

DOI:10.13364/j.issn.1672-6510.20230080

基于多任务学习的青光眼智能诊断

魏宏博¹, 武劲圆², 陈磊¹, 冯梓毅¹, 游国栋¹

(1. 天津科技大学电子信息与自动化学院, 天津 300222; 2. 天津市第一中心医院, 天津 300192)

摘要: 为了提高青光眼检测的准确率,降低青光眼的危害,本文提出一种基于多任务学习的青光眼智能诊断方法,将 U-Net 网络和 VGG16 网络结合,U-Net 网络和 VGG16 网络共用 U-Net 网络的编码器部分,通过 U-Net 网络得到杯盘比(cup-to-disc ratio, CDR),并将 CDR 作为眼底图像的特征之一输入 VGG16 网络,实现眼底图像的青光眼分类。实验使用 REFUGE 挑战数据集进行验证,网络模型在训练后得到的工作特性曲线下面积为 0.9788,且视盘和视杯的分割准确率分别达到 0.8745 和 0.9624,对比其他使用相同数据集的方法,本方法具有更高的青光眼分类准确率。

关键词: 青光眼诊断; 图像分割; 图像分类; 多任务学习

中图分类号: TP183 文献标志码: A 文章编号: 1672-6510(2024)02-0059-06

Intelligent Diagnosis of Glaucoma Based on Multi-Task Learning

WEI Hongbo¹, WU Jinyuan², CHEN Lei¹, FENG Ziyi¹, YOU Guodong¹

(1. College of Electronic Information and Automation, Tianjin University of Science & Technology, Tianjin 300222, China;
2. Tianjin First Central Hospital, Tianjin 300192, China)

Abstract: In order to enhance the accuracy of glaucoma detection and mitigate the risks associated with glaucoma, in this article we propose an intelligent diagnostic method for glaucoma based on multi-task learning. Our proposed method combines the U-Net and VGG16 networks, with the encoder part of the U-Net network being shared by both networks. By utilizing the U-Net network, the cup-to-disc ratio (CDR) is obtained from retinal images, and this CDR is used as one of the features input into the VGG16 network to achieve glaucoma classification for the retinal images. The proposed method was validated using the REFUGE challenge datasets. After training the network model, the area under the receiver operating characteristic curve (AUC) was measured to be 0.9788. Moreover, the segmentation accuracy for the optic disc and optic cup was found to be 0.8745 and 0.9624, respectively. In comparison to other methods using the same datasets, the proposed method in this article demonstrates higher accuracy in glaucoma classification.

Key words: diagnosis of glaucoma; image segmentation; image classification; multi-task learning

引文格式:

魏宏博,武劲圆,陈磊,等. 基于多任务学习的青光眼智能诊断[J]. 天津科技大学学报,2024,39(2):59-64.

WEI H B, WU J Y, CHEN L, et al. Intelligent diagnosis of glaucoma based on multi-task learning[J]. Journal of Tianjin university of science & technology, 2024, 39(2):59-64.

青光眼是一种慢性的眼科疾病,其导致的视力受损是不可逆的,目前还没有治愈青光眼的方法,但是早期发现可以大幅减少青光眼造成的视力受损^[1]。青光眼最常见的筛查方法有测量眼压、测量杯盘比

(cup-disc ratio, CDR)、光学相干断层扫描(optical coherence tomography, OCT)和视野测试。CDR 与眼压密切相关,通常眼压增加会导致视神经头损伤,并可能伴有 CDR 增加^[2]。与复杂的眼压测量相比,评

收稿日期: 2023-03-31; 修回日期: 2023-07-11

基金项目: 天津市科技支撑重点项目(17YFZCNC00230);天津市应用基础与前沿技术研究计划(自然科学基金)重点项目(13JCZDJC29100)

作者简介: 魏宏博(1998—),男,安徽亳州人,硕士研究生;通信作者: 游国栋,教授, yougdong@sina.com

估视神经头或杯盘比是早期发现青光眼的简单方法。OCT 图像可以提供杯盘比信息,但是 OCT 检测设备成本太高。眼底图像在能够获得杯盘比信息的同时更加经济实惠,因此通过在临床上观察眼底图像中视盘(optic disc, OD)和视杯(optic cup, OC)的大小,可以得到杯盘比信息,此信息可用于青光眼的初步筛查。

眼底图像以及视杯和视盘标注如图 1 所示。图 1(a)为青光眼患者的眼底图像,图 1(c)为非青光眼患者的眼底图像,图 1(b)和图 1(d)分别为对图 1(a)和图 1(c)进行视杯、视盘标注后的图像,其中绿圈内为视盘,黄圈内为视杯。对比非青光眼患者的眼底图像可以发现,青光眼的视杯明显较大,且两者的视盘大小相近。眼底图像作为一种经济有效的眼底疾病筛查工具,可以显示除青光眼外的多种眼底疾病的症状。

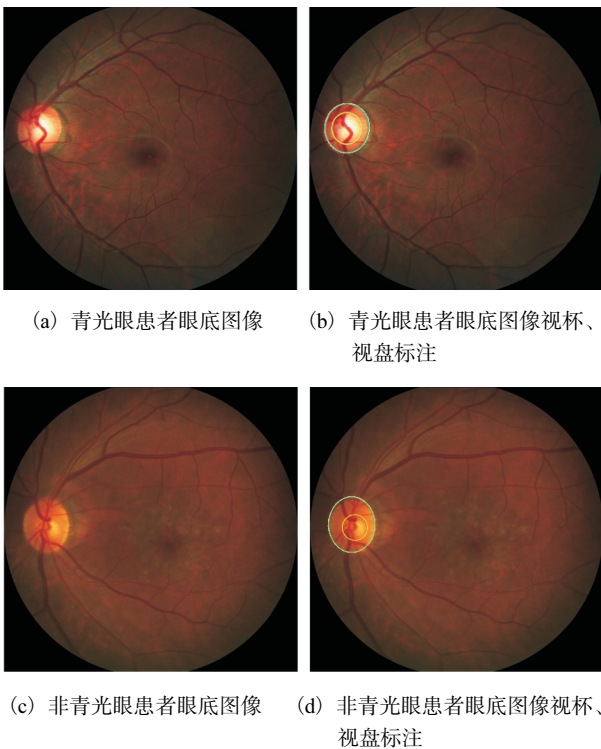


图 1 眼底图像以及视杯和视盘标注

Fig. 1 Fundus images and cup and disc labeling

由于青光眼的隐性患者较多,且眼科医生匮乏,所以使用深度学习检测青光眼成为一种主流的解决方案。为了得到眼底图像的杯盘比,需要将眼底图像的视杯和视盘分割出来,进而得到杯盘比。目前,大部分基于深度学习的医学图像分割算法是基于 U-Net 网络改进的^[3],与卷积神经网络(convolutional neural networks, CNN)相比,U-Net 网络更加适合医

学图像分割^[4]。为了得到精确的 CDR, Yu 等^[5]提出一种改进的 U-Net 网络模型,在下采样编码器部分采用 ResNet-34 模型,在视杯、视盘的分割中获得了 80% 以上的准确率。Juneja 等^[6]提出一种 G-Net 网络用来分割视盘和视杯,该网络使用了两个 U-Net 分别分割视杯、视盘,在视盘和视杯的分割上分别获得 95.8% 和 93.0% 的准确率。Wu 等^[7]在 U-Net 的基础上添加了维度交互的自注意力机制(DI-UNet),该方法将迁移学习和 U-Net 网络相结合,能够更好地提取图像的全局信息。为了能够更好地提取图像特征,以上方法都改变了 U-Net 的网络结构,并取得了较高的准确率。本文提出一种基于多任务学习的青光眼智能诊断系统,网络模型没有改变 U-Net 的基本网络结构,而是在原有的 U-Net 网络结构的基础上添加了 VGG16 网络,在对视杯、视盘分割的同时实现眼底图像的青光眼分类,分类和分割任务同时进行并且共享图像的特征,形成多任务学习,对眼底图像进行联合分类。

1 预备知识

1.1 U-Net 网络

U-Net 网络是为了解决生物医学图像分割问题而产生的^[8]。U-Net 网络结构如图 2 所示,其结构中包括编码器-解码器(encoder-decoder)。编码器部分为特征提取网络,由卷积和下采样组成;解码器部分为特征融合网络,由卷积和上采样组成。在解码器部分,还包括了跳跃连接,即将编码过程中产生的特征图连接到对应大小相同的解码部分,采用拼接操作对特征进行合并。

在 U-Net 网络中,采用 3×3 的卷积核进行特征提取,卷积的步长为 1,并且在每一层卷积完成后,运用特征融合将编码器生成的特征图融合到对应大小的解码器中。编码器的池化层用于降低图像的维度,而解码器使用上采样恢复图像维度,无论是编码器还是解码器,在改变图像维度时都使用 2×2 的卷积核,步长为 2。

1.2 VGG 网络

VGG 网络是一种传统的卷积神经网络^[9],在图像分割和图像分类中被大量应用。VGG16 网络结构如图 3 所示。

VGG16 网络模型由 13 个卷积层、3 个全连接层以及最大池化层组成。所有卷积层的过滤器都为 3×3 的卷积核且步长为 1,池化层的过滤器都为 2×2 的卷积核且步长为 2。输入层采用 $224 \times 224 \times 3$ 的图

像, 前两层都是卷积层, 采用 64 个过滤器进行 same 卷积, 目标输出为 $224 \times 224 \times 64$, 然后经过一个最大池化层, 使图像高度和宽度减半, 即输出为 $112 \times 112 \times 64$ 。VGG16 网络的 14 层和 15 层为全连接层, 包含了 4 096 个神经元。最后一层为输出层, 也就是 Softmax 层, 包含了 1 000 个神经元, 用于网络的分类

输出。

随着 VGG16 网络深度的增加, 卷积层的过滤器数量也在成倍增加, 而图像的尺寸在经过最大池化层后等比例减小。VGG16 网络和 U-Net 网络的下采样部分非常相似, 于是本文将 U-Net 网络和 VGG16 网络结合, 形成一种新的多任务学习。

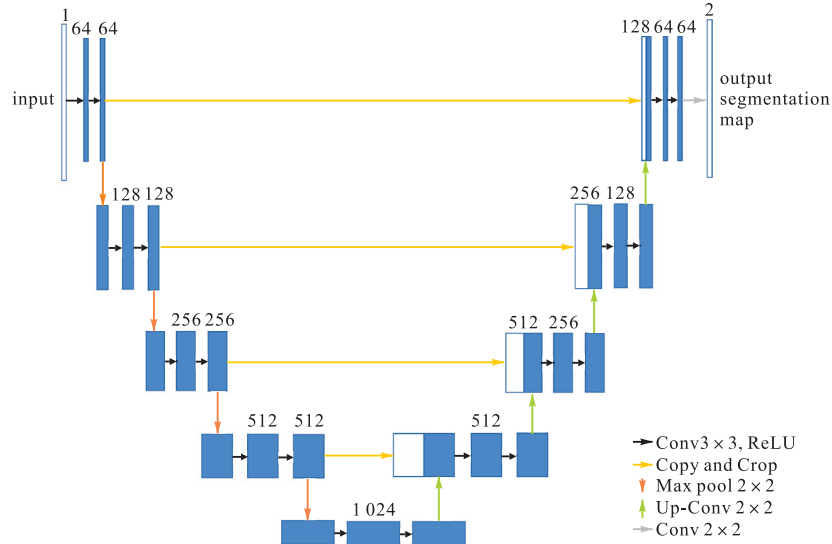


图 2 U-Net 网络结构

Fig. 2 Network structure of U-Net

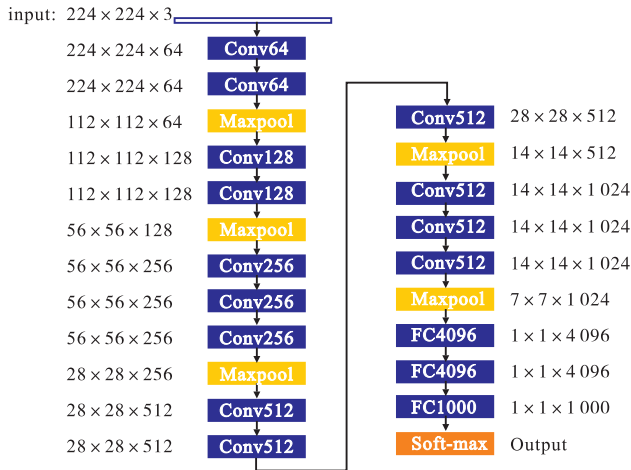


图 3 VGG16 网络结构

Fig. 3 Network structure of VGG16

2 网络模型

2.1 多任务学习

多任务学习 (multi-task learning, MTL) 是迁移学习的一个分支, 多任务学习通过挖掘不同任务之间的关系, 能得到额外的有用信息, 进而提高学习任务性能^[10]。多任务学习平等地处理多个任务, 当新任务开

始训练时, 新任务和旧任务将会同时进行训练, 这种学习方式需要大量计算, 耗费资源。

基于深度学习的多任务学习方法主要包括软参数共享方法和硬参数共享方法。本文采用硬参数共享方法, 即多个任务之间共享网络的相同隐藏层, 在网络靠近输出的部分开始分支完成不同的任务。多任务学习的网络结构如图 4 所示。

在 U-Net 网络提取眼底图像特征的同时, VGG16 网络共享 U-Net 网络的编码器部分。编码器提取的图像信息分别使用跳跃连接和全连接层输入 U-Net 网络的解码器和 VGG16 网络中, U-Net 网络实现图像视杯和视盘的分割, 进而得到 CDR; 将 CDR 作为图像的特征之一输入到 VGG16 网络的全连接层, 进而得到图像分类结果。

在本文方法中, 分类和分割同时训练, 利用图像级和像素级标签提供互补反馈。青光眼的诊断是一个二分类任务, 由 VGG16 网络完成, 训练采用交叉熵作为损失函数, 分类任务的损失定义为

$$L_c = -y_c \log(p_c) - (1 - y_c) \log(1 - p_c) \quad (1)$$

式中: L_c 为 VGG16 网络的损失, p_c 为 VGG16 网络的预测值, y_c 为对应的真实值。

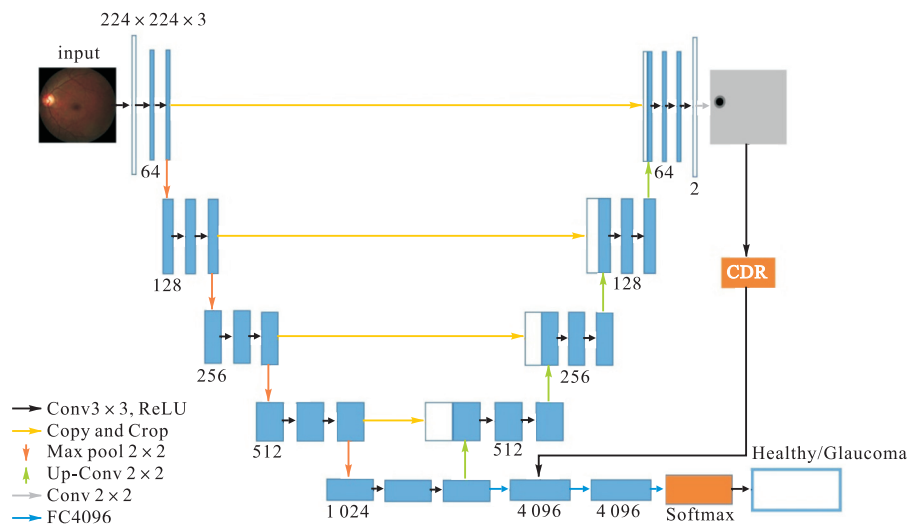


图 4 多任务学习的网络结构
Fig. 4 Network structure of multi-task learning

分割任务由 U-Net 网络完成,训练也采用交叉熵函数作为损失函数,分割任务的损失定义为

$$L_s = -\frac{1}{M} \sum_m \left[y_s^m \log(p_s^m) + (1 - y_s^m) \log(1 - p_s^m) \right] \quad (2)$$

式中: L_s 为 U-Net 网络的损失, p_s 为 U-Net 网络的预测值, y_s 为对应的真实值, M 表示图像中像素的数量。

为了避免过拟合,在分类网络中引入了 Dropout 正则化,它可以使所有隐藏层的节点在每一次迭代中都以一定的概率消失;通过避免在传播过程中对某些节点的强依赖,使反向传递的修正值能够比较均匀地分配在所有参数上。Dropout 设置为 0.3。

2.2 青光眼分类

在进行网络模型训练时,U-Net 网络得到了眼底图像视杯和视盘的分割图,进而得到 CDR,然后将 CDR 输入 VGG16 网络的全连接层,最后得到分类结果。CDR 输入之前 VGG16 网络全连接层输出为

$$y = f \left(\sum_{i=1}^n \omega_i x_i + b \right) \quad (3)$$

式中: y 为当前层的输出, f 为激活函数, ω_i 为第 i 个输入神经元的权重, x_i 为第 i 个输入神经元的输出, b 为偏置项, n 为前一层的神经元数量。

当 CDR 作为输入接入全连接层后,全连接层输出为

$$y = f \left(\sum_{i=1}^n \omega_i x_i + \eta C_p + b C_p \right) \quad (4)$$

式中: C_p 为 CDR 的网络预测值, η 为 CDR 在全连接层的权重。

全连接层将网络提取的特征信息映射到对应的标签上,输出维度为 2,即判定是否为青光眼。

3 实验

3.1 数据集

采用 REFUGE 挑战数据集对本方法进行实验验证。REFUGE 挑战数据集由 1 200 张带注释的视网膜图像组成,其中 121 张来自青光眼患者。该数据集中的视网膜图像以黄斑为中心,图片大小为 1 634 像素 × 1 634 像素或 2 124 像素 × 2 056 像素。数据集的默认分割为:400 张图像用于测试,800 张图像用于训练和验证。

3.2 评价指标

为了更好地评价网络性能,本文采用以下指标:受试者工作特性(receiver operating characteristic, ROC)曲线、工作特性曲线下面积(area under the curve, AUC)、视杯分割准确率(cup segmentation accuracy, CSA)、视盘分割准确率(disc segmentation accuracy, DSA)和杯盘比误差(γ)。ROC 曲线的横纵坐标分别是假阳性率和真阳性率。AUC 作为二分类模型评价指标,值在 0 ~ 1 之间,值越大表示分类越准确。

$$R_{TP} = \frac{N_{TP}}{N_{TP} + N_{FN}} \quad (5)$$

$$R_{FP} = \frac{N_{FP}}{N_{TN} + N_{FP}} \quad (6)$$

$$\gamma = |C_p - C_t| \quad (7)$$

式中: R_{TP} 为真阳性率, R_{FP} 为假阳性率, N_{TN} 为真阴性样本的数量, N_{TP} 为真阳性样本的数量, N_{FN} 为假阴性样本的数量, N_{FP} 为假阳性样本的数量, C_1 为 CDR 的真实值。

3.3 实验过程与结果

3.3.1 预处理

由于本网络模型的输入图像大小为 224 像素 \times 224 像素, 而数据集中图像有 1 634 像素 \times 1 634 像素和 2 124 像素 \times 2 056 像素两种大小, 所以需要数据集图像进行缩放。数据集的默认分割为训练集、测试集、验证集中图像数量相同, 但这并不是最合适的划分, 实验前将数据集重新划分, 训练集、测试集、验证集中图像数量的比值为 6 : 2 : 2。由于数据集中的青光眼图像占比较少(约占总数的 1/10), 所以需要对数据集进行数据增强, 数据增强后的青光眼图像占总数的 3/10。

3.3.2 实验过程

在经过预处理以后, 设置两组实验, 一组为多任务学习 (MTL), 即 U-Net 网络与 VGG16 网络同时训练; 另一组为单任务学习 (single task learning, STL), 即 U-Net 网络先训练, 得到眼底图像的 CDR, 然后 VGG16 网络进行训练, 对图像进行分类。两组实验都选择 Adam 优化器, 学习率设置为 0.000 1, batch_size 设置为 16, epoch 设置为 300。

3.3.3 实验结果

青光眼分类任务的 ROC 曲线如图 5 所示, 使用相同数据集的其他方法的结果与本实验的结果对比见表 2, 基于 MTL 和 STL 的眼底图像分割结果如图 6 所示。

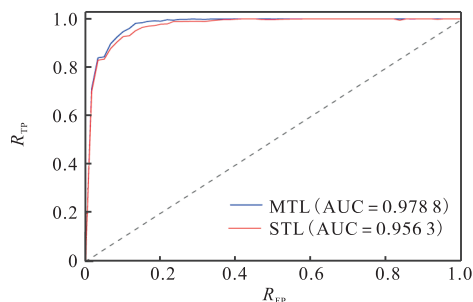


图 5 青光眼分类任务的 ROC 曲线

Fig. 5 ROC curve of glaucoma classification task

由图 5 可知, 本方法的 AUC 为 0.978 8, 而基于 STL 的方法仅为 0.956 3。与其他使用相同数据集的方案相比, 本方法的 AUC 值最高, 同时在视盘和视杯的分割准确率方面分别达到了 0.874 5 和 0.962 4,

杯盘比预测误差仅为 0.030 6, 文献[11]的视杯分割准确率达到 0.882 6, 表现最好。本方法在青光眼分类准确率、视盘分割和杯盘比的误差方面取得了最佳表现。此外, 对比 MTL 和 STL 两种方法, 基于 MTL 的方法在青光眼分类任务上表现更佳, 但如图 6 所示, 两种方法在视杯和视盘分割准确率方面的差异并不显著。

表 2 使用 REFUGE 挑战数据集的结果对比

Tab. 2 Comparison of results of using REFUGE challenge datasets

方法	AUC	CSA	DSA	γ
文献[11]方法	0.964 4	0.882 6	0.960 2	0.045 0
文献[12]方法	0.934 8	0.872 8	0.952 5	0.045 6
STL	0.956 3	0.869 6	0.959 8	0.035 5
MTL(本文)	0.978 8	0.874 5	0.962 4	0.030 6

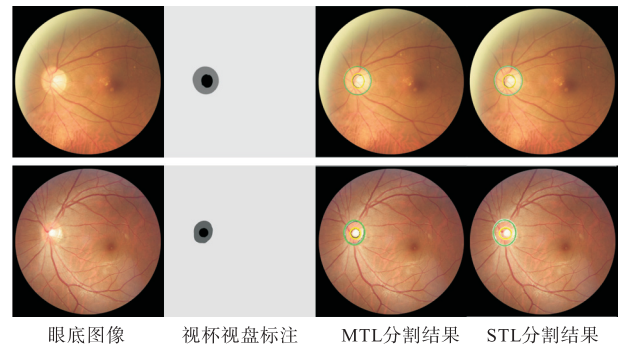


图 6 基于 MTL 和 STL 的眼底图像分割结果

Fig. 6 Fundus image segmentation results of MTL and STL

4 结 语

本文提出一种基于多任务学习的眼底图像青光眼分类模型。该模型由 U-Net 网络和 VGG16 网络组成, VGG16 网络共享 U-Net 网络的编码器部分, U-Net 网络在完成青光眼视杯、视盘的分割后, 将杯盘比信息输入到 VGG16 网络的全连接层, 最后由 VGG16 网络完成对青光眼的分类任务。VGG16 网络共享了 U-Net 网络提取的特征信息, 使分类任务的精度得以提高, 实验结果证明本方法具有较高的青光眼分类准确率。

参考文献:

- [1] COAN L, WILLIAMS B, VENKATESH M K A, et al. Automatic detection of glaucoma via fundus imaging and artificial intelligence: a review[J]. Survey of ophthalmology, 2023, 68(1): 17-41.
- [2] MEDEIROS F A, JAMMAL A A, THOMPSON A C.

- From machine to machine: an OCT-trained deep learning algorithm for objective quantification of glaucomatous damage in fundus photographs[J]. *Ophthalmology*, 2019, 126(4): 513–521.
- [3] LI X L, WANG Y Y, TANG Q S, et al. Dual U-Net for the segmentation of overlapping glioma nuclei[J]. *IEEE Access*, 2019, 7: 84040–84052.
- [4] BACCOUCH W, OUESLATI S, SOLAIMAN B, et al. A comparative study of CNN and U-Net performance for automatic segmentation of medical images: application to cardiac MRI[J]. *Procedia computer science*, 2023, 219: 1089–1096.
- [5] YU S, XIAO D, FROST S, et al. Robust optic disc and cup segmentation with deep learning for glaucoma detection[J]. *Computerized medical imaging and graphics*, 2019, 74: 61–71.
- [6] JUNEJA M, SINGH S, AGARWAL N, et al. Automated detection of Glaucoma using deep learning convolution network (G-net) [J]. *Multimedia tools and applications*, 2020, 79: 15531–15553.
- [7] WU Y L, WANG G L, WANG Z Y, et al. DI-Unet: DImensional interaction self-attention for medical image segmentation[J]. *Biomedical signal processing and control*, 2022, 78: 103896.
- [8] 苏润. 基于 U-Net 框架的医学图像分割若干关键问题研究[D]. 合肥: 中国科学技术大学, 2021.
- [9] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [EB/OL]. (2015–04–10) [2023–02–20]. <https://arxiv.org/abs/1409.1556>.
- [10] 李亚. 多任务学习的研究[D]. 合肥: 中国科学技术大学, 2018.
- [11] WANG S J, YU L Q, YANG X, et al. Patch-based output space adversarial learning for joint optic disc and cup segmentation[J]. *IEEE Transactions on medical imaging*, 2019, 38(11): 2485–2495.
- [12] ORLANDO J I, FU H, BREDA J B, et al. Refuge challenge: a unified framework for evaluating automated methods for glaucoma assessment from fundus photographs[J]. *Medical image analysis*, 2020, 59: 101570.

责任编辑: 周建军

(上接第7页)

- [35] XIE Y J, HILL C A S, XIAO Z F, et al. Silane coupling agents used for natural fiber/polymer composites: a review[J]. *Composites part A: applied science and manufacturing*, 2010, 41(7): 806–819.
- [36] ZANINI M, LAVORATTI A, LAZZARI L K, et al. Producing aerogels from silanized cellulose nanofiber suspension[J]. *Cellulose*, 2017, 24: 769–779.
- [37] SHANG Q Q, CHENG J W, HU L H, et al. Bio-inspired castor oil modified cellulose aerogels for oil recovery and emulsion separation[J]. *Colloids and surfaces A: physicochemical and engineering aspects*, 2022, 636: 128043.
- [38] LAITINEN O, SUOAJÄRVI T, ÖSTERBERG M, et al. Hydrophobic, superabsorbing aerogels from choline chloride-based deep eutectic solvent pretreated and silylated cellulose nanofibrils for selective oil removal[J]. *ACS Applied materials & interfaces*, 2017, 9(29): 25029–25037.
- [39] ZHANG X X, WANG H K, CAI Z Y, et al. Highly compressible and hydrophobic anisotropic aerogels for selective oil/organic solvent absorption[J]. *ACS Sustainable chemistry & engineering*, 2018, 7(1): 332–340.
- [40] HE X H, CHEN T T, JIANG T Y, et al. Preparation and adsorption properties of magnetic hydrophobic cellulose aerogels based on refined fibers[J]. *Carbohydrate polymers*, 2021, 260: 117790.
- [41] GU H B, ZHOU X M, LYU S Y, et al. Magnetic nano-cellulose-magnetite aerogel for easy oil adsorption[J]. *Journal of colloid and interface science*, 2020, 560: 849–856.

责任编辑: 周建军