

基于改进的 VGG16 网络金属表面缺陷图像分类研究

胡坤¹ 吴国庆¹ 胡祖辉^{2*} 王忠明¹

¹(南通大学电气工程学院 江苏 南通 226019)

²(南通大学信息科学技术学院 江苏 南通 226019)

摘要 针对工业生产中金属表面缺陷识别存在人力消耗大、效率不高等问题,提出改进的 VGG16 网络金属表面缺陷图像分类方法。以 VGG16 网络为基础,引入注意力机制 CBAM 增强特征学习能力,引入 Inception 网络结构拓宽网络宽度,以此增强模型非线性能力;对输入图像做数据增强处理,提高网络模型鲁棒性。实验证明,改进的网络模型在数据集 GC10-DET 上准确率达到 90.23%,在数据集 NEU-CLS 上准确率达到 98.84%。实验结果表明该方法在金属表面缺陷分类上具有良好的实际应用意义。

关键词 缺陷图像分类 VGG16 网络 注意力机制 Inception 网络

中图分类号 TP391

文献标志码 A

DOI:10.3969/j.issn.1000-386x.2024.06.026

METAL SURFACE DEFECT IMAGE CLASSIFICATION BASED ON IMPROVED VGG16 NETWORK

Hu Kun¹ Wu Guoqing¹ Hu Zuhui^{2*} Wang Zhongming¹

¹(School of Electrical Engineering, Nantong University, Nantong 226019, Jiangsu, China)

²(School of Information Science and Technology, Nantong University, Nantong 226019, Jiangsu, China)

Abstract Aimed at the problems of large labor consumption and low efficiency in metal surface defect recognition in industrial production, an improved VGG16 network metal surface defect image classification method is proposed. Based on the VGG16 network, the attention mechanism CBAM was introduced to enhance the feature learning ability, and the Inception network structure was introduced to broaden the network width, thereby enhancing the nonlinear ability of the model. The input image data was processed to improve the robustness of the network model. Through experimental verification, the improved network model has an accuracy rate of 90.23% on the data set GC10-DET, and an accuracy rate of 98.84% on the data set NEU-CLS. The experimental results show that this method has good practical application significance in the classification of metal surface defects.

Keywords Defect image classification VGG16 network Attention mechanism Inception network

0 引言

随着钢铁工业发展的不断深入,在机械制造、汽车生产、航空航天等领域对于对金属加工器件的质量要求不断提升,对材料、期间的检测技术及标准要求也越来越高。在金属加工生产的过程中,受生产环境、设备性能、工艺技术等因素的影响,可能会对金属表面造

成一定的缺陷,如气泡、孔洞、划痕等,若没有及时检测加以修复,这将对企业造成极大的经济损失,严重时会在工件使用过程中发生重大事故。因此为了提高工艺质量,保证产品竞争力,促进经济发展,金属缺陷识别检测技术尤为重要。基于机器视觉的表面缺陷检测当下研究的热点方向^[1],Pernkopf 等^[2]提出一种具有三位特征的钢块表面缺陷检测方法利用贝叶斯网络分类器提取缺陷特征进行分类。Ravikumar

等^[3]提出用直方图法提取机械零件表面缺陷并用决策树法进行分类。Zhang 等^[4]提出利用小波变化去除金属表面图像噪声,采用支持向量机的方法对缺陷进行分类。Park 等^[5]提出了一种卷积神经网络成像分析,识别检测几种零件表面缺陷。Wang 等^[6]提出一种云计算环境中使用 Faster R-CNN 算法的智能表面检测系统,更加有效地解决了识别复杂产品图像中的小缺陷问题。金属表面缺陷识别分类是机器视觉表面缺陷检测系统中重要的研究内容,如何提高识别算法的准确性和鲁棒性,一直是研究人员努力的方向。

本文以 VGG16 网络为基础对金属表面缺陷数据集做图像分类实验^[7]。数据集采用了 GC10-DET 金属表面数据集,针对该数据集首先进行图像预处理,对图像中缺陷特征进行提取,后因数据量小的问题,对数据集图像做了翻转、偏移、缩放等数据增强技术,提高了模型的鲁棒性;针对金属缺陷种类多,缺陷表现不一等问题,在 VGG16 网络中添加了 Inception 网络结构拓宽网络的宽度,由此来增强网络的非线性能力^[8],并且在卷积层之间引入了注意力机制 CBAM^[9],增强缺陷的权重参数,使之更好地提取缺陷特征,最终形成了改进的 VGG16 网络结构。金属表面缺陷图像分类准确率达到 90.23%,相比 VGG16 网络提高 3.63 百分点。在东北大学发布的 NEU-CLS 钢材表面缺陷数据集准确率达到 98.84%。

1 相关工作

1.1 注意力机制

对于卷积神经网络模型,影响其性能主要有三个因素:深度、宽度、基数。文献[9]提出了另外一种影响网络性能的因素,也就是注意力机制(CBAM, Convolutional Block Attention Module),注意力机制包括两方面内容:通道注意力(Channel Attention Module)、空间注意力(Spatial Attention Module)。通道注意力主要对不同通道特征图进行特征信息处理来得到特征明显的通道;空间注意力主要对特征图上不同特征区域进行特征信息处理得到特征明显的区域。通过引入注意力机制可以加强网络的表达能力,提高网络的学习能力,针对特征区域加强学习来提高网络学习的准确率^[10]。注意力机制结构如图 1 所示。

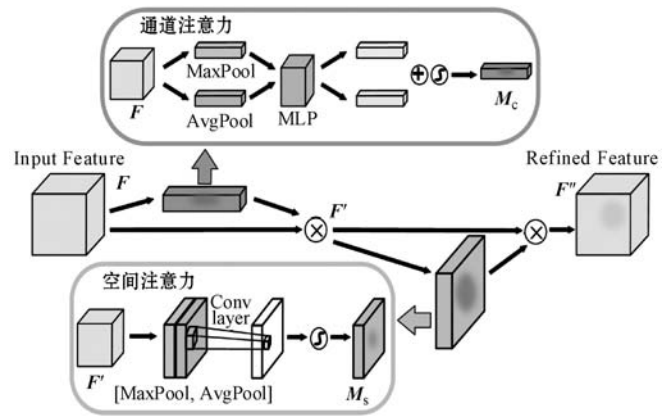


图 1 注意力机制结构

假设输入特征图为 $F \in \mathbf{R}^{C \times H \times W}$, F' 为执行通道注意力的特征图, $M_c \in \mathbf{R}^{C \times 1 \times 1}$ 为执行通道注意力的注意力图, F'' 为执行空间注意力的特征图, $M_s \in \mathbf{R}^{1 \times H \times W}$ 为执行空间注意力的注意力图,整个注意力过程公式如式(1)所示。

$$F' = M_c(F) \otimes F \quad (1)$$

$$F'' = M_s(F') \otimes F'$$

在通道注意力过程中,首先使用平均池化法压缩特征图空间维数得到 F_{avg}^c ,使用最大池化法汇聚特征图重要特征得到 F_{max}^c ,其次将 F_{avg}^c 与 F_{max}^c 传入多层感知机(MLP)得到注意力图 $M_c \in \mathbf{R}^{C \times 1 \times 1}$,按元素求和得到输出的特征向量,计算公式如式(2)所示, σ 表示 sigmoid 函数, $W_0 \in \mathbf{R}^{C/r \times C}$ 和 $W_1 \in \mathbf{R}^{C \times C/r}$ 为 MLP 权重。

$$M_c(F) = \sigma(MLP(AvgPool(F))) + MLP(MaxPool(F)) = \sigma(W_1(W_0(F_{avg}^c))) + W_1(W_0(F_{max}^c)) \quad (2)$$

在空间注意力过程中,同样使用平均池化法和最大池化法得到 F_{avg}^s 与 F_{max}^s 并把它们连接起来,再对特征区域通过一层卷积层得到空间注意力图 $M_s \in \mathbf{R}^{H \times W}$,计算公式如式(3)所示,其中 $f^{7 \times 7}$ 为 7×7 的卷积运算。

$$M_s(F) = \sigma(f^{7 \times 7}([AvgPool(F); MaxPool(F)])) = \sigma(f^{7 \times 7}([F_{avg}^s; F_{max}^s])) \quad (3)$$

通过通道注意力机制与空间注意力机制得到了更加重要的特征信息,提高了网络的学习效率。

1.2 Inception 网络结构

Inception 网络结构是一种高效表达特征的稀疏型结构,源自 GooLeNet 的网络结构。与传统网络不同的是,该网络通过增加网络的宽度来解决网络深度增加所带来的性能饱和问题,并且在保证分类准确率提升或保持不降的同时降低计算开销与内存开销^[11]。Inception 网络结构如图 2 所示,将上一层输入进行多尺度特征提取及池化,得到多个尺度的图像特征,最后再将特征进行叠加输出,为了减少 3×3 卷积层、 5×5 卷积层及 3×3 池化层所带来的计算问题,在相应网络

层前面加入一个 1×1 卷积层^[12]。这样增加了网络的宽度与尺度适应性,从而增加了网络的非线性能力^[13]。

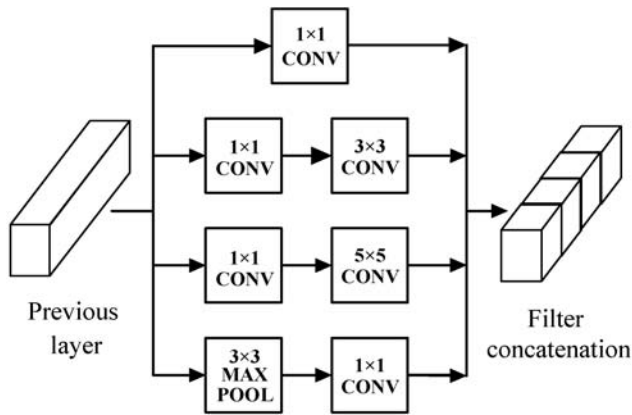


图 2 Inception 网络结构

2 改进的 VGG 网络模型

VGG 是一种适用于分类的卷积神经网络,扩展性很强,迁移到其他图片数据上的泛化性能非常好,并且 VGG 结构简洁,通过反复堆叠 3×3 的小型卷积核 (Convolution) 和 2×2 的最大池化层 (Maxpooling), 构建了 11 ~ 19 层神经网络;利用多个堆叠的更小尺寸的卷积核取得的效果优于更大尺寸的卷积核,两个 3×3 的卷积核串联相当于一个 5×5 的卷积核,3 个 3×3 的卷积核相当于 1 个 7×7 的卷积核,这样通过多个 3×3 的卷积核串联,可以增加网络的深度,增强网络的非线性能力,允许学习更复杂的特征,模型的分类能力更强^[14]。对于本文所采用的金属缺陷数据集,考虑 VGG 网络的优势,选取 VGG16 网络作为实验的基础网络,并在此基础上对 VGG16 网络做了一些改进,让网络模型学习特征的能力更强。改进后 VGG16 网络结构如图 3 所示。

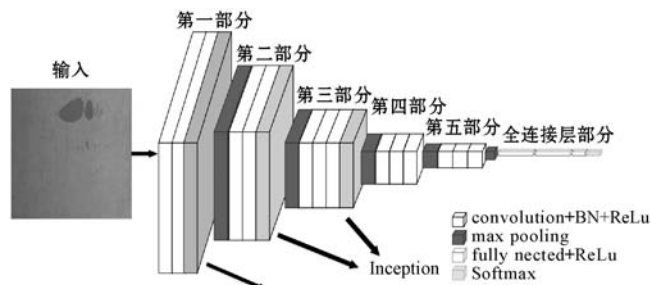


图 3 改进后 VGG16 网络结构

针对金属表面缺陷数据集缺陷的多样性以及图像数量有限的问题,引入注意力机制能够更好、更准确地学习到缺陷特征,从而提高缺陷识别的准确率。

针对 VGG 网络结构的单一性问题,以此加入 Inception 网络结构来拓宽网络宽度,在不增加网络的计算负担下,更好地表现金属表面缺陷的特征信息。

VGG16 网络按最大池化层为结点把卷积层分为 5 个部分,第一部分输入图像信息为 $224 \times 224 \times 64$,第二部分输入图像信息为 $112 \times 112 \times 128$,第三部分输入图像信息为 $56 \times 56 \times 256$,第四部分输入图像信息为 $28 \times 28 \times 512$,第五部分输入图像信息为 $14 \times 14 \times 512$ 。在第一部分卷积层后引入注意力机制 CBAM 模型解决缺陷特征表象不一导致学习困难的问题;在第二部分及第三部分卷积层后引入 Inception 网络结构拓宽网络宽度,增强网络的非线性能力,进一步加强网络对缺陷信息的学习。另外在每一层卷积层后再加上一层批量归一化层 (Batch Normalization, BN),以防止梯度消失或爆炸,加快训练速度,提高网络的泛化能力。网络层参数如表 1 所示。

表 1 网络层参数

层	感受野大小	填充	步长	通道数
(Conv1 + BN) × 2	3×3	1	1	64
CBAM				64
Maxpooling	2×2		2	128
(Conv2 + BN) × 3	3×3	1	1	128
Inception1	multiple			128
Maxpooling	2×2		2	256
(Conv3 + BN) × 3	3×3	1	1	256
Inception2	multiple			256
Maxpooling	2×2		2	512
(Conv4 + BN) × 3	3×3	1	1	512
Maxpooling	2×2		2	512
(Conv5 + BN) × 3	3×3	1	1	512
Maxpooling	2×2		2	512
Fc1				4 096
Fc2				4 096
Fc3				1 000
Softmax				10

改进后模型功能框如图 4 所示。首先采集数据集图像;其次对图像进行预处理,采用降采样方法把图像降成 VGG16 网络标准的 224×224 像素大小;然后对图像数据进行增强处理,利用 Keras 库中的 ImageDataGenerator 函数在每一次样本训练前对图形进行旋转、错切变换、缩放、翻转来增强数据,提高模型的泛化能力;接着利用注意力机制 CBAM 与 Inception 网络结构来提取图像信息的特征参数;紧接着通过 VGG16 网络机构融合学习到的参数信息,最后通过 Softmax 分类器来识别图像分类的结果。

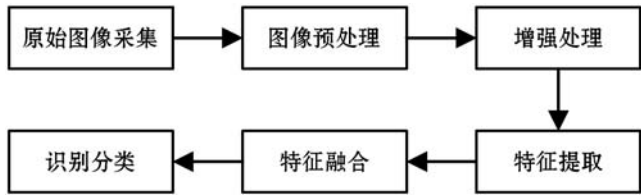


图4 模型功能框

3 缺陷图像分类实验

3.1 缺陷数据集

本文所采用的数据集来自天津大学提出的金属表面缺陷检测数据集 GC10-DET^[15], 该数据集图像来自实际工业操作中, 相机型号为 DALSA-CM-04K08A 的 CCD 摄像机线扫描得到的钢板表面缺陷, 缺陷包括冲孔、焊缝、月牙湾、水斑、油斑、丝斑、异物压入、压痕、严重折痕、腰折十种表面缺陷, 示意图如图 5 所示。

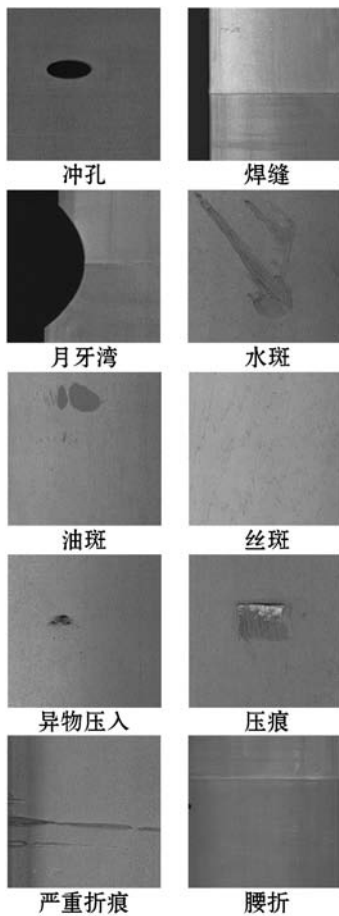


图5 缺陷示意图

冲孔:在带钢生产线上,带钢需要按照产品规格进行冲孔;机械故障可能会导致多余的冲孔,导致冲孔缺陷。

焊接:当带钢发生变化时,需要将带钢的两个线圈进行焊接,形成焊缝。

月牙湾:钢带在生产过程中,切割有时会产生缺

陷,就像半圈一样。

水斑:生产过程中干燥产生水斑。水斑对比度低,往往可能会错误检测为油斑等缺陷。

油斑:油斑通常是由机械润滑油污染引起的,会影响产品的外观。

丝斑:在带钢表面上出现的局部或连续的波状斑块,主要原因是辊体温度不均匀和压力不均匀。

异物压入:夹杂物是一种典型的金属表面缺陷,通常表现为细小的斑点状、鱼鳞状、条状、块状不规则分布在带钢表面上,并常伴有粗糙麻子状表面。有些夹杂物疏松易脱落,有些被压入板内。

压痕:压痕是钢板表面有斑点、片状或条状的周期性凸起或凹坑,主要是由工作辊或张力辊损坏引起的。

严重折痕:折痕是垂直的横向褶皱,其主要原因是开卷过程中带钢在移动方向上出现局部曲折。

腰折:缺陷部位有明显的褶皱,较普遍,有点像褶皱,说明缺陷局部变形过大,原因在于低碳。

在 GC10-DET 数据集中,图片大小为 400 KB 左右,图像为灰度图像,图片像素大小为 2 048 × 1 000, VGG 网络的标准输入像素大小为 224 × 224,因此对原数据集图像进行预处理,裁剪为像素大小为 224 × 224 的带有缺陷特征区域的图像,最终得到 1 949 幅缺陷图像。由于数据量有限的原因,采用图像随机旋转、图像缩放、图像平移的数据增强技术对训练数据集做增强处理,来防止学习中过拟合现象的出现。

3.2 实验环境及参数

本实验的深度神经网络搭建在基于 Ubuntu 18.04.4 LTS 系统下 TensorFlow 2.2 深度学习框架上,并采用基于 Python 3.6 版本语言编程实现,处理器为英特尔至强处理器 E5-2640 v4 @ 2.40 GHz × 10,内存为 8 GB, GPU 为 NVIDIA TITAN Xp, Cuda 版本为 10.1。

本次实验中,将数据集样本按 7:3 分为训练集与测试集。为了观察更加完整清晰,训练集设定迭代次数 Epoch 取为 200。根据实验设备的计算能力及样本训练大小 Batch size 对模型实验结果的影响最终确定 Batch size 取为 64。学习率使用 Adam 优化器使得在迭代过程中收敛速度加快,通过初始学习率 Learning_rate 为 0.1、0.01、0.001、0.000 1 和 0.000 01 的模型实验结果对比,最终确定 Learning_rate 取为 0.000 01。为防止过拟合问题,改善模型的鲁棒性,丢弃比率 Dropout 取为 0.2。

3.3 实验结果与讨论

本文针对金属表面缺陷图像分类问题进行研究,在现有方法的基础上提出改进的 VGG16 网络模型,并

与其他基础网络做比较。

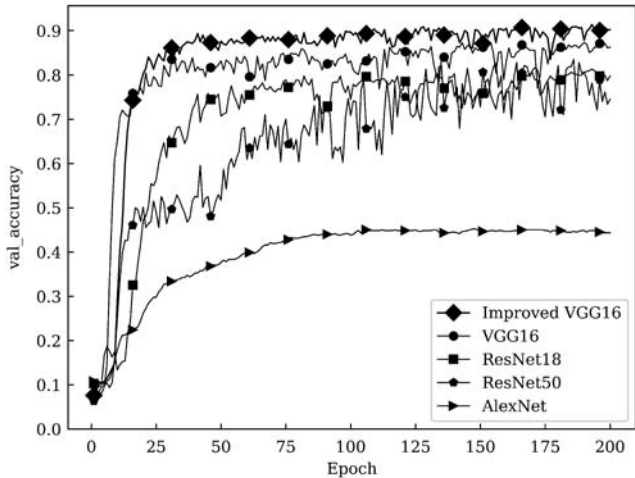


图 6 GC10-DET 测试集准确率对比图

实验结果如图 6 所示,改进的 VGG16 具有较快的收敛速度以及较高的准确性,最准确率达到 90.23%,相比 VGG16 提高 3.63 百分点,准确率远高于其他对比网络且波动性较小,对于金属表面缺陷图像分类具有一定的实际意义。

改进的 VGG16 在 GC10-DET 测试集各缺陷测试结果如表 2 所示。金属表面各缺陷分类结果精确度较高,其中异物压入缺陷与油斑缺陷精确度相对较低,通过观察数据集图像发现,原因在于异物压入缺陷表象不一,没有确定的缺陷特征,难于学习;而油斑缺陷与水斑缺陷极易混淆,两者之间仅是对比度上的差异,也存在学习困难的问题。压痕与严重折痕缺陷召回率相对较低,通过观察数据集发现,两类缺陷输入测试数量较少,存在数据不平衡现象,这对网络学习也造成一定难度。

表 2 GC10-DET 测试集各缺陷测试结果

缺陷	精确率	召回率	F1 分数	支持度
冲孔	0.96	1.00	0.98	74
焊接	0.91	0.86	0.89	87
月牙湾	0.92	0.98	0.95	60
水斑	0.90	0.92	0.91	78
油斑	0.81	0.82	0.82	62
丝斑	0.94	0.97	0.95	63
异物压入	0.82	0.86	0.84	49
压痕	0.89	0.77	0.83	22
严重折痕	0.85	0.75	0.80	44
腰折	0.98	0.96	0.97	45

图 7 为 GC10-DET 测试集混淆矩阵可视化热力

图,可以更直观地看到各类缺陷的测试情况。横坐标为测试结果,纵坐标为真实结果,从热力图中可以看到各类测试结果与真实结果极其吻合,数据集中在对角线上,颜色深浅表示集中程度,另外也可以直观看到数据不平衡现象。

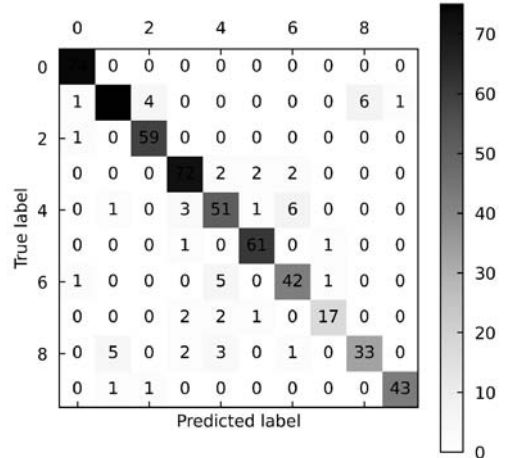


图 7 GC10-DET 测试集混淆矩阵可视化热力图

为了进一步验证改进后 VGG16 对于缺陷识别的有效性,同时选取了东北大学公开金属缺陷数据集 NEU-CLS 做实验验证^[16]。该数据集包括轧制氧化、斑块、开裂、点蚀表面、内含物和划痕六类缺陷,每种缺陷包含 300 幅 200×200 像素大小的图片,对比实验结果如表 3 所示。文献[17]采用一种基于小波变换和灰度共生矩阵的方法,利用二维 2 离散小波变换的带图像低频子带信息,再构建 4 个方向的灰度共生矩阵提取特征向量,最后用支持向量机识别分类,精度达到 96.17%^[17]。文献[18]提出一种深度残差网络 ResNet 模型引入软注意力机制的方法,改进后 ResNet50 达到 98.61%^[18]。而本文中改进的 VGG16 准确率达到 98.84%,相对其他方法准确率有一定提高。

表 3 NEU-CLS 数据集对比实验

网络模型	测试集准确率/%
文献[17]	96.17
文献[18]	98.61
AlexNet	95.22
ResNet18	96.85
ResNet50	92.45
VGG16	96.19
Improved VGG16	98.84

图 8 所示为改进的 VGG16 在 NEU-CLS 测试集混淆矩阵可视化热力图,可以看出在各类缺陷识别上准确率较高,其中第六类划痕缺陷准确率略有下降。

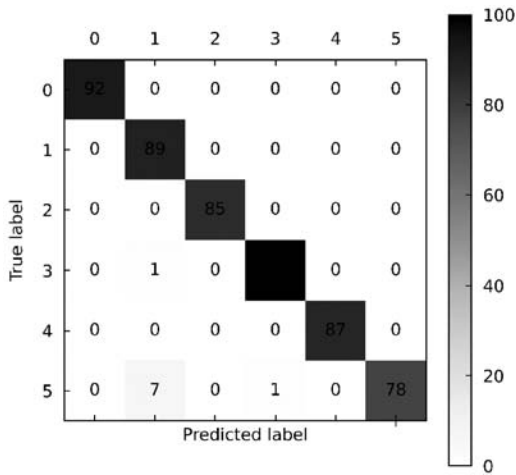


图8 NEU-CLS 测试集混淆矩阵可视化热力图

通过以上实验表明,本文改进后的 VGG16 在金属缺陷图像分类上有较高的准确性,但还有提升的空间;由于金属表面缺陷有一定的多样性,部分缺陷并非单一特性,这对缺陷分类提高了难度。

4 结 语

针对金属表面缺陷图像分类问题,提出了改进的 VGG16 金属表面缺陷图像分类方法,在基础 VGG16 网络模型卷积层间引入注意力机制 CBAM 与 Inception 网络结构,在不增加计算负担的情况下,实现了对多样性缺陷特征信息的学习性能提升,增强了网络的非线性能力,提高了模型的分分类效果。在数据增强后的金属表面缺陷数据集上做实验验证,并做其他基础网络的平行对比实验,结果表明,本文方法在数据集 GC10-DET 与数据集 NEU-CLS 上都得到良好的分类效果,相对其他方法也有一定提升,说明在本文模型在金属缺陷识别分类上有一定的应用价值。

考虑到金属表面缺陷数据量较少的原因,在后期工作中需要增加样本数据集,针对部分缺陷样本较少的问题,需要把数据不平衡的分类研究作为研究重点。

参 考 文 献

- [1] 汤勃,孔建益,伍世虔. 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.
- [2] Pernkopf F. Detection of surface defects on raw steel blocks using Bayesian network classifiers[J]. Pattern Analysis & Applications,2004,7(3):333-342.
- [3] Ravikumar S, Ramachandran K I, Sugumaran V. Machine learning approach for automated visual inspection of machine components[J]. Expert Systems with Applications,2011,38(4):3260-3266.
- [4] Zhang X W, Ding Y Q, Lv Y, et al. A vision inspection system for the surface defects of strongly reflected metal based on multiclass SVM[J]. Expert Systems with Applications,2011,38(5):5930-5939.
- [5] Park J K, Kwon B K, Park J H, et al. Machine learning-based imaging system for surface defect inspection[J]. International Journal of Precision Engineering and Manufacturing-Green Technology,2016,3(3):303-310.
- [6] Wang Y B, Liu M G, Zheng P, et al. A smart surface inspection system using faster R-CNN in cloud-edge computing environment[J]. Advanced Engineering Informatics,2020,43:1037-1040.
- [7] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[EB]. arXiv:1409.1556,2014.
- [8] Szegedy C, Liu W, Jia Y Q, et al. Going deeper with convolutions[EB]. arXiv:1409.4842,2014.
- [9] Woo S, Park J, Lee J Y, et al. CBAM: Convolutional block attention module[C]//15th European Conference on Computer Vision,2018:3-19.
- [10] 郑欣悦,黄永辉. 基于 VAE 和注意力机制的小样本图像分类方法[J]. 计算机应用与软件,2019,36(10):168-174.
- [11] 宋建锋,韦玥,苗启广,等. 压缩激励机制驱动的尿液细胞图像分类算法[J]. 西安电子科技大学学报,2020,47(2):39-45,59.
- [12] Alom M Z, Hasan M, Yakopcic C, et al Improved inception-residual convolutional neural network for object recognition [J]. Neural Computing & Applications, 2017, 32: 279-293.
- [13] Wang W, Zhang C W, Tian J G, et al. High-resolution radar target recognition via inception-based VGG (IVGG) networks[J]. Computational Intelligence and Neuroscience, 2020(5):1-11.
- [14] 杜超,刘桂华. 改进的 VGG 网络的二极管玻壳图像缺陷检测[J]. 图学学报,2019,40(6):1087-1092.
- [15] Lv X M, Duan F J, Jiang J, et al. Deep metallic surface defect detection: The new benchmark and detection network [J]. Sensors,2020,20(6):1562.
- [16] Song K C, Yan Y H. A noise robust method based on completed local binary patterns for hot-rolled steel strip surface defects[J]. Applied Surface Science,2013,285:858-864.
- [17] 单东日,童灿,乃学尚,等. 基于小波和灰度共生矩阵的带钢表面缺陷识别[J]. 制造技术与机床,2020(2):120-123.
- [18] 陆雅诺,陈炳才,陈德刚,等. 一种基于注意力模型的带钢表面缺陷识别算法[J]. 激光与光电子学进展,2021,58(14):242-250.