

改进LDS_YOLO网络的遥感飞机检测算法研究

吴杰¹, 高策¹, 余毅¹, 张艳超¹, 裴玉¹, 马少峰²

1. 中国科学院 长春光学精密机械与物理研究所, 长春 130033

2. 西昌卫星发射中心, 四川 西昌 615099

摘要:为解决遥感飞机检测算法网络计算复杂、检测精度低的问题,以主流网络YOLOv4为基础,从提高精度和简化模型两个方面进行改进研究,提出一种轻量级多尺度监督网络LDS_YOLO(light dense supervision YOLO)。针对遥感飞机目标细节信息提取不足的问题,改进三组多尺度融合预测层结构,在每一个支路第一次上采样前的四个卷积块之间设计密集连接方式,可以增强融合不同尺度飞机,丰富特征细节信息,提高预测准确率;针对目标特征关联度低的问题,引入一致性监督损失函数,通过监督分类网络辅助预测的同时提高检测精度;通过增加包含全局平均池化层、全连接层和特征映射层的轻量化模块,调整通道结构减少权重模型的特征冗余,降低网络参数量。在保证检测率的基础上将模型参数量降低为 3.6×10^6 ,计算量为77 MFLOPs,测试检测率比原始模型损失不到2.3%,速度达到17 frame/s;通过与主流检测算法进行对比,分析轻量化后算法模型的抗过拟合能力和鲁棒性,证明轻量化遥感飞机目标检测算法的有效性和可行性。

关键词:目标检测;卷积神经网络;一致性监督损失函数;多尺度融合;轻量化模型

文献标志码:A **中图分类号:**TP391.4 **doi:**10.3778/j.issn.1002-8331.2110-0185

Research on Improved LDS_YOLO Network Remote Sensing Aircraft Detection Algorithm

WU Jie¹, GAO Ce¹, YU Yi¹, ZHANG Yanchao¹, PEI Yu¹, MA Shaofeng²

1. Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China

2. Xichang Satellite Launch Center, Xichang, Sichuan 615099, China

Abstract: In order to solve the problems of complex network calculation and low detection accuracy of remote sensing aircraft detection algorithm, this paper improves the research from two aspects of improving accuracy and simplifying model based on the mainstream network YOLOv4, and proposes a lightweight multi-scale monitoring network LDS_YOLO (light dense supervision YOLO). Aiming at the problem of insufficient extraction of remote sensing aircraft target details, the structure of three groups of multi-scale fusion prediction layers is improved, and the dense connection mode is designed between the four convolution blocks before the first up-sampling of each branch, which can enhance the fusion of aircraft of different scales, enrich the feature details and improve the prediction accuracy. Aiming at the problem of low correlation degree of target features, the consistency supervision loss function is introduced to improve the detection accuracy while assisting prediction by supervision classification network. By adding lightweight modules containing global average pooling layer, full connection layer and feature mapping layer, the channel structure is adjusted to reduce feature redundancy of the weight model and reduce the number of network parameters. On the basis of ensuring the detection rate, the number of model parameters is reduced to 3.6×10^6 , and the computational cost is 77 MFLOPs. The test detection rate loss is less than 2.3% compared with the original model, and the speed is up to 17 frame/s. The anti-over-fitting ability and robustness of the lightweight algorithm model are analyzed by comparing with the mainstream detection algorithm, and the validity and feasibility of the lightweight remote sensing aircraft target detection algorithm are proved.

Key words: target detection; convolutional neural network; consistency monitoring loss function; multi-scale fusion; lightweight model

基金项目:吉林省科技发展计划(20190302086GX)

作者简介:吴杰(1995—),男,硕士,研究方向为模式识别,E-mail:2313050080@qq.com;高策(1981—),男,硕士,副研究员,研究方向为图像处理;余毅(1979—),男,博士,研究员,研究方为向靶场光电测量;张艳超(1985—),女,博士,副研究员,研究方向为数字图像处理;裴玉(1993—),男,硕士,助理研究员,研究方向为光电测控软件;马少峰(1982—),男,高级技师,研究方向为光学测量。

收稿日期:2021-10-15 **修回日期:**2021-03-04 **文章编号:**1002-8331(2022)15-0210-10

遥感图像的目标检测主要研究在图像中是否有目标物体存在,并确定其位置和类别。随着深度学习与遥感技术的发展,基于卷积神经网络的遥感飞机检测算法在民用和军事上都起着至关重要的作用。然而,与传统的目标检测不同,遥感飞机图像受天气环境、光照和背景等因素影响,如果算法要实现高精度的目标检测功能,需要计算机提供大量的算力和存储,降低了算法的检测速度,一定程度上限制了算法的应用。因此权衡算法检测速度与精度,针对遥感飞机设计一种满足模型参数少、检测精度高的网络是有必要的。

针对遥感目标检测技术,早期Liu等人^[1]用模板匹配的方法应用于遥感飞机的轮廓检测,但对于复杂背景的情况检测精度不高。随着深度学习的发展,对遥感目标检测主要有基于区域建议和基于逻辑回归的方法,前者代表有Faster RCNN^[2],具有较高的检测精度,但是遍历候选框导致速度较慢。基于逻辑回归的检测算法主要有SSD^[3]和YOLO系列,SSD在检测速度可以达到较高水平,但是缺陷在于小目标的定位,YOLOv4^[4]可以权衡检测精度和速度,但是两者指标都不是最佳。为了增加检测器性能,Zhang等人^[5]提出结合不同特征的卷积层网络提高对目标的检测精度,引入无锚的检测方法一定程度降低了遥感目标的漏检率,但是模型依旧较为复杂。

针对模型轻量化技术,卷积方式上Google提出了一种轻量级卷积神经网络MobileNet^[6],其核心思想通过卷积核分解来减少模型参数,实现在相同量级缩小模型计算量,小目标检测精度较低。网络结构上何凯明提出的残差模块,通过跳跃连接增加模型深度进而减少参数量^[7],然而由于连接的关联性太强容易产生过拟合,不利于遥感飞机目标的模型学习。Wang等人^[8]设计了一种轻量级的小型网络模型Scaled-YOLOv4,该网络简化了YOLOv4模型复杂度,极大降低了参数量,但是对复杂目标的检测精度有待提高。随后,Liu等人^[9]提出网络剪枝思想,训练时设置一个阈值然后将YOLOv4等网络模型中贡献度低于阈值的神经元裁剪,使得检测模型得到轻量化。

本文研究的遥感飞机检测算法从以上几点出发,结合遥感图像中飞机目标的特点,以网络YOLOv4模型为基础,通过改进多尺度融合预测网络和引入一致性监督网络,在网络结构和损失函数上对检测网络进行改进,结合更多的细节信息解决原网络在遥感飞机检测过程中精度不足的问题。轻量化方面分别在预测层设计密集连接方式,在分类层增加轻量化模块,加强目标特征学习并裁剪多余的卷积,解决基于卷积神经网络的遥感飞机检测算法中模型过大和速度慢的问题。

1 YOLOv4网络检测原理

以主流的YOLOv4网络为基础展开研究,应用对象

为遥感图像中的飞机这一类的简单目标,通过将输入图像调整为 416×416 大小,将待检测图像分成 13×13 形式的网格,再利用中心点与网格宽高预测出边界框宽高,求出三个候选边界框,进而得到最终预测结果。网络将三个边界框的得分进行筛选得出最大的预测置信值,然后该置信值对应的边界框会给相应的目标加上标签。结合目标栅格与图像左上角的边距(C_x, C_y)以及栅格所在候选框的宽高(P_w, P_h),预测出边界框的四个坐标值(T_x, T_y, T_w, T_h)和一个目标存在的可能性得分 T_o 。再由式(1)、(2)计算得出每个边界框与左上角的相对坐标(B_x, B_y)、宽高(B_w, B_h)和预测置信值 $C(confidence)$ ^[10-11]。具体计算公式如下:

$$P_r(obj) \times IOU(box, obj) = \sigma(T_o) \quad (1)$$

$$C = P_r(Object) \times IOU_{pred}^{truth} \quad (2)$$

C 的数值表示该处是否有目标,若置信度为0则认为该处不存在目标,若置信度不为0,则显示该处置信度的数值。 T_x, T_y 用sigmoid函数归一化处理,使其取值在0~1之间, σ 参数将归一化值转化为真实值, $P_r(obj)$ 为目标属于某一类的概率, $\sigma(T_o)$ 表示预测边界框的置信值,IOU表示真实标注框与算法检测框的交并比。如公式(3)所示,检测框精准度与IOU大小成正比。

$$IOU = \frac{area(B_{op} \cap B_{gt})}{area(B_{op} \cup B_{gt})} \quad (3)$$

式中, B_{op} 为预测框与标注框的交集, B_{gt} 表示预测框与标注框的并集。YOLOv4的损失函数由3个部分组成,分别为中心坐标预测、宽高坐标预测和置信度类别预测^[12]。

2 基于YOLOv4的网络结构改进

2.1 基于密集连接改进的多尺度融合预测

YOLOv4存在三个多尺度融合预测的网络层结构,决定着候选框检测的精准度,然而上采样和特征融合过程产生大量迭代参数,降低了检测的速度。因此本文设计在特征获取阶段和多尺度上采样后使用密集连接的网络结构,密集连接技术主要是借鉴跳跃连接的思想,在各个输入加上之前特征的输出,从而达到维度的扩充。假设输出特征图为 x_l ,原始特征图为 x_0 。两者之间的关系如公式(4)所示:

$$x_l = H_l([x_0, x_1, \dots, x_{l-1}]) \quad (4)$$

由式(4)可知,密集连接网络中各卷积层接收之前所有卷积层的特征图作为输入。遥感图像中飞机目标对比度较弱,有效增强了网络对飞机目标特征的提取能力,跳跃的网络连接方式不但可以增加网络深度,提高检测精度,还可以减少用于进行多尺度融合时产生的参数量,改进多尺度融合网络结构如图1所示。

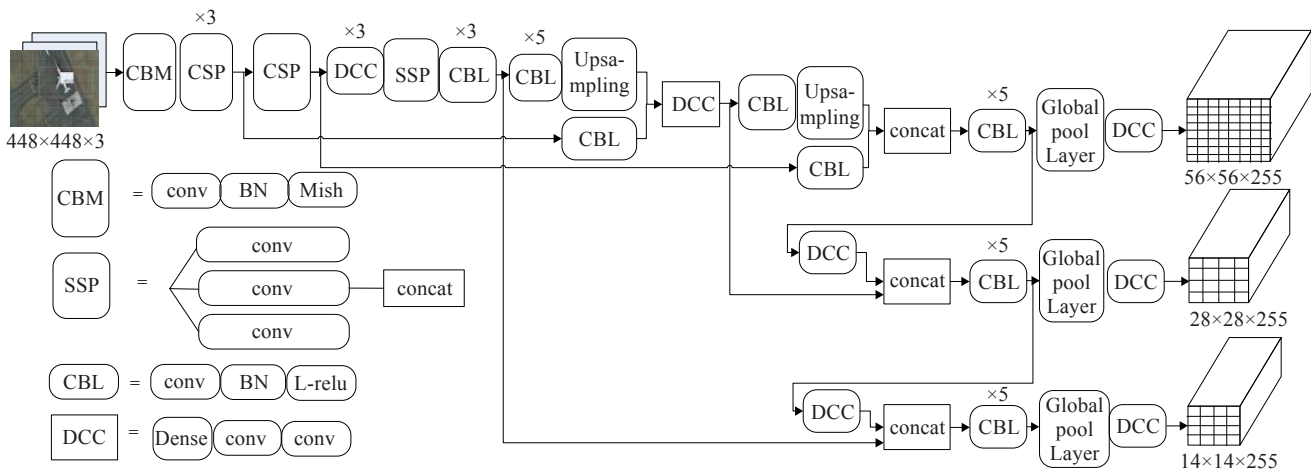


图1 改进多尺度融合网络结构图

Fig.1 Improved multi-scale fusion network structure diagram

图1中将输入图像调整为448×448,这样在预测层能够增大感受野,获得更多的预测信息,同时Mish激活函数曲线在负值不是直接截断,梯度下降较快。网络由特征金字塔的结构组成,通过上采样调整小尺度特征图分辨率,与上层大尺度预测层融合,分别在三组尺度预测层的支路设计密集连接的卷积形式,它们之间用由1×1大小的点卷积组成,作用是增加不同层之间飞机目标信息的关联度,增加密集连接后在进行上采样,可以在融合不同层细节信息的同时,去除多余的冗余。适当调整预测层前的卷积层,选择大小为14×14的最小尺度进行预测,减少模型过拟合严重的问题。

2.2 LDS_YOLO 模型参数量与计算量分析

对YOLOv4的改进集中在第三、四、五个卷积模块,并且特征图在这三个卷积模块中传播时深度保持不变,符合密集连接的要求,需要适度调整增长率,防止特征图经过Dense Block层后通道数会迅速增加在LDS_YOLO中,本文设置k=32,经过多次实验发现这样既能有效加强对特征的提取能力,又避免了冗余参数的大量产生。改进后的LDS_YOLO各层结构及参数如表1所示。

表1可知,改进后LDS_YOLO网络结构使得融合预测层和通道权重层相应地降低了参数量,网络预测层一减少了36 864-26 624=10 240个参数,同理相应地可以计算出网络预测层二减少的参数为3×3×32×32-(3×3×32×16+1×1×16×32)=4 096,网络预测层三由于融合了前两层减少参数更多为3×3×64×64-(3×3×64×32+1×1×32×64)=16 360。轻量化分类模块中的通道权重层减少的参数为3×3×64×64-(1×1×32×64)=34 816由此可得LDS_YOLO的参数含量共计40 779 912,相较于YOLOv4参数量减少了22.51%。

2.3 损失函数的选择

LDS_YOLO网络用逻辑回归方式来预测每个边界框中目标存在的可能性,大小作为预测的得分值,这一步通过设置合适的阈值来去掉得分低的边界框从而减少计算量。在训练过程中,网络用一致性监督损失函数^[13]损失来预测类别。一致性监督损失公式如式(5)、(6)所示:

$$Loss = - \sum_{i=1}^n y_i \lg y_i + (1 - y_i) \lg (1 - y_i) \tag{5}$$

$$\frac{\partial Loss}{\partial y} = - \sum_{i=1}^n \frac{y_i}{y_j} - \frac{1 - y_i}{1 - y_j} \tag{6}$$

表1 LDS_YOLO模型参数量和计算量

Table 1 LDS_YOLO model parameter amount and calculation amount

层名	输出尺寸	各层结构	计算量/10 ⁶	参数量/10 ⁶
Convolution Pooling	448×448×3	7×7 conv, stride 2 3×3 max pool	118	0.009 4
Dense Block(1)	448×448×32	[1×1 conv 3×3 conv]×6	1 000	0.532 0
TransitionLayer(1)	224×224×32	1×1conv, 2×2average pool	103	1.300 0
Dense Block(2)	224×224×64	[1×1 conv 3×3 conv]×12	559	0.618 0
TransitionLayer(2)	112×112×64	1×1conv, 2×2average pool	103	1.700 0
Dense Block(3)	112×112×128	[1×1 conv 3×3 conv]×24	549	0.331 0
TransitionLayer(3)	56×56×128	1×1conv, 2×2average pool	103	2.800 0
Prediction	7×7×1 000	[1×1 conv 3×3 conv]×12	319	0.024 0
Classification Layer		1 000 softmax	1	1.000 0
总计			2 850	8.000 0

其中,当且仅当 y_i 和 y_j 相等时, $Loss$ 为0;否则, $Loss$ 为一个正数。在预测网络设置之后增加一致性监督的网络,利用多尺度融合算法将改进YOLOv4的倒数第一、二、三、四层进行特征融合,构建 $\{C_1, C_2, C_3, C_4\}$ 主特征金字塔。通过上采样方式构建新的 $\{P_1, P_2, P_3, P_4\}$ 次金字塔用于每一层的目标预测,配合主金字塔的每个层通过 1×1 卷积进行通道降维,同时将上采样减少的特征表示为 $\{M_1, M_2, M_3, M_4\}$, 如图2所示。

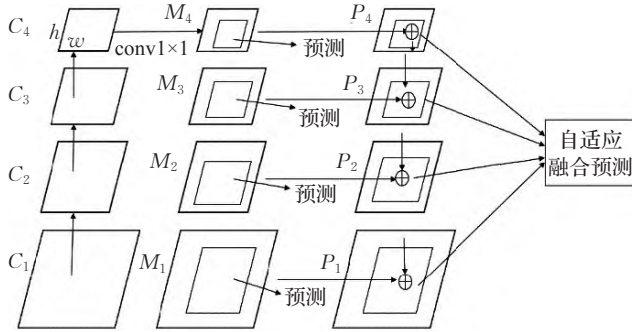


图2 一致性监督损失原理图

Fig.2 Schematic diagram of consistency monitoring loss

多个分类和回归连接到这些特征,以产生辅助损失,这些分类和回归的参数在不同次金字塔层上共享^[14]。为了稳定损失函数的变化,使用权重 λ 和 β 来平衡由一致性监督和原始损失产生的辅助损失。形式上, LDS_YOLO 最终损失函数如公式(7)所示:

$$L_{yolo} = \lambda(L_{cls,M}(p_M, t^*) + \beta[t^* > 0]L_{loc,M}(d_M, b^*)) + (L_{cls,P}(p, t^*) + \beta[t^* > 0]L_{loc,P}(d, b^*)) \quad (7)$$

p_M 、 d_M 和 p 、 d 分别是中间层和最终金字塔层的预测, t^* 和 b^* 分别是基本类别标签和回归目标, λ 是用于平衡辅助损耗和原始损耗的权重, β 是用来平衡分类和定位损失的权重。 $[t^* > 0]$ 的定义如公式(8)所示:

$$[t^* > 0] = \begin{cases} 1, & t^* > 0 \\ 0, & t^* = 0 \end{cases} \quad (8)$$

2.4 基于全局池化层的改进轻量化模块

2.4.1 LDS_YOLO 模型权重冗余分析

结合多尺度融合和密集连接的网络层以及基于全局平均池化层的网络轻量化模型都使得模型参数得到了减少,训练过程中每一个支路的 Dense 结构如图3所示。

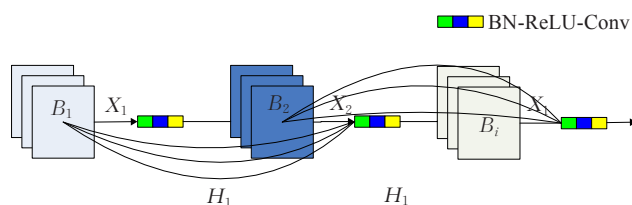


图3 Dense 模块结构

Fig.3 Dense module structure

Dense 模块中令 X_l 表示第 l 层的输出, H_l 表示一个非线性变化。第 l 层的输出结果如式(9)所示,其中 $[X_0, X_1, X_2, \dots, X_{l-1}]$ 表示将0到第 $l-1$ 层的输出特征叠加。Block 的参数量及计算量的计算由公式(10)和(11)表示:

$$X_l = H_l([X_0, X_1, X_2, \dots, X_{l-1}]) \quad (9)$$

$$Params_B_i = F \times F \times C_{out1} \times (C_{in0} \times L_i + C_{out2} \times \frac{L_i(L_i-1)}{2}) + F \times F \times C_{out3} \times C_{in3} \times L_i \quad (10)$$

$$FLOPs_B_i = H_i \times W_i \times Params_B_i \quad (11)$$

其中, B_i 表示第 i 个 Block, L_i 表示 B_i 中卷积层的总层数。

2.4.2 基于改进的特征通道权重的模型裁剪

遥感图像中的目标信息比普通图像环境更为复杂,飞机和周围建筑在形状上极易混淆,因此在特征分类的时候会增大图像分辨率和迭代次数,这样会增加特征分类模块模型参数量,本小节设计一种基于全局平均池化层的轻量化模块,结合特征权重映射在分类遥感飞机目标的同时,去掉模型冗余参数,减少模型大小。改进特征分类模块结构如图4所示。

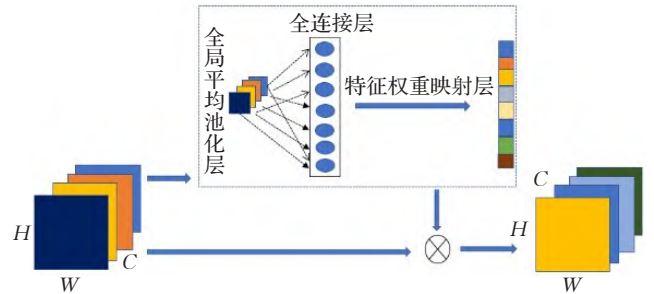


图4 基于全局池化层的轻量化分类模块

Fig.4 Improved pruning based on global pooling

模块主要包含三个部分:全局平均池化层(global average pool)、全连接层和特征权重映射层。其中全局平均池化层的作用是将每一个特征通道上的信息平均成一个值,对应的计算如公式(12)所示:

$$GAP(X) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W X_{i,j} \quad (12)$$

其中, H 、 W 、 C 分别代表输入卷积的高、宽和特征通道数, $X \in \mathbb{R}^{1 \times 1 \times C}$ 代表前一层的输出数据。全局平均池化层主要功能是获取模型通道上的信息,为得到特征通道间的相关性,需要基于全部数据集对模型进行训练。全连接层的作用是将全部数据集的目标信息进行一次全面的通道映射,达到最佳的分类效果。扩展型 Sigmoid 函数的作用是将数据限制在 $[0, 1]$ 范围内,方便设定合适的权重阈值进行裁剪,如公式(13)所示:

$$Y = \text{Sigmoid}(\varphi X) \quad (13)$$

假设在第 l 层对 LDS_YOLO 网络模型进行通道权重特征调整,则前面所有层的输出特征会由一个非线性

变换后得到第 l 层的输出特征,如公式(14)所示:

$$Z_l = X_l \times \text{Sigmoid}(\varphi \times \text{FC}(\text{GAP}(X_l))) \quad (14)$$

Z_l 表示 l 层的最后的输出结果,经过 Sigmoid 函数后的图像信息被限制在一定范围内,由全连接进行通道和特征的映射,通过设定合适的阈值将影响检测率较低的权重部分去除,在保证不丢失检测精度的前提下,可以学习数据集的全部特征并减少模型的参数。

3 实验结果与分析

3.1 数据集与实验环境

文章数据集采用了 NWPU VHR-10^[15] 遥感飞机数据集,受各个卫星所拍出的遥感图片质量差异,飞机的视角差异和高度差异等各种因素影响,包含了多种状况下的遥感航空图片,如图5所示。该数据集中遥感飞机的类别有战斗机、民用机等,环境有大目标、小目标,有建筑和空地等。针对数据集类别不平衡问题,设计场景数据增强方法,使类别均衡化,按样本数量均衡类别权重,利用上采样将少量的民用机和小目标的飞机数量增加,提高模型在复杂环境的鲁棒性。训练集为 23 000 张图像,测试集为 5 000 张。通过手动标注将大量数据汇集的图像标记为 xml 文档,标记后的文档中会存储图像中各个目标的类型、尺寸和方位等信息。

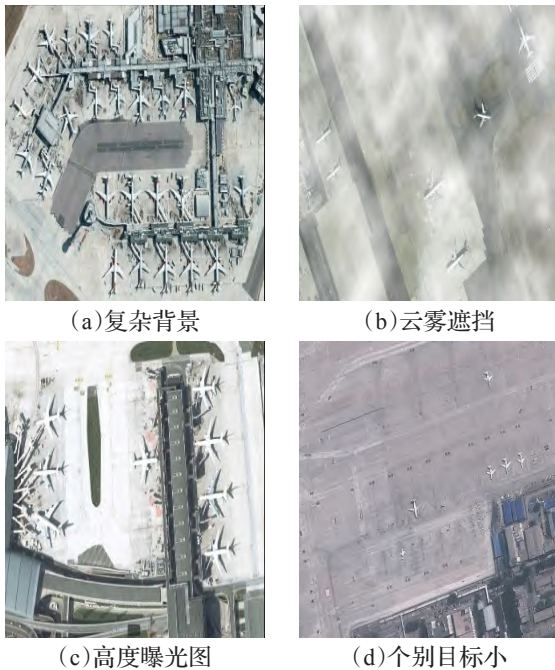


图5 遥感图像中低质量飞机数据集实例

Fig.5 Examples of low-quality aircraft data sets in remote sensing images

3.2 训练模型与性能分析

3.2.1 LDS_YOLO 模型训练分析

本次实验选用 NWPU VHR-10 遥感飞机数据集,主要用来分析 LDS_YOLO 基于密集连接改进的多尺度融合预测后对目标的检测能力,以及模型大小、计算量和

速度的优化程度,最终应用在无 GPU 的嵌入式设备树莓派 3B+ 上进行研究。其中训练集为 50 000 张图像,测试集为 7 692 张,训练迭代次数 batches 为 400 000 次, bath_size 设为 64,学习率设为 0.001,在训练过程图中每 20 000 次迭代保存一次此时的网络模型,为了证明本章所设计的算法可以完成对目标数据集的学习,利用 matplotlib 库将训练中的 loss、IOU 数据可视化,最终得到如图6所示的变化曲线。

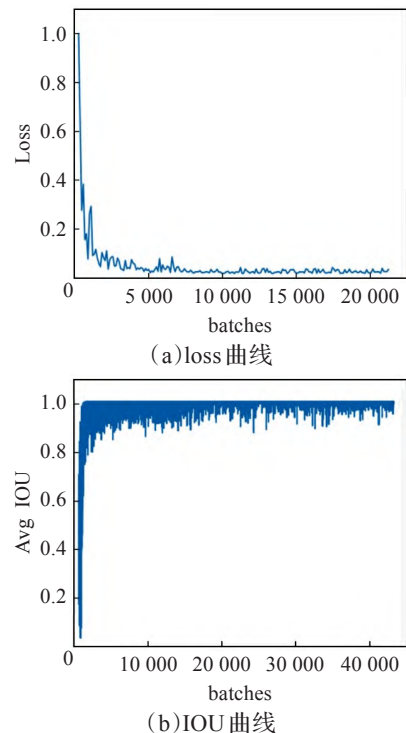


图6 训练过程中的 loss 曲线和 IOU 曲线

Fig.6 Loss curve and IOU curve during training

从图6可以看出,随着迭代次数的增加,模型的精度逐渐提高,从中选择检测率最高的权重可用于后续实验的研究,训练过程中 IOU 也逐渐趋于平稳,平稳值为 1,从而说明 LDS_YOLO 目标检测网络可以学习到遥感飞机数据集中的细节内容。

3.2.2 损失函数改进实验

为了验证论文中一致性监督损失函数的有效性,在 NWPU VHR-10 遥感飞机数据集上将改进的 LDS_YOLO 算法网络作为检测网络,分别使用原损失函数(二元交叉熵)以及一致性监督损失函数对网络进行训练。在训练的过程中对模型特征分类在二维平面内进行可视化,检测结果如图7所示,特征分布以(0,0)中心成放射状。

图7中不同颜色分别对应 NWPU VHR-10 数据集内的 10 种类别。实验发现相比较于原二元交叉熵损失函数,一致性监督损失函数的特征点分布明显使得不同类别之间的距离足够远,并且类别的特征分布变得更为紧凑,相同颜色点的分布近似压缩为一条直线,表明分类效果更好。

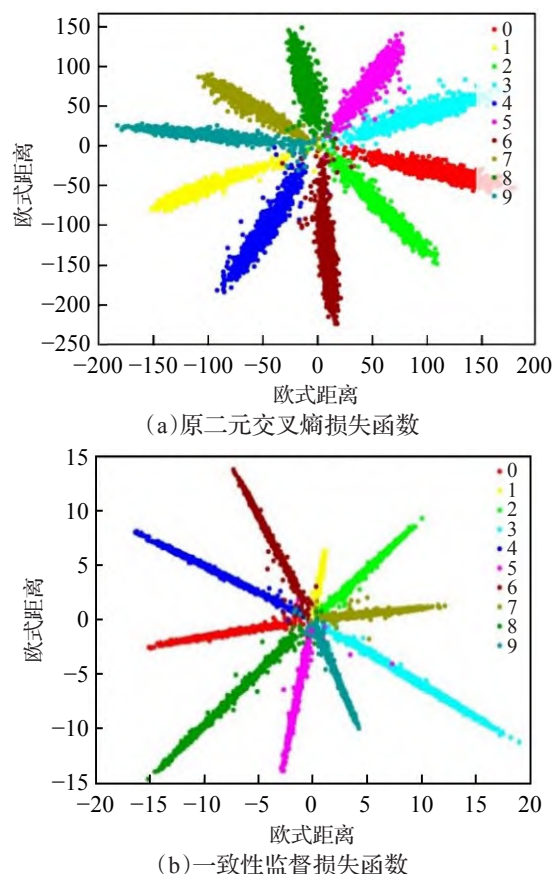


图7 不同损失函数的二维平面内可视化图

Fig.7 Two-dimensional in-plane visualization of different loss functions

3.2.3 改进剪枝后的损失函数实验

(1) 引入特征通道调整层后损失变化分析

本小节主要对LDS_YOLO轻量化算法中引入改进剪枝层的Dense模块进行实验(残差连接模块同样适用该剪枝方法)。首先使用改进前后的LDS_YOLO网络对遥感飞机数据集进行训练,训练过程中的损失函数变化曲线如图8所示。

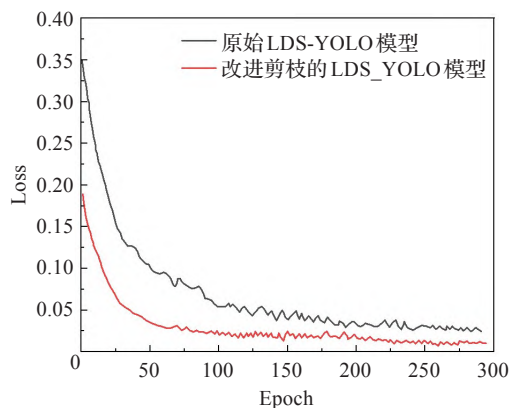


图8 损失函数曲线图

Fig.8 Loss function graph

图8中红色曲线在训练的过程中明显收敛得更快,且收敛后始终在黑色曲线下方,证明了在LDS_YOLO

网络引入改进剪枝层后不会影响到原模型的性能。

(2) 超参数 φ 实验分析

接下来分析应用了特征权重调整层后的LDS_YOLO模型,实验发现在公式(8)中,合适的 φ 值可以有效地将不同种类目标特征区间分开,这一特性方便剪枝时阈值的选取。本节对Sigmoid函数曲线特性进行实验,通过设置大量不同 φ 值进行分析,最终选择在[5, 8, 10, 100]之间做 φ 值的选择实验。使用不同 φ 值在LDS_YOLO网络模型训练时的损失值随着迭代次数而下降的变化曲线趋势如图9所示。

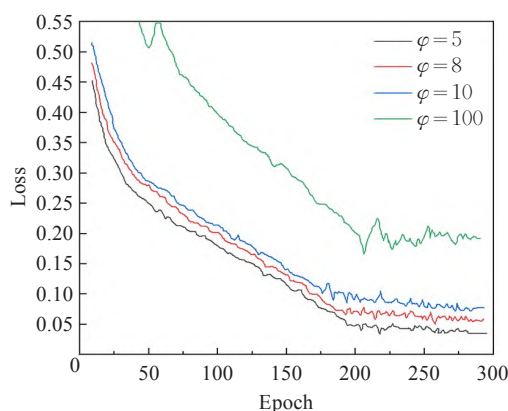


图9 损失函数趋势图

Fig.9 Loss function trend chart

图9中可以看到, φ 值过大会影响模型的收敛(以100为例,其他数值曲线未画出),无法准确地学习到目标的信息。数值在10以下模型才会收敛,而在实验所选取的使模型有效收敛的三个数值中, φ 值为5时所对应的曲线的收敛速度较其他值更快,且趋于平稳。

3.2.4 通道权重剪枝有效性对比实验

本小节实验以LDS_YOLO轻量化算法为例设置三组实验来说明通道权重剪枝的有效性。实验基于Python语言进行,LDS_YOLO模型在遥感飞机数据集上的剪枝实验结果如表2所示。

由表2可知,改进后的剪枝方法可适用于多种目标检测算法,相比于传统剪枝,参数量平均减少了9%,浮点运算平均减少了7%,说明模型在算力和存储空间上的需求都有所降低。检测速度方面剪枝后的模型在无GPU的树莓派3B+上提高了28%以上,证明了改进轻量化模块来减少模型参数的可行性。

3.3 轻量化目标检测算法模型综合实验

军事中经常需要将目标检测算法应用在无GPU的嵌入式平台,例如在无人机系统上对目标进行侦察,需要对目标检测模型在大小和速度上有较高的要求,这对计算机的存储空间和计算能力都是一个挑战。本节选用无GPU显卡的树莓派3B+作为算法的运行平台,展开对多尺度融合预测、损失函数和轻量化分类网络改进后的目标检测算法,进行综合仿真实验研究。

表2 三种算法模型在数据集上的剪枝结果

Table 2 Pruning results of three algorithm models on data set

模型	FLOPs/10 ⁶	Params/10 ⁶	Top-1 检测率/%	CPU 帧率/(frame/s)
YOLOv3(baseline)	265.38	135.0	68.33	1.2
YOLOv3(pruned)	242.38	121.0	57.54	3.7
YOLOv3(pruned-ours)	227.38	112.0	65.23	8.4
YOLOv4(baseline)	187.33	44.0	90.34	4.1
YOLOv4(pruned)	154.75	37.0	79.01	6.1
YOLOv4(pruned-ours)	129.12	28.0	88.75	9.2
LDS_YOLO(baseline)	125.51	8.0	89.95	10.0
LDS_YOLO(pruned)	104.16	6.0	79.12	14.3
LDS_YOLO(pruned-ours)	76.82	3.6	86.02	19.6

3.3.1 实验测试结果与对比分析

为了分析轻量化算法在性能上的变化,本文选用改进前的YOLOv4做对比实验,同样用遥感飞机数据集进行训练,选择效果最好的模型对目标进行测试,为了直观体现轻量化后模型在速度上的优势,左上角添加检测一张图像中目标所需的时间,三种检测方法的测试结果如图10所示。

图10分别为在不同场景的遥感飞机目标下三种算

法的检测效果,从中可以看出YOLOv4的检测速度在300~500 ms,换算成帧率平均在3 frame/s,检测率最高,在实验中存在一些漏检、误检;LDS_YOLO是专门设计用来放在无GPU的嵌入式上运行的,在牺牲1.3%~2.5%准确率的基础上,将检测速度提高为6~7 frame/s, LDS_YOLO_pruned_ours通过增加改进的剪枝模块,牺牲2.6%~3.5%准确率的基础上,将检测速度提高为9~10 frame/s。选取不同场景下的遥感飞机图像,选用准

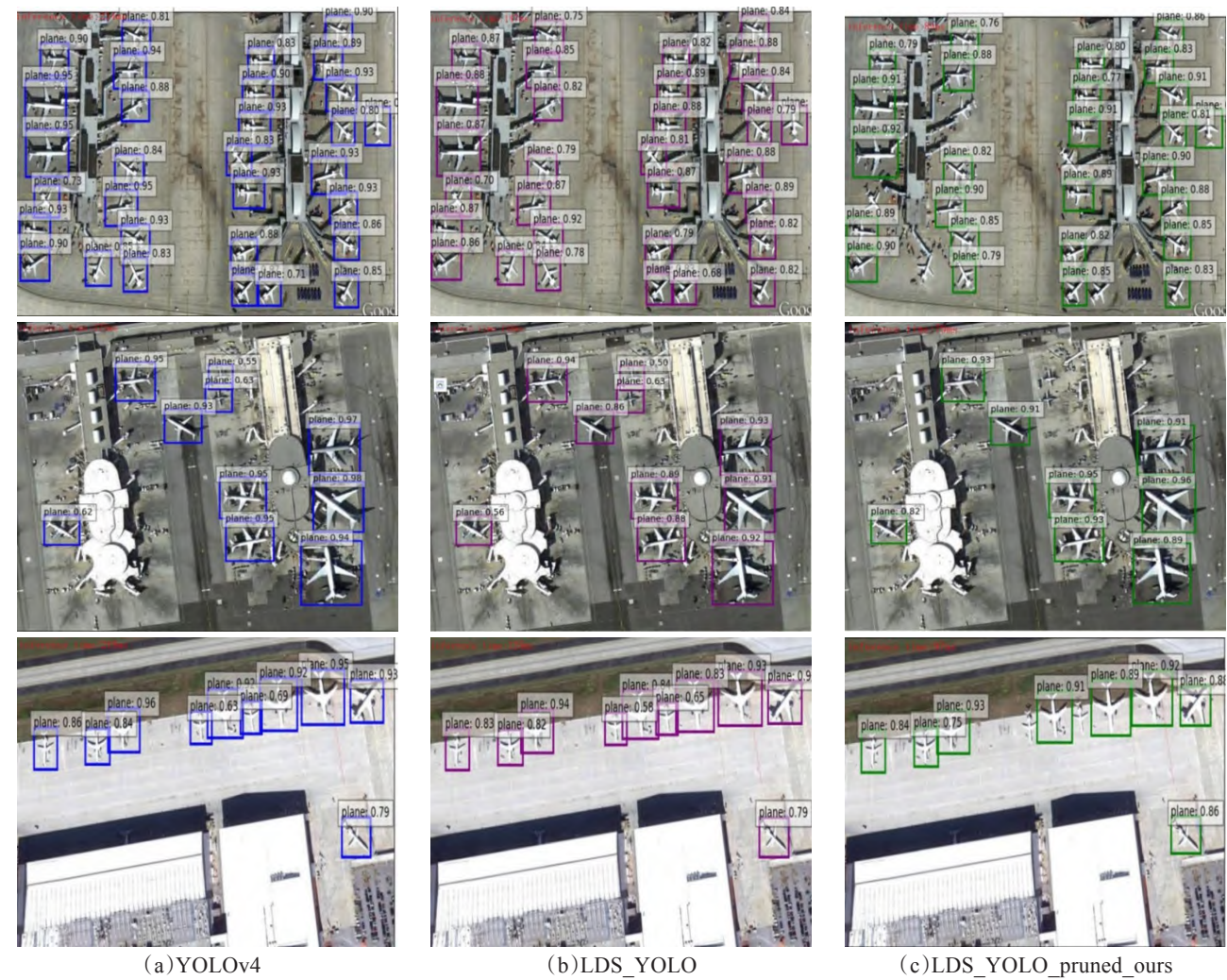


图10 改进前后检测算法测试集的部分效果图

Fig.10 Partial effect of test set of detection algorithm before and after improvement

表3 LDS_YOLO在低质量遥感图像下检测指标

Table 3 Detection indexes of LDS_YOLO in low-quality remote sensing images

图像种类	IOU			mAP			Recall		
	FasterRCNN	YOLOv4	LDS_YOLO	FasterRCNN	YOLOv4	LDS_YOLO	FasterRCNN	YOLOv4	LDS_YOLO
清晰图像	91.3	92.3	94.1	96.3	95.4	96.8	92.8	93.5	95.4
云雾遮挡	90.9	91.3	94.5	89.8	87.4	95.2	88.2	89.4	93.5
高度曝光	87.8	83.2	85.3	84.3	81.2	88.4	85.3	86.3	87.9
复杂背景	88.2	86.5	88.1	88.6	84.4	86.8	90.6	86.3	89.2
目标较小	81.4	82.4	84.6	83.5	86.4	88.9	84.5	87.4	90.1

确率较高但速度慢的Faster RCNN与YOLOv4和本文改进算法进一步比较在高曝光、目标小等复杂背景下改进算法的效果如表3所示。

由表3可以看出,改进后的LDS_YOLO算法在各个低质量遥感图像上的检测指标都存在一定程度上的提升,相比于YOLOv4算法LDS_YOLO算法mAP和召回率在小目标上提升最为明显,分别提高2.2%和4.1%,在云雾遮挡的环境下由于LDS_YOLO算法加入了一致性损失函数,可以更好地根据对比度信息在迭代过程中监督检测出飞机的形状,检测率可以高达95.2%。几种算法在高度曝光环境下检测结果影响最小,而复杂背景的目标检测在mAP上Faster RCNN效果较好,而本文算法存在进一步的改进空间。为了与其他算法比较,文章选取主流轻量化目标检测算法在相同数据集下进行对比实验分析,结果如表4所示。

表4 几种算法在数据下的对比实验分析

Table 4 Comparative experimental analysis of several algorithms under data

模型	FLOPs/10 ⁶	Params/10 ⁶	Top-1 检测率/%
VGG-19-pruned ^[16]	195.0	14.3	88.80
ResNet-56-pruned ^[17]	129.0	13.3	91.06
ShuffleNet(1.0) ^[18]	161.7	19.1	90.29
DenseNet-40-pruned ^[19]	190.0	16.6	86.81
YOLOv4	131.0	9.4	95.32
LDS_YOLO	110.0	10.1	94.32
LDS_YOLO-pruned	101.0	8.0	90.44
LDS_YOLO-pruned-ours	85.0	3.6	92.95

表4中性能参数包括Top-1检测率、模型的计算量FLOPs和参数量Params,通过在不同经典目标检测公共数据集上进行训练测试。可以看出,文章所设计的轻量化目标检测算法LDS_YOLO与其他轻量化算法比较,在计算量FLOP和参数量Params上最少,平均计算量为100 MFLOPs,模型参数量约为 1.0×10^7 ,检测率最高可以达到90%以上,同时改进后的剪枝操作,在检测率下降不超过1.5%的前提下,计算量和参数量进一步减少18%和21%,再次证明文章中轻量化目标检测算法LDS_YOLO和改进剪枝的LDS_YOLO在检测率高于主流算法的同时,计算量和参数量也都明显降

低,整体优于其他算法。为了证明经过优化得到的目标检测网络可以在遥感飞机数据集上取得很好的效果,下面对模型的抗过拟合能力、鲁棒性等进行实验分析。

3.3.2 模型抗过拟合能力的对比实验

为了分析轻量化LDS_YOLO的抗过拟合能力,文章选择在一个小训练集上进行训练,比较准确率和损失函数变化值^[20]。根据模型的准确率和损失值比较可以明显看出不同模型的抗过拟合能力差异。分别对YOLOv4、LDS_YOLO和LDS_YOLO_pruned_ours算法进行训练,实验结果如图11所示。

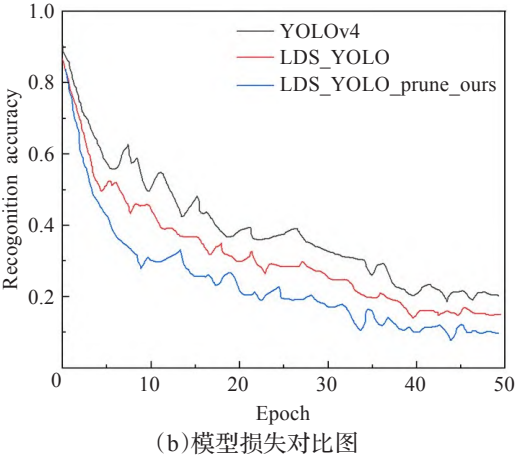
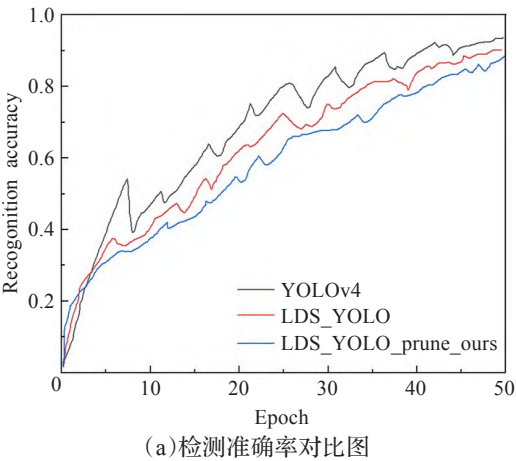


图11 模型对比实验结果图

Fig.11 Model comparison experiment results

图中Epoch表示遍历训练集的次数,Epoch与迭代次数的关系如公式(15)所示:

表5 难样本下各个算法鲁棒性数据

Table 5 Robustness data of each algorithm under difficult samples

难样本	准确率/%				
	YOLOv3	YOLOv4	LDS_YOLO	LDS_YOLO_prune	LDS_YOLO_prune_ours
旋转	66.5	92.9	88.2	79.3	86.3
平移	66.4	91.7	89.5	81.1	88.6
噪声	66.3	88.5	87.0	78.2	86.8
光照强	65.5	93.5	91.9	83.6	89.3
光照弱	64.7	91.1	89.6	77.5	88.6
复杂背景	67.4	88.4	86.9	72.1	82.0

$$\text{迭代次数} = \frac{\text{epoch} \times \text{数据量}}{\text{batch_size}} \tag{15}$$

由图 11 (a) 可知 YOLOv4、LDS_YOLO 和 LDS_YOLO_prune_ours 算法随着迭代次数的增加准确率逐渐上升,可见通过密集连接和通道剪枝的操作进行轻量化后的模型具有良好的学习能力;两者相比于 YOLOv4 在分类能力上分别损失了 1.2%和 2.3%。由图 11(b)可知, YOLOv4 的损失值整体在 0.41 以上,模型参数过多会导致损失值随着训练次数的增加而增加,在反向传播过程中均发生梯度爆炸,导致模型收敛变慢; LDS_YOLO 和 LDS_YOLO_prune_ours 的损失值随着训练的进行逐渐下降,可见轻量化后的模型具有良好的学习能力与抗过拟合能力。

3.3.3 轻量化后模型的鲁棒性实验

Madry 等人^[21]发现模型的容量越大,鲁棒性越好。CNN 具有一定程度的平移不变性和旋转不变性,这对检测系统来说是一个很好的特性。鲁棒性强的系统希望当图像受到一定程度的扰动时(比如旋转、平移、噪声、光照条件变化、复杂背景),系统能够容忍扰动而保持原本的输出,为了探究轻量化算法训练出的模型鲁棒性强弱,文章设置相应的难样本,分析模型检测目标的准确率,结果如表 5 所示。

结合实验结果和表 5 可以看出,当图像受到一定程度的扰动,包括旋转、平移、噪声、光照条件变化时,各个模型的检测率均会下降 1.3%左右,轻量化模型 LDS_YOLO 和通道剪枝 LDS_YOLO 没有随着图像扰动而使得检测率大幅度下降,仅仅在复杂背景下会出现 6%左右的下降,这一背景下其他成熟算法也都有不同程度的下降,原因是基于卷积神经网络的检测算法的不变性更多体现在旋转、平移和光照上,对于背景复杂、目标小等图像还是存在一定的不足。总体表明,轻量化算法训练出的模型对于大部分图像扰动,都展现出了很好的鲁棒性。

4 结论

本文针对目标检测网络在遥感飞机数据集上检测精度低、速度慢、测试模型大等问题设计了一种改进型

检测网络 LDS_YOLO。该网络通过多尺度融合预测层改进网络,结合改进损失函数,增加目标飞机信息关联度,提高检测精度并减少参数;利用包含全局平均池化层的轻量化分类网络对 LDS_YOLO 通道卷积进行裁剪,进一步减少网络模型大小。并对算法模型做实验性能分析,进一步证明了文章提出的目标检测算法的可行性。未来考虑在多尺度预测层设计自适应特征选取网络,加强主动判别重要特征能力,增加模型去冗余的速度。

参考文献:

[1] LIU G, SUN X, FU K, et al. Interactive geospatial object extraction in high resolution remote sensing images using shape-based global minimization active contour model[J]. Pattern Recognition Letters, 2013, 34(101): 1186-1195.

[2] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: towards real-time object detection with region proposal networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2016, 39(6): 1137-1149.

[3] LIU W, ANGUELOV D, ERHAN D, et al. SSD: single shot multibox detector[C]//European Conference on Computer Vision. Berlin: Springer International Publishing, 2016: 21-37.

[4] BOCHKOVSKIY A, WANG C Y, LIAO H Y M. YOLOv4: optimal speed and accuracy of object detection[C]//IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2020.

[5] ZHANG T, WANG G, ZHUANG Y, et al. Feature enhanced centernet for object detection in remote sensing images[C]//IGARSS 2020-2020 IEEE International Geoscience and Remote Sensing Symposium, 2020: 1639-1642.

[6] ZHANG X, ZHOU X, LIN M, et al. Shufflenet: an extremely efficient convolutional neural network for mobile devices[J]. arXiv:1707.01083, 2017.

[7] HE K, ZHANG X, REN S, et al. Deep residual learning for image conference recognition[C]//Proceedings of Computer Vision and Pattern Recognition (CVPR), 2016: 770-778.

[8] WANG C Y, BOCHKOVSKIY A, LIAO H Y M. Scaled-

- YOLOv4: scaling cross stage partial network[C]//2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition(CVPR),2021.
- [9] LIU X,WANG X,MATWIN S.Improving the interpretability of deep neural networks with knowledge distillation[C]//2018 IEEE International Conference on Data Mining Workshops(ICDMW).Singapore:IEEE,2018.
- [10] 许虞俊,李晨.基于YOLO优化的轻量级目标检测网络[J].计算机科学,2021,48(z2):265-269.
- XU Y J,LI C.Light-weight object detection network optimized based on YOLO family[J].Computer Science, 2021,48(z2):265-269.
- [11] RUSSAKOVSKY O,DENG J,SU H,et al.Imagenet large scale visual recognition challenge[J].International Journal of Computer Vision,2015,115(3):211-252.
- [12] REDMON J,DIVVALA S,GRIRSHICK R,et al.You only look once:unified,real-time object detection[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.Las Vegas,USA:CVPR,2016: 779-788.
- [13] WU D,LU S,JIANG M,et al.Using channel pruning-based YOLOv4 deep learning algorithm for the real-time and accurate detection fapple flowers in natural environments[J].Computers and Electronics in Agriculture, 2020,178(5):174-178.
- [14] 于娟,罗舜.基于YOLOv5的违章建筑检测方法[J].计算机工程与应用,2021,57(20):236-244.
- YU J,LUO S.Detection method of illegal building based on YOLOv5[J].Computer Engineering and Applications, 2021,57(20):236-244.
- [15] HUANG R,PEDOEEM J,CHEN C X.YOLO-LITE: a real-time object detection algorithm optimized for non-GPU computers[J].arXiv:1811.05588,2018.
- [16] 吴杰,段锦,赫立群,等.DS-YOLO网络在遥感图像中的飞机检测算法研究[J].计算机工程与应用,2021,57(1): 181-187.
- WU J,DUAN J,HE L Q,et al.Research on aircraft detection algorithm of DS-YOLO network in remote sensing image[J].Computer Engineering and Applications, 2021,57(1):181-187.
- [17] 陈春雪.卷积神经网络模型轻量化研究[D].广州:华南理工大学,2019.
- CHEN C X.Research of lightweight convolutional neural network[D].Guangzhou:South China University of Technology,2019.
- [18] 郑志强,刘妍妍,潘长城,等.改进YOLO V3遥感图像飞机识别应用[J].电光与控制,2019,26(4):28-32.
- ZHENG Z Q,LIU Y Y,PAN C C,et al.Application of improved YOLO V3 in aircraft recognition of remote sensing images[J].Electronics Optics & Control,2019, 26(4):28-32.
- [19] LIU Z,LI J,SHEN Z,et al.Learning efficient convolutional networks through network slimming[C]//Proceedings of the IEEE International Conference on Computer Vision(CVPR).San Diego:IEEE Computer Society,2017:2736-2744.
- [20] HE Y,ZHANG X,SUN J.Channel pruning for accelerating very deep neural networks[C]//Proceedings of the IEEE International Conference on Computer Vision.San Diego:IEEE Computer Society,2017:1389-1397.
- [21] MADRY A,MAKFLOV A,SCHMIDT L,et al.Towards deep learning models resistant to adversarial attacks[J].arXiv:1706.06083,2017.