

# 基于轻量 SE-U<sup>2</sup>-Net 的超声红外定量化识别

李志农, 童元, 吴恬

**摘要:** 针对裂纹区域小, 热扩散明显, 分割定量难度过大的问题, 提出基于轻量 SE-U<sup>2</sup>-Net 的超声红外定量化算法, 在 U<sup>2</sup>-Net 分割模型中引入 SE-Net 注意力机制加强对重要生热特征提取, 使用深度可分离卷积代替 U 型残差块中的普通卷积, 降低空间冗余, 减少数据处理的时间。该模型能够以高分辨率捕获局部和全局信息, 同时减少了神经网络的空间冗余和计算负担, 分割网络预测的效率相较于 U<sup>2</sup>-Net 提升了 34.32%, 定量结果与标定值的平均误差仅为 0.8%。

**Abstract:** Aiming at the problems of small crack area, obvious thermal diffusion, and excessive difficulty in segmentation and quantification, an ultrasonic infrared quantification algorithm based on lightweight SE-U<sup>2</sup>-Net is proposed. SE-Net attention mechanism is introduced into U<sup>2</sup>-Net segmentation model to enhance extraction of important heat generating features, and deep separable convolution is used to replace ordinary convolution in U-shaped residual block to reduce spatial redundancy and data processing time. The model can capture local and global information with high resolution, and reduce the spatial redundancy and computational burden of neural network.

## 1 引言

在红外热图像中, 物体的边缘和形态常常因为热扩散而模糊不清, 分割难度较大。随着近年来计算机视觉技术的不断发展, 图像语义分割已成为一个热门领域。它是计算机视觉中最基本的任务之一, 旨在将图像中的每个像素分配给对应的语义类别。深度学习作为一种强大的图像处理工具, 已经被广泛应用于图像语义分割任务中。本文从缺陷分割目标出发, 将基于深度学习的红外裂纹分割算法依次进行详解, 对模型总体框架和各个模块进行了深度分析, 并深入分析和研究 U-Net 和 U<sup>2</sup>-Net 分割算法的对生热区域分割的有效性; 在此基础上, 进行创新和改进, 提出了轻量 SE-U<sup>2</sup>-Net 网络, 实现了更精确快速的缺陷分割定量。

## 2 U<sup>2</sup>-Net 网络模型构建

U<sup>2</sup>-Net 是一个新颖的双层嵌套的 U 型网络结构, 主要包括混合不同大小的感受野的残差 U 块(Residual U-blocks, RSU)和连接 RSU 的外层 U 型结构<sup>[16]</sup>。其中, RSU 由三个部分构成: 输入卷积层、类似 U-Net 的编码器-解码器结构以及通过残差块连接输入层和中间层的残差连接层, RSU 结构图如

图(b)所示。RSU 和原始残差块 (ResNet) 的主要区别在于, RSU 用类似 U-Net 的结构, 替换了其中一层权重层, 并用通过加权层叠加的多层局部特征替换了原始特征。

U<sup>2</sup>-Net 整体结构由三个部分组成: 6 个编码阶段、5 个解码阶段以及采用深度监督策略的特征融合输出模块, 每阶段均由不同层数的 RSU 模块组成。U<sup>2</sup>-Net 整体结构如图所示。

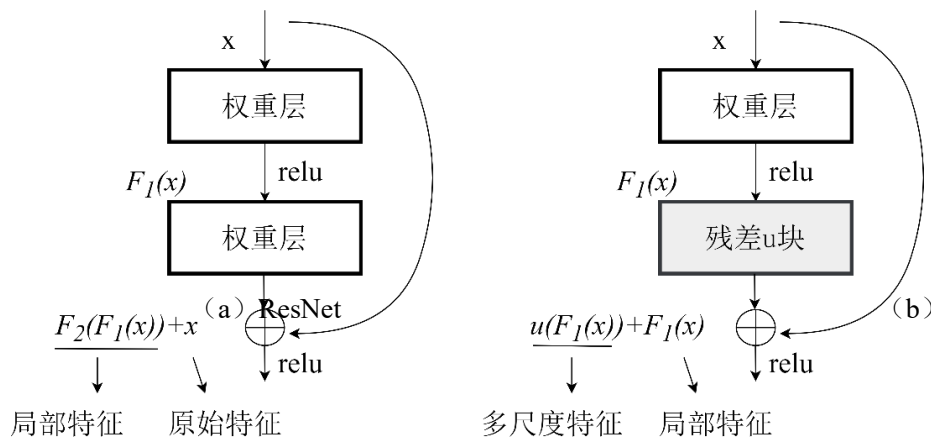


图 1

RSU 与残差网络对比

图中, 在 En\_1、En\_2、En\_3、En\_4 前 4 个阶段的编码器中, 分别使用了从 7 到 4 不同深度的 RSU 模块, 从而获得多尺度的特征信息; 在 En\_5、En\_6 两个阶段的编码器中, 使用了配置空洞卷积的 RSUF 模块, 将卷积上采样操作替换成扩展卷积, 使得中间层的特征图的尺寸与输入层一致, 每个编码器之间用下采样进行特征图的转换与传输; 解码器与左侧同位置编码器配置相同, 输入为上一级编码器输出与左侧对称解码器输出的级联合并, 进行更深层次的特征提取与融合;

最后部分是显著特征图融合模块，首先通过卷积层和 sigmoid 激活函数从 En\_6 编码器以及不同阶段解码器生成 S6、S5、S4、S3、S2、S1 显著图，并将这些显著图上采样到输入图像的大小；通过一个串联操作、 $1 \times 1$  的卷积层以及 sigmoid 函数将这些显著图融合，最终形成特征图输出。

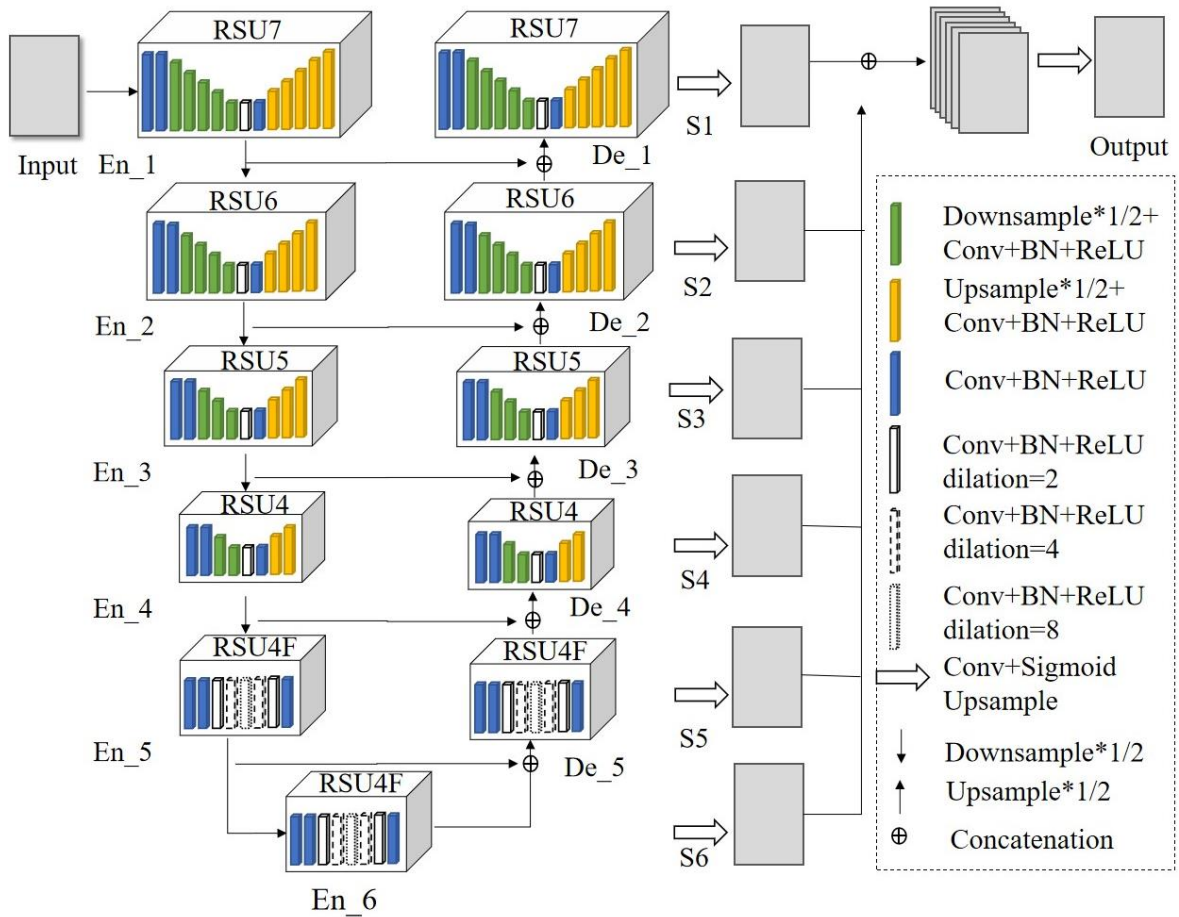


图 2 u<sup>2</sup>-net 网络结构图

### 3 SE-NET 注意力模块

注意力机制的核心是让神经网络能够聚焦于最需要关注的区域。在处理图像时，神经网络应该更加注重关注应该被注意的区域，而不是无所不包地处理，并且我们也不可能人工调整需要注意的地方。注意力机制就可以自适应地关注重要对象。通过对特征层进行加权，使得模型能够更好地理解图像。

Squeeze-and-Excitation Net(SE-Net)是将注意力机制应用到通道维度的代表，在通道域中，SE-Net 通过增强重要通道和抑制不重要通道的方式，显著提高了图像分类准确率，后续大量基于通道域的注意力机制均是基于此进行改进。SE-Net 模

型结构的核心由两部分构成，分别为压缩（Squeeze）、激活（Excitation）。其核心思想是模型根据损失 loss 可以自主学习每个特征通道的重要程度，依据特征通道的重要程度对特征维进行加权优化，从而提高模型的特征学习能力，以达到更好的训练效果如图所示。

（1）Squeeze 操作。对大小为  $C \times W \times H$  的特征图进行平均池化操作，将每个二维特征通道转变为一个数，使其拥有全局的感受野。映射关系如式（1）所示。

$$z_c = F_{sq}(u_c) = \frac{1}{HW} \sum_{i=1}^H \sum_{j=1}^W u_c(i, j) \quad (1)$$

式中：C、W、H 分别为特征图尺寸； $u_c$  为特征通道； $F_{sq}$  为 Squeeze 操作。

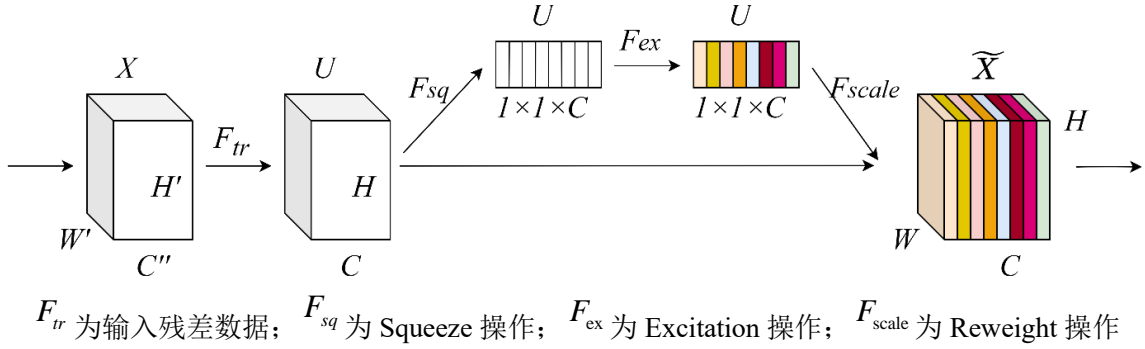


图 3 SENet 模型结构

（2）Excitation 操作。经过两个全连接层通过参数 W 表示不同特征通道的权重及通道相关性。映射关系如式（2）所示。

$$s = F_{ex}(z, W) = \sigma[g(z, W)] = \sigma[W_2 \delta(W_1, z)] \quad (2)$$

式中： $F_{ex}$  为 Excitation 操作定义； $z$  为 Squeeze 操作输出； $W_1$  和  $W_2$  为通道权重； $\delta(\cdot)$  为激活函数； $\sigma(\cdot)$  为归一化函数。

通过乘法将 Excitation 操作的权重加权到原输入特征上，得到输出并作为下一级的输入。映射关系如式（3）所示。

$$x_c = F_{scale}(u_c, s) = s_c u_c \quad (3)$$

式中： $F_{scale}$  为 Reweight 操作定义； $s_c$  为 Excitation 操作的输出矩阵通道。

#### 4 深度可分离卷积

深度可分离卷积(depthwise separable convolution, DS 卷积)是 2017 年谷歌公司的技术研发团队提出的适用于移动端的轻量级卷积神经网络 MoBileNetV1 模型中被广量的轻量 CNN 网络使用的卷积结构，该结构也被称为 depthwise 结构，该结构应用在深度学习模型中代替传统的普通卷积。传统卷积模型参数量大，内存需求大。对于形状为  $H \times W \times C$  的输入特征图，假设我们需要为该特征图提取 N 个特

征。将会有  $N$  个  $k \times k \times C$  卷积核来依次对信道维度和区域进行卷积。该过程图 4 所示，其参数为  $k \times k \times C \times N$ 。

深度可分离卷积由 DW 卷积(depthwise convolution)和 PW 卷积(pointwise convolution)组成，在卷积过程中，DW 卷积考虑图像大小，PW 卷积考虑图像深度。

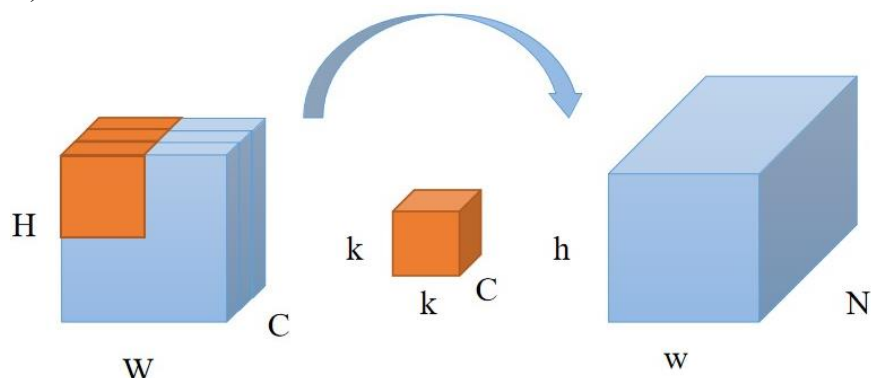


图 4 普通卷积

### (1) DW 卷积

在不改变通道深度的情况下利用二维卷积核对输入图像进行分组卷积，分别提取各个通道的空间特征，对于维度大小为  $H \times W \times C$  的输入图像，每个特征通道都有  $C$  个  $k \times k \times 1$  卷积核，输出特征图大小为  $h \times w \times C$ ，过程如图所示。

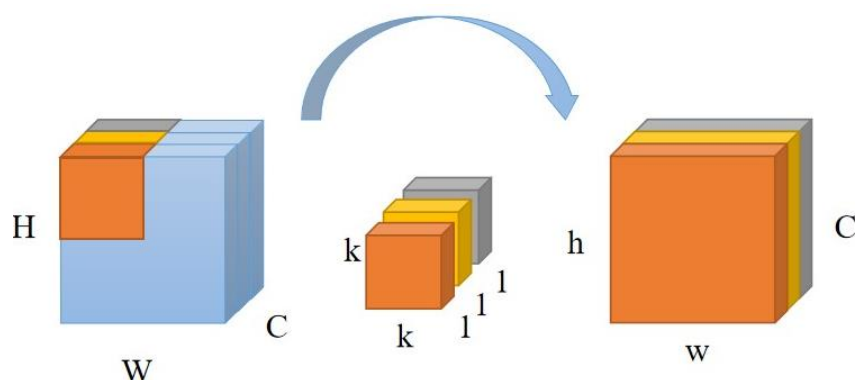


图 5 DW 卷积

### (2) PW 卷积

利用核函数逐点提取特征，增加特征图。假设特征的数量提取特征地图是  $N$ ，然后上一步输出特征图 ( $h \times w \times C$ ) 应该使用  $N$  个  $1 \times 1 \times C$  卷积核，最终输出大小为  $h \times w \times N$ ，如图 6 所示。

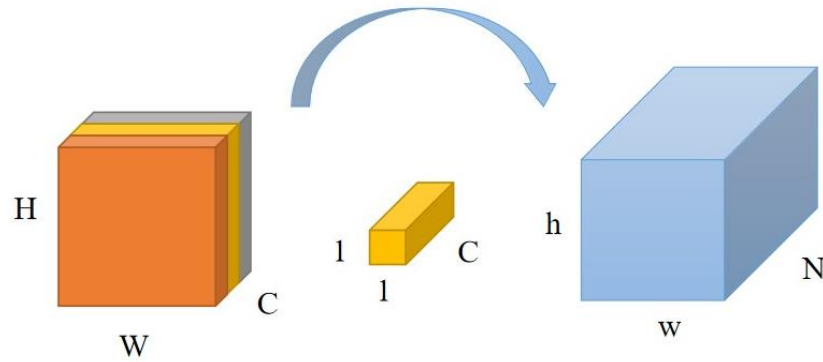


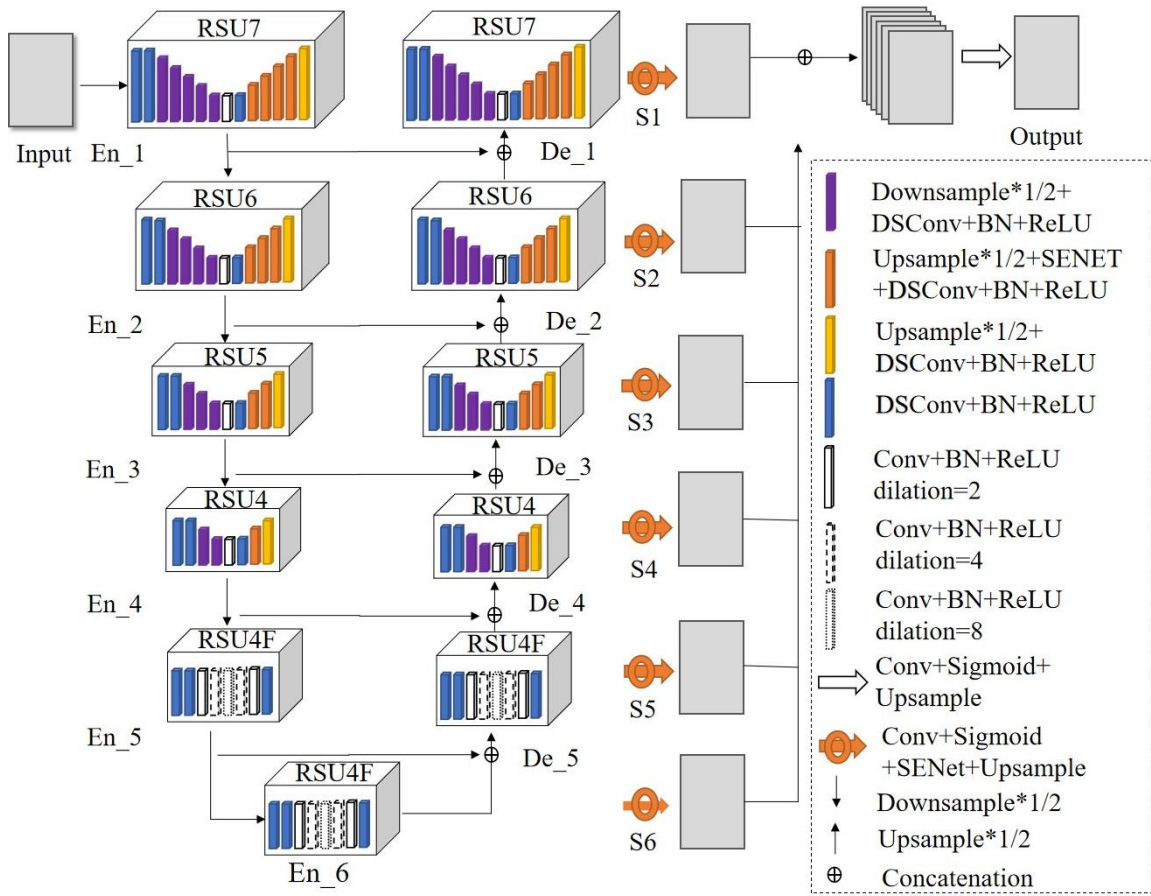
图 6 PW 卷积

与传统卷积参数相比较,使用 DS 卷积可使  $U^2$ -Net 的模型参数数量少 3 倍以上,大大减少了计算量,可有效降低模型训练成本。

## 5 轻量 SE- $U^2$ -Net 定量模型

$U^2$ -Net 中的 RSU 块属于残差结构,受 SE-ResNet 模块的启发,本论文将深度可分离卷积和 SE 注意力机制纳入到原始 RSU 中,提出了轻量 SE-RSU 模块。用 DS 卷积替换 RSU 中的普通卷积,利用 DS 卷积的优势,减少计算量,并且在编码阶段,使用嵌入 SE-Net 通道注意力模块的 RSU 残差块来提高特征提取的有效性,并在解码阶段,将获取到的编码阶段的有效特征层进行上采样操作,以逐像素地恢复原始图像的精度,并恢复细节信息。同时,通过跳跃连接将两部分获取的特征层进行特征融合,以达到更加准确的图像重建效果。将融合后的特征层进行卷积之后,嵌入 SE-Net 通道注意力模块,在前 3 次跳跃连接中,每进行一次特征融合后都添加一个 SE-Net 通道注意力模块。在外层 U 结构中,在每一个编码器输出与左侧对称解码器输出的级联合后,都添加一个 SE-Net 通道注意力模块,使模型在训练过程中能够始终关注重要特征,从而消除了上采样过程中发生的混叠效应,提高对图像序列中裂纹分割的准确性和鲁棒性。



图 7 轻量 SE-U<sup>2</sup>-Net 结构图

## 6 实验研究及对比分析

### 6.1 超声红外热像数据预处理

对于基于深度学习的图像分割算法而言，数据量的充足性是至关重要的，足够多的样本数量可以有效缓解数据量过少所导致的过拟合问题，并使得分割模型具备良好的泛化能力。为了提升超声红外图像分割算法的定量效果，需要对原始数据集进行扩增。常用的数据扩增手段有图像旋转、镜像、图像裁剪、缩放、图像亮度变化、加入噪声等。为有效保留图像的原有特征信息，在此采用图像旋转、镜像的方法对数据集进行扩增，如图 7 所示。最终获得每类 30 张，共计 300 张增强红外热图像。对每张红外热像在长度上进行基于缺陷真实长度的标定，在宽度上进去基于缺陷生热的标定，部分缺陷增强热图的定量标定图如图所示。

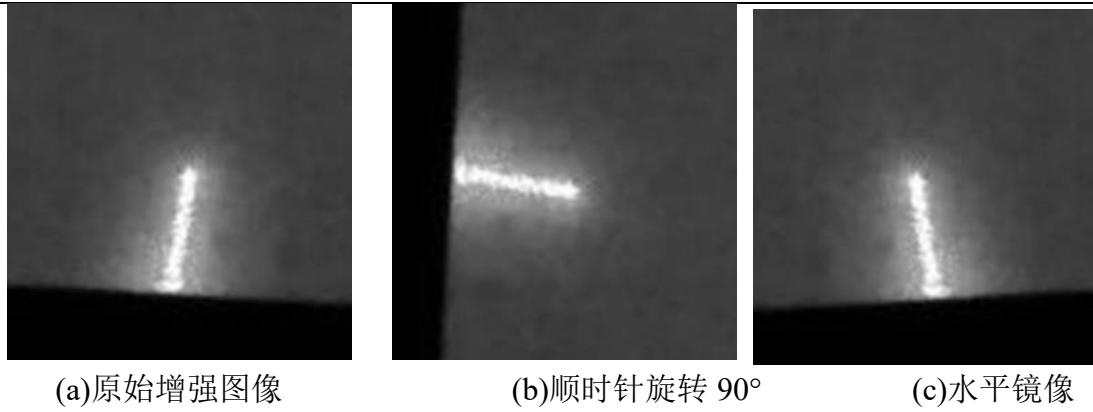


图 8 数据集扩增

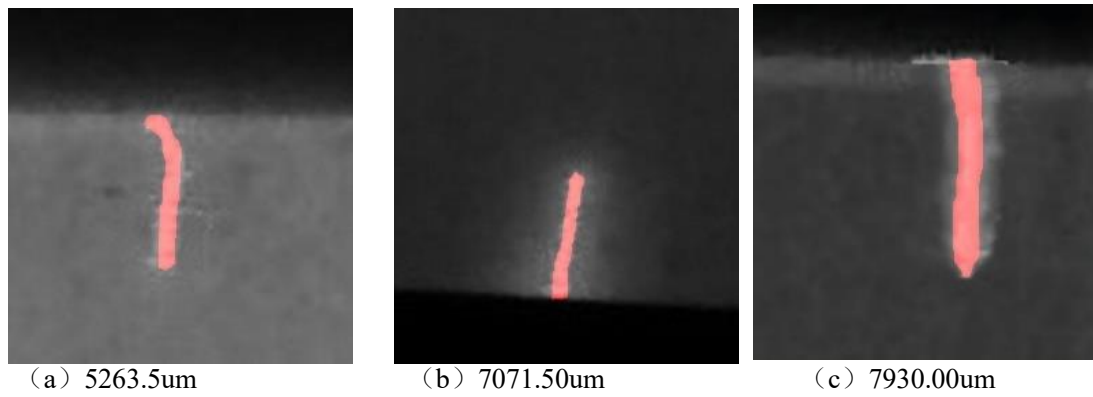


图 9 不同长度裂纹的标定结果

## 6.2 评价标准

通过使用实验结果与标准标注热像图的对比，对模型的性能表现和分割效果进行评估。使用 Dice 相似系数、交并比（IoU）、平均绝对误差（MAE）、召回率（Recall）作为评价指标。

Dice 相似系数。指的是模型分割出的缺陷生热区域 X 与标准标注的缺陷生热区域 Y 之间的一个重叠程度。Dice 的最大值为 1，值越接近 1，则模型越精准。

$$Dice = \frac{2 * (X \cap Y)}{X + Y} \quad (4)$$

- 1) 交并比（IoU）表示模型预测生热区域 X 与标定生热区域 Y 的交集和并集之比。IoU 的最大值为 1，值越接近 1，则区域推测越准确。

$$IoU = \frac{|Y \cap X|}{|Y \cup X|} \quad (5)$$



- 2) 平均绝对误差 (Mean Absolute Error, MAE) 表示预测生热区域  $x_i$  和标定值  $y_i$  之间绝对误差的平均值。范围为  $[0, +\infty)$ ，当预测值与标定值完全吻合时等于 0，即完美模型；误差越大，该值越大。

$$MAE = \frac{1}{m} \sum_{i=1}^m |y_i - x_i| \quad (6)$$

- 3) F<sub>1</sub> 分数 (F<sub>1</sub> Score) 是用来衡量二分类模型精确度的一种指标,同时兼顾了分类模型的准确率和召回率，它的取值区间为 0~1，值越大意味着模型越好。

$$F_1 = 2 \cdot \frac{precision \cdot recall}{precision + recall} \quad (7)$$

$$precision = \frac{TP}{TP + FP} \quad (8)$$

$$recall = \frac{TP}{TP + FN} \quad (9)$$

其中，TP (True Positive) 为预测准确的缺陷前景区域，FP (False Positive) 误判为缺陷前景的背景区域，FN (False Negative) 为误判为背景的缺陷前景区域，TN (True Negative) 为预测准确的背景区域。

为了验证不同模型的分割效果，分别使用 U-Net、U<sup>2</sup>-Net、轻量 SE-U<sup>2</sup>-Net 三种分割模型在同一数据集上进行实验比较。以 Dice 相似系数、IoU、MAE 和 F<sub>1</sub> 分数作为评价指标进行对比实验，实验结果去最佳值，如表所示。

表 1 不同分割模型的测试结果

网络	相似度系数 (Dice)	交并比 (IoU)	平均绝对误差 (MAE)	F <sub>1</sub> 分数 (F1-score)
U-Net	0.902	82.1%	0.007	0.901
U <sup>2</sup> -Net	0.962	92.8%	0.003	0.9
轻量 SE-U <sup>2</sup> -Net	0.981	96.3%	0.001	0.981

从表可以看出，U<sup>2</sup>-Net 的分割效率较 U-Net 有着大幅度的提升，U<sup>2</sup>-Net 的 Dice 相似系数为 0.962，IoU 为 92.8%，F<sub>1</sub> 分数为 0.963，较 U-Net 模型相比分别提升了 0.07，0.13，0.07，MAE 也从 0.007 提升到 0.003。这说明 U<sup>2</sup>-Net 模型的双层特征提取结构相较于 U-Net 的单层 U 型特征提取，能更好地捕捉的超声红外热像的多尺度特征，使得分割裂纹生热区域的结果更加精确。而轻量 SE-U<sup>2</sup>-Net 的 Dice 相似系数为 0.981，IoU 为 96.3%，F<sub>1</sub> 分数为 0.981，相较于 U<sup>2</sup>-Net 模型在 Dice 相似系数、IoU、F<sub>1</sub> 分数分别提升了 0.02，0.04，0.02，证明轻量 SE-U<sup>2</sup>-Net 模型加强

了对裂纹生热特征的关注度，提升了裂纹生热区域的分割准确率，提高了分割模型的鲁棒性。

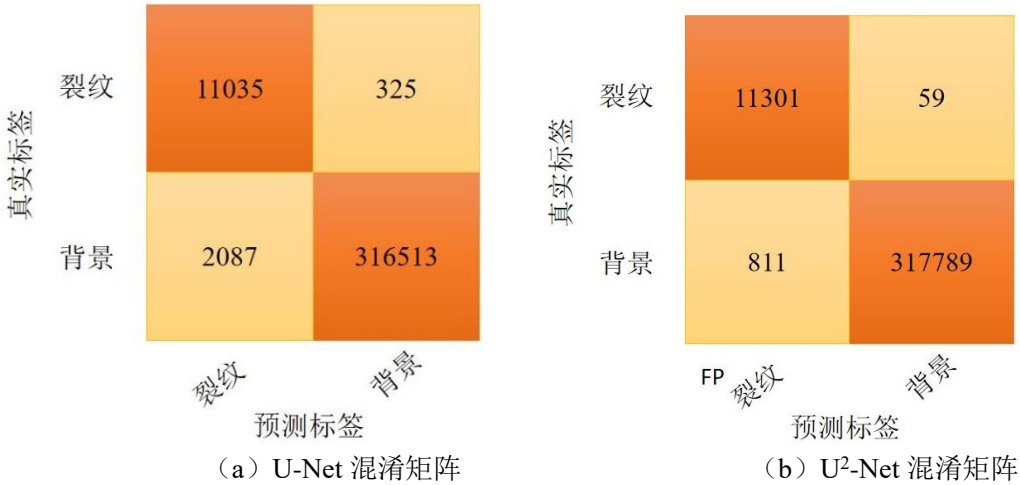
在相同数据集上进行实验，将 U<sup>2</sup>-Net 模型和轻量 SE-U<sup>2</sup>-Net 模型的分割性能以及分割效率做了对比，实验结果均取均值。每个模型均训练了 360 个 epoch。使用 PyTorch1.13.0 和 Python3.8 深度学习框架上进行的。

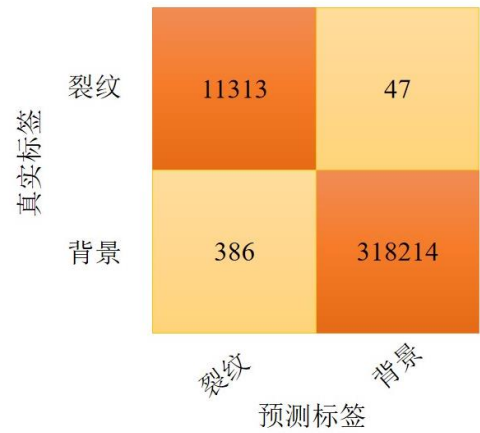
表 2 为轻量 SE-U<sup>2</sup>-Net 模型和 U<sup>2</sup>-Net 模型单张图像分割平均预测时间的对比结果。可以看出轻量 SE-U<sup>2</sup>-Net 模型单张图像的平均时间 82.1ms 比 U<sup>2</sup>-Net 模型单张图像的平均时间 125ms，降低了 42.9ms，分割预测的效率提升了 34.32%。证明了轻量 SE-U<sup>2</sup>-Net 模型能够降低网络计算量并提高计算效率，同时未降低了对特征的提取能力。

表 2 不同分割模型单张图像分割平均时间

实验模型	平均时间（ms）
U <sup>2</sup> -Net	125
轻量 SE-U <sup>2</sup> -Net	82.1

图 10 为 U-Net、U<sup>2</sup>-Net 和轻量 SE-U<sup>2</sup>-Net 的混淆矩阵比较图，混淆矩阵是一种特殊的情况分析矩阵，用于总结模型预测结果。它以矩阵形式展示，利用计数工具将预测正确和不正确的数值进行汇总，以判断模型是否将不同类别的事物混淆。从混淆矩阵可以清楚的看到裂纹生热和工件背景类别正确识别的像素量和错误识别的像素量。单个格里面的数据为对应位置的平均像素数。从下图中可以看出缺陷区域的正确预测像素数量是增加的，U<sup>2</sup>-Net 方法的对裂纹生热区域正确分割比 U-Net 方法多 625 个，而轻量 SE-U<sup>2</sup>-Net 的对背景的正确分割比 U<sup>2</sup>-Net 多 958 个。





(c) 轻量 SE-U<sup>2</sup>-Net 混淆矩阵

图 10 混淆矩阵比较图

7 定量结果

对疲劳裂纹进行分析研究，本章先对裂纹的生热面积进行了分割，得到了较好的分割效果，现在根据对疲劳裂纹的分割结果进行长度的定量化。得到不同分割模型的定量化结果及其误差如表所示。

可以看到 U-Net、U<sup>2</sup>-Net 和轻量 SE-U<sup>2</sup>-Net 三个模型对缺陷长度的定量都较为精准，误差均在 5%以内，且当裂纹却越长，模型对裂纹的长度定量误差越小，结果越精确，三个模型对长度定量的 MAE 分别为 1.39%、1.22%、0.8%。说明轻量 SE-U<sup>2</sup>-Net 对模型的分割不仅在生热宽度上精确度最高，在裂纹长度的定量上也最为精准。

表 3 模型缺陷长度定量结果

工件 编号	标定值	U-Net 模型值		U <sup>2</sup> -Net 模型值		轻量 SE-U <sup>2</sup> -Net 模型值	
	长度/um	长度/um	误差	长度/um	误差	长度/um	误差
01	3898.49	3827.05	-1.8%	3878.08	-0.5%	3857.67	-1.0%
02	5263.50	5386.95	2.3%	5487.96	4.2%	5218.61	-0.8%
03	5477.40	5396.17	-1.5%	5454.19	-0.4%	5465.80	-0.2%
04	6740.50	6671.15	-1.0%	6792.68	0.8%	6782.72	0.6%
05	6983.00	6879.16	-1.5%	6866.18	-1.6%	6944.06	-0.6%
06	7071.50	6892.47	-2.5%	6856.70	-3.0%	6802.96	-3.7%
07	7930.00	8050.33	1.5%	7966.10	0.4%	7881.87	-0.6%

08	8537.50	8619.21	0.9%	8591.64	0.6%	8507.25	-0.3%
09	9143.00	9183.46	0.4%	9169.97	0.3%	9102.55	-0.4%
10	9301.36	9250.69	0.5%	9263.28	0.4%	9288.45	0.1%

## 8 小结

为了降低内存使用、减少数据处理的时间,同时保持精度不下降,引入 SE-Net 注意力机制加强对重要生热特征提取,同时使用深度可分离卷积代替 U 型残差块中的普通卷积,减少空间冗余,得到轻量 SE-U<sup>2</sup>-Net 模型。该模型能够以多尺度捕获局部和全局信息,同时减少了神经网络的空间冗余和计算负担,分割网络预测的效率相较于 U<sup>2</sup>-Net 提升了 34.32%,定量平均误差仅为 0.8%。

## 参考文献

- [1] 姜千辉, 姜长胜, 葛庆平, 等. 红外热波序列图像的图像分割与三维显示[J]. 无损检测. 2008, 30(2): 100-103.
- [2] 李佳文, 李建, 杨杰. 基于梯度关系和局部连通性的红外小目标检测方法[J]. 上海航天(中英文), 2020, 37(05): 113-118.
- [3] 张晔, 朱鸿泰, 程虎, 等. 结合频域显著性分析和形态学滤波的红外小目标检测算法[J]. 激光与红外, 2022, 52(10): 1487-1493.
- [4] Luo Q, Gao B, Woo W, et al. Temporal and spatial deep learning network for infrared thermal defect detection[J]. NDT&E International, 2019, 108: 102164.
- [5] Kstner L, Ahmadi S, Jonietz F, et al. Classification of Spot-welded Joints in Laser Thermography Data using Convolutional Neural Networks[J]. Access IEEE, 2021, 9: 48303-48312.
- [6] Su Y, Tao F, Jin J, et al. Automated Overheated Region Object Detection of Photovoltaic Module With Thermography Image[J]. IEEE Journal of Photovoltaics, 2021(99): 1-10.
- [7] Su B, Chen H, Liu K, et al. RCAG-Net: Residual Channel-wise Attention Gate Network for Hot Spot Defect Detection of Photovoltaic Farms[J]. IEEE Transactions on Instrumentation and Measurement, 2021(99): 1-1.
- [8] 管宽岐, 蔺雨桐, 赵雨薇, 等. 基于深度学习的航拍光伏板红外图像热斑检测方法研究[J]. 电子测量技术, 2022, 45(22): 75-81.
- [9] Zhao S X, Han X, Favro L D, et al. Composite delamination depth profiling in sonic-IR imaging[J]. American Institute of Physics, 2012, 533: 1430.
- [10] Chen D, Naiming W U, Zhang Z. Defect Recognition in Thermosonic Imaging[J]. 中国航空学报(英文版), 2012(4): 657-662.
- [11] Holland S, Renshaw J. Physics-based image enhancement for infrared thermography [J]. NDT&E International 43 (2010) 440-445.

- 
- [12] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation[C]//International Conference on Medical image computing and computer-assisted intervention. Springer, Cham, 2015: 234-241.
- [13] Zhang, J, Chen K, Sergey G, et al. LCU-Net: A novel low-cost U-Net for environmental microorganism image segmentation[J]. Pattern Recognition: The Journal of the Pattern Recognition Society, 2021, 115(1): 107885
- [14] Smith, Abraham G, et al. Segmentation of roots in soil with U-Net[J]. Plant Methods. 2020 (16): 1-15.
- [15] Liu Z, Cao Y, Wang Y, et al. Computer vision-based concrete crack detection using U-net fully convolutional networks[J]. Automation in Construction, 2019, 104(AUG.):129-139.
- [16] Qin X, Zhang Z, Huang C, et al. U2-Net: Going deeper with nested U-structure for salient object detection.[J] Pattern recognition 106 (2020): 107404.
- [17] Zn A, Gz A, Hui Y B. A Review on the Attention Mechanism of Deep Learning[J]. Neurocomputing 452 (2021): 48-62.
- [18] Hu J, Shen L, Sun G. Squeeze-and-excitation networks[C]. Proceedings of the IEEE conference on computer vision and pattern recognition, Salt Lake City Utah: IEEE, 2018, 7132-7141.
- [19] Chollet F. Xception: Deep Learning with Depthwise Separable Convolutions[J]. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017: 1251-1258.

