

# 基于多通道的光学相干层析成像视网膜图像自动分类研究

陈思思, 陈明惠\*, 马文飞

上海理工大学医疗器械与食品学院上海介入医疗器械工程技术研究中心, 教育部医学光学工程中心, 上海 200093

**摘要** 光学相干层析成像(OCT)是获取眼部图像的主要技术手段之一。常见的眼科临床症状有软性玻璃疣, 糖尿病性黄斑水肿和脉络膜新生血管性疾病。根据 OCT 图像的特殊性, 通过利用神经网络对上述常见病变图像进行自动分类识别, 提出了基于 OCT 视网膜病变图像自动分类的轻量级卷积神经网络—GM-OCTnet, 其具有多通道、多尺度和相对轻量化等特点, 并且能实现高精度、低误诊或少漏诊。此外, 将所提出的模型与传统轻量级模型在 OCT 数据集中进行了比较。实验表明, 利用混合深度分离卷积和轻量型注意力机制替换原始深度卷积和压缩注意力机制, 相比传统的 GhostNet 模型平均准确率提高了 2% 左右, 验证了该方法在 OCT 视网膜图像自动分类识别的性能和有效性。

**关键词** 医用光学; 光学相干层析成像; 混合深度分离卷积; 轻量型注意力机制; 多通道; 多尺度

中图分类号 TP391

文献标志码 A

doi: 10.3788/CJL202148.2307001

## 1 引言

临床上, 软性玻璃膜疣(Drusen)产生的老年黄斑变性(AMD), 糖尿病性黄斑水肿(DME)和脉络膜新生血管性疾病(CNV)是引起视力丧失的原因, 这些眼底疾病的早期发现和治疗已成为各国关注的主要健康问题之一。目前主要利用光学相干层析成像(OCT)对中央视网膜进行三维成像, 然后眼科专家通过观察 OCT 图像进行眼科疾病分类识别<sup>[1-3]</sup>。这些眼科疾病和视网膜中央的黄斑区域有关, 具有相似性, 并且在分类时多数需要先进行分割和去噪<sup>[4-5]</sup>。因此, 对于专业医生来说, 人工分类是一项费时又具有挑战的任务。在本项工作中, 主要用轻量级多尺度的卷积神经网络对 OCT 视网膜图像进行自动分类, 实现眼科疾病的高分类准确率, 未来在临床中方便医生快速查看检测结果。

随着科技的发展, 人工智能进入了人们的视野, 利用机器学习的经典算法和其分支领域的深度学习架构对医学图像进行分类已经成为热门研究课题,

并逐渐进入了 OCT 视网膜图像的检测与分类识别领域<sup>[6-8]</sup>。2014 年, Srinivasan 等<sup>[9]</sup>创建 OCT 数据集, 利用支持向量机(SVM)对 OCT 视网膜图像进行分类识别, 将其分为 AMD 和 DME。2016 年, Wang 等<sup>[10]</sup>使用 Srinivasan 提供的 OCT 数据集, 利用顺序最小化算法(SMO)将 OCT 图像分为 AMD 和 DME。2018 年, Rasti 等<sup>[11]</sup>使用多尺度卷积神经网络对两个 OCT 数据集进行分类和识别, 并对 OCT 图像的局部信息和整体信息进行了组合。2019 年, Wang 等<sup>[12]</sup>提出了一种基于深度学习的自动检测 DME 和 AMD 病变的方法, 其中采用了两个公开的视网膜 OCT 数据集, 并使用了具有有效重用特征的网络模型来解决小数据集的问题, 增强了对不同数据集差异的适应性, 比较了几种网络模型, 并在具有预训练模型的网络上实现了转移学习, 最终得到了良好的分类精度。2020 年 6 月, Shaban 等<sup>[13]</sup>利用深度卷积神经网络对 OCT 图像和眼底摄影的视网膜病变严重程度进行分类筛查, 取得了良好的分类效果。

然而, 在目前的机器学习和深度学习中, 对 OCT

收稿日期: 2021-03-18; 修回日期: 2021-04-15; 录用日期: 2021-04-27

基金项目: 国家自然科学基金青年科学基金(61308115)、上海市科委产学研医项目(15DZ1940400)

通信作者: \*cmhui.43@163.com

图像进行分类识别时,许多研究人员只关注分类的高准确率,忽视了临床应用的可能性,因此造成网络模型的参数量、计算复杂度和浮点计算量(FLOPs)越来越高,同时模型在进行推理预测时需要耗费大量的时间。因此为了实现相对轻量化、速度快、多尺度的 OCT 图像自动分类的卷积神经网络,提出了一种用于 OCT 视网膜图像自动分类的卷积神经网络 GM-OCTnet,对 OCT 眼底疾病进行分类识别。

在此工作中,基于华为诺亚方舟实验室提出的 GhostNet 框架<sup>[14]</sup>思想,对其进行改进,提出了 GM-OCTnet,用于自动分类识别 CNV、DME、玻璃膜疣,以及正常视网膜图像,并且在新的 OCT 数据集上验证了算法的性能。主要研究包括:

- 1) 利用一种新的轻量空间注意力机制模块,提高网络对 OCT 图像多通道特征提取的能力;
- 2) 使用混合深度分离卷积层代替深度卷积层,实现了对 OCT 图像多尺度特征提取;
- 3) 比较了几个常用轻量级网络在 OCT 图像分类上的参数量、浮点计算量(FLOPs)、权重(Weight),以及精确度(Accuracy)。

## 2 实验方法与原理

视网膜病变图像自动分类包括三个步骤:1)预处理图像。对采集到的 OCT 图像进行图像裁剪,裁剪 OCT 图像的空白区域,填充边缘空白区域,利用双边滤波去噪等预处理方法克服图像背景噪声对分类准确率的干扰;2)划分数据集。将经过预处理后的图像划分为训练集、验证集和测试集;3)分类。利用训练好的模型对未经过分类训练的图像进行视网膜病变图像的分类,测试模型的性能。

利用混合深度分离卷积建立 GMNeck 模块,替换传统轻量级网络 GBNeck 模块,在整个网络中使用轻量空间注意力机制,实现该方法对 OCT 视网膜图像的分类。为了在临床 OCT 数据集上实现对 CNV、DME、玻璃疣,以及正常 OCT 图像的快速准确地分类,提出了 GM-OCTnet 网络,这是一种基于 OCT 图像特征提取的多尺度通道注意力机制的便捷式网络架构,如图 1 所示。其中,BN 为归一化层,AvgPool 为平均池化层,蓝色层表示使用混合深度分离卷积层代替普通深度卷积层,红色字体层为

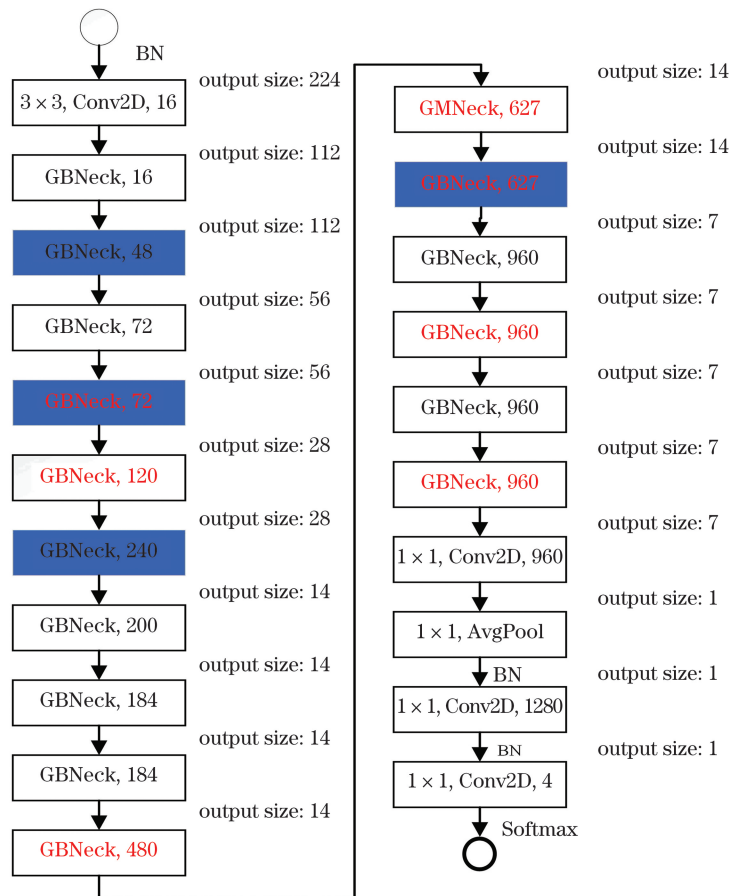


图 1 GM-OCTnet 网络的架构图  
Fig. 1 GM-OCTnet network architecture diagram

有轻量级空间注意力机制模块的层。

本研究提出了利用与卷积算子不同的轻量级空间注意力机制和轻量级卷积块来替换原来的卷积算子和压缩注意力机制(SE)<sup>[15-16]</sup>模块,再利用替换上述两个模块建立的 GM-OCTnet 网络的算法进行 OCT 视网膜图像自动分类。

### 2.1 轻量型子空间注意力机制

不同的 OCT 图像病变区域具有不同的尺度特征,因此为了更好地对不同 OCT 图像进行分类,大多数网络架构模型中采用注意力机制模块提高分类准确性,然而,现有的最先进的注意力机制导致计算参数高且内存开销大,因此不适合紧凑的卷积神经

网络。于是提出利用超轻量空间注意力机制<sup>[17]</sup>来替换传统的压缩注意力机制,以便在整个网络计算参数变化不大的同时提高分类准确性。

所使用的超轻量注意力机制的核心思想是为每个特征映射子空间推断不同的注意映射,以达到更好的特征提取效果,将提取的特征分成  $g$  组,对每组的子特征进行空间上的重新校准。将通道数进行分组,且每组的层包括  $1 \times 1$  的深度卷积层、 $3 \times 3$  的最大池化层、点卷积层、Softmax 层,以及通过元素乘法和加法得到中间特征图,利用融合运算将每组特征图进行融合,得到最终特征图如图 2 所示。

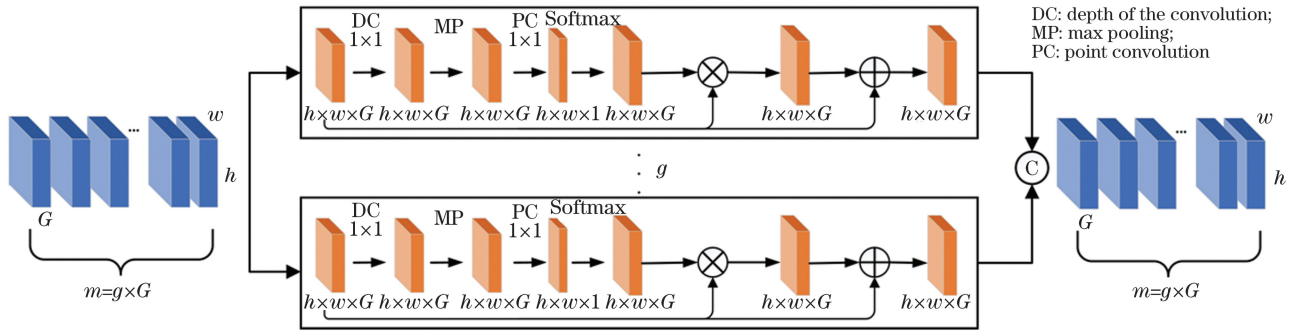


图 2 轻量型注意力机制模块

Fig. 2 Lightweight attention mechanism module

假设  $F \in \mathbb{R}^{m \times h \times w}$  是特征图的中间卷积层,其中  $m$  为通道数,  $m = g \times G$ ,  $G$  为分支子空间通道数,  $h$  为特征图的空间高度,  $w$  为特征图的空间宽度。若特征映射为  $g$  个互斥的组  $(F_1, F_2, \dots, F_n, \dots, F_g)$ , 则对应的特征映射计算公式如下

$$A_n = \text{softmax}\{\text{PC}^{1 \times 1}\{\text{MP}^{3 \times 3, 1}[\text{DC}^{1 \times 1}(F_n)]\}\}, \quad (1)$$

$$\hat{F}_n = (A_n \times F_n) + F_n, \quad (2)$$

$$\hat{F} = \text{concat}[(\hat{F}_1, \hat{F}_2, \dots, \hat{F}_n, \dots, \hat{F}_g)], \quad (3)$$

式中,  $\text{DC}(\cdot)$  为深度卷积,  $\text{MP}(\cdot)$  为最大池化,  $\text{PC}(\cdot)$  为点卷积。

### 2.2 GMNeck—轻量级卷积模块

GMNeck 模块里的混合深度分离卷积相比传统的卷积层参数,在选择不同分组卷积核数量时,具有不同的参数数量,因此在保持较高准确率和参数数量不会大幅度增加的情况下,需要选择合适的混合卷积分组数,此模块在步长等于 2 的时候,使用了混合深度分离卷积来改善原来大卷积核感受野大、准确率相对略高但参数数量相对较多和小卷积核感受野小、准确率低的问题<sup>[18]</sup>。具体地说,混合深度分离卷积是在一个卷积层中混合不同卷积核大小的深

度卷积。因为其可以替代普通卷积,所以在步长等于 2 的时候利用混合深度分离卷积以提高轻量级网络的准确率和效率。

在 GhostNet 模型中包含 GBNeck 模块,且在步长为 1 时只需两个 Ghost 模块,一个为了增加通道数量,一个为了减少通道数量以匹配快捷路径。在步长为 2 时,通过下采样层实现快捷路径,并且在两个 Ghost 模块之间插入步长为 2 的深度卷积。实际上,这里的 Ghost 模块中的主要卷积是逐点卷积,以提高效率。然而,本文的轻量级网络 GMNeck 实际上是在步长为 1 的情况下保持不变,将步长为 2 中的逐点卷积替换为混合深度分离卷积,以此减少计算参数和提高分类准确率。改变的部分如图 3 所示。

首先,把普通的卷积  $\mathbf{X}(N, H, W, C)$  设为形状为  $(H, W, C)$  的输入张量,  $N$  为数据样本数,  $H$  为空间高度,  $W$  为空间宽度,  $C$  为通道数。设  $\mathbf{G}(k, k, c, u)$  表示深度方向的卷积核,其中  $k \times k$  是核大小,  $c$  是输入通道大小,  $u$  是信道乘数。为了简单起见,这里假设内核宽度和高度是相同的  $k$ ,但是可以简单地推广到内核宽度和高度不同的情况。输出张量  $\mathbf{Y}(h, w, c, u)$  将具有相同的空间形状  $(h, w)$  和相乘

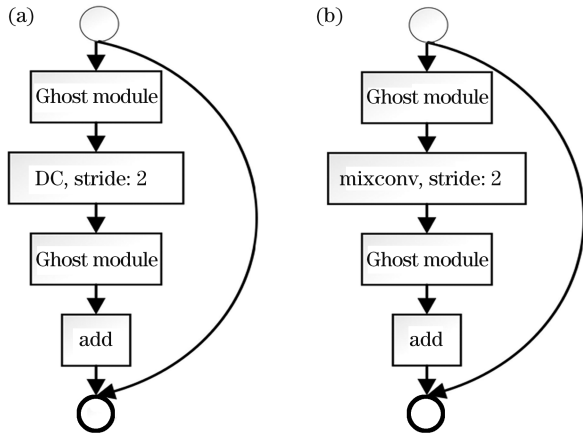


图 3 GBNeck 和 GMNeck 的比较。(a)传统 GBNeck 模块;(b)GMNeck 模块

Fig. 3 Comparison of GBNeck and GMNeck.

(a) Traditional GBNeck module; (b) GMNeck module

的输出通道大小  $u \times c$ , 每个输出特征映射值可表示为

$$Y_{x,y,z} = \sum_{-\frac{k}{2} \leq i \leq \frac{k}{2}, -\frac{k}{2} \leq j \leq \frac{k}{2}} X_{x+i,y+j,z/u} \cdot G_{i,j,z}, \forall z = 1, \dots, u \times c. \quad (4)$$

其次,对于混合卷积中,把原来的通道分为  $g$  个组,每个组的内核大小可分别为  $3, 5, 7, 9, \dots, 2g+1$ ,每组的通道大小有等划分和指数划分两种划分方式,例如给定总滤波器为 32,等划分时  $g$  组的混合卷积,通道分为  $\{32/g, 32/g, \dots, 32/g\}$ ,指数划分时  $g$  组的混合卷积,通道分为  $\{c_1 = 2^i, c_2 = 2^{i-1}, \dots, 32 - (c_1 + c_2 + \dots + c_g)\}$ ,总之两种划分总和等于滤波器总数。因此设输入张量为  $\langle \hat{X}^{(N,H,W,c_1)}, \dots, \hat{X}^{(N,H,W,c_g)} \rangle$ ,其中所有的  $\hat{X}$  具有相同的空间高度  $H$  和空间宽带  $W$ ,并且通道总数为  $c_1 + c_2 + \dots + c_g = c$ 。相应地, $g$  个组的深度卷积核为  $\langle \hat{G}(k_1, k_1, c_1, u_1), \dots, \hat{G}(k_g, k_g, c_g, u_g) \rangle$ 。因此, $t$  组输入张量和卷积核对应的输出特征映射的计算如下

$$\hat{Y}_{x,y,z}^t = \sum_{-\frac{k}{2} \leq i \leq \frac{k}{2}} -\frac{k}{2} \leq j \leq \frac{k}{2} \hat{X}_{x+i,y+j,\frac{z}{u}}^t \cdot \hat{G}_{i,j,z}^t, \quad \forall z = 1, \dots, u \times c_t. \quad (5)$$

最后,得到的输出张量为所有张量  $\langle \hat{Y}_{x,y,z_1}^1, \dots, \hat{Y}_{x,y,z_g}^g \rangle$  的融合,表达式为

$$Y_{x,y,z_0} = \text{concat}(\hat{Y}_{x,y,z_1}^1, \dots, \hat{Y}_{x,y,z_g}^g). \quad (6)$$

式中, $z_0 = z_1 + \dots + z_g = u \times c$  是最终输出通道大小。

混合深度分离卷积不仅可以实现单个运算,还

可以实现分组运算,并且该方法所用的混合深度分离卷积可分别分成不同组别,当混合深度分离卷积分组数为 1 时,就是普通卷积,为了获得合适的感受野,以及保持较高的准确率,混合深度分离卷积的分组数选择为 4,卷积核的大小选择为  $3 \times 3, 5 \times 5, 7 \times 7, 9 \times 9$ ,最终测试对 OCT 眼底图像的分类效果。

### 3 实验结果与分析

将 GM-OCTnet 方法应用于 OCT 图像分类识别任务,在收集到的数据集上进行预处理后,划分训练集、验证集和测试集,在保持训练集、验证集和测试集不变的情况下,执行了三次实验,并将测试平均值作为最终结果。最重要的是分别比较了在不同注意力机制组别下的模型,以及使用不同轻量级的模型,验证了此网络模型对 OCT 视网膜图像自动分类的性能。

#### 3.1 数据集和实验设置

使用该方法对眼科疾病的临床数据进行分类实验。数据集 1 (<https://data.mendeley.com/datasets/rscbjbr9sj/2>) 是通过 Spectralis SD-OCT 成像设备采集获得<sup>[19]</sup>,分别来自加州大学圣地亚哥分校 Shiley 眼科研究所、加州网膜研究基金会、医学中心眼科协会、上海第一人民医院和北京同仁眼科中心的成人 OCT 视网膜图像<sup>[20-23]</sup>,其中包括 4486 名患者的 OCT 图像,11348 张为 DME,37205 张为 CNV,8616 张为玻璃膜疣,26315 张为正常 OCT 图像。测试集来自 633 名患者的 1000 张图像,其中各个类别分别有 250 张 OCT 图像。OCT 数据集图片通过了初始图像的质量评估,模型训练和测试之前对数据集中每张图片进行预处理,裁剪掉白色无关区域,并且把图片裁剪成合适的大小,突出 OCT 图片的重点特征信息,得到最终的图像如图 4 所示,图 4(a)分别为 CNV、DME 和玻璃疣原始未经处理的图像,图 4(b)分别为 CNV、DME 和玻璃疣经过预处理后的图像。

为了进一步验证本文方法的性能,使用数据集 2 进行验证。该数据集在温州医科大学采集得到,其中训练集包括 24 名 AMD 患者图像 1485 张,23 名 DME 患者图像 1465 张,11 名干性 AMD 患者图像 1571 张,15 个正常人图像 3936 张,测试集来自 8 名患者的 640 张图像,各类别分别为 160 张 OCT 图像。将此数据集进行质量评估和数据集划分,因数据集 2 的空白区域多、图像噪点多,先利用算法进行填充、裁剪和使用双边滤波去噪处理,如图 5 所示,

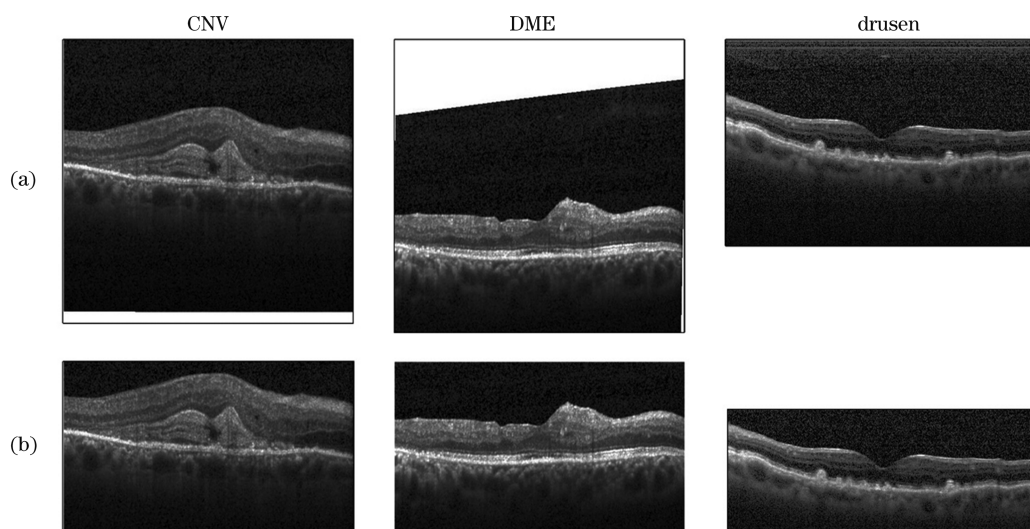


图 4 数据集 1 中 OCT 预处理图片。(a)预处理前图片;(b)预处理后图片

Fig. 4 OCT preprocessing images in dataset 1. (a) Image before preprocessing; (b) image after preprocessing

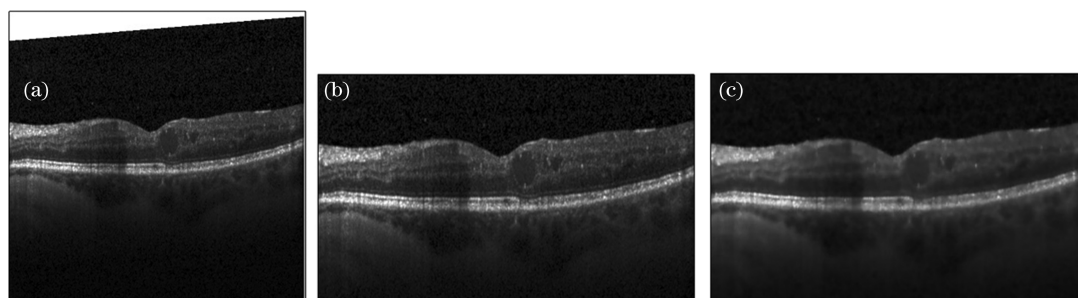


图 5 数据集 2 中 OCT 图像。(a)原始图像;(b)填充和裁剪后图像;(c)去噪后的图像

Fig. 5 OCT images in dataset 2. (a) Original image; (b) filled and cropped image; (c) image after denoising

再进行数据集划分,最后使用本方法对其进行分类,对比数据集 1 和数据集 2 的实验分类结果。

本实验选择 TensorFlow2.0<sup>[24]</sup> 作为深度学习开发框架,数据集的训练和验证在拥有 32 GB 显存的 Nvidia Tesla V100 GPU HPC 集群上进行,使用 Adam 优化器进行 OCT 分类,网络模型训练过程中刚开始的初始学习率设置为 0.001,在每 10 个周期后学习率乘以 0.8,训练的最大周期为 50。输入的训练集和验证集的图片尺寸为  $224 \times 224$ 。为了防止过拟合,我们应用了训练集数据增强方法,首先把 OCT

眼底图像进行随机旋转  $50^\circ$ ,其次对 OCT 图像进行水平和垂直方向上移动 0.2,最后将一半图像水平翻转,所有的验证实验都使用相同的训练参数。

### 3.2 分类结果

实验使用不同的模型 MobilenetV3<sup>[25]</sup>、ShufflenetV2<sup>[26]</sup>、GhostNet 和含有不同组别的超轻量空间注意力机制模块的 GM-OCTnet 对 OCT 图像进行分类,通过不同的评价指标评估提出的神经网络模型对四种 OCT 图像的分类效果。实验结果如表 1 所示,黑体代表性能指标最高。通过表格可

表 1 不同模型性能指标对比

Table 1 Comparison of performance indicators of different models

Model	Parameter / $10^6$	Accuracy /%	FLOPs / $10^6$	Weight / $10^6$
MobilenetV3 small	213	94.5	56	13.28
ShufflenetV2 1.0×	154	94.0	146	4.95
GhostNet 1.0×	259	93.5	166	10.20
GM-OCTnet( $g=1$ )	261	95.5	171	10.64
GM-OCTnet( $g=2$ )	261	94.5	171	10.64
GM-OCTnet( $g=4$ )	261	<b>96.1</b>	172	10.64
GM-OCTnet( $g=8$ )	261	94.9	172	10.64

以看出,虽然提出的神经网络模型的网络参数量相比 GhostNet 增加了  $2 \times 10^6$ , FLOPs 在 GhostNet 的基础上增加  $5 \times 10^6$ , 但不同组的 GM-OCTnet 网络分类准确度均比 GhostNet 高, 其权重文件大小比 MobileNetV3 的小  $2.64 \times 10^6$ , 最高准确度比 MobilenetV3 高 0.016, 可见, 使用不同分组的超轻量空间注意力机制模块优于传统的轻量级分类网络, 在超轻空间注意力机制模块分组数  $g=1$  时, 此模型相比 ShufflenetV2 的准确率提升了 1.5% 左

右。在分组数  $g=4$  时, 此模型相比原始 GhostNet 的分类准确率提升了 2.6%, 显示了所提 OCT 视网膜病变图像自动分类网络的优越性。

实验证明, 对传统 GhostNet 使用 7 个超轻量空间注意力机制模块能够达到最佳效果, 添加注意力机制模块的位置如图 1 所示。表 1 中显示, 在 Gm-OCTnet 网络分组  $g=4$  时, 浮点计算量相比传统 GhostNet 增加了  $6 \times 10^6$ , 但是准确率从原来 GhostNet 的 93.5% 提高到 96.1%。

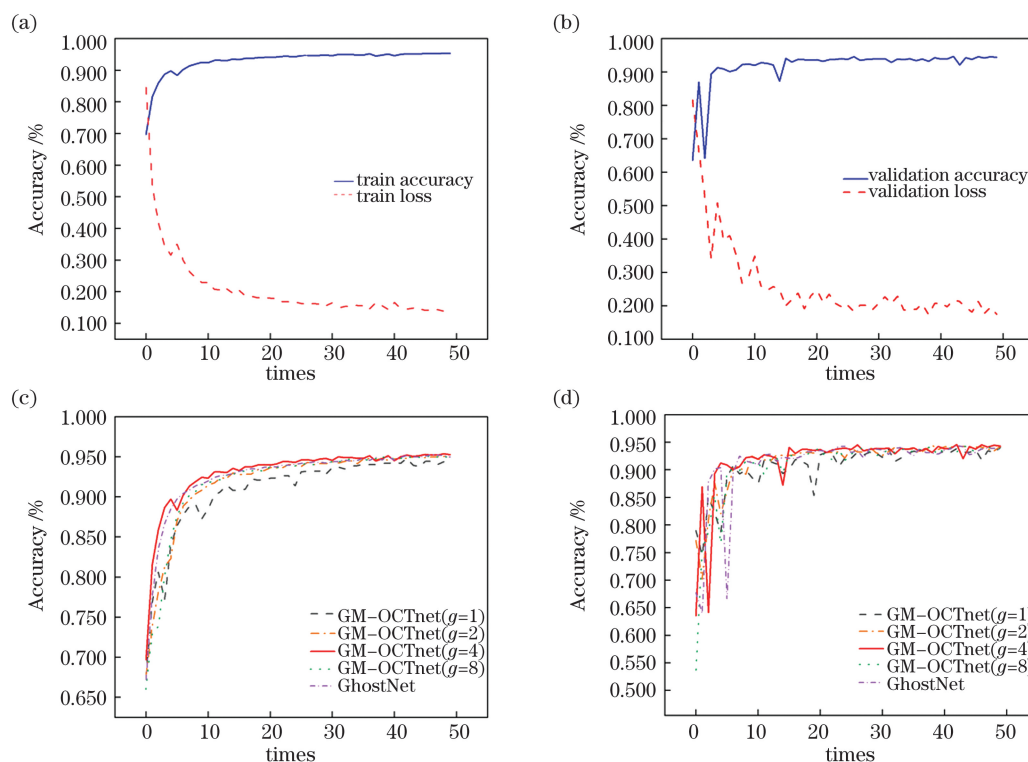


图 6 实验的准确性与损失的关系。(a) 50 次训练准确性和训练损失的关系;(b) 50 次验证准确性和验证损失的关系;(c) 5 种模型的训练准确性;(d) 5 种模型的验证准确性

Fig. 6 Relationship between the accuracy and the loss of the experiment. (a) Relationship between 50 times training accuracy and training loss; (b) relationship between 50 times validation accuracy and validation loss; (c) train accuracy of 5 models; (d) validation accuracy of models

为了更加直观地看到实验结果, 本实验显示了网络模型对 OCT 图像分类的训练准确度与训练损失关系和验证准确度与验证损失的关系。图 6(a) 表示在 50 个周期训练过程中, GM-OCTnet 模型在分组  $g=4$  时训练准确度随训练周期的增加逐渐升高, 且趋于稳定, 图 6(b) 表示在 50 个周期训练结束后, GM-OCTnet 模型在分组  $g=4$  时验证准确度随着训练周期的增加而升高, 且趋于稳定, 图 6(c) 表示传统模型和分组分别为  $g=1, g=2, g=4, g=8$  时, 5 种训练模型在 50 个周期过程的训练准确性, 显示出模型 GM-OCTnet 在分组  $g=4$  时的训练准确度随训练周期的增加而升高, 且优于其他模型,

图 6(d) 表示传统模型和分组分别  $g=1, g=2, g=4, g=8$  时, 5 种训练模型的 50 个周期过程的验证准确度, 显示出模型 GM-OCTnet 在分组  $g=4$  时的验证准确性比其他模型验证准确度高, 且达到最佳。

对比数据集 2 实验验证结果, 在数据集 2 上, 利用本算法进行自动分类后, 与数据集 1 的结果进行比较发现, 在分组数  $g=1$  时, 数据集 2 的分类准确度比数据集 1 的分类准确度高 0.018, 在分组数  $g=2$  时, 数据集 2 的分类准确度比数据集 1 的分类准确度高 0.03, 在分组数  $g=4$  时, 数据集 2 的分类准确度比数据集 1 的分类准确度高 0.015, 在分组数  $g=8$ , 数据集 2 的分类准确度比数据集 1 的准确

度高 0.021。如表 2 所示。对比实验证明,数据集 1 和数据集 2 利用本算法进行自动分类均达到了高分类准确度。

表 2 不同数据集分类效果对比

Table 2 Comparison of classification effects of different datasets

	Dataset1 accuracy	Dataset2 accuracy
GM-OCTnet( $g=1$ )	0.955	0.973
GM-OCTnet( $g=2$ )	0.945	0.975
GM-OCTnet( $g=4$ )	0.961	0.976
GM-OCTnet( $g=8$ )	0.949	0.970

## 4 结 论

GM-OCTnet 将传统 GhostNet 中简单的深度卷积替换为混合深度分离卷积,压缩激励模块利用轻量型空间注意力机制模块进行优化,实现了多尺度、多通道的 OCT 图像分类识别,在计算量相对传统 GhostNet 变化不大的情况下,有效提高了分类准确性。

测试和评估了基于 OCT 图像数据集的轻量型神经网络 GM-OCTnet 在分类诊断 CNV、DME、Drusen 和正常患者这四类眼科状态的效果。同时研究了不同模型在 OCT 图像分类的性能对比,使用准确性、参数量、计算量和权重文件作为评估标准,验证 GM-OCTnet 模型对 OCT 图像的分类有效性。为了进一步验证本算法的性能,利用两个数据集对本文所提算法性能进行验证。通过上述指标和准确性以及损失率的曲线图,可看出提出的 OCT 分类模型的优越性,若应用在临床上,可以提高专业眼科医生对患者眼科疾病的诊断效率,减少对患者的漏诊和误诊。实验证明,在使用 GM-OCTnet 对 OCT 图像分类相对于原始轻量型网络的分类准确率提高了 2% 左右。它的速度也快很多,只需要 0.01 s 就能准确分类不同的 OCT 图像。未来有了更多的训练数据,就有可能进一步提高自动分类的质量,达到更好的临床应用效果。

## 参 考 文 献

- [1] Huang Z Y, Hu Y C, Zhou C Q, et al. Swept source optical coherence tomography and its application in ophthalmology[J]. *Journal of Ophthalmology*, 2021, 36(1): 55-65.  
黄智宇, 胡毅成, 周传清, 等. 扫频光学相干断层成像技术及其在眼科的应用[J]. *眼科学报*, 2021, 36(1): 55-65.
- [2] Xu Y P, Du Y C, Chen F. Automatic layer segmentation of optical coherence tomography images in retinal vascular diseases[J]. *Journal of Shanghai Jiao Tong University (Medical Science)*, 2019, 39(6): 614-621, 613.  
许毓鹏, 杜宇辰, 陈凤娥. 视网膜血管性疾病光学相干断层成像图像的自动分层研究[J]. *上海交通大学学报(医学版)*, 2019, 39(6): 614-621, 613.
- [3] Farsi S, Chiu S J, O'Connell R V, et al. Quantitative classification of eyes with and without intermediate age-related macular degeneration using optical coherence tomography[J]. *Ophthalmology*, 2014, 121(1): 162-172.
- [4] Xie C X, Gao W R, Zhang Y. Image speckle noise suppression method based on amplitude difference probability density distribution of OCT imaging signal[J]. *Chinese Journal of Lasers*, 2020, 47(12): 1207004.  
谢晨霞, 高万荣, 张越. 基于 OCT 成像信号振幅差分概率密度分布的图像散斑噪声抑制方法[J]. *中国激光*, 2020, 47(12): 1207004.
- [5] Sidibé D, Sankar S, Lemaître G, et al. An anomaly detection approach for the identification of DME patients using spectral domain optical coherence tomography images [J]. *Computer Methods and Programs in Biomedicine*, 2017, 139: 109-117.
- [6] Hu T, Li W H, Qin X X. Semantic segmentation of polarimetric synthetic aperture radar images based on multi-layer deep feature fusion[J]. *Chinese Journal of Lasers*, 2019, 46(2): 0210001.  
胡涛, 李卫华, 秦先祥. 基于多层深度特征融合的极化合成孔径雷达图像语义分割[J]. *中国激光*, 2019, 46(2): 0210001.
- [7] 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, 1(4): 322-327.
- [8] Burlina P, Freund D E, Joshi N, et al. Detection of age-related macular degeneration via deep learning[C]// 2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI), April 13-16, 2016, Prague, Czech Republic. New York: IEEE Press, 2016: 184-188.
- [9] 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.
- [10] Wang Y, Zhang Y N, Yao Z M, et al. Machine learning based detection of age-related macular degeneration (AMD) and diabetic macular edema

- (DME) from optical coherence tomography (OCT) images [J]. *Biomedical Optics Express*, 2016, 7 (12): 4928-4940.
- [11] 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.
- [12] Wang D P, Wang L J. On OCT image classification via deep learning[J]. *IEEE Photonics Journal*, 2019, 11(5): 1-14.
- [13] Shaban M, Ogur Z, Mahmoud A, et al. A convolutional neural network for the screening and staging of diabetic retinopathy[J]. *PLoS One*, 2020, 15(6): e0233514.
- [14] Han K, Wang Y H, Tian Q, et al. GhostNet: more features from cheap operations[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), June 13-19, 2020, Seattle, WA, USA. New York: IEEE Press, 2020: 1577-1586.
- [15] Hu J, Shen L, Sun G. Squeeze-and-excitation networks [C] // 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, June 18-23, 2018, Salt Lake City, UT, USA. New York: IEEE Press, 2018: 7132-7141.
- [16] Kong F Q, Zhou Y B, Shen Q, et al. End-to-end multispectral image compression using convolutional neural network[J]. *Chinese Journal of Lasers*, 2019, 46(10): 1009001.  
孔繁镛, 周永波, 沈秋, 等. 基于卷积神经网络的端到端多光谱图像压缩方法[J]. *中国激光*, 2019, 46 (10): 1009001.
- [17] Saini R, Jha N K, Das B, et al. ULSAM: ultra-lightweight subspace attention module for compact convolutional neural networks[C]//2020 IEEE Winter Conference on Applications of Computer Vision (WACV), March 1-5, 2020, Snowmass, CO, USA. New York: IEEE Press, 2020: 1616-1625.
- [18] Tan M, Le Q V. Mixconv: mixed depthwise convolutional kernels[EB/OL]. (2019-07-22) [2021-03-15]. <https://arxiv.org/abs/1907.09595>.
- [19] Abràmoff M D, Lou Y, Erginay A, et al. Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning[J]. *Investigative Ophthalmology & Visual Science*, 2016, 57(13): 5200-5206.
- [20] Kermayn D S, Goldbaum M, Cai W J, et al. Identifying medical diagnoses and treatable diseases by image-based deep learning [J]. *Cell*, 2018, 172 (5): 1122-1131. e9.
- [21] Alsaih K, Lemaitre G, Rastgoo M, et al. Machine learning techniques for diabetic macular edema (DME) classification on SD-OCT images [J]. *Biomedical Engineering Online*, 2017, 16(1): 68.
- [22] Hussain M A, Bhuiyan A, Luu C D, et al. Classification of healthy and diseased retina using SD-OCT imaging and Random Forest algorithm [J]. *PLoS One*, 2018, 13(6): e0198281.
- [23] Alsaih K, Lemaitre G, Vall J M, et al. Classification of SD-OCT volumes with multi pyramids, LBP and HOG descriptors: application to DME detections[C] //2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), August 16-20, 2016, Orlando, FL, USA. New York: IEEE Press, 2016: 1344-1347.
- [24] Martín A, Agarwal A, Barham P, et al. Tensorflow: large-scale machine learning on heterogeneous distributed systems[EB/OL]. (2016-03-14)[2021-03-15].
- [25] Howard A, Sandler M, Chen B, et al. Searching for MobileNetV3 [C] // 2019 IEEE/CVF International Conference on Computer Vision (ICCV), October 27-November 2, 2019, Seoul, Korea (South). New York: IEEE Press, 2019: 1314-1324.
- [26] Yuan J Y, Nan X Y, Li C R, et al. Research on real-time multiple single garbage classification based on convolutional neural network [J]. *Mathematical Problems in Engineering*, 2020, 2020: 1-6.



# Research on Automatic Classification of Optical Coherence Tomography Retina Image Based on Multi-Channel

Chen Sisi, Chen Minghui\*, Ma Wenfei

*The Ministry of Education of Medical Optical Engineering Center,*

*Shanghai Institute for Interventional Medical Devices, School of Medical Instrument and Food Engineering,*

*University of Shanghai for Science and Technology*

## Abstract

**Objective** Vision loss is caused by age-related macular degeneration because of soft drusen, diabetic macular edema and choroidal neovascular disease. Early detection and treatment of these fundus diseases have emerged as a major health concern for all countries. Professional doctors often use retinal images from optical coherence tomography (OCT) to diagnose eye diseases. However, because there are several types of retinopathy images and the lesion area is similar, manually classifying OCT retina images is a time-consuming and difficult task. With the development of artificial intelligence, researchers began classifying medical images using classic machine learning algorithms and deep learning in its branch areas, eventually progressing to the automatic classification of OCT retinal images. Several researchers are only concerned with classification accuracy and ignore the possibility of clinical application. Consequently, the network model's parameter amount, computational complexity and floating-point operations (FLOPs) calculation amount are increasing, and the model is making inference predictions, which consumes a long time to complete. In this paper, a multi-channel, multi-scale lightweight convolutional neural network is proposed to automatically classify OCT retinal images for achieving high ophthalmic disease classification accuracy. In the future, doctors will be able to quickly view detection results in the clinic.

**Methods** In this study, a multi-channel OCT retinal image automatic classification deep neural network is used. The neural network model is based on the GM-OCTnet algorithm, which includes a light quantum spatial attention mechanism distinct from the convolution operator and a lightweight convolution block to replace the two modules in the original model for automatically classifying OCT retinal images. Image pre-processing, dataset division and classification using the model algorithm are the steps taken to achieve the automatic classification of the entire OCT retina. First, image cropping is performed on the collected OCT image, the blank area of the OCT image is cropped, the marginal blank area is filled and bilateral filter denoising and other pre-processing methods are used to overcome the interference of image background noise on the classification accuracy. Then, the pre-processed image is divided into three sets: training, validation, and test sets. Afterwards, the OCT images from the training set are trained using the proposed multi-channel lightweight deep neural network GM-OCTnet model algorithm. Following the test, the well-trained model is used to classify unclassified retinal images automatically. In addition, the results of this work on the automatic classification of OCT retinal images are validated by comparing the proposed model with three traditional lightweight models on the OCT data set, and different data sets are used to further validate the algorithm's performance.

**Results and Discussions** Different pre-processing methods were used to process the OCT images after evaluating the quality of two different data sets (Figs. 4 and 5). The experimental results show that when the number of groups is 4, the proposed multi-channel OCT automatic retina classification network achieves an average accuracy of 96.1%, which is 2.6% higher than that of the original model GhostNet in the automatic classification of OCT retina images. Its file size is  $2.64 \times 10^6$  smaller than that of MobileNetV3. The training and verification accuracies of the GM-OCTnet model when the grouping  $g = 4$  increase gradually with the increase of the training period and tend to stabilise, based on the relationship between the verification loss rate and the verification accuracy rate and the training loss rate and the training accuracy rate curve. When comparing the 50 training processes of different models, the proposed model is found to exhibit a higher accuracy rate than other models when the number of groups is equal to 4 and prioritises reaching the best value (Fig. 6). Overall, the model algorithm proposed in this paper has achieved high accuracy in the automatic classification of OCT retinal images. In addition, when the experimental results of two different datasets are compared, it is discovered that the automatic classification of datasets 1 and 2 achieved high classification accuracy using this algorithm (Tables 1 and 2).

**Conclusions** This study proposes a multi-channel, multi-scale lightweight network for automatically classifying OCT retinal images. The effect of the lightweight neural network GM-OCTnet based on the OCT image datasets in classifying and diagnosing the four types of ophthalmic conditions, i. e. choroidal neovascular disease(CNV), diabetic macular edema(DME), drusen and normal patients, were tested and evaluated. Two different datasets are used to further validate the performance of the algorithm proposed in this work. To validate the effectiveness of the GM-OCTnet model for OCT image classification, accuracy, parameter amount, calculation amount and weight file size are used as evaluation criteria. It is found the proposed OCT classification model has improved classification accuracy through experimental results. When used in the clinic, it can improve professional ophthalmologists' diagnosis efficiency for patients with ophthalmic diseases and also reduce missed diagnosis and misdiagnosis of patients.

**Key words** medical optics; optical coherence tomography; mixed depth separation convolution; lightweight attention mechanism; multi-channel; multi-scale

**OCIS codes** 170.4500; 170.3880; 170.5755