

# 融合残差 Inception 与双向 ConvGRU 的皮肤病变智能分割

顾敏杰<sup>1</sup>, 李 雪<sup>2</sup>, 陈思光<sup>1</sup>

(1. 南京邮电大学物联网学院, 南京 210003; 2. 南京医科大学附属妇产医院(南京市妇幼保健院)皮肤科, 南京 210004)

**摘要:** 由于皮肤病病灶的形状、颜色以及纹理差异极大,且边界不明确,使得传统深度学习方法很难对其进行准确分割。因此本文提出了一种融合残差 Inception 与双向卷积门控循环单元 (Convolutional gated recurrent unit, ConvGRU) 的皮肤病变智能分割模型。首先设计了一种云边协同的皮肤病变智能分割服务网络模型,通过该网络模型,用户可以获得快速、准确的分割服务;其次,构建了一种新的皮肤病变智能分割模型,通过融合残差 Inception 与双向 ConvGRU,该模型能融合不同尺度特征,提高模型特征提取能力,并能充分利用底层特征与语义特征之间的关系,捕获更丰富的全局上下文信息,取得更好的分割性能;最后,在 ISIC 2018 数据集上的实验结果表明,所提出的智能分割模型与近期提出的几种 U-Net 扩展模型相比,取得了更高的准确率与 Jaccard 系数。

**关键词:** 皮肤病;图像分割;残差网络;U型卷积神经网络;卷积门控循环单元

**中图分类号:** TP391 **文献标志码:** A

## Residual Inception and Bidirectional ConvGRU Empowered Intelligent Segmentation for Skin Lesion

GU Minjie<sup>1</sup>, LI Xue<sup>2</sup>, CHEN Siguang<sup>1</sup>

(1. School of Internet of Things, Nanjing University of Posts and Telecommunications, Nanjing 210003, China; 2. Department of Dermatology, Women's Hospital of Nanjing Medical University (Nanjing Maternity and Child Health Care Hospital), Nanjing 210004, China)

**Abstract:** The shape, color and texture of skin lesions are very different, and the boundaries are not clear, which makes it difficult for the traditional deep learning methods to segment them accurately. Based on the above challenge, this paper proposes a residual Inception and bidirectional convolutional gated recurrent unit (ConvGRU) empowered intelligent segmentation model for skin lesion. Specifically, a cloud-edge collaboration intelligent segmentation service network model for skin lesion is firstly designed. By this network model, users can obtain quick and accurate segmentation services. Furthermore, a novel intelligent segmentation model for skin lesion is developed. By integrating residual Inception and bidirectional ConvGRU, this model can fuse multi-scale features and make full use of the relationship between low-level features and semantic features. It improves the ability of the model to extract features

**基金项目:** 国家自然科学基金(61971235);中国博士后科学基金(2018M630590);江苏省“333高层次人才培养工程”;江苏省博士后科研资助计划(2021K501C);南京市妇幼保健院青年人才和南京邮电大学“1311”人才计划资助。

**收稿日期:** 2022-01-21; **修订日期:** 2022-09-16

and capture global context information, and leads to better segmentation performance. Finally, experimental results on ISIC 2018 dataset show that the proposed intelligent segmentation model achieves higher accuracy and Jaccard coefficient than several recently proposed U-Net extended models.

**Key words:** skin lesion; image segmentation; residual network; U-Net; convolutional gated recurrent unit (ConvGRU)

## 引 言

皮肤病是人类最普遍的疾病之一,而恶性皮肤病,如皮肤癌,是造成全世界人类死亡的重要原因之一,其中黑色素瘤是最致命的皮肤癌。2021年美国癌症协会的年度报告称,2021年美国将会有大约10万名新确诊的黑色素瘤患者<sup>[1]</sup>,如果不及早发现,黑色素瘤会深入皮肤,并攻击附近健康细胞。一旦黑色素瘤扩散到身体的其他部位,就变得难以治疗,其5年生存率低于15%<sup>[2]</sup>。Siegel等<sup>[3]</sup>研究发现,如果黑色素瘤能在早期阶段被诊断,5年生存率可达到92%,因此,恶性皮肤病的早期诊断至关重要。然而,良性皮肤病与恶性皮肤病在视觉上极其相似,这给皮肤病鉴别带来了极大困难。当前,皮肤病诊断需求急剧增加,主要依赖临床医生视觉评估来进行诊断,这种方法耗时烦琐,且易出错。

针对上述诊断需求量大与人工诊断易出错的问题,计算机辅助诊断(Computer-aided diagnosis, CAD)系统已普及于皮肤病的初步诊断。皮肤病CAD系统的执行过程分为4个步骤:预处理、分割、特征提取与分类。其中,分割是对皮肤病图像进行分析的基础工作,且分割结果对皮肤病诊断的影响最大<sup>[4]</sup>。分割可通过检查病灶的大小、形状、颜色与纹理来完成。然而,由于皮肤病种类繁多,且病灶的形状、大小和颜色各不相同,有些病灶边界不规则,有些病灶与皮肤之间的过渡很平滑,因此很难进行精确的分割。

随着近年来机器学习与深度学习的快速发展,它们在许多领域取得了良好的效果。与此同时,众多研究者开始将机器学习与深度学习应用于医学图像领域。目前,在图像分割领域,具有代表性的方法大概有3类:基于机器学习的方法、基于全卷积神经网络的方法和基于编码器-解码器架构的方法<sup>[5]</sup>。

基于机器学习的方法。Aljanabi等<sup>[6]</sup>利用人工蜂群算法寻找黑色素瘤分割的最优阈值,然后将得出的最优阈值用于黑色素瘤分割;Khan等<sup>[7]</sup>提出了用于检测与区分黑素瘤与良性痣的智能系统,首先使用高斯滤波器对皮肤病图像进行去噪,然后使用改进的 $k$ 均值聚类算法对病灶进行分割,最后应用支持向量机对皮肤病进行分类;Devi等<sup>[8]</sup>提出了一种基于模糊C均值聚类(Fuzzy C-means, FCM)的非皮肤镜图像病变分割方法,该方法利用局部极大值与欧氏距离检测图像直方图的二项分布特性,并使用直方图属性对FCM进行自动聚类选择,从正常皮肤中分割出病灶区域;文献[9]结合基于像素的马尔可夫随机场与随机区域合并算法,将似然函数分解为像素似然函数与随机区域合并似然函数的乘积,该方法提高了对边界不规则病变的分割精度。以上方法大多基于传统的机器学习算法,通常只能提取底层特征信息,而无法提取高级语义特征,因此,传统方法难以对复杂皮肤镜图像进行精确的分割。

基于全卷积神经网络的方法。为了克服传统机器学习方法在图像分割方面存在的问题,Long等<sup>[10]</sup>提出了全卷积神经网络(Fully convolutional network, FCN)。FCN是最早用于语义分割的深度学习方法,该网络将CNN中的所有全连接层都替换为卷积层,这使得它能接收任意大小的图像并输出相同大小的分割图。FCN经常被用于医学图像分割任务。Kaymak等<sup>[11]</sup>使用4种不同的FCN结构来对皮肤病图像进行分割,包括FCN-AlexNet、FCN-8s、FCN-16s与FCN-32s。Al-Masni等<sup>[12]</sup>提出一种新的FCN架构FrCN(Full resolution convolutional network),它直接学习输入数据每个像素的完整分辨率特征,实现了更精确的逐像素分割;Yang等<sup>[13]</sup>提出了一种随机森林与FCN相结合的皮肤病图像分割方

法,将随机森林与FCN的分割结果融合成最终的图像分割结果;文献[14]提出了一种基于对比度信息融合的FCN用于皮肤病变分割,该网络充分利用了架构中每个卷积块的对比度信息。尽管以上基于FCN的方法提升了皮肤病图像分割的准确率,但也存在一定的局限性,比如,没有考虑全局上下文信息,对图像细节敏感度较低等。

基于编码器-解码器架构的方法。基于深度学习的编码器-解码器体系结构被广泛应用于医学图像分割,其中U-Net<sup>[15]</sup>是目前最流行的用于医学图像分割的神经网络之一。U-Net由捕获上下文信息的编码器与提供精确分割定位的解码器组成。与FCN相比,U-Net的优势在于可以更多地感知到全局上下文信息,并且能够在训练样本缺少的情況下取得良好的分割性能。文献[16]使用空洞空间卷积池化金字塔(Atrous spatial pyramid pooling, ASPP)生成多尺度特征,并且使用Dense连接来组合编码器与解码器的特征图;文献[17]在U-Net的每一步编码步骤与解码步骤之间加入残差连接,保留了编码器阶段由于各层最大池化操作而丢失的信息;Iranpoor等<sup>[18]</sup>在U-Net编码器部分融合预先训练的ResNet网络进行特征提取,并将编码器部分的池化层替换成步长更大的卷积操作;Nazi等<sup>[19]</sup>使用U-Net从周围皮肤中分割出病灶区域,并通过融入dropout模块、数据增强与迁移学习,克服了神经网络应用于小样本医疗数据集的局限性;文献[20]使用10个标准卷积层、金字塔池模块与扩展卷积块代替U-Net中的编码路径,这使得模型能够学习更具有代表性的图像特征。尽管U-Net在医学图像分割领域取得了很大成功,但在特征提取与特征拼接方面存在一定提升空间。例如,原始U-Net使用3×3卷积进行特征提取,无法获得多尺度感受野;虽然U-Net编码器提取的特征具有丰富的图像细节信息,解码器提取的特征包含更多的高级语义特征,但是原始U-Net中的特征拼接操作无法充分利用细节信息与语义特征的依赖关系。

为了进一步提升分割模型特征提取与感知上下文信息能力,本文提出了一种融合残差Inception与双向卷积门控循环单元(Convolutional gated recurrent unit,ConvGRU)的皮肤病变智能分割模型。主要贡献包括:(1)设计了一种云边协同的皮肤病变智能分割服务网络模型,通过该网络,用户可以获得快速、准确的分割服务;(2)提出了一种基于U-Net的皮肤病变智能分割模型,通过融合设计的残差Inception模块以及双向ConvGRU模块,所设计出的编码器与解码器能够提取多尺度特征,并充分利用底层特征与语义特征的关系,以更好地感知全局上下文信息;(3)实验结果表明,提出的模型训练效率高,和现有模型相比具有更好的分割性能。

### 1 网络模型

本节构建了云边协同的皮肤病变智能分割服务网络,其网络架构如图1所示,由用户、边缘服务器、云服务器与医疗机构4部分组成。

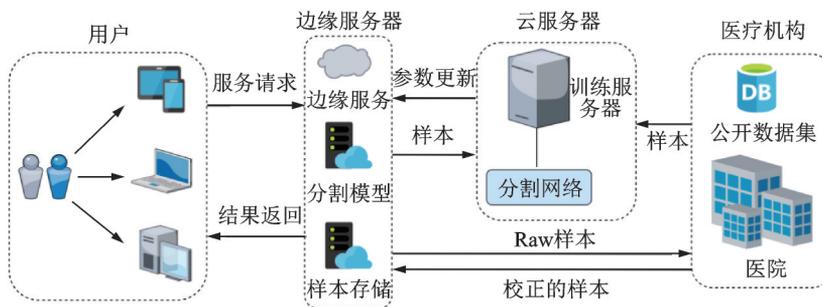


图1 皮肤病变智能分割服务网络模型

Fig.1 Intelligent segmentation network model for skin lesion

(1) 用户。由几种类型的用户终端组成,如平板电脑、智能手机、笔记本电脑与个人电脑等。主要考虑两类用户,包括普通患者与皮肤科医生。具体来说,普通患者可使用相机或手机拍摄皮肤病照片,通过将照片上传至客户端来生成分割服务请求;然后,客户端会把该请求发送给边缘服务器;最后,患者会接收到边缘服务器的服务响应,即获得相应的分割结果。相类似地,当皮肤科医生有辅助分割需求时,可将皮肤镜图片上传至客户端,获得由边缘服务器返回的分割结果,作为准确诊断的基础,反馈的分割结果可提高最终诊断的准确率。

(2) 边缘服务器。边缘服务器的主要功能是在接收到云服务器的分割模型参数后,为用户提供分割服务。当边缘服务器接收到用户的服务请求时,将图片输入至分割模型,然后把分割结果发送至用户。此外,边缘服务器会存储诊断业务中的相关样本。对于置信度高的分割结果,服务器会存储该结果及其原始图片;对于置信度低的结果,服务器会将其与原始图像发送至医疗机构校正,经过校正的样本将发送至边缘服务器进行存储。最后,边缘服务器中的新样本会作为扩充的训练样本用于后续的模式训练。

(3) 云服务器。由于模型训练对计算资源的高度依赖,因此其被部署在云端。模型在云服务器中基于医疗机构提供的数据集进行训练,并将训练参数上传到边缘服务器。同时,基于边缘服务器在提供服务过程中涉及到的新样本,模型能够进一步获得精度提升。

(4) 医疗机构。医疗机构主要有两个功能:①为云服务器中的模型训练模块提供训练样本,样本主要由两部分来源构成,包括国际皮肤成像合作组织(International skin imaging collaboration, ISIC)提供的公开数据集与医疗机构在提供诊断服务过程中收集的相关数据;②将边缘服务器中分割结果置信度较低的样本进行校正,然后将校正好的样本反馈给边缘服务器。

## 2 融合残差 Inception 与双向 ConvGRU 的皮肤病变分割模型

为了提高模型特征提取的准确度以及感知全局上下文的能力,本节融合残差 Inception 与双向 ConvGRU<sup>[21]</sup>,基于 U-Net 提出了一种新的皮肤病变智能分割模型。

### 2.1 分割模型架构

图 2 展示了本节所设计的融合残差 Inception 与双向 ConvGRU 的分割网络架构。与经典的 U-Net 类似,该模型整体架构由左侧的编码器与右侧的解码器组成,基于提升模型特征提取与捕获上下文信息能力,本部分重新设计了编码器与解码器,具体的设计如下:

(1) 编码器。编码器为典型的卷积神经网络结构,它包括两个卷积操作与一个用于下采样的  $2 \times 2$  最大池化。在每个下采样操作后,特征通道的数量将增加一倍。与原始 U-Net 不同的是,为了更好地提取空间特征,卷积操作使用了残差 Inception 模块,并且该设计没有对输入特征图的尺寸带来改变。

(2) 解码器。解码器包括一个用于上采样的  $2 \times 2$  的转置卷积、一个批标准化(Batch normalization, BN)操作、一个特征拼接操作与两个残差 Inception 模块。解码器将进行 4 次上采样操作,在每次上采样操作后,特征图的尺寸将增大一倍,特征通道的数量将减少一半。为了使网络能够更加有效地利用浅层特征与深层特征的关系,在原始 U-Net 的拼接操作中应用了双向 ConvGRU 模块。此外,本文在上采样后添加的一个批标准化操作用来提高神经网络的稳定性与训练速度。在最后一层,使用  $1 \times 1$  卷积核生成与原始输入图像大小相同的特征图。

### 2.2 残差 Inception 模块

基于 GoogLeNet 的 Inception 模块与 ResNet 的残差思想,本节设计了一种残差 Inception 模块代替

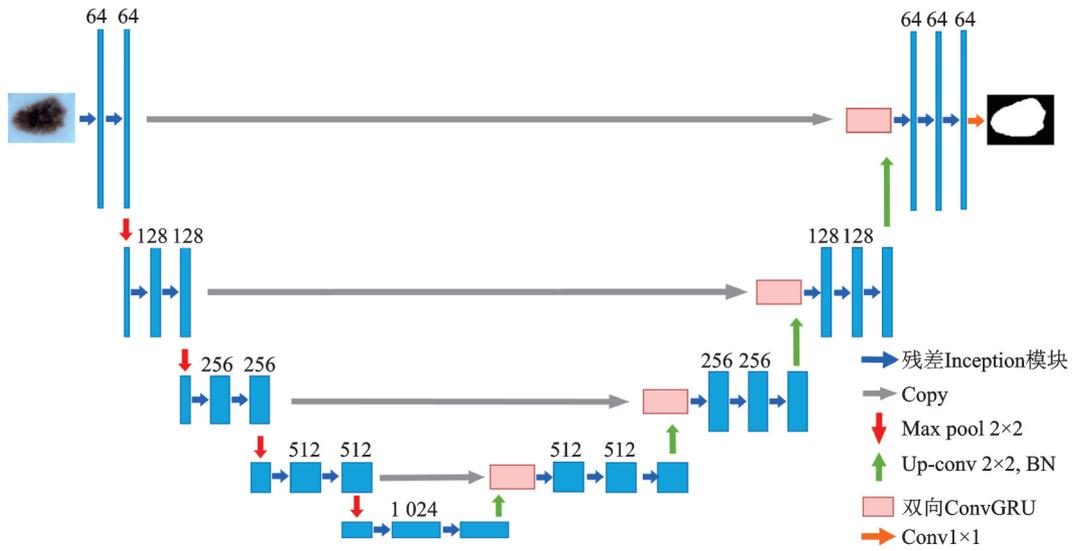


图2 皮肤病变分割模型架构

Fig.2 Model architecture for skin lesion segmentation

传统U-Net中的 $3 \times 3$ 卷积操作。一般来说,增加模型深度与宽度是提高其性能最直接的方法,但增加深度与宽度意味着参数的增加,具有大量参数的模型训练极其耗费计算资源,并且容易产生过拟合问题。为了解决这些问题,将Inception模块引入U-Net中。通过增加网络的深度与宽度以及减少参数,Inception模块可提高计算资源的利用率与图像特征提取的效率。Inception模块采用不同大小的卷积核来获得不同的感受野,最后对不同的特征图进行拼接,融合不同尺度的特征。特别地,为了解决随着网络加深而出现的网络退化问题,如梯度消失与梯度爆炸,本文在Inception模块中添加了残差连接。

设计的残差Inception模块如图3所示,包括3个 $1 \times 1$ 卷积核与3个 $3 \times 3$ 卷积核。它采用 $1 \times 1$ 卷积来减少输出特征图的通道数,这相当于减少参数的数量。与原始的Inception模块不同,设计的模块使用了两个 $3 \times 3$ 卷积核,而不是一个 $5 \times 5$ 卷积。两个 $3 \times 3$ 的卷积核与一个 $5 \times 5$ 的卷积核获得的感受野是等价的,但参数量会少25%左右,可极大减少计算量。前层的输出经过3组不同的卷积操作后,得到具有3组不同感受野的特征图。3组特征图进行拼接后,前层输出的特征图将会经过残差连接与之相加。残差连接的实现定义为

$$m = F(n, \{c_i\}) + n \tag{1}$$

式中: $n$ 与 $m$ 分别表示残差Inception模块的输入与输出; $F$ 函数由上述Inception操作组成; $c$ 表示其中的卷积操作。最后,每个残差Inception模块后面都有一个线性整流函数(Rectified linear units, ReLU)的激活函数。

### 2.3 双向ConvGRU模块

为了充分联系图像的低级特征信息与高级语义特征,设计的分割模型在原始U-Net的特征拼接操

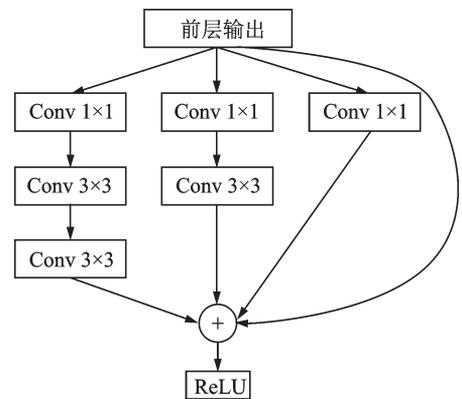


图3 残差Inception模块

Fig.3 Residual Inception module

作中应用双向 ConvGRU 模块。传统的 GRU 在状态转换中采用了全连接操作,它在处理序列数据时效果显著。但是由于它没有对空间数据进行编码,包含了太多的空间数据冗余,因此在处理图像数据时很难获得令人满意的效果。使用传统 GRU 处理图像数据主要存在两个问题:(1)参数数量太大;(2)可能会忽略像素间的空间相关性。

为了解决这个问题,设计的模型引入了 ConvGRU 结构,该结构使用卷积计算代替原始 GRU 的全连接层,这使其既可提取时序信息,又可有效提取空间特征,ConvGRU 的结构如图 4 所示。

ConvGRU 中的各部分作用如下:

(1) 更新门  $z_t$ 。  $z_t$  判断前一个时刻隐藏状态

$h_{t-1}$  对当前输入  $x_t$  的重要程度。  $h_{t-1}$  越重要,其被带入至当前时刻隐藏状态  $h_t$  的信息就越多。  $z_t$  的计算公式为

$$z_t = \sigma(w_{xz} * x_t + w_{hz} * h_{t-1}) \tag{2}$$

式中:  $w_{xz}$  与  $w_{hz}$  分别对应更新门  $z_t$  中  $x_t$  与  $h_{t-1}$  的权重;  $x_t$  为当前时刻的输入;  $*$  为卷积操作;  $\sigma$  为 sigmoid 激活函数,用于将  $z_t$  映射至  $0 \sim 1$ 。

(2) 重置门  $r_t$ 。  $r_t$  控制遗忘前一时刻状态信息的程度。  $r_t$  越大,表示遗忘得越少。  $r_t$  的计算公式为

$$r_t = \sigma(w_{xr} * x_t + w_{hr} * h_{t-1}) \tag{3}$$

式中  $w_{xr}$  与  $w_{hr}$  分别对应重置门  $r_t$  中  $x_t$  与  $h_{t-1}$  的权重。

(3) 候选隐藏状态  $h'_t$ 。 完成  $z_t$  与  $r_t$  的计算后,ConvGRU 将计算候选隐藏状态  $h'_t$  来用于最终隐藏状态  $h_t$  的计算。  $h'_t$  的计算公式为

$$h'_t = \tanh(w_{xh} * x_t + r_t \circ (w_{hh} * h_{t-1})) \tag{4}$$

式中:  $w_{xh}$  与  $w_{hh}$  分别对应候选状态计算中  $x_t$  与  $h_{t-1}$  的权重;“ $\circ$ ”为哈达玛积;  $\tanh$  为双曲正切激活函数,用于将数放缩至  $-1 \sim 1$ 。

(4) 最终隐藏状态  $h_t$ 。  $h_t$  保存当前时刻的信息,并将信息传递给下一个时刻。  $h_t$  计算公式为

$$h_t = (1 - z_t) \circ h'_t + z_t \circ h_{t-1} \tag{5}$$

式中更新门  $z_t$  决定从  $h'_t$  与  $h_{t-1}$  内容中保留多少信息。

传统的 ConvGRU 只处理正向的依赖关系,对历史信息的记忆不足,不能充分学习两个方向的数据依赖关系。为了使网络能够更好地考虑图像中时间序列中的所有信息,本文在 U-Net 中加入双向 ConvGRU 模块。如图 5 所示,双向 ConvGRU 包括一个前向 ConvGRU 与一个后向 ConvGRU。前向 ConvGRU 从左至右控制序列,后向 ConvGRU 从右至左控制序列。图 5 中的  $x_e$  与  $x_d$  分别为编码器输出的特征图以及对应解码器输出的特征图。前向隐藏状态  $\vec{h}_t$  与后向隐藏状态  $\overleftarrow{h}_t$  分别与对应权重相乘,并将结果相加即可得到最终的隐藏状态  $h_t$ ,即双向 ConvGRU 的最终输出结果  $y_t$  为

$$y_t = \tanh(\vec{w}_y * \vec{h}_t + \overleftarrow{w}_y * \overleftarrow{h}_t) \tag{6}$$

式中  $\vec{w}_y$  与  $\overleftarrow{w}_y$  分别对应输出层中前向 ConvGRU 与后向 ConvGRU 的权重。

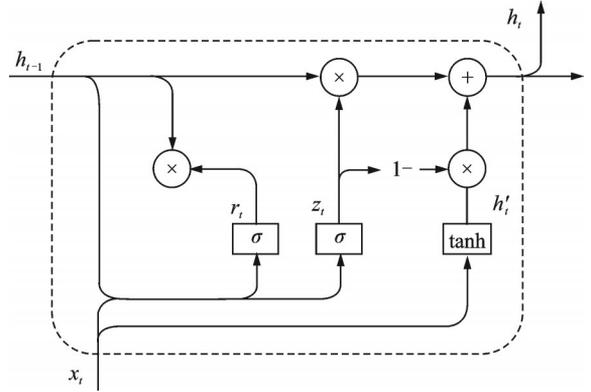


图 4 ConvGRU 结构

Fig.4 ConvGRU structure

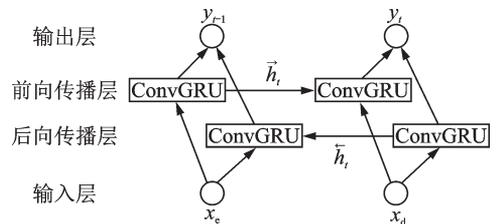


图 5 双向 ConvGRU 模块

Fig.5 Bidirectional ConvGRU module

### 3 实验结果与分析

首先介绍使用的数据集与评价指标,然后通过与其他方法的比较,验证所设计的皮肤病变分割模型的优越性。

#### 3.1 数据集与超参数

采用ISIC 2018<sup>[22]</sup>对模型进行训练与评估。ISIC 2018是国际皮肤成像合作组织发布的大型皮肤镜图像数据集,它由2 594张红绿蓝颜色系统(Red green blue, RGB)三通道皮肤镜病变图像构成,包括7种类别皮肤病:恶性黑色素瘤、黑色素细胞痣、基底细胞癌、光化性角化病、良性角化病、皮肤纤维瘤与血管病变。

数据集按7:1:2的比例随机划分为训练集、验证集与测试集。训练数据由原始图像与相应的真实标签(Ground truth)组成,真实标签为二值PNG格式图像,像素值分别为0与255。像素值0与像素值255分别表示目标病灶外与目标病灶内的区域。由于原始图像分辨率过高并且大小不一,因此所有图像的尺寸调整为 $256 \times 256$ 。

在模型训练过程中,使用自适应矩估计(Adaptive moment estimation, Adam)优化算法与二元交叉熵(Binary cross-entropy)来求解模型的最优参数。训练迭代数(Epoch)设置为60,批大小设置为8。此外,还应用了学习率动态调整策略。具体地,每当模型训练了6个Epoch而其性能提升未超过最小阈值 $1e-4$ ,学习率将会变为当前的0.1倍。当模型经过60轮迭代时训练结束,并保存性能最优的模型。

#### 3.2 评价指标

为了对模型进行准确评估,使用像素准确率(Pixel accuracy)与Jaccard系数作为模型分割性能的评判标准。Jaccard系数是图像分割任务中最常用的性能评价指标之一,它表示真实标签A与预测分割图B的交集面积,除以真实标签A与预测分割图B的并集面积。Jaccard系数与像素准确率的计算公式分别为

$$\text{Jaccard} = \frac{|A \cap B|}{|A \cup B|} \quad (7)$$

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \quad (8)$$

式中:TP、TN、FP和FN分别表示真阳性、真阴性、假阳性和假阴性。

#### 3.3 实验结果对比

图6显示了所设计模型在ISIC 2018训练数据集上准确率(Accuracy)与损失(Loss)随Epoch的变化曲线。从图中可看出,模型收敛速度非常快。网络在第一个Epoch结束就获得了0.8以上的准确率,并且随着Epoch的增加,Loss逐渐降低,Accuracy逐渐提高。在训练30个Epoch后,模型几乎全部收敛,这表明模型不仅分割性能优秀,而且训练效率高。

图7显示了设计模型与一些图像分割基准模型在ISIC 2018数据集上的性能对比。如图7所示,与基准模型相比,本文模型取得了最高的分割性能,主要有以下几点原因。虽然ResNet-50特征提取能力非常强,但是图像分割需要进行像素级的分类,这使得ResNet-50只能将每个像素周围的一个小区域作为输入,进行训练与预测。因此,ResNet-50在训练时需要大量的计算资源,并且由于输入区域的大小限制,网络只能提取局部特征信息。FCN对每个像素进行分类,这使得它无法有效地学习全局上下文信息。U-Net是FCN的一个扩展版本,它将网络编码器部分的特征拼接至解码器部分,避免了有效特征信息的丢失。而设计的模型在U-Net的基础上,使用了残差Inception模块来提取特征,并且在特征拼接操作中应用双向ConvGRU模块,使得网络能融合多种不同尺度的感受野,以及能够更好地感知图像

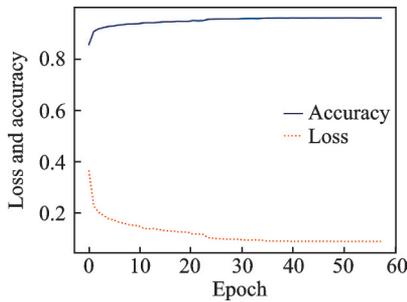


图6 准确率与损失随迭代次数的变化

Fig.6 Changes of accuracy and loss with iterations

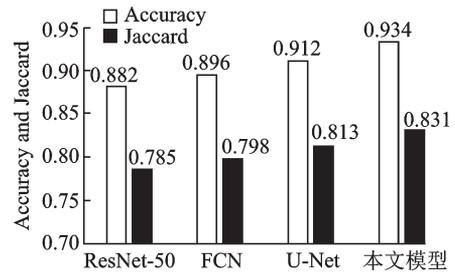


图7 不同模型的准确率与Jaccard系数对比

Fig.7 Comparison of accuracy and Jaccard coefficients of different models

的全局上下文信息,这使其在4种模型中分割性能最强,进一步验证了设计的残差Inception模块与在模型中添加双向ConvGRU的作用。

为验证本文模型中不同模块的有效性,本文进行了消融实验,原始U-Net、仅添加残差Inception的U-Net、仅添加双向ConvGRU的U-Net以及添加残差Inception和双向ConvGRU的U-Net的对比实验结果如图8所示。从图中可以看出,与原始U-Net相比,融合了本文提出模块的模型性能更好,这证明了残差Inception和双向ConvGRU的重大作用。

为更加形象地展现所提出模型的分割性能,图9显示了模型在ISIC 2018上的分割结果样例。对于图9中的所有皮肤病例,模型输出的分割结果与真实标签形状几乎一致,充分验证了模型具备对病灶区域进行精确分割的能力。

为进一步评估设计模型的分割性能,本文还与近期提出的几种U-Net扩展方案Res-Unet<sup>[23]</sup>、Dense U-Net<sup>[24]</sup>与R2U-Net<sup>[25]</sup>进行了对比,结果如表1所示。从表1可以看出,本文模型取得了最高的准确率与Jaccard系数,分别为0.934与0.831。为了增加模型的深度,从而提高模型的准确性,Res-Unet在U-Net中加入残差连接,取得了0.919的准确率和0.818的Jaccard系数。Dense U-Net在U-Net中加入Dense模块,取得了0.921的准确率和0.822的Jaccard系数。上述两个模型虽然缓解了训练过程中出现

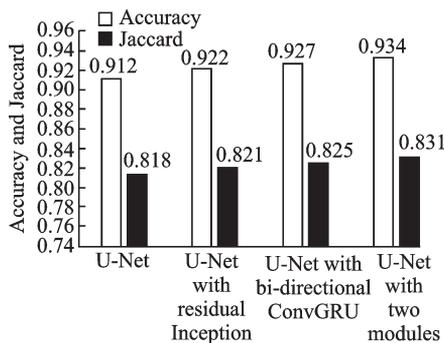


图8 原始U-Net与添加本文提出模块U-Net的准确率与Jaccard系数对比

Fig.8 Comparison of accuracy and Jaccard coefficients of original U-Net and U-Net with the proposed module

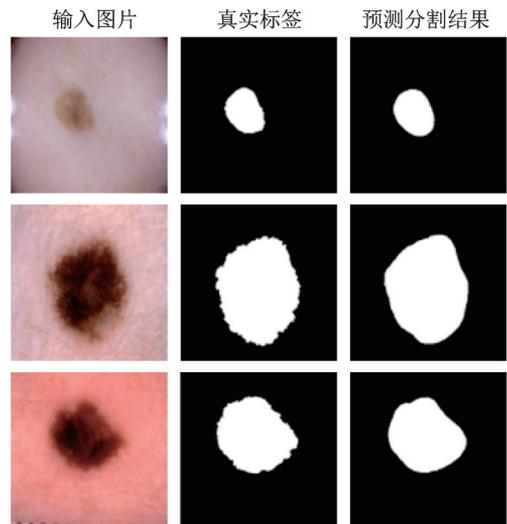


图9 测试图片样例

Fig.9 Examples of test images

的梯度消失问题,但在特征提取方面还有提升空间。R2U-Net 的准确率与 Jaccard 系数分别为 0.929 与 0.827,该模型在 U-Net 中加入循环残差卷积模块,增加了网络深度,从而提高了模型的特征提取能力,但是其利用上下文信息的能力依旧欠缺。与上述方法相比,本文模型不仅加入残差 Inception 模块提高了模型的特征提取能力,还使用双向 ConvGRU 使得网络能够更好地学习图像的时空上下文信息,因此取得了最高的性能指标,体现了本文模型的优势。

#### 4 结束语

本文提出了一种新的皮肤病变智能分割模型。设计了一种残差 Inception 模块,融合多尺度特征,使模型能学习更丰富的特征信息;在上采样阶段,通过融入双向 ConvGRU 模块,充分利用浅层特征与语义特征之间的关系,可提高模型感知全局上下文信息的能力,从而达到提高模型分割精度的目的。实验结果表明,与其他现有方法相比,所提模型显示了显著的性能优势。

#### 参考文献:

- [1] SIEGEL R L, MILLER K D, FUCHS H E, et al. Cancer statistics, 2021[J]. CA Cancer Journal for Clinicians, 2021, 71(1): 7-33.
- [2] WANG Y, LI Y, ZHANG Z, et al. Effect of doping microcapsules on typical electrical performances of self-healing polyethylene insulating composite[J]. Applied Sciences, 2019, 9(15): 3039.
- [3] SIEGEL R, MILLER K D, JEMAL A. Cancer statistics, 2018[J]. CA Cancer Journal for Clinicians, 2018, 68(1): 7-30.
- [4] FAN H, XIE F, LI Y, et al. Automatic segmentation of dermoscopy images using saliency combined with Otsu threshold[J]. Computers in Biology and Medicine, 2017, 85: 75-85.
- [5] MINAEE S, BOYKOV Y Y, PORIKLI F, et al. Image segmentation using deep learning: A survey[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2021, 44(7): 3523-3542.
- [6] ALJANABI M, ÖZOK Y E, RAHEBI J, et al. Skin lesion segmentation method for dermoscopy images using artificial bee colony algorithm[J]. Symmetry, 2018, 10(8): 347.
- [7] KHAN M Q, HUSSAIN A, REHMAN S U, et al. Classification of melanoma and nevus in digital images for diagnosis of skin cancer[J]. IEEE Access, 2019, 7: 90132-90144.
- [8] DEVI S S, SINGH N H, LASKAR R H. Fuzzy C-means clustering with histogram based cluster selection for skin lesion segmentation using non-dermoscopic images[J]. International Journal of Interactive Multimedia & Artificial Intelligence, 2020, 6(1): 26-31.
- [9] SALIH O, VIRIRI S. Skin lesion segmentation using stochastic region-merging and pixel-based Markov random field[J]. Symmetry, 2020, 12(8): 1224.
- [10] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, USA: IEEE, 2015: 3431-3440.
- [11] KAYMAK R, KAYMAK, UCAR A. Skin lesion segmentation using fully convolutional networks: A comparative experimental study[J]. Expert Systems with Applications, 2020, 161: 1-13.
- [12] AL-MASNI M A, AL-ANTARI M A, CHOI M T, et al. Skin lesion segmentation in dermoscopy images via deep full resolution convolutional networks[J]. Computer Methods and Programs in Biomedicine, 2018, 162: 221-231.
- [13] YANG T, PENG S, HU P, et al. Pigmented skin lesion segmentation based on random forest and full convolutional neural networks[C]//Proceedings of SPIE Conference on Optics in Health Care and Biomedical Optics. Guilin, China: SPIE, 2018:

表 1 不同 U-Net 扩展版本的准确率与 Jaccard 系数对比  
Table 1 Comparison of accuracy and Jaccard coefficients of different U-Net extension versions

分割模型	Accuracy	Jaccard
Res-Unet	0.919	0.818
Dense U-Net	0.921	0.822
R2U-Net	0.929	0.827
本文模型	0.934	0.831

514-520.

- [14] CHEN S, JI C, WANG R, et al. Fully convolutional network based on contrast information integration for dermoscopic image segmentation[C]//Proceedings of the 2020 5th International Conference on Mathematics and Artificial Intelligence (ICMAI). Chengdu, China: ACM, 2020: 176-181.
- [15] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[C]// Proceedings of the International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI). Munich, Germany: Springer, 2015: 234-241.
- [16] QAMAR S, AHMAD P, SHEN L. Dense encoder-decoder-based architecture for skin lesion segmentation[J]. *Cognitive Computation*, 2021, 13: 583-594.
- [17] VENKATESH G, NARESH Y, LITTLE S, et al. A deep residual architecture for skin lesion segmentation[C]//Proceedings of OR 2.0 Context-Aware Operating Theaters, Computer Assisted Robotic Endoscopy, Clinical Image-Based Procedures, and Skin Image Analysis.[S.l.]:[s.n.] 2018: 277-284.
- [18] IRANPOOR R, MAHBOOB A, SHAHBANDEGAN S, et al. Skin lesion segmentation using convolutional neural networks with improved U-Net architecture[C]//Proceedings of the 2020 6th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS). Mashhad, Iran: IEEE, 2020: 1-5.
- [19] NAZI A Z, ABIR T A. Automatic skin lesion segmentation and melanoma detection: Transfer learning approach with U-net and DCNN-SVM[C]//Proceedings of International Joint Conference on Computational Intelligence (IJCCCI). Budapest, Hungary: Springer, 2020: 371-381.
- [20] HAFHOUF B, ZITOUNI A, MEGHERBI A C, et al. An improved and robust encoder-decoder for skin lesion segmentation [J]. *Arabian Journal for Science and Engineering*, 2022, 47: 9861-9875.
- [21] SHI X, GAO Z, LAUSEN L. Deep learning for precipitation nowcasting: A benchmark and a new model[C]//Proceedings of the 31st International Conference on Neural Information Processing Systems (NIPS). Los Angeles, USA: MIT Press, 2017: 5617-5627.
- [22] CODELLA N, ROTEMBERG V, TSCHANDL P, et al. Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (ISIC) [EB/OL]. (2019-03-29). <https://doi.org/10.48550/arXiv.1902.03368>.
- [23] XIAO X, LIAN S, LUO Z, et al. Weighted Res-UNet for high-quality retina vessel segmentation[C]//Proceedings of the 9th International Conference on Information Technology in Medicine and Education (ITME). Raleigh, USA: ACM, 2018: 327-331.
- [24] WANG C, ZHAO Z, REN Q, et al. Dense U-Net based on patch-based learning for retinal vessel segmentation[J]. *Entropy*, 2019, 21(2): 168-182.
- [25] ALOM M Z, YAKOPCIC C, HASAN M, et al. Recurrent residual U-Net for medical image segmentation[J]. *Journal of Medical Imaging*, 2019, 6(1): 1-16.

## 作者简介:



顾敏杰(1997-),男,硕士研究生,研究方向:深度学习、智慧医疗等,E-mail: minjegu1997@163.com。



李雪(1986-),通信作者,女,副主任医师,研究方向:色素性皮肤病、基于AI的皮肤病诊断。



陈思光(1984-),通信作者,男,教授,硕士生导师,研究方向:边缘智能与智慧物联网等,E-mail: sgchen@njupt.edu.cn。

(编辑:王静)