

## 基于编码解码结构的微血管减压图像实时语义分割

白瑞峰 江山 孙海江 刘心睿

### Real-time semantic segmentation of microvascular decompression images based on encoder-decoder structure

BAI Rui-feng, JIANG Shan, SUN Hai-jiang, LIU Xin-ru

引用本文:

白瑞峰, 江山, 孙海江, 刘心睿. 基于编码解码结构的微血管减压图像实时语义分割[J]. *中国光学*, 2022, 15(5): 1055–1065.

doi: 10.37188/CO.2022–0120

BAI Rui-feng, JIANG Shan, SUN Hai-jiang, LIU Xin-ru. Real-time semantic segmentation of microvascular decompression images based on encoder-decoder structure[J]. *Chinese Optics*, 2022, 15(5): 1055–1065. doi: 10.37188/CO.2022–0120

在线阅读 View online: <https://doi.org/10.37188/CO.2022–0120>

## 您可能感兴趣的其他文章

### Articles you may be interested in

#### 基于太赫兹量子级联激光器的实时成像研究进展

Progress in real-time imaging based on terahertz quantum-cascade lasers

中国光学 (中英文). 2017, 10(1): 68 <https://doi.org/10.3788/CO.20171001.0068>

#### 无人机侦察多目标实时定位技术研究

Multi-targets real-time location technology for UAV reconnaissance

中国光学 (中英文). 2018, 11(5): 812 <https://doi.org/10.3788/CO.20181105.0812>

#### 光学卫星在轨动态场景实时匹配方法及试验

On-orbit dynamic scene real-time matching method and experiment of optical satellite

中国光学 (中英文). 2019, 12(3): 575 <https://doi.org/10.3788/CO.20191203.0575>

#### 基于改进的局部表面凸性算法三维点云分割

Improved local convexity algorithm of segmentation for 3D point cloud

中国光学 (中英文). 2017, 10(3): 348 <https://doi.org/10.3788/CO.20171003.0348>

#### 基于改进分水岭分割算法的致密荧光微滴识别

Recognition of dense fluorescent droplets using an improved watershed segmentation algorithm

中国光学 (中英文). 2019, 12(4): 783 <https://doi.org/10.3788/CO.20191204.0783>

#### 电磁编码超材料的理论与应用

Theory and application of coding metamaterials

中国光学 (中英文). 2017, 10(1): 1 <https://doi.org/10.3788/CO.20171001.0001>

文章编号 2097-1842(2022)05-1055-11

# 基于编码解码结构的微血管减压图像实时语义分割

白瑞峰<sup>1,2</sup>, 江山<sup>1\*</sup>, 孙海江<sup>1</sup>, 刘心睿<sup>1,3</sup>

- (1. 中国科学院 长春光学精密机械与物理研究所, 吉林 长春 130033;  
2. 中国科学院大学, 北京 100049;  
3. 吉林大学第一医院 神经肿瘤外科, 吉林 长春 130021)

**摘要:**针对真彩色微血管减压图像实时语义分割网络参数量大、语义分割精度低的问题,本文提出了一种适用于微血管减压场景的 U 型轻量级快速语义分割网络 U-MVDNet (U-Shaped Microvascular Decompression Network),该网络由编码解码结构构成。在编码器中设计了轻型非对称瓶颈模块 (LABM) 对上下文特征进行编码,解码器中引入了特征融合模块 (FFM),有效组合高级语义特征和低级空间细节。实验结果表明:对于微血管减压测试集,U-MVDNet 在单 NVIDIA GTX 2080Ti 上的参数量只有 0.66 M,平均交并比 (mIoU) 达到了 76.29%,速度达到 140 frame/s,且当输入图像尺寸为 640×480 时,U-MVDNet 在嵌入式平台 NVIDIA Jetson AGX Xavier 上实现了实时 (24 frame/s) 语义分割。本文方法未使用任何的预训练模型,参数量少且推理速度快,语义分割性能优于其他对比方法,在分割精度和速度上做到了良好的平衡。同时,还可以方便地在嵌入式平台上开发和应用,性能优越,易于部署。

**关键词:**微血管减压图像;编码解码;实时语义分割;U-MVDNet

中图分类号:TP394.1;TH691.9 文献标志码:A doi:10.37188/CO.2022-0120

## Real-time semantic segmentation of microvascular decompression images based on encoder-decoder structure

BAI Rui-feng<sup>1,2</sup>, JIANG Shan<sup>1\*</sup>, SUN Hai-jiang<sup>1</sup>, LIU Xin-rui<sup>1,3</sup>

- (1. *Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China;*  
2. *University of Chinese Academy of Sciences, Beijing 100049, China;*  
3. *Department of Neurosurgery, The First Hospital of Jilin University, Changchun 130021, China)*

\* *Corresponding author, E-mail: 617798169@qq.com*

**Abstract:** Aiming at the problems of large parameters and low semantic segmentation accuracy of real-time semantic segmentation networks for true-color microvascular decompression (MVD) images. This paper proposes a U-shaped lightweight fast semantic segmentation network U-MVDNet (U-Shaped Microvascular De-

收稿日期:2022-06-10; 修订日期:2022-07-05

基金项目:吉林省科技发展计划项目 (No. 20200404155YY, No. 20200401091GX); 白求恩医学工程与仪器中心 (长春) 项目 (No. Bqegczx2019047)

Supported by Jilin Province Science and Technology Development Plan Project (No. 20200404155YY, No. 20200401091GX); Bethune Center for Medical Engineering and Instrumentation (Changchun) (No. BQEGCZX2019047)

compression Network) for MVD scenarios, which consists of encoder-decoder structure. A Light Asymmetric Bottleneck Module (LABM) is designed in the encoder to encode context features. Feature Fusion Module (FFM) is introduced in the decoder to effectively combine high-level semantic features and underlying spatial details. Experimental results show that for the MVD test set, U-MVDNet achieves 0.66 M parameters, 76.29% mIoU (mean Intersection-over-Union), and 140 frame/s speed on NVIDIA GTX 2080Ti. And when input image size is  $640 \times 480$ , the real-time (24 frame/s) semantic segmentation is realized on NVIDIA Jetson AGX Xavier embedded development board. The proposed network has no pretrained model, fewer parameters, and fast inference speed. The semantic segmentation performance is superior to other comparison methods, and a good trade-off between segmentation accuracy and speed is achieved. Furthermore, U-MVDNet can also be easily developed and applied on embedded platform with superior performance and easy deployment.

**Key words:** microvascular decompression images; encoder-decoder; real-time semantic segmentation; U-MVDNet

## 1 引言

三叉神经痛和面肌痉挛是最常见的脑神经疾病,三叉神经痛主要表现为三叉神经分布区出现短暂的、类似电击的疼痛,轻微的触碰即可诱发<sup>[1]</sup>,面肌痉挛表现为单侧面部肌肉无痛性和反复性不自主抽搐<sup>[2]</sup>,严重影响患者的生活质量。微血管减压术 (Microvascular Decompression, MVD) 因手术创伤小,疗效显著,成为目前治疗三叉神经痛和面肌痉挛最常用的手术方式。

一般脑血管图像的获取途径有电子计算机断层扫描血管造影 (CTA)、数字减影血管造影 (DSA)、磁共振血管造影 (MRA) 等。其中,CTA 方案最为常见,NAZIR 等人<sup>[3]</sup>通过残差映射解决网络收敛问题,提出了一种高效的从 CTA 图像中自动分割脑血管的融合网络,但 CTA 的局限性在于小血管成像不够清晰,使得诊断准确率较低<sup>[4]</sup>。DSA 是诊断的“金标准”,为提高 DSA 中较小血管的分割准确性,PATEL 等人<sup>[5]</sup>提出了一种通过使用多分辨率输入来增强架构视野的专用卷积神经网络 DeepMedic,但是 DSA 是有创检查且存在放射性暴露等问题,且费用较为昂贵,不适合初期筛查,临床应用的局限性较大<sup>[6]</sup>。MRA 易受血管走向、血管壁钙化等的影响,且血管细小小时成像效果较差<sup>[7]</sup>,为改善小血管的分割效果,通过引入反向边缘注意网络关注缺失的脑血管边缘特征和细节信息<sup>[8]</sup>,或通过集成 3D U-Net、多尺度方法以及深度监督方法鉴别脑部时间飞跃磁共振血管成像

(TOF-MRA)中的小血管,引导网络中间层更好地生成判别性特征,有效地避免了梯度爆炸和梯度消失的问题,提高了模型的收敛性,从而实现对脑血管的分割<sup>[9-10]</sup>。与传统 CTA、DSA、MRA 等获取脑血管图像的方法相比,通过内窥镜获取真彩微血管减压 (MVD) 图像更简单方便,而且对真彩 MVD 图像的处理具有更大的发展前景。

由于脑血管自身结构的复杂性,早期脑血管分割算法主要基于强度模式识别,脑血管的分割方法主要有非局部均值分割方法<sup>[11]</sup>、阈值分割方法<sup>[12]</sup>、区域增长分割方法<sup>[13]</sup>和水平集方法<sup>[14-15]</sup>。随着卷积神经网络在不同领域的广泛应用<sup>[16]</sup>和持续优化<sup>[17]</sup>,卷积神经网络对医学图像的分割也呈现出了井喷式的发展,性能显著优于传统分割方法。大多数先进的语义分割网络是基于全卷积神经网络(FCN)<sup>[18]</sup>和 U-Net<sup>[19]</sup>网络的,其中,ZHANG M 等人<sup>[20]</sup>提出了一种基于 U-Net 的深度学习方法,结合预处理方法,在 DSA 图像中跟踪和分割脑血管。XIA L K 等人<sup>[21]</sup>将 Inception-ResNet 和维度扩展的 U-Net++相结合,实现了最大化多尺度特征的提取,对于 MRA 分割性能优异。MENG C 等人<sup>[22]</sup>设计了编码器-解码器架构,通过定义一个多尺度模块来分割不同直径的脑血管。

本文引入了真彩色 MVD 图像,MVD 图像中脑血管的分割面临着脑血管类内过于相似、脑血管边界与其他脑组织边界不够明显的问题,使得脑血管分割困难,脑神经亦是如此。针对上述问题,本文提出了一种新颖的基于深度学习的实时

语义分割网络, 称为 U-MVDNet, 该网络具有更加高效的编码结构和更加精细的解码结构。为训练和评估所提出的 U-MVDNet 网络效果, 与吉林大学第一医院积极展开合作, 构建了一个包含 3087 张 MVD 图像的数据集。这些图像具有标准的 Ground truth 注释, 为实时推断术中脑血管、脑神经的位置提供指导。借助于实时准确分割的帮助, 医生可以快速确定脑血管和脑神经的最佳位置, 节省了大量时间, 减轻术者的工作量, 有效辅助医生进行快速诊断, 达到专业医生甚至超越专业医生的水平。对本文方法与其他几种先进的实时语义分割方法进行比较。实验结果表明, 本文所提出的方法取得了良好的性能。

## 2 U-MVDNet 架构构建

本章首先对网络架构进行整体描述。然后, 详细介绍轻型深度瓶颈模块 (LABM) 和特征融合模块 (FFM), 并阐述这两个模块的有效性。

### 2.1 网络架构

基于轻型非对称瓶颈模块和特征融合模块设计了 U-MVDNet 体系结构, 如图 1 所示, 其中“C”表示级联操作。为了得到推理速度较快、mIoU 较高和轻量级的网络模型, 分析 U-MVDNet 模型的最优设计。U-MVDNet 的架构细节如表 1 所示。

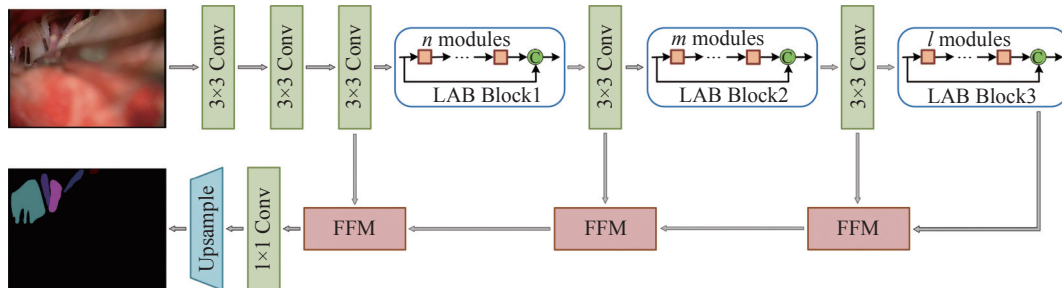


图 1 U-MVDNet 架构

Fig. 1 Architecture of U-MVDNet

表 1 U-MVDNet 架构细节

Tab. 1 Architecture details of proposed U-MVDNet

Layer	Operator	Mode	Channel	Output size
1	3×3 Conv	stride 2	32	256×256
2	3×3 Conv	stride 1	32	256×256
3	3×3 Conv	stride 1	32	256×256
4-5	n×LABM	dilated 2	32	256×256
6	3×3 Conv	stride 2	64	128×128
7-8	m×LABM	dilated 4	64	128×128
9	3×3 Conv	stride 2	128	64×64
10-12	l×LABM	dilated 8	128	64×64
13	1×FFM	-	128	64×64
14	1×FFM	-	64	128×128
15	1×FFM	-	32	256×256
16	1×1 Conv	stride 1	10	256×256
17	Bilinear interpolation	×2	10	512×512

在 U-MVDNet 网络模型中编码器由 3×3 卷积层和轻型深度瓶颈块 (LAB block) 构成。其中, 步长为 2 的 3×3 卷积层构成了下采样块。下采样操作缩小了特征图, 扩大了感受野, 能提取到更多的上下文信息。LAB block 中包含几个连续的 LABM, 便于稠密特征的提取。第一个、第二个和第三个 LAB block 分别由 n 个、m 个和 l 个 LABM 组成。为了加强空间关系和特征传播, 引入了模块拼接方法, 实现高级特征与低级特征间的融合。在 LABM 中引入扩张卷积, 获取到较大感受野和更复杂的特征, LAB block1、LAB block2 和 LAB block3 的扩张率分别为 2、4 和 8, 扩张率的选择符合本文特定的微血管数据集, 并且需获得适应于该数据集的感受野。

对于解码器, 本文使用 3 个 FFM 来聚合低级特征和高级特征, 逐步恢复分辨率。最后, 采用 1×1 卷积层和 2 倍上采样完成分割。相较于大多数语义分段模型中使用的解码器结构来说, 一般通过 4 倍或 8 倍上采样得到分割预测图, 本文采

用的 2 倍上采样,可以保留更多的特征信息,使边界信息更加完整,语义信息更加清晰。

本文模型属于端到端的深度学习体系结构,而且不依赖于任何主干,U-MVDNet 的内存消耗极低,参数量不足 66 万。

## 2.2 轻型非对称瓶颈模块 (LABM)

对于实时语义分割来说,在保证精度的前提下提升网络速度,就不可避免地要利用小的输入图像或轻量级的基础模型,本文通过对 ResNet<sup>[23]</sup> 中瓶颈模块 (见图 2(a)) 的观察,设计了轻型非对称瓶颈模块 (LABM)(见图 2(b)),LABM 兼具了瓶颈模块和因式卷积的优点。图 2 中的“w”表示输入通道数,“DCConv”表示深度扩张卷积。

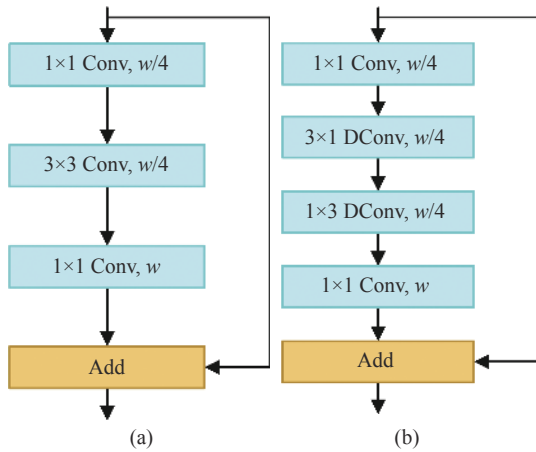


图 2 (a)ResNet 瓶颈设计及(b)轻型非对称瓶颈模块

Fig. 2 (a) ResNet bottleneck design and (b) LABM

图 2(b) 所示的瓶颈模块由  $1 \times 1$  卷积层、 $3 \times 1$  深度卷积层、 $1 \times 3$  深度卷积层和最终  $1 \times 1$  点卷积层组成,并在输入通道数与输出通道数相同时利用了残差连接。通道数在通过第一次  $1 \times 1$  卷积层时减少为原来的  $1/4$ ,之后通过  $3 \times 1$  和  $1 \times 3$  深度卷积层时通道数保持不变,最后由  $1 \times 1$  点卷积层恢复到原始通道数。与标准卷积相比,深度卷积减少了大量的计算成本,并与点卷积的组合非常有效。

在每个 LABM 的开头使用了  $1 \times 1$  卷积层,通道数减少了  $3/4$ ,与 ResNet 网络模型中的数千个通道相比,本文模型中的最大通道数只有 128 个,有效保存了大量的空间信息,减少了网络模型的运行时间和内存需求,提高了推理速度。

为进一步减少参数量,将卷积分解应用于深度卷积,得到更加轻量级的结构。ALVAREZ J 等人<sup>[24]</sup> 提出了标准卷积层可以用 1D 滤波器进行分

解。设  $\mathbf{W} \in \mathbb{R}^{C \times d^h \times d^v \times U}$  表示典型 2D 卷积层的权重, $C$  表示输入特征图的数量, $U$  表示输出特征图的数量, $d^h \times d^v$  表示每个特征图的内核大小 (通常  $d^h = d^v \equiv d$ )。设  $\mathbf{b} \in \mathbb{R}^U$ , 表示每个滤波器的偏置项的向量。分解层的第  $i$  个输出  $\mathbf{a}_i^l$  属于下式输入  $\mathbf{a}_c^0$  的函数,本文中为  $\mathbf{a}_c^0$ :

$$\mathbf{a}_i^l = \varphi \left( \mathbf{b}_i^l + \sum_{l=1}^L \bar{\mathbf{h}}_{il}^T * \left[ \varphi(\mathbf{b}_i^l) + \sum_{c=1}^C \bar{\mathbf{v}}_{lc} * \mathbf{a}_c^0 \right] \right), \quad (1)$$

其中,  $\bar{\mathbf{h}}_{il}^T$  和  $\bar{\mathbf{v}}_{lc}$  表示长度为  $d$  的向量; $L$  表示中间层滤波器的数量,  $\varphi(\cdot)$  指 PRelu<sup>[25]</sup> 激活函数。标准的  $n \times n$  深度卷积被  $n \times 1$  深度卷积和  $1 \times n$  深度卷积代替,对于  $N \times N$  的卷积核,非对称卷积将每个像素的计算复杂度从  $O(N^2)$  降低到  $O(N)$ 。同时,为了提取更加丰富的上下文信息,本文在深度卷积中引入了扩张率,将  $3 \times 1$  和  $1 \times 3$  的深度卷积改进为深度扩张卷积,使得在不降低特征图分辨率的情况下增加了感受野,便于从 MVD 图像中捕获到更复杂的特征。

## 2.3 特征融合模块 (FFM)

语义分割模型通常是由一个基于前端的编码器和一个基于后端的解码器网络组成的综合体。编码器由多个卷积层组成,用来获取图像的总特征和局部特征,其中的卷积层和池化层会逐渐缩减输入数据的空间维度,从而降低特征维度。解码器则是由多个反卷积或上池化组成,逐步恢复目标的细节和空间维度,将编码器学习到的低分辨率特征从语义上映射到高分辨率的像素空间,进行像素分类。通常,采用双线性上采样或者几个简单的卷积构成的解码器忽略了低层次信息,导致分割精度低。

高级特征包含语义信息,低级特征包含丰富的空间细节。由于语义层次和空间细节的差异,简单的低级特征和高级特征难以有效的融合。因此,可以在低级特征中引入语义信息,在高级特征中引入空间细节,实现增强特征融合。本文编码器由 LAB block 和下采样模块构成,通过该编码器得到了充分的语义信息,但是由于下采样操作导致大量空间位置信息丢失,为保证边界分割的质量,选择将包含丰富空间信息的低级特征嵌入到高级特征。

FFM 结构采用了 U 型结构, 如图 3 所示, 用来融合具有空间信息的低级特征和语义信息的高级特征。低级特征首先经过了  $1 \times 1$  卷积层, 并利用批量归一化处理来平衡特征尺度, 然后由通道平均池化压缩低级特征和 sigmoid 激活函数来生成单通道注意力图, 紧接着将其与  $3 \times 3$  卷积后的高级特征相乘。这里需要注意的是, 对于 U-MVDNet 中初始 FFM 的高级特征并不进行上采样处理, 而是直接与单通道注意力图相乘, 而对后续的 2 个 FFM 需要对高级特征进行上采样后, 再与单通道注意力图相乘。最后, 通过元素加法对高级特征与相乘后得到的加权特征进行融合。简而言之,  $F_L \in \mathbb{R}^{C_L \times H_L \times W_L}$  和  $F_H \in \mathbb{R}^{C_H \times H_H \times W_H}$  分别是 FFM 中输入的低级和高级特征。FFM 的计算如下:

$$F = \varphi(BN(\sigma(\text{AvgPool}(BN(f^{1 \times 1}(F_L)))))) \otimes f^{3 \times 3}(F_H) + f^{3 \times 3}(F_H)) \quad (2)$$

其中,  $\sigma$  表示 sigmoid 函数, AvgPool 表示平均池化操作, BN 表示批量归一化的处理,  $f^{1 \times 1}$  表示  $1 \times 1$  卷积层,  $f^{3 \times 3}$  表示  $3 \times 3$  卷积层,  $\otimes$  表示逐元素相乘,  $+$  表示逐元素相加,  $F \in \mathbb{R}^{C_L \times H_L \times W_L}$  表示最终的输出特征图。

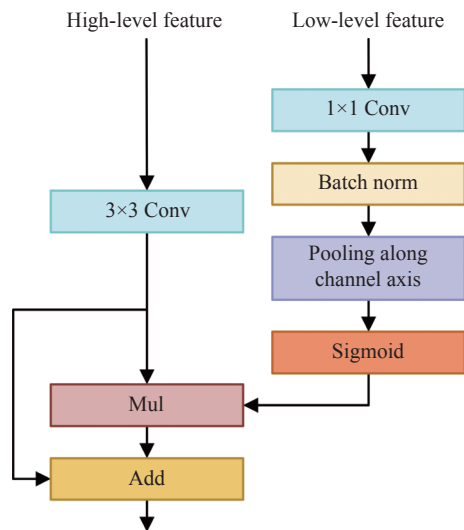


图 3 特征融合模块流程图

Fig. 3 Flow chart of feature fusion module

低级特征生成的空间注意力图反映了每个像素的重要性, 包含丰富的空间信息, 可指导特征学习, 利用空间细节来细化边界。FFM 模块提取了空间注意力图, 并将其嵌入到高级特征中, 之后与包含丰富语义信息的高级特征进行有效融合。

### 3 实验与分析

在本章中, 将在吉林大学第一医院提供的 MVD 数据集和 ISIC 2016<sup>[26]</sup> + PH2<sup>[27]</sup> 皮肤镜图像数据集上评估所提出的网络。首先, 介绍 MVD 数据集和 ISIC 2016 + PH2 数据集, 以及训练实施方案; 然后, 在 MVD 验证集上进行消融实验, 以证明所提出方法的有效性; 之后, 在 MVD 测试集上分别与其他实时语义分割网络进行对比, 以验证提出网络的分割效果, 同时在 ISIC 2016 + PH2 测试集上验证提出网络的泛化能力; 最后, 在嵌入式终端下完成对比, 并报告了 MVD 测试集下的分割精度和速度结果。

#### 3.1 数据集

##### 3.1.1 MVD 数据集

在医学领域, 医学图像普遍具有语义简单、数据量少和难以获取的特点, 而且, 针对基于深度学习的医学图像的分割需要医生进行专业的标注。本项研究涉及 60 例患者, 其中, 男 23 例, 女 37 例, 年龄为 40~70 岁。微血管减压数据是由 ZEISS 生产的 OPMI@ VARIO 700 型号手术显微镜采集得到。选择在释放脑脊液到缝合硬膜这段时间内获取的 3087 张 MVD 图像, 并由专家们手动标注。然后参考 PASVOL VOC 2012 数据集格式, 得到用于语义分割网络训练的 MVD 数据集。该数据集有 9 个类别 (当添加背景时, 包含 10 个类别)。类别名称及其相应颜色见表 2 (彩色版见期刊电子版)。

表 2 医学术语缩写及对应颜色

Tab. 2 Abbreviations of medical terms and corresponding color

简称	全称	对应颜色
cn5	三叉神经	
cn7	面神经	
cn9	舌咽神经	
cn10	迷走神经	
aica+cn7	小脑前下动脉及面神经	
pica+cn7	小脑后下动脉及面神经	
pica	小脑后下动脉	
aica	小脑前下动脉	
pv	岩静脉	

MVD 数据集属于微血管减压术中场景数据集, 包含了 3087 张精细注释的 MVD 图像。实验随机选取 1806 张用于训练, 973 张用于验证, 308 张用于测试, 图像的分辨率为  $768 \times 576$ 。

### 3.1.2 ISIC 2016 + PH2 数据集

ISIC 2016 + PH2 数据集包含了两个来自不同中心的样本, 用于评估皮肤病灶分割的准确性。其中, ISIC 2016 包含了 900 个训练样本和 379 个验证样本。PH2 数据集包含 200 个样本。本文将 ISIC 2016 数据集用于网络模型的训练和验证, PH2 数据集上完成网络模型测试。

## 3.2 实验设置

训练方案实施: 选用单 NVIDIA GTX 2080Ti, 并在 CUDA 10.1、cuDNN 7.6 和 Pytorch1.7 下完成网络训练, 训练参数如表 3 所示。训练中采用了随机水平翻转、随机高斯模糊、随机裁剪和标准化策略来增广数据, 且对提出的网络未进行任何预训练。在数据集上进行了 100 轮的训练, 最后使用相应的评估指标来评估分割精度。

表 3 训练参数

Tab. 3 Training parameters

Parameter name	Parameter selection		
Learning rate	Policy	Initialization	Power
	poly	0.16	0.9
Optimizer	Policy	Momentum	Weight decay
	SGD	0.9	$1 \times 10^{-4}$
Enter picture size	768 × 576		
Batch size	8		

## 3.3 消融实验

本节设计了一系列消融实验来证明本文提出网络 U-MVDNet 的有效性, 且消融实验都是在单 NVIDIA GTX 2080Ti 下完成的。这些消融研究基于 MVD 数据集上的训练集进行训练, 在 MVD 验证集上完成评估, 观察该网络中各成分的影响。

扩张率消融实验: 设计了具有 3 种不同扩张率的 LAB block, 分别为 LAB block1、LAB block2 和 LAB block3, 这 3 个 LAB block 构成本文的编码器。选取合适的感受野可以学习到更好的多尺度特征, 当感受野过大时, 会导致小目标的丢失。本文对 3 个 LAB block 分别设置了不同的扩张率进行比较, 扩张率分别为 2、4、8 和 4、8、16。如

表 4 所示, 针对 MVD 图像, 扩张率选择为 2、4、8 时更有效。

表 4 不同扩张率组合的 LABM 编码器结果

Tab. 4 Results of LABM encoder with different combinations of dilation rates

Name	Dilation rates	mIoU(%)
LABM_N2M2L4	2,4,8	72.35
LABM_N2M2L4	4,8,16	72.08

级联消融实验: 在 LAB block 中, 将 LAB block 的输入特征与输出特征进行级联操作。如表 5 所示, 其中, “Params”表示参数量, “FLOPs”指的是计算量。在  $n=2, m=2, l=4$  时, LAB block 中加入级联操作后准确度提高了 0.73%, 准确度达到 73.08%, 效果更优。级联操作应用于编码器, 有效增加了信息流。

表 5 不同设置下的 LABM 编码器结果

Tab. 5 Results of LABM encoder with different settings

Concatenation	Params(M)	FLOPs(G)	mIoU(%)
	0.30	2.81	72.35
√	0.54	4.03	73.08

编码器深度消融实验: 对 LAB block1、LAB block 2 和 LAB block 3 分别使用不同数量的 LABM 来改变编码器的深度。由表 6 可以看出,  $m$  和  $l$  的值对精确度的影响比  $n$  更大,  $m$  和  $l$  分别叠加 4 个 LABM 时可以获得精细的分割结果。当  $n, m$  和  $l$  都提高到 4 时, FLOPs 明显变大, 但分割精度降低了。因此, 在分割精度和计算复杂度之间做权衡, 最终选择  $n$  为 2,  $m$  为 4,  $l$  为 4。

表 6 输入尺寸为  $512 \times 512$  时, 不同深度的编码器结果

Tab. 6 Results of encoder with different depths when the input size is  $512 \times 512$

$n$	$m$	$l$	Params(M)	FLOPs(G)	mIoU(%)
2	2	2	0.52	3.95	72.35
2	2	4	0.54	4.03	73.08
2	4	4	0.55	4.11	73.84
4	4	4	0.55	4.20	73.37

解码器消融实验: 在 U-MVDNet 中, 通过 LAB block 来提取特征, 选择使用 FFM 来融合特征。

从表 7 可以看出, 沿通道平均池化操作能够提高准确性, 确保有效的获取空间细节。这表明通过 FFM 将空间细节嵌入到高级特征中可以有效提高精度, 得到更好的像素级预测。

表 7 不同构成要素的 FFM 解码器结果

Tab. 7 Results of FFM decoder with different components

FFM	Pooling	mIoU(%)
w/o	–	73.84
w		77.11
w	√	77.34

mIoU 性能消融实验: 为了探索扩张卷积对 mIoU 性能的影响, 设计了 2 个对比实验。在第一个实验中, 去除了 U-MVDNet 中的所有扩张卷积, 在另一个实验中, 将 U-MVDNet 开头的  $3 \times 3$  标准卷积, 设置为扩张率为 2 的扩张卷积。如表 8 所示, 当移除所有的扩张卷积时, mIoU 有明显的降低 (范围从 77.34% 到 75.61%)。当用扩张率为 2 的扩张卷积替换标准卷积时, mIoU 也出现了下降, 从 77.34% 降到 76.81%。实验结果证明, 扩张卷积对 mIoU 性能有显著影响。

表 8 U-MVDNet 的扩张率对 mIoU 的影响

Tab. 8 Effect of dilation of U-MVDNet on mIoU

Concatenation	mIoU(%)	Params(M)
U-MVDNet	77.34	0.66
U-MVDNet_w/o dilation	75.61	0.66
U-MVDNet_First $3 \times 3$ conv ( $r = 2$ )	76.81	0.66

表 9 MVD 测试集实验结果

Tab. 9 Experimental results on MVD test set

Method	Params(M)	Speed(frame·s <sup>-1</sup> )	mIoU(%)	cn5	cn7	cn9	cn10	aica+cn7	pica+cn7	pica	aica	pv
CGNet <sup>[28]</sup>	0.94	87.4	71.95	81.26	82.9	71.29	69.85	71.64	87.16	67.37	65.66	50.42
EDANet <sup>[29]</sup>	0.69	125	74.51	83.03	84.02	70.31	77.25	75.09	87.98	70.37	68.18	54.34
ContextNet <sup>[30]</sup>	0.88	163.3	75.81	82.14	84.15	74.91	78.08	76.67	87.84	72.08	69.77	56.65
U-MVDNet	0.66	140.8	76.29	82.25	85.45	74.8	76.91	76.32	87.85	74.08	69.83	59.12

如图 5 所示, 在第一列中, 只有本文所提出的 U-MVDNet 可准确定位“aica”的分割边界, 目标轮廓清晰。在第二列和第三列中, CGNet<sup>[28]</sup> 有明显的多像素混合问题, 并且在第二列中 CGNet<sup>[28]</sup>、

### 3.4 实验对比

为验证 U-MVDNet 的分割性能, 分别在 MVD 数据集和 ISIC 2016<sup>[26]</sup> + PH2<sup>[27]</sup> 数据集上进行实验。

#### 3.4.1 MVD 数据集测试实验

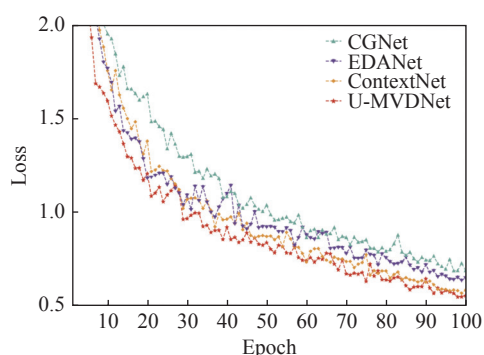
本文结合 LAB block 和 FFM 搭建了完整的网络, 并在 MVD 数据集上进行实验。首先, 使所有网络在 MVD 训练集、批量尺寸为 8、交叉熵损失函数和小批量 SGD 下完成 100 轮的训练。然后, 在单 NVIDIA GTX 2080Ti 下进行测试, 评估  $768 \times 576$  分辨率下的推理速度, 并与 CGNet<sup>[28]</sup>、EDANet<sup>[29]</sup> 和 ContextNet<sup>[30]</sup> 进行比较, 这 3 种网络都属于实时语义分割网络, 能够以小的参数量得到准确的分割结果。

从表 9 可以看出, U-MVDNet 拥有约 66 万个参数, 参数量最小, 接近于 EDANet<sup>[29]</sup>, 分割精度相较于 EDANet<sup>[29]</sup> 提高了 1.78%。除此之外, ContextNet<sup>[30]</sup> 是最快的实时网络之一, 比本文网络稍快, 然而, 该网络的分割精确度为 75.81%, 比 U-MVDNet 减少了 0.48%, 而且参数比 U-MVDNet 多出 22 万个。通过不同分割网络的比较可以发现, U-MVDNet 在保证精度的前提下仍具有较快的推理速度, 并在参数量和准确率上明显优于其他实时语义分割方法。同时, 为了便于观察, 本文截取了所有网络 100 轮的训练和验证损失, 如图 4(彩图见期刊电子版)所示, MVD 验证集上的可视化结果如图 5(彩图见期刊电子版)所示, 从训练和验证损失曲线可以看出, U-MVDNet 相比于其他网络, 损失衰减的更快更平稳。

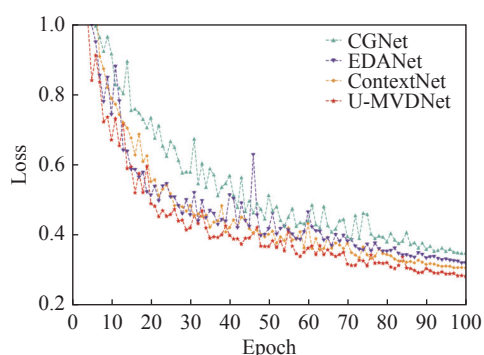
EDANet<sup>[29]</sup>、ContextNet<sup>[30]</sup> 没有准确定位“cn10”的分割边界, 存在明显的目标轮廓分割缺失。在第三列中, 只有 U-MVDNet 准确分割出了“aica”。第四列和第五列中, U-MVDNet 对“pica”的分割



优于其他网络, 分割更接近真值。综上所述, 本文提出的方法分割性能更高, 包含的特征信息更多。



(a) 训练损失曲线  
(a) Training loss curve



(b) 验证损失曲线  
(b) Validating loss curve

图 4 损失曲线图

Fig. 4 Loss curves

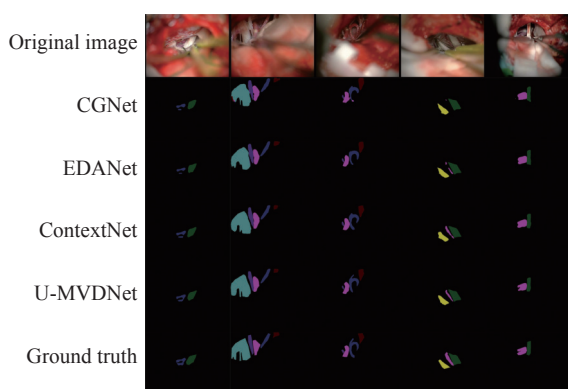


图 5 MVD 验证集上的可视化对比结果

Fig. 5 The visual comparison results of different methods on MVD validate set

### 3.4.2 U-MVDNet 先进性分析

本文使用 ISIC 2016 挑战赛的数据集和 PH2 数据集来进一步评估 U-MVDNet 的先进性。实验中, 将 ISIC 2016 数据集设置为训练集, PH2 数据集设置为测试集, 并在与 U-MVDNet 相同的条件下进行实验。同时, 使用 Jaccard 系数 (JAC)、

Dice 系数 (DIC)、准确性 (ACC)、特异性 (SPE)、灵敏度 (SEN) 来评估 U-MVDNet 对皮肤病灶的分割性能。

在测试集上, U-MVDNet 分别与 DeepLabv3<sup>[31]</sup> 和 CA-Net<sup>[32]</sup> 进行了比较, 结果如表 10 所示。U-MVDNet 的参数量仅为 DeepLabv3 和 CA-Net 参数量的 1/4 和 1/88, 但除了灵敏度之外, 其他的评价指标均达到了最优, 且具有最快的推理速度。图 6 (彩图见期刊电子版) 展示了皮肤病灶分割结果的典型案例。通过观察可以发现, U-MVDNet 的结果更接近于实际情况, 能够突出显示一些边界不明确的小区域, 具有稳定的分割性能。

表 10 ISIC 2016 + PH2 测试集实验结果

Tab. 10 Experimental results on ISIC 2016 + PH2 test set

Model	Params (M)	Speed (frame·s <sup>-1</sup> )	DIC (%)	JAC (%)	ACC (%)	SPE (%)	SEN (%)
DeepLabv3 <sup>[31]</sup>	58.2	98.7	88.6	81.2	91.9	89.1	95.9
CA-Net <sup>[32]</sup>	2.79	130.3	88.7	80.5	93.2	91.3	96.9
U-MVDNet	0.66	175.1	89.3	81.7	93.2	93.3	94.3

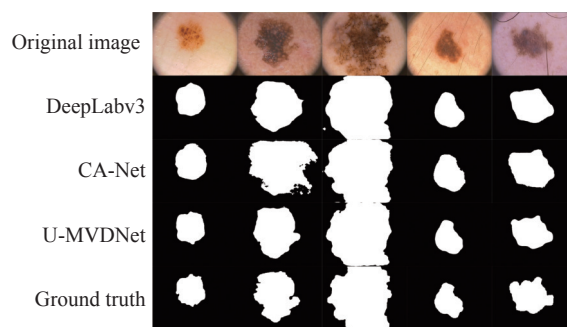


图 6 ISIC 2016 + PH2 测试集上的可视化对比

Fig. 6 The visual comparison results of different methods on ISIC 2016 + PH2 test set

### 3.5 嵌入式终端测试

本节将训练后的模型移植到嵌入式平台 NVIDIA Jetson AGX Xavier 中进行测试。平台的详细硬件参数如表 11 所示。在嵌入式终端中将本文方法与 CGNet<sup>[28]</sup>、EDANet<sup>[29]</sup>、ContextNet<sup>[30]</sup> 在不同分辨率的情况下进行对比, 并显示出各网络最终的分割精度。

使用 MVD 测试集在嵌入式终端上验证本文所提出 U-MVDNet 性能。从表 12 可以看出, 当分辨率为 640×480 时, 本文所提方法在 NVIDIA

Jeston AGX Xavier 上实现了 24.2 frame/s 的处理速度, 达到了实时处理速度, 并在分割精度上优于其他分割方法。

表 11 两种不同的硬件环境

Tab. 11 Two different hardware environments

	Jetson Xavier	服务器
GPU	Volta	GTX 2080Ti
CPU	8核Carmel ARM	8核i7-9700K
显存	32GB LPDDR4x	11GB GDDR6
显存带宽	136.5 GB/s	616 GB/s
CUDA核心	512	4352

表 12 不同分辨率下的测试结果

Tab. 12 Test results by different methods with different resolutions

Method	Size	Times/ms	Speed/frame·s <sup>-1</sup>	mIoU/%
CGNet <sup>[28]</sup>	640×480	65.7	15.2	70.31
	768×576	69.2	14.4	71.95
EDANet <sup>[29]</sup>	640×480	42.3	23.6	73.2
	768×576	45.2	22.1	74.18
ContextNet <sup>[30]</sup>	640×480	34.5	28.9	74.81
	768×576	36.1	27.7	75.81
U-MVDNet	640×480	41.5	24.2	75.76
	768×576	43.6	22.9	76.29

## 4 讨 论

以往的研究中, 很少对 MVD 中的血管和神经进行实时准确分割。本研究采用了新型的编码解码结构, 该结构具有分割速度快和分割精度高的特点, 可准确快速地完成对 MVD 中的血管和神经的分割。与以往的实时语义分割研究相比,

### 参考文献:

- [1] BENNETTO L, PATEL N K, FULLER G. Trigeminal neuralgia and its management[J]. *BMJ*, 2007, 334(7586): 201-205.
- [2] KIZILTAN M E, GUNDUZ A. Reorganization of sensory input at brainstem in hemifacial spasm and postparalytic facial syndrome[J]. *Neurological Sciences*, 2018, 39(2): 313-319.
- [3] NAZIR A, CHEEMA M N, SHENG B, et al.. OFF-eNET: an optimally fused fully end-to-end network for automatic dense volumetric 3D intracranial blood vessels segmentation[J]. *IEEE Transactions on Image Processing*, 2020, 29: 7192-7202.
- [4] 吴红宇, 郑波. 脑动静脉畸形CTA、DSA的影像学表现及诊断的对照性研究[J]. *中国 CT 和 MRI 杂志*, 2021,

编码器的结构更加简洁, 具有更少的卷积神经网络层数, 但能够获得更多的上下文信息, 解码器更有效地融合了高级和低级特征。

MVD 中, 术者通常是根据对照组和经验来判断血管和神经, 但是术中骨窗直径仅为 2.5 cm, 手术空间小, 在术中会有少量的出血, 以及随时出现的脑脊液, 影响术者操作, 为术者带来诸多不便, 导致术中可能出现许多不确定性因素, 甚至是风险的发生。本文提出的方法实现了脑血管和脑神经的快速准确分割, 减轻术者精神压力, 为术者快速决策和判断提供了依据, 有利于术中减少对脑脊液的释放, 避免过度牵拉神经及血管, 有效地减少了手术创伤, 减少了术后并发症的发生。

## 5 结 论

针对 MVD 图像, 为提高实时语义分割的速度和准确性, 本文提出了 U 型轻量级快速语义分割网络 U-MVDNet, 该网络由编码器和解码器构成。具体来说, 设计了一种新颖的轻型非对称瓶颈模块 (LABM) 来提取上下文信息, 基于该模块设计得到了编码器。解码器采用了特征融合模块 (FFM) 来聚合不同的特征。消融实验表明, LAB block 有效提取了上下文特征, FFM 有效地融合了深层上下文特征和浅层空间特征。选取 MVD 测试集进行测试和对比, 结果显示: U-MVDNet 在单 NVIDIA GTX 2080Ti 上以 140 frame/s 的速度获得了 76.29% 的 mIoU, 相较于其他对比网络, U-MVDNet 分割精度最高、参数量最少, 并且在嵌入式平台上能够实现实时 (24 frame/s) 语义分割, 分割精度也优于其他分割方法。本文所提出方法在一定程度上解决了微血管减压图像实时语义分割网络的参数量大、分割精度低的问题。

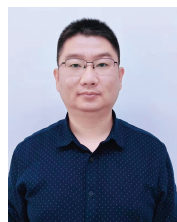
- 19(1): 36-37,52.
- WU H Y, ZHENG B. Analysis on imaging manifestations and diagnostic contrast of cerebral arteriovenous malformations in CTA and DSA[J]. *Chinese Journal of CT and MRI*, 2021, 19(1): 36-37,52. (in Chinese)
- [5] PATEL T R, PALIWAL N, JAISWAL P, *et al.*. Multi-resolution CNN for brain vessel segmentation from cerebrovascular images of intracranial aneurysm: a comparison of U-Net and DeepMedic[J]. *Proceedings of SPIE*, 2020, 11314: 113142W.
- [6] 王华. 磁共振血管成像与三维动脉自旋标记脑灌注成像技术诊断缺血性脑血管疾病一致性比较[J]. *实用医院临床杂志*, 2020, 17(1): 36-39.
- WANG H. Comparison of consistency of MRA and 3D-ASL cerebral perfusion imaging in the diagnosis of ischemic cerebrovascular diseases[J]. *Practical Journal of Clinical Medicine*, 2020, 17(1): 36-39. (in Chinese)
- [7] 徐冰洁. 电子计算机断层扫描联合核磁共振血管成像对脑血管疾病的诊断价值[J]. *临床合理用药杂志*, 2021, 14(32): 177-178.
- XU B J. Diagnostic value of computed tomography combined with nuclear magnetic resonance angiography in cerebrovascular diseases[J]. *Chinese Journal of Clinical Rational Drug Use*, 2021, 14(32): 177-178. (in Chinese)
- [8] ZHANG H, XIA L K, SONG R, *et al.*. Cerebrovascular segmentation in mra via reverse edge attention network[C]. *Proceedings of the 23rd International Conference on Medical Image Computing and Computer-Assisted Intervention*, Springer, 2020: 66-75.
- [9] GUO X Y, XIAO R X, LU Y Y, *et al.*. Cerebrovascular segmentation from TOF-MRA based on multiple-U-net with focal loss function[J]. *Computer Methods and Programs in Biomedicine*, 2021, 202: 105998.
- [10] HILBERT A, MADAI V I, AKAY E M, *et al.*. BRAVE-NET: fully automated arterial brain vessel segmentation in patients with cerebrovascular disease[J]. *Frontiers in Artificial Intelligence*, 2020, 3: 552258.
- [11] 陈星, 宋智洋, 周明全, 等. 面向脑血管分割的改进型非局部均值滤波算法研究[J]. *中国光学*, 2014, 7(4): 572-580.
- CHEN X, SONG ZH Y, ZHOU M Q, *et al.*. An improved non-local mean filter algorithm facing the cerebrovascular segmentation[J]. *Chinese Optics*, 2014, 7(4): 572-580. (in Chinese)
- [12] WANG R, LI CH, WANG J, *et al.*. Threshold segmentation algorithm for automatic extraction of cerebral vessels from brain magnetic resonance angiography images[J]. *Journal of Neuroscience Methods*, 2015, 241: 30-36.
- [13] BHUIYAN A, NATH B, CHUA J. An adaptive region growing segmentation for blood vessel detection from retinal images[C]. *Visapp: Second International Conference on Computer Vision Theory and Applications*, 2007: 404-409.
- [14] 王醒策, 张美霞, 武仲科, 等. 基于全局LBF水平集模型的脑血管层次粗分割[J]. *光学精密工程*, 2013, 21(12): 3283-3297.
- WANG X C, ZHANG M X, WU ZH K, *et al.*. Level coarse brain vessel segmentation based on global LBF model[J]. *Optics and Precision Engineering*, 2013, 21(12): 3283-3297. (in Chinese)
- [15] WANG J X, ZHAO SH F, LIU Z F, *et al.*. An active contour model based on adaptive threshold for extraction of cerebral vascular structures[J]. *Computational and Mathematical Methods in Medicine*, 2016, 2016: 6472397.
- [16] 陈晓冬, 艾大航, 张佳琛, 等. Gabor滤波融合卷积神经网络的路面裂缝检测方法[J]. *中国光学*, 2020, 13(6): 1293-1301.
- CHEN X D, AI D H, ZHANG J CH, *et al.*. Gabor filter fusion network for pavement crack detection[J]. *Chinese Optics*, 2020, 13(6): 1293-1301. (in Chinese)
- [17] 王春哲, 安军社, 姜秀杰, 等. 基于卷积神经网络的候选区域优化算法[J]. *中国光学*, 2019, 12(6): 1348-1361.
- WANG CH ZH, AN J SH, JIANG X J, *et al.*. Region proposal optimization algorithm based on convolutional neural networks[J]. *Chinese Optics*, 2019, 12(6): 1348-1361. (in Chinese)
- [18] LONG J, SHELHAMER E, DARRELL T. Fully convolutional networks for semantic segmentation[C]. *Proceedings of 2015 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, 2015: 3431-3440.
- [19] SHVETS A A, IGLOVIKOV V I, RAKHLIN A, *et al.*. Angiodysplasia detection and localization using deep convolutional neural networks[C]. *Proceedings of 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, IEEE, 2018: 612-617.
- [20] ZHANG M, ZHANG CH, WU X, *et al.*. A neural network approach to segment brain blood vessels in digital subtraction angiography[J]. *Computer Methods Programs in Biomedicine*, 2020, 185: 105159.
- [21] XIA L K, XIE Y X, WANG Q W, *et al.*. A nested parallel multiscale convolution for cerebrovascular segmentation[J].

- Medical Physics*, 2021, 48(12): 7971-7983.
- [22] MENG C, SUN K, GUAN SH Y, *et al.*. Multiscale dense convolutional neural network for DSA cerebrovascular segmentation[J]. *Neurocomputing*, 2020, 373: 123-134.
- [23] HUANG G, LIU ZH, VAN DER MAATEN L, *et al.*. Densely connected convolutional networks[C]. *Proceedings of 2017 IEEE Conference on Computer Vision and Pattern Recognition*, IEEE, 2017: 2261-2269.
- [24] ALVAREZ J, PETERSSON L. DecomposeMe: simplifying convnets for end-to-end learning[J]. *arXiv preprint arXiv*., 1606, 05426: 2016.
- [25] HE K M, ZHANG X Y, REN SH Q, *et al.*. Delving deep into rectifiers: Surpassing human-level performance on imagenet classification[C]. *Proceedings of 2015 IEEE International Conference on Computer Vision*, IEEE, 2015: 1026-1034.
- [26] CODELLA N C F, GUTMAN D, CELEBI M E, *et al.*. Skin lesion analysis toward melanoma detection: a challenge at the 2017 international symposium on biomedical imaging (ISBI), hosted by the international skin imaging collaboration (ISIC)[C]. *Proceedings of the 15th International Symposium on Biomedical Imaging*, IEEE, 2016.
- [27] MENDONÇA T, FERREIRA P M, MARQUES J S, *et al.*. PH<sup>2</sup>-A dermoscopic image database for research and benchmarking[C]. *Proceedings of the 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, 2013: 5437-5440.
- [28] WU T Y, TANG SH, ZHANG R, *et al.*. CGNet: a light-weight context guided network for semantic segmentation[J]. *IEEE Transactions on Image Processing*, 2020, 30: 1169-1179.
- [29] LO S Y, HANG H M, CHAN SH W, *et al.*. Efficient dense modules of asymmetric convolution for real-time semantic segmentation[C]. *Proceedings of the ACM Multimedia Asia*, ACM, 2019: 1.
- [30] POUDEL R P K, BONDE U, LIWICKI S, *et al.*. ContextNet: exploring context and detail for semantic segmentation in real-time[J]. *arXiv preprint arXiv*., 0455, 4: 2018.
- [31] CHEN L C, PAPANDREOU G, SCHROFF F, *et al.*. Rethinking atrous convolution for semantic image segmentation[J]. *arXiv preprint arXiv*., 1706, 05587: 2017.
- [32] GU R, WANG G T, SONG T, *et al.*. CA-Net: comprehensive attention convolutional neural networks for explainable medical image segmentation[J]. *IEEE Transactions on Medical Imaging*, 2021, 40(2): 699-711.

#### 作者简介:



白瑞峰(1994—),男,甘肃通渭人,博士研究生,2017年于兰州交通大学获得学士学位,主要从事智能医学图像处理方面的研究。E-mail: bairuifeng\_ucas@126.com



江山(1986—),男,吉林长春人,副研究员,硕士生导师,2010年、2013年于吉林大学分别获得学士、硕士学位,主要从事深度学习、高速目标跟踪处理方面的研究。E-mail: 617798169@qq.com