

改进 YOLOv4 的铁路沿线遥感影像地物检测方法

王阳萍^{1,2}, 韩淑梅^{1*}, 杨景玉^{1,2}, 党建武^{1,2}, 张占平¹

1. 兰州交通大学电子与信息工程学院, 甘肃 兰州 730070

2. 甘肃省人工智能与图形图像处理工程研究中心, 甘肃 兰州 730070

摘要 近年来, 高分遥感影像技术的快速发展为铁路沿线地物检测提供了一种重要技术手段。基于回归的一阶段目标检测方法 YOLOv4 具有检测精度高、速度快等优点, 但用于遥感影像检测时仍然存在部分细节特征信息丢失导致的小目标漏检, 以及进行大面积地物检测时效率低的问题。为此, 提出改进 YOLOv4 网络模型对遥感影像铁路沿线地物进行检测。首先, 设计由卷积、批量归一化和 Mish 激活函数组成的 CBM (convolution batch normalization mish) 模块, 并采用 DCBM (double CBM) 模块作为密集连接网络 (DenseNet) 的传输层用于 YOLOv4 网络特征提取以实现地物特征传递和信息重用, 增强小目标地物的检测能力, 降低漏检率; 然后针对 YOLOv4 在大面积检测时效率不高和模型参数空间较大的缺陷, 将压缩激励 SE (squeeze excitation) 通道注意机制用于骨干网中跨阶段局部单元 (cross stage partial, CSP) 的每个残差单元之后, 减少 SE 注意模块的重复调用次数, 使其能够在提高网络性能的同时降低模型参数量从而提高检测效率; 最后, 针对长条形状的铁路目标提取困难问题, 在网络结果输出之前引入改进的通道空间注意力机制 ICBA (improved convolutional block attention module) 保留原始特征信息, 解决铁路目标特征提取能力差的问题, 提高铁路中大尺度目标的检测效率。为验证所提方法的有效性, 选取 2 048 张分辨率为 1 920×1 080 的某段铁路沿线遥感影像地物样本数据, 将其中的铁路、房屋、楼宇建筑、农田和水池作为检测目标进行实验, 并与当前流行的目标检测方法进行对比。结果表明, 改进方法不仅增强了对小目标地物的检测能力, 提高了地物检测精度和速度, 而且提高了大尺度目标的检测效率。与 YOLOv4 算法相比, mAP 提高了 2.11%, 准确率提高了 2.93%, 召回率提高了 3.79%, 模型大小减少了 8.53%。所提方法为当前应用高速铁路沿线遥感影像地物快速精准检测提供了有效方法。

关键词 地物检测; 铁路沿线; 遥感影像; YOLOv4; 注意力机制

中图分类号: TP79 **文献标识码:** A **DOI:** 10.3964/j.issn.1000-0593(2022)10-3275-08

引言

我国高速铁路的大规模建设方便了人们出行的同时在铁路运维方面提出了新的挑战, 及时检测铁路沿线的地物目标已成为当前铁路维护人员需要解决的重要问题之一。而快速发展的高分辨率遥感技术^[1]以其覆盖范围大、地物纹理信息丰富、成像光谱波段多、成本低等优点为及时检测铁路沿线地物提供了一种快速、便捷、高效的解决方法。近年来基于卷积神经网络 CNN (convolutional neural networks) 的算法以

其鲁棒性强、检测精度高, 且能满足复杂背景下多尺度目标需求等特点, 被广泛应用于遥感影像目标检测^[2]。Wang 等^[3]提出用于大规模遥感影像中多类目标检测 FM-SSD 方法, 该方法能够在多个尺度上获得丰富的上下文信息, 但对同一特征图中的小目标识别效果较差且计算量大。Xu^[4]等针对遥感目标的特点, 提出一种轻量级特征增强网络 FE-YOLO, 该方法的检测速度大幅度提高, 但其特征提取能力弱, 无法达到检测精度要求。Li 等^[5]将 Faster R-CNN, YOLOv3 和 SSD 用于遥感影像农业温室地物检测研究中, 并从检测精度和速度方面进行对比, 得到 YOLO 算法在检测速度方面

收稿日期: 2021-08-23, 修订日期: 2022-03-01

基金项目: 国家自然科学基金项目 (61763025, 62067006), 甘肃省教育科技创新项目 (2021jybgbs-05), 甘肃省高校产业支撑计划项目 (2020C-19), 国家市场监督管理总局科技计划项目 (2019MK150), 兰州市科技计划项目 (2019-4-49, 2020-1-7), 天津大学-兰州交通大学自主创新基金合作项目, 甘肃省高等学校青年博士基金项目 (2021QB-053) 资助

作者简介: 王阳萍, 1973 年生, 兰州交通大学电子与信息工程学院教授 e-mail: 1328396793@qq.com

* 通讯作者 e-mail: 0219653@stu.lzjtu.edu.cn

优于其他两者。但 YOLOv3 对局部小块温室大棚检测效果较差。2020 年, Alexey Bochkovskiy 等^[6]基于 CSPNet 骨干网络、Mosaic 数据增强、Mish 激活、SPP 和 GIOU 损失提出 YOLOv4, 能够得到更快更精确的检测结果。YOLOv4 因其效率高在普通图像检测中得到更广泛的应用, 但当其用于遥感影像时存在小目标漏检和大尺度检测效率低的问题。为了更好地对铁路沿线地物进行检测, 我们对 YOLOv4 进一步改进, 采用 DCBM 模块改进 DenseNet 用于 YOLOv4 特征提取以实现特征重用, 增强小目标特征提取能力。同时在骨干网中 CSP 单元的每个残差单元之后连接一个 SE 模块, 使其在提高地物检测精度的同时降低模型参数量。为提高网络对铁路目标的检测效率, 提出一种改进的通道空间注意力模块 ICBAM, 并在网络预测结果输出之前引入该模块, 进一步提高网络性能。实验结果表明, 所提方法提高了地物目标的检测精度, 有效降低了模型大小和参数量, 具有实际的应用价值。

1 YOLOv4 目标检测算法

YOLOv4 网络将输入图像调整为固定大小, 并划分为 $S \times S$ 的网格单元作为输入, 基于回归获得边界框的位置及其所属类别, 从而实现端到端的目标检测。其在输入端使用 Mosaic 数据增强方法大大丰富检测数据集, 提高网络的鲁棒性。在骨干网方面, 采用 CSPDarknet53 网络结构增强 CNN 的学习能力, 使得在轻量化的同时保持准确性, 减少了计算量, 降低了内存成本。在 Neck 方面, 采用特征金字塔 SPP

模块的最大池化方式进行不同尺度的特征融合并采用 FPN 和 PAN 相结合的方式对不同主干层中不同的检测层进行特征聚合。在损失函数方面, 采用 CIOU 损失和 DIOU 损失相结合的方法同时考虑边界框宽高比的尺度信息、重叠面积和中心点距离信息, 更好地实现目标边界框回归, 得到更加准确的预测结果。

与 YOLOv3 算法相比, YOLOv4 算法在目前的公开数据集上 mAP 和 FPS 都有所提高, 在检测速度、模型性能、训练时间和硬件要求方面优于其他算法, 但仍然存在边界框定位不够精确、难以检测小目标物体、灵活性与检测速度等方面的不足。其主干网 CSPDarkNet53 主要由 CSP 单元组成, 而 CSP 单元包含了一些 ResNet 和 CBM 模块, 但这些 ResNet 包含了大量的参数计算, 负责 YOLOv4 网络的主要特征提取。将其用于铁路沿线遥感影像地物目标检测中, 存在部分地物目标信息丢失及模型参数空间大的问题。因此, 本工作针对铁路沿线遥感影像地物检测对其进行改进。

2 基于改进 DenseNet 网络的特征提取

遥感影像铁路沿线的地物目标具有小而密集分布的特点, 其在 YOLOv4 模型中经过进一步压缩后, 小目标地物容易出现细节特征信息丢失的现象, 产生小目标漏检问题。DenseNet 考虑到深层次的网络导致梯度消失且残差网络不能充分利用网络提取的特征问题, 通过连接所有特征图并进行通道合并, 实现多特征图的融合和特征重用, 从而提高准确度。使用 DenseNet 能使网络梯度的反向传播得到增强, 可

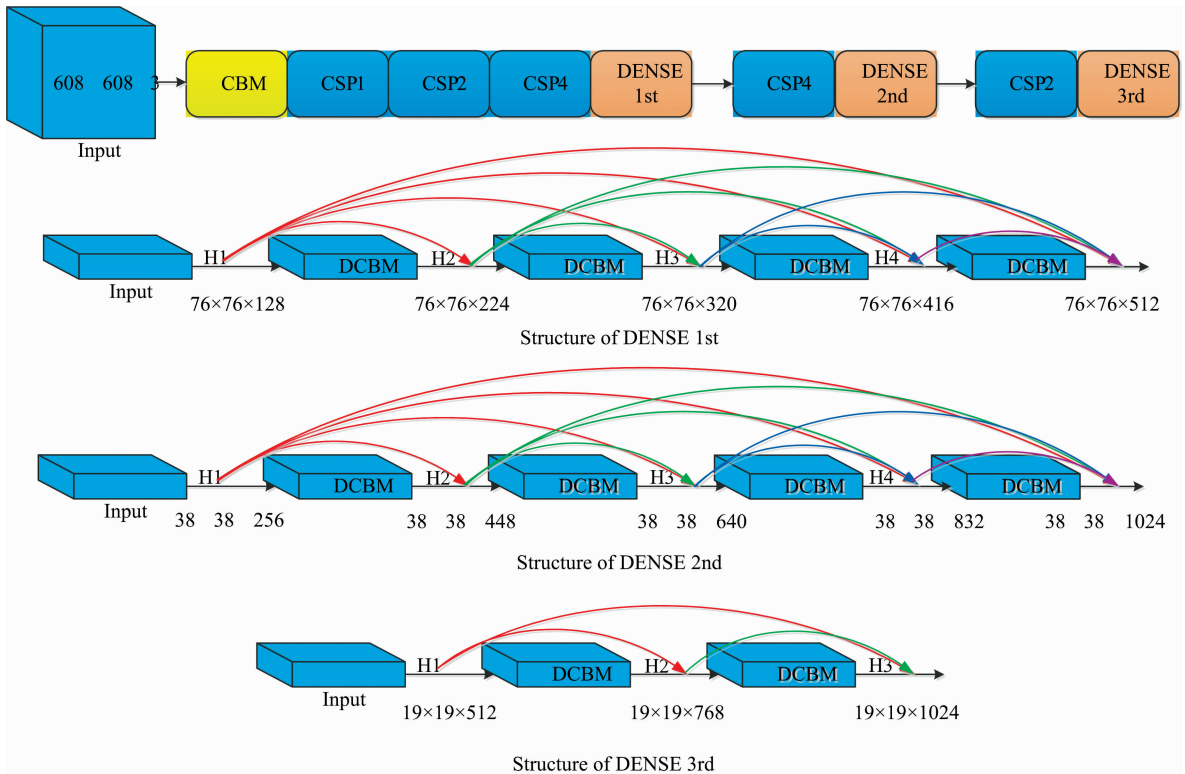


图 1 改进的 DenseNet 网络模型
Fig. 1 Improved DenseNet network model

以更好地利用已提取的地物特征信息，提高层间信息的透光率，间接降低模型参数。其每一层都可以接收到前面层的所有特征图， X_i 与前面的特征图关系可由式(1)表示

$$X_i = H_i[X_0, X_1, \dots, X_{i-1}] \quad (1)$$

本文设计出由卷积、批量归一化 BN 和 Mish 激活组成的 CBM 模块，而两个 CBM 模块级联成一个双 CBM(double CBM)模块，本文使用 DCBM 模块作为 DenseNet 的传输层，由 Conv(1×1×32)-BN-Mish-Conv(3×3×64)-BN-Mish 和 Conv(1×1×64)-BN-Mish-Conv(3×3×128)-BN-Mish 组成。考虑到使用大量 DenseNet 网络会增加网络层数从而导致特征冗余、降低检测速度，因此为前两个 CSP 模块设置了 4 层，为第三个 CSP 模块设置了 2 层。改进 DenseNet 的网络结构如图 1 所示。

其中，模块 DENSE 1st 中相邻层特征图增量为 96，模块 DENSE 2nd 中各相邻层特征图增量为 192，模块 DENSE 3rd 中相邻层特征图增量为 256。本模块的设计不但能减少网络对残差单元的依赖，而且使用改进的 DenseNet 网络作为特征提取网络的部分低分辨率层(CSP 单元中的部分残差单元层)，能有效解决深层网络传递时的特征消失问题，增加各层之间的特征传递，防止地物目标的细节特征信息丢失，一定程度上能够减少网络的参数量。

为保证网络其他层与 DenseNet 模块参数的正常传递，设置 DenseNet 模块输入输出特征图与原始 ResNet 模块的输入输出尺度相同。将改进 DenseNet 用于 YOLOv4 特征提取，结构如图 2 所示。

3 改进注意力机制

注意力机制有着即插即用的特点，被广泛用于目标检测领域^[7]。其参数量本身不高，但在卷积神经网络目标检测中，往往被多次调用，这将对网络造成一定负担^[8]，如何优化其结构，减少模块调用次数，对目标予以聚焦和关注，更快更准地获取注意力信息，将其用于铁路沿线地物目标检测有着重要作用。本工作提出一种改进的注意力机制用于 YOLOv4 进行地物检测。首先，融合通道注意 SE 模块和跨阶段局部单元 CSP 提出 SE-CSP 模块，降低模型大小，解决因注意力模块多次执行导致参数量大的问题，再用 ICBAM 注意模块聚焦原始特征信息，解决将 CBAM 直接用于 YOLOv4 导致对铁路目标提取能力差的问题。

3.1 融合通道注意模块跨阶段局部单元的 SE-CSP 模块

将 YOLOv4 模型用于铁路沿线遥感影像地物检测过程中，存在特征提取网络性能较差和模型参数量较大问题，无法提取到地物特征图通道之间的相关信息，从而降低了残差块对地物特征的提取能力。为增强通道特征之间的相关性以提高网络的特征提取能力，将 SE 模块嵌入到 ResNet 网络中，通过 SE 模块学习不同通道之间的相关性，并且筛选出针对通道的注意力，以达到较好的特征提取效果。

由于 YOLOv4 的骨干网 CSPDarknet53 每训练一次都需要执行 23 次残差块，且每次都需要通过使用 SE 模块显式地建立通道间的相互依赖关系，自适应地重新校准通道特征响

应，这将增加网络的计算量，同时影响网络的检测速度。为此，提出一种 SE-CSP 模块，通过分离残差块中的 Res Unit 与 SE 模块，并将 SE 模块拼接到 Res Unit 之后，保证 SE 模块建立通道依赖关系的同时将 SE 模块的执行次数减少至 5 次，从而保证在不降低网络提取特征能力的同时减少网络的计算量和 SE 模块的重复调用次数。本文提出的 SE-CSP 模块如图 3 所示。

| Type | Filters | Size | Output |
|-------------------|---------------|-------|---------|
| Convolutional | 32 | 3×3 | 608×608 |
| Convolutional | 64 | 3×3/2 | 304×304 |
| CrossStagePartial | | | |
| 1× | Convolutional | 32 | 1×1 |
| | Convolutional | 64 | 3×3 |
| | Residual | | 304×304 |
| Convolutional | 128 | 3×3/2 | 152×152 |
| CrossStagePartial | | | |
| 2× | Convolutional | 64 | 1×1 |
| | Convolutional | 64 | 3×3 |
| | Residual | | 152×152 |
| Convolutional | 256 | 3×3/2 | 76×76 |
| CrossStagePartial | | | |
| 4× | Convolutional | 128 | 1×1 |
| | Convolutional | 128 | 3×3 |
| | Residual | | 76×76 |
| 4× | Convolutional | 32 | 1×1 |
| | Convolutional | 64 | 3×3 |
| | DenseNet | | 76×76 |
| Convolutional | 512 | 3×3/2 | 38×38 |
| CrossStagePartial | | | |
| 4× | Convolutional | 256 | 1×1 |
| | Convolutional | 256 | 3×3 |
| | Residual | | 38×38 |
| 4× | Convolutional | 64 | 1×1 |
| | Convolutional | 128 | 3×3 |
| | DenseNet | | 38×38 |
| Convolutional | 1024 | 3×3/2 | 19×19 |
| CrossStagePartial | | | |
| 2× | Convolutional | 512 | 1×1 |
| | Convolutional | 512 | 3×3 |
| | Residual | | 19×19 |
| 2× | Convolutional | 128 | 1×1 |
| | Convolutional | 286 | 3×3 |
| | DenseNet | | 19×19 |

图 2 改进 DenseNet 的 YOLOv4 特征提取网络

Fig. 2 Improved DenseNet for YOLOv4 feature extraction network

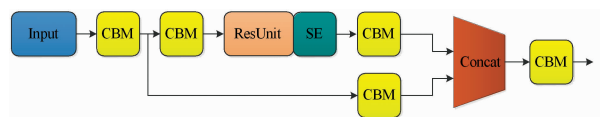


图 3 本文提出的 SE-CSP 模块

Fig. 3 SE-CSP module proposed in this article

3.2 改进 CBAM 注意力机制

铁路路基具有路线长、与周边环境接触面广、排列不规则等特点，导致对其提取困难。文献[9]为解决大尺度目标检测困难问题，在网络的输出端直接使用 CBAM 注意力模块，该方法在增强大尺度目标特征的同时削弱了对中小尺度

目标的提取能力。为此, 本工作提出一种改进的 CBAM 注意力机制 ICBAM。ICBAM 能够在保留原特征信息的同时增强对铁路目标的特征提取。所提 ICBAM 由 CBM 模块、通道注意模块和空间注意模块三部分组成, 其结构如 4 所示。

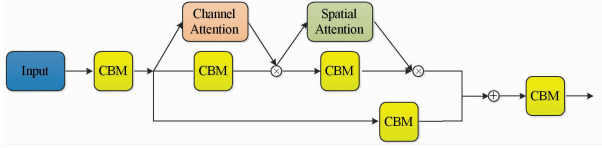


图 4 ICBAM 模块

Fig. 4 ICBAM module

给定输入特征图 $F \in R^{C \times H \times W}$, CBAM 依次推导出 1D 通道注意图 $M_c \in R^{C \times 1 \times 1}$ 和 2D 空间注意图 $M_s \in R^{1 \times H \times W}$ 。由此, ICBAM 注意力模块对特征图的整体处理过程为

$$F' = \delta(\delta(F)) \otimes M_c(\delta(F)) \quad (2)$$

$$F'' = \delta(F') \otimes M_s(F') \quad (3)$$

$$F''' = \delta(\delta(\delta(F))) F \oplus F'' \quad (4)$$

其中, $M_c(F)$ 和 $M_s(F')$ 分别表示在通道注意和空间注意上获取的权重, \otimes 表示逐元素乘法, \oplus 表示连接操作, δ 表示对输入特征图进行 CBM 操作, F' 是 F 在通道上的加权结果, F'' 是 F' 在空间上的加权结果, F''' 为 ICBAM 的最终输出结果。

通道注意 $M_c(F)$ 的具体计算公式为

$$\begin{aligned} M_c(F) &= \sigma(\text{MLP}(\text{AvgPool}(F)) + \text{MLP}(\text{MaxPool}(F))) \\ &= \sigma(W_1(W_0(F_{\text{avg}}^c)) + W_1(W_0(F_{\text{max}}^c))) \end{aligned} \quad (5)$$

式(5)中, “+”表示逐元素求和, “ σ ”表示 sigmoid 激活, MLP 是带有一个隐藏层的多层感知器, 由其组成共享网络。MLP 权重由 W_0 和 W_1 输入共享。其中 $W_0 \in R^{C/r \times C}$, $W_1 \in R^{C \times C/r}$ 。

空间注意 $M_s(F)$ 的具体计算公式为

$$\begin{aligned} M_s(F) &= \sigma(f^{7 \times 7}([\text{AvgPool}(F); \text{MaxPool}(F)])) \\ &= \sigma(f^{7 \times 7}([F_{\text{avg}}^s; F_{\text{max}}^s])) \end{aligned} \quad (6)$$

式(6)中, σ 表示激活函数, $f^{7 \times 7}$ 表示一个 7×7 卷积。

3.3 改进后的 YOLOv4 总体结构

为使 YOLOv4 用于铁路沿线遥感影像地物检测时能满足精度与实时性的要求, 对其做了相应改进。在骨干网方面, 通过改进 DensNet, 增强了特征重用, 减少了小目标地物的细节特征丢失现象。在改进注意力方面包括改进通道注意 SE 模块和改进通道空间注意 CBAM 模块两方面。在 SE 方面, 通过在每个 CSP 单元中的 ResUnit 之后连接一个 SE 模块, 与 CBM 模块共同组成 SE-CSPX 模块, 减少了 CSP 单元中残差块的重复调用次数, 在减少模型大小的同时提高了地物检测精度。在 CBAM 方面, 通过采用 CBM 模块、通道注意和空间注意模块组成 ICBAM 模块, 并在每个尺度预测结果输出之前引入 ICBAM, 在保留原始地物特征信息的同时, 降低了大尺度铁路目标的漏检率。引入改进注意力机制后的 YOLOv4 整体结构模型如图 5 所示。其中, SE-CSPX 模块中的 X 表示该模块重复使用的次数。本改进方法能够在降低模型参数的同时提升模型检测精度和模型训练效率, 具有高效可行性。

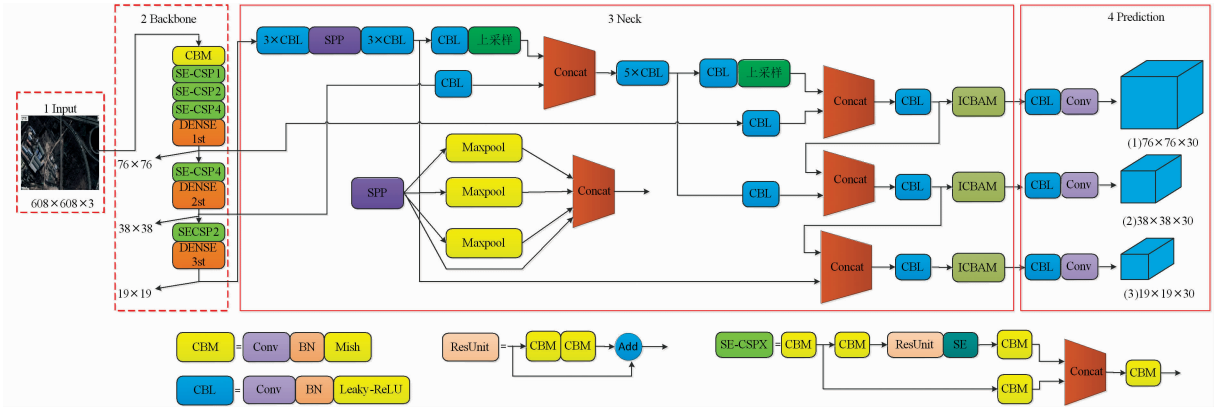


图 5 改进的 YOLOv4 网络模型

Fig. 5 Improved YOLOv4 network model

4 结果与讨论

4.1 数据集

考虑到现有公开数据集无法满足铁路沿线地物检测需求, 本文采集 Worldview I, Worldview II 和 QuickBird 卫星遥感影像数据, 采用 2 048 张铁路沿线数据制作地物检测数据集, 并用数据增强方法实现数据集扩充。数据包括铁路 (Rail)、房屋 (House)、楼宇建筑 (Building)、农田 (Farmland) 和水池 (Pool) 五种地物目标。其中铁路 764 张, 房屋

130 张, 楼宇建筑 220 张, 农田 562 张, 水池 372 张。通过旋转、镜像、改变亮度、高斯滤波、平移和缩放将数据集扩充为原来的 3 倍, 扩充后共包含地物目标图像 6 144 张。使用数据集的 70% 作为训练集, 10% 作为验证集, 20% 作为测试集。

算法运行环境: 操作系统 Windows 10, 平台 Keras, 语言 Python, CPU Intel (R) Xeon (R) 2.40 GHz, 显卡 16 G 的 NVIDIA Quadro P5000, 运行内存 128 GB。

4.2 参数设置

实验中设置批量训练数据样本数量 (Batch_size) 为 8, 图

像大小(Image_size)为 608×608 ，整个数据集训练次数(Epochs)为 100，动量(Momentum)为 0.9，初始学习率(Initial_learning_rate)为 0.001。为使训练模型达到较高精度，使用迁移学习策略在公开数据集 DOTA 上训练生成权重，然后通过共享权重方式在本数据集上进行训练。训练过程分两阶段，第一阶段冻结原始网络前 192 层训练得到模型权重，然后在此基础上解冻网络并调整 Batch_size 为 8，完成第二阶段训练。为防止固定学习率对训练精度的影响，当训练 5 次后模型精度不再发生变化时，使用余弦退火算法改变学习率，以得到最优模型。

4.3 消融实验结果

为分别验证改进 DenseNet 模块、SE-CSP 模块、ICBAM 模块对检测效果的影响，进行消融实验，部分实验结果如图 6 所示。其中，图 6(a)为遥感影像原图，图 6(b)为 YOLOv4 算法检测结果，图 6(c)为改进 DenseNet 用于原 YOLOv4 特征提取检测结果，图 6(d)为改进 DenseNet 和 SE 模块用于原 YOLOv4 检测结果，图 6(e)为改进 DenseNet、SE 模块和 CBAM 用于原 YOLOv4 检测结果。检测结果中红、深蓝、黄、淡蓝和粉色框分别代表铁路、水池、房屋、农田和高楼建筑物的预测框。

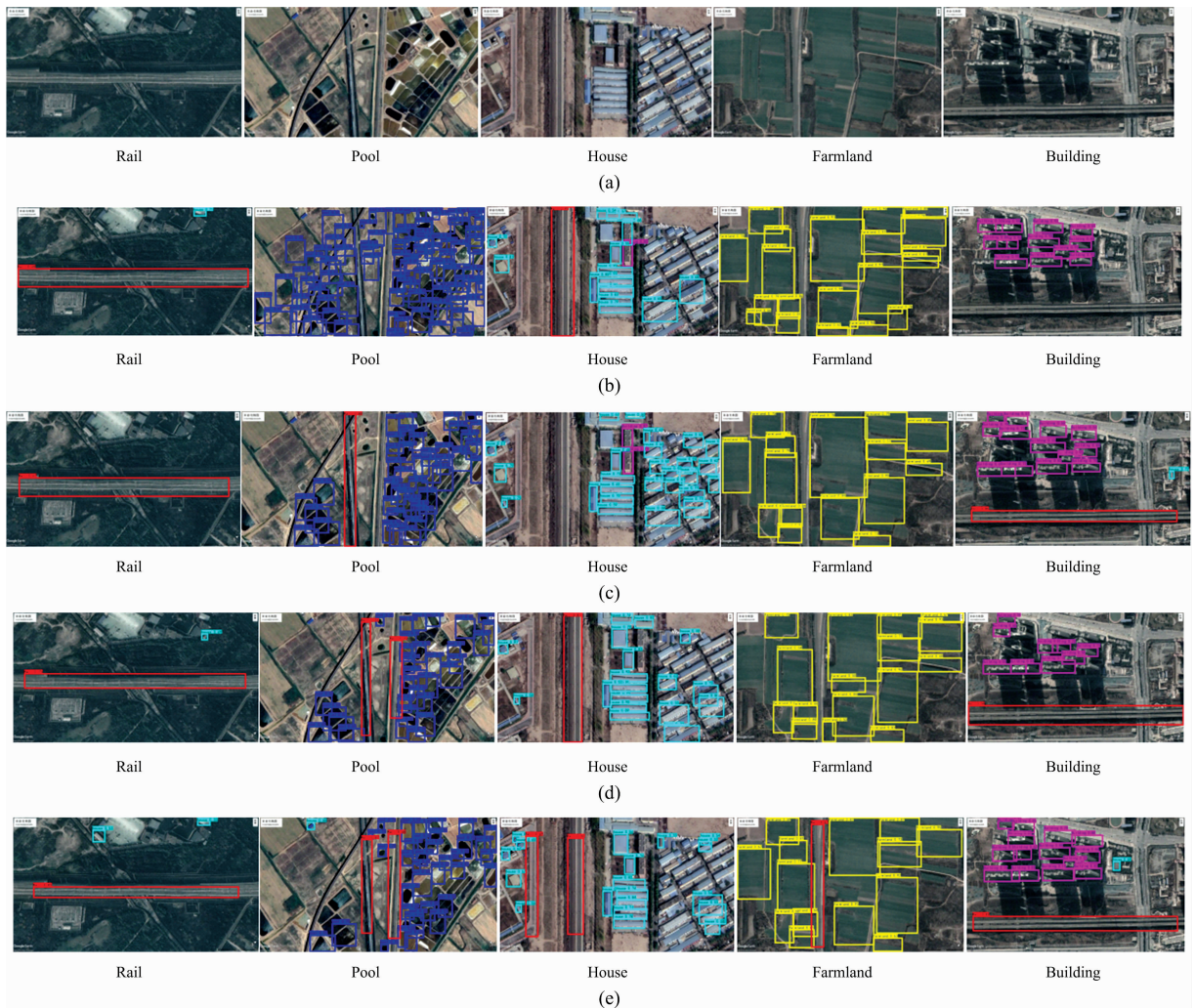


图 6 消融实验测试结果

(a): 遥感影像原图; (b): YOLOv4 算法检测结果; (c): 改进 DenseNet 用于原 YOLOv4 特征提取的检测结果; (d): 改进 DenseNet 和 SE 模块用于原 YOLOv4 检测结果; (e): 改进 DenseNet、SE 模块和 CBAM 模块用于原 YOLOv4 的检测结果

Fig. 6 Test results of ablation experiments

- (a): Original remote sensing images; (b): Detection using YOLOv4;
- (c): Improved DenseNet detection for original YOLOv4 feature extraction;
- (d): Improved DenseNet and SE modules for original YOLOv4 detection;
- (e): Improved DenseNet, SE and CBAM modules for original YOLOv4 detection

可以看出，通过使用改进 DenseNet 用于 YOLOv4 特征提取，对部分图像中铁路的检测精度有所提高，对少部分图

像中房屋和楼宇建筑的漏检率明显降低；通过在骨干网中 CSP 单元的每个 ResUnit 后连接 SE 模块，对大部分图像中

的铁路检测效果明显,通过在输出网络之前引入 ICBAM 注意力机制,对大多数图像中铁路的检测精度进一步提高,对少部分图像中农田楼宇和建筑的漏检率进一步降低。

通过客观评价分析,不同模块的改进过程中检测准确度、召回率、平均精度、模型大小和总时间对比分析结果见表 1。

表 1 消融实验结果分析

Table 1 Analysis of ablation experiments

| Method | Precision/% | Recall/% | mAP/% | Model Size/MB | Times/s |
|------------------------|-------------|----------|-------|---------------|---------|
| YOLOv4 | 80.66 | 79.02 | 81.65 | 250.43 | 0.058 |
| DenseNet | 81.35 | 80.53 | 82.37 | 233.18 | 0.049 |
| DenseNet and SE | 84.27 | 83.12 | 83.76 | 231.65 | 0.052 |
| DenseNet, SE and ICBAM | 83.59 | 82.81 | 83.36 | 229.06 | 0.052 |

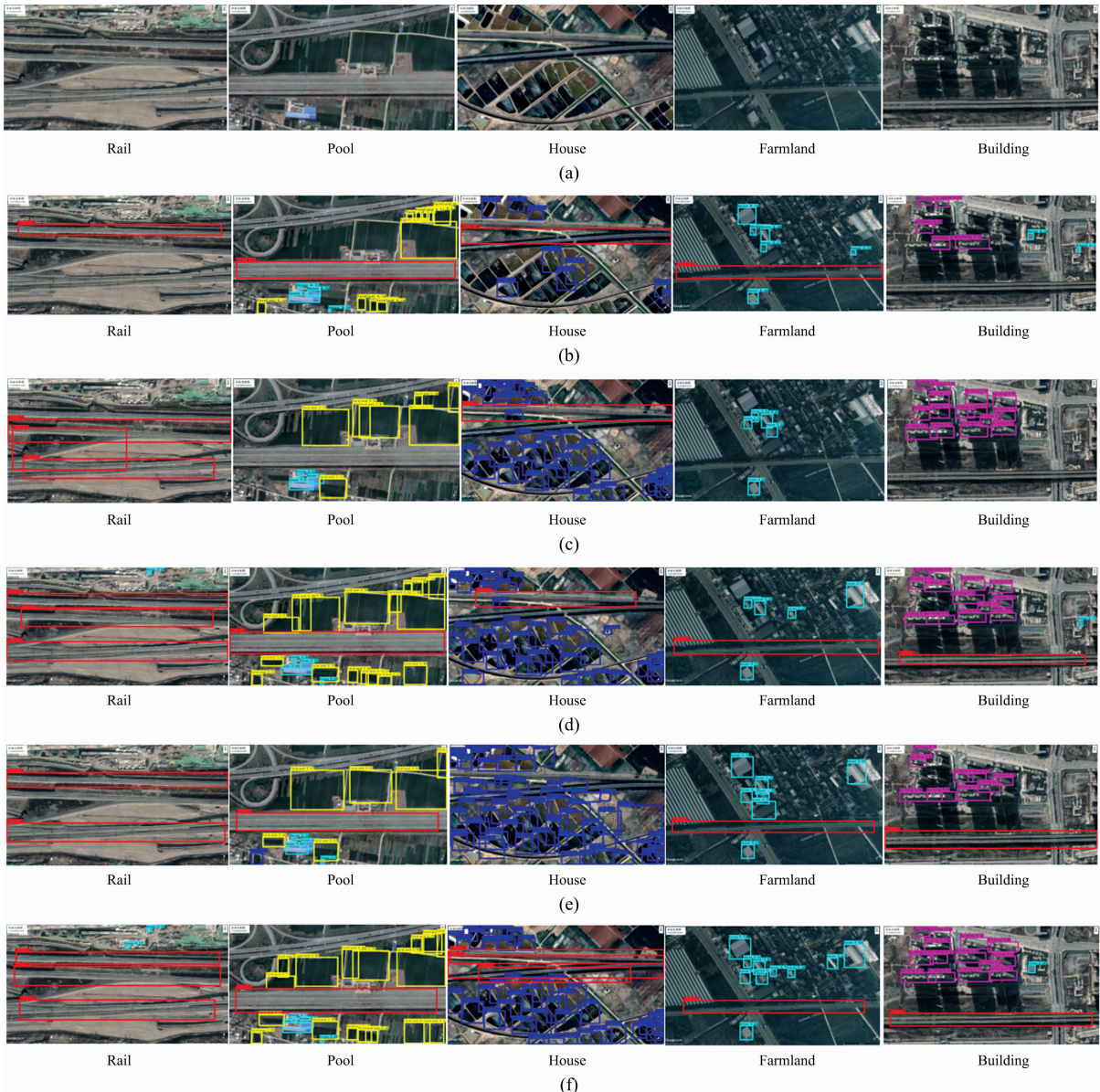


图 7 不同目标检测算法与改进算法实验结果

(a): 遥感影像原图; (b): YOLOv3 算法检测结果; (c): YOLOv3-UAV 算法检测结果^[10];
(d): YOLOv3-Ship 算法检测结果^[11]; (e): 原 YOLOv4 算法检测结果; (f): 本算法检测结果

Fig. 7 Experimental results of different target detection algorithms and improved algorithms

(a): Original remote sensing images; (b): Detection using YOLOv3 algorithm; (c): Detection using YOLOv3-UAV^[10];
(d): Detection using YOLOv3-ship^[11]; (e): Detection using original YOLOv4; (f): Detection using the algorithm proposed in this paper

经分析可知,通过使用 DenseNet 改进 YOLOv4 特征提取网络,模型参数量相比于原 YOLOv4 减少了 7.81%,模型大小降低了 6.89%,平均精度 mAP 提高了 0.72%。通过使用 SE-CSP 模块,参数量减少了 9.38%,模型大小进一步降低了 0.66%,mAP 进一步提高了 1.39%。通过在输出网络之前引入 ICBAM 注意力模块,检测速度没有明显降低,但对大尺度铁路目标的检测能力有所增强,而对其他目标精度损失的代价仅为 0.48%,每幅图像的平均检测时间为 0.052 s。结果表明这些改进效果可行,适当组合效果更好。

4.4 多种检测方法对比实验结果

为验证整体改进方法的有效性,将其与 YOLOv3, YOLO-UAV^[10], YOLO-Ship^[11] 以及 YOLOv4 算法进行比较,更直观的部分实验检测结果如图 7 所示。其中图 7(a)为遥感影像原图,图 7(b)为 YOLOv3 算法检测结果、图 7(c)为 YOLOv3-UAV

算法-文献[10]检测结果、图 7(d)为 YOLOv3-Ship 算法-文献[11]检测结果、图 7(e)为 YOLOv4 算法检测结果、图 7(f)为本改进方法的检测结果。更加具体的客观评价分析如表 2 所示。

表 2 表明,改进方法在准确率(Precision)、召回率(Recall)和 F1 得分方面优于其他算法。模型大小较原 YOLOv4 减少了 8.53%,平均检测精度 mAP 相对 YOLOv4 提高了 2.11%。对多目标、模糊目标和小目标的误检率降低。能够达到较高的精度和速度,能够有效检测复杂背景下的遥感影像地物目标。

5 结 论

提出改进 YOLOv4 的铁路沿线地物检测方法,与原 YOLOv4 算法相比,该方法检测精度更高、速度更快。针对小目标地物容易出现细节信息丢失现象产生的漏检问题,使用改进 DenseNet 增强对小目标检测能力;为解决 SE-ResNet 嵌入 YOLOv4 模型导致参数量较大问题,提出一种 SE-CSP 模块,使其在提高检测精度的同时减少模型大小,从而提高检测速度;针对铁路目标特征提取困难问题,提出一种 ICBAM 注意力机制,提高对铁路目标的检测效率。通过实验验证,得出改进方法在检测精度和速度方面都有所提高。但对于铁路目标,数据样本中只标记了水平和垂直方向的铁路。未来,可进一步开展对具有旋转角度的目标检测方法和性能提升的研究工作。

表 2 不同目标检测算法对比

Table 2 Comparison of different target detection algorithms

| Method | Precision /% | Recall /% | mAP /% | F1 | Model Size MB |
|-------------|--------------|-----------|--------|-------|---------------|
| YOLOv3 | 68.52 | 67.60 | 78.52 | 68.06 | 247 |
| YOLOv3-UAV | 76.37 | 73.20 | 80.38 | 74.75 | 247.36 |
| YOLOv3-Ship | 77.62 | 76.80 | 80.86 | 77.21 | 246.27 |
| YOLOv4 | 80.66 | 79.02 | 81.65 | 79.83 | 250.43 |
| 本算法 | 83.59 | 82.81 | 83.76 | 83.20 | 229.06 |

References

- [1] YIN Zi-yao, LI Jun-sheng, FAN Hai-sheng, et al(殷子瑶, 李俊生, 范海生, 等). Spectroscopy and Spectral Analysis(光谱学与光谱分析), 2021, 41(2): 494.
- [2] Konstantinidis D, Argyriou V, Stathaki T, et al. Computer Networks, 2020, 168: 107034.
- [3] Wang P, Sun X, Diao W, et al. IEEE Transactions on Geoscience and Remote Sensing, 2019, 58(5): 3377.
- [4] Xu D, Wu Y. Remote Sensing, 2021, 13(7): 1311.
- [5] Li M, Zhang Z, Lei L, et al. Sensors, 2020, 20(17): 4938.
- [6] Bochkovskiy A, Wang C Y, Liao H Y M. YOLOv4: Optimal Speed and Accuracy of Object Detection, arXiv: 2004. 10934, 2020.
- [7] Yin Fulian, Li Sitong, Ji Meiqi, et al. Applied Intelligence, 2022, 52(1): 19.
- [8] BIAN Ji-long, WANG Hou-bo, LI Jin-feng(边继龙, 王厚博, 李金凤). Journal of Beijing University of Posts and Telecommunications(北京邮电大学学报), 2021, 44(3): 27.
- [9] Li H, Deng L, Yang C, et al. IEEE Access, 2021, 16692: 16706.
- [10] Zhang Z, Liu Y, Liu T, et al. IEEE Geoscience and Remote Sensing Letters, 2019, 17(11): 1884.
- [11] Chen L, Shi W, Deng D. Remote Sensing, 2021, 13(4): 660.

Improved YOLOv4 Remote Sensing Image Detection Method of Ground Objects Along Railway

WANG Yang-ping^{1,2}, HAN Shu-mei^{1*}, YANG Jing-yu^{1,2}, DANG Jian-wu^{1,2}, ZHANG Zhan-ping¹

1. School of Electronic and Information Engineering, Lanzhou Jiaotong University, Lanzhou 730070, China

2. Gansu Engineering Research Center for Artificial Intelligence and Graphic & Image Processing, Lanzhou 730070, China

Abstract In recent years, the rapid development of high-resolution remote sensing technology has provided an effective technical means for detecting ground objects along the railway. The regression-based one-stage target detection method YOLOv4 has the advantages of high detection accuracy and fast speed. However, when it is used for remote sensing image detection, small targets are still missed due to the loss of some detailed feature information, and large-area ground object detection. Due to the problem of low efficiency, this paper improves the YOLOv4 network model to detect ground objects along the railway in remote sensing images. This paper improves the YOLOv4 network model to detect the ground features in remote sensing images along the railway. First, the CBM (Convolution Batch Normalization Mish) module is designed with composing of convolution, batch normalization, and Mish activation, and the DCBM (Double CBM) module is used for the transmission layer of the densely connected network (DenseNet) for the YOLOv4 network feature extraction. It can achieve feature transfer and information reuse and enhance small target feature detection capabilities. Then, to address the defects of YOLOv4 in the inefficiency of large area detection and the large space of model parameters, the SE (Squeeze Excitation) channel attention mechanism is used after each residual cell of Cross Stage Partial (CSP) in the backbone network to reduce the number of repeated calls of the SE attention module. Hence the performance of the network is improved while reducing the number of model parameters and improving detection efficiency. Finally, for the problem of difficult extraction of railroad targets in images, an improved channel space attention mechanism ICBAM (Improved Convolutional Block Attention Module) is introduced before the network result output to retain the original feature information. It can solve the problem of poor feature extraction ability of railroad targets, and improve the detection efficiency of large-scale targets. To verify the effectiveness of the proposed method, 1 676 remote sensing image samples are selected along a particular section of the railway with a resolution of $1\ 920 \times 1\ 080$. Railways, houses, buildings, farmland, and ponds in the data set are selected as targets for inspection, and some current popular target detection methods are compared. The experimental results show that the improved method enhances the detection ability of small targets, improves the accuracy and speed of detection, and improves the detection efficiency of large-scale targets. Compared with the YOLOv4 algorithm, the improved method mAP has increased by 2.11%, accuracy increased by 2.93%, the recall rate has increased by 3.79%, and the model size is reduced by 8.53%. The proposed method also provides an effective method for rapidly and accurately detecting ground objects in remote sensing images along the high-speed railway.

Keywords Ground target detection; Along with the railway; Remote sensing image; YOLOv4; Attention mechanism

(Received Aug. 23, 2021; accepted Mar. 1, 2022)

* Corresponding author