

文章编号: 2095-2163(2021)05-0163-07

中图分类号: TP391

文献标志码: A

# 基于改进的 AG-CNN 的视网膜 OCT 图像的黄斑病变识别方法

董喜超, 高志军, 董春游

(黑龙江科技大学 计算机与信息工程学院, 哈尔滨 150022)

**摘要:** 本文针对目前应用全局图像训练卷积神经网络可能会受到若干无关噪声区域的影响, 易导致视网膜 OCT 图像黄斑病变识别或诊断错误等问题, 提出了一种改进的注意力引导四分支卷积神经网络的视网膜 OCT 图像黄斑病变识别方法。采用改进注意力引导卷积神经网络框架, 通过集成全局分支、局部分支和层分割分支构成融合分支, 利用注意力热图对重要区域进行掩膜和训练, 减少视网膜 OCT 图像噪声的干扰和黄斑病变识别错误率, 通过与 VGG16 和 IDL 2 种方法在公开数据集上进行了实验验证比较。结果表明, 文中方法在视网膜 OCT 图像数据集上对于识别准确度和识别性能的提升具有显著性的作用。

**关键词:** 视网膜 OCT 图像; 黄斑病变; 改进注意力引导卷积神经网络

## Macular degeneration identification method of the OCT image with retina based on improved AG-CNN

DONG Xichao, GAO Zhijun, DONG Chunyou

(School of Computer and Information Engineering, Heilongjiang University of Science & Technology, Harbin 150022, China)

**[Abstract]** In view of the current application of global image training convolutional neural network may be affected by a number of irrelevant noise regions, which may easily lead to errors in the recognition or diagnosis of macular disease in retinal OCT images, an improved attention-guided four-branch convolutional neural network method was proposed to identify the macular degeneration in retinal OCT images. The attention-guided convolutional neural network framework was used to form fusion branches by integrating global branches, local branches and layer segmentation branches, and attention heat maps were used to mask and train important areas, which reduced the interference of retinal OCT image noise and macular disease recognition error rate, the proposed was compared with the two methods of VGG16 and IDL on the public data set. The results showed that the proposed method had a significant effect on the improvement of recognition accuracy and recognition performance on the retina OCT image data set.

**[Key words]** OCT images; macular degeneration; improved attention-guided convolutional neural network

## 0 引言

脉络膜新生血管(CNV)、糖尿病性黄斑水肿(DME)和玻璃膜疣(DRUSEN)是常见的视网膜黄斑病变。其中,CNV 出现的早期,无症状。随着其逐渐扩大、渗漏和破裂出血,可致视力减退、视物变形,出现黄斑中心或旁中心暗点,症状反复发作,黄斑部受到严重破坏,可造成永久性视力障碍<sup>[1]</sup>; DME 是导致糖尿病患者视力损害的常见原因,严重影响患者的视力功能和生活质量<sup>[2]</sup>; DRUSEN 患者会出现视力障碍,早期黄斑部多有玻璃膜疣侵犯,其视力尚可正常、有时有视物变形及中心视力下降<sup>[3]</sup>。视网膜光学相干断层扫描(OCT)成像具有无接触、无侵入和高分辨率成像特点,通过视网膜 OCT 图像对眼部患者进行自动和准确地识别是辅

助眼科医生诊断和分级黄斑病变的关键手段。然而在临床诊断中,眼科医生需要花费较长时间在 2 维视网膜 OCT 图像上手动识别这些病变,然后对眼部患者做出诊断决定。显然,这种人工分析对眼科医生来说既费时又费力,也容易受眼科医生主观因素影响,对患者产生漏诊或误诊现象。因此,为了有效地辅助眼科医生筛查和诊断视网膜黄斑病变,眼病视网膜 OCT 图像黄斑病变识别方法和分析技术成为近期学者们研究的热点和重点。

2011 年, Liu 等人利用局部二值模式(LBP)特征,对视网膜 OCT 图像的纹理信息和形状信息进行特征编码,随后结合支持向量机(SVM)实现了 OCT 图像的自动分类<sup>[4]</sup>; Srinivasan 等人通过计算 OCT 图像的多尺度方向梯度直方图(HOG)特征,识别了正常视网膜、年龄相关性黄斑变性和黄斑水

**基金项目:** 黑龙江省省属高等学校基本科研业务费科研项目(Hkqdg201911)。

**作者简介:** 董喜超(1994-),男,硕士研究生,主要研究方向:医学图像处理;高志军(1980-),男,博士,讲师,硕士生导师,主要研究方向:医学图像处理、机器学习;董春游(1962-),男,博士,教授,主要研究方向:人工智能。

收稿日期: 2021-02-04

肿<sup>[5]</sup>。

由于全卷积神经网络等深度学习方法能够实现端到端的自动学习,已被广泛有效地应用于医学图像处理中。2017年, Lee 应用深度学习方法有效地对正常和年龄相关性黄斑变性 OCT 图像进行了二分类<sup>[6]</sup>;2017年, Rasti 等人应用多尺度卷积神经网络实现了对黄斑病变 OCT 图像的分类<sup>[7]</sup>;2018年, Rasti 等人又引入了基于小波的 CNN 提取深部小波特征,利用随机森林分类器对黄斑 OCT 图像进行了分类<sup>[8]</sup>;2018年, Fauw 等人采用 U-Net 结构对视网膜 OCT 图像进行分割,将分割后的图像用于视网膜病变的诊断<sup>[9]</sup>;此外, Karri 等人提出了一种新的基于 CNN 的技术,可以用较少的训练 OCT 图像来识别黄斑病变,同时仍然显示出有效的结果<sup>[10]</sup>;2018年 Kermany 等人建立了一个基于深度学习框架(Image-Based Deep Learning, IDL)的诊断工具,用于筛查常见的和可致盲性的视网膜疾病的患者<sup>[11]</sup>。

现有的基于深度学习的视网膜 OCT 图像病变分类或识别方法多采用全局图像作为网络学习的输入,而眼部疾病有些发生在疾病特异性的(小)局部区域。应用全局图像训练卷积神经网络可能会受到若干无关噪声区域的影响,易导致错误的识别或诊断。

本文针对上述问题,受眼科医生阅读视网膜 OCT 图像时“先整体,再局部,最后再结合整体”的诊断思路启发,即先整体观察图像中是否有疑似病理区域,然后再集中精力察看有病理可能的局部区

域,最后在结合整张图像观察,确定病理结论。基于注意力引导卷积神经网络(AG-CNN)模型,在其全局分支、局部分支和融合分支基础上,又引入了层分割分支,提出了一种改进的 AG-CNN 的视网膜 OCT 图像的黄斑病变识别方法,从疾病特定区域学习,以避免噪声干扰,用全局分支和层分割分支来补偿由局部分支丢失的鉴别线索,从而实现对 CNV、DME 和 DRUSEN 等常见的视网膜黄斑病变的自动准确识别。

## 1 方法

### 1.1 AG-CNN 模型

2018年, Zhong 等人提出了 AG-CNN 算法,其模型框架如图1红色框所示,展示了一个以 ResNet-50 网络为主干的框架图,AG-CNN 由全局分支、局部分支和融合分支等3个分支组成。其中,全局分支和局部分支均分别由5个卷积块组成,具有批处理规范化和激活函数 ReLU,每一个都连接到一个最大池化层(Pool5)、一个完全连接(FC)层和一个激活函数 ReLU。除此以外,局部分支的输入是由全局分支生成的掩膜裁剪出的局部病变图片,将这两个分支连接到融合分支中,“BCE”表示二进制交叉熵损失。输入图像将添加到热图中进行可视化。

AG-CNN 利用注意力引导的掩膜推理过程定位显著区域,并学习判别特征进行识别。与依赖于边界框标注的方法相比,该方法只需要图像级的标签,而不需要任何额外的信息。

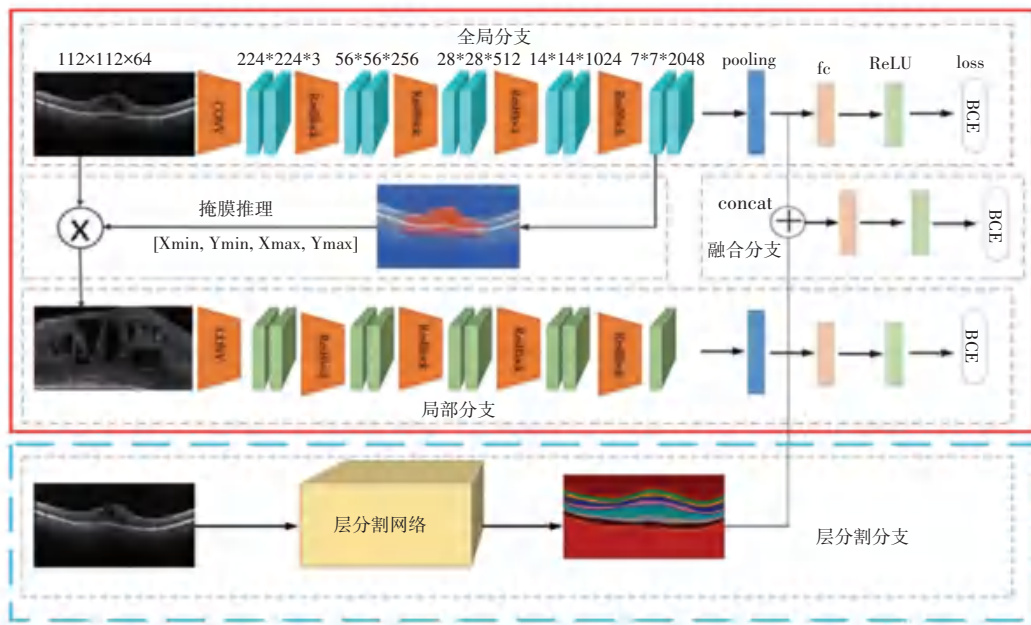


图1 改进的注意力引导卷积神经网络的总体框架

Fig. 1 Overall framework of the improved AG-CNN

用5维向量  $L = [l_1, l_2, \dots, l_c]$  的多维标签标记每个图像,其中  $l_c \in \{0, 1\}$ , 表示是否有病理,即1代表存在,0代表缺席,  $C = 5$ ,  $L$  的最后一个元素表示带有“*No Finding*”的标签。

### 1.1.1 全局分支

全局OCT图像作为全局分支的输入,在全局分支中训练ResNet-50的变体被作为主干模型<sup>[12]</sup>。其由5个下采样块组成,接着是一个全局最大池化层和一个用于分类的5维全连接(FC)层。最后,通过增加一个激活函数ReLU层对FC层的输出向量  $\bar{P}_g(c|I)$  进行非线性化和归一化处理,式(1):

$$\bar{P}_g(c|I) = \text{MAX}(0, P(c|I)), \quad (1)$$

式中,  $I$  表示全局图像;  $\bar{P}_g(c|I)$  表示  $I$  属于第  $c$  类的概率分数;  $c \in \{1, 2, \dots, C\}$ 。最小化二进制交叉熵(BCE)损失函数被用来优化全局分支的参数  $W_g$ , 式(2):

$$\zeta(W_g) = -\frac{1}{C} \sum_{c=1}^C l_c \log(\bar{P}_g(c|I)) + (1 - l_c) \log(1 - \bar{P}_g(c|I)). \quad (2)$$

式中:  $l_c$  表示第  $c$  类的真实标签,  $C$  表示病理数量。

### 1.1.2 局部分支

局部分支具有与全局分支相同的卷积网络结构。但这2个分支不共享权重,因为二者有不同的用途。 $\bar{P}_l(c|I_c)$  被用来表示局部分支的概率得分,  $W_l$  表示局部分支的参数。其中,  $I_c$  是本地分支机构的输入图像,执行与全局分支相同的规范化和优化。

### 1.1.3 融合分支

融合分支首先连接全局分支和局部分支,再由级联层连接到5维FC层以进行最终分类,概率得分  $\bar{P}_f(c|[I, I_c])$ ,  $W_f$  表示融合分支的参数,并用公式(2)对  $W_f$  进行优化。

注意力引导掩膜推理,一个二值掩膜被构造用来定位全局图像中用于识别的判别区域。通过对特征图进行阈值化操作而产生的,可以看作是一个注意力过程。该过程如下:给定一个全局图像,让  $f_g^k(x, y)$  表示最后一个卷积层输出的第  $k$  个信道中空间位置  $(x, y)$  的激活,其中  $k \in \{1, \dots, K\}$ ,  $K = 2048$ ;  $g$  表示全局分支。首先,在位置  $(x, y)$  取激活值  $f_g^k(x, y)$  的绝对值;其次,通过计算通道上的最大值生成注意力热图  $H_g$ , 式(3):

$$H_g(x, y) = \max(|f_g^k(x, y)|), k \in (1, \dots, K), \quad (3)$$

$H_g$  中的值直接表示识别激活的重要性,观察到图像的识别区域(病变区域)被激活。热图可以通过计算通道尺寸上的不同统计值来构建,如  $L_1$  距离为  $\frac{1}{K} \sum_{k=1}^K |f_g^k(x, y)|$  或  $L_2$  距离为  $\frac{1}{K} \sqrt{\sum_{k=1}^K (f_g^k(x, y))^2}$ 。

一个二进制掩膜  $M$  被设计来定位具有大激活值的区域。若热图中某个空间位置  $(x, y)$  的值大于阈值  $\tau$ , 则掩膜中相应位置的值被赋值为1, 否则为0, 式(4):

$$M_{(x,y)} = \begin{cases} 1, & H_g(x, y) > \tau, \\ 0, & \text{否则.} \end{cases} \quad (4)$$

其中:  $\tau$  表示控制参与区域大小的阈值。

$\tau$  越大,区域越小,反之亦然。利用掩膜  $M$ , 一个覆盖  $M$  中判别点的最大连通区域被画出,最大连通区域表示为水平和垂直轴上的最小和最大坐标  $[X_{min}, Y_{min}, X_{max}, Y_{max}]$ 。最后,从输入图像  $I$  中裁剪出局部判别区域  $I_c$ , 并将其调整到与  $I$  相同的大小。

## 1.2 改进的AG-CNN的视网膜OCT图像的黄斑病变识别模型

改进的AG-CNN如图1所示,由于CNV、DME和DRUSEN等常见的视网膜黄斑病变会导致患者视网膜层及其内层部分区域发生形变和厚度变化,故在原AG-CNN(红色框)的基础上引入层分割分支(蓝色框),并连接到融合分支,减少图像噪声对视网膜黄斑病变的错误识别。

新引入的层分割分支采用RelayNet算法实现,RelayNet使用卷积块的收缩路径来学习上下文特征的层次结构,然后使用卷积块的扩展路径来进行图像像素级的语义分割<sup>[13]</sup>。此外,引入将编码器输出中间到匹配解码器输入的跳过连接,以恢复在下采样期间丢失的空间信息。将视网膜OCT图像分割成7层,获得各层厚度信息。

改进的AG-CNN的视网膜OCT图像病变识别算法:

**步骤1** 输入: 输入图像  $I$ , 标签向量  $L$ , 阈值  $\tau$ ;

**步骤2** 初始化: 全局和局部分支权重;

**步骤3** 用  $I$  学习  $W_g$ , 计算  $\bar{P}_g(c|I)$ , 用公式(2)进行优化;

**步骤4** 计算掩膜  $M$  和边界框坐标  $[X_{min}, Y_{min}, X_{max}, Y_{max}]$ , 从  $I$  中裁剪出  $I_c$ ;

**步骤5** 使用  $I_c$  学习  $W_l$ , 计算  $\bar{P}_l(c|I_c)$ , 通过公式(2)进行优化;

**步骤6** 集中  $pool_g$  和  $pool_l$  学习  $W_f$ , 计算  $\bar{P}_f(c|$

$I, I_c$ ), 通过公式(2)进行优化;

步骤 7 输出:输出概率得分  $\bar{P}_f(c | I, I_c)$ 。

## 2 实验及结果

### 2.1 实验数据集、设置、训练过程和评价指标

#### 2.1.1 实验数据集

完整的 ImageNet 数据集有 1 400 多万幅图片, 涵盖 20 000 多个类别标注与超过百万的边界的框标注, 每个类别有 500 到 1 000 幅图片。

UCSD 数据集由 84 484 幅 OCT B 扫描组成。其中有 8 866 幅患有 DRUSEN, 37 455 幅患有 CNV, 11 598 幅患有 DME 和 26 565 幅正常的 OCT 图像。本文用 250 幅正常的、250 幅患有 DRUSEN、250 幅患有 CNV 和 250 幅患有 DME 的 OCT 图像作为测试集。

#### 2.1.2 实验设置

本文采用 pytorch 框架来实现改进的 AG-CNN 算法, 并在 GTX1050Ti GPU 上进行训练和验证。在动量为 0.9 的情况下, 采用随机梯度下降进行训练, 初始学习率设置为 0.01, 初始学习率随指数衰减而降低, 幂次为 0.9, 深度模型训练时间为 26 h。

#### 2.1.3 训练过程

第一阶段, 利用 ImageNet 预训练的全局分支网络进行微调,  $\bar{P}_g(c | I)$  由公式(1)归一化;

第二阶段, 当用阈值  $\tau$  进行掩模推理得到局部图像后, 将其送入局部分支进行微调,  $\bar{P}_l(c | I_c)$  也可通过公式(1)归一化, 微调局部分支时全局分支中的权重是固定的;

第三阶段, 分别让  $Pool_g$  和  $Pool_l$  代表全局分支和局部分支的 Pool5 层输出, 将其串联起来进行最后的微调, 并用公式(1)将概率分数  $\bar{P}_l(c | [I, I_c])$  归一化, 同样当微调融合分支的权值时, 前 2 个分支的权重是固定的。

#### 2.1.4 评估指标

本文采用准确度 (ACC)、灵敏度 (SE)、精确度 (PR)、特异性 (SP) 和混淆矩阵等评估指标评估算法的识别性能, 每一个都是针对独立类计算的。对于独立类的灵敏度是预测精度, 而对于每个类别标签, 其特异度的定义是相同的, 其中负样本是不在考虑类别中的样本。由于样本在不同类别之间的不平衡, 总体灵敏度 (OS)、总体精度 (OP) 和总体准确度 (OA) 等评估指标也被计算用于评估算法的识别性

能, 式(5) ~ 式(8):

$$ACC = \frac{n_{TP} + n_{TN}}{n_{TP} + n_{TN} + n_{FP} + n_{FN}} \times 100\% \quad (5)$$

式中:  $n_{TP}$  —— 将正样本识别为正样本数;

$n_{TN}$  —— 将负样本识别为负样本数;

$n_{FP}$  —— 将负样本识别为正样本数;

$n_{FN}$  —— 将正样本识别为负样本数。

$$SE = \frac{n_{TP}}{n_{TP} + n_{FN}} \times 100\%, \quad (6)$$

$$SP = \frac{n_{TN}}{n_{TN} + n_{FP}} \times 100\%, \quad (7)$$

$$PR = \frac{n_{TP}}{n_{TP} + n_{FP}} \times 100\%. \quad (8)$$

### 2.2 实验结果

VGG16 和 IDL 2 种方法与本文方法在 UCSD 数据集上的定量评估结果见表 1。首先, 在最后 3 列的 OS、OP 和 OA 等 3 个总体量化评估指标上, 改进 AG-CNN 分别为 97%、97% 和 98.5%, 比 IDL 分别略高 0.9%、0.9% 和 0.4%, 比 VGG16 显著性提高 20.9%、20.8% 和 7.1%; 其次, 具体比较其中较好的识别方法 IDL, 在正常类别中, 准确率提高 0.3%、灵敏度提高 0.7%、精确度提高 0.4% 和特异性提高 0.2%; 在 DME 类别中, 准确率提高 0.5%、灵敏度提高 0.9%、精确度提高 1.2% 和特异性提高 0.4%; 在 DRUSEN 类别中, 准确率提高 0.5%、灵敏度提高 0.9%、精确度提高 0.2% 和特异性提高 0.4%; 在 CNV 类别中, 准确率提高 0.5%、灵敏度提高 1.1%、精确度提高 0.8% 和特异性提高 0.3%。综上, 本文方法在 UCSD 数据集上的识别评估指标均有较高的提升, 表明文中的改进算法在视网膜 OCT 黄斑病变识别上有更好的识别性能。

表 1 3 种方法识别结果(百分数)

Tab. 1 The identification results with three methods (percentages)

方法	类别	ACC	SE	PR	SP	OS	OP	OA
VGG16	NORMAL	91.4	92.3	83.1	91.2			
	DME	92.5	70.9	74.3	96.3			
	DRUSEN	90.5	54.5	54.4	94.6			
	CNV	91.0	86.5	93.1	94.5	76.1	76.2	91.4
IDL	NORMAL	98.8	96.9	98.4	99.4			
	DME	98.0	97.1	94.8	98.2			
	DRUSEN	97.8	96.7	94.4	98.1			
	CNV	97.6	93.8	96.8	98.9	96.1	96.1	98.1
文中方法	NORMAL	99.1	97.6	98.8	99.6			
	DME	98.5	98.0	96.0	98.6			
	DRUSEN	98.3	97.6	95.6	98.5			
	CNV	98.1	94.9	97.6	99.2	97.0	97.0	98.5

为了研究有无训练过程对第一阶段实验结果的影响,本文进行了对比实验来验证训练过程第一阶段的性能表现,见表 2。比较实验评估结果,在最后 3 列的 *OS*、*OP* 和 *OA* 等总体量化评估指标上,有训练过程第一阶段实验结果分别为 97.0%、97.0% 和 98.5%,比无训练过程第一阶段实验结果分别提高 2.9%、2.9% 和 1.5%。具体来说,在正常类别中,准确率、灵敏度、精确度和特异性分别提高 1.6%、3.1%、2.8% 和 1.0%;在 DME 类别中,准确率、灵敏度、精确度和特异性提高 1.3%、1.8%、3.8%、1.4%;在 DRUSEN 类别中,准确率、灵敏度、精确度和特异性提高 2.1%、4.5%、3.6%、1.2%;在 CNV 类别中,准确率、灵敏度、精确度和特异性提高 1.1%、2.3%、2.0%、0.7%。综上,表明训练过程第一阶段对 OCT 图像识别效果提升起到显著性作用。

VGG16、IDL 和文中方法识别结果的混淆矩阵如图 2 所示。对角线(蓝色)是正确识别的个数,非对角线为错分到相应类别的个数。相比较而言,在正常人眼上,本文方法的正确识别的个数为 247,比 VGG16 和 IDL 分别多 39 个和 1 个;在 DME 上,本

文方法的正确识别的个数为 240,比 VGG16 和 IDL 分别多 54 个和 3 个;在 DRUSEN 上,本文方法的正确识别的个数为 239,比 VGG16 和 IDL 分别多 103 个和 3 个;在 CNV 上,本文方法的正确识别的个数为 244,比 VGG16 和 IDL 分别多 14 个和 2 个。可见本文方法在该 OCT 图像验证集上每类的正确识别均有所提升,具有更佳的识别效果。

表 2 有无训练过程第一阶段实验对比结果(百分数)

Tab.2 Comparative results of the first stage experiment with and without training process (percentages)

方法	类别	ACC	SE	PR	SP	OS	OP	OA
无	NORMAL	97.5	94.5	96.0	98.6			
	DME	97.2	96.2	92.8	97.5			
	DRUSEN	96.2	93.1	92.0	97.3			
	CNV	97.0	92.6	95.6	98.5	94.1	94.1	97.0
有	NORMAL	99.1	97.6	98.8	99.6			
	DME	98.5	98.0	96.6	98.9			
	DRUSEN	98.3	97.6	95.6	98.5			
	CNV	98.1	94.9	97.6	99.2	97.0	97.0	98.5

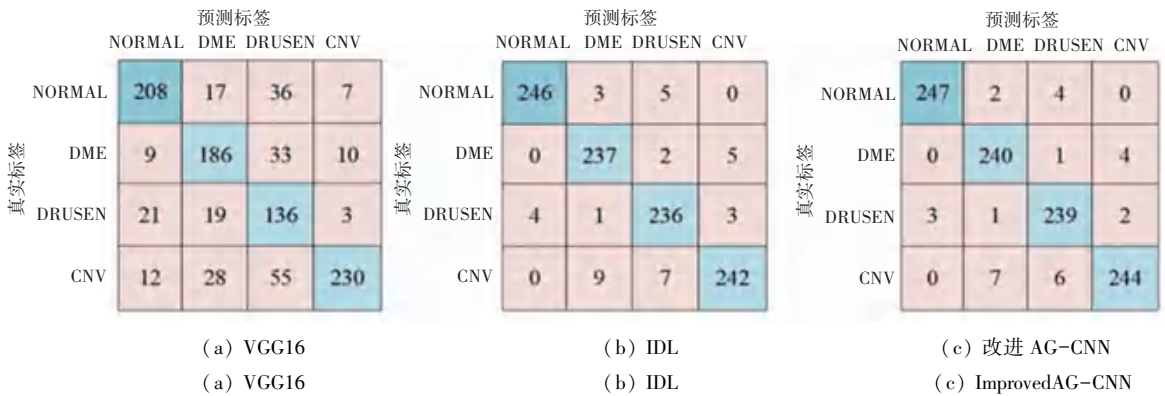


图 2 不同方法的混淆矩阵

Fig. 2 Confusion matrix of different methods

为了研究每个分支对实验结果的影响,本文进行了消融实验来验证全局分支、局部分支和文中新引入的层分割分支的性能表现,见表 3。基线 1 只有全局分支,基线 2 有全局分支、局部分支和融合分支,改进 AG-CNN 包含所有 4 个分支。比较基线 1 和基线 2 的结果,在最后 3 列的 *OS*、*OP* 和 *OA* 等总体量化评估指标上,基线 2 实验结果分别为 96.0%、96.2% 和 97.6%,比基线 1 实验结果分别提高 0.8%、1.1% 和 1.3%;在正常类别中,准确率、灵敏度、精确度和特异性提高 1.2%、1.1%、0.8%、0.6%;在 DME 类别中,准确率、灵敏度、精确度和特异性提高

1.6%、1.9%、1.0%、0.7%;在 DRUSEN 类别中,准确率、灵敏度、精确度和特异性提高 1.3%、1.0%、0.4%、0.7%;在 CNV 类别中,准确率、灵敏度、精确度和特异性提高 1.2%、1.0%、2.4%、1.8%。综上说明局部分支对 OCT 图像识别效果提升有显著性作用。比较基线 2 和改进 AG-CNN 的结果,在最后 3 列的 *OS*、*OP* 和 *OA* 等总体量化评估指标上,改进 AG-CNN 实验结果分别为 97.0%、97.0% 和 98.5%,比基线 2 实验结果分别提高 1.0%、0.8% 和 0.9%;在正常类别中,准确率、灵敏度、精确度和特异性提高 0.3%、0.3%、0.8%、0.4%;在 DME 类别中,准确率、灵

敏度、精确度和特异性提高 1.1%、1.6%、0.8%、0.4%；在 DRUSEN 类别中，准确率、灵敏度、精确度和特异性提高 1.0%、1.1%、0.6%、0.2%；在 CNV 类别中，准确率、灵敏度、精确度和特异性提高 1.1%、1.3%、1.0%、0.4%。综上所述层分割分支对 OCT 图像识别提升有进一步的作用。综上所述，全局分支、局部分支和层分割分支对提升 OCT 图像识别力均具有显著的作用。

VGG16、IDL 和文中方法在正常的、DRUSEN、CNV 和 DME 的测试集上的 ROC 曲线如图 3 所示。在所有类别里，文中方法（红色线）在 VGG16（蓝色点虚线）和 IDL（绿色虚线）上，围成的面积 AUC 更大，表明文中方法比 VGG16 和 IDL 在视网膜 OCT 图像识别力更强，和算法的识别性能更优。其中，本文方法在与较好的 IDL 方法相比，在正常类别中 AUC 提高 5.1%，在糖尿病性黄斑水肿类别中 AUC 提高 4.6%，在玻璃膜疣类别中 AUC 提高 5.8%，在脉

络膜新生血管类别中 AUC 提高 3.4%。

表 3 消融实验结果(百分数)

Tab. 3 Ablation results (percentages)

方法	类别	ACC	SE	PR	SP	OS	OP	OA
基线 1	NORMAL	97.6	96.2	97.2	98.6			
	DME	95.8	94.5	94.2	97.5			
	DRUSEN	96.0	95.5	94.6	97.6			
	CNV	95.8	92.6	94.2	97.0	95.2	95.1	96.3
基线 2	NORMAL	98.8	97.3	98.0	99.2			
	DME	97.4	96.4	95.2	98.2			
	DRUSEN	97.3	96.5	95.0	98.3			
	CNV	97.0	93.6	96.6	98.8	96.0	96.2	97.6
改进 AG-CNN	NORMAL	99.1	97.6	98.8	99.6			
	DME	98.5	98.0	96.0	98.6			
	DRUSEN	98.3	97.6	95.6	98.5			
	CNV	98.1	94.9	97.6	99.2	97.0	97.0	98.5

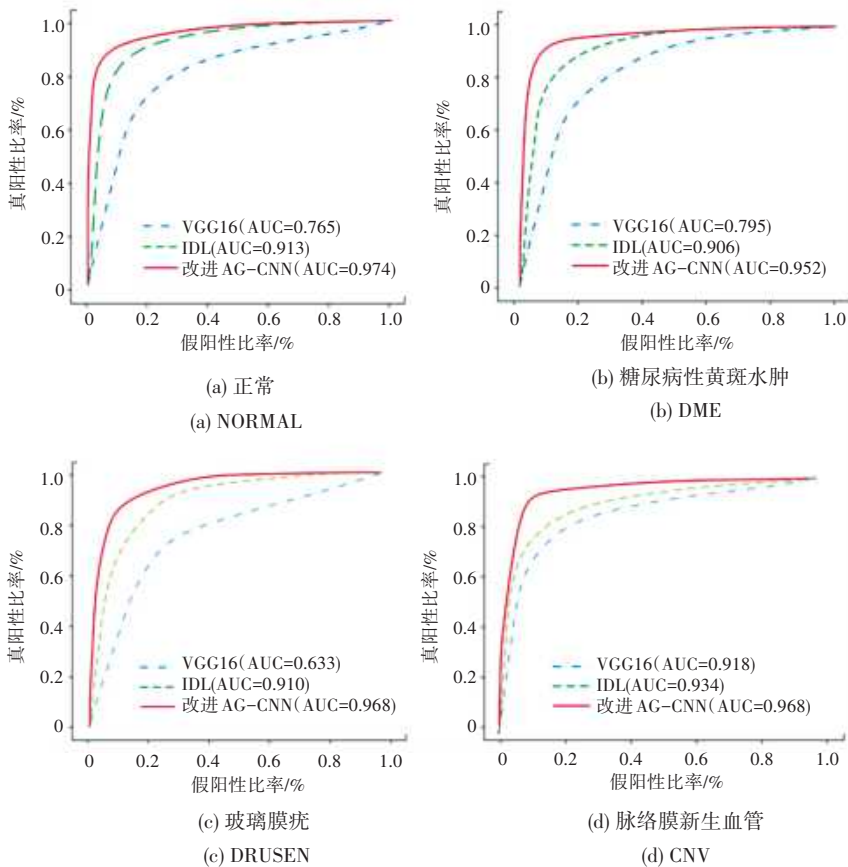


图 3 各类别的 ROC 曲线

Fig. 3 ROC curve of each category

尽管本文方法在公开数据集上的大部分 B 扫描图像上取得了比较令人满意的识别结果，但是对于一些 B 扫描图像，由于原 OCT 图像模糊或病变特征不显著，存在一些错误识别样本，如图 4 所示。图

4(a) 为将正常样本错分为 DRUSEN 病变样本；图 4(b) 为将 DME 病变样本错分为正常样本；图 4(c) 为 DRUSEN 病变样本错分为正常样本；图 4(d) 为 CNV 病变样本错分为 DRUSEN 病变样本。

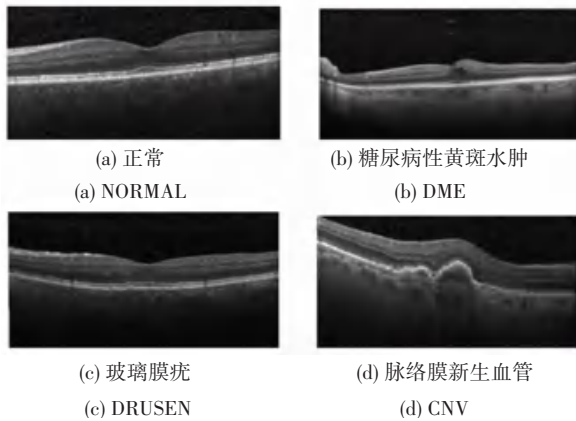


图4 识别错误部分样本

Fig. 4 Partial sample of identification error

### 3 结束语

为了避免全局图像训练卷积神经网络可能会受到若干无关噪声区域的影响,易导致错误的识别或诊断等问题,本文提出了一个注意力引导的4分支卷积神经网络方法,用于眼部黄斑疾病的识别。该网络通过考虑全局分支、局部分支、层分割分支中的全局、局部和各层厚度信息进行训练,与以往单纯依赖全局信息的方法不同,利用注意力热图对重要区域进行掩膜,用于训练局部分支。实验结果表明,本文提出的方法在视网膜OCT图像数据集上对于识别准确度和识别性能的提升具有显著的作用。当然,识别性能还需要进一步提高,今后可将病变分割分支引入到本文模型,提高识别的准确率和识别性能。

### 参考文献

- [1] 赵玥,刘剑,姚进. 光学相干断层扫描血管成像在脉络膜裂伤中的诊断价值[J]. 国际眼科杂志, 2020, 20(2):398-400.
- [2] 王卓实,林小琳,王欢,等. 黄斑水肿的微视野评估与视力的相关性研究[J]. 国际眼科杂志, 2020,20(10):124-150.

(上接第162页)

### 参考文献

- [1] 王义强,闰国琛,王晓军,等. 高速铣削工件表面粗糙度的预测[J]. 机械设计与制造,2014(11):131-133,137.
- [2] 张宝磊,李聪,覃金昌,等. 基于BP和GRNN神经网络的高速铣削TC4粗糙度预测[J]. 桂林航天工业学院学报,2015,20(4):502-505.
- [3] 胡博,潘霄,葛维春,等. 辅助弃风消纳的IVMD-LSTM电网典型时段负荷曲线拟合方法[J]. 可再生能源,2020,38(3):366-372.
- [4] 杨帆,周敏,全继民,等. 智能优化算法及人工智能在催化裂化模型分析中的应用进展[J]. 石油学报(石油加工),2020,36(4):878-888.
- [5] 于振,于万成. 一种基于BP神经网络的扭矩传感器静态校准系

- [3] 田莎,赵玥,姚进. 埋藏型视盘玻璃疣光学相干断层扫描血管成像的影像特征及其诊断价值[J]. 眼科学报,2020,35(4):219-227.
- [4] LIU Yuying, CHEN Mei, ISHIKAWA H, et al. Automated macular pathology diagnosis in retinal OCT images using multi-scale spatial pyramid and local binary patterns in texture and shape encoding[J]. Medical Image Analysis,2011,15(5):748-759.
- [5] SRINIVASAN P P, KIM L A, METTU P S, et al. Fully automated detection of diabetic macular edema and dry age-related macular degeneration from optical coherence tomography images[J]. Biomedical Optics Express,2014,5(10):3568-3577.
- [6] LEE C S, BAUGHMAN D M, LEE A Y. Deep Learning Is Effective for Classifying Normal versus Age-Related Macular Degeneration OCT Images[J]. Ophthalmology Retina, 2017:322-327.
- [7] RASTI R, RABBANI H, MEHRIDEHNAVI A, et al. Macular OCT Classification Using a Multi-Scale Convolutional Neural Network Ensemble[J]. IEEE Transactions on Medical Imaging, 2018, 37(4):1024-1034.
- [8] RASTI R, MEHRIDEHNAVI A, RABBANI H, et al. Automatic diagnosis of abnormal macula in retinal optical coherence tomography images using wavelet-based convolutional neural network features and random forests classifier[J]. Journal of Biomedical Optics, 2018, 23(3):1-10.
- [9] FAUW J D, LEDSAM J R, ROMERA-PAREDES B, et al. Clinically applicable deep learning for diagnosis and referral in retinal disease[J]. Nature medicine, 2018, 24(9):566.
- [10] KARRI S P K, CHAKRABORTY D, CHATTERJEE J. Transfer learning based classification of optical coherence tomography images with diabetic macular edema and dry age-related macular degeneration[J]. Biomedical Optics Express, 2017, 8(2):579.
- [11] 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.
- [12] HE K, ZHANG X, REN S, et al. Deep Residual Learning for Image Recognition[C]// IEEE Conference on Computer Vision & Pattern Recognition. IEEE Computer Society, 2016.
- [13] ROY A G, CONJETI S, KARRI S P K, et al. ReLayNet: retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks[J]. Biomedical Optics Express, 2017, 8(8):3627.

统[J]. 传感技术学报,2020,33(2):238-244.

- [6] 陈智雨,陆金桂. 基于ACO-BP神经网络的光伏系统发电功率预测[J]. 机械制造与自动化,2020,49(1):173-175,187.
- [7] 江毓,郑燕萍,张新,等. 基于改进BP神经网络的电机异响诊断[J]. 重庆理工大学学报(自然科学版),2020,34(1):242-246,262.
- [8] 董海龙,汤旻安,程海鹏. 非均衡交通流五岔路口交通信号智能控制研究[J]. 郑州大学学报(工学版),2017,38(1):68-73.
- [9] 胡红萍,白黄琴,白艳萍,等. 基于改进的遗传算法与人工神经网络的类流感的预测[J]. 中北大学学报(自然科学版),2019,40(6):481-488.
- [10] 黄丽华,李俊丽. 基于改进人工鱼群算法优化的BP神经网络预测控制系统[J]. 化工自动化及仪表,2019,46(8):610-614.
- [11] 李文龙,陈燕,赵杨,等. 采用神经网络和遗传算法优化磁粒研磨TC4弯管工艺参数[J]. 表面技术,2020,49(6):330-336.