，包括其研究背景与实用意义，并以 YOLOv7 目标检测算法为基础进行改进，从而进一步提升检测精度和检测速度。

第 1 章：绪论。介绍了当前环境下的交通道路检测技术，并对其背景和研究意义进行了简单分析，进而决定本文所使用的目标检测算法，并在此基础上对算法进行改进，从而实现道路目标检测技术的高精度、快速度。

第 2 章：面向交通道路的目标识别的理论分析。介绍了当前交通环境下的交通道路技术和传统的卷积神经网络的基础理论知识，并由此对常见的 One-stage 算法和 Two-stage 算法进行了对比。接着介绍了目标检测过程中对于结果的评价指标和 YOLO 系列算法，并在此基础上解释选用 YOLOv7 算法的原因。最后介绍了本文所使用的数据集及其处理过程。

第 3 章：针对交通道路常见目标的检测。简单介绍了 YOLOv7 的网络模型

结构。针对当下交通环境对行人与车辆检测的重要性进行了概述，采用NAM算法，在每个通道中都加入相应的权重，避免与 SE 和 CBAM 一样在每个通道中都添加完整的连接层或者卷积层。

第 4 章：针对夜间环境下的目标检测研究。首先介绍了在解决在夜间等特殊天气情况下对交通道路进行检测的重要性，接着在已改进的基础上对图像进行预处理操作，使用 ShuffleNetv2×1.5 网络结构，使用通道混洗的方式可以将不同通道的输入信息均匀打乱，以此解决不同群卷积之间所导致的信息不交互问题，最后还对损失函数和激活函数进行了优化改进。

第 5 章：针对雨雪雾环境下的目标检测研究。介绍了针对雨雪雾环境下目标检测的原理，接着通过分别使用暗通道先验算法和基于引导滤波的操作实现去雾和去雨雪的目的，接着在已经实现的去雨雪雾的基础上，通过使用 DIP 模块和 CNN-PP 模块的结合提高在雨雪雾环境下目标检测准确度。

第 6 章：结论与展望。对本文改进的算法进行总结并针对算法的部分不完美之处提出未来展望。

第2章 交通道路环境下目标检测的理论分析

2.1 卷积神经网络

卷积神经网络最初的作用是用来处理图像信息，但随着图像处理技术的不断发展，神经网络的局部连接和权重共享变得越来越重要，这也使得卷积神经网络在图像处理领域具有越来越重要的地位。和传统的目标检测算法相比，基于深度学习的卷积神经网络具有更强的缺陷检测方法，并通过端到端的检测，提高对图像特征的提取能力和目标检测的准确性。卷积神经网络的结构如图 2-1 所示。

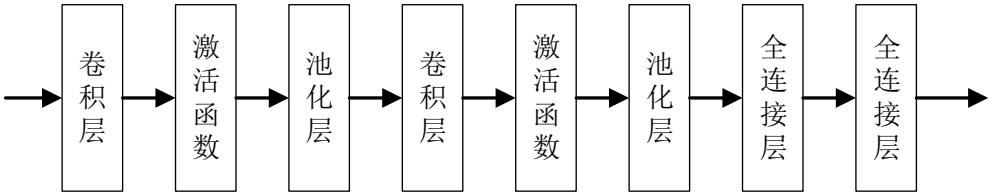


图 2-1 卷积神经网络结构图

首先我们需要把收集的样本数据转换成多种卷积神经网络的形态，这些卷积神经网络的形态就包括卷积、激活和池化。通过使用这些形态来构建多种新的模型，从而对卷积神经网络的形态进行切换，根据每种形态的特点来提炼出有用的信息，接着将这些特征信息传输到全连接层进行分类输出。

2.1.1 卷积层

卷积层作为卷积神经网络中非常重要的一个部分，由很多卷积核和偏置构成，其主要作用是通过卷积运算提取出目标图像的特征。在传统的目标检测算法中，卷积核又被称为滤波器，其权值也是固定不变的，一般会采用边缘算子、梯度算子或是拉普拉斯算子作为其权值算子，常用来对图像进行去噪和特征提取。但是在卷积神经网络中，卷积核的权值并不是固定不变的，这是因为多个卷积层进行串联或者并联后会产生参数更新，一张图片会同时存在很多个卷积核，卷积核的不同也会导致图像特征提取的效果不同。为了让卷积神经网络能够从海量数据中提取出有价值的信息，每个卷积核在执行卷积操作时，应当将其权重分配给所有的参与者，以实现更加丰富的特征图。卷积计算如公式（2-1）所示：

$$P(i,j)=\sum_{a=0}^{m-1}\sum_{b=0}^{n-1}Q(a,b)f(i-a,j-b) \tag{2-1}$$

其中，P表示卷积结果，Q表示卷积核，其大小为(m,n)，f表示被卷积矩阵。

卷积神经网络的卷积过程如图 2-2 所示。

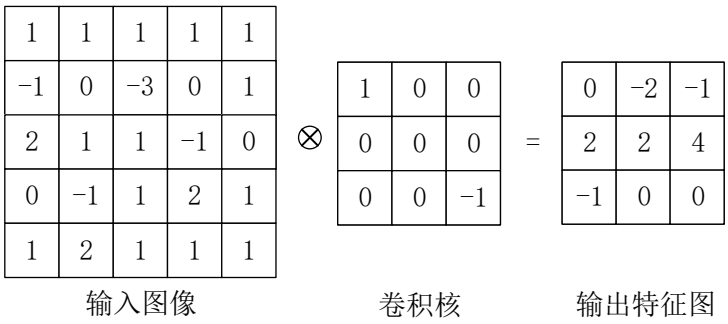


图 2-2 卷积神经网络

其中，⊗表示卷积操作。

卷积神经网络使用卷积核对相邻的神经元进行局部连接，通过这种操作可以使每个区域内的像素产生关联，进而减少网络模型的参数量，以此达到提高网络模型精确度的目的。

2.1.2 池化层

卷积层通过对输入图像进行卷积操作得到图像特征后，会将得到的特征数据传输到池化层进行下采样操作，这样可以在保留图像重要信息的同时减少参数和计算量，这不仅可以抑制过拟合，而且可以很好的提高运算速度。池化的原理就是对于输出的每一个通道的特征图的所有像素计算的一个平均值，常见的池化操作有最大池化和平均池化。池化操作的过程如图 2-3 所示。

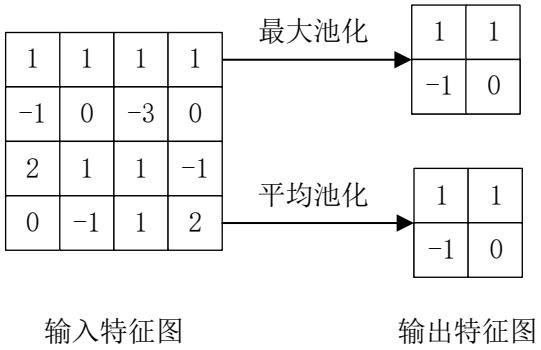


图 2-3 池化操作

其中，最大池化是取出输入特征图中每个区域的最大值，并将其作为池化后的结果，而平均池化则是对输入特征图进行区域划分，计算每个区域的平均值作

为结果。在实际的实验应用中，由于在反向传播的过程中，梯度特征会以全局形式的方式被平均分配到各个位置，因此平均池化的效果并没有最大池化好。

2.1.3 全连接层

全连接层是卷积神经网络模型最后的操作，其结构图如图 2-4 所示。

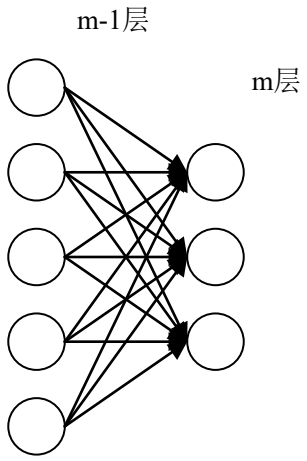


图 2-4 全连接层

卷积神经网络的卷积层通常被用于处理局部视野，它从输入特征图中提取出相关的信息，经过池化层的下采样操作，将结果传递给全连接层，全连接层则通过每一层的神经元与上一层的神经元建立完整的关系，从而使得不仅能够综合上一层的信息，而且还能够将这些信息进行分类，最终利用线性加权的方式来计算求和，从而实现对信息的有效处理。全连接层的计算如公式（2-2）所示：

$$Wx + b = z \quad \begin{cases} x = [x_0, x_1, \cdots, x_n]^T \\ b = [b_0, b_1, \cdots, b_m] \\ z = [z_0, z_1, \cdots, z_m] \end{cases} \tag{2-2}$$

其中，w表示$m \times n$的权重矩阵，x表示样本数据的特征，b表示偏置，z表示全连接层的输出。

2.1.4 激活函数

在处理多层卷积神经网络模型中，如何处理层与层之间的关系就显得尤为重要。为了保证上层输出与下层输入之间不是线性关系，Glorot 等[16]提出了激活函数，使得神经网络的层与层之间存在一个隐藏层，进而提高网络的检测能力和深层网络的特征表达能力。激活函数的计算流程如图 2-5 所示。

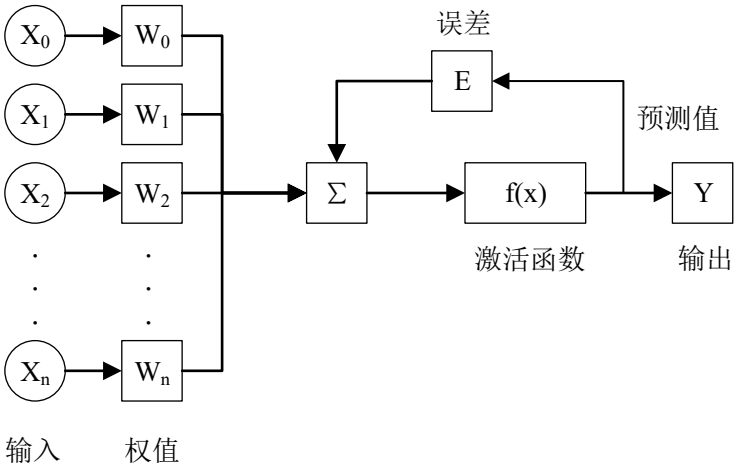


图 2-5 激活函数计算流程

激活函数作为一种解决卷积神经网络复杂非线性特征的重要工具，它可以把一个节点的输入变成一个可被接受的结果，从而实现对复杂结构的快速响应和优化。常见的激活函数有以下几种。

（1）Sigmoid 函数

Sigmoid 函数是非线性的，其形状为 S 型的曲线，常被用于处理二分类问题，取值范围为 0 到 1，当取值为 0 时，其表示为特别大的负数，当取值为 1 时，则表示为特别大的正数。Sigmoid 函数图像在 x-y 如图 2-6 所示。

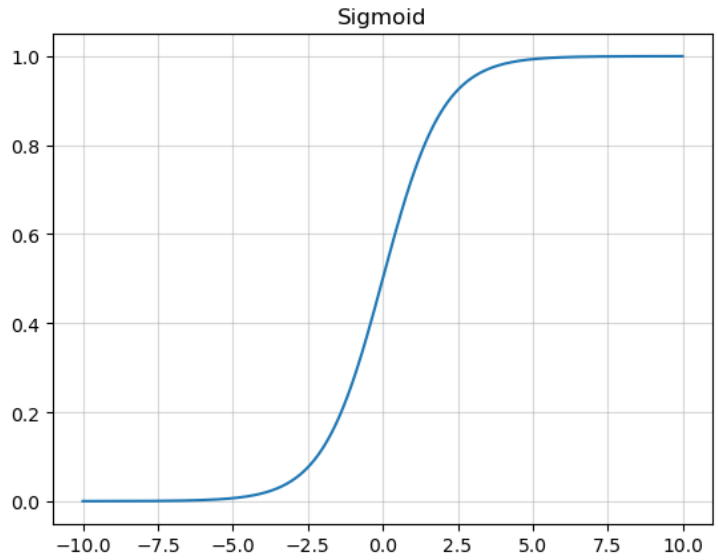


图 2-6 Sigmoid 函数

由图 2-6 中可以看出 Sigmoid 函数具有较好的梯度平滑性，且其可以快速地完成求导过程，这也使得其常被用于隐藏层。但是其存在的梯度弥散问题，会导致反向传播过程的训练效果降低。Sigmoid 的计算及其求导如公式（2-3）和

公式（2-4）所示：

$$Sigmoid(x) = \frac{1}{1+e^{-x}} \tag{2-3}$$

$$Sigmoid'(x) = \frac{e^{-x}}{(1+e^{-x})^2} = Sigmoid(x) \times [1 - Sigmoid(x)] \tag{2-4}$$

（2）Tanh 函数

Tanh 函数[17]是通过将 Sigmoid 函数向下平移和收缩得到的，和 Sigmoid 函数一样，Tanh 函数的输出范围为 0 到 1，且在输入值很大或很小的情况下，梯度几乎为 0。Tanh 函数图像如图 2-7 所示。

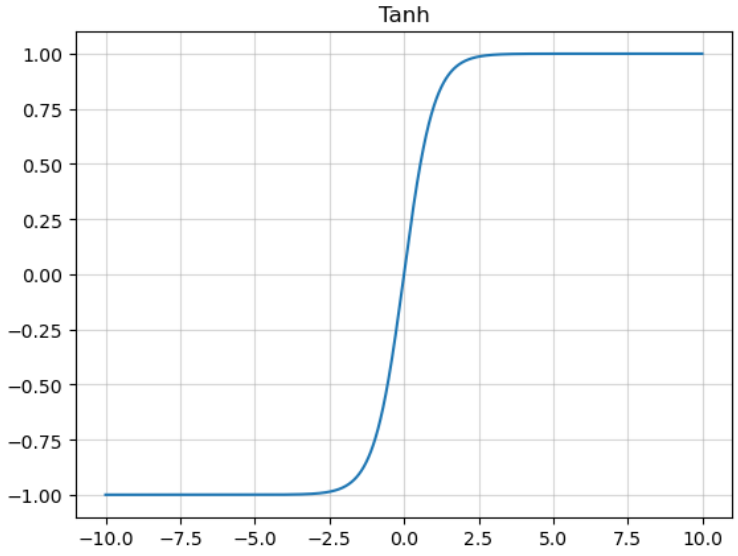


图 2-7 Tanh 函数

从图 2-7 中可以看出 Tanh 函数不仅保存了 Sigmoid 函数平滑性的特点，而且保证了输出具有正负的特点，这也使得其收敛速度快，不容易出现 loss 值震荡，但是仍未解决梯度弥散的问题。Tanh 函数的计算如公式（2-5）所示：

$$Tanh(x) = \frac{e^{x} - e^{-x}}{e^{x}+e^{-x}} \tag{2-5}$$

（3）ReLU 函数

为了解决梯度弥散的问题，Nair 等[18]提出了 ReLU 函数，其主要作用是避免当$x < 0$时会存在的神经元死亡情况，即更加有效的防止了梯度下降和反向传播的问题。除此之外，ReLU 函数不会受到其他激活函数中复杂指数函数的影响，并且可以使神经网络的计算成本更低、收敛速度更快。ReLU 函数图像如图 2-8 所示。

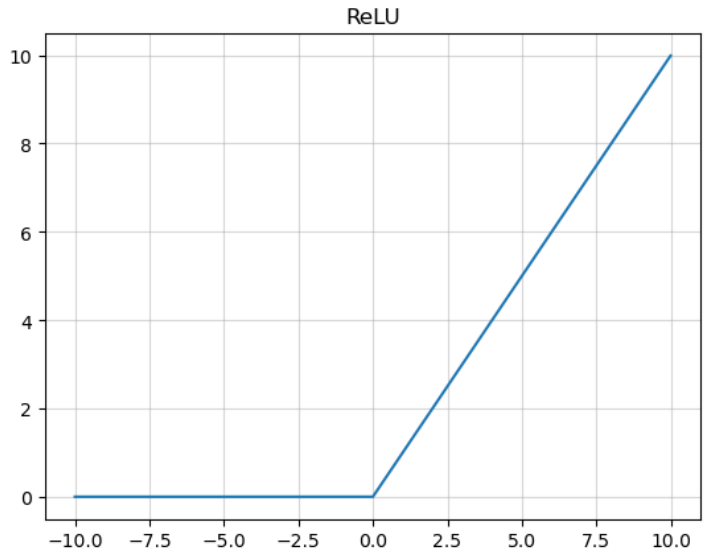


图 2-8 ReLU 函数

从图 2-8 中可以看出 ReLU 函数解决了梯度弥散问题，并且使得训练能快速收敛。当$x > 0$时，ReLU 函数的输入才能传播信号，也因此提高了网络的性能；但是当$x < 0$时，即使存在非常大的梯度传播，函数的输出也会停止。ReLU 函数的计算如公式（2-6）所示：

$$ReLU(x) = max\ (0, x) \qquad (2\text{-}6)$$

（4）LeakyReLU 函数

为了解决 RELU 函数在负输入情况下存在神经元死亡的问题，Naas 等[19]提出了 LeakyReLU 函数，通过手动训练的方法在负半轴增添一个很小的线性参数，解决 ReLU 负半轴硬饱和的问题。LeakyReLU 函数图像如图 2-9 所示。

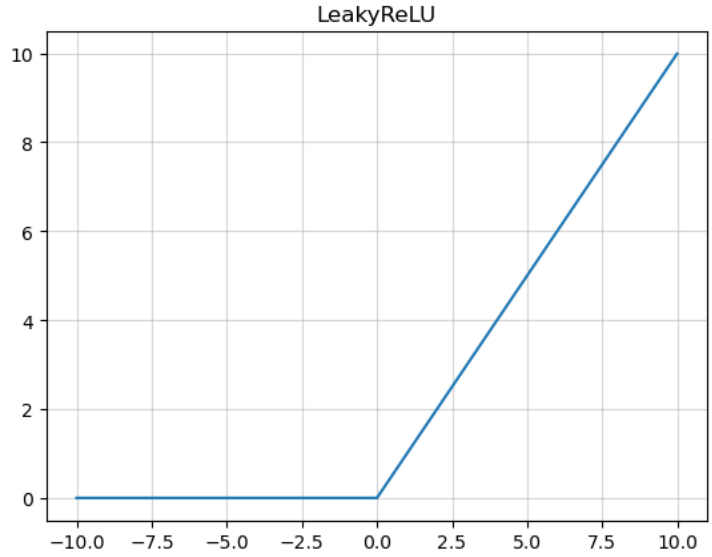


图 2-9 LeakyReLU 函数

其计算如公式（2-7）所示：

$$LeakyReLU(x) = \begin{cases} \alpha x, & x \leq 0 \\ x, & x > 0 \end{cases} \quad (2\text{-}7)$$

从图 2-9 和公式（2-7）中可以看出 LeakyReLU 函数在 x>0 时，其梯度值恒为 1；但当 x>0 时，该函数存在一个可手动设置数值的超参数$\alpha(\alpha \neq 0)$，这使得在 x<0 的情况下函数能够获得较小的梯度值，进而防止出现梯度弥散。

（5）Swish 函数

Swish 函数[20]是在 Sigmoid 函数的基础上进行改进得到的，通过输入简单的标量，就可以轻松替换以单个标量为输入的激活函数，避免更改隐藏容量或者参数数量，Swish 函数图像如图 2-10 所示。

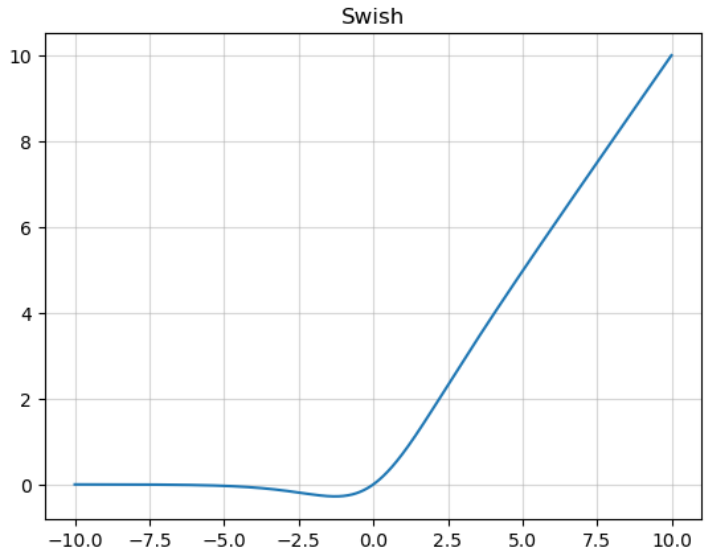


图 2-10 Swish 函数

从图 2-10 中可以看出 Swish 函数可以很好的防止在慢速训练期间梯度逐渐接近 0 并导致饱和的问题。但是由于其计算比较复杂，因此在一定程度上会影响模型的计算速度。除此之外，Swish 函数的参数 beta 需要进行学习，这也会导致模型需要更长的训练时间。

（6）Hard_Swish 函数

Hard_Swish 函数[21]解决了 Swish 函数计算量大的问题，该函数计算简单速度快等特点使其适用于嵌入式设备，其在保留 Swish 函数特点的同时，也修改了 Swish 函数中 Sigmoid 部分的饱和性，从而大大降低了梯度消失的可能性。Hard_Swish 函数图像如图 2-11 所示。

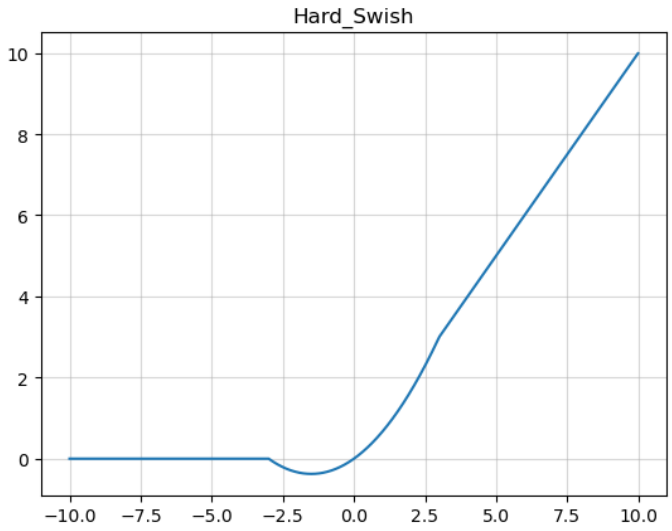


图 2-11 Hard_Swish 函数

从图 2-11 中可以看出 Hard_Swish 函数的形状和 Swish 函数十分接近。虽然 Hard_Swish 函数的性能和 Swish 函数没有明显差距，但是从部署的角度来看，Hard_Swish 函数采用分段的方法可以减少内存的访问量，从而大大降低延迟成本。Hard_Swish 函数的计算如公式（2-8）所示：

$$Hard\_Swish(x) = \begin{cases} 0, & x \le -3 \\ \dfrac{x^2 + 3x}{6}, & else \\ x, & x \ge 3 \end{cases} \tag{2-8}$$

为了更加形象的对比上述激活函数，因此将其放入图中进行对比，函数对比图如图 2-12 所示。

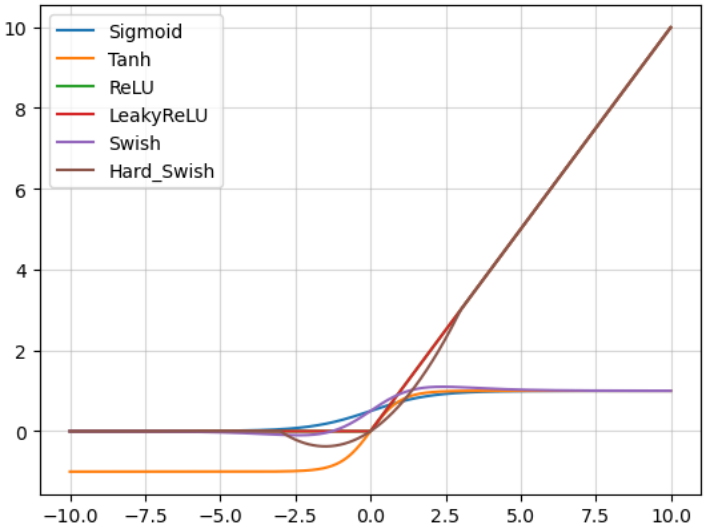


图 2-12 激活函数对比图

2.2 One-stage 目标检测算法

One-stage 目标检测算法的提出主要简化了候选框区域的过程，直接通过卷积特征网络进行特征提取和回归的过程，其流程如图 2-13 所示。

图 2-13 One-stage 流程图

下面是 One-stage 目标检测算法的详细设计思路：

（1）特征提取：首先使用一个预训练的卷积神经网络来提取图像的特征，这些特征可以通过网络的不同层级来获取，通常选择具有不同尺度和语义信息的特征图。

（2）Anchor生成：通过在特征图上的不同位置和尺度上生成一组候选框，用于表示可能包含目标的区域，这些 Anchors 通常是以不同宽高比和尺度进行定义的。

（3）目标分类和回归：通过分类器对每个 Anchor 判断其是否包含目标，并同时进行边界框的回归，分类器通常是一个多层感知机（Multilayer Perceptron, MLP）或卷积神经网络，用于预测目标类别的概率分布，回归器用于预测目标边界框的位置和大小。

（4）NMS 筛选：由于一个目标可能被多个 anchors 所覆盖，为了避免重复检测，需要进行非极大值抑制（Non Maximum Suppression, NMS）筛选，NMS 会根据预测的目标得分和重叠度来选择最佳的目标框。

（5）输出结果：最后根据分类器的预测结果和回归器的输出，得到最终的目标检测结果，通常会设置一个阈值来过滤低置信度的目标框。

常见的 One-stage 目标检测算法有 YOLO（You Only Look Once）系列算法和 SSD（Single Shot MultiBox Detector）算法。

2016 年，Liu 等[22]提出了使用单个深度神经网络检测图像中物体的 SSD 算法，该算法在预测时会将默认框中每个对象的类别进行生成分数，并调整检测框，便于更好的匹配图像中待检测对象的形状。同年，Redmon[23]提出了算法精度远超 R-CNN 的单阶段检测算法 YOLOv1，为工业方面的目标检测提供了技术支撑，但面对多目标密集或存在小目标的场景时，检测效果较差。在此基础上，Redmon[24]提出了 YOLOv2，在 YOLOv1 的基础上添加了预选框，使其计算量大

大降低，并且在一定程度上加强了对目标位置和大小的检测。为了提高关于小目标的检测效果，Redmon[25]又在 YOLOv2 的技术上提出了 YOLOv3，通过引入残差网络模块，使其不仅对小目标具有很好的检测效果，而且并没有牺牲检测速度和检测精度，但其对于密集目标的检测效果并不如人意。Bochkovskiy 等[26]在 2020 年提出了 YOLOv4，通过更换主干网络来增加感受野，通过牺牲训练时间来提高其检测精度，同时也降低了其泛化性。同年，Jocher[27]提出了 YOLOv5，相较于 YOLO 系列的前几个算法，YOLOv5 具有更快的检测速度、更好的检测效果和更稳定的鲁棒性。为了解决 YOLO 系列模型计算量的增加和正负采样不平衡的问题，Li 等[28]在 2022 年提出了 YOLOv6，该算法通过剔除锚框来降低模型计算量，以达到正负样本不平衡的问题，使得检测精度进一步提高。同年，Wang 等[29]提出了 YOLOv7，通过在输入阶段采用数据增强方法，使移动设备在保持检测速度的同时具有很好的检测精度，这也使得 YOLOv7 可以应用于交通道路领域。Gallagher[30]在 2023 年提出了 YOLOv8 网络，该网络有效地结合了高级特征和上下文信息以提高检测准确性，但在高度复杂或严重遮挡的场景中，其性能有所下降。Wang[31]在 2024 年提出了 YOLOv9，通过保持深度神经网络中的数据完整性和稳健梯度来防止数据退化。同年，清华大学提出了 YOLOv10[32]，它使用一个大型卷积核和一个部分自注意力模块，使模型能够在保持低计算成本的同时增强全局表示能力。随后，YOLOv11[33]亮相，提供卓越的性能和灵活性，有助于更准确、更高效地进行对象检测和跟踪任务。

YOLO 系列算法和 SSD 算法优缺点的对比如表 2-1 所示。

表 2-1 One-stage 常见算法的优缺点

目标检测算法	模型优点	模型缺点
SSD 算法	分层特征图预测不同尺度物体	模型对小物体检测准确率不高
YOLOv1 算法	单阶段目标检测模型新范式	模型难以预测密集目标和小物体
YOLOv2 算法	可预测多类物体	模型复杂度高且训练步骤多
YOLOv3 算法	采用独立的逻辑回归	模型训练时间常，泛化性差
YOLOv4 算法	结合多种策略方法	预测框误检率高
YOLOv5 算法	优化了主干网络	检测性能有所降低
YOLOv6 算法	采用混合通道策略	所采用的优化策略偏向小型网络
YOLOv7 算法	精度高，检测速度快	对小目标的检测效果不佳

续表 2-1

YOLOv8 算法	提高检测准确性	难以应对高度复杂的场景
YOLOv9 算法	防止数据退化	对设备的要求高
YOLOv10 算法	计算成本低，全局表示能力强	计算冗余，限制模型能力
YOLOv11 算法	更好的性能和灵活性	/

2.3 Two-stage 目标检测算法

Two-stage 主要通过一个完整的卷积神经网络来完成目标检测过程，所以会用到的是 CNN 特征，通过卷积神经网络提取对候选区域目标的特征的描述。Two-stage 算法的流程图如图 2-14 所示。

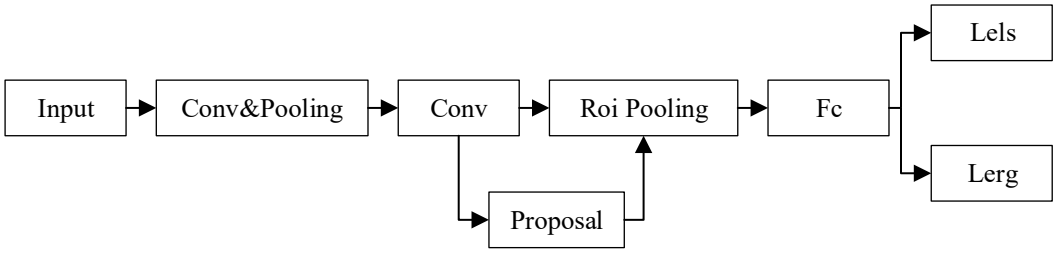


图 2-14 Two-stage 流程图

Two-stage 算法的训练过程主要分为两个步骤：对 RPN 网络的训练和对目标区域的网络训练。相较于传统的目标检测算法，Two-stage 算法不再依靠训练分类器，从而省略特征表示过程，由于整个过程都是由一个完整的神经网络完成，这使得其检测精度得到了提升，但是其检测速度却比 One-stage 算法慢。

Two-stage 常见算法的优缺点如表 2-2 所示。

表 2-2 Two-stage 常见算法的优缺点

目标检测算法	优点	缺点
R-CNN 算法	包含了卷积神经网络	耗费时间，计算量大
Fast R-CNN 算法	检测精度和检测速度有所提升	未充分挖掘样本
Faster R-CNN 算法	解决了获取 Proposal 耗时的问题	高层次特征信息损失较多
HyperNet 算法	提出了多层融合的特征	检测速度慢
FPN 算法	网络整体特征位置的信息丰富	不同尺度的表达能力不一致
Mask R-CNN 算法	特征提取能力强，检测效果好	检测和分割的整体耗时较长

2014 年，Cirshick 等[35]提出了轰动一时的 R-CNN，该算法在精度上相较于 One-stage 算法提升了 25%，同时也为其后续的两部算法奠定了基础。2015 年，Cirshick[35]在 R-CNN 的基础上，对 SPP 网络进行了修改，提出了 Fast R-CNN，

该算法在检测精度和检测速度上都有不少提升。2017年，Ren等[36]提出了Faster R-CNN，通过增加一个全新的 Region Proposal 卷积网络，解决了获取 Proposal 耗时的问题，但是由于其将训练过程划分为了两个部分，这也就使得其检测速度并不能满足实时性要求。此后在 Faster R-CNN 算法的基础上出现了很多变体，如具有更好网络特征的 HyperNet 算法[37]，具有更高精准度的 FPN 算法[38]，具有更完善 ROI 分类的 Mask R-CNN 算法[39]。

2.4 评价指标介绍

在目标检测的领域内，使用不同的算法检测同一目标时，往往会因为算法中间的差异性，导致检测效果和检测速度等的差异，因此就需要一种具有规范性的评价指标来对其进行一定程度的评价，但由于不同场景所导致的检测结果不同，就需要使用多个评价指标对检测算法进行不同方面的评价，再通过曲线形式将其表达出来。对于 YOLO 系列的评价指标，本文将其分为三类，分别是针对预测框的评价指标、针对分类的评价指标和针对目标检测模型的评价指标。

2.4.1 预测框的评价指标

在训练过程中，预测框的准确性起着十分重要的作用，通常用交并比（IOU）反映，其体现了标注框和与预测框的重合程度，用于衡量预测框的正确程度的计算如公式（2-9）所示：

$$\mathrm{IOU} = \frac{\text{预测框} \cap \text{标注框}}{\text{预测框} \cup \text{标注框}} \tag{2-9}$$

当 IOU 越接近 0 时，说明预测框和标注框的重叠部分越少；反之，预测框和标注框的重叠程度越高。

2.4.2 分类的评价指标

精确度（Precision）是指预测结果为正的样本数量占实际结果为正的比例，其计算如公式（2-10）所示：

$$Precision = \frac{TP}{TP + FP} \tag{2-10}$$

其中，当 Precision 越大时，预测结果为正的比例越大，误检越少。

召回率（Recall）则是指所有正样本数量占预测为正的比例，其计算如公式

（2-11）所示：

$$Recall = \frac{TP}{TP + FN} \tag{2-11}$$

其中，当 Recall 越大时，其预测结果中将正预测为负的比例就越低，漏检越少。

在对目标进行检测的过程中，需要对预测结果进行分类，通常其分类结果如表 2-3 所示。

表 2-3 样本分类标准

	实际结果		结果
预测结果	TP （Ture Positive）	FP （False Positive)	TP+FP （Predicted Positive）
	FN （False Negative）	TN （Ture Negative）	FN+TN （Predicted Negative）
结果	TP+FN （Actual Positive）	FP+TN （Actual Negative)	TP+FP +FN+TN

2.4.3 目标检测模型的评价指标

（1）P-R 曲线

P-R 曲线即为以 Precision 和 Recall 为坐标围成的曲线，其中不同颜色的曲线代表不同类别的P-R曲线，其中较粗的蓝色曲线为所有类别的平均P-R曲线。P-R 曲线与坐标轴围成面积通常被用作衡量模型预测结果好坏。

（2）F1-score

当使用不同模型时，其 Precision 和 Recall 也不相同，这也使得我们难以选择出其中最好的模型，因此我们使用 F1-score 来权衡各个模型之间的优劣，其计算如公式（2-12）所示：

$$F1 - score = \frac{1}{\frac{1}{2}\left(\frac{1}{Precision} + \frac{1}{Recall}\right)} \tag{2-12}$$

（3）AP

AP（Average Precision）作为样本中某一类别的平均精度，其通常被表示为由 P-R 曲线与坐标轴围成的面积，其计算如公式（2-13）所示：

$$AP = \int_0^1 P(R)dR \tag{2-13}$$

其中，$P(R)$为 P-R 曲线。

（4）mAP

mAP（mean Average Precision）作为所有类别平均精度的平均值，即所有类别的 AP 值的平均值，既可以反映整个模型的召回率，又可以体现出查准率对模型的影响，当 mAP 越大时，其所用模型的检测效果越好，其计算如公式（2-14）所示：

$$mAP = \frac{\sum_{I=1}^{C} AP_i}{c} \tag{2-14}$$

其中，c为被检测的类别数。

2.5 数据集介绍

为了针对各数据集的差异问题，本文使用的是多数据集训练，其中包括 BDD100K[40]、TT100K[41]、VOC2007[42]和 DAWN[43]数据集，以此实现本文算法的检测稳定性能。本文所用数据集的对比如表 2-4 所示。

表 2-4 数据集对比

		BDD100K	TT100K	VOC2012	DAWN
发布时间		2020 年	2021 年	2012 年	2008 年
图片张数		100000 张	10000 张	11540 张	1027 张
多样性	环境	√	×	√	√
	交通标志	×	√	×	×
	天气	√	√	×	√
目的		研究交通道路的多任务学习	针对各种环境下的交通标志检测	旨在对各种物体类别进行研究	强调了多样化的交通环境

从表 2-4 中可以看出，四种数据集都针对了不同环境下的目标检测，因此为了使研究效果更全面性，本文针对常见交通道路环境、夜间交通道路环境和雨雪雾交通道路环境制作了四种数据集，这四种数据集的具体数据如表 2-5 所示。

表 2-5 本文数据集的数据对比

数据集名称	交通道路环境类别	训练集数量	测试集数量	验证集数量	总计
A 数据集	常见	3500 张	1000 张	500 张	5000 张
B 数据集	夜间	2100 张	600 张	300 张	3000 张
C 数据集	雨雪雾	1400 张	400 张	200 张	2000 张
D 数据集	常见、夜间、雨雪雾	7000 张	2000 张	1000 张	10000 张

从表 2-5 中可以看出，相较于其他三个数据集，D 数据集不仅在环境类别上

更齐全，而且其数据集的数量也较多，因此 D 数据集更适用于本文的交通道路目标检测研究。

为了更好的显示本文使用数据集的有效性，将本文使用的 D 数据集与其他四种数据集进行对比。由于 TT100K 数据集更倾向于交通标志的检测，VOC2012 数据集中也并没有关于交通标志的检测，DAWN 数据集主要被用于雨雪雾天气下行人车辆的检测，故不做比较。

数据集中每张图片所出现人与车辆的数量对比结果如表 2-6 所示。

表 2-6 数据集属性对比

	BDD100K	**TT100K**	**VOC2012**	**DAWN**	**D 数据集**
Persons	129262	/	4087	476	8972
Vehicle	1021857	/	1161	7369	15739
Sign	343777	24717	/	/	4045

2.6 本章小结

本章首先介绍了当前交通环境下的交通道路技术，分析了卷积神经网络算法对交通道路技术目标检测的重要性。其次介绍了传统的卷积神经网络的基础理论知识，并由此对常见的 One-stage 算法和 Two-stage 算法进行了对比。接着介绍了目标检测过程中对于结果的评价指标。然后对 YOLO 系列算法进行了简单的介绍，并在此基础上解释选用 YOLOv7 算法的原因。最后介绍了本文所使用的数据集及其处理过程。为第 3 章、第 4 章和第 5 章中基于 YOLOv7 进行改进的研究内容做铺垫。

第3章 针对常见交通道路环境的目标检测

在交通道路环境中，针对常见目标的检测尤为重要，本章提出了 CTD-YOLOv7 算法。首先通过使用 NAM 算法增加 SENet 模块的识别准确率，进而对常见的交通道路场景中对行人车辆进行有效的检测识别。

3.1 YOLOv7 模型结构

相较于以 R-CNN 系列算法为主的二阶段算法，YOLO 系列算法为了提高检测的实时性，没有采用 proposal 策略。其所使用的直接回归思想可以直接将图像分割成若干个部分，再分别对每个部分进行检测，最后输出检测结果，并且整个操作过程都是建立在卷积和池化操作上。由于 YOLO 系列算法是将训练过程与检测过程融合到同一个网络中，因此其只用通过预测分析便可以得知目标的信息。

为了使目标检测算法能更好的应用于移动端，本文选用的是 YOLOv7 算法。YOLOv7 的网络模型如图 3-1 所示。

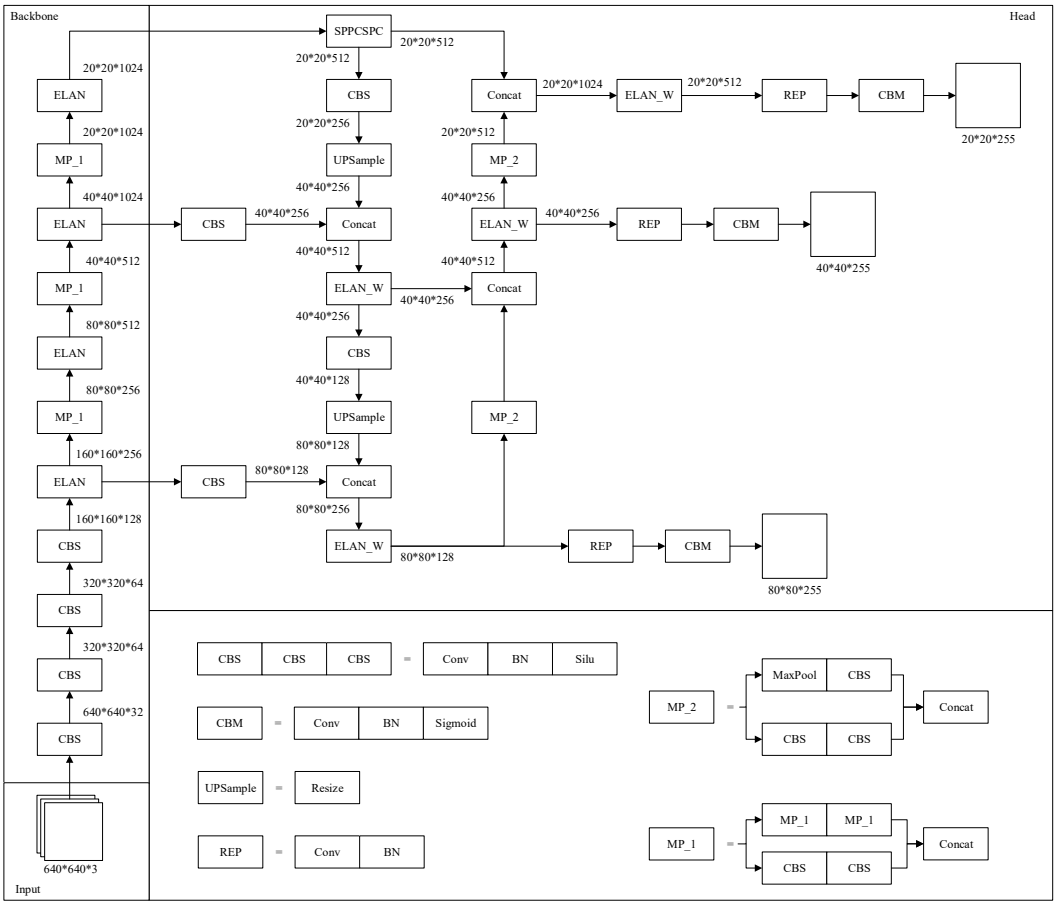


图 3-1 YOLOv7 模型图

其中，ELAN、ELAN-W 和 SPPCSPC 的结构如图 3-2 所示。

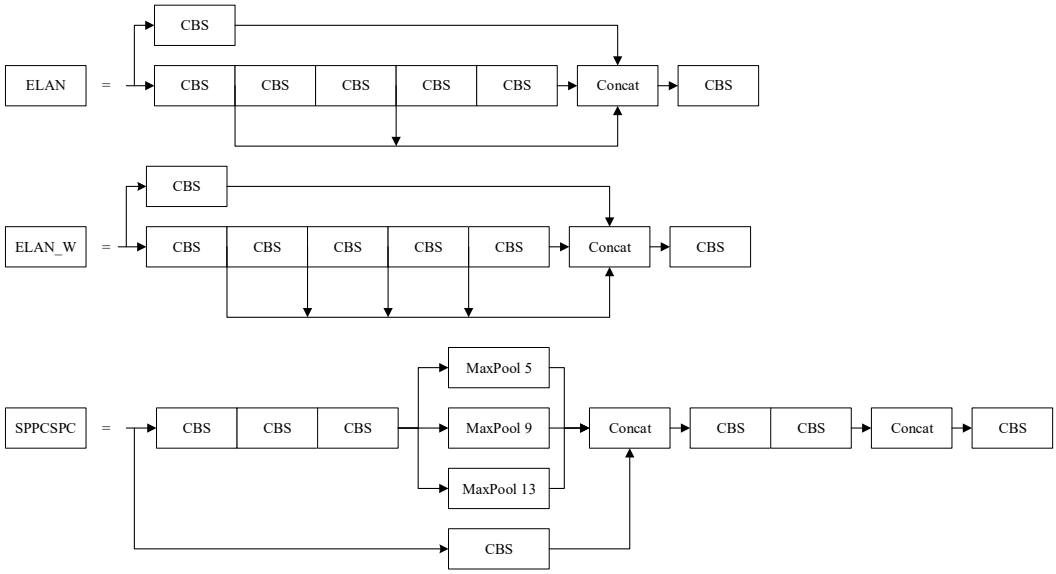


图 3-2 YOLOv7 部分模块结构图

Input 部分：该部分沿用了 YOLOv5 的预处理方式，使用 Mosaic 数据增强来实现对小目标的检测。

Backbone 部分：该部分主要使用 ELAN 和 MP 结构，ELAN 结构可以很好的控制梯度路径，从而实现深层网络有效学习和收敛的目的，而 MP 同时使用了 maxpooling 和 stride=2 的下采样卷积结构，使得网络模型在训练检测过程中保留了图像所有的特征，然后通过池化操作保留局部区域的重要特征信息。虽然消耗时间会有所增加，但是其检测效果会更加准确。

Neck&Head 部分：该部分在检测头部分依旧沿用了 YOLOv5 的 Anchor based 结构。

和 YOLO 系列的其他算法一样，YOLOv7 是通过 Pytorch 框架实现检测，也正是因为 Prtorch 系统的成熟性，故其可以更为轻松的应用于移动端，这也是本文在交通道路检测方面使用 YOLOv7 的主要原因之一。

3.2 针对车辆的目标检测

3.2.1 检测原理

在当前的交通环境中，车辆成为了道路流动的主体，这虽然象征着道路经济的快速发展，但也提高了道路安全隐患的风险。Cheno 等[44]构建了一种基于 HOG 特征向量的阴影区域解析框架，通过特征空间投影实现车辆阴影区域的假

设生成和类别判断。Bougharriou 等[45]在 Cheno 研究成果的基础上开发了完整的 HOG-SVM 检测系统，并建立了从特征提取到分类检测的端到端通道。Liao 等[46]在对夜间红外检测场景的研究中，设计了一种三阶车辆（大型车辆、中型车辆、小型车辆）尺度分类体系，通过内部参数自适应的调整策略优化 SVM 分类边界。宋璠等[47]提出了一种可变形的部件模型，通过局部梯度特征的空间弹性组合增强模型的表征能力，且具有更强的适应性和鲁棒性。张艳邦等[48]提出了一种基于超像素驱动的边缘增强算法，通过将超像素分割与边缘检测融合，构建了具有噪声抑制特性的背景模型，提高算法的计算效率。为了更好的降低道路事故风险，更精准的识别道路上的车辆就显得极为重要。本章对 YOLOv7 算法进行改进，提高算法的检测精度和检测速度，从而降低道路事故的风险。

3.2.2 SENet 模块的改进

在卷积神经网络中，图像在进行特征提取后会被分为空间特征和通道特征。空间特征是指在卷积神经网络前向传播的过程中通过卷积运算得到；通道特征则是卷积层中卷积核的个数以及后续相应的特征融合操作。SENet 模块[49]的提出就是赋予不同通道一个权值，在各通道维度上进行加权，含有不同通道权值的特征图通过加权函数对信息进行筛选，过滤掉不重要的信息，进而关注那些响应值较大、特征信息较丰富的特征图通道。首先 SENet 模块会分别对每个通道的特征图进行全局池化，接着对不同的通道进行压缩处理，然后通过特征激励操作得到不同通道维度的权重，最后将输入的特征图和得到的权重值进行内积得到输出结果特征图，实现特征图通道信息的关联。

SENet 模块的整体结构如图 3-3 所示。

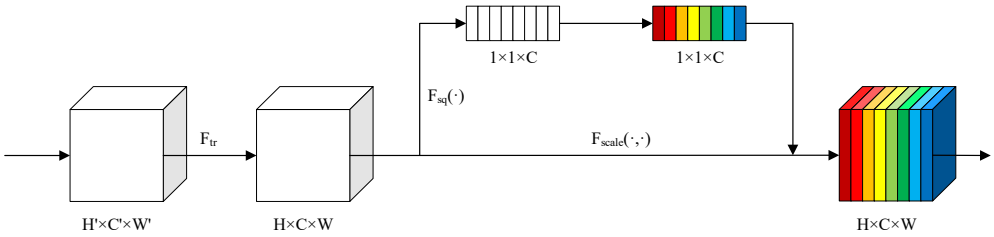


图 3-3 SENet 模块的整体结构图

在 SENet 模块中，首先会对特征图进行 Squeeze 操作，通过采用全局池化的操作将特征图缩小成 1x1 的特征图，并在每个通道内汇总各自对应的空间位置信息。其计算如公式（3-1）所示：

$$Z_c = F_{sq}(W_c) = \frac{1}{H \times W}\sum_{i=1}^{H}\sum_{j=1}^{W} u_c(i,j) \quad , \quad Z \in R^c \tag{3-1}$$

接着需要通过 Excitation 操作对通道的相关性进行联系，其计算如公式（3-2）所示：

$$S = F_{ex}(z,W) = \sigma\big(g(z,W)\big) = \sigma\big(W_2 ReLU(W_1 z)\big) \tag{3-2}$$

其中，$W_1 \in R^{\frac{C}{r}\times C}$，$W_2 \in R^{C\times\frac{C}{r}}$。

通过引入 bottleneck 架构可以有效地减少模型的复杂性，并增强泛化性。该架构由两个完整的层组成，第一个层负责降维，它的系数被称为r，并通过 ReLU 来激发，从而使得第二层重新回归初始的维度。通过计算每条通道的激活值，并与其对应的原始特征进行比较，可以得出x_c的综合分析结果，其计算如公式（3-3）所示：

$$x_c = Fscale(u_c, s_c) = u_c, s_c \tag{3-3}$$

SENet 模块的运行机制可以被视为一种深度学习，它利用权重系数来识别每条通道的特征，从而提高模型的准确性和可靠性。

3.2.3 NAM 注意力机制的改进

通过引入 NAM 注意力机制[50]可以提高权重，从而提升了整体的处理性能。这种方法通过采取批量归一化的方式，将不同的算法分别对应于不同的权重，从而大大减少了 SE 和 CBAM 的需求。NAM 是基于 CBAM 的架构，它改进了原有通道与空间注意力的结构，使它能够被应用于各种位置。此外NAM还使用批次统一的比率来评估通道的优劣，以此来决定其优先级，即如公式（3-4）所示：

$$B_{out} = BN(B_{in}) = \gamma\frac{B_{in} - \mu_\beta}{\sqrt{\sigma_\beta^2 + \varepsilon}} + \beta \tag{3-4}$$

其中，μ_β和σ_β分别为小批量β的均值和标准差；γ和β为可训练的仿射变换参数。

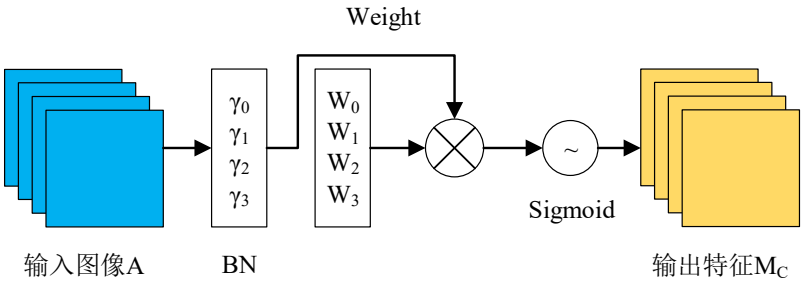


图 3-4 Channel 注意力机制

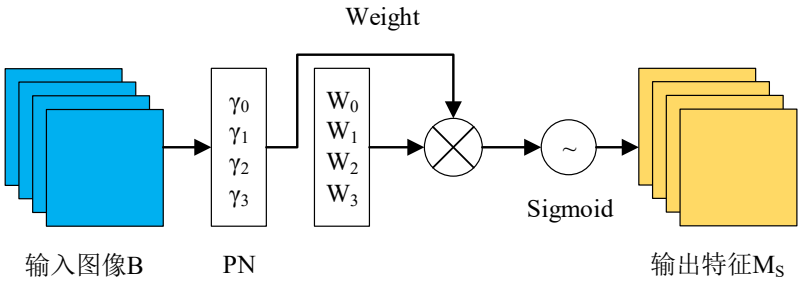


图 3-5 Spatial 注意力机制

NAM 中的 CAM 和 SAM 模块如图 3-4 和图 3-5 所示，其中M_C为 CAM 的输出特征；γ_i为每个通道的比例因子；ω_i为权重；M_S为 SAM 的输出；γ_i为比例因子。此外 NAM 还采取了 BN 的方式，即利用该比例因子来评估图像的重要性，这种方法被称作像素归一化。

3.3 实验环境

3.3.1 硬件环境

在基于深度神经网络的基础上对实验数据进行训练的过程中，需要先对待使用的数据集进行标记，并通过 GPU 来提高算法运行的训练速度和检测速度。本文使用的计算机为联想拯救者 Y7000P，处理器为 AMD Ryzen 7 6800H with Radeon Graphics 3.2GHz，内存 8G，显卡型号为 RTX3060，操作系统为 Windows11，使用的 CUDA 版本和算法框架分别为 10.0 和 Pytorch。

3.3.2 软件环境

本文使用 Anaconda 作为配置程序编译所需要的环境工具，Anaconda 作为 python 的一个开源版本，其包含了 conda、python 等众多依赖包，并且其可以通过安装指令以达到对安装包和环境的管理。

Conda 作为 Anaconda 中一个非常重要的工具，其常被用来对所需环境进行

管理，且不同于‘pip’指令，其可以在程序编译环境中对不同版本的 python 进行切换。

由于本文的研究建立在深度学习的基础上，故在对数据集进行训练的过程中需要使用 GPU 来加快训练速度。GPU 作为图形处理器，其可以对 2D、3D 的图像进行处理加速，然而想要充分发挥 GPU 的作用，就需要在计算机中安装对应版本的 CUDA 和 CuDNN。CUDA 是英伟达 GPU 推出的通用计算引擎，其为英伟达 GPU 进行图像处理的全部开发工具，CUDA 的软件框架主要由编程接口、runtime 库和设备驱动组成，其中编程接口常被用来提供应用开发库，以及解决大规模并行计算问题；runtime 则主要提供开发和运行所需要的相关组件和环境。CuDNN 又被称为神经网络的加速库，其性能好、易使用、内存消耗低等特点让我们在使用时可以将工作重心放在神经网络的设计实现上，同时 CuDNN 也支持各种算法。

Pytorch 作为本文的运行框架，其可以使用 GPU 加速计算 torch 中的张量，因此相较于其他的深度学习框架而言，pytorch 具有更好的灵活性，其直接构建的 python CAPI 可以支持 python 的直接访问，在保证原始代码完整性的前提下，降低了使用难度，除此之外 pytorch 对图像的动态效果具有很好的支持效果，其可以直接运行计算动态图的搭建和调整，此外 pytorch 框架的调试十分便捷，由于其可以在运行时直接生成动态图，因此可以在其使用时进行查看、停止等操作。

3.4 检测结果

为了验证改进后的算法具有更好的检测效果，本章使用数据集 A，并对 YOLOv5、YOLOv7 和 CTD-YOLOv7 进行比较，实验结果的分析如表 3-1 所示：

表 3-1 实验结果分析

算法	P	R	mAP@.5	mAP@.5:.95
YOLOv5	88.0	84.3	90.4	69.5
YOLOv7	88.5	85.1	92.1	74.0
CTD-YOLOv7	**91.1**	**86.1**	**94.4**	**74.6**

与 YOLOv5 算法和 YOLOv7 算法相比，改进后的 CTD-YOLOv7 算法表现出更优越的性能。与 YOLOv5 和 YOLOv7 相比，CTD-YOLOv7 的 Precision 提高了 3.1%和 2.6%，Recall 提高了 1.8%和 1.0%，mAP@.5 提高了 4.0%和 2.3%，

mAP@.95 提高了 5.1%和 0.6%，实验表明 CTD-YOLOv7 不仅保持了较高的检测准确率，而且保持了良好的稳定性。

为了更好地验证 CTD-YOLOv7 的有效性，通过实验对本章算法进行了对比，实验对比结果如表 3-2 所示：

表 3-2 对比实验

算法	FPS	mAP@.5
YOLOv5+SENet	38.237	90.4
YOLOv5+NAMAttention	38.250	89.8
YOLOv5+SENet+NAMAttention	37.346	90.4
YOLOv7+SENet	37.023	93.4
YOLOv7+NAMAttention	37.080	92.9
YOLOv7+SENet+NAMAttention	**36.727**	**94.4**

从表 3-2 中可以看出，与对比算法相比，本章改进的 CTD-YOLOv7 不仅提高了检测精度，而且检测速度更快，从而证实了本章算法的有效性。

为了更直观地证明改进算法的有效性，对表 3-2 中的算法进行了检测图对比。图 3-6 至图 3-13 显示了所提算法与对比算法的检测对比。

图 3-6 YOLOv5

图 3-7 YOLOv5+NAMAttention

图 3-8 YOLOv5+SENet

图 3-9 YOLOv5+NAMAttention+SENet

图 3-10 YOLOv7

图 3-11 YOLOv7+NAMAttention

图 3-12 YOLOv7+SENet

图 3-13 YOLOv7+NAMAttention+SENet

如图 3-7 至 3-13 所示，与对比算法的检测图相比，本章改进的 CTD-YOLOv7 在检测结果上具有更好的结果，避免了检测错误的问题，有效降低了漏检的概率。然而，本章改进的算法仍然无法很好的对小目标进行检测识别。

3.5 本章小结

本章首先对 YOLOv7 的网络模型结构进行了简单介绍。针对当下交通环境对行人与车辆检测的重要性进行了概述，简述了检测的重要性和检测的原理。为了增加 SENet 模块的识别准确率，采用 NAM 算法，即采用批量归一化的方法，在每个通道中都加入相应的权重，避免与 SE 和 CBAM 一样在每个通道中都添加完整的连接层或者卷积层，NAM 算法可以充分利用每条通道的特征，并且在每条通道中都加入相应的权重，可以使 SE 模块拥有较高的识别精确度。实验结果表明，本章提出的针对 SENet 模块和 NAM 注意力机制的改进可以在常见的交通道路场景中对行人车辆进行有效的检测识别。

第4章 针对夜间交通道路环境的目标检测

4.1 检测原理

近十年以来，随着计算机处理能力的增强，深度学习得以快速发展，并开始应用于很多领域，而目标检测则是众多领域中常被专家学者们研究的一个领域。王婧媛等[51]针对高密度人群检测中的特征重叠问题，提出基于 Mosaic 数据增强的级联训练框架，有效的提高人群基数的精度。章程军等[52]通过在主干网络加入 CSP-OSA（Cross Stage Partial-Optimized Stage Architecture）结构提高网络的特征提取能力，通过跨阶段梯度流优化与特征复用机制，实现整体性能的提升。王鹏等[53]聚焦多尺度特征融合，构建自底向上的路径聚合网络，不仅提高了整体网络的鲁棒性，而且具有更好的目标分辨能力。顾德英等[54]通过复合残差连接结构与自适应焦点损失的协同设计，提高了复杂场景下的目标检测速度。关昊等[55]针对夜间检测退化问题，使用 CBS（Convolution Batch norm SiLU）层替换原 Focus 结构，通过动态增强输入特征的照明不变性，从而提高对夜间目标检测的准确率。曹伊宁等[56]在 YOLOv7-tiny 架构中嵌入了无参 SimAM 注意力机制，通过能量函数驱动特征通道自适应加权，使得准确率、鲁棒性和泛化能力均有所提升。刘浩瀚等[57]在 YOLOv7-tiny 的基础上引入 ShuffleNet 网络结构来提升检测精度，通过特征通道重排增强跨维度信息交互，在保持模型计算量不变的前提下，提高了检测效果。

为了解决在夜间检测过程中存在的错检漏捡现象，本章在 CTD-YOLOv7 算法的基础上，提出了 NTD-YOLOv7 算法。首先对图像进行预处理，接着引入 GSConv 模块来减少训练过程中的参数量和计算量，同时选择 ShuffleNetv2×1.5 作为特征提取网络，使网络整体结构在减少参数量的同时能具有较高的跟踪精度，然后采用 hard-swish 激活函数来大幅度降低延迟成本，最后用 SIoU 函数替换 EIoU 函数，从而优化损失函数，提高鲁棒性和泛化能力。

4.2 图像预处理

对于夜间图像的模糊化、不清晰化，本文使用基于拉普拉斯算子的锐化方法对图像进行处理，通过对比局部灰度来判断暗色背景区域，对局部的暗色背景区域进行直方均衡化和 Gamma 变换处理，从而提高图像的清晰度。

该过程的主要流程如图 4-1 所示。

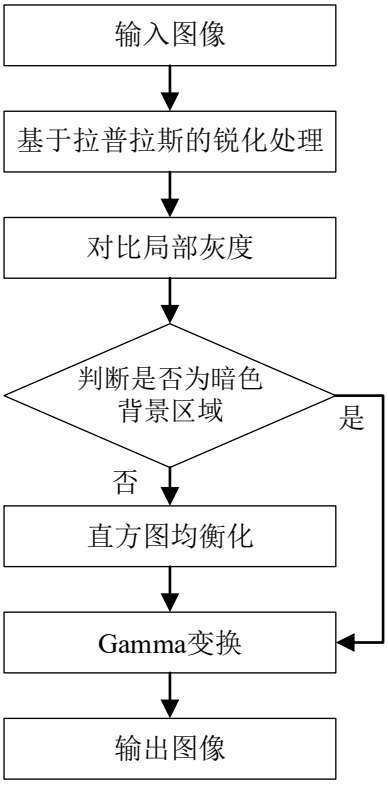


图 4-1 图像增强算法流程图

4.2.1 拉普拉斯算子

拉普拉斯算子作为一种简单的各向同性微分算子，其旋转不变性使其对图像中灰度值的计算具有便利性，拉普拉斯算子如公式（4-1）所示：

$$\nabla^2 f(x,y) = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} \tag{4-1}$$

其中，$f(x,y)$为图像中(x,y)点处的灰度值。

假设存在以(x,y)点为中心的 3×3 拉普拉斯算子的领域，其模型与计算如矩阵（4-2）和公式（4-3）所示：

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \tag{4-2}$$

$$\nabla^2 f(x,y) = \left(\sum_{i=1}^{3} \sum_{j=1}^{3} a_{ij} \right) \times f(x,y) \tag{4-3}$$

按一定比例，将上述公式中拉普拉斯变换的结果代入到原始的灰度图像中，得到的锐化公式如公式（4-4）所示：

$$g(x,y)=f(x,y)+c\times\nabla^2 f(x,y) \tag{4-4}$$

其中，$g(x,y)$表示锐化后的图像在点(x,y)处的灰度值，c表示增强系数，通常取 1。

4.2.2 暗色区域检测

假设存在一块大小为N、灰度级为L的图像，用n_i表示灰度级为r_i的像素点个数，灰度级r_i的归一化直方图H的定义如公式（4-5）所示：

$$H(r_i)=\frac{n_i}{N}\quad,\quad i=0,1,\cdots,L-1 \tag{4-5}$$

通过对比局部灰度均值方差和全局灰度均值方差，进而达到检测出夜间图像暗色区域的目的。具体的步骤如下：

（1）计算出全局灰度的均值E_g和方差σ_g^2，其计算如公式（4-6）所示：

$$\begin{cases} E_g=\sum_{i=0}^{L-1} r_i\times\frac{n_i}{N} \\ \sigma_g^2=\sum_{i=0}^{L-1}(r_i-E_g)\times\frac{n_i}{N} \end{cases} \tag{4-6}$$

（2）假定以点(x,y)为中心的区域$S(x,y)$，其大小为$M\times N$，则该邻域局部灰度的均值和方差的计算如公式（4-7）所示：

$$\begin{cases} E_S=\frac{1}{M^2}\sum_{x=1}^{M}\sum_{y=1}^{M} r(x,y) \\ \sigma_S^2=\frac{1}{M^2}\sum_{x=1}^{M}\sum_{y=1}^{M}[r(x,y)-E_S]^2 \end{cases} \tag{4-7}$$

（3）对比全局灰度和邻域灰度的均值和方差的大小，若满足下述条件，则该区域为暗色背景区域，即亮度较暗、对比度较低，其计算如公式（4-8）所示：

$$\begin{cases} E_S\le\frac{3}{5}E_g \\ \sigma_S^2\le\frac{3}{5}\sigma_g^2 \end{cases} \tag{4-8}$$

4.2.3 局部直方图均衡化

局部直方图均衡化是由全局直方图均衡化演变而来的，不同于全局均衡直方图，其不再是对整个图像进行操作，而是将图像分成若干个小局部块，进而

对每个小局部块进行直方图均衡化操作，最终将所有均衡化后的小局部块重新组合成最终所需要的输出图像。

局部均衡化直方图的转换函数的定义如公式（4-9）所示：

$$T(r_i)=\sum_{j=0}^{i}H(r_i)=\sum_{j=0}^{i}\frac{n_j}{N}\quad\begin{cases}i=0,1,\cdots,L-1\\0\le r_i\le 1\end{cases}\tag{4-9}$$

其优点是可以增强图像的局部细节，从而避免由全局直方图均衡化引起的过度增强或者过度失真的问题。

4.2.4 Gamma 变换

Gamma 变换又叫做幂律变换，是一种呈非线性关系的图像增强方法，主要用于修正整体图像，柔和图像的亮度。由于在夜间环境下，移动端采集到的图像大多处于低亮度，无法很好的通过人眼进行辨识，因此需要先对采集到的图像进行 Gamma 变换。

Gamma 变换是对输入图像灰度值进行的非线性操作，使输出图像灰度值与输入呈指数关系，其计算如公式（4-10）所示：

$$V_{out}=AV_{in}^{\gamma}\tag{4-10}$$

其中，V_{out}是输出的灰度值，A是灰度缩放系数，V_{in}是输入的灰度值，γ是伽马因子大小。

进行 Gamma 变换后，暗部细节和亮部细节都会发生相应的变化，如图 4-2 所示，当 Gamma<1 时，提升了暗部细节，压缩了亮部细节；当 Gamma>1 时，亮部细节得到了提升，而暗部细节被压缩；Gamma=1 时为原始图像。

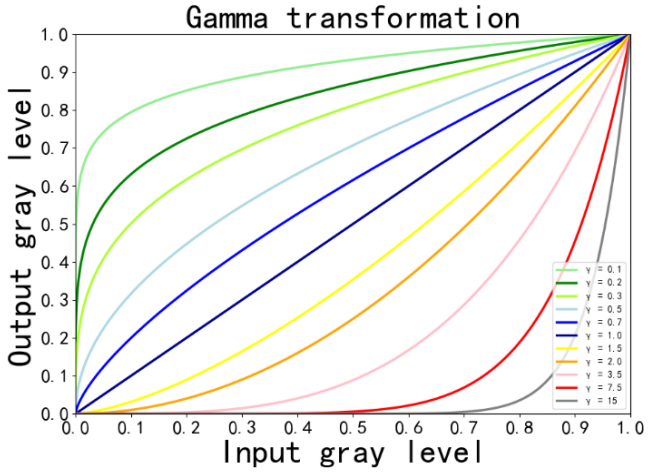


图 4-2 Gamma 细节变化图

实曲线 CRT Gamma 和虚曲线校正 CRT 亮度响应曲线如图 4-3 所示，通过校正的是最终近似斜率为 45°的响应曲线，此时 Gamma=1，但是为了使图像变得更加接近人眼感受的非线性响应，提高图像质量，故选择 Gamma=2.2 的映射。

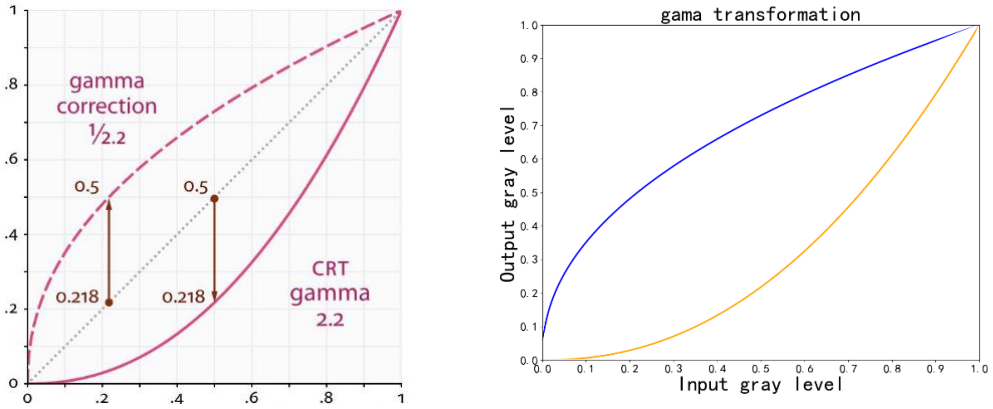


图 4-3 Gamma 实、虚曲线的响应曲线

CRT Gamma 曲线和校正 CRT 亮度响应曲线的计算如公式（4-11）所示：

$$\begin{cases} V_{out} = \dfrac{255}{255^{2.2}} \times V_{in}^{2} \\ V_{out} = \dfrac{255}{255^{1/2.2}} \times V_{in}^{2} \end{cases} \tag{4-11}$$

当 Gamma=2.2 时，计算亮度变化的幅度，映射到 Gamma 曲线后，从灰阶幅度 1 变化到 2 的亮度变化幅度为 37.04%。但从灰阶幅度 254 变化到 255，亮度变化只有 0.18%。Gamma 映射曲线和 Gamma 变化强度曲线如图 4-4 所示。

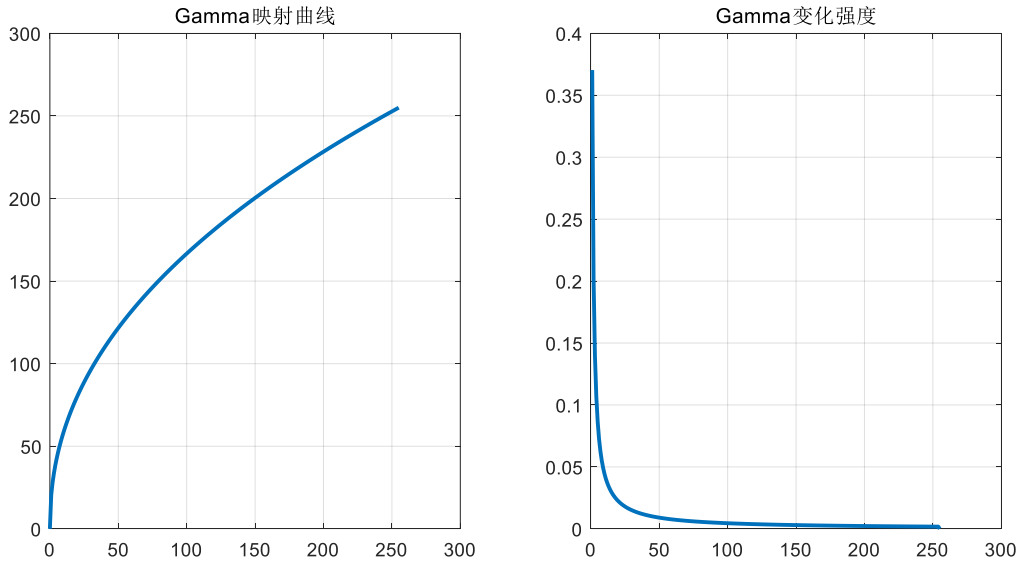


图 4-4 Gamma 映射曲线及变化强度

图 4-5 是处理前后的结果，可以看出通过 Gamma 非线性变换处理后的图片

具有更加丰富的颜色，这有利于后期对图像进行检测识别。

图 4-5 变换前后对比图

4.3 算法改进

4.3.1 GSConv 模块

对于夜间检测而言，检测精度和检测速度都是不可或缺的，因此需要在保证检测精度的前提下，降低网络结构的复杂性，以达到提高检测速度的目的。由于 GSConv[58]具有出色的性能和计算量的降低，常被用于视觉模型轻量化的研究中。为了较少网络的计算量，本章选择用 GSConv 模块代替原始网络 Neck 部分中的 Conv 模块。

GSConv 作为一种卷积神经网络的变体，它采用了自适应的图结构来进行图像分割和图像标注任务，相较于传统的 GCN 而言，GSConv 通过学习图结构来捕捉图像中的上下文信息，因此能够更好的适应不同图像任务的需求，并取得更好的性能效果。其作为深度可分离卷积模块（Depthwise Wise Convolution, DWConv[59]）、标准卷积模块（Standard Convolution, SConv[60]）和 Shuffle 模块[61]的组合，不仅有效的利用了 DWConv 模块的计算能力，同时也拥有了 SConv 模块的检测精度，其结构如图 4-6 所示。

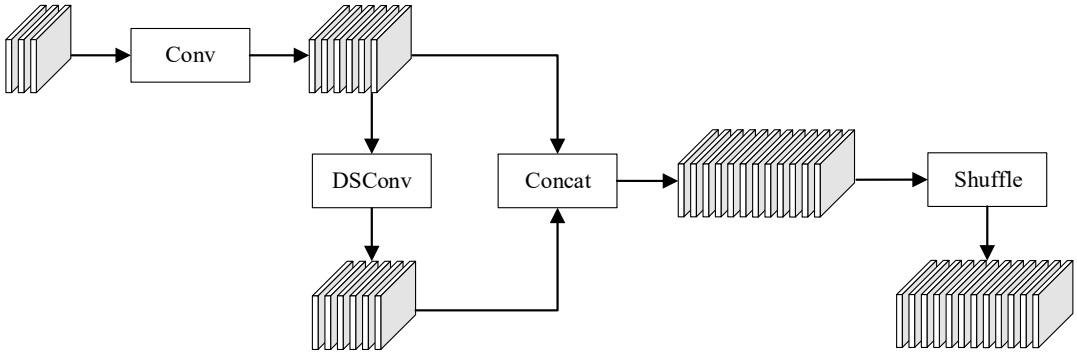


图 4-6 GSConv 模块

从图 4-6 中可以看出，GSConv 首先对普通卷积进行欠采样，然后使用

DWConv 深度卷积，并将 SConv 和 DSConv 的结果进行串接，最后进行 Shuffle 操作，使前两次卷积对应的通道号相邻，主要目的是使 DSConv 的输出尽可能接近 SConv。该模块采用 Shuffle 的方式，通过密集的卷积运算将 SConv 生成的信息渗透到 DSConv 生成的信息的各个部分中，可以大大减少 DSConv 缺陷对模型的负面影响，使其发挥 DSConv 的优势，即占用内存少，计算速度快。

4.3.2 ShuffleNetv2×1.5 网络结构

Zhang 等[62]在 2018 年提出了 ShuffleNet，作为一种先进的基于图像分类的轻量化卷积神经网络体系结构，ShuffleNet 由于其在计算量等方面的巨大优势，通常被应用于移动设备。

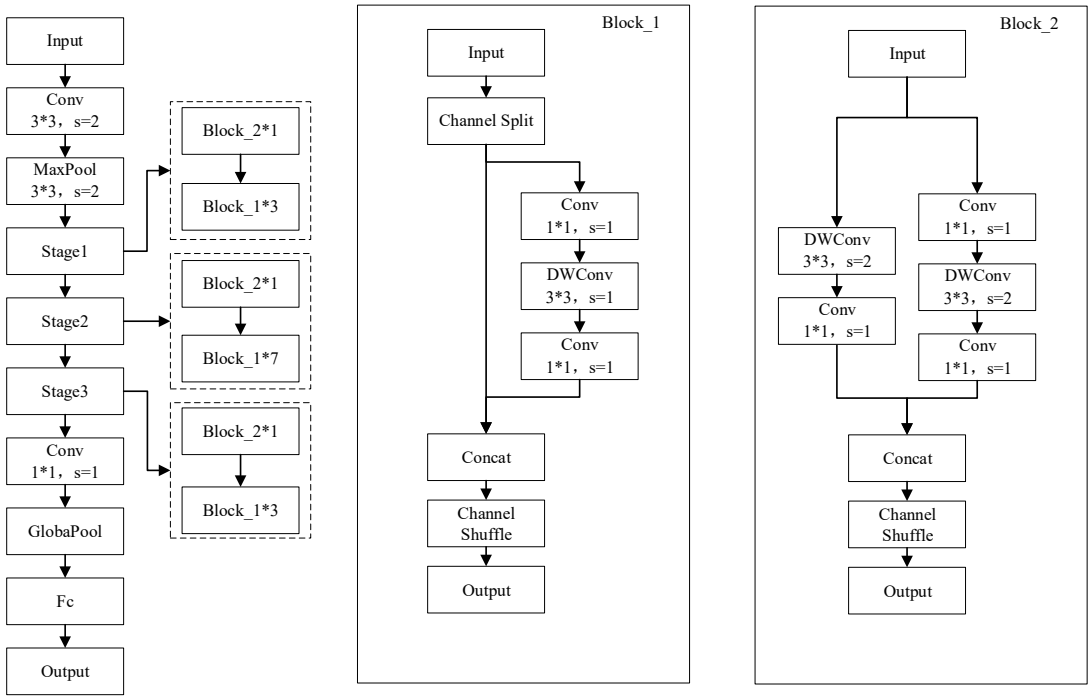


图 4-7 ShuffleNetv2 网络结构

ShuffleNet 网络结构是由群卷积（Group Convolution, GC）和深度可分离卷积所组成。这种组成在有效降低计算复杂度的同时，还能加强整体网络结构的深度，从而增强对图像特征的提取，提升检测精度。该网络采用了通道混洗的方式将不同通道的输入信息均匀打乱，以此解决不同群卷积之间所导致的信息不交互问题。ShuffleNetv2 由卷积层、池化层、三个阶段、全局平均池化和最终的全连接层组成，阶段 1 到阶段 3 是通过堆叠两个不同的残差块 Block_1 和 Block_2 形成的，然后使用全连接层上的 Softmax 函数输出 128 维特征向量。根据网络每层的输出通道数的不同，ShuffleNetv2 可分为四种不同大小的型号：

ShuffleNetv2×0.5、ShuffleNetv2×1.0、ShuffleNetv2×1.5 和 ShuffleNetv2×2.0。ShuffleNetv2×0.5模型最小，ShuffleNetv2×2.0模型最大。考虑到网络的深度和特征提取能力，本章选择 ShuffleNetv2×1.5 作为特征提取网络，ShuffleNetv2×1.5网络结构如图 4-7 所示。

4.3.3 激活函数的优化

激活函数作为在神经网络的神经元上运行的函数，其主要任务是负责将神经元的输入映射到输出端。

当一个神经网络的层数较多时，其模型在训练的时候会出现梯度消失和梯度爆炸的问题，且梯度消失和梯度爆炸的问题会随着网络层数的增加而变得越来越明显。梯度消失会导致网络和模型的训练效果下降。当深层网络和权值初始化太大时，则会导致梯度爆炸，在这种情况下，网络结构会变得不够稳定，训练的过程和结果也会变得不够稳定。梯度消失和梯度爆炸的本质都是因为梯度反向传播中的连乘效应，为了解决梯度的问题，本章选择采用优化激活函数的方法。

Hard_Swish 激活函数是用传统的 Swish 激活函数代替 ReLU 激活函数，其计算如公式（4-12）所示：

$$Hard\_Swish(x)=\begin{cases}0 & , \quad x\le -3\\ \frac{1}{6}x^2+\frac{1}{2}x & , \quad -3<x<3\\ x & , \quad x\ge 3\end{cases} \tag{4-12}$$

其对 x 的导数计算如公式（4-13）所示：

$$Hard\_swish'(x)=\begin{cases}0 & , \quad x\le -3\\ \frac{1}{3}x+\frac{1}{2} & , \quad -3<x<3\\ 1 & , \quad x\ge 3\end{cases} \tag{4-13}$$

Hard_Swish 激活函数求导简单，且能够有效防止训练时梯度逐渐接近零时导致的饱和现象发生，可以进一步提升网络模型的表达能力。

4.3.4 损失函数的优化

损失函数在深度学习中具有十分重要的作用，其作为模型的预测误差，常被用于衡量模型预测结果与实际标签之间的差异程度。在监督学习任务中，损失函数通常通过比较模型的预测结果与真实标签来计算误差。作为度量神经网

络预测信息与期望信息的距离，训练时主要包括三个方面的损失，分别为边界框置信度损失、类别损失和坐标损失。而其中坐标损失常使用 IoU（Intersection over Union）系列损失函数，IoU 是一种用于衡量预测框和真实框之间的重叠程度，其计算和损失函数如公式（4-14）和公式（4-15）所示：

$$IoU = \frac{|B \cap B_1|}{|B \cup B_1|} \tag{4-14}$$

$$L_{IoU} = 1 - IoU = 1 - \frac{|B \cap B_1|}{|B \cup B_1|} \tag{4-15}$$

其中，B为预测框的面积，B_1为真实框的面积，由于 IoU 具有非负性、同一性和对称性等特点，这使得输出的损失值始终处于 0 到 1 之间。

但 IoU 在现实情况中可能存在预测框与真实框不相交的情况，这会导致损失函数的输出恒为 1，继而无法继续学习。为了解决这一问题，侯志强等[63]提出了 GIoU，即加入了预测框和真实框的最小外接矩形，但当预测框和真实框处于包含状态时，GIoU 损失函数便无法确定两个框的位置关系。Zheng 等[64]提出的 DIoU 损失函数可以有效地减少损失收敛的时间，这是因为它将最小外接矩形的重叠面积标准化为一个最小的阈值，从而使得预测框和真实框之间的距离也可以达到最小。然而，这种损失函数并没有考虑纵横比。王永顺等[65]提出的 CIoU 进一步对模型的回归精度进行优化，但是对纵横比的权重设计存在一定的模糊化，且没有考虑样本难易程度的平衡问题。Zhang 等[66]在 GIoU 的基础上提出了 EIoU 损失函数，EIoU 通过计算重叠损失、中心距离损失和宽高损失，解决了 GIoU 损失函数在水平和垂直方向上的误差，提高了收敛的速度和回归的精度，但其在大规模数据的训练上会耗费大量的时间，且对于不同目标尺度的处理不够灵活。李功等[67]在 CIoU 的基础上提出了 SIoU，通过将边界框的参数视为整体进行回归，在不增加网络结构复杂度的情况下降低计算定位损失，加快模型的收敛速度。

SIoU 损失函数是一种由角度损失、距离损失和形状损失组成的损失函数，由于其较好的准确性和鲁棒性，因此本章选择 SIoU 作为损失函数。SIoU 损失函数如图 4-8 所示。

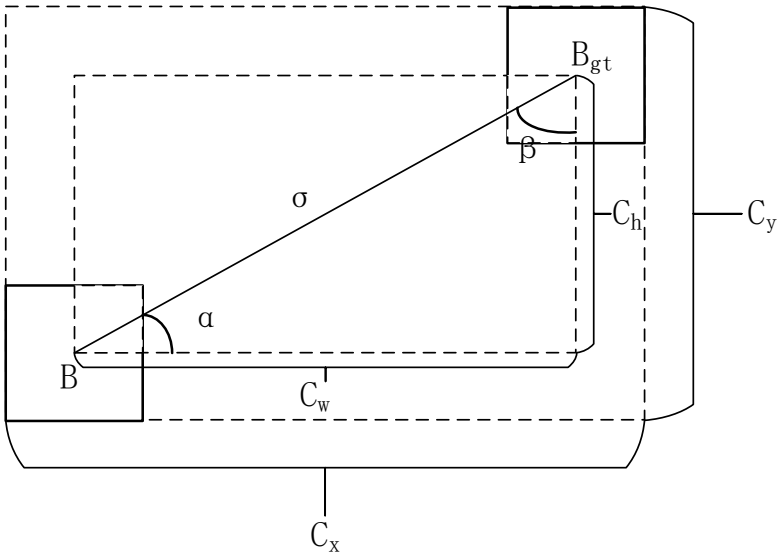


图 4-8 SIoU 损失函数

其中，B表示预测框的中心点，B_{gt}表示真实框的中心点，σ表示真实框中心点到预测框中心点的距离，C_x和C_y为真实框和预测框最小外接矩形的宽和高，C_h和C_w为真实框和预测框中心点的高度差和宽度差，α为平衡参数。

（1）角度损失

首先判断真实框中心点和预测框中心点的角度是否大于 45°，进而决定使用最小化α还是β，角度损失的计算如公式（4-16）所示：

$$\Lambda = 1 - 2 \times sin^2\left(arcsin(x) - \frac{\pi}{4}\right) = cos\left(2 \times \left(arcsin(x) - \frac{\pi}{4}\right)\right) \quad (4\text{-}16)$$

其中，x、σ和C_h的定义如公式（4-17）、公式（4-18）和公式（4-19）所示：

$$x = \frac{C_h}{\sigma} = sin\,\alpha \quad (4\text{-}17)$$

$$\sigma = \sqrt{C_h^2 + C_w^2} = \sqrt{\left(B_{Cx}^{gt} - B_{Cx}\right)^2 + \left(B_{Cy}^{gt} - B_{Cy}\right)^2} \quad (4\text{-}18)$$

$$C_h = max\left(B_{Cy}^{gt}, B_{Cy}\right) - min\left(B_{Cy}^{gt}, B_{Cy}\right) \quad (4\text{-}19)$$

（2）距离损失

距离损失代表预测框中心点与真实框中心点的距离，其计算如公式（4-20）所示：

$$\Delta = \sum_{t=x,y} \left(1 - e^{-\gamma \rho_t}\right) \quad (4\text{-}20)$$

其中，ρ_x、ρ_y和γ的定义如公式（4-21）、公式（4-22）和公式（4-23）所示：

$$\rho_x = \left(\frac{B_{Cx}^{gt} - B_{Cx}}{C_w}\right)^2 \tag{4-21}$$

$$\rho_y = \left(\frac{B_{Cy}^{gt} - B_{Cy}}{C_h}\right)^2 \tag{4-22}$$

$$\gamma = 2 - \Lambda \tag{4-23}$$

当α趋向于 0 时，距离损失大幅度降低；反之，当α趋向于 45°时，距离损失则会增大，γ被赋予时间优先的距离值会随着角度α的增大而增大。

（3）形状损失

SIoU 中的形状损失是一种用于度量预测框和真实框之间相似度的损失函数，其定义如公式（4-24）所示：

$$\Omega = \sum_{t=w,h} (1 - e^{-W_t})^{\theta} = (1 - e^{-W_w})^{\theta} + (1 - e^{-W_h})^{\theta} \tag{4-24}$$

其中，W_w和W_h的定义如公式（4-25）所示：

$$W_w = \frac{|w - w^{gt}|}{\max(w, w^{gt})} \quad , \quad W_h = \frac{|h - h^{gt}|}{\max(h, h^{gt})} \tag{4-25}$$

其中，(w, h)和(w^{gt}, h^{gt})分别表示预测框和真实框的宽度和高度，θ表示损失函数对形状损失的关注程度，由于过度关注形状损失会导致预测框移动的降低，因此θ的取值基本为 4。

最终，SIoU 损失函数的定义如公式（4-26）所示：

$$Loss_{SIoU} = 1 - IoU + \frac{\Delta + \Omega}{2} \tag{4-26}$$

4.4 检测结果

本章对数据集 B 进行实验，对比 YOLOv5 算法、YOLOv7 算法和本章改进的 NTD-YOLOv7 算法，实验结果分析如表 4-1 所示：

表 4-1 实验结果分析

模型	P(%)	R(%)	mAP@.5(%)	mAP@.5:.95(%)
YOLOv5	80.3	76.8	77.8	44.3
YOLOv7	84.6	77.2	79.6	43.6
改进 YOLOv7	**90.2**	**83.3**	**88.1**	**48.3**

与 YOLOv5 算法和 YOLOv7 算法相比，NTD-YOLOv7 算法表现出了优越的性能。NTD-YOLOv7 算法在精度 P 中分别提高了 4.9%和 5.6%，在 R 中分别提

高了 6.5%和 6.1%，在 mAP@.5 中提高了 9.3%和 8.5%，在 mAP@.95 中分别提高了 4.0%和 4.7%，表明 NTD-YOLOv7 算法不仅保持了较高的检测精度，同时也保持良好的稳定性。

图 4-9 显示了训练过程中的 P-R 曲线，从中可以看出训练后的模型具有良好的性能。

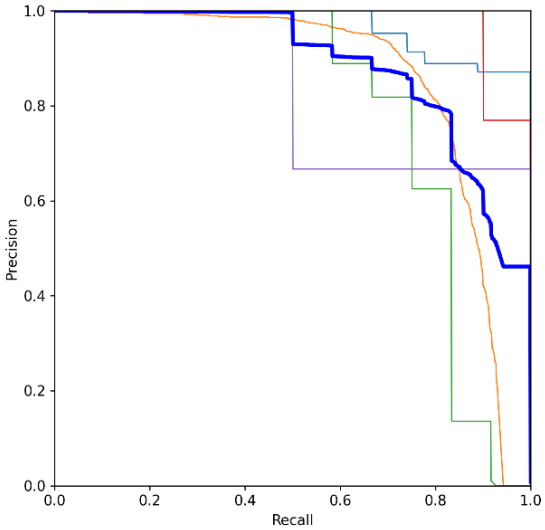
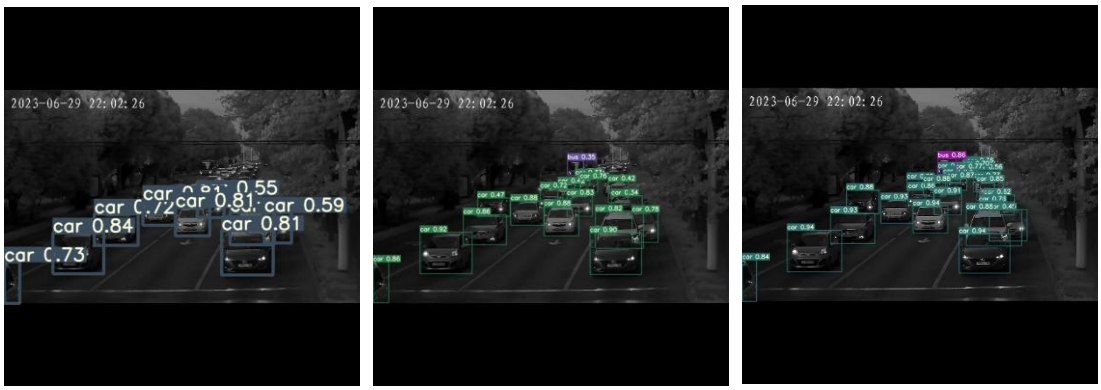


图 4-9 P-R curve

图 4-10 显示了 YOLOv5 算法、YOLOv7 算法和 NTD-YOLOv7 算法在明亮环境下的检测效果。



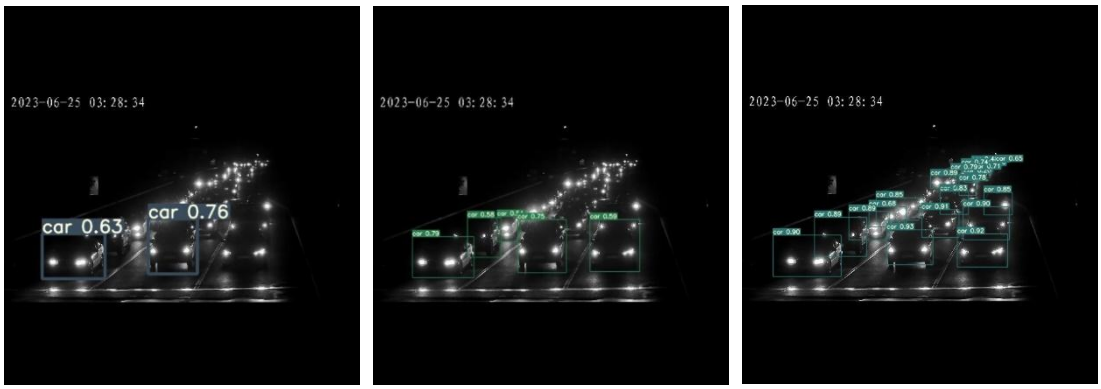
YOLOv5　　YOLOv7　　NTD-YOLOv7 算法

图 4-10 明亮环境下的效果对比图

从图 4-10 中可以看出，与 YOLOv5 算法和 YOLOv7 算法相比，NTD-YOLOv7 算法可以更有效地防止漏检、误检等问题的发生。因此，改进后的算法不仅具有更好的检测精度，而且适用于小目标检测场景。

图 4-11 显示了 YOLOv5 算法、YOLOv7 算法和 NTD-YOLOv7 算法在黑暗

环境下的检测效果。



YOLOv5　　YOLOv7　　NTD-YOLOv7 算法

图 4-11 黑暗环境下的效果对比图

从图 4-11 中可以看出，YOLOv5 算法和 YOLOv7 算法在黑暗环境中的检测效果都不好，其中漏检率较高，而改进后的 NTD-YOLOv7 算法基本解决了暗环境下的漏检问题。

表 4-2 ShuffleNetv2×1.5 的消融实验

模型	**P(%)**	**R(%)**	**mAP@.5(%)**	**mAP@.5:.95(%)**
YOLOv7	84.6	77.2	79.6	43.6
YOLOv7+ShuffleNetv2×1.5	**86.1**	**78.5**	**82.9**	**44.3**

关于损失函数的消融实验结果如表 4-3 所示：

表 4-3 损失函数的消融实验

模型	**CIoU**	**EIoU**	**SIoU**	**mAP@.5(%)**
YOLOv5	√			77.8
YOLOv5		√		80.1
YOLOv5			√	81.9
YOLOv7		√		79.6
YOLOv7			√	**88.1**

关于图像增强和 GSConv 的消融实验结果如表 4-4 所示：

表 4-4 图像增强和 GSConv 的消融实验

网络模型	图像增强	GSConv	mAP@.5(%)
YOLOv7			78.5
YOLOv7		√	84.2
YOLOv7	√		81.6
YOLOv7	√	√	**88.1**

从表 4-3 的实验结果可以看出，当 YOLOv5 采用不同的损失函数 CIoU、

EIoU 和 SIoU 时，mAP@.5 分别为 77.8%、80.1%和 81.9%；当 YOLOv7 采用 EIoU 损失函数时，mAP@.5 分别为 79.6%。与其他算法相比，本章改进的 YOLOv7 算法在 mAP@.5 上分别提高了 10.3%、8%、6.2%和 9.5%。

从表 4-4 的实验结果可以看出，在检测未使用 GSConv 模块的图像时，使用 GSConv 模块的检测精度比未使用 GSConv 模块的图像高 5.7%，而在检测未使用算法增强的图像时，使用 GSConv 模块的检测精度比未使用 GSConv 模块的图像高 6.5%。当我们不使用 GSConv 模块时，图像增强算法可以提高 3.1%的检测精度，而当我们使用 GSConv 模块时，图像增强算法可以提高 3.9%的检测精度。因此图像增强算法和 GSConv 模块都可以提高在夜间环境中对车辆检测的准确性。

4.5 本章小结

本章首先介绍了在夜间等特殊天气情况下对交通道路进行检测的重要性，并概述了近几年关于夜间目标检测的研究。其次在已改进的基础上对图像进行预处理操作，通过 Gamma 变换提高图像的清晰度，并采用直方图均衡化的方法避免由全局直方图均衡化引起的过度增强或者过度失真的问题。接着使用了 ShuffleNetv2×1.5 网络结构，其不仅适用于移动设备，而且其采用的通道混洗的方式可以将不同通道的输入信息均匀打乱，以此解决不同群卷积之间所导致的信息不交互问题，最后还对损失函数和激活函数进行了改进，从得出的检测结果看，改进后的 YOLOv7 算法对夜间目标具有较好的检测效果。

第5章 针对雨雪雾交通道路环境的目标检测

除了常见情况和夜间环境外，雨雪等特殊天气则是影响目标检测的一大因素。由于雨雪等特殊天气会造成图像模糊，导致行人车辆目标特征提取困难，检测精度差。对此，本章在 NTD-YOLOv7 的基础上，提出了 AWTD-YOLOv7 算法。首先介绍了图像清晰化的理论基础，接着使用较为广泛的暗通道去雾算法和基于引导滤波的去雨雪算法，然后使用 DIP 模块和 CNN-PP 模块的结合，提出了针对雨雪等特殊天气的目标检测方法。

5.1 检测原理

针对恶劣天气环境中的目标检测是交通道路领域需要克服的一个难题，其中包括对雾霾天气和雨雪天气的目标检测。目前去雾的主要算法方法有两类，一类是基于图像增强的去雾方法，该方法只是简单的加强了图像的细节信息，并没有很好的实现去雾效果；第二类是基于物理成像的去雾方法，该方法通过数学建模实现大气散射物理模型。

Tan[68]提出了基于局部对比度最大化的物理去雾模型，通过马尔可夫随机场优化实现大气散射系数估计，但是由于该方法中复原图像的对比度过高，导致 SSIM 指标下降。Tarel 等[69]通过利用中值滤波估计大气光成像函数，复原图像中的景物边缘出现光晕。He 等[70]提出暗通道理论，通过最小通道滤波推导透射率矩阵，结合物体成像模型推导出图像的透射率实现去雾。在 He[71]的基础上，陈书贞等[72]提出了混合暗通道补偿机制，联合亮度通道与色度约束构建自适应权重函数，以此解决有雾图像中出先大面积天空和白色物体时暗通道先验失效的问题。张晓刚等[73]针对过度去雾问题，开发基于 Canny-Guided 的边缘感知正则化项，通过梯度场动态调节透射率优化强度，改善了现有方法出现的过度去雾情况。陈丹丹等[74]改进了大气耗散函数的非均匀性建模，利用蒙特卡洛光线追踪模拟气溶胶分布，使得透射率估计更加准确。

目前图像去雨雪方法根据是否有时间特性分为视频去雨雪和单幅图像去雨雪。Guo 等[75]将全变分正则化与稀疏编码结合，构建三维时空张量分解模型，可以对视频的前景空间进行更有效的建模。Shijila[76]在 Guo 的基础上扩展 RPCA 框架，提出低秩背景-稀疏前景-噪声三分解模型，通过 Gaussian-Poisson 混合噪

声建模，很好的实现对视频中噪声的去除效果。李艳荻等[77]提出了超像素级时空梯度分析框架，采用 SLIC 超像素分割提取颜色/运动双模态梯度特征，结合光流轨迹聚类构建时空显著性图谱，从而构建超像素级的时空梯度图。Li 等[78]提出了分割和显著性约束方法，联合超像素分割与显著性注意力机制，该方法可以应对动态背景和显著性约束并能够检测到缓慢移动的对象，但受限于未引入噪声鲁棒性设计，所以在实际应用中可能会影响检测效果。

5.2 图像去雨雪雾研究

在面对雨雪雾环境时，交通道路采集到的实时图像并不如平时图像一样清晰，这也使得在雨雪雾环境下的检测结果并不好，因此对图像清晰化的预处理就显得极为重要。本章针对雨雪环境和雾天环境分别进行图像的清晰化处理，在处理雾霾天气时，本章基于暗通道先验的方法实现去雾，而在处理雨雪环境时，本章基于引导滤波的操作实现去雨雪。

5.2.1 基于暗通道先验去雾

近几年，雾霾已经变得越来越普遍，它不仅污染了环境，还给居民的健康带来极大的威胁，也严重影响了道路交通安全。由于雾霾的存在，空气中的微细颗粒物数量急剧上升，它们不仅削弱了光线的穿透能力，还使得许多数据的内容受到损害，从而给图像的后期处理带来极大的困难，因此采取适当的措施，如采用先进的技术，将受到污染的环境恢复原貌，显得尤为必要。为了很好的解决这一问题，本章决定使用暗通道先验的方法实现图像去雾的过程。

通常情况下，通过使用大气散射物理模型来表示有雾图像，其计算如公式（5-1）和公式（5-2）所示：

$$I(x) = J(x)t(x) + A\big(1 - t(x)\big) \tag{5-1}$$

$$t(x) = e^{-\beta d(x)} \tag{5-2}$$

其中，$I(x)$为有雾图像，β为散射系数，$d(x)$为图像的深度。

对于不包含天空区域的无雾图像，在没有任何外部遮挡的情况下，三个颜色的通道之间的差异会变得极小，甚至接近 0，晴朗天气条件下的图像$J(x)$的暗通道定义如公式（5-3）和公式（5-4）所示：

$$J^{dark}(x) = \min_{y \in \Omega(x)} \left(\min_{c \in \{R,G,B\}} J^c(y) \right) \tag{5-3}$$

$$J^{dark} \to 0 \tag{5-4}$$

其中，$\Omega(x)$为邻域窗口，c为图像中颜色通道。

为了能获得更准确的结果，需要对公式（5-1）进行局部最小值滤波操作，其结果如公式（5-5）所示：

$$t(x) = 1 - \min_{y\in\Omega(x)}\left(\min_{c\in\{R,G,B\}}\frac{J^c(y)}{A^c}\right) \tag{5-5}$$

为了让复原后的图像更接近人类的视觉体验，需要尽可能地减少雾气的存在，因此通过调节相关参数来精确控制减少雾气的程度，从而获得最佳的复原效果，因此通过约束参数$\omega(0 \le \omega \le 1)$来控制去雾的程度，其计算如公式（5-6）所示：

$$t(x) = 1 - \omega\min_{y\in\Omega(x)}\left(\min_{c\in\{R,G,B\}}\frac{J^c(y)}{A^c}\right) \tag{5-6}$$

其中，ω取 0.95。

如果仅仅使用公式（5-6）来粗略估算透射率，而未经过精细处理就直接用于图像去雾，那么图像中的边缘将会变得模糊不清，因此必须采取措施来优化透射率，以达到最佳的效果，其计算如公式（5-7）所示：

$$J(x) = \frac{I(x) - A}{\max(t(x), t_0)} + A \tag{5-7}$$

其中，t_0为一个限定常数，一般取 0.1。

本章使用的暗通道先验去雾方法的实验结果如图 5-1 所示。

（a）去雾前　　（b）去雾后

图 5-1 暗通道先验去雾效果对比图

从图 5-1 中可以看出，通过使用暗通道先验算法去雾，相比于原始图像，去雾后的图像具有更好的清晰度，可以更好的描绘出图像的细节部分，有利于后续的目标识别检测。

5.2.2 基于引导滤波的去雨雪操作

导向滤波作为一种图像处理技术，其在平滑图像细节的同时，还可以不影响图像的边缘信息。导向滤波利用引导图像对目标图像进行滤波操作，滤波后的图像不仅保留了目标图像的大体信息，而且也保留了引导图像的细节纹理，其计算如公式（5-8）所示：

$$q_i = a_k I_i + b_k \quad , \quad \forall i \in \omega_k \tag{5-8}$$

其中，q是滤波后图像的像素值；I是引导图像的像素值；i和k是像素索引值；a和b为系数。

对公式（5-8）两边进行梯度操作：

$$\nabla_q = a\nabla_I \tag{5-9}$$

为了得到系数a和b，需要对局部窗口内的代价函数进行最小化操作：

$$E(a_k, b_k) = \sum_{i\in\omega_k} ((a_k I_i + b_k - p_i)^2 + \varepsilon a_k^2) \tag{5-10}$$

通过最小二乘法得到最终的系数：

$$a_k = \frac{\frac{1}{|\omega|}\sum_{i\in\omega} I_i p_i - \mu_k \overline{p_k}}{\sigma_k^2 + \varepsilon} \tag{5-11}$$

$$b_k = \overline{p_k} - a_k \mu_k \tag{5-12}$$

其中，$|\omega|$为局部窗口ω_k内包含的像素点个数，μ_k为I在窗口内的期望，σ_k^2为对应的方差，$\overline{p_k}$是目标图像在窗口内的期望。

导向滤波图如图 5-2 所示。

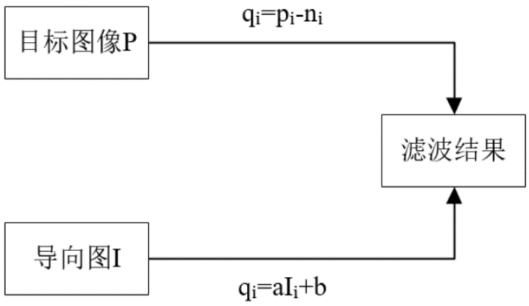


图 5-2 导向滤波图

在处理雨雪雾天气的图像时，可以将图像中的信息分为高频部分和低频部分，其中高频部分指的是噪声、雨雪和图像的边缘信息等。为了将图像进行去雨雪操作，需要将高频部分中的雨雪部分与边缘信息分隔开，然后将不含雨雪

的部分与低频部分进行结合，从而得到无雨雪的图像，其流程如图 5-3 所示。

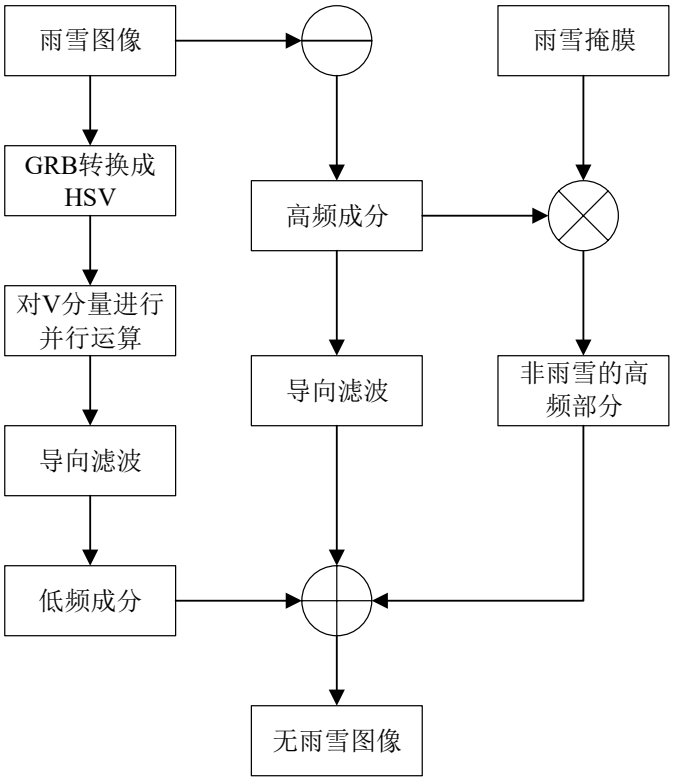
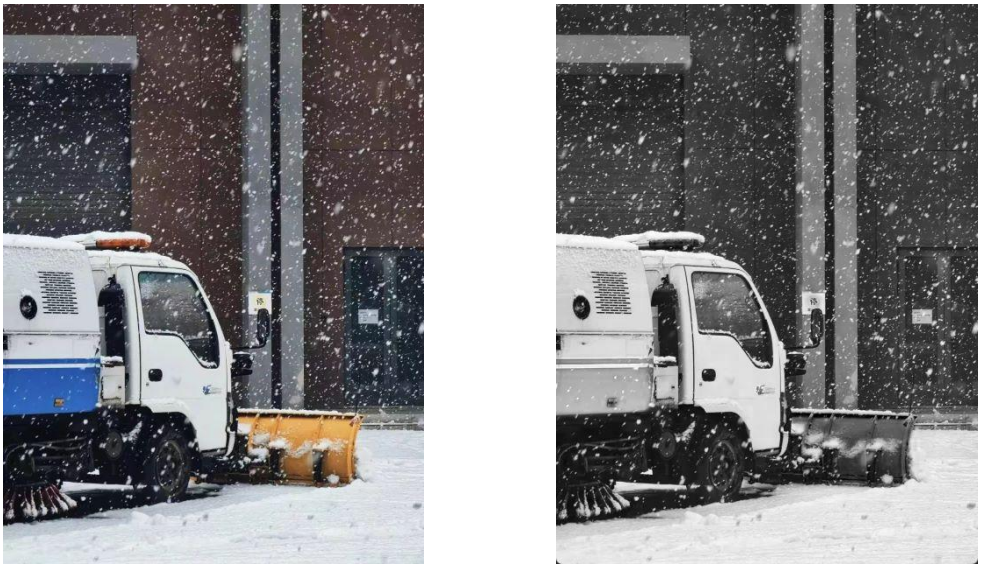


图 5-3 去雨雪流程图

去雨雪的具体操作步骤如下：

（1）颜色空间转换

将图像首先转换到 HSV 空间，图 5-4 为图像的颜色空间转换结果。

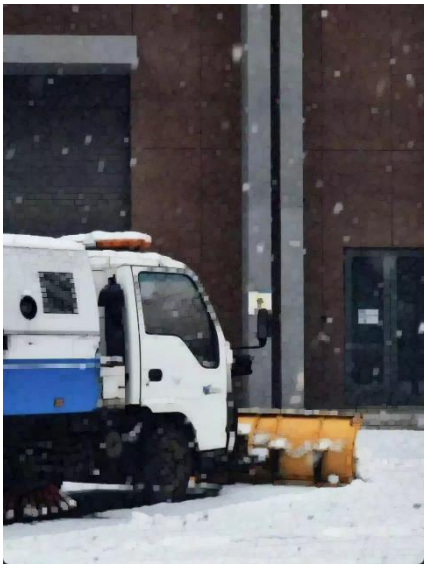


（a）原图像　　（b）V 空间图像

图 5-4 图像颜色空间转换图

（2）求解低频图像

利用开运算去除图像中的雨痕或雪花，使用导向滤波技术对图像进行平滑，将平滑后的分量进行合并得到最终的低频图像，其结果如图 5-5 所示。



（a）开运算图像

（b）低频成分图像

图 5-5 低频图像效果图

（3）求解高频部分

为了获取高频成分，利用原图像除去低频图像，即可得到具有雨雪和边缘信息的高频图像，其结果如图 5-6 所示。

图 5-6 高频图像

（4）剔除雨雪的高频成分

为保留图像的边缘信息，利用雨雪亮度高于背景亮度的特性生成非雨雪掩

膜。首先选取图像中三通道亮度值最大的像素值，然后与设定的阈值δ进行对比，再与高频成分相乘，可得到非雨雪的高频信息，其计算如公式（5-13）所示：

$$I_{\overline{rs}}(x,y)=\begin{cases}0 & , \ \max\limits_{c}\left(I^{c}(x,y)\right)\geq\delta \\ 1 & , \ else\end{cases} \tag{5-13}$$

其中，c表示颜色三通道，$c\in\{R,G,B\}$；x和y表示图像的位置坐标。

高频成分主要包括图像边缘成分和雨雪成分，其结果如图 5-7 所示。

（a）基于导向滤波的高频成分

（b）基于掩膜的高频成分

图 5-7 图像边缘的高频成分效果图

（5）生成去雨雪图像

将得到的低频成分与具有高频成分的边缘信息相结合，即可得到去雨雪的图像，去雨雪图像的效果图如图 5-8 所示。

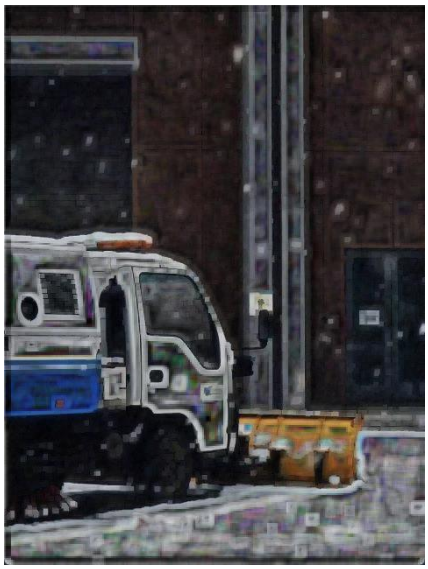


图 5-8 去雨雪效果图

从图5-8中可以看出，操作后的图像具有非常好的去雨雪效果，虽然会一定程度上损失图像的像素，但其去雨雪的效果有利于后续的目标识别检测。

5.3 算法改进

为了很好的在雨雪雾天气下对目标进行识别，只对图像进行清晰化处理还是远远不够的，因此本章在已改进的基础上加入一个可微图像处理（Differentiable Image Processing, DIP）模块，并且其参数由 CNN-PP 来进行预测，通过采用端到端的方式联合学习 CNN-PP 和 YOLOv7，从而确保 CNN-PP 可以适配 DIP，达到以弱监督的方式提高图像的检测效果。DIP 和 CNN-PP 的结构如图 5-9 所示。

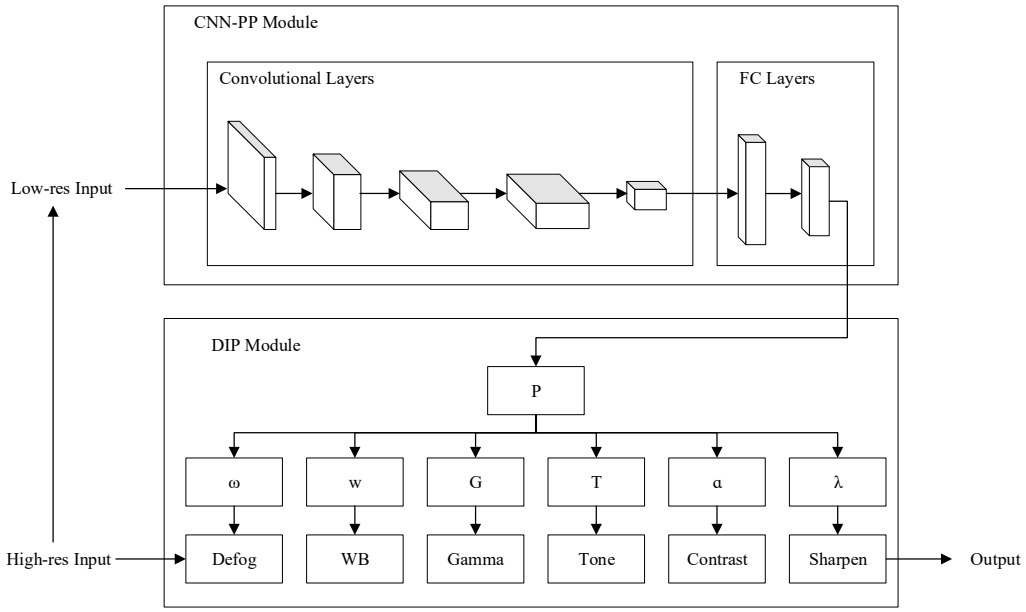


图 5-9 DIP 和 CNN-PP 结构图

如图 5-9 所示，整个流水线由一个基于 CNN 的参数预测器、一个可微分图像处理模块和一个检测网络组成。首先将输入图像的大小调整为 256×256，然后将其输入到 CNN-PP 中，接着通过得到的结果预测 DIP 的参数，最后将 DIP 模块滤波后的图像作为 YOLOv7 检测器的输入。

5.3.1 DIP 模块

DIP 模块由 6 个具有可调超参数的可微滤波器组成，包括 Defog、白平衡、伽玛、对比度、Tone 和 Sharpen。其中如 WB、伽玛、对比度和 Tone，可以表示为像素级滤波器。像素级滤波器通过将输入像素值$P_i = \{R, G, B\}$映射到输出像素值$P_0 = \{R_0, G_0, B_0\}$，其中$\{R, G, B\}$分别表示红色、绿色和蓝色三种颜色的通道

值，通过简单的乘法和幂运算可以得知三颜色的映射函数对输入图像和参数而言都是可微的，其输入映射函数和输出映射函数如公式（5-14）和公式（5-15）所示：

$$En(P_i) = P_i \times \frac{EnLum(P_i)}{Lum(P_i)} \tag{5-14}$$

$$P_0 = \frac{1}{T_L}\sum_{j=0}^{L-1} clip(L \times P_i - j, 0, 1)t_k \tag{5-15}$$

其中，$Lum(P_i)$、$EnLum(P_i)$和T_L的定义如公式（5-16）到公式（5-18）所示：

$$Lum(P_i) = 0.27R_i + 0.67G_i + 0.06B_i \tag{5-16}$$

$$EnLum(P_i) = \frac{1}{2}\left[1 - \cos\left(\pi \times Lum(P_i)\right)\right] \tag{5-17}$$

$$T_L = \{t_0, t_1, \cdots, t_{L-1}\} \tag{5-18}$$

5.3.2 CNN-PP 模块

在相机图像信号处理过程中，为了提高图像质量，通常会采用可调滤波器。然而这些滤波器的超参数需要专业人员通过目视检查手动调整，而这种方法既不方便又费时费力，因此寻找适合各种场景的超参数变得尤为重要。CNN-PP 通过了解图像的全局内容来预测 DIP 的参数，可以大大节省计算成本，对于给定任意分辨率的图像，可以通过使用双线性插值将其降采样到 256×256 分辨率。

CNN-PP 网络是作为一个复杂的多级结构，它由 5 个卷积块以及 2 个完整的连接层构建而成，可以检测到 DIP 模块的超参值。

5.4 检测结果

5.4.1 模块对比实验

自然图像质量评估器（Natural Image Quality Evaluator, NIQE）通过使用高斯混合模型（Gaussian Mixture Model, GMM）建立自然图像特征的概率分布，通常用于输入图像质量的评估。其计算如公式（5-19）所示：

$$NIQE = \sqrt{(v_1 - v_2)^{T-1}\left(\frac{\Sigma_1 + \Sigma_2}{2}\right)^{-1}} \tag{5-19}$$

其中，v_1、v_2和Σ_1、Σ_2分别为自然图像的多元高斯模型和失真图像的多元高斯模型的均值向量和协方差矩阵。

盲无参考图像控件质量评估器（Blind Referenceless Image Spatial Quality Evaluator, BRISQUE）使用支持向量机来学习图像质量与图像特征之间的关系。

为了验证所提算法的有效性，本文在多个数据集上与几种先进算法进行了定量和定性比较，包括 PSD[79]，EPDN[80]，MSCNN[81]，AOD-Net[82]，Dehaze-Net[83]，Refine-Net[84]。本文使用 SOTS 数据集，通过采用自然图像质量评估器和盲无参考图像空间质量评估器来衡量算法的性能，这两种评价指标的结果参数越低越好，且能够全面反映出算法的性能。实验结果如表 5-1 所示。

表 5-1 SOTS 上的性能对比

算法	NIQE↓	BRISQUE↓
PSD	3.190	21.326
EPDN	3.554	25.319
MSCNN	3.427	23.111
AOD-Net	3.616	29.185
Dehaze-Net	3.398	26.028
Refine-Net	3.363	20.739
本文算法	**3.255**	**21.147**

从表 5-1 可以看出，与现有的几种先进算法相比，本文算法在 NIQE 上仅比 PSD 算法高 0.065，效果位居第二，且比其余六种算法的平均值高 0.141；在 BRISQUE 上仅比 Refine-Net 算法高 0.412，效果也位居第二，且比其余六种算法的平均值高 3.138。实验结果进一步验证了该方法具有不错的效果。

5.4.2 目标检测实验结果

为了体现本章改进算法对常见交通环境、夜间交通环境和雨雪雾交通环境下的目标检测都具有更好的检测效果，本章对数据集 C 进行实验，对比原始 YOLOv5 算法、原始 YOLOv7 算法和 AWTD-YOLOv7 算法，其实验结果如表 5-2 所示。

表 5-2 检测结果分析

模型	P	R	mAP.5	mAP.95
YOLOv5	70.5	41.1	45.5	26.3
YOLOv7	71.8	52.8	57.3	36.3
AWTD-YOLOv7 算法	**75.3**	**69.4**	**69.2**	**42.3**

从表 5-2 中可以看出，AWTD-YOLOv7 算法在检测精度上比 YOLOv5 和

YOLOv7 高 23.7%和 11.9%。

为了更直观地证明改进算法的有效性，对表 5-1 中的算法进行了检测图对比。检测结果对比图如图 5-10 所示。

（a）YOLOv5

（b）YOLOv7

（c）AWTD-YOLOv7 算法

图 5-10 检测结果对比图

从图 5-10 中可以看出，YOLOv5 原模型算法存在较为严重的漏检；YOLOv7 原模型算法虽然具有较低的漏检率，但是却存在很明显的错检情况；本章改进的 AWTD-YOLOv7 算法不仅没有出现错检情况，而且具有较低的漏检率。因此本章改进的 AWTD-YOLOv7 算法更适用于雨雪雾等特殊天气情况下的

目标检测任务。

为了验证 AWTD-YOLOv7 算法不仅适用于雨雪雾等特殊天气下的检测，而且也适用于常见环境和夜间环境下的检测，本章对数据集 D 进行目标检测，通过对比 NTD-YOLOv7 算法和本章改进的 AWTD-YOLOv7 算法，其实验结果如表 5-3 所示。

表 5-3 改进算法的实验结果对比分析

模型	P	R	mAP.5	mAP.95
NTD-YOLOv7	74.7	60.2	63.9	39.8
AWTD-YOLOv7	**75.3**	**69.4**	**69.2**	**42.3**

从表 5-3 中可以看出，相较于 NTD-YOLOv7 算法，本章改进的 AWTD-YOLOv7 算法分别在 P、R、mAP.5、mAP.95 上提高了 0.6%、9.2%、5.3%、2.5%。

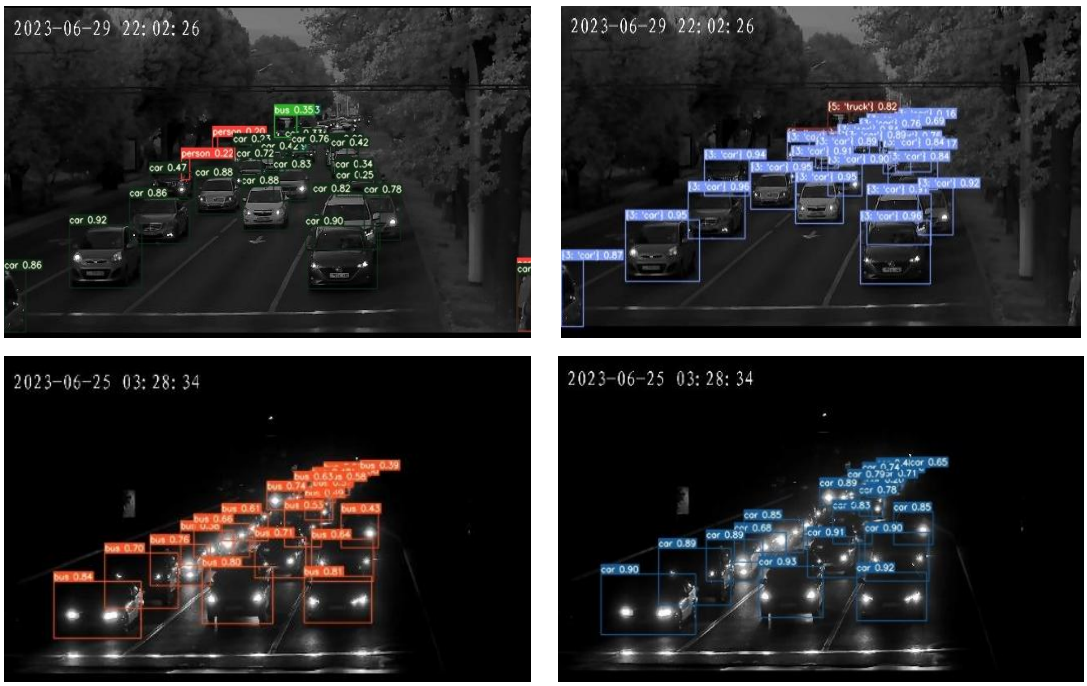
为了更好的展示 AWTD-YOLOv7 算法具有很好的检测效果，图 5-11 为常见交通环境下 NTD-YOLOv7 算法和 AWTD-YOLOv7 算法的检测效果对比图。

NTD-YOLOv7　　AWTD-YOLOv7

图 5-11 常见环境下 NTD-YOLOv7 和 AWTD-YOLOv7 检测效果对比图

从上面的结果对比图可以看出，AWTD-YOLOv7 算法在常见交通环境下仍然具有很好的检测效果。

图 5-12 为夜间交通环境下 NTD-YOLOv7 算法和 AWTD-YOLOv7 算法的检测效果对比图。

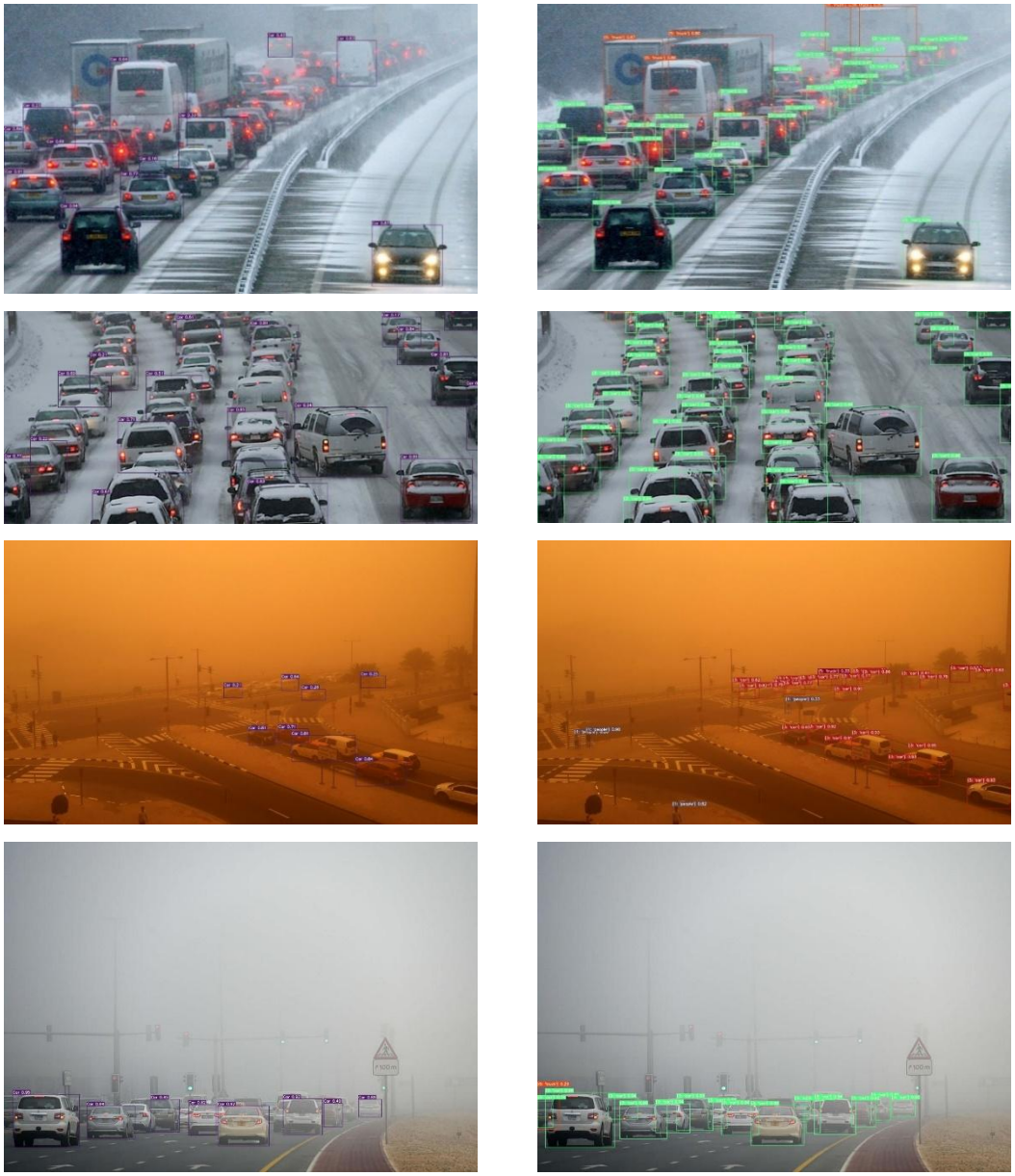


NTD-YOLOv7　　AWTD-YOLOv7

图 5-12 夜间环境下 NTD-YOLOv7 和 AWTD-YOLOv7 检测效果对比图

从上面的结果对比图可以看出，AWTD-YOLOv7 算法在夜间交通环境下仍然具有很好的检测效果。

图 5-13 为特殊天气交通环境下 NTD-YOLOv7 算法和 AWTD-YOLOv7 算法的检测效果对比图。



NTD-YOLOv7　　AWTD-YOLOv7

图 5-13 特殊天气下 NTD-YOLOv7 和 AWTD-YOLOv7 检测效果对比图

从图 5-13 的结果对比图可以看出，NTD-YOLOv7 算法存在较多的漏检情况，并不能很好的处理雨雪雾特殊天气交通环境下的目标识别任务，而 AWTD-YOLOv7 算法在雨雪雾特殊天气交通环境下则具有很好的检测效果。

通过分析图 5-11、图 5-12 和图 5-13，在面对常见交通环境、夜间交通环境和雨雪雾等特殊天气下的交通环境时，本章改进的 AWTD-YOLOv7 算法都具有很好的检测能力。

5.5 本章小结

为了解决在雨雪雾等更为恶劣的交通环境中的车辆行人检测，本章在 NTD-YOLOv7 算法的基础上进一步提出了 AWTD-YOLOv7 算法。首先介绍了针对雨雪雾环境下目标检测的原理，接着通过分别使用暗通道先验算法和基于引导滤波的操作实现去雾和去雨雪的目的，接着在已经实现的去雨雪雾的基础上，通过使用 DIP 模块和 CNN-PP 模块的结合提高在雨雪雾环境下目标检测准确度。实验结果表明，本章改进的算法不仅在雨雪雾环境下具有很好的检测效果，而且在面对常见环境和夜间环境的交通道路时仍然具有很好的检测效果。

第6章 结论与展望

6.1 结论

随着科学技术的快速发展和人民生活水平的提高，交通道路目标检测技术也逐渐收到广泛关注，但其非人工的技术也使得大多数人们对交通道路目标检测技术并不放心，因此交通道路目标检测技术也需要更好的智能性和安全性。本文针对交通道路目标检测在各种交通环境下所面临的问题，提出了一种基于YOLOv7 的目标检测算法。通过对常见交通环境、夜间交通环境和雨雪雾等复杂交通环境的分析，使用一种统一算法对三种不同环境的目标进行检测。

交通道路目标检测技术发展促进了卷积神经网络的进步，本文根据交通道路目标检测技术的需求和不同交通环境目标检测的复杂程度，选择了在检测精度、检测速度和移动设备需求上更具有优势的 YOLOv7 算法，并从以下三种不同交通环境对模型算法进行了改进研究：

（1）针对常见交通环境的目标检测，提出了CTD-YOLOv7算法。对SENet模块进行改进，注重各通道之间特征信息的关联性，提高对各个通道特征的辨别能力，选取NAM注意力机制，使用批量归一化的比例因子来表示权重的重要性，避免与 SE 和 CBAM 一样在每个通道中都添加完整的连接层或者卷积层；

（2）针对夜间环境的目标检测，在 CTD-YOLOv7 算法的基础上，提出了NTD-YOLOv7 算法。在 CTD-YOLOv7 的基础上对图像进行预处理操作，通过Gamma 变换等方法提高图像的清晰度，使用 ShuffleNetv2×1.5 网络结构来解决不同群卷积之间所导致的信息不交互问题，并对损失函数和激活函数进行了改进；

（3）针对雨雪雾等复杂环境的目标检测，在 NTD-YOLOv7 算法的基础上，提出了 AWTD-YOLOv7 算法。在 NTD-YOLOv7 的基础上，使用暗通道先验算法和基于引导滤波的操作实现去雾和去雨雪的目的，通过使用DIP模块和CNN-PP模块的结合提高目标检测的准确度。

最后实验结果表明，改进后的AWTD-YOLOv7算法相较于原始YOLOv7算法和其他算法，具有更高的检测精度、更快的检测速度和更好的检测效果，而且能在实际检测过程中减少对运动目标错检漏检的情况，证明本文改进的AWTD-YOLOv7算法具有实用性和有效性。

6.2 展望

随着技术的不断发展，交通道路目标检测技术也在不断更新，其中对交通道路安全的要求也越来越高，但是目前受到各种因素的影响，基于交通道路的目标检测技术并不能达到完美，而且其检测精度与检测速度都需要不断完善。本文通过对 YOLOv7 算法进行改进，并通过对比其他算法进行优化，虽然取得了一定成果，但仍然存在一些不足之处，还可以从以下几个方面开展进一步的研究：

（1）针对检测速度与检测精度的同步优化。本文虽然在检测精度与检测速度上有一定的提升，但是并不能很好面对一些突发情况，因此可以对网络模型的整体结构进行优化，在不损失检测精度与检测速度的前提下，使算法模型能够更好的适用于面对各种情况的移动端。

（2）数据集的扩展和实用性。通过使用各种交通道路下的图像数据集，使本文的算法具有更好的稳定性，以此来提升算法在各种环境下的检测能力。

（3）移动端设备的轻量化处理。本文针对雨雪雾等特殊环境的目标检测都是基于对图像预处理的前提，后续可以通过对网络模型的改进减去这一步骤，进而能使算法模型在提升检测精度和检测速度的同时，能够降低对移动设备软硬件的要求。

参考文献

[1] Arora N, Kumar Y, Karkra R, et al. Automatic Vehicle Detection System in Different Environment Conditions Using Fast R-CNN [J]. Multimedia Tools and Applications, 2022, 81(13): 18715-18735.

[2] Singh J P, Singh U P, Jain S. Model-based Person Identification in Multi-gait Scenario Using Hybrid Classifier [J]. Multimedia Systems., 2023, 29(03): 1103-1116.

[3] Zhang Y L, Lu Y, Zhu W Q, et al. Traffic Sign Detection and Recognition based on Multi-size Feature Extraction and Enhanced Feature Fusion Module [J]. Intelligent & Fuzzy Systems, 2023, 45(02): 2993-3004.

[4] Wei C, Han F, Fan Z Z, et al. Efficient License Plate Recognition in Unconstrained Scenarios [J]. Visual Communication and Image Representation, 2024, 104: 104314-104314.

[5] Xiong H, Yan Y, Sun L F, et al. Detection of Driver Drowsiness Level Using a Hybrid Learning Model based on ECG Signals [J]. Biomedical Engineering, 2023, 69(02): 151-165.

[6] Chen X Y, Xiang S M, Liu C L, et al. Vehicle Detection in Satellite Images by Hybrid Deep Convolutional Neural Networks [J]. Geoscience and Remote Sensing Letters, 2014, 11(10): 1797-1801.

[7] Ojala T, Pietikainen M, Maenpaa T. Multiresolution Grayscale and Rotation Invariant Texture Classification with Local Binary Patterns [J]. IEEE Transactions on Pattern Analysis and Machine Intelligent, 2002, 24(07): 971-987.

[8] David C, Travieso C M, Alonso J B. Thermal Face Verification Based on Scale-invariant Feature Transform and Vocabulary Tree: Application to Biometric Verification Systems [C]. Bio Inspired System and Signal Processing, 2012: 475-481.

[9] Dalal N, Triggs B. Histograms of Oriented Gradients for Human Detection [C]. Computer Vision and Pattern Recognition, 2005: 886-893.

[10] Bay H, Ess A, Tuytelaars T, et al. Speeded-Up Robust Features [J]. Computer

Vision and Image Understanding, 2008, 110(03): 346-359.

[11] Cortes C, Vapnik V. Support Vector Networks [J]. Machine Learning, 1995, 20(03): 273.

[12] Jin W J, Tang W H, Qian T, et al. Fault Diagnosis of High-Voltage Circuit Breakers Using Wavelet Packet Technique and Support Vector Machine [J]. Proceedings Journal, 2017, 1: 170-174.

[13] Freund Y. An Adaptive Version of the Boost by Majority Algorithm [J]. Machine Learning, 2001, 43(03): 293-318.

[14] 文学志, 方巍, 郑钰辉. 一种基于类 Haar 特征和改进 AdaBoost 分类器的车辆识别算法[J]. 电子学报, 2011, 39(05): 1121-1126.

[15] Felzenszwalb P, Girshick R, Mcallester D, et al. Visual Object Detection with Deformable Part Models [J]. Communications of the ACM, 2013, 56(09): 97-105.

[16] Glorot X, Bordes A, Bengio Y. Deep Sparse Rectifier Neural Networks [C]. Artificial Intelligence and Statistics, 2011, 15: 315-323.

[17] 李宏伟, 吴庆祥. 智能传感器中神经网络激活函数的实现方案[J]. 传感器与微系统, 2014, 33(01): 46−48.

[18] Nair V, Hinton G E. Rectified Linear Units Improve Restricted Boltzmann Machines [C]. Machine Learning, 2010: 807−814.

[19] Maas A L, Hannun A Y, Ng A Y. Rectifier Nonlinearities Improve Neural Network Acoustic Models [C]. Machine Learning, 2013: 456-462.

[20] Ramachandran P, Zoph B, Le Q V. Searching for an Activation Functions [EB/OL]. arXiv: 1710.05941, 2017. https://arxiv.org/abs/1710.05941.

[21] Andrew H, Mark S, Grace C, et al. Searching for MobileNetV3 [EB/OL]. arXiv: 1905.02244, 2019. https://arxiv.org/abs/1905.02244.

[22] Liu W, Angulov D, Erhan D, eg.al. SSD: Single Shot MultiBox Detector [C]. Computer Science, 2016, 9905: 21-37.

[23] Redmon J, Divvala S, Girshick R, et al. You Only Look Once: Unified, Real-Time Object Detection [C]. Computer Vision and Pattern Recognition, 2016: 779-788.

[24] Redmon J, Farhadi A. YOLO9000: Better, Faster, Stronger [C]. Computer Vision & Pattern Recognition, 2017: 6517-6525.

[25] Redmon J, Farhadi A. YOLOv3: An Incremental Improvement [C]. Computer Vision & Pattern Recognition, 2018: 1-4.

[26] Bochkovskiy A, Wang C Y, Liao H. YOLOv4: Optimal Speed and Accuracy of Object Detection [J]. Computer Vision & Pattern Recognition, 2020, 10934.

[27] Jocher G. YOLOv5 [BE/OL]. https://github.com/ultralytics/YOLOv5. 2022-11-22.

[28] Li H L, Li J, Wei H B, et al. Slimneck by GSConv: A Better Design Paradigm of Detector Architectures for Autonomous Vehicles [EB/OL]. arXiv: 2206.02424, 2022. https://arxiv.org/abs/2206.02424.

[29] Wang C Y, Bochkovskiy A, Liao H. YOLOv7: Trainable Bag of Freebies Sets New State of the Art for Real-Time Object Detectors [EB/OL]. arXiv: 2207.02696, 2022. https://arxiv.org/abs/2207.02696.

[30] Gallagher J. How to Train an Ultralytics YOLOv8 Oriented Bounding Box Model [EB/OL]. [2024-02-06]. https://blog.roboflow.com/train-yolov8-obb-model/.

[31] Wang C Y, Yeh I H, Liao H Y M. YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information[J]. arXiv:2402.13616, 2024.

[32] Wang A, Chen H, Liu L, et al. YOLOv10: Real-time End to End Object Detection[J]. arXiv:2405.14458, 2024.

[33] Khanam R, Hussain M. YOLOv11: An Overview of the Key Architectural Enhancements [EB/OL]. arXiv:2410. 17725. 2024.

[34] Cirshick R, Donahue J, Darrell T. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation [C]. Computer Vision and Pattern Recognition, 2014: 580-587.

[35] Cirshick R. Fast R-CNN [C]. Computer Vision, 2015: 1440-1916.

[36] Ren S Q, He K M, Cirshick R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks [J]. Pattern Analysis and Machine Intelligence, 2017, 39(06): 1137-1149.

[37] Kong T, Yao A B, Chen Y R, et at. HyperNet: Towards Accurate Region Proposal Generation and Joint Object Detection [C]. Computer Vision and Pattern Recognition, 2016: 845-85.

[38] Lin T Y, Dollar P, Girshick R, et al. Feature Pyramid Networks for Object Detection

[EB/OL]. arXiv: 1612.03144, 2016. https://arxiv.org/abs/1612.03144.

[39] He K M, Gkioxari G, Dollar P, et al. Mask R-CNN [C]. Computer Vision, 2017: 2980-2988.

[40] Yu F, Xian W, Chen Y, et al. BDD100K: A Diverse Driving Video Database with Scalable Annotation Tooling [EB/OL]. arXiv: 1805.04687, 2018. https://arxiv.org/abs/1805.04687.

[41] Zhu Z, Liang D, Zhang S H, et al. Traffic-Sign Detection and Classification in the Wild [C]. Computer Vision and Pattern Recognition, 2016: 2110-2118.

[42] VOC 2007 Dataset [EB/OL]. https://pjreddie.com/projects/pascal-voc-dataset-mirror. 2021-10-06.

[43] Kenk M A, Hassaballah M S. DAWN: Vehicle Detection in Adverse Weather Nature Dataset [EB/OL]. arXiv: 2008.05402, 2008. https://arxiv.org/abs/2008.05402.

[44] Cheon M, Lee W, Yoon C, et al. Vision-based Vehicle Detection System with Consideration of the Detecting Location [J]. Intelligent Transportation Systems, 2012, 13(03): 1243-1252.

[45] Bougharriou S, Hamdaoui F, Mtibaa A. Algorithmic Optimization of Vehicle Detection System based on Hybrid Metaheuristic with Machine Learning [C]. Automatic Control and Computer Engineering, 2022: 84-88.

[46] Liao J B, Yuan H X, Li Y B, et al. Research on Vehicle Target Recognition based on Infrared Images Processing [C]. International Conference on Optics Machine Vision, 2022, 12173: 304-309.

[47] 宋璠, 王世峰. 基于可变形部件模型 HOG 特征的人形目标检测[J]. 应用光学, 2016, 37(03): 380-384.

[48] 张艳邦, 张芬, 张姣姣. 基于 SVM 和背景模型的显著性目标检测算法[J]. 电子设计工程, 2022, 30(05): 17-21+27.

[49] Hu J, Shen L, Albanie S, et al. Squeeze and Excitation Networks [C]. Computer Vision and Pattern Recognition, 2022, 42(08): 2011-2023.

[50] Liu Y C, Shao Z R, Teng Y Y, et al. NAM: Normalization based on Attention Module [EB/OL]. arXiv: 2111.12419, 2021. https://arxiv.org/abs/2111.12419.

[51] 王婧媛, 方健. 基于 YOLOv5 的密集场所人数估计方法[J]. 吉林大学学报, 2021, 39(06): 682-687.

[52] 章程军, 胡晓兵, 牛洪超. 基于改进 YOLOv5 的车辆目标检测研究[J]. 四川大学学报, 2022, 59(05): 79-87.

[53] 王鹏, 王玉林, 焦博文等. 基于 YOLOv5 的道路目标检测算法研究[J]. 计算机工程与应用, 2023, 59(01): 117-125.

[54] 顾德英, 罗聿伦, 李文超. 基于改进 YOLOv5 算法的复杂场景交通目标检测[J]. 东北大学学报. 2022, 43(08): 1073-1079.

[55] 关昊, 黎海涛, 康衡等. 基于 YOLOv5 的 UAV 夜间目标检测技术研究[J]. 无线电通信技术, 2023, 49(02): 345-350.

[56] 曹伊宁, 李超, 彭雅坤. 基于改进 YOLOv7 的夜间行人检测算法[J]. 长江信息通信, 2022, 35(10): 57-60.

[57] 刘浩瀚, 樊一鸣, 贺怀清等. 改进 YOLOv7-tiny 的目标检测轻量化模型[J]. 计算机工程与应用, 2023, 59(14): 166-175.

[58] Li H L, Li J, Wei H B, et al. Slimneck by GSConv: A Better Design Paradigm of Detector Architectures for Autonomous Vehicles [EB/OL]. arXiv: 2206.02424, 2022. https://arxiv.org/abs/2206.02424.

[59] Sifire L, Mallat S. Rigid-Motion Scattering for Texture Classification [EB/OL]. arXiv: 1403.1687, 2014. https://arxiv.org/abs/1403.1687.

[60] Hu J Y, Liu B J, Peng S H. Forecasting Salinity Time Series Using RF and ELM Approaches Coupled with Decomposition Techniques [J]. Stochastic Environmental Research and Risk Assessment, 2019, 33(4-6): 1117-1135.

[61] Wang C Y, Liao H Y M, Wu Y H et al. CSPNet: A New Backbone That Can Enhance Learning Capability of CNN [C]. Computer Vision Pattern Recognition Workshops, 2020: 390-391.

[62] Zhang X Y, Zhou X Y, Lin M X, et al. Shufflenet: An Extremely Efficient Convolutional Neural Network for Mobile Devices [C]. Computer Vision and Pattern Recognition, 2018: 6848-685.

[63] 侯志强, 刘晓义, 余旺盛等. 使用 GIoU 改进非极大值抑制的目标检测算法[J].

电子学报, 2021, 49(04): 696-705.

[64] Zheng Z H, Wang P, Liu W, et al. Distance-IOU Loss: Faster and Better Learning for Bounding Box Regression [C]. Artificial Intelligence, 2020: 12993-13000.

[65] 王永顺, 贾文杰, 王晨飞等. 基于改进 YOLOv3 的车辆识别方法[J]. 激光与光电子学进展, 2021, 58(16): 240-247.

[66] Zhang Z H, Ren W, Zhang Z, et al. Focal and Efficient IoU Loss for Accurate Bounding Box Regression [EB/OL]. arXiv: 2101.08158, 2021. https://arxiv.org/abs/2101.08158.

[67] 李功, 赵巍, 刘鹏等. 一种用于目标跟踪边界框回归的光滑 IoU 损失[J]. 自动化学报, 2023, 49(02): 288-306.

[68] Tan R T. Visibility in Bad Weather from A Single Image [C]. Computer Vision and Pattern Recognition, 2008: 1-8.

[69] Tarel J P, Hautiere N. Fast Visibility Restoration from A Single Color or Gray Level Image [C]. Computer Vision, 2009: 2201-2208.

[70] He K M, Sun J, Tang X O. Single Image Haze Removal Using Dark Channel Prior [C]. Computer Vision and Pattern Recognition, 2009: 1956-1963.

[71] He K M, Sun J, Tang X O. Guided Image Filtering [J]. IEEE Transactions on Pattern Analysis & Machine Intelligence, 2013, 35(06): 1397-409.

[72] 陈书贞, 任占广, 练秋生. 基于改进暗通道和导向滤波的单幅图像去雾算法[J]. 自动化学报, 2016, 42(03): 455-465.

[73] 张晓刚, 唐美玲, 陈华等. 一种结合双区域滤波和图像融合的单幅图像去雾算法[J]. 自动化学报, 2014, 40(08): 1733-1739.

[74] 陈丹丹, 陈莉, 张永新等. 修正大气耗散函数的单幅图像去雾[J]. 中国图象图形学报, 2017, 22(06): 787-796.

[75] Guo X J, Wang X G, Yang L, et al. Robust Foreground Detection Using Smoothness and Arbitrariness Constraints [C]. Computer Vision, 2014, 8695(07): 535-550.

[76] Shijila B, Tom A J, George S N. Moving Detection by Low Rank Approximation and l 1-TV Regularization on RPCA Framework [J]. Journal of Visual Communication and Image Representation, 2018, 56: 188-200.

[77] 李艳荻, 徐熙平. 基于超像素时空特征的视频显著性检测方法[J]. 光学学报，2019, 39(01): 315-322.

[78] Li Y, Liu G C, Liu Q S, et al. Moving Object Detection Via Segmentation and Saliency Constrained RPCA [J]. Neurocomputing, 2019, 323: 352-362.

[79] Chen Z Y, Wang Y C, Yang Y, et al. PSD： Principled Synthetic-to-real Dehazing Guided by Physical Priors[C]. Computer Vision and Pattern Recognition. 2021: 7176-7185.

[80] Qu Y Y, Chen Y Z, Huang J Y, et al. Enhanced Pix2pix dehazing network[C]. Computer Vision and Pattern Recognition. 2019: 8152-8160.

[81] Ren W Q, Pan J S, Zhang H, et al. Single Image Dehazing Via Multi Scale Convolutional Neural Networks with Holistic Edges[J]. Computer Vision. 2020, 128(01): 240-259.

[82] Li B Y, Peng X L, Wang Z Y, et al. AOD-Net: All IN One Dehazing Network[C]. Computer Vision, 2017: 4780-4788.

[83] Cai B L, Xu X M, Jia K, et al. Dehaze Net: An End-to-end System for Single Image Haze Removal[J]. Image Processing, 2016, 259(11): 5187-5198.

[84] Zhao S Y, Zhang L, Shen Y, et al. Refine Net: A Weakly Supervised Refinement Frame Work for Single Image Dehazing[J]. Image Processing, 2021, 30: 3391-3404.

攻读硕士学位期间的研究成果

发表学术论文：

[1] **Zheng Bowen**, Lu Hucai, Zhu Shengbo, Chen Xinqiang and Xing Hongwei. Night Target Detection Algorithm Based on Improved YOLOv7[J]. Scientific Reports. 2024, 14(01): 15771. (**SCI 已检索**)

[2] **Zheng Bowen**, Lu Huacai and Li Ming. Autonomous Driving Target Detection Based on Improved YOLOv7[C]. *2024 3rd Electronic Information Engineering, Big Data, and Computer Technology*. 2024, 131811U: 1-7. (**EI 已检索**)

[3] **Zheng Bowen**, Lu Huacai and Li Peiyang. Research on Moving Target Detection Technology Based on Yolov5 Model[C]. *2023 2nd Artificial Intelligence and Intelligent Information Processing*. 2023: 305-309. (**EI 已检索**)

[4] **郑泊文**, 李佩阳, 陆华才. 面向复杂运动场景的目标检测技术研究[J]. *铜陵学院学报*. 2024, 23(01): 101-105.

致 谢

时光荏苒，如白驹过隙，我的硕士研究生生活即将画上句号。在这段宝贵的时光里，我收获了知识、成长与友谊，而这一切都离不开众多人的支持与帮助。此刻，我怀着无比感恩之心，向每一位在我求学历程中给予我关怀和帮助的人致谢。

首先，我要衷心感谢的是我的导师陆华才教授。从论文的选题、构思到撰写、修改，陆老师都给予了我悉心的指导。陆老师严谨的治学态度、精益求精的工作作风，一直激励着我不断努力，勇攀学术高峰。陆老师不仅在学术上对我严格要求，还在生活中给予我无微不至的关怀。他的言传身教，让我不仅学到了专业知识，更学会了如何做人、如何做学问。师恩如山，我将永远铭记在心。

其次，我要感谢我的父母和妹妹，他们是我最坚实的后盾。在我漫长的求学道路上，无论遇到多少困难和挫折，他们始终无条件的给予我支持和鼓励。我还要感谢我的舅舅，在我成长过程中，他总是在我最迷茫的时候指点教导我，给予我关心和帮助。

我要感谢我的同门李明虎和刘祥，他们在学习和生活中都给予了我很多帮助和照顾。我还要感谢我的师兄沈云畅、苏忠德、胡俊、汤祥宇和师姐杨冬雪、张彦，为我提供大量帮助和学习资源，教会我进行学术研究的方式方法。同时我还要感谢我的师弟陈新强、邢宏伟、刘闯、宋伟康、辛凯和师妹黄子甜，他们的活力和热情给实验室带来了新的气息。

我还要感谢研究生阶段认识的朋友们，感谢兄弟实验室的朱圣博、李明，感谢我的室友黄鹤鸣、周杨阳，感谢大家的相互鼓励和友好相处，让我的研究生生活充满快乐，祝他们在今后的工作生活中一帆风顺，一切都好。

再次向所有给予我帮助的人表示衷心的感谢！

尚德教学　唯实惟新