

DOI:10.19651/j.cnki.emt.2210853

# 基于 S-ASPP 和双注意力机制的磁瓦 外观缺陷检测算法

宁奉阁<sup>1</sup> 石金进<sup>2</sup>

(1. 三峡大学电气与新能源学院 宜昌 443002; 2. 福建工程学院交通运输学院 福州 350118)

**摘要:** 针对现有视觉缺陷检测算法对磁瓦外观识别精度低的问题,提出一种 S-ASPP 和双注意力机制的磁瓦外观缺陷检测算法。首先,为了解决模型参数量大,网络结构复杂所带来的检测效率低,算法部署硬件成本高等问题,提出采用轻量化 GhostNet 主干网络提取视觉底层特征;其次,设计 S-ASPP 模块引入密集连接和深度可分离卷积,以此在降低模型参数数量的同时,提高模型的预测速度。然后,为了解决特征提取过程中可能出现语义信息丢失的问题,设计双注意力模块,将 GhostNet 的中层特征送入双注意力模块并与 ASPP 处理后的高层特征进行拼接,使网络获取更多语义特征,提高分割精度。最后,为验证本文所提方法的有效性,在磁瓦数据集上与 DeepLabV3+, PSPNet, U-net 等 3 种算法进行对比,实验结果表明,基于可分离 ASPP 和双注意力机制的磁瓦缺陷检测算法的具有较高的识别精度,平均交并比达到 82.43%,类别平均像素准确率达到 93.08%。该算法平衡了识别精度与参数量之间的关系,综合性能最优。

**关键词:** 深度学习;缺陷检测;磁瓦;语义分割

**中图分类号:** TP391 **文献标识码:** A **国家标准学科分类代码:** 520.6040

## S-ASPP and dual attention mechanism based algorithm for tile appearance defect detection

Ning Fengge<sup>1</sup> Shi Jinjin<sup>2</sup>

(1. College of Electrical Engineering and New Energy, China Three Gorges University, Yichang 443002, China;

2. School of Transportation, Fujian University of Technology, Fuzhou 350118, China)

**Abstract:** Aiming at the issue of low accuracy of existing visual defect detection algorithms for magnetic tile appearance recognition, an algorithm based on S-ASPP and dual attention mechanism is proposed. Firstly, in order to solve the problems of low detection efficiency and high hardware cost for algorithm deployment caused by large amount of model parameters and complex network structure, a lightweight GhostNet backbone network is proposed to extract low-level visual features. Secondly, the S-ASPP module is designed to introduce dense connections and depthwise separable convolutions, so as to reduce the amount of model parameters and improve the prediction speed of the model. Next, in order to solve the problem that semantic information may be lost during feature extraction, a dual-attention module is designed. The middle-level features of GhostNet are feed into the dual-attention module and concatenate with the high-level features after ASPP processing, so that the network can acquire more semantic features and improve the segmentation accuracy. Finally, in order to verify the effectiveness of the proposed method, the magnetic tile defect detection algorithm is compared with DeepLabV3+, PSPNet and U-NET algorithms on the magnetic tile data set. Experimental results show that the magnetic tile defect detection algorithm based on separable ASPP and dual attention mechanism has high recognition accuracy, and the average crossover ratio reaches 82.43%. The average pixel accuracy of the category reached 93.08%. The algorithm balances the relationship between the recognition accuracy and the number of parameters and has the best comprehensive performance.

**Keywords:** deep learning; defect detection; magnetic tile; semantic segmentation

## 0 引言

磁瓦作为永磁电机的关键部件,用于产生稳定的磁场,

在实际生产过程中受加工设备、加工工艺、环境温度和人工操作失误等因素的影响难免会出现气孔、破损、裂纹、磨损和起层等多种缺陷。这些缺陷对电机的性能产生重要影

收稿日期:2022-07-27

响,可能会损坏电机,导致严重事故的发生。为此,需要在出厂前对磁瓦的质量进行检验,防止问题产品流入市场,引发重大事故。针对磁瓦外观缺陷的问题,已有相关人员开展深入研究。朱志勋等<sup>[1]</sup>提出一种同态滤波和Canny算法相结合的视觉缺陷检测方法,该方法利用同态滤波改善磁瓦图像表面光照不均的问题,提高图像对比度,然后使用改进Canny算子分割缺陷,进行形态学处理输出分割图片,该方法较好的分割多种磁瓦缺陷,但不能检测边缘信息较弱的磁瓦缺陷,算法效率较低;杨成立等<sup>[2]</sup>提出一种基于非下采样的Shearlet变换的缺陷检测方法,有效抑制磁瓦图像中的强干扰,准确识别缺陷,但是该方法无法区分垂直方向的磨削纹理和缺陷,容易导致漏检;Li等<sup>[3]</sup>提出了一种将非下采样Contourlet变换与Laws纹理特征测量相结合磁瓦缺陷检测方法,有效地消除磨削纹理,可以准确提取缺陷,但无法满足实时性要求;Yang等<sup>[4]</sup>提出一种基于NSST和EGLG的算法用于消除图片中噪声和纹理的干扰,提取磁瓦裂纹缺陷,但对于长度小于0.9 mm的缺陷检测效果不好;张露滨等<sup>[5]</sup>提出一种微型磁瓦表面在线缺陷检测方法,结合背景纹理抑制和掩膜技术有效抑制非缺陷区域纹理干扰,识别磁瓦划痕缺陷,但是该方法无法完全排除噪点干扰,对沿磁瓦纹理方向的缺陷检测效果不好。以上方法由于算法复杂,鲁棒性差,通常难以满足工业需求。

近年来卷积神经网络因其优异的特征学习能力,成为研究热点,被广泛应用在视觉外观缺陷检测领域。刘培勇等<sup>[6]</sup>提出一种基于多支路卷积神经网络缺陷检测模型,满足磁瓦表面缺陷的实时检测;祝礼佳等<sup>[7]</sup>提出一种基于NASNet网络的缺陷检测模型,该模型改进NASNet结构,结合双线性模型,有效提高磁瓦缺陷的识别精度,实现磁瓦缺陷自动分类;马燕婷等<sup>[8]</sup>提出一种基于YOLOv5网络的缺陷检测模型,将Transformer自注意力机制与BiFPN网络相结合,提高带钢表面缺陷的检测精度;Hu等<sup>[9]</sup>本文提出了一种UPM-DenseNet两阶段检测模型,通过定位网络生成原始图像数据,生成可能存在缺陷的区域,将其特征图作为分类网络的输入,通过分类网络完成磁瓦缺陷的在线识别;Lu等<sup>[10]</sup>提出一种多模态融合卷积神经网络,设计交叉注意力机制,实现特征信息交互,提高磁瓦缺陷检测效果;殷鹰等<sup>[11]</sup>开发智能化磁瓦内部缺陷检测设备,利用深度一维卷积网络提取磁瓦内部缺陷,有效地从声音信号中识别磁瓦缺陷。DeeplabV3+<sup>[12]</sup>作为经典语义分割网络,在VOC通用数据集上有着优异的分割效果,该网络采用编码器-解码器结构和空洞空间金字塔池化(atrous spatial pyramid pooling, ASPP)模块,通过对不同尺度的特征进行融合,实现物体的分割识别。例如,白欢欢<sup>[13]</sup>基于DeeplabV3+网络融入特征金字塔注意力和通道注意力机制对多种服饰进行分割。任鸿杰等<sup>[14]</sup>以DeepLabV3+网络为基础对玉米、薏米两种农作物进行分割;Zhao等<sup>[15]</sup>使

用电流互感器红外图像训练DeeplabV3+网络,识别效果明显优于其它网络。

以上研究表明,DeeplabV3+被广泛应用在各领域中,但由于自身结构复杂,低层特征信息在卷积过程中易出现像素点丢失导致识别精度低,对磁瓦缺陷的识别效果差。针对现有缺陷检测算法对磁瓦外观识别精度低的问题,本文提出一种基于可分离空洞空间金字塔池化模块(separable atrous spatial pyramid pooling, S-ASPP)和双注意力机制的磁瓦外观缺陷检测算法。首先,采用轻量级的GhostNet作为主干特征提取网络,减少模型参数量,提高网络训练速度;然后设计S-ASPP模块,引入密集连接和深度可分离卷积,增大感受野;最后,从GhostNet主干网络中引入中层特征信息送入双注意力机制模块并与可分离ASPP模块处理后的高层特征进行拼接,使网络获得更多的特征信息,进一步提高模型的分割精度。实验表明,该方法较DeeplabV3+参数量减少,检测精度有所提升,更好的权衡参数量与识别精度的关系。

## 1 网络模型

基于S-ASPP和双注意力机制的磁瓦表面缺陷检测模型主要通过两个部分相结合的方式实现:S-ASPP模块和双注意力机制模块,图1为所提网络模型结构。

该模型受DeeplabV3+网络的启发,采用编码器-解码器结构。选择轻量级的GhostNet网络模型作为主干网络。在编码器阶段中,输入图像经过GhostNet模型中的Ghost bottleneck进行特征提取,获得输出通道数分别为160,114和40的高层特征,中层特征和低层特征。将GhostNet输出的高层特征和中层特征分别送入可分离ASPP模块和双注意力机制模块,得到处理后的高层特征和中层特征,并将通道数调整为256。在解码器阶段,提取图像经过GhostNet模型输出的低层特征,通过卷积进行通道调整后,与编码器输出的高层特征和中层特征进行拼接操作进行特征融合。经过 $3 \times 3$ 卷积和上采样后,恢复空间信息,输出分割结果。

### 1.1 GhostNet网络

传统DeeplabV3+模型主干网络采用Xception利用3次深度可分离卷积提取特征,但是该方法使得通道数增多、网络复杂、参数量变大、对硬件要求高、部署设备成本高。采用轻量化的GhostNet网络<sup>[16]</sup>,该网络参数量少、运行速度快。GhostNet网络提出Ghost模块如图2所示,该模块将卷积运算与线性运算相结合,对于输入特征,先利用 $1 \times 1$ 的卷积获取必要的特征浓缩,然后通过深度可分离卷积获得特征浓缩的相似特征图,最后将经 $1 \times 1$ 的卷积处理的特征浓缩和经深度可分离卷积的特征浓缩的相似特征图堆叠在一起,获得完整的特征。在不改变输出特征大小的情况下,Ghost模块所需的参数量和计算复杂度远低于普通卷积,在维持相似性能的同时降低计算成本。

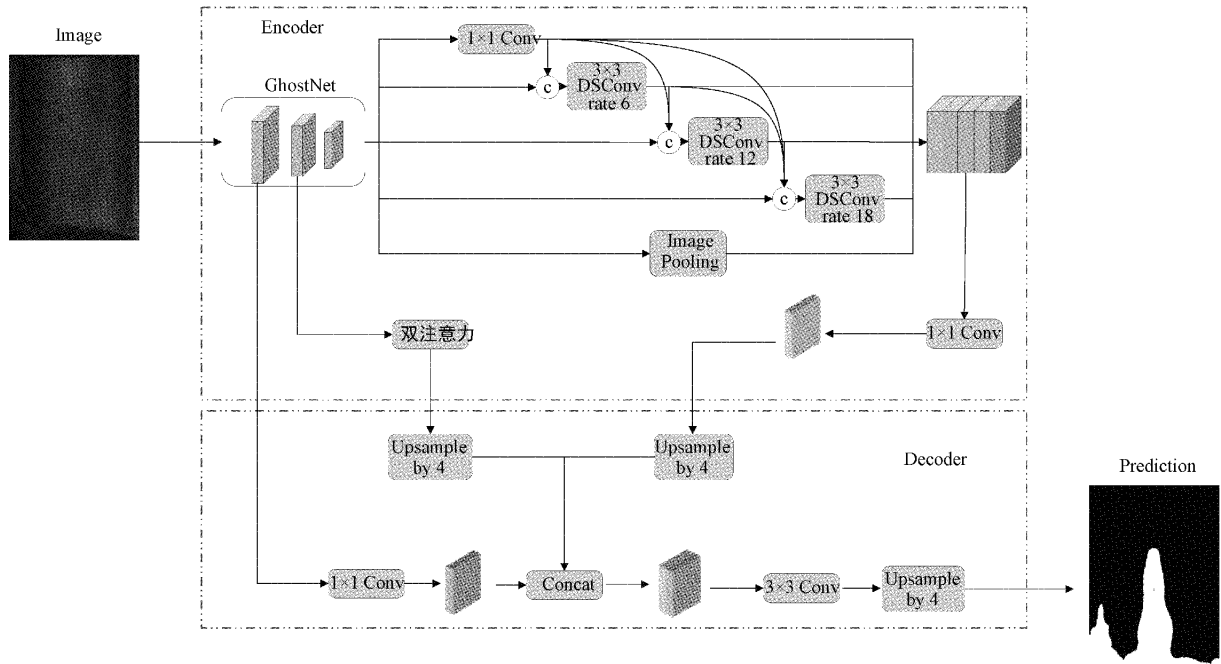


图 1 基于 S-ASPP 和双注意力机制的缺陷检测模型

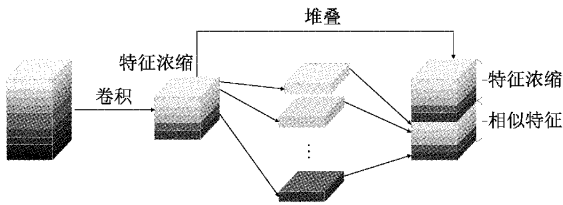


图 2 Ghost 模块结构

GhostNet 由一系列 Ghost bottleneck 组成,以 Ghost 模块为基础进行构建,具体网络参数如表 1 所示。GhostNet

表 1 GhostNet 网络结构

输入	卷积	升维	通道	SE	步长
$416 \times 416 \times 3$	Conv2d	—	16	—	2
$208 \times 208 \times 16$	G-bneck	16	16	—	1
$208 \times 208 \times 16$	G-bneck	48	24	—	2
$104 \times 104 \times 24$	G-bneck	72	24	—	1
$104 \times 104 \times 24$	G-bneck	72	40	1	2
$52 \times 52 \times 40$	G-bneck	120	40	1	1
$52 \times 52 \times 40$	G-bneck	240	80	—	2
$26 \times 26 \times 80$	G-bneck	200	80	—	1
$26 \times 26 \times 80$	G-bneck	184	80	—	1
$26 \times 26 \times 80$	G-bneck	184	80	—	1
$26 \times 26 \times 80$	G-bneck	480	112	1	1
$26 \times 26 \times 112$	G-bneck	672	112	1	1
$26 \times 26 \times 112$	G-bneck	672	160	1	2
$13 \times 13 \times 160$	G-bneck	960	160	—	1
$13 \times 13 \times 160$	G-bneck	960	160	1	1
$13 \times 13 \times 160$	G-bneck	960	160	—	1
$13 \times 13 \times 160$	G-bneck	960	160	1	1

的第一层为 16 个卷积核的标准卷积层,然后经过一系列 Ghost bottleneck,通道数逐渐增加。根据特征大小的不同,可分为步长为 2 和步长为 1 的两种 Ghost bottleneck。网络中除每个阶段最后一个 Ghost bottleneck 步长为 2,其他所有 Ghost bottleneck 步长为 1。Ghost bottleneck 具体结构如图 3 所示。

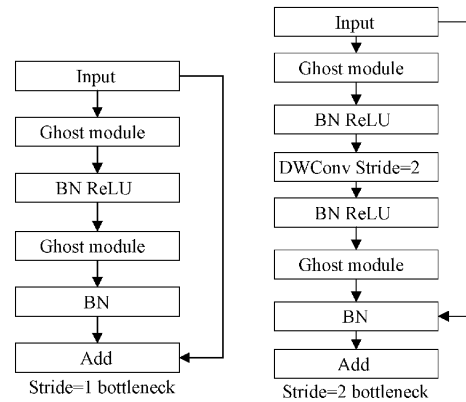


图 3 不同步长的 Ghost Bottlenecks 结构

1.2 S-ASPP 模块

S-ASPP 模块主要由密集连接和深度可分离卷积两部分构成,具体结构如图 4 所示。通过密集连接的结构将每一层的卷积的特征输出作为下层卷积特征输入的一部分,增强小目标信息的特征提取,扩大感受野。同时将空洞卷积替换成深度可分离卷积,以此来降低模型参数量。

1) 密集连接

针对磁瓦缺陷细小在卷积过程中由于像素信息小,易导致信息丢失,出现无法识别或误分类的现象。引入密集

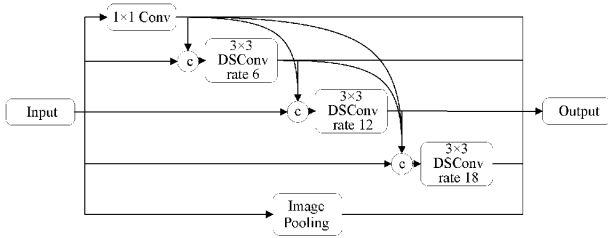


图4 S-ASPP模块结构

连接的思想,将每一层的卷积输出作为下层卷积输入的一部分,增强局部信息流之间的传递,实现小目标信息的加强提取。较普通ASPP模块该结构能获取更大的感受野。对于一个空洞率为 $r$ 卷积核大小为 $k$ 的空洞卷积来说,它能获取的感受野大小为:

$$R = (r - 1) * (k - 1) + k \quad (1)$$

对于采用密集连接的空洞卷积来说,它能获取的感受野大小为:

$$R = R_1 + R_2 - 1 \quad (2)$$

其中, $R_1$ 和 $R_2$ 分别为相邻空洞卷积获取的感受野。那么对于空洞率为6、12和18的ASPP模块来说,它能获取的最大感受野大小为:

$$R_{max} = \max(R_3^6, R_3^{12}, R_3^{18}) = R_3^{18} = 37 \quad (3)$$

对于采用密集连接的ASPP模块来说,它能获取的最大感受野大小为:

$$R_{max} = R_3^6 + R_3^{12} + R_3^{18} - 1 = 74 \quad (4)$$

由此可知,采用密集连接的ASPP模块比普通ASPP模块能获取更大的感受野,实现小目标信息的加强提取。

2) 深度可分离卷积

采用密集连接的ASPP模块虽然能扩大感受野,但是冗长的拼接操作增加模型计算量,为此本文引入深度可分离卷积来替换ASPP模块中的空洞卷积,以此降低模型参数量提高模型预测速度。

深度可分离卷积具体结构如图5所示,该模块将标准卷积分为逐通道卷积和逐点卷积两步,将输入特征的每个通道进行 $3 \times 3$ 的逐通道卷积并进行拼接,得到分离特征属性的特征信息,然后将每个通道进行 $1 \times 1$ 的卷积得到输出特征。与标准卷积相比深度可分离卷积使用更少的参数进行特征学习,性能却优于标准卷积,大大减少了模型参数量。

1.3 双注意力机制模块

尽管可分离ASPP模块扩大感受野面对小目标缺陷可以更有效的提取缺陷信息,但是DeeplabV3+网络高、低层特征信息拼接过程中,由于低层特征可能出现语义信息的丢失,造成复杂背景下缺陷的识别效果差。为了在复杂背景下实现磁瓦缺陷的精准检测,设计双注意力模块,将GhostNet的中层特征送入双注意力模块并与ASPP处理后的高层特征进行拼接,使网络获取更多特征信息,提高分割精度。

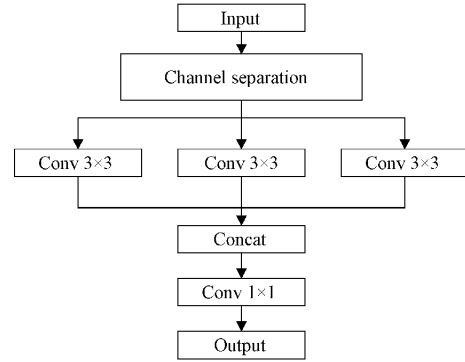


图5 深度可分离卷积结构

如图6所示,双注意力机制将经GhostNet处理的中层特征信息分别送入通道注意力机制(coordinate attention,CA)和空间注意力机制(spatial attention,SA),然后将二者特征信息相加融合,再次送入通道注意力机制中,形成融合位置信息和分类信息的注意力权重,最后将该权重与原始特征信息相乘,进行特征融合,得到最终的输出特征。

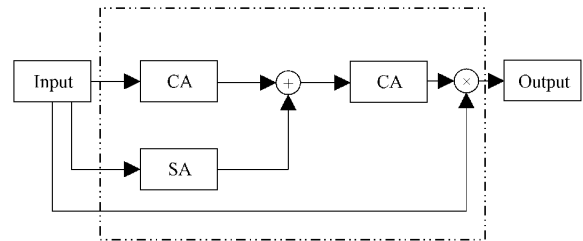


图6 双注意力模块结构

1) CA注意力机制

CA注意力机制是一种轻量化的注意力机制<sup>[17]</sup>,较传统SE注意力机制CA注意力机制将位置信息融入通道注意力中,避免使用2维全局池化造成的位置信息的丢失,同时保证模块的轻量高效。该注意力机制可分为坐标信息嵌入和坐标注意力生成两步,具体如图7所示。

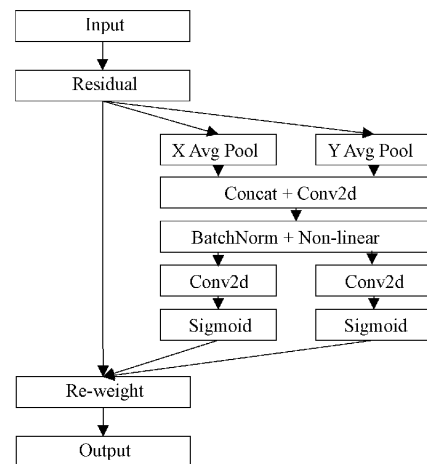


图7 CA注意力模块结构

## (1)坐标信息嵌入

当输入  $\mathbf{X}$  经过全局池化分解成水平方向和垂直方向的 1 维特征, 分别进行 1 维池化, 从而获得水平方向和垂直方向相关的位置信息。水平方向的 1 维特征如式(5)所示。

$$z_c^h(h) = \frac{1}{W} \sum_{0 \leq i \leq W} x_c(h, i) \quad (5)$$

垂直方向的 1 维特征如式(6)所示。

$$z_c^w(h) = \frac{1}{H} \sum_{0 \leq j \leq H} x_c(j, w) \quad (6)$$

通过两种变换获得两个空间方向的特征, 能让注意力模块捕捉到沿着一个空间方向的长期依赖关系, 并保存沿着另一个空间方向的精确位置信息, 使得网络能更准确地定位感兴趣的目标。

## (2)坐标注意力生成

将经过 1 维池化处理后的两个方向特征信息进行拼接, 然后进行归一化和非线性激活函数操作, 如式(7)所示。

$$f = \delta(F_1([z^h, z^w])) \quad (7)$$

其中,  $[z^h, z^w]$  表示沿空间维度的拼接操作,  $\delta$  表示非线性激活函数,  $F_1$  表示将两个特征进行拼接的操作,  $f$  表示空间信息在水平方向和垂直方向的的特征映射, 将  $f$  沿空间维数分解得到  $f^h$  和  $f^w$  两个单独张量, 将这两个张量进行卷积变换和非线性激活, 如式(8)和(9)所示。

$$g^h = \sigma(F_h(f^h)) \quad (8)$$

$$g^w = \sigma(F_w(f^w)) \quad (9)$$

其中,  $F_h$  和  $F_w$  分别表示对  $f^h$  和  $f^w$  进行的卷积变换操作,  $\sigma$  表示非线性激活函数 Sigmoid。将  $g^h$  和  $g^w$  分别作为输入  $\mathbf{X}$  在水平方向和垂直方向的注意力权重。最后, CA 注意力模块输出  $\mathbf{Y}$  可以写成:

$$y_c(i, j) = x_c(i, j) * g_c^h(i) * g_c^w(j) \quad (10)$$

## 2) SA 注意力机制

在空间注意力机制中, 输入特征信息经过  $1 \times 1$  的卷积将通道数降为 1, 将所有通道的特征聚合在一个通道上, 此时对该特征进行非线性激活操作, 并与输入特征信息相乘, 从而获得空间方向上的输出特征信息, 空间注意力结构如图 8 所示。

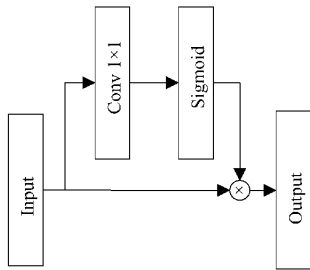


图 8 SA 注意力模块结构

## 2 实验结果及分析

## 2.1 实验环境

本实验基于 Windows10 操作系统, CPU 为 AMD

R7-5800, GPU 为 NVIDIA GeForce RTX 3060, 利用 Pytorch 框架实现所设计的网络。详细参数如表 2 所示。

表 2 实验环境配置

实验环境	配置
操作系统	Windows10
显卡	NVIDIA GeForce RTX 3060(12 G)
处理器	AMD R7-5800
学习框架	Pytorch

本实验加载主干网络的预训练权重加快模型收敛。设置网络初始学习率为  $5 \times 10^{-4}$ , Batch Size 为 8, 采用 Adam 优化器, 将图像裁剪成  $512 \times 512$  的大小, 训练 100 epoch。

## 2.2 实验数据集

本实验采用网络公开数据集 magnetic tile surface defects<sup>[18]</sup>, 该数据集包含气孔、破损、裂纹、磨损、起层和无缺陷共 6 种类别, 如图 9 所示。原始数据集磁瓦表面缺陷较少, 为此, 选取气孔、破损、裂纹、磨损和起层 5 种缺陷进行离线增强, 对每类缺陷采用旋转, 翻转, 裁剪, 尺度缩放等方式进行数据增强, 使得扩增后的数据集达到 1 568 张。实验磁瓦数据集采用 5 种缺陷样本, 共 1 568 张, 随机选取 1 411 张作为训练集, 157 张作为测试集。

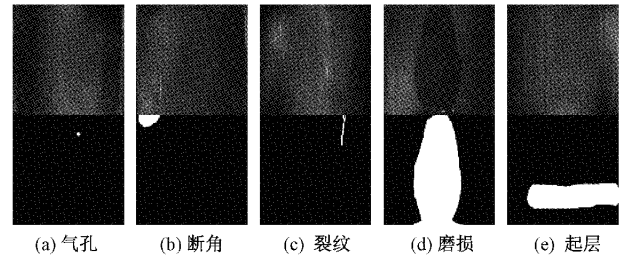


图 9 磁瓦表面缺陷示例

## 2.3 图像分割流程

本文采用的训练流程和分割流程如图 10 所示。训练之前先对数据集进行预处理并将数据集划分为训练集、验证集和测试集; 然后在训练阶段首先, 使用 GhostNet 预训练权重对主干网络进行初始化, 并结合验证集上的损失结果, 采用动态学习率衰减策略进行迭代更新; 最后在测试阶段将测试集输入到训练好的模型上进行测试, 得到磁瓦缺陷分割结果。

## 2.4 评价指标

本文采用类别平均像素准确率 (mean pixel accuracy, Mpa) 和平均交并比 (mean intersection over union, mIoU) 作为模型的评价指标。

$$Mpa = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij}} \quad (11)$$

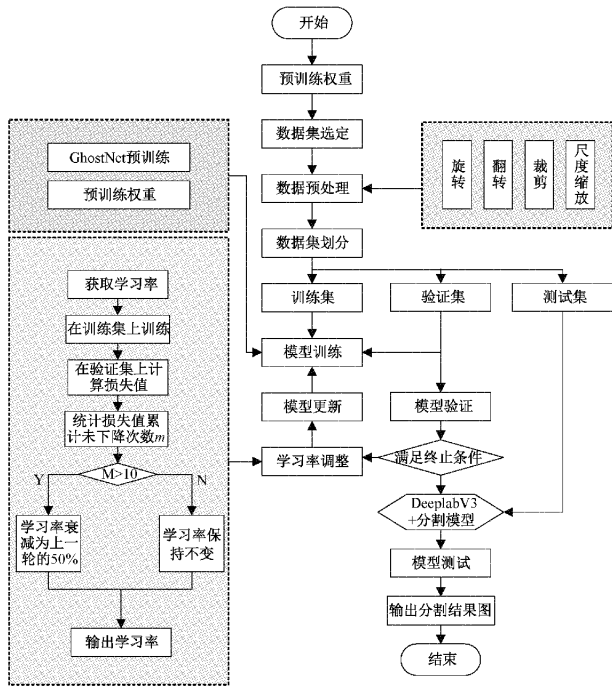


图10 磁瓦缺陷分割流程

$$mIoU = \frac{1}{k+1} \sum_{i=0}^k \frac{p_{ii}}{\sum_{j=0}^k p_{ij} + \sum_{j=0}^k p_{ji} - p_{ii}} \quad (12)$$

式中:  $k$  表示  $k$  个类别;  $p_{ii}$  表示正确预测  $i$  个类别的像素点数目;  $p_{ij}$  表示属于  $i$  类预测成  $j$  类的像素点数目;  $p_{ji}$  表示属于  $j$  类预测成  $i$  类的像素点数目。

### 2.5 实验结果对比

#### 1) 消融实验

为验证本文中各个模块的作用,分别训练5个不同网络进行对比实验。方案1为采用 MobilenetV2 为主干网络的 DeeplabV3+模型;方案2为采用 GhostNet 为主干网络的 DeeplabV3+模型;方案3为在方案2的基础上引入密集连接的 DeeplabV3+模型;方案4为在方案3的基础上引入深度可分离卷积的 DeeplabV3+模型;方案5为在方案4的基础上引入双注意力机制的 DeeplabV3+模型,也本文采用的最终模型。以上实验均采用同一数据集,并保证相同的训练策略,实验结果如表3所示。

表3 消融实验结果

方案	Parameter	训练时间	mIoU/%	Mpa/%
1	5 814 294	2 h 55 min	78.54	89.73
2	5 331 418	2 h 21 min	81.09	90.94
3	5 655 418	2 h 29 min	81.32	92.04
4	<b>4 281 178</b>	2 h 24 min	82.06	92.42
5	4 545 196	2 h 28 min	<b>82.43</b>	<b>93.08</b>

从表3可以看出,当逐步引入所改进的模块时,模型的各个指标均有所提升。对比方案1和方案2可以看出,采用 GhostNet 主干网络的 DeeplabV3+模型较采用 MobilenetV2 为主干网络的 DeeplabV3+模型 mIoU 值提升 2.55%, Mpa 值提升 1.21%, 同时参数量由 5 814 294 降为 5 331 418 实现网络轻量化;对比方案2和方案3可以看出,在原模型的基础上引入密集连接, mIoU 值和 Mpa 值均有所提升,表明密集连接加强 ASPP 模块特征提取的能力;对比方案3和方案4可以看出,将 ASPP 中的空洞卷积替换成深度可分离卷积,有效减少网络参数量和训练时间,同时保证了识别精度;对比方案4和方案5可以看出,通过将 GhostNet 的中层特征送入双注意力机制丰富特征信息, mIoU 值和 Mpa 值均有提升。

#### 2) 注意力机制对比实验

如表4所示,为验证本文所提双注意力机制的有效性,在采用 GhostNet 主干网络的 DeeplabV3+模型的基础上分别引入 SE 注意力机制<sup>[19]</sup>、CBAM 注意力机制<sup>[20]</sup>、CA 注意力机制和本文所提注意力机制对比。从表4可以看出,本文所提双注意力机制在 mIoU 值和 Mpa 值均高于其他注意力机制,结果表明较其他注意力机制所提双注意力机制在磁瓦数据集上的检测效果更优。

表4 消融实验结果

方案	注意力机制	Parameter	mIoU/%	Mpa/%
1	—	<b>5 331 418</b>	81.09	90.84
2	SE	5 591 034	81.48	92.18
3	CBAM	5 595 756	81.21	91.35
4	CA	5 595 402	81.66	91.53
5	本文方法	5 595 436	<b>82.11</b>	<b>92.42</b>

#### 3) 对比实验

表5给出了本文所提算法与其他算法在磁瓦数据集上的对比结果。由表5可以看出,本文方法在参数量、训练时间、mIoU 值和 Mpa 值的性能明显优于传统 PSPNet 和 U-net 网络模型;对比 DeeplabV3+算法,所提方法在模型参数量、训练时间、mIoU 值和 Mpa 值上有明显提升。总的来说,基于可分离 ASPP 和双注意力机制的缺陷检测算法更好的平衡了识别精度与参数量之间的关系,提高模型训练速度,综合性能更优。

表5 不同网络分割性能对比

算法	Parameter	训练时间	mIoU/%	Mpa/%
PSPNet	46 716 373	3 h58 min	77.47	89.17
U-net	43 932 996	5 h15 min	60.91	74.91
DeeplabV3+	5 814 294	2 h55 min	78.02	89.98
本文方法	<b>4 545 196</b>	<b>2 h28 min</b>	<b>82.43</b>	<b>93.08</b>

图 11 为各算法在磁瓦数据集的分割效果图。观察图 11 可知,可以明显看出本文方法的分割效果优于 PSPNet 算法和 U-net 算法;对比 DeeplabV3+ 算法可以看出本文方法的分割效果更加连续和平滑,在裂纹缺陷的分

割中本文方法呈现更好的边缘轮廓信息,边缘更加平滑,分割效果优于 DeeplabV3+ 算法。总之,对比其他算法本文方法对磁瓦数据集中的各类缺陷能够进行有效分割,分割结果更接近实际情况。

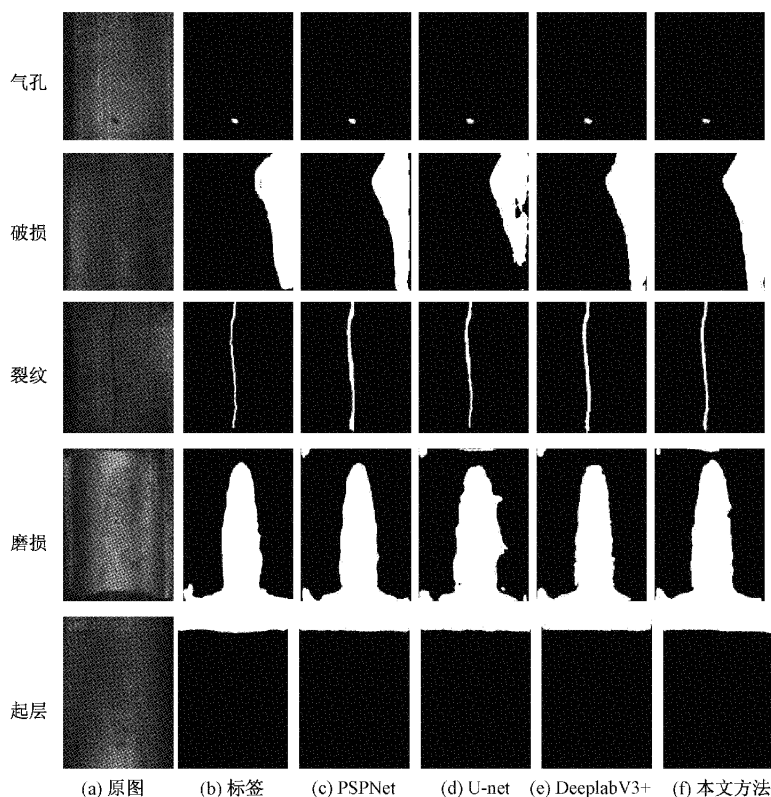


图 11 不同算法在磁瓦数据集上效果对比

### 3 结 论

本文针对现有视觉外观缺陷检测算法对磁瓦外观识别精度低的问题,提出基于 S-ASPP 和双注意力机制的磁瓦外观缺陷检测算法,首先,采用轻量化的 GhostNet 作为主干特征提取网络;其次,设计 S-ASPP 模块,引入密集连接和深度可分离卷积,使其扩大感受野的同时降低模型参数量,加快模型训练速度;最后,设计双注意力机制,将 GhostNet 中层特征信息送入改进双注意力结构并与 S-ASPP 处理后的高层特征进行拼接,使网络获取更多特征信息,提高分割精度。在磁瓦数据集上的实验数据表明,所提算法在降低模型参数的同时,保证识别精度,更好的权衡参数量与识别精度的关系。下一步将研究满足实时性和识别精度的轻量化网络,进一步提高工程实用性。

#### 参考文献

[1] 朱志勋, 赵磊, 李恒, 等. 基于改进的同态滤波与 Canny 算法的磁瓦表面缺陷检测算法研究[J/OL]. 激光与光电子学进展: 1-12 [2022-10-19]. <http://kns.cnki.net/kcms/detail/31.1690.tn.20210804.1202>.

040.html.

- [2] 杨成立, 殷鸣, 向召伟, 等. 基于非下采样 Shearlet 变换的磁瓦表面缺陷检测[J]. 工程科学与技术, 2017, 49(2): 217-224.
- [3] LI X, LIU Z, YIN G, et al. Ferrite magnetic tile defects detection based on nonsubsampling contourlet transform and texture feature measurement [J]. Russian Journal of Nondestructive Testing, 2020, 56(4): 386-395.
- [4] YANG C, LIU P, YIN G, et al. Crack detection in magnetic tile images using nonsubsampling shearlet transform and envelope gray level gradient[J]. Optics & Laser Technology, 2017, 90: 7-17, DOI: 10.1016/j.optlastec.2016.08.016.
- [5] 张露滨, 李俊峰, 沈军民. 微型磁瓦表面线缺陷视觉检测方法研究[J]. 光电子·激光, 2019, 30(9): 951-959.
- [6] 刘培勇, 董洁, 谢罗峰, 等. 基于多支路卷积神经网络的磁瓦表面缺陷检测算法[J/OL]. 吉林大学学报(工学版): 2022, DOI:10.13229/j.cnki.jdxbgxb20210855.

- [7] 祝礼佳, 刘桂华, 林杰. 基于卷积神经网络的磁瓦表面缺陷识别[J]. 制造业自动化, 2022, 44(3): 48-53.
- [8] 马燕婷, 赵红东, 阎超, 等. 改进YOLOv5网络的带钢表面缺陷检测方法[J/OL]. 电子测量与仪器学报: 1-8[2022-10-19]. <http://kns.cnki.net/kcms/detail/11.2488.TN.20220902.0823.002.html>.
- [9] HU C, LIAO H, ZHOU T, et al. Online recognition of magnetic tile defects based on UPM-DenseNet[J]. Materials Today Communications, 2022, 30: 103105, DOI: 10.1016/j.mtcomm.2021.103105.
- [10] LU H, ZHU Y, YIN M, et al. Multimodal fusion convolutional neural network with cross-attention mechanism for internal defect detection of magnetic tile [J]. IEEE Access, 2022, DOI: 10.1109/ACCESS.2022.3180725.
- [11] 殷鹰, 谢罗峰, 黄泰博. 基于深度学习的磁瓦内部缺陷声振检测方法[J]. 中国测试, 2020, 46(3): 32-38.
- [12] CHEN L C, ZHU Y, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]. Proceedings of the European conference on computer vision (ECCV), 2018, DOI: 10.48550/arXiv.1802.02611.
- [13] 白欢欢. 一种融合多注意力机制的服饰分割方法研究[J]. 国外电子测量技术, 2021, 40(7): 52-58.
- [14] 任鸿杰, 刘萍, 岱超, 等. 改进DeepLabV3+网络的遥感影像农作物分割方法[J]. 计算机工程与应用, 2022, 58(11): 215-223.
- [15] ZHAO T, YANG Y, SANG J, et al. Infrared image segmentation method of current transformer based on DeepLabv3 + network [C]. 2022 IEEE 5th International Electrical and Energy Conference (CIEEC), IEEE, 2022: 3111-3115, DOI: 10.1109/CIEEC54735.2022.9846719.
- [16] HAN K, WANG Y, TIAN Q, et al. Ghostnet: More features from cheap operations[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 1580-1589, DOI: 10.1109/CVPR42600.2020.00165.
- [17] HOU Q, ZHOU D, FENG J. Coordinate attention for efficient mobile network design[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2021: 13713-13722, DOI: 10.1109/CVPR46437.2021.01350.
- [18] HUANG Y, QIU C, YUAN K. Surface defect saliency of magnetic tile[J]. The Visual Computer, 2020, 36(1): 85-96.
- [19] HU J, SHEN L, ALBANIE S, et al. Squeeze-and-excitation networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8): 2011-2023.
- [20] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional block attention module[C]. Proceedings of the European Conference on Computer Vision (ECCV), 2018: 3-19, DOI: 10.48550/arXiv.1807.06521.

### 作者简介

宁奉阁, 硕士研究生, 主要研究方向为图像处理技术。

E-mail: 2694400910@qq.com

石金进(通信作者), 博士, 讲师, 硕士生导师, 主要研究方向为机器视觉。

E-mail: shijinjinhit@sina.com