

基于改进 YoloX 的输电通道工程车辆检测识别

张智坚¹, 曹雪虹¹, 焦良葆¹, 孟琳^{1,2}, 邹辉军¹

(1. 南京工程学院 人工智能产业技术研究院, 南京 211167;

2. 江苏省智能感知技术与装备工程研究中心, 南京 211167)

摘要: 针对输电通道下环境复杂, 各类工程车辆频繁损坏输电线路中所需解决的对工程车辆的检测识别问题, 在单阶段目标检测算法 YoloX 的基础上, 对 YoloX 算法中的损失函数进行修改, 平衡正负样本和难样本, 在网络中添加 CBAM 注意力机制, 将内部通道信息和位置信息结合, 提高特征的提取能力, 并通过修改强特征提取部分 Neck 中的 CspLayer 结构, 在保证检测速度的前提下, 提高模型的检测性能; 通过筛选亮度低的图片, 引入改进的 MSR 算法对图片进行亮度提升, 优化数据集; 实验结果表明, 提出的算法提高了检测的准确率, 与传统的 YoloX 算法相比, mAP 提高了 4.64%, 识别效果明显提升, 证明了新算法的有效性。

关键词: 目标检测; 工程车辆; YoloX; 注意力机制; MSR

Detection and Recognition of Transmission Channel Engineering Vehicles Based on Improved YoloX

ZHANG ZhiJian¹, CAO XueHong¹, JIAO LiangBao¹, MENG Lin¹, ZOU HuiJun¹

(1. AI Industrial Technology Research Institute, Nanjing Institute of Technology, Nanjing 211167, China)

(2. Jiangsu intelligent perception technology and equipment Engineering Research Center, Nanjing 211167, China)

Abstract: In view of the complex environment under the transmission channel and the frequent damage of various engineering vehicles on the transmission line, the problem of detection and identification of engineering vehicles needs to be solved. Based on the single-stage target detection algorithm YoloX, the loss function in the YoloX algorithm is modified to balance the positive or negative samples and difficult or easy samples, the CBAM attention mechanism is added in the network, the internal channel information and location information are combined to improve the feature extraction ability, and the CspLayer structure in the strong feature extraction part Neck is modified to improve the detection performance of the model under the premise of ensuring the detection speed. By screening the pictures with low brightness, the improved multi-scale retinex (MSR) algorithm is introduced to improve the brightness of the pictures and optimize the data set. Experimental results show that the proposed algorithm improves the detection accuracy. Compared with the traditional YoloX algorithm, the accuracy of the mAP is improved by 4.64%, and the recognition effect is significantly improved, which proves the effectiveness of the new algorithm.

Keywords: target detection; engineering vehicles; YoloX; attention mechanism; MSR

0 引言

随着我国经济的飞速发展, 城镇经济基础建设的不断加快, 大型工程器械得到了广泛的使用, 导致高压输电线路原有的安全通道环境被挤压。输电线路设备发生外力破坏的事件逐渐上升, 已经成为目前我国电网输电线路安全运行的主要隐患, 而单纯依靠电力巡检人员进行定期检查, 不仅效率低下, 而且更受到自然方面的影响^[1]。因此, 采用有效的目标检测方法对输电通道下的工程车辆进行精准的检测具有重大的意义。

在利用深度学习进行目标检测之前, 传统的目标检测主要是采用滑动窗口的方式。首先寻找图像中感兴趣的区域, 即选取可能包含检测对象的区域。然后对可能包含检测对象的区域进行特征提取, 例如尺度不变特征变换 (SIFT, scale-invariant feature transform) 等特征提取算法^[2]。最后利用支持向量机 (SVM, support vector machine) 等算法对提取好的特征进行检测分类^[3]。但是, 图片中物体尺度不一, 所占图片面积有大有小, 当利用一个固定的窗口进行滑动时, 物体所占图片面积越小, 候选框区域中背景区域的占比就越大; 物体所占图片面积越大,

收稿日期: 2022-03-09; 修回日期: 2022-04-11。

基金项目: 国家自然科学基金青年基金项目(61903183)。

作者简介: 张智坚(1997-), 男, 江苏扬州人, 硕士研究生, 主要从事智能电网及信息技术方向的研究。

曹雪虹(1964-), 女, 江苏苏州人, 教授, 硕士研究生导师, 主要从事无线通信系统、信息理论方向的研究。

焦良葆(1972-), 男, 安徽安庆人, 教授, 硕士研究生导师, 主要从事图像信号处理、视觉信息理解方向的研究。

引用格式: 张智坚, 曹雪虹, 焦良葆, 等. 基于改进 YoloX 的输电通道工程车辆检测识别[J]. 计算机测量与控制, 2022, 30(9): 67-73.

候选框区域只能包含物体的局部。因此在对候选区域中的物体进行特征提取时,一方面分类的时候可能产生误分类,另一方面可能对于同一个对象产生多个正确识别的结果。若设计各种尺寸的窗口,这样就会产生大量的计算,无法满足日趋复杂的输电通道场景的需求。

近年来,由于卷积神经网络在计算机视觉领域中得到了广泛的应用,给以深度学习为基础的目标检测算法提供了快速发展的机遇。基于深度学习的目标检测算法主要分为两类,第一类是基于候选区域的双阶段目标检测算法,第二类是基于回归问题的单阶段目标检测算法^[4]。双阶段的目标检测算法主要有 Fast R-CNN、Faster R-CNN、R-FCN 等,首先由特定的算法生成一系列待识别对象的候选框,再通过卷积神经网络对待识别对象进行分类。Girshick^[5]等提出 Fast R-CNN,利用卷积网络进行特征的提取分类以及边界框的回归,通过感兴趣池化层进行不同尺寸感兴趣区域的统一。Ren^[6]等提出 Faster R-CNN,使用区域预测网络(RPN)替代传统特征提取方法,提高了网络的训练速度。Dai^[7]等提出 R-FCN 方法,应用全卷积网络提升检测速度,通过位敏池化层,解决因图片自身变化所引起的属性改变问题。这类算法虽然准确度较高,但是计算量较大,导致检测速度较慢。单阶段的目标检测算法主要有 Yolo, SSD, Retina-Net 等,不需要产生候选框,直接对预测的目标物体进行回归。谢俊章^[8]等利用自适应激活函数改进 YOLOv4 的残差结构,提升了小型密集分布的车辆检测精度,但是模型结构复杂。于硕^[9]等利用 YOLOv5 网络,通过数据增强以及网络改进对小目标进行识别,提高了定位精度。Raj^[10]等以 SSD 目标检测算法为基础,应用 Canny 算子和 log 滤波器进行图像增强,提高了夜间车辆的检测精度。宋欢欢^[11]等利用 Retina-Net 算法,结合迁移学习,并结合模型压缩,对光照变化明显,车辆遮挡严重导致的检测精度下降问题有较大改善,但能识别车辆种类少。

综上,当前基于深度学习的目标检测算法大多数对样本的质量要求较高,一旦受到时间和气候的影响,如傍晚或者阴天光线不足的情况,会极大影响检测效果。此外,在保证检测速度的同时,往往精度会有大幅下降。

为了有效解决上述问题,本文通过对光线较暗的图片进行筛选,对 MSR 算法进行优化,提高图片亮度。以 YOLOX 算法为基础对工程车辆进行目标检测,通过修改网络的损失函数来平衡正负样本以及难易样本,通过添加 CBAM 注意力机制,结合内部通道信息和位置信息,提高 YOLOX 算法的提取能力,并通过修改 Neck 中 CspLayer 层,在保证检测速度的前提下,提高模型的检测性能。

1 YOLOX 网络介绍

1.1 YOLOX 网络结构

YOLOX 的具体结构如图 1 所示。

YOLOX 的网络结构主要由 4 个部分构成,分别为输入端、

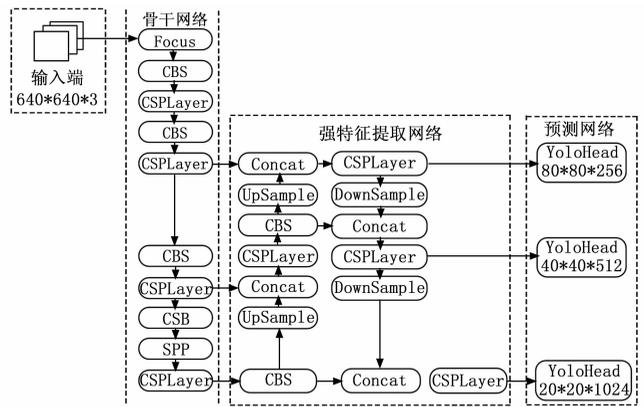


图 1 YOLOX 网络结构

骨干网络 Backbone、强特征提取网络 Neck 以及预测网络 Prediction 这 4 个部分构成。输入端采用 Mosaic 数据增强方法的同时,融合了 Mixup 数据增强方法,通过随机缩放、随机裁剪、随机排布、不同图片加权融合等方式对图片进行拼接,丰富了数据的多样性,可提升小目标的检测效果。主干网络 Backbone 中使用了 Focus、CSPLayer 等结构,Focus 结构通过对图片进行切片操作,将图片中每隔一个像素点取一个值,从而将宽高信息都集中到了通道空间,防止信息的丢失。CSPLayer 结构主要由组件 CBS(卷积层+归一化层+激活函数)和残差组件(由两个 CBS 结构构成)构成,通过对网络的深度进一步加深,来不断提高网络的特征提取和特征融合能力。在网络的 Neck 部分,通过利用特征金字塔 FPN+PAN 结构,将强语义特征和强定位特征进行融合,其中 FPN 结构以自顶向下的方式将深层的语义特征传到浅层,来增强多个尺度上的语义表达,而 PAN 结构则将浅层的定位信息传递到深层,来增强在多个尺度上的定位能力。此外,使用 CSPLayer 结构来进一步加强网络特征融合。在最后的 Prediction 部分,通过无锚点方式(Anchor-free)减少预测结果数量,完成初步的筛选,再利用 SimOTA 算法对预测结果进行精细化筛选,得到最终的预测结果。

1.2 改进策略

本文主要利用 YOLOX 算法研究工程车辆的检测,通过在网络中添加注意力机制,修改 CSPLayer 以及损失函数,对 YOLOX 算法的网络结构进行改进,使 YOLOX 算法能更适用于工程车辆的检测领域。

1.2.1 添加注意力机制 CBAM

Woo S^[12]等提出了一种轻量的注意力模块 Convolutional Block Attention Module,简称 CBAM,具体结构如图 2 所示。该模块主要由两个独立的子模块构成,分别是通道注意力模块(Channel Attention Module)和空间注意力模块(Spatial Attention Module),能够沿着空间和通道两个方向依次推断出注意力权重,然后将权重与原始特征图相乘,从而对特征进行自适应调整,能够更关注对目标物体的识别。

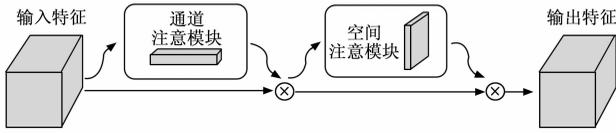


图 2 CBAM 结构图

CBAM 模块的实现主要分为两步, 第一步是经过 CA 模块, 该模块主要关注输入图片中有意义的信息, 压缩空间维度, 通道维度不变, 具体结构如图 3 所示。首先将输入的特征图 F 分别进行一个空间的全局最大池化 (Global Max Pooling) 和全局平均池化 (Global Average Pooling), 得到两个 $1 \times 1 \times C$ 的特征图。接下来将得到的两个特征图输入到一个两层的神经网络 (MLP), 第一层中神经元个数为 C/r (r 为减少率), 使用的是 Relu 激活函数, 第二层神经元个数为 C , 输出两个激活后的特征。最后, 将两个特征相加并利用 Sigmoid 激活函数, 获得一个权重系数 M_c , 将输入特征 F 与权重系数 M_c 相乘得到最后缩放后的新特征图 F' 。

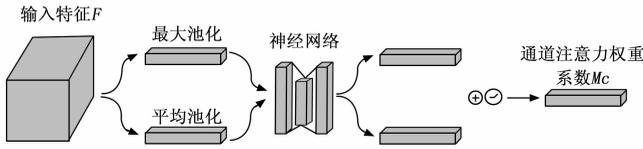


图 3 CA 模块结构图

第二步是经过 SA 模块, 该模块主要是关注对象的位置信息, 压缩通道维度, 空间维度不变, 具体结构如图 4 所示。首先利用一个基于通道的全局最大池化 (Global Max Pooling) 和全局平均池化 (Global Average Pooling) 对 CA 模块中输出的特征图 F' 进行处理, 得到两个 $H \times W \times 1$ 的特征图。接下来对这两个特征图基于通道进行一个拼接, 利用一个 7×7 卷积操作并使用 Sigmoid 激活函数, 获得一个权重系数 M_s 。最后将输入的特征 F' 与权重系数 M_c “相乘, 得到最终缩放后的新特征。

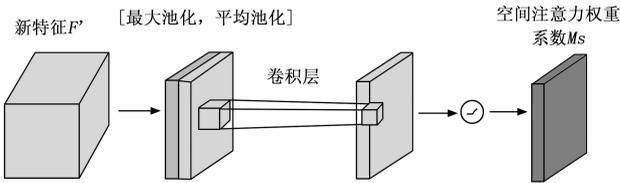


图 4 SA 模块结构图

通过以上两个步骤, CBAM 模块沿着两个独立的维度完成了通道信息和空间信息的融合, 完成自适应的特征优化。相比于注意力机制 SE 模块^[13]只考虑内部通道信息而忽略了位置信息的重要性, CBAM 模块结合了空间 (spatial) 和通道 (channel) 的注意力机制模块, 可以取得更好的效果。将此模块添加在 YoloX 网络中 Backbone 部分中 $20 \times 20 \times 1024$ 的 CSPLayer 层输出位置, 可提高 YoloX 算法对工程车辆的检测效果。

1.2.2 Neck 部分 CSPLayer 改进

在 YoloX 网络中, 骨干网络 Backbone 主要是对输入的

图片进行特征提取, 而 Neck 部分介于 Backbone 和 Prediction 之间, 主要是为了更好的提取 backbone 中的特征, 对接下来 Prediction 部分的定位和分类的结果有很大的影响^[14]。强特征提取网络 Neck 部分中的 CSPLayer 结构如图 5 所示, 通过利用残差结构, 对原来残差块的堆叠进行一个拆分, 一共分成左右两个部分, 其中主干部分继续对原来的残差块进行堆叠, 另一部分则经过少量处理直接连接到最后。图 5 中 (a) 先将两个部分进行连接, 然后再进行输入到 Partial Transition 层中, 采用这种做法会使得大量的梯度信息被重用, 有利于网络学习, (b) 中先将部分特征输入 Partial Transition 层中, 然后再进行连接, 这样做损失了部分的梯度重用, 但是计算复杂度比 (b) 图低。

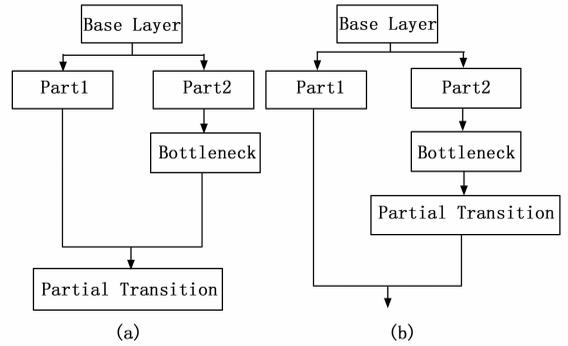


图 5 原始 CSPLayer

改进的 CSPLayer 通过将图 5 中 (a) 和 (b) 的结构进行结合, 具体结构如图 6 所示, 可以进一步提升学习能力, 但是也提高了一些计算复杂度, 所以需要进一步将 CSPLayer 中的传统卷积替换成深度可分离卷积, 减小计算复杂度。

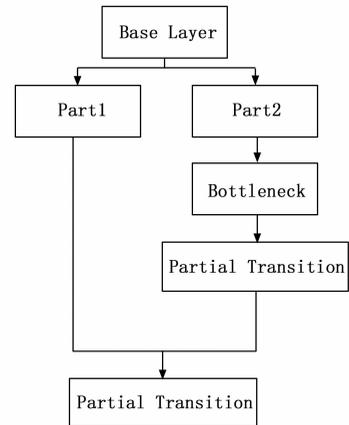


图 6 改进版 CSPLayer

对于传统卷积而言, 输入一个大小为 $D_F \times D_F \times M$ 的特征图, 输出一个大小为 $D_F \times D_F \times N$ 特征图, 其中卷积核的大小为 $D_K \times D_K \times M$, 则传统卷积的计算量 C 如公式 (1) 所示, 其中 $D_F \times D_F$ 为特征图的长和宽, $D_K \times D_K$ 为卷积核的长和宽, M 和 N 为通道数。

$$C = D_K \times D_K \times M \times N \times D_F \times D_F \quad (1)$$

而深度可分离卷积可以将卷积分为两个过程, 首先使用大小为 $D_K * D_K * 1$ 卷积核进行卷积, 这样每一个卷积核就会产生 M 个 $D_F * D_F$ 结果, 第二步使用 $1 * 1 * M$ 的卷积核, 这样就只得到一个结果, 一共有 N 个卷积核, 最终就得到了 $D_F * D_F * N$ 的结果, 深度可分离卷积的计算量 S 如公式 (2) 所示:

$$S = D_K \times D_K \times M \times D_F \times D_F + M \times N \times D_F \times D_F \quad (2)$$

通过将深度可分离卷积的计算量和传统卷积的计算量进行对比, 比值 K 如公式 (3) 所示, 计算量比传统的卷积核小很多, 卷积核个数 N 越大, 计算量减少的就更加明显。

$$K =$$

$$\frac{D_K \times D_K \times M \times D_F \times D_F + M \times N \times D_F \times D_F}{D_K \times D_K \times M \times N \times D_F \times D_F} = \frac{1}{N} + \frac{1}{D_K^2} \quad (3)$$

1.2.3 改善 obj 分支的损失函数

在 YoloX 网络中, Prediction 部分的 YoloHead 分为 3 个分支, 分别为 Cls 分支、Reg 分支以及 Obj 分支。Cls 分支主要用于判断每一个特征点所包含的物体种类, Reg 分支主要用于判断每一个特征点的回归参数, 回归参数调整后可以获得预测框, Obj 分支主要用于判断每一个特征点是否包含物体, 具体结构如图 7 所示。

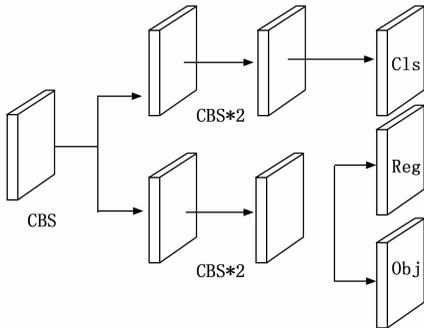


图 7 YoloHead 结构图

原始网络中 obj 分支采用 BCELoss 损失函数^[15], 具体如公式 (4) 所示。其中 M 为样本的类别数量, y_{ic} 代表符号函数 (即如果样本 i 的真实类别等于类别 c 则取值为 1, 否则取值为 0), p_{ic} 为观测样本 i 属于类别 c 的预测概率。BCELoss 损失函数虽然能解决梯度消失问题, 平衡正负样本, 但是对于难易样本不均衡的情况处理效果不佳, 而困难样本对网络的训练结果有一定的影响。

$$L = \frac{1}{N} \sum_{i=1}^N L_i = -\frac{1}{N} \sum_{i=1}^N \sum_{c=1}^M y_{ic} \log(p_{ic}) \quad (4)$$

Focal Loss 损失函数是基于二分类交叉熵损失的改进^[16], 既能平衡正负样本、也能平衡难易样本, 具体如公式 (5) 所示, 其中 α 用来调节正负样本不平衡, γ 用来调节难易样本不平衡, $y=1$ 表示正样本, $y=0$ 表示负样本。通过降低简单样本对于 loss 的影响、更多关注难样本, 避免大量简单负样本的 loss 在训练过程中占据 loss 的主导地位。此外通过利用调质因子 $(1-p)^\gamma$, 对于正样本来说,

预测值 p 趋向于 1, 说明是简单样本, $(1-p)^\gamma$ 就会趋近于 0, 损失函数值就会变小, 而预测值 p 趋向于 0, 说明是困难样本, $(1-p)^\gamma$ 就会趋近于 1, 损失函数值就会变大。所以这样对于简单样本而言, 权重就会降低, 可以减少简单样本的影响, 对于困难样本, 权重就会增大, 从而增加对困难样本的关注度, 让网络更倾向于使用困难样本进行参数的更新。

$$L = -\alpha(1-p)^\gamma \log(p) \quad y = 1$$

$$L = -(1-\alpha)p^\gamma \log(1-p) \quad y = 0 \quad (5)$$

2 Retinex 算法

2.1 常见 Retinex 算法

Retinex 是一种建立在科学实验和科学分析基础上的图像增强方法^[17], 主要包括单尺度的 SSR 算法 (Single Scale Retinex) 以及多尺度的 MSR 算法 (Multi-Scale Retinex) 等, 可以对不同类型的图片在动态范围类压缩、颜色等方面进行自适应的增强。

单尺度的 SSR 算法主要通过从原始图片中去除低频照射部分, 保留原始图片中的高频分量来进行图像增强^[18]。光照图像可以分为其入射图像和反射图像, 具体如公式 (6) 所示, 入射图像主要决定了其像素能达到的动态范围, 需要去除, 而反射图像是其本身的反射特性, 基本不会发生变化, 从而需要保留, 具体如公式 (7) 所示:

$$r(x, y) = \log R(x, y) = \log \frac{S(x, y)}{L(x, y)} \quad (6)$$

$$r(x, y) = \log S(x, y) - \log [F(x, y) \otimes S(x, y)] \quad (7)$$

$S(x, y)$ 表示原始图像, $r(x, y)$ 表示输出图像, $R(x, y)$ 表示反射图像, $L(x, y)$ 表示入射图像, $F(x, y)$ 表示中心环绕函数, 具体如公式 (8) 所示。其中 C 是高斯环绕尺度, C 越小, 动态压缩能力越强, 图像阴暗部分的细节可以得到更好的增强, C 越大, 颜色保真度越高, λ 是一个尺度, 需要满足的条件如公式 (9) 所示。中心环绕函数 $F(x, y)$ 通过使用低通函数, 对入射图像对应原始图像的低频部分进行估计, 并从原始图像中除去低频照射部分, 最终就会留下原始图像所对应的高频分量^[19]。

$$F(x, y) = \lambda e^{-(x^2+y^2)/C^2} \quad (8)$$

$$\iint F(x, y) dx dy = 1 \quad (9)$$

多尺度的 MSR 算法是在 SSR 算法的基础上发展来的, 可以同时保持图像高保真度与对图像的动态范围进行压缩的同时, 实现色彩增强、颜色恒常性、局部动态范围压缩、全局动态范围压缩^[20]。MSR 计算过程如公式 (10) 所示, 其中 K 是高斯中心环绕函数的个数。为了能够保持像 SSR 算法一样同时拥有高、中、低 3 个尺度的优点, 设置 K 的值为 3。因此 MSR 算法能较好的解决色彩和动态压缩范围的平衡问题, 但是还会出现边缘锐化不足等问题。

$$r(x, y) = \sum_k^K w_k \{ \log S(x, y) - \log [F_k(x, y) \cdot S(x, y)] \} \quad (10)$$

2.2 改进 MSR 算法

传统的 MSR 算法在提升图片亮度的同时, 会导致图片中对象的轮廓模糊^[21], 而在 YoloX 网络中, 图像中对象的轮廓对识别的效果影响很大, 为了更有效的增强整个图片, 提出一种基于边缘增强的多尺度 Retinex 算法 AttentionMSR-MSS。

由于图片中对象的轮廓模糊, 所以需要轮廓进行边缘强化, 引入拉普拉斯算子, 将拉普拉斯模板与原始图像进行卷积操作, 拉普拉斯模板如公式 (11) 所示, 通过将中心像素的灰度值与邻域内其他像素的平均灰度值进行对比, 如果中心像素的灰度值低于平均灰度值时, 就降低中心像素的灰度值, 如果中心像素的灰度值高于平均灰度值时, 就提高中心像素的灰度值, 然后将原始图像减去卷积结果得到图像边缘图像, 其过程如公式 (12) 所示, 最后将原始图片与边缘图像进行叠加, 获得边缘增强后的图片, 这样就可以保留图像背景的同时, 增强对象的轮廓细节信息。

$$K = \begin{pmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{pmatrix} \quad (11)$$

$$f(x, y) = S(x, y) - S(x, y) * K \quad (12)$$

由于 MSR 算法是利用去除低频部分, 保留高频部分来进行图像增强, 为了更可靠地增强整个图像, 打破它在增强局部细节方面的局限性, 将 AttentionMSR 算法与 MSS 算法^[22]进行结合, 通过在经过边缘增强的图像 f 中加入高频成分, 来扩展动态范围以提高局部可见性。首先将高斯核应用于经过边缘增强的图像 f , 获得处理后图像 B_1 、 B_2 和 B_3 , 具体如公式 (13) 所示, 其中 G_1 、 G_2 和 G_3 是高斯核。然后和原图做减法, 获得不同程度的细节信息 D_1 、 D_2 和 D_3 , 具体如公式 (14) 所示。再通过一定的组合方式把这些细节信息融合, 具体如公式 (15) 所示, 其中 w_1 、 w_2 和 w_3 分别为 0.5、0.5 和 0.25。最后, 将整体细节 D 添加到经过边缘增强的图像 f 中。

$$B_1 = G_1 * f \quad B_2 = G_2 * f \quad B_3 = G_3 * f \quad (13)$$

$$D_1 = f - B_1 \quad D_2 = f - B_2 \quad D_3 = f - B_3 \quad (14)$$

$$D = (1 - W_1 * \text{sgn}(D_1)) * D_1 + W_2 * D_2 + W_3 * D_3 \quad (15)$$

通过先对图片进行边缘强化, 然后在此基础上加入高频成分, 增强后的图像亮度明显提高, 图像细节信息也得到了提升, 图像边缘更加清晰, 具体对比图如图 8 所示, 其中 (a) 为原图, (b) 为使用 AttentionMSR-MSS 算法的效果图。



(a) 原图 (b) 改进后算法的效果图

图 8 效果对比图

3 实验

3.1 实验平台及参数配置

本次实验的实验平台配置如表 1 所示。

表 1 参数配置

软硬件平台	型号参数
操作系统	Ubuntu18.04
CPU	Intel(R) Xeon(R) Glod5118 CPU@2.3 GHz
GPU	GeForce RTX2080Ti
显存	24 G
内存	128 G
框架	Pytorch
编程环境	Python

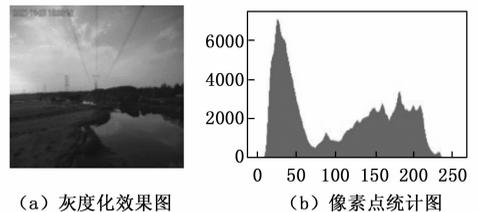
3.2 数据集准备

工程车辆的数据原始图是由某市公司提供的, 总计 12 303 张图片, 其中吊车的图片共有 5 184 张, 挖掘机的图片共有 4 893 张, 铲车的图片共有 710 张, 混凝土浇灌车的图片共有 604 张, 混凝土搅拌车的图片共有 597 张, 压路机共有 369 张, 其数据集类别数量如表 2 所示。

表 2 数据集类别及数量

类别名称	训练集	测试集	合计
crane	3 756	1 428	5 184
concrete_pour_truck	326	278	604
excavator	3 719	1 174	4 893
forklift	458	252	710
mixer_truck	421	176	597
roller	242	127	369

由于时间和气候的影响, 会存在傍晚或者阴天光线不足的情况, 所以需要对所有图片进行预处理, 提高图片的亮度。首先, 对输入图片进行灰度化处理, 通过设置灰度阈值规定 0~39 为偏暗像素点, 统计偏暗的像素个数, 统计过程如图 9 所示, 其中 (a) 图为灰度化后的图片, (b) 图为像素点统计图。然后计算偏暗像素点占图片总像素点的百分比 P , 根据实际情况设定 P 超过 20% 即表示图片偏暗。最后将 P 高于 20% 的图片筛选出来, 并通过改进版 MSR 算法进行亮度调节。



(a) 灰度化效果图 (b) 像素点统计图

图 9 像素点统计图

对于进行预处理后的图片, 采用 LabelImg 软件进行标注。操作界面如图 10 所示, 按照对应标签进行标注, 此标注为 crane, 标注完的信息以 xml 的格式保存在相同路

径下。



图 10 LabelImg 操作界面

3.3 实验结果

3.3.1 不同模块效果分析

为了更加直观的对比不同模块对检测效果的影响，本文通过消融实验，依次减少改进版 CspLayer 层、Focal loss 损失函数以及 CBAM 模块对改进版 YoloX 算法进行效果验证，实验结果如表 3 所示。

表 3 不同模块效果分析

CBAM	Focal loss	改进版 CspLayer	mAP/%
			92.12
✓	✓		93.28
✓		✓	96.23
	✓	✓	94.43
✓	✓	✓	96.76

由表可以看出，当所有模块全都加入时，对工程车辆检测效果最好，mAP 值比传统 YoloX 算法提升了 4.64%。与此同时，实验结果表明，将 CBAM 注意力机制与改进版 CspLayer 层结合对算法检测性能的提升起到了最为关键的作用。通过替换原始 obj 分支的损失函数，也能带来 0.53% 的 mAP 值的提升。

3.3.2 不同模型效果分析

本文通过将 YoloX 算法与其他目标检测算法进行对比，具体结果如表 4 所示。为了保证实验对比的有效性，模型的训练和测试都安排在同一台 GPU 服务器上进行操作，并且每次训练的过程保持一致。训练过程中设置的参数如下，输入的图片尺寸 Img-Size 统一缩放到 640×640，一次输入网络中的图片张数 Batch-Size 设置为 64，对于整个训练集的数据，迭代次数 Epochs 设置为 200 轮。测试时，设置输入的图片尺寸为 640，置信度 IOU 设置为 0.6。通过结果对比可知，YoloX 算法的检测效果最好，因此选择 YoloX 算法作为目标检测的基础算法，并对其进行改。

首先，将改进版的 YoloX 算法部署在 GPU 服务器中，并对 COCO 数据集进行训练，利用训练好的权重对 COCO 数据集进行检测，检测效果如图 11 所示，其中图 (a) 为原始 YoloX 算法的测试结果，图 (b) 为改进后 YoloX 算法的测试结果，可见原模型中未被识别的对象，改进后的模型

可以识别出来。



(a) 原始 YoloX 算法检测效果图 (b) 改进后 YoloX 算法检测效果图

图 11 COCO 数据集检测效果图

然后利用改进版的 YoloX 算法检验自建数据集的识别效果，通过训练自建工程车辆的数据集，获得最终工程车辆检测模型权重。改进后的 YoloX 检测网络相较于原 YoloX 检测网络模型 mAP 提高 4.64%，具体如图 12 所示，其中图 (a) 为原始 YoloX 算法的测试结果，图 (b) 为改进后 YoloX 算法的测试结果。

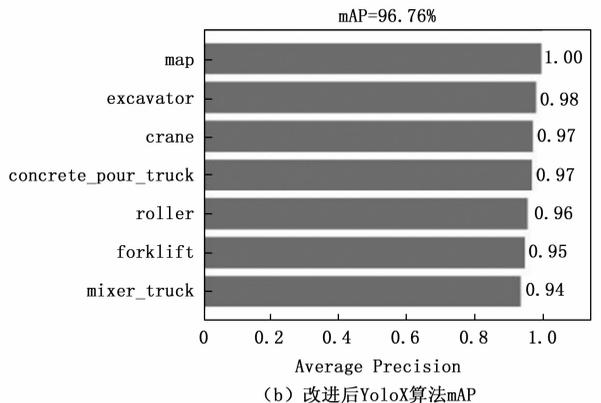
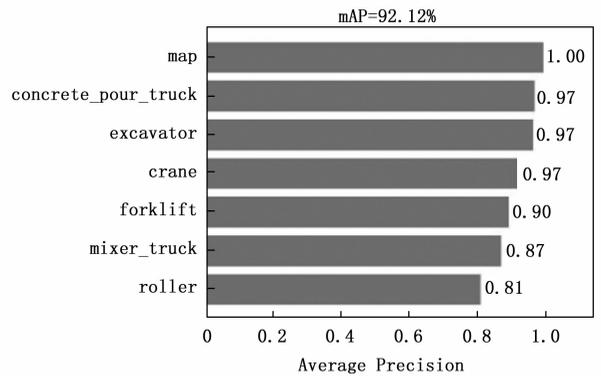


图 12 YoloX 算法测试结果对比图

改进后的 YoloX 算法在实际应用中的效果如图 13 所示，图 (a) 为来自摄像头拍摄的输电通道图片，图 (b) 为原始 YoloX 算法检测效果图，图 (c) 为改进后 YoloX 算法检测效果图。选取的部分图片中一共有 10 个工程车辆，原始 YoloX 算法共识别出 4 个工程车辆，改进后 YoloX 算法共识别出 7 个工程车辆。由对比图可知，改进后的 YoloX 算法训练出的模型在不同的场景中对工程车辆的检测效果较为理想。

表4 不同深度学习网络对比

算法	mAP/%	AP/%						
		crane	concrete_pour_truck	excavator	forklift	mixer_truck	roller	map
SSD	89.09	90.25	88.94	85.92	87.28	89.28	81.77	100.00
Yolov5	91.21	92.34	94.39	90.02	88.28	86.12	88.61	100.00
YoloX	92.12	92.28	97.21	97.26	90.36	87.56	81.87	100.00



(a) 来自摄像头拍摄的输电通道图片



(b) 原始YoloX算法检测效果图



(c) 改进后YoloX算法检测效果图

图13 工程车辆检测效果图

4 结束语

本文提出了一种改进版 YoloX 算法, 该方法通过改进 YoloX 算法的 CspLayer 结构, 添加 CBAM 注意力机制模块, 将 obj 分支的 BCELoss 损失函数替换为 Focal Loss 损失函数, 对工程车辆进行检测。与此同时, 通过利用改进 MSR 算法进行图片的亮度提升, 提高了数据集的质量, 更有利于提高了 YoloX 网络的识别。由于在实际大规模的应用场景中, 使用 YoloX 算法会存在一些误检, 影响检测效果, 后续工作可以进一步研究场景识别, 来进一步提升工程车辆的检测效果。

参考文献:

- [1] 康旭东. 电力线路巡检管理系统的设计与实现 [D]. 成都: 电子科技大学, 2017.
- [2] 葛群峰, 高兴波, 史旭华. 基于 YOLOv4 和 HOG 的闭环检测算法 [J]. 无线通信技术, 2021, 30 (2): 56-62.
- [3] 李军华, 权小霞, 汪宇玲. 多特征融合的瓷砖表面缺陷检测算法研究 [J]. 计算机工程与应用, 2020, 56 (15): 191-198.
- [4] 岳晓新, 贾君霞, 陈喜东, 等. 改进 YOLO V3 的道路小目标检测 [J]. 计算机工程与应用, 2020, 56 (21): 218-223.
- [5] GIRSHICK R. Fast R-CNN [C] // IEEE International Conference on Computer Vision. Santiago: ICCV, 2015: 1440-1448.
- [6] REN SHAOQING 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] DAI J, LI Y, HE K, et al. R-fcn: Object detection via region-based fully convolutional networks [C] // Advances in neural information systems. 2016: 379-387.
- [8] 谢俊章, 彭辉, 唐健峰, 等. 改进 YOLOv4 的密集遥感目标检测 [J]. 计算机工程与应用, 2021, 57 (22): 247-256.
- [9] 于硕, 李慧, 桂俊俊, 等. 复杂场景下基于 YOLOv5 的口罩佩戴实时检测算法研究 [J]. 计算机测量与控制, 2021, 29 (12): 188-194.
- [10] RAJ M, Chandan S. Pedestrian and Vehicle Detection Using Night-Vision Camera through CNN on Indian Roads [C] // International Conference on Advances in Computing, Communication Control and Networking. Department of Mechanical Engineering, Galgotias University, Greater Noida, India; Department of Mechanical Engineering, Galgotias University, Greater Noida, India;., 2018: 1136-1142.
- [11] 宋欢欢, 惠飞, 景首才, 等. 改进的 RetinaNet 模型的车辆目标检测 [J]. 计算机工程与应用, 2019, 55 (13): 225-230.
- [12] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional Block Attention Module [J]. Springer, Cham, 2018: 3-19.
- [13] 常惠, 饶志强, 赵玉林, 等. 基于改进 U-Net 网络的隧道裂缝分割算法研究 [J]. 计算机工程与应用, 2021, 57 (22): 215-222.
- [14] 王奔. 基于残差网络的显著物体检测方法及应用 [D]. 扬州: 扬州大学, 2020.
- [15] 史彩娟, 涂冬景, 刘靖祎. Re-GAN: 残差生成对抗网络算法 [J]. 中国图象图形学报, 2021, 26 (3): 594-604.
- [16] 万卓, 叶明, 刘凯. 基于改进 YOLOv4 的电机端盖缺陷检测 [J]. 计算机系统应用, 2021, 30 (3): 79-87.
- [17] 姜雪松, 姚鸿勋. 夜晚图像增强方法综述 [J]. 智能计算机与应用, 2020, 10 (3): 394-403.
- [18] 王卫星, 赵恒. 结合改进 Retinex 及自适应分数阶微分的雾霾公路交通图像增强 [J]. 光学精密工程, 2020, 28 (8): 1820-1834.
- [19] 刘志涛, 何自芬, 张印辉. 基于 YOLO 的雾天行人实时检测方法 [J]. 软件, 2019, 40 (10): 53-57.
- [20] 张娜, 韩美林, 杨琳. 雾霾图像清晰化处理算法的研究 [J]. 计算机与数字工程, 2019, 47 (6): 1478-1481.
- [21] 林昌, 陈武, 周海峰. 基于全局亮度自适应均衡化的海上图像带色彩恢复的多尺度 Retinex 算法 [J]. 科学技术与工程, 2020, 20 (33): 13712-13717.
- [22] KIM Y, KOH Y J, LEE C, et al. Dark image enhancement based on pairwise target contrast and multi-scale detail boosting [C] // IEEE International Conference on Image Processing. IEEE, 2015: 1404-1408.