

DOI:10.3969/j.issn.1672-2337.2025.01.009

基于轻量化卷积神经网络车载雷达图像目标识别方法

李家强, 汪星宇, 陈金立, 姚昌华

(南京信息工程大学电子与信息工程学院, 江苏南京 210044)

摘要: 针对车载毫米波雷达距离-方位图像细节模糊、目标占比小, 卷积神经网络模型复杂难以在端侧部署的问题, 本文提出了一种基于轻量化卷积神经网络 YOLOv5s 的车载雷达图像目标识别方法。首先结合 Ghost 卷积设计轻量化解耦头, 并行处理检测与分类问题; 其次设计融合注意力机制的 Concat_att 模块并引入更具边界框定位敏感性的网络损失函数 EIoU Loss, 充分提取特征图中小目标细节信息, 加速网络收敛, 提升网络精度; 最后通过 Slim 剪枝进一步压缩模型存储空间和计算量。实验结果表明, 当模型大小缩减至原始 YOLOv5s 网络的 76.8% 时, mAP@0.5 与 mAP@0.5:0.95 较原始网络分别提升了 2.7% 和 2.8%, 适用于小目标检测, 并能同时满足目标识别精度与实时性要求, 适合部署至车载嵌入式系统中。

关键词: 雷达图像; YOLOv5s; 轻量化; 注意力机制; 模型剪枝

中图分类号: TN957.51

文献标志码: A

文章编号: 1672-2337(2025)01-0082-10

引用格式: 李家强, 汪星宇, 陈金立, 等. 基于轻量化卷积神经网络车载雷达图像目标识别方法[J]. 雷达科学与技术, 2025, 23(1): 82-91.

LI Jiaqiang, WANG Xingyu, CHEN Jinli, et al. Object Recognition Method for Automotive Radar Images Based on Lightweight Convolutional Neural Network [J]. Radar Science and Technology, 2025, 23(1): 82-91.

Object Recognition Method for Automotive Radar Images Based on Lightweight Convolutional Neural Network

LI Jiaqiang, WANG Xingyu, CHEN Jinli, YAO Changhua

(School of Electronics and Information Engineering, Nanjing University of Information Science and Technology, Nanjing 210044, China)

Abstract: To address the issues of blurry details and small target proportions in automotive millimeter-wave radar images, as well as the complexity of convolutional neural network models that are difficult to deploy on the edge, an automotive radar image target recognition method based on lightweight convolutional neural network YOLOv5s is proposed. First, a lightweight decoupled head is designed by incorporating Ghost convolution, enabling parallel processing of detection and classification tasks. Next, the Concat_att module enhanced with attention mechanism is designed, and a more boundary-sensitive network loss function EIoU Loss is introduced to fully extract detailed information of small objects in feature maps, accelerating network convergence and improving accuracy. Finally, Slim pruning is applied to further compress the storage space of the model and reduce computational complexity. The experimental results indicate that when the model size is reduced to 76.8% of the original YOLOv5s network, the mAP@0.5 and mAP@0.5:0.95 are respectively improved by 2.7% and 2.8% compared to the baseline network. This method is suitable for small target detection and meets both the precision and real-time requirements of target recognition, making it appropriate for deployment in automotive embedded systems.

Key words: radar images; YOLOv5s; lightweight; attention mechanism; model pruning

0 引言

伴随社会不断发展, 汽车已逐渐成为人们主要的交通工具之一。然而, 汽车数量急剧增加致

使道路交通系统面临更大压力, 安全隐患不容忽视。自动辅助驾驶技术具备协助甚至替代驾驶员保证车辆安全运行的潜力, 能够降低交通事故发生率、减轻并改善道路拥堵状况, 是汽车行业未来

收稿日期: 2024-08-06; 修回日期: 2024-09-11

基金项目: 国家自然科学基金(No.62071238)

发展的显著趋势^[1]。自动驾驶系统包含相机、激光雷达、毫米波雷达、红外线等传感器^[2-3],通过复杂道路检测算法为其提供可靠的环境信息感知支持。其中,毫米波雷达以高分辨率、小尺寸、易于集成并可在全天时全天候条件下工作等优势,成为当前研究的热点。

毫米波雷达距离-方位图像相较于光学图像在成像原理、数据特征、应用场景和处理方法上存在显著差异。得益于卷积神经网络(Convolutional Neural Networks, CNN)在计算机视觉领域中的应用,针对雷达图像中复杂的空间结构和大量的像素信息,CNN可以从中提取特征,有效识别目标和背景之间的差异,为自动驾驶系统提供技术支持^[4]。基于CNN的目标识别算法可以分为二阶段(two-stage)算法和一阶段(one-stage)算法。二阶段算法首先生成候选区域以确保足够的准确率和召回率,再通过CNN预测目标的检测和分类信息,代表方法有R-CNN^[5],Fast RCNN^[6],Faster RCNN^[7],SPPNet^[8]等;而一阶段算法直接回归目标的检测和分类信息,代表方法有SSD^[9],YOLO^[10-13]系列算法等。

随着CNN模型层数加深、参数量增多,不可避免地减缓推理速度,致使功耗剧增。针对车载雷达图像目标识别任务,需兼顾考虑计算资源限制、功耗、模型部署便捷性和实时性要求。由此,对神经网络压缩的研究在推动其端侧部署中至关重要。卷积神经网络压缩旨在利用其网络参数与结构中的冗余,在保证模型性能的前提下,得到更精简、参数量更少的模型。常见的压缩方法有结构设计、量化、剪枝和知识蒸馏等。

广大学者对上述方面开展了大量研究。文献[14]提出基于毫米波雷达和视觉融合的多数据源深度学习物体检测网络MS-YOLO,提高了检测精度,但不可避免地受到自制数据集的局限性和模型复杂性的影响;Xu等人^[15]基于YOLOv5算法,提出一种轻量级SAR舰船检测器Lite-YOLOv5,通过轻量化模块设计和网络剪枝,减少计算量,并通过集成多个模块来提高检测性能,SAR图像虽然提供较高的分辨率,但其在实时性、处理简便性、成本效益与适用性上通常不及雷达距离-方位图像;

Luo等人^[16]采用基于网络架构搜索算法(Network Architecture Search, NAS),利用自动化搜索技术,设计并优化深度学习模型架构,提升模型性能和效率,然而,其通常针对特定结构设计,因而缺乏可迁移性。此外,高计算成本、弱泛化能力以及有限搜索空间均限制了其在实际应用中的部署。模型剪枝方法涵盖了多种策略^[17-20],通过制定不同的剪枝规则和损失函数,以在各个粒度上聚焦不同的重点,实现针对特定结构或普适性剪枝方案的优化。

针对上述问题,本文提出一种基于轻量化卷积神经网络车载雷达图像目标识别方法:首先结合Ghost卷积设计轻量化解耦头,解耦并行处理检测与分类任务;其次设计Concat_att结构,融合注意力机制,充分提取雷达图像细节信息;同时引入更具边界框定位敏感性的EIoU Loss,加速网络收敛;最后采用Slim模型剪枝方法,剪除卷积神经网络中不重要的通道以进一步压缩模型大小和计算量。实验结果表明,本文提出的方法在提升目标识别性能的同时减小模型的计算量和存储空间,且网络速度可以胜任实时任务需求,具有良好的应用前景。

1 YOLOv5s 网络结构

综合考量计算量与准确率,本文以YOLOv5s网络为基准,实现车载毫米波雷达距离-方位图像目标识别任务,原始网络结构如图1所示,由图片输入、主干网络、特征融合网络和检测头四部分构成。图片输入端采用K均值聚类算法(K-means)和马赛克数据增强(Mosaic Augmentation)等优化策略;主干网络以新型残差模块(C3)对图像特征多次提取,并通过空间池化金字塔(Spatial Pyramid Pooling Fast, SPPF)结构,加速网络处理。特征融合部分结合特征金字塔网络(Feature Pyramid Networks, FPN)与路径聚合网络(Path Aggregation Network, PANet),通过特征拼接模块融合不同层间特征图,网络多尺度预测能力得到提升。预测阶段,在三种不同尺度特征图上对目标位置与类别信息进行回归预测。

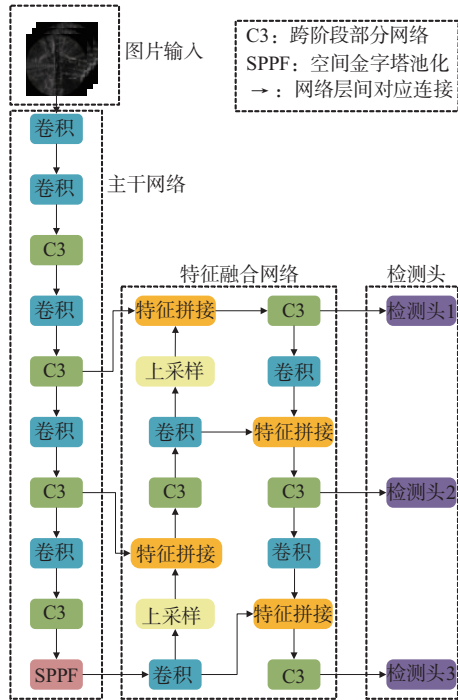


图1 YOLOv5s网络结构

2 卷积神经网络轻量化改进

2.1 结合 Ghost 卷积设计轻量化解耦头

YOLOv5 中检测头负责从特征图中预测目标的类别、边界框和置信度,实现目标的定位与分类,多种检测头结构如图 2 所示。理想情况下,检测与分类问题应彼此独立,但 YOLOv5s 网络中预测阶段检测头将检测和分类问题耦合,如图 2(a) 所示,导致车载毫米波雷达距离-方位图像中目标定位的不准确性增加,进而影响网络的性能和精度。

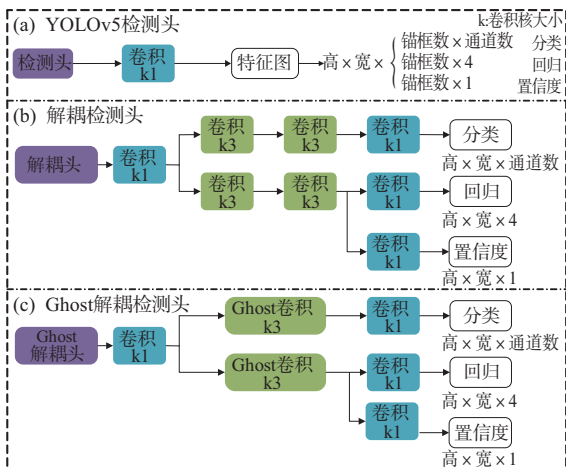


图2 Head 模块示意图

YOLOX^[21]通过解耦头(Decoupled Head, DH)的设计缓解上述问题,其结构如图 2(b)所示。其中分类用于预测目标类别信息;回归为预测框位置信息,通过 (x, y, h, w) 4 个坐标值表示;置信度用于区分前景或背景,判别预测框内是否包含物体。解耦头经特征整合后,通过并行的多分支卷积网络,将目标检测与分类任务解耦处理,以多次提取的特征获得相应结果,避免在雷达图像识别任务中受多尺度特征、噪声干扰和复杂的强度分布信息影响。但网络模型精度提升的同时,解耦头的引入不可避免地增加计算负担和模型大小,影响车载毫米波雷达图像目标识别任务的实时性能。

为解决上述问题,本文提出了 Ghost 解耦头(GhostDH)方法,旨在提升网络模型精度的同时,寻求计算量与性能之间的平衡。本文设计 GhostDH 通过两个关键步骤来优化原始解耦头的特征提取过程。首先如图 3 所示,使用 Ghost 卷积替换部分原始卷积操作,以提高计算效率。假定输入数据 $X \in \mathbb{R}^{c \times h \times w}$,其中 c 为输入特征图通道数, h 和 w 分别为高和宽。采用标准卷积,生成 n 张特征图,计算公式为

$$Y = X * f + b \quad (1)$$

式中 $*$ 代表卷积运算, b 是偏置项,输出特征图 $Y \in \mathbb{R}^{n \times h' \times w'}$ 具有 n 个通道, h' 和 w' 分别为卷积后的高和宽,滤波器 $f \in \mathbb{R}^{c \times k \times k \times n}$,其卷积核大小为 $k \times k$ 。通过式(1)可计算出标准卷积浮点操作数(FLOPs)为 $n \cdot h' \cdot w' \cdot c \cdot k \cdot k$ 。

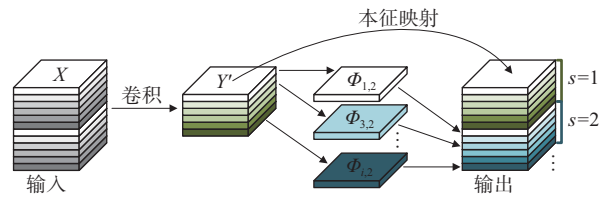


图3 GhostConv 示意图

相比,在 Ghost 卷积中,首先通过标准卷积生成 m 张特征图($m \leq n$),如式(2)所示:

$$Y' = X * f' + b' \quad (2)$$

输出特征图 $Y' \in \mathbb{R}^{m \times h' \times w'}$,滤波器 $f' \in \mathbb{R}^{c \times k \times k \times m}$,接着通过 Ghost 线性映射操作获得 $s - 1$ 组相似特征图,如式(3)所示:

$$y_{ij} = \Phi_{i,j}(y'_i), \forall i = 1, \dots, m, j = 1, \dots, s \quad (3)$$

式中, y_i' 是由标准卷积生成的第 i 张特征图, $\Phi_{i,j}$ 表示对于第 i 张特征图的第 j 次 Ghost 映射操作。Ghost 映射采用核为 $d \times d$ 的深度可分离卷积, 其卷积核大小与标准卷积核保持一致, 即 $d = k$ 。最后将两部分特征图进行拼接, 得到输出结果, 共 s 组特征图。

采用标准卷积与 Ghost 卷积的理论参数量压缩比 r_{compress} 如式(4)所示:

$$r_{\text{compress}} = \frac{n \cdot c \cdot k \cdot k}{\frac{n}{s} \cdot c \cdot k \cdot k + \frac{s-1}{s} n \cdot d \cdot d} \approx \frac{s \cdot c}{s+c-1} \approx s \quad (4)$$

式中, $s \ll c$ 即 Ghost 分组数远小于输出特征图通道数。同样, 二者之间的浮点数运算量之比 r_{FLOPs} 可以表示为

$$r_{\text{FLOPs}} = \frac{n \cdot h' \cdot w' \cdot c \cdot k \cdot k}{\frac{n}{s} \cdot h' \cdot w' \cdot c \cdot k \cdot k + \frac{s-1}{s} n \cdot h' \cdot w' \cdot d \cdot d} = \frac{c \cdot k \cdot k}{\frac{1}{s} \cdot c \cdot k \cdot k + \frac{s-1}{s} \cdot d \cdot d} \approx \frac{s \cdot c}{s+c-1} \approx s \quad (5)$$

分析可知, Ghost 卷积相较于标准卷积具有较低的浮点数运算量和模型大小, 以共享权值的方式, 高效提取雷达图像中目标空间关系、强度分布等重要信息。其次, 通过对解耦头的结构进行重新设计, 删除了原始解耦头中的部分卷积层, 简化特征提取过程, 减少了重复提取特征带来的冗余计算。改进后的结构如图2(c)所示。

GhostDH 通过解耦处理雷达图像目标检测与分类任务, 有效提高雷达目标检测与分类网络模型的精度, 增强泛化性、鲁棒性的同时, 具有 FLOPs 低和模型小的优势, 为车载计算资源有限的场景提供了一种有效的解决方案。

2.2 CBAM 注意力机制融入 Concat_att 模块

在神经网络中, 通常通过在某一层后添加注意力机制的方式, 自适应地关注特征图中不同部分的重要性, 以缓解通过多次下采样过程中特征图细节信息丢失问题, 理论上可提升网络性能。但许多情况下, 注意力机制的添加反而会降低网络精度, 甚至增加过拟合风险。针对上述问题, 本文设计了 Concat_att 模块, 通过融合网络中不同层

次的特征图, 强化提取雷达图像中小目标的细粒度信息。

Concat_att 模块结构如图4所示, 其在拼接整合网络模型不同层间特征图后, 通过 CBAM 注意力机制^[22]对拼接后特征图 X 进行进一步特征提取, 依次采用通道注意力和空间注意力模块, 先后在特征图的通道维度和空间维度上关注雷达图像中难以检测的小目标信息所在区域, 动态调整通道的权重, 提取更具代表性的特征 Y , 从而更好地定位和识别小目标。

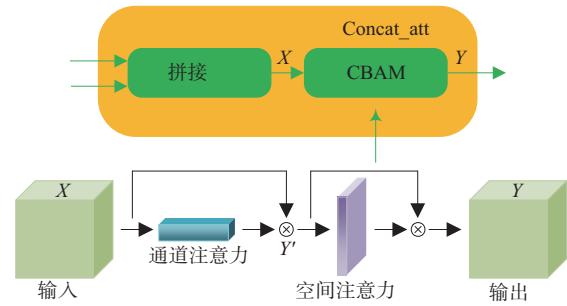


图4 Concat_att 模块

图5(a)、(b)分别代表 PANet 中上采样和下采样中所使用的 Concat_att 模块, 二者在特征整合阶段输入有所差异。

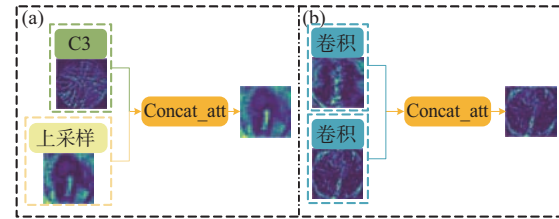


图5 上采样和下采样 Concat_att 模块

Concat_att 模块旨在解决雷达图像目标检测任务中小目标信息丢失的问题, 通过合理整合低层特征图和高层特征图, 以兼顾细节信息和语义信息。特征的重要程度与融合后特征图明暗变化呈正相关, 通过学习和选择性增强重要特征, 使网络更专注于雷达距离-方位图像中的小目标信息, 弱化无关的环境信息, 减少漏检与误检概率, 提升网络精度。

2.3 EIou 损失函数

在模型预测阶段, YOLOv5s 通过最小化损失函数, 使红色预测框逐渐贴近蓝色真实框, 实现目标检测, 如图6所示。传统 YOLOv5s 网络采用 CIou

(Complete Intersection over Union)作为边界框的损失函数,CIoU计算公式如式(6)~式(9)所示:

$$L_{\text{CIoU}} = 1 - \text{IoU} + \frac{\rho^2(b, b^{gt})}{(w^c)^2 + (h^c)^2} + \alpha\nu \quad (6)$$

$$\nu = \frac{4}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right)^2 \quad (7)$$

$$\alpha = \frac{\nu}{(1 - \text{IoU}) + \nu} \quad (8)$$

$$\text{IoU} = \frac{|B \cap B^{gt}|}{|B \cup B^{gt}|} \quad (9)$$

式中, IoU 表示预测框与真实框的重叠区域与并集区域之比, C 表示包围预测框 B 与真实框 B^{gt} 的最小外接矩形, b 表示预测框的中心点, w, h 分别为宽和高, b^{gt}, h^{gt}, w^{gt} 为对应真实框的相关信息, $\rho(\cdot)$ 表示欧几里得距离, α 为权重系数, ν 用于评估两个矩形框纵横比的一致性。CIoU 损失函数综合考量边界框重叠面积、中心点距离和边界框纵横比一致性, 但其存在如下问题: 当二者纵横比 ν 相同时, 即 $\{(w = kw^{gt}, h = kh^{gt}) | k \in \mathbb{R}^+\}$, 式(7)置零, 约束项梯度的消失不利于网络训练; 同时, 式(7)对 w, h 分别求偏导可以得到式(10)和式(11)。观察两式符号相反, 即网络训练过程中, 预测框无法在宽、高方向上以同样趋势变化, 使得网络对于雷达图像中的小目标不敏感, 进而导致网络收敛缓慢、性能不佳。

$$\frac{\partial \nu}{\partial w} = \frac{8}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right) * \frac{h}{w^2 + h^2} \quad (10)$$

$$\frac{\partial \nu}{\partial h} = -\frac{8}{\pi^2} \left(\arctan \frac{w^{gt}}{h^{gt}} - \arctan \frac{w}{h} \right) * \frac{w}{w^2 + h^2} \quad (11)$$

在此基础上, 针对雷达图像中小目标难以定位和识别的问题, 本文引入 EIoU (Efficient Intersection over Union)^[23] 损失函数, 如图 6 所示。EIoU 计算公式如下:

$$L_{\text{EIoU}} = 1 - \text{IoU} + \frac{\rho^2(b, b^{gt})}{(w^c)^2 + (h^c)^2} + \frac{\rho^2(w, w^{gt})}{(w^c)^2} + \frac{\rho^2(h, h^{gt})}{(h^c)^2} \quad (12)$$

EIoU 不仅包含中心点损失, 也将预测框和真实框纵横比一致性转化为宽度与高度的真实差异去考虑, 更好地适应雷达图像特征图中不同形状、大小的目标; 同时, EIoU 对于小目标的敏感度更高, 减少了损失函数受雷达图像分辨率限制和背景干扰影响; 此外, 相较于 CIoU, EIoU 损失计算简单且高效, 特别是在车载资源受限环境下, 如嵌入

式系统或边缘设备中, 使用 EIoU 损失函数可以更轻松实现目标识别任务, 提升模型的部署效率。

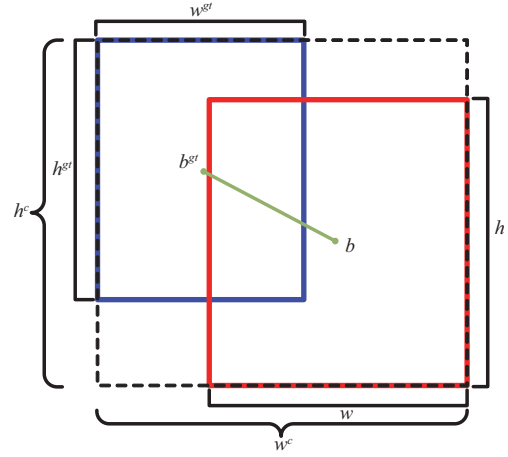


图 6 EIoU 示意图

2.4 模型剪枝

雷达图像特征图中, 不同通道包含不同的特征信息。有些通道在目标检测任务中对区分类别和定位目标更重要, 而其他通道可能包含冗余或不相关信息。针对 YOLOv5 模型存储空间、计算量大等问题, 本文使用 Slim 方法实现深度神经网络信道级稀疏, 利用批处理规范化层 (Batch Normalization, BN) 中的尺度信息来有效地识别和修剪网络中不重要的通道。在实际车载部署中, 模型剪枝减少模型尺寸的同时保持模型性能, 降低功耗并提高推理速度。

BN 层用于规范化神经网络模型中的输入数据, 具有加速网络训练、减小梯度消失、抑制过拟合等优势, 其表达式为

$$\hat{z} = \frac{z_{\text{in}} - \mu_B}{\sqrt{\sigma_B^2 + \varepsilon}}; z_{\text{out}} = \gamma \hat{z} + \beta \quad (13)$$

式中 z_{in} 和 z_{out} 为 BN 层的输入和输出, B 表示批次信息, μ_B 和 σ_B 分别表示当前输入批次的均值和标准差, γ 和 β 则是可训练的仿射变换参数, 二者初始值设置为 1 和 0, 分别对应尺度和位移信息。选用 BN 层的 γ 参数作为稀疏化裁剪的尺度因子, 其大小与对应通道重要性呈正相关。通过训练 γ 参数并对网络加以修剪, 剪除网络中不重要的通道, 从而减小计算量和模型大小。Slim 剪枝方法如图 7 所示。

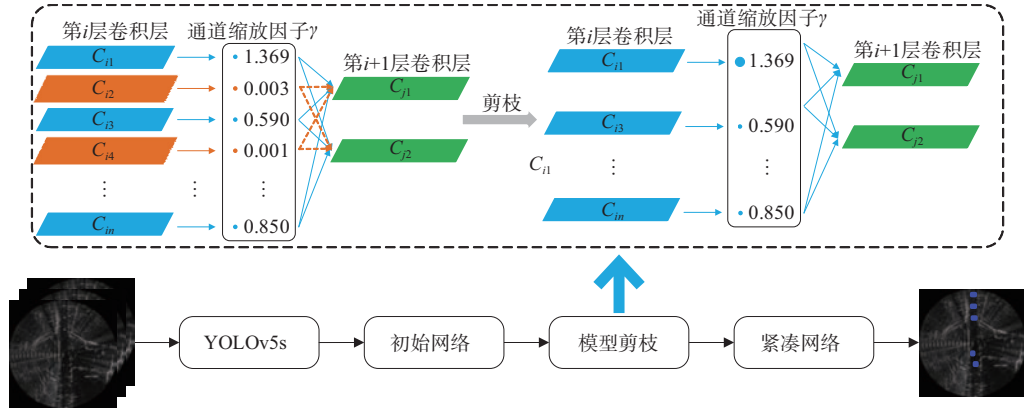


图7 Slim剪枝方法

网络模型稀疏训练中,损失函数为

$$L = \sum_{(x,y)} l(f(x,W),y) + \lambda \sum_{\gamma \in \Gamma} g(\gamma) \quad (14)$$

损失函数包含两个部分:第一部分 l 为网络正常训练损失,其中 (x,y) 为训练输入的雷达距离-方位图像和标签信息, W 为可训练权重;第二部分 g 为剪枝损失, γ 为可训练的仿射变换参数,对应尺度信息, λ 是稀疏性惩罚,用于平衡两项损失函数,在文中取 $\lambda = 10^{-4}$ 。本文采用L1正则化作为剪枝项损失,即 $g(\gamma) = |\gamma|$ 。

在网络修剪小比例因子通道后,经微调,循环往复,直至网络达到指定的计算量压缩比,即原始模型与剪枝后模型计算量之比,模型剪枝流程如图8所示。

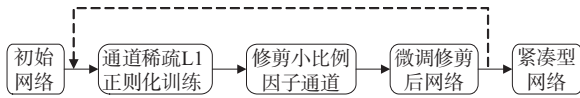


图8 模型剪枝流程

综上,本文所提方法的整体网络结构如图9所示,其中改进部分以红色标出。

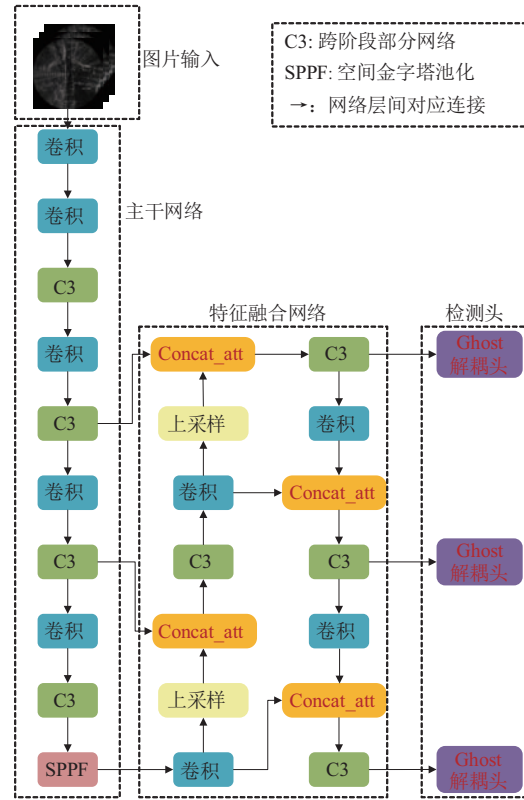


图9 本文方法整体网络结构

3 实验结果与分析

3.1 实验数据集与处理环境设置

本文使用的RADIATE数据集^[24]由英格兰赫瑞-瓦特大学发布。RADIATE数据集涵盖多种不同类型参与者在各种天气条件下和驾驶场景下的高分辨率雷达距离-方位图像。图10依次为城市道路、高速公路、夜晚道路与雾、雨、雪天气场景下的光学图像和对应笛卡尔坐标系下雷达距离-方位图像的标注示例,其中黄色、蓝色、粉色和红色矩

形框分别代表行人、汽车、货车和公交车,数据集中骑行者和卡车类别并未出现在示例图中。经处理,本文数据集共包含7000张图片,并以9:1的比例划分训练集和验证集。

本文实验环境:操作系统,Windows 11;CPU, Intel(R) Core(TM) i7-12700KF@3.6GHz, 32 GB运行内存;GPU, NVIDIA GeForce RTX 3070Ti, 8 GB显存。深度学习框架Pytorch1.10.0, 程序语言Python 3.7.0, 训练配置如表1所示。

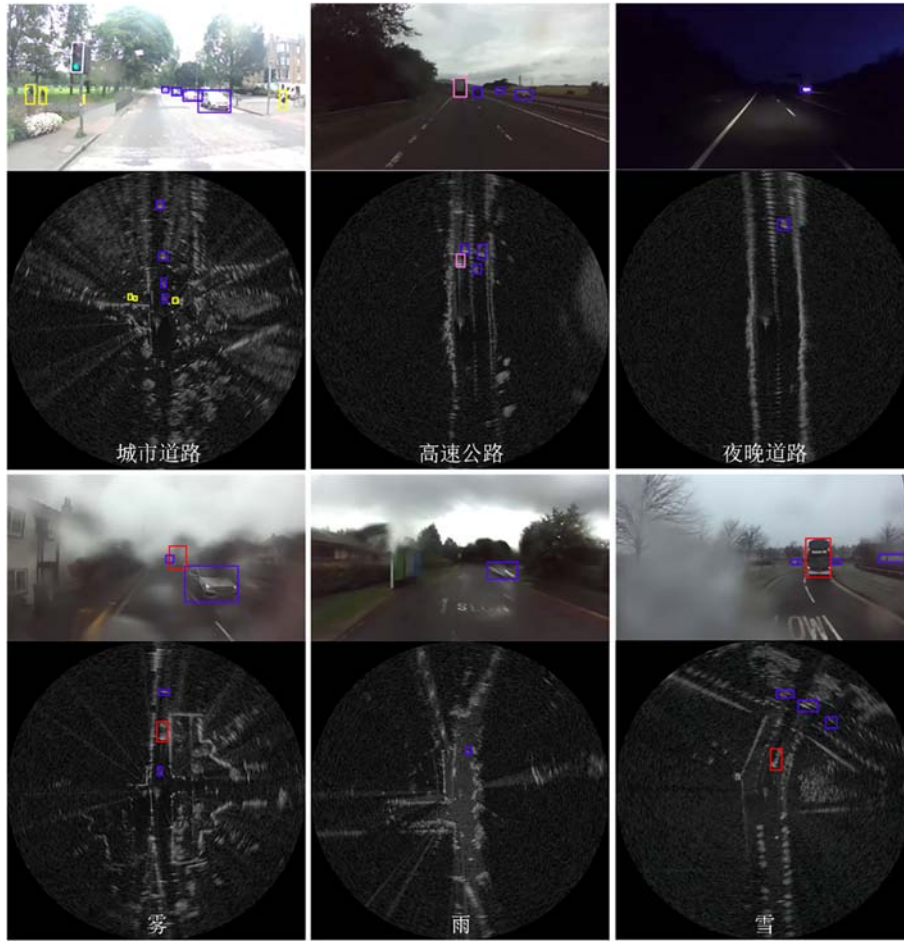


图10 RADIATE数据集示例

表1 模型训练配置

参数名称	参数值
输入图像尺寸	640×640
批处理大小	16
优化器	SGD
初始学习率	0.01
权重衰减系数	0.005
动量	0.937
训练轮次	300

3.2 实验数据处理结果与分析

在本文目标检测任务中,采用准确率(Precision)、召回率(Recall)和平均精度值(mean Average Precision, mAP)作为评估模型性能指标。准确率衡量目标检测模型的精确性,而召回率衡量模型的全面性,二者相互制约。通过绘制准确率-召回率(P-R)曲线,以曲线与 x 轴和 y 轴所围成的面

积作为平均精度(Average Precision, AP)的值, mAP 则表示所有类别 AP 的均值,用于综合评估目标检测模型的整体性能。 $mAP@0.5$ 表示在 IoU 阈值设置为 0.5 时各类别 AP 的平均值, $mAP@0.5:0.95$ 表示在不同 IoU 阈值下 AP 的平均值。各指标的计算公式分别为

$$Precision = \frac{TP}{TP + FP} \times 100\% \quad (15)$$

$$Recall = \frac{TP}{TP + FN} \times 100\% \quad (16)$$

$$AP = \int_0^1 P(R) dR \quad (17)$$

$$mAP = \frac{\sum_{i=1}^k AP(i)}{k} \quad (18)$$

其中 TP (True Positive) 代表被模型预测为正类的正样本, FP (False Positive) 代表被模型预测为正类的负样本, FN (False Negative) 代表被模型预测为

负类的负样本, k 代表目标检测的种类数量, 本文中 k 值为 6。衡量检测精度的同时, 采用模型大小 MB 和浮点数计算量 FLOPs 综合衡量网络模型的性能。通过考虑模型大小, 可以评估模型在存储和传输方面开销, 较小模型通常意味着更高的效率; 而浮点数计算量则可以反映模型在执行推断或训练时所需的计算资源, 较低 FLOPs 值通常意味着较低的计算成本和更高的推断速度。因此, 综合考虑模型大小和 FLOPs 可以提供更全面的性能评估, 帮助选择适合特定需求的网络模型。

为保证实验可控性, 所有方法均未搭载预训练权重。按照 2.1~2.3 节顺序依次对网络添加改进, 分析表 2 改进过程中消融实验结果可知: 相较于原始 YOLOv5s 网络, 平均精度值 mAP@0.5 和 mAP@0.5:0.95 分别提升了 4.4% 和 3.4%。分析表 3 网络分类准确度结果可知: 改进后, 针对小目标

如骑行者与行人两类, 检测与分类性能提升明显, 平均精度值 mAP@0.5 分别提升 18.3% 和 7.5%。同时, 模型大小与计算量的增加均在合理范围内, FPS 满足实时检测需求。在 Concat_att 模块中融合多种不同注意力机制, 实验结果如表 4 所示。本文方法性能的提升在于通过改进解耦头, 轻量化的同时解耦处理检测与分类过程, 更准确地预测不同尺度的特征图; 而设计的 Concat_att 模块, 融合 CBAM 注意力机制, 更加关注特征图中易丢失的小目标细节信息; 并通过改进 EIou 损失, 目标定位更精确, 提升了网络收敛速度和鲁棒性, 显著增强针对小目标的检测与分类性能。图 11(a)、11(b) 更直观地展示了平均精度的提升; 图 11(c) 中, 改进逐步优化使模型的准确率-召回率曲线趋向右上方, 与 x 轴、 y 轴围成的面积增大, 模型整体性能更好。

表 2 改进过程中消融实验结果

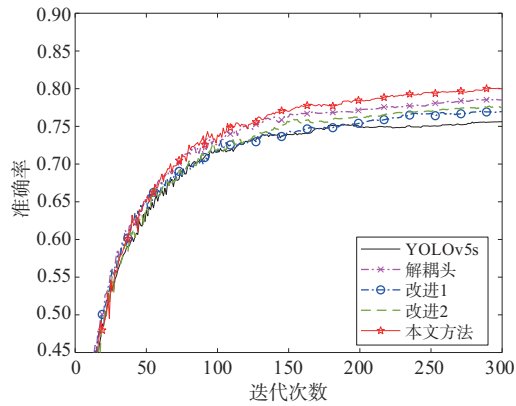
模型	DH	GhostDH	Concat_att	EIoU	模型大小/MB	计算量/GFLOPs	mAP@0.5/%	mAP@0.5:0.95/%
YOLOv5s					13.8	15.8	75.9	43.3
解耦头	√				28.0	56.2	78.5	45.7
改进 1		√			17.7	26.6	77.0	44.3
改进 2		√	√		17.9	26.7	77.6	45.1
本文方法		√	√	√	17.9	26.7	80.3	46.7

表 3 网络分类准确度结果

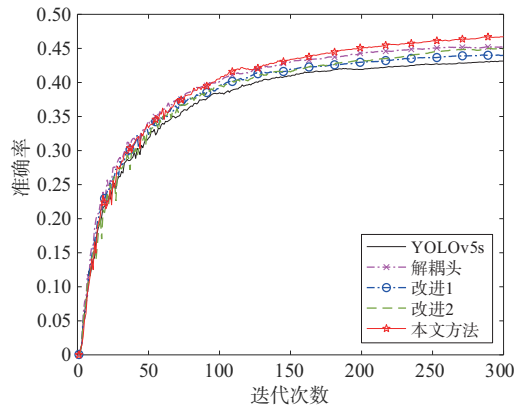
模型	汽车	货车	卡车	公交车	骑行者	行人	FPS(帧·秒 ⁻¹)
YOLOv5s	88.6	91.4	91.9	96.8	50.7	36.0	111
解耦头	88.5	91.6	92.4	97.1	58.7	42.6	87
改进 1	88.5	91.4	91.8	96.4	53.1	40.8	97
改进 2	88.8	91.0	91.1	96.3	58.5	39.9	88
本文方法	89.5	91.6	91.8	96.7	69.0	43.5	89

表 4 注意力机制对比

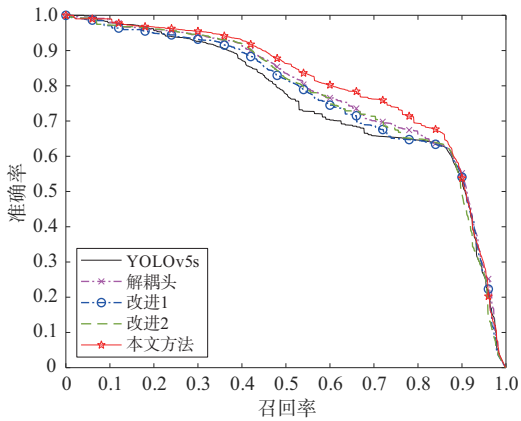
注意力机制	精确率 P/%	召回率 R/%	mAP@0.5/%	mAP@0.5:0.95/%
YOLOv5s	83.3	72.7	75.9	43.3
+SE	84.7	72.8	76.0	44.0
+SimAM	84.2	73.6	76.0	42.8
+CA	84.7	73.7	76.9	43.9
+ECA	84.2	71.9	75.4	43.6
+CBAM	84.3	74.2	77.0	44.9



(a) mAP@0.5



(b) mAP@0.5:0.95



(c) 准确率-召回率曲线

图 11 网络改进前后评价指标对比

可视化检测与分类结果如图 12 所示, 每行对应一张送入网络的雷达距离-方位图像。其中蓝色箭头和矩形框代表局部细节信息放大, 红色箭头代表漏检。通过分析原始网络与本文方法的检测与分类结果, 可以明显发现: 1) 预测框更加贴合实际目标, 包含更少冗余信息; 2) 相比原始网络, 本

文方法出现错检、漏检的概率降低, 且置信度得到一定的提升。

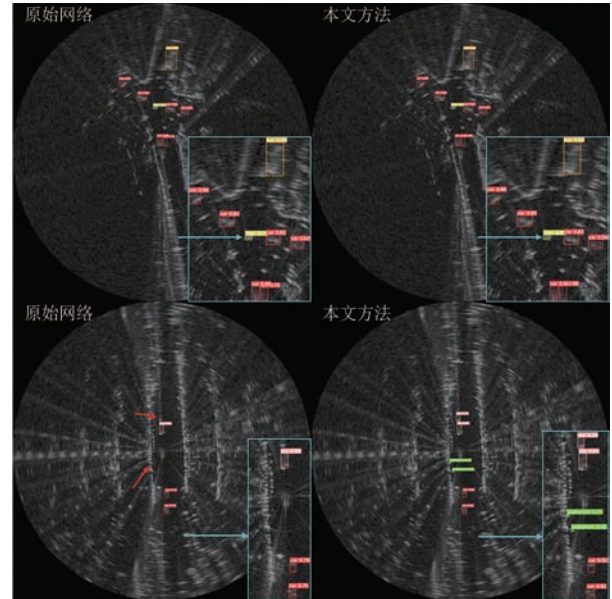


图 12 两幅不同图像可视化检测与分类结果

在本文方法的基础之上, 进一步对该模型进行剪枝处理, 分别设置剪枝加速比例为 1.5X、1.75X、2.0X 对本文方法进行剪枝处理, 实验结果如表 5 所示。当加速比例为 1.75X 时, 模型存储大小仅为原始 YOLOv5s 的 61.6%, 仍达到了比原始 YOLOv5s 更高的网络精度。分析表 6 与不同算法实验结果对比, 综合模型大小和精度方面考量, 当剪枝加速比例为 1.5X 时, 模型大小缩减至原始 YOLOv5s 网络的 76.8%, 且平均精度值 mAP@0.5 与 mAP@0.5:0.95 各有 2.7% 和 2.8% 的提升。本文所提基于轻量化卷积神经网络车载雷达图像目标识别方法具有优势。

表 5 模型剪枝实验结果

网络模型	模型大小/ MB	计算量/ GFLOPs	mAP@ 0.5/%	mAP@0.5: 0.95/%
YOLOv5s	13.8	15.8	75.9	43.3
本文方法	17.9	26.7	80.3	46.7
本文方法 1.5X 剪枝	10.6	17.8	78.6	46.1
本文方法 1.75X 剪枝	8.5	15.3	76.8	44.0
本文方法 2.0X 剪枝	6.9	13.4	70.9	38.3

表6 不同算法实验结果对比

网络模型	模型大小/ MB	计算量/ GFLOPs	mAP@ 0.5/%	mAP@0.5: 0.95/%	FPS (帧·秒 ⁻¹)
YOLOv3-tiny	16.7	12.9	57.7	26.7	142
YOLOv5-n	3.7	4.2	68.3	35.9	113
YOLOv5-s	13.8	15.8	75.9	43.3	111
YOLOv8-n	5.9	8.7	65.2	40.6	116
YOLOv8-s	21.4	28.6	71.0	45.1	95
本文方法	17.9	26.7	80.3	46.7	89
本文方法 1.5X剪枝	10.6	17.8	78.6	46.1	93

4 结束语

本文基于YOLOv5s卷积神经网络,针对雷达图像目标识别任务中的目标尺寸、空间分布特征及在资源有限平台端侧部署问题,提出一种轻量化改进网络:采用EIoU损失函数,设计注意力模块Concat_att,并结合Ghost卷积轻量化设计解耦头,提升识别精度并加速网络收敛,提升模型的泛化能力与鲁棒性,并采用模型剪枝进一步压缩模型存储空间和计算量。当模型大小缩减至原始YOLOv5s网络的76.8%时,较原始网络的mAP@0.5与mAP@0.5:0.95仍有2.7%和2.8%的提升,同时,检测速度可达93帧/秒,满足实时检测要求。综上所述,本文方法适用于雷达图像目标识别任务。

参考文献:

[1] MOZAFFARI S, AL-JARRAH O Y, DIANATI M, et al. Deep Learning-Based Vehicle Behavior Prediction for Autonomous Driving Applications: A Review[J]. IEEE Trans on Intelligent Transportation Systems, 2022, 23(1):33-47.

[2] 段宗明,刘明,吴博文,等.应用于智能驾驶的77 GHz毫米波汽车雷达收发机芯片[J].雷达科学与技术, 2021, 19(2):130-136.

[3] 施端阳,林强,胡冰,等.深度学习在雷达目标检测中的应用综述[J].雷达科学与技术, 2022, 20(6):589-605.

[4] SHEKAR A K, GOU L, REN L, et al. Label-Free Robustness Estimation of Object Detection CNNs for Autonomous Driving Applications [J]. International Journal of Computer Vision, 2021(4):1185-1201.

[5] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Wash-

ington, DC, USA:IEEE, 2014:580-587.

- [6] GIRSHICK R. Fast R-CNN[C]// the IEEE International Conference on Computer Vision, Santiago, Chile: IEEE, 2015:1440-1448.
- [7] REN Shaoqing, HE Kaiming, GIRSHICK R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2017, 39(6):1137-1149.
- [8] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition [J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2015,37(9):1904-1916.
- [9] LIU Wei, ANGUELOV D, ERHAN D, et al. SSD: Single Shot MultiBox Detector [C]// European Conference on Computer Vision, Amsterdam, The Netherlands:Springer, 2016:21-37.
- [10] REDMON J, DIVVALA S, GIRSHICK R, et al. You Only Look Once: Unified, Real-Time Object Detection[C]// the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA:IEEE, 2016:779-788.
- [11] REDMON J, FARHADI A. YOLO9000: Better, Faster, Stronger[C]// the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA:IEEE, 2017: 7263-7271.
- [12] REDMO J, FARHADI A. Yolov3: An Incremental Improvement [J] . arXiv, 2018. <https://arxiv.org/abs/1804.02767>.
- [13] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. Yolov4: Optimal Speed and Accuracy of Object Detection [J]. arXiv, 2020. <https://arxiv.org/abs/2004.10934>.
- [14] SONG Yunyun, XIE Zhengyu, WANG Xinwei, et al. MS-YOLO: Object Detection Based on YOLOv5 Optimized Fusion Millimeter-Wave Radar and Machine Vision [J]. IEEE Sensors Journal, 2022, 22(15):15435-15447.
- [15] XU Xiaowo, ZHANG Xiaoling, ZHANG Tianwen. Lite-YOLOv5: A Lightweight Deep Learning Detector for on-Board Ship Detection in Large-Scene Sentinel-1 SAR Images[J]. Remote Sensing, 2022, 14(4):1018.
- [16] LUO Jiqing, FANG Husheng, SHAO Faming, et al. Multi-Scale Traffic Vehicle Detection Based on Faster R-CNN with NAS Optimization and Feature Enrichment [J]. Defence Technology, 2021, 17(4):1542-1554.
- [17] LIN Mingbao, ZHANG Yuxin, LI Yuchao, et al. 1xN Pattern for Pruning Convolutional Neural Networks [J]. IEEE Trans on Pattern Analysis and (下转第100页)