

基于深度学习的车载导航导光板表面缺陷检测研究

王 昊, 李俊峰

(浙江理工大学机械与自动控制学院, 浙江 杭州 310018)

✉2426028009@qq.com; ljf2003@zstu.edu.cn



摘 要: 针对车载导航导光板表面缺陷像素值分布不均且普遍较小、背景复杂多变等特点, 提出了基于改进掩膜区域卷积神经网络(Mask Region-based Convolutional Neural Network, Mask R-CNN)模型检测车载导航导光板表面缺陷的检测方法。首先, 引入PinFPN模块改进原有Mask R-CNN的特征融合网络, 充分利用高低语义信息构成各级语义、位置信息兼备的共享特征层, 提升整体网络的检测精度; 其次, 通过引入跳层连接结构和SE(Sequence and Excitation)模块对网络的分割分支进行改进, 改善了传统Mask R-CNN网络语义信息获取不充分的问题; 最后, 通过在自建的车载导航导光板数据集上的一系列实验对比, 证明了本方法在检测精度和分割上的优势, 在自建数据集上的检测准确率达到95.3%, 满足工业检测的要求。

关键词: 缺陷检测; 深度学习; Mask R-CNN; 多尺度融合; SE模块

中图分类号: TP273 **文献标识码:** A

Research on Surface Defect Detection of Vehicle Navigation Light Guide Plate based on Deep Learning

WANG Hao, LI Junfeng

(School of Mechanical and Automatic Control, Zhejiang Sci-Tech University, Hangzhou 310018, China)

✉2426028009@qq.com; ljf2003@zstu.edu.cn

Abstract: Aiming at the uneven and generally small pixel value distribution and changeable background of surface defects of vehicle navigation light guide plate, this paper proposes a detection method for surface defects of vehicle navigation light guide plate based on improved Mask Region-based Convolutional Neural Network (Mask R-CNN) model. Firstly, PinFPN module is introduced to improve the feature fusion network of the original Mask R-CNN, and high and low semantic information is fully used to form a shared feature layer with both semantic and location information at all levels, so to improve the detection accuracy of the overall network. Secondly, the introduction of skip connection structure and SE (Sequence and Excitation) module improves segmentation branches of the network and insufficient acquisition of semantic information in traditional Mask R-CNN network. Finally, a series of comparative experiments are performed on the self-built data set of the vehicle navigation light guide plate, which proves that the proposed method has the advantages in detection accuracy and segmentation. The detection accuracy on the self-built data set reaches 95.3%, which meets the requirements of industrial detection.

Keywords: defect detection; deep learning; Mask R-CNN; multi-scale fusion; SE module

1 引言(Introduction)

导光板(LGP)是显示器背光模组的重要组成部分, 被广泛应用于车载和手机液晶屏中, 其质量好坏直接决定了液晶屏

成像的品质效果。然而, 在导光板生产加工过程中由于原材料、生产工艺及人为等因素, 不可避免会出现点压伤、线划伤、脏污等缺陷, 严重影响了用户使用。因此, 各大面板

厂商都必须在出厂前完成缺陷检测工作。但由于车载导航导光板背景纹理复杂、缺陷细微等特点，传统的人工提取特征后分类的方法^[1-3]，需要针对不同种类缺陷设计特定的检测方法且受光照等因素影响较大，导致算法稳定性较差，因此并不可行。

近年来，深度学习算法广泛应用于图像处理领域，如目标检测与实例分割，可以获取图像中的感兴趣目标的位置、类别信息甚至具体轮廓，非常适合工业场景中的缺陷检测任务。一些典型的目标检测算法如Faster R-CNN^[4]、SPP-Net^[5]、TDDNet^[6]、YOLO系列网络^[7-10]等被陆续应用于各种工业场景，包括磁瓦^[11]、铝型材^[12]、表面钢材^[13-14]等领域。李维刚等^[15]提出的基于改进的YOLOv3实现带钢表面缺陷检测，其主要改进在于提出了加权K-means聚类算法，优化先验框参数，融合浅层和深层的特征，生成更大尺度检测图层，在NEU-DET数据集上平均准确率达到80%。张晓等^[16]通过改进Faster R-CNN实现设备的定位检测工作，其主要改进在于引入了注意力机制，调整了初始锚框并改进了对锚框的融合方式，检测精度与收敛速度都有较大提升。李海培^[17]设计了一种改进的Fast R-CNN的钢轨表面缺陷检测方法，提出一种基于灰度垂直投影法完成了轨面区域提取，并改进网络训练策略，检测精度与实时性都有较大提升。CHEN等^[18]利用SSD、YOLO等网络来定位紧固件，并对目标检测网络获取的紧固件图片进行裁剪，最后利用分类网络确定紧固件是否缺失。但以上缺陷检测方法仅仅完成了对缺陷位置和类别信息的提取，忽略了缺陷形状及语义信息的重要性，仅仅能够做到缺陷检出，但后续为了将缺陷信息反馈到生产线以改进生产工艺，单纯的目标检测网络并不能满足要求。

针对这一情况，本文提出一种基于改进Mask R-CNN的车载导航导光板缺陷检测方法。首先，引入PinFPN模块改进原有Mask R-CNN的特征融合网络，充分利用高低语义信息构成各级语义、位置信息兼备的共享特征层，提升整体网络的检测精度；其次，通过引入跳层连接结构和SE(Sequence and Excitation)模块对网络的分割分支进行改进，改善了传统Mask R-CNN网络语义信息获取不充分的问题，实现对车载导航导光板缺陷的高效检测与分割，并通过一系列对比实验进行验证，证明了改进的Mask R-CNN网络应用于车载导航导光板缺陷分割和检测任务的有效性。

2 改进Mask R-CNN检测算法(Improved Mask R-CNN detection algorithm)

2.1 改进Mask R-CNN网络

Mask R-CNN是由HE等人^[19]在2017年提出的一种目标检测与实例分割框架，因其既能识别出目标的位置与类别信息，又可以生成目标的一个高质量mask而备受欢迎。Mask R-CNN实际上相当于是对Faster R-CNN的进一步拓展，相对于Faster R-CNN，其具体改进如下：

(1)使用RoIAlign代替了RoIPooling，解决了因RoIPooling导致特征图与原始图像无法对准的问题，使检测精度大幅提升，RoIAlign可以基本保留目标空间位置信息，实现像素级别的对准；

(2)Mask R-CNN在Faster R-CNN的基础上增加了一个分割分支，使其能够在原有目标检测的基础上进一步实现对目标的实例分割任务。

Mask R-CNN模型主要包括以下四个模块。模块1：主干特征提取网络(ResNet101)和特征融合网络即特征金字塔(Feature Pyramid Networks, FPN)生成公共特征层；模块2：RPN(Region Proposal Network)网络提取感兴趣区域并生成候选框；模块3：将由RPN生成的候选区域送入RoIAlign，通过双线性插值来进行对齐操作，具体来说就是将RoI映射成固定维数的特征向量；模块4：将上一模块的特征分别送入分类、边框回归、Mask预测三个分支，完成目标检测与分割任务。但是，传统Mask R-CNN的特征融合网络只对高层语义信息进行一次上采样，经过通道调整后与同维度的浅层语义信息进行粗略的特征融合，不能充分利用特征信息，难以准确识别缺陷类型和位置信息。为了实现车载导航导光板缺陷高精度检出，本文借鉴了PANet的优点，引入PinFPN^[20]改进了Mask R-CNN特征融合网络，充分融合高低语义信息以提升目标检测任务的精度，同时通过引入跳层连接结构和SE模块^[21]改进分割分支，使分割分支语义信息提取能力大大提升。改进后的Mask R-CNN结构如图1所示。

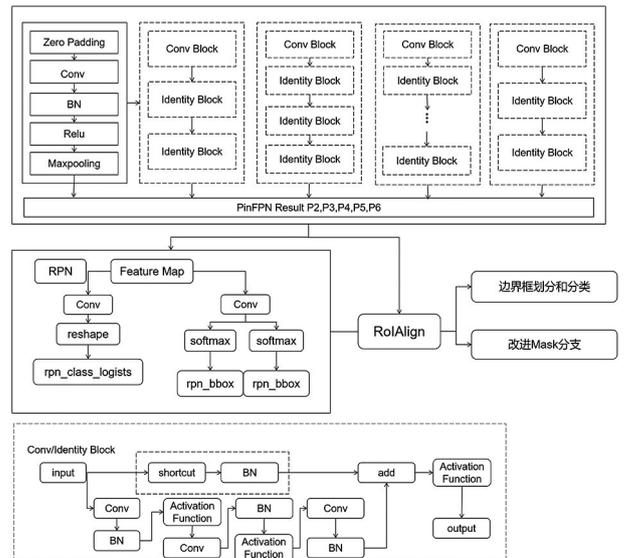


图1 改进后的Mask R-CNN结构图
Fig.1 Structure diagram of improved Mask R-CNN

2.2 改进的特征融合网络

本文在传统特征金字塔FPN的基础上引入了PinFPN模块，结构如图2所示，其中C1—C5分别代表从主干特征提取网络ResNet101各个卷积阶段获得的有着丰富信息资源的特征

层，并且随着主干卷积网络运算的不断进行，感受野不断加深，图像细粒度逐渐呈现下降趋势。M2、M3卷积运算较少，对小缺陷更加敏感是提升小物体目标检测精度的关键，相应的较大物体在特征金字塔中分辨率较低、感受野较大的M4和M5中加以获取。传统特征金字塔的缺点在于M4与M5没有充分融合浅层语义信息，导致位置信息十分不清晰，很难充分获取底层的细节信息，因此又添加了一路N2—N5自底向上的下采样路径，将底层位置信息向高层语义中的缺陷种类信息做了一定补充。但结合车载导航导光板的具体特点，点缺陷与线缺陷占比较大，也是影响检测精度整体水平的关键，大多数目标依赖N2与N3，但两者中的特征信息与M2、M3所含信息差别并不是很大，很难获得更充分的语义信息，对提高相应缺陷的检测精度没有很大帮助。因此，又继续增加了第三组特征融合，并添加一步跳层连接将M—列特征层与同维度P—列对应特征层进行堆叠，实现了对特征融合网络最终输出特征层的一次有效补充。P2、P3、P4、P5是多次融合的结果，有着十分丰富的特征信息，对提升检测精度有很大帮助。P6为分割分支的共享特征层，因其充分融合了高低语义信息，对提升分割效果也有着极大的促进作用。

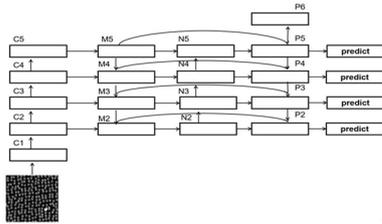


图2 PinFPN结构图

Fig.2 Structure diagram of PinFPN

特征金字塔利用式(1)实现不同尺度特征层上检测多尺度目标。

$$k = k_0 + \log_2 \left(\frac{\sqrt{wh}}{640} \right) \quad (1)$$

其中， k 指代宽高分别为 w 和 h 的目标区域； k_0 取4，表示特征金字塔层数；640表示用于模型训练的图像尺寸。

2.3 改进的Mask分支

Mask分支得到来自RoIAlign双线性插值后得到的RoI，对它们各自使用四次卷积和一个反卷积操作。针对每一个RoI，Mask输出 $N \times n \times n$ 维向量，其中 N 是类别数目， n 为特征图尺寸，对每一个大小为 n 的Mask进行编码，得到每一个RoI的 N 个类别的概率值，以实现实例分割任务。Mask使用全卷积操作来提取语义信息，对局部语义信息敏感度较好，但对各个像素进行分类，没有充分考虑像素与像素之间的关系，忽略了在通常基于像素分类的分割方法中使用的空间规整(Spatial Regularization)步骤，缺乏空间一致性即忽略

了上下文信息。传统的Mask分支如图3所示。

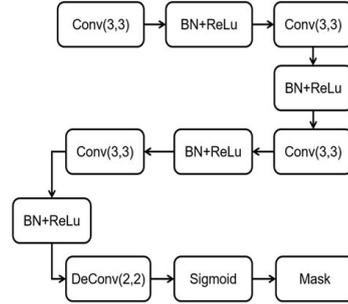


图3 Mask分支

Fig.3 Mask branch

由于车载导航导光板数据集中拥有数目可观的面缺陷，因此全局信息与局部信息同样重要，对车载图像各个尺度特征信息进行融合，可以一定程度上增大信息量，因此引入跳层连接结构，并引入多尺度注意力机制，以补全原有Mask分支上所缺乏的高层语义信息。Mask分支中卷积层和与之相堆成的反卷积层之间通过跳层连接将卷积层每一层的输出结果通过SE模块之后与上采样的结果实现堆叠，最终使用激活层构造分割Mask。通过改进的Mask分支，实现了不同尺度的特征层依照权重完成有效连接，使得分割分支获取了十分丰富的特征信息，大大提升了对目标区域的分割效果。改进后的Mask分支如图4所示。

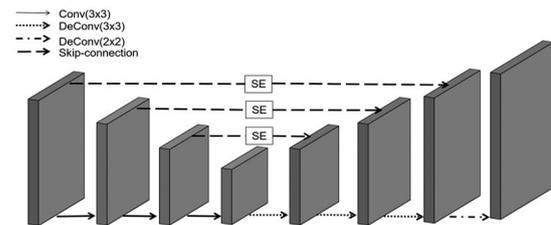


图4 改进后的Mask分支

Fig.4 Improved Mask branch

跳层连接结构能够有效融合高低语义信息，对车载导航导光板细节信息提取更加充分，但因为导光板缺陷本身亮度不均匀，特别是较大的面缺陷，其内部缺陷之间像素值差别较大，单纯进行一步多尺度融合并不能十分有效地分割出缺陷信息。因此，引入SE模块^[22]，该模块的优越性在于其关注通道(Channel)之间的关系，通过设置权值，使模型可以自动学习到不同的通道特征，使模型提升关键特征的权重，从而有效提升分割效果。SE模块结构如图5所示。

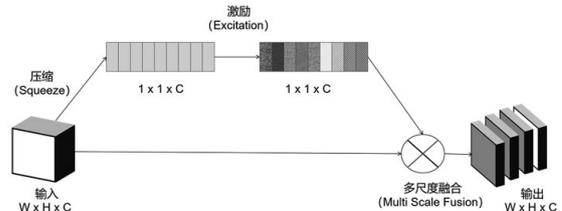


图5 SE模块

Fig.5 Squeeze and Excitation module

SE模块主要由Squeeze和Excitation两部分组成。 W 和 H

分别表示特征层宽与高， C 为通道数。第一步是压缩操作，具体操作是一步全局平均池化，压缩后得到 $1 \times 1 \times C$ 维特征向量；第二步是激励操作，由两个全连接组成，作用是降低通道个数从而降低计算量；最后一步是多尺度融合操作，是将SE模块计算所得通道权重分别与特征图各自对应通道相乘，得到具有注意力信息的有效特征层。

2.4 损失函数

Mask R-CNN是多任务模型，对于车载导航导光板缺陷而言，分类、回归、分割三个任务是同等重要的，但多任务学习网络对于不同任务的学习难易程度往往有较大差异，不同任务有着不同的收敛速度，分割任务的难度要远远大于目标检测，若简单地将损失函数权重进行相同设置，往往难以收敛，因此本实验采用加权损失函数，具体损失函数公式如下：

$$Loss = \delta L_{cls} + \varepsilon L_{bbox} + \theta L_{mask} \quad (2)$$

其中， L_{cls} 是分类误差、 L_{bbox} 是坐标误差、 L_{mask} 是分割误差； δ 、 ε 、 θ 分别为加权系数。

3 实验内容(Experimental content)

3.1 实验数据

车载导航导光板质量检测精度较高，工业面阵摄像头难以满足要求，本文检测系统采用16 k线扫相机呈现清晰的车载导航导光板图像。车载导航导光板缺陷分为以下几类：点缺陷、线缺陷和面缺陷。本实验采用亮度变化、背景替换等方法对数据进行补充。所有样本按上述方法处理后，考虑到车载导航导光板数据集中各种缺陷分布不均，数据集中各样本的比例大致设定为1:1:1，训练集和测试集的比例设定为7:3。具体数据集设置如表1所示。

表1 数据集设置

Tab.1 Dataset settings

类别	总数	训练集	测试集
点缺陷	2,897	2,028	869
线缺陷	2,936	2,055	881
面缺陷	2,856	1,999	857

3.2 实验参数设置

实验参数设置：初始学习率和最终学习率分别为0.01和0.001，批量处理大小为16，软硬件设置如表2所示。

表2 实验环境配置

Tab.2 Experimental environment configuration

软件/硬件	配置参数
CPU	Intel Core i7-8700k
GPU	RTX 2080Ti
操作系统	Windows 10 64 位
GPU加速器	CUDA 10.0 CUDNN 7.4.1.5
深度学习框架	Pytorch 1.2.0
Python	3.6

3.3 算法评价指标

本文采用目标检测网络常用的评价指标，如精度(AP)和平均精度(mAP)。AP能反映网络对各种缺陷的检测能力；mAP作为模型的综合性能评价指标，能够反映模型中各个类别的平均检测精度。

AP是准确率和召回率之间的定积分，具体表达式如下：

$$AP(x) = \int_0^1 precision_rate(x) drecall_rate(x) \quad (3)$$

其中， x 表示某种缺陷， $precision_rate$ 表示准确率， $recall_rate$ 表示召回率。

mAP是所有类型的平均AP，表达式如下：

$$mAP = \frac{1}{x} \sum_{x \in X} AP(x) \quad (4)$$

其中， x 是所有类型缺陷的集合。

准确率和召回率的具体表达式如下：

$$precision_rate = \frac{TP}{(TP + FP)} \quad (5)$$

$$recall_rate = \frac{TP}{(TP + FN)} \quad (6)$$

其中， TP 、 FP 和 FN 分别表示真正例、假正例和假负例的数目。此外， TN 代表真负例的个数，在本检测任务中没有实际意义。

3.4 结果与分析

为了方便与其他方法进行对比，文中将目标检测结果与分割结果分开进行对比讨论。

3.4.1 目标检测结果

为说明改进的Mask R-CNN在目标检测任务上的精度优势，使用相同的车载导航导光板缺陷数据分别与YOLOv3、Faster R-CNN、Mask R-CNN和本文改进的Mask R-CNN进行对比实验，表3显示了这些方法的缺陷检测结果。由表3可知，改进的Mask R-CNN具有更优秀的目标检测能力，目标检测精度较传统的Mask R-CNN有了明显提升，也验证了PinFPN模块能够有效地融合高低语义信息，提升检测精度，对点缺陷等像素值较小的目标提升效果尤为明显，较传统的Mask R-CNN网络点缺陷提升了近1.5%。而YOLOv3仅仅使用最简单的特征融合，仅完成一次高低语义信息融合，并且YOLOv3使用固定的感受野，很难对小物体目标做出有效检出，像素值较小的点缺陷与较浅的线缺陷在车载导航导光板复杂多变的背景干扰下更增大了识别难度。Faster R-CNN未进行额外的特征融合，仅将主干特征提取网络最后一层输出层作为公共特征层，生成建议框时使用RoIPooling造成特征信息的进一步缺失。加上二者都未对特征层进行充分的多尺度融合，因此检测精度都较低，达不到工业检测的要求。本文所引入的PinFPN模块，能够充分利用浅层语义信息，提升了小物体目标的检测精度，对大物体目标也保持较高的检测精度。

表3 不同检测方法的检测性能比较

Tab.3 Detection performance comparison of different detection methods

算法	mAP/%	AP/%		
		点缺陷	线缺陷	面缺陷
YOLOv3	87.1	86.5	81.1	93.7
Faster R-CNN	90.6	88.9	87.3	95.7
Mask R-CNN	94.2	94.2	90.6	97.7
本文方法	95.3	95.7	91.8	98.3

缺陷可视化结果如图6所示。很明显，改进的Mask R-CNN能够以更高的检测精度检出大部分缺陷，其他几种模型检测精度偏低，定位也不够准确，同一张测试图片下YOLOv3甚至出现了误检问题。

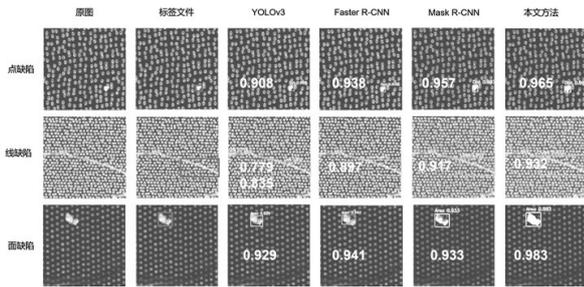


图6 不同模型的检测结果可视化

Fig.6 Test results visualization of different models

3.4.2 分割结果

传统的Mask R-CNN与改进的Mask R-CNN实例分割结果如图7所示。可以看出，改进的Mask分支较原始Mask R-CNN有着明显优势，分割效果更加明显，生成的掩膜覆盖得更加完整，与原始图像轮廓相差很小，体现了引入的注意力机制和多尺度机制的有效性。跳层连接结构可以通过复用浅层特征来提高多尺度目标(尤其是小目标)的分割能力；引入SE多通道注意力机制让网络自行判断各个通道特征信息的比重，可以更加充分地地完成有意义特征的筛选和定位。多尺度的特征融合机制配合注意力机制可以使得网络具备适应多尺度缺陷的检测能力。综上所述，本文改进的网络较传统的Mask R-CNN在准确性和通用性上有明显提升，均具有优势。

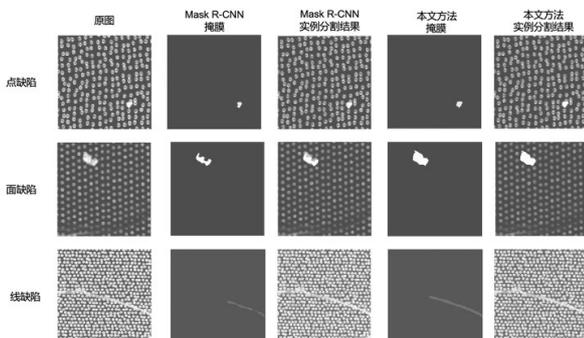


图7 实例分割结果

Fig.7 Instance segmentation results

4 结论(Conclusion)

针对车载导航导光板背景复杂、亮度多变等特点，提出了一种基于改进Mask R-CNN的车载导航导光板缺陷检测方法。为了高效利用主干特征提取网络中的高低语义信息，本文提出引入PinFPN模块，有效地提升了网络对小物体目标的检测能力。同时，针对传统Mask R-CNN分割效果较差的缺点，引入跳层连接和SE模块，实现了不同尺度的特征层依照权重实现有效连接，很大程度上提升了分割效果。在自建的车载导航导光板数据集上检测准确率达到95.3%，满足工业检测的要求，具有一定的应用前景。

参考文献(References)

- [1] 李俊峰,李明睿.基于机器视觉的导光板缺陷检测方法研究[J].光子·激光,2019,30(3):256-265.
- [2] 汤勃,孔建益,伍世度.机器视觉表面缺陷检测综述[J].中国图象图形学报,2017,22(12):1640-1663.
- [3] 李俊峰,李明睿.基于多方向Gabor滤波的导光板轻微线刮伤检测方法研究[J].光子·激光,2019,30(4):395-401.
- [4] REN S, HE K, 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.
- [5] HE K M, ZHANG X, REN S. Spatial pyramid pooling in deep convolutional networks for visual recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014, 37(9):1904-1916.
- [6] DING R, DAI L, LI G. TDD-net: A tiny defect detection network for printed circuit boards[J]. CAAI Transactions on Intelligence Technology, 2019, 4(2):110-116.
- [7] REDMON J, DIVVALA S, GIRSHICK R, et al. You only look once: Unified, real-time object detection[C]// IEEE. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016: 779-788.
- [8] REDMON J, FARHAD A. YOLO9000: Better, faster, stronger[J]. IEEE Conference on Computer Vision and Pattern Recognition, 2017, 65(30):7263-7271.
- [9] REDMON J, FARHADI A. YOLOv3: An incremental improvement[EB/OL]. (2018-04-08) [2021-12-25]. <https://arxiv.org/abs/1804.02767>.
- [10] BOCHKOVSKI A, WANG C Y, LIAO H M. YOLOv4: Optimal speed and accuracy of object detection[EB/OL]. (2020-04-23) [2021-12-25]. <https://arxiv.org/abs/2004.10934>.
- [11] XIE L, XIANG X, XU H, et al. FFCNN: A deep neural network for surface defect detection of magnetic tile[J]. IEEE Transactions on Industrial Electronics, 2020, 68(4):3506-3516.