

## 基于改进U-net的金属工件表面缺陷图像分割方法

王一 龚肖杰 苏皓

### Image segmentation method of surface defects for metal workpieces based on improved U-net

WANG Yi, GONG Xiaojie, SU Hao

引用本文:

王一, 龚肖杰, 苏皓. 基于改进U-net的金属工件表面缺陷图像分割方法[J]. 应用光学, 2023, 44(1): 86–92. DOI: 10.5768/JAO202344.0102004

WANG Yi, GONG Xiaojie, SU Hao. Image segmentation method of surface defects for metal workpieces based on improved U-net[J]. *Journal of Applied Optics*, 2023, 44(1): 86–92. DOI: 10.5768/JAO202344.0102004

在线阅读 View online: <https://doi.org/10.5768/JAO202344.0102004>

---

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### 基于U-net模型的航拍图像去绳带方法

Aerial image de-ropeing based on U-net model

应用光学. 2019, 40(5): 786–794 <https://doi.org/10.5768/JAO201940.0502001>

#### 基于注意力机制与图卷积神经网络的单目红外图像深度估计

Depth estimation of monocular infrared images based on attention mechanism and graph convolutional neural network

应用光学. 2021, 42(1): 49–56 <https://doi.org/10.5768/JAO202142.0102001>

#### 基于深度学习的无人车夜视图像语义分割

Semantic segmentation of night vision images for unmanned vehicles based on deep learning

应用光学. 2017, 38(3): 421–428 <https://doi.org/10.5768/JAO201738.0302007>

#### 色度亮度空间中的线束端子显微图像分割

Wire harness terminals microscopic image segmentation in chromaticity brightness space

应用光学. 2017, 38(6): 931–936 <https://doi.org/10.5768/JAO201738.0602005>

#### 基于深度学习的铸件X射线图像分割研究

Casting X-ray image segmentation based on deep learning

应用光学. 2021, 42(6): 1025–1033 <https://doi.org/10.5768/JAO202142.0602003>

#### 基于改进DeepLabv3+的无人车夜间红外图像语义分割

Semantic segmentation of nocturnal infrared images of unmanned vehicles based on improved DeepLabv3+

应用光学. 2020, 41(1): 180–185 <https://doi.org/10.5768/JAO202041.0106002>



关注微信公众号，获得更多资讯信息

文章编号: 1002-2082 (2023) 01-0086-07

# 基于改进 U-net 的金属工件表面缺陷图像分割方法

王 一<sup>1,2</sup>, 龚肖杰<sup>1</sup>, 苏 翱<sup>1,3</sup>

(1. 华北理工大学 电气工程学院, 河北 唐山 063210; 2. 唐山市金属构件产线智能化技术创新中心, 河北 唐山 063210;  
3. 唐山市半导体集成电路重点实验室, 河北 唐山 063210)

**摘要:** 针对金属工件表面小尺寸缺陷及受非均匀光照影响的图像缺陷难以分割的问题, 提出了一种改进的 U-net 语义分割网络, 实现金属工件表面缺陷图像的精确分割。首先, 在 U-net 网络中融入 CBAM (convolutional block attention module) 模块来提升图像中缺陷目标的显著度; 其次, 采用深度超参数化卷积 DO-Conv(depthwise over-parameterized convolutional) 代替网络中部分传统卷积, 增加网络可学习的参数数量; 然后, 采用 Leaky Relu 函数代替网络中部分 Relu 函数, 提高模型对负区间的特征提取能力; 最后, 采用中值滤波及非均匀光照的补偿方法进行图像预处理, 减弱非均匀光照对金属工件图像表面缺陷的影响。结果表明: 改进后的网络平均交并比、准确率和 Dice 系数指标分别达到 0.8335、0.9332、0.8674, 改进的网络显著提升了对金属工件表面缺陷图像的分割效果。

**关键词:** 表面缺陷; 图像分割; 语义分割网络; 卷积注意力模块; 深度超参数化卷积

中图分类号: TN209; TP391

文献标志码: A

DOI: [10.5768/JAO202344.0102004](https://doi.org/10.5768/JAO202344.0102004)

## Image segmentation method of surface defects for metal workpieces based on improved U-net

WANG Yi<sup>1,2</sup>, GONG Xiaojie<sup>1</sup>, SU Hao<sup>1,3</sup>

(1. College of Electrical Engineering, North China University of Science and Technology, Tangshan 063210, China;  
2. Tangshan Technology Innovation Center of Intellectualization of Metal Component Production Line, Tangshan  
063210, China; 3. Tangshan Key Laboratory of Semiconductor Integrated Circuits, Tangshan 063210, China)

**Abstract:** For the small-size defects of metal workpiece surface and the difficult segmentation of image defects due to non-uniform illumination, an improved U-net semantic segmentation network was proposed to achieve accurate image segmentation of surface defects for metal workpieces. Firstly, the convolutional block attention module (CBAM) was integrated into the U-net network to improve the significance of the defective targets in the image. Secondly, part of the traditional convolution in the network was replaced by depthwise over-parameterized convolution (DO-Conv) to increase the number of learnable parameters of the network. Then, the Leaky Relu function was used instead of the partial Relu function in the network to improve the feature extraction ability of the model for the negative intervals. Finally, the median filtering and non-uniform illumination compensation method were used for image preprocessing, so as to reduce the effect of non-uniform illumination on the surface defects of metal workpiece images. The results show that the improved network mean intersection over union, accuracy rate and Dice coefficient index reaches 0.833 5, 0.933 2 and 0.867 4, respectively. The improved network significantly improves the segmentation effect of surface defect images of metal workpieces.

**Key words:** surface defects; image segmentation; semantic segmentation network; convolutional block attention module; depthwise over-parameterized convolution

收稿日期: 2022-03-24; 修回日期: 2022-09-06

基金项目: 河北省高等学校科学技术研究项目 (ZD2022114); 唐山市科技计划项目 (21130212C)

作者简介: 王一 (1981—), 男, 副教授, 博士, 主要从事视觉检测与感知技术方面的研究。E-mail: [wangyi@ncst.edu.cn](mailto:wangyi@ncst.edu.cn)

通信作者: 龚肖杰 (1994—), 女, 硕士, 主要从事机器视觉、目标识别方面的研究。E-mail: [1692994031@qq.com](mailto:1692994031@qq.com)

## 引言

金属制品在实际生产生活中应用非常广泛, 金属工件质量的好坏直接影响产品的质量及其寿命, 甚至影响产品的实用性及安全性, 因此工件的缺陷检测至关重要<sup>[1]</sup>。传统基于机器视觉的缺陷分割技术有抗噪声差、精度不高并且容易造成过度分割、分割速度较慢等缺点。

随着人工智能技术的快速发展, 深度学习在产品表面缺陷识别领域的研究和应用越来越深入和广泛<sup>[2]</sup>。Wu 等<sup>[3]</sup>针对肺结节边界难以用一种模型分割的问题, 提出一种利用 3D-CRF(conditional random field)优化训练集样本输出的方法分割图像, 此方法可以降低模型的丢失率, 但是需要采集大量数据集进行训练。Tao 等<sup>[4]</sup>针对复杂工业环境下金属表面缺陷准确率低的问题, 设计了一种用于缺陷分割和定位的级联自动编码器, 实现了金属表面缺陷的自动检测, 提高了检测的准确性, 但是需要花费时间手动标记数据。Zhang 等<sup>[5]</sup>针对选煤厂清除煤炭中异物依靠人工效率低的问题, 设计一个嵌入注意力模块的卷积神经网络, 实现了实时从复杂背景中准确地分割出煤炭中的异物, 该方法提高了复杂环境中检测出异物的准确性, 并提高了异物检测的效率。Li 等<sup>[6]</sup>针对带钢表面缺陷的复杂性和多样性导致检测准确低的问题, 设计了一种全卷积的 YOLO 检测网络, 此网络样本集包含 6 类带钢表面缺陷, 可以同时预测缺陷区域的位置及尺寸信息, 该方法具有较好的检测效果, 但是检测精度还需要进一步提高。徐建桥等<sup>[7]</sup>针对轴承表面缺陷影响产品性能及寿命的问题, 提出先通过规范化拆分方法扩充样本数据集, 在利用 ResNet 网络训练检测模型, 该方法相比利用原图直接训练模型的检测效果更好。Xiao 等<sup>[8]</sup>针对肺结节的多样性难以准确检测与分割的问题, 设计具有多尺度特征提取能力的 3D-Res2Net 分层连接网络, 该方法不仅降低了梯度爆炸和消失的问题, 而且提高了肺结节检测和分割的精度及模型的训练速度。卢印举等<sup>[9]</sup>针对玻璃缺陷的透明性及多样性的问题, 设计一种玻璃缺陷图像分割算法, 该算法融合了改进的双特征高斯混合模型, 可以提高玻璃分割的准确性。Ye 等<sup>[10]</sup>针对圆柱形金属工件缺陷的复杂性及多样性导致检测准确率低的问题, 提出集成注意力的弱监督学习框架, 使用图像标签标注训练, 实现了金属工件的

分类与分割。曹富强等<sup>[11]</sup>针对当前图像分割算法对于工业铸件缺陷分割精度低的问题, 设计了一种 Effi-Deeplap 缺陷检测算法, 该算法在 DeepLabv3+基础上, 将原网络中 Xception 模块用 MBConv 替代, 并重新设计空洞空间金字塔池化中空洞卷积的扩张率, 从而提高小目标缺陷的分割精度, 该算法相比 DeepLabv3+具有更好的分割效果。

卷积神经网络虽然具有很强的泛化性, 可以自适应地分割不同环境的图像, 但是对于一些小目标分割效果不理想。为解决金属工件表面缺陷图像的精确分割, 提出一种改进的 U-net 型的表面缺陷分割算法, 在原网络结构中融入 CBAM 模块并在网络编码器结构中采用 DO-Conv 替换传统卷积, 增强特征的学习, 使其能够更高效地对缺陷特征进行提取。采用 Leaky Relu 激活函数代替部分 Relu 函数, 解决卷积中 Relu 函数的稀疏性问题。为解决非均匀光照对图像分割的影响, 设计图像采集视觉平台减少光照的影响, 并采用中值滤波及非均匀光照补偿算法处理采集的图像, 减少图像中局部反光及非均匀光照, 使得图像中光照分布均匀。通过实验证明, 改进后的网络平均交并、准确率及 Dice 系数相比原始网络均有所提高, 对金属工件表面图像缺陷分割效果更好。

## 1 网络结构

### 1.1 语义分割网络 U-net

U-net 网络是基于图像像素级分割的端到端的语义分割网络, 原用于医学细胞图像的分割, 因其所需数据量少、速度快、检测效果较好, 被广泛用于目标检测任务中。U-net 网络结构图如图 1 所示。

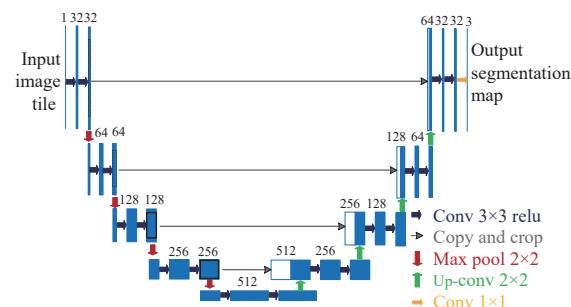


图 1 U-net 网络结构图

Fig. 1 Structure diagram of U-net network

U-Net 网络<sup>[12]</sup>主要包含两个部分: 左侧下采样的过程, 可称为编码器结构, 是模型深化的过程; 右侧上采样的过程, 该侧可称为解码器结构, 是分

辨率恢复的过程。左侧收缩路径经过 4 次下采样, 执行提取图像特征信息的工作, 该过程主要包括  $3 \times 3$  的卷积层、Relu 激活层与  $2 \times 2$  最大池化层, 每次下采样提取图像特征, 特征通道就增加一倍。下采样过程图像特征是从具体到抽象, 因此其图像分辨率会随之降低。右侧扩张路径经过 4 次上采样, 使分割图像恢复到原图像大小, 即在原图像基础上, 通过运用不同的插值算法插入新的元素从而达到图像放大的目的, 也就是一个反卷积过程。该过程包括作用于图像尺寸扩大的转置卷积、使特征图通道数减半的  $3 \times 3$  卷积、深层特征与浅层特征的融合过程和把通道映射到需要的目标类别数的  $1 \times 1$  卷积。运用反卷积和同维度特征图的拼接实现了分辨率的提升。

## 1.2 改进的语义分割网络

### 1.2.1 注意力机制模块 CBAM

CBAM<sup>[13]</sup>是双重轻量级注意力机制, 由通道和空间 2 种子模块组成, 能嵌入到任何 CNN(convolutional neural networks)架构中, 增强特征图有用的特征而忽略无关紧要的特征, 从而提高网络性能。在 U-net 网络中加入 CBAM 模块, 可以使网络提取出积极有效的特征, CBAM 注意力模块示意图如图 2 所示。

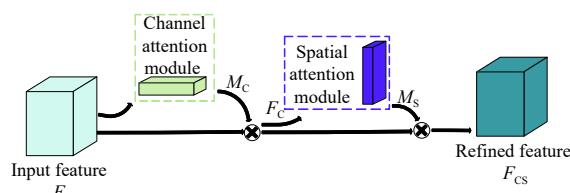


图 2 CBAM 结构示意图

Fig. 2 Structure diagram of CBAM network

在通道注意力模块内, 缺陷图像的输入特征图  $F$  先进行全局平均池化(average pooling)和全局最大池化(global average pooling)运算, 提取高层次特征, 得到 2 个特征描述后被转送到共享网络, 共享网络由含有隐含层的多层感知器组成, 将通过感知器的 2 个特征图进行叠加, 并由 sigmoid 函数激活, 得到通道注意力输出权值  $M_c$ 。通道注意力模块的计算公式为<sup>[14]</sup>

$$M_c(F) = \sigma \{ \text{MLP}[\text{MaxPool}(F)] + \text{MLP}[\text{AvgPool}(F)] \} \quad (1)$$

式中:  $F$  为输入特征图;  $\sigma$  为 sigmoid 函数; MLP 为多层次感知器; MaxPool( $F$ ) 与 AvgPool( $F$ ) 分别表示对特征图  $F$  进行全局最大池化和全局平均池化运算。

sigmoid 函数本质是指数分布族, 用于隐含层神经元输出, 取值范围在  $(0,1)$ , 是常见的 S 型函数, 具有平滑、易于求导的优点。其计算公式为

$$S(x) = \frac{1}{1+e^{-x}} \quad (2)$$

式中:  $x$  为输入变量;  $S(x)$  为变量  $x$  对应的输出。

在空间注意力模块内, 将通道注意力加权得到的特征图  $F_c$  分别沿通道轴做全局平均池化和全局最大值池化运算, 将得到的结果串联起来进行卷积降维操作, 再由 sigmoid 函数激活, 得到最终的空间注意图  $M_s$ 。将  $M_s$  与  $F_c$  相乘得到最终 CBAM 输出的显著特征图  $F_{CS}$ , 空间注意力模块如(3)式所示:

$$M_s(F_c) = \sigma \left\{ f^{7 \times 7} \{ [\text{AvgPool}(F_c); \text{MaxPool}(F_c)] \} \right\} \quad (3)$$

式中:  $\sigma$  为 sigmoid 函数;  $f^{7 \times 7}$  为  $7 \times 7$  的卷积核;  $F_c$  表示由通道注意力加权得到的特征图; AvgPool( $F_c$ ) 与 MaxPool( $F_c$ ) 分别表示对特征图  $F_c$  做全局平均池化和全局最大池化的运算。

### 1.2.2 深度超参数化卷积

DO-Conv 的提出主要是通过增加可学习的参数数量, 来提高在视觉任务中卷积神经网络的表现。DO-Conv 是通过结合传统卷积层与深度卷积层对输入特征图进行新型卷积操作, 其具体公式如(4)式所示:

$$O = W * (D \circ P) = (D^T \circ W) * P \quad (4)$$

式中:  $W$  为传统卷积层的卷积核;  $D$  为深度卷积层的卷积核;  $D^T$  为深度卷积层的卷积核的转置;  $P$  为输入特征区块;  $*$  表示卷积运算;  $\circ$  表示点乘运算。

特征组合方式为深度卷积层的卷积核  $D$  与输入特征区块  $P$  做卷积, 再与传统卷积层的卷积核  $W$  做运算; 卷积核组合方式为深度卷积层的卷积核的转置  $D^T$  与传统卷积层的卷积核  $W$  做卷积, 再与输入特征区块  $P$  做运算。两种组合方式输出结果相同, 但是 DO-Conv 的卷积核组合方式较为方便, 因此采用卷积核组合方式。

### 1.2.3 Leaky Relu 激活函数

Relu 是卷积神经网络常用的激活函数。当输入非负时, Relu 函数能防止模型训练过饱和, 但对输入非正部分有抑制。采用 Relu 的变体 Leaky Relu 函数<sup>[15]</sup>, 对输入为负时, 减轻 Relu 函数的稀疏性, Leaky Relu 函数公式如(5)式所示:

$$f(x) = \begin{cases} x, & x > 0 \\ ax, & x \leq 0 \end{cases} \quad (5)$$

式中:  $x$ 为上一层的输出, 即激活函数的输入值;  $f(x)$ 为激活函数的输出值;  $a$ 表示在 $(0, 1)$ 之间一个较小的常数。当输入小于0时, 给定一个较小的常数 $a$ , 可保证在此范围内有微弱的输出, 梯度能正常传播。

#### 1.2.4 损失函数

适用于多类别分割任务的 Lovasz Softmax 损失函数, 可以提高泛化能力, 它由交叉熵损失函数和指数损失函数组合而成, 其公式如(6)式所示:

$$\text{Loss}(f) = \frac{1}{|C|} \sum_{c \in C} \overline{\Delta J_c}(m(C)) \quad (6)$$

式中:  $C$ 为目标类别;  $\Delta J_c()$ 为模型训练产生的损失值;  $m(C)$ 为像素误差函数;  $\text{Loss}(f)$ 为由 $m(C)$ 构造的代替函数, 评价所有类别的平均损失。

#### 1.2.5 改进后的 U-net 网络

基于经典的 U型网络, 用 DO-Conv 替代网络中上采样操作之前的 Conv, 并将 Leaky Relu 函数作为 DO-Conv 结构中的激活函数, 从而提高网络提取特征的性能。将编码结构中经过卷积模块后的特征图送入 CBAM 模块中, 并将注意力模块输出的特征图与经过反卷积操作后维度相同的特征图拼接, 从而提高网络的抗干扰及其准确定位缺陷区域的能力。改进后的 U-net 网络结构如图 3 所示。图 3 中浅蓝色箭头和深蓝色箭头分别表示 $3 \times 3$  的 DO-Conv 和 Conv 操作, 紫色箭头表示 CBAM 操作, 灰色表示为通道相同特征图的拼接操作, 红色箭头和绿色箭头分别表示编码器 $2 \times 2$  的最大池化和解码器反卷积操作, 黄色箭头表示 $1 \times 1$  的 Conv 映射缺陷类别数操作。

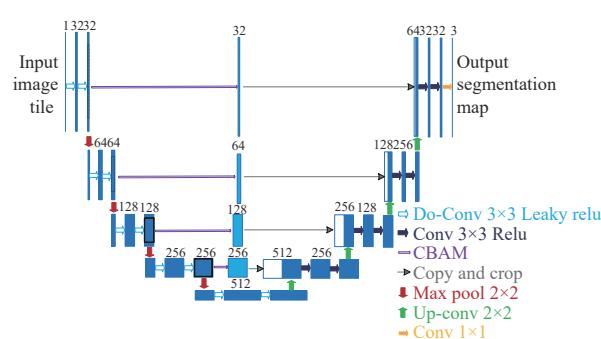


图 3 改进的 U-net 网络结构图

Fig. 3 Structure diagram of improved U-net network

## 2 实验与结果分析

### 2.1 实验设置与数据集

实验平台为基于 Windows 10 操作系统, 中央

处理单元为 Intel(R) Core(TM) i5-1035G1 CPU, 搭载了 NVIDIA MX350 独显, 配备 16G 内存, 使用 Anaconda4.10.1 版本的环境管理器, 编程软件为 Python3.8.8, 利用 Pytorch 深度学习框架完成实验。实验训练 epoch 设为 30, 学习率设置为 0.0001, 并选用 Adam 优化器和 Lovasz Softmax 损失函数进行实验。

图像采集视觉平台中工业相机为堡盟 CX 系列 VCXG-25.1 的 CMOS 相机, 其像元尺寸为 $4.8 \mu\text{m} \times 4.8 \mu\text{m}$ , 相机分辨率为 $1920 \text{ pixel} \times 1200 \text{ pixel}$ 。工业镜头选用 Computar 公司 M1614-MP2 型号, 光源选用亮度可调的条形 LED 灯。实验时将 LED 灯放置在检测对象的两侧, 由于凸轮轴是光滑曲面的轴类物体, 将灯光源方向与垂直方向形成 $45^\circ$ 的夹角的照明角度, LED 灯与凸轮轴顶部平齐, 采集图像时保持相同光照强度及角度。采取此种照明方式不仅能一定程度减弱因光源造成的高反光, 还可以增强缺陷与背景的对比度。采用大光圈、长焦距、小物距减少超景深问题的影响。采集装置一侧放置遮光布将自然光进行遮挡, 保证了环境的变量, 减少了自然光对图像的影响。图像采集视觉平台如图 4 所示。

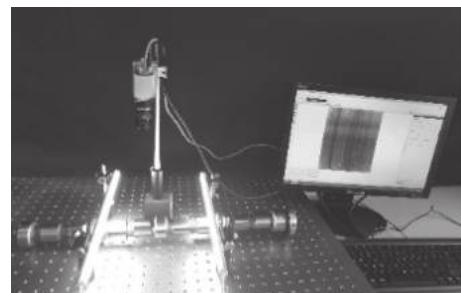


图 4 图像采集视觉平台

Fig. 4 Visual platform for image acquisition

实验选用凸轮轴作为实验对象, 根据实验条件对凸轮轴表面的污渍、划痕及凹坑 3 种缺陷图像进行采集, 并分别用 bl、sc、sca 字母代表。实验采集了 130 张污渍、100 张划痕及 100 张凹坑图像, 并用 labelme 软件进行标注。通过数据增强将测试图像扩增到 1782 张用于训练。采集 70 张污渍、54 张划痕、57 张凹坑及 17 张无缺陷图像用于测试。相机的分辨率为 $1920 \text{ pixel} \times 1200 \text{ pixel}$ , 拍摄的图像为横向长, 将图像方向进行了调整, 使得采集的图像为纵向长, 大小为 $1200 \text{ pixel} \times 1920 \text{ pixel}$ 。

针对非均匀光照对缺陷图像分割的影响, 使用中值滤波和非均匀光照补偿算法结合的方法处理图像, 使图像表面光照分布均匀, 从而减弱图像表

面光照不均匀及高反光现象。光照补偿算法主要思路:先求取原图的平均灰度,然后将图像分成 $108\times 108$ 块,求子块的平均亮度矩阵,再由平均亮度矩阵求得亮度插值矩阵及亮度分布矩阵,最后将原图与亮度分布矩阵做差,最后得到预处理后光照均匀的图像,污渍、划痕、凹坑图像预处理前后对比图如图 5 所示。可见,预处理之后的图像光照分布均匀,从而减少了对金属工件表面缺陷图像分割的影响。

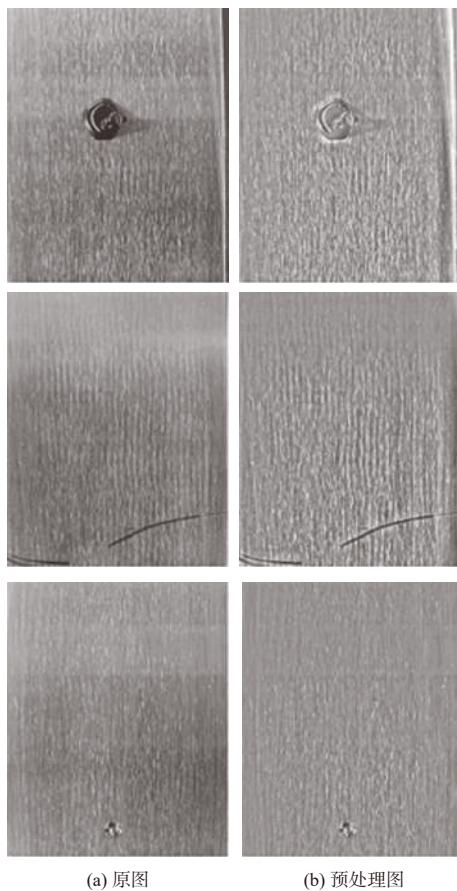


图 5 图像预处理前后对比图

Fig. 5 Comparison before and after image preprocessing

图像经过预处理后,图像尺寸大小不变,在图像进入网络之前利用 resize 函数将经预处理的图像缩放成尺寸为 512 pixel×512 pixel 的图像,即运用 resize 函数中输入图像 scr、输出图像大小 dsize 及双三次插值 INTER\_CUBIC 3 个参数,使其图像尺寸缩放为 512 pixel×512 pixel,则实际输入网络图像大小为 512 pixel×512 pixel。

## 2.2 评价指标

采用平均交并比 mIOU、准确率 Acc 及 Dice 相似度系数作为评价指标。交并比 IOU 是 2 个区域重叠部分与 2 个区域集合部分的比值, mIOU 是

各类对象交并比的均值。Dice 相似度系数是分割后缺陷图像与标准缺陷图像的重合度。mIOU、Acc、Dice 公式如下:

$$\text{IOU} = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} \quad (7)$$

$$\text{mIOU} = \frac{1}{N} \sum_{i=1}^N \text{IOU} \quad (8)$$

$$\text{Acc} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (9)$$

$$\text{Dice} = \frac{2\text{TP}}{\text{FP} + 2\text{TP} + \text{FN}} \quad (10)$$

式中:  $N$  为缺陷类别数; TP、FP、FN 与 TN 分别表示真正、假正、假负、真负。

## 2.3 实验结果与分析

在相同软硬件环境和数据集条件下,分别将经过图像预处理后含有 DO-Conv 的 DO-U-net 网络、含有 CBAM 的 Att-U-net 网络及同时含有 DO-Conv 和 CBAM 的 Att-DO-U-net 网络、原始的 U-net 网络与图像未预处理的 U-net 网络(简称 PW-U-net 网络)进行对比实验。各个模型训练阶段损失值变化曲线图如图 6 所示,图中的横坐标表示权重更新的次数。可以看出当模型训练到 Epoch 值为 28 时,基本趋于平缓,模型整体收敛性比较好,具有一定稳定性,训练比较成功。

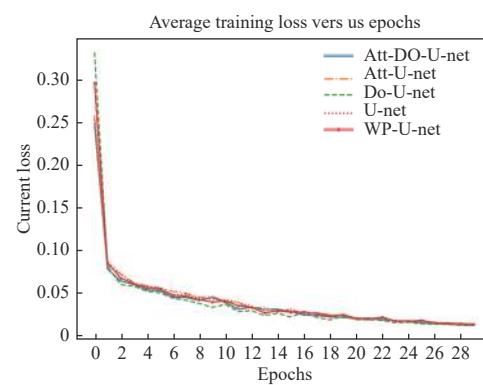


图 6 损失值变化曲线图

Fig. 6 Variation curves of loss values

通过实验,各个模型的平均交并比 mIOU 及准确率 Acc 如图 7、图 8 所示,各个模型的 mIOU 及 Acc 随着迭代次数逐步增加,最终达到较好的结果。PW-U-net 网络的 mIOU、Acc 指标分别为 0.7969、0.9008,原始 U-net 网络的 mIOU、Acc 指标分别为 0.8045、0.9088,Att-U-net 网络的 mIOU、Acc 指标分别为 0.8139、0.9148,DO-U-net 网络的 mIOU、Acc 指标分别为 0.8254、0.9171,Att-DO-U-net 网络的 mIOU、

Acc 指标分别为 0.8335、0.9332, U-net 网络的 mIOU、Acc 指标分别比 PW-U-net 网络高 0.076、0.080, 说明图像预处理可以提升分割效果及准确性。Att-DO-U-net 网络的 mIOU、Acc 指标分别比原始 U-net 网络提升了 0.029、0.024, 说明改进后的 Att-DO-U-net 网络更加关注图像的重要特征, 忽略次要信息, 可达到更好的分割效果。

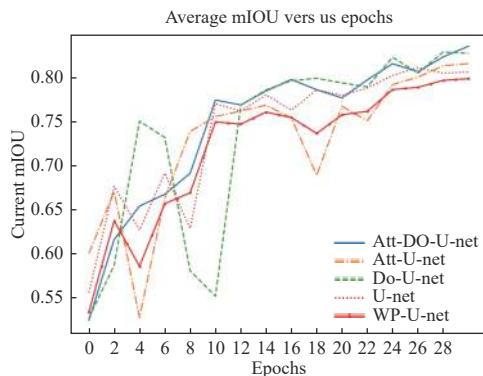


图 7 平均交并比图

Fig. 7 Variation curves of mean intersection over union

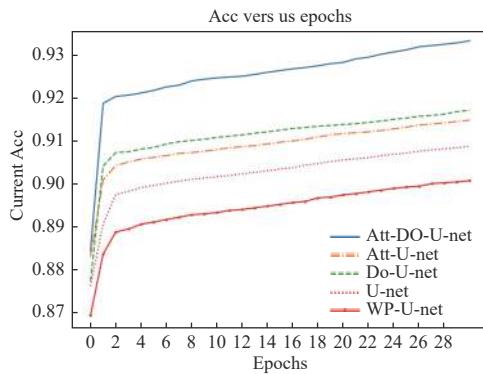


图 8 准确率

Fig. 8 Variation curves of accuracy rate

各模型 Dice 指标如表 1 所示。Dice 反应网络的分割能力, Dice 值越高则缺陷分割的区域与真实区域越接近, 分割效果越好。通过表 1 可知, Att-DO-U-net 网络的 Dice 达到 0.8674, 相比另外几个网络均有所提高, 分割效果更好。

表 1 Dice 指标对比

Table 1 Comparison of Dice indexes

No.	Approach	Dice
1	PW-U-net	0.8463
2	U-net	0.8489
3	DO-U-net	0.8516
4	Att-U-net	0.8512
5	Att-DO-U-net	0.8674

各个网络分割结果如图 9 所示。通过表 1 相似度系数与图 9 分割效果对比可以知道, PW-U-net 网进行分割的图像, 缺陷与边缘分割不准确, 被分割部位的边缘不光滑, 比较粗糙, 图像预处理后用 U-net 网络分割效果比未预处理的分割效果好一点。DO-U-net 和 Att-U-net2 个网络分割效果比原始 U-net 网络分割效果有所改善。本文提出同时加入注意力及 DO-Conv 的 Att-DO-U-net 网络分割的图像, 缺陷边缘更为光滑, 分割准确性也有所提升。

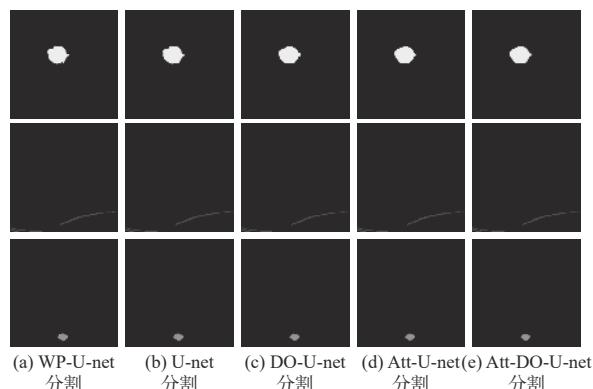


图 9 分割结果对比图

Fig. 9 Comparison of segmentation results

### 3 结论

本文提出嵌入 CBAM 及 DO-Conv 的 U-net 网络和中值滤波与非均匀光照补偿相结合的图像预处理方法, 实现了非均匀光照条件下金属工件表面小尺寸缺陷的精确分割。在此基础上, 设计了图像采集视觉平台, 采集金属工件图像。通过一系列的对比实验, 改进后的 Att-DO-U-net 网络相比原始 U-net 网络对缺陷图像的分割更为完整, 分割的缺陷边缘更为光滑。借此验证了本文提出的改进后的网络对金属工件表面缺陷图像分割效果相较于原始网络更好。

### 参考文献:

- [1] 赵鹤, 杨晓洪, 杨奇, 等. 融合注意力机制的金属缺陷图像分割方法[J]. 光电子·激光, 2021, 32(4): 403-408.  
ZHAO He, YANG Xiaohong, YANG Qi, et al. Metal defect image segmentation algorithm combined with attention mechanism[J]. Journal of Optoelectronics·Laser, 2021, 32(4): 403-408.
- [2] HE J W, ZHAO C D, GAO R Y, et al. Life prediction model of machine tool based on deep learning[J]. Inter-

- national Journal of Plant Engineering and Management, 2021, 26(1): 1-15.
- [3] WU W H, GAO L, DUAN H H, et al. Segmentation of pulmonary nodules in CT images based on 3D-UNEF combined with three-dimensional conditional random field optimization[J]. *Medical Physics*, 2020, 47(9): 4054-4063.
- [4] TAO X, ZHANG D P, MA W Z, et al. Automatic metallic surface defect detection and recognition with convolutional neural networks[J]. *Applied Sciences*, 2018, 8(9): 1575.
- [5] ZHANG K H, WANG W D, LUY Z Q, et al. Computer vision detection of foreign objects in coal processing using attention CNN[J]. *Engineering Applications of Artificial Intelligence*, 2021, 102: 104242-104253.
- [6] LI J Y, SU Z F, GENG J H, et al. Real-time detection of steel strip surface defects based on improved YOLO detection network[J]. *IFAC-Papers Online*, 2018, 51(21): 76-81.
- [7] 徐建桥, 吴俊, 陈向成, 等. 基于规范化样本拆分的轴承缺陷检测[J]. *应用光学*, 2021, 42(2): 327-333.  
XU Jianqiao, WU Jun, CHEN Xiangcheng, et al. Bearing defects detection based on standardized sample split[J]. *Journal of Applied Optics*, 2021, 42(2): 327-333.
- [8] XIAO Z T, LIU B W, GENG L, et al. Segmentation of lung nodules using improved 3D-UNet neural network[J]. *Symmetry*, 2020, 12(11): 1787.
- [9] 卢印举, 郝志萍, 戴曙光. 融合双特征的玻璃缺陷图像分割算法[J]. *包装工程*, 2021, 42(23): 162-169.  
LU Yinju, HAO Zhiping, DAI Shuguang. Glass defect image segmentation algorithm fused with dual features[J]. *Packaging Engineering*, 2021, 42(23): 162-169.
- 169.
- [10] YE M, ZHANG W W, CUI G H, et al. Surface defects inspection of cylindrical metal workpieces based on weakly supervised learning[J]. *International Journal of Advanced Manufacturing Technology*, 2022, 119(3/4): 1933-1949.
- [11] 曹富强, 王明泉, 张俊生, 等. 基于深度学习的铸件X射线图像分割研究[J]. *应用光学*, 2021, 42(6): 1025-1033.  
CAO Fuqiang, WANG Mingquan, ZHANG Junsheng, et al. Casting X-ray image segmentation based on deep learning[J]. *Journal of Applied Optics*, 2021, 42(6): 1025-1033.
- [12] HOU Y W, LIU Z Y, ZHANG T, et al. C-UNet: Complement UNet for remote sensing road extraction[J]. *Sensors*, 2021, 21(6): 2153.
- [13] FU H X, SONG G Q, WANG Y C, et al. Improved YOLOv4 marine target detection combined with CBAM[J]. *Symmetry*, 2021, 13(4): 623.
- [14] 王粉花, 张强, 黄超, 等. 融合双流三维卷积和注意力机制的动态手势识别[J]. *电子与信息学报*, 2021, 43(5): 1389-1396.  
WANG Fenhua, ZHANG Qiang, HUANG Chao, et al. Dynamic gesture recognition combining two stream three-dimensional convolution and attention mechanism[J]. *Journal of Electronics and Information*, 2021, 43(5): 1389-1396.
- [15] 高淑萍, 赵清源, 齐小刚, 等. 改进的MobileNet 的图像分类方法研究[J]. 智能制造学报, 2021, 16(1): 11-20.  
GAO Shuping, ZHAO Qingyuan, QI Xiaogang, et al. Research on the improved image classification method of MobileNet[J]. *CAAI Transactions on Intelligent Systems*, 2021, 16(1): 11-20.