



## 多尺度特征融合网络的视网膜OCT图像分类

韩璐, 毕晓君

引用本文:

韩璐, 毕晓君. 多尺度特征融合网络的视网膜OCT图像分类[J]. 智能系统学报, 2022, 17(2): 360–367.

HAN Lu, BI Xiaojun. Retinal optical coherence tomography image classification based on multiscale feature fusion[J]. *CAAI Transactions on Intelligent Systems*, 2022, 17(2): 360–367.

在线阅读 View online: <https://dx.doi.org/10.11992/tis.202111024>

## 您可能感兴趣的其他文章

### 基于注意力机制的显著性目标检测方法

Salient object detection method based on the attention mechanism

智能系统学报. 2020, 15(5): 956–963 <https://dx.doi.org/10.11992/tis.201903001>

### 基于注意力融合的图片描述生成方法

An image caption generation method based on attention fusion

智能系统学报. 2020, 15(4): 740–749 <https://dx.doi.org/10.11992/tis.201910039>

### 基于跳跃连接金字塔模型的小目标检测

Skip feature pyramid network with a global receptive field for small object detection

智能系统学报. 2019, 14(6): 1144–1151 <https://dx.doi.org/10.11992/tis.201905041>

### 层次化双注意力神经网络模型的情感分析研究

Hierarchical double-attention neural networks for sentiment classification

智能系统学报. 2020, 15(3): 460–467 <https://dx.doi.org/10.11992/tis.201812017>

### 基于双向消息链路卷积网络的显著性物体检测

Salient object detection based on bidirectional message link convolution neural network

智能系统学报. 2019, 14(6): 1152–1162 <https://dx.doi.org/10.11992/tis.201812003>

### 基于卷积神经网络的遥感图像分类研究

Classification of remote-sensing image based on convolutional neural network

智能系统学报. 2018, 13(4): 550–556 <https://dx.doi.org/10.11992/tis.201706078>

微信公众平台



关注微信公众号，获取更多资讯信息

DOI: 10.11992/tis.202111024

网络出版地址: <https://kns.cnki.net/kcms/detail/23.1538.TP.20211213.1720.004.html>

# 多尺度特征融合网络的视网膜 OCT 图像分类

韩璐<sup>1</sup>, 毕晓君<sup>2</sup>

(1. 哈尔滨工程大学 信息与通信工程学院, 黑龙江 哈尔滨 150001; 2. 中央民族大学 信息工程学院, 北京 100081)

**摘 要:** 目前基于深度学习的视网膜 OCT 图像分类方法存在网络特征提取能力低、小目标病变分类困难等问题。为此本文提出了一种双分支多尺度特征融合网络, 通过加入门控注意力机制, 利用深层特征作为选通信号传递给浅层特征, 在消除冗余特征的同时, 获得更细尺度的抽象信息。同时加入空洞空间金字塔模块, 实现在不降低特征图分辨率的同时增大感受野, 按不同比例有效捕获全局上下文信息, 提高了小目标病变分类精度。实验结果表明, 本文提出的方法在视网膜 OCT 图像分类任务中取得了较好效果, 分类准确率达 97.9%。

**关键词:** 视网膜; 光学相干断层扫描; 注意力机制; 空间空洞金字塔; 神经网络; 图像分类; 深度学习; 医学图像  
**中图分类号:** TP391.7   **文献标志码:** A   **文章编号:** 1673-4785(2022)02-0360-08

中文引用格式: 韩璐, 毕晓君. 多尺度特征融合网络的视网膜 OCT 图像分类 [J]. 智能系统学报, 2022, 17(2): 360-367.

英文引用格式: HAN Lu, BI Xiaojun. Retinal optical coherence tomography image classification based on multiscale feature fusion[J]. CAAI transactions on intelligent systems, 2022, 17(2): 360-367.

## Retinal optical coherence tomography image classification based on multiscale feature fusion

HAN Lu<sup>1</sup>, BI Xiaojun<sup>2</sup>

(1. College of Information and Communication Engineering, Harbin Engineering University, Harbin 150001, China; 2. School of Information Engineering, Minzu University of China, Beijing 100081, China)

**Abstract:** The retinal optical coherence tomography (OCT) image classification method based on deep learning has problems such as low ability of network feature extraction and difficult classification of small target lesions. Therefore, this paper proposes a dual branch multiscale feature fusion network. The gating attention mechanism is added to the vgg16 network, and the deep features are transmitted to the shallow features as gating signals. The redundant features are removed more fine-grained abstract information is obtained. Simultaneously, an atrous spatial pyramid pooling (ASPP) module is introduced to increase the receptive field and capture the global context information in various proportions without reducing the feature map resolution. The ASPP module increases the classification accuracy of small target lesions. The experimental results show that the proposed method has achieved good results in the retinal OCT image classification task, and the classification accuracy has reached 97.9%.

**Keywords:** retina; optical coherence tomography; attention mechanism; atrous spatial pyramid pooling; neural network; image classification; deep learning; medical image

近年来, 患有视网膜黄斑病变的患者数量明显增加, 并且随着病情的加深, 该病会对视力产生不可逆转的影响, 严重情况下会导致失明<sup>[1-3]</sup>。因此, 黄斑病变的早期发现和临床诊断至关重要

要, 通过适当的治疗和定期的筛查可以使黄斑病变引起的失明减少 90%。

光学相干断层扫描技术 (optical coherence tomography, OCT) 可以提供高分辨率的视网膜截面图像, 是目前用于检验视网膜疾病最为先进的技术手段, 具有非接触、无创、成像快等优点<sup>[4-7]</sup>,

收稿日期: 2021-11-13. 网络出版日期: 2021-12-21.

通信作者: 毕晓君. E-mail: [bixiaojun@hrbeu.edu.cn](mailto:bixiaojun@hrbeu.edu.cn).

医生通过对视网膜 OCT 图像的分析对眼底疾病做出诊断。然而,利用 OCT 技术对视网膜疾病进行人工诊断面临以下问题:视网膜黄斑病变患者逐年增加,医生面临巨大的阅片任务,依靠专业医生的诊断已经无法满足大量患者的诊疗需求;OCT 图像为灰度图像,个别病变特征不明显,医师诊断时有误诊和漏诊情况发生;个别地区医疗水平较差,导致大量患者在患病初期没有及时发现,造成疾病恶化<sup>[8-10]</sup>。

计算机辅助诊断技术<sup>[11-13]</sup>(computer-aided diagnosis, CAD)是解决这一问题的有效方法。早期的 CAD 技术使用基于手工特征的传统机器学习分类方法,尽管在视网膜 OCT 图像的分类领域取得了一些成果,但是存在严重依赖于人工设计特征,特征级别低,在网络训练过程中存在计算代价高、处理流程复杂等问题。近年来,深度学习迅速发展,并在计算机视觉领域表现优异。深度学习使用卷积神经网络在不同的层次上自动学习从低级到高级的图像特征,弥补了传统技术需要手工提取特征这一缺陷,成为解决视网膜 OCT 图像分类的主流算法。其中具有代表性的有 2017 年, Karri 等<sup>[14]</sup>提出了一种基于迁移学习的视网膜 OCT 图像分类方法。该方法通过微调预训练后的 GoogLeNet 网络,减小网络对大量数据的依赖,在训练数据有限的情况下实现对 DME<sup>[15-16]</sup>、AMD 和正常图像的分类,分类精度分别为 86%、89% 和 99%。2020 年,张添福等<sup>[17]</sup>提出了一种轻量化 OCT 图像分类网络。使用深度可分离卷积代替普通卷积层从而减小网络的参数。同时使用全局平均池化代替全连接层,提高空间鲁棒性,其网络准确率可达 97%。

以上研究对视网膜 OCT 分类任务做出了突出贡献,但是尚存以下两点问题:1) 视网膜 OCT 图像存在大量冗余,在特征提取过程中,显著病变特征容易被忽略,造成有用信息的丢失;2) 玻璃疣 (Drusen) 病变位置小且形态模糊,导致 Drusen 这类疾病的分类难度大,目前该类别的准确率尚需提高。

针对以上问题,本文主要做了以下 3 个方面的工作:1) 设计了一种双通道的多尺度融合网络,有效利用包含了丰富语义信息的深层特征以及包含纹理信息的浅层特征;2) 引入扩张卷积,通过在网络中加入一系列并行的扩张卷积,实现在不降低特征图分辨率的同时,增大感受野,按不同比例获得上下文信息;3) 引入门控注意力模块,利用深层特征作为选通信号传递给浅层特征,在消

除冗余特征的同时,获得更细尺度的细节信息。

## 1 双分支多尺度特征融合网络

2014 年,牛津大学著名研究组 Visual Geometry Group 提出 VGG 网络<sup>[18]</sup>,斩获该年 ImageNet 竞赛定位任务第一名和 Classification Task 分类任务第二名。该工作探索了卷积神经网络的深度对图像分类任务性能的影响,在固定网络架构中其他参数的同时,通过增加卷积层来平稳地增加网络深度,分类准确率获得显著提升。

VGG 网络由卷积层、最大池化下采样层和全连接层组成。其中,卷积层均采用卷积核大小为  $3 \times 3$ ,步距为 1,填充为 1 的卷积操作。VGG16 相比于 AlexNet,采用连续的几个  $3 \times 3$  卷积核代替 AlexNet 中的较大卷积核。两个  $3 \times 3$  卷积的堆叠层具有  $5 \times 5$  的感受野,3 个这样结构堆叠获得的感受野是  $7 \times 7$ 。使用小卷积核堆叠代替大卷积核可以整合非线性映射层,使决策函数更加具有判别性;在拥有相同感受野的前提下能够减少网络所需参数;同时增加网络非线性,让网络学习到更复杂、深层的特征。VGG 提出的网络深度从 11 层到 19 层不等,本文选择 VGG16 作为基线网络。

本文提出了一种基于改进 VGG16 的视网膜图像分类网络——双分支多尺度特征融合网络。主要改进点如下:

1) 在 VGG16 网络中引入针对医学图像的门控注意力机制模块,从而消除医学图像中的大量冗余,突出病变区域信息抑制图像中的无关区域;

2) 在 VGG16 网络中使用并行的扩张卷积在不减小特征图大小的同时,获得较大感受野,得到病变的细节信息,并与深度抽象特征融合,提高分类精度。

### 1.1 门控注意力机制

近年来,注意力机制作为一种可以即插即用在网络模型中的模块,在自然图像处理领域取得了良好的效果。其中最具代表性的工作有 2017 年 HU 等<sup>[19]</sup>提出的通道注意力机制、2018 年 Woo 等<sup>[20]</sup>提出的融合了通道注意力以及空间注意力的 CBAM 机制以及 2020 年 Wang 等<sup>[21]</sup>提出的改进通道注意力机制。这种机制在通道和空间两个维度加权生成注意力图,使网络能够关注重要的通道特征以及空间上的位置信息。

相比于自然图像,医学图像具有目标区域局部化这一特性。尤其是本文使用的眼部 OCT 图像,其病变位置均占整张 OCT 图像很小的区域。鉴于医学图像这一特性,若将通道和空间注意力



机制串联至网络中,会导致获得的加权注意力图谱单一。尽管这种机制串联在网络深处能够取得良好效果,但是网络的加深使特征图减小,导致相关病变的细节信息丢失,因此基于深度学习的视网膜 OCT 图像分类技术中采用通道和空间注

意力机制效果不佳。

本文引入了一种针对医学图像中目标局部化这一特性,重点关注医学图像中病变细节信息的门控注意力机制模块 (attention gate module, AG),如图 1 所示。

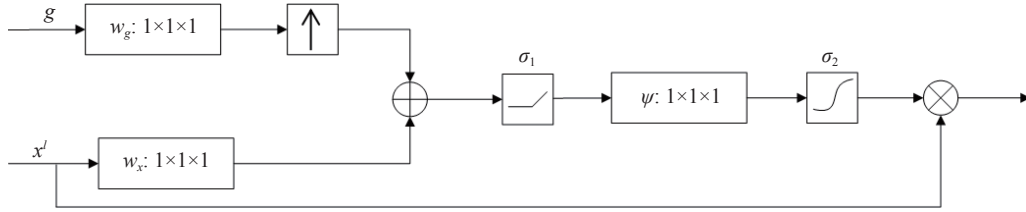


图 1 门控注意力模块

Fig. 1 Attention gate module

深层的粗糙特征包含目标对象的位置信息,并在全局范围内建立它们之间的关系。图 1 中:  $g$  代表网络中获取到的深层特征;  $x'$  代表特征提取过程中任意某一层获取的浅层特征,浅层特征中包含目标的细节信息,比如形状、大小等。浅层特征中包含着严重影响分类任务准确性的细节信息。AG 模块将深层特征和浅层特征融合并生成注意力图谱,然后将该注意力图谱与浅层特征相乘,用深层信息消除  $x'$  中与任务无关的特征内容,修剪冗余特征,突出显著目标区域。其公式为

$$q' = \psi^T(\sigma_1(W_x^T x' + W_g^T g + b_{xg})) + b_\psi \quad (1)$$

$$\alpha' = \sigma_2(q', g; \theta) \quad (2)$$

式中:  $\sigma_1$  是 RELU 非线性激活函数;  $\sigma_2$  是归一化 sigmoid 函数,将门控系数范围控制在  $[0,1]$ 。因此,AG 可以由以下参数描述:线性变换  $W_x$  和  $W_g$ ; 偏置  $b_\psi$ , 这里的线性变换采用  $1 \times 1 \times 1$  卷积实现。

本文在 VGG16 网络中采用 AG 模块,该网络经过 5 次下采样,得到大小为  $7 \times 7 \times 512$  的特征图,该特征图即为选通信号  $g$ 。由于影响分类准确率的浅层特征可能分布在网络的不同层次,因此,AG 模块将选通信号  $g$  提供的上下文全局信息分别与 VGG16 中第 9 层和第 13 层的浅层特征融合,在消除浅层特征中冗余的同时,获得更细尺度的病变抽象特征,进而融合多尺度特征。其中第 9 层和第 13 层的特征图大小分别为  $28 \times 28$  和  $14 \times 14$ 。最终网络得到  $14 \times 14$ 、 $28 \times 28$  以及  $7 \times 7$  等 3 种尺度的特征图,经过全局平均池化以及展平处理后,将 3 种尺度信息拼接起来,并通过分类层 (softmax) 进行分类。综上,本文通过在 VGG16 中加入 AG 模块,可以有效解决 OCT 图像中病变局部化这一问题,突出病变区域同时抑制背景噪声,让网络进一步挖掘到病变特征,提升分类准确率。

## 1.2 空洞空间金字塔模块

视网膜病变具有局部性,且病变区域在 OCT 图像中占据位置小。其中 Drusen 的此特点最为明显,其变位置小且模糊。此特点严重影响 OCT 图像的分类效果,目前提出的相关分类方法中,Drusen 类别的分类准确率最高为 92.5%,相比于视网膜 OCT 图像分类应用中其他 3 个类别的分类精度尚有待提高。

现阶段的经典分类网络通过加深网络深度,从而获得更大的感受野以及丰富的上下文信息。但是针对医学图像的特性,在分辨率小的特征图上进行分类将损失大量有用信息。为解决该问题,本文设计了双分支网络,在骨干网络经过 3 次下采样操作,加入并行扩张卷积空洞空间金字塔模块,以不同比例捕捉全局上下文信息。

扩张卷积 (dilated convolution) 由 Chen 等<sup>[22]</sup>于 2016 年提出,与普通卷积相比,扩张卷积引入扩张率这一参数,在基础卷积上加入间隔,卷积核各点间的间隔为扩张率减 1,如图 2 所示。

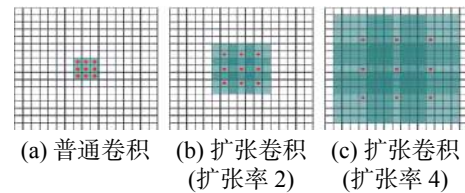


图 2 普通卷积和扩张卷积

Fig. 2 Ordinary convolution and dilated convolution

扩张卷积对应的卷积核实际大小以及感受野大小均大于普通卷积,但实际参数不变,以图 2 中不同扩张率的  $3 \times 3$  卷积为例,它们均只有 9 个点有参数,与普通的  $3 \times 3$  卷积参数相同,其余扩张位置的参数均为 0。扩张卷积对应的实际卷积核以及感受野大小计算公式为

$$RF_i = RF_{i-1} + (k-1) \times s \quad (3)$$

式中:  $k$  为卷积核尺寸;  $s$  是步长, 是上一层感受野大小。通过加入扩张卷积, 能够在不进行下采样操作的前提下, 同样获得更大的感受野。在大尺度特征图上实现细节信息的保留以及丰富上下文信息的获取。更好地保留了较小病变的形状以及轮廓特征, 有利于提升小目标分类精度。本文在 VGG16 网络中采用的空洞空间金字塔模块 (atrous spatial pyramid pooling, ASPP) 如图 3 所示。

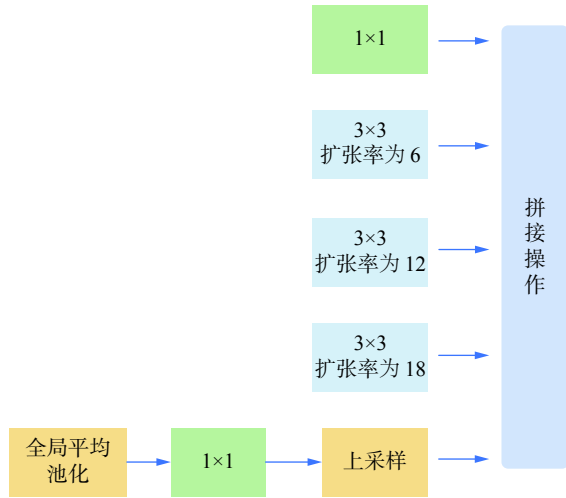


图 3 空洞空间金字塔模块

Fig. 3 Atrous spatial pyramid pooling module

骨干网络经过 3 次下采样后的特征图作为该模块的输入, 分别经过并行的  $1 \times 1$  卷积以及 3 个

扩张率分别为 6、12、18 的扩张卷积。为了融入全局上下文信息, 该模块采用了图像级特征, 通过对该模块的输入进行全局平均池操作, 并将得到的图像级特征输入到  $1 \times 1 \times 256$  的卷积核中, 然后经过双线性插值将特征上采样到所需的尺寸。最后将这 4 部分特征拼接输入到  $1 \times 1$  卷积进一步加强特征提取。

综上, 本文通过采用不同扩张率的空洞空间金字塔模块, 在不减小特征图大小的同时, 按不同比例捕捉上下文信息, 同时扩大感受野。在特征提取过程中, 不会因为过度下采样而损失病变信息, 同时又利用扩张卷积和图像级特征融合了全局信息, 较好地提升了网络对小目标 Drusen 病变的特征提取能力。

### 1.3 整体模块设计以及网络架构

基于门控注意力机制和空洞空间金字塔模块两个创新点, 本文提出了一种基于改进 VGG16 的双分支多尺度特征融合网络, 如图 4 所示。网络经过 3 次下采样后分成两个分支, 3 次下采样后得到的大小为  $28 \times 28 \times 512$  的特征图作为接下来两路分支的输入。一路分支继续下采样, 得到最深层的特征  $g$  作为选通信号, 为第 3 次、第 4 次下采样后的特征图提供上下文信息, 修剪浅层特征中的冗余信息, 突出病变区域显著特征。

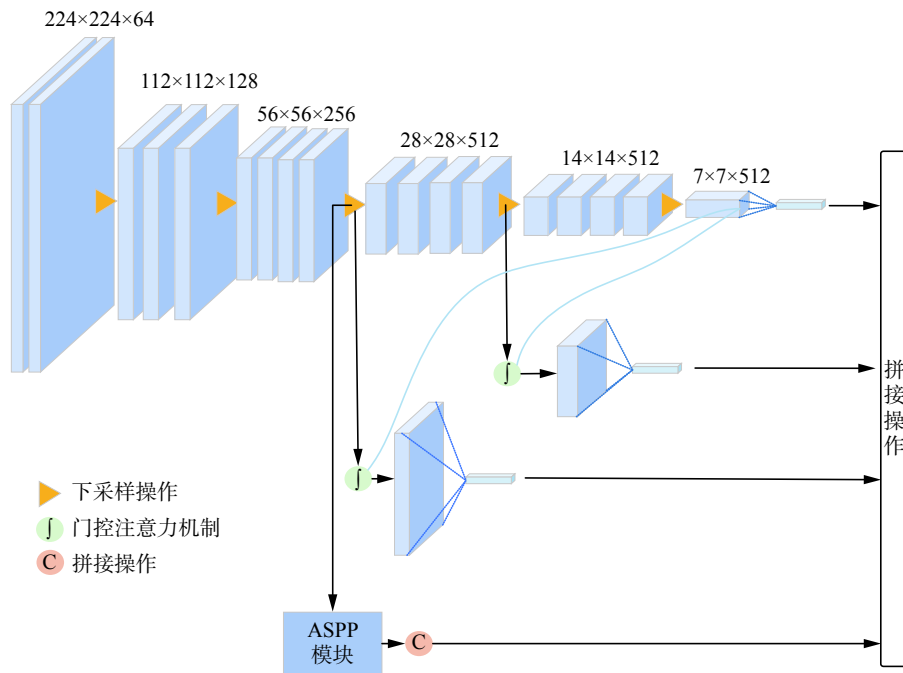


图 4 本文方法演示图

Fig. 4 Method demonstration diagram of this paper

另一路分支进入空间空洞金字塔模块, 分别进行不同扩张率的扩张卷积, 输出特征图大小为

$28 \times 28 \times 512$ , 尺度不变。在 AG 模块中, 选通信号  $g$  和第 3 次下采样后特征图融合得到的特征和空

间空洞金字塔模块的输出大小维度均相同,将两部分特征融合,进一步获得融合了选通信息以及多尺度信息的特征。并且,该融合后的特征为大尺度特征,实现了让网络在分辨率大的特征图上进行分类。为了让小目标病变获得良好的分类效果,需要网络获取丰富的病变区域信息,包括病变的形状、大小特征等。这种特征通常蕴藏在浅层网络中,但是由于病变区域过小,这些浅层特征会在特征提取过程中,损失大量病变区域的细节信息。本文提出的方法不仅融合了多尺度特征,还通过在大分辨率特征图上进行分类避免了细节特征的丢失,有效解决了现有方法对小目标病变分类效果不佳这一问题。

## 2 实验结果及分析

为验证本文提出方法的有效性与先进性,实验部分主要做了以下两个方面的工作:1)消融实验,本文提出的两个创新点分别引入实验以及最终的改进网络与基线网络的实验效果对比实验;2)与现有代表性算法的对比实验。

### 2.1 实验环境

本文使用的实验环境如表 1 所示。本实验采用 SGD 优化算法,一共训练 150 个 epoch,初始学习率设置为 0.001,学习率衰减采用指数衰减,衰减底数 gamma 设置为 0.98。

表 1 实验环境配置

Table 1 Experimental environment configuration

配件	参数
操作系统	Ubuntu 16.04
显卡	GeForce GTX 2080 Ti
内存/GB	16
开发工具	Pytorch1.2.0 Python3.6.2

### 2.2 数据集

本文使用的数据集是 Kaggle 平台提供的开源视网膜 OCT 病变图像,该数据集由加利福尼亚大学圣地亚哥分校(UCSD)于 2017 年公开。该数据集包含 4 种类别,分别是玻璃膜疣(Drusen)、脉络膜新生血管(CNV)、糖尿病黄斑水肿(DME)和正常类别,如图 5 所示。数据集包含训练集和测试集,训练集中 4 种类别分别包含 8 616、37 205、11 348、26 315 张图片。测试集由每类 250 张图片组成,共 1 000 张 OCT 图像。本文按照 8:2 的比例将训练集划分为训练集和验证集。

### 2.3 评价指标

本文使用准确率(Accuracy)、召回率(Recall)精确率(Precision)、特异性(Specifity)作为视网膜

OCT 分类任务的评价指标,具体公式如式(4)~(7)所示。

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

$$\text{Recall} = \frac{TP}{TP + FN} \quad (5)$$

$$\text{Precision} = \frac{TP}{TP + FP} \quad (6)$$

$$\text{Specifity} = \frac{TN}{FP + TN} \quad (7)$$

式中:TP 是将正样本正确分类的个数;TN 为将负样本正确分类的个数;FP 为将正样本分类错误的个数;FN 为将负样本分类错误的个数。本文中的视网膜分类任务属于多分类任务,这里的正样本是指定的某一特定类别,例如玻璃疣,而此时的负样本为除玻璃疣外的其他 3 种类别。同时,本文绘制了 4 种类别的混淆矩阵,可直观看出各类别的分类情况以及与基线网络分类情况的对比。

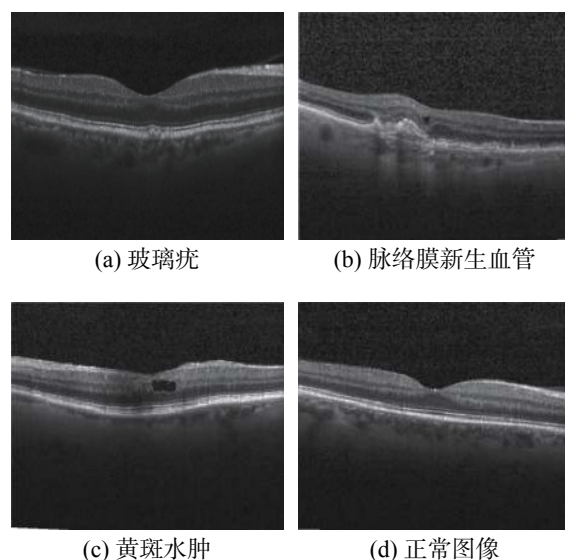


图 5 视网膜 OCT 图像示例

Fig. 5 Retinal OCT image example

### 2.4 实验分析

#### 2.4.1 算法的有效性验证

为了验证本文提出方法的有效性,这里对改进后的网络(引入 AG 模块和空间空洞金字塔模块)与只加入 AG 模块以及基线网络进行了消融实验,实验在同样的数据集、实验环境以及相同的网络参数配置上进行,实验结果如表 2 所示。

由表 2 可以看出,加入 AG 模块后,网络的识别率较基线网络提高了 1.9%,由此看出通过引入 AG 模块,网络更好地学习了病变区域特征,降低了大量背景冗余的影响。在此基础之上,加入本文的第二个改进点,网络的准确率进一步提高到 97.9%,较基线网络提高了 3.7%。其中 Drusen 病



变有了明显的提高, 提高了 1.5%。由该实验结果可以看出, 加入扩张卷积使网络在大尺度特征图上进行分类, 让小目标病变的细节信息不会随特征提取过程的深入而损失, Drusen 的识别效果显著地提升。

表 2 算法有效性验证实验  
Table 2 Algorithm validation experiment

方法	类别	精确率	召回率	特异性	总体准确率
VGG16	CNV	0.855	0.992	0.944	0.942
	DME	0.961	0.98	0.987	
	Drusen	0.973	0.852	0.992	
	Normal	1.0	0.944	1.0	
VGG16+AG	CNV	0.904	0.98	0.965	0.961
	DME	0.969	1.0	0.989	
	Drusen	0.978	0.892	0.993	
	Normal	1.0	0.972	1.0	
VGG16+ AG+ASPP	CNV	0.992	0.992	0.997	<b>0.979</b>
	DME	0.958	0.996	0.985	
	Drusen	<b>0.988</b>	<b>0.952</b>	<b>0.996</b>	
	Normal	0.98	0.976	0.993	

为了进一步直观地看出网络对 4 个类别的分类效果, 这里绘制了基线网络以及加入 AG 模块和加入 AG 模块、空间空洞金字塔模块的混淆矩阵, 如图 6 所示。在混淆矩阵中, 对角线上的数字代表每个类别正确分类的样本个数, 对角线上数值越大说明分类越准确。从图 6 可以看出, 加入 AG 模块后, 虽然各类别分类效果有所提高, 但是由于 Drusen 病变小而模糊, 且与 CNV 表现相似, 对 Drusen 的分类效果相比于其他 3 个类别差。再加入空间空洞金字塔模块后, Drusen 的分类效果有了明显的改善。

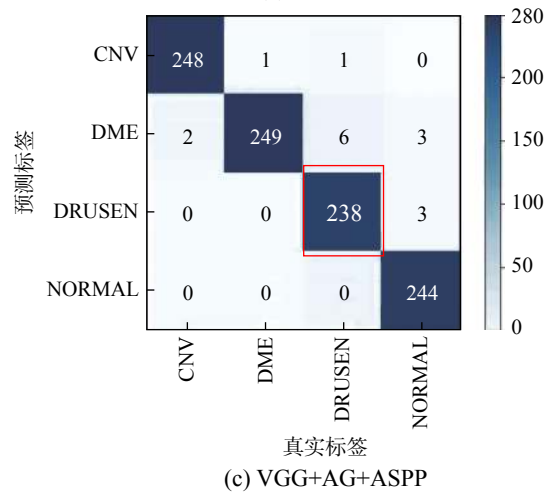
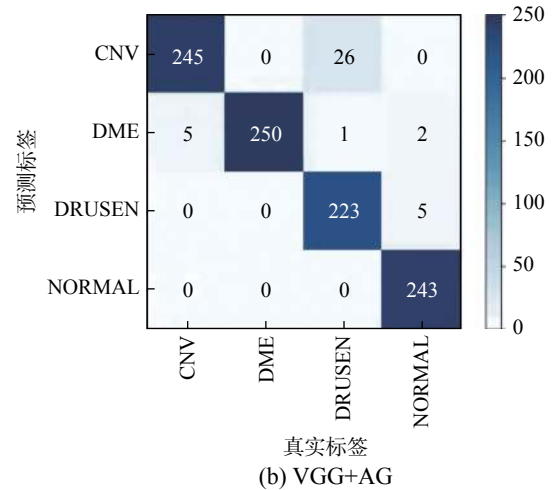
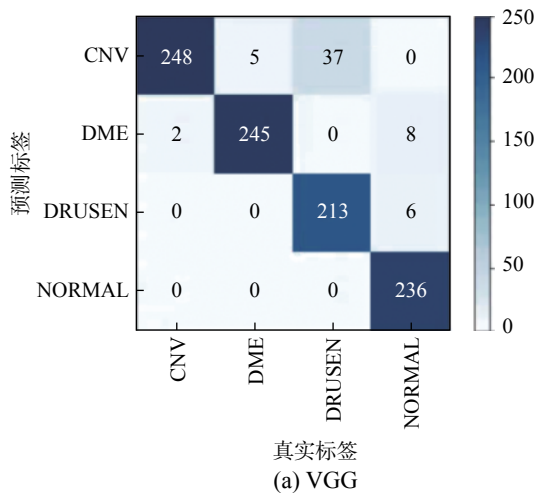


图 6 混淆矩阵  
Fig. 6 Confusion matrix

#### 2.4.2 算法的先进性验证

为验证本文方法的先进性, 将本文提出算法与现阶段具有代表性的基于深度学习的视网膜 OCT 分类算法进行对比, 对比结果如表 3 所示。首先, 从表 3 可以看出, 现有方法对 Drusen 的分类准确率不高, 该类别的最高准确率仅达 92.5%, 远低于另外 3 种类别的分类精度。这是因为现有方法针对小目标病变的特征提取能力不强, 在特征提取过程中, Drusen 这一病变的细节信息损失严重。针对这一问题, 本文通过加入门控注意力机制突出病变区域信息, 以及加入空间空洞金字塔模块减小网络下采样的次数, 在大分辨率的特征图上进行分类, 从而保留小目标病变区域的细节特征, 本文对 Drusen 病变的分类准确率较现有文献有了显著提升, 较文献 [17-24] 分别提高了 6.3% 和 9.8%。同时, 本文方法在 CNV、DME 两种类别病变上也获得了最好的分类效果。本文对 Normal 类别的分类准确率较文献 [17] 低 0.5%, 但本文方法的整体分类准确率依然是现有视网膜 OCT 图像分类任务中最高的, 相比于文献 [17] 提出的轻

量化视网膜 OCT 图像分类网络, 本文算法准确率提高了 0.9%, 较文献 [23] 提出的多层次可选择卷积分类方法准确率提高了 2.51%, 较文献 [24] 提出的迁移学习方法准确率提高了 1.4%, 较文献 [25] 提出的基于通道注意力机制的分类方法提升了 0.4%, 这充分验证了本文改进网络模型的先进性。

表 3 算法先进性验证实验

Table 3 Algorithm advanced verification experiment %

方法	Drusen	CNV	DME	Normal	整体识别准确率
文献[17]	92.5	96.9	95.8	98.5	97
文献[23]	—	—	—	—	95.39
文献[24]	89	98	92	96	96.5
文献[25]	—	—	—	—	97.5
本文算法	98.8	99.2	95.8	98	97.9

### 3 结束语

本文提出了一种应用于视网膜 OCT 图像分类任务的双分支多尺度特征融合网络。通过加入门控注意力机制模块, 让深层特征作为选通信号修剪浅层特征中的冗余信息, 突出 OCT 图像中的病变区域, 消除背景噪声的影响。同时引入空洞空间金字塔模块, 利用并行扩张卷积代替下采样过程, 在不降低特征图大小的前提下按不同比例捕捉上下文信息, 获得更大的感受野。本文提出的方法有效解决了现有方法中因 Drusen 病变位置小、形态模糊导致的该类别分类难度大、精度低的问题, 进一步提高了计算机辅助诊断的能力。针对目前视网膜黄斑病变患者多、医生诊断压力大以及医疗行业逐渐智能化的现状, 本文具有显著的研究价值。利用深度学习技术特征提取能力强、可处理大量数据的优势, 对本课题进一步深入研究, 可以让计算机辅助诊断技术提升至人类专家水平, 在实际应用中辅助人类医师更加高效、准确地诊断疾病, 同时可以挖掘大量医疗数据中的巨大价值, 实现医疗系统智能化的转变。

### 参考文献:

- [1] ROMERO-ARCA P. Current status in diabetic macular edema treatments[J]. *World journal of diabetes*, 2013, 4(5): 165–169.
- [2] 刘艳, 隋虹, 罗志忠, 等. 糖尿病性视网膜病变与糖尿病控制状态研究 [J]. *中国糖尿病杂志*, 2002, 10(1): 37–39, 36.  
LIU Yan, SUI Hong, LUO Zhizhong, et al. A study on the relationship between diabetic retinopathy and the state of diabetes mellitus control[J]. *Chinese journal of diabetes*, 2002, 10(1): 37–39, 36.
- [3] OTANI T, KISHI S, MARUYAMA Y. Patterns of diabetic macular edema with optical coherence tomography[J]. *American journal of ophthalmology*, 1999, 127(6): 688–693.
- [4] HUANG D, SWANSON E A, LIN C P, et al. Optical coherence tomography[J]. *Science*, 1991, 254(5035): 1178–1181.
- [5] DREXLER W, FUJIMOTO J G. State-of-the-art retinal optical coherence tomography[J]. *Progress in retinal and eye research*, 2008, 27(1): 45–88.
- [6] PIZURICA A, JOVANOVIĆ L, HUYSMANS B, et al. Multiresolution denoising for optical coherence tomography: a review and evaluation[J]. *Current medical imaging reviews*, 2008, 4(4): 270–284.
- [7] BEAUREPAIRE E, BOCCARA A C, LEBEC M, et al. Full-field optical coherence microscopy[J]. *Optics letters*, 1998, 23(4): 244.
- [8] GOH J K H, CHEUNG C Y, SIM S S, et al. Retinal imaging techniques for diabetic retinopathy screening[J]. *Journal of diabetes science and technology*, 2016, 10(2): 282–294.
- [9] 袁非. 糖尿病性视网膜病变及其早期检测 [J]. 国外医学眼科学分册, 1997(4): 219–224.
- [10] ADHI M, DUKER J S. Optical coherence tomography: current and future applications[J]. *Current opinion in ophthalmology*, 2013, 24(3): 213–221.
- [11] DOI K. Computer-aided diagnosis in medical imaging: Historical review, current status and future potential[J]. *Computerized medical imaging and graphics*, 2007, 31(4/5): 198–211.
- [12] ASIRI N, HUSSAIN M, AL ADEL F, et al. Deep learning based computer-aided diagnosis systems for diabetic retinopathy: a survey[J]. *Artificial intelligence in medicine*, 2019, 99: 101701.
- [13] KOPROWSKI R, TEPER S, WRÓBEL Z, et al. Automatic analysis of selected choroidal diseases in OCT images of the eye fundus[J]. *Biomedical engineering online*, 2013, 12: 117.
- [14] 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): 579–592.
- [15] HASSAN B, RAJA G, HASSAN T, et al. Structure tensor based automated detection of macular edema and central serous retinopathy using optical coherence tomography images[J]. *Journal of the Optical Society of*



- America A, *Optics, image science, and vision*, 2016, 33(4): 455–463.
- [16] CIULLA T A, AMADOR A G, ZINMAN B. Diabetic retinopathy and diabetic macular edema: pathophysiology, screening, and novel therapies[J]. *Diabetes care*, 2003, 26(9): 2653–2664.
- [17] 张添福, 钟舜聪, 连超铭, 等. 基于深度学习特征融合的视网膜图像分类 [J]. 激光与光电子学进展, 2020(24): 258–265.  
ZHANG Tianfu, ZHONG Shuncong, LIAN Chaoming, et al. Deep learning feature fusion-based retina image classification[J]. *Laser & optoelectronics progress*, 2020(24): 258–265.
- [18] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [EB/OL]. (2014-09-04)[2021-01-01]. <https://arxiv.org/abs/1409.1556>.
- [19] HU Jie, SHEN Li, SUN Gang. Squeeze-and-excitation networks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Salt Lake City, USA. IEEE, 2018: 7132–7141.
- [20] WOO S, PARK J, LEE J Y, et al. CBAM: convolutional block attention module[M]//Computer Vision-ECCV 2018. Cham: Springer International Publishing, 2018: 3–19.
- [21] WANG Qilong, WU Banggu, ZHU Pengfei, et al. ECA-net: efficient channel attention for deep convolutional neural networks[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Seattle, USA. IEEE, 2020: 11531–11539.
- [22] CHEN L C, PAPANDREOU G, KOKKINOS I, et al. DeepLab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFs[J]. *IEEE transactions on pattern analysis and machine intelligence*, 2018, 40(4): 834–848.
- [23] 朱纳, 李明. 多层次可选择核卷积用于视网膜图像分类 [EB/OL]. (2021-10-11)[2021-12-01]. <https://kns-cnki-net.wvpin.hrbeu.edu.cn/kcms/detail/50.1181.N.20211009.1903.006.html>.  
ZHU Na, LI Ming. Multi-level selective kernel convolution for retinal image classification[EB/OL]. (2021-10-11)[2021-12-01]. <https://kns-cnki-net.wvpin.hrbeu.edu.cn/kcms/detail/50.1181.N.20211009.1903.006.html>.
- [24] BHADRA R, KAR S. Retinal disease classification from optical coherence tomographical scans using multilayered convolution neural network[C]//2020 IEEE Applied Signal Processing Conference. Kolkata, India. IEEE, 2020: 212–216.
- [25] 于海琛. 基于 SE-Block 的视网膜疾病分类方法研究 [D]. 长春: 吉林大学, 2019.  
YU Haichen. Research on classification of retinal diseases based on SE-block[D]. Changchun: Jilin University, 2019.

#### 作者简介:



韩璐, 硕士, 主要研究方向为图像识别、深度学习。



毕晓君, 教授, 博士生导师, 主要研究方向为信息智能处理、数字图像处理、智能优化算法及机器学习。主持国家自然科学基金面上项目 2 项、科技部国际合作项目面上项目 1 项、教育部博士点基金项目 1 项、工业和信息化部海洋工程装备科研项目子项目 1 项、民品横向课题 1 项, 获国家专利 8 项。发表学术论文 170 余篇, 出版学术专著 3 部。