

基于 YOLOv7x 的接触网吊弦缺陷检测方法

王晓明¹, 陈智宇¹, 董文涛^{*1}, 姚道金¹, 黄贻凤²

(1. 华东交通大学 电气与自动化工程学院, 江西 南昌 330000;

2. 华东交通大学 土木建筑学院, 江西 南昌 330000)

摘要: 针对铁路运行时由于接触网吊弦缺陷造成的安全隐患问题, 提出一种基于改进 YOLOv7x 的接触网吊弦缺陷识别方法。首先在主干特征提取层的末端引入 Swin Transformer 网络替换原有的扩展高效层聚合网络模块, 提高网络掌握全局信息的能力; 再引用 SIOU(SCYLLA-IoU) 损失函数替换原网络的损失函数, 为预测框收敛过程添加方向惩罚机制; 最后使用 Coordinate Attention 注意力机制融合颈部层中的扩展高效层聚合网络模块, 增强颈部网络模块的全局感受野。实验仿真结果表明此改进算法训练出的模型精度达到 95.9%, 相较于原 YOLOv7x 算法检测精度提高了 4.7%, 检测速度也达到了 52 帧每秒; 本改进算法解决了在吊弦缺陷识别领域检测效率低下的问题, 在实际作业中能够提高接触网吊弦缺陷排查工作的效率。

关键词: 接触网吊弦缺陷检测; YOLOv7x; Swin Transformer; SIOU 损失函数; Coordinate Attention

中图分类号: U225.4

文献标志码: A

Detection method of catenary hanging string based on YOLOv7x

Wang Xiaoming¹, Chen Zhiyu¹, Dong Wentao^{*1}, Yao Daojin¹, Huang Yifeng²

(1. School of Electrical Automation and Engineering, East China Jiaotong University, Nanchang 330000, China;

2. School of Civil Engineering and Architecture, East China Jiaotong University, Nanchang 330000, China)

Abstract: Aiming at the safety hidden problem caused by overhead contact wire defects during railway operation, an improved YOLOv7x method for overhead contact wire defect identification was proposed. Firstly, Swin Transformer network is introduced at the end of the backbone feature extraction layer to replace the original extended and efficient layer aggregation network module, so as to improve the ability of the network to grasp global information. Then the SIOU(SCYLLA-IoU) loss function is used to replace the original network loss function, and the direction penalty mechanism is added to the convergence process of the prediction frame. Finally, Coordinate Attention is used to integrate the extended efficient layer in the neck layer to aggregate the network module, enhance the global receptive field of the neck network module. Experimental simulation results show that the accuracy of the model trained by the improved algorithm reaches 95.9%, which is 4.7% higher than that of the original YOLOv7x algorithm, and the detection speed reaches 52 frames per second. The improved algorithm solves the problem of low detection efficiency in the field of hanging strings defect identification, and can improve the efficiency of detection of hanging strings defect in practice.

Key words: Detection of catenary hanging string defects; YOLOv7x; Swin Transformer; SIOU Loss; Coordinate Attention

近年来国家高速发展，铁路线路网络遍布全国各地，铁路运输对国家的经济发展有着极其重要的作用。接触网是高速铁路的重要组成部分，受电弓通过接触网完成对列车的电力输送，接触网的各部分状态的好与坏时时刻刻影响着铁路的运行安全。吊弦作为接触网重要的一部分，对输电线路起到承重与支撑的作用，吸收和缓冲承力索及定位点所受的振动冲击，一旦发生断裂，将导致接触网几何参数发生变化，影响列车取流及安全运行^[1-3]，因此对吊弦状态进行常态化检测是必要的。往年一般采用人工对接触网各部分进行检测，这项检测方式十分耗费人力物力，检查效率也不尽人意，而机器视觉作为当前热门技术，在许多领域都有应用这项先进技术解决检测难题的高质量论文，同时也有许多学者将此技术带到了接触网缺陷检测方面。如 HAN 等^[4]提出一种基于 YOLOX 的针对绝缘子缺陷检测的视觉检测算法，屈志坚等^[5]提出一种基于 YOLOv3 的针对接触网鸟巢检测与识别的检测方法。基于深度学习的目标检测算法由于具有良好的泛化能力和从复杂背景中提取特征的能力，在接触网缺陷检测领域得到广泛应用。应用于接触网目标检测的算法一般分为两类：一阶段算法（One-State）和二阶段算法（Two-State）。二阶段目标检测算法大多使用 RPN(Region Proposal Network)来减少复杂背景对目标及其缺陷检测的干扰。其主要流程是对图像先进行区域生成，再通过卷积神经网络进行样本分类。刘杰等^[6]提出使用 Faster R-CNN 神经网络模型对吊弦缺陷图片数据进行训练。TAN 等^[7]提出了基于 Faster R-CNN 算法的吊弦缺陷检测模型，实现了吊弦的定位和识别以及弯曲松弛和断裂吊弦的故障检测。虽然上述二阶段改进算法可以使接触网部件检测的检测精度达到很可观的数值，但效率低、速度慢的问题是不容忽视的。

一阶段目标检测算法代表算法有 YOLO 系列和 SSD 等。它可以通过单次检测生成物体的位置坐标和类别概率，能够快速准确地完成检测任务。张妹等^[8]提出一种利用 SSD 定位并识别巡检图像中螺栓缺陷的检测方法，虽然 SSD 检测算法在速度上有很大提升，但是牺牲了过多的精度，并不适合用于缺陷排查。顾桂梅等^[9]提出使用 Transformer 模块替换原 YOLOv5s 算法主干网络末端的 C3 模块，以此来强化主干网络的全局特征信息提取能力，但该模块自注意力计算方式会带

来庞大而计算量，给模型检测速度带来较大影响。上述方法相较于二阶段目标检测方法能够大幅度提升检测速度，但是在检测精度上方面还有提升空间。

基于以上提到的不足，提出一种基于改进 YOLOv7x 的接触网吊弦缺陷检测方法；引入新的骨干网络 Swin Transformer 替换主干特征提取层末端的扩展高效层聚合网络(Extended Efficient Layer Aggregation Networks, ELAN)模块，通过自注意力的计算以此来提高模型的检测精度。引入 SIOU(SCYLLA-IoU)损失函数替换 YOLOv7x 所采用的 CIoU(Complete-IoU)，在原有的基础上引入了方向惩罚机制，相较于网络原损失函数 CIoU 提高了训练效率。最后在网络的颈部层引入 CA(Coordinate Attention)注意力机制，加强了通道与通道之间的联系。原网络经过改进后能够有效提高检测精度。

1 YOLOv7x 算法原理

YOLO 是当前目标检测领域最优算法之一，具有检测速度快，检测精度高且满足实时性的优势。而 YOLOv7^[10]作为目前 YOLO 最先进的目标检测算法，它在 5FPS 到 160FPS 的范围内在速度和精度上已经超过了所有的物体检测器。YOLOv7x 在 YOLOv7 的基础上进行了模型缩放，使得训练后的模型精度更高。

Yolov7x 网络分别由输入层，主干特征提取层和颈部层组成。主干特征提取层由多个不同参数的卷积标准化激活函数和池化函数拼接而成，用来提取输入的图像特征，供之后的网络使用，该层主要的两个模块 ELAN 与 MPConv 的结构如图 1a、图 1b 所示。颈部层不同于主干特征提取层多了池化模块(SPPCSPC)、反卷积模块(Unsample)如图 1c 所示，

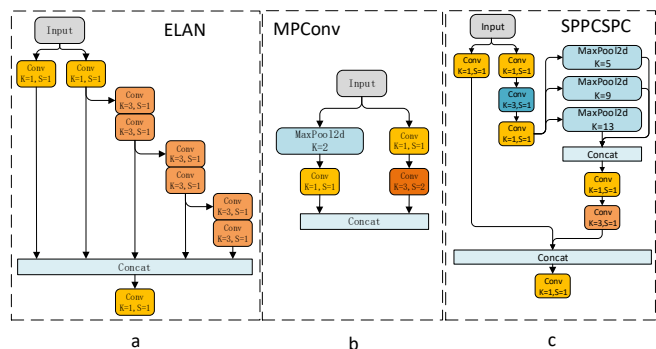


图 1 YOLOv7x 部分模块结构

Fig. 1 YOLOv7x partial module structure

该层通过上采样和下采样将不同层次的特征图融合在一起，生成具有多尺度信息的特征图，以此来提高检测的准确率。经过上述一系列卷积和最大池化后向检测头部分输出三种不同尺度的特征图像，与已经标注好的图片数据进行对比，计算

出此轮训练的损失函数，损失函数包括：定位损失、分类损失、置信度损失。根据训练情况调整模型参数，从而使模型更加精确。YOLOv7x 的整体网络结构如图 2 所示：

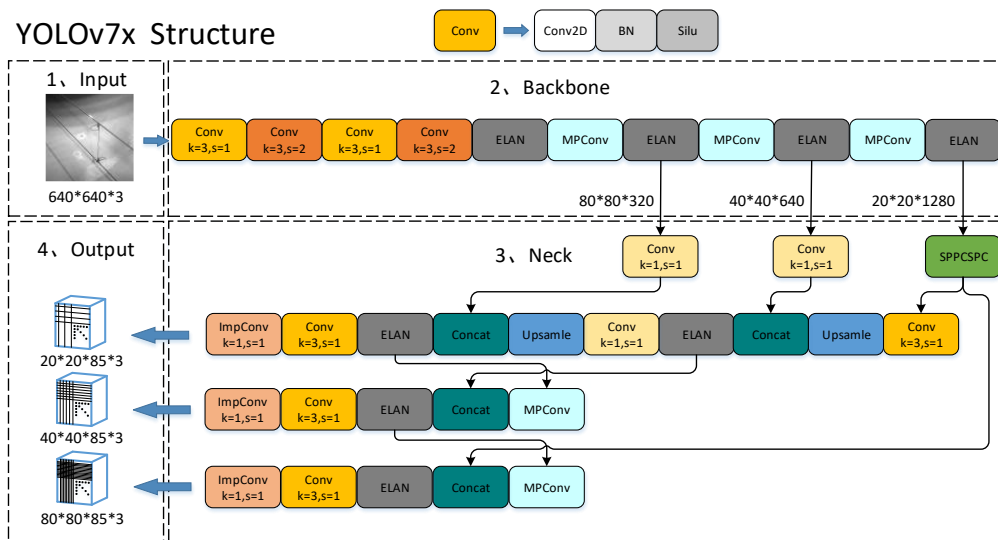


图 2 YOLOv7x 算法框架

Fig. 2 YOLOv7x algorithm framework

2 YOLOv7x 的改进

针对 YOLOv7x 网络结构提出以下几点改进：

- 1, 引入 Swin Transformer 网络替换主干特征提取层底部的 ELAN 模块，以牺牲些许检测速度的代价来提高模型的检测精度。
- 2, 替换网络原有的损失函数 CIoU 为 SIoU，有

助于模型快速收敛，提高模型的训练效率与检测精度。

- 3, 在颈部层添加 CA 注意力机制，增强网络模型特征学习能力能够提升模型精度，且参数量少不会带来过多的计算量。

改进后的网络框架如图 3 所示：

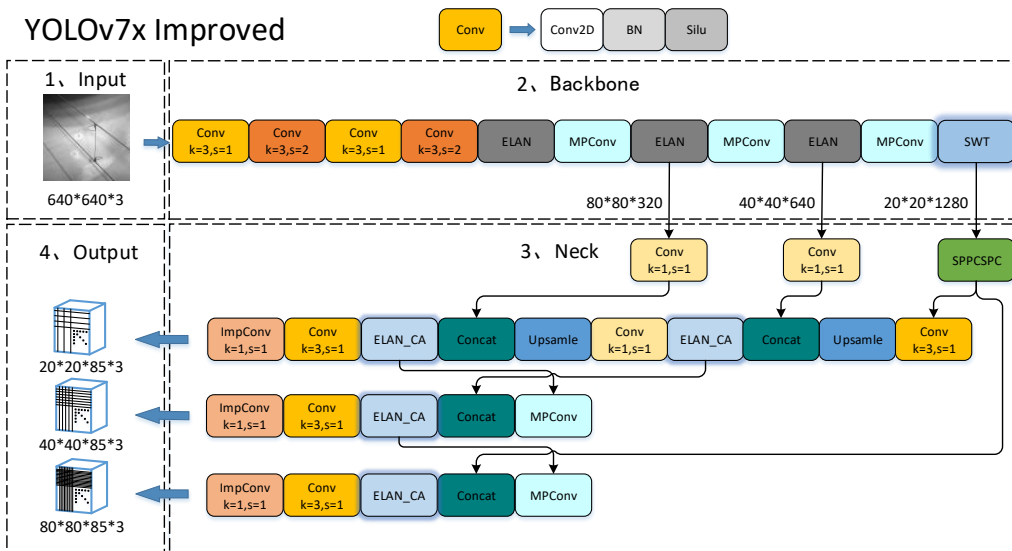


图 3 改进算法框架

Fig. 3 Improved algorithm framework

2.1 替换新骨干网络 Swin Transformer

由于 YOLOv7x 缺乏获取全局信息的能力, 网络提取目标特征的能力也会相应减弱^[11], 为弥补这方面不足, 在原骨干网络末端引入新的模块 SWT(Swin Transformer), SWT 是 LIU 等^[12]提出的一种新的视觉转换器, 它是基于 VIT^[13](Vision Transformer)改进后提出的一种更优的具有多头自注意力机制的计算方法, SWT 的网络如图 4 所示:

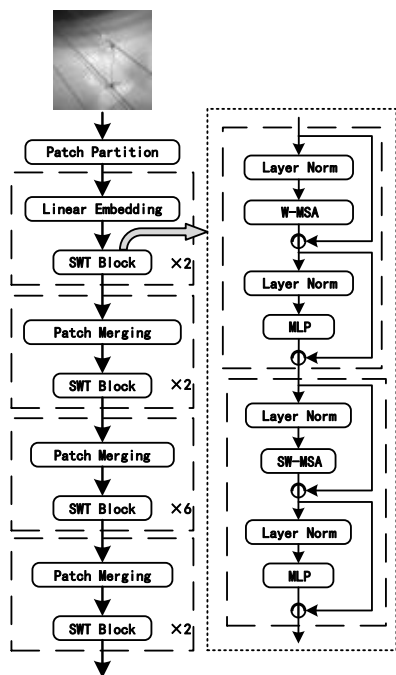


图 4 Swin Transformer 架构

Fig. 4 Swin Transformer framework

该网络使用多个自注意力层来代替 YOLOv7x 中的部分空间卷积层, 网络通过多头自注意力计算能够获得更多的上下文信息、建立全局的依赖关系, 扩大图像的感受野^[14]。不同于常规的 Transformer 结构, 例如前面所提到的 VIT 所采用的是全局自注意力计算方式, 在整个图像中计算全局自注意力, 计算复杂程度如式(1)所示。而 SWT 首先使用了类似于卷积操作的层次化构建方法, 将特征图划分为多个不重叠式的“小窗口”, 在每个小窗口内计算自注意力, 如图 4 中的 W-MSA(Window Multi-Head Self-Attention)多头自注意力模块, 该方法的计算复杂程度如式(2)所示。

以下两式是基于 $h \times w$ 的图像计算复杂度。

$$\theta^{W-MSA} = 4hwC^2 + 2M^2hwC \quad (1)$$

$$\theta^{SW-MSA} = 4hwC^2 + 2M^2hwC \quad (2)$$

式中, M 为表示为 $M \times M$ 像素大小的小窗口, C 代表特征图深度;

可见在实际计算过程中, 如果图片为 640×640 的大小, 全局自注意力计算方式的计算量将非常庞大, 在局部窗口内分别计算自注意力可以大大减少计算量。但是如果只是仅在每个独立的小窗口单独计算各自的自注意力会限制了网络建模能力, 无法为网络提供完整的全局感受野。图 4 中 SW-MSA(Shift-Window Multi-Head Self-Attention)多头自注意力模块解决了该缺陷, 算法流程如图 4 所示, 图 4a 对窗口进行移动得到了图 4b, 过程可以看作每个小窗口向右侧和下方移动了 $\frac{M}{2}$ 个

像素, 之后通过平移分割后的小窗口得到图 4c, 合并相邻同大小的小窗口, 这样又和图 4a 一样是 4 个 $M \times M$ 像素大小的窗口, 这种方式既能够使小窗口之间互相通信, 打破独立窗口对全局建模的限制性, 也能保证了计算量不变。之后在自注意力计算时通过掩码的方式隔绝不同区域的信息, 计算完成后复原图片, 至此完成了 SW-MSA 多头自注意力计算操作。

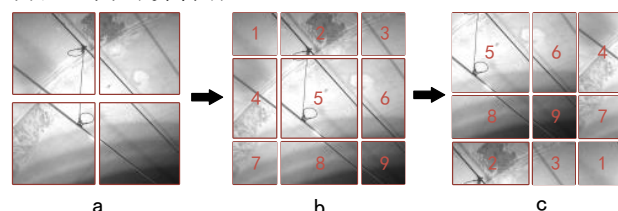


图 4 移动窗口工作原理

Fig. 4 Working principle of SWT

SWT 在计算网络的自注意力, 不同于原来的 Transformer 自注意力计算公式需要对每个头部添加相对位置偏执 B , 如式子(3)所示:

$$Attention(Q, K, V) = \text{soft max}\left(\frac{QK^t}{\sqrt{d_k}} + B\right)V \quad (3)$$

式中 $B \in R^{M^2 \times M^2}$, $Q, K, V \in R^{M^2 \times d}$ 分别代表查询、键和值向量矩阵; d 是查询向量矩阵维度, M^2 是每个小窗口的像素;

2.2 替换损失函数 SIoU

损失函数是模型改进与优化的重要指标, 针对不同模型选取一个合适的损失函数可以提高模型的学习速度、预测性能与检测精度。在实际检测过程中, 预测框的少量偏移与缩放对模型的精度都会有很大影响。YOLOv7x 采用的是 CIoU^[15] 损失函数, 相较于 IoU、GIoU、DIoU 损失函数加入了长宽比的惩罚项, 使得模型的评估更加准确, 但该损失函数收敛并没有考虑到预测框与目标框之间相对方向的因素, 这样会使预测框在训练的过程中盲目运动从而使得模型收敛速度慢效率变低, 而吊弦缺陷识别模型需要训练大量图像数据, 若在训练过程中预测框无法高效率的收敛, 将会造成大量的资源浪费并可能导致最终训练出的模型效果变差。针对这种问题本网络引入损失函数 SIoU^[16], 该损失函数基于 CIoU 损失函数引入了方向惩罚机制, 在预测框收敛过程中将所需回归之间的向量角度纳入考量标准, 重新定义了惩罚指标, 这样能够有效控制预测框回归的自由度, 相较于其他损失函数提高了训练效率。SIoU 损失函数方向惩罚机制原理如图 5 所示, 首先将预测框与目标框放在坐标轴上, 图中 B 方块为预测框, B^{GT} 为目标框。

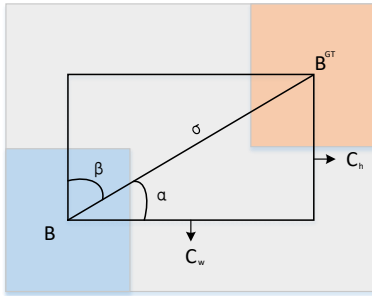


图 5 SIoU 损失函数角度惩罚方案

Fig. 5 SIoU loss function angle penalty scheme

为解决预测框在训练过程盲目运动的问题,

收敛过程首先尝试最小化 α , 如果 $\alpha \leq \frac{\pi}{4}$, 则最

小化 $\beta = \frac{\pi}{2} - \alpha$ 。为了实现这一点, 损失函数引入

公式(4)定义角度损失 Λ :

$$\Lambda = 1 - 2 \times \sin^2(\arcsin(x) - \frac{\pi}{4}) \quad (4)$$

式中 $x = \frac{c_h}{\sigma} = \sin(\alpha)$, $\sigma = \sqrt{(b_{c_x}^{gt} - b_{c_x})^2 + (b_{c_y}^{gt} - b_{c_y})^2}$,

$$c_h = \max(b_{c_y}^{gt}, b_{c_y}) - \min(b_{c_y}^{gt}, b_{c_y});$$

由于角度损失的变化, 需重新定义距离损失, 距离损失公式如式(5)所示:

$$\Delta = \sum_{t=x,y} (1 - e^{-\rho_t}) \quad (5)$$

$$\text{式中 } \rho_x = \left(\frac{b_{c_x}^{gt} - b_{c_x}}{c_w}\right)^2, \quad \rho_y = \left(\frac{b_{c_y}^{gt} - b_{c_y}}{c_h}\right)^2,$$

$$\gamma = 2 - \Lambda;$$

预测框形状损失函数定义如式(6)所示:

$$\Omega = \sum_{t=w,h} (1 - e^{-\omega_t})^\theta \quad (6)$$

$$\text{式中 } \omega_w = \frac{|w - w^{gt}|}{\max(w, w^{gt})}, \quad \omega_h = \frac{|h - h^{gt}|}{\max(h, h^{gt})},$$

形状损失函数中的 θ 是一个非常重要的自定义参数, 该参数控制着对形状损失的注意力程度, 将 θ 的值设置为 4。

整体 SIoU 损失函数定义为式(7)所示:

$$L_{box} = 1 - IoU + \frac{\Delta + \Omega}{2} \quad (7)$$

$$\text{式中 } IoU = \frac{|B \cap B^{GT}|}{|B \cup B^{GT}|};$$

SIoU 相较于原网络损失函数多了方向惩罚机制, 理论上替换后有助于改进模型。

2.3 添加注意力机制 CA

在视觉神经网络模型训练中, 计算机需要处理大量图像数据, 但每张图片信息中只有一部分是模型训练所需要的, 这样盲目的训练使得计算机资源浪费, 训练效率下降。为解决这类问题, 采用向 YOLOv7x 网络引入注意力机制的方法使计算机有选择性地对一些关键性的图片信息进行处理, 以此来提高模型的训练效率。目前流行的注意力机制有 SENet^[17] (Squeeze-and-excitation), CBAM^[18] (Convolutional Block Attention Module)等。SENet 将特征通道之间的关系纳入了考量范围, 降低了模型的计算成本。但缺点是 SENet 仅考虑了特征通道间的关系而忽略了位置信息的重要性, 位置信息对于在视觉任务中捕获对象结构至关重要, 后来的 CBAM 在 SE 的基础上解决了这一关键性问题, 但 CBAM 利用位置信息的方法是通过减少输入通道数, 然后通过一次卷积计算空间注意力, 如图 6a 所示, 而卷

积函数只能捕获局部关系, 是没有提取长距离间依赖关系的功能, 长距离间关系对视觉模型训练任务是至关重要的。

为了让模型训练时充分利用到通道与通道之间以及更远距离间的联系, 引入 CA 注意力机制^[15], 它的工作结构主要分为坐标信息嵌入和坐标注意力两部分。该注意力机制可以通过将位置信息嵌入到通道注意力机制中, 使网络能够捕获大范围的位置信息而不会带来过多的计算量, 结构如图 6B 所示。而且 CA 注意力机制方法轻便, 易于插入到 YOLOv7x 网络结构中。

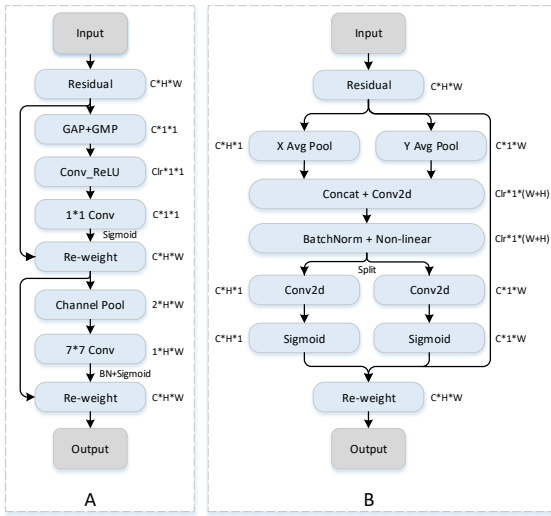


图 6 各注意力机制模块结构

Fig. 6 Structure of each attention mechanism module

注意力机制 CA 不同于 SE, 首先将全局池转化为一维特征编码, 经过输入给定 X , 沿水平坐标以及垂直坐标两个方向聚合特征, 生成方向感知图。通道 a 高度为 h 的输出可以表示为式(8)所示, 同理可得通道 a 宽度为 w 的输出可以表示为式(9)所示:

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i < W} x_c(h, i) \quad (8)$$

$$z_c^w(w) = \frac{1}{H} \sum_{0 \leq j < H} x_c(j, w) \quad (9)$$

其中 W 与 H 分别为水平和垂直两个空间范围的极值;

由上述可知注意力机制 CA 有能力使某个特征通道与相隔较远的特征通道建立联系, 并且可以沿水平方向保留其精确位置, 这有助于网络模型更精准的定位所需要的信息。

3 实验结果及分析

3.1 实验环境

为验证本文改进算法的可行性, 本文的对比实验是在同一个环境下进行, 编译环境为 pytorch, CPU 型号为 Intel(R) Xeon(R) Platinum 8157 CPU @ 2.30GHz, GPU 型号为 NVIDIA GeForce RTX 3090 24G 显存, 运行内存为 86G。

3.2 实验数据集来源及预处理

本实验数据集来源于南昌铁路局使用轨道检测车装置拍摄接触网各零件的高清图像集, 图片分辨率为 5120×5120 , 其中包含吊弦的图片一共有 1600 张, 通过 labelImg 软件对图片进行标注, 本文主要将吊弦状态分为正常吊弦(zcdx)、载流环断裂(zlhdl)、变形吊弦(bxdx)和断裂吊弦(dldx)这四种, 如图 7 所示:

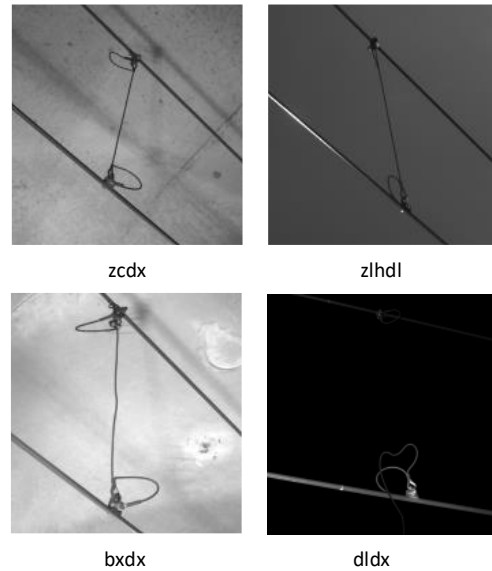


图 7 接触网吊弦状态

Fig. 7 Catenary suspension wire status

由于变形吊弦与断裂吊弦的图像数量较少, 为提高模型的泛化能力以及防止模型过拟合。在模型训练前通过数据增强方法, 对图片进行翻转, 镜像和旋转等操作, 使各类型吊弦图像数量比例合理。之后通过 python 代码按照 8:1:1 的比例随机划分出训练集, 测试集和验证集供模型训练, 至此数据集的制作完成。

3.3 实验设计及评价指标

为证明本文网络改进后在模型性能上有相应的提升, 本实验进行以下消融实验。通过模型训练获得的 $mAP@0.5$ (mean Average Precision) 来作为评判模型精确度的标准, 训练出的模型每秒检测出的图片数量作为该训练模型检测速度评判

标准,用FPS (Frames Per Second)表示。输入图像分辨率设置为640*640。训练轮次为200,步长设置为16。各实验结果如表1所示,消融实验在训练模型时使用YOLOv7x的预训练权重,因为整体网络结构与原YOLOv7x网络结构大体相似,两者可以共享大部分权重信息,这样可以大大减少训练时间。

表1 消融实验结果

Table 1 Ablation experiment results

YOLOv7x Improved			$mAP@0.5$	FPS
SWT	SIoU	CA		
×	×	×	91.2%	58.8
√	×	×	94.5%	55.2
×	√	×	91.6%	58.6
×	×	√	92.7%	54.3
√	√	×	94.7%	54.6
√	×	√	95.1%	52.9
×	√	√	93.0%	54.6
√	√	√	95.9%	52.3

注:√代表采用该改进方法,×代表未采用该改进方法

由上述消融实验结果可以证实:①在YOLOv7x的主干网络末端替换SWT后,模型精度提升3.3%,检测速度也相应的降低了3帧,这样的结果符合前文介绍所提到的SWT替换网络后的效果。②在替换新的网络损失函数SIoU后,检测精度提高了0.7%且检测速度大致不变。可见在添加新的方向惩罚机制后能够在不增加计算量的同时优化模型推理过程,提高检测精度,且添加SIoU后每一轮训练时间要快于原网络。③在颈部网络的ELAN模块融合CA注意力机制后相较于原网络精度提高了1.5%,检测速度降低了4帧,符合

预期。改进后的算法相较于原YOLOv7x算法检测速度略微降低了6帧,但是检测精度提升了4.7%。图6是本文改进算法与原YOLOv7x算法分别训练两百轮次的检测精度 $mAP@0.5$ 随训练轮次变化的曲线对比图。

其中红色曲线代表本文算法,用黑色曲线代表YOLOv7x算法。两种算法在85轮训练后精度都趋于稳定,原始算法 $mAP@0.5$ 在91%左右趋于稳定,最优效果达到的91.2%,改进后的算法 $mAP@0.5$ 稳定在95.6%左右,效果最好达到95.9%,由图8可知两种算法没有出现拟合或欠拟合现象且本文改进后算法要优于原始算法。

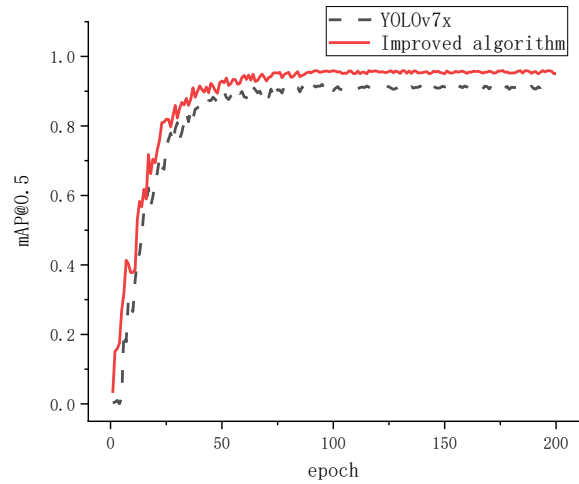


图8 精度对比曲线

Fig. 8 Accuracy comparison curve

为了验证实验结果的可靠性,分别采用表2中的五种算法训练出的模型权重检测同一组的图片,结果如图9所示。

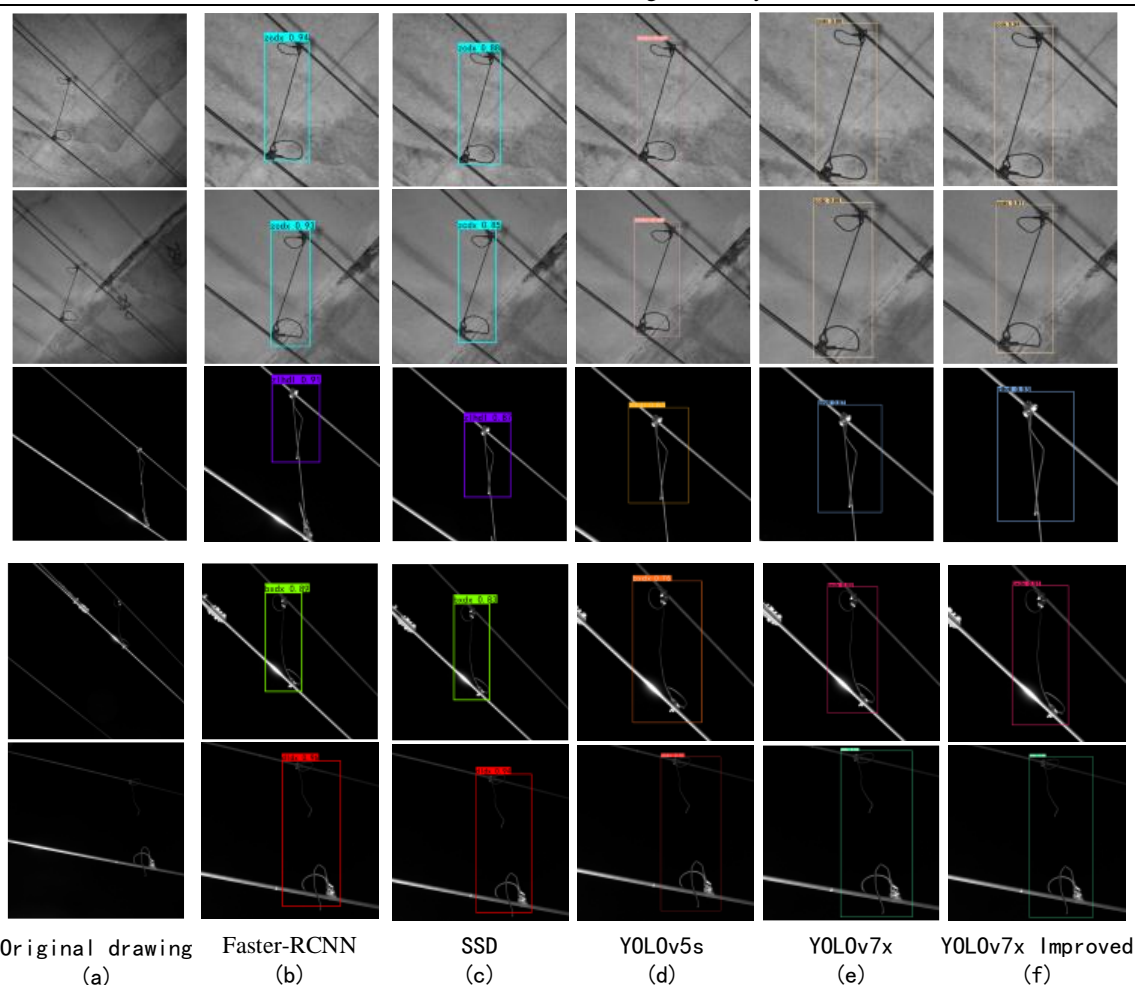


图 9 各类算法检测结果对比

Fig. 9 Comparison of detection results of various algorithms

由图 9 检测结果可知, 本文改进算法在检测置信度上要高于其他一阶段目标检测算法, 在检测部分图片的置信度要高于二阶段目标检测算法, 而且由表 2 可知本文检测算法在检测速度上要远高于二阶段目标检测算法。

为更确切的感受到本文算法的优越性, 本实验选取热门目标检测算法 Faster-RCNN、SSD 和 YOLOv5s 与本文改进行效果对比实验, 模型评判标准依旧不变, $mAP@0.5$ 作为模型的精度标准, FPS 作为模型的检测速度标准。训练结果如表 2 所示。

表 2 对比实验结果

Table 2 Comparison of experimental results

Algorithm	$mAP@0.5$	FPS
Faster-RCNN	94.4%	8.6
SSD	84.3%	62.4
YOLOv5s	88.9%	70.8

YOLOv7x	91.2%	58.8
YOLOv7x Improved	95.9%	52.3

由实验结果可以得出, 二阶段检测算法在检测速度上要远低于一阶段检测算法, 但是在检测精度上要优于一阶段检测算法。YOLOv5s 和 SSD 算法在速度上要优于 YOLOv7x 算法, 原因在于 YOLOv7x 网络结构更复杂, 运算量更大, 所以 YOLOv7x 在精度上要高于 YOLOv5s 及 SSD, 本文改进后的算法在检测精度上有明显优势, 检测速度达到了可观的效果。结合以上几点, 本文的改进算法更适合完成铁路上的接触网吊弦缺陷检测任务。

4 结论

本文研究了基于 YOLOv7x 算法的吊弦缺陷检测改进方法, 以实现接触网吊弦的高效率检测。以下是我们的结论:

- 1) 通过在原网络的主干特征提取层底部使用

Swin Transformer 模块替代 ELAN 模块,并在颈部层添加 CA 注意力机制,以及使用 SIoU 替换原网络的损失函数,可以改善模型并提高检测精度。

2) 实验结果显示,经过改进的 YOLOv7x 网络的检测精度达到了 95.9%,相较于原网络提高了 4.7%。在检测速度方面达到了 52 帧/s,领先于其他目标检测算法。

3) 研究结果表明,我们的改进算法在检测效率上有很大提升,满足了接触网吊弦缺陷检测的实时性需求。本改进算法并无法检测接触网吊弦是否出现断股情况,后期还将针对吊弦断股识别问题开展进一步研究。

参考文献

- [1] 吴文江,李响,高占凤,等.基于有限元法的高速铁路接触网吊弦动态特性研究[J].铁道学报,2020,42(07):65-72.
WU W J, LI X, GAO Z F, et al. Research on dynamic characteristics of dropper of high-speed railway catenaries based on finite element method[J]. Journal of the China Railway Society, 2020, 42(07): 65-72.
- [2] 洪晓斌,何永奎,周建熹等.承力索锚结线夹绞线区域损伤的超声导波监测[J].华南理工大学学报(自然科学版),2017,45(05):1-8.
HONG X B, HE Y K, ZHOU J X, et al. Ultrasonic guided wave monitoring of the damage in the area of the anchor and twisted wire of the bearing cable[J]. Journal of South China University of Technology (Natural Science Edition), 2017, 45(05): 1-8.
- [3] 黄盼盼,胡艳,武云龙.吊弦故障时弓网受流质量评估[J].电气化铁道,2023,34(02):43-46.
HUANG P P, HU Y, WU Y L. The arch net is evaluated by the flow quality when the hanging string fails[J]. Electrified railway, 2023, 34(02): 43-46.
- [4] HAN G J, LI T, LI Q, et al. Improved algorithm for insulator and its defect detection based on YOLOX[J]. Sensors, 2022, 22(16): 6186.
- [5] 屈志坚,高天姿,池瑞,等.基于改进的YOLOv3接触网鸟巢检测与识别[J].华东交通大学学报,2021,38(04):72-80.
QU Z J, GAO T Z, CI R. Nest detection and identification based on improved YOLOv3 overhead contact net [J]. Journal of East China Jiaotong University, 2021,38(04):72-80.
- [6] 刘杰,许建国,高春丽,等.基于机器视觉的接触网吊弦缺陷检测研究[J].铁道工程学报,2022,39(05):91-97.
LIU J, XU J G, GAO C L, et al. Machine vision-based research on the inspection of dropper defects of overhead contact line[J]. Journal of Railway Engineering Society, 2022, 39(05): 91-97.
- [7] TAN P, LI X F, WU Z G, et al. Multialgorithm fusion image processing for high speed railway dropper failure-defect detection[J]. IEEE Transactions on Systems, Man, and Cybernetics: Systems, 2021, 51(7): 4466-4478.
- [8] 张姝,王昊天,董骁翀,等.基于深度学习的输电线路螺栓检测技术[J].电网技术,2021,45(07):2821-2829.
ZHANG S, WANG H T, DONG X C, et al. Bolt detection technology of transmission lines based on deep learning[J]. Power System Technology, 2021, 45(07): 2821-2829.
- [9] 顾桂梅,贾耀华,温柏康.基于YOLOv5s的接触网吊弦线和载流环缺陷识别算法[J].铁道科学与工程学报,2023,20(03):1066-1076.
GU G M, JIA Y H, WEN B K. Defect identification algorithm of dropper line and current-carrying ring of catenary based on YOLOv5s[J]. Journal of Railway Science and Engineering, 2023, 20(03): 1066-1076.
- [10] WANG C Y, BOCHKOVSKIY A, LIAO H Y M, et al. YOLOv7: trainable bag-of-freebies sets new state-of-the-art for real-time object detectors[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver: IEEE. 2023: 7464-7475.
- [11] LU S L, LIU X Y, HE Z X, et al. Swin-Transformer-YOLOv5 for real-time wine grape bunch detection[J]. Remote Sensing, 2022, 14(22): 5853.
- [12] LIU Z, LIN Y T, GAO Y, et al. Swin transformer: hierarchical bision transformer using shifted windows[C]// Proceedings of The IEEE/CVF International Conference on Computer Vision. Montreal: IEEE. 2021: 10012-10022.
- [13] BOCHKOVSKIY A, BEYER L, KOLESNIKOV A, et al. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale[C]//International

Conference on Learning Representations.2021.

- [14] NIU Z, ZHONG G, Yu H. A review on the attention mechanism of deep learning[J]. Neurocomputing, 2021, 452: 48-62.
- [15] ZHENG Z, WANG P, LIU W, et al. Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression[J].arXiv,2019.DOI:10.1609/aaai.v34i07.6999.
- [16] ZHORA G. SIoU loss: more powerful learning for bounding box regression [EB/OL]. Arxiv Preprint. arXiv:2205.12740, 2022.
- [17] HU J, SHEN L, SUN G. Squeeze-and-excitation networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 7132-7141.
- [18] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional block attention module[C]//Proceedings of the European conference on computer vision (ECCV). 2018: 3-19.
- [19] HOU Q B, ZHOU D Q, FENG J S, et al. Coordinate attention for efficient mobile network design[C]// Proceedings of The IEEE/CVF Conference on Computer Vision and Pattern Recognition. Nashville: IEEE. 2021: 13713-13722.



第一作者: 王晓明 (1978-), 男, 副教授, 硕士生导师, 研究方向为人工智能, 机器视觉, 柔性传感器, 物联网技术及应用。E-mail: 2501@ecjtu.edu.cn



通信作者: 董文涛 (1987-), 男, 副教授, 硕士生导师, 研究方向为传感器与智能检测。E-mail : wentao_dong@163.com