

DOI: 10.19650/j.cnki.cjsi.J2513788

# 空间频域融合的视网膜血管分割方法\*

宋伟伟<sup>1</sup>, 许 茗<sup>1</sup>, 于晓升<sup>2</sup>, 李海星<sup>1</sup>, 王宏宇<sup>3</sup>

(1. 沈阳工业大学人工智能学院 沈阳 110870; 2. 东北大学机器人科学与工程学院 沈阳 110167;  
3. 北京大学未来技术学院 北京 100091)

**摘 要:** 视网膜血管的精确分割对于诊断多种眼部疾病至关重要, 它不仅有助于识别糖尿病、动脉硬化和心血管疾病等医学问题, 还能显著提升医生对患者疾病的诊断和治疗能力。现有的卷积神经网络方法虽然在空间域内通过卷积操作捕捉局部特征能力表现出色, 但在获取全局空间特征信息方面存在局限性。而频域方法虽能捕获图像的整体频谱分布和全局结构特征, 却因频域变换过程中的空间信息模糊化处理, 难以精确定位局部特征并保留高频细节信息。提出了一种空间频域融合的视网膜血管分割方法, 该方法结合了空间域和频域方法在获取局部和全局特征信息方面的优势。首先, 设计双支路的空间频域特征提取与融合模块, 在编码阶段融合频域和空间域的特征信息, 旨在减少下采样过程中丢失的细节特征。此外, 引入多尺度高通滤波器, 以提高模型定位血管边界和保持小血管连贯性的能力。最后, 通过空间频域自适应融合模块动态计算特征图各区域的融合权重, 提升小血管分割的准确性。在 DRIVE 和 CHASE\_DB1 这两个主流开源数据集上进行了性能测试, 其准确率分别为 96.9% 和 97.81%。实验结果表明该方法在血管分割的准确性、小血管的连贯性和应对病变的鲁棒性方面均展现出了竞争优势。

**关键词:** 视网膜血管分割; 空间频域自适应融合; 多尺度高通滤波器

**中图分类号:** TP391 TH79 **文献标识码:** A **国家标准学科分类代码:** 510.40

## Retinal vessel segmentation method based on fusion of frequency domain and spatial domain

Song Weiwei<sup>1</sup>, Xu Ming<sup>1</sup>, Yu Xiaosheng<sup>2</sup>, Li Haixing<sup>1</sup>, Wang Hongyu<sup>3</sup>

(1. School of Artificial Intelligence, Shenyang University of Technology, Shenyang 110870, China;  
2. Faculty of Robot Science and Engineering, Northeastern University, Shenyang 110167, China;  
3. Future Technology College, Peking University, Beijing 100091, China)

**Abstract:** Accurate segmentation of retinal blood vessels plays a vital role in diagnosing various eye diseases. It not only aids in identifying conditions such as diabetes, arteriosclerosis, and cardiovascular diseases but also enhances clinicians' ability to diagnose and treat patients effectively. While existing convolutional neural network (CNN) approaches excel at capturing local spatial features through convolutional operations, they face challenges in extracting global spatial information. Conversely, frequency domain methods can capture the overall spectral distribution and global structural features of images but struggle to precisely locate local details and preserve high-frequency information due to spatial information blurring during frequency transformation. To address these limitations, this study proposes a retinal blood vessel segmentation method based on spatial-frequency domain fusion, leveraging the strengths of both domains for local and global feature extraction. The approach features a dual-branch spatial-frequency feature extraction and fusion module in the encoding stage, designed to integrate frequency and spatial features and mitigate detail loss during downsampling. Additionally, a multi-scale Gaussian filter is incorporated to enhance the model's capability in accurately locating vessel boundaries and preserving continuity of small vessels. Finally, an adaptive spatial-frequency fusion module dynamically calculates fusion weights across feature map regions,

收稿日期: 2025-02-28 Received Date: 2025-02-28

\* 基金项目: 国家自然科学基金项目 (62206183) 资助

improving the precision of small vessel segmentation. Experiments conducted on two widely-used open-source datasets, DRIVE and CHASE\_DB1, demonstrate accuracy rates of 96.9% and 97.81%, respectively. Results indicate that the proposed method achieves competitive performance in segmentation accuracy, consistency of small vessel detection, and robustness in handling lesions.

**Keywords:** retinal vessel segmentation; frequency-space adaptive fusion; multi-scale high-pass filter

## 0 引言

视网膜作为眼睛的关键组成部分,其健康状况对于视觉功能至关重要。根据全球健康组织的调查数据,眼部疾病的发病率呈现上升趋势。在众多的眼部检查项目中,视网膜血管形态学参数,如血管的分支模式、角度、曲率、宽度和长度等,对眼部疾病的诊断和评估具有重要意义<sup>[1]</sup>。然而,医生手工分割视网膜血管往往耗时耗力。因此,开发一种能够快速且准确的自动分割视网膜血管方法,对于提升临床诊断效率具有重要的研究价值和实际应用意义。目前,视网膜血管分割领域仍然面临一系列挑战:首先,眼底图像中存在着视杯视盘等多种眼底组织与血管相重叠,增加了血管自动分割的难度<sup>[2]</sup>。其次,由于卷积神经网络其固有的局部卷积操作,随着训练数据的增加和网络层数的加深可能忽视全局结构,在血管的连接和分支处容易丢失细节,导致血管结构不连续和分割错误<sup>[3]</sup>。最后,受到各种病灶等噪声的干扰也会对整个分割性能带来一定的影响<sup>[4]</sup>。

为了应对上述挑战,设计了一种频域和空间域融合的网络用于视网膜血管分割。该网络由编码器和解码器两部分组成。为了应对血管细节丢失问题,编码器部分设计为频域支路与空间域支路,以双支路提取特征的形式强化特征提取。为了应对视网膜图像种眼底组织和病灶噪声干扰的问题,设计了空间域频域交互模块融合空间域信息与频域信息。频域和空间域是互补性的,在频域中使用高斯高通滤波减少低频干扰信息强化血管高频信息,利用空间域补充局部血管细节信息提高分割性能。创新点为:

1) 改进了 U-Net 结构,设计了一种新的双分支结构和空间频域自适应融合模块,可以充分挖掘和利用空间域和频域所包含的互补信息,将全局频域信息和空间域的局部特征信息自适应融合,以满足多层次血管细节特征提取的要求。

2) 设计了多尺度高斯高通滤波器,以增强眼底视网膜血管的纹理和边缘特征,有助于网络提取血管特征。

## 1 相关工作

### 1.1 基于眼底图像的血管分割方法

目前,国内外的研究者已提出了多种视网膜血管分

割方法。传统方法涵盖了线检测、多尺度滤波、粒子群优化算法、概率图模型和小波变换等。然而,这些传统方法对图像质量的要求较高,且受视网膜结构的影响较大,导致分割效果不尽如预期。随着深度学习技术的不断发展,越来越多的深度学习方法被用于视网膜血管分割任务,并取得了出乎意料成果。目前,视网膜血管分割主要分为两种类型:基于 U-Net 网络的方法和非基于 U-Net 网络的方法。在非基于 U-Net 网络的方法中,Shang 等<sup>[5]</sup>设计了一种叫做 DCNet 的新型 3 层扩张卷积网络,从不同级别的特征图中提取关键的特征用于分割。Li 等<sup>[6]</sup>提出全局 transformer 和双局部注意力网络,分别用于捕获像素间的长距离依赖关系和提取血管局部信息并巩固血管边缘细节。Du 等<sup>[7]</sup>巧妙地利用双阈值的方法来增强血管的连续性。Wu 等<sup>[8]</sup>提出了一种新的基于分流变压器的多编码器多解码器网络,逐层学习丰富的局部上下文特征信息。Tang 等<sup>[9]</sup>设计了一种可以提取多分辨率视网膜图像特征的双分支结构,将全局和局部特征融为一体。这些方法虽然增强了血管的连续性,但视网膜中的血管的分布非常复杂,血管和背景组织在颜色和亮度上非常相似,血管边缘的准确分割仍然非常困难。

随着 U-Net 在医学分割领域取得的突出成绩<sup>[10]</sup>,许多学者正在研究基于 U-Net 网络的视网膜血管分割模型。一种方法是加强血管特征的提取,通常使用扩大感受野和细化特征的方法。Ma 等<sup>[11]</sup>改进了 U-Net,在编码部分引入编码器融合模块来提取不同卷积的特征,在解码器部分使用有监督融合机制来捕获不同层的特征信息。Alom 等<sup>[12]</sup>将循环残差方法引入 U-Net 网络,利用加大网络深度和累积特征的方式加强特征提取。Ye 等<sup>[13]</sup>将多尺度特征和注意力机制融合,利用在多个尺度上学习到的通道注意力来增强多尺度特征。Liu 等<sup>[14]</sup>提出了一种细节增强模块来代替原始 U-Net 的简单跳跃连接,用于获取更多关于细微结构和细节的上下文信息。夏平等<sup>[15]</sup>在跳跃连接部分采用多尺度特征聚合模块。另一种方法是最大限度地保留血管信息,以实现血管特征的利用。Liu 等<sup>[16]</sup>使用残差网络作为主干网络加强获取上下文特征。苑玮琦等<sup>[17]</sup>提出了一种使用多尺度高斯滤波提取血管骨架的方法。Xie 等<sup>[18]</sup>提出了一种结构自适应层和残差网络,通过融合不同层特征提取血管细节特征。上述方法在实现血管分割方面取得了一定成效,但对于视网膜图像来说,由于其中存在大量粗细不同和直

径变化很大的血管,导致现有方法在对血管和周围组织的区分上仍面临困难,存在进一步提升空间。

## 1.2 基于傅里叶变换和频域的方法

傅里叶变换在图像处理领域提供了一种深刻且实用的方法论。通过傅里叶变换,图像中的高频信息得以突出,这些信息包含了图像的细节和纹理。利用傅里叶变换的特点,可以高效地进行多种图像处理任务。Rao等<sup>[19]</sup>提出了一种全局滤波网络,该网络能够在频域中学习长距离的空间依赖关系。此外,Fuoli等<sup>[20]</sup>为了提升图像的感知质量和恢复高频细节,设计了一种基于傅里叶变换的损失函数。Yu等<sup>[21]</sup>融合了频域和空间域的信息,频域和空间域相互引导。Huang等<sup>[22]</sup>设计了空间域和频域交互模块,使局部空间特征和全局频域信息互补学习。Wang等<sup>[23]</sup>提出了一种空间域和频域动态图的方法学习空间域和频域之间的关系。Xiao等<sup>[24]</sup>提出了一种新颖的频率感知的红外与可见光图像融合网络,使用频率感知模块在低频和高频信息中挖掘有意义的特征。更进一步,为了强调图像中的高频细节信息,使用滤波器可以在傅里叶变换中对低频成分进行抑制或减少,这样可以减少图像中的平滑区域或背景噪声。同时,滤波器能够保留高频信息,这些高频成分通常代表了图像中的细节、边缘和纹理<sup>[25]</sup>。

## 2 方法

### 2.1 总体网络体系结构

提出了一种新的基于 U-Net 的频域空间域交互的双分支结构网络 (frequency-spatial dual-domain network, FreNet)。其整体结构如图 1 所示, FreNet 由左侧的双编码器和右侧的解码器组成。对于输入图像  $x \in R^{H \times W \times C}$ ,  $H$  和  $W$  分别为输入图像的高度和宽度,  $C$  为通道数。当给定输入图像时,首先将其分别输入到编码器的两个子网络中,通过下采样来提取特征。频域分支利用多尺度高斯滤波器 (multi-scale gaussian high-pass filtering, MGHF) 捕获全局频域信息,而空间域分支则通过卷积操作提取局部细节。这两个分支的输出在空间频域自适应融合模块 (self-adaptive spatial frequency fusion module, SSF) 中合并,以整合多尺度特征,并传递给下一层的空间域分支。在解码器阶段,通过上采样和跳跃连接,将不同尺度的特征图恢复成对应的尺度,进一步细化血管信息的提取。

### 2.2 傅里叶变换和高斯高通滤波

在自然图像中,低频分量通常包含丰富的全局信息,而高频分量则携带着精细的纹理细节。傅里叶变换作为一种广泛应用于图像处理领域的技术,是获取频域信息

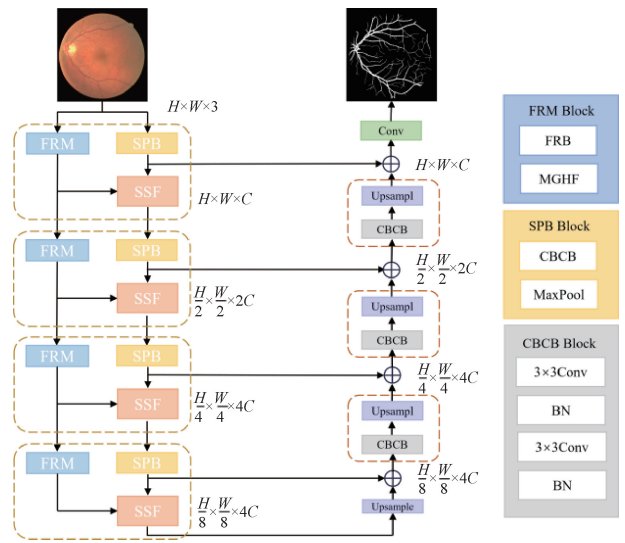


图1 FreNet网络结构

Fig. 1 Schematic diagram of the FreNet neural network architecture

的关键手段。给定单通道图像  $x$ , 图像  $x$  的傅里叶变换可表示为:

$$T_x(u, v) = \frac{1}{\sqrt{HW}} \sum_{h=0}^{H-1} \sum_{w=0}^{W-1} f(h, w) e^{-2j\pi \left( \frac{h}{H}u + \frac{w}{W}v \right)} \quad (1)$$

其中,  $f(h, w)$  是图像在空间域的表示, 是该图像在频域的表示。式(1)描述了如何将一个二维图像从空间域转换到频域。式(1)中的变量  $u$  代表频域中的水平分量,  $v$  代表垂直分量。  $h$  和  $w$  则代表空间域中的坐标。指数项  $\exp(-2j\pi(hu/H + wv/W))$  是复数指数函数, 它根据图像的尺寸  $H$  和  $W$  以及当前的空间域坐标  $(h, w)$  来计算频域中对应点的值。在频域中, 振幅分量  $A_x$  和相位分量  $P_x$  是  $x$  的两个重要分量。其表示方式如式(2)~(3)所示。

$$A_x(u, v) = \sqrt{R^2(x)(u, v) + I^2(x)(u, v)} \quad (2)$$

$$P_x(x)(u, v) = \arctan \left( \frac{I(x)(u, v)}{R(x)(u, v)} \right) \quad (3)$$

其中,  $R(x)$  和  $I(x)$  分别对应于  $T_x$  在频域中的实部和虚部。

傅里叶变换能够有效地通过相位分量捕捉图像的结构和细节, 振幅分量能够有效地描绘血管纹理。二者可以能够精确地提取视网膜血管信息, 将图像转换到频域可以更有效地处理图像中的高频信息。高频信息主要对应于图像的边缘、噪声和更细致的内容, 这对于提取血管边缘特征尤为重要。在频谱中, 通过抑制低频成分可以实现图像的锐化, 采用高斯高通滤波器在频域中提取图像特征, 通过减少低频信息来增强血管边缘的可见性。高斯高通滤波可表示为:



$$H(u, v) = 1 - e^{-\frac{D^2(u, v)}{2\sigma_0^2}} \quad (4)$$

其中,  $D^2(u, v)$  是距离频率矩形中心的距离;  $\sigma_0$  是关于中心的扩展度的度量, 越大的  $\sigma_0$  对应着越粗的血管, 越小的  $\sigma_0$  对应着血管越细。

### 2.3 空间域频域交互模块

FreNet 以 U-Net 网络为主体, 在编码器部分设计了双分支结构进行特征提取, 为简化描述称为双支路的空间频域提取与融合模块 (spatial-frequency feature analysis and fusion, SFAF), 如图 2 所示。  $F_{Fre}^n$  和  $F_{Spa}^n$  分别对应于第  $n$  个频域和空间域分支生成的结果, 如式 (5) 所示。

$$F_{Spa}^n, F_{Fre}^n = f_{SFAF}^n(F_{Spa}^{n-1}, F_{Fre}^{n-1}) \quad (5)$$

其中,  $f_{SFAF}^n$  是第  $n$  个 SFAF 的函数。

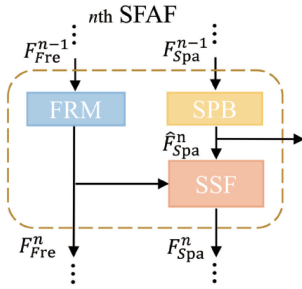


图2 SFAF 结构

Fig. 2 Structural diagram of SFAF module

SFAF 的作用是将通过卷积操作所获得的空间域特征和傅里叶变换得到的频率特征相结合, 输出相应的频率特征和融合后的空间域特征。在频域特征支路中通过多次提取频率信息, 逐步增强血管的整体走向和边缘特征。在空间域特征中经过多次特征提取后与频率信息融合, 有效地提取血管的连续性和小血管的细节。这种融合方法有助于更准确地识别和分析血管结构, 如式 (6) ~ (7) 所示。

$$F_{Fre}^n = f_{FRM}^n(F_{Fre}^{n-1}) \quad (6)$$

$$F_{Spa}^n = f_{SSF}^n(f_{SPB}^{n-1}(F_{Spa}^{n-1})) \quad (7)$$

其中, 在第  $n$  个 SFAF 处, 第  $n-1$  个 SFAF 产生的  $F_{Spa}^{n-1}$  和  $F_{Fre}^{n-1}$  分别为空间分支和频域分支。  $f_{FRM}^n$  和  $f_{SPB}^n$  分别对应频域分支和空间分支中的频域模块 (frequency module, FRM)、空间域模块 (spatial block SPB) 和空间频域自适应融合模块,  $F_{Fre}^{n-1}$  为在  $n-1$  个 SFAF 处提取的频域特征, 同理  $F_{Spa}^{n-1}$  为在  $n-1$  个 SFAF 处提取的空间域特征,  $F_{Spa}^n$  和  $F_{Fre}^n$  分别为第  $n$  次时提取的空间域信息和频域信息,  $\hat{F}_{Spa}^n$  为经过 SPB 提取的空间域特征。SPB 由两个  $3 \times 3$  卷积和  $2 \times 2$  的池化层组成。

FRM 主要包含频率提取模块 (frequency retrieval block, FRB) 和多尺度高斯滤波器, FRB 将输入通过傅里

叶变换分解为幅值分量  $A$  和相位分量  $P$ , 幅值  $A$  和相位  $P$  分别经过卷积后再分别与对应的输入相融合生成幅值  $\bar{A}$  和相位  $\bar{P}$ , 最后融合为复数形式输出, 如图 3 所示。

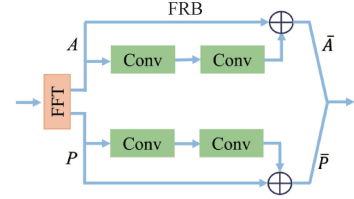


图3 FRB 结构

Fig. 3 Structural diagram of FRB module

叶变换分解为幅值分量  $A$  和相位分量  $P$ , 幅值  $A$  和相位  $P$  分别经过卷积后再分别与对应的输入相融合生成幅值  $\bar{A}$  和相位  $\bar{P}$ , 最后融合为复数形式输出, 如图 3 所示。

在频域分支中, 设计了多尺度高斯滤波器, 通过应用不同尺寸的高斯核对图像进行滤波和增强高频信息, 以突出血管边缘的细节。这些高斯核根据不同大小的输入图像, 动态的增强大小血管边缘。将不同滤波处理后图像进行拼接操作以加强血管特征。这种策略的目的是为了整合不同尺度的图像成分, 从而实现对血管特征的全面融合, 进而获得一个更为全面和详尽的血管表示。MGHF 如图 4 所示。  $F_{Fre}^{n-1}$  在经过 FRB 模块后, 由 MGHF 模块提取不同的高频信息后经过卷积操作恢复通道数, 最后使用傅里叶逆变换生成输出  $F_{Fre}^n$ 。

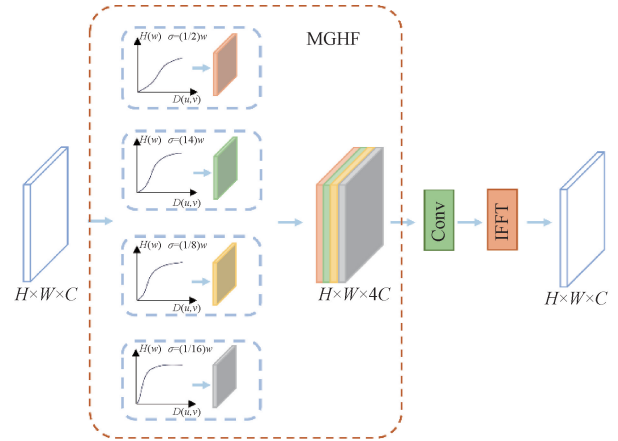


图4 MGHF 结构

Fig. 4 MGHF structure diagram

频域分支能够捕获全局感受野, 有助于关注长距离的依赖关系, 而空间分支则专注于提取局部纹理和边界等细节信息。全局依赖和局部依赖的互补性对特征提取至关重要, 它们共同促进了更准确的图像分割。为此, 提出了一种基于空间频域自适应融合模块的方法, 通过增强全局频率特征与局部空间特征的融合, 平衡局部和全局之间的依赖关系。该策略充分利用广泛覆盖的全局特征, 同时保留局部的细节特征, 从而在血管分割中实现更准确的结果, SSF 结构如图 5 所示。



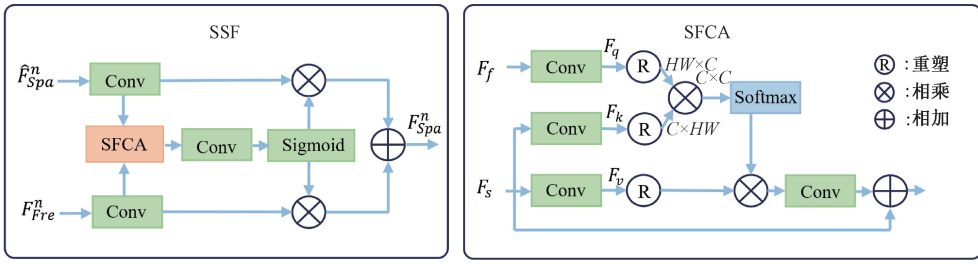


图5 SSF 结构

Fig. 5 SSF module structure diagram

该研究使用了一种基于自注意力机制的空间频域注意力模块 (spatial-frequency cross-attention module, SFCA), 利用全局频域信息和局部空间信息的互补性。该注意力模块由  $F_f$  和  $F_s$  为输入, 使用  $F_f$  通过卷积层后生成对应的查询矩阵  $F_q$ , 使用  $F_s$  分别通过不同的卷积层来获得键矩阵  $F_k$  和值矩阵  $F_v$ 。如式(8)~(9)所示:

$$\text{Attention}(F_k, F_q, F_v) = \text{Softmax}\left(\frac{F_q F_k^T}{\sqrt{d}}\right) V \quad (8)$$

$$F_{\text{Fuse}} = \text{Conv}(\text{Attention}(F_k, F_q, F_v)) + F_s \quad (9)$$

其中,  $d$  为超参数,  $F_{\text{Fuse}}$  为结果。在 SSF 中, 沿通道维度进行乘法操作可以实现对不同通道之间的关联性进行建模, 使得模型可以更好地理解不同通道之间的重要性, 提升特征表示的质量。此种方法的优点是能够在不引入过多参数和计算复杂度的情况下, 有效地捕获通道维度间的全局信息。通过在通道维度上进行乘法操作, 可以实现通道之间的交互和整合, 改善特征表示的质量, 提升模型的性能和泛化能力。

### 3 实验

#### 3.1 数据集与评价指标

该研究使用两个公开的数据集: DRIVE<sup>[26]</sup> 和 CHASE\_DB1<sup>[27]</sup> 来进行视网膜血管分割实验, 以评估该方法的性能。

DRIVE 数据库中有 40 幅彩色视网膜血管图像, 这些照片来自荷兰的一项糖尿病视网膜病变筛查计划, 数据集被分成训练集和测试集两组, 每组 20 张。CHASE\_DB1 数据集包含 28 张图像, 该数据集没有区分训练集和测试集, 该研究设定前 20 张图像训练集, 其余 8 张图像为测试集。

为了量化所提出模型的分割性能, 将分割结果与相应的专家标签进行比较, 并将每个像素的比较结果分为血管正确分类的真阳性 (true positive, TP), 非血管被归类为血管的假阳性 (false positive, FP), 血管被归类为非血管的假阴性 (false negative, FN) 和非血管归类为非血

管的真阴性 (true negative, TN)。使用准确性 (accuracy, ACC), 敏感性 (sensitivity, SE), 特异性 (specificity, SP) 和 F1 值作为测量指标。

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \quad (10)$$

$$SE = \frac{TP}{TP + FN} \quad (11)$$

$$SP = \frac{TN}{TN + FP} \quad (12)$$

$$F1 = \frac{2TP}{2TP + FP + FN} \quad (13)$$

ACC 为正确分类所占的比例, SE 为阳性样本中正确分类的比例, SP 为阴性样本中正确分类所占的比例, F1 是准确率与敏感性的加权平均, 取值越大表明模型性能越好。

#### 3.2 损失函数

视网膜血管分割任务是从眼底视网膜图像中准确地提取血管结构, 将血管结构与其他背景组织区分开来。选用了交叉熵损失函数和 Dice 损失函数组成的混合损失函数。

交叉熵损失函数能够衡量预测分割结果与真实标签之间的差异。交叉熵损失函数能在每个像素级别上进行调整, 对细小且关键的血管区域更加敏感, 提高血管区域和背景区域的分割精度, 交叉熵函数  $L_{CE}$  的计算如式(14)所示。

$$L_{CE} = -\frac{1}{N} \sum_{i=1}^N [y \log \hat{y} + (1 - y) \log (1 - \hat{y})] \quad (14)$$

其中,  $y$  表示标签值,  $\hat{y}$  表示模型计算结果的预测值,  $N$  表示样本数。

Dice 相似性系数可以评价两个样本之间的相似性, 对处理两个样本中的不平衡元素有很好的效果。在视网膜血管分割任务中, 血管像素所占比例不足, 这种类别不平衡可能导致模型过于偏向于背景类, 忽视血管, Dice 损失函数是基于预测结果与真实标签的重叠度来进行评估, 因此使用 Dice 损失函数来优化训练过程可以避免交叉熵损失可能导致的背景偏倚问题。  $L_{Dice}$  的计算如

式(15)所示。

$$L_{Dice} = 1 - \frac{2TP}{2TP + FP + FN}$$

(15)

在训练过程中结合交叉熵和 Dice 相似性系数损失函数形成混合训练损失  $L_{Total}$ ，将它们结合起来，确保模型在训练过程中得到均衡的优化，避免某种类型的损失主导训练过程，提高模型的泛化能力和分割的准确性，计算过程如式(16)所示。

$$L_{Total} = \alpha L_{CE} + \beta L_{Dice}$$

(16)

其中， $\alpha$  和  $\beta$  为权重参数，值都在  $[0,1]$  范围内，用于控制两部分损失的相对重要性。

### 3.3 性能评价

实验是在 NVIDIA RTX3090 平台上使用 PyTorch 框架实现的。该研究对网络进行了 200 轮的训练，初始学习率为 0.01，采用衰减策略。实验中使用随机翻转和随机切片的方法进行数据增强：首先，对图像做随机翻转，包括水平翻转和垂直翻转，然后从每张经过翻转的图像中随机切片出固定大小的图像块作为训练样本。该研究在 DRIVE 数据集和 CHASE\_DB1 数据集上将 FreNet 与经典算法和最近提出的算法进行比较。表 1 比较了不同算法使用 DRIVE 数据集的分割性能，表 2 比较了不同算法使用 CHASE\_DB1 数据集的分割性能。

表 1 不同分割方法在 DRIVE 数据集上的性能比较

Table 1 Comparison of segmentation performance across different methods on the DRIVE dataset

方法名称	ACC	SE	SP	F1
U-Net <sup>[10]</sup>	0.950 9	0.703 1	0.987 1	0.784 9
DEUNet <sup>[28]</sup>	0.956 7	0.794 0	0.981 6	0.827 0
R2UNet <sup>[12]</sup>	0.955 6	0.779 2	0.981 3	0.817 1
Nest U-Net <sup>[29]</sup>	0.951 2	0.806 0	0.986 9	0.786 3
MS-LSDNet <sup>[7]</sup>	0.959 7	0.674 5	0.987 2	
MFI-Net <sup>[13]</sup>	0.958 1	0.817 0	0.979 0	0.831 5
Wave-Net <sup>[14]</sup>	0.956 1	0.816 4	0.976 4	0.825 4
ResDO-UNet <sup>[16]</sup>	0.956 1	0.798 5	0.979 1	0.829 9
AFNet <sup>[30]</sup>	0.957 3	0.804 9	0.979 6	
Ma's model <sup>[11]</sup>	0.956 5	0.824 3	0.975 8	0.828 4
RCAR-UNet <sup>[31]</sup>	0.953 7	0.748 7	0.983 6	0.804 7
FreNet	0.969 0	0.838 0	0.985 5	0.821 1

如表 1 和 2 所示，提出的网络产生了积极的成果。在 DRIVE 数据集上，准确度指标最优为 96.90，敏感度指标最优为 83.80。在 CHASE\_DB1 数据集上，准确度最优为 97.81，敏感度指标最优为 84.57。该研究提出的模型更容易识别出小血管，天然的会增加血管识别错误的

表 2 不同分割方法在 CHASE\_DB1 数据集上的性能比较

Table 2 Comparison of segmentation performance across different methods on the CHASE\_DB1 dataset

方法名称	ACC	SE	SP	F1
U-Net <sup>[10]</sup>	0.951 7	0.636 5	0.982 0	0.699 5
R2UNet <sup>[12]</sup>	0.965 5	0.797 0	0.982 3	0.807 3
DEUNet <sup>[28]</sup>	0.966 1	0.807 4	0.982 1	0.803 7
Nest U-Net <sup>[29]</sup>	0.966 1	0.807 4	0.982 1	0.803 7
MFI-Net <sup>[13]</sup>	0.967 5	0.838 8	0.980 6	0.824 6
Wave-Net <sup>[14]</sup>	0.966 4	0.828 4	0.982 1	0.834 9
AFNet <sup>[30]</sup>	0.966 9	0.819 4	0.981 7	
OCE-Net <sup>[32]</sup>	0.967 8	0.813 8	0.982 4	0.819 6
Ma's model <sup>[11]</sup>	0.966 7	0.777 9	0.990 0	0.837 3
RCAR-UNet <sup>[31]</sup>	0.956 6	0.747 5	0.979 8	0.747 0
LMFR-Net <sup>[33]</sup>	0.974 0	0.838 5	0.984 4	0.829 1
FreNet	0.978 1	0.845 7	0.986 2	0.801 2

可能，所以该研究提出的模型在特异性上为次优。相较于其他模型，FreNet 引入频域空间域双支路进行特征提取与融合。频域支路能够捕捉到图像的全局特征，去除噪声平滑背景，突出血管的结构。空间域支路则专注于血管细节和边缘信息，通过频域空间域双支路方法，可以在多尺度上优化特征，减少仅对单一局部或全局特征的依赖，从而提高在小血管连续性方面和噪声干扰方面的分割精度。

对两个数据集进行了交叉验证，进一步验证该模型的可靠性及稳定性，如表 3 所示。交叉验证结果显示准确率有所下降，这可能由于不同的数据集采集眼底视网膜图像时的成像角度不同所造成的，CHASEDB1 数据集的图像主要是以视网膜黄斑为中心的区域采集的图像，而 DRIVE 数据集的图像主要从视网膜黄斑的边侧采集的图像。此外，CHASEDB1 数据集的分辨率较高，视网膜血管形状也较为简单，也影响了交叉验证的准确率。实验结果表明，所提出的模型仍具有良好的泛化能力。

表 3 数据集交叉验证结果

Table 3 Cross-validation results across datasets

训练集	测试集	ACC	SE	SP	F1
DRIVE	DRIVE	0.969 0	0.838 0	0.985 5	0.821 1
CHASEDB1	CHASEDB1	0.978 1	0.845 7	0.986 2	0.801 2
DRIVE	CHASEDB1	0.913 5	0.615 6	0.986 0	0.656 4
CHASEDB1	DRIVE	0.937 1	0.655 8	0.989 3	0.690 1

该研究提出的模型 FreNet 通过引入 FRB 模块、MGHF 模块和 SSF 模块来提高视网膜血管分割小血管连续性方面的能力。该研究在 DRIVE 数据集上进行消融实验,验证所设计的 3 个模块的有效性。为保证实验的公平性,所有对比方法均采用相同的训练策略和超参数设置,一共设计了 5 组实验,M1 为分别为在 Baseline 添加 FRB 模块的分割结果,M2 为添加 SSF 模块的分割结果,M3 为添加 MGHF 模块,如表 4 所示。

表 4 添加不同模块在 DRIVE 数据集上的性能比较  
Table 4 Performance comparison with different modules added on the DRIVE dataset

方法名称	ACC	SE	SP	F1
Baseline	0.958 7	0.753 7	0.982 6	0.815 7
+FRB (M1)	0.960 6	0.797 6	0.984 1	0.818 5
+FRB+SSF (M2)	0.966 8	0.814 6	0.985 6	0.820 7
+FRB+MGHF (M3)	0.968 7	0.826 3	0.985 5	0.819 5
FreNet (+FRB+MGHF+SSF)	0.969 0	0.838 0	0.985 5	0.821 1

通过观察表 4 的结果可知,相比于基础模型都实现了准确度和敏感度的显著提升,M1 引入的频域信息与空间域信息可形成互补,二者在不同语义尺度上,能捕捉视网膜血管的各类特征,使得模型在处理复杂形状和边缘时更为精确。M2 引入的 SSF 模块通过自适应融合空间频域特征,有效地学习了血管长距离语义信息,使血管分割的连贯性和整体性显著提高,表明这种特征融合策略是有效的。M3 加入 MGHF 模块滤除低频噪声,加强图像中的高频成分,增强模型对血管边界和轮廓信息的感知能力。尽管背景区域分割精度略有降低,但血管分割的精度和整体性能得到了很大的提高。根据表 4 实验数据可知,在所有模块的共同作用下,网络的综合表现性能最好。

从 DRIVE 数据集和 CHASE\_DB1 数据集中选取了几张图片来展示该研究的模型分割结果,并与 U-Net 等模型进行了比较,如图 6 和 7 所示。图 6 和 7 给出了一些特定血管的放大细节,相比于 AttU-Net 和 R2U-Net 方法,该研究的分割结果保留了更多细节,可以分割出细小的血管,并且在小血管的连贯性方面表现出更好的视觉分割结果。

该研究选取了存在干扰的眼底图像上血管分割结果,如图 8 所示。其中图 8 的干扰因素主要由病灶结构所引起的,可以看到所提出的方法相比其他方法在病灶处受到的干扰更少,血管的连贯性和分割结果更准确,验证了所提方法稳定性和鲁棒性。

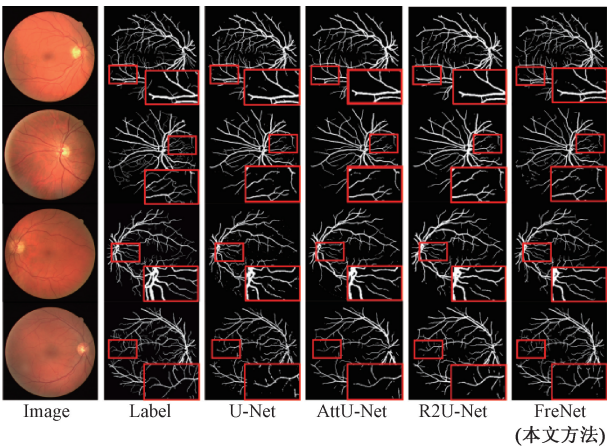


图 6 对 DRIVE 数据集上的血管分割结果进行分析  
Fig. 6 Analysis of blood vessel segmentation results on the DRIVE dataset

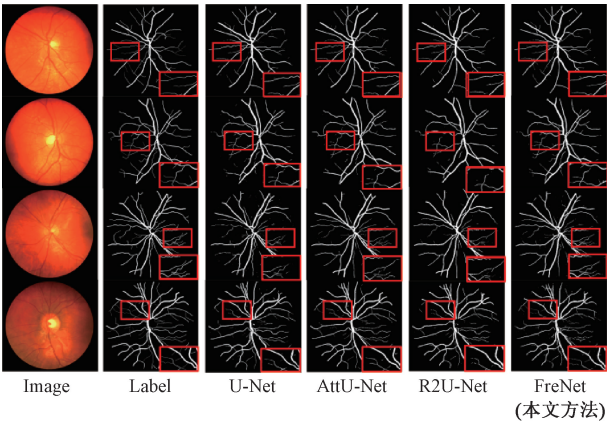


图 7 对 CHASE\_DB 数据集上的血管分割结果进行分析  
Fig. 7 Analysis of blood vessel segmentation results on the CHASE\_DB1 dataset

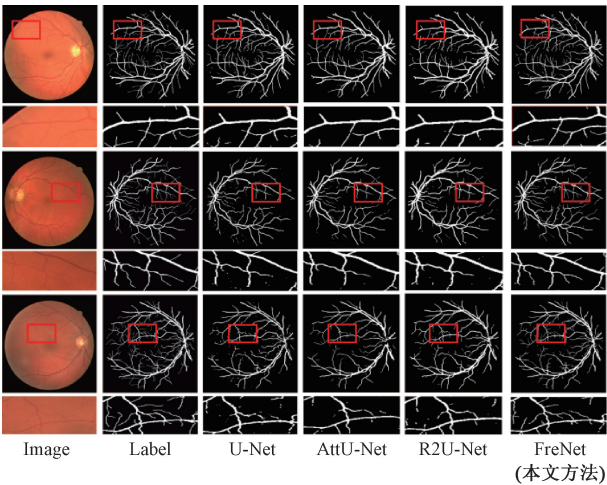


图 8 干扰图像分割结果对比  
Fig. 8 Comparison of segmentation results on interference images



可以注意到相比 U-Net,该研究所提出的 FreNet 模型在小血管、血管的连贯性方面和在病灶处血管分割的稳定性和鲁棒性方面更清晰,该研究方法的分割效果优于其他方法,进一步证明了该研究方法的优越性。

## 4 结 论

视网膜血管分割是医疗诊断中的一项任务,尽管已经进行了大量的研究,但仍然存在许多挑战,在该研究中,提出了一种空间域和频域相融合的方法。所提方法能够有效捕获多尺度上下文信息,融合不同层次的特征,如频域分支可以得到全局感受野,对血管的长距离特征信息的提取有益。空间域分支可以获得局部细节信息。在该网络中,设计的 MGHF 模块可以增强血管边缘和细节信息,获取多尺度的特征信息。通过使用 SSF 模块融合全局频率特性和局部空间特征,补充在卷积操作中获得的局部特征操作所带来的空间特征信息损失。在 DRIVE 和 CHASE\_DB1 数据集上,该研究的方法展示了强大的分割性能,并显著增强了分割小血管的能力。

## 参考文献

- [ 1 ] AGARWAL S, BHAT A. A survey on recent developments in diabetic retinopathy detection through integration of deep learning[J]. Multimedia Tools and Applications, 2023, 82(11): 17321-17351.
- [ 2 ] YE ZH P, LIU Y Q, JING T, et al. A high-resolution network with strip attention for retinal vessel segmentation[J]. Sensors, 2023, 23(21): 8899.
- [ 3 ] 张思杰,方翔,魏赋. 基于 GAN 的少样本视网膜血管分割研究[J]. 电子测量与仪器学报, 2021, 35(11): 132-142.  
ZHANG S J, FANG X, WEI F, et al. Research on retinal vascular segmentation based on GAN using few samples[J]. Journal of Electronic Measurement and Instrumentation, 2021, 35(11): 132-142.
- [ 4 ] 梁礼明,詹涛,雷坤,等. 多级自适应尺度的 U 型视网膜血管分割算法[J]. 电子测量技术, 2022, 45(13): 130-140.  
LIANG L M, ZHAN T, LEI K, et al. Multi-level adaptive scale U-shaped retinal blood vessel segmentation algorithm[J]. Electronic Measurement Technology, 2022, 45(13): 130-140.
- [ 5 ] SHANG ZH H, YU CH H, HUANG H, et al. DCNet: A lightweight retinal vessel segmentation network[J]. Digital Signal Processing, 2024, 153: 104651.
- [ 6 ] LI Y, ZHANG Y, LIU J Y, et al. Global transformer and dual local attention network via deep-shallow hierarchical feature fusion for retinal vessel segmentation[J]. IEEE Transactions on Cybernetics, 2022, 53(9): 5826-5839.
- [ 7 ] DU H W, ZHANG X Y, SONG G, et al. Retinal blood vessel segmentation by using the MS-LSDNet network and geometric skeleton reconnection method[J]. Computers in Biology and Medicine, 2023, 153: 106416.
- [ 8 ] WU CH W, GUO M, MA M, et al. Diverter transformer-based multi-encoder-multi-decoder network model for medical retinal blood vessel image segmentation[J]. Biomedical Signal Processing and Control, 2024, 93: 106132.
- [ 9 ] TANG X F, CHEN H, LI X R, et al. The joint learning of multi-resolution feature for multi-class retinal vessel segmentation[J]. Neurocomputing, 2024, 584: 127570.
- [ 10 ] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[C]. Medical Image Computing and Computer-Assisted Intervention, 2015: 234-241.
- [ 11 ] MA ZH D, LI X B. An improved supervised and attention mechanism-based U-Net algorithm for retinal vessel segmentation[J]. Computers in Biology and Medicine, 2024, 168: 107770.
- [ 12 ] 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): 014006.
- [ 13 ] YE Y W, PAN CH W, WU Y CH, et al. MFI-Net: Multiscale feature interaction network for retinal vessel segmentation[J]. IEEE Journal of Biomedical and Health Informatics, 2022, 26(9): 4551-4562.
- [ 14 ] LIU Y H, SHEN J, YANG L, et al. Wave-Net: A lightweight deep network for retinal vessel segmentation from fundus images[J]. Computers in Biology and Medicine, 2023, 152: 106341.
- [ 15 ] 夏平,何志豪,雷帮军,等. 密集级联卷积与自注意力特征聚合的视网膜血管分割算法[J]. 电子测量与仪器学报, 2024, 38(9): 36-44.  
XIA P, HE ZH H, LEI B J, et al. Retinal vascular segmentation algorithm based on full scale dense convolutional u-shaped networks[J]. Journal of

- Electronic Measurement and Instrumentation, 2024, 38(9):36-44.
- [16] LIU Y H, SHEN J, YANG L, et al. ResDO-UNet: A deep residual network for accurate retinal vessel segmentation from fundus images[J]. Biomedical Signal Processing and Control, 2023, 79: 104087.
- [17] 苑玮琦, 王安. 基于主曲率和主方向的多尺度视网膜血管骨架提取方法[J]. 仪器仪表学报, 2021, 42(6): 191-199.
- YUAN W Q, WANG AN. Multiscale retinal vascular skeleton extraction method based on principal curvature and principal direction[J]. Chinese Journal of Scientific Instrument, 2021, 42(6): 191-199.
- [18] XIE Y, SHANG J K, YANG Q, et al. ARSA-UNet: Atrous residual network based on Structure-Adaptive model for retinal vessel segmentation[J]. Biomedical Signal Processing and Control, 2024, 96: 106595.
- [19] RAO Y M, ZHAO W L, ZHU ZH, et al. Global filter networks for image classification[J]. Advances in Neural Information Processing Systems, 2021, 34: 980-993.
- [20] FUOLI D, VAN GOOL L, TIMOFTE R. Fourier space losses for efficient perceptual image super-resolution[C]. Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 2360-2369.
- [21] YU H, ZHENG N SH, ZHOU M, et al. Frequency and spatial dual guidance for image dehazing[C]. European Conference on Computer Vision. Cham: Springer Nature Switzerland, 2022: 181-198.
- [22] HUANG J, LIU Y J, ZHAO F, et al. Deep fourier-based exposure correction network with spatial-frequency interaction[C]. European Conference on Computer Vision, Cham: Springer Nature Switzerland, 2022: 163-180.
- [23] WANG Y, YU K, CHEN CH, et al. Dynamic graph learning with content-guided spatial-frequency relation reasoning for deepfake detection[C]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2023: 7278-7287.
- [24] XIAO G B, TANG ZH M, GUO H L, et al. FAFusion: Learning for infrared and visible image fusion via frequency awareness[J]. IEEE Transactions on Instrumentation and Measurement, 2024, 73: 1-11.
- [25] YU ZH, SUN B Y, LIU D, et al. STRASS dehazing: Spatio-temporal retinex-inspired dehazing by an averaging of stochastic samples[J]. Journal of Renewable Materials, 2022, 10(5): 1381.
- [26] STAAL J, ABRAMOFF M D, NIEMEIJER M, et al. Ridge-based vessel segmentation in color images of the retina[J]. IEEE Transactions on Medical Imaging, 2004, 23(4): 501-509.
- [27] OWEN C G, RUDNICKA A R, MULLEN R, et al. Measuring retinal vessel tortuosity in 10-year-old children: Validation of the computer-assisted image analysis of the retina program[J]. Investigative Ophthalmology and Visual Science, 2009, 50(5): 2004-2010.
- [28] WANG B, QIU SH, HE H G. Dual encoding U-Net for retinal vessel segmentation[C]. Medical Image Computing and Computer Assisted Intervention 2019: 84-92.
- [29] WANG CH, ZHAO Z Y, YU Y. Fine retinal vessel segmentation by combining nest U-Net and patch-learning[J]. Soft Computing, 2021, 25: 5519-5532.
- [30] LI D Y, PENG L X, PENG SH H, et al. Retinal vessel segmentation by using AFNet[J]. The Visual Computer, 2023, 39(5): 1929-1941.
- [31] DING W P, SUN Y, HUANG J SH, et al. RCAR-UNet: Retinal vessel segmentation network algorithm via novel rough attention mechanism[J]. Information Sciences, 2024, 657: 120007.
- [32] WEI X X, YANG K F, BZDOK D, et al. Orientation and context entangled network for retinal vessel segmentation[J]. Expert Systems with Applications, 2023, 217: 119443.
- [33] ZHANG W H, QU SH J, FENG Y W. LMFR-Net: Lightweight multi-scale feature refinement network for retinal vessel segmentation[J]. Pattern Analysis and Applications, 2025, 28(2): 44.

## 作者简介



宋伟伟, 2018 年于潍坊学院获得学士学位, 现为沈阳工业大学在读硕士研究生, 主要研究方向为医学图像处理。

E-mail: sutsww@163.com

**Song Weiwei** received his B.Sc. degree from Weifang University in 2018. Now he is a postgraduate student at Shenyang University of Technology. His main research

direction is medical image processing.



**许茗** (通信作者), 分别于 2012 年、2014 年、2021 年于东北大学获得学士、硕士和博士学位, 现为沈阳工业大学人工智能学院讲师, 主要研究方向为异常事件检测、医学图像处理和无监督学习。

E-mail: xuming@ sut. edu. cn

**Xu Ming** (Corresponding author) received his B. Sc. , M. Sc and Ph. D. degrees all from Northeastern University in 2012, 2014 and 2021, respectively. Now he is a master 's supervisor in the School of Artificial Intelligence, Shenyang University of Technology. His main research directions are abnormal event detection, medical image processing and unsupervised learning.



**于晓升**, 2006 年于英国贝特福德大学获得学士学位, 2007 年于英国利物浦大学获得硕士学位, 2014 年于东北大学获得博士学位, 现为东北大学机器人科学与工程学院副教授, 主要研究方向为视觉显著性检测, 多视图合成和医学图像处理等。

E-mail: yuxiaosheng@ mail. neu. edu. cn

**Yu Xiaosheng** received his B. Sc. degree from the University of Bedford in the United Kingdom in 2006, M. Sc degree from the University of Liverpool in the United Kingdom in 2007, and Ph. D. degree from Northeastern University in 2014. Now he is an associate professor at the School of Robot Science and Engineering at Northeastern University. His main research

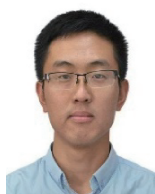
directions are visual saliency detection, multi-view synthesis and medical image processing.



**李海星**, 2014 年于东北农业大学获得学士学位, 2022 年于中国科学院大学获得博士学位, 现为沈阳工业大学人工智能学院讲师, 主要研究方向为计算机视觉, 医学图像处理。

E-mail: lhx@ sut. edu. cn

**Li Haixing** received her B. Sc. degree from Northeast Agricultural University in 2014, Ph. D. degree from the University of Chinese Academy of Sciences in 2022. Now she is a lecturer in the School of Artificial Intelligence of Shenyang University of Technology. Her main research directions are computer vision and medical image processing.



**王宏宇**, 2013 年于沈阳航空航天大学获得学士学位, 2015 年于东北大学获得硕士学位, 2019 年于东北大学获得博士学位, 现为北京大学博士后, 主要研究方向为结构光三维感知和医学图像处理。

E-mail: hongyuwang@ pku. edu. cn

**Wang Hongyu** received his B. Sc. degree from Shenyang Aerospace University in 2013, M. Sc degree from Northeastern University in 2015, and Ph. D. degree from Northeastern University in 2019. He is now a postdoctoral fellow at Peking University. His main research directions are structured light three-dimensional perception and medical image processing.