



联合超声甲状腺结节分割与分类的多任务方法研究

刘侠, 吕志伟, 王波, 王狄, 谢林浩

引用本文:

刘侠, 吕志伟, 王波, 王狄, 谢林浩. 联合超声甲状腺结节分割与分类的多任务方法研究[J]. 智能系统学报, 2023, 18(4): 764–774.
LIU Xia, LYU Zhiwei, WANG Bo, WANG Di, XIE Linhao. Multi-task method for segmentation and classification of thyroid nodules combined with ultrasound images[J]. *CAAI Transactions on Intelligent Systems*, 2023, 18(4): 764–774.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202203063>

您可能感兴趣的其他文章

基于注意力机制的显著性目标检测方法

Salient object detection method based on the attention mechanism

智能系统学报. 2020, 15(5): 956–963 <https://dx.doi.org/10.11992/tis.201903001>

基于小样本学习的LCD产品缺陷自动检测方法

An automatic small sample learning-based detection method for LCD product defects

智能系统学报. 2020, 15(3): 560–567 <https://dx.doi.org/10.11992/tis.201904020>

层次化双注意力神经网络模型的情感分析研究

Hierarchical double-attention neural networks for sentiment classification

智能系统学报. 2020, 15(3): 460–467 <https://dx.doi.org/10.11992/tis.201812017>

基于生成式对抗网络的道路交通模糊图像增强

Enhancement of blurred road-traffic images based on generative adversarial network

智能系统学报. 2020, 15(3): 491–498 <https://dx.doi.org/10.11992/tis.201903041>

基于双向消息链路卷积网络的显著性物体检测

Salient object detection based on bidirectional message link convolution neural network

智能系统学报. 2019, 14(6): 1152–1162 <https://dx.doi.org/10.11992/tis.201812003>

基于改进卷积神经网络的多标记分类算法

A multi-label classification algorithm based on an improved convolutional neural network

智能系统学报. 2019, 14(3): 566–574 <https://dx.doi.org/10.11992/tis.201804056>

DOI: 10.11992/tis.202203063

网络出版地址: <https://kns.cnki.net/kcms/detail/23.1538.TP.20230322.1840.010.html>

联合超声甲状腺结节分割与分类的多任务方法研究

刘侠¹, 吕志伟², 王波³, 王狄², 谢林浩²

(1. 哈尔滨理工大学自动化学院, 黑龙江 哈尔滨 150080; 2. 哈尔滨理工大学 黑龙江省复杂智能系统与集成重点实验室, 黑龙江 哈尔滨 150080; 3. 广东科学技术职业学院 计算机工程技术学院, 广东 珠海 519090)

摘要: 针对超声图像中甲状腺结节多尺度、结节边缘模糊、良恶分类不平衡问题, 提出一种联合超声甲状腺结节分割与分类的多任务方法。以全卷积网络作为主干共享网络, 将提取到的浅层特征共享给多任务分支网络, 在分割网络分支中, 先加入深层卷积块, 获取分割分支深层特征, 再对深层特征进行上采样。本文提出一种改进卷积注意力模块的多尺度卷积注意力模块, 将上采样结果与主干共享网络每个特征提取阶段经过带有多尺度卷积注意力模块跳跃连接后的特征张量进行拼接, 减少结节边缘模糊问题, 提高分割性能。同时将多尺度卷积注意力模块融入到分类分支中, 优化分类性能。实验结果表明: 本文所提多任务方法能有效提升分割和分类的精度, 较单任务深度学习网络具有更优的分割与分类性能, 能有效处理甲状腺结节多尺度、结节边缘模糊的问题, 降低良恶分类不平衡带来的影响。

关键词: 深度学习; 多任务学习; 甲状腺结节超声图像; 图像分割; 图像分类; 深层卷积块; 多尺度卷积注意力模块; 残差结构

中图分类号: TP391 文献标志码: A 文章编号: 1673-4785(2023)04-0764-11

中文引用格式: 刘侠, 吕志伟, 王波, 等. 联合超声甲状腺结节分割与分类的多任务方法研究 [J]. 智能系统学报, 2023, 18(4): 764-774.

英文引用格式: LIU Xia, LYU Zhiwei, WANG Bo, et al. Multi-task method for segmentation and classification of thyroid nodules combined with ultrasound images[J]. CAAI transactions on intelligent systems, 2023, 18(4): 764-774.

Multi-task method for segmentation and classification of thyroid nodules combined with ultrasound images

LIU Xia¹, LYU Zhiwei², WANG Bo³, WANG Di², XIE Linhao²

(1. School of Automation, Harbin University of Science and Technology, Harbin 150080, China; 2. Heilongjiang Provincial Key Laboratory of Complex Intelligent System and Integration, Harbin University of Science and Technology, Harbin 150080, China; 3. Computer Engineering Technical College, Guangdong Polytechnic of Science and Technology, Zhuhai 519090, China)

Abstract: Aiming at the problems of multi-scale thyroid nodules, blurred nodule edges, and unbalanced classification of benign and malignant thyroid nodules in ultrasound images, this paper proposes a multi-task method for segmentation and classification of thyroid nodules combined with ultrasound. The fully convolutional network is used as the backbone sharing network, and the extracted shallow features are shared to the multi-task branch network. In the branch segmentation networks, deep convolution blocks are added to obtain the deep features of the segmented branches, and then the deep features are up-sampled. An improved multi-scale convolutional attention module is proposed, which combines the up-sampling results with the feature tensor of each feature extraction stage of trunk sharing network after jumping connection with multi-scale convolution attention module, so as to reduce the fuzzy problem of nodule edge blurs and improve the segmentation performance. At the same time, a multi-scale convolutional attention module is integrated into the classification branch to optimize the classification performance. The experimental results show that the multi-task method proposed in this paper can effectively improve the accuracy of segmentation and classification, having better segmentation and classification performance than single-task deep learning network. It can effectively deal with the problem of multi-scale thyroid nodules and blurred nodule edges, and reduce the impact brought by unbalanced classification of benign and malignant.

Keywords: deep learning; multi-task learning; ultrasound image of thyroid nodule; image segmentation; image classification; deep layer convolutional block; multiscale convolutional block attention module; residual structure

收稿日期: 2022-03-31. 网络出版日期: 2023-03-23.

基金项目: 国家自然科学基金项目 (61172167); 黑龙江省青年科学基金项目 (QC2017076); 广东省教育厅青年创新人才类项目 (2019GKQNCX043); 广东省教育厅普通高校特色创新项目 (2019GKTSCX029); 广东省普通高校创新团队项目 (2021KCXTD079).

通信作者: 王波. E-mail: hust_wb@hrbust.edu.cn.

近年来, 甲状腺癌的发病率呈上升趋势, 及早发现、诊断和治疗是至关重要的。超声检查在甲状腺良恶性的鉴别诊断和评估方面有着安全、无创、诊断性好等优势, 在临床工作中得到了广泛

应用,是检测甲状腺结节的首选方法。甲状腺结节的形状和纵横比是良恶性诊断的重要特征,甲状腺结节的分割与分类对临床治疗有重要的意义。

深度卷积神经网络能极大地提高目标分类和分割的精度,广泛应用于生物医学。它使得特征检测能够被自动处理,从而改善传统模式分析技术中固有的特征检测的困难。以 UNet(U-Net)^[1]和 ResNet(residual network)^[2]为代表的卷积神经网络分别在医学图像分割和分类任务中取得了不错的效果,已被广泛应用于各自领域。注意力机制^[3]能使模型尽可能分配通道和空间中重要特征信息,在一定程度上提升了模型的效率和性能。

多任务方法能够同时学习多个相关联的任务,并在所有任务之间共享特征信息,从而提高每一个任务的性能。而甲状腺结节的分割与分类需要用到大量相同的结节特征,所以联合超声甲状腺结节分割和分类,通过多任务的方法增强特征信息共享能力,提高每个任务性能,对实现甲状腺结节超声图像的分割和分类具有重要的研究意义。目前,针对多任务网络常用的共享方式分为隐层参数的硬共享和软共享2种。硬共享方式通常在所有任务之间共享浅层特征,同时保留几个特定任务的输出层来实现。软共享方式每个任务都有自己的参数和模型,模型之间的所有信息进行共享。鉴于目前研究中出现的问题甲状腺结节多尺度、结节边缘模糊、良恶分类不平衡问题,本文提出一种联合超声甲状腺结节分割与分类的多任务方法。该方法以全卷积网络(fully convolutional networks, FCN)为主干共享网络,基于 UNet 解码层为分割分支网络和 ResNet34 为分类分支网络。主干共享网络对输入超声图像采用参数硬共享的方式进行浅层特征提取,将提取的特征共享给2个分支网络。在分割分支网络中,首先在主干共享网络后引入深层卷积块(deep layer convolutional block, DLCB),获取分割分支深层特征,其次对第1~4次特征提取阶段 Conv1~Conv4 获取到的浅层特征通过带有多尺度卷积注意力模块(multiscale convolutional block attention module, M-CBAM)的跳跃连接操作来保持以上4个阶段分割边缘特征轮廓,减少结节边缘模糊现象,其中多尺度卷积注意力模块用来解决结节多尺度问题,在分类分支残差模块前后结合 M-CBAM,通过 M-CBAM 和残差模块优化分类性能,降低良恶分类不平衡的影响。通过对比实验,证明本文所提的方法较单独分割和分类的深度学习网络具有更优的分割和分类性能。

1 甲状腺结节分割与分类相关工作

1.1 甲状腺结节分割

针对甲状腺结节分割的问题, Ronneberger 等^[1]在编码-解码结构上加入了跳跃连接,提出 UNet 模型,开启了深度学习在医学图像分割领域的篇章。吴俊霞等^[4]提出了一种用于超声图像甲状腺结节分割的条件分割对抗网络模型,在分割网络中引入一种多扩张率卷积块来实现结节区域的精准定位,分割出更为精确的二值掩膜。胡屹杉等^[5]针对甲状腺结节的边缘形态多样性,提出一种基于特征融合和动态多尺度空洞卷积的超声甲状腺分割网络,采用空间注意力机制来优化图像的低维特征,采用高低维特征融合保留重要特征。赵科甫等^[6]为了解决数据集规模小与质量差的问题,提出了一种以 UNet 为主干的网络,融合了不同的特征层,并使用了压缩和激励(squeeze and excitation, SE)注意力机制,进一步提高了分割性能。Chu 等^[7]针对甲状腺结节检测过程费时、特征提取困难等问题,提出了一种基于 UNet 的甲状腺结节检测方法,用于甲状腺结节的计算机辅助诊断,大大提高了小训练数据集对甲状腺结节的分割精度。Oktay 等^[8]提出了一种新的用于医学成像的注意门(attention gate, AG)模型 Attention-UNet, 该模型能够自动学习聚焦不同形状和大小的目标结构,抑制输入图像中的无关区域,同时突出对特定任务有用的显著特征。Zhou 等^[9]提出了一种医学图像分割体系结构 UNet++, 是一个深度监督的编码器-解码器网络,减少编码器和解码器子网络的特征映射之间的语义差距。Chen 等^[10]提出的模型 DeepLabv3+通过添加一个简单但有效的解码器模块来细化分割结果,将深度可分离卷积应用于空间金字塔池和解码器模块,形成更快、更强的编解码网络。Badrinarayanan 等^[11]提出具有编码-解码结构的 SegNet 用于图像分割,对称地进行下采样和上采样。基于深度学习的甲状腺结节分割方法无需人工干预,很大程度上提高了甲状腺结节分割速度,但很多模型采用多阶段的分割方式,进一步增加计算量。另外模型对于多尺度信息提取和注意力机制的使用不够充分,模型依靠数据数量较多。

1.2 甲状腺结节分类

针对甲状腺结节分类的问题, He 等^[2]针对更深层次的神经网络更难训练的问题,提出一个 ResNet 残差网络来简化训练,残差网络更容易得到优化,可以提高分类的准确性。Huang 等^[12]为

了更深入、更准确、更高效地进行训练,提出 DenseNet 卷积神经网络,缓解了消失梯度问题,加强了特征传递性和重要性,大大减少了参数数量。迟剑宁等^[13]提出一种基于深度网络和浅层纹理特征融合的甲状腺结节良恶性分类算法,将深度网络得到的深度特征与图像纹理特征融合得到高层次与低层次兼顾的融合特征,并针对数据集分布不平衡的问题采用了代价敏感随机森林分类器进行分类。邹奕轩等^[14]通过对3种不同网络深度、宽度及功能的卷积神经网络(convolutional neural networks, CNN)模型进行分类训练测试,表明 CNN 模型能够较好地应对甲状腺结节超声图像进行特征提取并分类。Moussa 等^[15]提出了一种基于 ResNet50 的卷积神经网络微调模型,能够将良性和恶性结节自动区分,提高了甲状腺结节分类的有效性。Wei 等^[16]提出了一种超声图像上甲状腺结节多中心分类的集成深度学习模型(ensemble deep learning model for multicenter classification of thyroid nodules on ultrasound images, EDLC-TN),用于超声图像中甲状腺结节的良恶性区分,在多医疗中心超声图像诊断甲状腺癌中具有普遍适用性。Xie 等^[17]设计了一种用于甲状腺结节良恶性分类的深度神经网络(deep neural networks, DNN),并提出了一种结合局部二值模式(local binary pattern, LBP)特征和源图像的新方法,提高了模型在相对小而有挑战性的数据集上的性能。验证了传统提取的底层特征可以作为卷积网络的输入,用于进一步的高级特征提取。Shi 等^[18]提出了一种基于知识引导的辅助增强方法,用于医学图像的合成。设计了术语和图像编码器来提取领域知识,然后将领域知识作为新的条件约束辅助分类器生成对抗网络(auxiliary classifier generative adversarial networks, ACGAN)框架,以合成高质量的甲状腺结节图像,具有较好的泛化性和鲁棒性,缓解医学领域数据不足的问题。基于深度学习的甲状腺结节分类方法能极大地提升分类的精度,但网络的训练十分依赖数据量,良恶性数据的不足以及网络层数的加深可能会产生类不平衡和过拟合现象,使得无法有效识别甲状腺结节良恶性,出现错识别和漏识别的情况。

1.3 多任务分割分类

与单任务图像分割和分类的方法相比,甲状腺结节的分割与分类需要用到大量相同的结节特征,所以联合超声甲状腺结节分割和分类,通过多任务的方法能增强特征信息共享能力,提高每个任务性能。Shen 等^[19]提出了一个多任务级联网络,利用检测和分割任务,实现良好的甲状腺

结节分割与分类预测。设计一个2步注意网络,利用分割结果获得更好的分类结果,然而第2步分类的准确率受第1步分割的影响较大。Amyar 等^[20]提出了一种新的多任务深度学习模型,用于从胸部 CT 图像中联合识别新冠肺炎患者和分割新冠肺炎病变,利用包含在多个相关任务中的有用信息来提高分割和分类性能。另一方面能够处理小数据的问题,但处理的数据库较小,需要在更大的数据库上测试方法,以确认其良好的性能。Kim 等^[21]提出了一个称为注意感知多任务卷积神经网络的模型,它通过端到端的训练自动学习适当的共享,注意机制抑制了无用的特征信息,使模型具有很好的鲁棒性。但模型结构复杂,所用时间较大。Cipolla 等^[22]提出了一种多任务深度学习方法,该方法通过考虑每个任务的同态不确定性来衡量多个损失函数,证明了对任务相关的同态不确定性建模可以提高模型的代表性。Wu 等^[23]开发了一个用于新冠肺炎诊断的联合分类和分割系统,实现了一个分割模型来发现新冠肺炎患者的 CT 图像中的细粒度病变区域。但模型较大,并且需要大量的数据集。叶剑锋等^[24]以全卷积神经网络为基础模型,提出一种基于损失边缘检测辅助任务和注意力机制的语义分割算法。该方法以边缘检测作为辅助任务,注意力机制设计使模型更加关注物体形状和边缘信息。通过多任务构造语义分割模型,提升分割的精度。Lyu 等^[25]提出了一种注意感知的多任务卷积神经网络模型,通过端到端训练自动学习合适的共享。模型引入注意力机制与残差快捷连接方法,保存模型有用信息。Mehta 等^[26]提出一种基于 UNet 模型的 Y-Net 网络,用于分割乳腺活检图像中不同类型的组织,同时预测特征图类别,识别图像类别重要区域。在 UNet 的基础上增加判别分类并行分支网络,支持卷积块模块化。多任务方法在医学图像领域有不错的效果,通过共享任务关联性特征,能够提高每一个任务的能力,但多任务结构比较复杂,且较难实现,只适用于关联性较强的任务。

甲状腺结节分割与分类关联性较强,所以设计一种同时进行分割与分类的多任务模型进行特征信息共享,提高分割分类性能,并在多任务模型的基础上对存在问题进行改进是关键。

2 联合分割与分类的多任务方法

针对超声图像的甲状腺结节多尺度、结节边缘模糊、良恶分类不平衡的问题,本文提出了一种联合多任务网络(joint multitasking network, JM-

Net) 模型。该模型以 FCN 为主干网络框架, 分支网络分别为基于 UNet 解码层的分割分支和 ResNet34 的分类分支。其中 UNet 解码层包含解码块、改进卷积注意力模块 (convolutional block attention module, CBAM) 的多尺度卷积注意力模块 M-CBAM 以及深层卷积块 (deep layer convolutional block, DLCB), ResNet34 在残差模块前后带有 M-CBAM。首先利用 FCN 对图像进行卷积池化

操作, 提取浅层特征, 将 FCN 的输出特征张量作为分割与分类分支网络的输入。其次分割分支网络先对输入特征张量通过深层卷积块, 提取深层特征, 再进行带有 M-CBAM 跳跃连接的上采样解码操作。分类分支网络同时对 FCN 输入特征张量进行 M-CBAM 和残差模块的操作。最后得到分割和分类的结果。该模型的整体框架如图 1 所示。

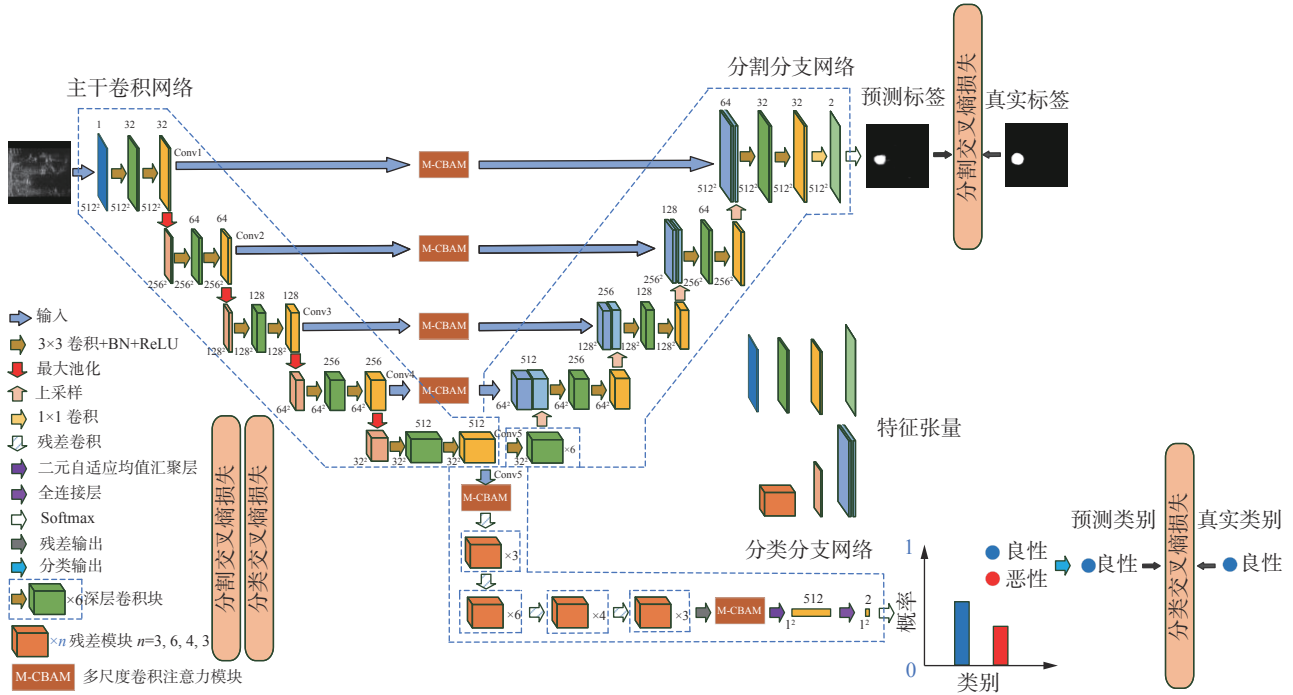


图 1 联合超声甲状腺结节分割与分类的多任务模型

Fig. 1 Multi-task model combining ultrasound thyroid nodule segmentation and classification

2.1 多尺度卷积注意力模块 M-CBAM

本文为减少图像边缘信息模糊、信息丢失以及甲状腺结节形态尺度多变的问题, 在 UNet 解码层跳跃连接中引入了改进 CBAM 的多尺度卷积注意力模块 M-CBAM。在 M-CBAM 通道注意力模块中, 为使多层感知器 (multilayer perceptron, MLP) 的全连接层 (fully connected, FC) 层能同时计算最大池化 (max-pooling, MaxPool) 和平均池化 (average pooling, AvgPool) 2 部分的特征, 将最大池化和平均池化先通过特征拼接进行求和, 通过 MLP 训练, 进而增强这 2 部分特征张量的关联性。在 M-CBAM 空间注意力模块中, 为了减少甲状腺结节形态尺度多变的问题, 将原来 7×7 的卷积核替换为多尺度卷积块 (multi scale convolutional block, MSCB), 通过引入多尺度卷积块 MSCB 来提高模型处理多尺度结节的能力。其中 M-CBAM 注意力机制执行步骤如下: 以第 1 次特征提取阶段 Conv1 为例, Conv1 获取到的输出特征张量浅层特征作为多尺度卷积注意力模块 M-

CBAM 的输入特征张量, 命名为 F , 通过通道注意力模块得到输出特征张量 F_1 , 通道注意力模块表达式为

$$F_1 = \delta(\text{MLP}(\text{MaxPool}(F) + \text{AvgPool}(F))) \otimes F \quad (1)$$

式中: F_1 为通过通道注意力模块得到输出特征张量, δ 为 Sigmoid 激活函数, MLP 为多层感知机, MaxPool 为最大池化, AvgPool 为平均池化, F 为输入特征张量, \otimes 为按位相乘。

$$F_2 = \delta(\text{MSCB}(\text{MaxPool}(F_1) + \text{AvgPool}(F_1))) \otimes F_1 \quad (2)$$

式中: F_2 为通过空间注意力模块得到的输出特征张量, δ 为 Sigmoid 激活函数, MSCB 为多尺度卷积块, MaxPool 为最大池化, AvgPool 为平均池化, F_1 为经过通道注意力模块得到的输出特征张量, \otimes 为按位相乘。将通过空间注意力模块得到输出特征张量 F_2 和输入特征张量 F 按位相乘得到最终经空间注意力和通道注意力改进的输出特征张量 F_3 , 通过通道和空间注意力模块得到的输出特征张量不仅可以保持边界边缘信息不丢失, 还可以提高特征信息利用率, 提高分割多尺度结节的

能力,最后与解码器进行上采样的高一级别特征图进行拼接,再经过 2 次 3×3 卷积、1 次 1×1 卷积

得到最后的分割结果。M-CBAM 多尺度卷积注意力模块网络结构如图 2 所示。

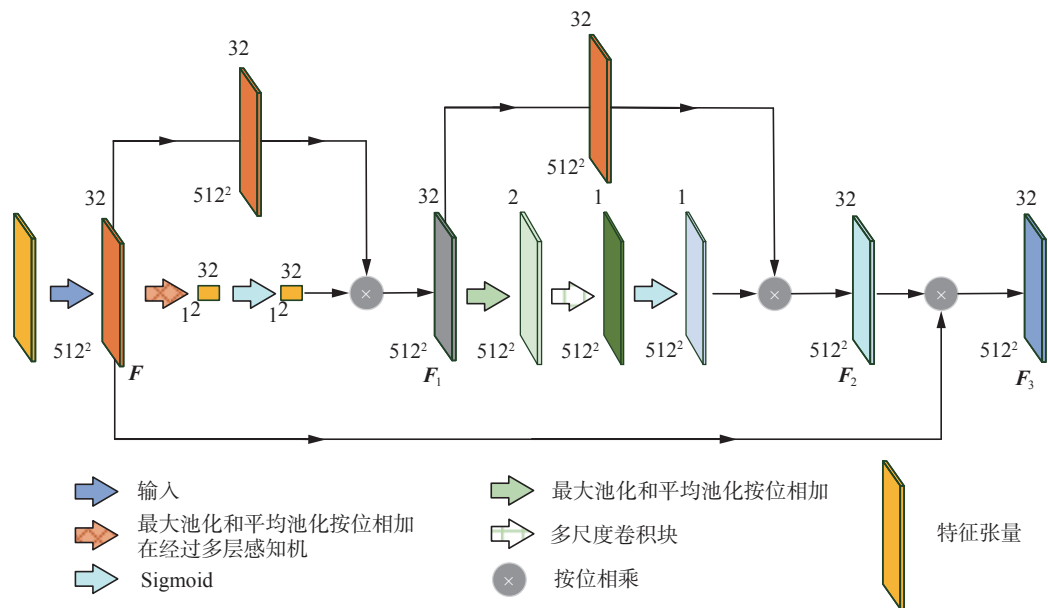


图 2 多尺度卷积注意力模块网络结构

Fig. 2 Multi-scale convolutional attention module network structure

2.2 融入 M-CBAM 的 ResNet34 分类网络

随着网络深度的加深,容易产生梯度爆炸和梯度消失等问题,可能发生网络退化现象。本文选择 ResNet34 网络作为分类分支,在 ResNet34 残差网络结构前后加入 M-CBAM 来提高分类信息的提取能力,降低良恶分类不平衡的影响。首先在残差结构前加入一个多尺度卷积注意力模块 M-CBAM,以第 5 次特征提取阶段 Conv5 获取到的输出特征张量浅层特征作为分类分支中多尺度卷积注意力模块 M-CBAM 的输入特征张量,输入特征张量为 B ,通过多尺度卷积注意力模块 M-CBAM 得到的输出特征张量为 B_1 。

利用 ResNet34 网络中残差网络结构防止梯度爆炸、梯度消失的问题。以一个残差模块为例,对 B_1 进行 2 个卷积层的卷积操作,作用是为了提取 B_1 在分类分支中的深层特征,增加分类特征信息利用率, B_1 通过 2 个卷积层后与 B_1 通过 1×1 卷积和 BN(batch normalization) 相加,相加之后再通过一个 ReLU 激活函数得到一个残差模块的输出特征张量,通过将输入信息绕道传到输出,能够保护信息的完整性。通过所有残差模块之后,在残差结构后加入一个多尺度卷积注意力模块 M-CBAM,通过 2 个 M-CBAM 和残差结构得到输出特征张量 B_3 ,其中第 1 个 M-CBAM 用来提取输入特征张量 B 中关键信息,减少无用浅层特征信息的影响,进而提高残差结构中残差模块的有效特征提取能力,得到输出特征张量 B_1 ,残差结构中

残差模块在获得有效的关键信息的基础上,用来提取 B_1 在分类分支中深层特征,增加分类深层特征信息利用率,得到输出特征张量 B_2 ,第 2 个 M-CBAM 提取 B_2 中关键信息,减少无用深层特征信息的影响。通过以上操作来提高分类性能,提取大量良恶关键特征信息来降低良恶分类不平衡的影响,残差模块的网络结构如图 3 所示。

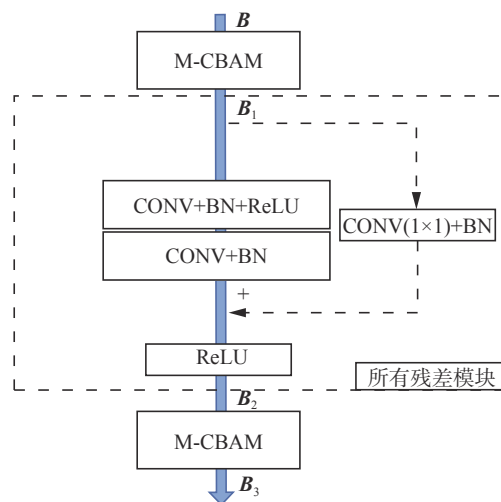


图 3 融入 M-CBAM 残差模块网络结构

Fig. 3 Integrate into the network structure of M-CBAM residual module

3 实验与分析

3.1 实验数据及参数设置

本文的实验数据由上海市胸科医院提供,一

共有 450 例甲状腺结节超声图像的样本以及专家标注轮廓。其中 304 例甲状腺结节超声图像有良恶性类别, 良性为 73 例, 恶性为 231 例, 所以挑选其中带有图像样本标签和类别的 304 例超声图像作为已有的数据集。去除隐私信息后, 将图片统一裁剪为 512×512 大小, 利用旋转、镜像、尺度变换等操作对训练集数据进行 9 倍数据增强, 增强后训练集共有 1647 例数据, 并按照 6:2:2 比例分别将良恶性数据随机划分并组合为训练集、验证集和测试集。

实验中网络的具体参数为: 输入图像大小为 $512 \times 512 \times 1$, Batch_size 大小为 16, 初始学习率设置为 0.0001, 模型优化器为自适应矩估计 (adaptive moment estimation, Adam), 分割和分类分支网络都使用交叉熵损失作为损失函数, 分割交叉熵损失 (segmentation cross entropy loss, SegCELoss) 和分类交叉熵损失 (classification cross entropy loss, ClaCELoss) 求和, 得到联合交叉熵损失函数 (joint cross entropy loss, JCELoss):

$$L_{JCELoss} = \alpha L_{SegCELoss} + \beta L_{ClaCELoss} \quad (3)$$

式中 α 和 β 各取 0.5。

最大迭代数为 100 次。在训练过程中, 如果验证集上连续 10 轮损失值没有下降, 则学习率减少为原来的 0.9 倍。实验的硬件环境为 Intel (R) Xeon(R) 4216 主频 2.10 GHz, 2 张 GeForce RTX 3070 显卡; 操作系统为 Ubuntu 20.04, 编程语言为 Python 3.8.5, 所有程序均在 Pytorch 框架下实现。

3.2 评价指标

从分割结果准确性方面, 将本文方法同 UNet、Attention-UNet、UNet++、Deeplabv3+、SegNet、多任务分割分支网络 (multi-task segmentation network, MSNet) 进行比较。评价指标包括骰子相似系数 (dice similarity coefficient, DSC)、重叠度 (intersection of union, IOU)、精准率 (positive predictive value, PPV)、召回率 (true positive rate, TPR)。

从分类结果准确性方面, 将本文所提方法与 ResNet34、ResNet50、ResNet101、DenseNet121、DenseNet169、多任务分类分支网络 (multi-task classification network, MCNet) 进行比较。评价指标包括准确度 (accuracy, ACC)、精确度 (PPV)、召回率 (TPR) 和受试者工作特征曲线 (receiver operating characteristic curve, ROC) 以及曲线下面积 (area under curve, AUC)。

其中, T_P 、 T_N 、 F_P 和 F_N 分别表示分类结果中真阳性、真阴性、假阳性和假阴性的像素数量, ROC 曲线是以真阳性率 (sensitivity, SENS) 为纵坐标、假阳性率 (1-specificity, 1-SPEC) 为横坐标

绘制的曲线。AUC 表示 ROC 曲线下面积, AUC 值越接近 1, 表示分类算法性能越好。分割分类指标的具体表达式为

$$D_{SC} = \frac{2T_P}{F_P + 2T_P + F_N} \quad (4)$$

$$I_{OU} = \frac{T_P}{F_P + T_P + F_N} \quad (5)$$

$$P = \frac{T_P}{T_P + F_P} \quad (6)$$

$$R = \frac{T_P}{T_P + F_N} \quad (7)$$

$$A_{CC} = \frac{T_P + T_N}{T_P + F_P + T_N + F_N} \quad (8)$$

式中: D_{SC} 为相似系数, I_{OU} 为重叠度, A_{CC} 为准确度。

3.3 结果与分析

3.3.1 分割结果分析

图 4 给出了不同网络的分割结果, 其中图 4(a) 是原始数据, 图 4(b) 是专家标注的金标准, 图 4(c)~4(i) 依次是 UNet^[1]、Attention-UNet^[8]、UNet++^[9]、Deeplabv3+^[10,8]、SegNet^[11]、MSNet、本文方法 JM-Net 的分割结果。

本文将 JMNet 与医学图像分割中具有代表性且具有编码-解码结构的主流模型进行了比较。图中线框标注了专家金标准、对比网络以及本文网络分割结果之间的部分差异。将分割结果从上到下共分为 8 组, 第 1~3 组甲状腺结节为良性结节, 4~8 组为恶性结节, 第 6 组恶性结节是由 2 张恶性结节图片拼接而成。针对图 4 中的分割情况, 第 1、3、8 组的 UNet、Attention-UNet、UNet++、SegNet 的分割结果中出现很多明显的过分割, 会将专家标注轮廓较远处的区域误认为是 ROI 区域, 划分到预测的结果中或者将较远处非 ROI 区域划分为 ROI 区域显示到轮廓外, 而 Deeplabv3+ 存在轻微的过分割, 但相对于其他对比网络分割效果来看, 效果好很多。第 2、4 组的对比网络普遍存在些许的过分割。第 5、6 组的 UNet、Attention-UNet、UNet++、MSNet 存在明显的欠分割。第 7 组的对比网络皆存在欠分割现象, 且 UNet、Attention-UNet 存在欠分割的同时还存在轻微过分割。

对比网络分割小型结节的分割效果较好, 分割中大型结节会出现些许的欠分割现象。从图 4(i) 可以看到本文网络在 M-CBAM 多尺度卷积块作用下, 分割多形态大小的甲状腺结节效果都比较好。分割结果的边缘轮廓保持在专家金标准附近区域, 并且形状和纵横比没有较大的误差, 从最后的结果来看, 本文提出的方法效果最好。

在包含良恶性及标签的甲状腺结节的超声图像上评价了包含本文网络在内的7种网络模型。表1给出了本文模型与对比模型在测试集上的分

割结果。结果表明,在数据集较少的情况下,UNet、Attention-UNet等网络都具有比较相似的分割能力,而本文网络则具有最佳的分割性能。

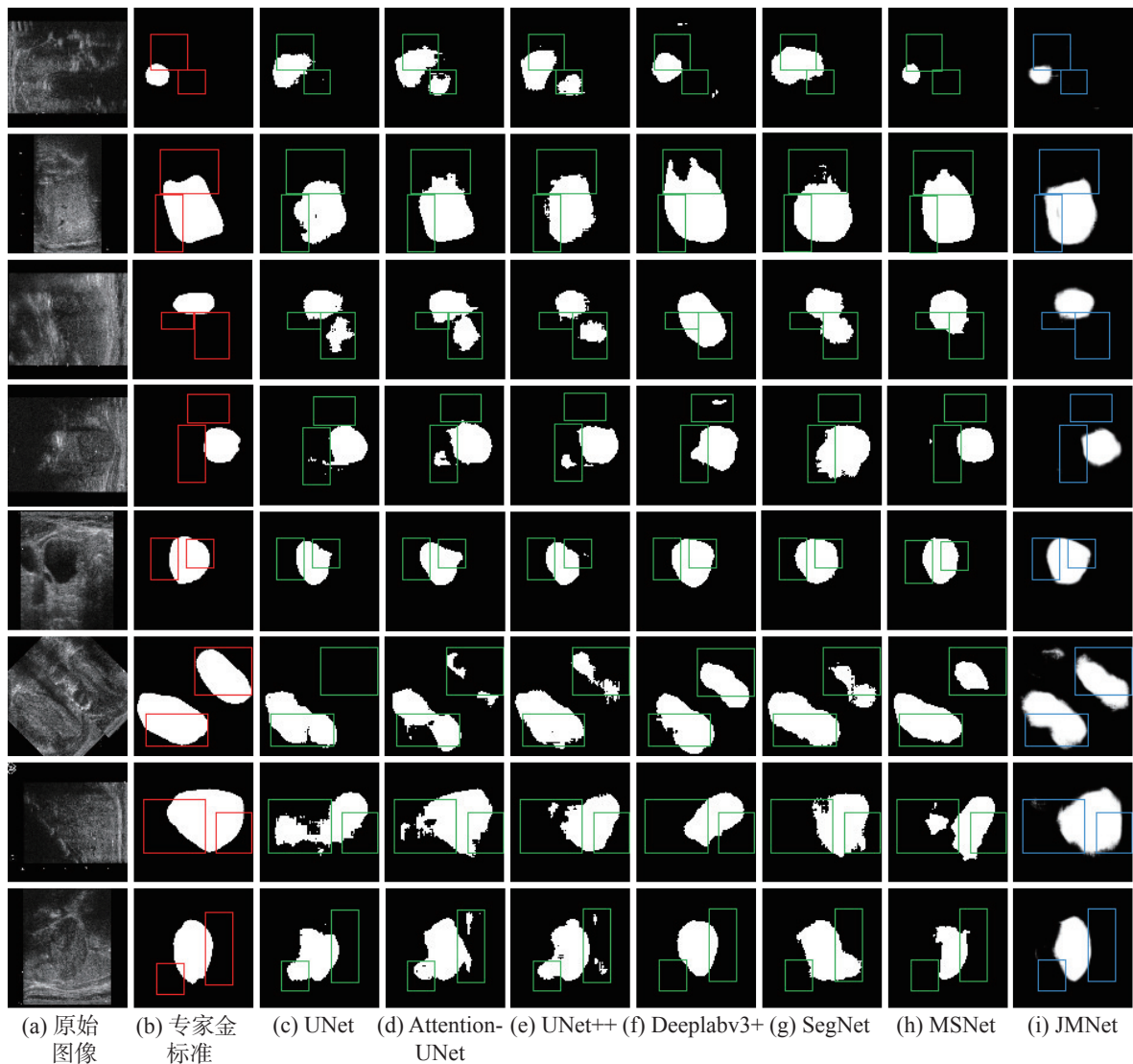


图4 不同网络分割结果对比

Fig. 4 Comparisons of segmentation results of different networks

表1 本文模型与其他网络模型在分割上的对比

Table 1 Comparisons between the proposed model and other network models on segmentation

网络	DSC	IOU	PPV	TPR
UNet	0.5929±0.0118	0.4624±0.0101	0.6546±0.0327	0.6856±0.0503
Attention-UNet	0.5826±0.0137	0.4563±0.0110	0.6477±0.0438	0.6739±0.0941
UNet++	0.5769±0.0159	0.4488±0.0170	0.6506±0.0558	0.6794±0.0600
DeepLabv3+	0.5732±0.0110	0.4396±0.0136	0.5883±0.0177	0.6621±0.0150
SegNet	0.5985±0.0203	0.4711±0.0228	0.6367±0.0390	0.6982±0.0246
MSNet	0.6179±0.0168	0.4937±0.0146	0.6938±0.0239	0.6774±0.0425
JMNet	0.6428±0.0037	0.5196±0.0032	0.6832±0.0125	0.7192±0.0026

本文方法利用主干共享网络FCN学习浅层特征,连接分割和分类2个分支网络进行浅层特

征的共享。利用M-CBAM注意力模块对特征的边缘部分进行保留,并减少尺度多形态问题,还

在一定程度上抑制了无用信息及噪声。共享的信息不同于单独网络提取的信息,会对分割及分类的效果同时进行有效的提升。图5给出了7个网络分割结果的雷达图,图中本文网络显示的环线基本全部在外围,说明各项指标均有较好的表现。

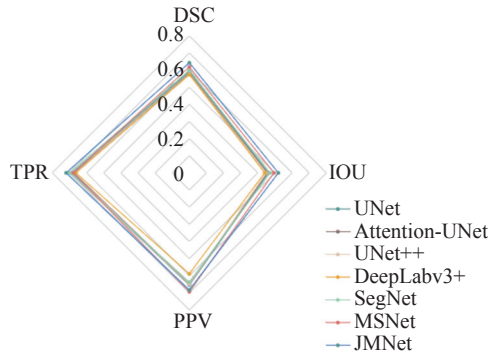


图5 不同网络分割指标雷达图

Fig. 5 Radar map of different network segmentation indicators

图6给出了7个网络在全体样本上进行测试获得的DSC和IOU的累积分布曲线,可以看到JMNet的累积分布曲线下的面积是最大的,测试样本中分割DSC为0的样本数量是最少的,曲线的平缓保持度优于其他网络。IOU的累积分布曲线同样表现出JMNet的分割效果是最好的,由此同样说明带有M-CBAM的多任务网络分割性能的提升体现在大多数的测试样本上。

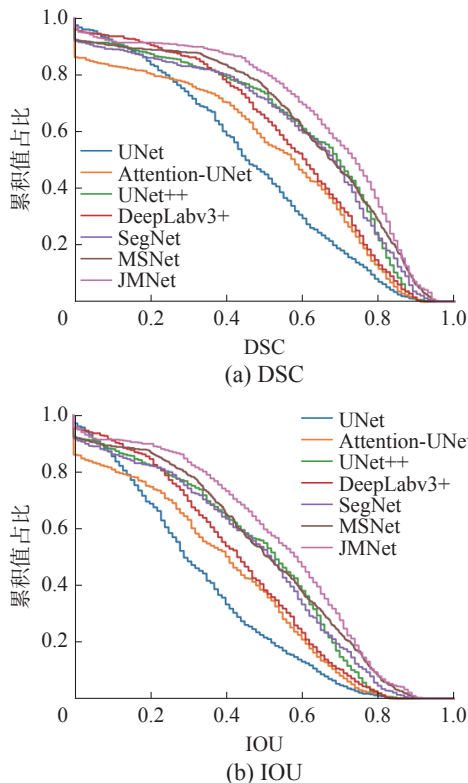


图6 各网络获取的指标累积分布曲线

Fig. 6 Cumulative distribution curves of indicators obtained by each network

表2给出了本文网络在测试集上进行5折交叉验证的结果(以DSC为例)。由表2可知,本文网络JMNet与其他分割模型相比,评价指标在每一次实验上都优于其他模型。结果表明,多任务网络模型JMNet可以有效地提高甲状腺结节的分割效果。

表2 5折交叉验证(DSC)
Table 2 Five-fold cross validation(DSC)

网络	第1折	第2折	第3折	第4折	第5折
UNet	0.5846	0.5762	0.5917	0.6074	0.6044
Attention-UNet	0.5864	0.5640	0.5694	0.5987	0.5944
UNet++	0.5518	0.5978	0.5819	0.5670	0.5858
DeepLabv3+	0.5798	0.5759	0.5631	0.5584	0.5887
SegNet	0.5641	0.5908	0.6145	0.5776	0.6296
MSNet	0.6124	0.5974	0.6119	0.6422	0.6255
JMNet	0.6402	0.6408	0.6406	0.6436	0.6490

表3给出了本文模型在测试集上进行消融实验的分割结果。本文在多任务网络的基础上进行消融实验,用于提升分割的性能。模型在跳跃连接处添加M-CBAM注意力模块时,指标DSC、IOU和TPR有大约5%的提升,PPV有将近10%的提升。模型在主干共享卷积后添加深层卷积块用于提取深层特征时,指标DSC和IOU有大约11%的提升,TPR有大约3%的提升,PPV提升将近18%。将M-CBAM注意力模块和深层卷积块都添加时,指标DSC和IOU有大约14%的提升,TPR提升大约9%,PPV提升将近18%。消融实验后,本文网络在分割性能上超过了单任务分割网络。

表3 分割消融实验
Table 3 Segmentation ablation experiment

M-CBAM	DLCB	DSC	IOU	PPV	TPR
—	—	0.5046	0.3759	0.5224	0.6000
√	—	0.5553	0.4226	0.6129	0.6333
—	√	0.6114	0.4938	0.7007	0.6334
√	√	0.6402	0.5276	0.7037	0.6879

3.3.2 分类结果分析

本文将JMNet与医学图像分类中具有代表性且具有残差结构的主流模型进行了比较。通过对不同分类网络进行测试,得到不同网络的ROC曲线,依次是ResNet34^[2]、ResNet50^[2]、ResNet101^[2]、Densenet121^[12]、DenseNet169^[12]、MCNet和本文方法JMNet的ROC曲线,ROC曲线结果如图7所示。

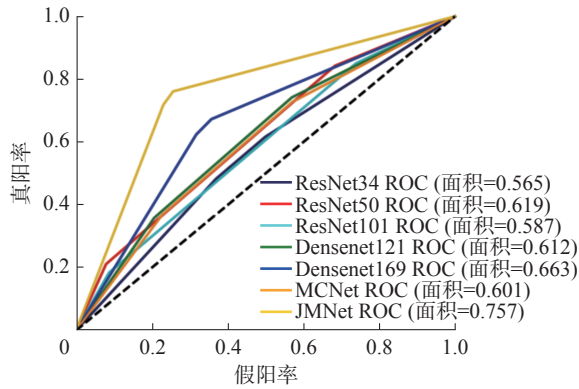


图7 对测试数据集采用不同分类算法分类的 ROC 曲线
Fig. 7 ROC curves classified by different classification algorithms were used for the test data set

由图7可知,分类网络的AUC面积依次是0.565、0.619、0.587、0.612、0.663、0.601、0.757。本文网络模型用多任务方法在分类分支残差模块前后结合M-CBAM,通过M-CBAM和残差模块优化分类性能。在数据集较少的情况下对甲状腺结节超声图像进行分类得到AUC的面积高于其他分类网络。

表4给出了分类数据集中训练、验证、测试的病例个数和样本个数,其中病例为原始数据,样本为数据扩增之后的数据。由于数据集样本有限,将数据分为6:2:2,为了确保验证集的良好数据相近,用于训练出最优的网络模型。将验证集中良性的数据集样本个数通过数据增强扩大为原来的9倍,而恶性的数据集样本个数扩大原来的3倍,形成了最终确定的数据集个数。

表4 数据集中训练、验证和测试样本分布表(以1组为例)
Table 4 Training, validation, and testing sample distribution tables in a dataset (for example, a set of examples)

数据集样本	总病例	总样本	恶性 病例	恶性 样本	良性 病例	良性 样本
训练	183	1647	139	1251	44	396
验证	61	273	46	138	15	135
测试	60	540	46	414	14	126

表5 本文模型与其他模型在分类上的比较

Table 5 Comparison of the model in this article with other models in classification

网络	ACC	AUC	PPV	TPR
ResNet34	0.6129±0.0385	0.5817±0.0329	0.5580±0.0246	0.5810±0.035
ResNet50	0.5615±0.0640	0.5997±0.0538	0.5780±0.0413	0.6010±0.0550
ResNet101	0.4917±0.0745	0.5500±0.0341	0.5510±0.0384	0.5510±0.0320
Densenet121	0.5378±0.1256	0.5707±0.0512	0.5710±0.0480	0.5710±0.0510
Densenet169	0.5015±0.1452	0.5904±0.0424	0.5880±0.0246	0.5880±0.0430
MCNet	0.6185±0.0713	0.5451±0.0370	0.5373±0.0377	0.5574±0.0366
JMNNet	0.7285±0.0256	0.7363±0.0401	0.6790±0.0278	0.7370±0.0400

测试结果的混淆矩阵如图8所示,由图8可知,在M-CBAM和残差模块的作用下,能降低良恶分类不平衡的影响,避免因恶性数据过多造成预测全为恶性的后果。良恶测试总样本数为540例,其中良性为126例,恶性为414例,恶性样本经过测试得到恶性结果为308例,恶性样本准确率为74.4%,良性样本经过测试得到良性结果为97例,良性样本准确率为77.0%,良恶性准确率平均达到76%左右。

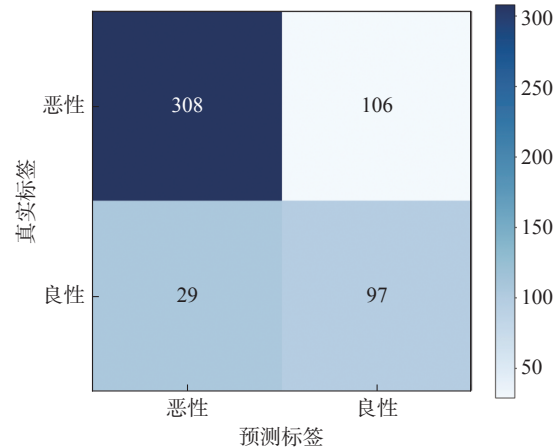


图8 混淆矩阵
Fig. 8 Confusion matrix

表5给出了本文模型与其他分类网络在测试集上的分类结果。由表5可知,在ACC、AUC、PPV和TPR分类指标中,其他分类网络的性能指标不及本文模型,可能是因为在数据集稀少的情况下,其他网络不能充分地学习特征。本文网络模型得到的ACC相较于其他模型有10%~20%的提升,AUC有10%~20%的提升,PPV有10%~15%的提升,TPR有10%~18%的提升。通过与其他分类网络进行比较,可以反映出本文提出的方法在数据集较少的情况下,也能对甲状腺结节超声图像进行较为准确的癌变风险评估。在一定程度上减轻分类不平衡的困扰,并提高分类的性能。结果表明,在数据集较少的情况下,本文网络具有最佳的分类性能。

表6给出了本文网络和对比分分类网络在测试集上进行5折交叉验证的结果(以ACC为例)。由表6可知,本文网络与其他分类网络相比,在每折验证中都显著优于其他分类网络。结果表明,多任务网络可以有效地提高甲状腺结节的分类效果。

表6 5折交叉验证(ACC)
Table 6 Five-fold cross validation(ACC)

网络	1折	2折	3折	4折	5折
ResNet34	0.5611	0.6000	0.6704	0.6407	0.5907
ResNet50	0.4574	0.6537	0.5648	0.5426	0.5889
ResNet101	0.5510	0.3944	0.5204	0.4130	0.5796
Densenet121	0.7352	0.4574	0.4611	0.4019	0.6333
Densenet169	0.4167	0.3444	0.3944	0.6519	0.7000
MCNet	0.6093	0.5185	0.6815	0.6926	0.5907
JMNet	0.7704	0.6981	0.7056	0.7333	0.7352

表7给出了本文网络在测试集上进行消融实验的分类结果。由表7可知,本文在多任务网络的基础上进行消融实验,用于提升分类的性能,当单独添加多尺度卷积注意力模块M-CBAM和深层卷积块DLCB时,对分类的效果没有提升,反而有所下降,同时添加2个模块时,分类的性能才有所提升。多任务学习的任务影响是相互的,分类能影响分割,分割也能影响分类,分割分支网络和分类分支网络的损失函数都用的交叉熵损失函数,分割分类网络同时进行训练,主干网络相同,所以在主干网络部分有大量的共享参数,共享的参数越多,过拟合的概率就越小。将分割和分类损失函数加权求和,使分割与分类的损失对共享层的影响大致相同,得到加权求和后分割与分类总损失,通过设置迭代次数,进行一定迭代次数的训练,一次训练完成之后将结果模型进行保存,开始下一次的训练,并将每次训练的结果进行比较,保存最优模型,当分类模块错误分类或者分割模块错误分割时,损失就会增加,会在主干共享网络去调整共享参数,使每个任务损失达到平衡,相互提高关联任务的性能。

表7 分类消融实验
Table 7 Classification ablation experiment

M-CBAM	DLCB	ACC	AUC	PPV	TPR
—	—	0.7481	0.7447	0.6850	0.7450
√	—	0.7407	0.7481	0.6910	0.7450
—	√	0.7111	0.7011	0.6550	0.7000
√	√	0.7704	0.7840	0.7151	0.7850

4 结束语

本文提出一种联合多任务网络JMNet用于超

声图像的甲状腺结节分割和分类。通过主干共享网络FCN将分割和分类网络联合,并将浅层特征信息共享给联合网络,提高浅层特征的关联性,并提出一种改进CBAM的多尺度卷积注意力机制M-CBAM,将M-CBAM引入到分类网络中,进一步提升了网络的分割和分类的精度。通过实验证明本文所提出的方法较单任务分割以及分类网络能更加有效地提高分割分类性能。如何将任务进行扩充以更好地应用于其他医学图像任务是下一步的研究方向。

参考文献:

- [1] RONNEBERGER O, FISCHER P, BROX T. U-net: convolutional networks for biomedical image segmentation[M]. Cham: Springer International Publishing, 2015: 234–241.
- [2] HE Kaiming, ZHANG Xiangyu, REN Shaoqing, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2016: 770–778.
- [3] HU Jie, SHEN Li, SUN Gang. Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 7132–7141.
- [4] 吴俊霞, 强彦, 王梦南, 等. 基于条件分割对抗网络的超声甲状腺结节分割[J]. 太原理工大学学报, 2023,54(2): 3920–398.
- WU Junxia, QIANG Yan, WANG Mengnan, et al. Ultrasonic thyroid nodule segmentation based on segmentation adversarial network[J]. Journal of Taiyuan University of Technology, 2023,54(2):3920–398.
- [5] 胡屹杉, 秦品乐, 曾建潮, 等. 基于特征融合和动态多尺度空洞卷积的超声甲状腺分割网络[J]. 计算机应用, 2021, 41(3): 891–897.
- HU Yishan, QIN Pinle, ZENG Jianchao, et al. Ultrasound thyroid segmentation network based on feature fusion and dynamic multi-scale dilated convolution[J]. Journal of computer applications, 2021, 41(3): 891–897.
- [6] 赵科甫, 张蕾. 基于卷积神经网络的甲状腺结节超声图像分割[J]. 现代计算机, 2021(15): 54–60.
- ZHAO Kefu, ZHANG Lei. Segmentation of thyroid nodules on ultrasound image using convolution neural network[J]. Modern computer, 2021(15): 54–60.
- [7] CHU Chen, ZHENG Jihui, ZHOU yong. Ultrasonic thyroid nodule detection method based on U-Net network[J]. Computer methods and programs in biomedicine, 2021, 199: 105906.
- [8] OKTAY O, SCHLEMPER J, FOLGOC L L, et al. Attention U-net: learning where to look for the pancreas[EB/OL]. (2018–04–11)[2022–03–31]. <https://arxiv.org/abs/1804.03999>.
- [9] ZHOU Zongwei, RAHMAN SIDDIQUEE M M, TA-

- JBAKSH N, et al. UNet++: A nested U-net architecture for medical image segmentation[M]. Cham: Springer International Publishing, 2018: 3–11.
- [10] CHEN L C, ZHU Yukun, PAPANDREOU G, et al. Encoder-decoder with atrous separable convolution for semantic image segmentation[C]//Computer Vision - EC-CV 2018: 15th European Conference. New York: ACM, 2018: 833–851.
- [11] 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*, 2017, 39(12): 2481–2495.
- [12] HUANG Gao, LIU Zhuang, VAN DER MAATEN L, et al. Densely connected convolutional networks[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2017: 2261–2269.
- [13] 迟剑宁, 于晓升, 张艺菲. 融合深度网络和浅层纹理特征的甲状腺结节癌变超声图像诊断[J]. *中国图象图形学报*, 2018, 23(10): 1582–1593.
- CHI Jianning, YU Xiaosheng, ZHANG Yifei. Thyroid nodule malignantrisk detection in ultrasound image by fusing deep and texture features[J]. *Journal of image and graphics*, 2018, 23(10): 1582–1593.
- [14] 邹奕轩, 周蕾蕾, 赵紫婷, 等. 基于卷积神经网络的甲状腺结节超声图像良恶性分类研究[J]. *中国医学装备*, 2020, 17(3): 9–13.
- ZOU Yixuan, ZHOU Leilei, ZHAO Ziting, et al. Study on the classification of benign and malignant thyroid nodule in ultrasound image on the basis of cnns[J]. *China medical equipment*, 2020, 17(3): 9–13.
- [15] MOUSSA O, KHACHNAOUI H, GUETARI R, et al. Thyroid nodules classification and diagnosis in ultrasound images using fine-tuning deep convolutional neural network[J]. *International journal of imaging systems and technology*, 2020, 30(1): 185–195.
- [16] WEI Xi, GAO Ming, YU Ruiguo, et al. Ensemble deep learning model for multicenter classification of thyroid nodules on ultrasound images[J]. *Med sci monit*, 2020, 26: e926096.
- [17] XIE Jiahao, GUO Lehang, ZHAO Chongke, et al. A hybrid deep learning and handcrafted features based approach for thyroid nodule classification in ultrasound images[J]. *Journal of physics:conference series*, 2020, 1693(1): 012160.
- [18] SHI Guohua, WANG Jiawen, QIANG yan, et al. Knowledge-guided synthetic medical image adversarial augmentation for ultrasonography thyroid nodule classification[J]. *Computer methods and programs in biomedicine*, 2020, 196: 105611.
- [19] SHEN Xueda, OUYANG Xi, LIU Tianjiao, et al. Cascaded networks for thyroid nodule diagnosis from ultrasound images[M]. Cham: Springer International Publishing, 2021: 145–154.
- [20] AMYAR A, MODZELEWSKI R, LI H, et al. Multi-task deep learning based CT imaging analysis for COVID-19 pneumonia: classification and segmentation[J]. *Computers in biology and medicine*, 2020, 126: 104037.
- [21] KIM S J, KIM H S. Multi-tasking U-net based paprika disease diagnosis[J]. *Korean institute of smart media*, 2020, 9(1): 16–22.
- [22] CIPOLLA R, GAL Y, KENDALL A. Multi-task learning using uncertainty to weigh losses for scene geometry and semantics[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 7482–7491.
- [23] WU Yuhuan, GAO Shanghua, MEI Jie, et al. JCS: an explainable COVID-19 diagnosis system by joint classification and segmentation[J]. *IEEE transactions on image processing: a publication of the IEEE signal processing society*, 2021, 30: 3113–3126.
- [24] 叶剑锋, 徐轲, 熊峻峰, 等. 基于注意力机制和辅助任务的语义分割算法[J]. *计算机工程*, 2021, 47(9): 203–209, 216.
- YE Jianfeng, XU Ke, XIONG Junfeng, et al. Semantic segmentation algorithm based on attention mechanism and auxiliary task[J]. *Computer engineering*, 2021, 47(9): 203–209, 216.
- [25] LYU Kejie, LI Yingming, ZHANG Zhongfei. Attention-aware multi-task convolutional neural networks[J]. *IEEE transactions on image processing*, 2020, 29: 1867–1878.
- [26] MEHTA S, MERCAN E, BARTLETT J, et al. Y-net: joint segmentation and classification for diagnosis of breast biopsy images[M]. Cham: Springer International Publishing, 2018: 893–901.

作者简介:



刘侠, 教授, 主要研究方向为医学图像处理、生物特征识别。主持国家自然科学基金项目 2 项、黑龙江省自然科学基金项目 2 项, 授权发明专利 3 项。发表学术论文 40 余篇。



吕志伟, 硕士研究生, 主要研究方向为医学图像分割和机器学习。



王波, 副教授, 主要研究方向为模式识别、医学图像分析与处理、计算机视觉。主持或参与国家自然科学基金项目、黑龙江省自然科学基金项目等 5 项。发表学术论文 30 余篇。