

DOI:10.16356/j.1005-2615.2021.05.009

基于深度学习的先进陶瓷零件实时缺陷检测系统

马晨凯¹, 吴毅慧², 傅华奇³, 业宁¹

(1. 南京林业大学信息科学技术学院, 南京 210037; 2. 南京林业大学汽车与交通工程学院, 南京 210037;
3. 黑龙江工程学院电气与信息工程学院, 哈尔滨 150050)

摘要: 传统先进陶瓷零件检测与分类的主流方法为纯机械尺寸过滤和人工判断, 为解决其成本高、失误率高和损坏率高等问题, 提出了基于深度学习的多目标实时检测分类模型 (Multi-object real-time detection and classification model, MRDC)。该模型以 YOLOv3 为基础, 使用 SKNet 作为注意力机制进行特征重构提高精确度, 配合灰度图快速转化算法与跳帧检测方法提高检测速度, 可实现实时缺陷检测。对实际生产中的先进陶瓷零件进行采集训练, 多批次采集图像数据, 每批数据含多个陶瓷零件的 1 000 张图像, 平均精确率均值达到 99.19%, 用先进陶瓷零件生产线视频检验, 识别分类的正确率达到 100%, 可以保证每分钟检测 450~550 个零件。多目标实时检测分类模型拥有识别速度更快、识别准确率更高和零件不易损坏等优点, 可极大地节约生产原料与人力成本, 减少废品产出。

关键词: 先进陶瓷材料; 目标检测; 机器学习; 视频图像识别

中图分类号: TP391.4; TP183; TM619.23 **文献标志码:** A **文章编号:** 1005-2615(2021)05-0726-09

Real-Time Defect Detection System for Advanced Ceramic Parts Based on Deep Learning

MA Chenkai¹, WU Yihui², FU Huaqi³, YE Ning¹

(1. College of Information Science and Technology, Nanjing Forestry University, Nanjing 210037, China;
2. College of Automotive and Transportation Engineering, Nanjing Forestry University, Nanjing 210037, China;
3. College of Electrical and Information Engineering, Heilongjiang Institute of Technology, Harbin 150050, China)

Abstract: The mainstream detection and classification methods of traditional advanced ceramic parts are pure mechanical size filtering and manual judgment. To solve the problems of high cost, high error rate and high damage rate, a multi-object real-time detection and classification model (MRDC) based on deep learning is proposed. The model is based on YOLOv3, uses SKNet as an attention mechanism for feature reconstruction to improve accuracy, and cooperates with grayscale map fast transformation algorithm and frame skipping detection method to improve detection speed, which can realize real-time defect detection. Image data in multiple batches are collected, and each batch of data contains 1 000 images of multiple ceramic parts. The mean average precision reaches 99.19% when the advanced ceramic parts in actual production are collected and trained. The correct rate of recognition and classification reaches 100% when the production line video of advanced ceramic parts is used for inspection, which can guarantee to detect 450—550 objects per minute. The MRDC model has the advantages of faster recognition, higher recognition accuracy, and less damage to objects, which can greatly save production raw materials and labor costs, reduce scrap output, and protect the environment more.

Key words: advanced ceramic material; target detection; machine learning; video image recognition

收稿日期: 2020-10-28; 修订日期: 2020-12-08

通信作者: 业宁, 男, 教授, E-mail: yening@nifu.edu.cn。

引用格式: 马晨凯, 吴毅慧, 傅华奇, 等. 基于深度学习的先进陶瓷零件实时缺陷检测系统[J]. 南京航空航天大学学报, 2021, 53(5): 726-734. MA Chenkai, WU Yihui, FU Huaqi, et al. Real-time defect detection system for advanced ceramic parts based on deep learning[J]. Journal of Nanjing University of Aeronautics & Astronautics, 2021, 53(5): 726-734.

先进陶瓷采用新型制备工艺制成,含有先进陶瓷的整体产品具有安全性、绝缘性、隔热性和稳定性的优异性能,被广泛应用于航天、机械、电子和化工等高精度领域。根据市场研究机构 Research and Markets 的预测^[1],全球先进陶瓷市场规模到2021年将超过1 000亿美元,在国内先进陶瓷中仅1个品种的氧化铝陶瓷配件就达到了15亿件的使用量。由于需求量庞大,提升先进陶瓷正品率将极大地提高经济效益。为进一步降低先进陶瓷的次品数量,保证高精度产品的安全性,需要寻找能够更准确地对先进陶瓷零件的成品进行判断的方法。目前生产企业主要采用传统方式对先进陶瓷零件进行成品判断,包括纯机械尺寸过滤和人工判断两种检验方式,但这些方式都存在成本高、失效率高和损坏率高等问题。例如采用机械筛选时仅能筛除出尺寸不合格的零件而无法将存在破损的先进陶瓷零件一同筛除,还需要进一步采用人工筛除,不仅增加了成本,也存在一定程度的遗漏和损坏。

随着人工智能技术的发展,零件检测已经逐渐从依靠纯机械方式发展为利用计算机图像自动识别^[2]。柳云鹤等提出了基于遗传算法匹配训练实际机械零件图像样本,使用K近邻(K-nearest neighbor, KNN)分类器对其进行分类,不需要大量图像样本,并有着较高的识别率^[3]。匡逊君等利用了HU不变矩和仿射不变矩,提取了机械零件图像的组合矩特征,并应用支持向量机对零件分类,该方法无需大量样本,具有很好的泛化能力^[4]。岳晓峰等使用粒子群算法改进径向基函数神经网络,避免了网络结构和参数选取不当时出现其收敛慢甚至无法收敛的情况,使得分类的成功率得到很大提高^[5]。Hinton等首次使用深层卷积神经网络 AlexNet 分类大规模图像^[6]。Girshick等提出了R-CNN(Region convolutional neural network),并在目标检测上运用了该方法^[7],随后的Fast R-CNN^[8]和Faster R-CNN^[9]也极大地提高了检测效率。YOLO基于单独的End to end网络将目标检测转化为回归问题,合并了候选区和对象识别,在不遗失准确率的同时进一步提高了检测速度^[10]。

先进陶瓷零件外形有很多种,有内圆外六边形的螺母状、存在多凹槽的不规则条状以及中心至底面以长方体挖空的棱柱状等。本文选择的先进陶瓷样本外形是由上下两个空心圆柱组成,其中上半部分空心圆柱内径比下半部分空心圆柱外径稍小,下半部分内、外径差为上半部分内、外径差的

2~3倍。

在实际生产中,先进陶瓷零件的筛选对边框标注的精度要求不高,但是对识别速度和准确性有很高的要求。本文提出了一种多目标实时检测分类模型(Multi-object real-time defect detection and classification model, MRDC),利用优化后的灰度图转换算法对图像先期处理,采用SKNet模块进行特征重构,并使用了基于Darknet-53网络的YOLOv3^[11]对图像全局信息进行评估,经过多次训练后可以准确地检测先进陶瓷零件的位置并判断其是否为成品。

本文率先使用基于深度学习的人工智能技术识别分类先进陶瓷零件,在极大地提高了速度与准确率的同时实现了无接触分类,避免了目标检测与分类过程中可能对先进陶瓷零件造成的损伤。

1 背景知识

1.1 YOLO 算法原理

准确度和速度直接决定了目标检测模型的优劣^[12],随着研究的不断深入,深度卷积神经网络在精度上有着很大的提升^[13],Faster R-CNN等算法可以在识别前选出可能存在目标的候选区,在检测精确度和分类准确率上表现极佳,但是效率过低^[14]。YOLO不使用候选区,所以目标检测速度快,泛化性能较好;其输入一般选择416像素×416像素和RGB三通道图像,经过CNN神经网络的变换将得到1个7×7×30的张量,即49个30维的向量,其中存放了2个Bounding box的位置、2个Bounding box的置信度以及各个对象的概率。其中Bounding box的置信度为

$$\text{Confidence} = \text{Pr}(\text{Object}) \times \text{IOU}_{\text{pred}}^{\text{truth}} \quad (1)$$

式中:Pr(Object)表示对象存在的概率;IOU_{pred}^{truth}表示Bounding box与真实Bounding box的交并比,即两者的近似程度。

损失实为经过神经网络训练后的输出值与实际的目标值之间的偏差。YOLO的损失函数一般使用以上两者误差的平方和作为样本的整体误差,包括边框中心点误差、边框高度与宽度误差、边框内有无对象时的置信度误差以及各个对象的分类误差^[15],一般与30维向量中的一部分相对应,即

$$\text{Loss} = \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} \left[(x_i^j - \hat{x}_i^j)^2 + (y_i^j - \hat{y}_i^j)^2 \right] + \lambda_{\text{coord}} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{\text{obj}} \left[(\sqrt{w_i^j} - \sqrt{\hat{w}_i^j})^2 + \right.$$

$$\begin{aligned}
 & (\sqrt{h_i} - \sqrt{\hat{h}_i})^2 \Big] + \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{obj} (C_i - \hat{C}_i)^2 + \\
 & \lambda_{noobj} \sum_{i=0}^{S^2} \sum_{j=0}^B I_{ij}^{noobj} (C_i - \hat{C}_i)^2 + \\
 & \sum_{i=0}^{S^2} I_{ij}^{obj} \sum_{c \in \text{classes}} (P_i^j(c) - \hat{P}_i^j(c))^2 \quad (2)
 \end{aligned}$$

式中： x, y, w, C, P 为网络预测值； $\hat{x}, \hat{y}, \hat{w}, \hat{C}, \hat{P}$ 为标注值； I_{ij}^{obj} 仅可取值1或0，表示第*i*个网格的第*j*个 anchor box 是否对此 object 负责。

YOLOv1 简单地将图片划分为 49 个网格，但是这样会导致每个网格生成的 Bounding box 较少，从而使其在邻近目标和小目标的检测中表现较差。此外 YOLOv1 还会有更多的定位误差，往往最后目标框选的精度并不高^[16]。

1.2 YOLOv3 算法

YOLOv3 在 YOLOv1 的基础上基于残差网络 Resnet^[17] 搭建了有 53 个卷积层的 Darknet-53 网络以用于特征提取工作，此网络的组成部分为一系列尺度为 1×1 和 3×3 的卷积层^[18]，CBR 模块作为 YOLOv3 的基本组件之一可分为 CBR_1 和 CBR_3，对应着上述 1×1 和 3×3 的卷积核尺寸^[19]。Resnet 思想具体体现为建立 Residual 残差模块，即

在一些层之间设置快捷链路，并将快捷链路和 2 个卷积层组成 1 个残差组件，如图 1 所示，其中 $F(x)$ 为残差函数， $K(x)$ 为经残差网络处理后的输出值， x 为输入值。

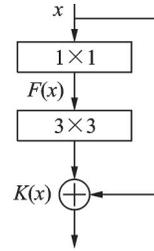


图1 残差网络结构图

Fig.1 Structural diagram of residual network

则残差网络函数可表示为

$$F(x) = K(x) - x \quad (3)$$

YOLOv3 将输入图像映射到 3 个尺度的输出张量，如果输入为 $416 \text{像素} \times 416 \text{像素}$ 、RGB 三通道图像，一般 3 种尺度分别为 13×13 、 26×26 、 52×52 ，那么总共具有 $13 \times 13 \times 3 + 26 \times 26 \times 3 + 52 \times 52 \times 3 = 10647$ 个预测。在每个预测中，对象存储在其中的 80 维，边框坐标与置信度分别存储在其中的 4 维和 1 维，共 $80 + 4 + 1 = 85$ 维，如图 2 所示。

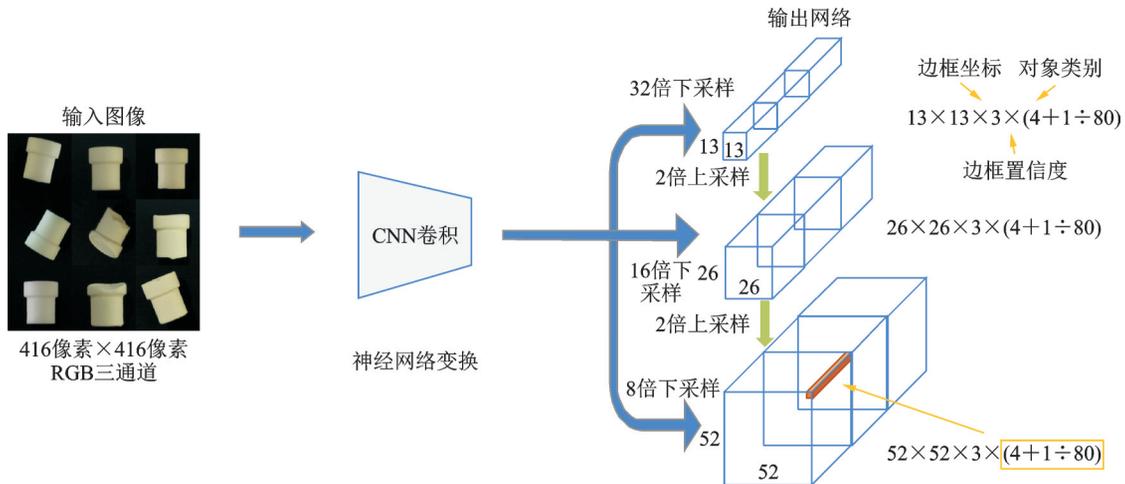


图2 YOLOv3 目标检测流程

Fig.2 YOLOv3 object detection process

不同尺度的感受野大小不同，每个尺度通过 K-means 聚类得到 3 个先验框，以 COCO 数据集为例，9 个先验框的具体尺寸如表 1 所示。

YOLOv3 同样用 ReLU 作为激活函数，输入图像经过 5 次步长为 2 的卷积层进行下采样，提升了平均精确率均值 (Mean average precision, mAP)，并且一定程度上解决了 YOLOv1 在小物体检测中表现较差的问题。

在边界框的预测上，YOLOv3 采用 Adam 优化器，预测对象类别采用单独的 Logistic 分类器取代传统的 Softmax 分类器^[20]，支持多标签对象同时预测框的类别、置信度和预测框。当物体的中心存在对应的网格时，就利用其对物体进行预测。

如图 3 所示，先验框用蓝色虚线表示，预测框用红色实线表示，其中： P_w 和 P_h 为先验框的宽和高； b_w 和 b_h 为实际预测框的宽和高。若从中心点

表 1 感受野尺寸表
Table 1 Receptive field size

特征图	感受野	先验框尺寸
13×13	大	116×90
		156×198
		373×326
26×26	中	30×61
		62×45
		59×119
52×52	小	10×13
		16×30
		33×23

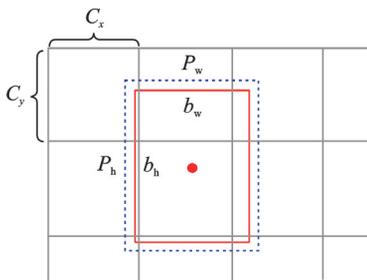


图 3 预测框示意图

Fig.3 Prediction boxes indicate

至网格最近的横坐标距离为 $\sigma(t_x)$,最近的纵坐标距离为 $\sigma(t_y)$,可以得到预测框表达式为

$$b_w = P_w e^{t_w} \quad (4)$$

$$b_h = P_h e^{t_h} \quad (5)$$

$$b_x = \sigma(t_x) + c_x \quad (6)$$

$$b_y = \sigma(t_y) + c_y \quad (7)$$

每个预测框使用非极大值抑制(Non-maximum suppression, NMS)算法,即选取得分最高的输出并将与其重叠的去掉, C_i 存在于第 j 个 Bounding box 的可能性如式(8)所示,其中 Score 为各边界框置信度下的条件概率。最后根据特征图与原图的映射关系在原图上输出预测框,从而实现先进陶瓷零件的定位。

$$\text{Score}_{ij} = P(C_i|\text{Object}) \times \text{Confidence}_j \quad (8)$$

2 多目标实时检测分类模型

本文提出一种效率与准确性较高的多目标实时检测分类模型 MRDC。该模型在上述 YOLOv3 算法的基础上添加了 SKNet(Selective kernel network)注意力机制,并结合灰度图转化算法的思想,采用跳帧追踪检测。因此,MRDC 模型在继承 YOLOv3 高速准确特点的同时可在视频目标检测与分类中进一步提升效率。

2.1 灰度图转化算法优化

一般使用的样本均为 sRGB 空间图像,将图像转为灰度图有多种算法,例如:直接将 RGB 求和取均值(式(9)),Photoshop 中经典的 Adobe RGB(式(10)),经典的非线性空间心理学灰度图转化公式(式(11))^[21]。

$$\text{Gray} = \frac{R + G + B}{3} \quad (9)$$

$$\text{Gray} = (0.2973R^{2.2} + 0.6274G^{2.2} + 0.0753B^{2.2})^{-2.2} \quad (10)$$

$$\text{Gray} = 0.299R + 0.587G + 0.114B \quad (11)$$

式中: R, G, B 取值为 0~255 之间的整数,分别代表红色、绿色和蓝色的颜色成分;Grey 为灰度值。

本文在式(11)算法的基础上进行优化。在实际训练和识别中,因式(11)需要进行大量的浮点数运算导致其效率大幅度降低。因此,本文将其放大 100 倍并使用移位代替除法,然后采用去尾法近似。由式(12~14)结果可得到式(15)的各项参数。在 16 位运算下,7 位精度的式(15)比式(11)精度更高,运算速度更快。

$$0.299 \times 2^7 = 38.272 \approx 38 \quad (12)$$

$$0.587 \times 2^7 + (38.272 \bmod 1) = 75.408 \approx 75 \quad (13)$$

$$0.114 \times 2^7 + (75.408 \bmod 1) = 15 \quad (14)$$

$$\text{Gray} = (38R + 75G + 15B) \gg 7 \quad (15)$$

为进一步提高效率,MRDC 模型采用式(15)对图像进行先期快速灰度图转化。

2.2 基于 SKNet 的特征重构

SKNet 为轻量的网络结构^[22],使用非线性方法使得感受野的尺寸根据激励因素自动变化,可达到随着输入尺度的不同自动地调整接受区域大小的目的。

SKNet 先进行 Split 操作,采用分组卷积的方法对于不同大小的卷积核进行完整卷积得到 \tilde{U} 和 \hat{U} 。随后进行 Fuse 操作融合信息,根据式(16)得到 U ,有

$$U = \tilde{U} + \hat{U} \quad (16)$$

然后输出 a_c 和 b_c 两个矩阵,以实现权重选择。最后进行 Select 操作,按照式(17)融合不同尺寸的 Feature map,有

$$V_c = a_c \cdot \tilde{U} + b_c \cdot \hat{U}, \quad a_c + b_c = 1 \quad (17)$$

大量使用残差网络的 YOLOv3 未使用 Darknet-53 的全连接层,如图 4 所示,在主干网络 3 个尺度的输出端后分别加入 SK 注意力机制模块,重新赋予权重以实现特征重构。

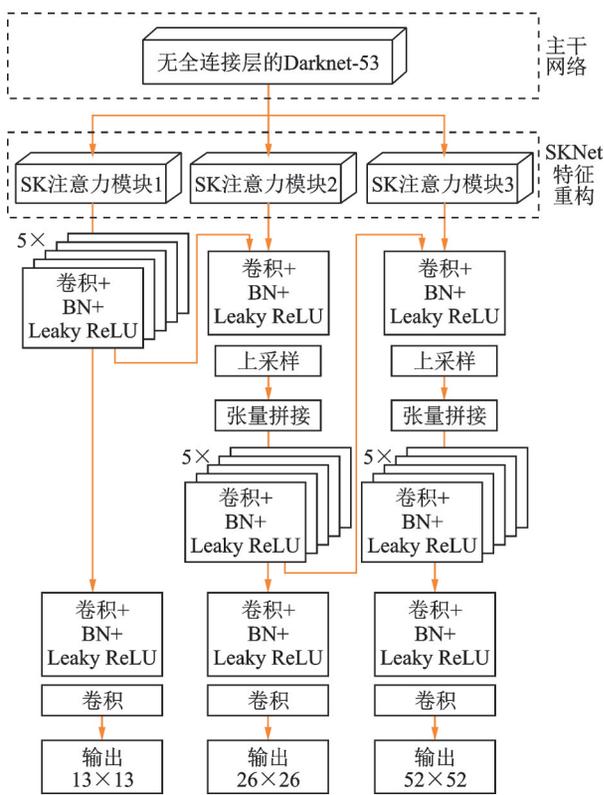


图4 SKNet特征重构图

Fig.4 SKNet feature reconstruction

2.3 MRDC 模型设计

结合对灰度图转化算法的优化,基于SKNet的特征重构与跳帧追踪检测算法,MRDC模型可在YOLOv3的基础上进一步提高准确率。模型步骤如下:

(1)采用优化后的灰度图转化算法对原始样本数据进行维数约减。

(2)自动判断是否需要归一化处理。为了适配更多的数据集,使MRDC模型有更好的扩展性,当数据集尺寸不一致时对其进行归一化处理。

(3)采用加入注意力机制后的YOLOv3算法对处理后的数据进行训练,利用SKNet进行特征重构,在实验中调整各项参数以得到最小损失的模型并进行预测,在经过非极大值抑制后得到预测结果,整体流程如图5所示。

(4)对视频进行预测时,每5帧取其中的第1帧,在对取到的第1帧预测前,将其从BGR格式转为灰度图,然后使用训练得到的模型进行预测。

(5)为保持预测结果并实现追踪目标,对其中第1帧预测得到的预测框坐标 (x_{ul}, y_{ul}) 、 (x_{dr}, y_{dr}) 与先进陶瓷零件在两帧之间平移的像素 Δx 进行计算,如式(18~19)所示,当 $n=4$ 时可分别得到另外

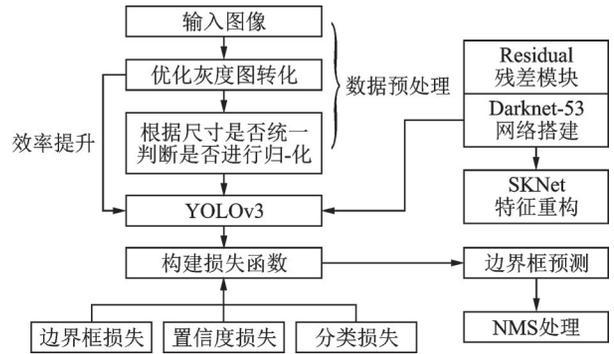


图5 模型训练与预测流程图

Fig.5 Model training and prediction process

4帧的框选坐标 (x'_{ul}, y'_{ul}) 、 (x'_{dr}, y'_{dr}) 。

$$x'_{ul} = \sum_{i=1}^n x_{ul} + \Delta x_i \quad (18)$$

$$x'_{dr} = \sum_{i=1}^n x_{dr} + \Delta x_i \quad (19)$$

将每一帧视频重新拼接得到输出视频,预测视频具体流程如图6所示。

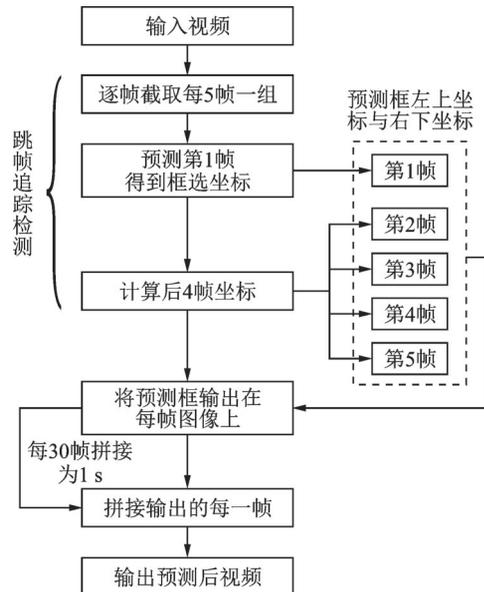


图6 视频预测流程图

Fig.6 Video forecasting process

3 实验过程和结果

基于YOLOv3与SKNet注意力机制的MRDC模型使用Python语言,并利用Keras框架搭建了Darknet-53网络;硬件为处理器i7-9700K、内存16 GB和显卡NVIDIA RTX 2070 Super;操作系统为Ubuntu 16.04 LTS。

3.1 实验准备

受齐齐哈尔市某特种陶瓷厂委托,采用MRDC模型解决实际先进陶瓷零件生产中的无接触筛选问题。但在其目标检测分类的研究中,未发

现公开数据集。

3.1.1 图像数据集建立

多批次运用数码相机采集数据,每次得到1 000张900像素×900像素的数量在1~9之间不等、位置散乱的先进陶瓷零件图,并将其制作为数据集,图例如图7所示。

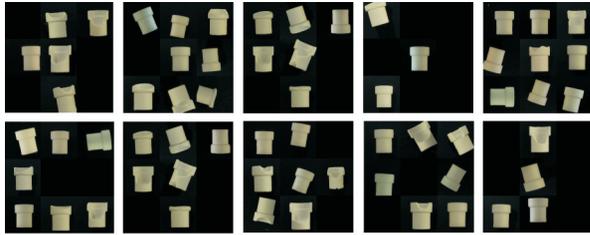


图7 数据集示例图

Fig.7 Sample data set diagram

使用开源工具labelImg将图片中的每个零件位置进行标注并设定分类,在全部标注后可得到包含了对象的位置和分类的1 000个XML格式的标注文件,最终得到先进陶瓷零件目标检测分类数据集。

3.1.2 视频采集

采集得到数段宽高比为900像素×900像素、格式为mp4的视频,用于实时测试模型效果。为逐列展示视频中的零件,每隔25~35帧对视频进行截图,如图8所示,后1张截图的第1、2列零件与前1张截图的2、3列零件相同。

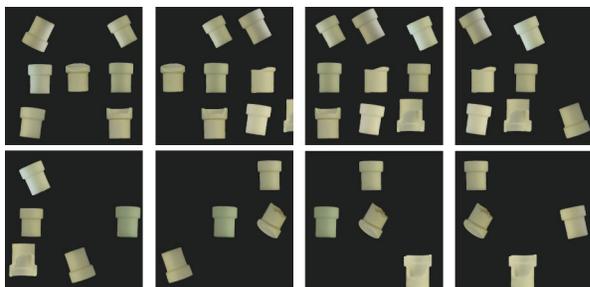


图8 视频示例图

Fig.8 Video examples

使用Cisco公司发布的开源H.264编码器Openh264对视频进行编码与解码。从视频截取的图像均为BRG格式,而训练所使用的样本为灰度图,使用在灰度图样本下训练得到的模型对视频进行目标检测与分类将直接导致准确率大幅下降。因此需要将每1帧图像从BRG格式快速转化为灰度图进行目标检测和分类。

3.2 数据预处理与SKNet部署

将每个数据标注文件中存放的多个先进陶瓷零件的位置信息与图像中的先进陶瓷零件一一对应,使用十次交叉验证方法将数据集分为10份,并

将其中的1份作为测试集,剩下的9份作为训练集。同时对数据使用随机旋转角度、随机水平翻转以及随机裁剪等增强策略,读取的bbox生成对应的3种尺寸的存储类别和真实框的中心宽高置信度的Label,Feed之前将图片Resize为416像素×416像素,最后将其转化为灰度图输入。

在Darknet-53的主干网络后加入3个SK注意力机制模块后得到使用SKNet进行特征重构后的模型。分别对原始YOLOV3模型和特征重构后的模型进行实验,检视两次实验的Loss,比对实验所得mAP,判定进行特征重构对多目标实时检测分类模型的价值。

3.3 模型训练与测试

在训练前修改YOLOv3的相关配置与参数,设置学习率为0.000 1,当Epoch达到100时结束训练,设置Batch_size为4,使用10%的数据作为测试集。调整SKNet的各项参数,设置额外路径数量SKNet_M为2,每个路径的基数SKNet_G为32, Fuse操作参数数量SKNet_r为16,参数含义与数值如表2所示。

表2 部分实验参数表

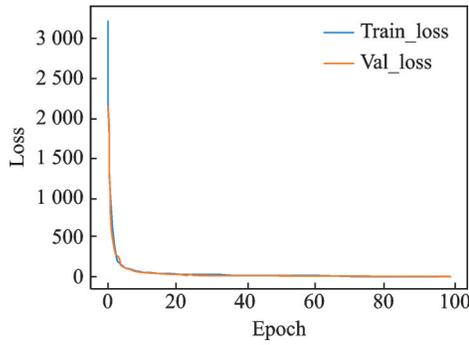
Table 2 Partial experimental parameters

参数名	参数含义	参数值
Epoch	迭代次数	100
Learning_rate	学习率	1e-4
Batch_size	单次训练样本数	4
Val_split	测试集比率	0.1
SKNet_M	额外路径数量	2
SKNet_G	单个路径基数	32
SKNet_r	Fuse操作参数	16

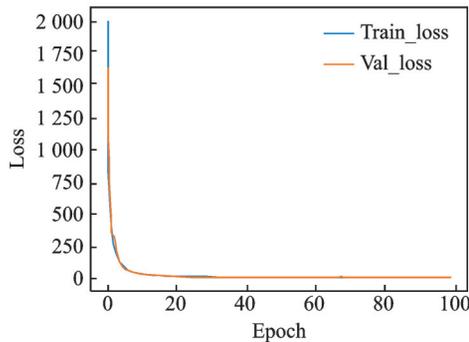
3.4 实验结果

学习率 Learning rate 直接影响模型训练结果,若学习率过高会导致无法优化,而学习率过小易导致优化过慢,出现掉入局部最优的情况^[23]。学习率设置为0.000 1,每次训练得到训练 Loss 和验证 Loss,经过100次训练后发现训练 Loss 和验证 Loss 都逐渐趋近于0,其中第88次训练得出的模型效果较好,两次训练得到的 Loss 如图9所示。特征重构后模型的部分训练结果如表3所示。

在训练100次后选择损失较低的第88次训练得到的模型,由于先进陶瓷零件对框选的精度要求不高,但是对分类判断的准确性要求较高,所以精确率 Precision 相比于召回率 Recall 更有参考意义,但由于最后精确性较高,故以平均精确率均值 mAP 值作为判断标准。多批次采集图像数据,每次采集含多个陶瓷零件的图像1 000张作为数据



(a) Loss value of original model



(b) Loss value of optimized model

图9 损失函数曲线对比

Fig.9 Loss function comparison

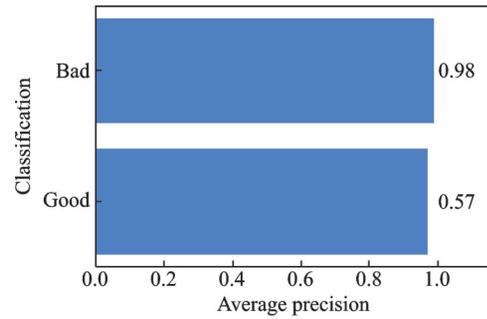
表3 损失值详情表

Table 3 Details of loss values

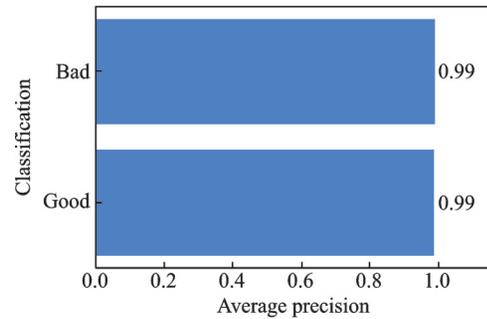
训练次数	学习率	训练Loss	验证Loss
1	1e-4	2 008.094 6	1 661.706 9
2	1e-4	408.425 1	345.792 8
3	1e-4	237.299 2	337.524 5
4	1e-4	163.146 7	148.025 8
5	1e-4	128.500 2	108.083 7
6	1e-4	100.291 2	88.648 2
7	1e-4	83.377 5	91.307 1
10	1e-4	71.504 9	72.533 9
20	1e-4	29.863 3	30.137 2
30	1e-4	25.028 3	23.990 0
50	1e-4	23.909 6	22.583 4
70	1e-4	21.078 3	19.899 4
88	1e-4	18.840 0	18.431 9

集,经过验证后得出原模型的 mAP 值为 97.57%,加入 SKNet 特征重构后模型的 mAP 值为 99.19%,这说明本文模型很好地实现了对目标的检测与分类,两种模型的平均精确率 AP 值如图 10 所示。

使用数据集外的多个图片样本对模型进行验证。测试后的结果表明,本文提出的 MRDC 模型时间效率较高,可以准确地对图片样本进行目标检测与分类。随机选择 1 个图片样本进行可视化测试,保持硬件配置不变,经过 32 ms 左右得到如图 11 所示结果,8 个个体的坐标均输出正确,其本身



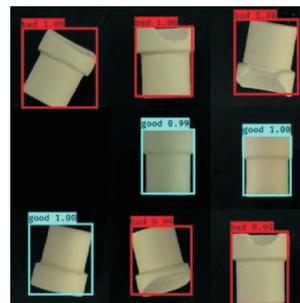
(a) AP of original model



(b) AP of optimized model

图10 平均精确率

Fig.10 Average precision



Found 8 boxes for img
 bad 0.99 (660, 665) (834, 886)
 bad 0.99 (360, 649) (545, 856)
 bad 1.00 (373, 57) (547, 268)
 bad 1.00 (670, 36) (864, 262)
 bad 1.00 (34, 62) (262, 302)
 good 0.99 (392, 362) (553, 557)
 good 1.00 (696, 380) (842, 562)
 good 1.00 (54, 642) (248, 856)
 5.7101977999999995

图11 图片样本可视化结果图

Fig.11 Visualization results of sample images

被框选的位置无误,分类准确率为 100%。

使用视频样本对模型进行验证。经测试可知,MRDC 模型可以高效准确地对视频中出现的每一个先进陶瓷零件进行框选、分类和追踪,可以保证每分钟检测 450~550 个零件。每 25~35 帧截取图片,如图 12 所示,视频中的先进陶瓷零件位置标记与分类准确率为 100%。

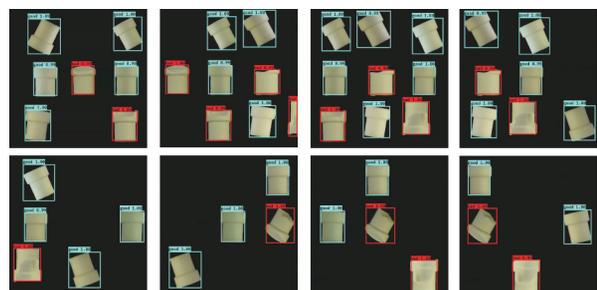


图12 视频样本可视化结果图

Fig.12 Visualization results of sample video

4 结 论

本文针对先进陶瓷零件无接触检测分类问题,基于YOLOv3算法与SKNet注意力机制提出了多目标实时检测分类模型MRDC,可对先进陶瓷零件图像进行快速灰度图转化,实时追踪视频中所有零件的位置并加以分类,克服了靠机械先期过滤加人工筛选方式成本高、准确性低和零件损伤率高的问题,可高效准确地对大批量的先进陶瓷零件进行目标检测与成品判断并可视化输出。在测试中,平均精确率可达到99.19%;批量测试中,由生成的图片样本可视化结果图和视频样本可视化结果图可知零件标注框位置准确、追踪及时,因此MRDC模型可以很好地投入生产实践。下一步将继续研究如何提高先进陶瓷零件的平均精确率,以及在更复杂的视频背景下检测分类更多样本的同时保证精确率,使得先进陶瓷零件无接触检测分类可以更好地适用于更加复杂的大批量生产实践中。同时,也将基于现有的多目标实时检测分类模型进一步优化生产流水线,配置自动分拣装置,实现不合格零件的自动剔除,改变实际生产现状,达到在每条生产线上减少6名分拣员的同时实现全天候24h自动化生产的目的,这有助于更高效准确地对次品零件进行二次回收,节约成本,减少废品产出,提高了经济效益的同时保护了环境,有助于先进陶瓷产业的发展。

参考文献:

- [1] 张脉官,姚素媛.谈高性能陶瓷辊棒的发展趋势[J].佛山陶瓷,2020,30(5):5-8,20.
ZHANG Maiguan, YAO Suyuan. Development trend of high performances ceramic rollers[J]. Foshan Ceramics, 2020, 30(5): 5-8, 20.
- [2] 李若尘,朱悠翔,孙卫民,等.基于深度学习的木材缺陷图像的识别与定位[J].数据采集与处理,2020,35(3):494-505.
LI Ruochen, ZHU Youxiang, SUN Weimin, et al. Recognition and localization of wood defect image based on deep learning[J]. Journal of Data Acquisition and Processing, 2020, 35(3): 494-505.
- [3] 柳云鹤,李苏.基于遗传算法的2D机械零件分类识别[J].传感器与微系统,2020,39(8):148-151.
LIU Yunhe, LI Su. Classification and recognition of 2D mechanical parts based on GA[J]. Transducer and Microsystem Technologies, 2020, 39(8): 148-151.
- [4] 匡逊君,赵文倩,李明富.基于图像不变矩和SVM的机械零件分类识别[J].制造业自动化,2012,34(15):65-68,72.
KUANG Xunjun, ZHAO Wenqian, LI Mingfu. Classifying and recognizing of mechanical parts based on the image invariant moment and SVM[J]. Manufacturing Automation, 2012, 34(15): 65-68, 72.
- [5] 岳晓峰,刘天.基于粒子群算法的RBF神经网络零件表面缺陷分类[J].电子技术与软件工程,2019(4):18-20.
YUE Xiaofeng, LIU Tian. Particle swarm algorithm based RBF neural network for part surface defect classification[J]. Electronic Technology & Software Engineering, 2019(4): 18-20.
- [6] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(5786):504-507.
- [7] GIRSHICK R, DONAHUE J, DARRELL T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation[C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Columbus: IEEE Computer Society, 2014: 580-587.
- [8] GIRSHICK R. Fast R-CNN[C]//Proceedings of International Conference on Computer Vision. Santiago: IEEE Press, 2015: 1440-1448.
- [9] REN S, HE K, GIRSHICK R, et al. Faster R-CNN: Towards real-time object detection with region proposal networks [C]//Proceedings of Advances in Neural Information Processing Systems. Montreal: MIT Press, 2015: 91-99.
- [10] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. Las Vegas: IEEE Computer Society, 2016: 779-788.
- [11] REDMON J, FARHADI A. YOLOv3: An incremental improvement [EB/OL]. (2018-04-08) [2020-05-28]. <https://arxiv.org/abs/1804.02767>.
- [12] 阮激扬.基于YOLO的目标检测算法设计与实现[D].北京:北京邮电大学,2019.
RUAN Jiyang. Design and implementation of object detection algorithm based on YOLO[D]. Beijing: Beijing University of Posts and Telecommunications, 2019.
- [13] 曹香滢,孙卫民,朱悠翔,等.基于科优先策略的植物图像识别[J].计算机应用,2018,38(11):3241-3245.
CAO Xiangying, SUN Weimin, ZHU Youxiang, et al. Plant image recognition based on family priority strategy[J]. Journal of Computer Applications, 2018, 38(11): 3241-3245.
- [14] 张婷婷,章坚武,郭春生,等.基于深度学习的图像目标检测算法综述[J].电信科学,2020,36(7):92-106.
ZHANG Tingting, ZHANG Jianwu, GUO Chunsh-

- eng, et al. A survey of image object detection algorithm based on deep learning[J]. Telecommunication Science, 2020, 36(7): 92-106.
- [15] 杨建伟,涂兴子,梅峰璋,等. 基于深度学习优化YOLOv3算法的芳纶带检测算法研究[J]. 中国矿业, 2020, 29(4): 67-72.
YANG Jianwei, TU Xingzi, MEI Fengzhang, et al. Research on aramid band detection algorithm based on deep learning optimization YOLOv3 algorithm [J]. China Mining Magazine, 2020, 29(4): 67-72.
- [16] 黄健,张钢. 深度卷积神经网络的目标检测算法综述[J]. 计算机工程与应用, 2020, 56(17): 12-23.
HUANG Jian, ZHANG Gang. Survey of object detection algorithms for deep convolutional neural networks [J]. Computer Engineering and Applications, 2020, 56(17): 12-23.
- [17] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]//Proceedings of IEEE Conference on Computer Vision and Pattern Recognition. [S.l.]: IEEE Computer Society, 2016: 770-778.
- [18] 徐守坤,倪楚涵,吉晨晨,等. 基于YOLOv3的施工现场安全帽佩戴的图像描述[J]. 计算机科学, 2020, 47(8): 233-240.
XU Shoukun, NI Chuhan, JI Chenchen, et al. Image caption of safety helmets wearing in construction scene based on YOLOv3[J]. Computer Science, 2020, 47(8): 233-240.
- [19] 冯加明,储茂祥,杨永辉,等. 改进YOLOv3算法的车辆信息检测[J]. 重庆大学学报, 2020, 61(6): 1-12.
FENG Jiaming, CHU Maoxiang, YANG Yonghui, et al. Vehicle information detection based on improved YOLOv3 algorithm[J]. Journal of Chongqing University, 2020, 61(6): 1-12.
- [20] 唐标,程志万,李博,等. 基于改进的YOLOv3在电力场景中的人群识别[J]. 电子器件, 2020, 43(3): 563-568.
TANG Biao, CHENG Zhiwan, LI Bo, et al. Crowd behavior recognition based on improved YOLOv3 in electric power surveillance[J]. Journal of Electron Devices, 2020, 43(3): 563-568.
- [21] 刘美. 彩色图像灰度化算法的研究与实现[D]. 长春: 长春理工大学, 2019.
LIU Mei. Research and implementation of color-to-gray algorithm[D]. Changchun: Changchun University of Science and Technology, 2019.
- [22] 李彬,王平,赵思逸. 基于双重注意力机制的图像超分辨重建算法[J]. 图学学报, 2021, 42(2): 206-215.
LI Bin, WANG Ping, ZHAO Siyi. Image super-resolution reconstruction based on dual attention mechanism[J]. Journal of Graphics, 2021, 42(2): 206-215.
- [23] 周明浩,朱家明. 基于卷积神经网络的机械零件表面缺陷分类[J]. 农业装备与车辆工程, 2019, 57(11): 19-23.
ZHOU Minghao, ZHU Jiaming. Classification of surface defects of mechanical parts based on convolution neural network[J]. Agricultural Equipment & Vehicle Engineering, 2019, 57(11): 19-23.

(编辑:刘彦东)