

DOI: 10.11830/ISSN.1000-5013.202403033



多特征融合的焊缝图像多标签分类算法

牛顿¹, 林宁², 林振超², 黄凯², 王合佳¹, 郑力新¹

(1. 华侨大学 工学院, 福建 泉州 362021;

2. 福建省特种设备检验研究院, 福建 泉州 362021)

摘要: 为了实现焊缝缺陷的准确分类, 提出一种多特征融合的焊缝图像多标签分类算法。首先, 通过残差神经网络(ResNet-50)提取图像的特征信息, 根据得到的特征图构建图结构, 提出关联度引导邻域传播(RDGNP)算法优化图结构; 然后, 使用图卷积神经网络(GCN)提取图结构的特征信息, 并设计特征融合模块实现图像特征和图结构特征的结合; 最后, 得到多标签分类结果。实验结果表明: 文中算法能够有效地实现焊缝缺陷的多标签分类, 在 X 射线焊缝缺陷数据集上的性能有明显提升。

关键词: 多标签分类; 全局相关性; 图像特征; 图结构特征; 特征融合

中图分类号: TP 391.41; TU 229 **文献标志码:** A **文章编号:** 1000-5013(2024)04-0514-10

Weld Image Multi-Label Classification Algorithm Based on Multi-Feature Fusion

NIU Dun¹, LIN Ning², LIN Zhenchao²,
HUANG Kai², WANG Hejia¹, ZHENG Lixin¹

(1. College of Engineering, Huaqiao University, Quanzhou 362021, China;

2. Fujian Special Equipment Inspection and Research Institute, Quanzhou 362021, China)

Abstract: In order to achieve accurate classification of welding defects, a weld image multi-label classification algorithm based on multi-feature fusion is proposed. Firstly, feature information of images is extracted by a residual neural network (ResNet-50), and the graph structure is constructed based on the obtained feature maps. An algorithm named relation degree guided neighborhood propagation (RDGNP) is proposed to optimize the graph structure. Then, the feature information of the graph structure is extracted using graph convolutional neural network (GCN), and a feature fusion module is designed to achieve the combination of image features and graph structure features. Finally, multi-label classification results are obtained. Experimental results show that the proposed method can effectively realize the multi-label classification of welding defects, and its performance on the X-ray welding defects dataset is significantly improved.

Keywords: multi-label classification; global relevance; image feature; graph structure feature; feature fusion

随着现代工业的迅速发展, 工业焊接在高压、高负荷环境下的广泛应用对焊件的质量提出严格要求。在焊接过程中, 焊件质量容易受到环境和人为因素等多方面影响, 产生裂纹、气孔、夹渣、未熔合和未焊透等缺陷^[1]。为了避免工程事故的发生, 必须对焊件进行全面的缺陷检测与分类。

传统的人工观察焊缝图像方法一般由专业检验员执行, 然而这种方法存在多种主观因素, 效率较

收稿日期: 2024-03-20

通信作者: 郑力新(1967-), 男, 教授, 博士, 主要从事图像分析、机器视觉、深度学习方法的研究。E-mail: zlx@hqu.edu.cn.

基金项目: 福建省科技计划项目(2020Y0039); 福建省泉州市科技计划项目(2020C042R)

低, 检验过程中容易出现漏检和误检等问题^[2]。近年来, 随着大数据的普及和计算能力的提升, 深度学习算法迅速崛起, 以深度学习为基础的焊缝图像处理方法被广泛应用。Say 等^[3]利用数据增强技术对公共数据集 GDXray 进行数据扩增, 通过卷积神经网络(CNN)模型识别出 6 种焊缝缺陷类别, 取得了 92.0% 的准确率。Kumaresan 等^[4]通过 VGG-16 模型实现焊缝缺陷分类, 平均准确率达到 90.0%。Totino 等^[5]基于迁移学习方法, 使用 SqueezeNet 模型实现 4 种焊缝缺陷的分类任务, 准确率达到 93.3%。经典的卷积神经网络采用局部感受野和权值共享, 以分层的方法提取局部特征^[6]进行图像特征表示, 但是该方法仅针对图像的局部信息进行分析, 未能全面考虑到图像的全局信息。由于焊缝区域可能存在环境复杂和缺陷形状各异、分布随机等情况, 要求网络从全局综合分析并处理图像的特征信息, 而卷积神经网络从图像中提取的局部特征会丢失其全局信息, 使卷积神经网络难以提升应用于多标签焊缝图像的分类准确率。

Bruna 等^[7]提出图神经网络(GNN), 通过向神经网络中引入图结构, 能够很好地获取特征图的全局结构信息, 弥补神经网络对图像全局相关性的不敏感。Kipf 等^[8]提出一种创新的图卷积神经网络(GCN), 将卷积神经网络扩展到图结构上。Wang 等^[9]构建一种图引导卷积神经网络(GCNN), 利用训练样本之间的相似度构建图结构, 结果表明, GCNN 可以有效地提高 VGG-11 的特征提取能力, 解决基础模型在分析类间相似性和类内差异性时的不足, 使其具有更优的分类效果。Balcioglu 等^[10]使用 GCN 进行金属螺母零件表面缺陷的分类任务, 其准确率相较于 2D-CNN 提高了 2.955 4%。

目前, 图神经网络大多数应用于表面缺陷的分类, 表面缺陷往往具有较大的尺寸, 且形状较为简单直观, 而内部缺陷通常具有尺寸微小、结构复杂等特点, 利用 GCN 实现焊缝内部缺陷分类任务不足以充分理解图像的细节特征信息, 因此, 对焊缝内部缺陷的分类研究仍然较少。综合考虑当前焊缝缺陷分类算法的不足和图像分类算法的发展, 结合分析焊缝图像的局部关键信息和全局结构信息, 进行焊缝缺陷的多标签分类任务^[11], 从而代替人工实现高效焊件分类工作, 本文提出一种多特征融合的焊缝图像多标签分类算法。

1 多特征融合的焊缝图像多标签分类算法

多特征融合的焊缝图像多标签分类算法框架, 如图 1 所示。该框架主要包括图像特征提取模块、图结构重构模块、图结构特征提取模块和特征融合模块 4 个部分。首先, 在残差神经网络(ResNet-50)中引入 Res2Net 模块以提取焊缝图像的多尺度特征信息; 其次, 根据焊缝图像的特征图构建初始的图结构, 并提出关联度引导邻域传播(RDGNP)算法, 进一步细致优化图结构; 接着, 通过 GCN 提取图结构的特征信息; 最后, 将图像特征与图结构特征输入特征融合模块进行结合, 经过多标签分类器得到焊缝图像所含缺陷类别。

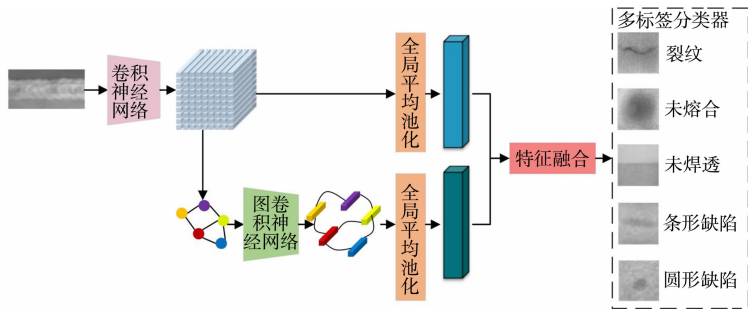


图 1 多特征融合的焊缝图像多标签分类算法框架

Fig. 1 Framework of multi-label classification algorithm for welding images based on multi-feature fusion

1.1 图像特征提取模块

焊缝缺陷往往存在尺寸微小的情况, 导致卷积神经网络无法很好地提取缺陷区域的特征信息。为了丰富焊缝缺陷的语义信息, 提高图结构表达特征之间相似性信息的准确性, 采用 ResNet-50 作为基础模型用于焊缝图像的特征提取。同时, 在 ResNet-50 的每一个 BottleNeck Block^[12]中引入 Res2Net 模块^[13], 从粒度级别提取焊缝缺陷的多尺度特征, 得到多层次的焊缝缺陷特征组合, 增大每个网络层的感

受野,进一步增强网络对焊缝缺陷的特征表达能力,以满足网络对焊缝图像感兴趣区域的关注需求。图像特征提取模块的结构,如图 2 所示。

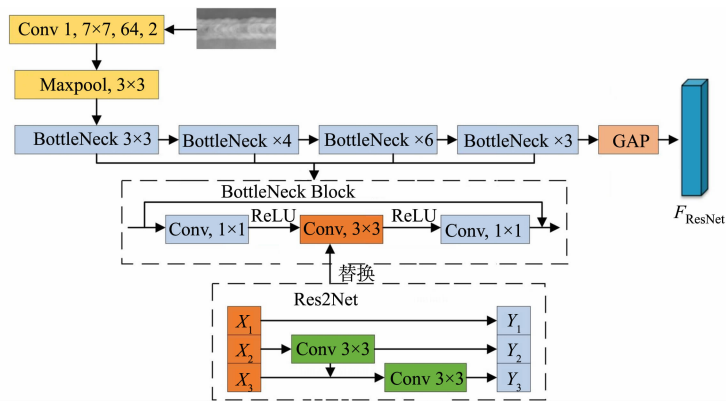


图 2 图像特征提取模块结构图

Fig. 2 Structure diagram of image feature extraction module

在 Bottleneck Block 结构中,使用 Res2Net 模块替换中间的 3×3 卷积,具体地,经过 1×1 卷积后,将输入的特征图划分为 T 个子集,记为 $X_t, t=1, 2, \dots, T$,各子集尺寸大小相同,通道数为输入特征图通道数的 $1/t$ 。在每一个子集中,除第一个通道外,每一个通道都包含前一个通道的特征信息;接下来,将 T 个子集输出的特征图进行拼接,然后进行 1×1 卷积,得到多层次的焊缝缺陷特征组合,这样能够产生具有多种尺度和不同感受野的特征,从而改善模型提取焊缝缺陷特征能力的不足。

为了进行 ResNet-50 的预训练模型迁移,进一步优化模型的学习能力,在原模型基础上,去除用于控制图像维度的平均池化层及用于分类的全连接层,可以得到焊缝图像的特征提取器。将焊缝图像作为图像特征提取器的输入,可得到焊缝图像的特征图为

$$F=f_{\text{ResNet}}\left(I ; \theta_{\text{ResNet}}\right) \in \mathbf{R}^{H \times W \times C} . \quad(1)$$

式(1)中: F 为包含 Res2Net 的 ResNet-50 提取的特征图; f_{ResNet} 为焊缝图像的特征提取器,即包含 Res2Net 的 ResNet-50 模型; I 为输入的焊缝图像; θ_{ResNet} 为图像特征提取器的参数; H, W, C 分别为特征图的高度、宽度和通道数。

由图 2 可知:在图像特征提取模块的末端,引入一个全局平均池化(GAP)层,用于处理最后一个 Bottleneck Block 输出的特征图,最终提取到图像特征 $F_{\text{CNN}} \in \mathbf{R}^C$ 。

1.2 图结构重构模块

1.2.1 图结构初构模块 虽然图像特征提取模块利用包含 Res2Net 的 ResNet-50 提取焊缝图像的局部关键特征,但是未能学习到其中的全局结构特征,模型无法很好地学习和处理形状各异、分布随机的焊缝缺陷。图结构能够有效地描述数据间的复杂关系,对焊缝图像的特征图进行特征重构,通过构建图结构连接邻居特征以增强单个特征,从而更好地捕捉图像中的上下文信息。焊缝图像作为图像特征提取模块的输入,经过提取多尺度特征得到特征图后,将焊缝图像的每个特征 $f_i \in \mathbf{R}^C$ 视为一组无序的节点 $V=\left\{v_1, v_2, \cdots, v_i, \cdots, v_n\right\}(i=1, 2, \cdots, n)$ 进行图结构的构建。为了更加快速地构建焊缝图像的初始图,距离度量方法采用欧式距离函数,引入 K 近邻(KNN)算法构建 KNN 图作为表示特征之间相似性关系的图结构。

为了提高模型的推理速度,从特征图中划分出一个子集进行图结构的构建。设子集特征个数为 S ,那么,子集的特征图可以重写为 $F_s=\left[f_1, f_2, \cdots, f_s, \cdots, f_S\right]$,其中, $s=1, 2, \cdots, S, f_s \in \mathbf{R}^C$ 为所选子集的单个特征。在子集包含的特征样本中,通过 KNN 算法为每一个节点 v_s 找到与其最近邻的 k 个节点,即 v_s 的 K 近邻节点 $N_k\left(v_s\right)$,并在 v_s 与其邻居节点 v_m 之间添加一条边 e_s^m ,则构建出子集的 KNN 图 $g_s=\left(v_s, \epsilon_s\right)$,其中, v_s 为子集中每个特征 f_s 组成的节点集合, ϵ_s 为子集中所有节点之间的边集合。

搜索特征图的所有剩余特征样本来完善子集的 KNN 图,从而构建初始图,适应焊缝缺陷的复杂结构,并降低计算复杂度。对于子集中的每一个节点 v_s ,引入爬山(HC)算法^[14]搜索除子集外的所有剩余特征样本,分别与 v_n 进行比较,并将比较后的特征样本存储在 2 个优先级队列 Q 和 R 中。在每次迭代

中, Q 中存储 v_s 的前 k 个最近邻居节点, 搜索到的邻居节点根据与 v_s 的距离按升序排名, 不断更新距离更近的邻居节点, 并保持最近邻居节点的个数为 k 。 R 中存储 v_s 的最近邻居节点个数不固定, 但是优先级顺序与 Q 相同。 R 除了用于存储最近邻居节点, 也存储最近邻居节点的 K 距离邻域。

由于 HC 算法只是考虑当前符合最近邻居条件的节点, 容易陷入局部最优, 为了保证算法更加全面地选择特征样本, 使用反向 K 近邻(RKNN)算法^[15], 保证选择的特征样本与原特征样本具有良好的相似性, 避免选择相异的特征样本。通过搜索 RKNN 图 \bar{G} 找到与 v_s 距离最近的 r_0 , 并根据 RKNN 图分析 r_0 的近邻关系, 从中选择出与 v_s 最相似的邻居节点, 并将其加入到 Q 和 R 中, 进行最近邻居节点更新。接着, 将 KNN 图 g_s 加入到 Q 中进行最近邻居节点更新。当 Q 不再更新或者 R 已经为空时, 迭代结束, 得到 v_i 的 K 近邻节点 $N_k(v_i)$ 。由此构建焊缝图像的初始图 $g=(V, \epsilon)$, 其中, ϵ 为初始图中所有节点之间的边集合。

1.2.2 图结构优化模块 在采用 HC 算法进行搜索的过程中, 仅仅考虑部分足够接近节点 v_s 的邻居节点 r_0 的邻域用于更新 v_s 的最近邻居。然而, 在 R 中存储的所有最近邻居节点的邻域中, 可能存在一些节点在搜索期间未与 v_s 进行比较, 导致遗漏访问特征样本的问题。为此, 选择利用 R 中每一个最近邻居节点 r 的邻域进一步更新 v_s 的最近邻居节点, 以更加充分地完善图结构。

提出 RDGNP 算法, 通过描述节点之间的关联度(RD)^[16], 引导邻域的传播, 其表达式为

$$\text{Rel}(O, P) = \frac{\overline{O_P} + \overline{P_O}}{k} \quad (2)$$

式(2)中: $\text{Rel}(\cdot)$ 表示关联度函数; $\overline{O_P}$ 表示节点 O 在节点 P 的按升序排名的邻居节点集合中的排名位置; $\overline{P_O}$ 表示节点 P 在节点 O 的按升序排名的邻居节点集合中的排名位置。 $\text{Rel}(\cdot)$ 的值越小, 节点之间的关联度越高, 则说明二者具有相似性关系的可能性越大。且当 $\text{Rel}(O, P) \geq 1$ 时, O 与 P 的关联度很低, 则特征样本之间不具有相似性。

在每次迭代中, 将待传播的最近邻居节点 r 加入到工作队列 W 中, 若其邻居节点满足条件 $\text{Rel}(v_s, p) < 1$ 。这说明 p 与 v_s 具有一定相似性, 故将其也加入到工作队列 W 中, 即实现邻域的传播。不断重复上述步骤, 直到工作队列为空, 即没有新的邻居节点可用于传播。通过迭代地传播邻居信息, 不断提升 KNN 图的质量, 从而提高特征之间相似性关系的描述准确性。利用关联度反映节点之间的相似性信息, 算法更加针对性地选择具有相似性的特征样本进行传播。最终构建焊缝图像的图结构 $G=(V, E, \mathbf{A})$, 其中, E 为图结构中所有节点之间的边集合; \mathbf{A} 为图结构的邻接矩阵, $\mathbf{A} \in \mathbf{R}^{n \times n}$, 若 $v_i \in N_k(v_j)$ 或 $v_j \in N_k(v_i)$, 则 $A_{i,j}=1$, 否则 $A_{i,j}=0$ 。

1.3 图结构特征提取模块

为了学习图结构中各节点之间的邻居信息, 以挖掘不同复杂特征之间蕴含的上下文关系, 解决焊缝图像中缺陷形状各异、分布随机的问题, 采用 GCN 作为图结构的特征提取器, 用于焊缝图像的全局结构特征的提取。图结构特征提取模块结构, 如图 3 所示。GCN 通过聚合邻居节点的特征信息来实现上下文信息的传递, 由于构建的图结构中每一个节点均代表单个特征样本, 其邻接矩阵不能充分准确地表达该图结构, 因此, 采用邻接矩阵和节点矩阵共同进行表达。然而, 将邻接矩阵和节点矩阵相乘后, 矩阵会丢失节点自身的特征信息, 故对模型中各节点进行自连接处理, 包含自连接的邻接矩阵表示为

$$\tilde{\mathbf{A}} = \mathbf{A} + \mathbf{I}_n \quad (3)$$

式(3)中: $\tilde{\mathbf{A}}$ 表示包含自连接的邻接矩阵; \mathbf{I}_n 表示单位矩阵。

信息传递的过程需要保证特征的分布状态不受影响, 故进行归一化操作。引入 $\tilde{\mathbf{A}}$ 的度矩阵 $\tilde{\mathbf{D}}$, $\tilde{\mathbf{D}}$ 对角上的每一个元素的计算公式为

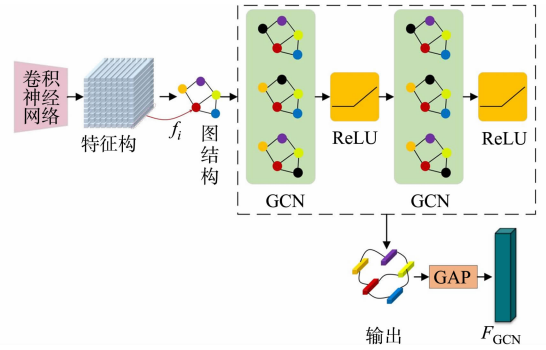


图 3 图结构特征提取模块结构图
Fig. 3 Structure diagram of feature extraction module for graph structure

$$\widetilde{\boldsymbol{D}}_{i,i} = \sum_{j=1}^n \widetilde{\boldsymbol{A}}_{i,j} \circ \quad (4)$$

则 GCN 的第 l 层的传播方程为

$$h^{(l+1)} = \sigma(\widetilde{\boldsymbol{D}}^{-\frac{1}{2}} \widetilde{\boldsymbol{A}} \widetilde{\boldsymbol{D}}^{-\frac{1}{2}} h^{(l)} \omega^{(l)}). \quad (5)$$

式(5)中: $h^{(l)}$ 和 $h^{(l+1)}$ 分别为 GCN 第 l 层的输入和输出,且 $h^{(0)}=V$; $\omega^{(l)}$ 为 GCN 训练的权重参数; $\sigma(\cdot)$ 为 GCN 的非线性激活函数,为了减少模型的计算量,使用 ReLU 函数。

同样地,在图结构特征提取模块的末端,引入 GAP 层用于处理 GCN 输出的特征图,聚合所有节点特征,最终提取到图结构特征 $F_{\text{ResNet}} \in \mathbf{R}^C$ 。

1.4 特征融合模块

特征融合模块通过捷径分支操作将图像特征提取模块输出的特征图和图结构特征提取模块输出的特征图进行 Concat 融合。

传统 GCN 模型最终的激活函数采用 Softmax 函数,主要适用于单标签分类问题。为了实现多标签图像分类任务,将最终的激活函数替换为 Sigmoid 函数。该函数的输出向量的各分量之间没有总和约束,直接表达了各类别本身的可能性。随着激活函数的更改,损失函数相应替换为二元交叉熵损失函数,其数学表达式为

$$\text{Loss} = -\frac{1}{N} \sum_{I=0}^{N-1} \frac{1}{c} \sum_{a=1}^c b_{I,a} \ln(y_{I,a}) + (1 - b_{I,a}) \ln(1 - y_{I,a}). \quad (6)$$

式(6)中:Loss 为二元交叉熵损失函数; N 为焊缝图像数; c 为焊缝缺陷类别数; $b_{I,a}$ 为焊缝图像 I 第 a 个子标签的真实值; $y_{I,a}$ 为焊缝图像 I 第 a 个子标签的网络输出经过 Sigmoid 函数的预测值。

2 实验结果与分析

2.1 实验环境

实验采用的操作系统为 Windows 11,处理器为 12th Gen Intel(R) Core(TM) i7-12700H 2.30 GHz,显卡为 NVIDIA GeForce RTX 3060 Laptop GPU,并通过编程语言 Python 3.8 和深度学习框架 Pytorch 2.0.0 进行训练。

将训练图像分辨率统一调整为 1 280 px×700 px,以保持输入模型的图像尺寸一致,训练过程中使用的优化器为 Adam,学习率设置为 0.001,批次大小设置为 16,并根据验证结果保存最优的模型参数。

2.2 评价指标

精确率(P)表示真实正样本中被正确分类的样本占有所有样本的比例,其计算公式为

$$P = \frac{\text{TP}}{\text{TP} + \text{FP}}. \quad (7)$$

式(7)中:TP 表示真实正样本中被正确分类的样本;FP 表示真实负样本中被错误分类的样本。

召回率(R)表示真实正样本中被正确分类的比例,其计算公式为

$$R = \frac{\text{TP}}{\text{TP} + \text{FN}}. \quad (8)$$

式(8)中:FN 表示真实正样本中被错误分类的样本。

F_1 表示 P 和 R 的调和平均数,用来衡量分类器的综合性能。 F_1 值越大,代表模型性能越好,其计算公式为

$$F_1 = \frac{2 \times P \times R}{P + R}. \quad (9)$$

平均精准度(P_A)表示模型在每个类别上的性能表现,其计算公式为

$$P_A = \int_0^1 P(R) dR. \quad (10)$$

均值平均精度(P_{mA})表示 P_A 的平均值,用来衡量模型在所有类别上的性能表现。 P_{mA} 值越大,代表模型性能越好,其计算公式为

$$P_{mA} = (\sum P_A)/c.$$

(11)

选取 P, R, F_1 和 P_{mA} 4 种评价方法对焊缝图像多标签分类模型的性能进行评估。

2.3 实验数据集

数据集采用 WELDX 自建数据集和 GDXray 公开数据集^[17]。使用 VIDAR 型工业胶片扫描仪对真实工业场景中采集的 X 射线焊缝缺陷底片样本进行扫描并转换为数字图像,由此自制了 X 射线焊缝缺陷 WELDX 数据集。数据集按照行业标准 NB/T 47013.1—2015^[18]将焊缝缺陷分为 5 种类别,分别是裂纹、未熔合、未焊透、条形缺陷和圆形缺陷,共计 1 251 张图像,其中,裂纹、未熔合和未焊透缺陷尺寸较大,条形缺陷和圆形缺陷尺寸微小,且各缺陷尺寸大小均不一。经过人工裁剪、筛除不符合要求的图像后,共获得 528 张图像,各图像的分辨率大小不等。使用 LabelImg 软件标注出 X 射线焊缝图像中缺陷的类别和位置信息,并通过翻转、平移及对比度变化等数据增强方法扩充原始样本,最终获得 2 112 张 X 射线焊缝图像。

为了确保训练结果的普遍性和可靠性,将 WELDX 数据集划分为训练集、验证集和测试集,其中,训练集包含 1 372 张 X 射线焊缝图像,验证集包含 529 张 X 射线焊缝图像,测试集包含 211 张 X 射线焊缝图像。将 GDXray 数据集作为外部验证集来评估模型在其他焊缝缺陷检测场景中的泛化能力。X 射线焊缝缺陷数据集样本示例,如图 4 所示。

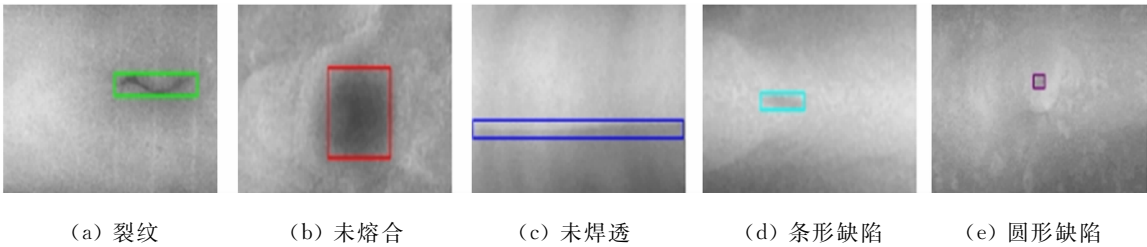


图 4 X 射线焊缝缺陷数据集样本示例

Fig. 4 Sample example of X-ray welding defect dataset

2.4 评估实验

为了评估文中算法的有效性,对多种基于不同建模方式的传统算法输出结构进行修改并对比输出结果,挑选的传统算法包括 ResNet-50, ResNet-101, SqueezeNet^[19], VGG-16^[20], SCConv-R50^[21], RefConv^[22]和 GCN。其中,ResNet-50 和 ResNet-101 分别是 ResNet 算法 50 层和 101 层的网络模型。比较 8 种不同算法在 WELDX 数据集上的表现,结果如表 1 所示。

表 1 不同传统算法在 WELDX 数据集上的表现对比

Tab. 1 Performance comparison of different traditional algorithms on WELDX dataset

算法	$P/\%$	$R/\%$	$F_1/\%$	$P_{mA}/\%$
ResNet-50	72.5	57.6	64.2	72.3
ResNet-101	71.7	56.4	63.3	71.4
SqueezeNet	71.6	55.8	62.7	69.7
VGG-16	71.5	56.1	63.1	70.1
SCConv-R50	72.2	56.7	63.5	71.9
RefConv	72.1	57.2	63.8	72.0
GCN	73.3	59.2	65.8	74.5
文中算法	76.7	67.1	72.0	78.1

在传统深度卷积神经网络模型中,ResNet 算法在 WELDX 数据集上的性能表现更加优异,通过引入残差模块,该算法能解决随着神经网络深度增加而出现的梯度弥散问题。相比 ResNet-101 算法,ResNet-50 算法的 F_1 值提高了 0.9%, P_{mA} 值提高了 0.9%,说明 ResNet-50 算法的拟合度更好,网络层数更深的模型可能会增加过拟合的风险,从而影响模型的性能。由表 1 可知:ResNet-50 算法比 SqueezeNet, VGG-16, SCConv-R50, RefConv 算法的 F_1 值分别提高了 1.5%, 1.1%, 0.7%, 0.4%, 且 P_{mA} 值分别提高了 2.6%, 2.2%, 0.4%和 0.3%。因此,文中算法选取 ResNet-50 作为基础网络,对输入焊缝图像进行特征提取。

从表 1 也可以看出,相比其他传统算法,文中算法在 WELDX 数据集上取得了更优的焊缝图像多标签分类结果;与 ResNet-50 基础网络相比,文中算法的 F_1 和 P_{mA} 分别提高了 7.8%,5.8%,说明采用图卷积神经网络学习特征节点之间的上下文相似性关系可以帮助网络提升多标签图像分类的性能;与 GCN 相比,文中算法的 F_1 和 P_{mA} 分别提高了 6.2%,3.6%,说明采用包含 Res2Net 的 ResNet-50 算法能够充分发挥其强大的特征提取能力,深入挖掘 X 射线焊缝图像中的细节信息。相比其他传统算法,文中算法在 WELDX 数据集上的表现(P, R, F_1 和 P_{mA} 4 个评价指标)均有较大提升,说明在焊缝图像多标签分类任务中,该算法表现出一定的效果。

选取 RMIC^[23], ML-GCN^[24], F-GCN^[25], FAN^[26], CFMIC^[27] 和 ViGh^[28] 6 种不同的多标签图像分类算法与文中算法进行对比,进一步评估文中算法的有效性。不同分类算法在 WELDX 数据集上的表现对比,如表 2 所示。

RMIC 算法以强化学习的原始多标签问题为基础,融合图像特征与前序标签信息,从而建立新的标签预测状态。与 RMIC 算法相比,文中算法的 F_1 和 P_{mA} 分别提高了 3.8%和 1.8%,说明采用图卷积神经网络构建图结构来学习上下文信息可以取得更好的效果。

ML-GCN 算法利用标签作为节点,标签相关性作为边来构建图结构,边的权重基于标签共现概率,并通过 GCN 更新图结构。因为加入标签共现数据的信息,在性能上得到较大提升。然而,采用标签共现关系会使标签类别较少的样本被正确分类的可能性下降,而标签类别较多的样本被正确分类的可能性上升。文中算法通过 HC 算法和 RKNN 算法在多尺度特征的基础上构建图结构能够增强对图像特征之间相似性关系的学习, F_1 和 P_{mA} 分别提高了 2.5%和 1.3%。

表 2 不同多标签图像分类算法在 WELDX 数据集上的表现对比

Tab. 2 Performance comparison of different multi label image classification algorithms on WELDX dataset				
算法	$P/\%$	$R/\%$	$F_1/\%$	$P_{mA}/\%$
RMIC	72.6	64.2	68.2	76.3
ML-GCN	74.4	65.2	69.5	76.8
F-GCN	75.3	66.7	70.8	77.0
FAN	72.9	66.5	69.7	77.2
CFMIC	76.2	66.8	71.3	77.5
ViGh	76.4	66.9	71.5	77.7
文中算法	76.7	67.1	72.0	78.1

F-GCN 算法基于 GCN 采用多模态分解双线性池化,将图像信息和标签嵌入信息进行融合。该算法通过端到端方式进行训练,相较于 ML-GCN 算法,更强调标签语义信息和图像特征信息的有效结合,以促进对图像特征的标签语义引导学习。FAN 在 ML-GCN 基础上引入一种稳健的协方差池化,以汇总卷积特征,从而形成全局图像表示,捕获图像特征。文中算法引入 RDGNP 算法,针对不同类别提取的图像特征可以更加有效地提取图结构特征,从而进一步提升模型的性能。由此,相较于 F-GCN 算法,文中算法的 F_1 和 P_{mA} 分别提高了 1.2%和 1.1%。相较于 FAN 算法,文中算法的 F_1 和 P_{mA} 分别提高了 2.3%和 0.9%。

CFMIC 算法结合注意力机制和 GCN,并引入跨模式融合模块,以捕获局部和全局标签的依赖关系。尽管该算法取得良好的效果,但是其包含的跨模式融合模块会增加模型的复杂度和规模。相比 CFMIC 算法,文中算法的复杂度更低,训练过程中需要的代价更少,能够更好地满足实际工程应用中实时性的需求,其 F_1 和 P_{mA} 分别提高了 0.7%和 0.6%。

ViGh 算法基于输入图像构建图结构,采用注意力机制(HDmA)和 GNN 获取图像的上下文信息,从而构建全局关系。文中算法采用 GCN 和 RDGNP 算法建立图像特征的上下文关系的同时,使用包含 Res2Net 的 ResNet-50 提取图像的细节特征信息,两者结合实现图像的多标签分类,丰富特征的表达能力,与 ViGh 算法相比,文中算法的 F_1 和 P_{mA} 分别提高了 0.5%和 0.4%。

实验表明,文中算法通过结合焊缝图像的多尺度特征和图结构特征,综合考虑特征细节信息和上下文信息,使模型更加适用于焊缝图像的多标签分类任务,尽管真实工业场景中的焊缝缺陷分类困难,模型性能依然表现良好。然而,这并不代表文中算法在其他焊缝缺陷检测场景中的可靠性。

为了评估文中算法的泛化性,采用文中算法对外部验证集(GDXray)中的图像进行预测,最终得到 F_1 和 P_{mA} 分别达到 61.5%和 67.1%。虽然 GDXray 数据集是在其他检测场景中利用不同的扫描方式获得的焊缝图像,但是文中算法也能相对准确地将焊缝缺陷进行分类,证明其具有良好的泛化能力。

2.5 消融实验

为了评估文中算法中不同模块在焊缝图像多标签分类任务中的有效性,在 WELDX 数据集上对文中算法的不同模块进行消融实验,消融实验结果如表 3 所示。

表 3 文中算法中不同模块在 WELDX 数据集上的消融实验结果

Tab. 3 Ablation experimental results of different modules of proposed algorithm on WELDX dataset

模块	$P/\%$	$R/\%$	$F_1/\%$	$P_{mA}/\%$
ResNet-50	72.5	57.6	64.2	72.3
ResNet-50+Res2Net	71.3	60.8	65.6	74.6
ResNet-50+GCN	74.2	61.9	67.5	75.1
ResNet-50+GCN+RDGNP	75.0	62.5	68.2	76.4
文中算法	76.7	67.1	72.0	78.1

由表 3 可知:当仅使用 ResNet-50 进行特征提取时,由于网络感受野有限,存在忽略图像中微小缺陷的风险,并且只能获取 X 射线焊缝图像的局部信息,而未能捕获全局信息;通过引入 Res2Net 到焊缝图像的多标签分类网络模型中,特征提取阶段替换了具有不同感受野大小的卷积核,这样,每个卷积核在粒度级别上更好地捕捉焊缝缺陷的特征信息,有效缓解了焊缝缺陷尺寸微小的问题,因此,该模型在提取特征时增强了对小目标的敏感程度,从而进一步提高了模型的多标签分类准确率;与仅采用 ResNet-50 的模型相比,采用 ResNet-50+Res2Net 的模型 F_1 和 P_{mA} 分别提高了大约 1.4%和 2.3%。

添加 Res2Net 模块能够提高网络针对焊缝缺陷的关注程度,然而,对于焊缝图像的特征提取依然缺乏全局相关性的描述,而焊缝缺陷存在形状各异的特点,因此更重要的是,基于预训练 ResNet-50 提取的特征图构建图结构后,引入 GCN 学习图结构的上下文信息,从而提取焊缝图像的全局结构特征。通过融合 ResNet-50 输出的图像特征和 GCN 输出的图结构特征,有效结合 X 射线焊缝图像中的局部关键信息和全局结构信息,使特征表达更加充分准确。与仅采用 ResNet-50 提取图像特征相比,采用多特征融合的模型 F_1 和 P_{mA} 分别提高了大约 3.3%和 2.8%。

在小样本数据集中,一般采用 KNN 算法直接构建图结构,能够更好地理解和学习网络,但 X 射线焊缝图像中往往包含噪声信息,且焊缝缺陷的分布随机,可能影响模型的稳定性。通过引入 RDGNP 算法,利用关联度计算方法来强化对构建图结构过程中特征间关系的感知能力,能够更加准确地捕捉特征之间的相似性关系。相较于 ResNet-50+GCN 初始融合模型,ResNet-50+GCN+RDGNP 模型的 F_1 和 P_{mA} 分别提高了大约 0.7%和 1.3%。

由此可以看出,加入 Res2Net 模块、GCN 模块和 RDGNP 模块均会提升焊缝图像多标签分类的 4 个评价指标,针对性地解决焊缝内部缺陷存在的尺寸微小、结构复杂、分布随机和形状各异等问题,说明文中算法中的各模块都能促进焊缝图像多标签分类的实现,并且各模块共同作用的效果更为显著,其中,ResNet-50+GCN 融合模型的性能影响最大。

通过实验明确模型的结构,在图结构重构模块中,针对特征节点的最近邻居节点的数量(k)进行消融实验,目的在于确定合适的最近邻居节点数量,以在特征信息之间实现有效的交换。不同邻居节点数量在 WELDX 数据集上对模型性能的影响,如表 4 所示。

表 4 不同邻居节点在 WELDX 数据集上对模型性能的影响

Tab. 4 Influence of different neighboring nodes on model performance on WELDX dataset

k	$P/\%$	$R/\%$	$F_1/\%$	$P_{mA}/\%$
3	73.0	63.5	67.9	74.7
5	70.8	66.5	68.7	75.4
7	76.7	67.1	72.0	78.1
9	73.4	66.9	70.1	76.8
11	70.5	64.9	67.7	74.2

由表 4 可知:当最近邻居节点的数量 k 为 7 时,模型的性能达到了最佳状态,表明在该配置下,特征节点之间的信息交流得以最优化,实现了对图结构的有效重构。当邻居节点数量太少时,交换的信息受限,降低了模型的表达能力;而当邻居节点数量过多时,一些无用特征的引入可能会干扰整体学习过程。因此,在图结构重构模块中将 k 设置为 7,有助于平衡特征信息的交换与噪声的干扰,为模型的整体性能提供有效的优化策略,以确保 GCN 学习图结构的特征信息前实现更优的重构效果。

3 结 论

针对现有传统卷积神经网络只能提取焊缝图像的局部特征,而无法捕获焊缝图像的全局特征进行有效分类的问题,提出一种多特征融合的焊缝图像多标签分类算法,通过包含 Res2Net 的 ResNet-50 提取焊缝图像的多尺度特征,基于此构建图结构,并引入 RDGNP 算法进一步优化图结构,再利用 GCN 学习图结构蕴含的空间特征,以获取焊缝图像的上下文信息。最后,将卷积神经网络提取的图像特征和图卷积神经网络提取的图结构特征结合,以预测最终的分类结果。在 X 射线焊缝缺陷 WELDX 数据集上的实验结果表明,相较于其他算法,文中算法在 4 个评价指标上均优于传统算法,证明其在焊缝图像多标签分类任务中的有效性,为解决因内部缺陷存在形状各异、尺寸微小、分布随机和结构复杂等特点导致缺陷难以实现准确分类的问题打下良好的算法基础。

然而,算法仍然存在一些不足,如对于焊缝图像的特征提取还有待进一步提高,数据集的建立需要进一步扩充等。因此,后续的研究工作将考虑引入注意力机制等增强网络对感兴趣区域的特征表达,收集更多 X 射线焊缝图像用于网络的训练,以提高焊缝图像多标签分类效果。

参考文献:

[1] DING Kai, NIU Zhangqi, HUI Jizhuang, *et al.* A weld surface defect recognition method based on improved MobileNetV2 algorithm[J]. *Mathematics*, 2022, 10(19): 3678. DOI:10. 3390/math10193678.

[2] XU Hao, YAN Zhihong, JI Bowen, *et al.* Defect detection in welding radiographic images based on semantic segmentation methods[J]. *Measurement*, 2022, 188: 110569. DOI:10. 1016/j. measurement. 2021. 110569.

[3] SAY D, ZIDI S, QAISAR S M, *et al.* Automated categorization of multiclass welding defects using the X-ray image augmentation and convolutional neural network[J]. *Sensors*, 2023, 23(14): 6422. DOI:10. 3390/s23146422.

[4] KUMARESAN S, AULTRIN K S J, KUMAR S S, *et al.* Deep learning-based weld defect classification using VGG16 transfer learning adaptive fine-tuning[J]. *International Journal on Interactive Design and Manufacturing*, 2023, 17(6): 2999-3010. DOI:10. 1007/s12008-023-01327-3.

[5] TOTINO B, SPAGNOLO F, PERRI S, RIAWELC: A novel dataset of radiographic images for automatic weld defects classification[J]. *International Journal of Electrical and Computer Engineering Research*, 2023, 3(1): 13-17. DOI:10. 53375/ijecer. 2023. 320.

[6] 张智慧, 林耀进, 张小清, 等. 基于类别一致性的层次特征选择算法[J]. *闽南师范大学学报(自然科学版)*, 2022, 35(4): 41-49. DOI:10. 16007/j. cnki. issn2095-7122. 2022. 04. 007.

[7] BRUNA J, ZAREMBA W, SZLAM A, *et al.* Spectral networks and locally connected networks on graphs[C]// *International Conference on Learning Representations*. Banff: [s. n.], 2014: 1-14. DOI:10. 48550/arXiv. 1312. 6203.

[8] KIPF T N, WELING M. Semi-supervised classification with graph convolutional networks[C]// *International Conference on Learning Representations*. Toulon: [s. n.], 2017: 1-14. DOI:10. 48550/arXiv. 1609. 02907.

[9] WANG Yucheng, GAO Liang, GAO Yiping, *et al.* A graph guided convolutional neural network for surface defect recognition[J]. *IEEE Transactions on Automation Science and Engineering*, 2022, 19(3): 1392-1404. DOI:10. 1109/tase. 2022. 3140784.

[10] BALCIOGLU Y S, SEZEN B, ÇERASI C C, *et al.* Machine design automation model for metal production defect recognition with deep graph convolutional neural network[J]. *Electronics*, 2023, 12(4): 825. DOI:10. 3390/electronics12040825.

[11] 周忠眉, 孟威. 多角度标签结构和特征融合的多标签特征选择[J]. *闽南师范大学学报(自然科学版)*, 2021, 34(1): 64-71. DOI:10. 16007/j. cnki. issn2095-7122. 2021. 01. 011.

[12] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, *et al.* Deep residual learning for image recognition[C]// *Proceed-*

- ings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas; IEEE Press, 2016; 770-778. DOI:10.1109/cvpr.2016.90.
- [13] GAO Shanghua, CHENG Mingming, ZHAO Kai, *et al.* Res2Net: A new multi-scale backbone architecture[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 43(2): 652-662. DOI:10.1109/TPAMI.2019.2938758.
- [14] HAJEBI K, ABBASI-YADKORI Y, SHAHBAZI H, *et al.* Fast approximate nearest-neighbor search with k -nearest neighbor graph[C]// Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence. Barcelona; AAAI Press, 2011; 1312-1317. DOI:10.5591/978-1-57735-516-8/IJCAI11-222.
- [15] CHEEMA M A, LIN Xuemin, ZHANG Wenjie, *et al.* Influence zone: Efficiently processing reverse k nearest neighbors queries[C]// IEEE 27th International Conference on Data Engineering. Washington D C; IEEE Press, 2011; 577-588. DOI:10.1109/ICDE.2011.5767904.
- [16] LIU Yongli, ZHAO Congcong, CHAO Hao. Density peak clustering based on relative density under progressive allocation strategy[J]. Mathematical and Computational Applications, 2022, 27(5): 84. DOI:10.3390/mca27050084.
- [17] MERY D, RIFFO V, ZSCHERPEL U, *et al.* GDXray: The database of X-ray images for nondestructive testing[J]. Journal of Nondestructive Evaluation, 2015, 34(4): 42. DOI:10.1007/s10921-015-0315-7.
- [18] 全国锅炉压力容器标准化技术委员会. 承压设备无损检测: NB/T 47013.1—2015[S]. 北京: 新华出版社, 2015.
- [19] IANDOLA F N, HAN Song, MOSKEWICZ M W, *et al.* SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size[C]// International Conference on Learning Representations. Toulon; [s. n.], 2017; 1-13. DOI:10.48550/arXiv.1602.07360.
- [20] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition[EB/OL]. (2015-04-10)[2024-02-10]. <https://doi.org/10.48550/arXiv.1409.1556>.
- [21] LI Jiafeng, WEN Ying, HE Lianghua. SCConv: Spatial and channel reconstruction convolution for feature redundancy[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Vancouver; IEEE Press, 2023; 6153-6162. DOI:10.1109/cvpr52729.2023.00596.
- [22] CAI Zhicheng, DING Xiaohan, SHEN Qiu, *et al.* Refconv: Re-parameterized refocusing convolution for powerful convnets[C]// International Conference on Learning Representations. Vienna; [s. n.], 2024; 1-17. DOI:10.48550/arXiv.2310.10563.
- [23] HUANG Xun, BELONGIE S. Arbitrary style transfer in real-time with adaptive instance normalization[C]// Proceedings of the IEEE International Conference on Computer Vision. Venice; IEEE Press, 2017; 1501-1510. DOI:10.1109/iccv.2017.167.
- [24] CHEN Zhaoming, WEI Xiushen, WANG Peng, *et al.* Multi-label image recognition with graph convolutional networks[C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Long Beach; IEEE Press, 2019; 5177-5186. DOI:10.1109/cvpr.2019.00532.
- [25] WANG Yangtao, XIE Yanzhao, LIU Yu, *et al.* Fast graph convolution network based multi-label image recognition via cross-modal fusion[C]// Proceedings of the 29th ACM International Conference on Information & Knowledge Management. New York; ACM, 2020; 1575-1584. DOI:10.1145/3340531.3411880.
- [26] LI Yaning, YANG Liu. More correlations better performance: Fully associative networks for multi-label image classification[C]// 25th International Conference on Pattern Recognition. Milan; IEEE Press, 2021; 9437-9444. DOI:10.1109/icpr48806.2021.9412004.
- [27] WANG Yangtao, XIE Yanzhao, ZENG Jiangfeng, *et al.* Cross-modal fusion for multi-label image classification with attention mechanism[J]. Computers and Electrical Engineering, 2022, 101: 108002. DOI:10.1016/j.compeleceng.2022.108002.
- [28] PANG Wenkai, TAN Zhi. A steel surface defect detection model based on graph neural networks[J]. Measurement Science and Technology, 2024, 35(4): 046201. DOI:10.1088/1361-6501/ad1c4b.

(责任编辑: 黄晓楠 英文审校: 陈婧)