

基于轻量级卷积网络的视网膜病变自动检测

王凌霄¹, 杨军¹, 王文赛¹, 李婷^{1,2*}

¹中国医学科学院生物医学工程研究所, 天津 300192;

²北京脑科学与类脑研究中心, 北京 102206

摘要 光学相干断层扫描技术是临床中检测视网膜病变的主要手段,但人工诊断的模式存在主观性强、效率低的问题,为此提出一种轻量级卷积神经网络用于视网膜病变的自动检测。所提网络由两种模块组成,第1种模块将空洞卷积与深度可分离卷积相结合以降低参数量;第2种模块利用分解卷积方法,通过将常规卷积层分解成多层不对称卷积的方式延展深度。两种模块交叉组合构成特征提取器,使用Softmax函数作为分类器,获得了44层深、参数量为9.2 MB的轻量级模型。所提网络在测试集上的准确率、敏感性、特异性、接收者操作特征曲线下面积分别达到0.980、0.954、0.987和0.997。可视化结果表明,模型诊断依据与眼科专家相一致。这些结果表明,所提网络能够准确地实现视网膜疾病的自动检测。

关键词 医用光学; 光学相干断层成像; 卷积神经网络; 视网膜病变; 图像分类

中图分类号 TP391.4

文献标志码 A

doi: 10.3788/LOP202259.0617017

Automatic Detection of Retinal Diseases Based on Lightweight Convolutional Neural Network

Wang Lingxiao¹, Yang Jun¹, Wang Wensai¹, Li Ting^{1,2*}

¹Institute of Biomedical Engineering, Chinese Academy of Medical Sciences, Tianjin 300192, China;

²Chinese Institute for Brain Research, Beijing, Beijing 102206, China

Abstract One major method for detecting retinopathy in clinics is optical coherence tomography. However, this manual diagnostic model is affected by strong subjectivity and low efficiency. Therefore, this paper proposes a lightweight convolutional neural network for the automatic detection of retinopathy. The proposed network consists of two modules. The first module combines atrous convolutions and depth wise separable convolutions to reduce the number of parameters; the second module uses the decomposition convolution method to extend the depth by decomposing the conventional convolution layer into multilayer asymmetric convolution. Both modules are combined to form a feature extractor, and the Softmax function is used as the classifier to obtain a lightweight model with 44 layers deep and 9.2 MB parameters. The accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve of the proposed network on the test set are 0.980, 0.954, 0.987, and 0.997, respectively. The visualization results show that the diagnostic basis of the model is consistent with that of ophthalmologists. These results show that the proposed network can accurately automate retinal disease detection.

Key words medical optics; optical coherence tomography; convolutional neural network; retinal disease; image classification

收稿日期: 2022-01-13; 修回日期: 2022-01-28; 录用日期: 2022-02-10

基金项目: 国家自然科学基金面上项目(81971660)、天津杰出青年基金(20JCJQIC00230)、北京脑科学与类脑研究中心科研合作项目(2020-NKX-XM-14)、京津冀基础研究合作专项(19JCZDJC65500(Z))、中国医学科学院医学与健康科技创新工程(2021-I2M-1-042, 2021-I2M-1-058)

通信作者: *liting@bme.cams.cn

1 引言

人体视网膜病变可能会导致严重视力丧失甚至失明^[1],其健康状况受到多种病理因素影响,典型病变有糖尿病性黄斑水肿(DME)、脉络膜新生血管(CNV)和玻璃疣(DRUSEN)。这些病变的临床诊断依赖于对不同程度的积液、囊肿、渗出物及形态异常等结构的检测^[2-4]。如果能及时地发现病灶并采取治疗措施,就能够预防或延缓视网膜病变的恶化,从而最大程度保护患者的视力。

临床中用于视网膜检查的手段通常有眼底照相、荧光素眼底血管造影(FFA)、光学相干断层扫描(OCT)等。眼底照相易受屈光介质浑浊度的影响,难以辨别早期微小病变。FFA无法可视化单个视网膜层结构,主要用于糖尿病视网膜病变的筛查。OCT具有高分辨率、无侵入性的特点^[5],能够清晰显示多层视网膜细胞层横截面结构,被作为DME等视网膜病变诊断的金标准。该技术在角膜厚度^[6]和视网膜神经纤维层厚度的测量中具有不可替代的优势,在临床中也常用于青光眼的诊断^[7]。然而,视网膜疾病临床诊断主要基于眼科医生对OCT图像序列中每幅图像的主观分析,这种诊断模式效率低下,且观察者之间和观察者内部的巨大差异往往影响诊断结果的准确性。因而发展一种高性能的OCT图像自动分析技术对于视网膜疾病的准确、及时诊断至关重要。

卷积神经网络(CNN)作为一种深度学习模型,可快速查看大量数据并自动提取图像特征,已被用于视网膜OCT图像病变检测和分类的研究^[8-11]。Kermany等^[12]通过基于Inception V3模型的迁移学习方法,获得了一种能够区分DME、CNV、DRUSEN和正常情况4种类型的OCT图像的分类算法,且整体准确率为96.6%。连超铭等^[13]对Inception V3预训练模型进行微调,进一步提高了分类准确率。Li等^[14]提出一种基于VGG 16预训练模型的方法,准确率达到98.6%。Rasti等^[15]设计了一种深度集成模型,该模型通过检测不同尺度的视网膜病变,实现了高性能的OCT图像分类。Fang等^[16]提出了一种注意力机制卷积神经网络,该网络利用局部病变区域的信息改进了视网膜OCT图像的分类过程。虽然这些方法表现出了不错的分类性能,但由于模型结构复杂性相对较高、体量较大,训练中容易受到过拟合的影响,且庞大的参

数量也对计算机性能提出了较高的要求。这些“重量级”特性限制了模型的可移植性,进而降低了广泛应用的可能性。

为此,本文提出了一种高效的轻量级卷积神经网络。所提网络能够在大幅降低参数量和计算复杂度的同时保持高准确性,且轻量化特点便于模型的训练和移植,为临床落地提供了可行性。

2 研究方法

2.1 数据集处理

本实验组使用公开的UCSD数据集^[12,16-17]验证模型的性能。该数据集以TIFF格式提供了109309幅视网膜OCT二维图像,这些图像均来自成年患者,图像标签由眼科专家严格标注并作为分类任务的金标准,其中训练集有108309幅图像,测试集有1000幅图像。本实验将原数据集中的测试集图像仍作为测试集,并按9:1的比例将训练集随机划分为新的训练集和验证集。数据集的具体划分如表1所示。

表1 视网膜OCT图像数据集划分

Category	Training dataset	Validation dataset	Testing dataset	Total
CNV	33485	3720	250	37205
DME	10214	1134	250	11348
DRUSEN	7755	861	250	8616
NORMAL	46026	5114	250	51140
Total	97480	10829	1000	109309

在重新划分的训练集上对分类模型进行参数拟合,并在验证集上对参数进行监督优化,最后在测试集上评估模型的性能。为了增加数据集的异质性,防止训练中出现过拟合现象,对训练集图像采用了两种数据增强策略:1)通过旋转图像引入数据扰动;2)通过对图像进行水平翻转保持左右眼样本数据量的平衡。图1分别展示了数据集中CNV、DME、DRUSEN、正常4种视网膜OCT的图像,图中箭头和十字形表示病变部位。

2.2 模型设计

2.2.1 空洞-深度可分离卷积模块

深度可分离卷积层(SeperableConv)通过将通道的卷积与空间的卷积完全分离,能够在不增加参数量的前提下提高特征学习能力^[18]。然而,由于网络宽度的增加,深度可分离卷积层的参数量并未减

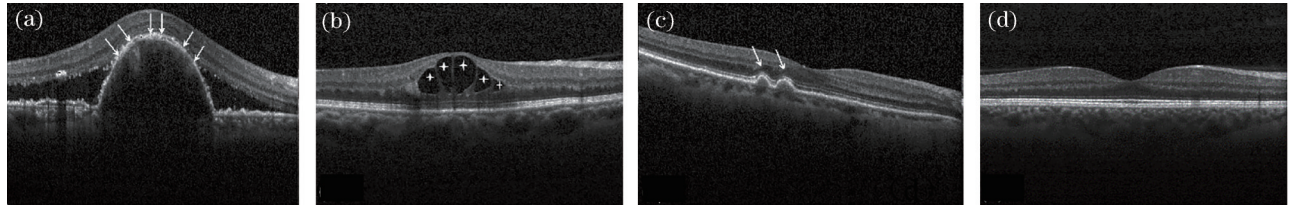


图1 样本图像示例。(a)脉络膜新生血管;(b)糖尿病性黄斑水肿;(c)玻璃疣;(d)正常
Fig. 1 Sample image examples. (a) CNV; (b) DME; (c) DRUSEN; (d) NORMAL

小,因而仍有较高的计算复杂度。空洞卷积通过设置扩张率在小尺寸卷积核元素间插入空格来扩大卷积范围,能够在减小参数量和计算复杂度的同时获得与大尺寸卷积核相同的“视野”和特征学习能力。为了平衡模型的计算复杂度与性能,本实验组将空洞卷积核运用到深度可分离卷积层中,设计了空洞-深度可分离卷积(Dilated SepConv)模块,如[图2(a)]所示,垂直点划线左侧为原深度可分离卷

积模块,右侧为所设计的空洞-深度可分离卷积模块,所设计模块中的卷积层使用扩张率为2的空洞卷积核(2×2, Rate is 2)替换原模块中的常规卷积核(3×3),卷积前通过激活函数(ReLu)激活,卷积后通过批规范化(BatchNorm)标准化输出,由最大池化层(Max Pooling)对前一个卷积层的输出进行下采样以减小特征图尺寸。所设计模块保留了残差连接,并调整了1×1卷积层的通道数。

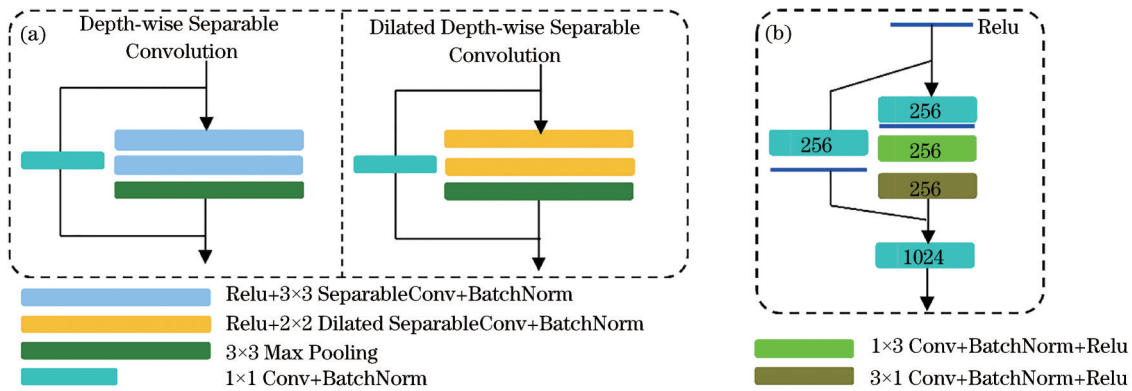


图2 模块结构示意图。(a)空洞-深度可分离卷积模块;(b)残差单元

Fig. 2 Schematic diagrams of module structure . (a) Dilated SepConv ; (b) Residual Unit

2.2.2 残差单元

卷积神经网络的深度对特征学习性能有极大影响,如何合理地增加网络深度来提升模型性能正成为深度学习领域的一个热点^[19-21]。Inception V3模型^[22]基于Inception模块搭建构构成,通过将常规卷积层(单层)分解为不对称卷积层(多层)的方式延展网络深度,并结合基于多分支并行的多层特征融合技术,在视网膜OCT图像病变检测任务中具有非常不错的表现^[12-13]。然而,多分支并行的结构特点也使得Inception V3模型的宽度和参数量大幅增加,在模型较深的情况下,不利于训练,甚至会导致模型性能的下降。

为此,本实验组借鉴Inception模块延展网络深度的方式,同时剪去模块中的多个分支,并对剩余分支中的各个卷积层通道数进行调整,同时在不对称卷积层后引入批规范化和激活函数,并引入残差连接^[20],

获得了改进的残差单元(Residual Unit),如[图2(b)]所示,图中矩形框内的数字表示调整的卷积层通道数。

本实验组将4个空洞-深度可分离卷积模块和8个残差单元模块以交叉堆叠的方式构成特征提取器,使用Softmax损失函数作为分类器,对CNV、DME、DRUSEN、NORMAL 4种视网膜情况进行检测,最终的病变检测模型如图3所示。视网膜OCT图像输入网络后,由特征提取器对其进行特征识别,并由分类器输出该图像从属于4种情况的概率值。

2.3 评价方法

为评估模型对4种视网膜OCT图像的分类性能,采用准确率(Accuracy)、灵敏度(Sensitivity)、特异性(Specificity)、接收者操作特征曲线(ROC)及接收者操作特征曲线下面积(Auc)作为评价指标,表

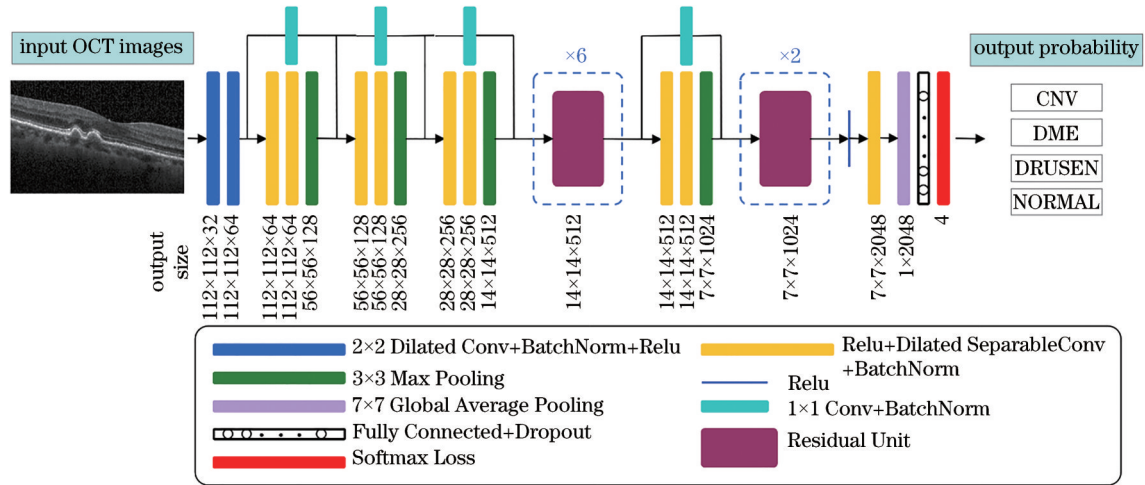


图 3 所提模型结构示意图

Fig. 3 Schematic diagram of proposed model structure

达式分别为

$$R_{\text{Accuracy}} = \frac{N_{\text{TP}} + N_{\text{TN}}}{N_{\text{TP}} + N_{\text{TN}} + N_{\text{FP}} + N_{\text{FN}}}, \quad (1)$$

$$R_{\text{Sensitivity}} = \frac{N_{\text{TP}}}{N_{\text{TP}} + N_{\text{FN}}}, \quad (2)$$

$$R_{\text{Specificity}} = \frac{N_{\text{TN}}}{N_{\text{TN}} + N_{\text{FP}}}, \quad (3)$$

式中: N_{TP} 、 N_{TN} 、 N_{FP} 、 N_{FN} 分别代表真阳性、真阴性、假阳性、假阴性图像的数量。ROC 和 Auc 根据测试结果自动绘制。

此外,为评估模型的整体分类性能,将每种病变的分类准确率加和求均值获得整体准确率(Ov-Acc),按相同方式获得整体灵敏度(Ov-Se)和整体特异性(Ov-Sp),并与眼科专家和其他模型的结果进行比较。

3 结果分析

3.1 测试结果

表 2 为测试结果,表中详细给出了所提模型对每种情况的准确率、敏感性和特异性。所提模型在正常和 DME 病变检测中具有不错的表现,分类准

表 2 所提模型对各类视网膜图像的分类结果

Table 2 Classification results of proposed model for various retinal images

Category	Accuracy	Sensitivity	Specificity
CNV	0.966	1.000	0.947
DME	0.986	0.944	1.000
DRUSEN	0.968	0.872	1.000
NORMAL	1.000	1.000	1.000

确率分别达到了 0.986 和 1。另外,所提模型对 CNV 病变检测的灵敏度,对 DME、DRUSEN 病变检测的特异性值为 1,高灵敏度和高特异性分别与临床中的低漏诊率与低误诊率相对应,表明所提模型可以有效检测出具有 CNV 病变的视网膜 OCT 图像,并准确地将 DME 或 DRUSEN 病变图像与其他病变图像区分开来。

ROC 图以假阳率(FPR)为横轴、真阳率(TPR)为纵轴,曲线距离左上角越近,表明模型分类性能越优。实验结果如图 4 所示,其中[图 4(b)]为[图 4(a)]方框区域的局部放大。

从图 4 可以看出,所提模型对每种情况的 ROC 及平均 ROC 曲线均非常靠近左上角,表明模型对 4 种视网膜 OCT 图像特征均有较强的识别能力。进一步计算 Auc 值以量化分类表现,所提模型对各类情况的 Auc 值均高于 0.99,平均 Auc 值为 0.997,其中对正常视网膜的 Auc 值达到了 1。

本实验组进一步计算了整体准确率(Ov-Acc)、整体敏感性(Ov-Se)和整体特异性(Ov-Sp)以评估模型的整体分类性能,并与眼科专家和其他模型的评估结果进行了比较,结果如表 3 所示,其中粗体表示最优参数。从表 3 可以看出,所提模型对 4 类视网膜 OCT 图像检测的整体特异性达到了 0.987,优于 Inception V3 模型的检测结果;准确率达到 0.980,仅略低于 VGG 16 模型的 0.986;敏感性达到 0.954。实验结果表明,所提模型与 VGG16、Inception V3 等典型模型的性能相当,并在准确率和特异性方面优于眼科专家的评估表现。

表 3 中还给出了模型卷积层数量(Layers)和参

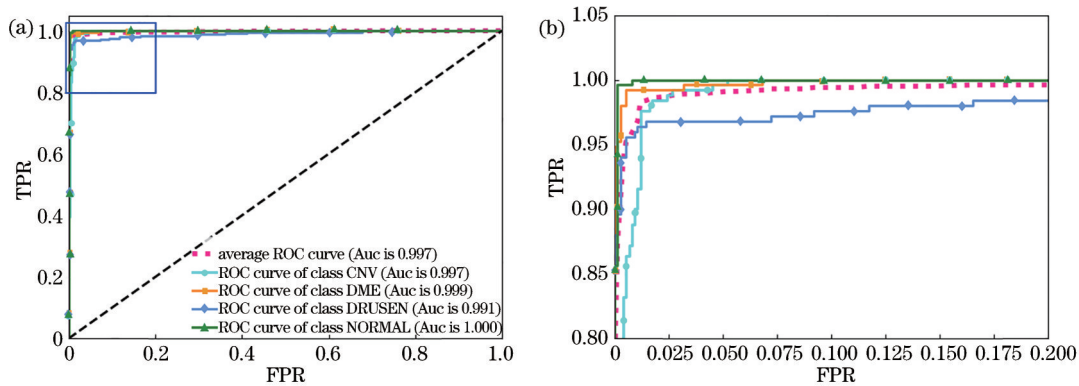


图 4 所提模型的 ROC 曲线。(a) ROC 曲线;(b)局部放大图

Fig. 4 ROC curves of proposed model. (a) ROC curve; (b) partial enlarged view

表 3 不同分类模型整体评价结果

Table 3 Overall evaluation results of different classification models

Model	Ov-Acc	Ov-Se	Ov-Sp	Layers	Parameter Count /MB	Ratio
Human Expert ^[12]	0.921	0.994	0.943			
VGG 16 ^[14]	0.986	0.978		16	138.3	8.6
RongheNet ^[23]	0.972	0.970		6	54.5	9.1
Inception V3 ^[12]	0.966	0.978	0.974	42	23.6	0.6
Proposed model	0.980	0.954	0.987	44	9.2	0.2

数量(Parameter Count),并计算了参数量与卷积层数量的比率(Ratio),比率越小表示每卷积层运算占用的参数量越少、计算损耗越低,在模型分类性能相当的情况下,模型越高效。从表 3 可以看出,在所有模型中 VGG 16 模型的参数量最高,为 138.3 MB,庞大的参数量导致该模型计算复杂度高且检测速度慢。RongheNet 模型^[23]的参数量与 VGG 16 模型相比有了显著下降,然而,该模型在仅使用 6 个卷积层结构的情况下仍占用了 54.5 MB 的参数量,因此具有最低的模型效率。所提模型的卷积层数、参数量、比率分别为所在列的最高、最低、最低值,其中参数量仅为 9.2 MB,仅为 Inception V3 模型参数量的 38%,是 VGG 16 模型参数量的 6.7%。

以上分析表明,所提模型实现了轻量化,并在仅使用 VGG 16 模型 6.7% 的参数量的条件下,保持了与 VGG 16 模型同等水平的准确率,在模型精度与计算复杂度之间取得了较好的平衡。这得益于所设计的模型结构,空洞-深度可分离卷积模块有效降低了模型参数,并保持了较好的特征学习能力,多个堆叠的残差单元加深了模型深度,进一步提升了模型的性能。另外,所提模型在 GeForce GTX1050 Ti 显卡上对测试集中每幅图像的诊断时

间为 0.027 s,而使用迁移学习方法的 Inception V3 预训练模型对测试集每幅图像的预测时间约为 0.026 s^[23]。所提模型在诊断速度方面可与预训练模型相媲美,满足了实时诊断的需求。

3.2 模型决策可视化

深度学习模型常被看作“黑匣子”,为了了解模型的诊断过程,本实验组使用类激活热力图(CAM)对模型的决策过程进行可视化,以观察模型的诊断是否基于关键的临床特征。从测试集样本中各随机选择 1 例模型诊断与实际类别一致的 CNV、DME、DRUSEN 病变图像,使用类激活热力图将模型诊断过程中对图像的主要关注区域可视化,这些区域的特征即为模型做出分类决策的主要依据。[图 5(a)~(c)]分别是所提模型对 CNV、DME、DRUSEN 图像的分类激活热力图,最右侧色度条表示对模型诊断的影响程度,从上到下依次降低。图像表明,所提模型识别 CNV 病变时的主要依据是视网膜新生血管区域,识别 DME 病变时的主要依据是视网膜内积液区域,识别 DRUSEN 病变图像时的主要依据是视网膜层物质堆积起伏区域。可视化结果表明,所提模型能够准确地识别各类视网膜 OCT 病变图像的临床特征,并与眼科专家的诊断依据相一致,具备良好的可靠性。

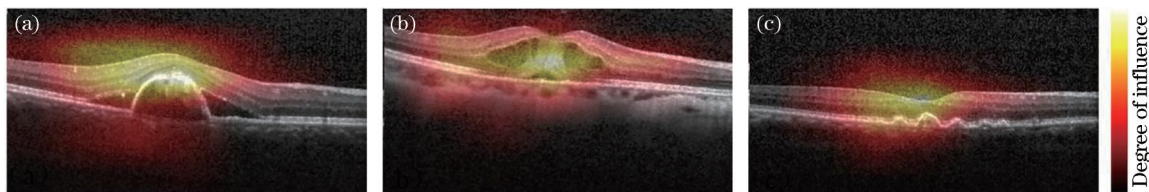


图5 类激活热力图。(a)脉络膜新生血管;(b)糖尿病性黄斑水肿;(c)玻璃疣

Fig. 5 CAM. (a) CNV; (b) DME; (c) DRUSEN

4 结 论

基于一种高效的轻量级卷积神经网络实现了视网膜OCT图像病变的自动检测。为了降低参数量和计算复杂度,基于深度可分离卷积设计了空洞-深度可分离卷积模块,同时通过堆叠多个改进残差单元的方式延展模型深度,进一步提高特征学习性能。所提模型在测试集上的整体准确率、敏感性、特异性分别达到0.980、0.954、0.987,参数量仅为9.2 MB,每幅图像的诊断为0.027 s。实验结果表明,所提模型较好地平衡了精度和计算复杂度,同时满足实时诊断的需求。所提模型能够降低移动端及医院部署的难度,为移动端实时检测和临床应用提供了可行性。

致谢 感谢国家自然科学基金面上项目(N81971660)、天津杰出青年基金(20JCJQIC00230)、北京脑科学与类脑研究中心科研合作项目(2020-NKX-XM-14)、京津冀基础研究合作专项(19JCZD-JC65500(Z))、中国医学科学院医学与健康科技创新工程(2021-I2M-1-042,2021-I2M-1-058)的支持。

参 考 文 献

- [1] 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.
- [2] 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.
- [3] Freund K B, Yannuzzi L A, Sorenson J A. Age-related macular degeneration and choroidal neovascularization [J]. *American Journal of Ophthalmology*, 1993, 115(6): 786-791.
- [4] Abdelsalam A, del Priore L, Zarbin M A. Drusen in age-related macular degeneration: pathogenesis, natural course, and laser photocoagulation-induced regression[J]. *Survey of Ophthalmology*, 1999, 44(1): 1-29.
- [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] Huang Q, Xia L K. Comparison of central corneal thicknesses measured with anterior segment optical coherence tomography and ultrasound pachymetry[J]. *Chinese Journal of Laser Medicine & Surgery*, 2018, 27(6): 370-374.
黄青,夏丽坤. AS-OCT和A超测量中央角膜厚度的比较[J]. *中国激光医学杂志*, 2018, 27(6): 370-374.
- [7] Jeoung J W, Choi Y J, Park K H, et al. Macular ganglion cell imaging study: glaucoma diagnostic accuracy of spectral-domain optical coherence tomography [J]. *Investigative Ophthalmology & Visual Science*, 2013, 54(7): 4422-4429.
- [8] Benaddy M, El Meslouhi O, Es-saady Y, et al. Handwritten tiffinagh characters recognition using deep convolutional neural networks[J]. *Sensing and Imaging*, 2019, 20(1): 1-17.
- [9] 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): 035005.
- [10] de Fauw J, 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): 1342-1350.
- [11] Perdomo O, Rios H, Rodríguez F J, et al. Classification of diabetes-related retinal diseases using a deep learning approach in optical coherence tomography[J]. *Computer Methods and Programs in Biomedicine*, 2019, 178: 181-189.
- [12] Kermany 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): 1200-1209.

- 1122-1131.
- [13] Lian C M, Zhong S C, Zhang T F, et al. Transfer learning-based classification of optical coherence tomography retinal images[J]. *Laser & Optoelectronics Progress*, 2021, 58(1): 011702.
连超铭, 钟舜聪, 张添福, 等. 光学相干断层扫描视网膜图像的迁移学习分类[J]. *激光与光电子学进展*, 2021, 58(1): 011702.
- [14] Li F, Chen H, Liu Z, et al. Fully automated detection of retinal disorders by image-based deep learning[J]. *Graefe's Archive for Clinical and Experimental Ophthalmology*, 2019, 257(3): 495-505.
- [15] 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.
- [16] Fang L Y, Wang C, Li S T, et al. Attention to lesion: lesion-aware convolutional neural network for retinal optical coherence tomography image classification [J]. *IEEE Transactions on Medical Imaging*, 2019, 38(8): 1959-1970.
- [17] Li F, Chen H, Liu Z, et al. Deep learning-based automated detection of retinal diseases using optical coherence tomography images[J]. *Biomedical Optics Express*, 2019, 10(12): 6204-6226.
- [18] Chollet F. Xception: deep learning with depthwise separable convolutions[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, July 21-26, 2017, Honolulu, HI, USA. New York: IEEE Press, 2017: 1800-1807.
- [19] He K M, Zhang X Y, Ren S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 770-778.
- [20] Szegedy C, Liu W, Jia Y Q, et al. Going deeper with convolutions[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition, June 7-12, 2015, Boston, MA. New York: IEEE Press, 2015: 1552-1560.
- [21] He K M, Zhang X Y, Ren S Q, et al. Identity mappings in deep residual networks[M]//Leibe B, Matas J, Sebe N, et al. *Computer vision-ECCV 2016. Lecture notes in computer science*. Cham: Springer, 2016, 9908: 630-645.
- [22] Szegedy C, Vanhoucke V, Ioffe S, et al. Rethinking the inception architecture for computer vision[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition, June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE Press, 2016: 2818-2826.
- [23] Zhang T F, Zhong S C, Lian C M, et al. Retina image classification based on deep learning feature union[J]. *Laser & Optoelectronics Progress*, 2020, 57(24): 241025.
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, 57(24): 241025.