

面向钢丝绳微弱损伤智能识别的多尺度注意力网络

曲诚, 陈景龙, 常元洪, 周子桐
(西安交通大学机械工程学院, 710049, 西安)

摘要: 针对复杂恶劣工况下客运索道钢丝绳表面损伤程度微弱、识别难度高,且损伤样本数量较少的问题,提出了一种融合多尺度特征提取与注意力机制的钢丝绳微弱损伤智能识别方法。首先引入生成对抗网络对运行图像去除模糊,并使用透视变换和随机裁剪扩充样本数量,接着采用多尺度卷积神经网络提取损伤图像的整体特征与局部特征,对小尺寸特征图上采样后与大尺寸特征图拼接以融合不同尺度的特征,并添加注意力机制对关键特征进行增强,最终经预测模块输出损伤的类别和坐标。通过钢丝绳损伤模拟实验台上采集的损伤图像对所提方法进行验证,结果表明:与现有方法比较,该方法在保证平均精度均值(mAP)不降低的同时大幅度缩短了训练时间,体现了该方法的有效性。

关键词: 智能识别;微弱损伤;注意力机制;钢丝绳

中图分类号: TH17 **文献标志码:** A

DOI: 10.7652/xjtuxb202107016 **文章编号:** 0253-987X(2021)07-0141-10



OSID 码

Multi-Scale Attention Network for Intelligent Identification of Weak Damage on Wire Ropes

QU Cheng, CHEN Jinglong, CHANG Yuanhong, ZHOU Zitong
(School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an 710049, China)

Abstract: Aiming at the problems of weak surface damage, high identification difficulty and small number of damage samples of passenger ropeway under complex and severe working conditions, an intelligent wire rope weak damage identification method based on multi-scale feature extraction and attention mechanism is proposed. Firstly, the generative adversarial network is introduced to remove the ambiguity of the images under running condition, the perspective transformation and random clipping are used to expand the number of samples. Then the multi-scale convolution neural network is used to extract the overall features and local features of the damage images, and the small-scale feature map samples are collected and then spliced with the large-scale feature map to join the features of different scales. On this basis, the key features are enhanced by the attention mechanism. Finally, the coordinates and category of the damage are output by the prediction module. The proposed method is verified by collecting the damage images on the wire rope damage simulation test bench. The results show that compared with the existing methods, the proposed method greatly shortens the training time while the mean average precision (mAP) is not reduced, which reflects the validity of the proposed method.

Keywords: intelligent identification; weak damage; attention mechanism; wire rope

钢丝绳作为客运索道缆索系统的核心部件之一,在客运索道的运行中发挥着重要作用^[1]。由于客运索道往往架设于山峰、滑雪场等地,钢丝绳长期工作在雨雪、大风等恶劣的环境中,极易导致断丝、磨损、绳股松散等微弱损伤的发生。受环境因素和运行振动的影响,钢丝绳的微弱损伤特征往往难以被及时发现^[2],且一旦客运索道发生事故,会造成严重的经济损失和负面的社会影响。因此,开展客运索道钢丝绳微弱损伤早期识别方法的研究具有重要的工程意义和应用价值。

目前应用较广的钢丝绳表面损伤识别方法主要是基于漏磁原理^[3]。基于漏磁原理的钢丝绳损伤检测示意图如图1所示,首先使用永磁铁对钢丝绳进行励磁,当钢丝绳出现表面损伤时,损伤部位的横截面积会减小,使内部的磁力线从损伤部位漏出,此时如果附近存在传感器,就会检测到漏磁场,从而识别出损伤^[4]。这种检测方法虽然目前应用较为广泛,但是存在装置结构复杂、检测效率低、易受环境干扰出现误差等不足^[5]。

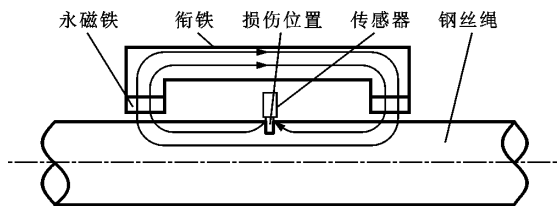
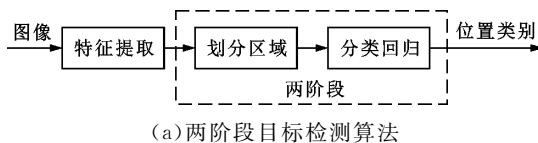


图1 基于漏磁原理的钢丝绳损伤检测示意图

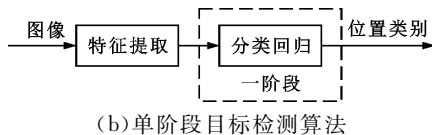
Fig.1 Schematic diagram of wire rope damage detection based on magnetic leakage principle

近年来,随着目标检测技术的快速发展,为钢丝绳的表面损伤识别带来了新的发展方向。自2014年 Girshick 提出的 RCNN(Region CNN)方法^[6]首次将深度学习引入目标检测任务以来,已有多国学者对此类算法展开研究。目前基于深度学习的目标检测算法主要分为两阶段方法和单阶段方法,目标检测算法流程如图2所示,两类方法首先都使用卷积神经网络提取图像特征,在此基础上,两阶段方法先划分大量可能存在目标的候选区域,再使用分类与回归网络预测目标的类别及位置,单阶段方法则在特征提取后直接进行类别与位置的预测。

基于目标检测的智能检测算法现已被广泛应用于人脸识别、行人检测等工业领域中。Zhao 等针对苹果采摘机器人识别目标具有种类多、光照变化大、存在遮挡等问题,提出了基于 YOLOv3 深度卷积神经网络的苹果定位方法,实现了对复杂环境下苹果



(a)两阶段目标检测算法



(b)单阶段目标检测算法

图2 目标检测算法流程图

Fig.2 Flow chart of object detection algorithm

的检测识别^[7]。Zhang 等融合特征提取能力较强的 Inception 网络^[8]和深度残差网络^[9],设计出新的网络作为目标检测算法 Faster-RCNN^[1]的特征提取网络,并应用于动车故障检测中^[11]。Zhang 等以 FCN 为基础网络,提出了可变形池化核,并加入了批量标准化层和空洞卷积层以提升网络性能,实现了对遥感图像中尺寸差异较大的目标检测^[12]。

虽然基于深度学习的目标检测模型在多个领域都取得了一定效果,但当前的研究仍然面临两大挑战,一方面是复杂恶劣运行条件下的微弱损伤识别难度大,钢丝绳损伤的面积通常较小,损伤特征本就微弱,同时钢丝绳在高速运行时产生的振动,以及雨、雪、霜等干扰又会一定程度地掩盖损伤特征,导致损伤更加难以识别。一方面是小样本下的损伤识别精度低,训练目标检测模型需要使用大量的样本数据,然而对于客运索道钢丝绳等类似设备,由于损伤发生的时间、地点均为未知,给损伤样本的获取带来了较大困难,而小样本条件下训练出的模型识别效果较差,无法满足工程需求。为了解决复杂恶劣工况下钢丝绳表面损伤程度微弱、识别难度高,且损伤样本数量较少的问题,实现对钢丝绳微弱损伤的快速识别与准确定位,本文提出了一种融合多尺度特征提取与注意力机制的损伤识别方法。首先采用生成对抗网络对运行图像去除模糊,并使用透视变换和随机裁剪扩充样本数量;接着在使用多尺度卷积神经网络进行特征提取的基础上,融合 SE 注意力机制对关键特征进行增强;最终输出钢丝绳的损伤位置及损伤种类。通过对钢丝绳损伤模拟实验台上采集的损伤图像进行识别来验证本文提出方法的有效性。

1 融合多尺度特征提取与注意力机制的损伤识别网络

本文提出的融合多尺度特征提取与注意力机制

的 YOLOv3-tiny-SE 模型总体结构如图 3 所示,模型主要由预处理模块、特征提取模块和预测模块组成。预处理模块通过生成对抗网络对钢丝绳运行图像进行去模糊,并使用透视变换、随机裁剪扩充样本数量;特征提取模块通过使用添加注意力机制的多尺度卷积神经网络提取图像中的整体与细节特征,并通过上采样和拼接操作以融合不同尺度的特征;预测模块通过将特征图中的信息转化为输出坐标以

实现对图像中损伤位置及种类的识别。

1.1 预处理模块

构建高准确率的损伤识别模型,通常需要大量有效的训练样本对模型进行全面的训练,然而在高速运行条件下获取到的钢丝绳图像会出现模糊现象,导致模型难以提取有效特征,影响后续的分类与定位,因此在预处理模块首先采用生成对抗网络对图像进行去模糊处理。

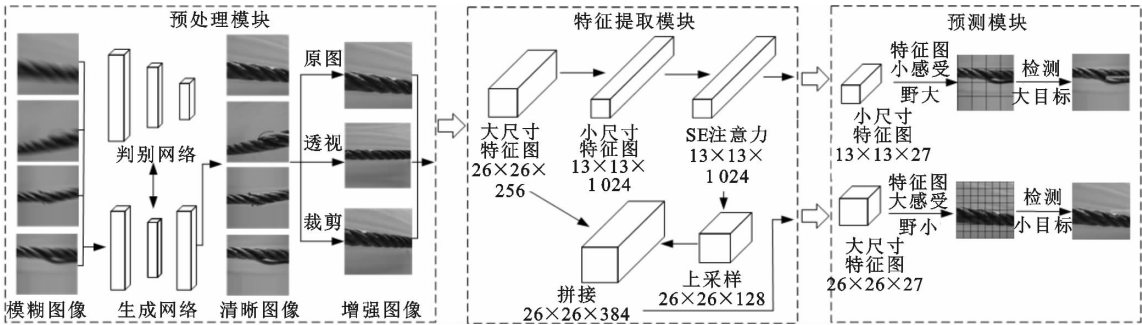


图 3 YOLOv3-tiny-SE 模型总体结构

Fig. 3 Overall structure of YOLOv3-tiny-SE model

实际的钢丝绳处于持续运行状态,损伤可能发生在钢丝绳表面的任何位置,同时实验模拟时的拍摄角度与实际监控的角度可能不完全一致,因此为了使数据集中的样本更符合实际情况,提高模型的泛化能力,在预处理模块采用透视变换和随机裁剪两种方法对去模糊后的图像进行数据增强。

1.1.1 去除模糊 模糊图像可以看作清晰图像与模糊核卷积后的结果^[13],模糊过程可表示为

$$I^B = k * I^S + N \tag{1}$$

式中: k 为模糊核,其大小由拍摄设备的移动速度等因素决定; $*$ 为卷积操作; I^B 为模糊图像; I^S 为清晰图像; N 为随机噪声。

传统的图像去模糊方法通常基于滤波优化,即在已知卷积核 k 的情况下,通过反卷积操作对模糊图像进行复原,然而在实际应用时,大部分情况下模糊核的信息是未知的,因此传统方法需要对模糊场景做出预先假设,以简化复原难度,目前基于滤波优化的图像去模糊方法在特定条件下可以取得较好的效果,但面对复杂多变的模糊情况时仍存在局限性。

近年来随着深度学习技术的快速发展,神经网络等方法逐渐被用于解决去模糊问题。与基于滤波优化的去模糊方法相比,基于深度学习的智能去模糊方法具有更广的适用性,使用时只需要获取成对的模糊图像与清晰图像,在模糊场景与模糊核信息未知的情况下,即可学习模糊图像与清晰图像之间

的关系并复原其他模糊图像,因此本文采用了深度学习方法中的生成对抗网络对图像去除模糊。

基于生成对抗网络的钢丝绳运行图像去模糊流程如图 4 所示,主要由生成网络和判别网络两部分组成。生成网络负责对模糊图像进行重构,生成网络结构如图 5 所示,生成网络主要由卷积模块、9 个残差模块和反卷积模块组成。首先通过卷积模块对图像进行下采样,接着通过残差模块还原图像的细节并经反卷积模块复原图像尺寸,最后将还原的细节特征与原模糊图像相加,即可得到重构图像。

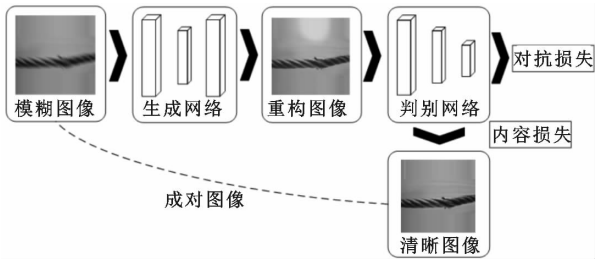


图 4 基于生成对抗网络的钢丝绳运行图像去模糊流程图

Fig. 4 Deblurring diagram of wire rope image based on generative adversarial network

从模糊图像到重构图像的过程可表示为

$$I^S = G_{\theta_G}(I^B) \tag{2}$$

式中: G_{θ_G} 代表生成网络。

使用生成网络得到重构图像后,接着使用判别网络评估重构图像的真实程度,判别网络结构如图

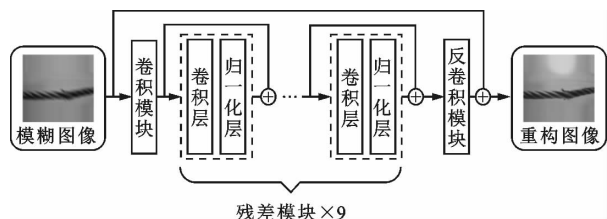


图5 生成网络结构

Fig. 5 Structure of generative network

6所示,首先通过卷积层与池化层提取图像特征,再经过全连接层判断图像是真实的清晰图像还是由生成网络生成的重构图像,最后输出图像的标签。

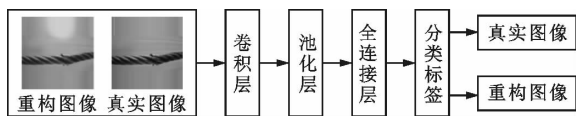


图6 判别网络结构

Fig. 6 Structure of discriminative network

判别网络的判别流程可表示为

$$l = D_{\theta D}(I^R, I^S) \quad (3)$$

式中: I^R 为真实的清晰图像; l 为判别网络输出的图像标签; $D_{\theta D}$ 代表判别网络。

在训练生成网络与判别网络的过程中会产生对抗损失与内容损失^[14],其中对抗损失可表示为

$$L_{adv} = \sum_{n=1}^N -D_{\theta D}(G_{\theta G}(I^B)) \quad (4)$$

内容损失可表示为

$$L_{con} = \frac{1}{WH} \sum_{i=1}^W \sum_{j=1}^H (G_{\theta G}(I^B(i, j)) - I^S(i, j))^2 \quad (5)$$

式中: W 和 H 代表图像的长和宽。

通过交替的训练生成网络与判别网络,降低对抗损失和内容损失,在训练结束后使用生成网络对运行图像进行去模糊处理,可以提高图像的清晰度。

1.1.2 透视变换 透视变换可表示为

$$\mathbf{Y} = \mathbf{M}\mathbf{X} \quad (6)$$

式中 $\mathbf{X}$ 和 $\mathbf{Y}$ 分别是变换前后的图像, $\mathbf{M}$ 是变换矩阵,表达式为

$$\mathbf{M} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{bmatrix} \quad (7)$$

在进行透视变换时,首先确定变换前后图像中4个顶点的坐标;接着,以变换前顶点的坐标 (u, v) 作为输入 $\mathbf{X}$, 变换后顶点的坐标 (x, y) 作为输出 $\mathbf{Y}$, 求出变换矩阵中的参数;最后将原图像所有点的坐标与变换矩阵进行相乘,即可得到透视变换后的图

像^[15]。

1.1.3 随机裁剪 数据增强中常见的随机裁剪流程是先设定裁剪图像的尺寸,然后在原始图像中随机选取一点,以该点为中心,设定尺寸为边长进行裁剪,此方法适用于一般的图像分类任务,但对于本文的研究对象,如果只采用这种方法,裁剪图像可能会缺失损伤部分,影响后续对损伤位置的标注。

为解决此问题,本文提出如图7所示的改进的随机裁剪方法,首先确定原始图像中的目标区域,对于本文而言目标区域即为钢丝绳的损伤区域,接着计算目标区域4个顶点的坐标,然后按照设定的裁剪尺寸对原始图像进行裁剪,得到若干裁剪图像和裁剪图像的顶点坐标,进一步根据目标区域和裁剪图像的顶点坐标计算 $U(A, B)$, $U(A, B)$ 可表示为

$$U(A, B) = \frac{\text{area}(A) \cap \text{area}(B)}{\text{area}(A)} \quad (8)$$

式中: A 代表目标区域; B 代表裁剪图像; area 代表计算面积; $\cap$ 代表取交集。若 $U(A, B)$ 为1,说明裁剪图像包含全部目标区域,保留该裁剪图像,若 $U(A, B)$ 小于1,说明该裁剪图像只包含部分目标区域或不包含目标区域,舍弃该裁剪图像。该方法可以节省裁剪后筛选有效图像的时间,以便后续对损伤位置进行标注。

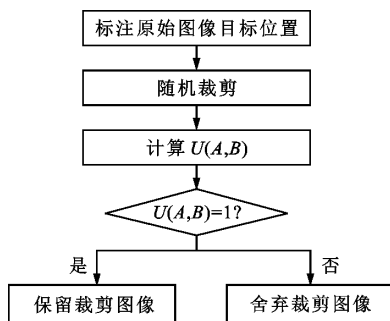


图7 改进的随机裁剪方法流程

Fig. 7 Process of improved random clipping method

1.2 特征提取模块

1.2.1 多尺度特征提取 特征提取模块用于提取输入图像的整体轮廓特征与局部细节特征,保留图像中的关键信息,流程如图3中间部分所示,图像输入到网络后,根据输入图像的尺寸与所需计算量,首先经过5个卷积核尺寸为 3×3 的卷积层和4个最大池化层进行初步特征提取。经过初步特征提取后,进一步进行深度特征提取,对小尺寸特征图上采样并与大尺寸特征图进行拼接,融合后的特征图可表示为

$$f_3 = \text{concat}\{\text{up}[F(f_1)], f_2\} \quad (9)$$

式中: f_1 和 f_2 分别代表小尺寸和大尺寸的特征图; F 代表卷积池化操作; up 代表上采样; concat 代表拼接; f_3 代表融合后的特征图^[16]。此时的特征图充分融合了浅层特征与深层特征,浓缩了原始图像的丰富信息。

1.2.2 SE 注意力机制 在前一小节中,经过多尺度特征提取后的特征图通道数已到了 1 024,此时的特征图虽然包含了图像中丰富的整体轮廓信息与细节特征信息,但这些信息分布在不同通道的不同位置,若直接与预测模块连接,可能会导致网络训练较多轮次才能达到平衡状态,为了解决此问题,本方法在特征提取模块中添加了 SE 注意力机制^[17],用于学习深度特征图各通道的重要程度和通道之间的相关性,对包含关键信息的通道赋予更高权重,以此改善网络表现^[18]。

SE 注意力机制的结构如图 8 所示,对于多通道特征图,首先通过全局平均池化浓缩特征图每个通道的信息,大小为 $W \times H \times c$ 的特征图经压缩后大小为 $1 \times 1 \times c$,压缩过程可表示为

$$z_c = \frac{1}{WH} \sum_{i=1}^H \sum_{j=1}^W f_c(i, j) \quad (10)$$

式中: $f_c(i, j)$ 是压缩前特征图的每个元素, z_c 是压缩后的特征图。

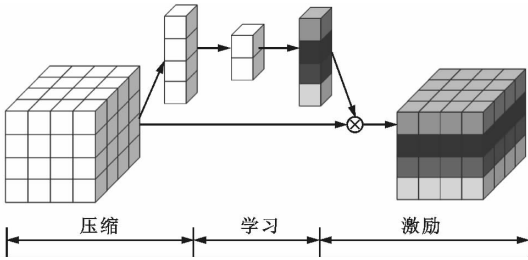


图 8 SE 注意力机制

Fig. 8 SE attention mechanism

对特征图压缩后,使用两层全连接层对特征图不同通道之间的相关性及各通道的重要性进行建模,进一步使用 sigmoid 激活函数对全连接层的输出进行归一化,此过程可表示为

$$s = \sigma(W_2 \delta(W_1 z_c)) \quad (11)$$

式中: s 是特征图每个通道的权重; W_1 和 W_2 分别代表两层全连接层的权重; $\delta(\cdot)$ 是 ReLU 激活函数; $\sigma(\cdot)$ 是 sigmoid 激活函数。两种激活函数可分别表示为

$$\delta(t) = \max(0, t) \quad (12)$$

$$\sigma(t) = \frac{1}{1 + e^{-t}} \quad (13)$$

两层全连接层使用不同的激活函数,原因在于第一层全连接层用于评判特征图不同通道之间的差异,因此使用 ReLU 激活函数保留差异,而第二层全连接层用于分配每个通道的权重,因此使用 sigmoid 激活函数将每个通道的权重限制在 0 和 1 之间,避免出现异常值。

网络在经过训练学习到优化后特征图每个通道的权重值后,对初始特征图的每个通道进行激励操作以增强关键特征,抑制干扰信息。激励操作可表示为

$$o = sf \quad (14)$$

式中: f 和 o 是输入和输出的特征图。

1.3 预测模块

经过融合注意力机制的特征提取模块后,输出特征图已经包含了目标的位置及类别信息,预测模块负责将特征图中的信息转化为输出坐标,如图 9 所示,特征图的每个网格对应 3 个先验框(虚线框),每个先验框对应一个预测框,因此每个网格对应 3 个预测框,模型的优化目标是使预测框尽可能匹配真实目标^[19]。

预测框的信息存在于特征图的深度方向上,每个预测框由 3 类信息组成,如图 10 所示,第 1 类是位置信息,由 x 、 y 、 w 、 h 4 个值组成,代表先验框与预测框之间的转换系数;第 2 类是类别信息,类似 one-hot 编码,预测类别的类别值接近 1,其他类别值接近 0;第 3 类为置信度信息,代表目标中心落在特征图该网格中的概率。综上所述,特征图的总深度为

$$d = b(n + 5) \quad (15)$$

式中: b 为每个网格对应的预测框个数; n 代表总类别数; 5 代表 4 个位置信息与 1 个目标置信度,本文目标为识别 4 种钢丝绳表面损伤,且每个网格对应 3 个预测框,因此预测模块的特征图总深度为 $3 \times (4 + 5) = 27$ 。

预测值需要通过进一步转换以得到实际坐标,转换过程如图 11 所示,虚线框是先验框,实线框是预测框,从先验框到预测框的转换过程可表示为

$$b_x = c_w + \sigma(x) \quad (16)$$

$$b_y = c_h + \sigma(y) \quad (17)$$

$$b_w = p_w e^w \quad (18)$$

$$b_h = p_h e^h \quad (19)$$

式中: b_x 和 b_y 是预测框的中心坐标; b_w 和 b_h 是预测框的长和宽; c_w 和 c_h 是特征图网格左上角的坐标; p_w 和 p_h 是先验框的长和宽(在对训练图像进行标注时,标注每张图像上目标区域的位置与长和宽,并

对训练集所有图像目标区域的长和宽使用 K -means 算法聚类,得到最可能出现的目标区域长和宽,将此值作为先验框的长和宽)。

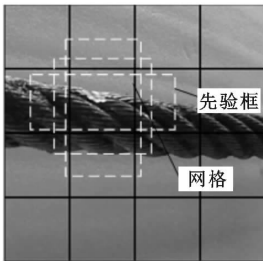


图 9 先验框与网格的对应关系

Fig. 9 Relation between anchor box and grid

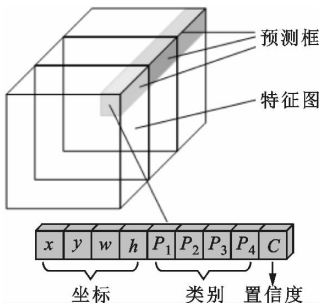


图 10 预测框信息组成

Fig. 10 Information of predictive box

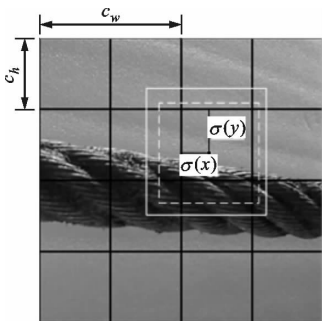


图 11 坐标转换过程

Fig. 11 Process of coordinate conversion

2 实验验证

2.1 图像数据获取

在如图 12 所示的钢丝绳损伤模拟实验台上采集损伤钢丝绳图像,实验台主要由损伤钢丝绳、摄像头和计算机组成。使用砂轮和尖嘴钳对钢丝绳加工了如图 13 所示的外部磨损、绳股松散、绳芯外露和断丝 4 种表面损伤。摄像头通过连接线与计算机连接,以 15 帧/s 的速度录制视频并在计算机端显示实时画面。采集图像时,将损伤钢丝绳固定于支架上,使用摄像头对钢丝绳进行录像,并移动支架以模

拟钢丝绳的实际运行状态,采集视频后,使用软件截取视频中的钢丝绳图像,得到原始图像。

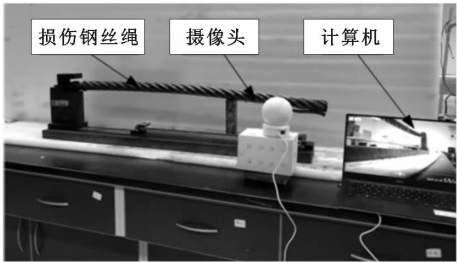


图 12 钢丝绳损伤模拟实验台

Fig. 12 Simulation test bench for wire rope damage

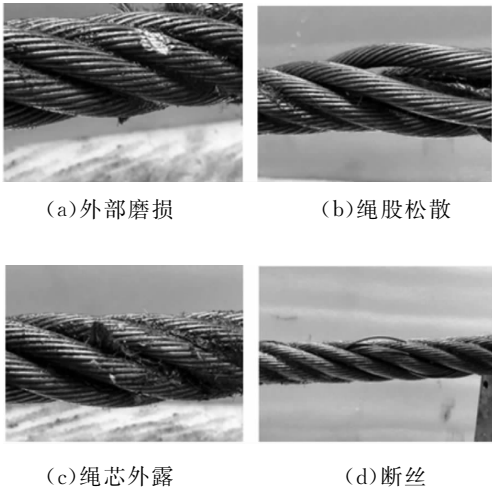


图 13 实际钢丝绳表面损伤

Fig. 13 Surface damages of actual wire rope

2.2 图像数据预处理

获取原始图像后,应用预处理模块对运行图像去除模糊并进行数据增强,构造损伤数据集,接着标注数据集中损伤图像的位置和类别并划分训练集和测试集,训练集和测试集分别包含 4 种损伤形式的 350 张图像和 100 张图像,表 1 给出了训练集和测试集的样本分布。

表 1 训练集与测试集样本分布

Table 1 Sample distribution of training set and test set

损伤形式	训练样本数	测试样本数
外部磨损	100	40
断丝	100	20
绳芯外露	100	20
绳股松散	50	20
合计	350	100

2.3 实验环境及网络训练参数

本实验在 Windows10 系统下使用 Python3.7 编程环境,基于 Pytorch 库搭建网络模型,在训练时

使用 CUDA 并行计算架构以加速训练进程。

训练时图像数设置为 4,采用自适应矩估计(Adam)优化器,共训练 250 轮,初始学习率为 0.005,当训练轮数达到 125 和 175 时学习率分别衰减为原来的 1/10 以减小后期的波动,使损失函数进一步收敛以接近最优解。

2.4 评估指标

为了评估去模糊的效果,本文引入峰值信噪比(P_{SNR})以衡量图像之间的相似度, P_{SNR} 可表示为

$$P_{SNR} = 10\lg\left(\frac{Z_1^2}{E}\right) \tag{20}$$

式中: Z_1 是图像的最大像素值,本文中该值为 255; E 代表两张图像所有对应像素点差值的平方和。 P_{SNR} 越大,代表两张图像相似度越高。

为了评估预测模块的预测效果,本实验采用目标检测领域常用的平均精度(AP)和平均精度均值(mAP)^[20]。对目标检测网络进行测试时,根据真实标签与预测标签是否匹配,预测结果可分为真阳性(T_P)、假阳性(F_P)、假阴性(F_N)、真阴性(T_N),其中真阳性代表预测结果为真且预测正确,假阳性代表预测结果为真但预测错误,假阴性、真阴性的定义以此类推。按照上述定义,精度和召回率可表示为

$$P = \frac{T_P}{T_P + F_P} \tag{21}$$

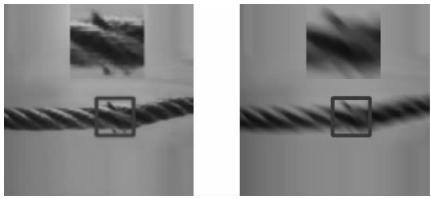
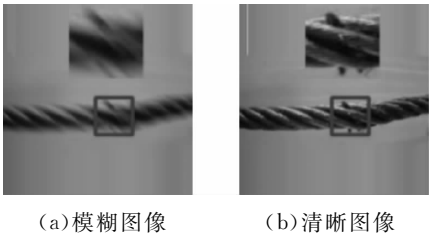
$$R = \frac{T_P}{T_P + F_N} \tag{22}$$

进一步,平均精度是精度随召回率变化曲线的积分,平均精度均值是所有类别平均精度的均值。

2.5 实验结果与分析

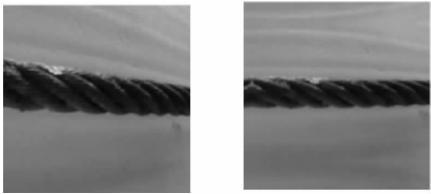
本小节分别给出了预处理模块与预测模块的实验结果与结果分析。

2.5.1 预处理模块结果分析 图 14 给出了去除模糊前后图像的对比,经中值滤波去模糊后的图像与真实清晰图像的 P_{SNR} 为 25.4 dB,而使用生成网络去除模糊后的图像与真实清晰图像的 P_{SNR} 达到了 27.4 dB,证明了生成网络在去除图像模糊方面的有效性。图 15 给出了透视变换前后图像的对比。



(c)生成网络去模糊后 (d)中值滤波后
图 14 去除模糊前后图像的对比

Fig. 14 Comparison of the images before and after deblurring



(a)透视变换前 (b)透视变换后
图 15 透视变换前后图像的对比

Fig. 15 Comparison of the images before and after perspective transformation

2.5.2 预测模块结果分析 使用实验获取的模糊钢丝绳绳损伤图像数据集对 YOLOv3、YOLOv3-tiny 和本文提出的 YOLOv3-tiny-SE 方法进行训练,图 16 给出了训练过程中平均精度均值的变化趋势,可以看出,至训练结束时,3 种方法的平均精度均值都已经基本稳定,且 YOLOv3-tiny-SE 的平均精度均值要高于其他两种方法。

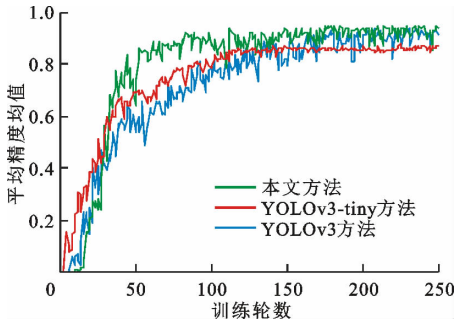


图 16 训练过程中平均精度均值的变化趋势
Fig. 16 Trend of mAP in the training process

图 17 给出了 3 种方法对不同种类损伤的识别效果对比,可以看出,本文方法对每种损伤的识别效果都要优于其他两种方法。

使用包含透视变换图像与不包含透视变换图像的数据集分别训练 3 种方法,训练结果对比如表 2 所示,从表中数据可以看出,数据集中包含透视变换图像时,3 种方法训练后的平均精度均值都有了明显的提升,证明了透视变换有助于提升识别效果。

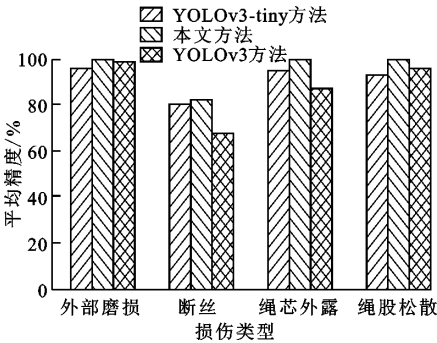


图 17 3 种方法对不同种类损伤的识别效果对比
Fig. 17 Comparison of the identification effect of three methods for different damages

表 2 有无透视变换的平均精度均值对比
Table 2 Comparison of mAP with or without perspective transformation

网络结构	训练集数据	平均精度均值/%
YOLOv3	仅原始图像	70.4
	原始图像+透视变换	86.6
YOLOv3-tiny	仅原始图像	83.0
	原始图像+透视变换	91.0
本文方法	仅原始图像	84.2
	原始图像+透视变换	93.7

为了说明关键参数选择对训练结果的影响,在对本方法方法进行训练时,分别设置了不同的图像数与初始学习率,首先将图像数分别设置为 2、4 和 8,不同图像数对训练结果的影响如图 18 所示,可以看出,当图像数设置为 2 时,平均精度均值曲线在训练的后期仍然存在一定波动,使模型的稳定性无法得到保证,而设置为 8 时,平均精度均值始终较低,因此,经过对比选择后,本文方法选择将图像数设置为 4。

图 19 对比了初始学习率分别设置为 0.05、

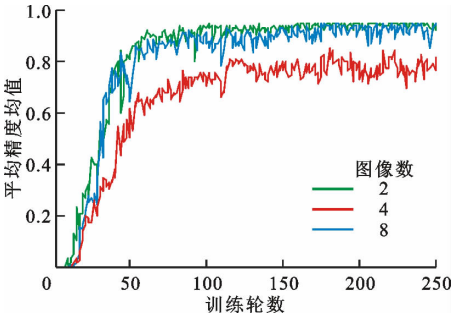


图 18 不同图像数对训练结果的影响
Fig. 18 Influence of different number of images on training results

0.005 和 0.000 5 时的训练结果,可以看出,当初始学习率较小时,权重每轮更新的幅度较小,经多轮训练直至训练结束时平均精度均值仍然没有稳定;当初始学习率较大时,权重每轮更新的幅度也会变大,导致训练后期精度曲线仍然存在小幅震荡,影响最终精度,因此在综合考虑训练速度及结果的稳定性后,本文方法选择将初始学习率设置为 0.005。

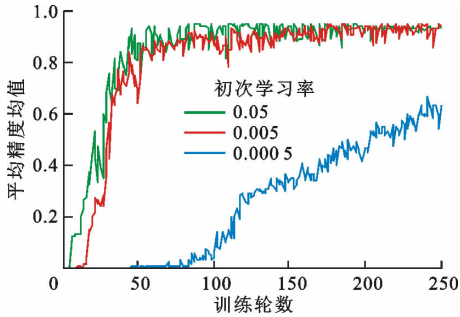


图 19 不同初始学习率对训练结果的影响
Fig. 19 Influence of different learning rate on training results

图 20 给出了本文方法对实际损伤图像的检测结果,可以看出,该方法可以有效地识别出钢丝绳图像中不同损伤的类型与位置。

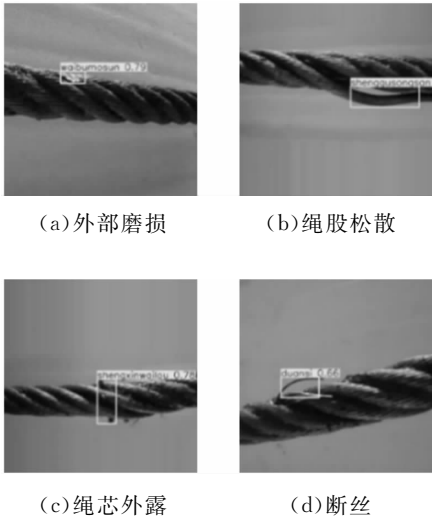


图 20 本文方法对实际损伤图像的检测结果
Fig. 20 Detection results of the proposed method

图 21 给出了 3 种方法的训练时间的对比,与 YOLOv3 方法训练时间达到 4.223 h 相比,由于模型结构更简单,训练一轮所需要更新的参数较少,本文方法的训练时间只有 0.928 h,缩短了近 80%。与 YOLOv3-tiny 方法相比,本文方法在训练时间上基本持平,但识别准确率更高,可见本文方法在性能与效率上都具有良好的表现。

通常来说,复杂的网络模型具有更强的特征提

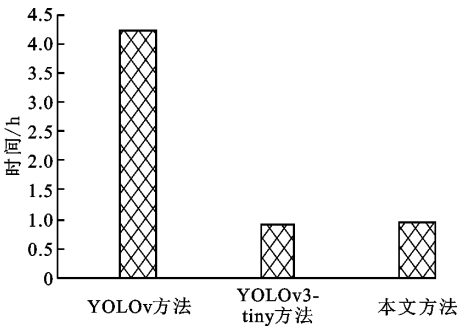


图 21 3 种方法的训练时间的对比

Fig. 21 Comparison of training time for three methods

取能力,预测准确率更高,但在本实验中,YOLOv3 方法的检测效果低于结构更简单的 YOLOv3-tiny 方法和本文方法,经过分析,主要原因在于虽然 YOLOv3 方法的层数多、结构复杂,在面对类别数多、尺寸差异大的目标检测时效果好,但本实验中钢丝绳图像的损伤面积相对较小,在经过 YOLOv3 方法的多层特征提取后,图像的关键特征信息会发生丢失,因此导致 YOLOv3 方法的检测效果欠佳。

3 结 论

针对复杂恶劣工况下钢丝绳表面损伤程度微弱、识别难度高,且损伤样本数量较少的问题,提出了一种融合多尺度特征提取与注意力机制的损伤识别网络,并在实验室钢丝绳损伤图像数据集上进行了验证,取得了较好的效果,得出如下结论。

(1)为了解决工程实际中有效样本较少的问题,引入了生成对抗网络用于运行图像去模糊,在此基础上,使用改进的随机裁剪和透视变换方法进行数据增强,在扩充样本数量的同时保留了损伤特征,节省了后续筛选有效样本的时间。

(2)构建了融合注意力机制的 YOLOv3-tiny-SE 方法用于钢丝绳表面损伤识别。针对实验获取的钢丝绳模糊损伤图像进行分析与测试,并与现有 YOLOv3 和 YOLOv3-tiny 方法进行对比,实验结果表明,与现有方法相比,本文方法的训练时间缩短了 80%,且平均精度均值可以达到 93.7%,验证了本文方法在钢丝绳损伤识别方面的优越性和有效性。

虽然本文方法在保证精度的前提下提高了效率,但该方法的检测效果受参数设置的影响较大,通过对比不同情况下的实验效果可以对参数进行筛选优化,但是效率较低。在后续研究中,会采用神经架构搜索(NAS)技术,通过强化学习来优化网络结构与参数,提高参数选择的效率。

参考文献:

[1] 罗原,衣宝龙,王洪欣. 客运索道振动频谱分析与故障诊断应用研究 [J]. 起重运输机械, 2018, 1: 159-161.

LUO Yuan, YI Baolong, WANG Hongxin. Study on vibration spectrum analysis and fault diagnosis of passenger cableway [J]. Hoisting and Conveying Machinery, 2018, 1: 159-161.

[2] 陈若蒙,梁少霞,钱剑雄. 基于失效模式分析的客运架空索道风险控制方法研究 [J]. 中国特种设备安全, 2019, 35(12): 27-30.

CHEN Ruomeng, LIANG Shaoxia, QIAN Jianxiong. Research on risk control method of passenger aerial cableway based on failure mode analysis [J]. China Special Equipment Safety, 2019, 35(12): 27-30.

[3] 张恩超. 基于钢丝绳漏磁检测定量算法研究 [D]. 哈尔滨: 哈尔滨工业大学, 2013: 7-8.

[4] ZHOU Ping, ZHOU Gongbo, ZHU Zhencai, et al. A review of non-destructive damage detection methods for steel wire ropes [J]. Applied Sciences, 2019, 9 (13): 2771-2787.

[5] SUN Huixian, ZHANG Yuhua, LUO Feilu. Texture defect detection of wire rope surface with support vector data description [C] // Proceedings of the Chinese Conference on Pattern Recognition. Piscataway, NJ, USA: IEEE, 2009: 721-725.

[6] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2014: 580-587.

[7] 赵德安,吴任迪,刘晓洋,等. 基于 YOLO 深度卷积神经网络的复杂背景下机器人采摘苹果定位 [J]. 农业工程学报, 2019, 35(3): 164-173.

ZHAO Dean, WU Rendi, LIU Xiaoyang, et al. Apple positioning based on YOLO deep convolutional neural network for picking robot in complex background [J]. Transactions of the Chinese Society of Agricultural Engineering, 2019, 35(3): 164-173.

[8] SZEGEDY C, LOFFE S, VANHOUCKE V, et al. Inception-v4, inception-resnet and the impact of residual connections on learning [C] // Proceedings of the 31st AAAI Conference on Artificial Intelligence. Palo Alto, CA, USA: AAAI, 2017: 4278-4284.

[9] HE Kaiming, ZHANG Xiangyu, REN Shaoqing. Deep residual learning for image recognition [C] // Proceedings of the IEEE Conference on Computer Vision

- and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2016: 770-778.
- [10] REN Shaoqing, HE Kaiming, 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, 2015, 39(6): 1137-1149.
- [11] 张江勇. 基于深度学习的动车关键部位故障图像检测 [D]. 成都: 电子科技大学, 2019: 60-61.
- [12] 张宗军. 基于深度学习的遥感图像地面目标检测算法研究 [D]. 太原: 中北大学, 2020: 57-58.
- [13] KUPYN O, BUDZAN V, MYKHAILYCH M, et al. DeblurGAN: blind motion deblurring using conditional adversarial networks [C] // Proceedings of the 31st IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2018: 8183-8192.
- [14] GULRAJANI I, AHMED F, ARJOVSKY M, et al. Improved training of Wasserstein GANs [C] // Proceedings of the 31st Annual Conference on Neural Information Processing Systems. Cambridge, MA, USA: NIPS, 2017: 5767-5777.
- [15] 许彦平. 基于三维重构与透视变换的汽车 A 柱盲区可视化方法研究 [D]. 长春: 吉林大学, 2019: 24-28.
- [16] REDMON J, FARHADI A. YOLOv3: an incremental improvement [EB/OL]. [2020-11-01]. <https://arxiv.org/abs/1804.02767>.
- [17] JIE Hu, LI Shen, SAMUEL A, et al. Squeeze-and-Excitation networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 42(8): 2011-2023.
- [18] WOO S, PARK J, LEE J, et al. CBAM: Convolutional block attention module [C] // Proceedings of the European Conference on Computer Vision. Berlin, Germany: Springer Press, 2018: 3-19.
- [19] JOSEPH R, ALI F. YOLO9000: better, faster, stronger [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway, NJ, USA: IEEE, 2017: 6517-6525.
- [20] JU Moran, WANG Zhongbo, LUO Haibo, et al. The application of improved YOLO V3 in multi-scale target detection [J]. Applied Sciences, 2019, 9(18): 3775-3789.

[本刊相关文献链接]

- 吴守军, 陈健, 冯辅周, 等. 改进的熵特征在行星齿轮裂纹故障识别中的应用. 2021, 55(6): 61-68. [doi: 10.7652/xjtuxb202106008]
- 吴春志, 吴守军, 冯辅周, 等. 一种具有强抗噪性的深度学习故障诊断模型. 2021, 55(4): 61-68. [doi: 10.7652/xjtuxb202104007]
- 阚寒冰, 陈尧, 贾志杰, 等. 融合历史轨迹的智能汽车城市复杂环境多目标检测与跟踪算法. 2018, 52(10): 132-140. [doi: 10.7652/xjtuxb201810019]
- 陈保家, 汪新波, 严文超, 等. 采用品质因子优化和子带重构的共振稀疏分解滚动轴承故障诊断方法. 2018, 52(4): 70-76. [doi: 10.7652/xjtuxb201804011]
- 邵卫东, 李志刚, 李军. 二维时间发展混合层噪声的直接数值模拟. 2017, 51(1): 38-44. [doi: 10.7652/xjtuxb201701007]
- 杨源, 库涛, 查宇飞, 等. 快速多特征金字塔的尺度目标跟踪方法. 2016, 50(10): 49-56. [doi: 10.7652/xjtuxb201610008]
- 王富平, 水鹏朗. 采用多尺度方向微分比率的角度点检测算法. 2016, 50(4): 68-75. [doi: 10.7652/xjtuxb201604011]
- 贾永红, 谢志伟, 张谦, 等. 采用独立阈值的遥感影像变化检测方法. 2015, 49(12): 12-18. [doi: 10.7652/xjtuxb201512003]
- 裴世源, 徐华. 非均质复合材料力学性能的不确定性多尺度计算方法. 2015, 49(10): 8-13. [doi: 10.7652/xjtuxb201510002]
- 高燕华, 刘玉欢, 喻罡. 多尺度非参数化水平集的超声心动图分割. 2013, 47(2): 53-57. [doi: 10.7652/xjtuxb201302009]

(编辑 武红江)