

基于改进 ResUnet 网络结构的金属表面缺陷分割方法

周子杰¹, 李琳¹, 薛殿龙¹, 李家军¹, 黄伟龙², 陈德阳²

(1. 中国航发 哈尔滨东安发动机有限公司, 哈尔滨 150000;

2. 重庆大学 机械与运载工程学院, 重庆 400044)

摘要: 针对金属表面缺陷因尺寸和形状多样化导致检出率低, 检出形状差异大等问题, 提出了一种基于改进 ResUnet 的表面缺陷语义分割算法, 此方法以 ResUnet 网络结构为基础架构, 使用先进的 ConvNeXt-T 改进骨干网络, 提取更具代表性的特征; 在解码阶段添加全局上下文模块, 增强模型的全局上下文建模能力; 同时融合解码阶段的多级多尺度特征, 使模型更适应缺陷尺寸和大小多变的钢铁表面缺陷, 提升缺陷识别精度; 在谢韦尔钢铁公司提供的钢铁表面缺陷数据集上开展了所提缺陷识别算法的有效性定量和定性验证, 与对比方法相比, 该方法的 mDice 值达到了最高的 0.778 4, 且单张图像在 GPU 上的运行时间只需 14.7 ms, 同时缺陷分割结果与标签最接近, 说明该方法具有较好的鲁棒性、准确性和高效性。

关键词: 缺陷分割; 多尺度特征; 全局上下文; 自注意力; UNet

Metal Surface Defect Segmentation Method Based on Improved ResUnet Network Structure

ZHOU Zijie¹, LI Lin¹, XUE Dianlong¹, LI Jiajun¹, HUANG Weilong², CHEN Deyang²

(1. Harbin Dongan Engine Co., Ltd., AECC, Harbin 150000, China;

2. College of Mechanical and Vehicle Engineering, Chongqing University, Chongqing 400044, China)

Abstract: Metal surface defects has the characteristics of diverse shapes and sizes, which result in low detection rates and significant differences in detected shapes. To address this issue, a metal surface defect semantic segmentation algorithm based on improved ResUnet is proposed. Take the ResUnet network architecture as a basic architecture, adopt an advanced ConvNeXt-T to improve the backbone network, and extract more representative features; Add a global context module during the decoding phase to enhance the global context modeling capability of the model; Fuse multi-level and multi-scale features in decoding stage, make the model more suitable for steel surface defects with variable sizes and dimensions, and improve defect recognition accuracy. The effectiveness of the proposed defect recognition algorithm is quantitatively and qualitatively validated on the steel surface defect dataset provided by Xavier Steel Company. Compared with the comparison method, this method has reached the maximum mDice of 0.778 4, and the running time of a single image on GPU is only 14.7 ms. The defect segmentation results are closest to the labels, indicating that this method has good robustness, accuracy, and efficiency.

Keywords: defect segmentation; multi-scale features; global context; ConvNeXt-T; ResUnet

0 引言

在钢铁产品的制造或加工等过程中, 由于技术的限制, 会不可避免的在其表面产生裂纹、划痕等缺陷, 这些缺陷的存在会严重影响钢铁在使用过程中的性能,

因此, 在生产线上对钢铁产品的缺陷进行检测识别至关重要^[1-2]。由于表面缺陷的多样性与复杂性, 导致目前在表面缺陷自动检测领域在技术上还面临诸多的挑战。传统图像处理法利用图像局部异常反映的原始特性来分割和检测缺陷, 大致有阈值法^[3]、结构法^[4]、滤波

收稿日期:2024-05-24; 修回日期:2024-08-28。

作者简介:周子杰(1993-),男,硕士研究生,工程师。

引用格式:周子杰,李琳,薛殿龙,等. 基于改进 ResUnet 网络结构的金属表面缺陷分割方法[J]. 计算机测量与控制, 2025, 33(9):191-199.

器法^[5]、基于模型^[6]等方法,但当光照、摄像机角度和背景发生改变时,缺陷检测的性能受到影响很大。基于传统机器学习的钢铁表面缺陷识别方法主要分为基于纹理特征的方法和基于形状特征的方法。文献[7]设计了一种判别流形正则化局部描述器算法,获取钢铁表面缺陷特征,并通过空间中定义的流形距离完成匹配,实现钢铁表面图像中缺陷的分类。文献[8]提出一种基于混合染色体遗传算法的分类模型,结合图像的几何、形状、纹理和灰度特征对带钢表面缺陷进行识别和分类。基于传统机器学习的方法通常仅对钢铁表面缺陷进行分类,没有实现对缺陷的定位和分割功能。近年来,伴随着深度学习技术的发展,越来越多的工业场景开始引入深度学习技术,不少研究者使用深度学习代替传统的机器学习方法来检测钢铁表面缺陷。文献[9]提出一种基于 U-Net 和 FPN 的编码器-解码器架构的分割网络,实现图像像素级钢材表面缺陷的自动化检测,该算法在时间成本和分割精度方面都具有不错的性能。文献[10]提出一种带背景约束的全局低秩非负重建算法,融合二维显著图像的检测结果和三维的轮廓信息,实现钢轨表面缺陷的检测。文献[11]提出一种卷积变分自编码器,并用其生成缺陷图像,然后使用这些生成的图像训练提出的 CNN 分类器,以实现高精度的金属表面缺陷检测。

基于深度学习的图像分类方法能对各表面缺陷图像进行分类,但不能区分缺陷区域与非缺陷区域。语义分割应用于钢铁表面缺陷识别领域,可以将钢铁表面图像分割成不同的区域,并将每个像素分配到不同的类别中(背景和各类缺陷),这使得语义分割技术在钢铁表面缺陷的识别上具有很大的潜力。文献[12]提出一种钢铁表面缺陷分割算法,通过采用稠密连接策略、改进网络损失函数和训练数据增强等方法,提升了钢铁表面缺陷分割的精度,但其对样本量较少、尺寸较小的缺陷分割效果较差,经常出现漏检等情况。文献[13]提出一种多任务模型用于钢铁表面缺陷分割和缺陷严重性评估,通过模型融合和过程参数的扩展,提高了缺陷分割的准确性,不过相较于单任务模型,该模型相对更加复杂,不利于实际生产中的实施。文献[14]提出一种钢铁表面缺陷检测的链式空洞空间金字塔池化网络(CASPP-Net),通过引入改进的全局注意力模块和残差边界细化模块有效提升了缺陷分割的精度。语义分割技术在钢铁表面缺陷识别领域中的应用是一种快速、准确、高效的缺陷检测方法,具有巨大的应用前景,它可以大大提高缺陷检测的准确性和效率,为钢铁生产企业提供强有力的支持,同时也为计算机视觉技术的发展和应用提供了重要的参考和借鉴。

1 ResUnet 网络结构

ResUnet^[15]网络是一种将 ResNet 网络和 U-Net 网络结合的图像分割网络,ResNet^[16]网络提出了恒等跨层连接,能有效避免网络加深时的退化现象,且训练时收敛更快,U-Net^[17]网络广泛应用于医学图像分割领域,其使用多个上下文连接的通道和上采样模块,实现了低层特征与高层特征的融合,适用于大尺寸、小样本的图像分割。ResUnet 综合了两者的优点,残差单元能简化网络的训练,残差块的跨层连接以及低层与高层通过跳跃连接的信息融合能促进特征信息的传播,提升特征对目标的表达能力。钢铁表面缺陷通常具有缺陷尺寸变换大、缺陷样本较少、各缺陷类别样本不均衡等特点,ResUnet 能较好的适应钢铁表面缺陷识别场景下的这些特点,可用于钢铁表面缺陷的分割。

ResUnet 网络的一般结构如图 1 所示。对于一张 3 通道的输入缺陷图像,首先采用带残差连接的两层 3×3 卷积对其初步编码以提取特征,紧接着通过 4 个编码块与解码块,编码块和解码块结构如图 2 所示。编码块采用核尺寸为 2×2 的最大池化对特征进行下采样,使提取的特征从扁平状变得更加立体,特征包含的语义信息更完整;解码块使用插值规模因子为 2 的双线性插值对高层特征进行上采样,使特征图逐步恢复至输入图像大小,之后与低层特征以特征通道维度拼接的方式融合在一起。编码块和解码块均采用连续两次的 3×3 卷积对输入特征进一步融合,同时使用残差连接将初始特征以恒等变换的方式融合到输出特征中,增强深层网络的纹理、颜色等低级特征的表达能力,保证网络在多级特征提取之后仍保留位置等细节信息。每一次下采样都将特征图大小减半,特征维度

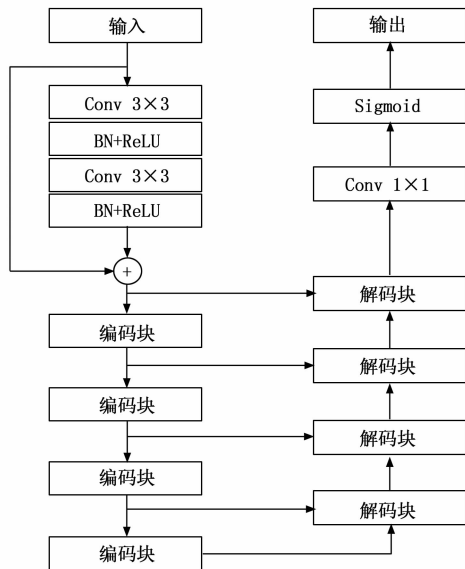


图 1 ResUnet 网络结构

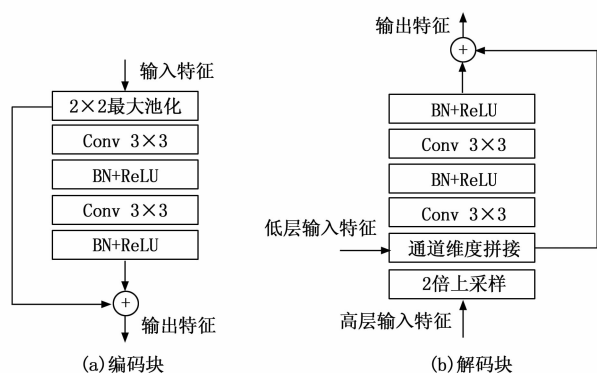


图2 ResUnet 网络的编码块和解码块结构

加倍;每一次上采样都将特征图大小加倍,特征维度减半,最终获得原始输入图像大小的特征图。特征提取完成之后,使用 1×1 卷积进行逐像素分类,输出各缺陷与背景分离的缺陷分割图。

2 基于改进 ResUnet 的表面缺陷分割

2.1 网络架构及损失函数

相较于计算机视觉领域的自然图像分割,工业领域的钢铁表面缺陷分割任务较为简单,缺陷场景的背景复杂度低,但缺陷的边缘等细节信息难以捕捉,不同缺陷之间存在类内差异大且类间不明显的问题,因此,缺陷的高级语义信息和低级特征信息都很重要。

使用 ConvNeXt^[18] 网络改进 ResUnet 网络的特征提取网络,在解码阶段引入全局上下文模块,增强网络的全局上下文建模能力,同时在网络进入分割层之前,融合解码阶段的多级特征,增强高级语义信息和低级特征信息的复用度,能有效提升网络对缺陷分割的性能。改进之后的 ResUnet 网络如图3所示。

提出的改进 ResUnet 网络首先使用步长为4的 4×4 卷积将输入图像下采样4倍并使用层规范化(LN)操作对数据进行归一化,这样能降低特征图的大小,有效降低模型的计算量。紧接着特征图将经过特征编码、特征解码和多级特征融合3个阶段,在特征编码阶段,利用 ConvNeXt 编码块对特征进行初级编码以获得能够有效表达不同缺陷和背景的特征;在特征解码阶段,除了使用和 ResUnet 一样的残差块解码并融合特征之外,还添加了全局上下文模块,以增强缺陷分割网络的全局建模能力;在多级特征融合阶段,使用 1×1 卷积调整各级特征的通道数至一致,同时将低分辨率的高级语义特征上采样到与最大分辨率特征大小一致,之后将各级特征通过通道维度拼接的方式组合在一起,再经过一层 1×1 卷积融合这些特征,获得用于缺陷分割预测的最终特征。最后使用 1×1 卷积对特征图进行逐像素预测并上采样到原图大小,获得与原图逐像素对应的缺陷分

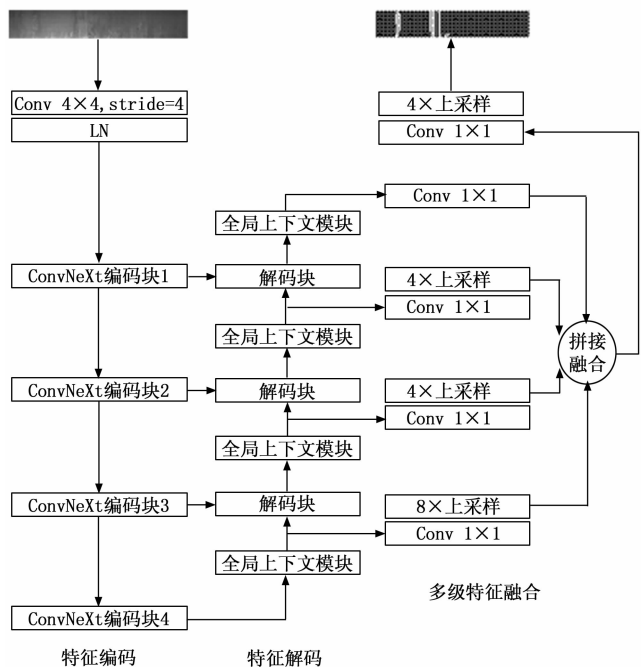


图3 改进的 ResUnet 网络结构

割输出图,上采样方法为双线性插值法。

图像分割问题可以归结为逐像素的分类问题,因此图像分割神经网络最常使用的损失函数为交叉熵损失函数,其计算如式(1)和(2)所示:

$$L_{ce} = - \sum_{k=0}^K \sum_{i=1}^W \sum_{j=1}^H t_{ij} \log_2 p_k(s_{ij}) \quad (1)$$

$$p_k(s_{ij}) = P_{\text{softmax},k}(s_{ij}), k = 0, 1, \dots, K \quad (2)$$

式中, W 、 H 分别表示图像的宽和高, K 表示缺陷的总类别数, t_{ij} 表示像素的真实类别标签, s_{ij} 表示像素的预测类别标签, $P_{\text{softmax},k}()$ 表示使用 softmax 函数计算像素预测为类别 k 的概率。

为改善类别不平衡给网络训练带来的影响,使用加权的交叉熵损失函数,如式(3)所示,其中, w_k 表示类别 k 的权重系数:

$$L_w = - \sum_{k=0}^K w_k \left[\sum_{i=1}^W \sum_{j=1}^H t_{ij} \log_2 p_k(s_{ij}) \right] \quad (3)$$

2.2 ConvNeXt 特征提取网络

ConvNeXt 是文献 [18] 提出的一种深度神经网络结构,它是以残差网络模型 ResNet-50 为基准,对照先进的骨干网络 Swin^[19] 的结构修改而成的,具有强大的特征提取能力,在图像分类和物体识别方面性能表现优异。ConvNeXt 的主要结构如图3的特征编码部分所示,由4个阶段的 ConvNeXt 编码块组成,ConvNeXt 编码块由基本模块 ConvNext Block 堆叠而成,除第一阶段外,在使用多个 ConvNeXt Block 对特征进行编码提取前,先使用核大小为 2×2 ,步长为2的卷积对特征进行2倍下采样,这样4个阶段的特征相较于原图大小的下采样倍数分别为4,8,16,32。

ConvNeXt 与 ResNet 基本模块结构的对比图如图 4 所示。相较于 ResNet Block 的残差结构, ConvNeXt Block 采用倒残差结构, 在第一层使用卷积核大小为 7×7 的逐通道卷积 (depthwise Convolution, dConv), 紧接着使用两层 1×1 卷积对特征升维和降维, 同时融合不同通道维度的特征, 在第一层 1×1 卷积之前使用 LN 对卷积之后的特征做归一化, 在第二层卷积之前使用 GELU 激活函数以提高网络的非线性建模能力。

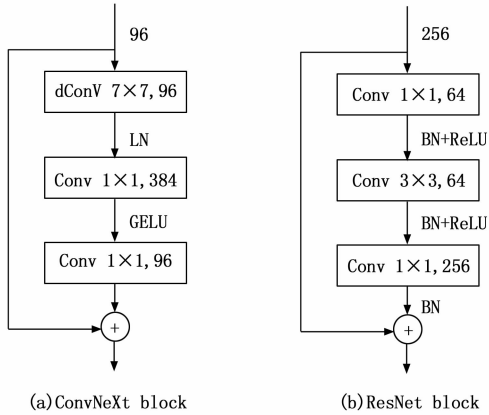


图 4 基本模块对比图

如图 5 所示为普通卷积和逐通道卷积的卷积过程图, 逐通道卷积能够有效降低模型的参数量, 普通卷积的输出特征的每个维度由所有输入特征维度卷积之后相加所得, 而逐通道卷积的输入特征的通道维度与输出特征的通道维度一致, 且一个维度的输入特征只使用一个卷积核运算获得对应维度的输出特征。

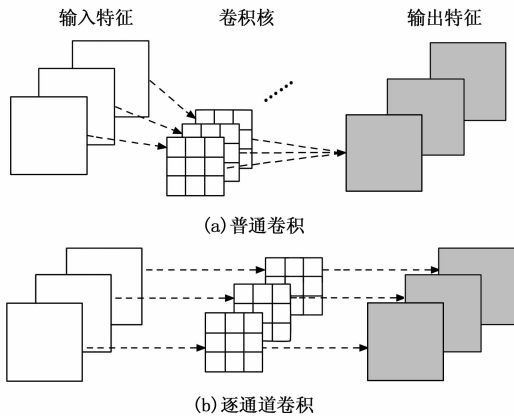


图 5 卷积过程图

假设某卷积的输入和输出特征通道数均为 C , 卷积核大小为 $k \times k$, 则使用普通卷积的参数量 N_{Conv} 和逐通道卷积的参数量 N_{dConv} 分别如式 (4) 和式 (5) 所示, 不难看出, 逐通道卷积的参数量是普通卷积参数量的 $1/C$:

$$N_{\text{Conv}} = C \times (k \times k) \times C \quad (4)$$

$$N_{\text{dConv}} = C \times (k \times k) \quad (5)$$

ConvNeXt 网络使用逐通道卷积能够极大的降低模型的参数量, 由于在逐通道卷积之后又使用了两层 1×1 卷积实现升维与降维的同时使特征不同维度之间进行信息交互, 弥补了逐通道卷积在不同特征维度之间没有信息交互的不足, 因此, ConvNeXt Block 的结构既轻量, 又有强大的特征提取能力。为了对标 Swin 模型结构, ConvNeXt 模型推出了 T/S/B/L 四个版本, 使用的为 ConvNeXt-T 版本, 其 4 个阶段的 Block 堆叠次数分别为 3、3、9、3 次, 输出特征维度分别为 96、192、384、768。

2.3 全局上下文模块

为了增强网络对全局上下文信息的感知和建模能力, 在特征解码阶段添加全局上下文模块, 通过添加全局上下文注意力^[20]的方式给特征图加权融合上下文信息, 以获得更具代表性的特征, 计算如式 (6) 所示:

$$z_i = x_i + W_{v2} \text{ReLU} \left\{ \text{LN} \left[W_{v1} \sum_{j=1}^{N_p} \frac{\exp(W_k x_j)}{\sum_{m=1}^{N_p} \exp(W_k x_m)} x_j \right] \right\} \quad (6)$$

其中: x_i 和 z_i 分别表示某特征像素的输入和输出值, N_p 表示输入特征图的像素位置总数, W_k 、 W_{v1} 、 W_{v2} 为 3 个转换矩阵, 用于计算自注意力机制里面的键矩阵或值矩阵, LN 和 ReLU 表示层归一化和激活函数。

全局上下文模块的网络结构如图 6 所示。首先使用 1×1 卷积把特征通道数降维为 1 以融合输入, 同时使用函数获得逐像素概率图, 再以此概率给输入逐像素添加权重, 完成上下文建模功能, 之后使用两次 1×1 卷积对含上下文信息的特征图进行升维与降维, 实现不同特征通道之间的信息交互, 最终以逐像素相加的方式在输入特征上融合上下文信息, 从而提高网络对图像全局信息的感知能力。算法的流程图如图 7。

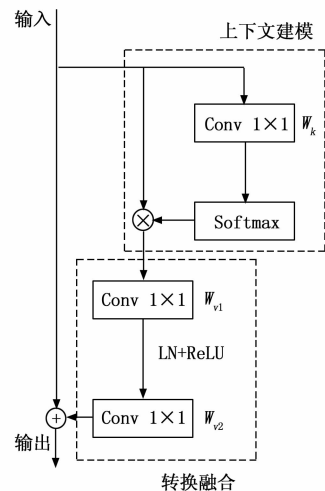


图 6 全局上下文模块结构

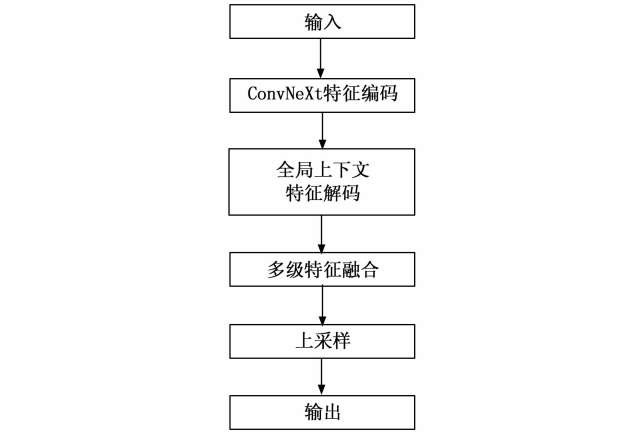


图 7 算法流程图

3 实验验证与结果分析

3.1 实验平台与数据集

验证实验在 Windows10 64 位的操作系统下进行, CPU 为 Intel Core i9-13900K, 使用 3.8 版本的 Python 语言环境, 以 Pytorch 版本为 1.13.1 的深度学习框架搭建的网络模型, 后端使用 CUDA 11.8 软件开发工具包来使用 GPU 加速, GPU 加速器使用的是 GeForce RTX 3090, 显存为 24G。

使用的数据集为 Kaggle “Severstal: Steel Defect Detection” 比赛中的数据集^[20], 该数据集包含 6 666 张有缺陷的图像, 图像分辨率为 1 600×256, 包括 4 种缺陷, 分别是麻点含 769 张、开裂含 195 张、划痕含 4 759 张、斑块含 516 张, 其余 427 张图像同时含两种以上缺陷类别, 数据集具体情况如图 8 所示。由图 8 可以看出, 此数据集存在严重的类不平衡现象, 采用数据增强和修改损失函数中各个类别的权重来降低数据不均

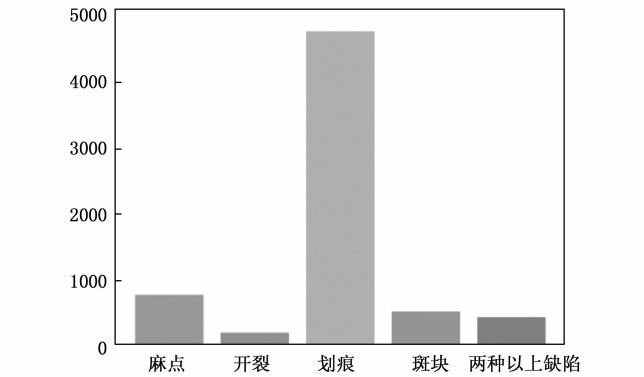


图 8 数据集所含缺陷类别图像数

3.2 缺陷分割性能评价指标

评价指标用于评价模型性能的好坏。最基本也是最重要的指标为 TP (True Positive)、FN (False Negative)、FP (False Positive)、TN (True Negative), 对某

一类缺陷而言, TP 代表预测类别与实际类别都是该类缺陷的像素数, FN 代表实际类别是该类缺陷而被模型预测为其他类别或背景的像素数, FP 代表实际类别不是该类缺陷但模型错误的将其预测为该类缺陷的像素数, TN 代表预测类别和实际类别都不是该类缺陷的像素数, 这 4 个指标组成的混淆矩阵如表 1 所示。

表 1 目标类别缺陷的像素分割混淆矩阵

实际类别	预测类别	
	目标类别缺陷	其他
目标类别缺陷	TP	FN
其他	FP	TN

精确度 (Precision) 也称精度, 查准率, 阳性预测值 (PPV, positive predictive value), 表现为预测类别为目标类别缺陷的像素中, 有多少是真实的目标类别缺陷, 其计算公式为如式 (7) 所示:

$$Precision(PPV) = \frac{TP}{TP + FP} \tag{7}$$

召回率 (Recall) 也称真阳性率 (TPR, true positive rate), 灵敏度, 查全率, 表现为真实目标类别的缺陷有多少被正确预测了出来, 其计算公式如式 (8) 所示:

$$Recall(TPR) = \frac{TP}{TP + FN} \tag{8}$$

精确度和召回率是两个相互制约的指标, 如果模型一味追求高精确度, 可能会导致较低的召回率, 出现大量缺陷未被检测出的问题; 而如果模型过于着重于召回率时, 可能会使模型的精确率过低, 造成很多非缺陷被误判为目标缺陷的问题。因此, 在实际使用过程中需要权衡两者的关系, 也即 F-Score 指标, 其计算公式如式 (9) 所示:

$$F-Score = (1 + \alpha^2) \frac{PPV \times TPR}{\alpha^2 PPV + TPR} \tag{9}$$

其中: α 为加权系数, 当召回率和精确度同等重要时, 取 $\alpha=1$, 此时也称该指标为 F_1 系数; 当精确度比召回率更重要时, 取 $\alpha<1$; 当任务的召回率比精确度更重要时, 取 $\alpha>1$ 。旨在准确无误的分割出图像中存在的不同缺陷, 认为精确度和召回率同等重要, 因此使用 F_1 指标, 也叫做 Dice 相似系数, 其计算公式分别如式 (10) 和式 (11) 所示:

$$F_1 = \frac{2 \times PPV \times TPR}{PPV + TPR} \tag{10}$$

$$Dice = \frac{2TP}{2TP + FP + FN} \tag{11}$$

对于多类别缺陷分割而言, 数据不平衡将导致样本量多的缺陷类别的 Dice 相似系数更高, 而样本量少的缺陷类别的 Dice 相似系数偏低, 因此, 使用平均 Dice 相似系数 (mDice) 作为缺陷分割算法的评价指标, 计

算公式如式 (12) 所示, 其中, K 表示缺陷总类别数, 当 $i=0$ 时表示对背景 (非缺陷部分) 分割的 Dice 指标。此外, 使用模型参数量和单张图片推理的计算量及推理时间作为评价算法实时性的指标:

$$mDice = \sum_{i=0}^K Dice_i$$

(12)

3.3 模型训练及超参数设置

由于图像数据集所含缺陷的样本较少, 故以迁移学习的方式提高分割模型的准确性及稳定性, 使用 ImageNet 数据集上预训练的 ConvNeXt-T 骨干网络模型的参数作为使用模型的骨干网络的初始参数。将数据集以 8 : 2 的比例划分为训练集和测试集, 其中, 训练集与测试集中各个类别的图像数量如表 2 所示。

表 2 实验数据集划分详情

缺陷类别	训练集	测试集	总计
仅含麻点	616	153	769
仅含开裂	156	39	195
仅含划痕	3 808	951	4 759
仅含斑块	413	103	516
含两种以上缺陷	342	85	427
总计	5 335	1 331	6 666

设置批量处理参数 batch_size 为 8, 训练周期为 80 个周期 (epoch), 迭代次数约 52 000 次。初始学习率 learning_rate 设置为 1e-4, 以线性学习率衰减 (Linear learning rate decay) 的策略逐步将学习率衰减从而更好地调整模型参数, 前 1 500 次迭代采用学习率预热的方式逐步增大学习率到初始值, 以免模型训练之初性能不稳定, 模型优化器采用 AdamW 优化器, 它比 Adam 优化器具有更好的稳定性, 一阶矩估计 beta_1 和二阶矩估计 beta_2 分别取 0.9 和 0.999, weight decay 参数设置为 0.05。由于数据集存在严重的类别不平衡现象, 因此, 采用训练时随机裁剪、水平翻转、垂直翻转、亮度 and 对比度变换的方式对数据进行增强, 同时采用加权交叉熵损失函数来缓解数据不平衡的影响, 其中, 背景、麻点、开裂、划痕、斑块的权重分别设为 0.5、1.5、1.7、0.8、0.7。图 9 显示了训练过程中损失的变化, 最终损失稳定在 0.03 附近。

3.4 实验结果与分析

除训练了改进 ResUnet 网络之外, 为了验证方法的有效性与先进性, 还使用了 U-Net、ResUnet、FCN、OCRNet^[23] 网络进行对比实验。各模型算法性能对比结果如表 3 所示, 为了同时验证模型的准确性和高效性, 除使用平均相似系数指标之外, 还计算了各模型的参数量、计算量以及单张图片的运行时间。

由表 3 可以看出, 原始的 U-Net 网络和 ResUnet 网络虽然参数量较少, 但计算量比较大, 这是因为 U-Net

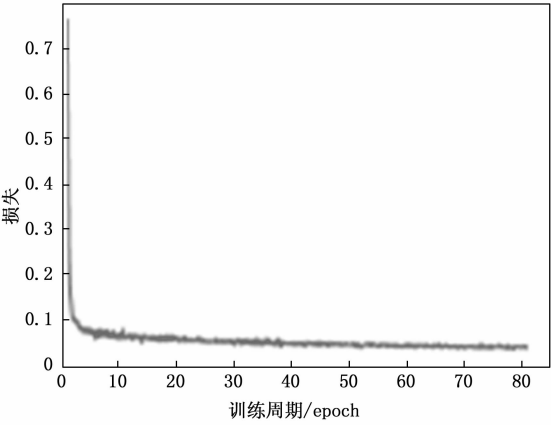


图 9 改进的 ResUnet 网络训练损失

表 3 各模型算法性能对比

方法	参数量/M	计算量/GFLOPs	mDice	GPU/(ms\图像)
U-Net ^[17]	28.99	317.13	0.657 0	20.9
ResUnet ^[15]	30.39	334.28	0.663 3	24.6
FCN ^[22]	40.13	309.34	0.757 5	19.6
OCRNet ^[23]	36.51	239.32	0.759 1	17.7
本文	36.02	87.75	0.778 4	14.7

网络设计的初衷是为了实现医学图像领域的二值分割, 不适用于多类别分割, 且它们的特征提取是从原图直接开始的, 所操作的特征图尺寸太大将导致计算量过大, 同时模型运行时间也比较长, 其平均相似系数不足 0.7, 分割结果差且实时性不强; 自然图像领域效果较好的 FCN 网络和 OCRNet 网络使用了广泛应用的深度残差网络作为特征提取的骨干网络, 能够获得对不同缺陷更好表示的特征, 且骨干网络特征提取之前先将原始输入图像下采样了 4 倍, 在保证准确分割的同时降低计算量, 提高模型运行速度, 其平均相似系数分别达到 0.757 5 和 0.759 1; 该方法由于采用了先进的骨干网络 ConvNeXt-T 作为特征提取网络, 且在解码阶段添加全局上下文模块, 增强了网络特征的全局表达能力, 同时融合了解码阶段的多级多尺度特征, 以更加适应缺陷形状及大小多变的情况, 最终实现了 0.778 4 的, 单张图片运行时间也仅需 14.7ms, 相较于原始的 ResUnet 网络在系数上提升了 11.51%, 单张图片运行时间快了 10 ms, 在实现高缺陷分割准确率的同时实现了更快的检测速度。

为了验证所提改进点的有效性, 以采用 ConvNeXt-T 骨干网络的 ResUnet 网络为基准, 通过添加不同改进策略的方式进行消融实验, 实验结果如表 4 所示, 其中 GC 代表解码阶段添加全局上下文 (Global Context) 模块, FF 代表解码阶段对多级多尺度特征的特征融合

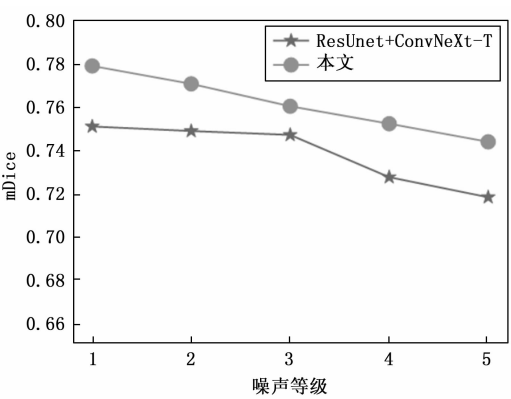
(Feature Fusion)。由表 4 可以看出，基准网络实现了 0.750 3 的值，单张图像的推理时间仅需 12.1 ms；在解码阶段添加全局上下文模块之后，有助于模型对全局上下文的建模能力，增强了特征的全局表示能力，使值提升到 0.760 1；仅在网络解码阶段融合多级多尺度特征，能够更适应特征对不同尺寸与形状的缺陷的表达能 力，对各类缺陷的识别效果更好，其分割指标相较于基准网络有明显提升；同时给基准网络添加全局上下文模块和融合解码阶段的多级多尺度特征，其分割结果最好，值比基准网络提高了 2.81%，且在 GPU 上单张图像的推理时间仅增加了 2.6 ms，证明所提改进策略是有效的。

表 4 改进点消融实验

方法	mDice	GPU 时间 /(ms\图像)
ResUNet + ConvNeXt-T	0.750 3	12.1
ResUNet + ConvNeXt-T + GC	0.760 1	12.6
ResUNet + ConvNeXt-T + FF	0.767 0	14.1
ResUNet + ConvNeXt-T + GC + FF	0.778 4	14.7

为了验证所提方法的鲁棒性和抗噪性，通过给实验数据集添加噪声的方式进行实验，采用不同等级的高斯噪声对原始数据集进行加噪，可以模拟低光照环境下的低质量的图像传输和压缩等情况，各加噪等级的均值都设为 0，方差分别设为 0.000 5、0.001、0.005、0.01，共四组噪声情况，使用加噪后的数据集分别对基准网络 ResUNet + ConvNeXt-T 和的最终改进方法进行实验，实验结果如图 10 所示，其中噪声等级 1 表示没有添加任何噪声的原始数据集，噪声等级 2、3、4、5 分别对应给原始数据集加入 4 个方差逐渐增大的高斯噪声。从图中可以看出，在加入的噪声等级越来越高的情况下，模型的缺陷识别性能越来越低，说明数据集的数据质量对结果存在一定的影响，但总体而言，模型的缺陷识别性能下降较为平缓，每两个噪声等级之间的下降均在 1% 左右，而基准网络虽然在前两个噪声等级下性能保持较好，但后两个噪声等级使其性能下降较为明显，噪声等级 4 下的模型性能超过了所提方法的两倍，说明所提方法具有更好的抗噪性能，对不同质量的数据集的鲁棒性较强。

图 11 所示为模型对输入图像预测的各个缺陷类别的激活图，其中，黑色表示特征图激活之后的概率值较低，而白色表示特征图激活之后的概率值较高。由可视化结果可以看出，所提方法能够较好的识别输入图像中可能存在的缺陷，在原图只存在划痕缺陷的情况下，能比较精确的识别出划痕缺陷所在的位置，且没有误识别出其他缺陷（如麻点、开裂和斑块），证明了所提方法的有效性。



注：上图中噪声等级从1至5分别表示均值为0，方差为0、0.0005、0.001、0.005、0.01的高斯噪声

图 10 不同噪声等级下的模型性能

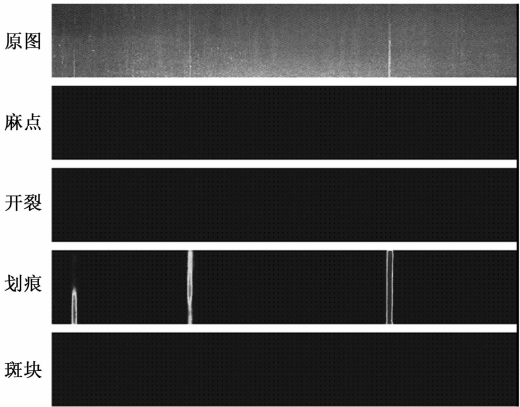


图 11 模型预测的各缺陷激活图

图 12 所示为四类缺陷分割的主观视觉效果，第一排为原始输入图像，第二排为缺陷标签图像，第三排为使用所提方法预测的每一类缺陷的分割结果。通过观察可以看出，所提出的模型对各种不同类别的缺陷都表现出了较好的性能，且缺陷的分割结果比人工标定的标签图像更加平滑，更符合缺陷在真实加工或生产环境中的分布规律。

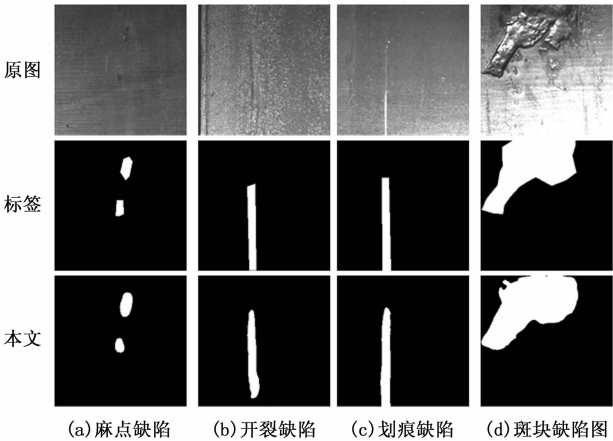


图 12 缺陷分割效果

图 13 为不同对比方法在分割同时包含多种类别缺陷的图像时的分割结果对比图, 示例图像的标签中, 左 45°斜线填充区为开裂缺陷, 右 45°线填充区为划痕缺陷, 水平线填充区为斑块缺陷。由对比图可以明显看出, U-Net、ResUnet 方法受背景干扰比较严重, 分割结果存在碎片化的情况, 且部分区域分割类别与实际不符, 错检严重; FCN、OCRNet 方法的分割结果较前者更细致, 但仍有比较明显的漏检和错检情况; 该方法的分割结果与标签最接近, 相较于其他方法, 漏检和错检情况最少, 最不容易被背景所干扰, 证明所提出的方法具有较好的鲁棒性和先进性。

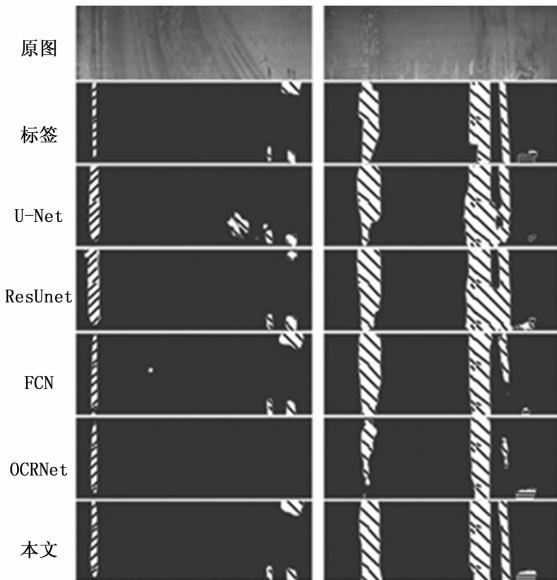


图 13 多类别缺陷分割对比示例

4 结束语

提出了一种基于改进 ResUnet 的表面缺陷分割算法, 该方法以 ResUnet 网络结构为基础架构, 使用先进的 ConvNeXt-T 改进骨干网络, 提取更具代表性的特征; 在解码阶段设计并添加全局上下文模块, 增强模型的全局上下文建模能力; 同时融合解码阶段的多级多尺度特征, 使模型更适应缺陷尺寸和大小多变的钢铁表面缺陷, 提升缺陷识别精度。在谢韦尔钢铁公司提供的钢铁表面缺陷数据集上进行实验, 通过定量和定性分析的方法验证了所提缺陷识别算法的有效性, 与对比方法相比, 所提方法的值达到了最高的 0.778 4, 且单张图像在 GPU 上的运行时间只有 14.7 ms, 同时通过主观视觉效果分析, 验证了所提方法的缺陷分割结果与标签最接近, 说明所提方法具有较好的鲁棒性、准确性和高效性。

参考文献:

[1] CAO J, YANG G, YANG X. A pixel-Level segmentation

convolutional neural network based on deep feature fusion for surface defect detection [J]. *Ieee Transactions On Instrumentation and Measurement*, 2021, 70: 1–12.

- [2] YANG J, LI S, WANG Z, et al. Using deep learning to detect defects in manufacturing: A comprehensive survey and current challenges [J]. *Materials*, 2020, 13 (24): 5755.
- [3] CAO J, ZHANG J, WEN Z, et al. Fabric defect inspection using prior knowledge guided least squares regression [J]. *Multimedia Tools and Applications*, 2015, 76 (3): 4141–4157.
- [4] AMINZADEH M, KURFESS T. Automatic thresholding for defect detection by background histogram mode extents [J]. *Journal of Manufacturing Systems*, 2015, 37: 83–92.
- [5] ZHOU X, WANG Y, XIAO C, et al. Automated visual inspection of glass bottle bottom with saliency detection and template matching [J]. *IEEE Transactions On Instrumentation and Measurement*, 2019, 68 (11): 4253–4267.
- [6] XI J, SHENTU L, HU J, et al. Automated surface inspection for steel products using computer vision approach [J]. *Applied Optics* (2004), 2017, 56 (2): 184–192.
- [7] ZHAO J, PENG Y, YAN Y. Steel surface defect classification based on discriminant manifold regularized local descriptor [J]. *IEEE Access*, 2018, 6: 71719–71731.
- [8] HU H, LIU Y, LIU M, et al. Surface defect classification in large-scale strip steel image collection via hybrid chromosome genetic algorithm [J]. *Neurocomputing*, 2016, 181: 86–95.
- [9] QIAN K. Automated detection of steel defects via machine learning based on real-time semantic segmentation [C] // *Proceedings of the 3rd International Conference on Video and Image Processing*, 2019: 42–46.
- [10] NIU M, SONG K, HUANG L, et al. Unsupervised saliency detection of rail surface defects using stereoscopic images [J]. *IEEE Transactions on Industrial Informatics*, 2020, 17 (3): 2271–2281.
- [11] YUN J P, SHIN W C, KOO G, et al. Automated defect inspection system for metal surfaces based on deep learning and data augmentation [J]. *Journal of Manufacturing Systems*, 2020, 55: 317–324.
- [12] 李 原, 李燕君, 刘进超, 等. 基于改进 Res-UNet 网络的钢铁表面缺陷图像分割研究 [J]. *电子与信息学报*, 2022, 44 (5): 1513–1520.
- [13] NEVEN R, GOEDEME T. A multi-branch U-Net for steel surface defect type and severity segmentation [J]. *Metals*, 2021, 11 (6): 870.
- [14] ZHENG Z, HU Y, ZHANG Y, et al. CASPPNet: A chained atrous spatial pyramid pooling network for steel

defect detection [J]. Measurement science and technology, 2022, 33 (8): 085403..

[15] ZHANG Z, LIU Q, WANG Y. Road extraction by deep residual U-Net [J]. IEEE Geoscience and Remote Sensing Letters, 2018, 15 (5): 749–753.

[16] 程镨杰, 杨 耘, 李龙威, 等. 基于深度可分离卷积的轻量化残差网络高光谱影像分类 [J]. 光学学报, 2023: 1–20.

[17] 魏 欣, 李 镨, 关 欣. 基于 3D U-Net 的轻量级脑肿瘤分割网络 [J]. 光电子·激光, 2022, 33 (12): 1338–1344.

[18] LIU Z, MAO H, WU C, et al. A convnet for the 2020s [C] //Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 11976–11986.

[19] LIU Z, LIN Y, CAO Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows [C] //

Proceedings of the IEEE/CVF international conference on computer vision, 2021: 10012–10022.

[20] DEMIR K, AY M, CAVAS M, et al. Automated steel surface defect detection and classification using a new deep learning-based approach. NEURAL COMPUTING & APPLICATIONS, 2022, 35 (11): 8389–8406.

[21] CAO Y, XU J, LIN S, et al. Gcnet: Non-local networks meet squeeze-excitation networks and beyond [C] //Proceedings of the IEEE/CVF international conference on computer vision workshops, 2019.

[22] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation [C] //Proceedings of the IEEE conference on computer vision and pattern recognition, 2015: 3431–3440.

[23] 陈德阳, 唐 智, 何牧耕. 基于 OCR-UNet 的金属表面缺陷分割 [J]. 组合机床与自动化加工技术, 2023 (11): 169–173.

(上接第 167 页)

时复用、设置串口中断等方法实现了基于 PXI 总线架构的靶弹测试发射控制装备一体化设计，在控制硬件投入规模基础上，设计了软硬件多重安全控制和飞行器模飞动态测试功能，提升了装备使用效益，可为靶弹技术阵地准备和发射控制提供有力支撑。实际应用结果表明，本系统方案能够有效保障靶弹的全系统测试、射前检查、模拟/实弹发射任务，系统运行稳定、安全、可靠。

参考文献：

[1] 周旦辉, 等. 靶弹系统发展的某些新特点 [J]. 现代防御技术, 2007, 35 (2): 27–30.

[2] 闫淑群, 黎玉刚, 等. 基于 PXI 总线的导弹自动测试系统设计 [J]. 计算机测量与控制, 2011, 19 (8): 1025–1028.

[3] 曹怀志, 张 安. 空射巡航导弹火控系统总体构思 [J]. 电光与控制, 2005, 12 (1): 22–25.

[4] 开湘龙. 基于 PXI 总线的导弹测试系统研制 [D]. 哈尔滨工程大学, 2014: 2–5.

[5] 王 储. 某型号导弹测发控系统研制 [D]. 哈尔滨: 哈尔滨工业大学, 2013: 2–7.

[6] 毛 宁, 刘艳华, 马丽媛. 陆军武器火控系统的发展趋势 [J]. 火力与指挥控制, 2016, 41 (8): 6–9.

[7] 高 亢, 齐 鑫, 纪 浩. 基于 PXI 总线的导弹自动测试系统设计 [J]. 工业控制计算机, 2018, 31 (11): 1–2.

[8] 王 旭, 荣 刚, 等. 低成本高集成一体化测发控系统设计研究 [J]. 宇航计测技术, 2023, 43 (6): 76–82.

[9] 赵 明, 王 雷. 地(舰)空导弹发射控制系统安全性设

计 [J]. 现代防御技术, 2016, (44) 3: 179–185.

[10] 种 山, 杜 江. 导弹动态测试技术及应用 [J]. 现代防御技术, 2009, (37) 3: 41–43.

[11] 樊水康, 郭会兵, 等. PXI 计算机在一体化多模式发射控制系统中的应用 [J]. 火力与指挥控制, 2013, 38 (7): 138–140.

[12] 翟 磊. 空空导弹发射控制技术 [J]. 制导与引信, 2015, 36 (4): 6–8.

[13] 刘 佳, 樊水康. 一种火箭弹发射电路设计 [J]. 火力与指挥控制, 2014, 39 (s): 131–133.

[14] National instruments corporation. NI-SWITCH User Manual [Z]. 2024: 980–992.

[15] 张学英, 易 航, 等. 运载火箭测发控系统通用化设计 [J]. 导弹与航天运载技术, 2012, 4: 15–19.

[16] 刘 栋, 李兴城, 等. 多管火箭炮地面发射控制装置 [J]. 兵工自动化, 2015, 34 (1): 12–14.

[17] National instruments corporation. NI-DAQmx [Z]. 2021.

[18] 刘君华. 虚拟仪器语言 LabWindows/CVI 教程 [M]. 北京: 电子工业出版社, 2001.

[19] 段亚博, 等. 一种软硬结合确保火工品测试安全的方法 [J]. 兵工自动化, 2023, 6 (42): 27–29.

[20] 周金芳, 李兴鹏, 陈 浩, 等. 某型导弹自动驾驶仪动态测试系统的设计 [J]. 电子产品世界, 2017, 2–3 合刊: 56–59.

[21] 杜 江, 付京来, 郑建辉. 基于测发一体化技术的导弹发射控制系统设计 [J]. 计算机测量与控制, 2010, 18 (3): 605–607.

[22] National instruments corporation. LabWindows/CVI Real-Time Module [Z]. 2013.