

# 双阶段噪声自适应超声图像分割网络<sup>\*</sup>

王霞霞 张 焕 吴兴隆 黄 青 徐国平

(武汉工程大学计算机科学与工程学院 武汉 430205)

**摘要:** 在医学图像分割领域,不同设备或参数的噪声干扰对模型的泛化产生负面影响。为解决由不同噪声源导致的分割模型性能下降问题,提出了一种自适应噪声污染的双阶段高效分割模型 NANet。该模型主要包括两部分:去噪和分割。其中去噪部分采用无监督的 U 型自动编码器结构,结合频域可学习的去噪模块,以缓解噪声引起的目标域差异。分割部分基于 DeepLabv3+ 架构构建模块,聚焦于提高噪声污染超声图像中目标的分割准确度和泛化能力。实验测试了 4 个不同的超声数据集在受到均匀、高斯和椒盐噪声污染下,模型的分割性能。实验结果表明,所提出的 NANet 在臂丛神经超声数据集的原始图像、受均匀噪声、高斯噪声以及椒盐噪声污染的图像数据上分割的 Dice 系数分别为 79.71%、78.51%、79.75% 和 79.1%。对比实验显示, NANet 在不同噪声类型的超声图像上的分割性能明显优于传统分割方法,其中在高斯和椒盐噪声污染的超声图像上,准确率相比 U-Net、DeepLabv3+ 和 AttUNet 方法均提高超过 10%。结果验证了所提出的 NANet 在不同噪声环境中的鲁棒性和泛化能力。

**关键词:** 医学图像;图像分割;图像去噪;噪声自适应

中图分类号: TP389.1; TN919.81 文献标识码: A 国家标准学科分类代码: 520.20

## Two-stage noise adaptation network for medical image segmentation

Wang Xiaxia Zhang Xuan Wu Xinglong Huang Qing Xu Guoping

(School of Computer Science and Engineering, Wuhan Institute of Technology, Wuhan 430205, China)

**Abstract:** In the field of medical image segmentation, noise interference from different devices or parameters negatively affects the generalization of the model. To solve the problem of segmentation model performance degradation caused by different noise sources, a two-stage efficient segmentation model NANet with adaptive noise pollution is proposed. The model mainly consists of two parts: denoising and segmentation. The denoising part uses an unsupervised U-shaped autoencoder structure combined with a frequency-domain learnable denoising module to mitigate the noise-induced target-domain differences. The segmentation part builds modules based on the DeepLabv3+ architecture, focusing on improving the segmentation accuracy and generalization ability of targets in noise-polluted ultrasound images. Experiments were conducted to test the segmentation performance of the model on four different ultrasound datasets contaminated with uniform, Gaussian and pretzel noise. The experimental results show that the Dice coefficients of the proposed NANet for segmentation on the original image of the brachial plexus neural ultrasound dataset, the image data contaminated by uniform, Gaussian, and pretzel noises are 79.71%, 78.51%, 79.75%, and 79.1%, respectively. Comparison experiments show that the segmentation performance of NANet on ultrasound images with different noise types is significantly better than that of traditional segmentation methods, in which the accuracy is improved by more than 10% compared to U-Net, DeepLabv3+, and AttUNet methods on both Gaussian and pretzel noise images. The results validate the robustness and generalization ability of the proposed NANet in different noise environments.

**Keywords:** medical images; image segmentation; image denoising; noise adaptation

## 0 引言

近年来,计算机辅助诊断技术在各类医学图形分割任务中得到广泛研究<sup>[1]</sup>,例如磁共振成像、计算机断层扫描、超声等。深度学习技术依赖于高质量的训练数据和其强大的特征学习能力,在计算机辅助技术中取得了显著成果<sup>[2]</sup>。

收稿日期:2024-04-09

\* 基金项目:东省数字孪生人重点实验室项目(2022B1212010004)、智能机器人湖北省重点实验室项目(HBIRL202202)、中央高校基本科研业务费专项(PA2023IISL0095)、武汉工程大学第十五届研究生教育创新基金(CX2023319)、湖北省自然科学基金(2022CFB804)项目资助

由于成像设备本身、环境光线条件和信号传输等因素影响,真实世界的成像条件通常会在采集的图像中引入过多的噪声类型和电平<sup>[3]</sup>。

图像中存在的噪声给深度学习模型的训练和性能带来了多方面的挑战,包括特征学习的困难<sup>[4]</sup>和模型过拟合<sup>[5]</sup>等问题。传统方法尝试将数字图像滤波技术应用于医学图像去噪。中值滤波不受相邻像素极值的影响,在去除椒盐噪声方面表现出色。袁乐民<sup>[6]</sup>的研究表明中值滤波在医学图像去噪中效果显著,但对高密度噪声的抑制效果较差。张雪峰等<sup>[7]</sup>提出的基于分数阶的中值滤波算法能有效去除椒盐噪声并保留边缘细节,但对高斯噪声不敏感。然而,当中值滤波器邻域较大时图像会变得过于平滑且边缘细节无法保留,因此无法满足需求。小波滤波的多分辨率和时频局部化特性使其能够在去除噪声的同时很好地保留图像边缘细节。Wu 等<sup>[8]</sup>使用小波滤波有效缓解了数据分布梯度分布不准确的问题,但难以检测不同类型的噪声信号。陈军<sup>[9]</sup>研究了小波技术在抑制医学超声图像高斯噪声中的应用,但在噪声密度较高时图像边缘会消失。刘源等<sup>[10]</sup>将小波与非局部均值滤波融合有效抑制图像中的相干噪声,但对大噪声的抑制效果不佳。双边滤波<sup>[11]</sup>作为一种非线性滤波不仅考虑位置影响,还能同时保留边缘和降低噪声。在超声图像去噪中比各向异性扩散滤波保留了更多边缘信息。然而,加权平均法会过度平滑许多关键诊断信息。

此外,目前通用的分割架构在不同噪声图像上也会受到较高限制,导致模型在训练时过度拟合噪声数据忽略真实的图像特征;或者噪声导致分割边界模糊,使得模型难以准确地分割目标。经典的分割架构(U-Net<sup>[12]</sup>、DeepLabv3+<sup>[13]</sup>、LinkNet<sup>[14]</sup>以及 Transformer<sup>[15]</sup>)在受不同噪声污染的超声图像数据和原始低噪数据上测试时,噪声的引入会降低模型分割的性能。Wang 等<sup>[16]</sup>在分割算法中引入噪声模型,通过滤波或优化方法进行噪声鲁棒的分割;Aksam 等<sup>[17]</sup>通过非局部均值滤波器(non-local means filter, NLM)结合图像的纹理和相似性信息来进行噪声抑制和分割;Boykov 等<sup>[18]</sup>基于图像先验知识来约束分割结果,以减少噪声影响;Fu 等<sup>[19]</sup>和 Woo 等<sup>[20]</sup>则利用注意力机制提高分割网络对噪声的鲁棒性。但是随着噪声密度和种类的增加,这些方法的分割性能均会出现不同程度的下降。

为了解决上述提到的噪声影响分割的问题,本文构建了一种新型的双阶段基准网络(baseline network, BaseNet),并在此基础上提出了一个频域可学习的去噪模块,最终提出了一个双阶段噪声自适应的超声医学图像分割模型(noise adaptive network, NANet)。具体来说,该模型包含 2 个阶段:去噪和分割。在去噪阶段采用 U 型的自编码器架构,并在编码器和解码器中加入频域可学习的去噪模块(learnable frequency fourier transform, LFFT)以缓解特征图中不同频域的噪声影响。在分割阶段将去噪阶段得到的特征图作为输入,采用 DeepLabv3+ 网络进行精确

分割,保证模型的稳定性并提取更加细致的特征细节以提高模型的分割性能。为了验证提出网络的稳定性和有效性,本文选取了 4 种不同病灶类型的超声图像数据集。在保持训练集不变的情况下,在测试集的图像中分别加入均匀、高斯和椒盐噪声来模拟真实的环境干扰,并利用训练好的模型分别对其进行测试。实验结果证明了本文提出的模型在处理噪声超声图像时的优越性和有效性。

## 1 双阶段噪声自适应图像分割网络

本文提出了一种用于超声图像分割任务的双阶段噪声自适应网络-NANet,它可以对不同类型和数量的噪声目标域进行精确分割。图 1 展示了 NANet 网络的具体结构,该网络包含 2 个组成阶段:全局去噪和局部分割。网络的具体组成结构如下所述。

### 1.1 总体架构

NANet 网络的 2 个阶段分别使用无监督的 U 型自编码器架构和 DeepLabv3+ 作为基本构建块,并在去噪阶段加入 LFFT 模块。具体结构如图 1 所示。

#### 1) 全局去噪

在全局去噪阶段(图 1 上部分),本文使用无监督的 U 型自编码器作为基准。网络输入通过  $3 \times 3$  卷积层提取特征,然后经过 4 个编码器层处理,每个层包括最大池化和卷积操作。本文引入了一个 LFFT 模块抑制不同频域的噪声,从而减小其对特征图的影响,提高噪声图像的质量。解码器层也包含 4 个层,每个层都由双线性上采样和卷积组成,用于逐层恢复特征图。最后,将得到的特征图输入到分割阶段进行精确分割。通过这一去噪阶段的设计, NANet 网络能有效处理输入图像的噪声,减少干扰,从而获得更干净、准确的特征图,提升整体分割性能。

#### 2) 局部分割

在局部分割阶段(图 1 下部分),本文采用 DeepLabv3+ 网络作为分割网络,主要是利用其稳定性以及空洞空间金字塔模块(atrous spatial pyramid pooling, ASPP)增加感受野以提取特征。在编码器中,使用 MobileNetV2<sup>[21]</sup> 作为基准,利用多个转置残差块( Inverted Residual)进行特征提取,以获得更加稠密的特征。具体结构如图 2 左侧所示。先使用  $1 \times 1$  卷积操作来提升通道数,并加入深度可分离卷积进行特征提取。再使用  $1 \times 1$  卷积减少通道数,使其与输入通道数相同。最后,将残差块的输入与输出相加,以实现特征的残差学习。

此外,还增加了带有图像级别特征的 ASPP 模块,通过不同的空洞率获取多尺度的卷积特征。在解码器端,将编码器的输出特征进行 4 倍的上采样,与基准中拥有相同尺寸的浅层特征进行通道维度的拼接。在特征拼接后,对拼接结果进行  $3 \times 3$  卷积操作,进一步细化特征,并进行上采样得到最终的输出。采用 DeepLabv3+ 结构可以保持分割结果的稳定性,并且能够很好地恢复物体分割的细节,例如

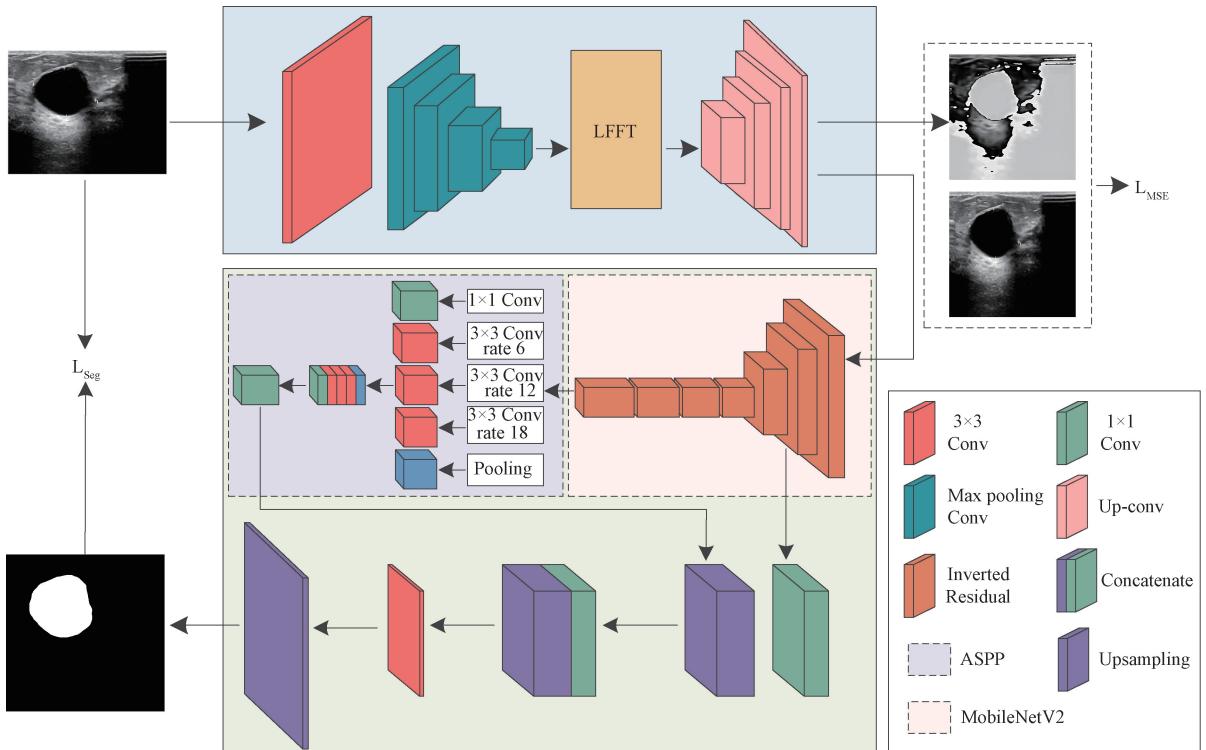


图 1 双阶段噪声自适应网络框架

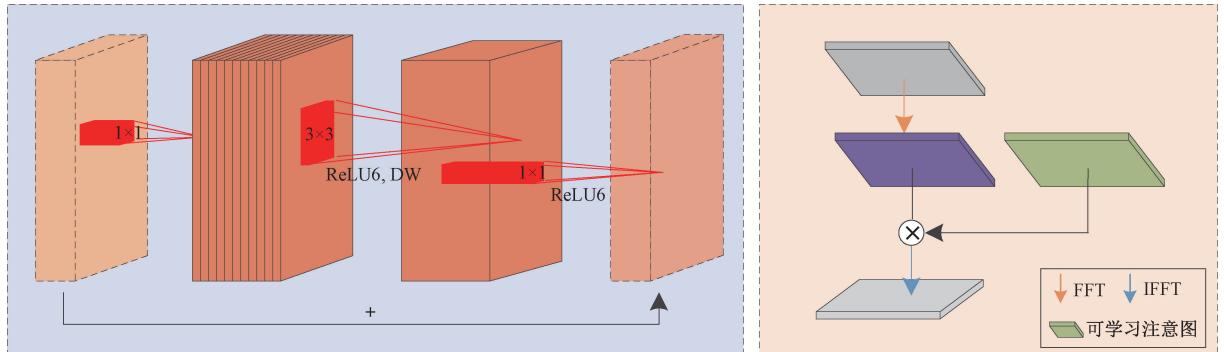


图 2 转置残差块和 LFFT 模块结构

形状、边缘等。这些设计使得 NANet 网络在处理超声医学图像分割任务时具有较好的性能和鲁棒性。

## 1.2 LFFT 模块

在典型的自编码器结构中,往往不能很好地消除不同类型的噪声影响。因此,本文引入 LFFT 模块通过快速傅里叶变换<sup>[22]</sup>对整体进行设计。具体的结构如图 2 右侧所示。其主要思想是通过可学习参数学习一个权重,将其映射到傅里叶空间特征图中,对不同频域的噪声进行抑制。首先,沿着空间维度对特征图执行二维快速傅里叶变换,可表示为:

$$P = \mathcal{F}[X] \in \mathbb{C}^{H \times W \times C} \quad (1)$$

$$\mathcal{F} = \mathbf{F}(\mathbf{u}, \mathbf{v}) = \sum_{M_x=0}^{M_x-1} \sum_{M_y=0}^{M_y-1} f(x, y) e^{-j2\pi(\frac{ux}{M_x} + \frac{vy}{M_y})} \quad (2)$$

其中,  $\mathcal{F}[\cdot]$  代表二维傅里叶变换。 $\mathbf{F}(\mathbf{u}, \mathbf{v})$  是二维傅里叶变换的结果,所在坐标系被称为频域,由  $u = 0, 1, 2, \dots, M_x - 1$  和  $v = 0, 1, 2, \dots, M_y - 1$  定义的矩阵  $\mathbf{M}_x \times \mathbf{M}_y$  称为频域矩阵,  $u, v$  是频域变量。 $f(x, y)$  所在坐标系被称为空间域,由  $x = 0, 1, 2, \dots, M_x - 1$  和  $y = 0, 1, 2, \dots, M_y - 1$  定义的  $\mathbf{M}_x \times \mathbf{M}_y$  矩阵称为空间域矩阵。频域矩阵中每个点的都代表了一个频率为  $u, v$  的函数,这些函数在空间域的组合即为原函数  $f(x, y)$ 。

通过一个可学习的参数化映射  $W \in \mathbb{C}^{H \times W \times C}$  乘以  $P$  得到频谱:

$$\mathbf{P}' = W \otimes P \quad (3)$$

最后,将得到的频域采用二维逆傅里叶变换反转回空间域:

$$\mathbf{X}' = \mathcal{F}^{-1}[\mathbf{P}'] \quad (4)$$

$$\mathcal{F}^{-1} = f(x, y) = \frac{1}{M_u \times M_v} \sum_{u=0}^{M_u-1} \sum_{v=0}^{M_v-1} \mathbf{F}(u, v) e^{-j2\pi(\frac{ux}{M_u} + \frac{vy}{M_v})} \quad (5)$$

LFIT 模块是一种可学习的频域滤波器, 可以调整全局特征的特定频域部分, 从而减少图像噪声干扰。

### 1.3 损失函数

损失函数可以对分割结果产生关键影响, 本文将模型的整体损失分为 2 个部分: 1) 重构损失; 2) 分割损失。重构损失使用 MSE 均方误差将模型去噪阶段输出的图像与训练集中对应的原始输入图像进行逐个像素点比较, 计算对应损失。计算方法如下:

$$L_{MSE} = \frac{1}{n} \sum_{i=1}^n (f(x_i) - y_i)^2 \quad (6)$$

式中:  $y_i$  为图像的真实值,  $f(x_i)$  为图像在处的预测值,  $x_i$  为输入图像的像素值,  $n$  为图像的像素总数。

分割损失采用 Cross-Entropy 损失和 Dice 损失的结合, 其占比是一致的。具体表示如下:

$$L_{Seg}^{CE} = -\frac{1}{N} \sum_i^M y_{ic} \log(p_{ic}) \quad (7)$$

$$L_{Seg}^{Dice} = 1 - \frac{2 | X \cap Y |}{| X | + | Y |} \quad (8)$$

$$L_{Seg} = 0.5 \times L_{Seg}^{CE} + 0.5 \times L_{Seg}^{Dice} \quad (9)$$

式中:  $N$  表示样本数,  $M$  表示分割类别数,  $y_{ic}$  表示第  $c$  类第  $i$  个样本的真实标签,  $p_{ic}$  表示第  $c$  类第  $i$  个样本的预测概率;  $X$  表示真实分割图像的像素标签,  $Y$  表示模型预测分割图像的像素类别。

本文综合考虑了重构损失和分割损失, 并将重构损失的权重调整为 0.01, 以平衡其较大的数值。这样的调整是为了避免模型过于偏向去噪, 从而确保分割效果更加理想。模型整体损失如下:

$$L = 0.01 \times L_{MSE} + L_{Seg} \quad (10)$$

## 2 实验结果与分析

### 2.1 数据集

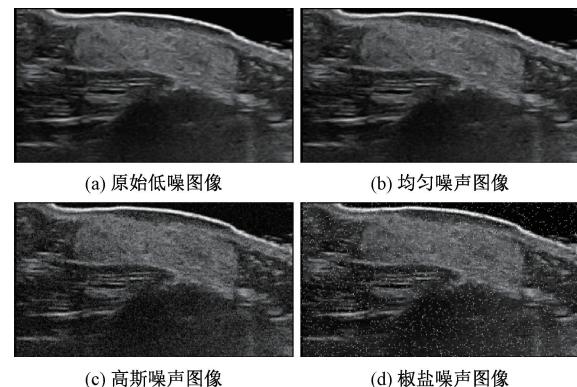
本文对提出的方法在 4 个超声数据集上进行了评估。在保持所有数据集的训练集不变的情况下, 在测试集上添加了不同类型和数量的噪声, 得到了 3 个不同目标域的测试集, 具体如图 3 所示。

#### 1) 乳腺癌病灶超声数据集(BUSI)

该数据集是 2018 年采集的, 包括 600 例年龄在 25~75 岁的女性乳房超声波图像<sup>[23]</sup>。由 780 张图像组成, 平均图像大小为  $500 \times 500$  像素, 被分为正常、良性、恶性三类。本文训练采用良性与恶性两类中随机选取的 390 张图像作为训练集, 62 张图像作为验证集, 195 张图像作为测试集。

#### 2) 臂丛神经超声数据集(Nervese)

该数据集主要来自于 101 名患者, 由专业的麻醉师完



(a) 原始低噪图像 (b) 均匀噪声图像

(c) 高斯噪声图像

(d) 椒盐噪声图像

图 3 不同噪声污染的超声图像

成标注, 用于检测臂丛神经和周围解剖组织。本文训练随机选取 1300 张作为训练集, 324 张作为验证集, 697 张作为测试集。

#### 3) 甲状腺癌病症超声数据集

该数据集收集自武汉同济医院, 所有数据均由医学专家团队标注。本文选取 240 张作为训练图像, 21 张作为验证集, 110 张作为测试集。

#### 4) 血管癌病症超声数据集

血管瘤图像及其注释标签源自重庆医科大学附属儿童医院, 并已作为公共数据集发布<sup>[24]</sup>。该数据集包含 215 张原始图像和标签, 通过镜像翻转和仿射映射进行数据增强。本研究选择 100 张图像作为训练集, 15 张作为验证集, 100 张作为测试集。

### 2.2 评估指标与实验环境

为了评估本文提出方法的性能, 使用 Dice 系数、精确率和召回率来衡量分割性能。实验基于 Python 3.8 和 PyTorch 1.9.1 实现, 使用 Adam 作为学习率调整策略, 初始学习率均设置为 0.001, 批处理大小为 6。所有实验均采用相同数据增强操作进行预处理。在乳腺癌、甲状腺和血管瘤超声数据集中, 图像均被裁剪为  $512 \times 512$ , 训练周期为 300。在臂丛神经超声数据集中, 训练图像输入大小为  $420 \times 580$ , 训练周期为 200。所有模型在 Nvidia4090 GPU 上训练, 该 GPU 内存容量为 24 GB。优化器采用随机梯度下降 SGD, 动量设置为 0.9, 权值衰减值为  $1 \times 10^{-4}$ 。

### 2.3 对比实验

#### 1) 乳腺癌病灶超声数据集(BUSI)

在乳腺癌病灶超声数据集上, 本文对所提出的方法进行了测试。表 1 展示了在原始低噪测试集和加入均匀、高斯和椒盐噪声的测试集上, 不同分割网络的结果。经典的分割模型(如 U-Net、DeepLabv3+、LinkNet)在原始测试集上表现良好, 但在测试不同噪声域的数据集时性能显著下降。相比之下, 本文提出的基准网络 BaseNet 和 NANet 网络在 3 个不同的噪声域中表现更为稳定, Dice 值分别稳定在 68% 和 69% 左右。这表明本文提出的双阶段网络在处理不同噪声图像时具有较强的鲁棒性和泛化性能。

表 1 乳腺癌病灶超声数据集中比较不同模型在不同噪声条件下的分割结果

%

噪声类型	U-Net	Deep	Link	Swin	Trans	Base	NA-Net
原始图像	Dice	61.69	<b>71.3</b>	68.83	40.59	50	68.9
	Pre	57.82	67.16	64.7	33.54	42.38	65.63
	Recall	79.57	85.21	83.4	81.88	81.3	83.27
均匀图像	Dice	36.1	55.08	68.31	26.94	44.62	68.66
	Pre	29.26	51.48	64.08	18.3	35.13	65.05
	Recall	73.8	76.07	83.06	92.91	84.87	83.02
高斯图像	Dice	29.59	27.18	67.91	19.36	31.01	68.08
	Pre	28.21	19.15	63.63	11.98	21.19	63.37
	Recall	45.47	78.43	82.21	98.25	90.34	83.16
椒盐图像	Dice	17.24	21.67	67.1	16.99	24.06	67.54
	Pre	10.92	14.29	63.22	10.19	15.31	64.05
	Recall	78.01	81.51	81.54	99.19	93.85	82.18

## 2) 脊髓神经超声数据集

表 2 验证了本文提出网络在脊髓神经超声数据集上的有效性。结果显示, U-Net 和 AttUNet 网络在不同噪声域的测试集上性能持续下降; 而 DeepLabv3+ 和 LinkNet 网络相对较为稳定。基准网络在原始测试集上并未占据优势, 但在高斯

和椒盐噪声上性能有所提升。NA-Net 网络在 4 种不同测试集上的结果均高于其他网络, 在高斯噪声测试集上的分割结果甚至超过原始测试集, 并且在椒盐噪声图像上, 相比其他分割性能最高的 LinkNet 网络, Dice 值提升了 5%。这突显了 NA-Net 模型在噪声超声图像分割方面的有效性和优越性。

表 2 脊髓神经超声数据集中比较不同模型在不同噪声条件下的分割结果

%

噪声类型	U-Net	Deep	AttU	Link	Base	NA-Net
原始图像	Dice	75.14	77.79	77.9	77.06	77.65
	Pre	70.27	73.4	74.12	74.77	73.3
	Recall	84.31	85.25	85.11	82.26	83.44
均匀图像	Dice	71.03	74.74	71.3	78.31	<b>78.71</b>
	Pre	70.54	82.52	74.68	76.12	75.11
	Recall	76.55	71.02	72.82	72.82	85.77
高斯图像	Dice	58.6	68.39	50.33	77.07	76.23
	Pre	67.69	80.34	65.02	79.61	72.17
	Recall	57.7	62.19	46.2	77.58	83
椒盐图像	Dice	56.14	67.27	38.53	74.12	77.5
	Pre	62.43	81.1	57.6	78.68	74.11
	Recall	58.15	60.54	33.21	72.23	83.47

## 3) 甲状腺癌病症超声数据集

为了验证提出模型在超声图像上的稳定性, 对甲状腺癌病症超声数据集进行了实验, 结果如表 3 所示。与经典模型(CNN 类和 Transformer 类)相比, 提出的网络在不同噪声超声图像上表现出了稳定且优越的分割性能。与最佳模型相比, 在均匀、高斯和椒盐测试集上分别提升了 8.83%、56% 和 55%。尽管在原始测试集上与 DeepLabv3+ 和 LinkNet 网络相比, 提出的网络分割结果略有下降, 可能是在去噪阶段丢失了一些细节信息; 但在其他噪声测试集上的分割结果明显优于其他模型。原始超声图像上的轻微下降在可接受范围内。

## 4) 血管癌病症超声数据集

血管癌病症超声数据集上的实验结果如表 4 所示。在该数据集上, 基准网络在 3 种不同噪声域的测试集上表现十分稳定, 其分割结果均优于其他模型。随着 LFFT 模块的加入, NA-Net 网络取得了更优的分割结果。与分割结果最好的公共模型相比, 在原始和均噪声域的测试集上表现十分稳定, 其分割结果均高于匀测试集上有 1% 左右的提升, 在高斯和椒盐数据集上 Dice 值提高了 8%。NA-Net 网络在不同噪声图像上均取得了不错的成果, 证明了提出网络的通用性和稳健性。

## 2.4 消融实验

为了验证 LFFT 模块在模型中不同位置的影响, 在去

表 3 甲状腺癌病症超声数据集中比较不同模型在不同噪声条件下的分割结果

%

噪声类型		U-Net	Deep	Link	Swin	Trans	Base	NAN
原始图像	Dice	76.11	82.36	<b>82.73</b>	40.16	46.79	78.33	79.79
	Pre	71.16	76.64	81.09	30.27	36.69	72.25	73.51
	Recall	90.51	90.67	87.83	93.18	95.19	91.61	91.98
均匀图像	Dice	24.01	68.83	64.76	19.27	27.91	75.23	<b>77.66</b>
	Pre	15.54	61.64	59.51	11.82	18.85	66.61	70.83
	Recall	98.74	90.59	85.66	98.89	98.4	94.18	92.75
高斯图像	Dice	15.62	4.52	19.89	15.43	17.42	62.76	<b>75.97</b>
	Pre	9.21	13.41	12.4	9.08	10.45	51.66	68.39
	Recall	99.79	5.09	95.54	1	99.98	96.27	93.24
椒盐图像	Dice	15.5	0.22	15.53	15.42	16.63	46.01	<b>72.55</b>
	Pre	9.13	0.16	9.15	9.08	9.9	35.7	63.29
	Recall	99.99	0.88	99.99	1	99.99	96.6	94.61

表 4 血管癌病症超声数据集中中比较不同模型在不同噪声条件下的分割结果

%

噪声类型		U-Net	Deep	Link	Swin	Trans	Base	NAN
原始图像	Dice	54.16	70.44	66.2	47.63	15.48	69.17	<b>71.5</b>
	Pre	41.7	64.71	56.91	35.47	16.34	62.39	65.8
	Recall	92.8	84.44	89.81	91.24	20.33	83.53	83.56
均匀图像	Dice	46.57	67.67	58.04	46.51	14.87	69.52	<b>71.2</b>
	Pre	33.33	63.33	46.43	34.36	15.69	63.83	66.02
	Recall	93.67	81.25	89.86	92.1	19.88	82.22	82.29
高斯图像	Dice	30.71	61.21	36.48	37.03	10.74	69.03	<b>70.14</b>
	Pre	19.45	57.49	24.41	24.95	12.83	65.89	65.32
	Recall	96.89	74.54	93.5	95.24	16.41	77.99	81.11
椒盐图像	Dice	44.83	61.44	51.13	45.94	10.46	66.96	<b>69.65</b>
	Pre	34.05	58.15	40.85	34.12	17.37	63.1	63.8
	Recall	83.34	73.81	81.64	90.52	10.22	77.19	82.8

噪阶段的编码器层中分别添加了 LFFT 模块,结果如表 5 所示。对比基准网络发现,在前三层的编码器后添加 LFFT 模块,分割结果并没有明显提升。可能是因为浅层

获取的特征图中包含大量的形状、纹理、颜色等细节信息,在进行 LFFT 处理后更多的特征信息丢失,导致分割结果下降。

表 5 LFFT 模块在模型中不同位置的影响结果

%

模型	LFFT 位置	原始图像			均匀图像			高斯图像			椒盐图像					
		Dice	Pre	Recall	Dice	Pre	Recall	Dice	Pre	Recall	Dice	Pre	Recall			
BaseNet	×	×	×	×	69.05	65.41	83.05	<b>69.19</b>	65.33	83.68	67.88	62.7	85.28	67.58	62.82	84.34
BaseNet_v0	√	√	√	√	68.9	65.63	83.27	68.66	65.05	83.02	68.08	63.37	83.16	67.54	64.05	82.18
BaseNet_v1	√	×	×	×	68.85	64.7	83.93	68.84	64.83	83.8	67.45	63.45	83.32	67.67	64.41	80.67
BaseNet_v2	×	√	×	×	68.94	64.2	84.55	68.99	64.41	84.51	67.45	62.78	83.45	67.95	63.24	83.27
BaseNet_v3	×	×	√	×	68.91	63.72	83.51	68.79	63.42	83.73	67.49	61.95	83	67.5	62.11	82.85
NANet	×	×	×	√	<b>69.54</b>	64.7	83.84	69.03	64.37	83.65	<b>68.1</b>	63.16	83.58	<b>68.57</b>	64.22	82.46

此外,还进行了不同噪声类型和数量的消融研究,结果如图 4 所示。在不同类型噪声域的测试集上,随着噪声数量的增加,大多数分割模型的结果会随之下降,且下降

幅度较大。特别是在椒盐噪声图像中,与其他分割模型相比,本文提出的网络受噪声数量变化影响较小,变化幅度较弱。

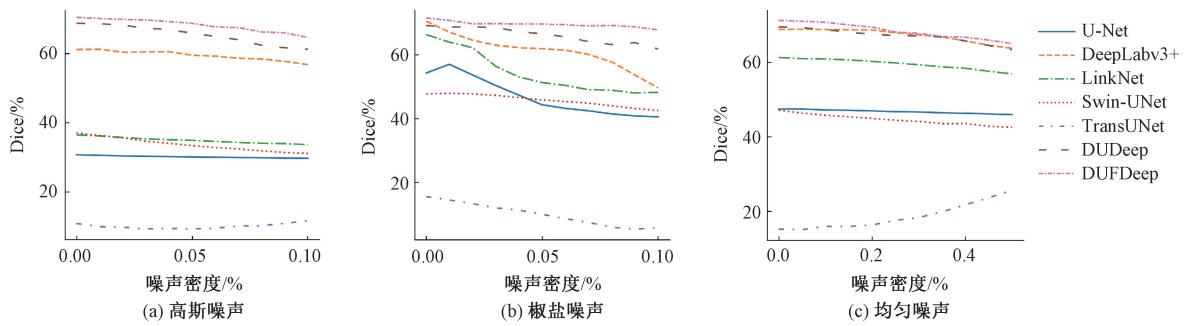


图 4 不同噪声以及数量在不同模型的分割性能对比不同噪声以及数量在不同模型的分割性能比较结果

## 2.5 可视化结果

为了进一步评估本文提出的模型在有噪声环境下的分割准确率,进行了可视化分析。对所提出的模型与现有

的分割模型在不同噪声测试集上进行了对比,具体结果如图 5、6 所示。

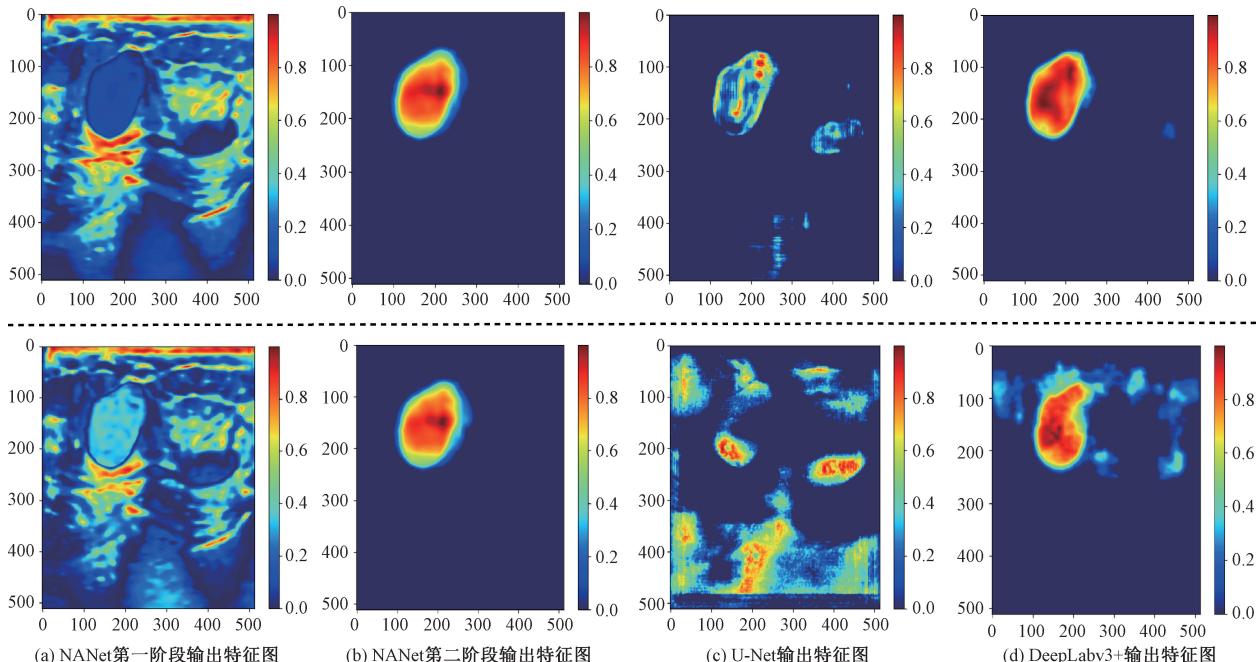


图 5 在乳腺癌病灶超声数据集上网络的特征热力图

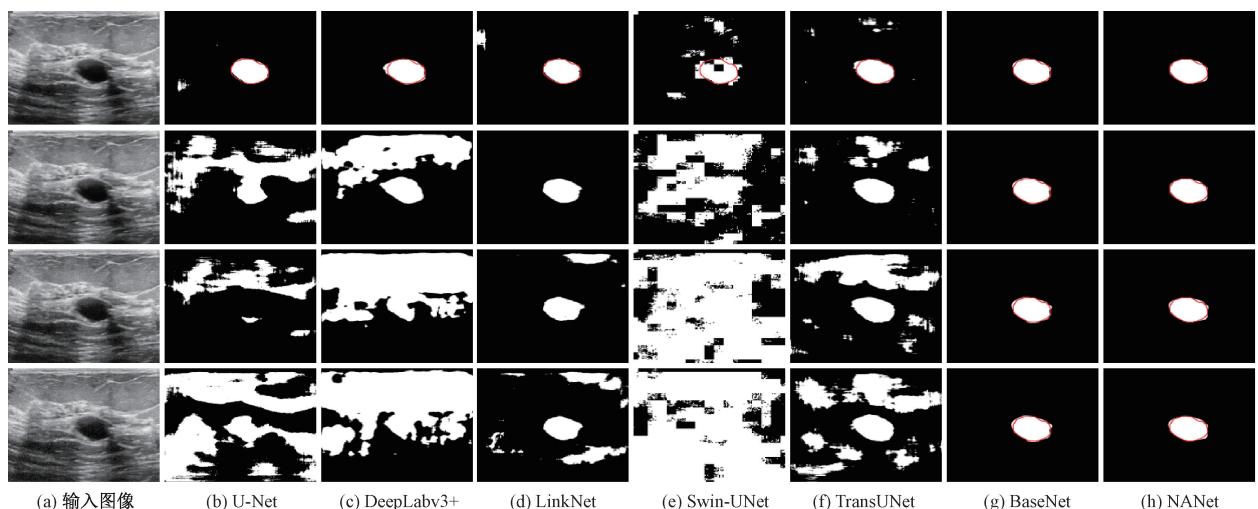


图 6 不同网络在不同噪声测试集中的结果可视化图

图 5 展示了网络在不同层的特征热力图。第 1 行为原始测试集,第 2 行为椒盐噪声测试集。与 U-Net 网络在解码器端上采样后得到的特征图(图 5(c))相比,NA-Net 网络在去噪阶段对应层得到的输出(图 5(a))包含更多细节信息,且定位更加精确。在分割阶段,与 DeepLabv3+ 网络在解码器端上采样后得到的特征图(图 5(d))相比,NA-Net 网络在去噪阶段对应层得到的输出(图 5(b))分割边缘更加精确,且受噪声的影响较弱。

在图 6 中,不同行从上至下分别代表原始低噪测试集、均匀噪声测试集、高斯噪声测试集和椒盐噪声测试集。从该图可以直观地看出 U-Net、DeepLabv3+ 等网络在噪声测试集上分割性能明显下降,噪声较大程度地影响了模型在定位、分割等方面的判断。而本提出的模型(g、h 列)在分割结果明显高于之前的网络,其中 NA-Net 网络(h 列)的边缘细节分割更精确,表明 LFFT 模块的有效性(红色框表示标注图)。

### 3 结 论

本文提出了一种用于超声医学图像分割的双阶段噪声自适应网络,旨在提升有噪声环境下的分割准确率和稳定性。所提出的模型架构主要包括去噪和分割 2 个阶段。其中去噪阶段由自编码器和 LFFT 模块组成,旨在抑制不同频域的噪声影响。而分割阶段基于 DeepLabv3+ 架构,实现对去噪阶段特征图的精确分割。通过在 4 个不同病灶的超声数据集上进行实验,结果表明本文提出的模型在不同噪声污染的测试集上均表现出了更为稳定的分割性能。此外,消融实验进一步验证了 LFFT 模块在噪声自适应分割任务中的有效性。本文的研究成果在实际医学图像处理应用中具有重要的参考价值。在未来的研究中,本文将致力于将该网络运用到其他噪声污染下的图像分类和目标检测任务中。

### 参考文献

- [1] 张驰名,王庆凤,刘志勤,等.基于深度迁移学习的肺结节辅助诊断方法[J].计算机工程,2020,46(1):271-278.
- [2] MESSAOUDI H, BELAID A, SALEM D B, et al. Cross-dimensional transfer learning in medical image segmentation with deep learning[J]. Medical Image Analysis, 2023, 88: 102868.
- [3] ZHANG T, CHENG J, FU H, et al. Noise adaptation generative adversarial network for medical image analysis[J]. IEEE Transactions on Medical Imaging, 2019, 39(4): 1149-1159.
- [4] 宋永献,王祥祥,夏文豪.基于 VMD-HDNLN 的下肢肌电噪声信号处理方法研究[J].电子测量技术,2023, 46(14): 53-58.
- [5] 杨智翔,孙玉宝,白志远,等.高光谱图像去噪的稀疏光谱 Transformer 模型[J].电子测量技术,2024,47(1): 150-158.
- [6] 袁乐民.基于中值滤波算法的医学影像图像除噪设计实现研究[J].贵州大学学报(自然科学版),2020, 37(2): 42-45, 50.
- [7] 张雪峰,闫慧.基于中值滤波和分数阶滤波的图像去噪与增强算法[J].东北大学学报(自然科学版),2020, 41(4): 482-487.
- [8] WU W, WANG Y, LIU Q, et al. Wavelet-improved score-based generative model for medical imaging[J]. IEEE Transactions on Medical Imaging, 2023, DOI: 10.1109/TMI.2023.3325824.
- [9] 陈军.基于小波分析的医学影像图像除噪设计及仿真研究[J].贵州大学学报(自然科学版),2020,37(5): 78-81, 88.
- [10] 刘源,吴思进,李伟仙,等.融合小波与非局部均值滤波的 DSPI 相位图去噪[J].电子测量技术,2023,46(20): 110-119.
- [11] ZHANG Y, WANG M, XIA X, et al. Medical image fusion based on quasi-cross bilateral filtering [J]. Biomedical Signal Processing and Control, 2023, 80: 104259.
- [12] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation [C]. Medical Image Computing and Computer-Assisted Intervention-MICCAI 2015, 2015.
- [13] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. Deeplab: Semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected crfs [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, 40(4): 834-848.
- [14] CHAURASIA A, CULURCIELLO E. Linknet: Exploiting encoder representations for efficient semantic segmentation[C]. St. Petersburg, FL: 2017 IEEE visual communications and image processing (VCIP), IEEE, 2017.
- [15] CAO H, WANG Y, CHEN J, et al. Swin-unet: Unet-like pure transformer for medical image segmentation [C]. European Conference on Computer Vision, Cham: Springer Nature Switzerland, 2022.

- [16] WANG G, LIU X, LI C, et al. A noise-robust framework for automatic segmentation of COVID-19 pneumonia lesions from CT images [J]. IEEE Transactions on Medical Imaging, 2020, 39 (8): 2653-2663.
- [17] AKSAM IFTIKHAR M, JALIL A, RATHORE S, et al. Robust brain MRI denoising and segmentation using enhanced non-local means algorithm [J]. International Journal of Imaging Systems and Technology, 2014, 24(1): 52-66.
- [18] BOYKOV Y, FUNKA-LEA G. Graph cuts and efficient ND image segmentation [J]. International Journal of Computer Vision, 2006, 70(2): 109-131.
- [19] FU J, LIU J, TIAN H, et al. Dual attention network for scene segmentation[C]. State of California: IEEE/ CVF Conference on Computer Vision and Pattern Recognition, 2019.
- [20] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional block attention module[C]. Munich: European Conference on Computer Vision(ECCV), 2018.
- [21] HOWARD A, ZHMOGINOV A, CHEN L C, et al. Inverted residuals and linear bottlenecks: Mobile networks for classification, detection and segmentation [C]. Salt Lake City: CVPR, 2018.
- [22] DUHAMEL P, VETTERLI M. Fast Fourier transforms: a tutorial review and a state of the art[J]. Signal Processing, 1990, 19(4): 259-299.
- [23] AL-DHABYANI W, GOMAA M, KHALED H, et al. Dataset of breast ultrasound images[J]. Data in Brief, 2020, 28: 104863.
- [24] 张晋炜,王冬,尹书月,等.超声诊断婴幼儿血管瘤进展[J].中国介入影像与治疗学,2018,15(2): 116-118.

### 作者简介

王霞霞,硕士研究生,主要研究方向为计算机视觉、医学图像分析等。

E-mail: wxxvva@163. com

张炫,硕士研究生,主要研究方向为计算机视觉、医学图像分割等。

E-mail: zhang81588@163. com

吴兴隆,博士,副教授,主要研究方向为计算机视觉、医学图像分析等。

E-mail: xwu@wit. edu. cn

黄青,博士,副教授,主要研究方向为计算机视觉、医学图像分析等。

E-mail: hustsy2008@163. com

徐国平(通信作者),博士,讲师,主要研究方向为计算机视觉、医学图像分析等。

E-mail: xugp@wit. edu. cn