光电技术及应用

DOI:10.16818/j.issn1001－5868.20240717001

基于混合注意力机制的钢材表面缺陷分类算法

应何杰[1,3]，赖联锋[1,3*]，任学航[1]，熊铃铃[2]，薛章起[1]

（1. 宁德师范学院 机电工程学院，福建 宁德 352100；2. 宁德师范学院 信息工程学院，福建 宁德 352100；3. 宁德师范学院 新能源汽车产业节能技术研究协同创新中心，福建 宁德 352100）

摘 要： 针对钢材缺陷特点及当前分类算法对复杂模型与非启发式注意力机制的依赖，提出一种基于卷积块注意力模块（Convolutional Block Attention Module，CBAM）混合注意力架构的钢材表面缺陷分类算法（Multi-Attention based Classification Method，MACM）。首先，通过灰色关联度分析实现启发式通道注意力，增强算法可解释性并缓解层间信息丢失对通道注意力判定的影响。其次，采用紧凑双线性池化实现空间特征融合，增强抽象特征之间非线性复杂交互信息的捕获。最后，基于CBAM架构实现通道和空间注意力的结合。实验结果表明，MACM算法在实现较好性能的同时相较现有SOTA算法更为轻量，验证了MACM算法在提升分类准确性、降低模型复杂度方面的有效性。

关键词： 钢材缺陷分类；深度学习；注意力机制；灰色关联度分析；紧凑双线性池化

中图分类号： TP391.41 **文章编号：** 1001－5868(2025)01－0090－08

Classification of Steel Surface Defects Based on Multi-Attention Mechanism

YING Hejie[1,3], LAI Lianfeng[1,3], REN Xuehang[1], XIONG Lingling[2], XUE Zhangqi[1]

(1. College of Mechanical and Electrical Engineering Ningde Normal University, Ningde 352100, CHN; 2. College of Information Engineering Ningde Normal University, Ningde 352100, CHN; 3. The Collaborative Innovation Center of Ningde Normal University Ningde Normal University, Ningde 352100, CHN)

Abstract: In view of the characteristics of steel defects and the current classification methods' dependence on complex models and non-heuristic attention mechanisms, this paper proposes a multiattention-based classification method (MACM) for steel surface defect classification based on the CBAM (Convolutional Block Attention Module) hybrid attention architecture. Firstly, heuristic channel attention is achieved through gray correlation analysis to enhance the interpretability of the algorithm and overcome the impact of interlayer information loss on channel attention determination. Secondly, compact bilinear pooling is used to achieve spatial feature fusion, enhancing the capture of nonlinear complex interaction information between abstract features. Finally, the combination of channel and spatial attention is achieved based on the CBAM architecture. The experiments show that the MACM method not only performs well but is also lighter than existing state-of-the-art methods, validating the effectiveness of the MACM algorithm in improving classification accuracy and reducing model complexity.

Key words: classification of steel defects; deep learning; attention mechanism; grey correlation analysis; compact bilinear pooling

收稿日期：2024－07－17.

基金项目：新能源汽车产业节能技术研究协同创新中心项目（2022ZX02）.

* **通信作者：**赖联锋

0　引言

钢铁是工业生产的关键基石，随着各行业的不断发展，对钢材质量的要求也不断提升。特别是在带材的生产制造过程中，不同的制造工艺往往会导致钢材表面出现各种缺陷，进行影响其最终性能[1-3]。因此，为能够在生产过程中针对性优化钢材加工工艺以提升钢材质量，对钢材表面缺陷进行分类检测至关重要。

当前对钢材表面加工缺陷的分类检测方法主要有人工检测、传统图像处理和基于深度卷积神经网络（Deep Neural Networks，DNN）的计算机视觉分类检测。其中人工检测分类依赖于技术人员经验，并且受限于人眼分辨能力，无法有效识别出尺寸较小、外形差异较低的缺陷，因此该方案当前仅在少数特定场合有应用[4]。传统图像处理方案通常基于LBP，SIFT或小波变换等图像特征提取方法手动提取缺陷特征，而后结合SVM、浅层神经网络等方法实现对钢材表面缺陷的分类检测[5-8]。相对人工分类检测而言，传统图像处理方案在一定程度上提高了分类检测效率和准确率，但考虑到钢材表面缺陷类型存在类间差异较小，而类内相对差异较大的特点[9]，手动设置的缺陷特征提取难以全面覆盖所有潜在的缺陷特征，因而当外部环境变化（如光线等）导致钢材表面缺陷成像出现差异时，此类分类算法仍然存在着鲁棒性欠优、泛化性不足的问题[4,10]。

近年来，随着DNN技术的发展，基于DNN的计算机视觉分类检测方案在表面缺陷检测领域得到广泛应用，并表现出逐渐替代人工检测和传统图像处理方案的趋势[11-12]。当前对钢材表面缺陷识别的研究主要从算法优化和处理数据集来展开，其中算法优化的目的是开发或引入新的图像特征提取结构，以增强分类效果。对此，江培营等[1]在Inception-ResnetV2模型基础上嵌入卷积块注意力模块（Convolutional Block Attention Module，CBAM）注意力机制，提出一种基于注意力机制的改进Inception-ResnetV2模型的钢板表面缺陷图像分类方法。曾亮等[2]通过构建多尺度多输入网络并针对性融合不同尺度特征，提出一种基于多方位感知的带钢表面缺陷长尾分类方法。Kursat等[13]提出一种并行训练残差和注意力结构以提取深层特征，并在匹配从邻域分量分析和Relief算法获得的索引的基础上进行筛选，最后使用支持向量机算法进行分类。为增强模型特征捕捉能力，Zhao等[14]将融合注意力模块与并行多尺度模块相结合，实现对重要特征信息的集中和低相关信息的保留，并由此提出一种轻量化分类网络。Liu等[15]提出一种轻量化的并行卷积神经网络结构，利用多尺度图像特征融合以获取更好的分类精度。

针对当前钢材表面缺陷数据集规模较小且构建数据集过程耗时长的问题，提出通过不同的算法实现数据集或算法优化以降低数据集对分类模型的影响。Zhang等[3]提出一种双流（Dual-Stream）神经网络结构，包括样本生成流及分类训练流，结合迁移学习方法以解决当前钢材表面缺陷数据量较为匮乏导致分类精度难以提升的问题。张立等[4]出于相同的考量采用迁移学习方法，基于预训练模型结合Inception-V3结构对钢材表面缺陷分类检测展开研究。Jain等[16]引入生成对抗网络（GANs），训练生成器根据随机噪声合成新的表面缺陷图像并进一步用于分类算法的训练。针对当前优化算法主要优化分类模型自身，Zhao等[17]提出一种SDDA数据增强方法以实现图像进行多尺度变换，所提出的方法针对模型输入数据进行优化，因此只会增加训练时间，而不会增加模型的推理时间。

综上所述，已有大量研究者对钢材的表面缺陷分类进行深入研究和探索。但是，当前基于DNN的计算机视觉分类检测方案在图像特征提取过程中普遍存在依赖于卷积层堆叠，以获取高维度的特征信息的现象，这对分类模型的构建具有正面效果的同时，也对硬件提出了更高的要求，不利于算法的实际部署。虽然，也有研究者尝试通过引入注意力机制以增强有效特征提取，从而达到降低模型深度的目的[1,13-14]，但主要采用的仍是非启发式注意力机制，该方法存在可解释性低的问题。

针对上述问题，本文提出一种基于CBAM混合注意力架构的钢材表面缺陷分类算法（Multi-Attention Based Classification Method，MACM）。该算法首先通过引入灰色关联度分析（Grey Relation Analysis，GRA）实现启发式通道注意力机制，以针对性筛选特征通道维度信息，从而降低亮度、对比度等客观因素对钢材表面缺陷分类造成的影响。在此基础上基于紧凑双线性池化（Compact Bilinear Pooling，CBP）构建空间注意力机制进一步提取特征空间信息，以增强算法对小微缺陷特征的

捕获能力。最后参考CBAM方案结构实现通道注意力和空间注意力的融合。通过利用注意力机制对缺陷特征通道及空间两个方面的筛选,实现对钢材表面缺陷细节特征信息的挖掘,从而提升缺陷的分类精度。同时通过有效的特征提取降低对模型复杂度的要求,提升算法的实际应用价值。

1 相关理论及研究方案

1.1 灰色关联度分析

GRA被广泛用于确定多个变量特征之间的关系,它属于灰色系统理论的一部分。该方法为探究影响事物发展的主要因素,为定量描述事物或因素之间的相互关系提供了理论依据[18]。

其算法如下:给定比较序列$\boldsymbol{X}_1,\boldsymbol{X}_2,\cdots,\boldsymbol{X}_m$与参考序列$\boldsymbol{X}_0$,GRA采用公式(1)描述参考序列曲线与比较序列曲线各点之间的差异:

$$\xi_m(i)=\frac{\min\limits_m\min\limits_i\left|\boldsymbol{x}_0(i)-\boldsymbol{x}_m(i)\right|}{\left|\boldsymbol{x}_0(i)-\boldsymbol{x}_m(i)\right|+\alpha\max\limits_m\max\limits_i\left|\boldsymbol{x}_0(i)-\boldsymbol{x}_m(i)\right|}+\frac{\alpha\max\limits_m\max\limits_i\left|\boldsymbol{x}_0(i)-\boldsymbol{x}_m(i)\right|}{\left|\boldsymbol{x}_0(i)-\boldsymbol{x}_m(i)\right|+\alpha\max\limits_m\max\limits_i\left|\boldsymbol{x}_0(i)-\boldsymbol{x}_m(i)\right|} \tag{1}$$

式中,$\xi_m(i)$表示第m个比较序列第i个观测值$\boldsymbol{x}_m(i)$与对应参考序列值$\boldsymbol{x}_0(i)$的相对差值,又称为$\boldsymbol{X}_m$对$\boldsymbol{X}_0$在i时刻的关联系数。$\alpha\in[0,1]$表示分辨系数,其值常取为0.5。$\min\min\left|\boldsymbol{x}_0(i)-\boldsymbol{x}_m(i)\right|$为两级的最小差,$\max\max\left|\boldsymbol{x}_0(i)-\boldsymbol{x}_m(i)\right|$为两级的最大差。

基于上述计算,每个比较序列都可以对应求出n个关联系数,信息相对分散难以进行有效比较,因此在实际应用中,通常对比较序列的n个关联系数做平均值处理,如公式(2)所示:

$$\boldsymbol{r}_m=\frac{1}{n}\sum_{i=1}^{n}\xi_m(i) \tag{2}$$

式中,$\boldsymbol{r}_m$为所求第m个比较序列$\boldsymbol{X}_m$与观测序列$\boldsymbol{X}_0$的灰色关联度。

结合上述分析可知,与常见的相关系数、主成分分析等关联性计算方法类似,GRA实质上也是通过比较参考序列与比较序列曲线之间几何形状的差别来判定二者之间的相关性[18]。但GRA的优势在于它不需要完整的数据,可以利用部分信息来评估数据序列之间的关联性,考虑到DNN不同层级之间由于卷积、池化等操作会造成信息丢失现象[19-20],GRA的这一特性可以弥补这一现象造成的问题从而能够更准确地分析DNN不同层级内特征通道之间的相关性。

1.2 紧凑双线性池化

双线性池化(Bilinear Pooling, BP)是一种针对多模态数据的池化方法,它主要用于处理包含两种不同类型特征的输入数据,其核心思想是通过双线性映射来捕获不同模态特征之间的交互,从而提取更丰富的特征表示[21]。因此,采用双线性池化有助于多模态特征的提取和融合,从而可以为模型的搭建提供更丰富的特征信息。

BP对图像$\boldsymbol{I}$的特征$\boldsymbol{a}$和$\boldsymbol{b}$在每个位置$\boldsymbol{l}$使用矩阵外积对特征$\boldsymbol{a}_l\in\boldsymbol{R}^M$和$\boldsymbol{b}_l\in\boldsymbol{R}^N$进行组合,即在位置$\boldsymbol{l}$处特征$a_l$和$b_l$的组合得到的双线性特征为:

$$\boldsymbol{\phi}(\boldsymbol{l},\boldsymbol{I})=\boldsymbol{a}_l\boldsymbol{b}_l^T\in\boldsymbol{R}^{M\times N} \tag{3}$$

式中,M,N表示对应特征的通道数。为了最终融合特征$\boldsymbol{a}$和$\boldsymbol{b}$得到图像的双线性池化特征,还要用池化函数聚合图像中所有位置的双线性特征,其中最常采用的是和池化,从而可得到双线性池化特征输出,如公式4所示:

$$\boldsymbol{\phi}(\boldsymbol{I})=\sum\nolimits_{l\in L}\phi(\boldsymbol{l},\boldsymbol{I})=\sum\nolimits_{l\in L}\boldsymbol{a}_l{\boldsymbol{b}_l}^T\in\boldsymbol{R}^{M\times N} \tag{4}$$

记 $\boldsymbol{A}=[a_1,a_2,\ldots,a_L]$, $\boldsymbol{B}=[b_1,b_2,\ldots,b_L]$,则公式(4)可用矩阵表示为:

$$\boldsymbol{\phi}(\boldsymbol{I})=\boldsymbol{A}\boldsymbol{B}^T\in\boldsymbol{R}^{M\times N} \tag{5}$$

结合上述分析可知,通过BP算法所得的双线性特征是高维的($M\times N$),这对含BP算法的模型的训练提出了较高的要求。为此,引入CBP策略,该策略在降维的同时可以得到更优于BP的融合效果[22]。具体而言CBP在BP的融合阶段,即上述公式(4)或公式(5)的阶段采用低维映射取代外积计算,从而达到简化池化过程、降低最终输出特征维度的目的。

BP或CBP相对于当前特征融合所采用的直接线性叠加等手段,更有助于捕获特征之间的非线性复杂交互,具备为模型提供更丰富信息的潜力。

1.3 基于CBAM架构的复合注意力机制

1.3.1 CBAM

CBAM是一种结合通道注意力与空间注意力的混合注意力模块,其结构如图1所示[23]。

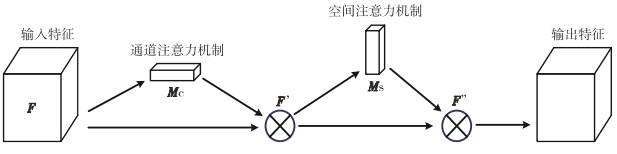


图1 CBAM结构

如图所示，该模块依次对输入特征做通道注意力运算和空间注意力运算。其中通道注意力权重 $\boldsymbol{M}_{\mathrm{C}}$ 的计算式为：

$$\boldsymbol{M}_{\mathrm{C}}(F)=\sigma\left(\mathrm{MLP}\left(MaxPool(F)+AvgPool(F)\right)\right) \tag{6}$$

式中，σ 为 sigmoid 激活函数，MLP 表示多层感知器，F 表示输入特征。这里的最大池化和平均池化是指在特征每个通道内的池化操作。

空间注意力权重 $\boldsymbol{M}_{\mathrm{S}}$ 的求算如公式(7)所示：

$$\boldsymbol{M}_{\mathrm{S}}(F)=\sigma\left(f^{7\times7}\left(\left[AvgPool(F);\ MaxPool(F)\right]\right)\right) \tag{7}$$

式中，$f^{7\times7}$ 表示卷积核大小为 7×7 的卷积，此处最大池化和平均是指沿着特征通道方向的池化操作。

结合公式(6)可知，CBAM 对于特征通道间的关联性采用非启发式机制，此方法依赖于模型训练效果和模型结构设计，并且降低了模型的可解释性。为此使用 GRA 算法实现对通道间关系权重的计算，其作为一种有效的统计工具，适用于处理存在信息缺失或不确定性的情况，通过量化特征序列间的相似性来确定它们之间的关联度[18]。与原有的非启发式方法相比，该方法不仅提高了模型的可解释性，而且减少了分析通道关联对特征信息完整性的依赖。

而对于空间注意力机制，结合公式(7)可知，CBAM 将两个通道池化的结果直接叠加后再以卷积增加该过程的非线性。然而，随着模型深度的增加，深层特征随之更为抽象，简单的相加对特征间的复杂交互的捕获能力有限，并且相加操作不会选择性地强化或抑制某些特征，可能同时混杂了重要和次要信息。而 BP 通过特征的外积来捕获特征间的交互能够揭示更复杂的相互关系，为后续网络提供更丰富的特征信息。因此，如前文所述为避免 BP 可能导致的高计算成本和维度灾难，使用 CBP 来融合两个通道池化所得特征，以更好地提取特征的空间信息。

1.3.2　MACM

结合上述讨论，所提出的基于 CBAM 结构的 MACM 算法其通道注意力方案如图 2 所示。

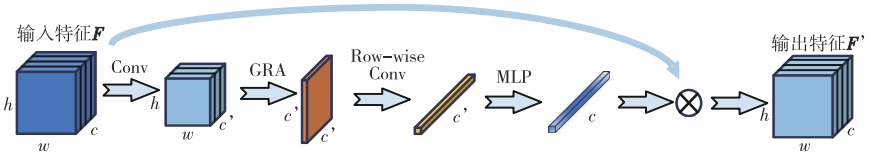


图 2　基于 GRA 的通道注意力

对于输入特征 F 首先采用 1×1 卷积进行通道降维，从 c 降到 c' 以减少计算量。而后采用 GRA 分析特征各通道之间的关联性得到关联性矩阵，该矩阵中的第 i 行第 j 列表示了第 j 个通道与第 i 个通道之间的关联度。再通过行卷积(Row-Wise Conv)与只有一个隐藏层的 MLP 得到维度为 $1\times1\times c$ 的权重向量。最后，将权重以相乘的方式逐通道加权到输入特征 F 得到最后带有注意力权重的输出特征 F'。

所提出的结合 CBP 的空间注意力方案如图 3 所示。对于经过通道注意力加权所得的特征 F'，首先如 CBAM 中一样分别沿通道方向做最大池化和平均池化，这里为了从不同的角度捕获更多的特征信息，采用了分组池化的方法，将原特征 F' 分为两组进行池化，因而所得池化后特征的通道维度 c' 实际等于 2。其次再使用 CBP 对两种池化所得的特征进行融合，并对融合后的特征进行 3×3 卷积操作得到通道维度为 1 的空间权重矩阵。最后同样将权重

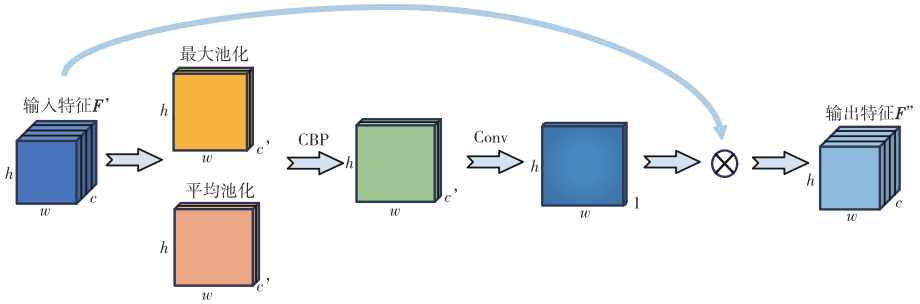


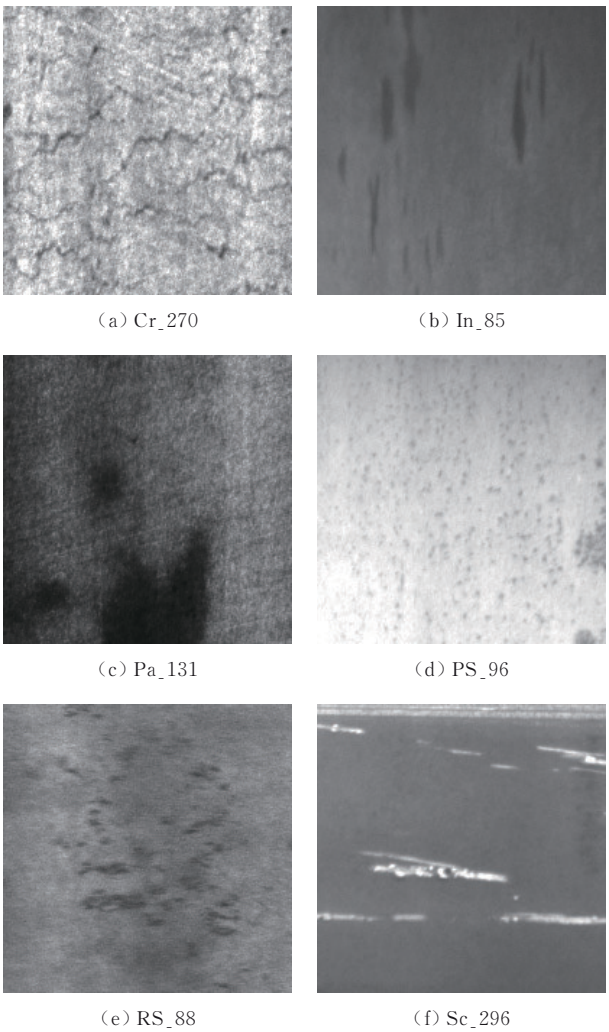
图 3　基于 CBP 的空间注意力

矩阵以相乘的方式逐元素加权到输入特征F'得到最后带有注意力权重的输出特征F''。

2 实验结果与分析

2.1 实验数据集及评价标准

所使用的数据集是由东北大学发布的钢材表面缺陷数据集NEU-DET[9]，其中包含裂纹(Crazing, Cr)、夹杂(Inclusion, In)、斑点(Patches, Pa)、点蚀(Pitted Surface, PS)、铁鳞(Rolled-in Scales, RS)和划痕(Scratches, Sc)共6类钢材表面缺陷类型，每种类型的典型图像如图4所示。每个类别有300张图像，大小均为200×200像素。实验中按照8:1:1的比例将数据集随机划分为训练集、验证集和测试集。



(a) Cr_270 (b) In_85

(c) Pa_131 (d) PS_96

(e) RS_88 (f) Sc_296

图4 6种钢材表面缺陷典型图像

采用精确率(Precision)、召回率(Recall)和准确率(Accuracy)三个常用指标对模型性能进行评价。其计算公式如式(8)~(10)所示：

$$Precision=\frac{TP}{TP+FP} \tag{8}$$

$$Recall=\frac{TP}{TP+FN} \tag{9}$$

$$Accuracy=\frac{TP+TN}{TP+TN+FP+FN} \tag{10}$$

式中，TP为正样本正确分类数，TN为非正样本正确分类数，FP为非正样本误分类数，FN为正样本误分类数。

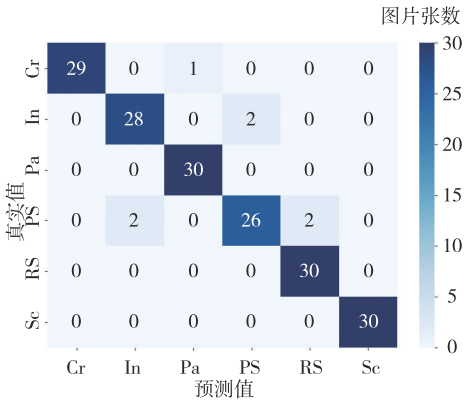
2.2 消融实验

首先采用消融实验验证所提出MACM算法各部分的有效性。实验对象分别如下：(1)为突出所提出算法对钢材表面缺陷细节特征信息的挖掘能力，以resnet18这一相对较浅的网络结构作为基线模型，记为base；(2)在base方案的基础上添加CBAM混合注意力机制，记为base+CBAM；(3)在base+CBAM方案基础上将CBAM中的通道注意力替换为基于GRA的通道注意力，记为base+CBAM(GRA)；(4)在base+CBAM方案基础上将CBAM中的空间注意力替换为基于CBP的空间注意力，记为base+CBAM(CBP)；(5)在base+CBAM方案基础上将CBAM中的通道注意力和空间注意力同时替换为基于GRA和基于CBP的注意力机制，记为base+CBAM(MACM)。

图5所示为在测试集上进行的消融实验中，6种缺陷检测方案的混淆矩阵。混淆矩阵对角线数据表明，对于Cr，Pa和RS三类缺陷各类算法方案都展现出相近的分类性能，而对于In，PS和Sc三类缺陷不同算法却表现出一定的差异。结合图4所展示的6种钢材表面缺陷典型图像可知，Cr，Pa和RS相对于In，PS和Sc而言，前三类缺陷与背景存在较明显差异且缺陷自身尺寸较大，因此对这几种缺陷的分类效果一定程度上体现了不同算法对小微缺陷特征的提取挖掘能力。混淆矩阵数据表明：相对于base方案，引入注意力机制可以有效提升算法模型对缺陷小微特征的提取能力，使用基于GRA与CBP的注意力机制后其小微特征提取能力可得到更进一步提升。

消融实验各模型在测试集上的精确率、召回率指标结果如表1和表2所示。可见在融入CBAM注意力机制后，相对于base方案对不同缺陷类型的精确率和召回率都得到了一定的提升，其中对于In缺

陷的精确率和对 Sc 缺陷召回率的降低结合图 5 不难看出是 Sc 缺陷误判为 In 缺陷导致的。该现象一方面是 In 与 Sc 缺陷二者的形状、尺寸等特征相近导致算法的误判；另一方面是引入 CBAM 注意力机制后，注意力权重的叠加在增加对主要特征的关注的同时一定程度上也削弱了 In 与 Sc 缺陷抽象特征的差异，从而令算法一定程度上混淆了这两个缺陷。但总体而言引入注意力机制之后有效降低了对 Pa，RS 缺陷的误判，同时增加了 In 缺陷 TP 的数量，整体相对 base 算法起到了提升的效果。



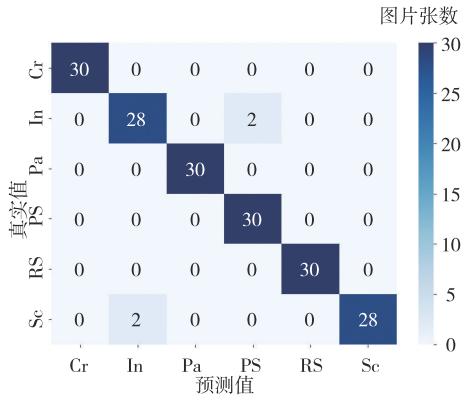
（a）base



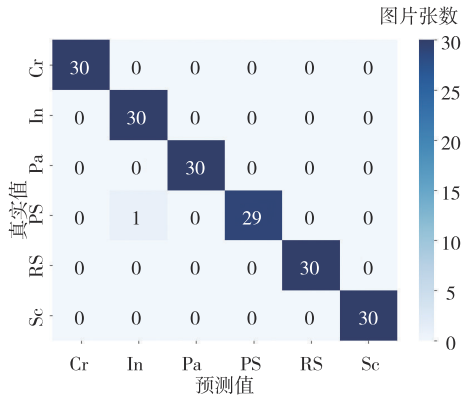
（b）base＋CBAM



（c）base＋CBAM（GRA）



（d）base＋CBAM（CBP）



（e）base＋CBAM（MACM）

图 5　不同方案在测试集上的混淆矩阵

表 1　消融实验不同方案精确率对比结果

算法	缺陷类型					
	Cr	In	Pa	PS	RS	Sc
base	1.00	0.933 3	0.967 7	0.928 6	0.937 5	1.00
base＋CBAM	1.00	0.857 1	1.00	1.00	1.00	1.00
base＋CBAM（GRA）	1.00	0.935 5	1.00	0.965 5	1.00	1.00
base＋CBAM（CBP）	1.00	0.933 3	1.00	0.937 5	1.00	1.00
base＋CBAM（MAB）	1.00	0.967 7	1.00	1.00	1.00	1.00

表 2　消融实验不同方案召回率对比结果

算法	缺陷类型					
	Cr	In	Pa	PS	RS	Sc
base	0.966 7	0.933 3	1.00	0.866 7	1.00	1.00
base＋CBAM	1.00	1.00	1.00	0.933 3	1.00	0.90
base＋CBAM（GRA）	1.00	0.966 7	1.00	0.933 3	1.00	1.00
base＋CBAM（CBP）	1.00	0.933 3	1.00	1.00	1.00	0.933 3
base＋CBAM（MAB）	1.00	1.00	1.00	0.966 7	1.00	1.00

在base+CBAM方案的基础上引入基于GRA或基于CBP的注意力机制后，结合表1、表2和图5可知，引入基于GRA的通道注意力（base+CBAM（GRA）方案）通过对通道特征的加权，强化了亮度、对比度等通道特征从而提升了对In和Sc缺陷的分类能力。引入基于CBP的空间注意力（base+CBAM（CBP）方案）增强了空间特征的获取从而提升了算法对In和Ps缺陷的分类能力。基于CBAM结构对两种算法的结合（base+CBAM（MAB）方案）实现了上述效果的有效结合，使得所提出的MACM算法对Ps和Sc缺陷的检测都得到了提升。

表3通过准确率展现了不同方案的整体性能比较。数据显示引入CBAM注意力机制后算法缺陷分类准确率得到提升，结合提出的基于GRA的通道注意力和基于CBP的空间注意力机制可进一步提升算法性能。基于CBAM结构对两者的结合令算法对钢材表面缺陷的分类能力达到最优。

表3 消融实验不同方案准确率对比结果

算法	准确率/%
base	96.11
base+CBAM	97.22
base+CBAM(GRA)	98.33
base+CBAM(CBP)	97.77
base+CBAM(MAB)	99.44

通过上述消融实验，验证了所提出的基于GRA的通道注意力和基于CBP的空间注意力算法的有效性及鲁棒性，证明MACM算法对钢材表面缺陷细节特征信息的挖掘和提取能力。同时，所提出算法在较浅层神经网络上的性能表现，也验证了该算法在挖掘有效细节特征信息以提升算法性能的能力，以及其在降低模型复杂度、提升算法应用潜力的贡献。

2.3 对比分析

为进一步评估所提出的MACM算法，将其与现阶段同样基于NEU-DET数据集的SOTA分类算法进行对比，对比结果如表4所示。

表4 对比结果

方法	准确率/%
文献[15]	98.89
文献[3]	99.70
文献[4]	99.80
文献[14]	99.60
文献[5]	99.10
本文	99.44

由表4可知，所提出的MACM算法相对于文献[3]、[4]、[14]而言准确率分别低0.26%，0.36%，0.16%。然而，文献[4]以基于ImageNet的预训练模型结合改进AlexNet构建算法架构，相较所提出的基于resnet18的MACM方案模型更为复杂，对硬件友好性较差。同样，文献[3]中采用双流神经网络结构；文献[4]中使用并行残差-注意力网络结构并以支持向量机作为分类算法。两者相较所提出的MACM算法其算法模型结构更为复杂，但其整体效果提升却有限（分类准确率差异小于0.4%）。因此，相较于当前的SOTA算法，所提出算法能够在更简单的模型中表现出较好的性能，同时对硬件要求更为友好，具有较好的实际应用潜力。

3 结论

针对钢材表面缺陷特点及当前主流算法对复杂模型和非启发式注意力机制的依赖。本文提出一种基于混合注意力架构的钢材表面缺陷分类算法。实验显示，所提出的MACM算法中基于GRA的通道注意力机制和基于CBP的空间注意力机制单独或同时应用均可实现模型性能的提升，验证了所提出算法的鲁棒性和有效性。同时与当前SOTA算法相比MCAM算法在较浅层神经网络（resnet18）上也表现出较好的性能水平，进一步验证了MACM算法的有效性。

实验结果较好验证了所提出算法的性能，但仍存在一些问题有待探索。如当前所用的NEU-DET数据集体量较小，算法是否具有更为广泛的适用性仍需要进一步验证。

参考文献：

[1] 江培营，陶青川，艾梦琴. 基于注意力机制和深度学习的钢板表面缺陷图像分类[J]. 计算机应用与软件，2021，38(9)：214-219.
Jiang Peiying，Tao Qingchuan，Ai Mengqin. Steel surface defect image classification based on attention mechanism and deep learning[J]. Computer Applications and Software，2021，38(9)：214-219.

[2] 曾 亮，胡 谦，吴 霓，等. 基于多方位感知的带钢表面缺陷长尾分类方法[J]. 仪表技术与传感器，2023(11)：116-120，126.
Zeng Liang，Hu Qian，Wu Ni，et al. Strip steel surface defect long-tail classification method based on multi-directional perception[J]. Instrument Technique and Sensor，2023(11)：116-120，126.

[3] Zhang J, Li S, Yan Y, et al. Surface defect classification of steel strip with few samples based on dual-stream neural network[J]. Steel Research International, 2022, 93(5): 2100554.

[4] 张　立, 王桂棠, 陈建强, 等. 迁移学习在热轧钢带表面缺陷分类中应用研究[J]. 机械设计与制造, 2022(5): 220-224.
Zhang Li, Wang Guitang, Chen Jianqiang, et al. Study on the classification of surface defects in hot-rolled steel strips by transfer learning[J]. Machinery Design & Manufacture, 2022(5): 220-224.

[5] 王官宗, 朱建鸿. 基于ILLBP和ISSA的带钢表面缺陷识别[J]. 钢铁研究学报, 2023, 35(7): 896-906.
Wang Guanzong, Zhu Jianhong. Recognition of strip surface defects based on ILLBP and ISSA[J]. Journal of Iron and Steel Research, 2023, 35(7): 896-906.

[6] Samsudin S S, Arof H, Harun S W, et al. Steel surface defect classification using multi-resolution empirical mode decomposition and LBP [J]. Measurement Science and Technology, 2021, 32(1): 015601.

[7] 朱晓珺, 韩　林, 邹香玲. 基于SIFT特征与多层BP神经网络的钢板缺陷检测算法[J]. 组合机床与自动化加工技术, 2017(10): 54-56, 61.
Zhu Xiaojun, Han Lin, Zou Xiangling. Steel plate defection detection algorithm based on SIFT feature and multi-layer BP neural network [J]. Modular Machine Tool & Automatic Manufacturing Technique, 2017(10): 54-56, 61.

[8] Zhang Z F, Liu W, Ostrosi E, et al. Steel strip surface inspection through the combination of feature selection and multiclass classifiers[J]. Engineering Computations, 2021, 38(4): 1831-1850.

[9] Bao Y, Song K, Liu J, et al. Triplet-graph reasoning network for few-shot metal generic surface defect segmentation [J]. IEEE Transactions on Instrumentation and Measurement, 2021, 70: 5011111.

[10] 王　杰, 潘　凤, 张艳莎, 等. 融合带权非局部模块的铝型材表面缺陷分类[J]. 计算机与现代化, 2023(5): 86-92.
Wang Jie, Pan Feng, Zhang Yansha, et al. Surface defect classification of aluminum profiles with weighted non-local modules[J]. Computer and Modernization, 2023(5): 86-92.

[11] Tang B, Chen L, Sun W, et al. Review of surface defect detection of steel products based on machine vision[J]. IET Image Processing, 2023, 17(2): 303-322.

[12] 伍　麟, 郝鸿宇, 宋　友. 基于计算机视觉的工业金属表面缺陷检测综述[J]. 自动化学报, 2024, 50(7): 1261-1283.
Wu Lin, Hao Hongyu, Song You. A review of metal surface defect detection based on computer vision [J]. Acta Automatica Sinica, 2024, 50(7): 1261-1283.

[13] Demir K, Ay M, Cavas M, et al. Automated steel surface defect detection and classification using a new deep learning-based approach [J]. Neural Computing and Applications, 2023, 35(11): 8389-8406.

[14] Zhao Y, Sun X, Yang J. Automatic recognition of surface defects of hot rolled strip steel based on deep parallel attention convolution neural network [J]. Materials Letters, 2023, 353: 135313.

[15] Liu Y, Yuan Y, Balta C, et al. A light-weight deep-learning model with multi-scale features for steel surface defect classification[J]. Materials, 2020, 13(20): 4629.

[16] Jain S, Seth G, Paruthi A, et al. Synthetic data augmentation for surface defect detection and classification using deep learning[J]. Journal of Intelligent Manufacturing, 2022, 33(4): 1007-1020.

[17] Feng X, Luo L, Gao X. SDDA: A method of surface defect data augmentation of hot-rolled strip steel [J]. Journal of Electronic Imaging, 2022, 31(3): 033002.

[18] 邓聚龙. 灰色系统基本方法: 汉英对照[M]. 武汉: 华中科技大学出版社, 2005.
Deng Julong. The Primary Mothods of Grey System Theory [M]. Wuhan: Huazhong University of Science & Technology Press, 2005.

[19] 周秀文. 灰色关联度的研究与应用[D]. 长春: 吉林大学, 2007.
Zhou Xiuwen. The study on the grey relational degree and its application[D]. Changchun: Jilin University, 2007.

[20] Sun Y, Xia C, Gao X, et al. Aggregating dense and attentional multi-scale feature network for salient object detection[J]. Digital Signal Processing, 2022, 130: 103747.

[21] Lin T Y, RoyChowdhury A, Maji S. Bilinear CNN models for fine-grained visual recognition [C]// 2015 IEEE International Conference on Computer Vision (ICCV). IEEE, 2015: 1449-1457.

[22] Gao Y, Beijbom O, Zhang N, et al. Compact bilinear pooling [C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2016: 317-326.

[23] Woo S, Park J, Lee J Y, et al. CBAM: Convolutional block attention module [C]// Computer Vision - ECCV 2018. Cham: Springer International Publishing, 2018: 3-19.

作者简介：

应何杰(1993—)，男，福建省宁德市人，助教，中级工程师，主要从事深度学习、计算机视觉工业应用研究；

赖联锋(1982—)，男，福建省宁德市人，博士，教授，硕导，主要从事摩擦学及深度学习领域研究。