

中图法分类号: TP301 文献标识码: A 文章编号: 1006-8961(2024)03-0741-14

论文引用格式: Xu W W, Xu L F, Li B K, Zhou X, Lyu N and Zhan S. 2024. TransAS-UNet: regional segmentation of breast cancer Swin Transformer and of UNet algorithm. Journal of Image and Graphics, 29(03):0741-0754(徐旺旺, 许良凤, 李博凯, 周曦, 律娜, 詹曙. 2024. TransAS-UNet:融合 Swin Transformer 和 UNet 的乳腺癌区域分割. 中国图象图形学报, 29(03):0741-0754)[DOI:10.11834/jig.230130]

TransAS-UNet:融合 Swin Transformer 和 UNet 的乳腺癌区域分割

徐旺旺^{1,2}, 许良凤^{1,2}, 李博凯^{1,2}, 周曦^{1,2,3}, 律娜^{4*}, 詹曙^{1,2}

1. 合肥综合性国家科学中心人工智能研究院, 合肥 230601; 2. 合肥工业大学计算机与信息学院, 合肥 230601;
3. 安徽水利电力职业技术学院, 合肥 231603; 4. 安徽医科大学第一附属医院, 合肥 230022

摘要: **目的** 乳腺癌在女性中是致病严重且发病率较高的疾病, 早期乳腺癌检测是全世界需要解决的重要难题。如今乳腺癌的诊断方法有临床检查、影像学检查和组织病理学检查。在影像学检查中常用的方式是 X 光、CT (computed tomography)、磁共振等, 其中乳房 X 光片已用于检测早期癌症, 然而从本地乳房 X 线照片中手动分割肿块是一项非常耗时且容易出错的任务。因此, 需要一个集成的计算机辅助诊断 (computer aided diagnosis, CAD) 系统来帮助放射科医生进行自动和精确的乳房肿块识别。**方法** 基于深度学习图像分割框架, 对比了不同图像分割模型, 同时在 UNet 结构上采用了 Swin 架构来代替分割任务中的下采样和上采样过程, 实现局部和全局特征的交互。利用 Transformer 来获取更多的全局信息和不同层次特征来取代短连接, 实现多尺度特征融合, 从而精准分割。在分割模型阶段也采用了 Multi-Attention ResNet 分类网络对癌症区域的等级识别, 更好地对乳腺癌进行诊断医疗。**结果** 本文模型在乳腺癌 X 光数据集 INbreast 上实现肿块的准确分割, IoU (intersection over union) 值达到 95.58%, Dice 系数为 93.45%, 与其他的分割模型相比提高了 4%~6%, 将得到的二值化分割图像进行四分类, Accuracy 值达到 95.24%。**结论** 本文提出的 TransAS-UNet 图像分割方法具有良好的性能和临床意义, 该方法优于对比的二维图像医学分割方法。

关键词: 乳腺癌; 深度学习; 医学图像分割; TransAS-UNet; 图像分类

TransAS-UNet: regional segmentation of breast cancer Swin Transformer and of UNet algorithm

Xu Wangwang^{1,2}, Xu Liangfeng^{1,2}, Li Bokai^{1,2}, Zhou Xi^{1,2,3}, Lyu Na^{4*}, Zhan Shu^{1,2}

1. Institute of Artificial Intelligence, Hefei Comprehensive National Science Center, Hefei 230601, China;
2. School of Computer Science and Information Engineering, Hefei University of Technology, Hefei 230601, China;
3. Anhui Water Conservancy and Electric Power Technical College, Hefei 231603, China;
4. The First Affiliated Hospital of Anhui Medical University, Hefei 230022, China

Abstract: Objective Breast cancer is a serious and high-morbidity disease in women. Early detection of breast cancer is an important problem that needs to be solved all over the world. The current diagnostic methods for breast cancer include clinical, imaging, and histopathological examinations. The commonly used methods in imaging examination are X-ray,

收稿日期: 2023-03-24; 修回日期: 2023-07-07; 预印本日期: 2023-07-14

* 通信作者: 律娜 lvna@ahmu.edu.cn

基金项目: 安徽省高校协同创新项目 (GXXT-2022-041); 合肥市自然科学基金项目 (2021008)

Supported by: The University Synergy Innovation Program of Anhui Province (GXXT-2022-041); Hefei Natural Science Foundation (2021008)

computed tomography (CT), and magnetic resonance imaging. etc., among which mammograms have been used in early cancer to detect; however, manually segmenting the mass from the local mammogram is an very time-consuming and error-prone task. Therefore, an integrated computer aided diagnosis (CAD) is needed to help radiologists perform automatic and precise breast mass identification. **Method** In this work, we compared different image segmentation models based on the deep learning image segmentation framework. At the same time, on the based UNet structure, we adopt the Swin architecture to replace the downsampling and upsampling processes in the segmentation task, to realize the interaction between local and global features. At the same time we use a Transformer to obtain more global information and different hierarchical features to replace short connections and realize multi-scale feature fusion to achieve accurate segmentation. In the segmentation model stage, we also use so as Multi-Attention ResNet classification network to identify the classification of cancer regions Better diagnosis and treatment of breast cancer. During segmentation the Swin Transformer and atrous spatial pyramid pooling (ASPP) modules are used to replace the common convolution layer through analogy with the UNet structure model. The shift window and multiple attention are used to achieve the integration of feature information inside the image slice and extract information complementarity between non-adjacent areas. At the same time, the ASPP structure can achieve self-attention of local information with an increasing receptive field. A Transformer structure is introduced to correlate information between different layers to prevent the loss of shallow layers of important information during downsampling convolution. The final architecture not only inherits advantages Transformer's in learning global semantic associations, but also uses different levels of characteristics to preserve more semantics and more details in the model. As the input dataset of classification networks, binarized images obtained by the segmentation model can be used to identify different categories of breast cancer tumors. Based on ResNet50, this classification model adds multi-type attention modules and overfitting operations. squeeze-and-excitation (SE) and selective kernel (SK) attention can optimize network parameters, so that it only pays attention to the differences in segmentation regions improving the efficiency of the model. Thus proposed model by us achieved accurate segmentation of the lump on the breast cancer X-ray dataset INbreast, and we also compared it with five segmentation structures: UNet, UNet++, Res18_UNet, MultiRes_UNet, and Dense_UNet. After the segmentation model, a more accurate binary map of the cancer region was obtained. Problems, such as feature information blending of different levels and self-concern of the local information of the convolutional layer, exist in up-sampling and downsampling based on the UNet structure. Therefore, the Swin Transformer structure, which has a sliding window operation and hierarchical design, is adopted. Window Attention is shifted mainly by the Window Attention module and the Shifted window attention module, which enables the input feature graph to be sliced into multiple windows. The weight of each window is shifted in accordance with the shifted self-attention, and the position of the entire feature graph is shifted. It can realize the information interaction within the same feature graph. In upsampling and downsampling, we use four Swin Transformer structures. and in the process of fusion, we use the pyramid ASPP structure to replace the common feature graph channel addition operation, which can use multiple convolution check feature graphs and channel fusions, and the given input can be sampled in parallel with cavity convolution at different sampling rates. Achieve multiple scale capture image context information is obtained. In order to better integrate high- and low-dimensional spatial information, we propose a new multi-scale feature graph fusion strategy and use a Transformer with skip connections to enhance spatial domain information representation. Each cancer image was classified into normal, mass, deformation, and calcification according the introduction of the INbreast dataset. Each category was labeled and then sent to the classification network. The classification model we adopted takes ResNet50 as the baseline model. On this basis, two different kinds of attention, i. e., SE and SK, are added. SK convolution replaces 3×3 convolution at every bottleneck. Thus, more image features can be extracted at the convolutional layer. Meanwhile SE belongs to channel attention, and each channel can be weighted before the pixel value is outputted. Three methods, namely, Gaussian error gradient descent, label smoothing, and partial data enhancement, are introduced to improve the accuracy of the model. **Result** In the same parameter environment, the intersection over union (IoU) value reached 95.58%. Dice coefficient was 93.45%, which was 4%–6% higher than that of the other segmentation models. The binary segmentation image is classified into four categories, and the Accuracy reached 95.24%. **Conclusion** Experiments show that our proposed TransAS-UNet image segmentation method demonstrates good performance and clinical significance which is superior to those of other 2D image medical segmentation methods.

Key words: breast cancer; deep learning; medical image segmentation; TransAS-UNet; image classification

0 引言

癌症是世界人口死亡率最高的主要原因,而在女性中乳腺癌是最常见的癌症疾病(Bray 等, 2018),世界卫生组织报告称,每年约有 62.67 万女性死于癌症相关疾病,2018 年诊断出 200 多万新病例(Bray 等, 2018)。而在早期对患者病情的发现与治疗,可以大大降低死亡率和医疗成本,这对患者和医疗资源紧缺的国家来说具有至关重要的作用。尽管医疗水平不断发展,但现存的医疗手段仍存在很多弊端。主要使用的检查手段包括乳腺 X 射线、磁共振成像、热成像和超声波成像,这些成像结果只能确定癌症可疑区域,并不能判断患者是否患癌和癌症级别(Saad 等, 2016)。在早期乳腺癌治疗中, X 光检查被认为是最有效的方法(George 和 Sankar, 2017),可用于检测癌症预后早期良性和恶性肿块,然而肿瘤检测是基于对乳房感兴趣区域(region of interest, ROI)的精确分割。因此,乳房 X 线计算机辅助诊断中乳房肿块的分割在医学图像的定量和定性分析中起着重要作用,它对早期癌症的分析和治疗有直接影响。

早期 Zhang 等人(2010)基于从像素的共现矩阵计算出的能量描述符(Haralick 描述符),使用纹理分析来识别乳房 X 线照片中的可疑肿块; Ke 等人(2009)引入了小波变换模最大值,以自动方式识别乳房 X 线片中的肿块,并通过提取其轮廓来细化检测到的肿块; Jai-Andaloussi 等人(2013)首先使用迭代阈值算法在乳房 X 光检查中发现可疑肿块的感兴趣区域。然后从应用 BEMD (bivariate empirical mode decomposition) 二维经验模式分解方法获得的第 1 模式中提取感兴趣区域的轮廓,最后通过提取的轮廓来识别肿块; Yang 等人(2023)提出了一种改进的多策略差分进化算法,采用基于新型差分进化、二维 Kapur 熵和二维直方图的多级阈值图像分割框架来分割肿块。

在传统的机器学习方法中,主要采用有监督和无监督方法分割乳腺癌肿块,从输入图像学习不同具有代表性的特征来学习和预测癌症。Salih 和 Kamil(2018)提出了一种基于模糊集和阈值技术的乳房 X 光图像分割方法,该方法产生了 86.0% 的

Dice 系数、66.0% 的召回率和 20.0% 的准确度; Cardoso 等人(2017)提出了乳房 X 光图像中的肿块分割,并与深度和定制特征进行比较,结合 CRF (conditional random fields) 和 SSVM (structured support vector machine) 的深度学习特征的第 2 和第 3 模型。

然而,随着深度学习在医学影像上的发展, Sun 等人(2020)开发了一种新的注意力引导密集上采样网络 AUNet,用于整个乳房 X 光检查中的乳房肿块分割;经典的医学分割网络 UNet (Ronneberger 等, 2015)主要采用具有跳跃连接的对称编码—译码结构;按照这一技术路线, UNet++ (Zhou 等, 2018)、Res-UNet (ResNet-UNet) (Xiao 等, 2018) 和 UNet3+ (Huang 等, 2020) 等各种延伸网络用于医学成像模式的图像和视频分割。Cho 等人(2022)设计了一个乳腺肿瘤集成分类网络 (breast tumor ensemble classification network, BTEC-Net) 来对是否包含乳腺肿瘤进行分类; Iqbal 和 Sharif (2023) 提出一种结合数据扩展网络 (dynamically expandable networks, DEN)、概率映射生成网络 (probabilistic mapping generation, PMG) 和 U 型金字塔扩张网络 (U-shaped pyramid-dilated network, PDF-UNet) 的半监督学习方案; Nelson 和 Krishna (2023) 开发一种使用基于深度学习方法的自动核分割算法,利用星凸多边形方法的像素级分析获得更准确的分割结果; Chen 等人(2023)开发了一种新的细化残差卷积网络,该网络利用不同膨胀率来捕捉更多的全局信息。

在分割阶段,使用残差特征选择 UNet (residual feature selection-UNet, RFS-UNet) 对 BTEC-Net 分类,为异常的图像进行专属分割, He 等人(2023)提出了一种混合 CNN-Transformer 网络 (hybrid CNN-Transformer network, HCTNet) 来提高超声图像中乳腺病变的分割。但是实际的临床医学图像分辨率不高,并且上下文不相连,很少关注不同区域内部和相邻之间的联系,忽略整个图像的全局特征和远程信息交互,导致对肿瘤分割存在误差。在本次实验中,不仅将 Transformer 结构与普通的卷积层之间作为模块相连,更关注在不同采样之间使用 Transformer 来融合特征信息,高效地将两个模块相结合,从而在模型训练中各自体现网络的优势,识别的乳腺肿块很少有分割和分类相继进行,对分割的不同类型的肿瘤

缺少评定。本文提出 TransAS-UNet (Transformer atrous spatial pyramid pooling Swin UNet) 网络, 实现不同层次的特征信息交融和卷积层局部信息的自我关注, 同时提出一种新的多尺度特征图融合策略, 采用具有跳跃连接的 Transformer 增强了空间域信息表示。实验表明, 本文提出的分割算法优于对比方法, 性能指标得到提升。

本文的主要贡献如下: 1) 提出新的高效分类网络 TransAS-UNet, 在上下采样阶段基于 UNet 分割网络引入 Swin 和 ASPP (atrous spatial pyramid pooling), 可以实现对不同层次的深浅特征信息交融, 增加卷积层的感受野来达到局部信息的自我关注。2) 为了准确地捕获空间域信息, 通过在跳跃连接上引入 Transformer 来增强空间域信息, 构建了一种新的多尺度特征图融合策略, 以增强高维和低维空间信息的融合。3) 在分割出乳腺癌的肿块上采用第 2 阶段分类网络, 基于 ResNet50 (residual network 50) 上使用 SK (selective kernel) 卷积代替模型中间卷积层, 并结合 SE (squeeze-and-excitation) 对输入图像的通道分配权重, 使用 ERF (error function) 梯度下降和标签软化等方法实现精准分类。4) 本文模型在乳腺癌 X 光数据集 INbreast 上实现肿块的准确分割, 交并比 (intersection over union, IoU) 值达到 95.58%, Dice 系数为 93.45%, 与其他分割模型相比提高了 4%~6%, 将分割的肿块进行四分类, Accuracy 值达到 95.24%。在许多临床医学分割和分类应用中, 本文方法比大多数模型具有更高的精度。

1 方法

实验中, 基于 UNet 结构的上下采样存在不同层次的特征信息交融 (徐磊等, 2023) 和卷积层局部信息的自我关注等问题, 本文采用了 Swin Transformer 结构, 该结构具有滑窗操作和层级设计, 主要由 Window Attention 模块和 Shifted Window Attention 模块组成, 可以实现将输入乳腺 X 光特征图切片成多个窗口, 根据自我注意力来分配每个窗口的权重, 同时也能对整个特征图的窗口位置进行移动, 实现同一特征图内部的信息交互。本文在上下采样过程中各使用 4 个 Swin Transformer 结构, 在切片融合中使用金字塔结构 ASPP (Chen 等, 2018) 替换普通的特征图通道相加操作, 能够运用多个卷积核对原图卷

积得到不同的特征图再通道相融, 给定的输入以不同采样率的空洞卷积并行采样, 实现多个比例捕捉图像的上下文信息。为了更好地融合高维和低维空间信息, 本文提出了一种新的多尺度特征图融合策略, 采用具有跳跃连接的 Transformer 增强了空间域信息表示。对于乳腺 X 光图像, 需要识别的肿瘤区域相对于整个乳腺部位占比较小, 通过增加局部信息关注可以有效训练。经过分割网络对乳腺癌肿块分割得到二值图像作为第 2 阶段分类网络的输入, 在基线模型 ResNet50 (He 等, 2016) 上融入 SK 和 SE 多种注意力, 通过非线性方法融合来自不同核的特征, 进而实现感受野不同尺寸的调整, 在数据过拟合中添加 ERF 梯度下降和标签软化方法。图 1 展示了本文实验采用的模型框架。

1.1 Conv 模块

为了处理复杂的医学图像分割任务, 本文使用卷积块作为编码器和解码器的每一层。该模块类似于小型 UNet 结构, 这样可以提取更多细节。卷积层进行的处理就是卷积运算 (文可等, 2023), 对于输入数据卷积运算以一定间隔滑动滤波器的窗口, 将各个位置上滤波器的元素和输入的对应元素相乘再求和, 并将这个结果保存到输出的对应位置。将这个过程在所有位置都进行一遍, 就可以得到卷积运算的输出。

在分割结构中, 卷积模块主要实现两个步骤: 对整幅图像的切片处理 (侯佩和齐亚莉, 2021) 和嵌入向量。首先利用二维卷积将 stride, kernel size 设置为 patch_size 大小, patch_size 在本次实验中前期预处理为 4, 这样经过卷积操作形成 4×4 个尺寸大小为 $(H/4, W/4)$ 特征图, 乳腺癌 X 光数据集中大部分的肿瘤较小并且像素值较大, 通过切片操作可以加快模型的训练和减小单次训练参数。接着对四维向量 $(N, C, H/4, W/4)$ 利用降维函数实现维度重建 $(N, C, H/4 \times W/4)$, 再交换通道维度和特征图维度的位置, 最后经过归一化函数计算每幅特征图的均值和方差。具体卷积模块流程如图 2 所示。

1.2 STA (Swin Transformer and ASPP) 网络

随着 Transformer 在自然语言处理领域的发展, 其在计算机视觉方面也得到应用, 同时也存在很多挑战。一方面视觉实体变化大, 在不同场景下视觉 Transformer 性能未必很好; 另一方面图像分辨率高、像素点多, Transformer 基于全局自注意力的计算导致

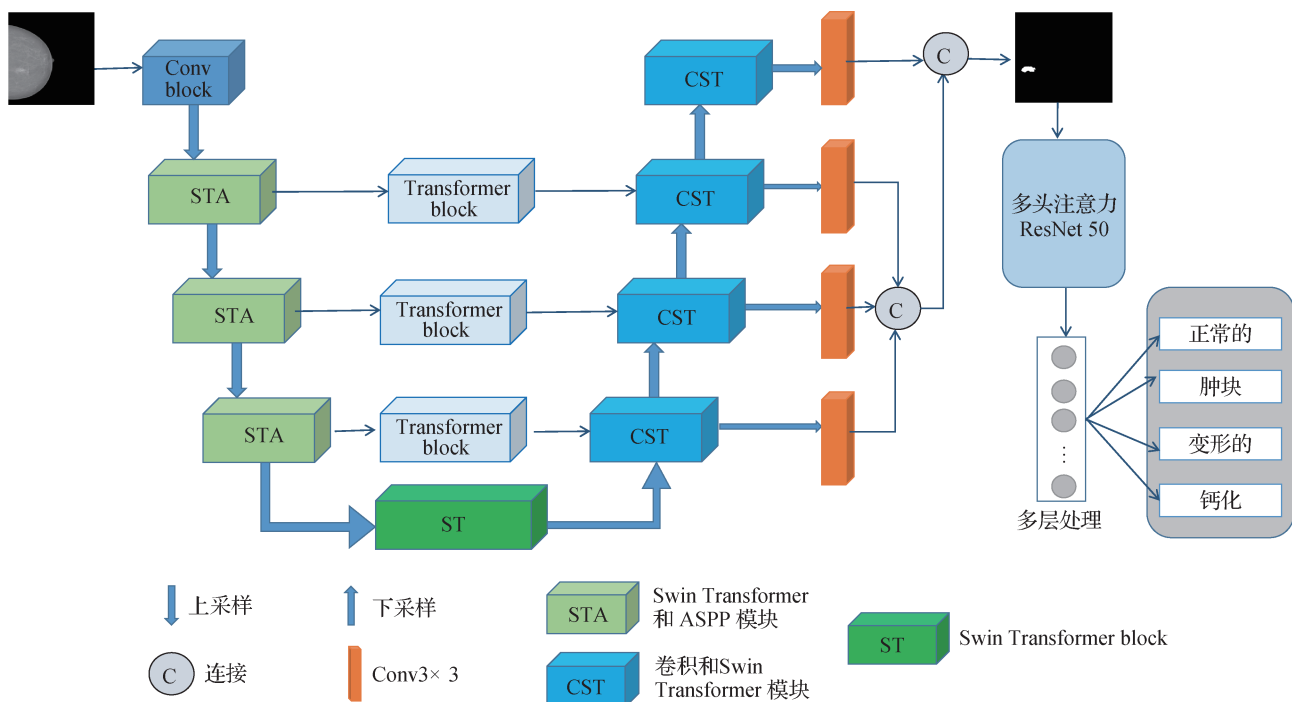


图 1 实验分割和分类的网络框架

Fig. 1 Network framework for experimental segmentation and classification

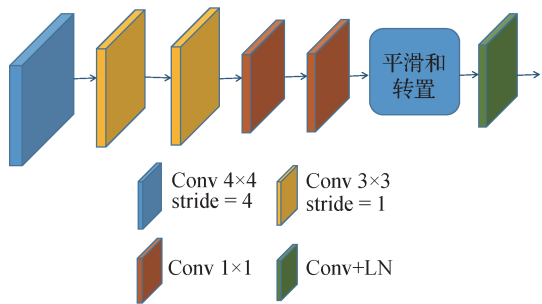


图 2 卷积模块

Fig. 2 Conv block

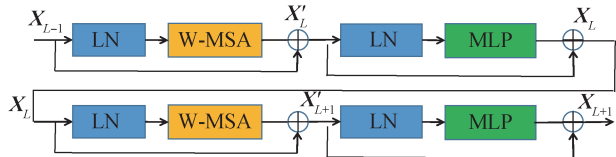


图 3 Swin Transformer 模块

Fig. 3 Swin Transformer block

$$S = F \left(\frac{QK^T}{\sqrt{d}} + B \right) V \tag{1}$$

计算量较大。Swin Transformer(Liu 等 2021)基于以上问题采用滑动窗口的方法来实现局部性的自我注意力,为了解决传统的Transformer都是基于全局来计算注意力的计算量大问题,Swin Transformer则将注意力的计算限制在每个窗口内,进而减少了计算量。利用层级设计将两级Transformer相连,可以使模型适合不同的视觉实体。图3显示了Swin Transformer结构,由归一化(layer normalization, LN)层、残余短连接、多层感知层(multi-layer perceptron, MLP)、基于窗口的多头自我注意力和基于移位的多头自我注意力组成。

多头自我注意力可以通过提供编码远距离依赖或异构交互的能力来补充骨干网络(Cao 等, 2019),与Transformer全局注意力计算方式不同的是加入了相对位置编码,计算为

式中, Q, K, V 分别表示 query 矩阵、key 矩阵和 value 矩阵, B 中的值取自偏置矩阵, F 代表 softmax 函数, S 代表计算权重。普通的切片注意力值计算是将其分为多个有序的窗口来分配权重,为了更好地与其他窗口进行信息交互,在 Swin Transformer 第 2 阶段是采用移位窗口,经过打乱切片的特征图可以学习到不相邻区域之间的信息。

如图 4 所示,乳腺 X 光片经过移位窗口操作带来的效果是窗口数增加。左边是没有重叠的 Window Attention,右边是将窗口进行移位的 Shift Window Attention。移位后的窗口包含了原本相邻窗口的元素,但也带来窗口的个数翻倍,由原本 4 个窗口变成了 9 个窗口。而最后计算方式是通过特征图移位,并给 Attention 设置 mask 来间接实现的,能在保持原有的窗口个数下实现结果等价。对于乳腺 X

光图通过移动方式得到9个窗口数,对每个窗口计算自注意力方式采用循环移位操作,如图5所示,整幅图像分为4部分,经过移位左上角切片移动到右下角,使得相邻的部分进行信息交互。

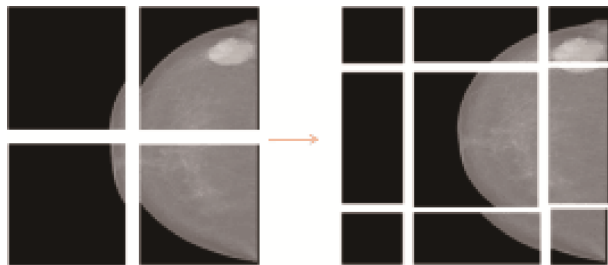


图4 移位窗口的实现方式

Fig. 4 Shift window implementation

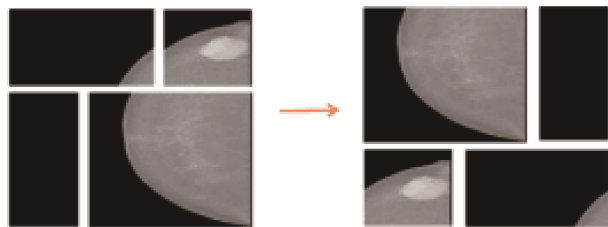


图5 移位循环操作

Fig. 5 Shift cycle operation

本文在编码和解码阶段都使用两个连续的 Swin Transformer,降低了普通分割网络 UNet 中卷积层的局限性,使网络更好地让全局和远程语义信息交互,既能学习到癌症肿块和其他肌肉之间的不同,又能寻找自身的特征信息。在进行不同层次特征信息提取之前,采用空洞卷积金字塔结构 ASPP(Zhao 等,2017)将不同切片之间的信息融合再送入下一个 Swin Transformer block。在每个阶段开始前做降采样,用于缩小分辨率,调整通道数,进而形成层次化的设计,同时也能节省一定运算量。图6是 ASPP 金字塔模块结构。该网络将 Swin Transformer 中的 patch merging 中的全连接层替换为4种卷积核进行通道数相加,其中包括普通的 1×1 卷积、padding 和 dilation 为6,核大小为 3×3 的卷积、padding 和 dilation 为12,核大小为 3×3 的卷积、padding 和 dilation 为18,核大小为 3×3 的卷积,最后将这4层的输出进行 concat,并用 1×1 卷积层降维至给定通道数,得到最终输出。在 3×3 的空洞卷积核中设置多个膨胀因子,目的是获取不同尺度的感受野,实现自由的多尺度特征提取。在下采样过程中,为了满足分割

网络的输入尺寸要求,通常采用切片的方式输入网络,而每个切片在 Swin Transformer 中自我注意力计算量过大,会设置不同的窗口分别计算,ASPP就是对多个窗口空洞卷积,在输入尺寸多变的情况下实现输出固定。

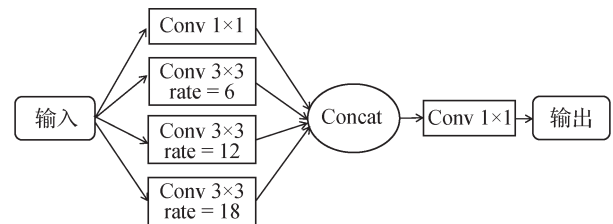


图6 ASPP结构

Fig. 6 ASPP structure

1.3 Transformer 模块

原始的 Transformer 应用在自然语言处理领域,解决了长距离的信息不能有效地提取和记忆,导致信息大量丢失和在引入 Attention 之后对关系的捕捉不足而出现翻译效果不理想等问题,Transformer 引入 self-attention 的机制将3种关系全部进行学习,提出 multi-head attention 机制,分别学习对应的3种关系,使用了全 Attention 的结构;对于词语的位置,Transformer 使用 positional encoding 机制进行数据预处理,增大了模型的并行性,取得了更好的实验效果。在本次实验中,不仅将 Transformer 结构与普通的卷积层之间作为模块相连,更关注在不同采样之间使用 Transformer 来融合特征信息,高效地将两个模块结合,从而在模型训练中各自体现网络的优势。之前的研究特征融合是通过简单地拼接特征来实现的,但特征金字塔(feature pyramid network, FPN)(Lin 等,2017)通过从网络架构的不同层提取不同尺度的特征来进行预测,从而为多特征融合创建了一种新的特征金字塔方法。此外,FastFCN(faster fully convolutional network)(Wu 等,2019)通过在卷积后组合不同大小的特征而更具语义。对于网络来说,高级和低级特征信息是互补的,需要高效地结合各自的优势。本文还在 Transformer 的输入部分引入位置编码赋予每个切片,然后利用 dropout 防止模型的过拟合,再将变换器块添加到跳跃连接,防止编码器部分在更丰富的全局卷积特征的情况下将特征图分辨率降得太低。图7显示了 Transformer 模块的具体实现过程。通过这种设计,防止乳腺癌 X 光图像经过

下采样过程像素值较高的癌症区域信息丢失,也是为了对正常乳腺组织和非正常组织的特征差异对比,从而实现有效分割。

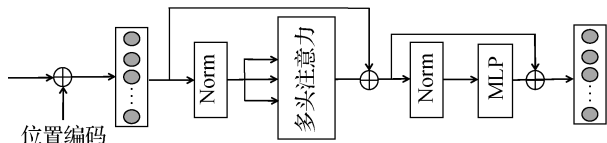


图7 Transformer 结构

Fig. 7 Transformer block

该模型中首先对输入的数据进行嵌入,将嵌入后的结果融入位置信息后输入到网络中,得到 X_{L-1} 。输入信息输入到注意力层中来捕获多维度的上下文信息,得到的张量参数经过短连接操作与输入信息融合,计算为

$$X_L' = M[N(X_L)] + X_{L-1} \quad (2)$$

$$X_L = M(X_L') + X_L \quad (3)$$

式中, M 代表多头注意力(multi-head attention, MHA)机制模块, X_L 和 X_{L-1} 代表 Transformer 模型的输出, N 和 M 是归一化和多层感知机操作,式中自注意力计算为

$$M = F\left(Q \times \sqrt{\frac{K}{d}}\right) \quad (4)$$

$$A = M \times V \quad (5)$$

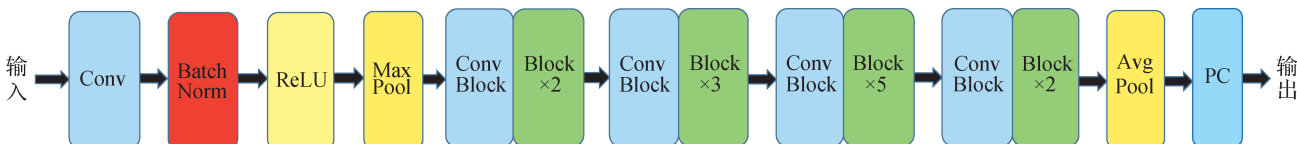


图8 ResNet50网络结构

Fig. 8 ResNet50 network structure

SE模块主要是由压缩和激励组成,在输入图像的通道位置上赋予一定的注意力权重,克服卷积运算局部感受野和通道所占的重要性不同带来的缺点,借助于卷积运算实现全局提取特征。SE模块结构如图9所示。

与SE模块相同的是,SK也是通道注意力,如图10所示,但是生成的注意力权重方式有差异,实现对不同尺寸的输入特征自适应调整感受野大小。在加入模型中,SK卷积是通过替换普通的卷积使网络在训练时选择合适的卷积尺寸,而SE则通过给予通道权重。SK主要分为分割、融合和选择操作(Xiang等, 2019)。

式中, Q 类比为询问, K 类比为索引, V 类比为回答, d 是特征值, F 代表通过 softmax 函数将原始权重标准化, A 是自注意力权重。

1.4 基于 ResNet50 多注意力网络

本文基线模型 ResNet50(He等, 2016)的网络结构如图8所示,网络中包含了49个卷积层、1个全连接层。该网络与其他模型相比不仅能够解决梯度消失问题,而且具有较少的参数数量和计算量。同时考虑到处理复杂的图像信息时,实验中加入了SE(Hu等, 2017)通道注意力机制,可以让网络聚焦于感兴趣的癌症区域,提高分类效果。然而标注的卷积神经网络中每个神经元的感受野是相同的,为了使神经元对于不同尺寸的输入信息能够自适应地调整其感受野的大小,本文采用SK(Xiang等, 2019)卷积替换普通 3×3 卷积,通过非线性的方法融合来自不同核的特征,进而实现感受野不同尺寸的调整。在数据处理方面,本文对部分类别较少的样本进行随机翻转和裁剪,再进行所有的数据归一化操作。同时还进行了标签软化,将真实的硬标签平滑处理以及将损失函数分配不同的注意力权重来平衡正负样本。在学习率衰减方式上,本文对比了随机梯度下降和高斯误差调度器。

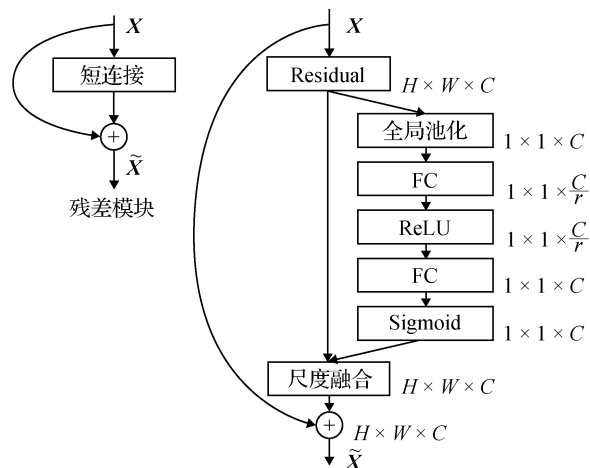


图9 SE网络结构

Fig. 9 SE network architecture

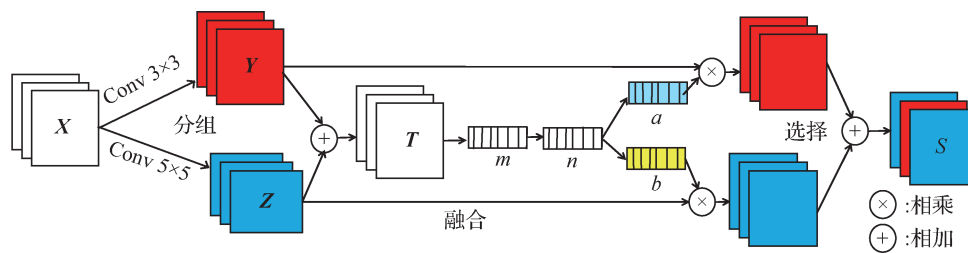


图 10 SK网络结构

Fig. 10 SK network architecture

本文在分割结果数据集上进行多模型对比,以 ResNet50 为基线模型,然后添加 SK 和 SE 注意力方法来对比实验指标,SK 卷积主要在每个 BottleNeck 替换 3×3 卷积,将每层相同大小感受野的卷积层转化为神经元对于不同尺寸的输入信息进行自适应调整感受野大小(马吉权等,2022),从而在卷积层可以提取更多的图像特征。通过在此基础上增加注意力 SE,同时引入高斯误差梯度下降法、标签平滑、部分数据增强技术和损失函数赋予权重 4 种方法来提高模型的精度。

2 实验

2.1 数据集

INbreast 数据库是一个乳腺摄影数据库,其中的图像是从大学医院(葡萄牙波尔图的乳腺癌中心,圣若昂医院)的乳腺癌中心获取的。INbreast 共有 115 例(410 幅图像),其中 90 例来自双乳女性(每例 4 幅图像),而 25 例来自乳房切除术患者(每例 2 幅图像)。包括几种类型的病变(肿块、钙化、不对称和变形),专家还以 XML(extensible markup language)格式提供了精确的轮廓。像素尺寸为 70 mm,对比度分辨率为 14 位。图像矩阵是 $3\,328 \times 4\,084$ 像素或 $2\,560 \times 3\,328$ 像素,图像以 DICOM(digital imaging and communications in medicine)格式保存。数据库包括正常乳房 X 光片、肿块乳房 X 光图、钙化乳房 X 光照片、乳房变形、不对称和多个发现的图像,如图 11 所示。

良性钙化大于与恶性肿瘤相关的钙化,通常较粗糙,为圆形缘光滑,更容易看到。与恶性肿瘤相关的钙化通常很小。结构畸变被定义为正常乳房 X 线摄影模式的局部中断,通常表现为星形畸变,没有明显的肿块。不对称性缺乏肿块的凸状外边界,可以

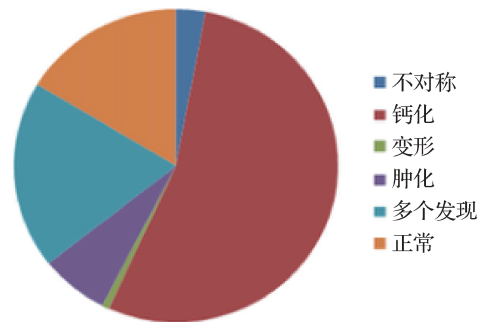


图 11 INbreast 数据库中的不同类型所占的比例饼状图

Fig. 11 Pie chart of the proportion of different types in the INbreast database

用 3 种方式表示:大小不对称、局部不对称和全局不对称。对所有数据集先进行 DCOM 到 PNG 的格式转化,然后作为模型训练的原图,根据其提供的不同类型癌症分割图作为标签,同时读取 XML 文件以确定每个癌症区域的精准位置。为了验证模型的有效性和在训练过程中是过拟合还是欠拟合,在实验中对数据集按 7:3 划分为训练集和测试集。图 12 显示了该数据集中的一部分乳腺癌 X 光图像。

2.2 评价指标

在分割和分类实验中,本文使用不同的评价指

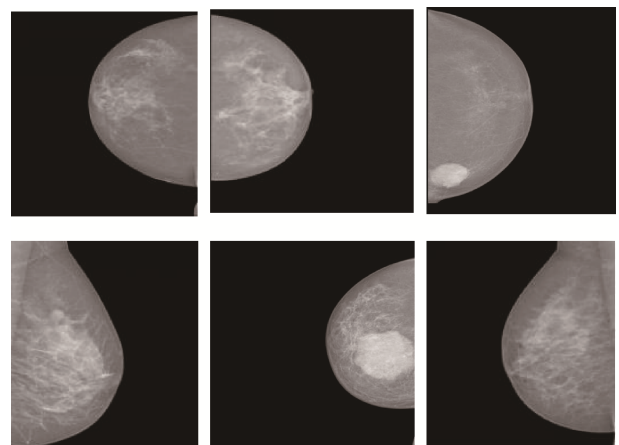


图 12 INbreast 部分乳腺癌 X 光图

Fig. 12 INbreast partial breast cancer X-ray

标来衡量网络模型的性能,同时与其他相关模型进行对比。主要评价指标为 Dice 系数、准确率、精确率、召回率和混淆矩阵。在评价指标中有 4 个主要变量, TP (true positives) 是真正类、 TN (true negatives) 是真负类、 FN (false negatives) 是假负类、 FP (false positive) 是假正类。

医学图像研究中最常用的指标是骰子系数 (Dice), 是一种集合相似度度量函数, 用于确定两个样本的相似程度, 在医学图像分割上它从边界获取细粒度信息至关重要, 其计算方法为

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

准确率可以衡量预测正确的正负样本占总样本的比例, 具体为

$$f_{\text{acc}} = \frac{TP + TN}{TP + FN + FP + TN} \quad (7)$$

为了解决样本不平衡性使得准确率不能完全评价模型的优劣性的局限, 本文采用部分数据增强方式使准确率具有代表性, 在多分类中准确率可以代表模型对所有类别的整体评价。

同时, 本文将交并比 (IoU) 作为分割评价标准, 对于图像分割而言是计算预测掩码和真实掩码之间的交并比, 具体为

$$f_{\text{IoU}} = \frac{TP}{TP + FP + FN} \quad (8)$$

为了衡量像素点集之间的相似度, 本文采用双向豪斯距离, 式 (9) 中 A 和 B 分别代表两组点集合, 为了排除离群点之间的影响, 使用 95% 分位最大值距离作为评价指标, 用 HD_{95} 表示, 其计算式为

$$h(A, B) = \max_{a \in A} \{ \min_{b \in B} \|a - b\| \} \quad (9)$$

$$H(A, B) = \max(h(A, B), h(B, A))$$

在分类模型中, 本文还将精确率和召回率用来衡量模型的有效性。精确率指当模型预测为正样本, 实际为正样本的概率, 计算式为

$$f_{\text{pre}} = \frac{TP}{TP + FP} \quad (10)$$

召回率是当真实标签为正样本, 预测为正样本所占的比例, 计算为

$$f_{\text{recall}} = \frac{TP}{TP + FN} \quad (11)$$

2.3 实验结果

2.3.1 肿瘤区域分割

对乳腺癌肿瘤区域分割的过程中, 实验中将转

化的图像格式和二值化标签送入分割网络, 通过 4 次采样图像分辨率减少和通道数成倍增加, 网络特征图从轮廓、形状等浅层特征到纹理、肿瘤的部位, 编码阶段以 STA 网络结构作为提取特征图的主要方法, 关注肿瘤的全局特征信息, 以增加卷积层的感受野, 获取更多的局部自我关注。在上采样过程中还原图像的分辨率, 与编码阶段同层次的特征信息进行交融送入 CST (convolution and Swin Transformer) 模块中, 每层输出的特征图都可以作为分割图像的结果, 因此本文利用多尺度融合策略避免重要信息的丢失。最后的肿瘤分割效果可以准确地对 X 光图像区域进行分割, 同时与其他分割模型 UNet、UNet++、Res18_UNet、MultiRes_UNet (Ibtehzah 和 Rahman, 2020)、Dense_UNet (Kaku 等, 2019) 5 种网络进行对比, 分割效果图更准确, 边缘轮廓更清楚, 如图 13 所示。

为了网络训练更加鲁棒, 模型参数更加准确, 本文在网络模型中设置的训练轮数为 1 500, 但大多数模型在前 1 000 轮就已经取得比较好的分割结果, 后续的训练结果反而变差。模型的其他参数都保持一致, 其中 batch_size 为 4, 无监督, 学习率为 0.001, 下降方式为余弦函数图像下降法, 优化器为随机梯度下降法。输入数据做统一预处理, 主要是归一化、随机旋转 90 度、上下和水平翻转、亮度增强等数据增强方法, 损失函数为交叉熵。本文首先对比了不同模型在 INbreast 数据集上训练集和测试集的结果, 以 7:3 划分总的数据集, 以 IoU 为指标可以看出, 训练数据在模型中表现良好, 但在测试集上 IoU 与训练集相差 10%~15%, 如表 1 所示。

经过多次实验得出, 指标的差异性原因是模型的过拟合, 如图 14 所示, 模型的过拟合问题与模型结构关系不是很明显, 所以重要原因是模型设置的参数和优化过拟合的方法。本文在此基础上改变学习率的变化方式, 以 0.1 为初始值, 设置每 10 轮指标未得到提升, 就开始学习率衰减, 使用的衰减方式为 ERF 下降法, 同时在计算损失函数时, 输出像素值利用标签软化的方式将其不局限于 0 和 1, 缩小损失值, 在数据增强上增加多种方法来减少过拟合现象。

利用上述方法解决过拟合问题, 本文在不同模型上重新训练, 以 Dice 和 IoU 作为评价标准, 结果如表 2 所示。可以看出, TransAS-UNet 相比于其他 5 种模型, IoU 指标有 4%~6% 的提高, 达到 95.58%,

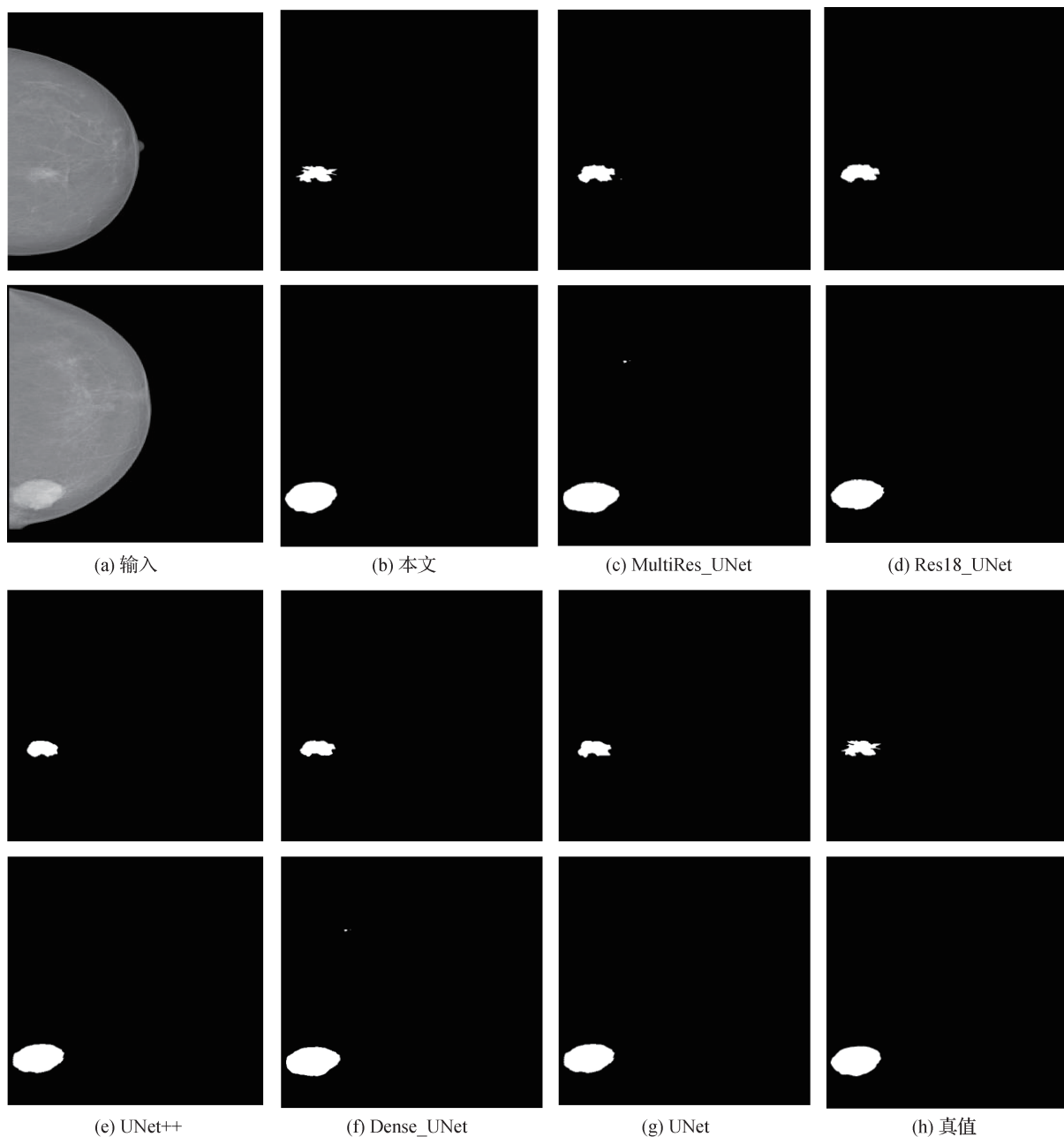


图 13 不同模型分割对比图

Fig. 13 Comparison of segmentation results of different models

((a)input; (b)our; (c)MultiRes_UNet; (d)Res18_UNet; (e)UNet++; (f)Dense_UNet; (g)UNet; (h)ground truth)

Dice 系数为 93.45%，在降低 HD_{95} 方面相比于其他模型也具有优势。从实验中可以看出，本文方法能够取得最佳结果，利用 Swin-Transformer、金字塔结构 ASPP、将 Transformer 作为同层之间的短连接以及构建多尺度融合策略能促进乳腺癌肿瘤区域的分割。本文方法优于对比方法。

2.3.2 实验对比

为了探究视觉大模型 SAM (segment anything

model) 在乳腺癌 X 光数据集的有效性，本文将其放在同等设备环境中进行模拟，结果如表 3 所示。可以发现，IoU 及其他指标不如 TransASwin-UNet 模型，说明 SAM 模型 (Kirillov 等, 2023) 在涉及小目标分割上存在弊端，同时对于尚未见过的乳腺癌 X 光图像缺少泛化能力。

同时，为了进一步验证模型的有效性和泛化能力，选用 MIAS 乳腺癌 X 光数据集进行对比实验，该

表1 不同模型IoU对比

Table 1 Comparison of IoU among different models

模型	Train_IoU	Test_IoU
UNet	0.940 0	0.850 0
UNet++	0.931 6	0.779 0
Res18_UNet	0.944 2	0.799 6
MultiRes_UNet	0.934 2	0.782 0
Dense_UNet	0.934 2	0.778 2
TransAS-UNet	0.965 8	0.873 3

注:加粗字体表示各列最优结果。

数据集包含 161 名患者的左右乳房图像,由 322 幅图像组成,主要包含 3 种类型:正常 208 个、良性 63 个和恶性(异常)51 个。选用在 INbreast 数据集上效果较好的 UNet、Res18_UNet、Dense_UNet 3 种模型进行对比,最终对比实验结果如表 4 所示。由实验结果可以看出,本文 TransAS-UNet 模型在提升 IoU 和减少 HD₉₅ 距离上效果明显,说明该算法在乳腺癌数据集上具有良好的分割性能,通过结合局部信息重要特征的关注和全局信息之间的交互可以实现肿瘤的有效分割。

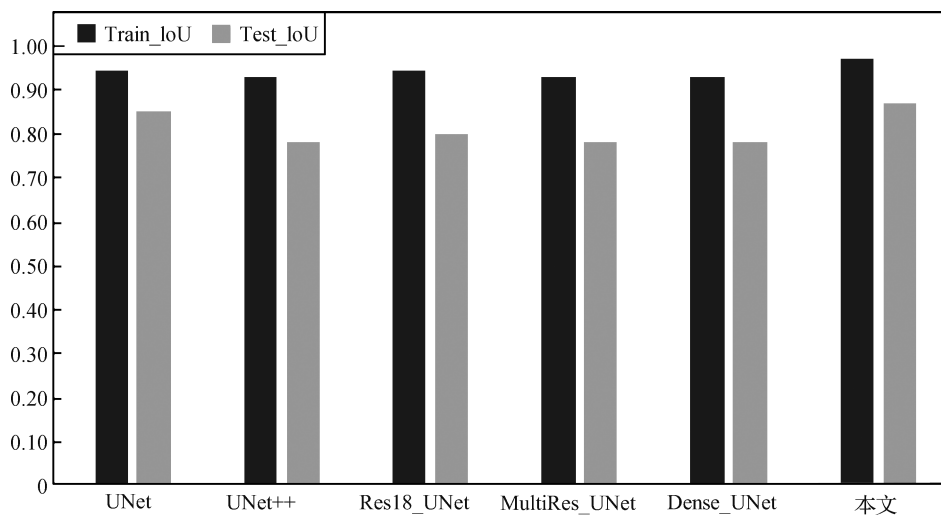


图 14 不同模型在训练集和测试集上 IoU 对比

Fig. 14 Comparison of IoU among different models in training set and test set

表 2 不同模型测试集上的指标对比

Table 2 Comparison of indicators of different models on test set

模型	IoU	Dice	HD ₉₅
UNet	0.911 2	0.916 6	22.145
UNet++	0.894 5	0.900 0	25.542
Res18_UNet	0.912 0	0.912 6	31.128
MultiRes_UNet	0.900 1	0.892 3	20.827
Dense_UNet	0.902 2	0.886 9	26.379
TransAS-UNet	0.955 8	0.934 5	20.242

注:加粗字体表示各列最优结果。

表 3 提出的模型与大模型 SAM 实验效果对比

Table 3 Comparison of experiment results between the proposed model and the large model SAM

模型	IoU	Dice	HD ₉₅
SAM	0.927 6	0.905 5	36.558
TransAS-UNet	0.955 8	0.934 5	20.242

注:加粗字体表示各列最优结果。

表 4 在 MIAS 数据集上实验效果对比

Table 4 Comparison of experimental effects on MIAS dataset

模型	IoU	HD ₉₅
UNet	0.916 8	28.448
Res18_UNet	0.933 0	16.220
Dense_UNet	0.925 3	25.826
TransAS-UNet	0.966 4	17.454

注:加粗字体表示各列最优结果。

2.3.3 肿瘤类型识别

经过分割模型得到较准确的癌症区域二值图,根据 INbreast 数据集介绍可以找出每幅癌症图像所属的类别,本文将其分为正常、肿块、变形和钙化,对每个类别给予标签,再送入分类网络(阮旭凌等,

2022)。本文采用的分类模型以 ResNet50 为基线模型,通过在此基础上增加两种不同的注意力 SE、SK, SK 卷积主要在每个 BottleNeck 上替换 3×3 卷积,将每层相同大小感受野的卷积层转化为神经元,对于不同尺寸的输入信息自适应调整感受野大小,从而在卷积层可以提取更多的图像特征;而 SE 属于通道注意力,可以在像素值输出之前对各个通道赋予权重。同时本文引入了高斯误差梯度下降法、标签平滑和部分数据增强技术 3 种方法来提高模型的精度。高斯误差梯度下降法解决随机梯度在梯度下降时具有随机性的问题,每次迭代的损失函数虽然朝着全局最优但并不是局部最优,而我们既要寻找全局最优又要局部最优;标签平滑是利用软标签解决预测错误导致误差大、增加模型的泛化能力和解决过拟合问题;部分数据增强技术缓解四分类中样本不平衡,减少少量样本的预测误差。

本文的实验结果在 INbreast 公开数据集上采用 SK-SE-ResNet50 准确率达到 95.24%,精确度达到 94.80%,召回率为 95.80%。由表 5 可见,增加注意力可以促进网络更快提取特征来区分不同类别的差异。

表 5 不同模型分类指标对比
Table 5 Comparison of classification indexes of different models

模型	/%		
	准确率	精确度	召回率
ResNet50(SGD)	87.34	88.47	86.39
SK-ResNet50	89.68	90.38	88.49
SK-SE-ResNet50	90.23	91.24	90.18
SK-SE-ResNet50(ERF)	93.65	93.49	93.23
SK-SE-ResNet50(标签平滑)	92.89	94.80	93.40
SK-SE-ResNet50(部分数据增强技术)	95.24	93.89	95.80

注:加粗字体表示各列最优结果。

3 结论

本文提出了一种新的基于 UNet 和 Transformer 的 TransAS-UNet 结构的乳腺癌区域分割和分类方法,该结构主要分为肿瘤区域的分割和肿瘤类型的识别。在分割过程中类比于 UNet 结构模型,将 Swin

Transformer 和 ASPP 模块替换普通的卷积层,利用移位窗口和多头注意力达到图像切片内部的特征信息交融和不相邻区域之间的信息互补,同时 ASPP 结构可以在增加感受野的情况下达到局部信息的自我关注。在不同层之间本文引入 Transformer 结构来关联信息,防止在下采样卷积中浅层重要信息的丢失,最终的体系结构不仅继承了 Transformer 在学习全局语义关联方面的优势,还使用了不同级别的特征,以使模型保留更多的语义和更多的细节。分割模型得到的二值化图像作为分类网络的输入数据集,本文可以实现不同类别的乳腺癌肿瘤的分类,该分类模型以 ResNet50 为基础,添加多类型注意力模块和防止过拟合操作,SE、SK 两种注意力可以优化网络参数,使其只关注分割区域差异,从而提高模型的效率。

本文模型在乳腺癌 X 光数据集 INbreast 上实现肿块体的准确分割,并与 UNet、UNet++、Res18_UNet、MultiRes_UNet、Dense_UNet 5 种分割结构进行对比,在同一个参数环境中,IoU 值达到 95.58%,Dice 系数为 93.45%,与其他的分割模型相比提高了 4%~6%,本文将分割的肿块进行四分类,Accuracy 值达到 95.24%。

在未来,将聚焦于实现乳腺癌分割和分类网络一体化,网络参数和结构可以不用二阶段分步训练,提高网络的效率。同时,将在分割数据集预处理上采用 ROI,这样可以避免无用特征信息的多次提取。为了验证模型的广泛性,后面的工作也会在其他癌症数据集上进行实验。相信利用深度学习技术可以实现医学上更多难关突破,减少医生的工作量,给临床带来更大的价值。

参考文献 (References)

- Bray F, Ferlay J, Soerjomataram I, Siegel R I, Torre L A and Jemal L. 2018. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: A Cancer Journal for Clinicians, 68(6): 394-424 [DOI: 10.3322/caac.21492]
- Cao Y, Xu J R, Lin S, Wei F Y and Hu H. 2019. GCNet: non-local networks meet squeeze-excitation networks and beyond//Proceedings of 2019 IEEE/CVF International Conference on Computer Vision Workshop. Seoul, Korea (South): IEEE: 1971-1980 [DOI: 10.48550/arXiv.1904.11492].
- Cardoso J S, Marques N, Dhungel N, Carneiro G and Bradley A P.

2017. Mass segmentation in mammograms: a cross-sensor comparison of deep and tailored features//Proceedings of 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China: IEEE: 1737-1741 [DOI: 10.1109/ICIP.2017.8296579]
- Chen G P, Dai Y and Zhang J X. 2023. RRCNet: refinement residual convolutional network for breast ultrasound images segmentation. *Engineering Applications of Artificial Intelligence*, 117: #105601 [DOI: 10.1016/j.engappai.2022.105601]
- Chen L C, Papandreou G, Kokkinos I, Murphy K and Yuille A L. 2018. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 40 (4): 834-848 [DOI: 10.48550/arXiv.1606.00915]
- Cho S W, Baek N R and Park K R. 2022. Deep learning-based multi-stage segmentation method using ultrasound images for breast cancer diagnosis. *Journal of King Saud University—Computer and Information Sciences*, 34 (10): 10273-10292 [DOI: 10.1016/j.jksuci.2022.10.020]
- George M J and Sankar S P. 2017. Efficient preprocessing filters and mass segmentation techniques for mammogram images//Proceedings of 2017 IEEE International Conference on Circuits and Systems (ICCS) Thiruvananthapuram, India: IEEE: 408-413 [DOI: 10.1109/ICCS1.2017.8326032]
- He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition//Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, USA: IEEE: 770-778 [DOI: 10.1109/CVPR.2016.90]
- He Q Q, Yang Q J and Xie M H. 2023. HCTNet: a hybrid CNN-Transformer network for breast ultrasound image segmentation. *Computers in Biology and Medicine*, 155: #106629 [DOI: 10.1016/j.combiomed.2023.106629]
- Hou P and Qi Y L. 2021. Automatic region segmentation method of breast tumors based on deep neural networks. *Journal of Biomedical Engineering Research*, 40(3): 241-245 (侯佩, 齐亚莉. 2021. 基于深度神经网络的乳腺肿瘤自动区域分割方法. *生物医学工程研究*, 40(3): 241-245) [DOI: 10.19529/j.cnki.1672-6278.2021.03.03]
- Hu J, Shen L, Albanie S, Sun G and Wu E H. 2017. Squeeze-and-excitation networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42 (8): 2011-2023 [DOI: 10.4850/arXiv.1709.01507]
- Huang H M, Lin L F, Tong R F, Hu H J, Zhang Q W, Iwamoto Y, Han X H, Chen Y W and Wu J. 2020. UNet 3+: a full-scale connected UNet for medical image segmentation//Proceedings of ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). Barcelona, Spain: IEEE: 1055-1059 [DOI: 10.1109/ICASSP40776.2020.9053405]
- Ibtehaz N and Rahman M S. 2020. MultiResUNet: rethinking the U-Net architecture for multimodal biomedical image segmentation. *Neural Networks*, 121: 74-87 [DOI: 10.1016/j.neunet.2019.08.025]
- Iqbal A and Sharif M. 2023. PDF-UNet: a semi-supervised method for segmentation of breast tumor images using a U-shaped pyramid-dilated network. *Expert Systems with Applications*, 221: #119718 [DOI: 10.1016/j.eswa.2023.119718]
- Jai-Andaloussi S, Sekkaki A, Quellec G, Lamard M, Cazuguel G and Roux C. 2013. Mass segmentation in mammograms by using bidimensional empirical mode decomposition BEMD//Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Osaka, Japan: IEEE: 5441-5444 [DOI: 10.1109/EMBC.2013.6610780]
- Kaku A, Hegde C V, Huang J, Chung S, Wang X Y, Young M, Radmanesh A, Lui Y W and Razavian N. 2019. DARTS: DenseUnet-based automatic rapid tool for brain segmentation [EB/OL]. [2023-03-24]. https://arxiv.org/pdf/1911_05567.pdf
- Ke L, He W and Kang Y. 2009. Mass auto-detection in mammogram based on wavelet transform modulus maximum//Proceedings of 2009 Annual International Conference of the IEEE Engineering in Medicine and Biology Society. Minneapolis, USA: IEEE: 5760-5763 [DOI: 10.1109/IEMBS.2009.5332615]
- Kirillov A, Mintun E, Ravi N, Mao H Z, Rolland C, Gustafson L, Xiao T T, Whitehead S, Berg A C, Lo W Y, Dollár P and Girshick R. 2023. Segment anything [EB/OL]. [2023-03-24]. <https://arxiv.org/pdf/2304.02643.pdf>
- Xiang L, Wang W H, Hu X L and Yang J. 2019. Selective kernel networks//Proceedings of 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Long Beach, USA: IEEE: 510-519 [DOI: 10.1109/CVPR.2019.00060]
- Lin T Y, Dollár P, Girshick R, He K M, Hariharan B and Belongie S. 2017. Feature pyramid networks for object detection//Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, USA: IEEE: 2117-2125 [DOI: 10.4850/arXiv.1612.03144]
- Liu Z, Lin Y T, Cao Y, Hu H, Wei Y X, Zhang Z, Lin S and Guo B N. 2021. Swin Transformer: hierarchical vision Transformer using shifted windows [EB/OL]. [2023-03-24]. <https://arxiv.org/pdf/2103.14030.pdf>
- Ma J Q, Zhao S M and Kong F H. 2022. Semantic image segmentation by using multi-scale strip pooling and channel attention. *Journal of Image and Graphics*, 27(12): 3530-3541 (马吉权, 赵淑敏, 孔凡辉. 2022. 多尺度条形池化与通道注意力的图像语义分割. *中国图象图形学报*, 27(12): 3530-3541) [DOI: 10.11834/jig.210359]
- Nelson A D and Krishna S. 2023. An effective approach for the nuclei segmentation from breast histopathological images using star-

- convex polygon. *Procedia Computer Science*, 218: 1778-1790 [DOI: 10.1016/j.procs.2023.01.156]
- Ronneberger O, Fischer P and Brox T. 2015. U-net: convolutional networks for biomedical image segmentation//*Proceedings of the 18th International Conference on Medical Image Computing and Computer-Assisted Intervention*. Munich, Germany: Springer: 234-241 [DOI: 10.1007/978-3-319-24574-4_28]
- Ruan X L, Liu Q, Guo Z H and Yan J F. 2022. Research on breast cancer prediction model. *Journal of Medical Informatics*, 43(5): 34-39 (阮旭凌, 刘琦, 郭志恒, 晏峻峰. 2022. 乳腺癌预测模型构建研究. *医学信息学杂志*, 43(05): 34-39)
- Saad G, Khadour A and Kanafani Q. 2016. ANN and Adaboost application for automatic detection of microcalcifications in breast cancer. *The Egyptian Journal of Radiology and Nuclear Medicine*, 47(4): 1803-1814 [DOI: 10.1016/j.ejrn.2016.08020]
- Salih A M and Kamil M Y. 2018. Mammography image segmentation based on fuzzy morphological operations//*Proceedings of the 1st Annual International Conference on Information and Sciences (AiCIS)*. Fallujah, Iraq: IEEE: 40-44 [DOI: 10.1109/AiCIS.2018.00020]
- Sun H, Li C, Liu B Q, Liu Z Y, Wang M Y, Zheng H R, Feng D D and Wang S S. 2020. AUNet: attention-guided dense-upsampling networks for breast mass segmentation in whole mammograms. *Physics in Medicine & Biology*, 65(5): #055005 [DOI: 10.18550/arXiv.18.10151]
- Wen K, Jin X, An H, He J and Wang J. 2023. CentroidNet: a lightweight, fast nuclei centroid detection model for breast Ki67 scoring. *Journal of Image and Graphics*, 28(04): 1119-1133 (文可, 金旭, 安虹, 何杰, 王珏. 2023. CentroidNet: 轻量快速的乳腺癌 Ki67 细胞核中心点检测模型. *中国图象图形学报*, 28(4): 1119-1133) [DOI: 10.11834/jig.211207]
- Wu H K, Zhang J G, Huang K Q, Liang K M and Yu Y Z. 2019. Fast-FCN: rethinking dilated convolution in the backbone for semantic segmentation [EB/OL]. [2023-03-24]. <https://arxiv.org/pdf/1903.11816.pdf>
- Xiao X, Lian S, Luo Z M and Li S Z. 2018. Weighted res-UNet for high-quality retina vessel segmentation//*Proceedings of the 9th International Conference on Information Technology in Medicine and Education (ITME)*. Hangzhou, China: IEEE: 327-331 [DOI: 10.1109/ITME.2018.00080]
- Xu L, Song H H and Liu Q S. 2023. Super-resolution reconstruction of binocular image based on multi-level fusion attention network. *Journal of Image and Graphics*, 28(4): 1079-1090 (徐磊, 宋慧慧, 刘青山. 2023. 多层次融合注意力网络的双目图像超分辨率重建. *中国图象图形学报*, 28(4): 1079-1090) [DOI: 10.11834/jig.211119]
- Yang X, Wang R, Zhao D, Yu F H, Heidari A A, Xu Z Z, Chen H L, Algarni A D, Elmannai H and Xu S L. 2023. Multi-level threshold segmentation framework for breast cancer images using enhanced differential evolution. *Biomedical Signal Processing and Control*, 80(2): #104373 [DOI: 10.1016/j.bspc.2022.104373]
- Zhang Y, Tomuro N, Furst J and Raicu D S. 2010. Image enhancement and edge-based mass segmentation in mammogram//*Proceedings Volume 7623, Medical Imaging 2010: Image Processing*, 11(2). San Diego, USA: SPIE: 1452-1459 [DOI: 10.1117/12.844492]
- Zhao H S, Shi J P, Qi X J, Wang X G and Jia J Y. 2017. Pyramid scene parsing network//*Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, USA: IEEE: 6230-6239 [DOI: 10.48550/arXiv.162.01105.]
- Zhou Z W, Rahman Siddiquee M M, Tajbakhsh N and Liang J M. 2018. UNet++: a nested U-Net architecture for medical image segmentation. *Computer Vision and Pattern Recognition [EB/OL]*. [2023-03-24]. <https://arxiv.org/pdf/1807.101165.pdf>

作者简介

徐旺旺,男,硕士研究生,主要研究方向为深度学习与医学影像交叉领域。E-mail: wwxu@mail.hfut.edu.cn

律娜,通信作者,女,副主任医师,主要研究方向为医学影像和机器学习应用。E-mail: lvna@ahmu.edu.cn

许良凤,女,副教授,主要研究方向为深度学习与医学影像交叉领域。E-mail: 1903723035@qq.com

李博凯,男,硕士研究生,主要研究方向为微表情检测定位及识别。E-mail: 18756526100@163.com

周曦,男,讲师,主要研究方向为计算机视觉人工智能机器学习应用。E-mail: 45937753@qq.com

詹曙,男,教授,主要研究方向为深度学习与医学影像交叉领域。E-mail: shu_zhan@hfut.edu.cn