

# 基于改进 Swin-Unet 的糖尿病性黄斑水肿分割方法

汪 扬<sup>1</sup> 卓广平<sup>1</sup> 阙玉常<sup>1</sup> 刘国强<sup>1</sup> 张光华<sup>2</sup>

WANG Yang ZHUO Guangping KAN Yuchang LIU Guoqiang ZHANG Guanghua

## 摘要

糖尿病性黄斑水肿 (DME) 是糖尿病患者最常见的失明原因之一, 光学相干断层扫描技术 (OCT) 有助于糖尿病视网膜病变的早期检测和预防, 视网膜 OCT 图像中的 DME 区域分割常被用于定量评估黄斑区的水肿程度和病变区域的面积。针对糖尿病性黄斑水肿分割存在的小目标区域漏分割以及部分水肿区域分割不准确的问题, 提出了一种改进 Swin-Unet 的 DME 分割方法, 提高了区域分割的准确性。首先, 将 Swin-Unet 的瓶颈替换为由快速傅里叶卷积 (FFC) 块构成的频域特征提取模块, 用于提取视网膜 OCT 图像中所包含的光谱域信息; 其次, 对语义分割后的区域进行形态学图像处理技术进行处理, 提升结果的可视化效果; 最后, 在 OCT2017 和杜克大学数据集上对模型进行了验证。实验结果表明, 与 Swin-Unet 相比, 所提出的方法在 OCT2017 数据集上 dice 相似系数、交并比、召回率和精确率分别提升了 3%、2%、2% 和 3%, 在杜克大学数据集上分别提升了 5%、4%、6% 和 5%, 改进后的 Swin-Unet 模型可以提升对 DME 分割的精度, 为医生提供更为可靠的诊断依据。

## 关键词

深度学习; 糖尿病性黄斑水肿; OCT 图像分割; Swin-Unet; 快速傅里叶卷积

doi: 10.3969/j.issn.1672-9528.2024.01.005

## 0 引言

糖尿病性黄斑水肿 (diabetic macular edema, DME) 是一种严重的糖尿病视网膜病变, 是指黄斑区发生的视网膜水肿。DME 是糖尿病患者最常见的失明原因之一, 对患者生活和社会经济都造成了严重影响<sup>[1]</sup>。

光学相干断层扫描 (optical coherence tomography, OCT) 技术是一种基于光学干涉原理的非侵入式成像技术, 它兼顾了安全性高、成像速度快、纵向分辨率高、横向分辨率高等优点, 可以实现对生物组织的高精度成像, 对于人眼的视网膜成像有十分重要的应用<sup>[2]</sup>。在 DME 的诊断和治疗过程中, 视网膜 OCT 图像中的水肿区域分割被用于定量评估黄斑区的水肿程度和病变区域的面积, 以及监测治疗效果<sup>[3]</sup>。

在早期, DME 区域的分割主要采用人工分割的方式, 由于优秀眼科医生的稀缺, 分割效果不尽如人意, 同时也会耗费大量人力物力。随着科技水平的发展, 目前的 OCT 图像分割方法主要分为基于传统图像处理的方法和基于深度学习的方法。基于传统图像处理的方法主要包括阈值法、边缘检测和区域生长等, 这些方法的优点在于计算速度快、易于实现, 但是在处理复杂情况下的图像时存在一定的局限性。

近些年来, 越来越多的基于深度学习的医学图像分割方法被提出, Long 等人<sup>[4]</sup>提出的全卷积网络 (fully convolutional neural network, FCN) 在图像分割领域取得了显著的效果, 也为之后的研究奠定了极大的基础。U-Net<sup>[5]</sup>在 FCN 网络的基础上, 通过采用 U 型结构, 使用编码器、解码器、瓶颈处理以及跳跃连接将语义信息与空间信息相结合, 在医学图像分割领域得到了人们的信赖。当前, 越来越多的基于卷积神经网络的分割模型被应用到 OCT 图像的 DME 分割任务中去。Venhuizen 等人<sup>[6]</sup>提出了一种基于 U-Net 的两阶段全卷积神经网络, 包括两阶段架构, 第一阶段用于提取特征, 第二阶段用于水肿分割, 用来减少背景类对分割效果的影响。Liu 等人<sup>[7]</sup>提出了一种基于多尺度输入、多尺度侧输出和双注意机制的增强嵌套 U-Net 结构, 在多层分割和多流体分割任务上取得了优异的分割性能。Gao 等人<sup>[8]</sup>在 U-Net++ 中使用 ResNet<sup>[9]</sup>作为骨干网, 重新设计了跳跃连接结构, 对合成结构进行了改进, 对 OCT 图像中的水肿区域进行了分割, 获得了良好的分割效果。Farshad 等人<sup>[10]</sup>通过改进 U-Net 提出了一种能够同时提取光谱域特征与空间域特征的双编码器网络 Y-Net, 在 OCT 图像中的水肿区域分割和视网膜层分割任务中取得了十分不错的效果。由于 transformer 自身的注意力机制具有强大的远程上下文信息建模能力, 可以很好地应用到医学图像分割中<sup>[11]</sup>。近些年, 一系列基于 transformer

1. 太原师范学院计算机科学与技术学院 山西晋中 030619

2. 太原学院智能与自动化系 山西太原 030032

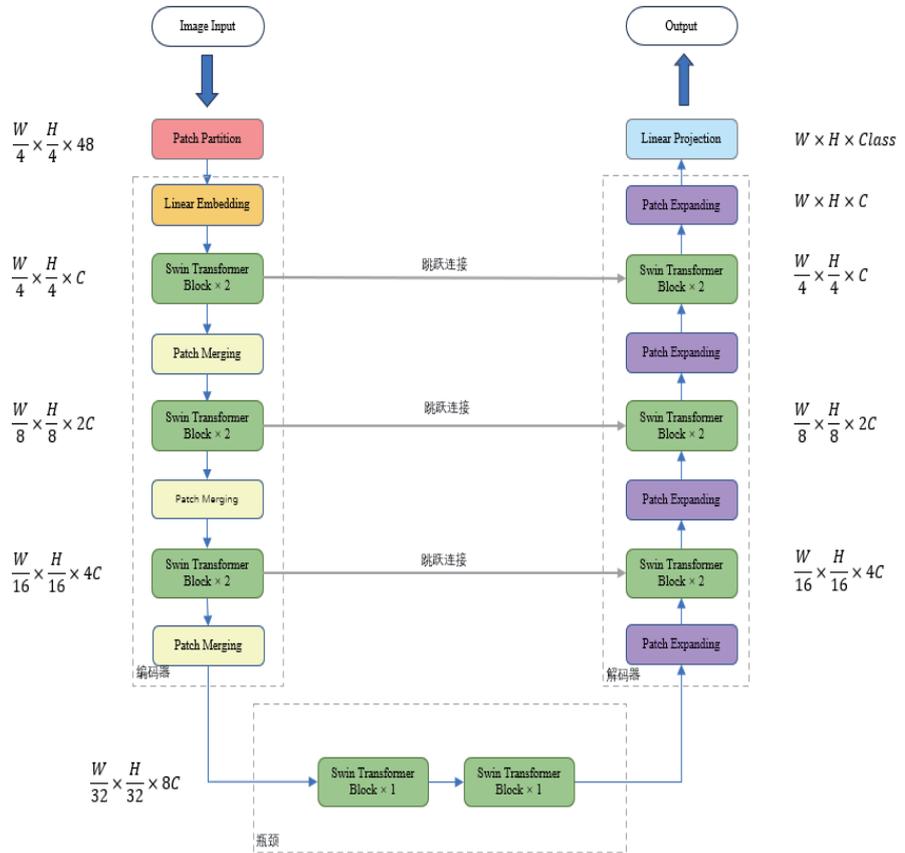
[基金项目] 山西省自然科学基金面上项目 (201801D121147)

的医学图像分割方法被研究人员们提出。Chen 等人<sup>[12]</sup>提出了 transUNet, 它采用 CNN 与 transformer 相结合的编码器结构, 兼具 transformer 和 U-Net 的优点。Hu 等人<sup>[13]</sup>将 U-Net 中的传统卷积替换为 Swin transformer<sup>[14]</sup> 块, 提出了一种全 transformer 的 U 型网络 Swin-Unet, 保留了 U-Net 中的跳跃连接, 很好地保留了低层次的特征信息。研究人员们已经开始尝试使用基于 transformer 的方法进行 DME 分割, 如 Gao 等人<sup>[15]</sup>将 Swin-Unet 网络中的瓶颈部分的 Swin transformer 块替换为 ResNet34 的第五层网络, 在黄斑水肿流体分割任务上取得了不错的成果。

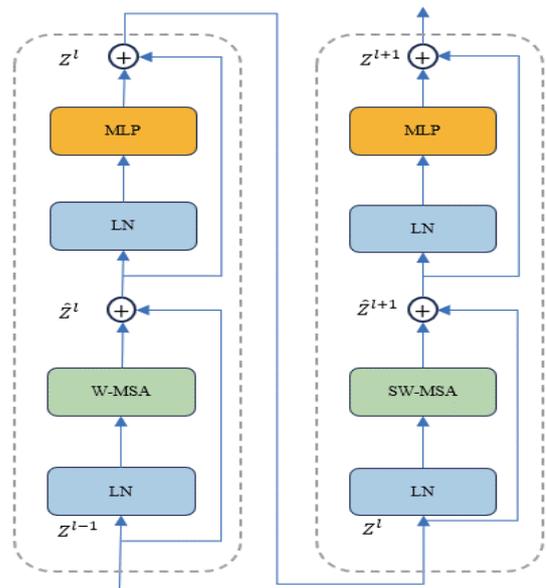
视网膜 OCT 图像中的水肿区域分割在医生诊断的过程中起到了十分重要的作用, 但现有的方法对存在的小目标区域不能有效地进行识别和分割。同时, 有专家指出, OCT 图像中存在光谱特征<sup>[9]</sup>, 这些光谱特征可以帮助模型更好地进行分割, 现有的空间神经网络可能会遗漏这些信息<sup>[16]</sup>。对以上问题, 本文对 Swin-Unet 进行相关改进, 提出了一种同时利用光谱域特征与空间域特征的分割方法 FSwin-Unet (FFC-Swin-Unet)。

### 1 Swin-Unet 模型简述

Swin-Unet 是一种纯 transformer 的 U 型编码器 - 解码器结构, 专门用于医学图像分割。如图 1 (a) 所示, 其整体结构由编码器、瓶颈、解码器和跳跃连接组成。Swin-Unet 首先将医学图像分割成大小为 4×4 的非重叠 patch, 将输入图像转换为序列后输入编码器。其次, 采用 patch 合并层进行下采样降低空间分辨率, 同时增加特征维数, 对称的解码器通过 patch 扩展层执行上采样, 恢复特征图的空间分辨率, 同时减少特征维数。Swin-Unet 继承了 U-Net 中的跳跃连接, 将图像中各个层次的信息进行融合, 提高了网络模型的学习能力。其基本处理单元为 Swin transformer 块, 如图 1 (b) 所示, 每个处理单元为两个连续的 Swin transformer 块, 每个 Swin transformer 块由 LayerNorm (LN) 层, 多头注意力模块、残差连接以及具有 GELU 非线性的 2 层 MLP 构成, 基于窗口的多头自注意 (W-MSA) 模块和基于移位窗口的多头自注意 (SW-MSA) 模块分别应用于这两个连续的 Swin transformer 块。



(a) 整体结构示意图



(b) Swin transformer 块结构示意图

图 1 Swin-Unet 模型结构示意图

在这种窗口划分机制下, 每个处理单元的流程可表示为:

$$\hat{Z} = W\_MSA(LN(Z^{l-1})) + Z^{l-1} \tag{1}$$

$$Z^l = MLP(LN(\hat{Z})) + \hat{Z} \tag{2}$$

$$\hat{Z}^{l+1} = MLP(LN(Z^l)) + Z^l \tag{3}$$

$$Z^{l+1} = MLP(LN(\hat{Z}^{l+1})) + \hat{Z}^{l+1} \quad (4)$$

式中： $\hat{Z}^l$ 与 $Z^l$ 分别为表示第 $l$ 个模块的(S)W-MSA 模块和 MLP 模块的输出。

其自注意力的计算公式为：

$$Attention(Q, K, V) = SoftMax(\frac{QK^T}{\sqrt{d}} + B)V \quad (5)$$

式中： $Q, K, V \in \mathbb{R}^{M^2 \times d}$ 分别表示查询矩阵 (Query)、键矩阵 (Key) 和值矩阵 (Value)。 $M^2$ 为窗口中的 patch 数量。 $d$ 表示查询矩阵或键矩阵的维度。 $B$ 的值取自偏置矩阵  $\hat{B} \in \mathbb{R}^{(2M-1) \times (2M-1)}$ 。

Swin-Unet 很好地实现了局部到全局的自关注，实现了像素级的自我关注，与传统的卷积操作相比，它能够更好地学习全局的语义信息，提高医学图像分割的性能。

## 2 Swin-Unet 相关改进

### 2.1 整体结构

本文模型基于改进的 Swin-Unet 模型，对 DME 患者的视网膜 OCT 图像中的黄斑水肿区域和图像背景区域进行两级语义高精度的分割操作，改进后模型的整体结构如图 2 所示。

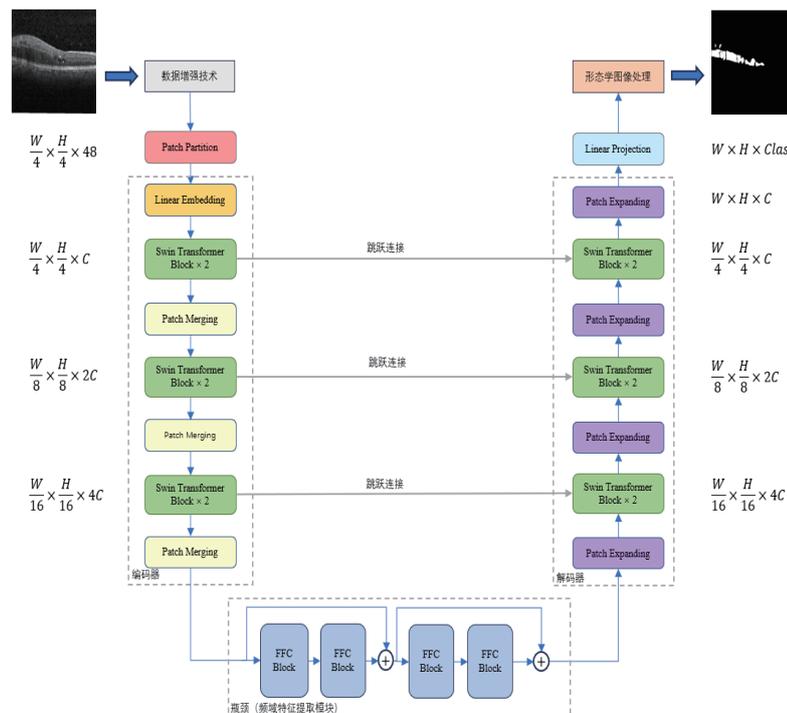


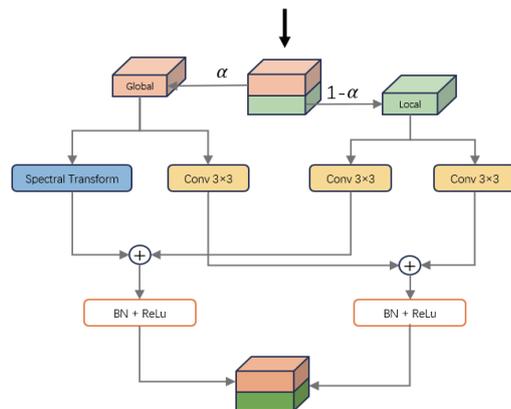
图 2 FSwin-Unet 模型结构示意图

本文将 Swin-Unet 瓶颈处的两个连续 Swin transformer 块替换为一个频域特征提取模块，该模块从不同的频率分布中分离出特征，这使得改进后的模型能够关注特征中更为重要的频率范围，识别 OCT 图像中的高频斑点<sup>[17]</sup>，使网络具有提取 OCT 图像中的光谱特征的能力。同时，本文采用形态学

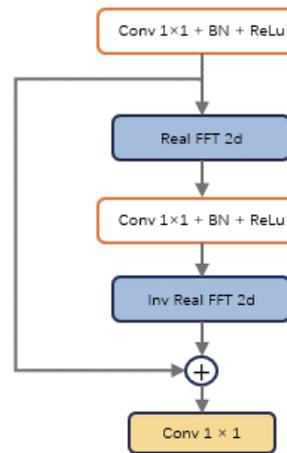
图像处理技术<sup>[18]</sup>对分割后的图像进行后处理操作，使分割结果达到最佳的可视化效果。

### 2.2 频域特征提取模块

频域特征提取模块主要由四个快速傅里叶卷积 (fast fourier convolution, FFC) 块构成，其中每两个 FFC 块进行残差连接。FFC 是一种基于快速傅里叶变换<sup>[19]</sup> (fast fourier transform, FFT) 的卷积方法，它的结构如图 3 (a) 所示，FFC 通过非局部感受野和跨尺度融合改善传统卷积的局部性和固定尺度的限制，在单个操作单元中整合了局部、半全局和全局分支，实现了对不同尺度的综合处理。其中 $\alpha$ 定义了特征在全局和局部傅里叶单元中的百分比，之前的研究认为 $\alpha$ 的值为 0.5 时，可以得到较好的结果，故本文将 $\alpha$ 的值定义为 0.5。该结构主要通过 Special transform 块实现对 OCT 图像中包含的光信号转换到频域表示，从频域的角度分析图像的频率成分。这样可以提取图像中的频域特征，例如纹理、边缘、斑块等，其为分割任务提供更丰富的特征信息。如图 3 (b) 所示，首先将空间特征采用 FFT 转换到频域，然后在频域做卷积处理，最后通过逆 FFT 转换回去，以此达到频域特征的提取。



(a) 整体结构示意图



(b) Special transform 结构示意图

图 3 快速傅里叶卷积结构示意图

### 2.3 形态学图像处理

为了消除网络分割后产生的部分病变区域不均匀、边缘不平滑的问题，本文采用形态学方法中的开运算对分割后的图像进行后处理，开运算包含腐蚀和膨胀两种操作，主要用于消除图像中的斑点噪声，平滑分割后的水肿区域边缘，增加可视化效果。

### 2.4 损失函数

本文提出的是一种端到端的深度学习框架，主要采用 dice 损失函数和交叉熵损失函数组合的联合损失函数进行模型的训练。损失函数的计算公式为：

$$L_{dice}(y, \hat{y}) = 1 - \frac{2y\hat{y} + \varepsilon}{y + \hat{y} + \varepsilon} \quad (6)$$

$$L_{ce}(y, \hat{y}) = -\sum_{c=0}^C y_c \log \hat{y}_c \quad (7)$$

$$L = \lambda L_{dice} + (1 - \lambda)L_{ce} \quad (8)$$

式中： $y$  为分割标签， $\hat{y}$  为分割预测图， $C$  为像素类别， $\lambda$  为 dice 损失函数所占的权重。

## 3 实验与分析

### 3.1 实验数据

本文使用公开的数据集 OCT2017<sup>[20]</sup> 以及杜克大学 OCT 数据集<sup>[21]</sup> 分别进行实验。其中 OCT2017 数据集共包含 CNV、DME、DRUSEN 和 NORMAL 四类视网膜 OCT 数据。本文选取了 500 张类别为 DME、大小为  $512 \times 512$  的视网膜 OCT 图像数据作为实验的数据集，使用 LabelMe 软件对水肿区域进行标注，并且通过数据增强技术，将数据扩充到 6000 张，根据 7:2:1 的比例对数据集进行划分，得到训练集 4200 张，验证集 1200 张，测试集 600 张。杜克大学 OCT 数据集共包含 10 位患者的 110 张图像组成，每张图像的大小为  $496 \times 768$ ，在这 110 张图像中，分别由两位专家对水肿区域进行了标注。本文采用专家 1 的标注作为实验标准，通过数据增强技术将数据集扩充至 4000 张，根据 7:2:1 的比例将数据集划分，得到训练集 2800 张，验证集 800 张，测试集 400 张。

### 3.2 实验环境及参数设置

本文实验的操作平台为 Linux 操作系统，使用两张型号为 NVIDIA GeForce RTX 2080Ti 的显卡作为算力支持，实验平台使用 PyCharm2020，采用 PyTorch 1.5 框架对模型进行训练、验证和测试。优化器采用随机梯度下降优化器，其冲量设置为 0.9，迭代次数为 500 次，每次迭代进行的批处理大小为 16，初始学习率为 0.01，权值衰减设为 0.000 1。

### 3.3 定量评价指标

为了定量评估模型的分割效果，本文使用 Dice 相似系数、

交并比 (intersection over union, IoU)、召回率 (Recall) 和精确率 (Precision) 等评价指标对 OCT 图像分割的性能进行评估。Dice 相似系数计算预测结果与真实标签的重叠像素的比例，用于评估两个集合的相似程度；IoU 衡量了模型预测的区域与真实标签的重叠程度；Recall 用于评估分类或检测模型在正样本中正确识别出的比例，它衡量了模型对于真正例的识别能力；Precision 用于衡量分类模型预测为正例的样本中，真正为正例的比例，它可以衡量模型的预测准确性。

Dice、IoU、Recall 和 Precision 的计算分别为：

$$Dice = \frac{2(X \cap Y)}{X + Y} \quad (9)$$

$$IoU = \frac{|Y \cap X|}{|Y \cup X|} \quad (10)$$

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

$$Precision = \frac{TP}{TP + FP} \quad (12)$$

式中： $X$  表示模型预测的水肿区域的面积； $Y$  为标注的水肿区域面积。 $TP$  (true positive) 表示类别为水肿区域的像素点被正确预测的数量， $FP$  (false positive) 表示非水肿区域像素点被错误预测为正例的数量， $FN$  (false negative) 表示水肿区域像素点被错误预测为负例的数量。

### 3.4 数据结果分析

本文比较了 FSwin-Unet 与 Swin-Unet 在 OCT2017 数据集和杜克大学数据集上的分割性能。两个数据集上的分割结果对比图，如图 4、图 5 所示，与 Swin-Unet 相比，本文提出的方法分割出的结果图更接近原始标签，对于一些小目标水肿区域，也没有遗漏太多细节。同时，可以清楚地观察到，FSwin-Unet 分割出的结果与 Swin-Unet 相比，分割图边缘更加平滑，可视化效果更佳。

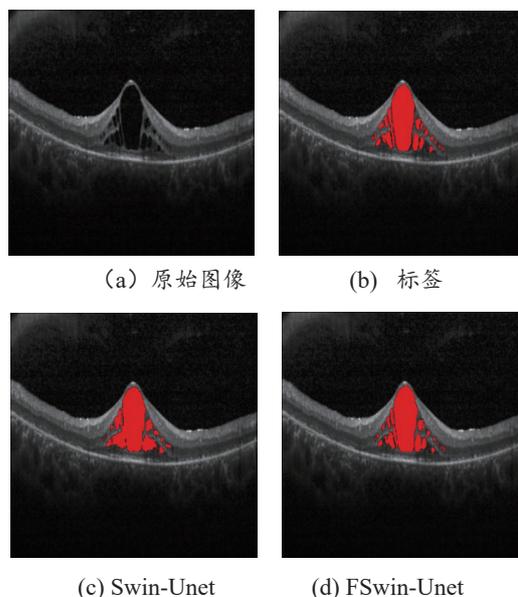


图 4 OCT2017 数据集分割对比图

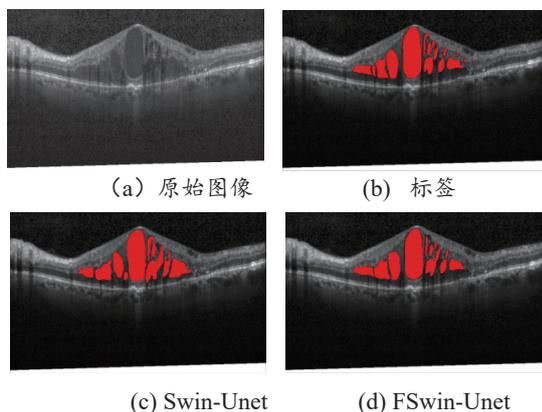


图 5 杜克大学数据集分割对比图

表 1、表 2 为 FSwin-Unet 与 Swin-Unet 在两个数据集上的性能对比。从表 1 可以看出，FSwin-Unet 算法在 OCT2017 数据集上的 Dice 相似系数、IoU、Recall 和 Precision 分别为 0.93、0.87、0.92 和 0.94。与 Swin-Unet 相比，各项指标分别提高了 0.03、0.05、0.03 和 0.03。从表 2 可以看出，FSwin-Unet 算法在杜克大学数据集上的 Dice 相似系数、IoU、Recall 和 Precision 分别为 0.83、0.80、0.85 和 0.83。同 Swin-Unet 相比，各项指标分别提高了 0.05、0.04、0.06 和 0.05。以上分析表明，FSwin-Unet 在 OCT2017 数据集与杜克大学数据集上的各项指标均高于 Swin-Unet。

表 1 OCT2017 数据集分割结果

模型	Dice	IoU	Recall	Precision
Swin-Unet	0.90	0.87	0.91	0.88
FSwin-Unet	0.93	0.89	0.93	0.91

表 2 杜克大学数据集分割结果

模型	Dice	IoU	Recall	Precision
Swin-Unet	0.78	0.76	0.79	0.78
FSwin-Unet	0.83	0.80	0.85	0.83

Swin-Unet 与 FSwin-Unet 的混淆矩阵对比图如图 6、图 7 所示。从预测黄斑水肿的平均像素数来看，可以清楚地观察到改进后的模型在 OCT2017 数据集上预测准确的黄斑水肿平均像素数增加了 65 个，同时，水肿区域错分为背景的像素数减少了 76 个，背景被错分为水肿的像素数减少了 86 个。

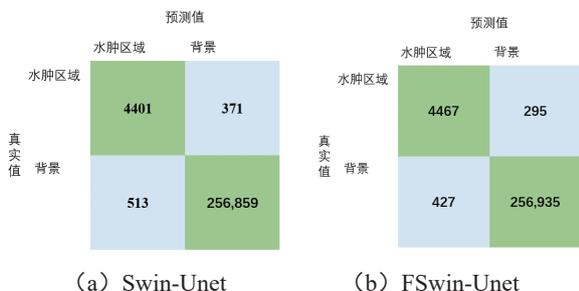


图 6 OCT2017 数据集混淆矩阵对比图

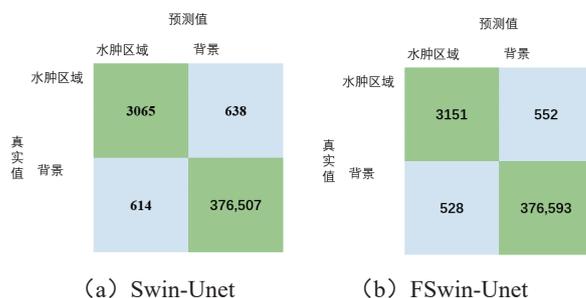


图 7 杜克大学数据集混淆矩阵对比图

在杜克大学数据集上，FSwin-Unet 相较于 Swin-Unet 预测准确的黄斑水肿平均像素数增加了 86 个，水肿区域错分为背景的像素数减少了 85 个，背景被错分为水肿的像素数减少了 61 个。

本文在杜克大学数据集上进行了消融实验，实验结果如表 3 所示。在加入了形态学处理后，模型的 dice 相似系数、IOU 和 Precision 等指标分别提升了 0.01、0.01 和 0.01；将 Swin-Unet 瓶颈替换为以残差形式连接的 2 个 FFC 块后，四项指标分别提升了 0.03、0.02、0.02 和 0.02；将 Swin-Unet 的瓶颈替换为以残差形式连接的 4 个 FFC 块时，各项指标分别提升了 0.04、0.04、0.05 和 0.05。结果表明，本文对于 Swin-Unet 分割 DME 的改进是切实有效的。

表 3 杜克大学数据集消融实验结果

模型	Dice	IoU	Recall	Precision
Swin-Unet	0.78	0.76	0.79	0.78
Swin-Unet+形态学处理	0.79	0.77	0.80	0.79
Swin-Unet+2*FFC 块	0.81	0.78	0.81	0.80
Swin-Unet+4*FFC 块	0.82	0.80	0.84	0.83

#### 4 结论

本文基于 Swin-Unet 网络模型提出了一种结合频域特征与空间特征的 FSwin-Unet 模型，引入快速傅里叶卷积提取 OCT 图像中包含的光频信息，用于分割视网膜 OCT 图像中的 DME 区域。与 Swin-Unet 模型相比，FSwin-Unet 在 OCT2017 数据集上的 Dice 相似系数、IoU、Recall 和 Precision 等评价指标分别提升了 0.03、0.02、0.02 和 0.03，在杜克大学 OCT 数据集上分别提升了 0.05、0.04、0.06 和 0.05。实验结果表明，本文提出的改进对于提升 Swin-Unet 的分割表现切实有效，采用 FSwin-Unet 可以有效地分割视网膜 OCT 图像中的 DME 区域，提高分割精度，相较于原始的 Swin-Unet 模型，分割出的结果图边缘平滑性更好。下一步将研究对 DME 分割更优的算法模型，优化模型结构、降低模型的时间复杂度、提高模型处理数据的速度，同时尝试研究对视网膜 OCT 图像中的不同内积液进行分类与分割。

## 参考文献:

- [1] MILLER K, FORTUN J A. Diabetic macular edema: current understanding, pharmacologic treatment options, and developing therapies[J]. The Asia-pacific journal of ophthalmology, 2018, 7(1): 28-35.
- [2] HUANG D, SWANSON E A, LIN C P, et al. Optical coherence tomography[J]. Science, 1991, 254(5): 1178-1181.
- [3] MOURA J, SAMAGIO G, NOVO J, et al. Joint diabetic macular edema segmentation and characterization in OCT images[J]. Journal of digital imaging, 2020, 33: 1335-1351.
- [4] LONG J, SHELHAMER E, DATTELL T. Fully convolutional networks for semantic segmentation[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2015: 3431-3440.
- [5] RONNEBERGER O, FISCHER P, BROX T. U-Net: convolutional networks for biomedical image segmentation [C]//Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference. Cham: Springer, 2015: 234-241.
- [6] VENHUIZEN F G, VAN G B, LIEFERS B, et al. Deep learning approach for the detection and quantification of intraretinal cystoid fluid in multivendor optical coherence tomography[J]. Biomedical optics express, 2018, 9(4): 1545-1569.
- [7] LIU W, SUN Y, JI Q. Mdan-UNet: multi-scale and dual attention enhanced nested U-Net architecture for segmentation of optical coherence tomography images[J]. Algorithms, 2020, 13(3): 60.
- [8] GAO Z, WANG X, LI Y. Automatic segmentation of macular edema in retinal OCT images using improved U-Net++[J]. Applied sciences, 2020, 10(16): 5701.
- [9] ZHOU Z, RAHMANV S M M, TAJBAKHS N, et al. UNet++: a nested U-Net architecture for medical image segmentation[C]//Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop. Cham: Springer International Publishing, 2018: 3-11.
- [10] FARISJAD A, YEGANEH Y, GEHLBACH P, et al. Y-Net: A spatospectral dual-encoder network for medical image segmentation[C]//International Conference on Medical Image Computing and Computer-Assisted Intervention. Cham: Springer, 2022: 582-592.
- [11] 傅励瑶, 尹梦晓, 杨锋. 基于 Transformer 的 U 型医学图像分割网络综述 [J]. 计算机应用, 2023, 43(5): 1584-1595.
- [12] CHEN J, LU Y, YU Q, et al. TransUNet: transformers make strong encoders for medical image segmentation[EB/OL]. (2021-02-08)[2023-06-29]. <https://arxiv.org/abs/2102.04306>.
- [13] CAO H, WANG Y, CHEN J, et al. Swin-Unet: UNet-like pure transformer for medical image segmentation[C]//European conference on computer vision. Cham: Springer, 2022: 205-218.
- [14] LIU Z, LIN Y, CAO Y, et al. Swin transformer: Hierarchical vision transformer using shifted windows[C]//Proceedings of the IEEE/CVF international conference on computer vision. Piscataway: IEEE, 2021: 10012-10022.
- [15] GAO Z, CHEN L. Research on semantic segmentation method of macular edema in retinal OCT images based on improved Swin-Unet[J]. Electronics, 2022, 11(15): 2294.
- [16] CHI L, JIANG B, MU Y. Fast fourier convolution[J]. Advances in neural information processing systems, 2020, 33: 4479-4488.
- [17] SCHMITT J M, XIANG S H, YUNG K M. Speckle in optical coherence tomography[J]. Journal of biomedical optics, 1999, 4(1): 95-105.
- [18] DI R C, DEMPSTER A, KHAN S, et al. Analysis of infected blood cell images using morphological operators[J]. Image and vision computing, 2002, 20(2): 133-146.
- [19] RAO K R, KIM D N, HWANG J J. Fast fourier transform: algorithms and applications[M]. Dordrecht: Springer, 2010.
- [20] SUVOROV R, LOGACHEVA E, MASHIKHIN A, et al. Resolution-robust large mask inpainting with fourier convolutions[C]//Proceedings of the IEEE/CVF winter conference on applications of computer vision. Piscataway: 2022: 2149-2159.
- [21] KERMANY D S, GOLDBAUM M, CAI W, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning[J]. Cell, 2018, 172(5): 1122-1131.
- [22] Vision and Image Processing (VIP) Laboratory. Available online: [https://people.duke.edu/~sf59/Chiu\\_BOE\\_2014\\_dataset.htm](https://people.duke.edu/~sf59/Chiu_BOE_2014_dataset.htm) (accessed on 15 August 2014).

## 【作者简介】

汪扬 (1997—), 男, 黑龙江大庆人, 硕士研究生, 研究方向: 深度学习、计算机视觉。

卓广平 (1972—), 通信作者, 男, 山西太原人, 研究生导师, 博士, 副教授, 研究方向: 大数据分析挖掘、认知与智能等。

阚玉常 (1996—), 男, 山东临沂人, 硕士研究生, 研究方向: 深度学习、计算机视觉。

刘国强 (1997—), 男, 山西吕梁人, 硕士研究生, 研究方向: 深度学习、计算机视觉。

张光华 (1986—), 男, 山西太原人, 博士, 副教授, 研究生导师, 研究方向: 智能医疗。

(收稿日期: 2023-08-09)