

ERA-UNet:一种芯片引线键合多特征提取算法

张小国,丁丁,王士强,刘亚飞

(东南大学 仪器科学与工程学院,南京 210096)

E-mail: xgzhang@seu.edu.cn

摘要: 芯片引线键合 X 射线图像中引线、焊球和晶粒特征灰度差异小,阈值分割和边缘检测等传统图像处理方法鲁棒性差且操作复杂。目前,已有较多基于深度学习的语义分割方法,但在多类别、小目标和多尺度分割性能方面仍存在挑战和局限,且对高分辨率 X 射线芯片图像而言,这些网络的模型复杂度较高。针对上述问题,提出了基于 U-Net 改进的芯片引线键合多特征提取网络(ERA-UNet)算法,首先引入平滑卷积预下采样模块以减少 GPU 内存占用和计算量,然后设计残差多尺度特征融合模块以加强特征提取,并改进跳跃连接实现邻近多级特征融合。最后,构建了芯片引线键合语义分割数据集,并设计了对比及消融实验验证算法性能。实验结果表明,ERA-UNet 网络在自建数据集上的 MIoU 达到了 93.05%,相比于其他先进网络具有更优的分割性能,实现了对引线键合多特征的高精度实时提取。

关键词: 芯片检测;引线键合;语义分割;U-Net;特征融合;残差连接

中图分类号: TP391

文献标识码: A

文章编号: 1000-1220(2026)02-0487-08

ERA-UNet: a Multi-feature Extraction Algorithm for Chip Wire Bonding

ZHANG Xiaoguo, DING Ding, WANG Shiqiang, LIU Yafei

(School of Instrument Science and Engineering, Southeast University, Nanjing 210096, China)

Abstract: In chip wire bonding X-ray images, the grayscale differences between wire, solder ball, and grain features are small, making traditional image processing methods, such as thresholding and edge detection, ineffective and complex. Although many deep learning-based semantic segmentation methods have been developed, challenges remain in multi-class, small object, and multi-scale segmentation, especially for high-resolution chip images where model complexities are high. To address these issues, an improved U-Net-based chip wire bonding multi-feature extraction network (ERA-UNet) is proposed. Firstly, a smoothing convolution early down-sampling module is introduced to reduce GPU memory usage and computational load. Then, a residual multi-scale fusion module is designed to enhance feature extraction, and skip connections are modified to achieve adjacent multi-level feature fusion. Finally, a chip wire bonding semantic segmentation dataset is constructed, and extensive comparison and ablation experiments are performed to validate the algorithm's performance. Experimental results show that the ERA-UNet network achieves an MIoU of 93.05% on the self-built dataset, outperforming other advanced networks in segmentation performance, and achieves high-precision real-time extraction of wire bonding features.

Keywords: chip detection; wire bonding; semantic segmentation; U-Net; feature fusion; residual connection

0 引言

引线键合(Wire Bonding, WB)是芯片封装技术中一个重要的步骤^[1],是一种将金属引线连接到芯片焊盘上的技术。焊点虚焊、铝丝断裂、引线变形、焊球面积不足和间距异常等缺陷可能出现在引线键合工序中,而这些问题会直接影响芯片的质量、寿命和应用可靠性^[2]。

传统基于人工视觉检测、光学成像检测、超声显微成像检测和红外热成像检测等方法都存在一定缺陷^[3],而基于 X 射线检测的方法精度高、成像速度快并且操作简便^[4],同时作为一种无损检测,目前已得到了广泛应用。但芯片引线键合图像中需要提取的引线、焊球和晶粒的特征与背景灰度差异小,芯片缺陷种类多样且背景复杂,阈值分割和边缘检测等传

统缺陷检测方法,主要通过图像的浅层特征进行检测,导致其操作复杂,且鲁棒性差。

近年来,基于深度学习的芯片检测算法已被工业界广泛应用,主要包括目标检测算法和图像分割算法。作为一种优秀的目标检测算法,YOLO 算法在其发展历程中不断改进网络架构,显著提高了检测速度和精度^[5]。但在芯片引线键合缺陷检测中,语义分割可以精确捕捉引线、焊球和晶粒的特征,从而提升检测的准确性和鲁棒性。

2014 年提出的 FCN 网络^[6]首次将卷积神经应用在语义分割任务中,构建了端到端、像素到像素的全卷积语义分割网络。另一个经典的网络是 U-Net 网络^[7],使用了对称的 U 型结构和跳跃连接。U-Net 网络能够以较少的图像进行训练,并得到较高的分割精度,已广泛应用于医学图像等领域中。类似

的深度学习方法也被成功应用于无人机航拍影像的语义分割中^[8],可以看到 U-Net 及其衍生模型在多个领域中的广泛应用和潜力。目前,已有较多基于 U-Net 改进的算法。UNet++^[9]采用密集的跳跃连接,通过特征叠加的方式整合不同层次的特征,提高了分割的准确率。U²Net^[10]将设计的RSU(Residual U-block)模块替代 U-Net 中的 VGG 块,融合不同大小的感受野以捕获更多的上下文信息。还有利用注意力机制改进的工作 Attention U-Net^[11]。

与医学图像领域相同,都基于 X 射线成像,近年来 U-Net 及其改进算法在芯片检测中得到了广泛应用。吴忠卿^[12]针对芯片气泡特征不明显而较难提取的问题提出 SSMob-UNet 网络,引入空间注意力机制和空间金字塔池化层,同时利用焦点损失函数解决背景样本不平衡问题。杨宇飞^[13]提出了一种基于 U-Net3+ 改进的检测算法,通过设计门控注意力模块和 CNN-Transformer 混合架构提升芯片封装缺陷的特征提取能力。但这两个算法都不是专门针对引线键合研究的。刘复铭^[14]在 UNet++ 模型的基础上对上采样和下采样操作、编码解码路径和损失函数进行了一系列改进,提升了模型对芯片引线的分割性能。李宇^[15]使用键合区域分割子网络和多任务缺陷检测子网络构成端到端的级联网络,在选择分割网络时还通过五折交叉验证对比了 U-Net、Attention U-Net、U-Net++、U-Net3+ 的分割性能,实验证明 U-Net 与其他网络对引线键合区域分割准确度相差不大,但运行效率更高。

本文聚焦于芯片封装内部结构特征的分割,包括引线、焊球及晶粒。这些特征与医学图像分析中的眼底血管、肝脏等具有相似性,而 U-Net 在该领域表现出色。尽管已有众多基于 U-Net 的改进算法,并且部分已被尝试应用于芯片检测领域,但在处理多类别、小目标、多尺度特征分割时仍面临挑战。此外,这些网络在处理高分辨率芯片图像的分割任务时,模型的复杂度往往较高。

针对芯片图像通常背景复杂、缺陷特征差异小、芯片类型多样等问题,以及工业上对检测精度和速度的需求,本文在 U-Net 的基础上提出一种基于 ERA-UNet 的芯片引线键合多特征提取算法,引入预下采样模块(Early DownSampling, EDS)以减少 GPU 内存占用和计算量,设计残差多尺度特征融合模块(Residual Multi-scale Fusion Block, RMF)以加强特征提取,并改进跳跃连接实现邻近多级特征融合(Adjacent Multi-level Feature Fusion, AMFF),最后在自建芯片引线键合语义分割数据集上进行实验,并与其他网络进行对比分析。

1 ERA-UNet 网络架构

本文提出的算法基于 U-Net 网络改进,ERA-UNet 的整体结构如图 1 所示。

1) 为了实现输入特征图的下采样,ERA-UNet 在 U-Net 网络的编码器输入端引入预下采样 EDS-b 模块,相应地,在解码器输出端进行上采样操作,以减少 GPU 内存占用和计算量;

2) 基于 U-Net 特征提取模块在处理多类别、小目标、多尺度特征提取时存在的局限性,本文设计了一种残差多尺度特征融合模块 RMF 的基础架构,在此架构的基础上设计分组密集连接金字塔结构 GDASP,并分别引入注意力机制和快速金

字塔池化,以加强对图像的有效提取及传播;

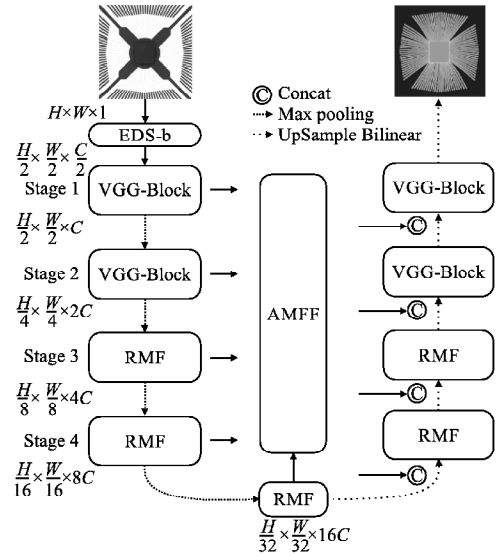


图 1 ERA-UNet 网络结构图

Fig. 1 Structure of ERA-UNet network

3) 为增强解码器恢复信息能力,本文提出了一种邻近多级特征融合模块 AMFF,采用共享注意力机制和分组机制实现相邻层级的语义特征融合。

1.1 预下采样 EDS 模块

基于 X 射线成像的芯片图像分辨率高、尺寸大,普通神经网络在直接处理这些高分辨率图像时往往存在 GPU 内存占用和浮点运算数(Floating Point Operations, FLOPs)过高的问题,对模型的实际部署造成障碍。为了提高模型推理速度,减少 GPU 内存占用和计算量,本文引入并改进了预下采样(Early DownSampling, EDS)模块^[16],为与原模块区分,本文将原 EDS 模块称为 EDS-a,将改进的 EDS 模块称为 EDS-b。EDS 模块改进前后的结构如图 2 所示。相较于 EDS-a 模

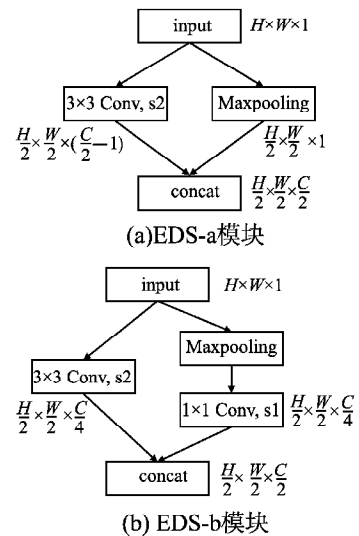


图 2 EDS 模块结构图

Fig. 2 Structure of EDS module

块,EDS-b 模块中的一个分支将通道数转为设定值 C 的 1/4,

并通过步距为 2 的 3×3 卷积将尺寸减半;另一个分支则采用最大池化操作将尺寸减半,再通过步距为 1 的 1×1 平滑卷积将通道数转为设定值 C 的 $1/4$,最后在通道维度上拼接得到最终输出.在池化下采样后增加平滑卷积层,能够有效抑制高频噪声,并在减小特征图尺寸的同时保留更多的细节信息,这有助于改善网络对边缘、角点、纹理、颜色等低级特征的学习和更好地捕捉上下文信息.该模块被嵌入在网络编码器的输入端,因此为了在上采样时恢复原先的尺度,在解码器输出端也增加一个对应的上采样块.

1.2 残差多尺度融合模块

1.2.1 残差多尺度融合模块架构设计

U-Net 网络的原始结构层数相对简洁,其特征提取模块采用了经典的 VGG 卷积块^[17].尽管增加网络深度能够增强性能,但这往往伴随着参数量和计算量的增加,并可能诱发训练过程中的梯度消失,对模型性能产生不利影响^[18].特别是在处理高分辨率、大尺寸的 X 射线芯片图像,以及工业芯片检测中对精度和实时性的要求,需要在提升准确性的同时控制计算量的增加.为解决该问题,ResNet 网络^[19]引入了残差连接,通过将残差块的输入加入到深层细节中,以减轻模型因网络深度增加而导致的退化问题.

其次,残差模块能有效地传递低层特征到高层,增强了网络对低级特征的学习能力,从而提高了网络对图像细节和语义信息的捕捉能力,减少了因信息丢失而导致的性能下降.至今,ResNet 已有多种变体,如图 3 所示,主要通过用更复杂的模块替换简单卷积块,以提取更多信息.

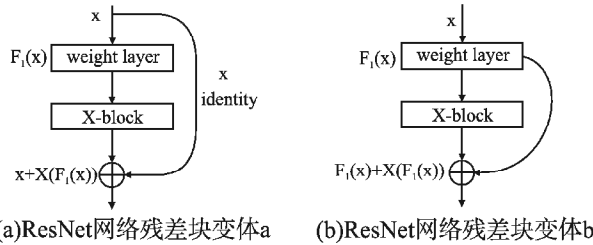


图 3 残差模块架构

Fig. 3 Structure of residual module

图 3(a)中,当网络模型层数过多而产生梯度消失或爆炸时,残差趋近于零,此时残差块的输出 $H(x)$ 等于 x ,残差块表现为一个恒等映射,避免了模型性能的急剧下降.相比之下,图 3(b)中残差块变体 b 的输出 $H(x)$ 将局部语义特征信息 $F_1(x)$ 和多尺度或更丰富的语义特征信息 $X(x)$ 进行了结合,还可以根据数据集图像特征信息定制设计 X-block.

为了增强模型性能,本文在该残差块变体 b 的基础上进行了优化,将残差连接引入 U-Net 中,如图 4(b)所示,将 U-Net 网络的部分 VGG 卷积块替换为残差结构,从而得到残差多尺度融合模块 (Residual Multi-scale Fusion Block, RMF) 的基础架构.通过引入残差连接,模型能够更有效地保留并传递各层次的特征信息,不仅结合了局部语义特征信息,包括图像中的边界等特征,还能够获得更大的感受野和更丰富的全局特征信息,减少高级语义特征信息的丢失,从而提升了语义信息的完整性.

1.2.2 分组密集连接金字塔模块

在当前卷积神经网络的设计架构中,小卷积核在运算效率和内存消耗方面表现较好,但常受限於较小的感受野,无法捕获到更全面的特征信息.而对于浅层高分辨率特征图,较

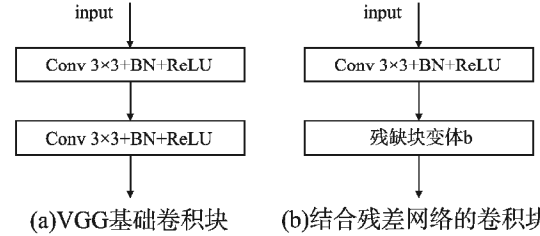


图 4 U-Net 卷积块结构改进

Fig. 4 Improvement of U-Net convolutional block structure

大的感受野能够捕获到更多的全局特征,因此在传统 U-Net 网络中使用的 ASPP^[20]的基础上, Yang 等人提出了 DenseASPP^[21].在保持膨胀因子不变的情况下, DenseASPP 有效扩大了感受野范围,并显著提升了从高分辨率图像中提取全局特征的能力.然而,这种更密集的级联设计也相应地增加了参数量和运算量,本文基于 DenseASPP 的架构进行了改进,借鉴 Res2Net^[22]的思想,引入了分组 (Groups) 拆分操作,并将特征图通道拼接的处理方式改为相加,以此降低运算量和参数量.这一改进结构被称为分组密集连接金字塔结构 (Groups Densely connected Atrous Spatial Pyramid, GDASP),如图 5 所示.其中,将特征图通道拼接改为特征图相加的轻量化设计思路可能会导致不同尺度的信息融合不足,造成信息的丢失,从而影响分割精度,但是特征图相加操作能让模型更容易聚焦于重要的特征,并且能够在保留不同尺度信息的同时,避免由于拼接而导致的通道数过大,增加计算负担,这对于大尺寸的 X 射线芯片图像检测至关重要.

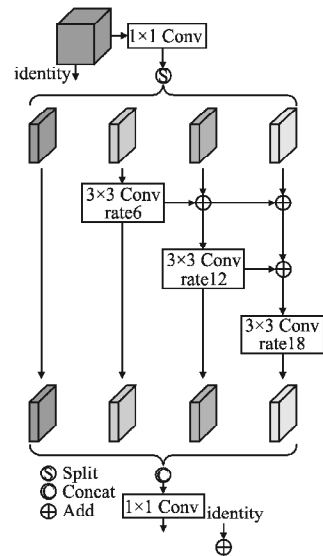


图 5 分组密集金字塔结构图

Fig. 5 Structure of grouped dense pyramid

这样的密集连接结构提供了更密集的尺度采样,同时也降低了参数量和运算量.下面分别计算 ASPP、DenseASPP 和 GDASP 的参数量:

$$P_{ASPP} = 3 \times (c_i \times k^2 \times c_o) \quad (1)$$

$$c_l = c_i + n \times (l-1)$$

$$P_{DenseASPP} = \sum_{l=1}^L (c_l \times 1^2 \times c_i + c_l \times K^2 \times n) \quad (2)$$

$$c_l = c_i \times \left(\frac{c_b}{32.0} \right)$$

$$P_{GDASPP} = (c_i \times 1^2 \times 4c_l) + (4c_l \times 1^2 \times c_o) + \sum_{l=1}^L (c_l \times K^2 \times c_l) \quad (3)$$

式中, c_i 和 c_o 代表初始输入通道和输出通道, c_l 表示每一层的输入通道, n 表示当前层之前的每个层的输出通道, 这里设 $n = c_i$, 即每个 1×1 的普通卷积的输出通道为 c_i . c_b 为一个标准值, 由于 3.2.2 节的实验中证明通道分成 4 份且每一份的通道数为 26 的分组方式效果较好, 因此这里假设等于 26.

为了进行更直观的对比, 设定输入与输出通道数均为 128, 层数均为 3. 通过上述计算公式可以得到, ASPP 的参数量为 442368, DenseASPP 的参数量为 540672, GDASP 的参数量为 398528, 即分组密集连接金字塔模块在扩大感受野时也实现了参数量的缩减.

将该分组密集连接金字塔模块嵌入图 4(b) 中的残差块变体 b 位置处, 使得残差多尺度融合模块能够结合残差连接保留局部特征信息的优点, 同时, GDASP 模块能够进一步扩大感受野, 从多个尺度提取不同的特征, 获取更丰富的全局特征信息, 从而得到更完整的语义特征信息.

1.3 邻近多级特征融合 AMFF 模块设计

当前, 在 U-Net 网络中针对跳跃连接 (Skip Connection) 的改进主要围绕着多尺度、注意力机制和残差连接等方面. UNet++ 采用嵌套结构, 设计的跳跃连接融合了长连接和短连接相融合, 以聚合不同语义尺度的特征; AttentionUNet 提出了交叉注意力模块, 使网络注意力系数更能突显局部区域的特征; MultiResUNet^[23] 引入了残差结构, 改善同一级别的编码器特征和解码器特征之间的语义差距问题, 使连续的同质区域更加均匀; M2SNet^[24] 引入了多尺度卷积滤波器以获得丰富的感受野和高效性能. 受上述网络启发, 针对待分割目标尺度差异大和边界模糊等问题, 为更好地捕捉细节信息, 改善多尺度分割结果, 本文提出了邻近多级特征融合模块 (Adjacent Multi-level Feature Fusion, AMFF). 通过引入注意力机制帮助模型聚焦于关键区域, 更有效地处理边界模糊或背景复杂的任务, 同时采用分组机制在局部范围内进行信息聚合, 为注意力机制提供更加细化的特征表示, 以此增强多级上下文语义特征的聚合能力以及解码器恢复细节和空间关系的能力, 提升模型性能. AMFF 模块结构如图 6 所示, 为方便展示, 该结构图仅为第 i 层的结构.

假设输入一个图像, 经过编码器进行特征提取后, 能够得到 M 张特征图, 将每张特征图表示为 f_i^0 , 其中 $1 \leq i \leq M$. 如图 6 所示, 在处理第 i 级特征图 f_i^0 时, 首先采用注意力机制, 得到 i 张注意力图 f_i^1 , 此注意力机制可以采用如 CBAM、SE 和 CA 等常见模块, 通过加权的方式增强对重要区域的特征响应, 同时抑制无关区域的干扰, 从而提高网络对感兴趣区域的特征表达能力, 包括对模糊边界的分割. 本文通过实验, 选择在此芯片图像分割任务中效果最好的 CBAM 模块. 由于网络中每一层的注意力权重是全局共享的, 因此在计算其他层的 AMFF 时, 无需重新生成新的注意力权重, 这一步有效减少了内

存占用.

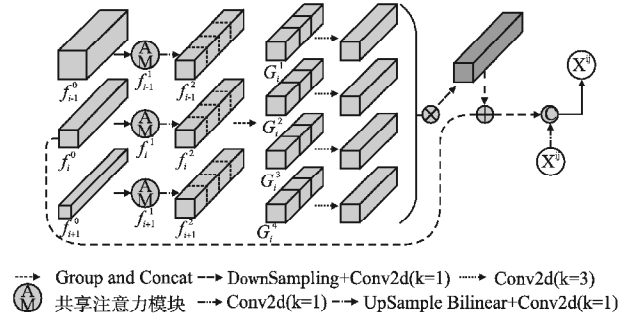


图 6 AMFF 模块结构图

Fig. 6 Structure of AMFF module

为了提升多尺度特征的提取能力, 同时确保各级别特征图能够融合复杂的细节信息和丰富的语义信息, 该模块对相邻多级特征图采用分组融合. 由于邻近的多级特征图之间尺度存在差异, 在分组融合前, 需对特征图 f_{i-1}^1 和 f_{i+1}^1 分别进行池化和上采样, 以确保特征图 f_{i-1}^1 , f_i^1 和 f_{i+1}^1 的尺寸一致, 记为 f_{i-1}^2 , f_{i+1}^2 和 f_i^2 , 随后进行分组融合步骤: 首先, 将尺寸对齐后的 3 个特征图各自划分为 4 份, 接着从划分后的 f_{i-1}^2 , f_i^2 和 f_{i+1}^2 中各取一份并沿着通道维度进行拼接, 以此得到 4 个特征图, 分别记为 G_i^1 , G_i^2 , G_i^3 和 G_i^4 .

$$G_i^n = [c^n(D(f_{i-1}^2)), c^n(D(f_i^2)), c^n(D(f_{i+1}^2))] \quad (4)$$

其中, D 、 U 分别表示下采样和上采样, c^n 表示通道拆分成 4 份后的第 n 份, $[\cdot]$ 表示在通道维度上的拼接操作.

为了融合更多高级语义信息和低级细节信息, 对分组融合后的特征图 G_i^1 , G_i^2 , G_i^3 和 G_i^4 分别使用平滑卷积. 随后, 通过哈达玛乘积 (Hadamard Product) 将 4 个特征图依次相乘, 得到该层级 AMFF 的输出. 相较于拼接, 哈达玛乘积因减少了通道数而提高了计算效率, 能有效增强特征图共有特征并抑制无关特征. 最后, 在将 AMFF 的输出与解码器上采样的特征图融合前, 本方法还引入了一条恒等映射通道, 直接将输入特征图的信息传递给解码器, 避免模型退化, 同时有助于提高模型性能和收敛速度.

2 数据集构建

在深度学习中, 高质量的数据集是至关重要的. 无论是从质量还是数量上, 数据集都会对模型的训练效果产生影响. 由于目前国内外缺少关于芯片内部封装的 X 射线图像的公开数据集, 本文与无锡某公司合作, 构建了芯片引线键合语义分割数据集, 该数据集包含芯片型号 10 种, 原始图像分辨率包括 1024×1024 和 1536×1536 两种.

由于芯片封装内部的各种元件对 X 射线的吸收效果相近, 导致原始图像的待分割目标与背景的灰度差异小, 图像特征不明显, 如图 7(a) 所示, 因此在标注前应对原始灰度图像进行灰度拉伸, 将原始图像转换为 8 位灰度图像.

如图 7 所示, 将图像灰度拉伸至整个灰度空间后, 不仅利于肉眼观察和数据集标注, 在模型训练时, 还极大减少了内存占用空间, 能够提高数据加载和处理的速度.

如图 8 所示, 可以观察到每张图像中, 芯片引线、焊球及

晶粒的特征繁多复杂,不易标注,加之可用数据样本数量不多,导致原始数据集样本较少,存在过拟合风险.为了提高模

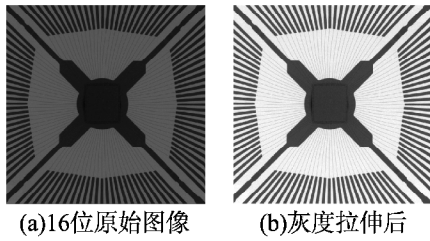


图7 图像灰度拉伸

Fig. 7 Image grayscale stretching

型的训练效果并尽量规避过拟合问题,本文通过数据增强(Data Augmentation)技术以扩充原始数据集.具体地,基于原图分别进行旋转、翻转、添加噪声、亮度与对比度增强等操作,从一张原图生成24张数据增强后的图像.随后将这些图像按60%、20%、20%的比例进行随机划分,最终构建了一个

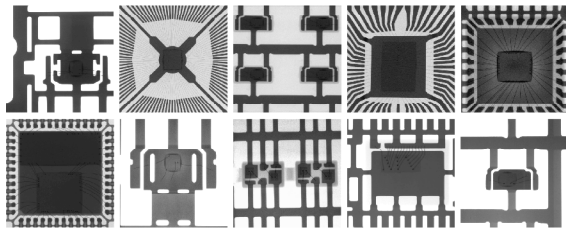


图8 引线键合语义分割数据集

Fig. 8 Dataset of wire bonding semantic segmentation

包含 1572 张训练集、524 张验证集和 524 张测试集的语义分割数据集,其中每类数据集均包含了芯片图像及其对应的掩码标签图像.

表1 数据集中各个型号芯片图像数量

Table 1 Number of images for each chip model in the dataset

型号	各型号总数	训练集	验证集	测试集
C1	445	267	89	89
C2	200	120	40	40
C3	550	330	110	110
C4	150	90	30	30
C5	375	225	75	75
C6	225	135	45	45
C7	150	90	30	30
C8	150	90	30	30
C9	250	150	50	50
C10	125	75	25	25
总数	2620	1572	524	524

数据集中各型号芯片图像数量与数据集划分情况如表 1 所示.

3 实验设计与结果分析

为验证本文提出的预下采样 EDS-b 模块、基于残差网络的多尺度融合模块 RMF 和邻近多级特征融合模块 AMFF 等

改进所带来的模型性能提升,本文将结合自建数据集进行消融实验,并与其他经典的语义分割网络进行实验对比分析.

3.1 实验环境和训练策略

本文在深度学习服务器上进行训练,具体软硬件环境配置如表 2 所示.

表2 实验环境配置

Table 2 Experimental environment setup

实验环境	配置
CPU	AMD EPYC 7642
GPU	NVIDIA GeForce RTX 3090 显存:24GB
操作系统	Ubuntu 18.04
深度学习框架	PyTorch 1.10.2
CUDA	11.3

在训练网络模型的过程中,由于输入图像尺寸较大,GPU 显存有限,部分模型可能会出现显存溢出的问题,因此统一设定输入图像尺寸为 1536×1536 ,批处理大小(batch size)为 2,迭代次数为 300 次.

实验使用带动量的随机梯度下降法(Stochastic Gradient Descent,SGD)进行训练,动量项设置为 0.9.学习率采用余弦退火学习率策略,呈余弦函数逐渐衰减.损失函数采用 Lovász-Softmax Loss^[25],该损失函数将 IoU 计算函数进行平滑延拓,使其成为可导的函数,能够直接优化损失,从而为以 IoU 和 MIoU 作为评价指标的模型带来更高效的训练和更优的泛化能力.

以下消融实验和对比实验中,针对不同的网络结构,训练策略均保持一致.

3.2 消融实验

为验证 ERA-UNet 网络的有效性,本文设计消融实验以验证本文提出的预下采样 EDS 模块、基于残差网络的多尺度融合模块 RMF 和邻近多级特征融合模块 AMFF 等改进对模型性能带来的影响.

3.2.1 预下采样 EDS 模块

为验证改进的 EDS 模块对模型性能的影响,实验在基础 U-Net 网络模型上进行.实验结果如表 3 所示,其中,EDS-a 为 Enet 中的 EDS 模块,EDS-b 为本文改进后的 EDS 模块.

表3 预下采样 EDS 模块改进对模型性能的影响

Table 3 Impact of improved EDS module on model performance

模型	模块	MIoU (%)	Params(M)	FLOPs(G)
U-Net	w/o EDS	89.78	7.85	506.49
	EDS-a	90.49	7.86	129.52
	EDS-b	90.69	7.86	129.48

通过表 3 可以看出,在基础网络 U-Net 中加入预下采样 EDS 模块均可提升模型性能,同时浮点运算数大幅减少,可以显著提升计算效率.其中,相比于不加 EDS 模块的原 U-Net 网络,加入 EDS-a 模块后 MIoU 提升了 0.71%,加入改进的模块 EDS-b 后 MIoU 提升了 0.91%,与此同时浮点运算数只有原来的 25%.

3.2.2 分组密集连接金字塔模块

本实验在残差多尺度融合模块的基础上进行,将图 4(b)

中的残差块变体 b 替换为分组密集连接金字塔模块 (GDASP). 本实验对比验证 GDASP 模块中采用不同的分组的数量和不同的膨胀因子组合对模型性能的影响, 实验结果如表 4 所示, 其中 Groups 表示分组组合, 其中 $26w \times 4s$ 表示通道分成 4 份且每一份的通道数为 26.

表 4 分组密集连接金字塔模块对模型性能的影响

Table 4 Impact of grouped dense connection pyramid module on model performance

模型	Groups	Rates	MIoU (%)
U-Net	$26w \times 4s$	[1, 1, 1, 1]	91.58
	$18w \times 6s$	[1, 1, 1, 1, 1, 1]	91.43
	$14w \times 8s$	[1, 1, 1, 1, 1, 1, 1, 1]	91.44
	w/ EDS-b	$26w \times 4s$	[1, 6, 12, 18]
	$18w \times 6s$	[1, 1, 6, 12, 18, 24]	92.10
	$14w \times 8s$	[1, 1, 1, 1, 6, 12, 18, 24]	92.01

由表 4 可知, 当分组组合为 $26w \times 4s$, 且膨胀因子组合为 [1, 6, 12, 18] 时 MIoU 最高, 达到了 92.21%, 相比没有膨胀卷积提高了 0.63%; 当增加分组数时, 浮点运算数减少, 但 MIoU 也随之降低, 多余的分组并不会带来额外的性能提升. 因此本文选用分组组合为 $26w \times 4s$ 和膨胀因子组合为 [1, 6, 12, 18] 的 GDASP 模块.

3.2.3 邻近多级特征融合模块 AMFF

为了选择合适的注意力机制模块, 在 GDASP 模块后分别引入 SE, ECA, SK, CBAM 和 CA 注意力模块, 实验结果如表 5 所示.

通过表 5 中的实验数据可知, 加入注意力模块能够提升 MIoU, 尤其是采用 CBAM 注意力模块时, MIoU 达到了最高值 91.97%. 此外, 加入 CBAM 对参数量和计算量的影响较小, 因此本文选择 CBAM 作为注意力机制.

为了验证将跳跃连接改进为邻近多级特征融合模块

AMFF 对模型性能的影响, 进行了几组测试: 将跳跃连接替换为 AMFF, 分别测试单独使用 AMFF 和同时使用 AMFF 与残差多尺度融合模块 RMF 的效果, 还包括 identity 恒等映射通道的消融实验和共享注意力模块对 AMFF 的影响, 实验结果如表 6 所示.

表 5 不同注意力机制对模型性能的影响

Table 5 Impact of different attention mechanisms on model performance

模型	注意力模块	MIoU (%)	Params (M)	FLOPs (G)
UNet w/ EDS-b w/ GDASP	×	91.58	14.40	176.10
	SE	91.81	14.45	176.12
	ECA	91.61	14.40	176.12
	SK	91.86	18.84	206.43
	CBAM	91.97	14.45	176.13
	CA	91.82	14.56	176.16

表 6 AMFF 模块对模型性能的影响

Table 6 Impact of AMFF module on model performance

模型	共享注意力	恒等映射	MIoU (%)
U-Net	×	×	91.18
w/EDS-b	×	✓	91.65
w/o RMF	✓	✓	92.30

从表 6 可以看出, 单独引入恒等映射时, MIoU 相对提高了 0.47%; 当同时引入共享注意力模块时, MIoU 相对提高了 1.12%. 实验证明了两者的有效性.

将上述预下采样 EDS-b 模块、残差多尺度融合模块 RMF 和邻近多级特征融合模块 AMFF 一同加入网络, 并对芯片多特征提取中不同种类特征的分割效果进行评估, 最终实验效果如表 7 所示.

表 7 不同模块对模型效果的影响

Table 7 Impact of different modules on model performance

模型	EDS-b	RMF	AMFF	IoU (%)				MIoU (%)
				bg	die	wire	ball	
U-Net	×	×	×	98.85	93.44	82.38	84.45	89.78
	✓	×	×	99.49	97.27	82.83	83.16	90.69
	✓	✓	×	99.53	98.03	86.43	84.98	92.25
	✓	×	✓	99.45	97.76	85.97	86.02	92.30
	✓	✓	✓	99.58	98.23	87.01	87.38	93.05

表 7 中, bg、die、wire 和 ball 分别代表背景、晶粒、引线 and 焊球. 综上实验可知, 得益于 RMF 模块能够增大感受野和获取更丰富上下文信息的优点, 在 EDS-b 的基础上施加 RMF 模块能够使得 MIoU 提升 1.56%; 而 EDS-b 与 AMFF 的组合模式能够降低模型复杂度的同时实现邻近多级特征融合, 因此 MIoU 值从 90.69% 提升至 92.30%. 同时, 针对芯片不同特征分割的结果可知, 不同模块对分割效果都有一定提升作用, 但各自的提升重点有所不同, EDS-b 模块注重晶粒特征的提升, RMF 模块注重引线和焊球的均衡提升, AMFF 模块对焊球提取效果的提升较为显著. 本文改进的基于 U-Net 的 EDS-b

+ RMF + AMFF 网络模型能够综合利用 RMF 模块增强特征提取能力以及 AMFF 模块增强解码器恢复信息能力, 使得 MIoU 最终达到了 93.05%, 相比于基础的 U-Net 模型提升了 3.27%, 实现了对芯片引线键合多特征分割精度的有效提升.

3.3 对比实验

为进一步验证本文提出的 ERA-UNet 网络模型的分割性能, 将引入经典的语义分割模型 U-Net^[7]、SegNet^[26]、DeepLabv3+^[20]、Attention U-Net^[11]、UNet++^[9] 和 U²Net^[10] 进行对比, 在白芯芯片引线键合语义分割数据集上进行的对比实验结果如表 8 所示.

表 8 不同算法在测试集上的性能对比
Table 8 Comparison of the performance of different algorithms on the test set

模型	MIoU(%)	Params(M)	FLOPs(G)
U-Net	89.78	7.85	506.49
SegNet	89.97	29.44	1446.08
DeepLabv3+	87.31	39.75	534.32
Attention U-Net	88.36	8.73	604.10
UNet++	90.01	9.16	1255.39
U ² Net	87.80	44.05	1359.03
ERA-UNet	93.05	19.92	228.67

从表 8 看出,本文提出的 ERA-UNet 网络在 MIoU 上指

标最好,优于其他先进的语义分割模型如 UNet++ 等,证明了本文提出的 ERA-UNet 网络的有效性.并且,ERA-UNet 具有更好的计算速度,更适用于半导体工业监测装备的实际生产.

为了更加直观地对比改进后的模型对芯片内部封装各个特征的提取分割效果,图 9 展示了 ERA-UNet 网络和基础 U-Net 的推理结果,由于芯片型号较多且图像尺寸较大,在数据集中选取了具有代表性的型号芯片进行展示,同时对图片进行局部放大.

从可视化结果可以看出,ERA-UNet 分割性能整体优于 U-Net 和 UNet++,其分割边界更精细,能够更好地捕获小尺度语义特征,避免了 U-Net 和 UNet++ 分割出现的引线断裂等问题.

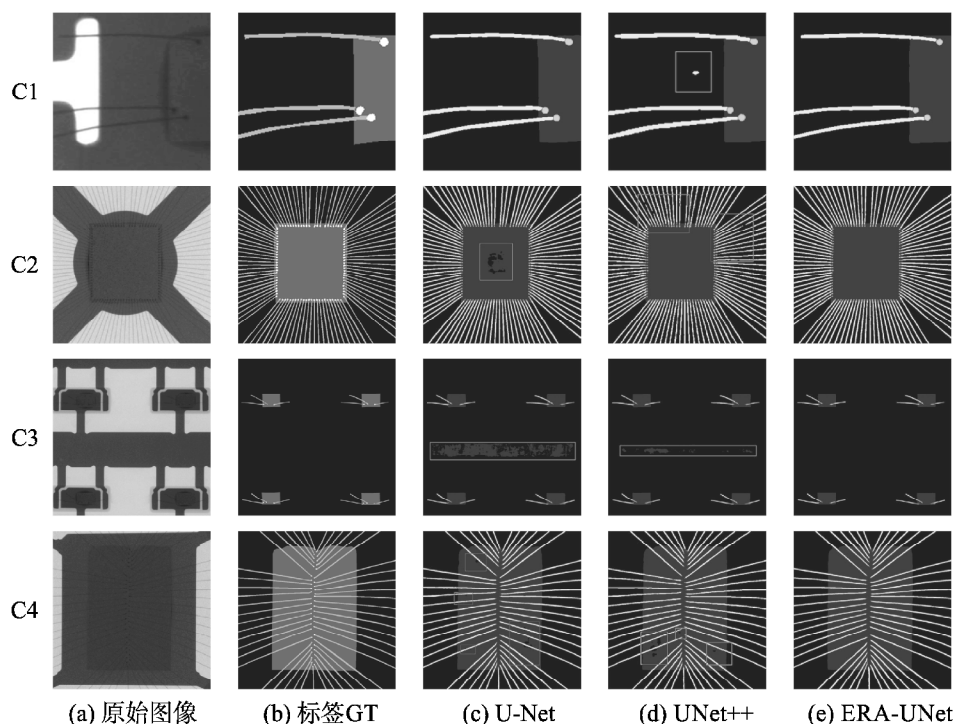


图 9 不同网络分割结果比较

Fig. 9 Comparison of different network segmentation results

4 结论

本文提出的基于 ERA-UNet 的芯片引线键合多特征提取算法,实现了芯片 X 射线引线键合图像中引线、焊球和晶粒特征的精确提取.通过引入预下采样模块以减少了 GPU 内存占用和计算量,基于分组密集连接金字塔结构 GDASP 设计的残差多尺度特征融合模块有效加强了特征提取,并改进了跳跃连接实现邻近多级特征的深度融合,从而在整体上提高了语义分割精度.

消融实验证明了本文中所改进的 EDS-b、RMF 和 AMFF 模块对提升模型性能的作用.此外,与其他先进的网络模型的对比实验表明了本文算法不仅可以实现对引线键合多特征的高精度提取,达到了最佳的分割精度;并且 ERA-UNet 具有更好的计算速度,更适用于半导体工业监测装备的实际生产,本文算法已经应用于半导体检测的实际生产.

References:

- [1] Chen Y, Chien C, Wang C. Empirical study of multi-objective parameter optimization in wire bonding process [C]//14th International Microsystems, Packaging, Assembly and Circuits Technology Conference (IMPACT), IEEE, 2019:134-137.
- [2] Wang C, Chiu J. Shear force classification before wire bonding based on probe mark 2-d images using machine learning methods [J]. IEEE Transactions on Semiconductor Manufacturing, 2022, 35 (2): 210-219.
- [3] XU L H. X-ray detection system for solder defects of ceramic packaging chips based on object detection [D]. Wuxi: Jiangnan University, 2023.
- [4] Du Plessis A, Yadroitsev I, Yadroitsava I, et al. X-ray microcomputed tomography in additive manufacturing: a review of the current technology and applications [J]. 3D Printing and Additive

- Manufacturing, 2018, 5(3):227-247.
- [5] XU Y W, LI J, DONG Y F, et al. Survey of development of yolo object detection algorithms [J]. Journal of Frontiers of Computer Science and Technology, 2024, 18(9):2221-2238.
- [6] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 39(4):640-651.
- [7] 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 (MICCAI), 2015:234-241.
- [8] LUO X D, WU Y Q, CHEN J L. Research progress on deep learning methods for object detection and semantic segmentation in UAV aerial images [J]. Acta Aeronautica et Astronautica Sinica, 2024, 45(6):235-264.
- [9] Zhou Z, Rahman Siddiquee M M, Tajbakhsh N, et al. UNet++: a nested u-net architecture for medical image segmentation [C]//Proceedings of Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support (DLMIA/ML-CDS), 2018:3-11.
- [10] Qin X, Zhang Z, Huang C, et al. U2-Net: going deeper with nested U-structure for salient object detection [J]. Pattern Recognition, 2020, 106:107404, doi:10.1016/j.patcog.2020.107404.
- [11] Oktay O, Schlemper J, Folgoc L L, et al. Attention u-net: learning where to look for the pancreas [J]. arXiv preprint arXiv:1804.03999, 2018.
- [12] LI K, WU Z Q, JI Y, et al. Detection method of weld bubble defect in chip X-ray image based on improved U-Net network [J]. Journal of Huazhong University of Science and Technology (Natural Science Edition), 2022, 50(6):104-110.
- [13] YANG Y F. X-ray inspection system of sealing chips defect based on deep semantic segmentation [D]. Wuxi: Jiangnan University, 2023.
- [14] LIU F M. Research on semiconductor chip wire extraction and morphology detection method based on semantic segmentation [D]. Nanjing, Southeast University, 2023.
- [15] LI Y. Study on quality inspection method for wire bonding in integrated circuit based on X-ray imaging [D]. Nanjing: Southeast University, 2023.
- [16] Paszke A, Chaurasia A, Kim S, et al. Enet: a deep neural network architecture for real-time semantic segmentation [J]. arXiv preprint arXiv:1606.02147, 2016.
- [17] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition [J]. arXiv preprint arXiv:1409.1556, 2014.
- [18] Bi L, Kim J, Kumar A, et al. Stacked fully convolutional networks with multi-channel learning: application to medical image segmentation [J]. The Visual Computer, 2017, 33(6):1061-1071.
- [19] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016:770-778.
- [20] Chen L, Zhu Y, Papandreou G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation [C]//Proceedings of the European Conference on Computer Vision (ECCV), 2018:801-818.
- [21] Yang M, Yu K, Zhang C, et al. Densenet for semantic segmentation in street scenes [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018:3684-3692.
- [22] Gao S, Cheng M, Zhao K, et al. Res2net: a new multi-scale backbone architecture [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2019, 43(2):652-662.
- [23] Ibtihaz N, Rahman M S. MultiResUNet: rethinking the U-Net architecture for multimodal biomedical image segmentation [J]. Neural Networks, 2020, 121:74-87, doi:10.1016/j.neunet.2019.08.025.
- [24] Zhao X, Jia H, Pang Y, et al. M2SNet: multi-scale in multi-scale subtraction network for medical image segmentation [J]. arXiv preprint arXiv:2303.10894, 2023.
- [25] Berman M, Triki A R, Blaschko M B. The iou loss: a tractable surrogate for the optimization of the intersection-over-union measure in neural networks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2018:4413-4421.
- [26] 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, 2015, 39(12):2481-2495.

附中文参考文献:

- [3] 徐林海. 基于目标检测的陶封芯片焊缝缺陷 X 射线检测系统 [D]. 无锡: 江南大学, 2023.
- [5] 徐彦威, 李 军, 董元方, 等. YOLO 系列目标检测算法综述 [J]. 计算机科学与探索, 2024, 18(9):2221-2238.
- [8] 罗旭东, 吴一全, 陈金林. 无人机航拍影像目标检测与语义分割的深度学习研究方法研究进展 [J]. 航空学报, 2024, 45(6):235-264.
- [12] 李 可, 吴忠卿, 吉 勇, 等. 改进 U-Net 芯片 X 线图像焊缝气孔缺陷检测方法 [J]. 华中科技大学学报 (自然科学版), 2022, 50(6):104-110.
- [13] 杨宇飞. 基于深度语义分割的芯片熔封缺陷 X 射线检测系统 [D]. 无锡: 江南大学, 2023.
- [14] 刘复铭. 基于语义分割的半导体芯片引线提取和形态检测方法研究 [D]. 南京: 东南大学, 2023.
- [15] 李 宇. 基于 X 射线成像的集成电路引线键合质量检测方法的研究 [D]. 南京: 东南大学, 2023.