

改进SSD的飞机遥感图像目标检测

王浩桐, 郭中华

引用本文:

王浩桐, 郭中华. 改进SSD的飞机遥感图像目标检测[J]. 液晶与显示, 2022, 37(1): 116–127.

WANG Hao-tong, GUO Zhong-hua. Improved SSD based aircraft remote sensing image target detection[J]. *Chinese Journal of Liquid Crystals and Displays*, 2022, 37(1): 116–127.

在线阅读 View online: <https://doi.org/10.37188/CJLCD.2021-0203>

您可能感兴趣的其他文章

Articles you may be interested in

基于注意力与特征融合的光学遥感图像飞机目标检测

Attention and feature fusion for aircraft target detection in optical remote sensing images

液晶与显示. 2021, 36(11): 1506–1515 <https://doi.org/10.37188/CJLCD.2021-0088>

一种高精度的卷积神经网络安全帽检测方法

High precision detection method of safety helmet based on convolution neural network

液晶与显示. 2021, 36(7): 1018–1026 <https://doi.org/10.37188/CJLCD.2020-0309>

基于特征融合的轻量级SSD目标检测方法

Lightweight SSD object detection method based on feature fusion

液晶与显示. 2021, 36(10): 1437–1444 <https://doi.org/10.37188/CJLCD.2021-0007>

基于注意力掩模融合的目标检测算法

Object detection algorithm based on attention mask fusion

液晶与显示. 2019, 34(8): 825–833 <https://doi.org/10.3788/YJYXS20193408.0825>

基于改进Faster R-CNN的交通标志检测算法

Traffic sign detection algorithm based on improved Faster R-CNN

液晶与显示. 2021, 36(3): 484–492 <https://doi.org/10.37188/CJLCD.2020-0195>

文章编号:1007-2780(2022)01-0116-12

改进 SSD 的飞机遥感图像目标检测

王浩桐¹, 郭中华^{1,2*}

(1. 宁夏大学 物理与电子电气工程学院, 宁夏 银川 750021;

2. 宁夏沙漠信息智能感知重点实验室, 宁夏 银川 750021)

摘要:针对当前飞机遥感图像目标检测算法的精度和实时性不能兼顾的问题,提出了基于 SSD(Single Shot MultiBox Detector)的飞机遥感图像目标检测算法。首先使用经过改进后的深度残差网络替换 SSD 的骨架网络,对于特征图之间缺少特征信息关联和特征通道间缺少差异性权重值的问题,设计了一种含有特征感受野增强模块与注意力机制模块的新型特征金字塔网络。该网络用以融合不同层级的特征信息和训练特征通道间的权重系数,使得深层网络和浅层网络都得到结构层次丰富的融合特征,为后续网络的分类与定位提供了良好的前提。另外,在改进后的 SSD 算法中还使用了聚焦分类损失函数来解决正负样本不平衡的问题。在飞机遥感数据集上进行相关实验,精度均值达到 92.45%,每秒帧率为 35.6。结果表明,改进后的 SSD 算法能够同时兼顾高检测精度和实时性。

关键词:目标检测;遥感图像;特征融合;感受野增强;注意力机制

中图分类号:TP394.1;TH691.9 **文献标识码:**A **doi:**10.37188/CJLCD.2021-0203

Improved SSD based aircraft remote sensing image target detection

WANG Hao-tong¹, GUO Zhong-hua^{1,2*}

(1. School of Physics and Electronic and Electrical Engineering, Ningxia University, Yinchuan 750021, China;

2. Ningxia Key Laboratory of Desert Information Intelligent Perception, Yinchuan 750021, China)

Abstract: Aiming at the problem that the accuracy and real-time performance of current aircraft remote sensing image target detection algorithms can not be balanced, an aircraft remote sensing image target detection algorithm based on single shot multibox detector(SSD) is proposed. Firstly, the improved deep residual network is used to replace the skeleton network of SSD. For the problems of lack of feature information association between feature maps, and different channels from feature maps have no weight values, this paper designs a new feature pyramid network with feature receptive field enhancement module and attention mechanism module. The network is used to fuse the feature information of different levels and the weight coefficients between the training feature channels, so

收稿日期:2021-08-05;修订日期:2021-09-09.

基金项目:宁夏自然科学基金(No.2020AAC03026);宁夏大学研究生创新研究项目(No.GIP2020075)

Supported by Ningxia Natural Science Foundation (No.2020AAC03026); Ningxia University Graduate Innovative Research Project Foundation (No.GIP2020075).

* 通信联系人, E-mail: guozhh@nxu.edu.cn

that both the deep network and the shallow network can obtain fusion features with rich structural levels, which provides a good prerequisite for the classification and positioning of the subsequent network. In addition, the focus classification loss function is also used in the improved SSD algorithm to solve the problem of imbalance between positive and negative samples. Related experiments are carried out on the aircraft remote sensing data set, and the average accuracy reaches 92.45%, and the frame per second is 35.6. The results show that the improved SSD algorithm can balance high detection accuracy and real-time performance at the same time.

Key words: target detection; remote sensing image; feature fusion; attention mechanism; receptive field enhancement

1 引言

随着遥感技术的发展,遥感图像目标检测在军事和民用领域有着广泛应用,其中光学飞机遥感图像目标检测已经成为遥感领域的重要研究方向之一^[1]。飞机遥感图像目标检测的核心任务是判断遥感图像中是否存在飞机目标并对其进行检测、特征提取、分类和定位。传统的遥感图像目标检测算法需要手工定义特征,如模板匹配或机器学习等方法^[2]。此类方法定义的特征鲁棒性差,对尺度多变和背景复杂的遥感图像难以进行有效的目标检测。现今,得益于硬件 GPU 的高性能计算,基于深度学习的目标检测模型成为遥感图像领域的研究热点^[3]。目前基于卷积神经网络^[4](Convolutional Neural Network, CNN)的目标检测算法主要分为两类,一类是以 Fast R-CNN^[5]和 Faster R-CNN^[6]为代表的两阶段目标检测算法,另一类是以 YOLO^[7](YOU Only Look Once)和 SSD^[8]为代表的单阶段目标检测算法。

以上算法常用于检测日常生活中的通用物体,对尺度大小不一和背景复杂的遥感图像难以进行有效的飞机目标检测。如果直接将现有的深度学习目标检测算法应用到遥感影像上进行飞机目标检测,显然是不能得到等同于自然光学图像的检测精度。因此,对常用物体的目标检测算法进行改进,变得更加适合于遥感图像的目标检测算法已成为了近年来的热点问题之一^[9]。文献[10-11]使用 Faster-RCNN 算法对遥感图像进行目标检测,该算法取得的检测精度较好,但检测效率偏低,基本不能满足实时检测的条件。文献[12-13]在使用 SSD 算法的

基础上,结合特征金字塔网络^[14](Feature pyramid networks, FPN)结构的优势,得到一种基于特征融合的遥感图像目标检测算法。该算法在满足实时检测条件的同时也一定程度上提高了遥感目标的检测精度,但是对于小尺度的飞机遥感目标来说,检测精度一般较差。对以上 4 篇目标检测文献的优缺点进行分析可知,高检测精度与实时检测往往不能同时兼顾,可能有如下两个原因:其一,飞机遥感图像经常存在场景复杂多变和目标尺度大小不一的情况。虽然使用了 FPN 进行多尺度特征融合,但是特征融合的程度并不能满足飞机遥感目标的高精度检测,还增加了网络的训练参数,降低了模型的检测效率。其二,在模型训练的过程中默认了所有特征图通道的重要程度都是一致的,缺少通道之间的关联性会导致算法最终的检测精度不够理想。

为实现飞机遥感目标的高精度检测,本文首先使用改进后的深度残差网络^[15](Deep residual network, ResNet)替换了 SSD 算法的 VGG^[16](Visual geometry group)特征提取网络,用以提升算法对飞机遥感图像的特征提取能力。而后设计了一种新型特征金字塔网络用以实现不同层级的特征图之间的特征融合。本文设计的新型网络可以提高算法对不同尺度飞机目标的特征提取能力,对小尺度飞机目标的检测精度有较大的提升。同时本文还使用了注意力机制网络^[17](Squeeze-and-excitation networks, SENet)和聚焦分类损失函数^[18](Focal classification loss),用以提升网络学习的有效性,改进后的算法仅需较少的时间代价即可带来较高的检测精度提升。

2 预备工作

2.1 SSD 目标检测网络

SSD 是一种端对端的单阶段目标检测算法,该算法在检测效率和检测精度上有着较好的平衡性。除此之外,该算法的优势主要有 3 个方面:其一,采用多种尺度特征进行目标检测。Conv4_3、FC7 层在原图上的感受野较小,主要用于检测较小目标物体,而 Conv8_2、Conv9_2、Conv10_2、Conv11_2 在原图上的感受野较大主要用来检测较大目标物体;其二,设置不同高宽比例的锚框用以更好地匹配目标;其三,多种数据增强的方法使得算法对输入不同大小和不同形状的目标具有更好的鲁棒性。本文以输入尺寸为 300×300 的 SSD 网络为例,如图 1 所示。

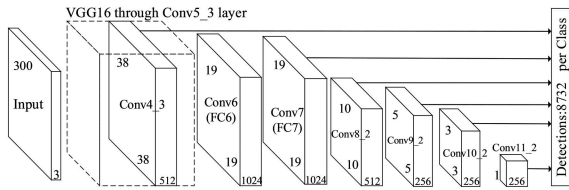


图 1 SSD300 网络结构图

Fig.1 SSD300 algorithm framework diagram

2.2 特征金字塔网络

特征金字塔网络是一种将不同层级的特征图通过自上而下的方式进行特征信息融合的网络。本文所提到的原始 FPN 是以 SSD 的骨架网络输出的 6 个特征图为基础进行融合的结果。不同层级的特征图融合的优势在于弥补了特征层之间的语义差距。特征金字塔网络结构如图 2 所示。 $0.5\times$ 和 $2\times$ 分别表示特征图下采样(缩小)2 倍和

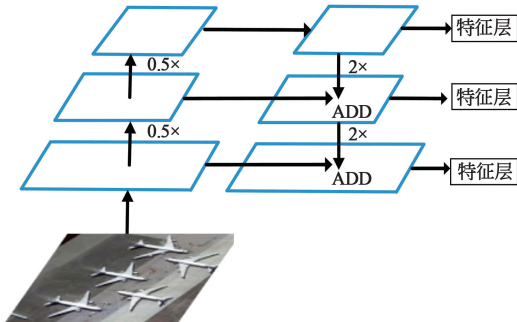


图 2 特征金字塔网络

Fig.2 Feature pyramid network structure

特征图上采样(扩大)2 倍,特征金字塔网络结构如图 2 所示。

3 改进的 SSD 目标检测算法

3.1 骨架网络的替换

ResNet 是一种用于特征提取的卷积神经网络。因为其具有高效性和实用性广泛用于计算机视觉目标检测领域。以往的特征提取网络经常会面临网络层数过深而导致性能退化的问题。ResNet 的不同之处在于提出了残差结构来解决网络退化的问题,残差结构如图 3(a)所示。

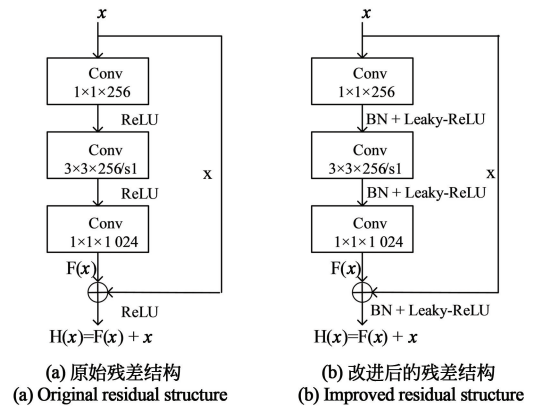


图 3 残差结构

Fig.3 Residual structure

在残差结构中,当输入为 x 时,原始特征记为 $H(x)$,残差特征记为 $F(x)$ 。则网络学习到的残差特征可以表示为:

$$F(x) = H(x) - x. \quad (1)$$

与原始特征 $H(x)$ 相比,深度残差网络更容易学习到残差特征 $F(x)$ 。在理想状态下,当 $F(x)$ 为 0 时,ResNet 仅仅只是完成了恒等映射的任务,网络的学习能力不会因层数的增多而下降。实际上 $F(x)$ 并不会等于 0 或者一直等于 0,残差结构总会通过输入学习到新的特征,网络的学习性能自然会更好^[19]。ResNet101 与 VGG16 两种特征提取网络以 Faster-RCNN 作为目标检测算法的成绩如表 1 所示^[15]。

本文改进后的残差结构如图 3(b)所示,相比于原残差结构增加了批量归一化(Batch Normalization, BN)^[20]和 Leaky-relu^[21]。BN 的作用是为了对输出特征图进行归一化处理,加快网络的

表 1 使用 Faster R-CNN 的两种特征提取网络在 PASCAL VOC 2007/2012 数据集上进行目标检测 (mAP)

Tab.1 Faster R-CNN's two feature extraction networks used to perform target detection on the PASCAL VOC 2007/2012 data set (%)

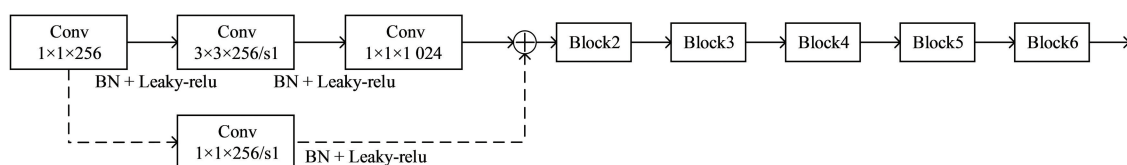
Training data	Test data	VGG16	ResNet101
07+12	VOC 07 test	73.2	76.4
07++12	VOC 12 test	70.4	73.8

收敛速度,同时也能一定程度上解决梯度爆炸或者梯度消失导致网络无法训练的问题。ReLU 激活函数的特点是当输入小于 0 时,输出结果恒为 0。这样的结果会导致网络可能会丢失一部分特征信息,个别情况下会导致对应权重无法更新从而无法学习到有效的特征。而 Leaky-relu 激活函数即使在输入小于 0 时,仍然会有不为 0 的输出,使得模型能够持续更新权重和学习特征。

通过以上分析可知,ResNet 相比于 SSD 算法中的骨架网络(VGG16)有着更好的特征提取能力。在模型大小与计算量方面,ResNet50 在分类精度与 ResNet101 相差不多的情况下,计算量远小于后者。因此在改进的过程中,首先将 SSD 原本的骨架网络 VGG16 替换为网络层数相对更

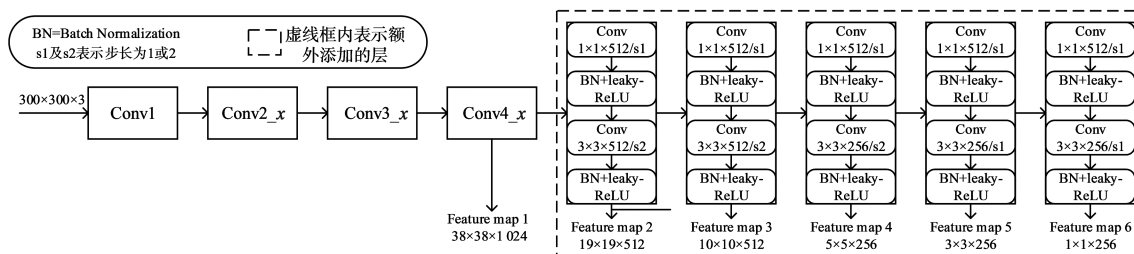
深且计算量相对较低的 ResNet50。为了让 ResNet50 网络适合作为 SSD 算法的骨架网络,现对 ResNet50 网络做部分调整,删除和添加一些层作为网络输出的特征图。

在 SSD 算法中有 6 个尺度和通道数都不相同的输出特征图。为了与 SSD 算法的输出特征图相匹配,对于输入尺寸为 $300 \times 300 \times 3$ 的 ResNet50,对 Conv4_x 的第一个残差结构进行修改,将短路连接(Shortcut connection)上的卷积核的步长由 2 修改为 1,Conv4_x 的输出就变成了 $38 \times 38 \times 1024$ 。将其作为 SSD 算法中的第二个输出特征图(Feature map 2),与修改前的输出结果($38 \times 38 \times 512$)相比,可以发现输出特征图的通道数增加了一倍,得益于残差结构的优点,输出特征图可以学习到的特征信息量变得更加丰富。改进后的 Conv4_x 结构如图 4(a)所示,Block(x:2,3,4,5,6)由图 3(b)组成。当输入尺寸为 $300 \times 300 \times 3$ 时,Conv5_x 的特征输出层尺寸和 SSD 算法的第 2 个特征输出层尺寸不一致,故删除 ResNet50 网络里 Conv4_x 之后的层结构,额外设计了 5 个输出特征图 Feature map(2~6)。在每次缩小输出特征图尺寸前使用 1×1 的卷积核重新聚合前层的特征信息并调整通道数,之后再次用 3×3 的卷积核进行输



(a) Conv4_x 层

(a) Conv4_x layer



(b)改进后的 ResNet50

(b) Improved ResNet50

图 4 骨架网络的替换

Fig.4 Skeleton network replacement

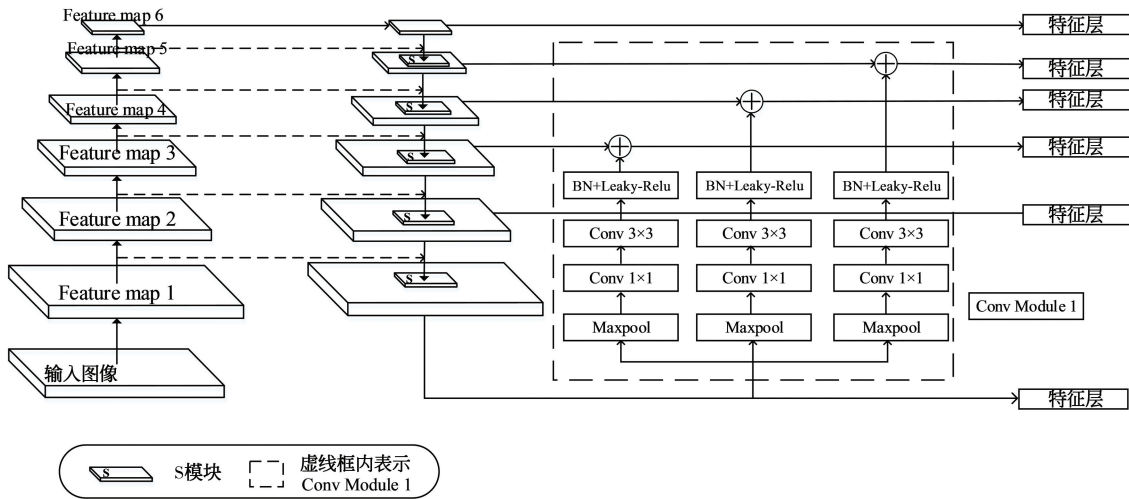
出特征图尺寸的缩小,循环此操作达到多尺度输出特征图的结果。

Feature map(1~6)是改进后的残差模块与额外添加层的输出结果,因而可以在保证模型训练收敛的同时使特征提取网络具有更好的语义抽象能力,改进后的 ResNet50 如图 4(b)

所示。

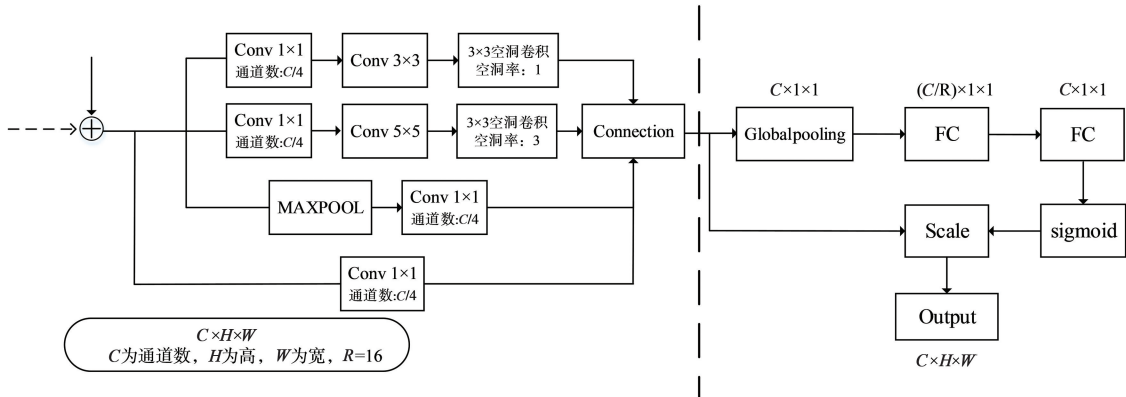
3.2 新型特征金字塔网络

本文提出了一种新型特征金字塔网络,如图 5(a)所示。该网络除却 FPN 经典的自上向下的卷积网络以外,还有 S 模块与 Conv Module 1 模块两个部分组成。



(a)改进后的 SSD 网络结构图

(a) Diagram of improved SSD network structure



(b)S 模块

(b) S module

图 5 改进后的 SSD 算法

Fig.5 Improved SSD algorithm

(1)S 模块(特征感受野增强模块+注意力机制模块)

特征感受野增强模块如图 5(b)虚线左侧所示,该模块由并行卷积核和短路连接两部分组成。在并行卷积核的部分中,在使用 3×3 , 5×5 , Maxpool 卷积核进行不同尺度目标的特征感受野增强之前,首先使用 1×1 卷积核进行通道数的减

少,这样的操作减少了网络的训练参数,变相地提高了网络的效率。另外,空洞卷积的好处不仅在于可以学到更多的特征信息,还可以增大飞机目标的感受野。不同的空洞率还附带有下采样的效果,使网络学习到的特征信息更加丰富^[22]。并行卷积核的输出与短路连接的 1×1 直连卷积核连接,进行通道之间的拼接特征融合(Connection)。

将特征感受野增强模块添加到不同层级的特征图之间,不仅可以提升特征图对飞机目标的感受野,还可以增强网络对小尺度飞机目标的检测能力。

在传统 FPN 结构中,对于同级特征图部分仅通过简单线性叠加方式进行特征融合,而注意力机制可对特征图内部不同位置特征施以差异性权重值,有选择性地激活任务相关区域,抑制次要信息以提升模型效果^[23]。为此,本文设计带有 SENet 的 FPN 结构模块,其具体操作模块如图 5(b)虚线右侧所示。首先通过全局平均池化得到每个通道上的均值,使二维的特征通道变成一个具有全局感受野的实数。而后通过一个 FC(全连接层)将特征维度降低到输入的 $1/16(R=16)$,经过 ReLu 激活后,再通过一个 FC 层升回到原来的维度。这样的操作比直接用一个 FC 层更具有优势:(1)可以具有更多的非线性且更好地拟合通道间复杂的相关性;(2)极大地减少了参数量和计算量。降维和升维操作之后再通过一个 Sigmoid 激活函数得到归一化的权重,最后通过一个 Scale 的操作将归一化后的权重加权到每个通道的特征上。

(2) Conv Module 1

结构为自顶向下的 FPN 是为了解决高低阶特征图中语义与位置信息间的矛盾,实现对高低阶不同尺度特征有效融合。但是传统的 FPN 仅含有自顶向下的结构,没有自下而上的结构,换言之就是低阶特征图含有高阶特征图的特征信息,而高阶特征图中不含有低阶特征图的特征信息。为了使高阶特征图包含经过 S 模块增强后的低阶特征信息,本文设计了 Conv Module 1 模块。将通过此模块处理后的特征信息与高阶特征图 (Feature map 3, Feature map 4, Feature map 5) 相加得到新的深层特征图,保证不同阶数的特征图含有特征信息的完整性。其中 Maxpool 层池化后的宽和高等于与之相加特征图的宽和高。 1×1 卷积核是为了调整通道数,同时也为了减少卷积操作需要的训练参数。 3×3 卷积核的卷积模式为“same”,即不改变特征图的尺寸大小。BN 和 Leaky-ReLU 的引入使得 Conv Module 1 具有一定的非线性和加快网络收敛速度的特性。Conv Module 1 具体如图 5(a)虚线框内的结构所示。

3.3 改进 SSD 的损失函数

针对模型训练过程中遥感图像正、负样本划分不平衡的问题,本文将使用聚焦分类损失函数对网络模型进行优化训练,公式具体表示为:

$$L(x, c, p, l, g) = \frac{1}{N} [L_{fl}(x, c, p) + aL_{loc}(x, l, g)] \quad (2)$$

式中: N 表示与真实框相匹配的候选框的个数, x 为输入图像, c 为目标类别, p 为预测类别概率, l 为候选框, g 代表真实的标签框, a 为两者的权重。 $L_{fl}(x, c, p)$ 和 $L_{loc}(x, l, g)$ 分别为聚焦分类损失和边界框回归损失,其中 $L_{loc}(x, l, g)$ 借鉴 SSD 算法中的边界框回归函数,表示为:

$$L_{loc}(x, l, g) = - \sum_{i \in Pos}^N \sum_{m \in \{cx, cy, w, h\}} x_{ij}^k \text{smooth}_{L1}(l_i^m - g_j^m) \quad (3)$$

式(3)中定位损失 L_{loc} 是候选框与真实标注框之间的 smooth_{L1} 损失。其中 g_j^m 为真实标注框的值, l_i^m 为候选框的值。

式(2)中的 $L_{fl}(x, c, p)$ 结合聚焦分类损失函数采用交叉熵计算损失,表示为:

$$L_{fl}(x, c, p) = - \sum_{i \in pos}^N x_{ij}^k p \ln(c_i^k) - \sum_{i \in neg} (1 - p) \ln(c_i^0) \quad (4)$$

式中: c_i^0 表示正确且类别为背景预测框的概率, c_i^k 为利用 Softmax 函数计算的概率值。

4 实验与结果分析

4.1 数据集与实验环境

本文采用西北工业大学开源的遥感图像数据集 NWPU VHR-10^[24] 进行飞机遥感图像目标检测实验,数据集包含尺寸大小不同的图片并且按照 7:3 的比例来随机划分训练集和验证集。依据数据集对目标面积大小划分的标准 (Small object: $\text{area} < 32^2$; Medium object: $32^2 < \text{area} < 96^2$; Big object: $\text{area} > 96^2$) 对采集的光学遥感数据集进行划分,由于一些飞机在小分辨率的情况下只有极少的像素个数,已很难通过人眼确定是否为飞机,因此,仅对像素个数超过 6 个的目标进行划分。图 6 为本文数据集的样本示例,面积类型从左到右分别为 Small object, Medium

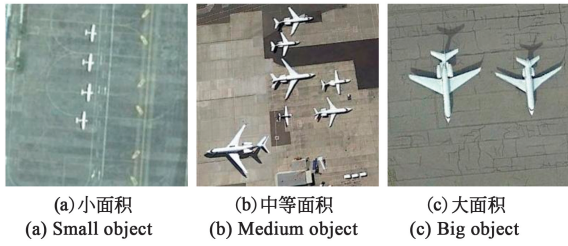


图 6 本文数据集的样本示例

Fig.6 Sample example of the data set in this article

object, Big object.

本文的实验环境选用 Pytorch 开源框架进行训练、测试,基于 ubuntu16.04 操作系统,CPU 为 Inter Corei9-9900K,3.60 GHz,显卡为 NVIDIA GTX2080。编译环境为 torch-1.6、torchvision-0.7,cuda10.1,python 3.6 语言编程。

4.2 评价标准

对于飞机这单一种类的遥感图像,本文采用精度均值 (Average precision, AP) 和每秒帧率 (Frames per second, FPS) 作为算法的评价标准。其中精度均值 (AP) 表示为 $AP = \int_0^1 p(r) dr$, 0 和 1 表示准确率 (Precision) 和召回率 (Recall) 绘制的曲线在坐标轴上的范围,准确率与召回率的定

义式如下:

$$\text{Precision} = \frac{TP}{TP + FP}, \quad (5)$$

$$\text{Recall} = \frac{TP}{TP + FN}, \quad (6)$$

式(5)和(6)中 TP 为正样本中的正例,FP 为负样本中的正例,FN 为负样本中的负例。此外,FPS 决定了算法的效率,其数值的大小反映了检测速度的快慢^[25]。

4.3 实验结果分析

为了初步评估改进 SSD 算法的性能,本文选取了验证集中 3 张具有代表性的图片,并使用几种经典目标检测算法和改进后的算法进行横向结果比较。比较结果如图 7 所示,其中 3 张具有代表性的图片如图 7(a)所示。第一张图片是飞机遥感目标存在于复杂环境背景下的情况;第二张图片是同时含有复杂环境背景和飞机遥感目标密集分布的情况;第三张图片是同时含有飞机遥感目标密集分布和飞机遥感目标尺度极小的情况。

由图 7 几种目标检测算法检测结果可知,对于图 7(a)中的第一张图片,除了文本改进的算法以外,其他目标检测算法均有不同程度的漏

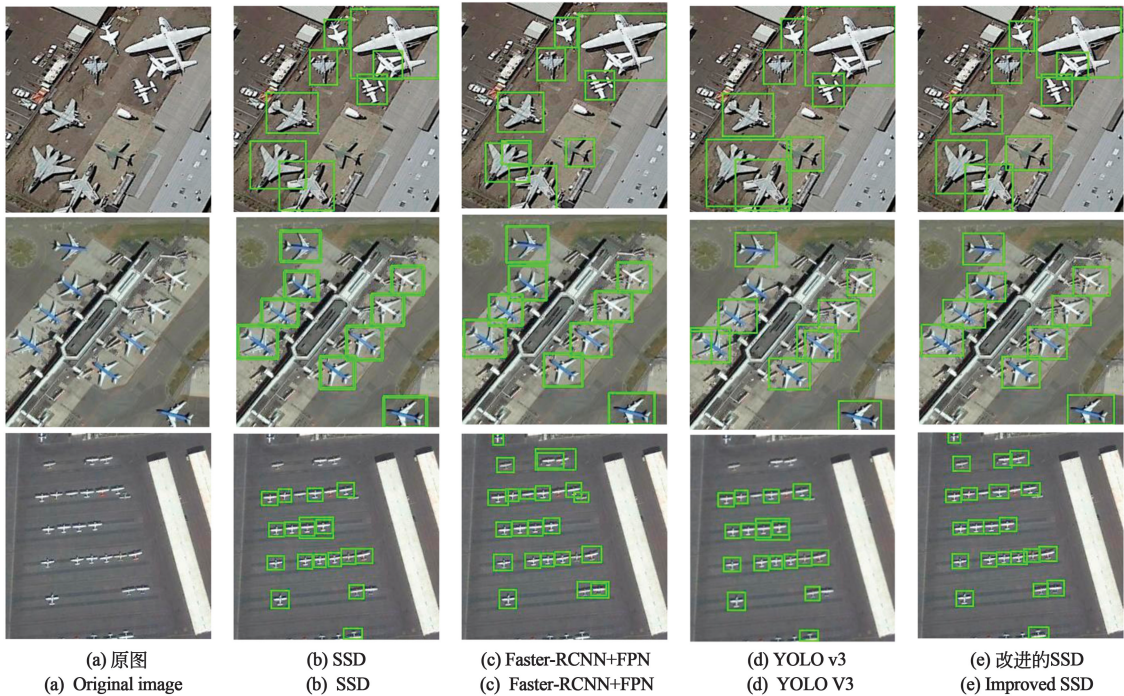


图 7 检测结果对比

Fig.7 Comparison of test results

检法以外,其他目标检测算法均有不同程度的漏检和误检发生。对于图 7(a)中的第二张图片,只有 YOLO v3 算法出现了误检的情况,其余 3 种算法检测结果基本一致,即都能正确识别出全部飞机目标。对于图 7(a)中的第三张图片,由于飞机遥感目标存在尺寸极小且密集的情况,包括本文算法在内,4 种算法均不能正确检测出全部目标。但是改进后的算法虽然没有全部检测出飞机目标,但检测结果明显要优于其他 3 种算法。对以上不同算法检测结果的差异分析可知,造成误检和漏检的主要原因是特征图之间缺少相应的联系,即浅层网络缺少深层网络重要的语义信息,深层网络缺少浅层网络丰富的细节信息。除此之外,特征图不同通道间的重要程度也缺少定量的权重系数。改进后的 SSD 算法综合考虑了各个特征图之间的关联性,设计了 S 模块和 Conv Module 1 模块。使得各层级的特征图携带更加完整的特征信息以便可以较好地识别出复杂环境背景下尺度大小不一的飞机遥感目标。另外,注意力机制网络 SENet 可以帮助网络学习特征图不同通道间的重要程度,在一定程度上提高算法的检测精度。

以上 3 张飞机遥感图片的检测结果多为极端条件下的检测结果,例如飞机遥感目标尺度极小且密集分布的极端条件。在更为实际的情况下,飞机遥感目标并不是只存在于极端条件中,对于生活中较容易接触到的典型场景也是飞机遥感目标检测的主要用地。本文列举了 4 种生活中含有飞机遥感目标的典型场景并进行相关的飞机目标检测实验。典型场景分别是民用机场(图 8(a))、包含住宅区背景的民用机场(图 8(b))、机场跑道(图 8(c))和包含其他交通工具背景(图 8(d))的场景。改进后的 SSD 在各个典型场景下的检测结果如图 8 所示。

场景 2 和场景 4 具有较为复杂的地面背景,可以简单归类为极端条件中的复杂背景一类。场景 2、场景 3 和场景 4 具有尺度较小和较为密集分布的特点,可以简单归类为极端条件中的密集分布和极小尺度一类。改进后的算法从不同特征层的特征信息的完整性的角度出发,增加了特征层彼此的关联性,一定程度上增加了飞机目标检测的鲁棒性。由图 8 的检测结果分析可知,改进后的算法不仅适用于极端条件下的飞机遥感目标



图 8 典型场景下的检测结果

Fig.8 Detection results in typical scenarios

的检测,对贴近生活中的典型场景也具有较好的检测效果。

为了进一步评估改进 SSD 算法的性能,本文使用 4.2 节的评价标准计算得到了该算法的检测精度与检测速度,并与其他几种目标检测算法进行横向结果对比^[26]。对比结果如表 2 和图 9 所示。

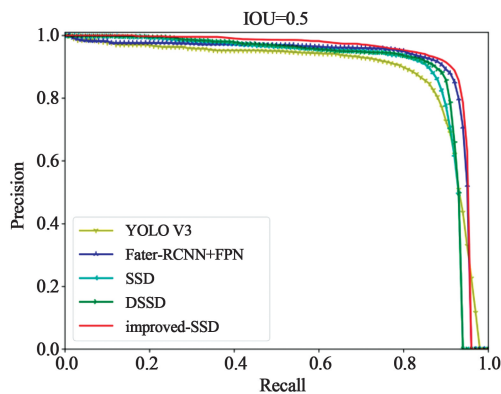
从表 2 的实验结果中可以分析到,本文提出的改进后的 SSD 算法在精度均值(AP)的指标上优于 YOLO、YOLO v2、YOLO v3、SSD、DSSD 和 Faster-RCNN+FPN 算法。从检测精度上看,当 $IOU=0.5$ 且 area 为 all 时(all 为包含所有面积类型的目标),改进后的 SSD 算法相比于 SSD 算法提高了 3.73%,相比于 Faster-RCNN + FPN、YOLO、YOLO v2、YOLO v3 和 DSSD 分别提高了 1.10%, 14.23%, 8.09%, 4.31%, 1.87%。当 $IOU=0.5:0.95$ 且 area=small 时,改进后的 SSD 算法的精度均值为 44.58%,较改进前提升了 4.93%。显然,改进后的 SSD 算法对小型目标的检测精度相比于其他算法提升幅度更为明显,证明改进后的 SSD 算法对小型目标检测更具有优势。从检测速度上看,改进 SSD 的检测速度和改进前相比基本没有太大的变化,和其他单阶段目标检测算法相比仍然可以满足实时检测的需求^[27]。

表 2 检测精度和检测速度

Tab.2 Detection accuracy and detection speed

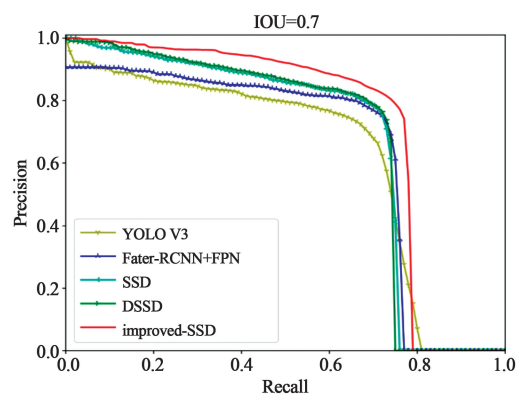
Method	Backbone	AP/ %		FPS/(frame · s ⁻¹)
		(IOU=0.5, area=all)	(IOU=0.5 : 0.95, area=Small)	
Faster-RCNN+FPN	VGG16	91.35	40.56	3.4
YOLO	GoogleNet	78.22	28.64	49.6
YOLO v2 ^[28]	DarkNet19	84.36	33.46	42.8
YOLO v3 ^[29]	DarkNet53	88.14	35.22	38.6
SSD	VGG16	88.72	39.65	41.3
DSSD ^[30]	ResNet101	90.58	41.25	20.5
改进 SSD	ResNet50	92.45	44.58	35.6

图 9 为 IOU=0.5 和 IOU=0.7 时各目标检测算法的 PR 曲线。从检测结果可以看出,改进后的 SSD 算法在准确率和召回率两个性能指标均优于其他算法,对于边界框的回归自然就越准确。



(a) 几种经典目标检测算法的 PR 曲线 (IOU=0.5)

(a) PR curves of several classic target detection algorithms



(b) 几种经典目标检测算法的 PR 曲线 (IOU=0.7)

(b) PR curves of several classic target detection algorithms

图 9 几种经典目标检测算法的 PR 曲线

Fig.9 PR curves of each algorithm

5 消融实验与目标检测结果

5.1 新型 FPN 的消融实验

为探究新型 FPN 与原始 FPN 对 AP 值的影响,本文设置 SSD、SSD+FPN 以及 SSD+FPN (新型)3 组实验来探究其影响。其余设置均与 SSD 算法相同。实验结果如表 3 所示,3 组相比于二组 AP 值提升了 2.62%,1.83%。

表 3 新型特征金字塔消融实验结果

Tab.3 New feature pyramid ablation experiment results

Method	AP/ %
	(IOU=0.5, area=all)
SSD	88.74
SSD+FPN	89.53
SSD+FPN (新型)	91.36

实验证明新型特征金字塔网络通过特征融合和注意力机制的方式有效弥补了不同特征图之间的信息空缺并选择性地激活和抑制特征图通道间的主要次要信息。既增强了低阶特征图的语义,也丰富了高阶特征图的细节,又对特征图内部不同位置特征施以差异性权重值,有效提升了改进后的算法对飞机遥感目标的检测精度^[31]。

5.2 骨架网络替换和聚焦分类损失函数的消融实验

为探究骨架网络替换和聚焦分类损失函数对 AP 值的影响,本文分别设 3 组模型对照,一组为

SSD+FPN(新型),二组为 SSD+FPN(新型)+Focal classification loss,三组为 SSD+FPN(新型)+Focal classification loss+ResNet50(改进

后)。其余设置均与 SSD 算法相同。实验结果如表 4 所示,第三组相比于一组和二组 AP 值提升了 1.09% 和 0.86%。

表 4 骨架网络替换与损失函数的消融实验结果

Tab.4 Skeleton network replacement and loss function ablation experiment results

Method	AP/% (IOU=0.5, area=all)
SSD+FPN(新型)	91.36
SSD+FPN(新型)+Focal classification loss	91.59
SSD+FPN(新型)+Focal classification loss+ResNet50(改进后)	92.45

实验证明,使用改进后的 ResNet50 作为算法的骨架网络相比 SSD 算法中的 VGG16 网络有着更好的特征提取能力。聚焦分类损失函数在一定程度上解决了网络训练时正负样本不均衡的问题,对检测精度的提高有一定的作用。

5.3 目标检测结果

改进后的 SSD 算法对飞机遥感图像数据集的部分检测结果如图 10 所示,在复杂背景(图 10(a))、小尺度且密集分布的遥感目标(图 10(b))、极小尺度且密集(图 10(c))和曝光背景(图 10(d))的情况下,该算法均可以精确检测出相关的

目标物体,证明了改进后的 SSD 在不同环境背景且目标面积类型为 all 时均有较好的检测效果。

6 结 论

对于 SSD 算法在飞机遥感图像领域目标检测效果不佳的问题,本文对 SSD 算法做出了一定的改进。使用改进后的 ResNet50 作为算法的骨架网络,与改进前的骨架网络相比有着更加优秀的特征提取能力。设计了新型特征金字塔网络,该网络用于不同特征图之间的特征融合,增加了特征图之间的相互联系,保证了不同特征图上含有特征信息的完整性。还使用了注意力机制监督网络的学习,在一定程度上提高了算法在复杂环境背景下进行飞机遥感目标检测的检测精度。使用了 Focal classification loss 作为改进后算法的损失函数,与改进前的损失函数相比,Focal classification loss 更加注重解决分类问题中数据类别不平衡以及判别难易程度差别的问题。

从实验结果来看,改进后的 SSD 对不同场景下且面积类型不同的飞机遥感目标的 AP 值为 92.45%,FPS 为 35.6。证明了改进后的算法在检测精度上要优于 SSD 算法,对比其他几种经典的目标检测算法,改进后的算法在检测精度和检测速度方面有着更好的平衡性。

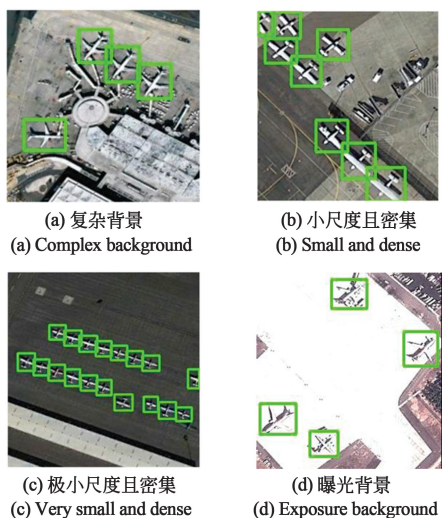


图 10 改进后的 SSD 在不同场景下的部分检测结果

Fig.10 Some test results of the improved SSD in different scenarios

参 考 文 献:

- [1] 张艳,社会娟,孙叶美,等.基于改进 SSD 算法的遥感图像目标检测[J].计算机工程,2021,47(9):252-258,265.

- ZHANG Y, DU H J, SUN Y M, *et al.* Object detection in remote sensing images based on improved SSD algorithm [J]. *Computer Engineering*, 2021, 47(9): 252-258, 265. (in Chinese)
- [2] 林娜,冯丽蓉,张小青.基于优化 Faster-RCNN 的遥感影像飞机检测[J].遥感技术与应用,2021,36(2):275-284.
LIN N, FENG L R, ZHANG X Q. Aircraft detection in remote sensing image based on optimized faster-RCNN [J]. *Remote Sensing Technology and Application*, 2021, 36(2): 275-284. (in Chinese)
- [3] 王浩桐,郭中华.锚框策略匹配的 SSD 飞机遥感图像目标检测[J/OL].计算机科学与探索,(2021-07-26)[2021-08-18].<http://kns.cnki.net/kcms/detail/11.5602.TP.20210726.1310.002.html>.
WANG H T, GUO Z H. Target detection of SSD aircraft remote sensing images based on anchor frame strategy matching [J/OL]. *Journal of Frontiers of Computer Science and Technology*, (2021-07-26)[2021-08-18]. <http://kns.cnki.net/kcms/detail/11.5602.TP.20210726.1310.002.html>. (in Chinese)
- [4] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [C]//*Proceedings of the 25th International Conference on Neural Information Processing Systems*. Lake Tahoe, Nevada: Curran Associates Inc., 2012: 1097-1105.
- [5] GIRSHICK R. Fast R-CNN [C]//2015 *IEEE International Conference on Computer Vision (ICCV)*. Santiago, Chile: IEEE, 2015: 1440-1448.
- [6] REN S Q, HE K M, 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*, 2017, 39(6): 1137-1149.
- [7] REDMON J, DIVVALA S, GIRSHICK R, *et al.* You only look once: unified, real-time object detection [C]//*Conference on Computer Vision and Pattern Recognition*. Las Vegas, USA: IEEE, 2016: 779-788.
- [8] LIU W, ANGUELOV D, ERHAN D, *et al.* SSD: single shot MultiBox detector [C]//14th *European Conference on Computer Vision*. Amsterdam, The Netherlands: Springer, 2016: 21-37.
- [9] 董旭彬,赵清华.改进 Mask R-CNN 在航空影像目标检测的研究应用[J].计算机工程与应用,2021,57(8):133-144.
DONG X B, ZHAO Q H. Research and application of improved Mask R-CNN in aerial image target detection[J]. *Computer Engineering and Applications*, 2021, 57(8): 133-144. (in Chinese)
- [10] 李靛,贺柏根,霍家道.FPN 在遥感图像检测中的应用[J].指挥控制与仿真,2019,41(2):130-136.
LI L, HE B G, HUO J D. Application of FPN in remote sensing image [J]. *Command Control & Simulation*, 2019, 41(2): 130-136. (in Chinese)
- [11] 叶子琦,王晓寅,张恒瑞.基于 Faster R-CNN 的遥感图像目标检测技术[J].电子世界,2020(8):63-64.
YE Z Q, WANG X Y, ZHANG H R. Remote sensing image target detection technology based on faster R-CNN [J]. *Electronics World*, 2020(8): 63-64. (in Chinese)
- [12] YU X, FU L R, DAI B S, *et al.* Soybean leaf morphology classification based on FPN-SSD and knowledge distillation [J]. *Journal of Northeast Agricultural University (English Edition)*, 2020, 27(4): 9-17.
- [13] 辛鹏,许悦雷,唐红,等.全卷积网络多层特征融合的飞机快速检测[J].光学学报,2018,38(3):0315003.
XIN P, XU Y L, TANG H, *et al.* Fast airplane detection based on multi-layer feature fusion of fully convolutional networks [J]. *Acta Optica Sinica*, 2018, 38(3): 0315003. (in Chinese)
- [14] LIN T Y, DOLLÁR P, GIRSHICK R, *et al.* Feature pyramid networks for object detection [C]//*Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, USA: IEEE, 2017: 936-944.
- [15] HE K M, ZHANG X Y, REN S Q, *et al.* Deep residual learning for image recognition [C]//*Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition*. Las Vegas, USA: IEEE, 2016: 770-778.
- [16] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [J]. *arXiv*: 1409.1556, 2014.
- [17] 李克文,李新宇.基于 SENet 改进的 Faster R-CNN 行人检测模型[J].计算机系统应用,2020,29(4):266-271.
LI K W, LI X Y. Pedestrian detection model based on improved faster R-CNN with SENet [J]. *Computer Systems & Applications*, 2020, 29(4): 266-271. (in Chinese)
- [18] LIN T Y, GOYAL P, GIRSHICK R, *et al.* Focal loss for dense object detection [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2020, 42(2): 318-327.
- [19] 邹慧海,侯进.基于改进 SSD 算法的道路小目标检测研究[J/OL].计算机工程,(2021-06-09)[2021-07-13].
<https://kns.cnki.net/kcms/detail/31.1289.tp.20210608.1800.003.html>.

- ZOU H H, HOU J. Research on road small target detection based on improved SSD algorithm [J/OL]. *Computer Engineering*, (2021-06-09)[2021-07-13]. <https://kns.cnki.net/kcms/detail/31.1289.tp.20210608.1800.003.html>. (in Chinese)
- [20] 刘建伟,赵会丹,罗雄麟,等.深度学习批归一化及其相关算法研究进展[J].自动化学报,2020,46(6):1090-1120.
LIU J W, ZHAO H D, LUO X L, *et al.* Research progress on batch normalization of deep learning and its related algorithms [J]. *Acta Automatica Sinica*, 2020, 46(6): 1090-1120. (in Chinese)
- [21] XU J, LI Z S, DU B W, *et al.* Reluplex made more practical: leaky ReLU [C]//*Proceedings of the 2020 IEEE Symposium on Computers and Communications (ISCC)*. Rennes, France: IEEE, 2020: 1-7.
- [22] 邢永鑫,孙游东,王天一.基于改进 SSD 算法对奶牛的个体识别[J/OL].计算机工程与应用,(2020-12-09)[2021-08-04]. <http://kns.cnki.net/kcms/detail/11.2127.tp.20201208.1537.024.html>.
XING Y X, SUN Y D, WANG T Y. Individual recognition of dairy cow based on improved SSD algorithm [J/OL]. *Computer Engineering and Applications*, (2020-12-09)[2021-08-04]. <http://kns.cnki.net/kcms/detail/11.2127.tp.20201208.1537.024.html>. (in Chinese)
- [23] 胡志伟,杨华,娄甜田.采用双重注意力特征金字塔网络检测群养生猪[J].农业工程学报,2021,37(5):166-174.
HU Z W, YANG H, LOU T T. Instance detection of group breeding pigs using a pyramid network with dual attention feature [J]. *Transactions of the Chinese Society of Agricultural Engineering*, 2021, 37(5): 166-174. (in Chinese)
- [24] 西北工业大学.NWPU VHR-10[EB/OL].(2014-04-24)[2021-09-09].<https://hyper.ai/datasets/5422>.
Northwestern Polytechnical University. NWPU VHR-10 [EB/OL]. (2014-04-24)[2021-09-09]. <https://hyper.ai/datasets/5422>. (in Chinese)
- [25] 李娜,姜志,王军,等.基于 Faster R-CNN 的仪表识别方法[J].液晶与显示,2020,35(12):1291-1298.
LI N, JIANG Z, WANG J, *et al.* Instrument recognition method based on Faster R-CNN [J]. *Chinese Journal of Liquid Crystals and Displays*, 2020, 35(12): 1291-1298. (in Chinese)
- [26] 唐悦,吴戈,朴燕.改进的 GDT-YOLOV3 目标检测算法[J].液晶与显示,2020,35(8):852-860.
TANG Y, WU G, PIAO Y. Improved algorithm of GDT-YOLOV3 image target detection [J]. *Chinese Journal of Liquid Crystals and Displays*, 2020, 35(8): 852-860. (in Chinese)
- [27] 任彬,王宇庆,丛振,等.基于 MPSOC 的航空图像目标检测系统设计[J].液晶与显示,2021,36(7):1006-1017.
REN B, WANG Y Q, CONG Z, *et al.* Design of aerial image target detection system based on MPSOC [J]. *Chinese Journal of Liquid Crystals and Displays*, 2021, 36(7): 1006-1017. (in Chinese)
- [28] REDMON J, FARHADI A. YOLO9000: better, faster, stronger [C]//*Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition*. Honolulu, USA: IEEE, 2017: 6517-6525.
- [29] REDMON J, FARHADI A. YOLOv3: an incremental improvement [J]. *arXiv*: 1804.02767, 2018.
- [30] FU C Y, LIU W, RANGA A, *et al.* DSSD: deconvolutional single shot detector [J]. *arXiv*: 1701.06659, 2017.
- [31] 李明山,韩清鹏,张天宇,等.改进 SSD 的安全帽检测方法[J].计算机工程与应用,2021,57(8):192-197.
LI M S, HAN Q P, ZHANG T Y, *et al.* Safety helmet detection method of improved SSD [J]. *Computer Engineering and Applications*, 2021, 57(8): 192-197. (in Chinese)

作者简介:



王浩桐(1995—),男,宁夏青铜峡人,硕士研究生,2018年于重庆邮电大学获得学士学位,主要从事计算机视觉方面的研究。E-mail: 562287297@qq.com



郭中华(1973—),男,宁夏银川人,博士,教授,2008年于西北工业大学获得博士学位,主要从事图像处理、机器视觉方面的研究。E-mail: guozhh@nxu.edu.cn