

文章编号: 1006-3080(2025)02-0242-08

DOI: 10.14135/j.cnki.1006-3080.20240305003

基于改进 U-Net3+的相控阵超声图像语义分割

毛鑫玥, 王慧锋, 周家乐, 顾震, 颜秉勇

(华东理工大学信息科学与工程学院, 上海 200237)

摘要: 超声相控阵成像已广泛应用于聚乙烯燃气管道的焊接缺陷检测中, 随着机器视觉技术的快速发展, 利用机器辅助或自动化分析超声图像能极大地提高缺陷检测速度, 减少人为判断失误的发生。在基于超声图像的焊接缺陷检测技术中, 图像语义分割精度对缺陷类别和严重等级的判定至关重要。本文在 U-Net3+网络的基础上提出一种融入残差及注意力机制的改进模型, 并应用于电熔焊接缺陷检测的相控阵超声图像语义分割。首先, 改进模型通过在编码器各层之间采用残差结构来提升编码器的图像特征提取能力; 其次, 通过在跳跃连接中引入卷积块注意力模块 (Convolutional Block Attention Module, CBAM), 加强模型对原始图像信息的利用率, 使模型更易聚焦于原始图像中的有效区域。实验结果表明, 改进后的模型在电熔焊接超声图像上具有良好的分割效果, 在 Dice、mIoU 两项指标上, 相比 U-Net 分别提升了 8.81% 和 12.84%; 相比 U-Net3+ 的分割效果分别提升了 1.09% 和 1.81%。

关键词: 相控阵超声图像; 图像语义分割; U-Net3+; 注意力机制; 残差网络

中图分类号: TP391.4

文献标志码: A

聚乙烯 (Polyethylene, PE) 管道以其耐腐蚀、韧性好等优点, 在燃气行业得到广泛的应用, 已成为城市燃气管道的首选^[1], 然而燃气泄漏事件时有发生, 严重威胁着人民的生命财产安全。其中, 燃气管道的焊接缺陷是造成燃气泄漏的主要因素。电熔焊接方式是燃气管道的主要连接方式之一, 常用于连接 DN90 以下 (公称直径小于 90 mm) 的聚乙烯管道^[2]。超声相控阵检测已广泛应用于聚乙烯燃气管道的焊接缺陷检测中。相控阵探头发射的超声波, 在传播过程中遇到不同的声阻抗界面时, 会出现一定声强的反射信号, 通过解码反射信号, 就能重构出一张反映管材内部结构的超声图像^[3]。

传统人工检测方法中, 检测人员通过测量、分析超声图像中特征线、电阻丝线和底面回波线等不同类型区域的位置、形态可以判别出焊接处内部是否存在冷焊/过焊、夹杂、错位、孔洞等常见缺陷和缺陷等级^[4]。但人工判断焊接缺陷的方法面临主观性强、

测量操作不规范、耗时严重等问题。

图像语义分割是计算机视觉的一个重要领域, 通过对图像中的每个像素进行分类, 能有效识别出不同物体的位置和形态, 被广泛应用于物体分类^[5]、场景理解^[6]、缺陷检测^[7]等任务中, 并取得了良好效果。通过对电熔焊接超声图像进行语义分割, 能够快速把特征线、电阻丝线、底面回波线等不同区域标记出来, 方便计算机后续计算和判断不同区域的位置和形态, 从而实现智能缺陷检测, 提高检测速度, 减少人工判断错误的目的。提高图像语义分割的精确性和鲁棒性成为电熔焊接智能缺陷检测的关键。

U-Net^[8] 是用于图像语义分割的经典网络结构, 跳跃连接是其标志性结构, 擅长捕捉上下文信息, 适用于高分辨率的精确分割, 在 2015 年的 ISBI 细胞跟踪挑战中, U-Net 以巨大的优势击败了其他新兴技术, 此后便在医学领域被广泛应用。2018 年, Oktay 等^[9] 设计了注意力门控结构, 提出 Attention U-Net,

收稿日期: 2024-03-05

基金项目: 青年科学基金项目 (61906068)

作者简介: 毛鑫玥 (1999—), 女, 河南郑州人, 硕士生, 主要研究方向为计算机视觉。E-mail: y30210935@mail.ecust.edu.cn

通信联系人: 周家乐, E-mail: zhou.jiale@ecust.edu.cn

引用本文: 毛鑫玥, 王慧锋, 周家乐, 等. 基于改进 U-Net3+的相控阵超声图像语义分割 [J]. 华东理工大学学报 (自然科学版), 2025, 51(2): 242-249.

Citation: MAO Xinyue, WANG Huifeng, ZHOU Jiale, et al. Semantic Segmentation of Phased Array Ultrasound Images Based on Improved U-Net3+[J]. Journal of East China University of Science and Technology, 2025, 51(2): 242-249.

取得了较好的效果,成为医学图像分割领域中的经典模型。同在2018年提出的U-Net++^[10]则针对经典U-Net的跳跃连接机制进行了扩增,在解码器部分将来自深层与浅层编码器的不同尺度特征进行了整合,提高了网络抓取特征的能力。2020年提出的U-Net3+^[11]则进一步针对U-Net++参数量大、计算效率低的问题,去掉了U-Net++中使用的稠密卷积块,提出全尺度跳跃连接,使模型结构更简洁。与此同时,基于U-Net的改进网络模型也不断被提出,应用于不同任务中。2022年,胡剑锋等^[12]提出了一种使用残差模块改进U-Net得到的ResU-Net模型进行焊接图像的语义分割,实现了焊接缺陷的自动识别;Zhou等^[13]提出一种使用通道融合注意力模块(CFAM)和空间群注意力模块(SGAM)两种注意力改进的CFSG U-Net,对肋骨的CT图像进行检测,得到了粒度更细的分割结果,可用于辅助医生进行骨折诊断;2024年,王艺妮等^[14]将Swin-Transformer模块和混合注意力机制引入U-Net的跳跃连接中,提出了DH-Swin U-Net,对人体关节的医学图像进行语义分割,在真实医学图像数据集中提高了分割精度。

除了U-Net家族外,也有其他语义分割经典模型常被使用,如DeepLab^[15-17]系列、SegNet^[18]等,其中DeepLab系列为深度卷积网络结构,使用空洞金字塔池化(ASPP)模块捕捉不同尺度信息,模型复杂度较高,常用于大规模场景的语义分割;而SegNet使用编码器-解码器结构,但未使用跳跃连接,模型较为简单,适用于计算条件受限的场合,对图像细节的捕捉能力不如U-Net网络家族。因此,本文选择U-Net网络家族作为基础模型进行进一步的改进研究。

虽然各种经典网络框架及其改进算法在近年来得到了极大的发展,但由于电熔焊超声图像存在不同部分颜色差异大、边缘不清晰、形状复杂等问题,使用传统的U-Net难以得到高精确度的语义分割图像。因此,针对以上问题,本文基于U-Net家族网络中的U-Net3+模型,提出了一种融入残差结构与注意力机制的改进模型,能够有效提高电熔焊接超声图像语义分割的精确度。

1 融入残差及注意力的U-Net3+模型

电熔焊接的超声图像主要由特征线、电阻丝线和底面回波线3个部分组成,它们的特征信息是判断管道内部焊接接面是否存在缺陷的依据。扫描得到的焊接超声图像示例及各主要特征区域名称如图1所示。

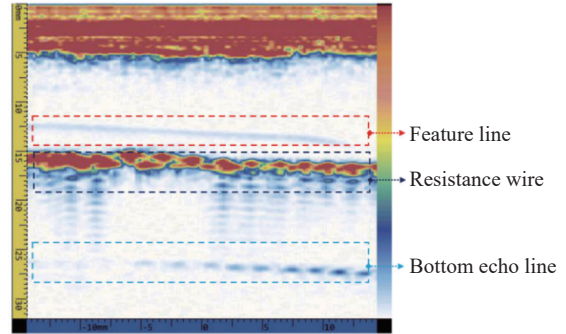


图1 相控阵超声图像

Fig. 1 Phased array ultrasound images

1.1 语义分割网络的整体架构

传统的U-Net结构由编码器-解码器和跳跃连接构成,具体的网络深度可以根据使用场景进行改动,具有一定的灵活性。每层编码器输出的特征图输入对应的解码器,与上一层解码器的输出特征图融合,即跳跃连接,该U-Net的标志性结构为它提供了捕捉丰富的多尺度上下文信息的能力。U-Net++、U-Net3+强化了这种结构,将跳跃连接扩展到所有编码器与解码器之间,使用跨层的全尺度跳跃连接进一步提升了该长处。

本文提出的电熔焊超声图像语义分割模型结构如图2所示。基于U-Net的编码器-解码器框架和跳跃连接结构,我们提出了一种由残差模块和注意力机制改进的网络模型,在编码器部分引入残差结构,将普通卷积块替换为残差块组,残差块为网络层数的加深提供了基础,使得算法能够通过更深层的卷积网络提取图像特征的细节。同时,我们在跳跃连接部分选择性地引入注意力机制,平衡网络对图像局部特征的捕捉能力和对全局信息的感知能力。

图2中, $X_e(i=1,2,3,4,5)$ 表示第*i*层编码器生成的特征图, $X_d(i=1,2,3,4)$ 表示第*i*层解码器生成的特征图,编码器与解码器之间的箭头表示跳跃连接,其中虚线为使用了注意力的跳跃连接,实线为直接融合的跳跃连接,不同尺寸的特征图使用上/下采样进行匹配。编码器之间的卷积层改为一组残差块,4个残差块组分别由3、4、6、4个基本残差块构成。

1.2 残差结构

在深度神经网络中,层数的加深通常能够提升网络提取特征的能力,取得更好的结果,但在网络深度增加到一定程度后,网络会出现梯度消失、性能退化等现象,准确度不再增加甚至下降。经典语义分割框架U-Net的解码器和编码器通常各有4~5层,如果为提升准确率单纯地增加编码器深度,也存在着这样的问题。本文中,在U-Net中引入残差结构的具

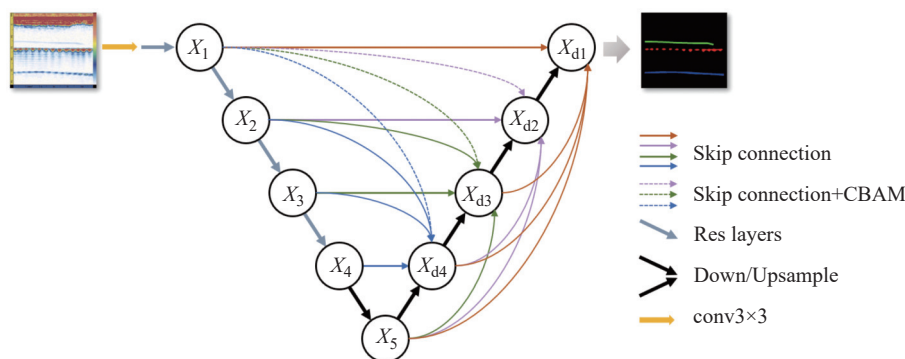


图 2 融入残差及注意力的 U-Net3+模型

Fig. 2 Improved U-Net3+model based on residual and attention

体方式是将 U-Net 编码器部分的卷积块替换为残差块^[19]。残差模块提出的快捷连接 (Shortcut connection), 其基本原理在于将每个残差块的输出与输入相加, 形成一个“残差学习”框架, 能够更有效地使模型学习深层的特征表示, 使更深、更复杂的神经网络成为可能。

残差部分由一系列残差模块组成。对每一个残差块, 其输入和输出关系如下:

第 $l+1$ 层的输出 x_{l+1} 为:

$$x_{l+1} = x_l + F(x_l, W_l) \quad (1)$$

其中, x_l 为第 l 层残差块的输出, $F(\cdot)$ 为残差部分。

将多个残差块连接起来, 就构成了一个残差块组。对于包含多层结构的残差块组, 从浅层 l 到深层 L 残差块的输出 x_L 为:

$$x_L = x_l + \sum_{i=l}^{L-1} F(x_i, W_i) \quad (2)$$

残差块组中通过一次步长为 2 的卷积来改变张量尺寸, 实现降采样。进行降采样时, 为了与主路特征图的尺寸对齐, 快捷连接支路上也要进行降采样。将若干残差块和一个降采样残差块叠加, 就构成了一个残差层, 在本文中用在两层编码器之间, 进行特征提取。图 3 示出了残差模块的结构。

1.3 注意力模块

注意力机制常被用于卷积神经网络的性能优化^[20-22], 它模仿生物系统中选择性地关注和处理信息的过程, 通过对神经网络提取到的特征图赋予权重, 加强对关键信息的提取能力。本文用于改进 UNet3+ 模型的卷积注意力模块 (Convolutional Block Attention Module, CBAM) 通过融合空间维度和通道维度的注意力权重, 来抑制不重要的信息, 提高语义分割结果的精确度。模块结构如图 4 所示, 输入特征图分别经过全局最大池化与全局平均池化, 分别从空间维度和通道维度被压缩, 得到两个不同维度的特征描述。

CBAM 模块^[23]主要包含通道注意力 (Channel attention module) 和空间注意力 (Spatial attention module) 两个部分。这两个模块可以并行连接或顺序连接, 研究表明顺序连接的效果更好^[23]。在 CBAM 模块的运算过程中, 大小为 $H \times W \times C$ 的特征图 F 输入后, 先分别经过最大池化和平均池化, 对输入特征在空间维度上进行压缩, 得到尺寸为 $1 \times 1 \times C$ 的通道注意力图, 再经过共享的卷积层, 将两个输出相加并取 Sigmoid 后, 即得到通道注意力权重值。通道注意力模块的输入输出关系可表达为:

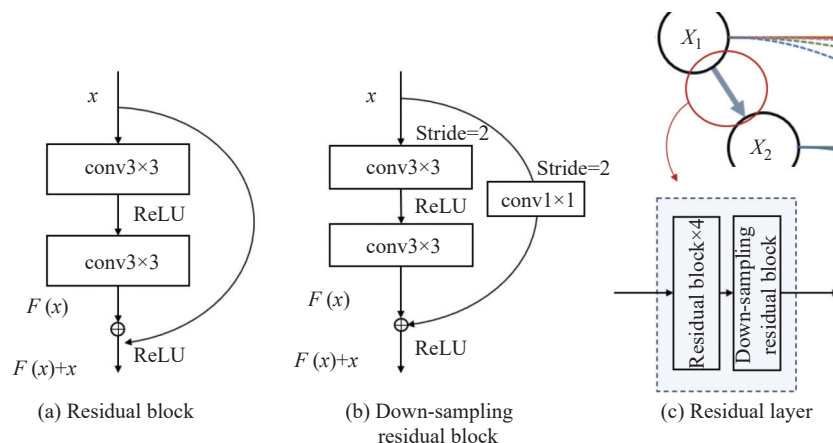


图 3 残差模块结构

Fig. 3 Structure of residual block

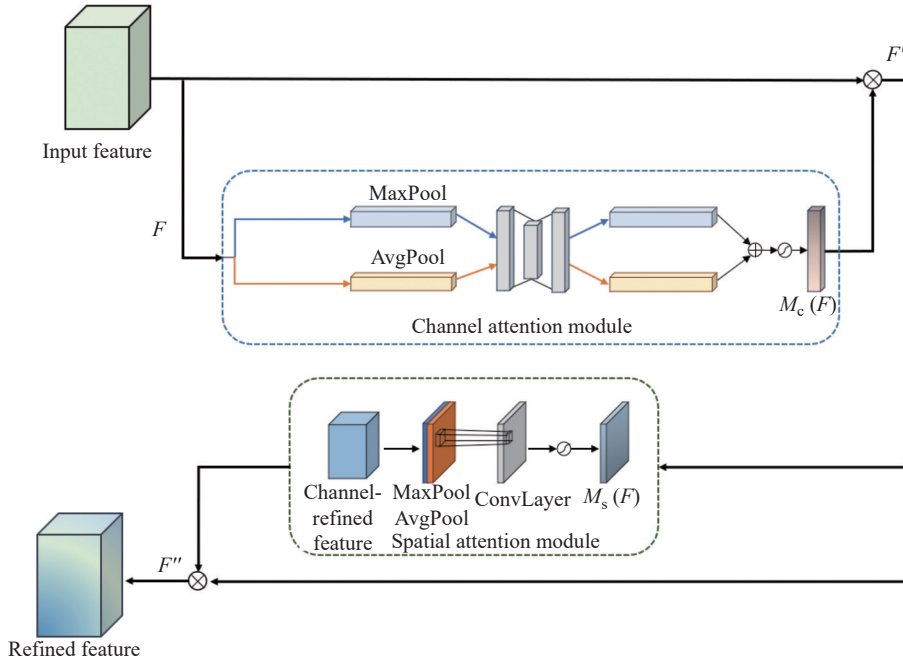


图4 CBAM注意力模块结构

Fig. 4 Structure of CBAM attention block

$$M_c(F) = \sigma\{\text{MLP}[\text{AvgPool}(F)] + \text{MLP}[\text{MaxPool}(F)]\} \quad (3)$$

本文所使用的MLP连接为两个 1×1 的卷积层。

通道注意力权重 M_c 与输入特征相乘后,得到通道注意力模块的输出 F' 。 F' 作为空间注意力模块的输入,经过平均池化、最大池化和卷积层,生成一个尺寸为 $1 \times H \times W$ 的空间注意力权重 M_s ,用于编码需要聚焦的空间位置。空间注意力机制可表达为:

$$M_s(F') = \sigma(f^{7 \times 7}([\text{AvgPool}(F'); \text{MaxPool}(F')])) \quad (4)$$

其中, σ 表示Sigmoid操作, 7×7 表示卷积核的大小。

空间注意力权重 M_s 与 F' 相乘后,输出CBAM模块的最终输出 F'' 。

CBAM模块结合了通道注意力和空间注意力的优势,经过它处理的特征图能够聚焦于图像的有效部分,从通道和空间位置两个角度抑制多余信息,提高网络提取特征的效率。本文将应用于不同尺度的跳跃连接结构中,用于增强网络提取有效特征的效率。

2 实验设计与分析

2.1 实验环境设置

本实验软件环境:编译环境为Python3.9.12、OpenCV 4.7.0.72、torch 1.12.1、CUDA 11.6,操作系统为Windows 10;硬件环境:Intel Core i9-10900KF CPU及NVIDIA GeForce RTX 3090 24G显卡。

2.2 数据集与评估指标

2.2.1 实验数据及预处理 实验所用的图像数据来源于真实的电熔焊工业应用环境,图像由Phascan超声相控阵仪器检测电熔焊接接头的管道焊接面得到,图像大小为 513×447 。训练集共200张,其中包括正常类型的数据49张,冷焊缺陷类型的数据41张,过焊缺陷类型数据19张,孔洞缺陷类型的数据28张,电阻丝错位缺陷类型的数据33张,熔合面夹杂缺陷类型的数据30张;测试集共61张,其中正常类型的数据15张,冷焊缺陷类型的数据14张,过焊缺陷类型数据4张,孔洞缺陷类型的数据16张,电阻丝错位缺陷类型的数据5张,熔合面夹杂缺陷类型的数据7张。数据集中各类型的图像如图5所示。

2.2.2 数据预处理 由于直接从工业生产中得到的图像样本较少,为丰富数据样本,向训练提供足够的支撑,需要对数据集进行数据增强预处理。常用的图像数据增强方法包括旋转、翻转、亮度变化、加入噪声等。本文主要使用水平翻转、加入高斯噪声、加入椒盐噪声这3种方式,其中椒盐噪声的添加比例为0.3。扩增后的训练集共有800张图片。数据增强结果如图6所示。

电熔焊超声图像的结构在水平方向上基本具有一致性,因此可进行水平翻转,而不破坏其图像信息的合理性。在实际生产中,扫描成像不一定具有很高的清晰度,因此需要考虑算法对噪声的鲁棒性,所以使用椒盐噪声和高斯噪声增强训练数据。

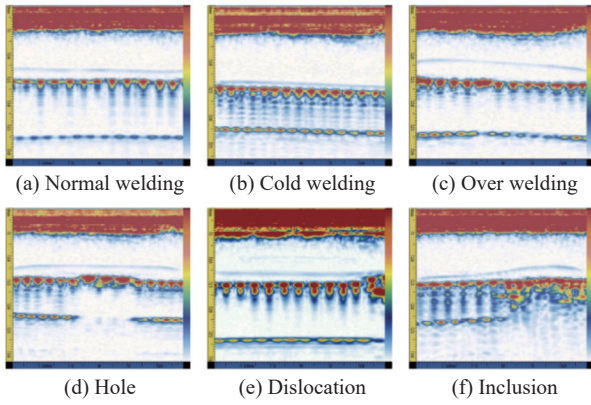


图 5 不同种类电熔焊缺陷的超声图像数据集

Fig. 5 Ultrasonic image dataset of different welding defects

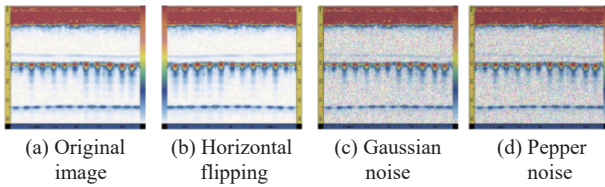


图 6 数据增强结果图

Fig. 6 Images of data enhancement result

2.2.3 评价指标 本实验采用两个广泛应用于语义分割的衡量指标,平均交并比 (mIoU)、Dice 系数作为结果的评价指标。两个指标的取值都在 0~1 之间,越接近 1 表明结果与标签越相似,分割效果越好。mIoU 及 Dice 系数的数学表达式分别如式 (5)、(6) 所示:

$$\text{mIoU} = \frac{1}{k+1} \sum_{i=0}^k \frac{\text{TP}_i}{\text{TP}_i + \text{FN}_i + \text{FP}_i} \quad (5)$$

$$\text{Dice} = \frac{2 \times \text{TP}}{2 \times \text{TP} + \text{FP} + \text{FN}} \quad (6)$$

其中, k 为样本总数, TP 表示被正确分割的像素数量, FP 表示被错误分割的像素数量, FN 表示被错误

预测背景的像素数量。

2.2.4 超参数设置 训练采用 Adam 优化器,学习率设置为 0.001,在训练过程中自适应调整; batch size 设置为 1,训练轮数为 20 个 epoch。选择二元交叉熵损失函数 (BCEWithLogitsLoss) 作为模型训练的损失函数。

2.3 结果与分析

2.3.1 对比实验 在相同的实验条件下,将本文模型运行得到的分割结果与其他常用的分割模型 U-Net、Attention U-Net、U-Net3+ 的运行效果进行了对比。表 1 中列出了实验的数据指标结果对比,图 7 中列出了图像分割结果对比。

表 1 不同模型在电熔焊超声图像数据集上的结果对比

Table 1 Comparison of different models on the ultrasonic image dataset of electrofusion welding

Method	Dice/%	mIoU/%
U-Net	82.05	70.57
Attention U-Net	87.03	77.26
U-Net3+	89.77	81.60
U-Net3+ + Res+ CBAM	90.86	83.41

从表 1 中可以看出,在电熔焊超声图像数据集中,本文提出的模型得到的 Dice 和 IoU 参数分别为 90.86% 和 83.41%,取得了较为良好的结果,在两个参数上都有提升。由于在编码器阶段引入了残差模块,并采用了注意力机制和多层次跳跃连接,模型提取边缘细节的能力得到增强,并在类别区分的准确性上有所提高,能够定位超声图像的不同区域,分类结果更接近标签图像。相比于未经改进的经典模型 U-Net,本文算法在 Dice 系数和交并比分别提高了

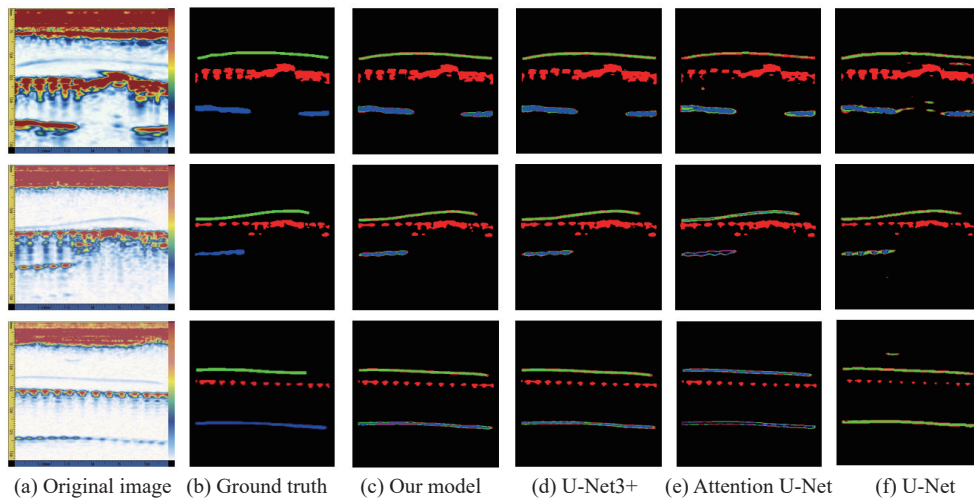


图 7 图像分割结果对比

Fig. 7 Comparison of segmentation results

8.81% 和 12.84%; 相比于经典的注意力改进 U-Net 模型 Attention U-Net, 本文算法在两个参数上分别提高了 3.83% 和 6.15%; 相比于 U-Net3+, 本文算法分别提高了 1.09% 和 1.81%, 这说明本文模型加强了网络对超声图像细节的提取能力, 提高了算法分割精度。

可视化结果及局部细节对比如图 7、图 8 所示。从图中可以看出, 本文提出的模型得到的分割结果边界更光滑, 形状更准确, 并且对比 U-Net 和 Attention U-Net, 在像素类别的判断上更加准确, 说明改进后的模型通过对空间信息的利用, 提高了图像

中上下文信息传递的有效性。

从图 8 中可以看出, 本文模型得到的语义分割图在结构细节处噪点少, 形状清晰, 能够准确标记出细小结构, 这有助于在后续将输出图像用于缺陷的判断中时提升判断的准确度(如孔洞缺陷的判断需要计算连通域个数, 电阻丝错位的判断需要计算各连通域的质心坐标, 准确的结构划分能够为此提供判断基础)。对于成像质量较差, 较为模糊的图像, 本文的方法也能够抵抗噪声, 得出达到使用要求的图像结果。

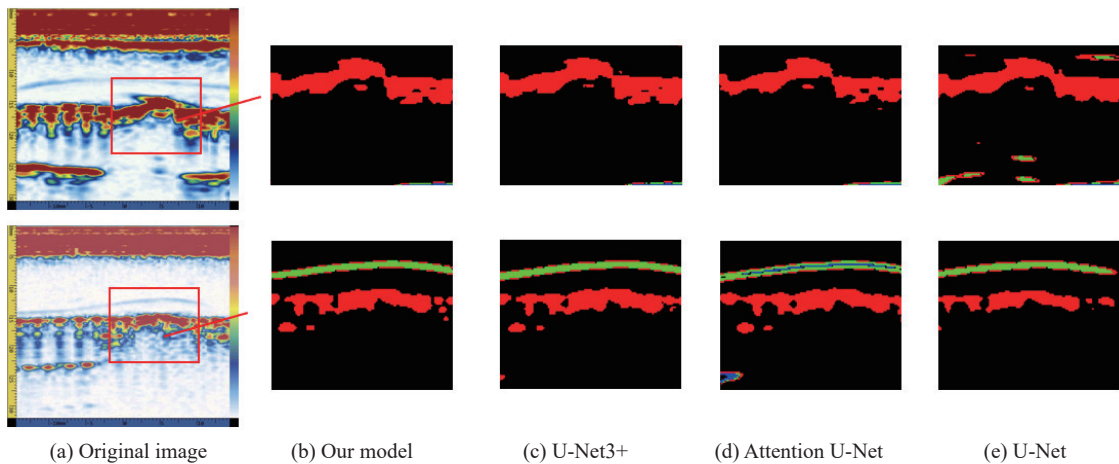


图 8 分割结果局部细节

Fig. 8 Local details of segmentation results

2.3.2 消融实验 为验证本文模型所采用的模块的有效性, 在相同的实验参数设置下设计并进行了 5 组消融实验。通过对消融实验结果的对比, 发现多层次跳跃连接和 CBAM 注意力的引入都能够直接提升网络性能, 并且在与残差模块共同使用时有更好的效果, 不同组合的数据指标对比如表 2 所示。CBAM 模块使网络侧重于目标对象区域, 对来自背景的干扰信号有更强的稳定性。多层次跳跃连接结构则对解码器部分提供了多尺度的特征, 增强模型对全局信息的感知。残差层的引入加深了编码器部分的网络深度, 在单独使用时, 网络层数的叠加使提取到的特征偏重局部信息, 但与 CBAM 注意力共同使用时, 综合了高维特征和全局上下文信息, 得到了比单独使用 CBAM 更大的效果提升。

2.3.3 特征图分析 通过特征图可视化, 对注意力模块在空间信息上的过滤作用进行分析, 各阶段图像如图 9 所示。其中, 图 9(a) 为编码器输出 X_1 经下采样后, 输入各跳跃连接支路的特征图像, 图 9(b) 为注意力权重图, 图 9(c) 为注意力模块处理后的输出特征图。图中较亮的像素表示该点权值较高。观察经

注意力过滤前的特征信息可以看出, 残差层和降采样得到的特征图基于原始图像的像素强度, 大致表现出了特征线、电阻丝线、底面回波线 3 个区域的分层, 如图 9(a) 所示, 但在这 3 个有效区域之外, 其中也包含了来自背景和超声扫描边界的干扰信息, 并且提取到的边界较为粗糙。CBAM 注意力模块学习到的权重如图 9(b) 所示。由图 9(c) 可看出, 所生成的注意力权重图比起输入特征图, 具有更精细的结构, 表明注意力机制根据位置而非单纯的像素数值关注到了重要特征, 使过滤后的输出特征变得更精细、更明确, 边缘更清晰的同时降低了周边背景噪声

表 2 不同模块的消融实验

Table 2 Ablation experiments on different modules

Method	Dice/%	mIoU/%
U-Net3+	89.77	81.60
U-Net3+ + Res	88.32	79.58
U-Net + Res + CBAM	90.13	82.19
U-Net3+ + CBAM	90.30	82.49
U-Net3+ + Res+ CBAM	90.86	83.41

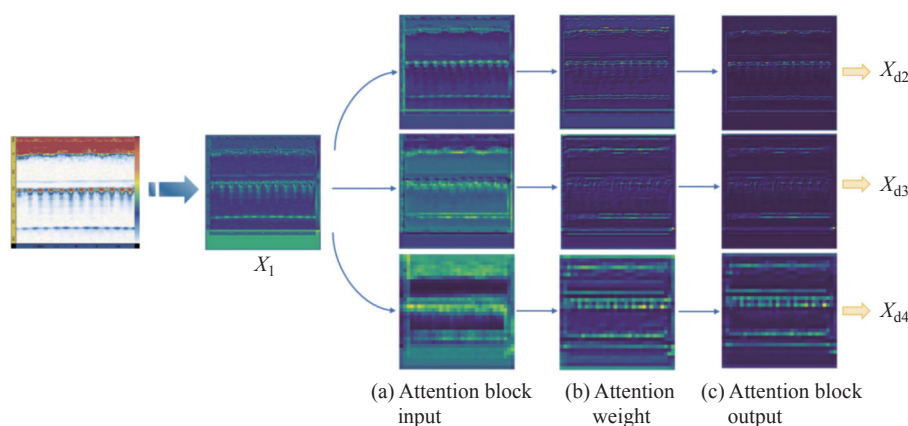


图 9 注意力模块输入/输出可视化

Fig. 9 Visualization of attention block input/output

的影响。

3 结束语

为进一步提升电焊超声图像的语义分割精度,保障缺陷类别分类的准确度,本文在 U-Net3+网络基础上提出一种融入残差及注意力机制的改进模型,在编码器各层之间采用残差结构,并在跳跃连接中引入 CBAM。与传统的 U-Net相比,多层次的跳跃连接增强了模型对多尺度特征的融合能力;与 U-Net3+相比,本文提出的模型所引入的残差结构和 CBAM 注意力机制提升了模型对局部细节的提取能力。实验结果表明,改进后的模型在电熔焊接超声图像上具有良好的分割效果,相比其他模型,在分割精度上得到了提升,验证了所提出的改进方法的有效性。然而,超声图像中仍存在可供继续发掘的信息,如图形各区域之间的位置关系等,未来将探索并研究更有效的模型设计方法,在保障分割精确度的同时控制参数量,提升计算速度,增强模型实用性。

参考文献:

- [1] 华晔,武志军,熊志敏,等. 聚乙烯 (PE) 燃气管道近期发展动态 [J]. 全面腐蚀控制, 2014, 28(4): 34-40, 44.
- [2] 高加辉,郭宇强,陈允杰. 浅谈聚乙烯 (PE) 管在城市燃气管网中的应用 [C]//中国燃气运营与安全研讨会 (第十一届)暨中国土木工程学会燃气分会 2021 年学术年会论文集 (上册). 中国: 中国土木工程学会燃气分会, 2021: 795-798.
- [3] 伏喜斌. 聚乙烯管道电熔焊接及超声相控阵检测技术 [J]. 焊接, 2012(7): 49-51, 71.
- [4] 黄跃鑫,钟舜聪,伏喜斌,等. 聚乙烯管道电熔接头的超声相控阵成像及缺陷特征 [J]. 焊接学报, 2018, 39(2): 119-123, 134.

- [5] RICARDO P, TIAGO B, LUÍS G, *et al.* A deep learning-based global and segmentation-based semantic feature fusion approach for indoor scene classification[J]. *Pattern Recognition Letters*, 2024, 179: 24-30.
- [6] 李钰,袁晴龙,徐少铭,等. 基于感知注意力和轻量金字塔融合网络模型的室内场景语义分割方法 [J]. 华东理工大学学报 (自然科学版), 2023, 49(1): 116-127.
- [7] LU Q, LIN J, LUO L, *et al.* A supervised approach for automated surface defect detection in ceramic tile quality control[J]. *Advanced Engineering Informatics*, 2022, 53(2): 101692.
- [8] RONNEBERGER O, FISCHER P, BROX T, U-Net: Convolutional networks for biomedical image segmentation[C]// *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*. Cham: Springer, 2015: 234-241.
- [9] OKTAY O, SCHLEMPER J, FOLGOC L L, *et al.* Attention U-Net: Learning where to look for the pancreas[EB/OL]. (2018-4-11) [2018-5-29]. <https://doi.org/10.48550/arxiv.1804.03999>.
- [10] ZHOU Z, SIDDIQUEE M M R, TAJBAKHS N, *et al.* UNet++: A nested U-Net architecture for medical image segmentation[EB/OL]. (2018-7-18)[2018-12-10]. <https://arxiv.org/pdf/1807.10165>.
- [11] HUANG H M, LIN L F, TONG R F, *et al.* UNet 3+: A full-scale connected UNet for medical image segmentation[C]// *2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Spain: IEEE, 2020: 1055-1059.
- [12] 胡剑峰,周家乐,王慧锋,等. 一种电熔焊接缺陷自动识别方法的研究与应用 [J]. 华东理工大学学报 (自然科学版), 2022, 48(4): 526-533.
- [13] ZHOU Z Y, FU Z H, JIA J C, *et al.* Rib fracture detection with dual-attention enhanced U-Net[J]. *Computational and Mathematical Methods in Medicine*, 2022(2022): 8945423.
- [14] 王艺妮,时宏伟. 基于 DH-Swin Unet 的医学图像分割算

- 法[J]. 计算机系统应用, 2024, 33(3): 202-212.
- [15] CHEN L C, PAPANDREOU G, KOKKINOS I, *et al.* Semantic image segmentation with deep convolutional nets and fully connected CRFs[J]. (2014-12-24)[2015-2-28]. <https://doi.org/10.48550/arxiv.1412.7062>.
- [16] CHEN L C, PAPANDREOU G, KOKKINOS I, *et al.* DeepLab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 40: 834-848.
- [17] CHEN L C, PAPANDREOU G, SCHROFF F, *et al.* Rethinking atrous convolution for semantic image segmentation[J]. (2017-6-17)[2017-12-5]. <https://arxiv.org/abs/1706.05587>.
- [18] BADRINARAYANAN V, KENDALL A, CIPOLLA R. SegNet: A deep convolutional encoder-decoder architecture for image segmentation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 39: 2481-2495.
- [19] HE K M, ZHANG X Y, REN S Q, *et al.* Deep residual learning for image recognition[C]// 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). USA: IEEE, 2016: 770-778.
- [20] QU T Q, SUN Q, QIAN Z, *et al.* Fault diagnosis for PV arrays considering dust impact based on transformed graphical features of characteristic curves and convolutional neural network with CBAM modules[J]. Applied Energy, 2024, 355: 1-29.
- [21] LYU Z, WANG C, SUN X, *et al.* Real-time ship detection system for wave glider based on YOLOv5s-lite-CBAM model[J]. Applied Ocean Research, 2024, 144(13): 103833.
- [22] CHEN L J, YAO H D, FU J Y, *et al.* The classification and localization of crack using lightweight convolutional neural network with CBAM[J]. Engineering Structures, 2023, 275: 115291.
- [23] WOO S, PARK J, LEE J Y, *et al.* CBAM: Convolutional Block Attention Module[C]// Computer Vision-ECCV 2018. Cham: Springer, 2018: 3-19.

Semantic Segmentation of Phased Array Ultrasound Images Based on Improved U-Net3+

MAO Xinyue, WANG Huifeng, ZHOU Jiale, GU Zhen, YAN Bingyong

(School of Information Science and Engineering, East China University of Science and Technology,
Shanghai 200237, China)

Abstract: Ultrasonic phased array images have been widely used in the welding defect detection of polyethylene gas pipelines. With the rapid development of machine vision technology, using machine assisted or automated analysis of ultrasonic images can greatly improve the defect detection speed and reduce the occurrence of human judgment errors. In the welding defect detection technology based on ultrasound images, the accuracy of image semantic segmentation is crucial for determining the defect categories and severity levels. This paper proposes an improved model incorporating residual modules and attention mechanism based on the U-Net3+ network, and applies it to the semantic segmentation of phased array ultrasonic image for defect detection in electrofusion welding. Firstly, the improved model enhances the image feature extraction ability of the encoder by adopting a residual structure between each layer of the encoder. Secondly, by introducing Convolutional Block Attention Module (CBAM) in the skip connection, the model's utilization of original image information is strengthened, making it easier for the model to focus on effective regions in the original image. The experimental results show that the improved model has good segmentation performance on ultrasonic images of electrofusion welding, with improvements of 8.81% and 12.84% in Dice and mIoU indicators compared to U-Net, respectively. Compared to U-Net3+, the segmentation performance is improved by 1.09% and 1.81%, respectively.

Key words: phased array ultrasound images; image semantic segmentation; U-Net3+; attention mechanism; residual network

(责任编辑: 李娟)