文章编号:1671-4598(2025)05-0062-07 DOI:10.16526/j.cnki.11-4762/tp.2025.05.008 中图分类号:TP391.4 文献标识码:A

基于改进 ResNet50 金属表面缺陷检测模型

丁冠华[1,2],姚 旭[1,2]
(1. 延安大学 物理与电子信息学院,陕西 延安 716000
2. 陕西省能源大数据智能处理省市共建重点实验室,陕西 延安 716000)

摘要:针对传统的金属表面缺陷检测存在的成本高、效率低,难以满足现代制造业对高精度和高效率的需求的问题,提出了一种基于 ResNet50 和改进可形变卷积的金属缺陷识别模型;该模型在 ResNet50 基础上设计了可形变卷积的残差连接结构,增强了对复杂缺陷形状和位置的适应能力,并在残差连接前引入卷积注意力模块以更好地捕捉显著缺陷特征;通过将该结构嵌入全局平均池化层之前,形成了新的网络架构 ResNet50_DCAB;在金属缺陷数据集上的实验结果显示,ResNet50_DCAB 模型的验证集最高准确率达到 98.6%,初始准确率为 97.2%,均优于常用深度学习模型;结果表明,ResNet50_DCAB 在金属缺陷检测任务中表现出较高的识别精度和鲁棒性,具有较大的工业应用潜力。

关键词:金属表面缺陷检测;ResNet;残差网络;注意力机制;可形变卷积

Metal Surface Defect Detection Model Based on Improved Resnet50

DING Guanhua[1,2], YAO Xu[1,2]
(1. School of Physics and Electronic Information, Yan'an University, Yan'an 716000, China;
2. Shaanxi Provincial Key Laboratory of Energy Big Data Intelligent Processing, Yan'an 716000, China)

Abstract: To address the issues of high cost and low efficiency in traditional metal surface defect detection, it is difficult for modern manufacturing to meet the requirements of high precision and high efficiency. A metal defect recognition model based on ResNet50 and improved deformable convolution is proposed. The model incorporates a residual connection structure with deformable convolution based on ResNet50 network, enhancing its adaptability to complex defect shapes and positions. Additionally, a convolutional attention module is introduced before the residual connection to better capture the features of significant defects. By embedding this structure before the global average pooling layer, a new network architecture, named ResNet50_DCAB, is formed. Experimental results on a metal defect dataset show that the ResNet50_DCAB model achieves a maximum validation accuracy of 98.6% and an initial accuracy of 97.2%, outperforming commonly used deep learning models. The results demonstrate that the ResNet50_DCAB model exhibits high recognition accuracy and robustness in metal defect detection tasks, making it highly promising for industrial applications.

Keywords: metal surface defect detection; ResNet network; residual network; attention mechanism; deformable convolution

0 引言

金属表面缺陷检测在工业制造过程中起着至关重要的作用,直接影响到产品的质量、性能和使用寿命。特别是在航空航天、汽车制造、船舶制造、电子设备和建筑材料等领域,对金属表面的质量有着严格的要求,即便是微小的缺陷,如裂纹、划痕、腐蚀、凹坑或异物污染,都可能对金属部件的结构完整性和功能性能产生严重影响。因此,金属表面缺陷的准确检测和及时修复对于保障产品质量和安全至关重要。传统的检测方法通常依赖于人工目视检查和简单的光学手段,但这些方法存在劳动强度大、主观误差高的不足,并且在复杂且大规模的生产线中效率较低,难以满足现代工业对高精度和高效率的要求。随着工业自动化和智能制造技术的不断进步,基于图像处理和机器学习的自动化检测技术逐渐成为主流,并被认为是提高检测准确性和效率的重要

收稿日期:2024-11-18; **修回日期**:2025-01-23。
基金项目:国家自然科学基金(52365069);研创项目(YCX2024095)。
作者简介:丁冠华(1999-),男,大学本科。
通讯作者:姚 旭(2001-),男,硕士研究生。
引用格式:丁冠华,姚 旭.基于改进 Resnet50 金属表面缺陷检测模型[J].计算机测量与控制,2025,33(5):62-68,78.

途径。

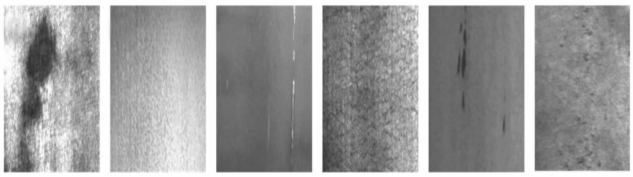
近年来，随着图像处理和计算机视觉技术的发展，机器学习和深度学习被广泛应用于工业领域。在金属表面缺陷检测领域，文献［1］提出了一种基于卷积神经网络（CNN，convolutional neural network）的快速且鲁棒的缺陷检测模型[1]。使用一个改进的 CNN 架构来处理金属表面图像，通过层层卷积操作自动提取图像的特征，减少对人工设计特征的依赖。文献［2］提出了一种改进的实时金属表面缺陷检测模型，基于 YOLOv5 并加入了通道混洗机制[2]。通过深度卷积和通道混洗来优化特征提取，提高了对小缺陷的检测效果。文献［3］通过使用支持向量机（SVM，support vector machine）作为分类模型，将提取的纹理特征输入 SVM 中，以区分正常表面和含有缺陷的表面[3]，文献［3］提出的“纹理特征提取＋SVM”方案是一种经典的传统机器学习方法，在数据量小、资源受限的场景下具备优势，且方法简单、易实现。然而，随着任务复杂度和数据规模的提升，该方法的局限性会逐渐显现。对于大规模工业应用，更现代化的深度学习方法可能更适合。文献［4］基于传统的支持向量机方法，提出了一种改进算法，结合图像预处理与特征提取，通过优化核函数和超参数来提高检测精度[4]。文献［5］提出了一种基于改进的 VGG16 网络的金属表面缺陷图像分类方法[5]。改进的 VGG16 网络通过引入多尺度特征融合和通道注意力机制，提升了网络在金属表面缺陷分类中的特征提取能力。金属表面缺陷检测的效率及准确率直接关系到国家工业生产的效率及利益，基于计算机视觉的缺陷检测更是重要研究内容，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。

尽管已有研究在金属缺陷检测方面取得了一定的进展，但面对缺陷种类多样、复杂且随机分布的特性，目前的方法仍然存在局限性。为了更好地解决这些问题，本文提出了一种基于改进 ResNet50 的金属缺陷检测模型。该模型通过引入可形变卷积（DCN，deformable convolution），增加卷积的空间自适应性，增强了对复杂缺陷形状的适应能力。此外，受 ResNet50 中 Bottleneck 结构的启发，本文设计了一个卷积核大小为 3×3 的可形变卷积的 Bottleneck 结构，并在残差连接之前引入了卷积注意力模块（CBAM，convolutional block attention module），以进一步强调特征的显著性。通过将可形变卷积与注意力机制结合，构成一种新的注意力模块——可形变卷积注意力模块（DCAB，deformable convolutional attention block）。该模块不仅增强了对复杂形变特征的提取能力，还有效解决了特征捕获中的不均衡问题，显著提升了金属缺陷检测的精度和鲁棒性。为了验证该方法的有效性，本文对改进后的 ResNet50 网络进行了超参数调优，并与其他主流模型进行了对比，实验结果表明，所提出的模型在金属缺陷检测任务中具有更高的准确性和更好的泛化性能。

1　数据集增强与划分

1.1　数据集介绍

数据集采用由东北大学研究团队公开的 NEU-DET[6]（Northeastern University Detection of Surface Defects）金属表面缺陷数据集。数据集由金属表面图像构成，每张图像都包含一个或多个表面缺陷。每个缺陷都属于特定的类别，并带有相应的标签。该数据集共有 6 个类别 1 800 张图片，分别为氧化铁皮（Rs）300 张，裂纹（Cr）300 张，夹杂物（In）300 张，乳化液斑迹（Pa）300 张，麻点（Ps）300 张，划痕（Sc）300 张。每张图片分辨率均为 200 pixels×200 pixels。数据集各类别缺陷图像如图 1 所示。



(a)Ps　(b)Pa　(c)Sc　(d)Cr　(e)In　(f)Rs

图 1　金属表面缺陷数据集

1.2　小样本数据集 NEU-DET 的增强与划分

针对该数据集数据样本过少，可能存在过拟合的风险。为此对数据集进行扩充是首要解决的问题[7]。首先按照 8：2 的比例将数据集划分为训练集和测试集，其次通过镜像翻转、平移、缩放、旋转以及添加高斯噪声等操作，将原始训练集中每个类别的数据扩充到 700 张，共 4 200 张图片。并将扩充后的数据集按照不同缺陷类别，放入不同文件夹中，方便后续训练。扩充后的训练集在一定程度上提高了模型的鲁棒性和泛化性[8]。经数据增强和划分后的数据集数量如表 1 所示。

表 1　金属表面缺陷数据集数量分布表

缺陷类型	氧化铁皮	裂纹	夹杂物	乳化液斑迹	麻点	划痕
扩充前	300	300	300	300	300	300
扩充后	760	760	760	760	760	760

2　算法设计

2.1　ResNet50

随着深度学习的发展，人们发现更深的网络能够提取更多的特征，于是开始不断地加深网络的深度，但是随之而来也产生了新的问题。传统的卷积神经网络的输入输出是直接连接，一味地加深网络深度很容易会出现梯度爆炸或梯度消失的现象，使得模型难以收敛。针对这一问题，He 等人[9]在 2015 年提出了残差网络（Res-

Net，residual network），通过引入残差块（Residual Block），使得训练非常深的网络成为可能，且性能显著提升。残差块公式如式所示：

$$y = F(x,\{W_i\}) + x \tag{1}$$

其中：F（x，$\{W_i\}$）表示卷积层的输出，表示卷积后得到的输出；x 是残差块的输入（来自上一层的特征图）；y 是残差块的输出。这里将输入 x 直接加到卷积层的输出上就形成了残差块。由于 ResNet50 在性能和计算复杂度之间取得了良好的平衡，在图像分类任务中得到了广泛的应用，故本文选取 ResNet50 做为主干网络。ResNet50 的残差块结构如图 2 所示。

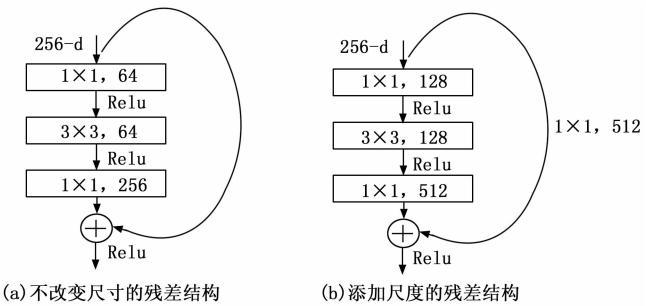


图 2 ResNet50 基本残差结构

ResNet50 由多个残差块堆叠而成，总共包含 50 层，其中包含 1 个卷积块、4 个残差块和 1 个输出。卷积块是由一个 7×7 大小的卷积层、批归一化层（BN，batch normalization，）、激活层（ReLU）、以及一个 3×3 的最大池化层组成；残差块的组成部分包括恒等残差块（即不改变尺寸的残差结构）与卷积残差块（即添加尺度的残差结构）。每个残差块中的卷积残差块数相同而恒等残差块数不同，其中卷积残差块数均为 1，恒等残差块数依次为 2、3、5 和 2。恒等残差块与卷积残差均由尺寸大小为 1×1、3×3、1×1 的卷积核组成[10]。

2.2 可形变卷积

可形变卷积是一种增强卷积神经网络感受野和建模复杂几何形变的有效方法，最早由 Dai 等人[11]在 2017 年提出。传统的卷积操作采用固定的空间位置进行采样，然而在面对和处理复杂的图像变化或物体几何形状变化时，固定采样可能无法捕捉图像中的局部变化。为解决这一问题，可形变卷积在标准卷积的基础上引入了形变偏移量，使得卷积核的采样点能够根据输入图像的特征自适应地调整位置，从而有效提升网络的特征提取能力[12]。标准卷积与可形变卷积采样点如图 3 所示。图 3 中（a）是常见的 3×3 卷积方式，（b）是采样可形变卷积，加上偏移量之后的采样点变化，而（c）和（d）则是可形变卷积的特殊变化情况。普通卷积操作可用下式表示：

$$y(p_0) = \sum_{p_n \in R} w(p_n) \cdot x(p_0 + p_n) \tag{2}$$

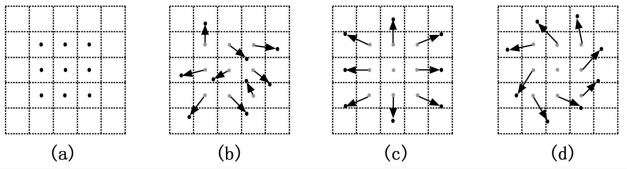


图 3 可形变卷积不同采样点

其中：y（p_0）表示输出特征图的某个位置，p_n 表示区域内任意位置，p_0 表示中心位置，w 为卷积核的权重，x 和 y 分别表示输入和输出。而在可形变卷积中，卷积核的位置不再固定，而是增加了一个偏移量 $\triangle p_n$，即可形变卷积的公式如式（3）所示：

$$y(p_0) = \sum_{p_n \in R} w(p_n) \cdot x(p_0 + p_n + \Delta p_n) \tag{3}$$

2.3 注意力机制

注意力机制近年来被广泛应用于提升卷积神经网络的表现[13]，CBAM 是一种轻量且有效的注意力模块，能够嵌入到现有的卷积神经网络中，从而增强网络的特征提取能力。CBAM 由 Woo 等人[13]在 2018 年提出，模块通过逐步对特征图进行通道注意力和空间注意力的增强，以便让网络能够更精准地关注到更重要的特征区域。CBAM 模块由两个子模块构成：通道注意力模块（Channel Attention Moudle）[14]和空间注意力模块（Spatial Attention Moudle），而 CBAM 则是将这两个模块按顺序对输入特征图施加注意力操作，最终输出经过优化的输出特征图。CBAM 的设计理念是，先通过通道维度筛选出重要的特征通道，再通过空间维度选择关键的空间区域。这种逐步的注意力机制提高了网络对关键信息的关注度。注意力机制模块如图 4 所示。

在图 4 中若输入特征图为 $F \in R^{C\times H\times W}$，$M_c$（$F$）为执行通道注意力后输出特征图，其输出结果再与输入相乘得到 M_c（F）· F。此时输入经过通道注意力模块，得到了通道维度的特征关注加强。M_c（F）· F 做为下个阶段空间注意力模块的输入，M_s［M_c（F）］· F 为执行空间注意力后的输出特征图。此时输入特征图经过了通道维度和空间维度的特征加强。整个注意力过程的公式如式（4）所示：

$$F' = M_s(M_c(F)) \cdot F \tag{4}$$

在通道注意力模块中，其主要作用式通过挖掘各个通道的上下文信息，来为每个通道分配不同的权重，筛选出对当前任务重要的特征。首先，对输入特征图 F（i，j）在空间维度上并行地进行全局平均池化和全局最大池化，得到两个大小为 C 的描述向量，分别记为 F^c_{avg}和 $F^c_{\max}$，其公式如式（5）（6）所示。将这两个生成的描述向量通过一个共享全链接层，按元素求和，计算出通道注意力的权重，通道注意力的权重函数如式（7）所示。其中，W_0 和 W_1 表示全连接层的权重，δ 为 Re-

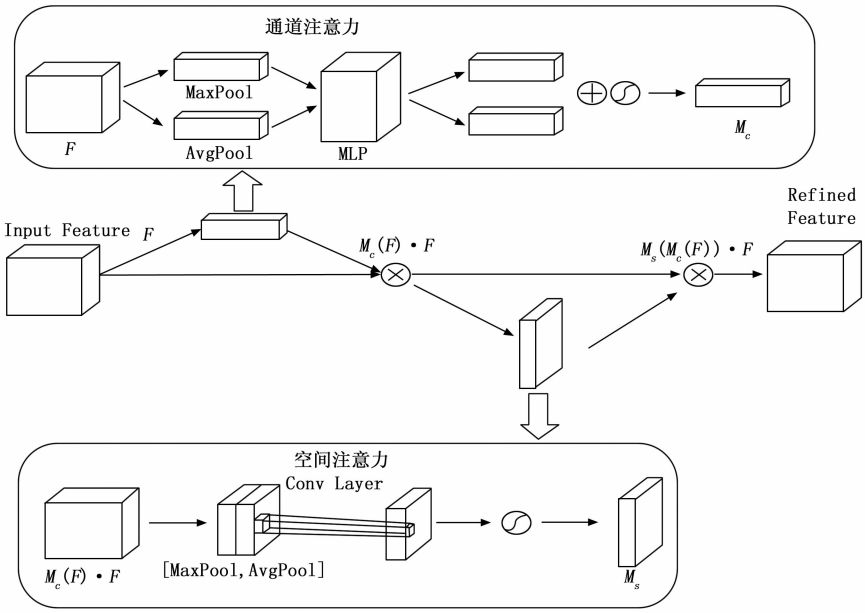


图 4　注意力机制

LU 激活函数，σ 是 Sigmoid 激活函数。最后的注意力权重通过逐元素相乘作用于原始图像中：

$$F^c_{\text{avg}} = \frac{1}{H\times W}\sum_{i=1}^{H}\sum_{j=1}^{W}F(i,j) \tag{5}$$

$$F^c_{\max} = \max_{i=1,\cdots,H,j=1,\cdots,W} F(i,j) \tag{6}$$

$$M_c(F) = \sigma\{W_1\{\delta[W_0(F^c_{\text{avg}})]\} + W_1\{\delta[W_0(F^c_{\max})]\}\} \tag{7}$$

在空间注意力模块中，其作用是根据通道注意力模块的输出特征图，进一步对空间维度的特征信息加以强调关注。首先，同样也是在通道维度上对输入特征图进行全局平均池化和全局最大池化，得到两个通道特征图，记为 F^s_{avg} 和 $F^s_{\max}$，公式如式（8）～（9）所示。将这两个特征图在通道维度上进行连接，生成一个 2 通道的特征图。接着，使用一个 7×7 的卷积操作生成空间注意力图。空间注意力公式如式（10）所示。其中，[;] 表示特征图在通道维度上进行拼接操作，$f^{7\times7}$ 表示 7×7 的卷积操作，σ 是 Sigmoid 激活函数。最终生成的空间注意力图将对输入特征图的每个空间位置进行加权：

$$F^s_{\text{avg}} = \frac{1}{C}\sum_{k=1}^{C}F_k \tag{8}$$

$$F^s_{\max} = \max_{k=1,\cdots,C} F_k \tag{9}$$

$$M_s(F) = \sigma\{f^{7\times7}[(F^s_{\text{avg}};F^s_{\max})]\} \tag{10}$$

CBAM 模块通过引入通道注意力和空间注意力机制，提升了卷积神经网络对重要特征的聚集能力，提升了网络学习的效率。

3　算法整体框架

3.1　DCAB 注意力模块

在原始的 ResNet50 网络中，其中的 3×3 卷积核在图像上的滑动窗口均为固定的，在金属缺陷数据集上由于复杂多变的表面缺陷，网络无法根据图像局部的几何变化自适应地调整感受野，这可能导致对复杂形变物体的特征提取不足。根据这一问题，提出引入可形变卷积，来增强对于局部几何变形的适应能力，更高效地提取复杂形变物体的特征。为了能够更加强调可形变卷积学习到的局部特征，且针对金属表面缺陷数据集的多样性以及数量有限等问题，引入注意力机制能够更好、更加准确地学习到缺陷特征，从而提升网络学习的效率[15]。受到 ResNet 网络中 BottleNeck 结构的启发，设计了一种基于可形变卷积的 BottleNeck 结构，能在空间中自适应地提取输入特征，同时在该残差连接之前引入通道注意力和空间注意力的混和模块，即 CBAM 模块，以加强可形变卷积中学习到的局部变化特征，和提升模型对目标特征的敏感度和表达能力。DCAB 模块结构如表 2 所示。

表 2　DCAB 模块结构

层级	输出通道数	操作说明
conv1	in_planes/2	1×1 卷积层
offset_conv	18（offset）	用于生成形变卷积的偏移量
deform_conv(conv2)	in_planes/2	可形变卷积 3×3
conv3	in_planes×2	1×1 卷积扩展通道数
CBAM	in_planes×2	CBAM 模块不会改变通道数

在表 2 中，DCAB 模块执行主要包括两个流程；输入经过 1×1 卷积降维以此降低计算复杂度，同时压缩和提取特征。之后，使用 3×3 可形变卷积来动态调整卷积核位置，以捕捉到输入图像中的更多局部特征。经过形变卷积后，再通过一个 1×1 卷积来恢复原来的通道数，为了强调可形变卷积中学习到的局部特征，输出再经过 CBAM 中的通道注意力和空间注意力模块，增强图像在通道和空间维度的注意力权重。最后将输入特征图与 CBAM 模块的输出连接，以学习网络中的残差。DCAB 公式如式（11）所示，其中：x 为输入，y 为输出，$w(p_n)\cdot x(p_0+p_n+\Delta p_n)$ 是可形变卷积操作，$M_c(F)$ 是通道注意力生成的权重，$M_s(F')$ 是空间注意力生成的权重：

$$y(p_0) = \sum_{p_n\in R} w(p_n)\cdot x(p_0+p_n+\Delta p_n)\cdot M_c(F)\cdot M_s(F') + x \tag{11}$$

3.2　ResNet50 _ DCAB 模型

ResNet50 是广泛应用于卷积神经网路的模型之一，

其通过残差链接有效地解决了梯度消失和梯度爆炸等问题。然而，尽管 ResNet50 在许多视觉任务上都有出色的表现，但其固定的层次结构和卷积核可能限制了其在复杂特征任务中的泛化能力和特征表达能力。因此为了进一步提升 ResNet50 网络的性能，在 ResNet50 的最后一个卷积层（即残差块 5）后加入改进的 DCAB 模块，加大网络的复杂特征的提取力度，增强 ResNet50 的表达能力[16]。输入特征图在经过连续的残差结构卷积操作后，图像数据会被进行归一化处理，并应用全局自适应平滑池化，这一过程有助于提升神经网络模型的精度以及加速模型的收敛速度。由于在该阶段的特征数据仍然处于多维形式，无法直接输入全连接层，因此必须首先进行一维化处理。最终，经过处理的特征数据将被输入全连接层，并通过 Softmax 分类器生成对应缺陷类别识别结果[17]。ResNet50 _ DCAB 网络结构如表 3 所示。

表 3 ResNet50_DCAB 网络结构

层级	感受野大小	输出通道数	操作说明
输入层	3	224×224	原始输入图像
卷积层 Conv1	64	112×112	7×7,stride=2
最大池化层 MaxPool	64	56×56	3×3,stride=2
Layer1(ResNet)	256	56×56	3 个 Bottleneck 块
Layer2(ResNet)	512	28×28	4 个 Bottleneck 块
Layer3(ResNet)	1 024	14×14	6 个 Bottleneck 块
Layer4(ResNet)	2 048	7×7	3 个 Bottleneck 块
Layer5(DCAB)	4 096	7×7	3 个 Bottleneck 块（DCAB）
全局平均池化	4 096	1×1	自适应平均池化
全连接层	num_classes(6)	1×1	最终分类输出

4 实验结果与分析

4.1 实验环境及参数

本文实验的深度学习模型搭建在 Window11 操作系统下的 Pytorch 2.0.1 深度学习框架上，并且采用 python 3.9 版本语言编程实现。本实验硬件环境中，配备了高性能的 Intel Core i5-13490F 处理器和显存为 16 GB 的 NVIDIA GeForce RTX 4060 Ti GPU，结合大容量的 16 GB 内存和 1 TB SSD 存储，能够为深度学习实验提供良好的支持。强大的并行计算能力与高速存储有助于提升模型训练的效率，缩短实验周期。

CNN 模型的参数配置对模型训练的结果至关重要。本文模型设置批处理大小（Batch _ size）为 64，为使模型充分收敛，设置训练集迭代次数（Epoch）为 100 轮。学习率使用 Adam 优化器，使其在学习过程中加速收敛。本实验使用在 ImageNet 上预训练的 ResNet50 网络模型，以提高网络的收敛速度并缩短训练时间，提高模型训练效率，为避免对预训练权重进行过大的调整，导致模型丢失在大规模数据集上预训练的优势，因此选取初始学习率为 0.000 1。为防止过拟合问题，改善模型的鲁棒性，丢弃比率为 0.2。

4.2 学习率对模型的影响

学习率是深度学习中最重要的超参数之一，它直接影响模型的训练过程和最终性能[18]。学习率决定了模型在每次梯度下降时权重更新的幅度，从而对模型的收敛速度、稳定性以及最终的泛化能力产生重要影响。过小的学习率可能会导致模型无法有效地逃离局部最优点，导致模型训练缓慢，模型表现无法达到最优。而过大的学习率可能导致模型错过最优点，甚至在梯度下降过程中反复跳过最优解，最终无法收敛，导致训练效果很差。一个恰当的学习率能够帮助模型有效利用梯度信息，在较少的训练轮次内达到收敛，提高训练效率。为探究不同学习率对于 ResNet50 _ DCAB 模型的性能影响，在基于迁移学习的基础上分别将学习率设置为 0.001、0.000 1、0.000 01 以及 0.000 001。用上述 4 种不同的初始学习率对增强后的 NEU-DET 数据集进行训练和测试，并设置迭代轮数为 100 轮。表 4 为不同学习率下模型的最终验证集准确率和训练损失。

表 4 不同学习率下验证集准确率及损失

学习率	最终训练损失	最终验证准确率	收敛速度（epoch 数）	平均验证准确率
0.001	1.09	0.956	22	0.919
0.000 1	1.06	0.981	9	0.978
0.000 01	1.09	0.956	4	0.949
0.000 001	1.11	0.944	5	0.942

学习率对模型的训练损失影响较小，但对验证集准确率和泛化能力有显著影响。综合实验结果显示，当学习率为 0.000 1 时，模型的验证集准确率达到 98.1%，同时具有较低的训练损失和良好的收敛特性，展现出最佳的综合性能。相比之下，过高或过低的学习率都会导致模型泛化能力下降，难以在验证集中达到最佳表现。不同学习率下的损失曲线和准确率曲线如图 5 和图 6 所示，进一步直观验证了适中学习率 0.000 1 的优越性，从而为模型的学习率选择提供了有力支持。因此，学习率为 0.000 1 被最终选定为最佳学习率，以实现模型在训练和验证阶段的性能平衡。

4.3 DCAB 消融实验

DCAB 模块是由 DCN 卷积和 CBAM 组合而成，为了验证两个模块是否都对 ResNet50 的识别准确率起到提升作用，因此设计了 DCAB 的消融实验：将以预训练的 ResNet50 网络模型作为基准，将普通卷积（NC，Normal Convolution）、DCN 卷积、普通卷积加 CBAM 模块以及 DCAB 分别作为 layer5 插入到 ResNet50 最后

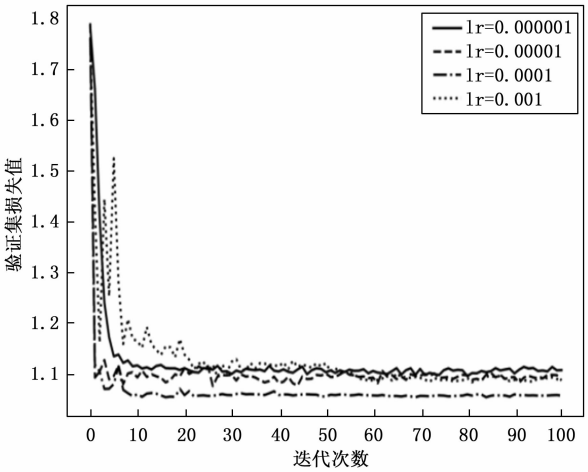


图 5　不同学习率对损失值的影响

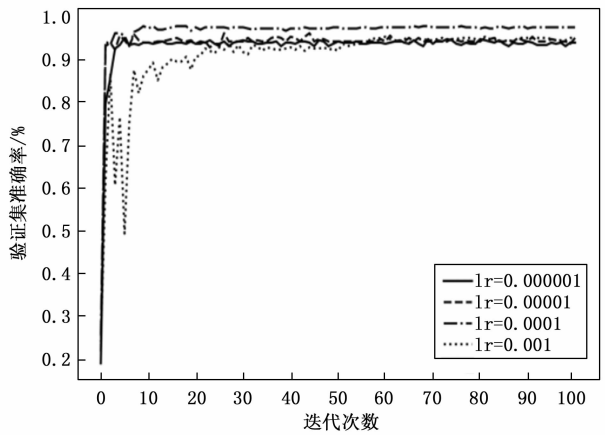


图 6　不同学习率对准确率的影响

一层，并于未添加的 ResNet50 网络对比，研究各模块对网络性能提升的影响。为了确保对比的公平性，训练迭代次数均为 100 轮，学习率为 0.0001，批处理大小为 64，使用 Adam 优化器计算。此次消融实验评价指标为验证集平均准确率、验证集损失、精确率、召回率以及 F1-score。实验结果如表 5 引入不同模块对于 ResNet50 模型性能对比[19]。

表 5　不同模块对于 ResNet50 模型的性能影响

模型	平均准确率	验证集损失	精确率	召回率	F1-score
ResNet50	0.970	0.062	0.984	0.983	0.983
ResNet50＋NC	0.967	1.056	0.986	0.986	0.986
ResNet50＋DCN	0.976	1.057	0.987	0.986	0.986
ResNet50＋NC＋CBAM	0.972	0.120	0.984	0.983	0.983
ResNet50＋DCN＋CBAM	0.978	1.056	0.987	0.986	0.986

从消融实验的结果可以看出，标准 ResNet50 在较简单的任务中已经表现出优异的分类性能。引入普通卷积或可形变卷积虽然能够提高模型的特征表达能力，但在验证集上的损失较高，表明其泛化能力可能有所下降。加入 CBAM 后，无论与普通卷积还是可形变卷积结合，模型的分类性能都有所改善，尤其是在验证集上的表现更加稳定，验证了 CBAM 对增强特征选择的重要性。ResNet50＋DCN＋CBAM 模型在各项指标上表现最优，表明可形变卷积和 CBAM 的结合能够有效提升模型在复杂任务中的性能，同时确保较好的泛化能力。

4.4　DCAB 有效性验证

DCAB 模块由于引入了残差链接，将此模块插入其他主流网络中时也不会因为增加网络层数而引起梯度爆炸或梯度消失的问题，因此 DCAB 模块具有一定的适用性。为了进一步验证 DCAB 模块的有效性，将在 AlexNet，GoogleNet 以及 DenseNet 等主流网络模型中添加 DCAB 模块，并于未添加 DCAB 模块的 AlexNet，GoogleNet 以及 DenseNet 网络对比。通过控制变量法，将各项超参数调整一致，只改变网络模型来观察 DCAB 模块对于各种网络的性能影响。为了全面评估添加与未添加 DCAB 的网络性能，如表 6 所示实验列出了各个模型的初始准确率、最高准确率、平均准确率、参数量以及浮点运算量等参数。

表 6　不同网络添加 DCAB 性能对比

模型	初始准确率	最佳准确率	平均准确率	参数量/10^6	浮点运算量/10^9
AlexNet	0.589	0.897	0.882	57.028	0.710
AlexNet＋DCAB	0.592	0.917	0.895	59.345	0.729
GoogleNet	0.428	0.900	0.873	5.975	1.568
GoogleNet＋DCAB	0.397	0.939	0.903	41.488	1.904
DenseNet	0.761	0.919	0.896	9.850	3.102
DenseNet＋DCAB	0.778	0.914	0.896	45.362	3.437

由表 6 可知，AlexNet 在加入 DCAB 模块后，初始准确率从 0.589 略微上升至 0.592，最佳性能从 0.897 上升至 0.917，平均准确率从 0.882 提升至 0.895。虽然参数量和浮点运算有所增加，但这部分计算资源的增幅相对较小，表明 DCAB 在资源效率和性能提升之间取得了良好的平衡。GoogleNet 加入 DCAB 模块后虽在初始准确率上下降了 0.031，这可能表明 DCAB 模块在训练的初始阶段影响了 GoogleNet 的优化路径。然而，模型的最佳准确率从 0.900 提升至 0.939，平均准确率也从 0.873 提升至 0.903。说明 DCAB 模块通过进一步优化学习过程，提升了网络的最终性能。其参数量和计算量的增加较为明显，但从性能提升来看是值得的。

4.5　ResNet50 _ DCAB 与其他 CNN 对比

为了验证 ResNet50 _ DCAB 改进网络的有效性，将 ResNet50 _ DCAB 网络与一些主流图像分类网络模型 AlexNet、GoogleNet、DenseNet、ResNet34 以及 ResNet50 在金属缺陷数据集上的训练结果进行对

比[20]。训练前将各模型超参数调整一致，迭代次数均为 100 轮，学习率为 0.000 1。表 7 是不同 CNN 网络的性能评估对比，评价指标为初始准确率、最高准确率、参数量以及浮点运算量。

表 7 ResNet50_DCAB 与不同 CNN 网络性能对比

模型	初始准确率	最佳准确率	参数量 /10^6	浮点计算量 /10^9
AlexNet	0.589	0.897	57.028	0.710
GoogleNet	0.428	0.900	5.975	1.568
DenseNet	0.761	0.919	9.850	3.102
ResNet34	0.656	0.911	21.287	3.681
ResNet50	0.953	0.983	23.546	3.908
ResNet50_DCAB	0.972	0.986	113.045	4.536

从表 7 的结果可以看出，除 AlexNet 的最佳验证集准确率为 0.897 外，其余模型的最佳准确率均超过了 0.9，表明在金属缺陷检测任务中，较为先进的深度学习网络普遍具有较好的性能。其中，ResNet50 _ DCAB 表现出了最为出色的识别效果，其最佳验证集准确率达到了 0.986，显著优于其他模型。具体而言，ResNet50 _ DCAB 相较于 AlexNet、GoogleNet、DenseNet、ResNet34 和原始 ResNet50 的验证集最优准确率分别提升了 8.9、8.6、6.7、7.5 和 0.3 个百分点，充分表明了所提出改进网络在识别复杂缺陷特征方面的显著优势。

此外，ResNet50 _ DCAB 的初始准确率为 0.972，相较于其他网络模型的初始准确率也具有明显的领先优势。这个结果说明，ResNet50 _ DCAB 在训练的早期阶段已经能够快速获得较高的性能，这对于加快模型的收敛速度具有重要意义。

虽然添加了 DCAB 模块使得网络的准确率得到了提升，但也使得网络结构变得更加复杂，参数量和计算量均有所增加。这种复杂性导致模型在实际应用中可能需要更多的计算资源和存储空间，因此在特定应用场景中，需要权衡模型性能与计算开销之间的关系。尤其是在嵌入式系统或对计算能力有限的设备上，进一步的优化，如模型压缩、参数剪枝或知识蒸馏，可能是必要的。

5 结束语

根据 ResNet50 网络中的残差链接结构，提出了一种基于可形变卷积的残差链接结构，由于残差链接，该模块的插入不会引起网络的梯度爆炸以及梯度消失问题，并且在残差链接前加入 CBAM 模块，以加强模块对于重要特征的捕捉能力。结合 DCN 和 CBAM 后的模块为 DCAB。

在 NEU-DET 数据集中通过消融实验证明了 DCAB 模块的有效性，通过迁移学习的 ResNet50 网络在验证集上的平均准确率为 0.970 而 ResNet50 _ DCAB 网络的平均准确率提升至了 0.978；DCAB 与其他网络结合也均取得了提升，进一步验证了 DCAB 模块的适用性；最后将 ResNet50 _ DCAB 网络与 AlexNet、GoogleNet、Densenet、ResNet34 以及 ResNet50 相比，最优准确率分别提升 0.089、0.086、0.067、0.075 以及 0.003。

DCAB 模块具有一定的适用性和有效性，但是在引入 DCAB 模块后网络参数随之增加，使网络变得臃肿，在后续工作中需要通过蒸馏剪枝等方法来权衡网络性能和参数之间的关系。

参考文献：

[1] WANG T，CHEN Y，QIAO M，et al. A fast and robust convolutional neural network-based defect detection model in product quality control [J]. The International Journal of Advanced Manufacturing Technology，2018，94：3465－3471.

[2] YASIR S M，AHN H. Faster metallic surface defect detection using deep learning with channel shuffling [J]. Computers，Materials & Continua，2023，75 (1)：1847－1861.

[3] REN R，HUNG T，TAN K C. A generic deep-learning-based approach for automated surface inspection [J]. IEEE transactions on cybernetics，2017，48 (3)：929－940.

[4] 李 明，刘 悦. 基于支持向量机的金属表面缺陷检测算法 [J]. 世界有色金属，2016，(2)：120－121.

[5] 胡 坤，吴国庆，胡祖辉，等. 基于改进的 VGG16 网络金属表面缺陷图像分类研究 [J]. 计算机应用与软件，2024，41 (6)：175－180.

[6] ZHANG J，KANG X，NI H，et al. Surface defect detection of steel strips based on classification priority YOLOv3-dense network [J]. Ironmaking & Steelmaking，2021，48 (5)：547－558.

[7] LV X，DUAN F，JIANG J，et al. Deep metallic surface defect detection：The new benchmark and detection network [J]. Sensors，2020，20 (6)：1562.

[8] 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.

[9] HE K，ZHANG X，REN S，et al. Deep residual learning for image recognition [C] //Proceedings of the IEEE conference on computer vision and pattern recognition. 2016：770－778.

[10] 刘源洞，何茂征，黄益斌，等. 基于 ResNet50 和改进注意力机制的船舶识别模型 [J]. 计算机应用，2024，44 (6)：1935－1941.

[11] DAI J，QI H. Deformable convolutional networks [C] //Proceedings of the IEEE international conference on computer vision. 2017：764－773.

(下转第 78 页)