分类号：TP391
密　级：公开

学校代码：10363
学　　号：2210920110

安徽工程大学
Anhui Polytechnic University

硕士学位论文

题目 基于深度学习的交通标志检测方法研究

论 文 作 者　杜云龙
指 导 教 师　强俊 副教授 刘涛 工程师
学科（专业）　计算机技术
研 究 方 向　人工智能及应用

论文提交日期：2024 年 5 月 23 日

分类号：TP391　　单位代码：10363
密　级：公开　　学　　号：2210920110

题　　目　基于深度学习的交通标志检测方法研究

英文并列

题　　目　RESEARCH ON TRAFFIC SIGN DETECTION METHODBASED ON DEEP LEARNING

学生姓名：杜云龙
指导教师：强俊副教授/刘涛工程师
专　　业：计算机技术
研究方向：人工智能及应用

论文答辩日期：2024 年 5 月 24 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：杜云龙

日期：2024 年 5 月 23 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：杜云龙　　指导教师签名：

日期：2024 年 5 月 23 日　　日期：2024 年 5 月 23 日

基于深度学习的交通标志检测方法研究

摘 要

由于我国汽车保有量的快速增加，给道路交通安全带来了前所未有的压力。自动驾驶技术的出现，能为驾驶员提供决策信息以及有效减轻驾驶员在驾驶过程中的负担，从而降低交通事故发生的风险。交通标志检测作为自动驾驶技术的重要组成部分发挥了重要的作用。基于此，本文的主要工作内容如下：

（1）针对国内场景对交通标志数据集进行分析与改进。首先针对一些数据类别数据量少的问题，使用数据增强方式增加这些数据类别的数据量，使数据集更加均衡化。其次结合实际场景中可能出现的遮挡、光照变化等不同状况，对部分数据进行此种情况的图像扩充，以此来增加图像的多样性。

（2）针对 SSD 算法存在的对目标检测精度低和目标特征信息不足等问题，提出一种基于 SSD-RFC 的单步多目标检测算法。首先，用残差网络替换原有的 VGG-16，提高对特征的提取能力。随后，采用高效轻量的特征融合模块来融合高低层的特征信息。最后，在残差网络的卷积层末端添加 CBAM 注意力机制模块，以此减少特征在网络模型中产生的损失。经实验验证，上诉改进提高了算法的检测性能，较原网络 mAP 提升了 4.9%。但将改进后的算法应用在包含大量小型交通标志的 TT100K 数据集时发现检测精度出现了一定程度的下降。

（3）针对现有算法对现实场景拍摄的小型交通标志图像存在不易检测、背景复杂，模型复杂度高等问题，提出一种基于 YOLOv7-CCα的目标检测算法。首先在预测网络中引入集中式特征金字塔 CFP，加强特征的层内调节能力，有效保留边缘区域信息；其次，通过整合位置信息到通道注意力机制中的 CA 模块，可以提升模型对目标特征在不同位置上的关注度进而降低冗余检测，在几乎不带来计算开销的前提下有效提升网络的精度；最后，对于 CIOU 损失函数引入α-IOU，增加高 IOU 对象的损失和梯度的权重，来提高边界框回归的准确性。所提算法的 mAP 分别比原网络、SSD-RFC 提升了 4.6%、7.8%，同时模型参数量较 SSD-RFC 减少了 5.9M。

由实验结果可知，所提出的算法具有良好的检测精度和检测速率，有效解决了现有算法对交通标志检测存在的问题和不足。因此，本文研究内容具有一定的

理论意义和应用价值。

关键词：自动驾驶，深度学习，交通标志检测，SSD，YOLOv7

RESEARCH ON TRAFFIC SIGN DETECTION METHOD BASED ON DEEP LEARNING

ABSTRACT

Due to the rapid proliferation of vehicles in our country, unprecedented pressure has been exerted on road traffic safety. The advent of autonomous driving technology offers promising avenues for furnishing drivers with decision-making cues and alleviating the cognitive burden experienced during driving, thereby mitigating the risk of traffic accidents. Notably, traffic sign detection, constituting a pivotal facet of autonomous driving technology, assumes paramount significance. Against this backdrop, this study delineates its primary objectives as follows:

(1) Analyzing and improving the traffic sign dataset for domestic scenarios. Firstly, addressing the issue of insufficient data quantity for certain data categories by using data augmentation techniques to increase the volume, thus achieving a more balanced dataset. Secondly, considering potential scenarios such as occlusion and changes in lighting conditions in real-world situations, augmenting some data to simulate these conditions and enhance the diversity of images.

(2) Addressing the low accuracy and insufficient feature inform

ation in object detection of the SSD algorithm, proposing a single-s tep multi-object detection algorithm based on SSD-RFC. Initially, re placing the original VGG-16 with a residual network to enhance fe ature extraction capability. Then, utilizing an efficient and lightweig ht feature fusion module to integrate features from high and low la yers. Finally, adding a CBAM attention mechanism module to the e nd of the residual network's convolutional layers to reduce feature l oss within the network model. Experimental results show that the mentioned improvements enhance the algorithm's detection performa nce, with a 4.9% increase in mAP compared to the original networ k. However, when applying the improved algorithm to the TT100K dataset containing a large number of small traffic signs, a certain degree of decline in detection accuracy was observed.

(3) Addressing challenges such as difficulty in detecting small t raffic sign images captured in real-world scenes, complex backgroun ds, and high model complexity in existing algorithms, proposing an object detection algorithm based on YOLOv7-CCa. Firstly, introduc ing a centralized feature pyramid (CFP) into the prediction network to enhance the intra-layer adjustment capability of features and eff ectively retain edge area information. Secondly, integrating positiona l information into the channel attention (CA) module to increase th e model's focus on target features at different positions, thereby red

ucing redundant detections and effectively improving the network's accuracy with almost no additional computational overhead. Finally, introducing α-IOU into the CIOU loss function to increase the weight of losses and gradients for high IOU objects, enhancing the accuracy of bounding box regression. The proposed algorithm achieves mAP improvements of 4.6% and 7.8% compared to the original network and SSD-RFC, respectively, while reducing the model parameters by 5.9M compared to SSD-RFC.

According to the experimental results, the proposed algorithm has good detection accuracy and speed, effectively solving the problems and shortcomings of existing algorithms in traffic sign detection. Therefore, the research content of this article has certain theoretical significance and application value.

KEY WORDS: autonomous driving, deep learning, traffic sign detection, SSD, YOLOv7

目 录

第 1 章 绪论

1.1 研究背景及意义

近年来，随着我国国民生活水平的不断提高，汽车已经成为众多家庭不可或缺的常用交通工具。然后，我们在享受汽车带给我们生活便利的同时也面临着日趋严重的交通安全问题。根据公安部统计，截至 2023 年 9 月，全国私家车保有量 4.3 亿辆，全国汽车保有量超过百万辆的城市数量达到了 94 个，机动车驾驶人数超过 5.2 亿人。汽车保有量的不断攀升，使交通拥堵程度不断加剧，道路交通事故频发[1,2]，对人民的生命财产安全带来了巨大损失，阻碍了社会的进一步发展。

部分交通事故是由于车辆驾驶员在驾驶过程中反应不及或操作不当，而未按交通标志指示驾驶车辆。这是因为在驾驶机动车过程中，交通标志牌从进入驾驶员视野到离开是一个由远及近的过程，在此过程中交通标志牌在驾驶员的视野中逐渐由小变大。若交通标志处于小状态时检测识别系统便能对其进行有效检测，这样就能够有效降低驾驶员因疲劳、注意力不集中等因素而错过重要交通标志的可能性，为驾驶员预留更多的反应时间，从而减少交通事故。智能交通系统[3,4,5]的引入为解决这一问题提供了可能性，智能交通系统是利用先进的信息、通信、感知和控制技术以及交通管理原则对交通系统进行综合优化和智能化管理的系统，其在城市交通管理、交通安全保障和公共交通系统优化中都发挥了及其重要的作用。

交通标志检测技术[6]作为智能交通系统的重要技术组成部分可以帮助车辆实时识别并理解路标，辅助驾驶员及时做出正确的决策，提高行车安全性[7]。其首先通过捕获车载摄像头对交通标志图像，随后对捕获的图像进行图像处理以及模式识别，若成功识别到交通标志，智能交通系统会及时向驾驶员反馈交通信息，以此来向驾驶员进行提醒和预警。因此，对于交通标志检测的研究越来越受到广大研究者的关注，成为一个重要的研究方向。

对于交通标志的检测算法来说，准确性和实时性同等重要，一个检测精度差的交通标志检测算法会给驾驶员带来误导，影响正确的决策，容易引发交通事故。

而在车辆高速行驶的过程中，交通场景在不断的发生变化，交通标志检测算法若不能及时的对其进行检测也就失去了检测的意义。因此，在交通标志检测算法中，检测精度和实时性都决定了其能否运用在实际的驾驶场景中。

综上所述，探究具有高精确度、低时延的交通标志检测算法对车辆的安全驾驶具有非常重要的意义。

1.2 国内外研究现状

1.2.1 基于传统方法的交通标志检测

通常分为以下两种：

（1）基于颜色信息检测：基于颜色特征检测主要涉及以下步骤：①颜色空间转换：将图像从 RGB 空间转换为其他颜色空间，如 HSV[8]、HSI[9]、YCbCr[10]等。这些颜色空间可以更好的表示颜色信息。②阈值化：根据目标颜色的特征，在选择的颜色空间中设置阈值，将图像分割为目标颜色和非目标颜色的区域。③对分割后的图像进行形态学操作，如腐蚀、膨胀、开运算、闭运算等，以去除噪声或连接目标区域从而实现对目标的识别和检测。

这种方法的检测过程较为简单，但其极易受到环境变化的影响，这导致了该方法对交通标志的检测结果存在较大的误差。研究者们对其存在的不足进行了大量改进，GAO[11]等通过改进颜色模型来提高对颜色信息的提取能力，随后使用 FOSTS（Foveal system for traffic signs)模型对交通标志的特征进行检测。

Benallal[12]等人在研究过程中发现当改变光照环境时，RGB 空间也会发生变化，所以在交通标志分类之前需要对 RGB 组件进行对比。

（2）基于形状信息检测：基于形状特征的检测有 Hough 变换[13]、相似度检测[14]、边缘检测[15]、距离变换匹配[16]等方法，主要涉及以下步骤：①边缘检测和轮廓提取：通过对图像边缘的检测获取目标的轮廓信息，这些轮廓可以通过形态学操作来进一步处理和增强。②形状特征描述与匹配：针对提取到的轮廓，可以计算其形状特征，将形状特征与预定义的模块进行匹配以识别图像中的目标。③后处理：通过对检测结果进行如非极大值抑制、去除重叠部分等得到最终的形状检测结果。

Garcia[17]等人使用 Hough 变换来定位道路标志信息，并且使用算法来改进特征信息，有效地获取道路标志信息并最终使用检测网络来检测道路标志类别的

特征。Loy 等[18]通过考虑交通标志的对称性对交通信号进行数学建模，从而能够有效检测旋转、遮蔽等质量不高的交通标志图。

传统方法的交通标志检测算法有效的前提条件是要有大量的先验知识，这样才能对具有鲜明特征的交通标志进行良好的检测。但随着自动驾驶的发展，实际场景中面对的交通标志通常具有多尺度、目标小的特点，同时由于遮挡、天气等影响，传统方法检测交通标志的方法显得力不从心，达不到日益增长的检测需求。

1.2.2 基于深度学习的交通标志检测

近年来，由于传统目标检测算法的诸多局限性使其难以满足自动驾驶在实际应用中对交通标志检测的要求，随着机器学习[19]的发展，基于深度学习[20]的目标检测算法应运而生。基于深度学习的目标检测有两阶段和单阶段两类。

（1）两阶段目标检测算法：指将检测过程分为两阶段，首先对输入的图像进行检测得到目标候选框，然后对得到的目标候选框进行分类，从而完成对目标的检测和识别。

文献[21]基于对比度限制自适应直方图均衡（CLAHE）的图像增强和神经网络作为多类分类器提出了一种两步 TSR 方法。该方法采用三种架构，即 LeNet、VggNet 和 ResNet 进行分类。最后在 GTSDB（German traffic sign recognition benchmark）数据集上进行分类测试，得到了比较好的分类精度，实验结果还表明与其他类似方法相比，提出的 CLAHE 的图像增强和基于 ResNet 的分类器组成的架构有助于获得更好的分类精度。

文献[22]提出了一种改进 Faster-RCNN[23]的交通标志检测算法。首先通过重构 Faster-RCNN 的骨干网络以及对区域候选网络进行改进达到网络框架轻量化的目的。其次将 scSE 注意力机制与 GSCOnv 卷积进行融合，使网络能够在面对目标的多尺度特征有良好的检测能力。然后将锚选框的尺寸进行更新、对每个目标子区域进行双线性插值、采用平衡 L1 损失函数等一系列改进方法对网络进行优化。最终不仅减小了模型权重，更明显提升了检测精度。

文献[24]提出了以一种改进 Libra R-CNN 的交通标志检测算法。其在 Libra R-CNN 网络的锚框提取阶段，采用 GA-RPN 生成锚框，这样可以生成更为精确和更多样化的样本，从而减少交通标志图像样本中由于复杂背景影响和存在小目标不容易定位的问题，达到提升检测精度的目的。

（2）单阶段目标检测算法：指能够直接从图像中检测出目标位置和类别的算法，只需要一次前向传播就可以完成整个检测过程。

文献[25]提出了一种改进 SSD（Single Shot MultiBox Detector）[26]的交通标志算法。首先采用比原始 SSD 算法的主干网络 VGG-16[27]更深层次的 ResNet50[28]进行替换，从而使弱目标特征的表征能力得到增强，同时将 RFB 模块添加到 SSD 的额外添加层中，以此来提升网络对小目标的感受野。其次，使用 Bi-FPN 加权双向金字塔将图像的深度和浅层特征进行有效融合。最后使用 K-meas++算法改变窗口的大小，从而匹配交通标志图像中存在的小目标。实验数据表明改进后有效提高了原网络的 mAP[29]。

文献[30]提出了一种基于 YOLOv4[31]的改进交通标志检测方法。在 YOLOv4 算法结构上去掉了 19×19 的大感受野检测层,加入 152×152 的尺度检测层;在算法中引入注意力机制, 在主干网络提取的 3 个特征层后加入高效通道注意力(Efficient Channel Attention, ECA)模块,使网络关注有用信息,提高算法检测能力;同时为了加快网络的收敛, 利用 K-means 聚类算法重新生成网络的先验框。该方法的平均准确率达到 84.95%，较 YOLOv4 提高了 4.58%。

文献[32]提出了一种基于 YOLOv5[33]的交通标志检测算法用于解决对交通标志检测存在漏检、误检等问题。通过设计一种由 CBAM[34]、SPCConv、C3 相结合的 CSC3 模块，提升网络的特征提取能力，降低参数量。同时，增加小目标检查头和跨层连接，通过重构 Concat 连接提升模型的检测精度。最后引入 EIOU，解决漏检、误检问题。相较于原始网络，mAP 提高了 8.3 个百分点，参数量下降了 10.1%，检测速度达到了 74FPS。

文献[35]提出了一种改进 YOLOv7[36]的轻量化交通标志检测算法。为增大感受野，在骨干网络引入大核模块，其次在检测颈部融合坐标注意力、随机池化等方法，在构建通道注意力的同时又能捕捉准确位置，此外，提出了集中综合深度可分离卷积模块，能够更好的提取图像特征。实验结果表明网络在较低参数量和计算量的情况下实现了较高的检测效率。

1.3 交通标志检测的难点

当前，对于交通标签的检测主要面临着以下几个问题：

（1）复杂多变的环境条件：道路上的光照、天气、道路标志损坏或遮挡等因素都会对交通标志检测造成不利影响。

（2）多样性的交通标志类型：交通标志检测类型各异，有着不同形状、颜色和图案，这对检测算法提出了较高的泛化能力要求。

（3）实时性要求高：在自动驾驶、智能交通管理等场景下，交通标志检测需要具备较高的实时性。

（4）精度要求高：交通标志往往较小且与周围环境相似，容易被误检或漏检。

（5）应用场景复杂：道路上会存在各种干扰物体、车辆、行人等，这会干扰交通标志的检测，增加检测的复杂度。

（6）图像特征分布散：交通标志经常出现在图像的边缘区域，这导致检测网络对交通标志的特征提取不够充分。

针对以上问题，本文在基准目标检测框架上做出改进，通过实验验证与对比来证明了所提出方法的优越性。

1.4 研究内容及主要工作

在智能交通系统中，交通标志检测是其重要的研究内容之一。然而，在实际场景中，交通标志牌存在着以下问题：在图像中所占比例小，图像分辨率差，特征不宜凸显，且易受到环境变化的影响。这些问题都会导致检测精度和速度不足以满足智能驾驶需求。此外，由于在实际工程中需要考虑到模型部署，这对模型的规模也进行了一定的限制。本文通过分析当前网络模型存在的不足，并指出影响检测效果的关键因素，提出基于 SSD-RFC 网络的交通标志检测方法以及基于YOLOv7-CCa网络的小型交通标志目标检测方法。

具体研究内容如下：（1）针对现有的单阶段多锚框的 SSD 算法在交通标志的检测中存在的漏检率高、背景复杂、准确率较低等问题，本文提出一种基于 SSD-RFC 网络的交通标志检测算法。它采用了三个步骤来提高模型的检测精度和速率。首先在 SSD 网络中引入残差网络 ResNet50 模块，增强特征提取能力。

接着通过添加 CBAM 注意力机制来减少特征在网络模型中产生的损失。在此基础上，使用高效轻量的特征融合模块将高低层的特征信息进行融合，进行几何信息和语义信息的互补。相比于原网络，该方法能够有效的提高模型对交通标志检测的精度，实现实时准确的检测目标。（2）针对现有算法在交通标志小目标图像检测上存在检测精度低、模型复杂度高等问题，提出基于 YOLOv7-CCα模型的交通标志检测方法。首先将集中式金字塔 CFP 添加到预测网络，提高特征的层内调节能力，降低模型的参数量。其次添加 CA 注意力机制，学习不同通道之间的相关性来动态调整权重，提高模型对重要特征的关注度。最后优化分类及定位的损失函数，提高边界框的回归精度。该方法有助于智能交通系统在交通标志较远时（即小目标状态）就能够实现精准检测，从而给驾驶员预留更多的反应时间，降低交通事故和违规行为的发生率。

1.5 章节安排

论文共分为五章，各章节内容如下：

第 1 章：绪论。

绪论首先阐述了交通标志检测研究的现实背景和意义，其次从传统方法和深度学习方法两个方向对当前交通标志检测进行了现状介绍，接着总结了交通标志检测中存在的一些难点，最后对本文的研究内容和章节结构进行了安排。

第 2 章：目标检测算法相关理论基础。

该章主要对目标检测网络算法的理论基础知识进行了概述，然后着重介绍了经典的两大类目标检测模型，分析不同的目标检测算法存在的优缺点。其次对目标检测结果的评估标准进行了详细阐述。最后对交通标志数据集进行介绍，说明本文选择数据集的依据。

第 3 章： 基于 SSD 网络的交通标志检测算法研究。

本章首先对 SSD 网络进行介绍，并对其在交通标志检测上存在的不足之处进行分析。针对 SSD 算法对交通标志检测存在的精度低、漏检率高等问题，提出了三个改进方法，一是将 SSD 目标检测网络的主干网络 VGG16 替换为 ResNet50。二是引入新型特征融合模块加强浅层特征的语义信息。三是引入注意力机制使网络关注重要的特征区域。最后通过实验验证改进措施是合理且有效果的。

第 4 章：基于改进 YOLOv7 网络的交通标志检测算法研究。

针对现有算法对交通标志图像中存在的小型目标存在着提取信息不充分、模型参数量大、数据集缺乏细分类等问题，提出一种基于 YOLOv7 网络的交通标志检测算法。其使用集中式特征金字塔减少模型参数量，同时加强对图像边缘小目标的特征提取，并采用 CA(Coordinate Attention)注意力机制将目标的位置信息添加到通道注意，接着引入损失函数提高边界框的回归精度。实验结果表明本章所提算法具有良好的检测精度和检测速度，并且参数量较小，能够满足在移动设备上进行部署的要求。

第 5 章：总结与展望。

对全文的研究工作进行总结，并对于本文的核心内容——网络改进部分进行详细阐述。此外，对交通标志图像目标检测中存在的不足进行了分析，对未来交通标志检测的研究方向进行了展望。

第 2 章 目标检测算法相关理论基础

2.1 目标检测概述

目标检测是计算机视觉领域中一项及其重要的任务，旨在识别图像或视频中出现的特定物体，并标定其位置。随着深度学习的迅速发展，目标检测取得了显著的进步，在自动驾驶[37]、辅助驾驶[38]、安全[39]等诸多领域都得到了广泛的应用。

目标检测的核心目标是实现对图像中各种目标物体的准确识别和定位。这一任务通常设计将图像中的物体进行分类，并在图像中标注出它们的位置，以方便进一步的分析和应用。在目标检测中，包含以下几个关键步骤：①特征提取：利用图像处理和计算机视觉计算，从输入图像中提取特征。②候选区域提取：在图像中提取可能包含目标的候选区域，以减少后续处理的计算量。③目标分类：对每个候选区域进行目标分类，确定该区域内是否包含感兴趣的目标。④位置精确定位：对于被分类为目标的候选区域，进一步通过回归算法来精确确定目标的边界框位置。⑤后处理与模型评估：通过引入一些后处理技术来提高检测结果的准确性和稳定性，最后根据一些评价指标对训练好的模型进行性能评估。

目标检测作为计算机视觉领域的重要研究方向之一，不断引入新的算法和技术，取得了显著的进展。深度学习技术的发展为目标检测提供了强大的支持，使其在各个领域中都得到了广泛的应用。

2.2 卷积神经网络

卷积神经网络(Convolutional Neural Network，CNN）[40]是一种主要用于处理视觉数据的深度学习网络。核心结构主要包含卷积层、池化层、激活层、全连接层。

（1）卷积层

卷积层主要负责对输入数据的特征提取。与最初的图像处理特征提取方式不同，传统方法通常通过固定的计算方式，来获取图像中的特征。相比之下，卷积层能够获取其中更有代表性的信息用来提供给网络决策。

卷积层通过卷积核在图像上滑动并进行内积运算，卷积核即离散的二维滤波器。具体操作如图 2-1 所示，以3×3 的卷积核为例，卷积核中的值代表着权重，若输入一个4×4 的矩阵，这些权重在卷积操作中用来提取特征。卷积操作是通过将卷积核与输入矩阵的每个位置进行逐元素相乘，并将所有乘积结果相加得到下一特征图的对应元素值。这一过程从输入矩阵的左上角开始，按照从左向右、从上到下的顺序进行计算。在这个过程中，卷积核的元素值保持固定，即实现了权值共享。移动步长为 1 的情况下，下一特征图的大小为 2x2。这就是一次卷积的操作过程。

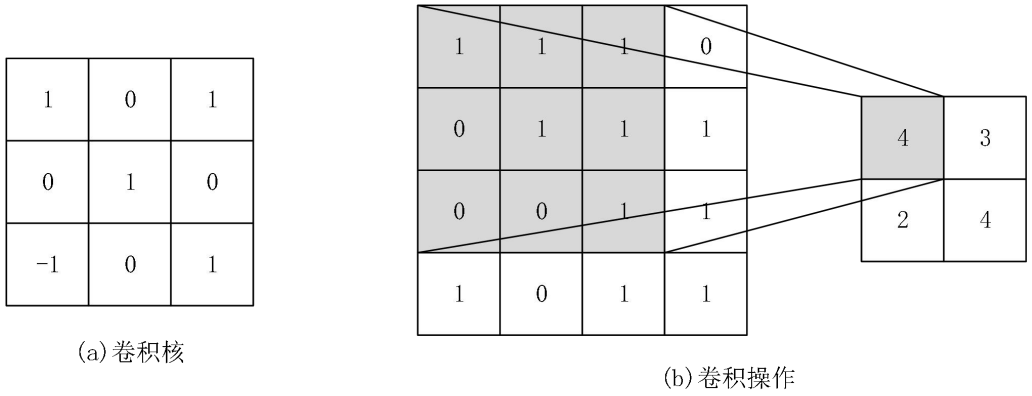


图 2-1 卷积操作示意图

Fig. 2-1 Convolution operation diagram

在多通道卷积中，对于每个通道都会应用对应的卷积核进行卷积操作。这样处理后，输出的特征图尺寸通常会比输入图像的尺寸小，这是由于卷积运算导致了信息的压缩和特征的提取，通道数会增加。这一过程有助于网络更好地理解输入数据的复杂特征。在卷积层中，有三个关键参数。卷积核大小（Kernel Size）是用来捕捉特征的滤波器，其大小决定了模型在输入数据中查看的区域大小，即在输入中看到的区域大小。通过调整卷积核大小，可以获得不同复杂程度的特征信息。二是步长（Stride）：步长（Stride）表示卷积核在特征图上进行滑动时的移动距离。较大的步长会导致输出特征图尺寸减小，而较小的步长则会增加输出特征图的尺寸。三是填充（Padding）：填充是为了在卷积操作前后保持特征图的尺寸不变。

（2）池化层

池化操作是从输入数据的局部区域中选择最大值或平均值。这有助于捕获输

入数据的关键特征，并减少数据的维度，减少计算量。通常，池化层会与卷积层交替使用，以逐渐减小特征图的尺寸，提取更高级别的抽象特征。

池化层主要包括最大池化（Max Pooling）和平均池化（Average Pooling）。最大池化：对于每个池化区域，在输入区域中选择最大的像素值，然后将该值发送到下一层。这种操作有助于保留区域内的主要特征，同时减小图像尺寸。平均池化： 对于每个池化区域，在输入区域中计算像素值的平均值，然后将平均值发送到下一层。平均池化有助于获取区域内的整体信息，同时减小图像尺寸。

池化过程示意图如图 2-2 所示

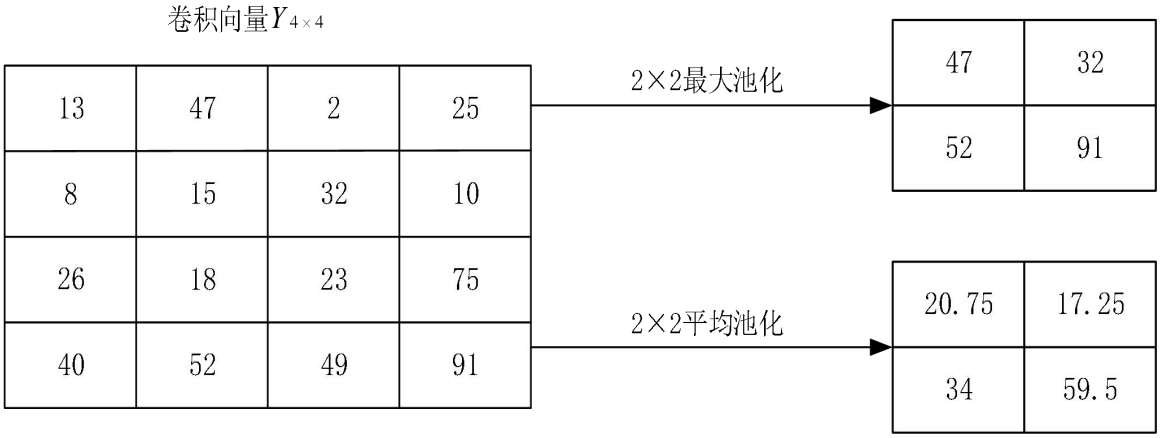


图 2-2 平均和最大池化计算示例

Fig. 2-2 Example diagram of average and maximum pooling calculations

池化层公式如(2-1),(2-2)所示：

$$Y^k = \xi^k down(I^{k-1}) + v^k \tag{2-1}$$

$$I^k = f(Y^k) \tag{2-2}$$

ξ^k 、v^k 分别代表权重矩阵以及偏置矩阵， $down()$ 为池化函数

（3）全连接层

全连接层用于将前一层的特征图展开成一个向量，然后通过权重矩阵的线性变换和激活函数来进行特征提取和非线性变换。

全连接层主要有以下作用：①特征提取：通过学习权重矩阵，可以将输入特征映射到更高维的特征空间，从而提取更加抽象和复杂的特征。②非线性变换：全连接层中通常会使用激活函数对线性变换的结果进行非线性映射，增加神经网络的表达能力。③输出预测：最后的全连接层用于将特征映射到输出类别，并产

生模型的最终预测结果。计算公式如(2-3)、(2-4)：

$$Y^k = w^k I^{k-1} + b^k \tag{2-3}$$

$$I^k = f(Y^k) \tag{2-4}$$

其中，w^k、b^k 分别代表全连接层的权重和偏置。

（4）激活层

激活层的功能是对提取到的特征进行非线性处理。在神经网络中若不进行非线性操作，每个神经元的输出将是线性的，这样的网络可能会收敛缓慢且缺乏足够的表达能力。为了加速模型的收敛速度并提高表现力，引入非线性激活函数是一种常见的做法。以下是一些广泛使用的非线性激活函数，其函数图像如图 2-3 所示：

1、 Sigmoid 函数，Sigmoid 函数将输入映射到 0 和 1 之间，常用于输出层进行二分类问题的概率估计。Sigmoid 函数公式如(2-5)：

$$\sigma(x) = y = \frac{1}{1+e^{-x}} \tag{2-5}$$

2、 Tanh 函数

Tanh 函数将输入映射到-1 和 1 之间，相比于 Sigmoid，它的输出范围更广，有时用于隐藏层。公式如(2-6)：

$$\tanh(x) = \frac{(e^{x} - e^{-x})}{(e^{x} + e^{-x})} = 2 \times sigmoid(2x) - 1 \tag{2-6}$$

3、 ReLU 函数

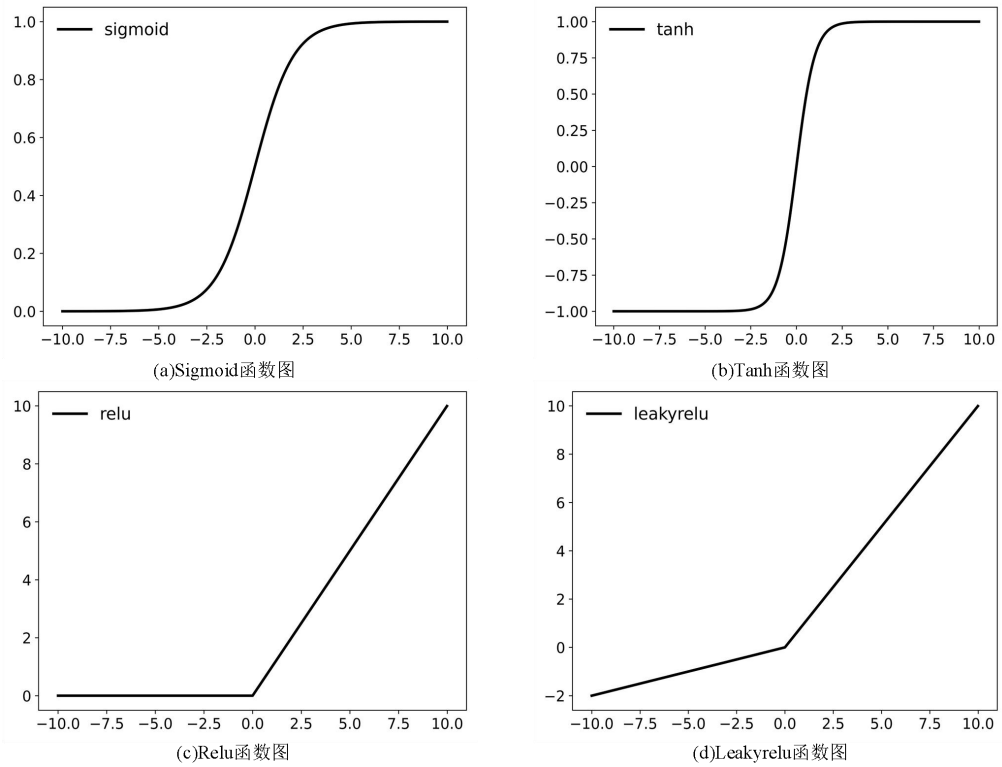
由于 ReLU 函数只有线性关系，ReLU 函数在前向传播和反向传播都有较快的计算速度。当输入值为正时，不存在梯度消失现象，当输入值为负时，梯度为 0，会产生梯度消失现象。公式如(2-7)：

$$\mathrm{Re}\,LU(x) = \max(0, x) \tag{2-7}$$

4、 LeakyReLU 函数

LeakyReLU 允许小于 0 的输入有一个小的梯度，避免了 ReLU 可能产生的梯度消失现象。使得网络能够更好的处理负数输入。函数公式如(2-8)：

$$Leaky\,\mathrm{Re}\,LU(x) = f(x) - \max(\alpha x, x) \tag{2-8}$$



(a)Sigmoid函数图　(b)Tanh函数图

(c)Relu函数图　(d)Leakyrelu函数图

图 2-3 非线性激活函数

Fig. 2-3 Nonlinear activation function

2.2.1 卷积神经网络训练

卷积神经网络的训练过程包含前向传播和反向传播

前向传播（Forward Propagation）：输入数据通过网络的各个层，经过卷积层、池化层等，得到网络的输出。这是数据从网络的输入层一直传播到输出层的过程，计算中间层的激活值，并产生最终的预测结果。

反向传播（Backward Propagation）：通过计算实际输出与期望输出之间的误差，使网络更新网络参数。误差传播的过程中，使用梯度下降等优化算法来更新每一层的权重和偏差，使得网络的输出逐渐逼近期望输出。这是通过对损失函数进行微分来计算梯度，并利用梯度信息对网络参数进行调整的过程。

这两个过程是交替进行的，通过不断调整网络的权重和偏置，使得网络的输出逐渐接近期望的输出。训练过程旨在最小化损失函数，使得网络能够更好地泛化到未见过的数据。这种迭代的过程使得卷积神经网络能够逐渐调整参数，提高对输入数据的表示能力，从而提高模型的性能。训练过程的成功依赖于合适的损失函数、优化算法和适当的数据集。训练过程如图 2-4 所示。

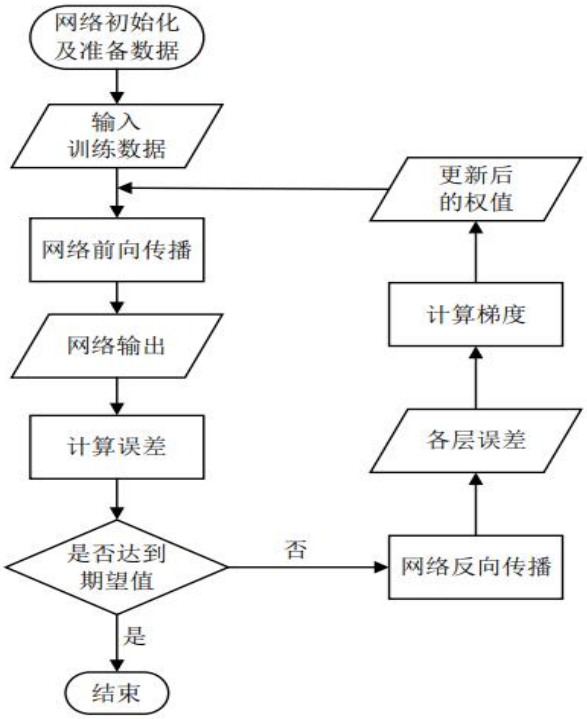


图 2-4 卷积神经网络训练过程

Fig. 2-4 Convolutional neural network training process

2.2.2 常见的优化器

优化器是一种用于调整模型参数以最小化损失函数的算法，可以进行参数更新、加速收敛等。常见的有以下几种：

（1）SGD(Stochastic Gradient Descent)

SGD 每次迭代随机选择一个样本进行梯度计算和参数更新。SGD 的缺点是更新过于频繁，可能会引起震荡。如式(2-9)、(2-10)，

$$\delta_t = \frac{\partial E(\theta_t)}{\partial \theta_t} \tag{2-9}$$

$$\theta_{t+1} = \theta_t - \alpha \cdot \delta_t \tag{2-10}$$

式中，δ_t 为 t 时刻权重参数 θ_t 对损失函数 E 的梯度，θ_{t+1} 为 $t+1$ 时刻的权重参数，α 学习率。

（2）AdaGrad(Adaptive Gradient)

AdaGrad 在 SGD 的基础上引入了二阶动量的思想，并采用自适应学习率的策略。AdaGrad 的独特之处在于针对每个参数使用不同的学习率，并根据参数历史梯度的平方根来进行调整。这样可以更加有效地更新参数，适应不同参数的变化情况。AdaGrad 的更新规则可以表示为：

$$v_t = \sum_{t=1}^{t} \delta_t^2 \tag{2-11}$$

$$\theta_{t+1} = \theta_t - \frac{\alpha}{\sqrt{v_t}} \cdot \delta_t \tag{2-12}$$

式中，v_t表示 t 时刻的二阶动量。

虽然 AdaGrad 的自适应学习率能够使每个参数在训练中拥有不同的学习率，从而更好地适应数据的特点，但它也存在一些问题。其中之一是由于二阶动量的累积，学习率的分母中的项会随着时间增长而单调递增。这可能导致学习率衰减过快，使得训练提前结束，尤其是对于稀疏数据集。

（3）RMSprop(Root Mean Square Prop)

RMSProp 优化器为了解决 AdaGrad 学习率衰减过激进的问题而提出的改进算法。RMSProp 对二阶动量的累积采用了指数加权移动平均的方式，而不是像 AdaGrad 一样使用过去所有梯度的累计。具体如式(2-13)、(2-14)，

$$G_{t,i} = \beta \cdot G_{t-1,i} + (1-\beta) \cdot (\nabla J(\theta_{t,i}))^2 \tag{2-13}$$

$$\theta_{t+1,i} = \theta_{t,i} - \frac{\alpha}{\sqrt{G_{t,i} + \varepsilon}} \cdot \nabla J(\theta_{t,i}) \tag{2-14}$$

其中：

- $G_{t,i}$是参数$\theta_{t,i}$对应的二阶动量，通过指数加权移动平均计算。
- $\nabla J(\theta_{t,i})$是参数$\theta_{t,i}$对应的梯度。
- α 是学习率。
- β 是用于计算指数加权移动平均的超参数，通常取接近 1 的值。
- ϵ 是为了防止分母为零而加入的小常数。

RMSProp 的改进在于引入了一个衰减系数 β，它控制了二阶动量的累积速度。通过使用指数加权移动平均，RMSProp 能够自适应地调整学习率，更好地适应不同参数的训练情景，并减缓学习率的衰减速度。这有助于提高算法的稳定性和性能。

（4）Adam(Adaptive Moment Estimation)

Adam 优化器是一种结合了随机梯度下降的一阶动量和 RMSProp 的二阶动量思想，并对两者进行了修正的优化算法。它同时考虑了梯度的一阶矩估计和二阶

矩估计，从而能够自适应地调整学习率，适应不同参数的训练情况。这种优化算法在训练过程中表现出较好的性能和收敛速度。Adam 的更新规则如下：

$$m_t = \beta_1 \cdot m_{t-1} + (1-\beta_1) \cdot \nabla J(\theta_t) \tag{2-15}$$

$$v_t = \beta_2 \cdot v_t - 1 + (1-\beta_2) \cdot (\nabla J(\theta_t))^2 \tag{2-16}$$

$$\hat{m}_t = \frac{m_t}{1-\beta_1^t} \tag{2-17}$$

$$\hat{v}_t = \frac{v_t}{1-\beta_2^t} \tag{2-18}$$

$$\theta_{t+1} = \theta_t - \frac{\alpha}{\sqrt{\hat{v}_t + \varepsilon}} + \hat{m}_t \tag{2-19}$$

其中：

- m_t 和 v_t 分别是一阶矩估计和二阶矩估计，用于估计梯度和梯度平方的指数加权移动平均。
- β_1 和 β_2 是用于计算指数加权移动平均的超参数，通常取接近 1 的值。
- α 是学习率。
- $\hat{m}_t$ 和 $\hat{v}_t$ 是修正后的一阶矩估计和二阶矩估计，通过对 m_t 和 v_t 进行偏差校正。
- ε 是为了防止分母为零而加入的小常数。

Adam 优化器的引入修正了一阶和二阶动量的偏差，使得在训练过程中可以更稳定地调整学习率，适应不同的参数更新情景。Adam 优化器以其结合了梯度一阶和二阶矩估计的特性，在实际深度学习训练中表现出色，并被广泛采用，为模型的收敛速度和性能提供了有效的支持。

2.3 目标检测模型

2.3.1 基于区域提议策略的两阶段目标检测算法

两阶段目标检测算法可分为两个阶段，首先是对输入图像进行多次检测，从而生成目标候选框，然后对这些目标候选框进行检测、分类。

Girshick 等人[41]提出的 R-CNN 是一个经典的两阶段目标检测算法，R-CNN 极大推动了深度学习在目标检测领域的发展。（图 2-5）。

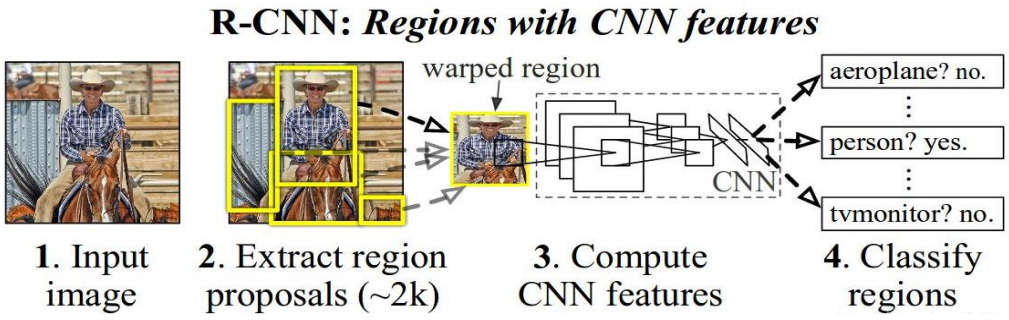


图 2-5 R-CNN 网络流程图

Fig. 2-5 R-CNN network flow chart

R-CNN 的工作流程有：①候选区域生成：使用选择性搜索[42]等算法来生成可能包含目标的候选区域。②特征提取：对生成的的候选区域进行图像特征提取，通常在预训练的卷积神经网络上进行前向传播实现。③分类：将每个候选区域的特征输入到支持向量机(SVM）[43]中，以确认该区域是否存在感兴趣的目标类别。④边界框回归：对候选区域进行边界框回归，使其定位目标在图像中的位置。⑤后处理：对分类和位置预测结果进行后处理，如非极大值抑制,以消除重叠的候选区域和提高检测结果的准确性。

Fast R-CNN[44]是 R-CNN 的改进版本，通过引入一些优化实现了分类和边界框的同步，以此来提高目标检测的速度和准确性。Fast R-CNN 主要在以下几方面进行了优化：①共享卷积特征：Fast R-CNN 在整个图像上只进行一次卷积操作，然后通过候选区域池化层来提取候选区域的特征，避免了 R-CNN 中重复计算的问题。②端到端训练：Fast R-CNN 从特征提取到分类和回归的整个过程一块进行反向传播，不需要像 R-CNN 那样分别训练分类器和边界框回归器。③引入了候选区域池化层：Fast R-CNN 可以在单个前向传播中提取所有候选区域的特征，相比于 R-CNN 需要对每个候选区域进行独立的前向传播速度更快。

Ren 等又对 Fast R-CNN 进行优化提出了 Faster R-CNN，它引入了区域提议网络（Region Proposal Network，RPN）[45]，它负责生成候选区域。从而不在依赖于外部的区域提取方法。其工作流程如下：

①特征提取：输入图像通过卷积神经网络进行前向传播，得到特征图；

②RPN：RPN 通过特征图上的滑动窗口生成锚点，并通过学习预测这些锚点的目标性得分以及边界框的调整值；

③RoI 池化：RPN 生成的候选区域通过 RoI 池化层映射到固定大小的特征上，用于后续的目标分类与定位；

④目标分类与定位：RoI 池化得到的特征用于目标分类和定位，与 RPN 一同构成了整个端到端的目标检测系统。目标分类使用 softmax 分类器，同时引入边界框回归层，用于微调目标边界框。

Faster R-CNN 算法的结构如图 2-6 所示：

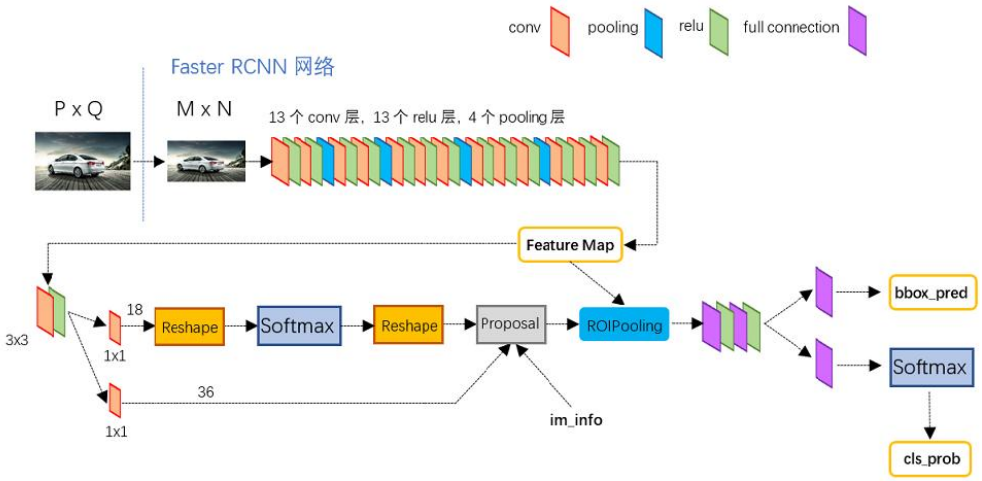


图 2-6 Faster R-CNN 网络结构

Fig. 2-6 Faster R-CNN network structure

其中的 RPN 结构通过在输入图像的特征图上滑动窗口，为每个窗口位置生成多个固定宽高比的锚点（anchors）。然后，通过学习来预测这些锚点是目标或非目标，进而生成候选区域。这一过程的并行性使得 RPN 非常高效。RPN 产生的候选区域被用于后续的目标分类与定位，同时，RPN 和 Fast R-CNN 共享相同的特征图，使得整个系统更加紧凑和高效。

2.3.2 基于回归框架的单阶段目标检测算法

单阶段目标检测算法能够直接从图像中检测出目标位置和类别，只需要一次前向传播就可以完成整个检测过程。主要的单节段目标检测算法包括 YOLO（You Only Look Once）系列和 SSD。

YOLO 系列的第一版由 Redmon 等人提出，即 YOLOv1[46]。随后加强版 YOLOv2[47]也被推出。紧接着，在 YOLOv2 的基础上，Redmon 等人进一步提出了速度更快、检测精度更高的 YOLOv3[48]网络。YOLOv3 整合了当时视觉领域的一些最先进的改进方法和技巧，使得该网络在目标检测任务中取得了显著的进展。

不久 Alexey 等在 YOLOv3 的基础上提出了 YOLOv4 网络模型。YOLOv4 结合了大量已有的研究成果。主干网络采用 CSPNet（Cross-Stage Partial Networks）[49]结构，将输入特征图划分为两个路径，以增加特征的传递效率；引入 SA

M（Spatial Attention Module）和 PPM（Path Aggregation Module）模块，更有效的捕捉空间特征；采用 CIoU（Complete Intersection over Union）[50]损失函数、Mish 激活函数、多尺度检测等优化模型的训练。YOLOv4 使用 NMS（非极大值抑制）进行后处理，可能导致在密集目标区域产生冗余的边界框，从而影响检测精度。其网络流程图 2-7 所示。

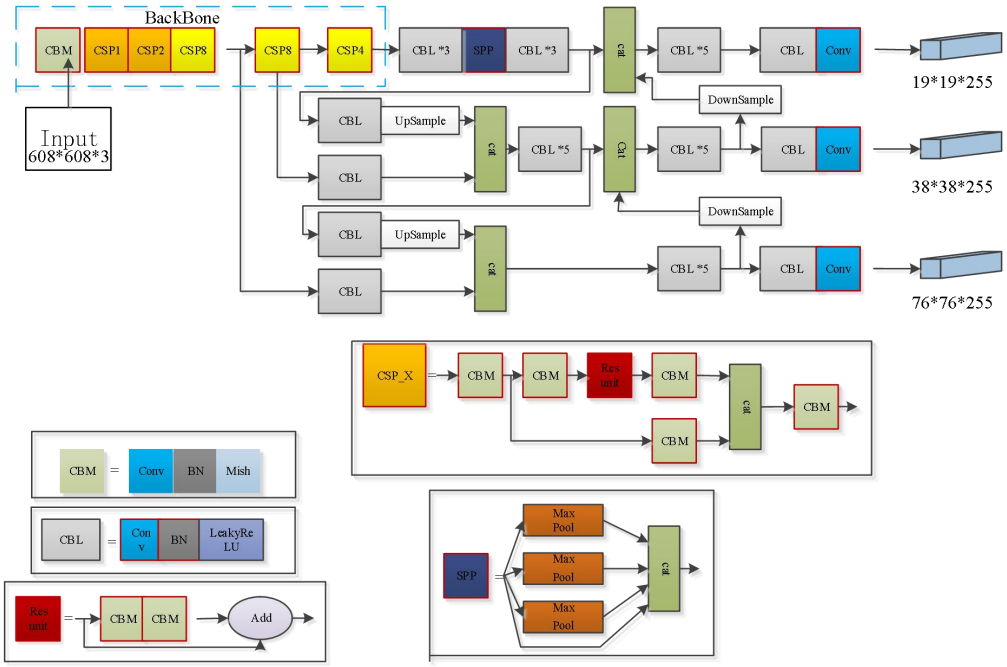


图 2-7 YOLOv4 网络结构

Fig. 2-7　YOLOv4 network structure

不久后，Zhang 等在 YOLOv4 的基础提出了 YOLOv5，相较于 YOLOv4，YOLOv5 使用了 Mosaic 数据增强、Adaptive Anchor Box、自适应图片缩放、Focus 结构，跨阶段局部网络结构、空间金字塔池化结构以及 FPN+PAN 结构等。这些改进使得 YOLOv5 在速度和精度方面都取得了提升。YOLOv5 系列包含四个模型，分别是 YOLOv5s、YOLOv5m、YOLOv5l 和 YOLOv5x。这四种模型的基本框架相同，但其规模和复杂度逐渐增加。YOLOv5s 的网络结构如 2-8 所示。

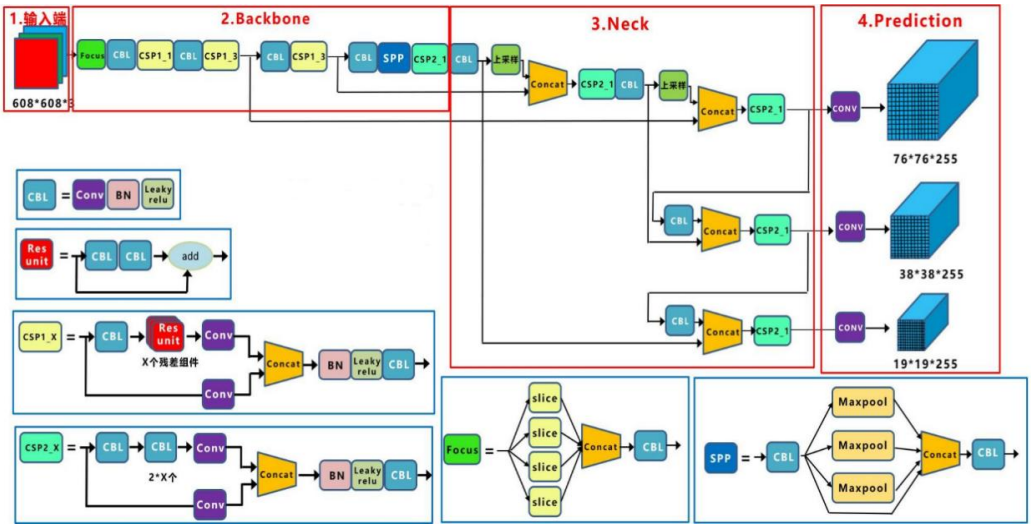


图 2-8　YOLOv5s 网络结构

Fig. 2-8　YOLOv5s network structure

YOLO 系统的快速发展，不断涌现了 YOLOX[51]、YOLOv6、 YOLOv7 等改进算法，不断刷新 YOLO 系列的检测性能。

SSD 是一种基于深度学习的目标检测算法，由 Google 于 2016 年提出。SSD 通过在不同尺度的特征图上预测边界框和类别，实现了多尺度目标检测。它在每个特征图上使用不同大小和宽高比的锚框来预测目标，这锚框覆盖了不同位置和比例的目标，使其能够很好的适应各种大小和形状的目标物体。随后，对于每个锚框，SSD 预测其所属的目标类别和边界框的位置偏移量。分类部分使用 sotfmax 函数进行目标类别的分类，位置回归部分则用来准确调整锚框的位置，以使其更准确的框出目标的位置。

SSD 在神经网络中引入多个不相同尺度的特征图，然后通过卷积和池化来获取特征图。这种方式有效地捕捉了不同尺度物体的特征信息，从而实现了对多尺度目标的准确检测。然后，将特征的位置视为候选框，对每个候选框进行分类和回归来完成目标的检测和定位。在损失计算上，SSD 使用多任务损失函数来优化分类和位置回归，损失函数包括分类损失和边界框回归损失，这提高了优化分类和定位的准确性。SSD 在训练阶段使用标注的数据对网络进行训练，并通过反向传播更新网络参数，在推理阶段将输入图像传递给网络，并应用非极大值抑制(NMS)来筛选出最终的检测结果。

2.4 目标检测的评估标准

目标检测算法通常包含以下评价指标：检测精度 f_P、平均精度 f_{AP}、均值平均精度 f_{mAP}。计算公式为：

$$f_P = \frac{T_P}{T_P + F_P} \tag{2-1}$$

$$f_{AP} = \frac{\sum f_P}{Num(Total\_Objects)} \tag{2-2}$$

$$f_{mAP} = \frac{\sum f_{AP}}{Class\_num} \tag{2-3}$$

式(3-6)中，T_P 表示被模型预测为正类的正样本数量；F_p 表示被模型预测为正类的负样本数量；式(2-3)中，$Class\_num$ 为目标类别数。

同时，本文选取模型参数量 Params 评估模型复杂度，选取帧率 FPS（Frames Per Second）评估模型检测速度，其代表一秒时间内处理的帧数，单位是帧每秒（f/s）。模型参数量 Params 计算公式如(2-4)：

$$Params = C_o \times (k_w \times k_h \times C_i + 1) \tag{2-4}$$

其中，C_o、C_i 分别表示输出通道数和输入通道数，K_w 和 K_h 分别表示卷积核的宽和高。

2.5 数据集

本文选取我国的两个主流交通标志数据集 CCTSDB（CSUST Chinese Traffic Sign Detection Benchmark）[52]和 TT100K（Tenhua-Tencent100k）[53]

（1）CCTSDB

CCTSDB 数据集由湖南省重点实验室提供。CCTSDB 数据集涵盖了多种交通场景，共包含 10800 张交通标志目标图像，其中包括约 40000 多个交通标志。CCTSDB 数据具有较好的分布特性，但因样本仅覆盖禁止标志、指示标志、警告标志三大类，因此在实际应用中有一定局限性。此外，其图像中交通标志的占比较大，因此该数据集适合应用在对中大型交通标志的检测场景中，对于小目标的检测效果不够好。如图 2-9 所示。

图 2-9　CCTSDB 交通标志类别

Fig. 2-9　CCTSDB traffic sign category

（2）TT100K 数据集

TT-100K 数据集由腾讯联合清华大学提供，通过 6 台高清宽视角单反摄像机拍摄，展示了中国各个城市的道路状况。由于拍摄场地、光照、气象等因素的影响，图像也会呈现出不同的形态。它共包含 221 个不同的类别（见图 2-10），在目前的道路标志识别领域中是比较完整的。同时，交通标志占图像的比例大部分都很小（约 0.2%），相较于 CCTSDB 更适合用于小型交通标志的研究场景中

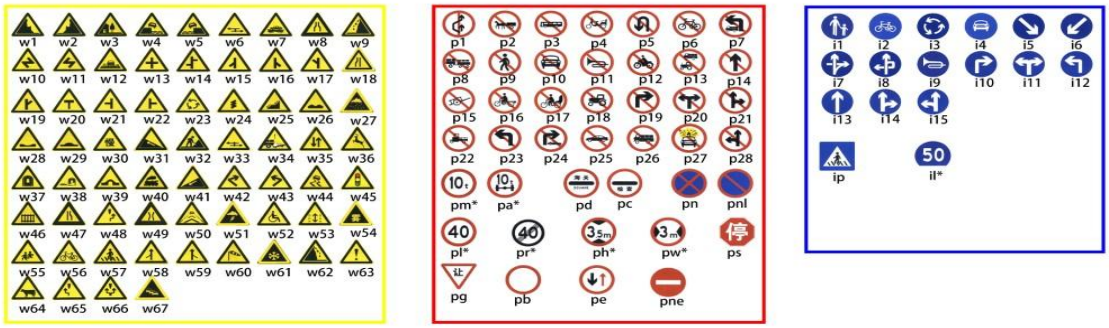


图 2-10 TT-100K 交通标志类别

Fig. 2-10　TT-100K traffic sign category

2.6 小结

本章分析了卷积神经网络的各个网络层次，接着着重阐述了两种目标检测模型：基于区域提议策略的两阶段目标检测模型和基于回归框架的单阶段目标检测模型。最后对目标检测中的评价指标和本文采用的两类数据集进行了详细介绍。

第 3 章 基于残差网络的 SSD 目标检测改进算法

本章针对 SSD 在交通标志检测中存在的识别精度不高，高、低层特征图信息缺少互补等问题进行了改进,主要工作有以下三个方面：首先，引入残差网络 ResNet50 取代 SSD 原始网络中的 VGG16。其次，构建特征金字塔。最后，添加注意力机制。通过仿真模拟实验，该方法相较于 SSD 算法对交通标志的检测精度有了明显的提升。

3.1 SSD 目标检测算法

SSD 网络使用 VGG16 作为基础的卷积神经网络来提取图像特征，如图 3-1 所示。SSD 中只用到了 VGG16 的前五层卷积，在特征提取后面添加多个卷积层来取代后面的全连接层。输入的图像在经过 SSD 网络后会被处理成一系列特征图，每个特征图上都会产生先验框，先验框包含的信息作为目标检测和定位的重要依据。

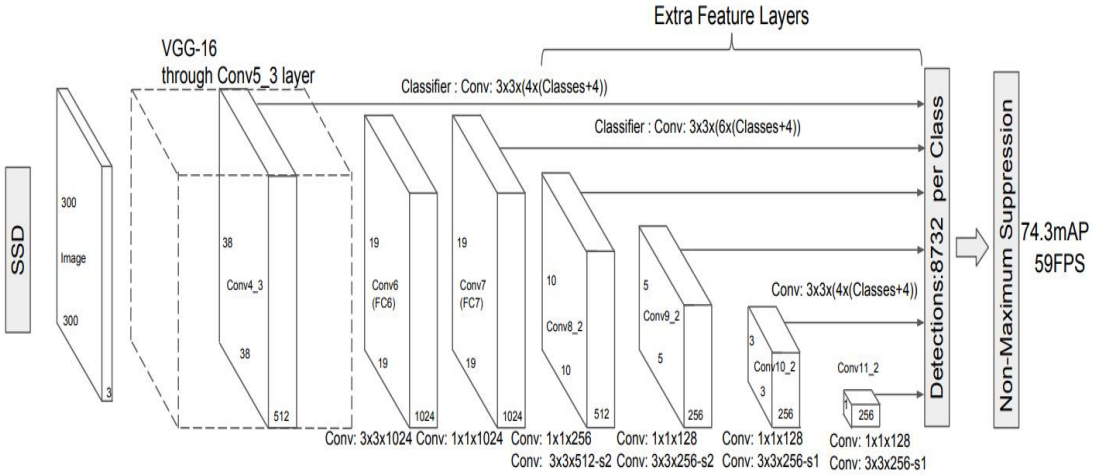


图 3-1 SSD 结构图

Fig. 3-1 SSD Architecture

SSD 采用了多尺度映射，为每个特征图设计了不同数量、尺度和宽高比的先验框。先验框的生成方法如(3-1)式：

$$s_k = s_{\min} + \frac{(s_{\max} - s_{\min})}{m-1}(k-1), k \in [1, m] \tag{3-1}$$

其中，s_k 表示第 k 个特征图中先验框的尺度，m 是使用特征图的数量，而 $s_{\min}$

和 $s_{\max}$ 表示设置比例的最大值和最小值。

先验框的宽高比例通常取 $\{1,2,3,1/2,1/3\}$。确定尺度和宽高比后，先验框的大小由式(3-2)、(3-3)得。

$$w_k^a = s_k\sqrt{a_r} \tag{3-2}$$

$$h_k^a = s_k / \sqrt{a_r} \tag{3-3}$$

其中，w_k^a 和 h_k^a 分别为先验框的宽和高。对宽高比例为 1 的先验框，会增一个尺度 $s_k^{'}$，其计算方法如式(3-4):

$$s_k^{'} = \sqrt{s_k s_{k+1}} \tag{3-4}$$

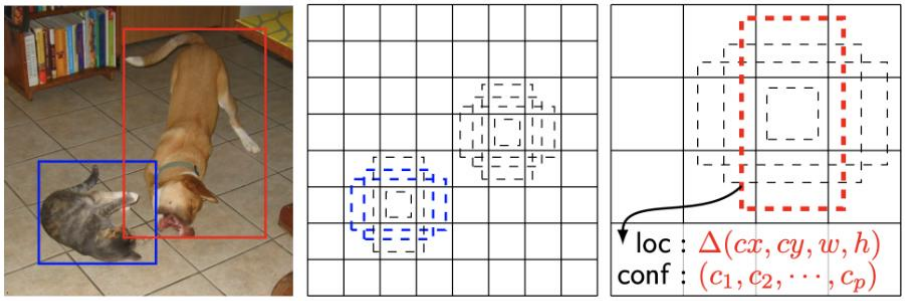


图 3-2 先验框生成过程

Fig. 3-2 The prior frame generation process

总体来看，SSD 通过结合多尺度特征提取、多尺度先验框和单次前像传播等方法减少了神经网络计算量，提高了算法检测性能和速度。

3.2 模型改进

3.2.1 残差网络 ResNet

SSD 网络采用 VGG-16 作为检测网络模型。VGG 采用了连续的 3×3 卷积核进行叠加，同时通过最大池化操作来减小卷积核的大小。网络结构如图 3-3 所示。

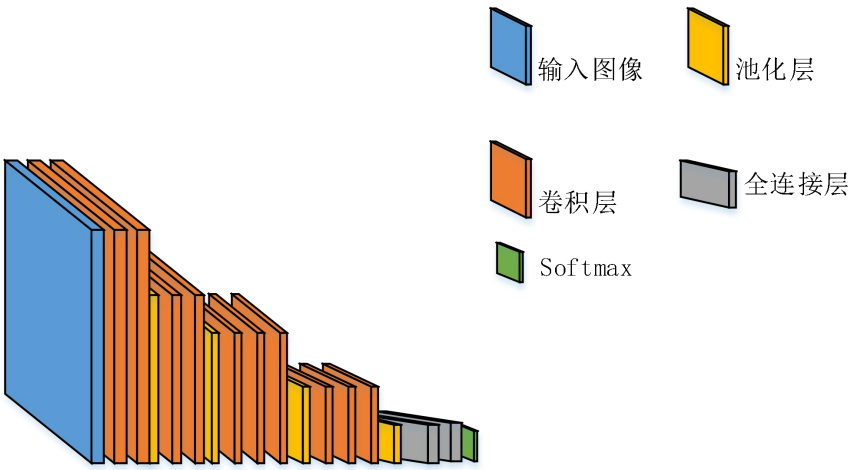


图 3-3 VGG16 网络结构图

Fig. 3-3　VGG16 network architecture diagram

VGG-16 采用固定的输入尺寸输入图片，它对目标的多尺度检测是使用大小不同的卷积核和最大池化来实现。VGG-16 能够通过加深其网络深度来提高性能，但加深 VGG-16 的网络深度肯能会发生梯度爆炸，进而降低网络性能。

为了解决梯度爆炸的问题，本文采用 ResNet50 网络替换 VGG-16。ResNet50 模型是由 50 个卷积层组成，其中包括了多个残差块。这些残差块通过跨层连接的方式使得信息可以更加轻松地在网络中传递，从而避免了梯度消失的问题。相比于原 SSD 使用的 VGG16 模型，ResNet50 可以训练更深的网络，提高网络的精度和性能。残差网络如 3-4 所示。

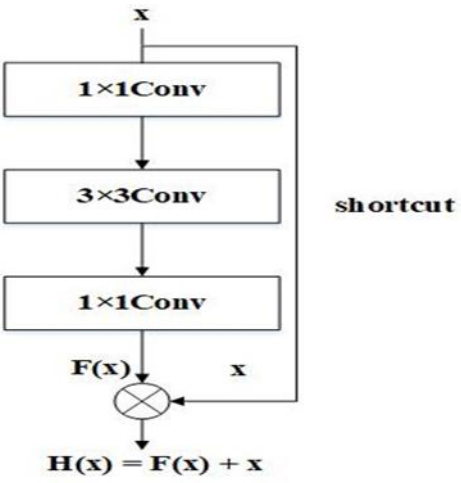


图 3-4 残差网络

Fig. 3-4　Residual Network

$$F(x)=H(x)-x \tag{3-5}$$

在极端情形下，即残差结构的映射结果 $F(x)=0$，残差结构仅执行恒等映射，

网络此时仍然保留了之前迭代的学习能力。因此，即使网络模型的深度增加，网络的学习性能也不会降低。同时由于残差网络的残差分支可以被用来学习新的特征，所以几乎不会出现极端情况。此外，残差结果的卷积层会进行批量归一化操作来使网络的收敛速度提升。

综上所述，残差结构可以用在因为网络模型深度过深而产生的梯度爆炸的情况中，用来解决网络可能产生的退化问题。

3.2.2 特征金字塔

在交通标志检测中，大多数图像中同时包含着近距离和远距离不同尺度的交通标志目标。这对网络模型的检测性能提出了新的挑战。图像输入到网络模型后会生成低层和高层两种不同的层次信息，图像的低层特征包含了大量的位置细节信息，但缺乏足够的特征语义信息，而高层特征的语义信息较为丰富，但目标位置信息比较模糊。因此，怎么将这两层的特征融合得到既有清晰的位置细节信息也有丰富的语义信息的特征成为了解决问题的关键。

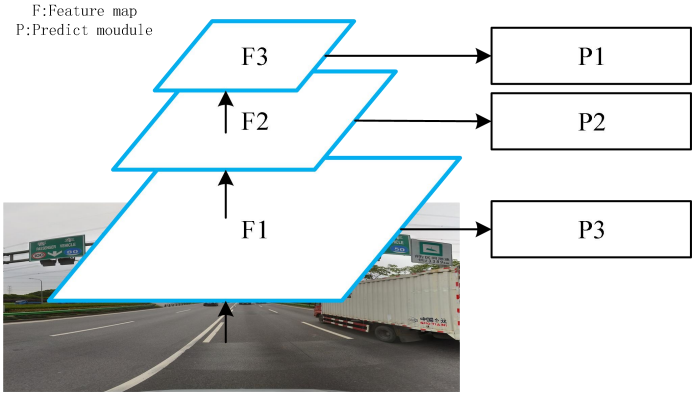


图 3-5 原 SSD 特征融合方式

Fig. 3-5 Feature fusion mode of the original SSD

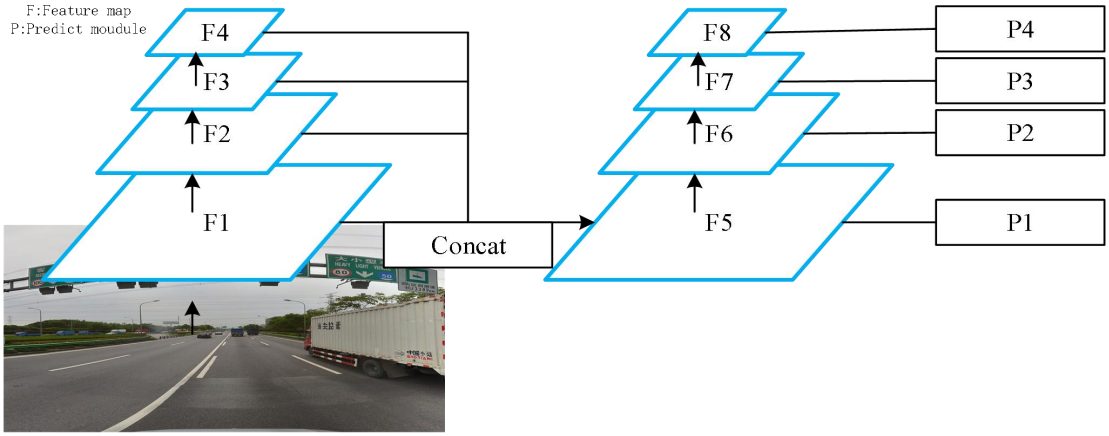


图 3-6 改进特征融合方式

Fig. 3-6 Improved feature fusion mode

而特征金字塔的引入为解决这一问题提供了一种可靠途径，特征金字塔能够将高层特征的丰富语义和低层特征高分辨率信息进行融合，通过融合高低层的特征有效解决不同尺度的问题，提高目标检测的性能。

图 3-5 展示了原始的 SSD 预测层结构，通过在多个尺度的特征层上直接进行预测。然而，各个层之间完全没有联系，因此浅层的高分辨率特征由于语义信息不足导致不能有效检测目标。

本文采用如图 3-6 结构的特征融合模块，该特征融合模块将把网络的 F2、F3、F4 三个输出特征进行特征融合，再通过 F5、F6、F7、F8 四个特征提取层，实现在不同尺度上的预测。相较于原特征融合方法，仅在特征提取网络最后进行 Concat 连接，不仅降低了模型的参数量和计算量，并防止了特征冗余造成的模型性能下降。

3.2.3 选择性注意力模块

深度学习中，注意力机制可有效提升网络特征提取的性能。Hu[54]等人在 2018 年设计了通道注意力的有效机制，提出了 SE-NET（Squeeze-and-Excitation Net-works）,建立起了特征图中的空间相关性。ECA-Net（Efficient Channel Attention Networks）将 SE-Net 中的 MLP 模块替换成一维卷积的形式，有效地减少了参数计算量。之后，Woo 等人提出了 CBAM 机制，其在 SE-Net 或 ECA-Net 的通道注意力机制中增加了空间注意力机制。CBAM 的通道注意力机制用于对每个通道进行加权，以强调对于目标分类任务有用的特征。空间注意力机制用于对空间位置进行加权，以保留有用的特征区域，而忽略不重要的区域。相比 SE-Net 和 ECA-Net，CBAM 机制将两者结合帮助目标检测网络再复杂环境中更高效地检测目标，如图 3-7 所示。

在此章中，将 CBAM 注意力机制加入 ResNet50 的最后一个卷积层的输出上，用来捕获全局上下文信息，并且能够进一步增强模型对目标的特征提取的能力。

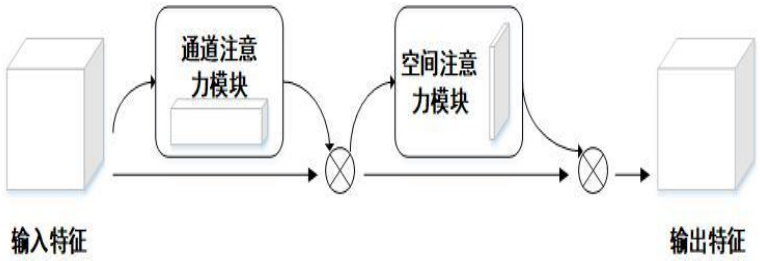


图 3-7 CBAM 注意力机制

Fig. 3-7 CBAM Attention Mechanism

3.2.4 改进后的检测模型

本文提出的改进 SSD 模型如图 3-8 所示。输入图片尺寸为 300×300。为使 ResNet50 能够更好的适应网络模型，本文对 ResNet50 网络进行了局部的修改，去掉 Conv5、平均池化层和 Softmax 层，不但能够大大降低模型的参数量，并且对模型性能也不会产生大的影响。

通过在 ResNet 的后三个模块 Conv2_x 、 Conv3_x 、 Conv4_x 的输出特征图后连接一个高效的特征金字塔。其网络流程如下：①将 Conv2_x 输出的 256×75×75 的特征图通过一个1×1 卷积降维到128×75×75 。②将 Conv3_x 模块输出的512×38×38 的特征图先通过一个1×1 的卷积降维到128×38×38 ，再通过一个两倍的双线性插值（Bilinear Interpolation）上采样将特征图宽高扩大两倍，得到128×75×75 的特征图。③将 Conv4_x 输出的1024×19×19 的特征图在经过 CBMA 注意力机制用来捕获全局上下文信息后，通过一个1×1 的卷积降维到 128×19×19 大小，再通过一个四倍的双线性插值上采样将特征图宽高扩大四倍，得到128×75×75 大小的特征图。④最后把这三个尺度相同的特征图进行特征融合，得到384×75×75 的特征图。该特征图具有三个尺度的语义信息，大大增强了低层特征的语义。

在得到 Conv2_x 、 Conv3_x 、 Conv4_x 三个尺度的语义信息的特征图之后，使用1×1 的卷积和 BN 层对该特征图进行降维，从而得到128×75×75 尺度的特征图。改进后的网络模型最终生成了128×75×75 、 256×38×38 、 512×19×19 、 256×10×10 、 512×5×5 、 256×3×3 、 512×1×1 七个不同尺度的预测特征图，送入到检测器中进行检测。

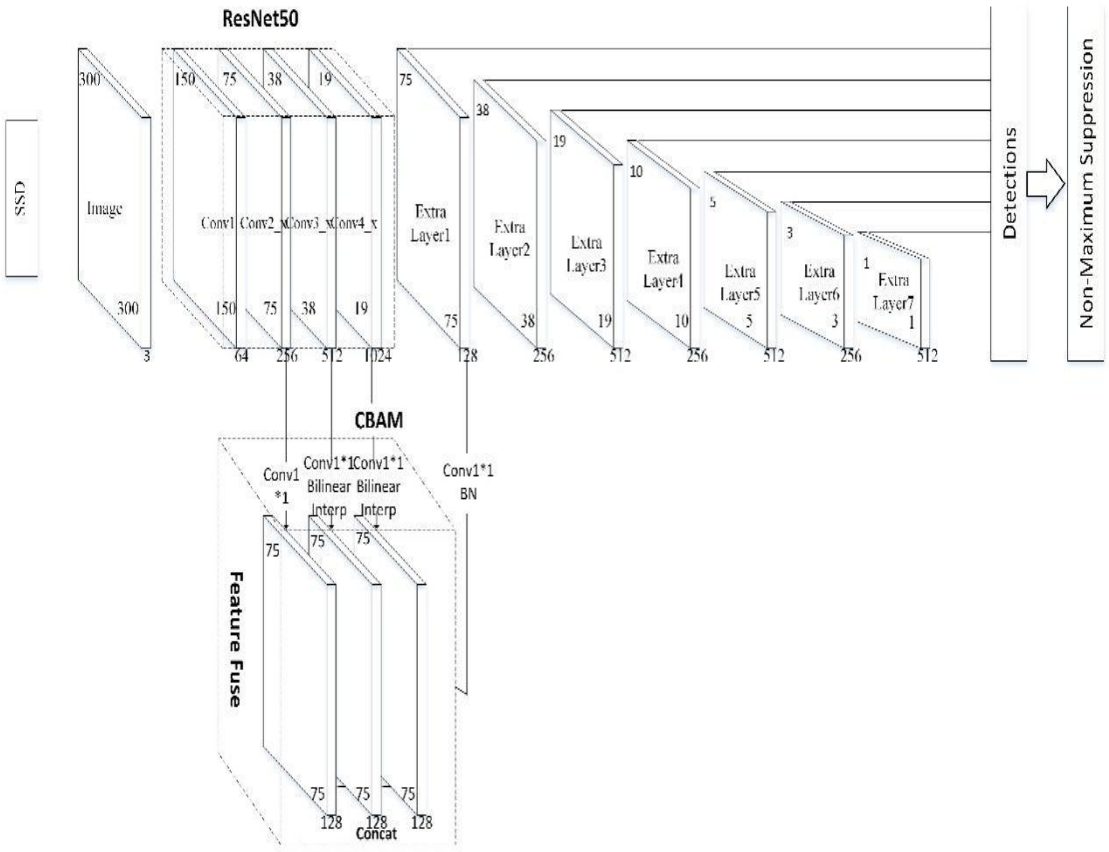


图 3-8 改进后的网络结构图

Fig. 3-8 Improved Network Architecture

3.3 实验与结果分析

3.3.1 环境及参数设置

实验采用了带有 Ubuntu18.04 的 Linux 计算机。Intel(R) Core(TM) i9-10940 xCPU @3.30GHz，64GRAM；NVIDIA GeForce RTX3080 GPU。软件环境：Python3.7，PyTorch1.9.0，CUDA11.0。

详细参数如下：

（1）lr=0.01。

（3）batch size=8。

（4）Number of epochs=300。

（5）Optinmizer=SGD、Weight decay=5e-4、Gamma=0.1。

3.3.2 数据扩充

本章采用 CCTSDB 数据集。CCTSDB 是中国交通标志数据基准，该数据包含了 15734 张真实场景图像。分为三个类别：mandatory、prohibitory、warning。

为降低相似图像对实验的干扰，删去 CCTSDB 中相似图像 1904 张，并在其中选取合适的 2164 张图像进行数据增强操作，对符合条件的图片进行 GridMask、Solarize_add、Posterize 扩充，使其具有更多的真实场景中的可能情况。训练集与测试集按照 9：1 划分，对本文改进的算法进行评估。

3.3.3 实验结果分析与对比

本小节将本文改进的 SSD 检测算法与其他先进的方法进行比较，其中包括基于 Anchor-free 的 CenterNet，YOLO 系列的 YOLOv3、YOLOv4，Faster R-CNN 以及原 SSD。实验结果如表 3-1 所示。

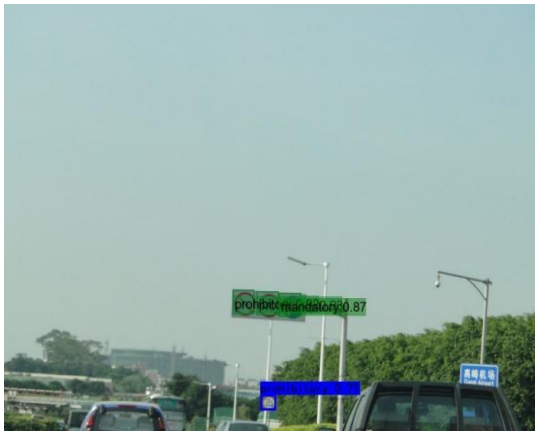
可以看出，改进的 SSD 算法在各个评价指标上表现出良好的结果。CenterNet 的检测精度为 87.9%，而 FPS 仅为 22，检测时间较久；YOLOv3 和 YOLOv4 在准确率方面达到了 87.1%和 87.7%，检测速度也较快但参数量增加较多；Faster RCNN 的检测精度为 86.7%，参数量却有 137.09M，这不利于模型的部署；原 SSD 虽然参数量小，检测速度快，但检测精度仅为 85.6%。文献[55]在提高检测精度后，检测速度大幅下降，FPS 仅为 29。改进的 SSD 算法与改进的 YOLOx[55]相对比，在参数量小幅增加的情况下，mAP 值达到 90.5%，且 FPS 达到了 35。表明模型有较高的检测速度和检测精度，满足移动端实时检测。

各算法的检测效果如图 3-9、图 3-10 所示。从图 3-9 可以看出，当交通标志在无遮挡情况下时原 SSD 未能识别全部目标，CenterNet 与 YOLOv3 的识别结果的置信度分别为 0.56 和 0.71，低于改进后 SSD 的 0.88。而 YOLOv4 和 Faster R-CNN 在无遮挡情况下与改进后的 SSD 检测效果近似。从图 3-10 可以看出，当交通标志存在遮挡情况时除改进 SSD 算法外均存在错检、漏检。

表 3-1 不同模型性能结果对比
Table 3-1 Performance Comparison of Different Models

Model	f_{mAP}	f_{AP}			FPS/帧	Params/M
		Mandatory	Prohibitory	Warning		
CenterNet	0.879	0.859	0.884	0.895	22	32.66
YOLOv3	0.871	0.876	0.856	0.881	67	61.94
YOLOv4	0.877	**0.915**	0.853	0.864	47	64.36
Faster RCNN	0.867	0.849	0.863	0.872	45	137.09
SSD	0.856	0.864	0.838	0.865	**86**	**26.28**
改进 YOLOX[54]	0.886	-	-	-	29	27.98
OURS	**0.905**	0.905	**0.902**	**0.907**	35	44.12

(a) CenterNet



(b) YOLOv3

(c) YOLOv4

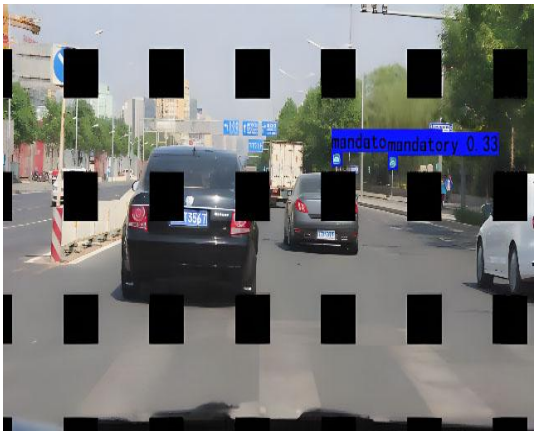
(d) Faster R-CNN

(e) SSD

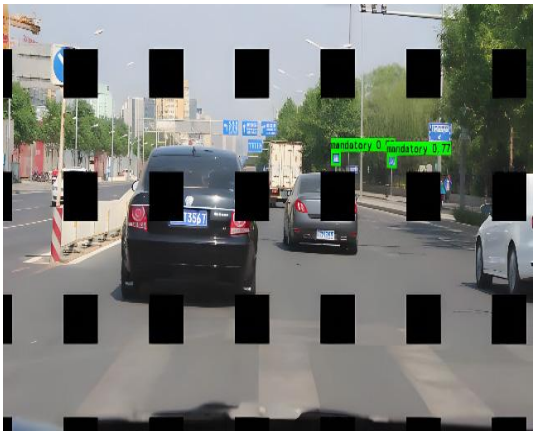
(f) OURS

图 3-9 不同模型的无遮挡检测效果

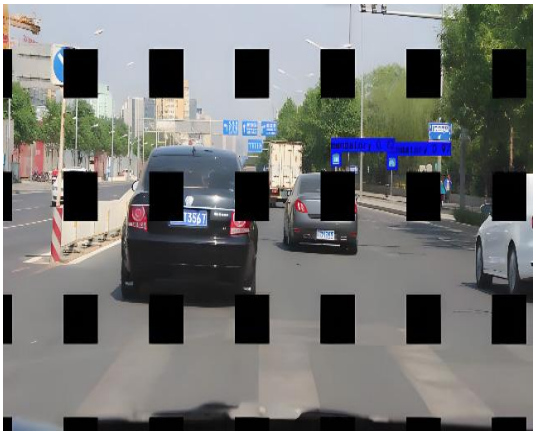
Fig.3-9 Unobstructed Detection Performance of Different Models



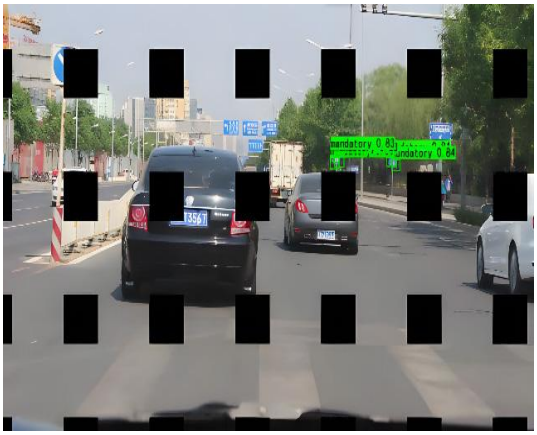
(a) CenterNet



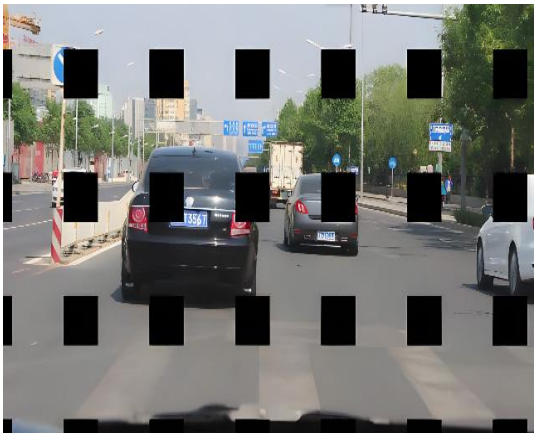
(b) YOLOv3



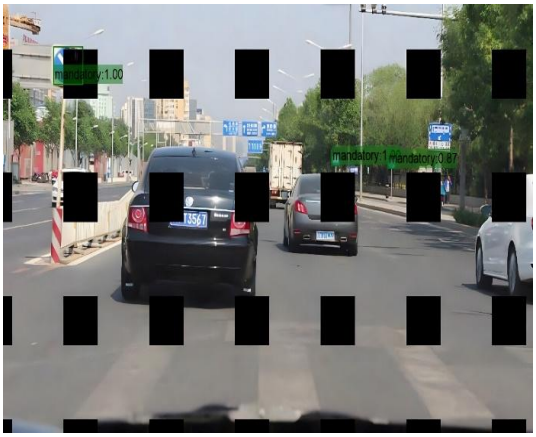
(c) YOLOv4



(d) Faster R-CNN



(e) SSD



(f) OURS

图 3-10 不同模型的有遮挡检测效果

Fig. 3-10 Obstructed Detection Performance of Different Models

3.3.4 消融实验

为了验证所提出的改进措施是否有效，设计了一组消融实验，通过组合所提及修改网络、设计特征融合方式、添加 CBAM，来验证对网络的影响。具体实验性能见表 3-2。

可以看出，通过改进主干网络，参数量小幅增加，mAP 提高了 1.3%，说明 ResNet50 能够更好的提取目标特征。在 ResNet50 网络的基础上添加 CBAM 模块使 mAP 提高了 0.57%，这表明 CBAM 模块使特征提取网络更加关注于重点区域，进一步提高了网络的特征提取能力。此外，通过加入轻量融合模块到 ResNet50 网络中，该模块整合了高层和底层特征，有效改善了对小目标检测的性能，mAP 显著提升至 89.38%。

对比 SSD 和模型 ResNet50 可知，ResNet50(R)、CBAM(C)、Fusion Feature(F)对于模型均增加了参数量，但性能都得到了不同程度的提升。RFC-SSD 在应用了全部改进方法后实现了最佳的检测性能，在 mAP 上取得了最优的结果，与基线网络相比，提升了 4.93%，并且满足了检测的实时性。根据消融实验结果可知，本文的改进方案对网络性能的提升是有效的。

表 3-2 消融实验
Table 3-2 Ablation Experiment

Model	f_{mAP}	f_{AP}			FPS	Params
		Mandatory	Prohibitory	Warning		
SSD	0.856	0.864	0.838	0.865	**86**	**26.28**
SSD+R	0.869	0.880	0.856	0.870	72	29.80
SSD+R+C	0.874	0.890	0.858	0.875	68	35.52
SSD+R+F	0.893	0.898	0.890	0.892	63	37.35
SSD+R+F+C	**0.905**	**0.9059**	**0.903**	**0.907**	35	44.12

3.3.5 在小型交通标志图像上的检测

随着智能交通系统的不断发展，在一些场景中对交通标志检测提出了更高的要求，不仅需要对交通标志的具体类别进行有效检测，更需要在交通标志距离驾驶员较远距离时（即小目标状态）就能够对其进行识别，从而给驾驶员预留出更多的反应时间，降低事故发生率。同时为了更容易将检测系统部署在移动设备中，需要对模型的参数量进一步的优化。

而本章所采用的 CCTSDB 数据集中的交通标志所占图像的比例都比较大，同时，CCTSDB 仅对交通标志进行了禁止、指示和警告三类划分。为验证所改进的 SSD-RFC 算法对小型交通标志检测的能力，对包含大量小型交通标志的 TT100K 数据集进行检测，检测结果如表 3-3、表 3-4。

表 3-3 SSD-RFC 在 TT100K 数据集上的分类识别结果
Table 3-3 Classification and Recognition Results of SSD-RFC on TT100K Dataset

Class	P	R	mAP
i2	0.861	0.835	0.913
i4	0.897	0.903	0.952
i5	0.935	0.93	0.97
il100	0.837	0.727	0.767
il60	0.969	0.966	0.992
il80	0.809	0.926	0.919
io	0.792	0.835	0.832
ip	0.861	0.887	0.923
p10	0.795	0.554	0.718
p11	0.819	0.825	0.879
p12	0.835	0.522	0.736
p19	0.654	0.491	0.684
p23	0.943	0.75	0.862
p26	0.859	0.839	0.913
p27	1	0.817	0.872
p3	0.889	0.825	0.914
p5	0.917	0.902	0.963
p6	0.834	0.5	0.614
pg	0.895	1	0.931
ph4	0.881	0.476	0.672
ph4.5	0.742	0.758	0.78
pl100	0.861	0.961	0.967
pl120	0.848	0.922	0.967
pl20	0.63	0.401	0.458
pl30	0.764	0.713	0.824
pl40	0.818	0.697	0.831
pl5	0.83	0.811	0.88

pl50	0.771	0.559	0.762
pl60	0.877	0.631	0.817
pl70	0.887	0.464	0.622
pl80	0.873	0.708	0.856
pm20	0.891	0.724	0.88
pm30	0.665	0.75	0.82
pm55	0.787	0.667	0.799
pn	0.891	0.949	0.944
pne	0.954	0.924	0.96
po	0.782	0.58	0.729
pr40	0.871	0.923	0.98
w13	0.676	0.458	0.536
w55	0.624	0.629	0.732
w57	0.878	0.792	0.877
w59	0.679	0.865	0.829
all	0.83	0.748	0.83

表 3-4 主流检测模型在 TT100K 数据集上的检测性能

Table 3-4: Detection performance of mainstream detection models on the TT100K dataset

Model	f_{mAP}	Params	FPS
SSD	0.668	**26.2**	**84.2**
Faster-RCNN	0.695	137.1	4.3
YOLOv3	0.581	61.9	29.6
YOLOV4	0.742	64.4	41.7
YOLOV7	**0.862**	37.4	76.1
SSD-RFC	0.83	44.1	32

通过表 3-3 可以看出，SSD-RFC 对部分类别的交通标志保持了较好的检测性能，如 i4、i5、il60、pl100 等，但对于一些类别无法进行有效检测，如 p19、ph4、pl20、w13 等。这是由于这些类别的交通标志过小且特征较复杂，并且 SSD-RFC 采用的残差网络中的很多信号是由 shortcut 层进行传递，这会使网络模型的有效深度没有得到太多提升，网络的有效感受野也与先前变化不大。

通过表 3-4 可以看出，SSD-RFC 的 mAP 相较于 Faster-RCNN、YOLOv3、YOLOv4 仍然保持着领先。但在 mAP、Params、FPS 三个指标上均低于 YOLOv7，这体现了 YOLOv7 对小型交通标志目标检测的潜力，下一章将通过对 YOLOv7 的详细分析与改进来实现对小型交通标志目标的有效检测。

3.4 小结

在本章中，对 SSD 的优缺点进行了详细分析，并针对其不足提出了三个方面的改进方法。一是使用残差网络 ResNet50 来替换 VGG-16，增加特征提取网络的深度。二是通过对常见的特征金字塔进行分析并选择合适的特征金字塔进行引入。三是通过添加注意力机制使网络更加关注目标区域。通过实验数据验证了改进后网络的合理性。最后将所改进算法应用在包含大量小型交通标志的 TT100K 数据集中，通过实验数据分析，发现 SSD-RFC 对小型交通标志的检测效果不尽人意。随后通过与主流模型的检测效果进行对比，发现 YOLOv7 在小型交通标志的检测上存在一定的潜力。下一章将会针对 YOLOv7 在交通标志上的检测展开论证。

第 4 章　基于改进 YOLOv7 的小型交通标志检测

针对上一章中问题，本章通过改进 YOLOv7 实现了有效改进，改进措施有以下几点：首先，采用含有大量交通标志小目标的 TT100k 数据集，并通过对数据集进行数据增强使数据均衡化。其次，将集中式特征金字塔 CFP[56]以引入到预测网络中，降低模型的参数量，同时使特征的层内调节能力得到增强。接着，在主干网络中引入注意力机制 CA[57]。最后，添加α-IoU[58]损失函数。通过仿真模拟实验，提出的新算法不仅能够实现对交通标志小目标的有效检测，还减少的模型参数量使其满足在移动设备中部署的条件。

4.1 YOLOv7 目标检测算法

YOLOv7 的网络结构由 Input、Backbone、和 Head 组成。Input 首先将图像大小改变为$3\times640\times640$，然后将图像输入到 Backbone 中进行特征提取，最后特征图通过 FPN+ PAN[59,60]结构进行多尺度特征融合，将融合后产生的三个尺度的特征图输入到检测头中进行检测，得到 3 个不同尺度的结果，最后在经过 NMS 后得到检测结果。网络结构如图 4-1。

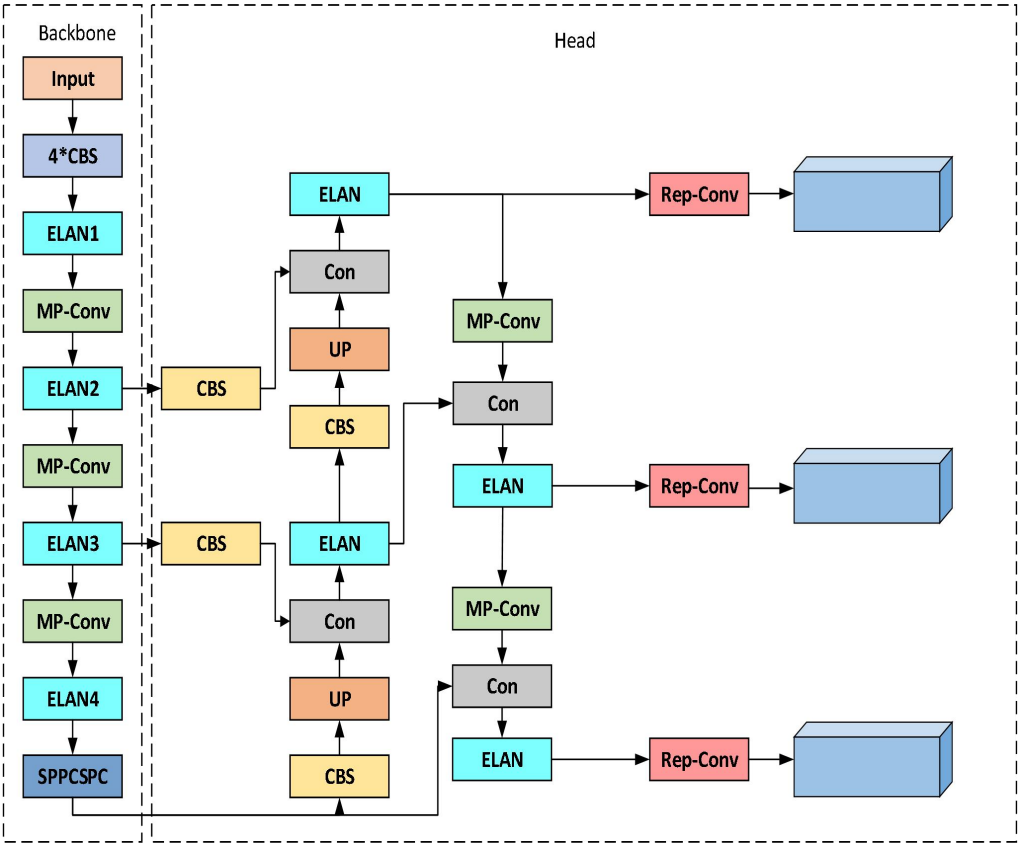


图 4-1 YOLOv7 网络结构图

Fig. 4-1 YOLOv7 Network structure

（1）Backbone

在 Backbone 部分，尺寸大小为3×640×640的图像经 4 个 CBS 模块后得到一个120×160×160尺寸的特征图，接着再通过 ELAN 模块会生成一个160×160×256尺寸的特征图，最后经 3 个 MP-Conv 模块和 ELAN 模块交替操作，生成三个尺寸分别为512×80×80、1024×40×40、1024×20×20的特征图。

如图 4-2 所示，CBS 模块由 Conv 层、BN 层、和 SiLU 函数组成。首先浅蓝色的 CBS 是一个1×1的卷积，stride=1，其次稍深蓝色的 CBS 是一个3×3的卷积，stride=1，深蓝色的 CBS 是一个3×3的卷积，stride=1。

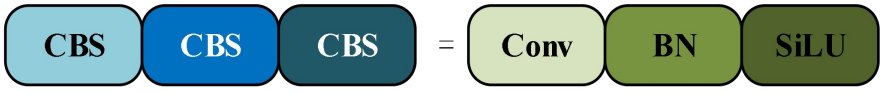


图 4-2 CBS 结构

Fig. 4-2 CBS structure

SiLU 激活函数有利于提高深层次网络的检测效果，其函数如式(4-1)所示，图像如(4-3)所示。

$$SiLU(x) = x \cdot Sigmoid(x) \tag{4-1}$$

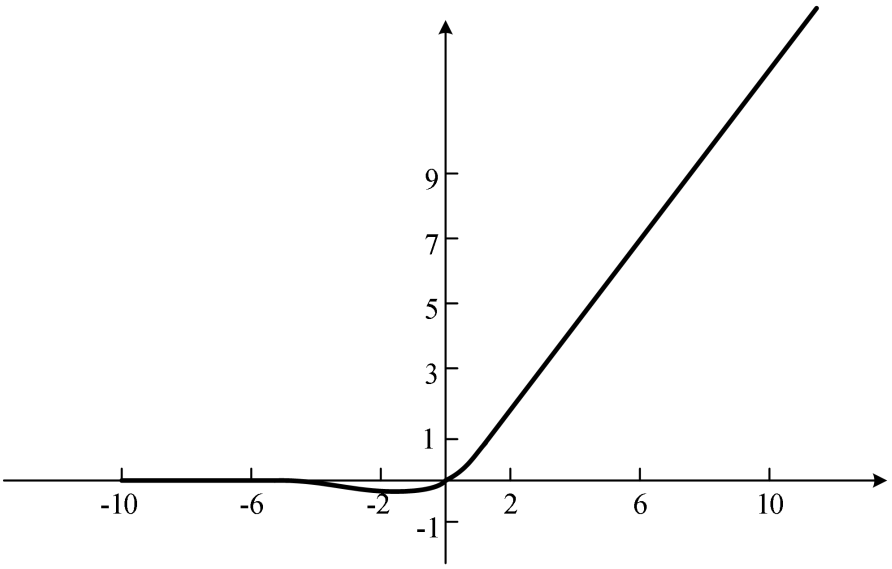


图 4-3 SiLU 激活函数图像

Fig. 4-3 SiLU activation function image

ELAN 模块是一个高效的网络结构，如图 4-4 所示。其通过控制最短和最长的两条梯度路径来使网络能够学习到更多的特征。第一条路径经过一个1×1 的卷积模块改变通道数，第二条路径先经过一个1×1 的卷积模块改变通道数，随后经过 4 个 CBS 进行特征提取。最后将特征进行进行叠加来得到最终的特征提取结果。

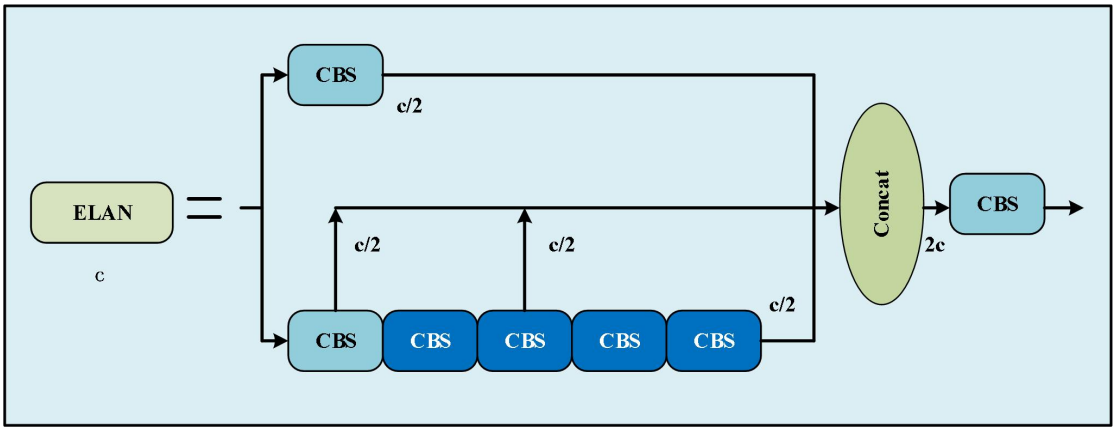


图 4-4 ELAN 结构图

Fig. 4-4 ELAN structure

MP-1 模块有两条路径，作用是进行下采样，如图 4-5 所示。第一条路径先经过 Maxpool 进行下采样，在经过 CBS 进行通道数的改变，第二条路径先经过 CBS 改变其通道数，然后再经过一个 CBS 进行下采样，最后将两条路径的结果进行叠加得到最终的结果。

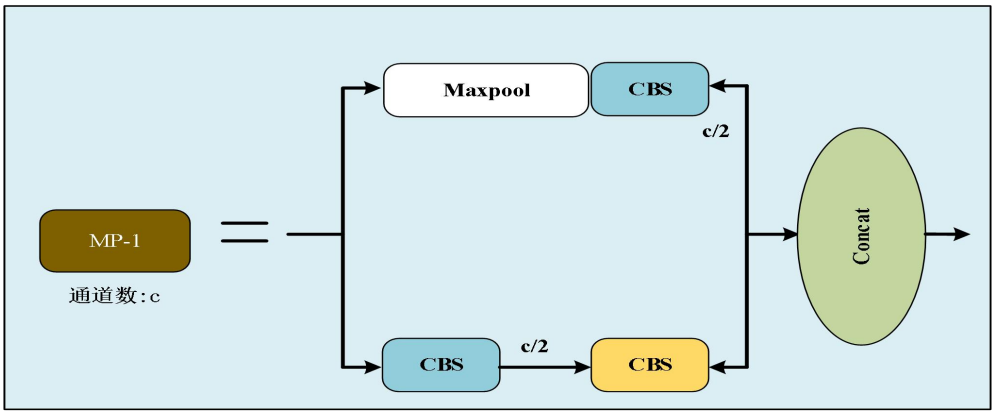


图 4-5 MP-1 结构图

Fig. 4-5　MP-1 structure

（2）Head

YOLOv7 通过 FPN+PAN 进行多尺度特征融合。其中，FPN 负责由上至下传输语义信息，而 PAN 负责由下至上传输位置信息，在此基础上，实现低层和高层的特征融合，从而来提升目标的检测效果。

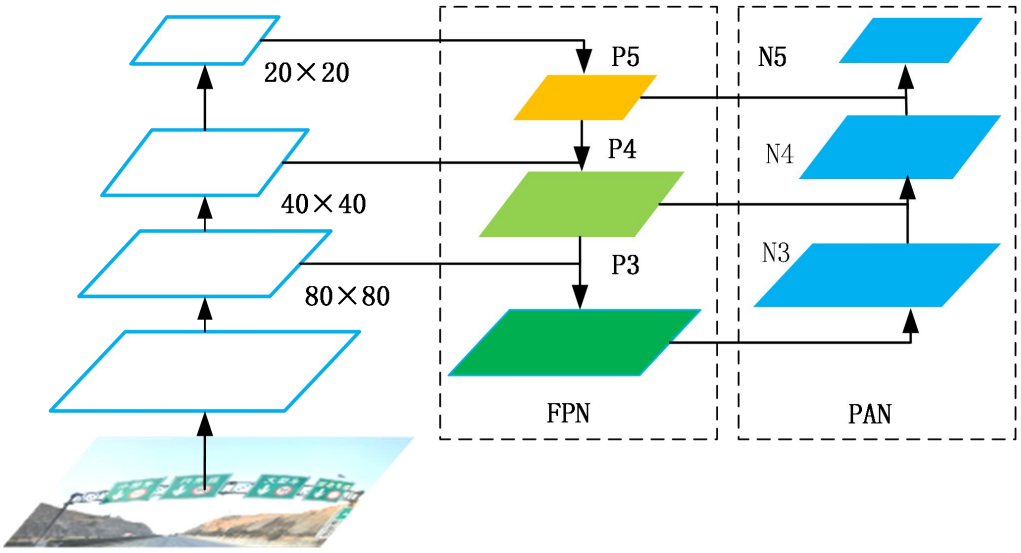


图 4-6 FPN+PAN 结构图

Fig. 4-6　FPN+PAN structure

SPPCSPC 通过最大池化来获得不同的感受野，从而增大网络的感受野。其网络结构如图 4-7 所示。在第一条路径中，四个不同尺度的最大池化有四种不同的感受野来区别不同尺度的目标，在第二条路径中，经过 CBS 改变通道数的变化。最后将两条路径进行合并，这样不仅能够减少计算量，还可使检测速度更快。

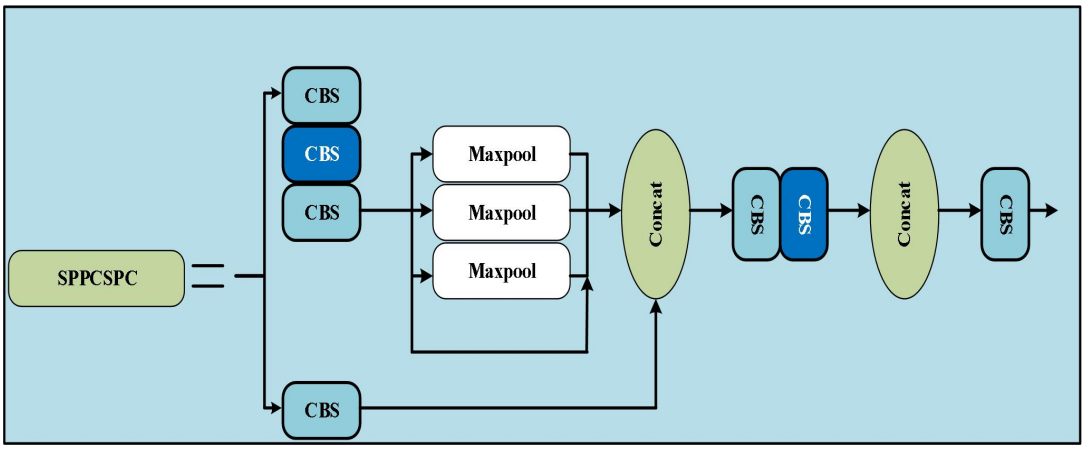


图 4-7 SPPCSPC 结构图

Fig. 4-7 SPPCSPC structure

4.2 模型改进

4.2.1 CFP 集中式特征金字塔

由于在交通标志图像中交通标志经常出现在图像的边缘区域，所以在进行交通标志识别时必须充分考虑到输入图像的边界信息，但是这种信息会随着网络的层数的增加而不断减少。

在 YOLOv7 模型中，由 RepConv 和 E-ELAN 模块进行对输入图像特征的提取。但由于其都将重点放在了相同层间的特征的相互作用，这会导致 YOLOv7 对处于边界的特征信息得不到有效的提取。为解决上诉问题，选用 CFP 作为特征融合方式，该方法不但能够捕捉到全局的长距离依赖，而且能够进行整体、差异化的特征表达，从而达到更好的空间特征调整。CFP 结构如图 4-8 所示。

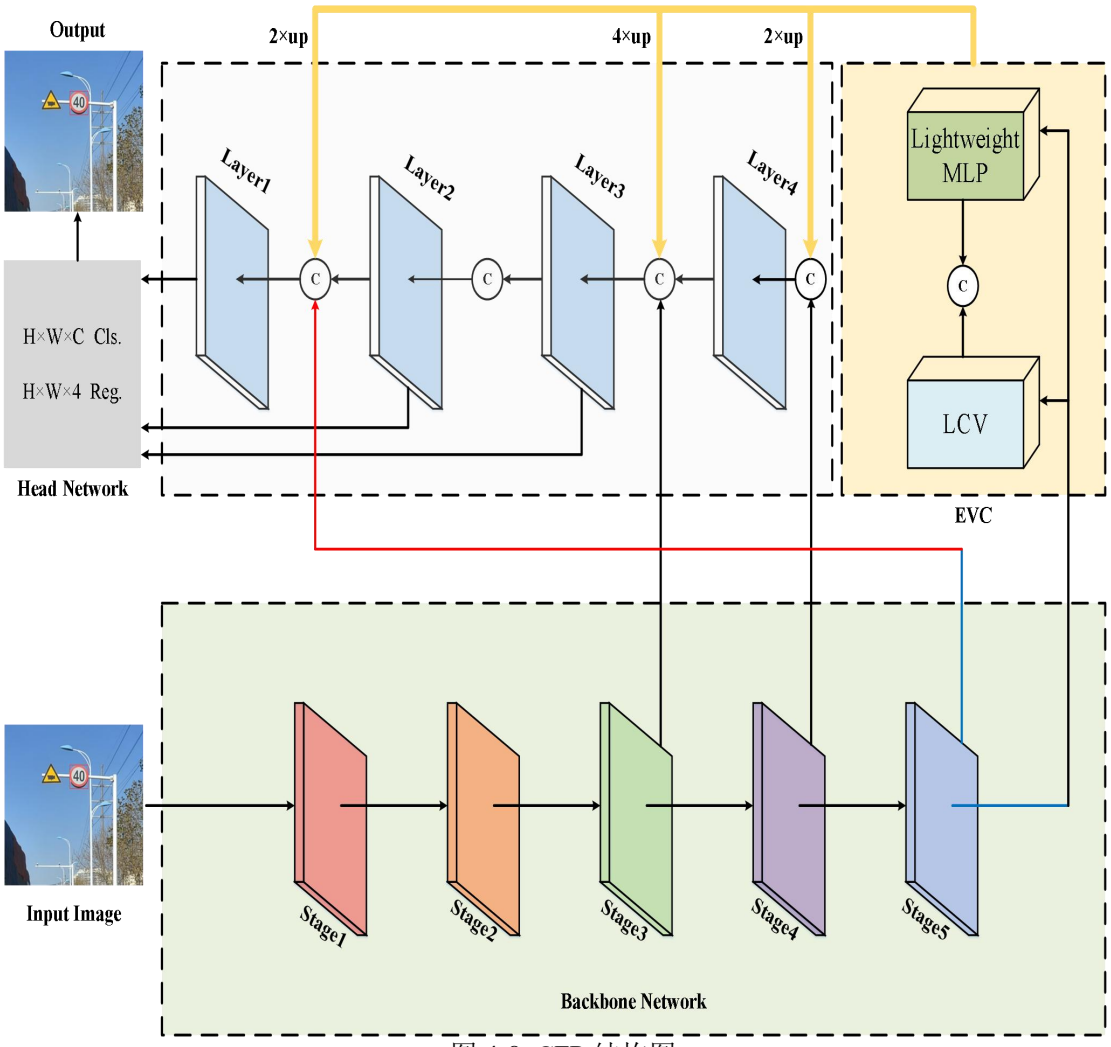


图 4-8 CFP 结构图

Fig. 4-8 CFP structure

在 CFP 中，引入了一个空间显示视觉中心(Explicit Visual Center，EVC)。EVC 由多层感知机（Multilayer Perceptron，MLP）和可学习视觉中心机制（Learnable Visual Center，LVC)组成。MLP 能够捕获长期依赖关系，LVC 则可以聚合图像边界信息。这两个模块的特征图沿通道维度拼接后同时具有这两个模块的优点，使得检测模型能够学习到全方位的特征表示。

4.2.2 CA 注意力机制

随着网络层数的加深，很容易造成图像细节特征的丢失，造成模型错误率偏高。基于此，采用 CA 注意力机制对神经网络的性能进行增强，CA 注意力机制能够通过学习不同通道之间的相关性来动态调整不同通道的权重，从而提高网络对重要特征的关注度，达到优化网络的目的。CA 模块结构如图 4-9 所示。

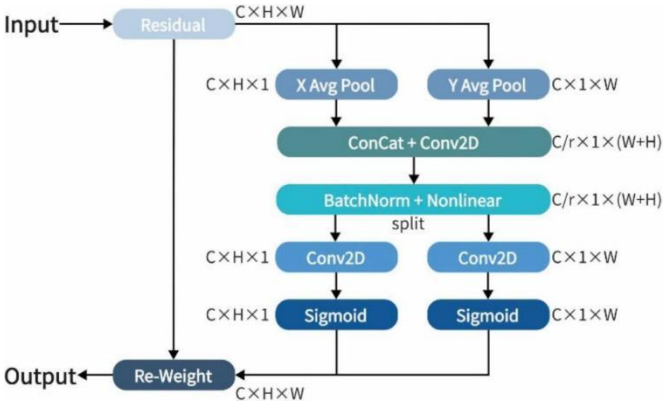


图 4-9 CA 模块结构图

Fig. 4-9 CA structure

CA 从高度 h 和宽度 w 两个方向对特征图平均池化，从而得到两个特征图。如式(4-2)(4-3)。

$$Z_c^h(h)=\frac{1}{W}\sum_{0\le i<W}X_c(h,i) \tag{4-2}$$

$$Z_c^w(w)=\frac{1}{H}\sum_{0\le j<H}X_c(j,w) \tag{4-3}$$

随后将获得全局感受野的 h 和 w 方向的特征图进行拼接，再通过1×1的卷积进行降维，最后将经过 BN 处理的特征图 F_1 送入 Sigmoid 激活函数得到特征图 f，如式 4-4。

$$f=\varphi(F_1([z^h,z^w])) \tag{4-4}$$

接着将 f 按照原先的 h 和 w 进行1×1的卷积，得到通道数与原来相同的特征图 F_h 和 F_w，然后再经过 Sigmoid 激活函数分别得到特征图在 h 和 w 上的注意力权重 g^h 和 g^w，如式(4-5)(4-6)所示：

$$g^h=\sigma(F_h(f^h)) \tag{4-5}$$

$$g^w=\sigma(F_w(f^w)) \tag{4-6}$$

最后对 g^h 和 g^w 在原始特征图上进行乘法加权计算，得到最终带有注意力权重的特征图，如式(4-7)所示：

$$y_c(i,j)=x_c(i,j)\times g_c^h(i)\times g_c^w(j) \tag{4-7}$$

4.2.3 损失函数

YOLOv7 采用 CIoU 损失函数计算回归损失，其不但考虑到了预测框与真实框之间的重叠面积、中心点距离，还考虑到了长宽比关系。CIoU 损失函数如式(4

-8~10)所示。

$$v=\frac{4}{\pi^2}(\arctan\frac{w^{gt}}{h^{gt}}-\arctan\frac{w}{h})^2 \tag{4-8}$$

$$\alpha=\frac{v}{(1-IoU)+v} \tag{4-9}$$

$$L_{CIOU}=IoU-(\frac{p^2(b,b^{gt})}{c^2}+\alpha v) \tag{4-10}$$

v是预测框与真实框的相似因子，α是权重函数，c是预测框与真实框的最小重叠区域对角线距离。

但由于其相似因子并不是真正的长宽比，因此在网络的训练过程中，当预测框收敛至真实长宽比时，CIoU 会成为模型进一步优化的阻碍。针对此问题，将α-IoU 引入，通过调节功率参数α可提高边界框的回归精度。其函数如式(4-11)所示。

$$L_{\alpha-IoU}=\frac{1-IoU}{\alpha},\alpha>0 \tag{4-11}$$

α-CIoU 对于 IoU 和惩罚项使用相同的功率α，可得$L_{\alpha-\mathrm{CIoU}}$，如式：

$$L_{\alpha-CIoU}=1-(IoU^{\alpha}-(\frac{\rho^2(B,B^{gt})}{c^2})^2)+(\mu v)^{\alpha} \tag{4-12}$$

4.2.4 改进后的 YOLOv7

改进后的 YOLOv7 结构图如图 4-10 所示。改进后的模型较原模型在参数量有所下降的情况下提升了检测精度，提高了检测能力。首先在预测网络内加入 CFP 模块来捕获长程特征依赖关系，提升神经网络的层内特征的交互能力，使模型获得了更加有效的特征表示。其次，通过引入 CA 注意力机制，使其兼顾横向和纵向的注意力，同时保留样本内目标的位置信息，提高模型对特征的表达能力。最后，对 CIOU 引入α-IoU，以此来提高边界框的回归精度，同时提高了模型的泛化能力。

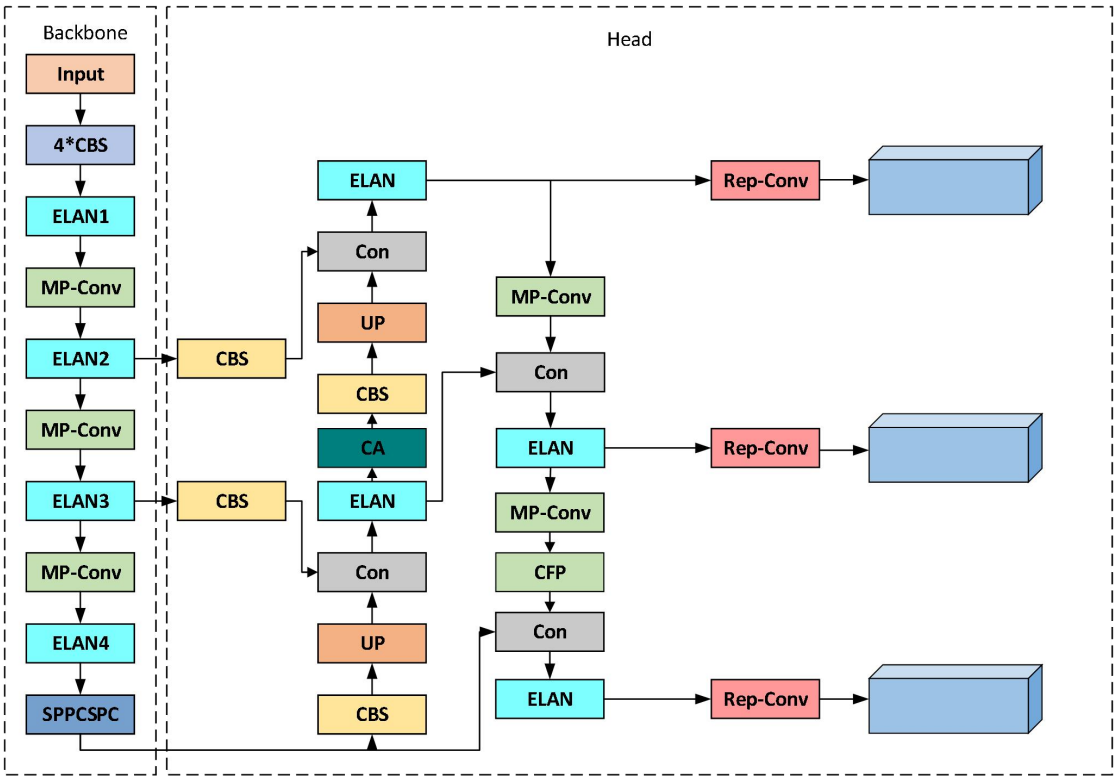


图 4-10 改进后的 YOLOv7 结构图

Fig.4-10 The improved structure of YOLOv7

4.3 实验结果及分析

4.3.1 实验环境及参数设置

受 GPU 显存限制，将 batch_Size 设置为 4，其余实验环境与第三章保持一致。

4.3.2 数据集类别均衡化

本章采用 TT100K 进行实验。该数据集中所有图像的尺寸大小都为 2048*2048 像素。由于一些类别的数据量过少，因此本章数据集中筛选出具有代表性的 42 类交通标志进行实验，共 9460 副图像。

通过对数据集分析发现各类标签分布不均衡如图 4-11 所示，通过挑选包含有数量小于 200 的交通标志类别的图像以及一些网络图像进行标签工作，以此使数据集进行丰富，促进数据集中不同类别的均衡化。如图 4-12 所示。通过对上诉数据集的筛选和剔除无法使用的样本后，最终得到 11249 张可用交通标志样本，按训练集：测试集=9:1 进行划分。

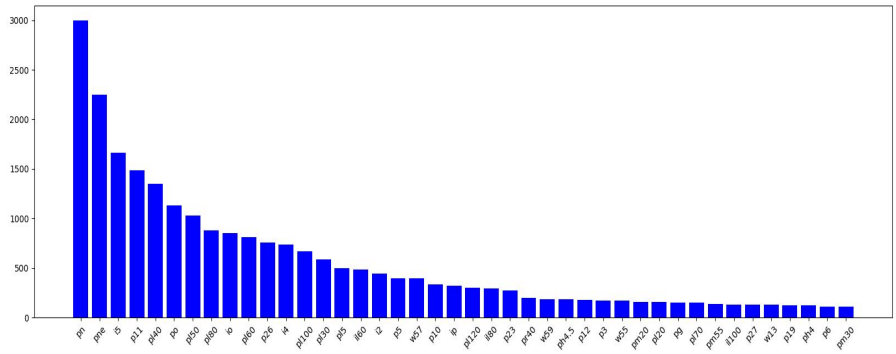


图 4-11 tt100K 标签数量

Fig. 4-11 tt100K Number of labels

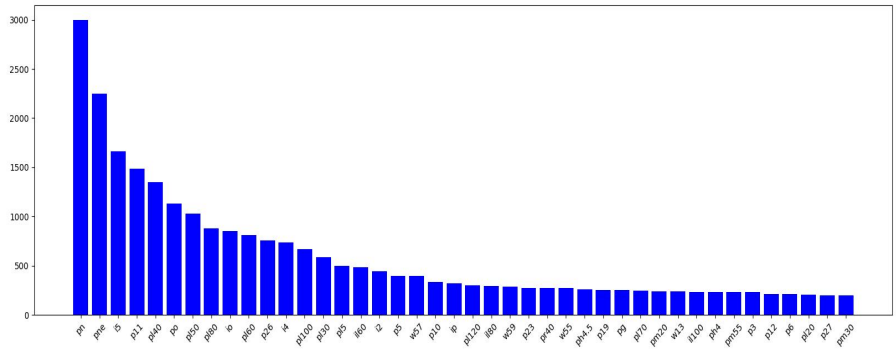


图 4-12 tt100K 数据均衡化后的标签数量

Fig. 4-12 Number of labels after tt100K data equalization

4.3.3 与主流网络模型进行对比

为进一步验证本章所提出的算法的有效性，将本章改进后的 YOLOv7 检测算法与其他先进方法进行比较。这些网络包括经典的单阶段以及两阶段目标检测模型，例如 Faster-RCNN、SSD、YOLOv3、YOLOv4,以及近年来的一些交通标志检测模型，例如井方科[61]等人基于多尺度特征融合改进的 STS-YOLO 和曲宸阳等人[62]改进的 YOLOv7。

从表 4-2 中可知，本文算法相较于基线 YOLOv7 算法，总样本平均精度提升了 4.6 个百分点，参数量仅增加 0.8M。在主流单阶段算法中，本文算法在各项指标上都明显优于 Faster-RCNN、YOLOv3、YOLOv4 算法，尽管 SSD 检测速度最快，但检测精度过低。同时，本章算法的总样本平均精度不仅比文献[61]、文献[62]分别高出 3.1、4.3 个百分点，检测速度也分别高 8.4、7 帧每秒。而对比文献[62]中改进的 YOLOv7 算法，本文在参数量上也具有优势。

表 4-2 与主流模型的对比试验
Table 4-2 Comparative experiments with mainstream models

Model	f_{mAP}	Params/m	FPS/帧
SSD	0.668	26.2	**84.2**
Faster-RCNN	0.695	137.1	4.3
YOLOv3	0.581	61.9	29.6
YOLOV4	0.742	64.4	41.7
YOLOV7	0.862	37.4	76.1
STS-YOLO	0.877	**6.2**	55.7
改进 YOLOv7	0.865	42.4	57.1
OURS	**0.908**	38.2	64.1

4.3.4 消融实验结果分析

为了验证在 TT100k 数据集尚设计消融实验验证本文改进的有效性，结果如表 1 所示。实验对比了分析了基线 YOLOv7 算法与（1）在 YOLOv7 中添加 CA 注意力机制模块；（2）在 YOLOv7 中采用 CFP 集中式特征金字塔；（3）对 YOLOv7 中的 CIOU 损失函数添加α-IoU。

为了验证所提出的改进措施是否有效，设计了一组消融实验，通过组合所提及的引入 CFP 集中式特征金字塔、CA 注意力机制以及添加α-IoU，来验证对网络的影响。具体实验数据见表 4-3。

从表 4-3 可以看出，基线 YOLOv7 算法的 mAP 值为 86.2%。添加 CA 注意力机制后，在参数量略微增加的情况下，mAP 提高了 2.3%。这表明 CA 模块在略微增加模型参数量的情况下大幅提升了模型的检测能力。

其次，在引入α-IoU 损失函数，通过获得更准确的边界框回归精度，实现模型对目标更精确的定位，对交通标志的 mAP 提高了 1.1%。

表 4-3 消融实验
Table 4-3 Ablation experiment

Model	f_{mAP}	Params/m	FPS/帧
YOLOv7	0.862	37.4	76.1
YOLOv7+CA	0.885	37.9	73.2
YOLOV7+CFP	0.881	35.1	68.3
YOLOV7+α-IoU	0.873	37.4	71.8
YOLOV7+CA+CFP+α-IoU	0.908	38.2	64.1

最后，在使用 CFP 特征金字塔模块后，模型在 mAP 提高了 1.9%的同时，参数量还有所下降。这表明 CFP 中的 MLP 和 LVC 模块生成的特征图学习到了全方位的特征表示，实现了对边界特征信息的有效提取，从而提升了平均检测精度。

综上，在各自应用以上三种方法时模型性能都得到了不同程度的提升，将三种改进方法同时添加到算法中在参数量小幅增加的情况下，mAP 提高了 4.6%。虽然改进后的算法在 FPS 性能上有所下降，但 64.1 帧每秒能够满足行车过程中对交通标志实时检测的要求。根据消融实验的结果可知，以上三种改进方案对模型的性能提升都是有效的。

4.3.5 与 SSD-RFC 网络模型进行对比

为验证改进 YOLOv7 相比于上一章所提出的 SSD-RFC 在交通标志小目标检测上的优越性以及模型是否更易部署在移动设备中，对 YOLOv7 与 SSD-RFC 在 TT100k 数据集中进行实验结果对比。

表 4-4 展示了 SSD-RFC 和 YOLOv7-CCα对 TT100K 数据集 42 个类别的检测结果。可以看出，YOLOv7-CCα在其中的 40 种交通标志的检测精度优于 SSD-RFC，其中在 ill100、p10、p19、ph4 等类别中检测精度得到了大幅度提高，这反映了对数据集进行均衡化处理是十分必要的。同时，从表中数据可以看出，一部分类别的识别精度已经接近饱和，如 i5、i160、p5 等，而有另外一部分类别的识别精度仍存在提高空间，如 io、p10、ph4.5 等，主要原因是这些类别的交通标志特征较为复杂，需要更多的图像样本进行模型训练和特征学习。

表 4-4 YOLOv7-CCα与 SSD-RFC 在各类别上的 AP 对比
Table 4-4 YOLOv7-CCα Comparison of APs with SSD-RFC in various categories

Class	SSD-RFC	YOLOV7-CCα
i2	0.913	**0.939**
i4	0.952	**0.959**
i5	**0.97**	0.968
il100	0.767	**0.953**
il60	0.992	**0.996**
il80	0.919	**0.96**
io	0.832	**0.89**
ip	0.923	**0.941**
p10	0.718	**0.891**
p11	0.879	**0.904**
p12	0.736	**0.889**
p19	0.684	**0.824**
p23	0.862	**0.927**
p26	0.913	**0.94**
p27	0.872	**0.985**
p3	0.914	**0.909**
p5	0.963	**0.964**
p6	0.614	**0.825**
pg	0.931	**0.954**
ph4	0.672	**0.838**
ph4.5	0.78	**0.867**
pl100	0.967	**0.982**
pl120	0.967	**0.993**
pl20	0.458	**0.749**
pl30	0.824	**0.912**
pl40	0.831	**0.915**
pl5	0.88	**0.942**
pl50	0.762	**0.884**
pl60	0.817	**0.91**
pl70	0.622	**0.79**
pl80	0.856	**0.924**
pm20	0.88	**0.947**
pm30	0.82	**0.862**
pm55	0.799	**0.958**
pn	0.944	**0.947**
pne	0.96	**0.963**
po	0.729	**0.818**
pr40	0.98	**0.99**
w13	0.536	**0.781**
w55	0.732	**0.832**
w57	0.877	**0.903**
w59	**0.829**	0.798

表 4-5 评估了 SSD-RFC 和 YOLOv7-CCα在多个指标上的表现，这些指标包括 mAP、Params 和 FPS。YOLOv7-CCα的 mAP、Params 和 FPS 相较于 SSD-RFC 均有了明显提升。mAP 提高了 7.8%，Params 减少了 5.9m，FPS 提升了 32.1。

SSD-RFC 性能下降的原因主要在于该数据集共包含 42 类交通标志，远多于 CCTSDB 的三类，同时该数据集比 CCTSDB 具有更加复杂的背景噪声，这使得模型的各项指标出现了不同程度的下降。在这种情况，本章提出的 YOLOv7-CCα具有更加优异的检测精度、更高的检测速率以及更加轻量化的结构，这使得本章算法能够在自动驾驶领域实现更好的应用。

表 4-5 与 SSD-RFC 的 Params 与 FPS 对比
Table 4-5 Comparison of Params and FPS with SSD-RFC

Model	mAP	Params/m	FPS/帧
SSD-RFC	0.83	44.1	32
YOLOv7-CCα	**0.908**	**38.2**	**64.1**

4.3.6 可视化检测结果

在训练完成后，选取部分图像进行测试并进行可视化，如图 4-13 所示（示例图中部分交通标志过小，已用红框进行标识）。选取的图像包含单目标、多目标、远距离、中距离、近距离等多种情形且图像场景均为高速公路、日常街道等生活中常见的场景。在无干扰因素场景下，如图 4-13（a）、（b）中，改进模型检测道路图像中所有的交通标志。在弱光和复杂街景情况下，如图 4-13（c）、（e）中，检测模型能够检测到绝大多数交通标志，但也存在着漏检情况。图 4-13（d）中交通标志在图像中极小，但仍然得到了有效检测且置信度达到了 0.81。图 4-13（e）中图 4-13（f）中三个交通标志紧密排列也全都得到了成功识别。综上所述，单目标图像在近景和远景中都具有良好的检测效果，对于近距离多目标也具有可靠的性能。

图 4-13 交通标志提取示例

Fig. 4-13 Example of traffic sign extraction

4.4 小结

本章针对第三章所提算法在小型交通标志中检测效果不佳的问题提出了基于 YOLOv7 的网络改进模型。通过以下三种方法对模型进行了改进，首先是将 CFP 集中式特征引入到预测网络中，提高特征的层内调节能力，同时也能降低模型的参数量。其次添加 CA 注意力机制，学习不同通道之间的相关性来动态调整权重，提高模型对重要特征的关注度。最后优化分类及定位的损失函数，提高边界框的回归精度。由实验分析得知，通过综合运用以上三种改进方法，模型对小型交通标志的检测能力得到了有效提升。

第 5 章 总结和展望

5.1 论文总结

随着国内汽车行业的蓬勃发展，相信自动驾驶技术也会迎来大规模的应用。而交通标志检测作为自动驾驶技术的必要一环，能为驾驶员以及自动驾驶系统提供信息以供决策。因此，本文在交通标志检测方向上展开了研究，主要内容如下：

（1）分析了交通标志检测在自动驾驶中的关键作用和地位，接着通过对国内外对交通标志的研究现状进行了分析，最后梳理出当前交通标志检测面临的难点。

（2）介绍了卷积神经网络的工作流程和主要构造，对当前主流的基于区域提议策略的两阶段和基于回归框架的单阶段目标检测模型进行了论述，并对目标检测的评估标准和交通标志数据集进行了说明。

（3）通过分别对中大型交通标志检测以及小型交通标志检测的研究，提出了基于 SSD-RFC 网络和 YOLOv7-CCα网络的交通标志检测算法。对于中大型交通标志目标，通过改进 SSD 实现了有效检测，主要工作在以下几个方面。首先对数据集进行扩充以提高泛化性，在此基础上，引入残差网络 ResNet50 作为新的骨干网络，并将其与 CBAM 注意力机制相结合，使其能够更有效的提取目标特征信息。最后，采用新型特征融合模块来让浅层和深层的特征信息进行融合，使浅层和深层的特征信息有效互补。改进后的模型通过实验验证，证明了改进的正确性和有效性。随后为验证所改进算法对小型交通标志的检测是否有效，使用包含有大量小型交通标志数据集的 TT100K 进行验证，结果显示出 SSD-RFC 算法对小型交通标志的检测性能且鲁棒性不足，容易出现错检和漏检等问题。基于以上问题，通过主流网络实验对比，提出一种基于 YOLOv7-CCα网络的交通标志检测算法。该方法首先对数据集的样本进行均衡化。其次，引入集中式特征金字塔 CFP，在降低模型的参数量的前提下提高了对目标特征的提取能力。接着，通过添加 CA 注意力机制使模型更加关注感兴趣的区域。最后，对定位损失函数进行改进，提高边界框的回归精度。改进后的模型通过仿真实验模拟，表明了所提算法对小目标的检测性能有很大优势。

5.2 展望

随着我国汽车的自动驾驶的飞速发展，对交通标志的检测越来越受到更多人的关注。然后，由于数据样本本身存在着诸多的检测难点，仍有许多问题有待于深度探讨，以下总结的未来研究内容。

（1）实验数据集的丰富。在对密集小目标的检测中，本文仍然无法完全避免对交通标志的错检、漏检。这不仅有交通标志类别过多的原因，也因为数据集中各交通标志类别数量分布极不均衡，这都影响了对交通标志检测的效果。未来的研究可以对现有数据集进行扩充或通过不断开发新的数据集来满足交通标志检测需求。

（2）更加高效轻量的网络。虽然本文提出的检测方法减少了一定的参数量，但随着自动驾驶的快速发展，未来对检测的速率要求必然是越来越高。因此未来的研究方向可以聚集于在不影响检测精度的前提下尽可能的压缩模型参数量，以不断提高模型的检测速度，更好的满足在实际场景中的应用。

（3）迁移学习。当前交通标志数所覆盖的应用场景及对象种类较少。随着自动驾驶技术的进一步发展与应用，交通标志信息会越来越复杂。然而，对每一个数据集都进行大规模的深度学习训练是不太现实的。所以，后续研究可进一步探究将已有的模型运用于新的样本上，以达到更好的检测效果。

参考文献

[1] Khorshidi A, Ainy E, Soori H. Iranian road traffic injury project: Assessment of road traffic injuries in Iran in 2012[J]. Injury Prevention, 2016, 22: 2016.

[2] Davidson P, Dharmaratne S. Disastrous but preventable: Road traffic accidents[J]. Health Care for Women International, 2016, 37: 706.

[3] Haydari A, Yılmaz Y. Deep reinforcement learning for intelligent transportation systems: A survey[J]. IEEE Transactions on Intelligent Transportation Systems, 2020, 23(1): 11-32.

[4] Zhang J, Wang F, Wang K, et al. Data-Driven Intelligent Transportation Systems: A Survey[J]. IEEE Transactions on Intelligent Transportation Systems, 2021, 12(4): 1624-1639.

[5] 张帅,古玉锋,凌浩,等.基于高频车站及时间窗的立体轨道交通系统智能调度算法[J].计算机应用研究, 2023, 41(5):1-7.

[6] Liang Z, Shao J, Zhang D, et al. Traffic sign detection and recognition based on pyramidal convolutional networks[J]. Neural Computing and Applications, 2020, 32: 6533-6543.

[7] Lv Z, Li Y, Feng H, et al. Deep learning for security in digital twins of cooperative intelligent transportation systems[J]. IEEE transactions on intelligent transportation systems, 2021, 23(9): 16666-16675.

[8] Malik R, Khurshid J, Ahmad S N. Road Sign Detection and Recognition using Color Segmentation, Shape Analysis and Template Matching[C].International Conference on Machine Learning & Cybernetics. IEEE, 2007: 3556-3560.

[9] 申中鸿，付梦印，杨毅，等. 基于背景像素突变检测的交通标志图像分割[J]. 计算机应用研究, 2012, 29(9): 3531-3535.

[10] Chakraborty S, Deb K. Bangladeshi road sign detection based on YCbCr color model and DtBs vector[C].2015 International Conference on Computer and Information Engineering, 2016: 158-161.

[11] Gao X W, Podladchikova L, Shaposhnikov D, et al. Recognition of traffic signs based on their colour and shape features extracted using human vision models[J]. Journal of Visual Communication and Image Representation, 2006, 17(4): 675-685.

[12] Lahmyed R, Ansari M E, Kerkaou Z. Automatic road sign detection and

recognition based on neural network[J]. Soft Computing, 2022: 12(2): 1-22.

[13] Yakimov P, Fursov F. Traffic signs detection and tracking using modified Hough transform[C].12th International Joint Conference on e-Business and Telecommunications, Colmar, 2015: 22-28.

[14] Saxena P, Gupta N, LASKAR, et al. A study on automatic detection and recognition techniques for road signs[J]. International Journal of Computational Engineering Research, 2015, 5(12): 24-28.

[15] Kaur S, Singh I. Comparison between edge detection techniques[J]. International Journal of Computer Applications, 2016, 145(15): 15-18.

[16] Boujemaa K S, Bouhoute A, Boubouh K, et al. Traffic sign recognition using convolutional neural networks[C].2017 International Conference on Wireless Networks and Mobile Communications, 2017: 1-6.

[17] Garcia-Garido M A, Sotelo M A, Martin-Gorostiza E. Fast traffic sign detection and recognition under changing lighting conditions[C].2006 IEEE Intelligent Transportation Systems Conference, Toronto, ON, Canada. IEEE, 2006. 811-816.

[18] Loy G, Barnes N. Fast shape based road sign detection for a driver assistance system[C].2004 IEEE International Conference on Intelligent Robots and Systems. IEEE, 2004, 1:70-75.

[19] Fabian P, Ga V, Alexandre G, et al. Scikit-learn: Machine learning in python[J]. Journal of Machine Learning Research, 2011, 12: 2825-2830.

[20] Wang J, Zhang T J, Cheng Y, et al. Deep learning for object detection: A survey[J]. Computer Systems: Science & Engineering, 2021, 38(2): 165-182.

[21] Dubey U, Chaurasiya R K. Efficient traffic sign recognition using CLAHE-based image enhancement and ResNet CNN architectures[J]. International Journal of Cognitive Informatics and Natural Intelligence (IJCINI), 2021, 15(4): 1-19.

[22] 李学军，权林霏，刘冬梅，等.基于 Faster-RCNN 改进的交通标志检测算法[J]. 吉林大学学报(工学版), 2023, 16(5):1-10.

[23] Ren S Q, He K M, Girshick R, et al. Faster RCNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39(6): 1137-1149.

[24] 赵子婧,刘宏哲,曹东璞.基于 Libra R-CNN 改进的交通标志检测算法[J].机械工程学报,2021,57(22):255-265.

[25] 赵友章,吕进.基于改进 SSD 的交通标志检测算法[J].电子测量技

术,2023,46(07):151-158.

[26] Liu W, Anguelov D, Erhan D, et al. Ssd: Single shot multibox detector[C].Computer Vision – ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11 – 14, 2016, Proceedings, Part I 14. Springer International Publishing, 2016: 21-37.

[27] Chen X, Luo X, Wu Z, Qin X, Shang J, Li B, Wang M, Wan H. A VGGNet-Based Method for Refined Bathymetry from Satellite Altimetry to Reduce Errors[C]. Remote Sensing. 2022, 14(23):5939.

[28] Alyaa J. Jalil, Naglaa M. Reda. Infrared Thermal Image Gender Classifier Based on the Deep ResNet Model[C]. Advances in Human-Computer Interaction. 2022, 2022(3852054):11-17.

[29] Li K, Huang Z, Cheng Y C, et al. A maximal figure-of-merit learning approach to maximizing mean average precision with deep neural network based classifiers[C].2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, 2014: 4503-4507.

[30] 尹宋麟,谭飞,周晴等.基于改进 YOLOv4 模型的交通标志检测[J].无线电工程,2022,52(11):2087-2093.

[31] Bochkovskiy A, Wang C Y, Liao H Y M. Yolov4: speed and accuracy of object detection[J]. arXiv preprint.2020, 2004: 10934.

[32] 杨祥,王华彬,董明刚.改进 YOLOv5 的交通标志检测算法[J].计算机工程与应用,2023,59(13):194-204.

[33] Zhang H, Tian M, Shao G, et al. Target detection of forward-looking sonar image based on improved YOLOv5[J]. IEEE Access, 2022, 10: 18023-18034.

[34] WOO S, PARK J, LEE J Y. CBAM: convolutional block attention module[C].Proceedings of the European Conference on Computer Vision (ECCV), 2018:3-19.

[35] 李禹纬,付锐,刘帆.改进 YOLOv7 的轻量化交通标志检测算法[J].太原理工大学学报,2024,55(01):195-203.

[36] Wang C Y, Bochkovskiy A, Liao H Y M. YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C].Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. 2023: 7464-7475

[37]耿焕同，李嘉兴，蒋骏，等.面向自动驾驶的高精度实时语义分割算法架构[J].计算机科学,2024,32(03):1-12.

[38]钟明霞,姜柏军.红外热成像技术在 ADAS 系统中的应用[J].激光与红外,2022,52(05):721-725.

[39]易海涛,李博,刘旗等.基于改进 SSD 的轧制设备手部安全检测方法研究[J].制造业自动化,2023,45(02):22-29.

[40]Chen N, Chen Y, Blasch E, et al. Enabling smart urban surveillance at the edge[C].2017 IEEE International Conference on Smart Cloud (SmartCloud). IEEE, 2017: 109-119.

[41]Girshick R, Donahue J, Darrell T, et al. Region-based convolutional networks for accurate object detection and segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 38(1):142-158.

[42]He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[C].Proceedings of the IEEEconference on computer vision and pattern recognition, 2016: 770-778.

[43]Chapelle O, Haffner P, Vapnik V N. Support vector machines for histogram-based image classification[J]. IEEE transactions on Neural Networks, 1999, 10(5): 1055-1064.

[44]Girshick R. Fast R-CNN[C].IEEE International Conference on Computer Vision (ICCV), 2016: 1440-1448.

[45]Ren J, Yu Z Y, Liu J B, et al. Robust tracking using region proposal networks[J]. arXiv preprint arXiv:1705.10447, 2017.

[46]Redmon J, Divvala S, Girshick R, et al. You only look once: Unified real-time object detection[C].IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016: 779-788.

[47]Redmon J, Farhadi A. YOLO9000: Better, faster, stronger[C].IEEE conference on Computer Vision and Pattern Recognition, 2017: 6517-6525.

[48]Li G F, Liu X, Tao B, et al. Research on ceramic tile defect detection based on YOLOv3[J]. International Journal of Wireless and Mobile Computing, 2021, 21(2): 128-133.

[49]Wang C Y, Liao H Y M, Wu Y H, et al. CSPNet: A new backbone that can enhance learning capability of CNN[C].Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops, 2020: 390-391.

[50]ZHENG Z, WANG P, REN D, et al. Enhancing geometric factors in model learning and inference for object detection and instance segmentation[J]. IEEE Transactions on Cybernetics, 2021, 52(8): 8574-8586.

[51]Zhang M, Wang C, Yang J, et al. Research on Engineering Vehicle Target Detection in Aerial Photography Environment based on YOLOX[C].2021 14th International Symposium on Computational Intelligence and Design (ISCID), 2021: 254-256.

[52]Zhang J M, Huang M T, Jin X K, et al. A real-time Chinese traffic sign detection algorithm based on modified YOLOv2[J]. Algorithms, 2017, 10(4): 127.

[53]Zhe Z, Liang D, Zhang S, et al. Traffic-Sign Detection and Classification in the Wild[C].2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016.

[54]HU J, SHEN L, ALBANIE S. Squeeze-and-excitation networks[C].Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018:7132-7141.

[55]霍爱清, 南思媛, 胥静蓉.改进YOLOX的弱光线道路交通标志检测[J].电子测量技术, 2023, 46(06):62-67.

[56]Quan Y, Zhang D, Zhang L, et al. Centralized feature pyramid for object detection[J]. IEEE Transactions on Image Processing, 2023:37-46-.

[57]Hou Q, Zhou D, Feng J. Coordinate attention for efficient mobile network design[C].Proceedings of the IEEE/CVF conference on computer vision and pattern recognition. 2021: 13713-13722.

[58]He J, Erfani S, Ma X, et al. alpha-IoU: A family of power intersection over union losses for bounding box regression[J]. Advances in Neural Information Processing Systems, 2021, 34: 20230-20242.

[59]Liu S, Qi L, Qin H F, et al, Path Aggregation Network for Instance Segmentation[C].2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018: 8759-8768.

[60]Lin T, Dollár P, Girshick R, et al. Feature pyramid networks for object detection[C].2017 IEEE Conference on Computer Vision and Pattern Recognition, 2017: 936-944.

[61]井方科,任红格,李松.基于多尺度特征融合的小目标交通标志检测[J].激光与光电子学进展，2023, 29(07):1-13.

[62]曲宸阳,程艳云.基于改进YOLOv7的交通标志检测算法[J].微电子学与计算机，2023, 29(07):1-11.

攻读学位期间发表的学术成果

已发表的论文：

[1] 杜云龙,强俊，王洪铭等.应用于交通标志的单步多目标检测方法研究[J]. 重庆工商大学学报(自然科学版)。

专利申请：

[1] 一种优化SSD检测模型训练方法及小目标检测方法[P]. 安徽省：CN114821265A,2022-07-29.

参与项目：

[1] 安徽省高校优秀拔尖人才培育资助项目（gxyqZD2021123）

[2] 印刷品印刷质量图像检测系统的研发（HX-2022-11-129）

[3] 文检三维微距测量分析系统的研发（HX-2021-12-134）

致 谢

在毕业论文即将完成之际，我衷心感谢所有在我研究生生涯中给于我帮助、支持和鼓励的人。

我的导师强俊对我论文的晚餐进行了全面细致的指导，从论文的选题、拟定提纲、研究工作的展开一直到论文的完成都倾注了导师的大量的心血。在此，我向导师致以崇高的敬意和衷心的感谢。

其次，我要感谢实验室的同学们。在研究生阶段，我们共同度过了许多难忘的时光。在学术上，我们互相学习、互相启发；在生活中，我们互相照顾、互相支持。还有我的家人，他们在学习和生活上都给予我无微不至的关心和支持，才有了我现在的论文研究成果。

最后，向在我这三年学习和生活中曾给予我支持和教导、扶持和帮助的老师和同学们表示深深的感谢。