

doi:10.11835/j.issn.1000.582X.2023.09.014

基于 STA-YOLOv5 的水利建造人员安全帽佩戴检测算法

李顺祥¹, 蒋海洋², 熊伶¹, 黄才生¹, 蒋有高¹, 邓曦³, 王楷², 张鹏²

(1. 重庆市西部水资源开发有限公司 重庆 401329; 2. 重庆大学 自动化学院, 重庆 400030; 3. 中国建筑科学研究院有限公司, 北京 100013)

摘要: 在大型水利建设工程现场, 存在高空坠物、塔吊转动、墙体坍塌等问题, 对于建造人员人身安全造成巨大威胁, 佩戴安全帽是保护建造人员的有效措施, 作为工程作业中的安全管理, 对建造人员进行安全帽佩戴的精确检测很有必要。针对现有安全帽检测算法在大型水利建造场景下对小且密集的安全帽目标存在漏检、检测精度较低等问题, 提出一种基于 STA-YOLOv5 的安全帽佩戴检测算法, 该算法将 Swin Transformer 和注意力机制引入到 YOLOv5 算法中, 提高模型对安全帽的识别能力。实验结果表明, STA-YOLOv5 算法具有更精确检测结果, 识别准确率达到 91.6%, 较原有的 YOLOv5 算法有明显提升。

关键词: 水利建造; 安全帽佩戴检测; STA-YOLOv5; Swin Transformer; 注意力机制

中图分类号: TP391.4

文献标志码: A

文章编号: 1000-582X(2023)09-142-11

Hydraulic construction personnel safety helmet wearing detection algorithm based on STA-YOLOv5

LI Shunxiang¹, JIANG Haiyang², XIONG Ling¹, HUANG Caisheng¹, JIANG Yougao¹,
DENG Xi³, WANG Kai², ZHANG Peng²

(1. Chongqing Western Water Resources Development Co., Ltd., Chongqing 401329, P. R. China;
2. College of Automation, Chongqing University, Chongqing 400030, P. R. China; 3. China Academy of Building Research, Beijing 100013, P. R. China)

Abstract: At the site of large-scale water conservancy construction projects, there are problems such as falling objects, tower crane rotation, wall collapse, which pose a great threat to the personal safety of construction personnel. Wearing safety helmets is an effective measure to protect construction personnel. Therefore, it is necessary for construction personnel to carry out accurate detection of helmet wearing as a safety management in engineering operations. Aiming at the problems of missing detection and low detection accuracy of small and dense helmet targets in large-scale hydraulic construction scenarios, a helmet wearing detection algorithm based

收稿日期: 2023-05-12

基金项目: 重庆市水利科技资助项目(渝西水司文[2021]19号), 重庆市技术创新与应用发展专项重点项目(cstc2021jcsx-gksbX0032)。

Supported by Chongqing Water Science and Technology Project (No. 19, Shuisiwen [2021], West Chongqing) and the Technology Innovation & Application Development Projects of Chongqing(cstc2021jcsx-gksbX0032).

作者简介: 李顺祥(1969—), 男, 高级工程师, 主要从事水利工程建设管理、智慧水利等方向研究, (E-Mail)281849629@qq.com。

通信作者: 王楷(1981—), 男, 博士, 副教授, (E-mail)kaiwang@cqu.edu.cn。

on STA-YOLOv5 is proposed, which introduces Swin Transformer and attention mechanism into YOLOv5 algorithm to improve the identification ability of the model to the helmet. The experimental results show that the STA-YOLOv5 algorithm has more accurate detection results, and the recognition accuracy reaches 91.6%, which is significantly improved compared with the original YOLOv5 algorithm.

Keywords: water conservancy construction; helmet wearing detection; STA-YOLOv5; Swin Transformer; attention mechanism

近年来,随着大型水利建设工程建设项目的迅速发展,对建造现场的安全管理提出了更高要求。安全帽是施工过程中最基本的安全保护装备之一,佩戴安全帽可有效保护人身安全,避免因高空坠物、塔吊转动等意外事故造成伤害。因此,在建造现场的安全管理工作中,佩戴安全帽是不可缺少的施工要求之一,及时检测工人是否佩戴安全帽具有较大意义。

随着图像处理与目标检测的快速发展,很多学者对安全帽检测方法进行了研究。张博^[1]等使用 open-pose 定位人体头部区域,再使用 Faster RCNN 来检测安全帽,最后分析空间关系判断是否佩戴安全帽;朱晓春^[2]等基于改进 Darknet 网络的 YOLO v3 算法实时检测施工人员是否佩戴安全帽,检测速度得到明显提升;胡晓栋^[3]等使用更加轻量的 MobileViTv2 对安全帽的特征进行提取;王玲敏^[4]等人将注意力机制加入 YOLOv5 算法中,并将特征金字塔替换为加权双向特征金字塔(BiFPN),提高了网络对密集物体的检测能力;沈希忠^[5]等在 YOLOX 基础上,采用轻量级的 GhostNet 替换 CSPDarknet 作为 backbone,同时扩展特征融合网络的特征层输入尺度,提高网络对小目标的检测能力,并使用深度可分离卷积代替常规卷积,降低网络参数。Vishnu^[6]等将 CNN 应用于摩托车驾驶员头盔佩戴检测,准确率高达 92.87%,该方法不能进行多个目标检测。Hao^[7]等人使用基于颜色的混合描述提取安全帽特征,用分层支持向量机对安全帽进行分类检测。

在大型施工场景中,如渝西水资源配置工程,是重庆历史上投资规模最大、覆盖面最全和受利人群数量最多的民生工程(如图 1 所示),该工程以提高管理能力、工程质量和工作效率为中心,包括智慧建管、智慧设计、智慧运维和智慧工地 4 大板块,其中安全管理尤为重要。



图1 渝西水资源配置工程

Fig. 1 Yuxi Water Resources Allocation Project

针对大型水利建设工程,由于存在钢筋、水泥、输电线路、工地桥架等各种设备,使得施工场景图像涵盖范围较广泛且包含较多物体,图像中工人佩戴的安全帽目标较小。且由于摄像头本身受到光照强度、拍摄角度、拍摄距离、天气等因素的影响,导致在安全帽检测^[8-10]时更容易出现准确度低、漏检率大、错检率高等问题。为了提高大型施工场景中安全帽检测的性能,提出一种引入 Swin Transformer 和注意力机制的 YOLOv5 算法,即 STA-YOLOv5 算法,使安全帽检测的定位目标更加快速、精确。

1 算法整体框架

1.1 YOLO系列目标检测算法对比分析

YOLO系列为单步目标检测框架,对输入图像直接进行分类概率回归和包围框坐标回归实现目标检测。Redmon J^[11]等人在2016年提出基于回归的目标检测算法YOLOv1,实现了目标识别与定位的统一,先后提出改进版本YOLOv2^[12]和YOLOv3^[13],准确率得到了进一步提高。YOLOv4是对YOLOv3的改进,网络结构主要包括主干特征提取网络CSPDarknet53、空间金字塔池化(SPP^[14])、路径聚合网络(PANet^[15]),同时使用多种数据增强技术来提升检测性能。YOLOv5^[16-17]使用C3Darknet作为主干网络从输入图像中提取丰富的信息特征,使用PANet作为Neck聚合特征,模型检测层与YOLOv3相同,Conv卷积模块的激活函数采用SiLU^[18]函数。因此,相比于其他YOLO系列算法,选择YOLOv5作为安全帽佩戴检测算法主网络的原因在于其实时性能优异、简单易用、多尺度检测能力,YOLO系列算法对比如表1所示。此外,YOLOv5通过设置depth_multiple和width_multiple 2个参数调节主干网络的深度和宽度并划分出4个量级模型:YOLOv5s、YOLOv5m、YOLOv5l、YOLOv5x。

表1 YOLO系列算法对比

Table 1 Comparison of YOLO series algorithms

算法	优点	缺点
YOLOv1	图像划分为网格单元,检测速度快	对密集目标和小目标检测效果不佳
YOLOv2	使用聚类产生锚框,提高分类精度	使用预训练,迁移难
YOLOv3	借鉴残差网络,实现多尺度检测	模型复杂,对中、大目标检测效果差
YOLOv4	在检测精度与速度间实现了平衡	检测精度有待提高
YOLOv5	模型尺寸小,灵活性高,检测速度快	性能有待提高

1.2 STA-YOLOv5算法整体框架

通过设计STA-YOLOv5网络结构解决安全帽是否佩戴及检测过程误检、漏检率高的问题。STA-YOLOv5网络结构如图2所示。整体框架主要由Input、Att、Backbone、Neck和Head等5部分组成。其中,输入Input为三通道的RGB图像。主干网络Backbone使用Swin Transformer^[9]网络,通过对特征图分解成小块进行信息跨层交流,有效捕获全局信息和局部信息之间的关系,增强提取特征。注意力模块Att通过引入融合通道和空间注意力机制,对提取特征更深层次学习,增强模型的检测能力。特征融合层Neck采用路径聚合网络(path aggregation network, PANet),对Backbone和Att输出的有效特征图进行融合,实现不同特征层的信息交融。PANet更好融合了浅层与深层的特征信息,充分提取到网络中各层次的特征,加强特征提取,得到更丰富的特征信息。输出检测部分Head有3个YOLO Head检测器,输出不同尺度特征图进行目标预测。

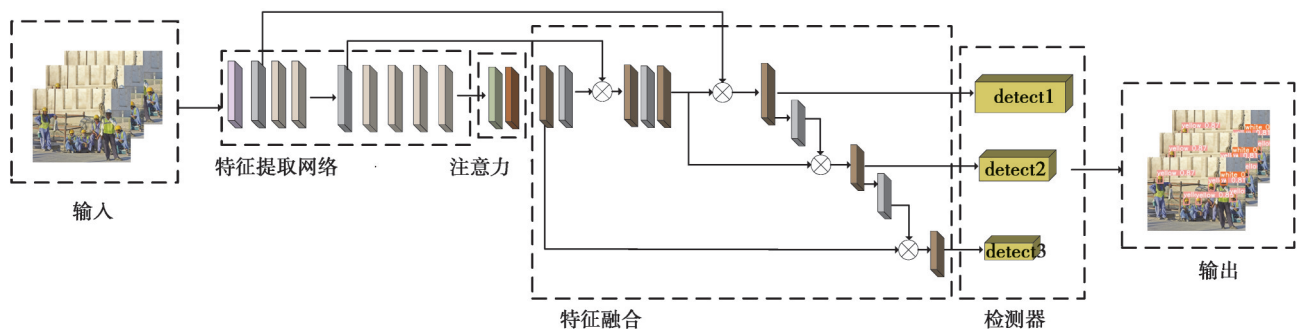


图2 STA-YOLOv5整体框架

Fig. 2 STA-YOLOv5 overall framework

2 基于 Swin Transformer 的安全帽特征提取网络

Swin-Transformer^[20-21]算法利用窗口多头注意力机制(windows multi-head self-attention,W-MSA)和窗口移动多头注意力机制(windows shift multi-head self-attention,WS-MSA)实现层级式 Transformer,可作为安全帽的特征提取网络,在应用到 YOLOv5 的算法结构中。Swin Transformer 整体网络结构如图 3 所示,采用类似 CNN 的层次化模型构建方法。

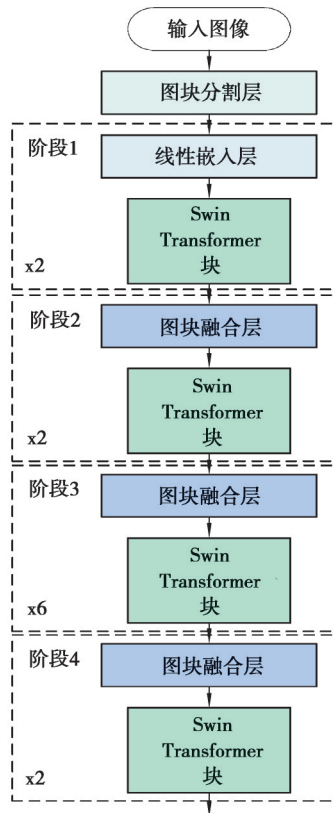


图 3 Swin Transformer 整体网络结构

Fig. 3 Overall network structure of Swin Transformer

整个网络结构主要由 4 个结构相似的阶段模块组成,除第一个是由线性嵌入层+ Swin Transformer 块组成之外,其余均是由图块融合层+ Swin Transformer 块组成。除第一个阶段外,每个阶段都会先通过图块融合层缩小输入的分辨率,对特征进行下采样操作扩大感受野,以获取全局信息。图块融合层是该环节中关键步骤(如图 4 所示)。在每个阶段开始前进行下采样操作,缩小图像的分辨率,调整通道数形成层次化设计,同时节省运算量。图块融合层通过在行方向和列方向间隔 2 个单位选取元素来进行二倍下采样操作,将得到的图像拼接在一起后展开,此时图像的高和宽缩小了 1/2,但是通道加深了 4 倍,再通过一个 1×1 的卷积对图像深度进行调整变为原来的 2 倍^[22]。图块融合层会增加一定计算量,但不会丢失信息,而 Pooling 往往会丢失一部分信息,这也是该网络提升准确率的一个因素。

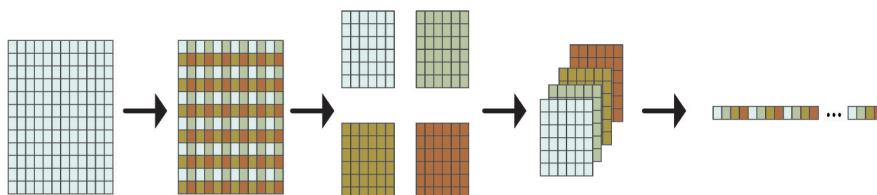


图 4 图块融合层

Fig. 4 Patch Merging

3 基于融合通道和空间注意力机制的安全帽检测

YOLOv5网络存在着不同程度的目标区域关注不全和对无关背景区域产生无效关注的问题,增大对安全帽检测难度。工人集体作业时,出现人员堆积密集,使安全帽形状不规则,边缘存在棱角分明或是圆润平滑的差异等问题,从而使检测变得困难。注意力机制通过给不同部分的特征图赋予权重选择部分特征图,抑制无用信息,以达到选择更优特征的目的。对此,研究设计一种融合通道和空间注意力(如图5所示),以加强对重要特征的利用,减少重要细节特征损失。

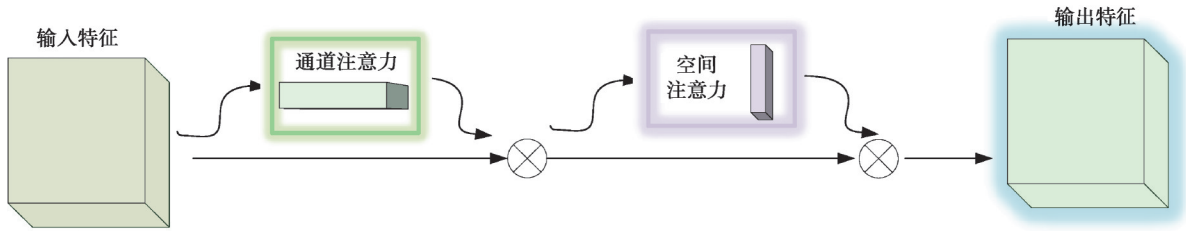


图5 融合通道和空间的注意力模块

Fig. 5 Attention module of fusion channel and space

图5由输入特征、通道注意力模块、空间注意力模块和输出特征组成。输入特征为 F_{in} ,输出特征为 F_{out} 。

$$F_{Mid} = M_c(F_{in}) \otimes F_{in}, \quad (1)$$

$$F_{out} = M_s(F_{Mid}) \otimes F_{Mid}, \quad (2)$$

输入特征 $F \in R^{C \times H \times W}$,然后经通道注意力模块 M_c 得到新特征 F_{Mid} 如式(1), F_{Mid} 经过空间注意力模块得到输出特征 F_{out} 。

3.1 SENet通道注意力

网络训练过程中,图像中安全帽的数量可能会出现类别不平衡、无关信息干扰多的情况,引入SENet模块解决此问题。SENet^[21-23](squeeze and excitation network)是一种应用通道注意力机制的深度卷积神经网络(CNN),自适应调整每个通道的权重,突出对当前任务最有用的特征通道,抑制对任务不重要的通道^[24]。

SENet的通道注意力机制通过压缩和激励2个关键步骤来实现。

1)压缩(Squeeze):通过全局平均池化操作对输入特征图进行降维,将每个通道的特征图压缩为单一值,如式(3)所示。这有助于捕获每个通道的全局感受野信息,并得到通道维度上的全局特征描述

$$Z_c = F_{sq}(U_c) = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W U_c(i,j). \quad (3)$$

2)激励(Excitation):使用一个多层感知机(MLP)来对每个通道的全局特征进行建模。通过2个全连接层(fully connected layer, FC)和函数ReLU,依次为FC- ReLU-FC,建立通道之间的相关性。再通过Sigmoid激活函数得到 C 个 $[0,1]$ 之间的权重 s ,如式(4)。MLP的目标是学习一种激励函数(excitation function),它基于全局特征生成每个通道的权重。这些权重用于对原始特征进行重新加权,增强有用特征通道的表示能力

$$s = F_{ex}(z, W) = \sigma(g(z, W)) = \sigma(W_2 \delta(W_1 z)). \quad (4)$$

最后通过乘法将通道加权到先前特征上,完成在通道维度上对原始特征的重标定,如式(5)。

$$\hat{x}_c = F_{scale}(u_c, s_c) = s_c \cdot u_c. \quad (5)$$

在整个SENet结构中,压缩和激励操作是串联在一起的(如图6所示)。

3.2 多尺度可变形卷积的空间注意力

针对由于出现拍摄角度、拍摄距离等问题,采用可变形卷积(deformable convolutional Networks, DCN)^[24-25]增强几何变换能力,构造了空间注意力模块如图7所示。同时选择了3种卷积核大小 $1 \times 1, 3 \times 3, 5 \times 5$ 以并行挖掘不同尺度的特征。为了得到权重信息,还需进一步对其进行 3×3 卷积,再经过sigmoid激活函数,将得到的权值大小映射到 $[0,1]$ 。

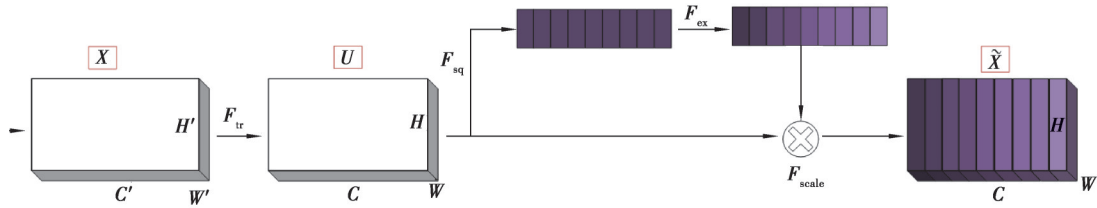


图 6 SENet 结构图

Fig. 6 SENet structure diagram

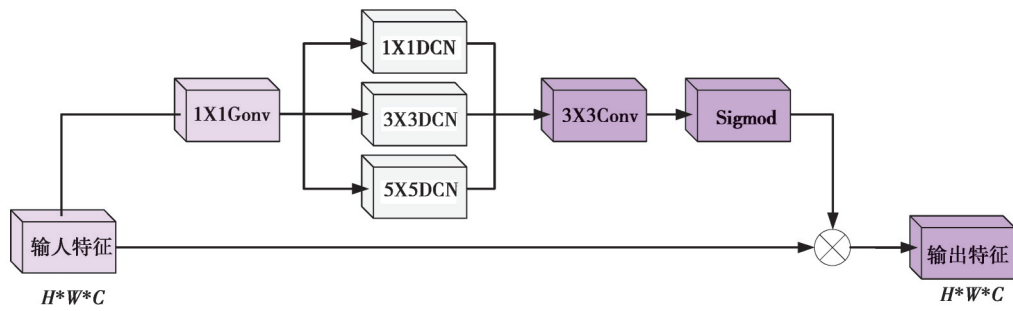


图 7 空间注意力模块

Fig. 7 Spatial attention module

卷积网络对大尺寸多形变目标的建模存在固有缺陷,因为卷积网络只对输入特征图的固定位置进行采样^[26]。可变形卷积可以提高模型对形变目标的建模能力,使用平行卷积层学习 offset 偏移,使得卷积核在输入特征图上的采样点发生偏移,可集中于人们感兴趣的区域或目标,即对卷积核中每个采样点的位置都增加了偏移量,可实现在当前位置附近随意采样而不局限于常规卷积的规则采样点。如图 8 所示为常规卷积和可变形卷积采样点的对比,其中,图 8(a)为常规 3×3 大小卷积核的采样方式,图 8(b)为可变形卷积的普通采样方式,图 8(c)与图 8(d)为可变形卷积采样的特殊方式。

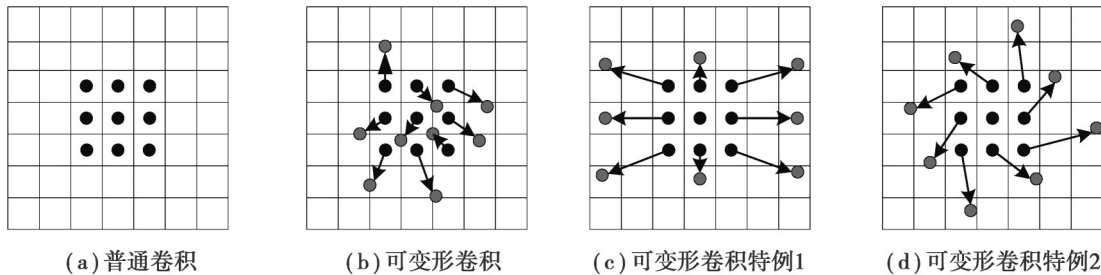


图 8 常规卷积和可变形卷积采样点对比

Fig. 8 Comparison of sampling points between conventional convolution and deformable convolution

3.3 安全帽检测算法

在 Head 检测层根据传递来的安全帽图像特征对边界框和类别进行预测,通过非极大值抑制(non maximum suppression,NMS)^[27]消除冗余预测框,最终输出置信度最高的预测类别,并返回边框坐标。

首先将安全帽图像划分成为 3 个 S*S 的网格。最后每个单元网格负责检测落入其中心的安全帽图像,同时输出多个预测框及其置信度。每个边界框包含参数 $(t_x, t_y, t_w, t_h, t_c)$,其中 (t_x, t_y) 为候选框的中心坐标, (t_w, t_h) 为候选框的中心点。 t_c 为置信度。在位置预测上,假设 Anchor Box 设置在每个网格单元的左上角,坐标位置为 (c_x, c_y) ,宽度和高度为 (p_w, p_h) ,最终生成的预测坐标为 (b_x, b_y, b_w, b_h) ,其公式为

$$bx = (2\sigma(t_x) - 0.5) + c_x, \tag{6}$$

$$b_y = (2\sigma(t_y) - 0.5) + c_y, \quad (7)$$

$$b_w = p_w (2\sigma(t_w))^2, \quad (8)$$

$$b_h = p_h (2\sigma(t_h))^2, \quad (9)$$

其中: (b_x, b_y) 为预测边界框的中心坐标; (b_w, b_h) 为预测框的高度和宽度。在输出多个预测框之后, 将抛弃掉置信度低的预测框并通过非极大抑制获得安全帽位置。

4 实验与分析

4.1 安全帽图像数据集

目前安全帽佩戴数据集主要有 SHWD 和 GDUT-HWD 2 种。而 SHWD 安全帽佩戴数据只有 2 个类别, head 和 helmet, 即安全帽佩戴与否, 类别较少。通过识别安全帽颜色, 管理层可以更好地监督和管理工地的工人活动, 也可区分不同的工种或职责。而在水利建造工地中, 受工人着装与场景特定色彩影响, 通过 rgb 信息获得安全帽颜色干扰较大。为使模型更具实用性, 选择具有 5 个类别的 GDUT-HWD 数据集, 分别是 Red (佩戴红色安全帽)、Blue (佩戴蓝色安全帽)、Yellow (佩戴黄色安全帽) 和 White (佩戴白色安全帽)、None (未佩戴安全帽), 这不仅可以识别是否佩戴安全帽, 同时还能检测佩戴安全帽的颜色。GDUT-HWD 数据集数量有 3 174 张, 包含 18 893 个实例, 还具有场景变化、视角变化、照明变化等, 使训练的模型适用于真实场景。安全帽数据集部分样本如图 9 所示。

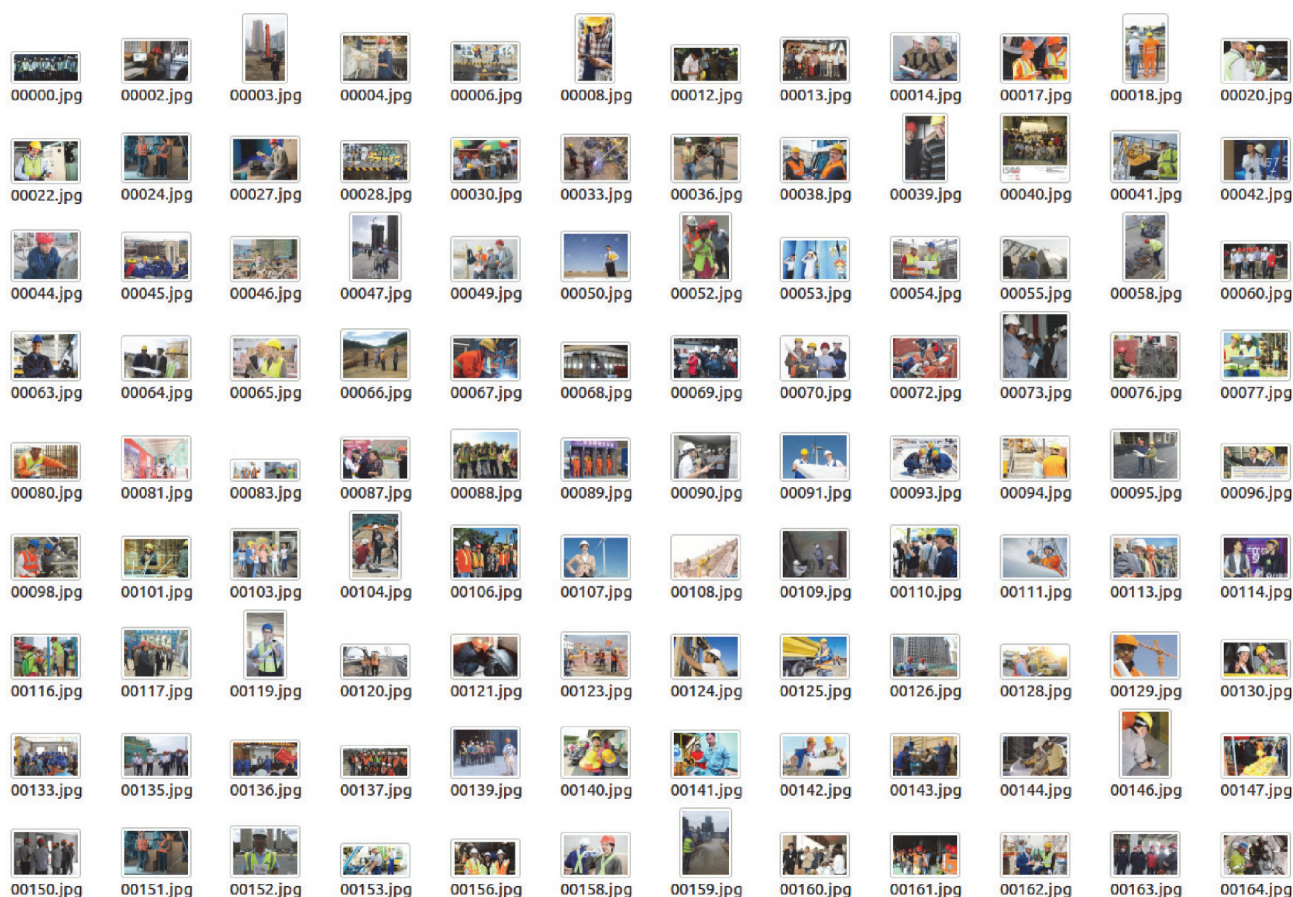


图9 安全帽数据集

Fig. 9 Hard hat data set

4.2 实验环境及配置

针对安全帽图像进行检测实验, 具体实验的环境及配置如下所示:

1) 操作系统: Ubuntu16.04 操作系统;

- 2)编译器:Pycharm 2020;
- 3)处理器: Intel®Xeon(R) CPU E5-2650 v4 @ 2.20 GHz×24;
- 4)内存:1 T;
- 5)显卡为:RTX2080Ti;
- 6)显存容量:12 GB
- 7)编译语言:Python 3.7.5

4.3 评价指标

4.3.1 平均精度 mAP

精度 AP 代表每个类别测试模型的性能,而 mAP 代表所有类别的测试模型性能,是所有 AP 的平均值。 mAP 公式为

$$mAP = \frac{1}{n} \sum_{q \in R} AP(q)。 \quad (10)$$

4.3.2 准确率 precision

准确率 $precision$ 表示为正确预测的正样本数量与预测为正样本的样本数量之间的比例。 $precision$ 公式为

$$precision = \frac{TP}{TP + FP}, \quad (11)$$

式中 TP (true positives) 为被识别正样本,实际上也是正样本的数量; FP (false positives) 为被识别负样本,实际上是负样本的数量。

4.3.3 FLOPs

FLOPs 指浮点运算数,理解为计算量,用来衡量算法/模型的复杂度。

4.4 实验结果与分析

4.4.1 算法模型主要实验参数表

STA-YOLOv5 模型主要参数设置如表 2 所示。

表 2 STA-YOLOv5 模型主要实验参数表
Table 2 Main experimental parameters of the STA-YOLOv5 model

超参数	变量
Img	3 174
训练集	2 857
测试集	317
Batch Size	8
Epoch	500
Optimizer	Adam
Learning rate	0.000 1

4.4.2 安全帽图像检测

水利建造人员安全帽佩戴检测效果如图 10 所示。

4.4.3 不同算法模型 loss

使用 $CIoU^{[28]}$ 作为 Bounding box 的损失函数,对于每个网络模型在 300 个批次训练过程当中,保存其所有的 loss 值,再对 loss 值进行可视化。为权衡安全帽检测速度与准确度,适应水利建造工地实时性与准确性需求,选用 YOLOv5 与 STA-YOLOv5 模型进行对比,不同模型的 loss 曲线对比如图 10 所示。由图 11 可知,在训练过程中两模型都没有出现过拟合与欠拟合现象,验证了模型改进的可行性。



图 10 水利建造人员安全帽佩戴检测效果图

Fig. 10 Hydraulic construction personnel safety helmet wearing test effect

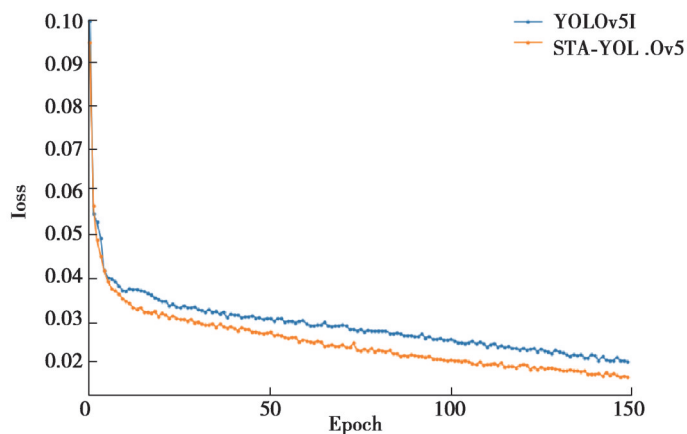


图 11 不同算法模型 loss 图

Fig. 11 loss diagram of different algorithm models

4.4.4 算法模型测试结果与分析

在实验中使用100张安全帽图像数据进行测试实验, YOLOv5l、STA-YOLOv5模型的测试实验结果如表3所示。

表 3 模型测试结果

Table 3 Model test results

指标 模型	mAP	Precesion	FLOPs	类别 mAP				
				Blue	Yellow	White	Red	None
STA-YOLOv5	0.914	0.916	128.5GFLOPs	0.952	0.932	0.954	0.815	0.918
YOLOv5l	0.891	0.896	108.3GFLOPs	0.932	0.922	0.942	0.731	0.903

由实验结果可知, STA-YOLOv5跟 YOLOv5l相比, 虽然模型复杂度稍显大, 但平均精确率与精确率分别提升了2.3个百分点和2个百分点, 且识别不同种类安全帽的平均精确率都有一定程度提升, 验证了 STA-YOLOv5模型能显著提升检测精确度, 满足实际复杂大型工地场景下安全帽佩戴的检测需求。

5 结语

为解决实际大型工地环境下安全帽是否佩戴问题,提出引入 Swin Transformer 和注意力机制的 YOLOv5 算法对安全帽进行检测。首先,使用更容易捕获图像全局信息和局部信息之间关系的 Swin Transformer 作为网络的特征提取器,使网络对安全帽特征的提取能力得到显著增强;其次,引入融合通道和空间注意力机制,使 YOLOv5 算法可在保持速度优势的同时提高定位精度和减少漏检安全帽的情况。通过将 2 种结构结合起来的 STA-YOLOv5 算法,成功实现对工人是否佩戴安全帽的实时检测任务。实验结果表明,该方法具有高效、准确、稳定等优点,可以应用于各类大型水利建筑行业。

参考文献

- [1] 张博,宋元斌,熊若鑫,等.融合人体关节点的安全帽佩戴检测[J].中国安全科学学报,2020,30(2):177-182.
Zhang B,Song Y B,Xiong R X,et al.Helmet-wearing detection considering human joint[J].China Safety Science Journal,2020,30(2):177-182.(in Chinese)
- [2] 朱晓春,王欣,马国力,等.改进 YOLO v3 算法的安全帽佩戴检测[J].南京工程学院学报(自然科学版),2020,18(4):23-26.
Zhu X C,Wang X,Ma G L,et al.Safety helmet wearing detection for improved YOLO v3 algorithm[J].Journal of Nanjing Institute of Technology (Natural Science Edition),2020,18(4):23-26.(in Chinese)
- [3] 胡晓栋,王国明.基于改进 YOLOv5 的安全帽检测算法研究[J].计算机时代,2023(6):76-81.
Hu X D,Wang G M.Research on helmet detection algorithm based on improved YOLOv5[J].Computer Era,2023(6):76-81.(in Chinese)
- [4] 王玲敏,段军,辛立伟.引入注意力机制的 YOLOv5 安全帽佩戴检测方法[J].计算机工程与应用,2022,58(9):303-312.
Wang LM,Duan J,Xin L W.YOLOv5 helmet wear detection method with introduction of attention mechanism[J].Computer Engineering and Applications,2022,58(9):303-312.(in Chinese)
- [5] 沈希忠,戚成.改进 YOLOX 的安全帽佩戴检测算法[J].浙江工业大学学报,2023,51(3):289-297.
Shen X Z,Qi C.Improved safety helmet wearing detection algorithm of YOLOX[J].Journal of Zhejiang University of Technology,2023,51(3):289-297.(in Chinese)
- [6] Vishnu C,Singh D,Mohan C K,et al.Detection of motorcyclists without helmet in videos using convolutional neural network [C]//2017 International Joint Conference on Neural Networks (IJCNN).May 14-19,2017.Anchorage,AK,USA:IEEE,2017:3036-3041.
- [7] Hao W A,Jza B.An intelligent vision based approach for helmet identification for work safety[J].Computers in Industry,2018,100:267-277.
- [8] 江新玲,杨乐,朱家辉,等.面向复杂场景的基于改进 YOLOX 的安全帽检测算法[J].南京师大学报(自然科学版),2023,46(2):107-114.
Jiang X L,Yang L,Zhu J H,et al.Safety helmet detection algorithm based on improved YOLOX_s for complex scenes[J].Journal of Nanjing Normal University (Natural Science Edition),2023,46(2):107-114.(in Chinese)
- [9] 李达,刘辉.针对小目标的 YOLOv5 安全帽检测算法[J].现代信息科技,2023,7(9):9-13.
Li D,Liu H.YOLOv5 helmet detection algorithm for small targets[J].Modern Information Technology,2023,7(9):9-13.(in Chinese)
- [10] 陈光,乔梁,黄晓明,等.基于目标跟踪的行为识别方法研究:以安全帽佩戴识别检测为例[J].黑龙江科学,2023,14(8):50-52.
Chen G,Qiao L,Huang X M,et al.Research on behavior recognition method based on target tracking:through taking helmet wear identification detection as an example[J].Heilongjiang Science,2023,14(8):50-52.(in Chinese)
- [11] Redmon J,Divvala S,Girshick R,et al.You only look once: unified, real-time object detection[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).June 27-30,2016.Las Vegas,NV,USA:IEEE,2016:779-788.
- [12] Redmon J,Farhadi A.YOLO9000: better, faster, stronger[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition(CVPR).July 21-26,2017.Honolulu,HI,USA:IEEE,2017:6517-6525.
- [13] Redmon J,Farhadi A.YOLOv3: an incremental improvement[EB/OL].2018:arXiv:1804.02767.https://arxiv.org/abs/1804.02767.
- [14] He K M,Zhang X Y,Ren S Q,et al.Spatial pyramid pooling in deep convolutional networks for visual recognition[J].IEEE

- Transactions on Pattern Analysis and Machine Intelligence, 2015, 37(9): 1904-1916.
- [15] Liu S, Qi L, Qin H F, et al. Path aggregation network for instance segmentation[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. June 18-23, 2018. Salt Lake City, UT, USA: IEEE, 2018:8759-8768.
- [16] 郑楚伟, 林辉. 基于 Swin Transformer 的 YOLOv5 安全帽佩戴检测方法[J]. 计算机测量与控制, 2023, 31(3): 15-21.
Zheng C W, Lin H. YOLOv5 helmet wearing detection method based on swin transformer[J]. Computer Measurement & Control, 2023, 31(3): 15-21. (in Chinese)
- [17] 乔炎, 甄彤, 李智慧. 改进 YOLOv5 的安全帽佩戴检测算法[J]. 计算机工程与应用, 2023, 59(11): 203-211.
Qiao Y, Zhen T, Li Z H. Improved helmet wear detection algorithm for YOLOv5[J]. Computer Engineering and Applications, 2023, 59(11): 203-211. (in Chinese)
- [18] Elfving S, Uchibe E, Doya K. Sigmoid-weighted linear units for neural network function approximation in reinforcement learning[J]. Neural Networks, 2018, 107:3-11.
- [19] Cao X, Zhang Y W, Lang S, et al. Swin transformer based YOLOv5 for small-object detection in remote sensing images[J]. Sensors, 2023, 23(7): 3634.
- [20] Wen X, Li B, Wang X W, et al. A Swin transformer-functionalized lightweight YOLOv5s for real-time coal-gangue detection [J]. Journal of Real-Time Image Processing, 2023, 20(3): 47.
- [21] 稽文, 刘全金, 黄崇文, 等. 基于 Swin-Transformer 的 YOLOX 交通标志检测[J]. 无线电通信技术, 2023, 49(3): 547-555.
Ji W, Liu Q J, Huang C W, et al. YOLOX traffic sign detection based on Swin-Transformer[J]. Radio Communications Technology, 2023, 49(3): 547-555. (in Chinese)
- [22] 任钰. 基于 Faster R-CNN 的小目标检测研究与应用[D]. 安庆: 安庆师范大学, 2022.
Ren Y. Research and application of small target detection based on faster R-CNN[D]. Anqing: Anqing Normal University, 2022. (in Chinese)
- [23] 张雪明, 茅健. 嵌入 SENet 的卷积神经网络的零件缺陷检测方法[J]. 农业装备与车辆工程, 2023, 61(1): 94-98.
Zhang X M, Mao J. A method of parts defect detection based on convolutional neural network embedded in SENet[J]. Agricultural Equipment & Vehicle Engineering, 2023, 61(1): 94-98. (in Chinese)
- [24] Dai J F, Qi H Z, Xiong Y W, et al. Deformable convolutional networks[C]//2017 IEEE International Conference on Computer Vision (ICCV). October 22-29, 2017. Venice, Italy: IEEE, 2017: 764-773.
- [25] 卢俊哲, 张铖怡, 刘世鹏, 等. 面向复杂环境中带钢表面缺陷检测的轻量级 DCN-YOLO[J]. 计算机工程与应用, 2023, 59(15): 318-328.
Lu J Z, Zhang C Y, Liu S P, et al. Lightweight DCN-YOLO for strip surface defect detection in complex environments[J]. Computer Engineering and Applications, 2023, 59(15): 318-328. (in Chinese)
- [26] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. June 18-23, 2018. Salt Lake City, UT, USA: IEEE, 2018: 7132-7141.
- [27] Neubeck A, Van Gool L. Efficient non-maximum suppression[C]//18th International Conference on Pattern Recognition (ICPR'06). August 20-24, 2006. Hong Kong, China: IEEE, 2006: 850-855.
- [28] Zheng Z H, Wang P, Liu W, et al. Distance-IoU loss: faster and better learning for bounding box regression[J]. Proceedings of the AAAI Conference on Artificial Intelligence, 2020, 34(7): 12993-13000.

(编辑 侯 湘)