

基于改进 YOLOv3 网络的齿轮缺陷检测

张广世¹, 葛广英^{1*}, 朱荣华¹, 孙群²

¹聊城大学物理科学与信息工程学院, 山东 聊城 252059;

²聊城大学机械与汽车工程学院, 山东 聊城 252059

摘要 为解决工业制造中齿轮缺陷检测难的问题,提出一种基于改进的 YOLOv3 网络的缺陷检测方法。首先构建齿轮缺陷图像数据集,包括图像采集与扩充和缺陷标注;其次采用密集连接网络(DenseNet)结构代替原有的网络结构,提高特征提取能力;最后增加网络预测尺度,提高对于小尺寸缺陷的检测能力。利用齿轮缺陷图像对该方法进行验证,发现所提方法的平均精确率均值比 YOLOv3 网络提高了 3.87%,对齿轮缺失部分的精确率提高了 5.7%。与 YOLOv3 网络相比,所提方法在齿轮缺陷检测上有一定的先进性和有效性。

关键词 图像处理; 缺陷检测; 特征提取; 预测尺度; 平均精确率

中图分类号 TN919.8

文献标志码 A

doi: 10.3788/LOP57.121009

Gear Defect Detection Based on the Improved YOLOv3 Network

Zhang Guangshi¹, Ge Guangying^{1*}, Zhu Ronghua¹, Sun Qun²

¹ College of Physics and Information Engineering, Liaocheng University, Liaocheng, Shandong 252059, China;

² College of Mechanical and Automotive Engineering, Liaocheng University, Liaocheng, Shandong 252059, China

Abstract In this study, we propose an improved YOLOv3 network detection method to solve the problem that gear defects are difficult to detect in industrial manufacturing. First, a gear defect image database is constructed by performing various activities, including image acquisition and expansion and defect labeling. Second, the feature extraction ability is improved using the DenseNet network structure instead of the original network structure. Finally, the small-size defect detection ability is improved by increasing the network prediction scale. When compared with the YOLOv3 network, the mean average precision and the missing-part precision of the gear increased by 3.87% and 5.7%, respectively, using the proposed method. This experiment demonstrates that the proposed method exhibits several advantages and that the gear defects can be effectively detected.

Key words image processing; defect detection; feature extraction; prediction scale; average precision

OCIS codes 100.4996; 150.1135; 100.5010

1 引言

近年来,随着中国工业的快速发展,对设备中零件质量的要求越来越高。塑料齿轮质轻、价廉,传动噪声小,已被广泛应用于机械工业领域,其质量可直接决定整个机械系统的性能。对于齿轮中存在的缺失、污点等表面缺陷,传统的人工检测方法不但效率较低,且容易产生视觉疲劳,造成误判^[1]。因此,众多研究人员提出了各种方法来检测表面缺陷,这些方法主要分为三大类:机器视觉检测法、机器学习检

测法以及深度学习检测法。

机器视觉检测方法通常是先利用工业相机采集图像,然后使用图像处理算法得到结果。例如王文成^[2]对齿轮图像进行去噪、分割等预处理后,再与圆形蒙版图像对比,并设定相关阈值来检测齿轮缺失部分;伍玉琴等^[3]将齿轮图像进行二值化后,再与标准模板进行对比,以检测待测齿轮缺陷。这类方法的优点在于检测速度较快,缺点为可检测缺陷类型单一,泛化能力弱。

机器学习检测法通常先提取输入图片的图像特

收稿日期: 2019-08-30; 修回日期: 2019-09-25; 录用日期: 2019-10-31

基金项目: 中央引导地方科技发展专项资金

* E-mail: 406381534@qq.com

征,再使用模式识别方法进行判断。例如吴进仪^[4]将提取的缺陷图像多种特征输入到支持向量机和BP神经网络中进行缺陷检测;Er-Raoudi等^[5]首先提取缺陷图像的离散小波特征,然后使用主成分分析法进行降维,将处理之后的特征输入到人工神经网络中进行识别。但这些方法所提取的特征受人为因素干扰过大,会造成部分图像信息丢失,影响最终的检测率。

近年来,随着大数据与人工智能时代的来临,深度学习在图像分类^[6-7]、目标检测^[8]领域有着广泛的应用。深度学习下的目标检测算法主要分为两类:一类是先使用卷积神经网络来产生候选区域,然后将候选区域进行分类和回归运算,代表性算法为R-CNN(regions with convolutional neural networks)^[9]以及对于R-CNN的改进算法Fast R-CNN^[10]和Faster R-CNN^[11]等;另外一类算法是使用回归算法直接预测不同目标的类别与位置,其代表性算法为SSD(single shot multibox detector)^[12]以及YOLO(you only look once)^[13]网络。众多学者根据自己的研究对象以及这两种网络的特点进行了有针对性的改进。彭玉青等^[14]对YOLO网络的层数和参数进行调整,并使用空间金字塔池化(SPP)方法进行多尺度输入;Kim等^[15]对YOLOv3方法进行改进,同样采用SPP方法,并使用5个不同尺度的特征图以提高特征提取能力;常海涛等^[16]对Faster R-CNN进行改进,调整了锚的尺寸与数量,并对输入图像进行数据增强以提高最终识别效果。

本文将使用深度学习的方法对塑料齿轮的表面污痕和缺失部分进行检测,对YOLOv3网络进行改进,使用密集连接网络(DenseNet)代替原有模型中的DarkNet-53网络,以提高特征提取能力;构建齿轮缺陷数据集并进行数据增强,以防止过拟合;通过增加预测尺度来提高对于小尺寸缺陷的检测能力。

2 YOLOv3 网络

YOLOv3^[17]网络是在YOLO^[13]和YOLOv2^[18]网络的基础上发展而来的,与Faster R-CNN相比,YOLO网络将目标检测问题转换成回归问题,不再使用卷积神经网络对多个候选区域进行特征提取,而是直接通过回归来生成每个类的边界框坐标和概率,因此该算法的计算速度要远远超过Faster R-CNN。具体来讲,YOLO网络首先将输入图片分为 $S \times S$ 个单元格,如果待检测目标的

中心落在某个单元格,那么该单元格负责预测这个目标。每个单元格将预测 B 个边界框以及边界框的置信度。除此以外,每个单元格还需要给出由该单元格负责的边界框属于第 i 个类别的概率。

YOLOv3网络则是在YOLO网络的基础上进行改进。首先,YOLOv3使用了基于特征金字塔网络(FPN)的多尺度预测方法,通过选取多个尺度的特征图,显著提高了原始YOLO网络对于小目标的检测准确率;其次,更换原有YOLO特征提取网络,采用DarkNet-53网络来提取输入图像特征。YOLOv3具体架构如图1所示。

如图1所示,原始网络模型中的DarkNet-53网络与ResNet^[19]网络类似,都是采用残差单元来解决反向传播中的梯度消失问题,除此以外,该网络结构没有使用全连接层,而是采用了全卷积结构,可大幅度提高计算速度。

在YOLO网络结构中,损失函数可表示为

$$L_{\text{loss}} = E_{\text{coord}} + E_{\text{iou}} + E_{\text{cls}}, \quad (1)$$

式中: E_{coord} 为坐标误差; E_{iou} 为交并比(IOU)误差; E_{cls} 为分类误差。

$$E_{\text{coord}} = \lambda_c \sum_{i=1}^{S^2} \sum_{j=1}^B \alpha_{ij}^{\text{obj}} [(x_i - \bar{x}_i)^2 + (y_i - \bar{y}_i)^2] + \lambda_c \sum_{i=1}^{S^2} \sum_{j=1}^B \alpha_{ij}^{\text{obj}} [(w_i - \bar{w}_i)^2 + (h_i - \bar{h}_i)^2], \quad (2)$$

式中: λ_c 为坐标误差的权重; S^2 为输入图像的单元格数量; B 为每个单元格所预测的边界框数量; α_{ij}^{obj} 的取值为1或0,即第 i 个单元格的第 j 个边界框中是否存在检测目标,存在则为1,否则为0; (x_i, y_i) 、 w_i, h_i 分别为预测边界框的中心坐标、高度和宽度; (\bar{x}_i, \bar{y}_i) 、 \bar{w}_i, \bar{h}_i 分别为实际边界框的中心坐标、高度和宽度。参考原始YOLOv3网络,本研究将 λ_c 、 S 和 B 分别设置为5、7和9。

$$E_{\text{iou}} = \sum_{i=1}^{S^2} \sum_{j=1}^B \alpha_{ij}^{\text{obj}} (C_i - \bar{C}_i)^2 + \lambda_{\text{noobj}} \sum_{i=1}^{S^2} \sum_{j=1}^B \alpha_{ij}^{\text{obj}} (C_i - \bar{C}_i)^2, \quad (3)$$

式中: λ_{noobj} 为交并比误差的权重; C_i 为预测置信度; \bar{C}_i 为实际置信度。参考原始YOLOv3网络, λ_{noobj} 参数值设置为0.5。

$$E_{\text{cls}} = \sum_{i=1}^{S^2} \sum_{j=1}^B \alpha_{ij}^{\text{obj}} \cdot \sum_{c \in \text{classes}} [p_i(c) - \bar{p}_i(c)]^2, \quad (4)$$

式中: c 为检测目标所属的种类; $p_i(c)$ 为单元格中目标属于类别 c 的实际概率; $\bar{p}_i(c)$ 为预测概率。

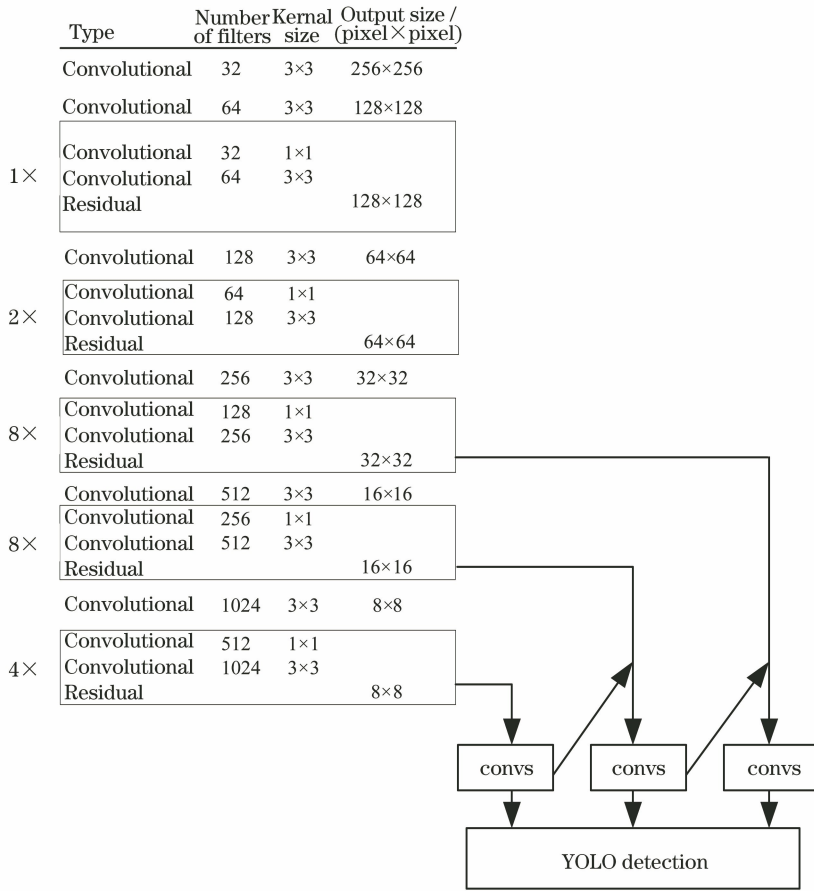


图1 YOLOv3 结构图

Fig. 1 YOLOv3 network structure

3 改进 YOLOv3 网络

3.1 具体改进措施

为了使 YOLOv3 模型适应于塑料齿轮的缺陷检测,将对原始结构进行如下修改。

3.1.1 引入 DenseNet 网络

在神经网络训练过程中,由于卷积和降采样操作的存在,特征信息会逐渐丢失。与基于 ResNet 的 DarkNet-53 网络相比,DenseNet 网络可更有效地提取特征信息^[20],从而解决层数过多导致的梯度消失问题。ResNet 和 DenseNet 的特征信息传播方式可分别表示为

$$x_{ResNet, l} = H_l(x_{l-1}) + x_{l-1}, \quad (5)$$

$$x_{DenseNet, l} = H_l(x_0, x_1, \dots, x_{l-1}), \quad (6)$$

式中: l 表示神经网络的层数; $H(\cdot)$ 为非线性转化函数; x_l 表示第 l 层的输出。对于 ResNet 网络,第 l 层的输出等于第 $l-1$ 层的输出加上 $H_l(x_{l-1})$,即对第 $l-1$ 层的输出进行非线性转化后的结果;而 DenseNet 网络是将第 0 层到第 $l-1$ 层的输出都进行非线性转化,然后在维度上进行连接作为第 l 层

的输出。相比于 ResNet 网络,DenseNet 网络可实现特征重用,提升网络计算效率,因此本研究采用 DenseNet 网络的思路进行图像特征提取。

3.1.2 增加预测尺度

YOLOv3 网络中使用多尺度检测方法来解决对小目标识别效果差的问题。首先,将原始输入图像进行卷积,将得到 $13 \text{ pixel} \times 13 \text{ pixel}$ 的特征图作为预测;然后,对其进行上采样,使得尺寸变为 $26 \text{ pixel} \times 26 \text{ pixel}$,与卷积过程中尺寸为 $26 \text{ pixel} \times 26 \text{ pixel}$ 的特征图在维度上相连接,变为新的 $26 \text{ pixel} \times 26 \text{ pixel}$ 的特征图,并以此为基础进行第二次预测;最后,将得到的 $26 \text{ pixel} \times 26 \text{ pixel}$ 尺寸的特征图进行上采样,与卷积过程中的 $52 \text{ pixel} \times 52 \text{ pixel}$ 尺寸的特征图相连接,形成新的尺寸为 $52 \text{ pixel} \times 52 \text{ pixel}$ 特征图,并进行第三次预测。通过这种方式,使得 $13 \text{ pixel} \times 13 \text{ pixel}$ 、 $26 \text{ pixel} \times 26 \text{ pixel}$ 、 $52 \text{ pixel} \times 52 \text{ pixel}$ 的特征图分别对应大尺寸、中尺寸、小尺寸的检测目标,很好地解决了小目标识别效果差的问题。

在塑料齿轮缺陷检测中,存在较小尺寸的齿轮

缺失区域,使用原始 YOLOv3 网络中的 3 种预测尺度不能满足要求,因此,本研究将原始尺度预测结构扩展成 4 种尺度,将 52 pixel×52 pixel 的特征图进行上采样,与卷积过程中的 104 pixel×104 pixel 尺寸的特征图连接,进行第 4 次预测,该特征图是输入图像的 4 倍下采样尺寸,其感受野适合检测齿轮缺陷中尺寸较小的缺失位置。

3.2 整体结构图

改进之后的总体网络结构如图 2 所示。从图 2

可以看出,所提算法首先将输入图片作卷积和最大池化操作,生成 64 个 104 pixel×104 pixel 尺寸的特征图,然后使用密集块(dense block)和过渡层(transition layer)模块提取特征。密集块中的增长率(growth rate)设置为 32,即密集块中每一层卷积输出的特征图数量为 32;同时借鉴了 DenseNet-BC^[18]的思想,在原有的 3×3 卷积块中增加 1×1 卷积块,这样密集块中每一层卷积操作可表示为 BN+ReLU+1×1-conv+BN+ReLU+3×3-conv。

Layer	Operation	Number of channels	Output size / (pixel×pixel)
Convolution	7×7 conv, stride is 2	64	208×208
Polling	3×3 maxpool, stride is 2	64	104×104
Dense block (1)	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 6$	256	104×104
Transition layer (1)	1×1 conv	128	104×104
	2×2 average pool, stride is 2		52×52
Dense block (2)	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 12$	512	52×52
Transition layer (2)	1×1 conv	256	52×52
	2×2 average pool, stride is 2		26×26
Dense block (3)	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 24$	1024	26×26
Transition layer (3)	1×1 conv	512	26×26
	2×2 average pool, stride is 2		13×13
Dense block (4)	$\begin{bmatrix} 1 \times 1 \text{ conv} \\ 3 \times 3 \text{ conv} \end{bmatrix} \times 16$	1024	13×13
Classification layer	7×7 global average pool		1×1
	1000D fully-connected softmax		

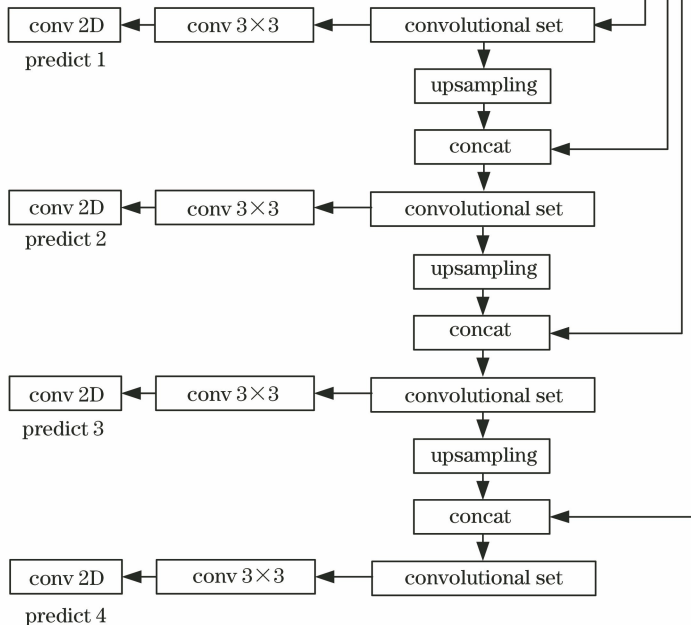


图 2 改进后的网络结构

Fig. 2 Improved network structure

1×1 卷积块为瓶颈层(bottleneck layer),该层的卷积核数量增长率为密集块的4倍,其作用在于降低特征数量,提高计算效率。过渡层包括 1×1 卷积操作和 2×2 平均池化操作,具体结构为BN+ReLU+ 1×1 -conv+ 2×2 -average pooling,该层的主要作用是连接相邻的密集块,并降低特征图尺寸。同时,过渡层也可以压缩模型,本研究将过渡层的压缩系数(compression rate)设置为0.5。

本研究增加了一种尺度预测,最终尺度分别为 $13 \text{ pixel} \times 13 \text{ pixel}$ 、 $26 \text{ pixel} \times 26 \text{ pixel}$ 、 $52 \text{ pixel} \times 52 \text{ pixel}$ 、 $104 \text{ pixel} \times 104 \text{ pixel}$,所设置的尺度预测方法和YOLOv3类似。以 $13 \text{ pixel} \times 13 \text{ pixel}$ 特征图与 $26 \text{ pixel} \times 26 \text{ pixel}$ 特征图融合为例,首先提取出卷积过程中的 $13 \text{ pixel} \times 13 \text{ pixel}$ 特征图,将其输入到convolutional set模块中,该模块的具体操作为 1×1 -conv+ 3×3 -conv+ 1×1 -conv+ 3×3 -conv+ 1×1 -conv,其中conv具体操作为conv 2D+BN+leaky ReLU。从该模块输出后,一方面将其进行下一步的 3×3 conv和conv 2D操作来对目标进行预测,另一方面将输出图像进行 1×1 -conv操作后上采样,与特征提取获得的 $26 \text{ pixel} \times 26 \text{ pixel}$ 大小的特征图进行拼接。其他尺度预测方法与该方法类似。

此外,原始YOLOv3网络是在ImageNet-1000数据集上进行训练,该数据集不包含本研究所需要的齿轮缺陷检测图片,为了提升检测效果,本研究采用k-means方法对锚框(anchor box)重新进行聚类,以适应齿轮缺陷检测的需要。

4 实验与结果分析

4.1 实验平台

本实验在Windows 10环境下完成,计算机内存为16 GB,CPU为Inter Core i3 8100 3.6 GHz,GPU为GTX1660ti,显存为6 GB;python版本为3.6.5,并同时安装了CUDA 10.0和cudnn 7.4.2以支持GPU的使用,安装了opencv 3.4、tensorflow 1.13等第三方库来支持神经网络的运行。

4.2 数据集

齿轮缺陷数据集的构建可分为2个部分:图像采集与扩充、数据标注。

4.2.1 图像采集与扩充

所使用的带有缺陷的齿轮共计200个,缺陷种类为污痕和缺失。采用工业相机对齿轮进行拍摄,每个图片包含2~6个带有缺陷的齿轮。图像采集设备为陕西维视数字图像技术有限公司生产的型号

为MV-EM120M/C的CCD面阵工业相机,镜头型号为MP系列的AFT-1214MP百万像素工业镜头。通过采集设备共获得450幅原始图像,且均为 $1280 \text{ pixel} \times 960 \text{ pixel}$ 的三通道彩色图像。

为了防止训练过程中出现过拟合,对图像进行数据扩增。通过对原始图像进行 90° 、 180° 和 270° 的旋转以及对对比度和亮度调整的方式来扩充图像集;还通过添加高斯白噪声的方式进一步扩充图像样本数量。扩充后的图像共4500幅,训练集和测试集的图像数量比例为9:1,分别为4050幅和450幅。采用机器视觉软件Halcon来实现上述过程。

4.2.2 数据标注

与用于图像分类的数据集不同,用于缺陷检测的图像数据集需要提供图像中缺陷位置的坐标,为此本研究使用labelImg软件来标注缺陷位置。该软件具体操作如下:首先使用该软件打开样本图像,在图像中用矩形框出缺陷位置,并且对不同类型的缺陷标记相应的标签,标记完成后,该软件会生成样本图像对应的XML文件,XML文件包含所对应的图像路径、标签和缺陷坐标等信息;然后编写格式转换程序,将XML文件转换为包含路径、标签和坐标的txt格式文件。

4.3 模型评价

为方便叙述,将改进后的网络命名为YOLOv3_Dense4,采用原有预测尺度但使用DenseNet的网络命名为YOLOv3_Dense3,原始网络命名为YOLOv3。YOLOv3_Dense4网络训练过程中的各项参数如下:动量参数为0.9,权值衰减系数为0.005,学习率为0.001,批尺寸为2,共训练200个epoch。对训练集的交并比和损失函数值进行记录,随着epoch数量的增加,训练集的损失函数和交并比的变化趋势如图3所示。

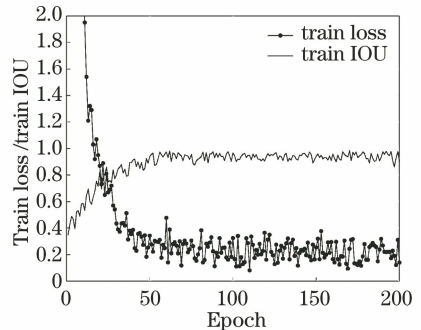


图3 YOLOv3_Dense4网络训练集损失函数和交并比曲线
Fig. 3 Loss function and IOU curves of train dataset of YOLOv3_Dense4 network

由图 3 可知:训练集中的损失函数值在训练开始时候较高,训练到第 50 个 epoch 左右趋于收敛,其值稳定在 0.1~0.4 范围内;与此同时,交并比也在第 50 个 epoch 趋于收敛,稳定在 90%~95% 范围内,这说明 YOLOv3_Dense4 网络的训练效果是有效的。本研究同时给出训练集的损失函数和交并比曲线图(图 4),以便于查看神经网络的性能。

由图 4 可知:测试集的损失函数值随着训练的进行,在第 125 个 epoch 左右趋于收敛,其值稳定在 0.2~0.6 范围内;交并比也同样在第 125 个 epoch 左右趋于收敛,收敛后的曲线中心值稳定在 90% 左右。对比图 3 和图 4 可知,测试集和训练集交并比和损失函数值随着训练的进行均趋于收敛,测试集的收敛速度更慢,但综合测试集的表现来看,所提

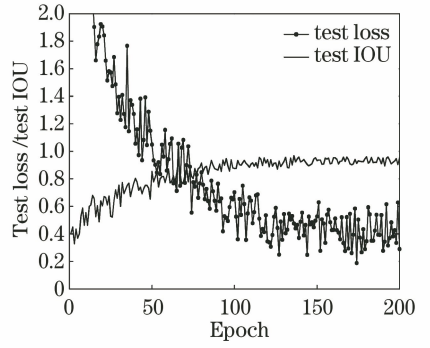


图 4 YOLOv3_Dense4 网络测试集损失函数和交并比曲线
Fig. 4 Loss function and IOU curves of test dataset of YOLOv3_Dense4 network

网络模型是有效的。训练完成后,调用生成的权重文件对缺陷齿轮进行检测,结果如图 5 所示。

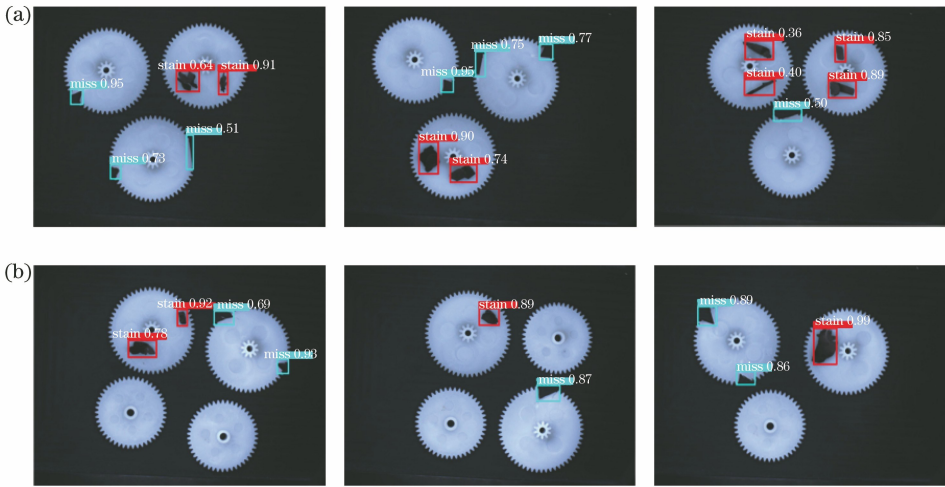


图 5 齿轮缺陷检测结果。(a)不同缺陷齿轮检测;(b)缺陷齿轮和无缺陷齿轮检测

Fig. 5 Defection results of defect gear. (a) Detection results of gears with different defects; (b) detection results of defect gears and flawless gears

在图 5 中,缺失位置和污痕位置分别以 miss 和 stain 标注,并使用矩形框标出缺陷位置,所标注的数字为相应的置信度。由图 5 可知,对于不同缺陷齿轮检测以及在没有缺陷齿轮中检测出带有缺陷的齿轮这两种情况,本文算法都取得不错的效果。

4.4 算法性能对比

为验证本文方法的有效性,将 YOLOv3、YOLOv3_Dense3、YOLOv3_Dense4 的性能进行对比,具体评价标准为单类精确率(P)、召回率(R)、 F_1 值、平均精确率均值(R_{mAP})和单幅图像平均检测时间(time),具体公式如下

$$P = P_{TP} / (P_{TP} + P_{FP}), \quad (7)$$

$$R = P_{TP} / (P_{TP} + P_{FN}), \quad (8)$$

$$F_1 = 2PR / (P + R), \quad (9)$$

$$R_{mAP} = (P_{stain} + P_{miss}) / 2, \quad (10)$$

式中: P_{TP} 表示被检测出的缺陷数量; P_{FP} 表示没有缺陷却给出缺陷检测框的数量; P_{FN} 表示有缺陷却没有标记缺陷位置的数量; P_{stain} 和 P_{miss} 分别表示对污痕和缺失的检测精确率。此外,为了对比不同网络性能,拍摄了 150 张带有缺陷齿轮图片进行缺陷检测,检测结果如表 1 所示。

由表 1 可知:相对于原始的 YOLOv3 网络,采用 YOLOv3-Dense4 结构的网络 R_{mAP} 值提高了 3.87%;相比于 YOLOv3_Dense3, YOLOv3_Dense4 的 R_{mAP} 提高了 0.83%;相比于 YOLOv3 网络, YOLOv3_Dense4 网络对于齿轮缺失部分的检测能力明显提高,其精确率提高了 5.70%,召回率提高了 17.48%;相对于 YOLOv3_Dense3 结构, YOLOv3_

表 1 不同方法性能对比

Table 1 Performance comparison of different methods

Net	Class	Number of defects	P_{TP}	P_{FP}	P_{FN}	$P / \%$	$R_{mAP} / \%$	$R / \%$	$F_1 / \%$	Time / s
YOLOv3	Stain	786	725	21	61	97.18	94.58	92.24	94.65	0.098
	Miss	572	451	39	125	91.98		78.15	84.50	
YOLOv3_Dense3	Stain	786	756	10	30	98.69	97.62	96.18	97.42	0.107
	Miss	572	475	17	97	96.54		83.04	89.29	
YOLOv3_Dense4	Stain	786	765	6	21	99.22	98.45	97.33	98.27	0.104
	Miss	572	547	13	25	97.68		95.63	96.64	

Dense4 网络的精确率和召回率也有了 1.14% 和 12.59% 的提高;使用 DenseNet 的网络结构单张图片检测时间要稍微高于 YOLOv3 结构,但总体相差不大。综上所述,本研究采取 4 种尺度的检测方法是有效的。

此外,为了查看本文算法在无缺陷齿轮中检测出带有缺陷的齿轮的能力,另拍摄了 100 张图片,其中每个图片包含 3 个无缺陷齿轮和 2 个带有缺陷的齿轮。所采用的评价标准为

$$r = w/n, \quad (11)$$

式中: r 为误检率; w 为无缺陷齿轮被判断为有缺陷齿轮的数量; n 为无缺陷齿轮总数。结果显示,仅有

表 2 不同光照强度下 YOLOv3_Dense4 性能对比

Table 2 Performance comparison of YOLOv3_Dense4 under different light intensities

Light intensity /lx	Class	Number of defects	P_{TP}	P_{FP}	P_{FN}	$P / \%$	$R_{mAP} / \%$	$R / \%$	$F_1 / \%$
394.93	Stain	256	248	4	8	98.41	97.80	96.88	97.64
	Miss	218	208	6	10	97.20		95.41	96.30
321.61	Stain	256	247	3	9	98.80	98.00	96.48	97.63
	Miss	218	209	6	9	97.21		95.87	96.54
249.28	Stain	256	241	3	15	98.77	98.21	94.14	96.40
	Miss	218	207	5	11	97.64		94.95	96.28

对比表 2 中的各项参数可以看出,在 3 种不同的光照强度下,YOLOv3_Dense4 网络检测结果相差并不大,只有对于 249.28 lx 下的污痕检测的召回率相对较低,比其他光照强度下的召回率小 2%~3%,具体表现为未检测出污痕缺陷的数量增多,但

4 个无缺陷齿轮被错误标记为有缺陷齿轮,误检率为 1.3%,可满足工业检测的要求。

4.5 不同光照强度下算法性能实验

在实际的工业应用中,光照强度对最终缺陷检测的结果有着很大的影响,为此,对 YOLOv3_Dense4 网络在不同光照强度下的具体性能进行测试。其实验条件如下:在暗室中拍摄图片,避免其他光源影响;环形光源位于齿轮正上方 10 cm 左右,相机镜头位于齿轮正上方 30 cm 左右,且保持光圈大小不变,仅改变环形光源的光照强度。在 3 种不同的光照强度下各拍摄多张缺陷齿轮图片,然后进行缺陷检测。具体检测结果如表 2 所示。

总体误差在可接受范围以内。不同光照强度下的结果如图 6 所示。从图 6 可以看出,在这 3 种光照强度下,缺陷齿轮中的污痕和缺失都可以检测出来,结合表 2 可知,YOLOv3_Dense4 网络在不同的光照强度下都有着较好的检测效果。

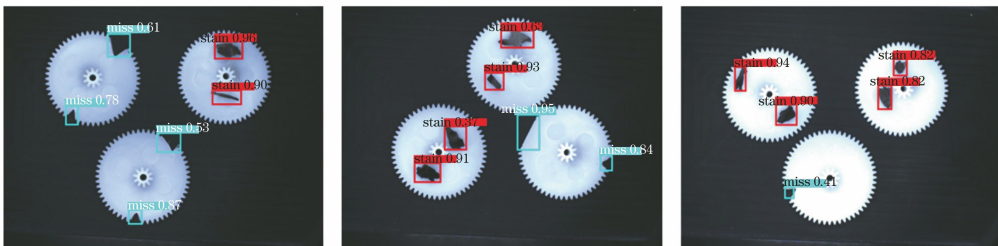


图 6 不同光照强度下齿轮缺陷的检测结果。(a) 249.28 lx;(b) 321.61 lx;(c) 394.93 lx

Fig. 6 Detection results of gear defects under different light intensities. (a) 249.28 lx; (b) 321.61 lx; (c) 394.93 lx

5 结 语

使用一种改进的 YOLOv3 网络对工业生产中的塑料齿轮进行缺陷检测,通过构建齿轮缺陷数据集来提高神经网络的泛化能力以及避免过拟合;使用 DenseNet 代替 YOLOv3 网络中的 DarkNet-53 网络,可更有效地提取图像特征;通过扩大神经网络预测尺度来提高对小尺寸缺陷的检测能力。由实验结果可知,所提方法取得了不错的效果,具体表现在平均精确率均值的提高、对尺寸缺失部分的检测精确率和召回率的提高。下一步工作是继续调整所提算法的结构,加快检测速度。

参 考 文 献

- [1] Li S B, Yang J, Wang Z, et al. Review of development and application of defect detection technology[J/OL]. Acta Automatica Sinica[2019-07-24]. <https://doi.org/10.16383/j.aas.c180538>.
李少波, 杨静, 王铮, 等. 缺陷检测技术的发展与应用研究综述[J/OL]. 自动化学报[2019-07-24]. <https://doi.org/10.16383/j.aas.c180538>.
- [2] Wang W C. Design of gear defect detection system based on the Halcon[J]. Journal of Mechanical Transmission, 2014, 38(9): 60-63, 98.
王文成. 基于 Halcon 的齿轮缺陷检测系统设计[J]. 机械传动, 2014, 38(9): 60-63, 98.
- [3] Wu Y Q, Xu H Y, Cheng X J. Research on gear defect recognition based on machine vision[J]. Coal Mine Machinery, 2019, 40(4): 170-172.
伍玉琴, 徐海元, 成晓军. 基于机器视觉的齿轮缺陷识别研究[J]. 煤矿机械, 2019, 40(4): 170-172.
- [4] Wu J Y. Research on lining defect detection system based on machine vision[D]. Hangzhou: Zhejiang University of Technology, 2019.
吴进仪. 基于机器视觉的衬布缺陷检测系统研究[D]. 杭州:浙江工业大学, 2019.
- [5] Er-Raoudi M, Diany M, Aissaoui H, et al. Gear fault detection using artificial neural networks with discrete wavelet transform and principal component analysis[J]. Journal of Mechanical Engineering and Sciences, 2016, 10(2): 2016-2029.
- [6] Zhuo D, Jing J F, Zhang H H, et al. Classification of chopped strand mat defects based on convolutional neural network [J]. Laser & Optoelectronics Progress, 2019, 56(10): 101009.
卓东, 景军锋, 张缓缓, 等. 基于卷积神经网络的短切毡缺陷分类[J]. 激光与光电子学进展, 2019, 56(10): 101009.
- [7] Yuan L S, Lou M Y, Liu Y Q, et al. Palmvein classification based on deep neural network and random forest [J]. Laser & Optoelectronics Progress, 2019, 56(10): 101010.
袁丽莎, 娄梦莹, 刘娅琴, 等. 结合深度神经网络和随机森林的手掌静脉分类[J]. 激光与光电子学进展, 2019, 56(10): 101010.
- [8] Zhao H, An W S. Image salient object detection combined with deep learning [J]. Laser & Optoelectronics Progress, 2018, 55(12): 121003.
赵恒, 安维胜. 结合深度学习的图像显著目标检测[J]. 激光与光电子学进展, 2018, 55(12): 121003.
- [9] Girshick R, Donahue J, Darrell T, et al. Rich feature hierarchies for accurate object detection and semantic segmentation [C] // 2014 IEEE Conference on Computer Vision and Pattern Recognition, June 23-28, 2014, Columbus, OH, USA. New York: IEEE, 2014: 580-587.
- [10] Girshick R. Fast R-CNN [C] // 2015 IEEE International Conference on Computer Vision (ICCV), December 7-13, 2015, Santiago, Chile. New York: IEEE, 2015: 1440-1448.
- [11] Ren S Q, He K M, 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, 2017, 39(6): 1137-1149.
- [12] Liu W, Anguelov D, Erhan D, et al. SSD: single shot multibox detector[M] // Computer Vision-ECCV 2016. Cham: Springer International Publishing, 2016: 21-37.
- [13] 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. New York: IEEE, 2016: 6517-6525.
- [14] Peng Y Q, Zhao X S, Tao H F, et al. Hand gesture recognition against complex background based on deep learning[J]. Robot, 2019, 41(4): 534-542.
彭玉青, 赵晓松, 陶慧芳, 等. 复杂背景下基于深度学习的手势识别[J]. 机器人, 2019, 41(4): 534-542.
- [15] Kim K J, Kim P K, Chung Y S, et al. Performance enhancement of YOLOv3 by adding prediction layers with spatial pyramid pooling for vehicle detection[C] // 2018 15th IEEE International Conference on Advanced Video and Signal Based

- Surveillance (AVSS), November 27-30, 2018, Auckland, New Zealand. New York: IEEE, 2018: 1-6.
- [16] Chang H T, Gou J N, Li X M. Application of Faster R-CNN in image defect detection of industrial CT[J]. Journal of Image and Graphics, 2018, 23(7): 1061-1071.
常海涛, 苟军年, 李晓梅. Faster R-CNN 在工业 CT 图像缺陷检测中的应用[J]. 中国图象图形学报, 2018, 23(7): 1061-1071.
- [17] Redmon J, Farhadi A. YOLOv3: an incremental improvement [EB/OL]. [2019-07-24]. <https://arxiv.org/abs/1804.02767>.
- [18] 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. New York: IEEE, 2017: 6517-6525.
- [19] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition [C] // 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 770-778.
- [20] Huang G, Liu Z, Laurens V D M. Densely Connected Convolutional Networks [C] // IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 2261-2269.