

基于深度学习特征融合的视网膜图像分类

张添福¹, 钟舜聪^{1*}, 连超铭¹, 周宁¹, 谢茂松²

¹福州大学机械工程及自动化学院, 福建 福州 350108;

²福建医科大学附属第一医院, 福建 福州 350000

摘要 针对光学相干层析视网膜图像进行人工分类诊断时存在漏检、效率低等问题, 提出一种基于深度学习技术构建联合多层特征的卷积神经网络分类算法。首先通过均值漂移和数据归一化算法对视网膜图像进行预处理, 并结合损失函数加权算法解决数据不平衡问题; 其次使用轻量深度可分离卷积替代普通卷积层, 降低模型参数量, 采用全局平均池化替换全连接层, 增加空间鲁棒性, 并联合不同卷积层构建特征融合层, 加强层间特征流通; 最后使用 SoftMax 分类器进行图像分类。实验结果表明, 该模型在准确率、精确率、召回率上分别达到 97%、95%、97%, 缩短了识别时长, 所提方法在视网膜图像分类诊断中具有良好的性能。

关键词 图像处理; 卷积神经网络; 视网膜图像; 特征融合; 加权损失函数

中图分类号 TP391.4 **文献标志码** A

doi: 10.3788/LOP57.241025

Deep Learning Feature Fusion-Based Retina Image Classification

Zhang Tianfu¹, Zhong Shuncong^{1*}, Lian Chaoming¹, Zhou Ning¹, Xie Maosong²

¹School of Mechanical Engineering and Automation, Fuzhou University, Fuzhou, Fujian 350108, China;

²First Affiliated Hospital of Fujian Medical University, Fuzhou, Fujian 350000, China

Abstract Aiming at the problems of missed detection and low efficiency in manual classification and diagnosis of optical coherence tomography retina images, a deep learning-based convolutional network classification algorithm is proposed to construct joint multilayer features. First, retinal images are preprocessed using the mean shift and data normalization algorithm. The loss function weighting algorithm is combined to solve the data imbalance problem. Second, a lightweight deep separable convolution rather than an ordinary convolution layer is used to reduce the number of model parameters. Global average pooling replaces fully connected layers to increase spatial robustness, and different convolutional layers are used to build feature fusion layers to enhance feature circulation between layers. Finally, the SoftMax classifier is used for image classification. Experimental results show that the model can achieve 97%, 95%, and 97% in accuracy, precision, and recall, respectively, thereby reducing the recognition time. The proposed deep learning feature fusion-based method performs well in the classification and diagnosis of retinal images.

Key words image processing; convolutional neural network; retinal image; feature fusion; weighted loss function

OCIS codes 100.3008; 100.4996; 100.6950

1 引言

糖尿病性视网膜病变是一种常见的眼科疾病, 它是糖尿病引起的微血管并发症之一, 也是一种不可逆致盲性疾病^[1]。截至 2017 年, 全球约有 4.51 亿

人患有糖尿病, 预计到 2045 年, 这个数字将增加到 6.93 亿^[2]。在中国, 糖尿病患者中患视网膜病变者占比高达 23%^[3]。

针对光学相干层析(OCT)获得的医学图像采用人工的方式进行病变诊断, 存在许多问题, 如医师

收稿日期: 2020-04-24; 修回日期: 2020-05-23; 录用日期: 2020-06-09

基金项目: 国家自然科学基金(51675103)、福建省科技计划项目(2019I0004)

* E-mail: zhongshuncong@hotmail.com

阅片较慢、诊断过程中容易出现漏检和误检等^[4]。而糖尿病视网膜病变(DR)患者数量较多,基层地区医疗资源匮乏,导致大量患者无法及时接受治疗,因此视网膜病变早期诊断尤为重要。近年来,卷积神经网络(CNN)在计算机视觉领域表现优异,借助深度学习技术,构建计算机辅助诊疗系统,可以帮助医师对疾病进行初步筛查,甚至能够精确诊断出疾病类型^[5]。深度学习模型能够提取图像更深层次的数据特征,弥补人类对于图像特征认知不足的缺陷,可以有效避免专业医师主观性漏诊和误诊情况的发生。目前,国内外基于深度学习进行视网膜 OCT 图像分类的成果有:王翀等^[6]针对视网膜 OCT 图像构建了一种多分类器联合决策的卷积神经网络分类模型,最终模型分类准确率、精确率、召回率分别达到 89.6%、90.8%、88.6%;Bhowmik 等^[7]采用迁移 Inception-V3 模型的方法对 OCT 视网膜进行分类,分类准确率可达 92.6%;于海琛等^[8]基于 SE-Block 构建视网膜 OCT 图像分类模型,最终分类准确率和精确率分别为 94.1%、90.6%。与此同时,深度学习在其他图像领域识别任务中也取得了不错的分类性能^[9-14]。

基于深度学习方法对医学图像进行智能诊断,均取得了不错的效果^[15],但训练模型时仍存在对设备要求较高、泛化能力弱、分类时长较长等缺点,且其分类准确率仍有待提高。本文提出联合不同卷积层特征进行 OCT 图像分类的方法,将高层特征和低层特征融合,提升了模型对图像的理解力,同时采用深度可分离卷积压缩模型大小,减少了可训练参

数量,解决了训练难度大、诊断识别时长较长等问题。本文主要工作如下:1)对视网膜图像进行预处理,去除散斑噪声;2)设置类别权重,重新定义损失函数,解决数据不平衡问题;3)在模型中加入轻量深度可分离卷积层,替代原本的普通卷积层,以减小模型尺寸,降低可训练参数量;4)调整模型结构及其相关参数,通过可视化热力图优选卷积层,进而构建特征融合层,提高了模型分类识别能力。最终模型的实时诊断时间大大缩短,模型尺寸大大降低,其在分类识别能力上已达专业医师的诊断水平。

2 本文方法

首先,在 AlexNet 的基础上构建 CNN 模型,但不加载模型预训练的权重,这是因为视网膜 OCT 数据集与 ImageNet 数据集不具有特征相似性。模型前两层的训练参数较大,占有 70% 以上的参数量,使用轻量深度可分离卷积层替换原本的普通卷积层,可以减小可训练参数和模型尺寸,然后在卷积层后加入批归一化(BN)层,将前一层的输出数据空间分布进行标准化,更好地提取图像数据更深层次的数据特征,并在卷积层中加入 L2 正则化,降低模型复杂度,减小过拟合的可能性。其次,采用全局平均池化(GAP)替换全连接层,使得输入数据的空间变化的鲁棒性更强。最后,融合不同层级特征,提升模型对于图像语义特征的表达能,并使用 SoftMax 分类器进行分类诊断。将构建的卷积神经网络定义为 RongheNet 模型,如图 1 所示。以下将介绍构建模型中用到的方法。

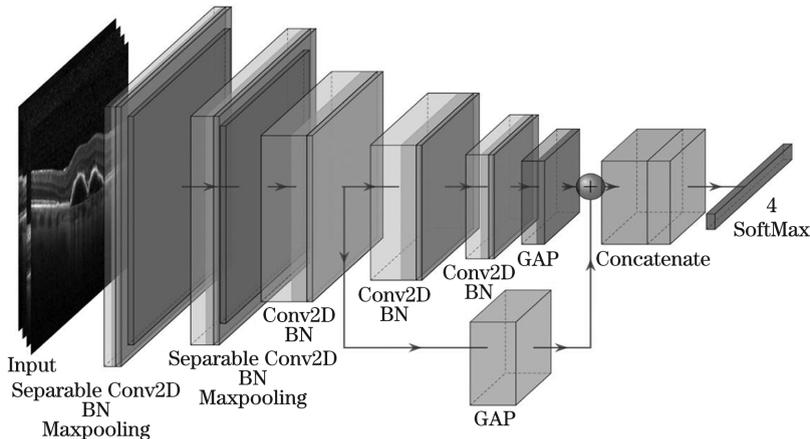


图 1 卷积神经网络模型

Fig. 1 Convolutional neural network model

2.1 卷积神经网络

卷积神经网络使用局部感受野和权值共享机制,大幅减少了训练参数,使得深度网络模型的训练

得以实现。CNN 模型能够自发地从大量的训练样本中提取图像数据特征,可以解决由人为构造特征不当而导致的分类效果不佳的问题。

CNN 主要有卷积层、池化层以及全连接层。卷积层中卷积核遍历特征图的局部区域进行卷积运算,提取所有位置上更为抽象的局部抽象特征。RongheNet 模型中使用的卷积核尺寸为 3×3 , 5×5 , 11×11 , 激活层均使用 Relu 激活函数。池化层采用最大池化法,池化层能够保留特征图的有效特征并减小其尺寸,从而有效控制过拟合现象,提高模型训练速度。

深度可分离卷积不同于普通卷积,将卷积过程分为两步,如图 2 所示,先对图像各个通道分别进行卷积运算,提取每个通道的特征,再使用 1×1 卷积核对各个通道提取特征进行综合,使用低维矩阵相乘,避免使用高维矩阵相乘,减少了模型训练参数,提升了模型在设备上的训练速度,压缩了模型大小,缩短了分类识别时长。

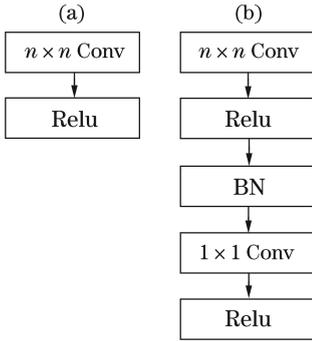


图 2 卷积层结构对比。(a)普通卷积;(b)深度可分离卷积
Fig. 2 Convolutional layer structure comparison. (a) Ordinary convolution; (b) depth separable convolution

RongheNet 模型使用 Softmax 作为输出决策层,在训练过程中通过反向传播不断优化交叉熵损失函数值,以提升模型分类性能,最终实现图像的分类识别。当数据之间分布差别较大时,由神经网络提取到的数据特征分布差异也较大,并且数据特征经过激活函数的非线性运算后,数据分布很可能发生改变,导致模型泛化能力差,因此需在模型中多次归一化数据分布。批归一化层^[16]首先对输入数据 $\mathbf{x} = (x^{(1)}, x^{(2)}, \dots, x^{(n)})$ 进行归一化处理,即

$$\hat{x}^{(n)} = \frac{x^{(n)} - \mu_x}{\sqrt{\sigma_x^2 + \epsilon}}, \quad (1)$$

式中: μ_x 为输入数据的均值; σ_x^2 为标准差; 为了避免分母为 0, 在分母中加上接近 0 的值 ϵ 。若只将数据归一化至 $(-1, 1)$ 区间, 则某些激活函数无法达到非线性变换的目的, 将会削弱网络的性能, 因此加入缩放因子 $\gamma^{(n)}$ 和移位因子 $\beta^{(n)}$ 进行数据重构, 即

$$y^{(n)} = \gamma^{(n)} \hat{x}^{(n)} + \beta^{(n)}. \quad (2)$$

在 RongheNet 模型的卷积层后加入 BN 层, 以解决数据分布变化的问题, 提升训练精度。

正则化的目的是限制参数数量, 避免模型因过于复杂而发生拟合现象。L2 正则化则是直接在原来损失函数 L_{in} 的基础上加上权重参数 w 的平方和, 其中 λ 是正则化参数, 表达式为

$$L = L_{in} + \lambda \sum_j w_j^2. \quad (3)$$

数据的特征融合^[17]分为两种方式: 数据特征融合和决策特征融合。数据特征融合是将多层特征进行融合, 以此提高模型对图像的表达能。决策特征融合是将多个分类器预测结果进行决策融合。RongheNet 模型使用数据特征融合, 联合模型中多层卷积层构造特征融合层, 以提升模型对 OCT 视网膜图像的分类性能。

2.2 预处理方法

为解决视网膜图像数据集中数据不平衡的问题, 对于数据量较少的类别, 赋予其较大的权重。在前向传播和反馈传播时, 补偿对于数据量少类别的损失函数^[18]值, 使模型更加关注此类数据特征的学习, 避免模型出现总体准确率高, 但对于特定类别诊断分类性能极差的情况。权重设置公式为

$$\alpha_i = \frac{C_{all}}{CC_i}, \quad (4)$$

式中: C 是类别数; C_{all} , α_i 和 C_i 分别为数据总数、第 i 类类别权重和其数据总数。

OCT 视网膜图像中存在大量的散斑噪声, 使用均值漂移算法对图像进行平滑滤波处理, 中和色彩分布相近的颜色, 弱化面积较小的散斑噪声, 以降低图像干扰噪声, 突出病变关键部位, 提升模型分类准确率。

3 分类实验

3.1 数据集简介

本文使用的数据集是加州大学圣地亚哥分校的 Kermany 等^[19]在 2018 年发布的数据集。该数据集包括脉络膜新生血管 (CNV)、糖尿病性黄斑水肿 (DME)、玻璃膜疣 (DRUSEN)、正常 (NORMAL) 4 类, 如图 3 所示, 共计 83484 张图像。

本文将数据集按 7 : 2 : 1 的比例随机划分成训练集 (train)、验证集 (validation)、测试集 (test), 划分后数据集之间没有任何交叉, 数据分布如表 1 所示。

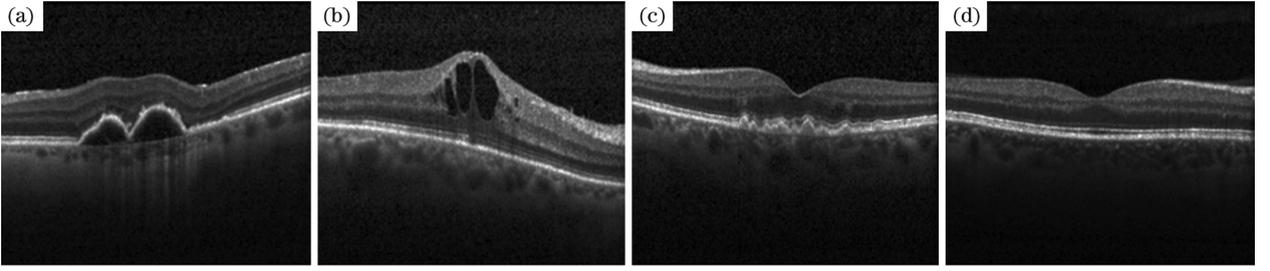


图3 数据集样本图。(a) CNV;(b) DME;(c) DRUSEN;(d) NORMAL

Fig. 3 Sample graphs of the dataset. (a) CNV; (b) DME; (c) DRUSEN; (d) NORMAL

表1 OCT数据集划分情况

Table 1 Division of OCT dataset

Image type	Number of images			
	Train	Validation	Test	Total
CNV	26043	7441	3721	37205
DME	7943	2270	1135	11348
DRUSEN	6031	1723	862	8616
NORMAL	18420	5263	2632	26315
Total	58437	16697	8350	83484

3.2 数据集预处理

当样本噪声多、不同类别数据不平衡时,需对其进行预处理操作,以提高模型提取特征的能力。由于OCT图像中存在大量的散斑噪声,影响模型对训练数据的特征提取,本研究使用均值漂移算法去除视网膜图像中的散斑噪声,如图4所示,并将数据集图像全部裁剪为299 pixel×299 pixel的图像。

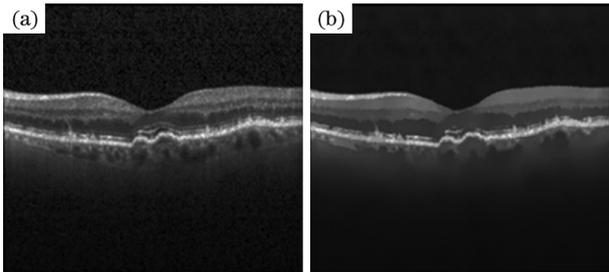
图4 数据预处理。(a)OCT视网膜原图像;
(b)均值漂移去散斑图像

Fig. 4 Image preprocessing. (a) Original image of OCT retina; (b) mean-shift removed speckle image

OCT视网膜数据集各类别数量分布不平衡,本次研究赋予数量较少类别较大的权重,以此解决数据不平衡的问题。各类别权重计算如(4)式所示。计算可得CNV, DME, DRUSEN, NORMAL的权重分别为0.561, 1.840, 2.422, 0.7931。加权损失函数公式为

$$L = -\frac{1}{M} \sum_{n=1}^M \sum_{i=1}^C \alpha_i \cdot y_i^{(n)} \cdot \ln(\hat{y}_i^{(n)}), \quad (5)$$

式中, M 是一个batch的样本数, $y_i^{(n)}$ 和 $\hat{y}_i^{(n)}$ 分别代表实际值和预测值。如(5)式所示,数据量少的类别通过赋予较大的类别权重,增大其在损失函数中的比重,通过反向传播,调整参数,可以提升对于数据量少类别的分类准确率。

3.3 评价指标

本研究采用多种评价指标对模型进行评估,对于多类别分类问题,将当前预测类别定义为正类,其他类别定义为负类。则预测结果可分成4种情况:正类预测为正类(TP, R_{TP}),正类预测为负类(FN, R_{FN}),负类预测为负类(TN, R_{TN}),负类预测为正类(FP, R_{FP})。

定义4种评价指标,如分类准确率^[20](Accuracy, A_{cc})、精确率(Precision, P)、召回率(Recall, R_{recall})、平均准确率(MA, A_M),其具体表达式为

$$A_{cc} = \frac{R_{TP} + R_{TN}}{R_{TP} + R_{TN} + R_{FP} + R_{FN}}, \quad (6)$$

$$P = \frac{R_{TP}}{R_{TP} + R_{FP}}, \quad (7)$$

$$R_{recall} = \frac{R_{TP}}{R_{TP} + R_{FN}}, \quad (8)$$

$$A_M = \frac{1}{C} \sum_{i=1}^C a_i, \quad (9)$$

式中 a_i 为每个类别的准确率。

3.4 设置实验

实验在 Intel i5-8500 CPU, 内存为 12 G, GeForce GTX1650 显卡上进行, 基于 Tensorflow 深度学习框架, 利用 Cuda10.0 和 CUDNN10.0 加速模型训练。

在 AlexNet 的基础上, 将 5 个卷积层特征图个数改为 128、256、512、512、256, 2 个全连接层的个数改为 128、64, 分类器神经元改为 4, 定义其为 SlimNet 模型。在 SlimNet 模型基础上, 在卷积层后都加上批归一化层, 并进行 L2 正则化操作, 其定义为 BnL2Net 模型。将前两层普通卷积层替换成轻量可分离卷积层, 去除 BnL2Net 的两个全连接层, 使用全局池化层, 其定义为 GAPNet 模型。在第 3 个卷积层后加上全局平均池化层, 并联合最后一层卷积层构建特征融合层, 其定义为 RongheNet 模型。

本次研究设置了 4 组对照实验, 以此验证不同

网络模型对分类性能的影响。

实验一: 验证类别权重对模型性能的影响。

1) 直接训练 SlimNet 模型; 2) 基于 SlimNet 模型进行类别权重设置, 重新定义交叉熵损失函数后进行训练。

实验二: 验证 BN 层和 L2 正则化对模型的影响。1) 对 SlimNet 模型进行训练; 2) 对 BnL2Net 模型进行训练。

实验三: 验证 GAP 层相比于全连接层对模型性能的影响, 并检验轻量深度可分离卷积层对模型参数数量和分类识别时长的影响。1) 训练 GAPNet 模型; 2) 训练 BnL2Net 模型;

实验四: 验证联合不同卷积层特征对模型的影响。1) 训练 GAPNet 模型; 2) 训练 RongheNet 模型。

4 实验结果分析

对照实验得到的验证集准确率和损失函数曲线, 如图 5 所示。分类指标对比如表 2 所示。

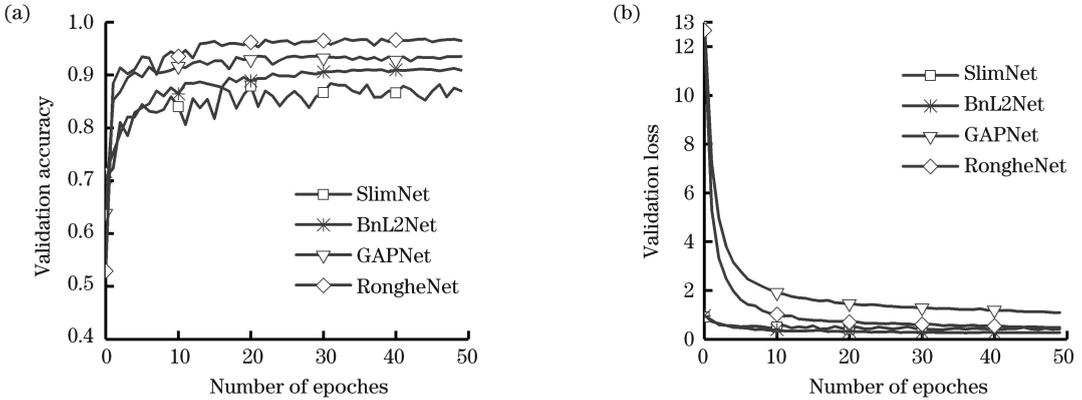


图 5 验证结果曲线。(a)验证准确率曲线;(b)验证损失曲线

Fig. 5 Validation result curves. (a) Validation accuracy curve; (b) validation loss curve

表 2 不同训练方法的分类指标对比

Table 2 Comparison of classification indexes of different training methods

unit: %

Network type	Category weight	Accuracy	MA	Precision	Recall
SlimNet	No	89.9	82.5	88.0	82.0
SlimNet	Yes	90.2	87.2	85.0	87.0
BnL2Net	Yes	93.0	91.6	92.0	93.0
GAPNet	Yes	94.9	93.3	92.0	93.0
RongheNet	Yes	97.2	97.0	95.0	97.0

由表 2 可知, SlimNet 模型直接训练时, 虽然总体分类准确率可达 89.9%, 但是平均分类准确率仅达到 82.5%, 由图 6(a) 混淆矩阵可知,

SlimNet 模型对于 DME 和 DRUSEN 的分类准确率为 78.9% 和 58.8%, 分类准确率较低。视网膜图像数据不平衡, 未使用加权交叉熵损失函数

时,模型主要学习数据量大的 NORMAL 和 CNV 类别特征,预测正确样本大多属于这两个类别,总体分类准确率高,但对于 DME 和 DRUSEN 的分类性能差。加上类别权重后,由图 6(b)混淆矩阵可知,对于 DME 和 DRUSEN 的分类准确率分

别由 78.9%和 58.8%提升至 87.8%和 77.5%,平均准确率由 82.5%提升至 87.2%,这说明设置类别权重、重定义交叉熵损失函数,能够有效解决样本数据不平衡问题,有助于提升模型的性能。

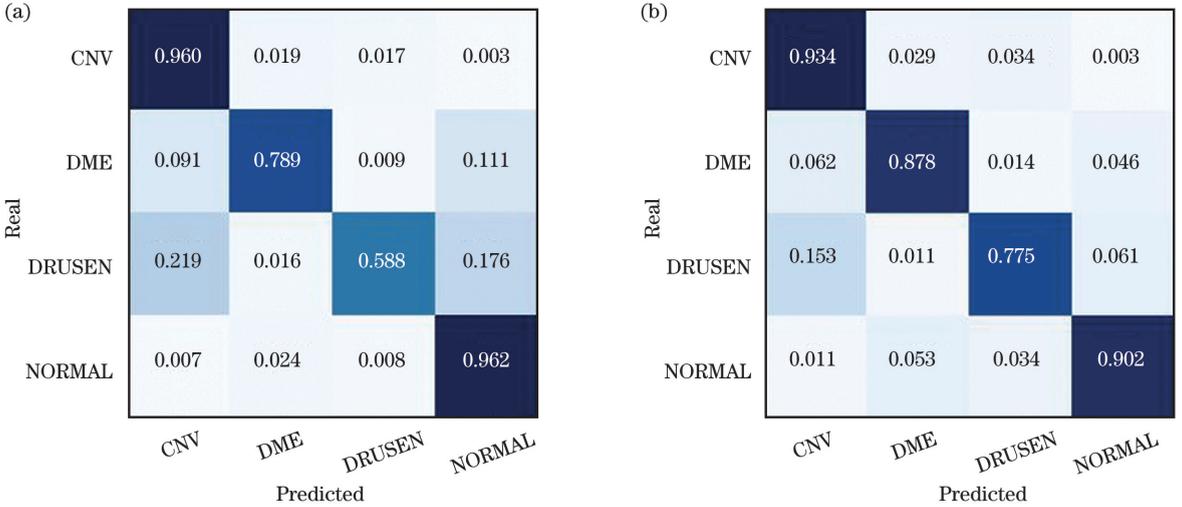


图 6 混淆矩阵。(a)未设置加权损失函数混淆矩阵;(b)设置加权损失函数混淆矩阵

Fig. 6 Confusion matrix. (a) Confusion matrix without weighted loss function; (b) confusion matrix with weighted loss function

在 BnL2Net 模型中引入批归一化层,将数据分布归一化。由图 5(a)验证准确率曲线可知, BnL2Net 模型中曲线波动小,较平稳,在训练至第 10 个 epoch 时,模型准确率上升较慢,趋近于最优状态;由图 5(b)验证集交叉熵的损失函数可知, BnL2Net 模型中的损失曲线稳步下降,最终在第 10 个 epoch 的下降趋势缓慢且趋近于 0,此时网络已经收敛,模型趋近于该网络下的最优性能。由表 2 可知, BnL2Net 模型的平均准确率比 SlimNet 提高 4.4 个百分点,分类准确率达到 93.0%,这说明基于 BN 层对数据进行归一化操作后,不仅使训练更平稳,还能提升模型的性能。

由表 2 可知, GAPNet 模型比 BnL2Net 平均准确率高 1.7 个百分点,精确率和召回率达到 92.0%和 93.0%。在 BnL2Net 模型中,全连接层对 feature map 作全排列,丢失了空间信息,并且其可训练的参数较多,容易造成过拟合。GAPNet 模型采用 GAP 层替换全连接层,减少了训练参数,更好地保留卷积结构,提升了对输入空间变化的鲁棒性,且使用深度可分离卷积减少训练参数,缩短了识别时长,使网络训练变得更容易。

在深层卷积神经网络中,低层次卷积层提取的主要是基础特征,层次越高的卷积核提取的特征越

抽象,也越能体现图像的语义信息。热力图中区域越亮,表示该卷积层提取该区域特征的概率越大,本次研究通过可视化热力图(见图 7),选取特征融合层。图 7 左边为原始图像,右边为 5 个卷积层对应的热力图,由图 7 可知,第 3 层卷积层在病变区域的亮度最大,这说明该层卷积层对于病变部位最感兴趣,病变部位特征的提取对于模型的正确分类十分重要,因此 RongheNet 模型将第 3 层卷积层提取的特征与最后一层卷积层提取的特征进行融合,以提高网络对图像的表达力。

由图 8 混淆矩阵可知, RongheNet 模型相比于

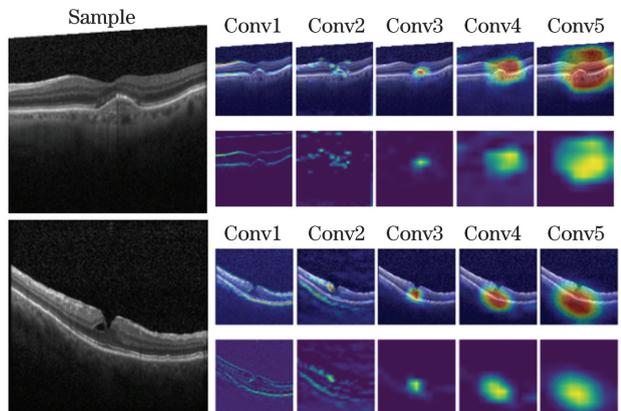


图 7 可视化热力图

Fig. 7 Visual heat maps

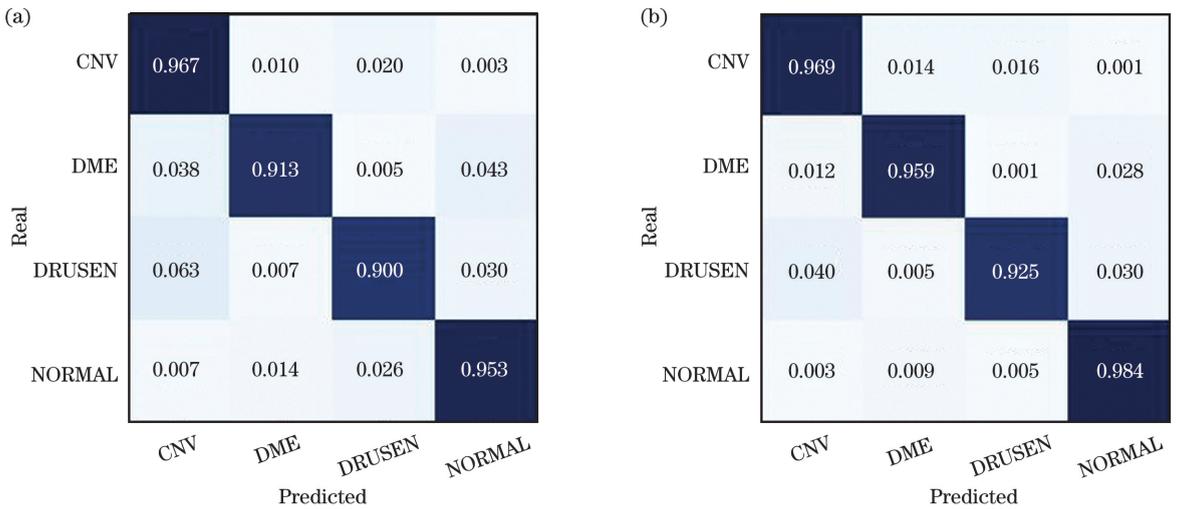


图 8 混淆矩阵。(a)GAPNet 模型的混淆矩阵;(b)RongheNet 模型的混淆矩阵

Fig. 8 Confusion matrix. (a) Confusion matrix for GAPNet model; (b) confusion matrix for RongheNet model

GAPNet 模型对于 DME 和 DRUSEN 的分类准确率分别提升了 4.6% 和 2.5%。

由表 2 可知,平均准确率达到 97.0%,精确率和召回率分别达到 95.0% 和 97.0%,这说明将模型提取的高层次特征和低层次特征进行融合,能够有效提升模型的性能。

进一步,为了验证模型的性能,将 RongheNet 模型与现有的 OCT 视网膜图像分类模型进行对比。由表 3 可知,RongheNet 模型在平均准确率和精确率、召回率上均优于现有 OCT 视网膜模型取得的分类性能。在 8616 张测试集上进行分类诊断,RongheNet 模型仅需 63 s,而经过迁移微调后的经典 Inception-V3 模型需要 231 s。由此可知,本次研究所提模型,不仅分类性能好,且能大幅缩短识别时长,实现实时诊断,解决专业医师匮乏、患者不能得到及时诊治的问题。

表 3 不同模型性能对比

Table 3 Performance comparison of different models unit: %

Method	MA	Precision	Recall
Wang <i>et al</i> ^[6]	89.6	90.8	88.6
Bhowmik <i>et al</i> ^[7]	92.6	94.0	94.0
Yu <i>et al</i> ^[8]	94.1	94.6	90.6
Ours	97.0	95.0	97.0

4 结 论

利用深度学习技术实现了 OCT 视网膜图像的自动分类。通过观察热力图,优选感兴趣卷积层,

创新性地联合不同层级特征构建特征融合层,使用轻量深度可分离卷积层减小模型大小,降低参数量,缩短识别时长,所构建的模型分类准确率达到 97%,召回率和精确率分别达到 97% 和 95%,基于 8616 张测试集进行分类诊断时仅需 63 s,模型占用内存空间仅为 54.5 MB,具有良好的分类性能且能满足实时诊断需求。后续的研究中可以将该模型构建思想应用到其他医学图像分类中,通过尝试构建更深的网络模型和融合更多层特征,进行分类诊断。

参 考 文 献

- [1] Wang J R, Yang Y. Research progress of related factors of type II diabetic retinopathy[J]. Journal of Kunming Medical University, 2019, 40(4): 131-135.
王金瑞, 杨莹. 2 型糖尿病视网膜病变相关因素研究进展 [J]. 昆明医科大学学报, 2019, 40(4): 131-135.
- [2] Cho N H, Shaw J E, Karuranga S, et al. IDF diabetes atlas: global estimates of diabetes prevalence for 2017 and projections for 2045 [J]. Diabetes Research and Clinical Practice, 2018, 138: 271-281.
- [3] Laddha A P, Kulkarni Y A. Tannins and vascular complications of diabetes: an update[J]. Phytomedicine, 2019, 56: 229-245.
- [4] Yang L, Shen X. Research on correlation between diabetic retinopathy and dry eye[J]. International Eye Science, 2018, 18(4): 744-747.
杨玲, 沈玺. 糖尿病性视网膜病变与干眼的相关性研究 [J]. 国际眼科杂志, 2018, 18(4): 744-747.

- [5] Bi K, Wang Y. Advances in the application of computer aided diagnosis in ultrasound medicine[J]. *Oncoradiology*, 2019, 28(5): 296-300.
毕珂, 王茵. 计算机辅助诊断技术在超声医学中的应用进展[J]. *肿瘤影像学*, 2019, 28(5): 296-300.
- [6] Wang C, He X X, Fang L Y, et al. Automatic classification of retinal optical coherence tomography images via convolutional neural networks with joint decision[J]. *Chinese Journal of Biomedical Engineering*, 2018, 37(6): 641-648.
王翀, 何兴鑫, 方乐缘, 等. 基于联合决策卷积神经网络的光学相干断层扫描图像自动分类[J]. *中国生物医学工程学报*, 2018, 37(6): 641-648.
- [7] Bhowmik A, Kumar S, Bhat N. Eye disease prediction from optical coherence tomography images with transfer learning [C] // *Communications in Computer and Information Science*, May 15, 2019. Rome: ICFNN, 2019: 104-114.
- [8] Yu H C. Research on classification of retinal diseases based on SE-block[D]. Changchun: Jilin University, 2019.
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.
- [9] Voets M, Møllersen K, Bongo L A. Replication study: development and validation of deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs [EB/OL]. (2015-02-11) [2020-05-25]. <https://arxiv.org/abs/1803.04337>.
- [10] Zhang R G. Retina vessel segmentation based on multi-scale and multi-path fully convolutional neural network[J]. *Laser Journal*, 2020, 41(2): 194-198.
张润谷. 基于多尺度多路径FCN的视网膜血管分割[J]. *激光杂志*, 2020, 41(2): 194-198.
- [11] Zhuo D, Jing J F, Zhang H H, et al. Classification of chopped strand mat defects based on convolutional neural network [J]. *Laser & Optoelectronics Progress*, 2019, 56(10): 101009.
卓东, 景军锋, 张缓缓, 等. 基于卷积神经网络的短切毡缺陷分类[J]. *激光与光电子学进展*, 2019, 56(10): 101009.
- [12] Meng T, Liu Y H, Zhang K Y. Algorithm for pathological image diagnosis based on boosting convolutional neural network[J]. *Laser & Optoelectronics Progress*, 2019, 56(8): 081001.
孟婷, 刘宇航, 张凯昱. 一种基于增强卷积神经网络的病理图像诊断算法[J]. *激光与光电子学进展*, 2019, 56(8): 081001.
- [13] Feng Y, Yi B S, Wu C Y, et al. Pulmonary nodule recognition based on three-dimensional convolution neural network [J]. *Acta Optica Sinica*, 2019, 39(6): 0615006.
冯雨, 易本顺, 吴晨玥, 等. 基于三维卷积神经网络的肺结节识别研究 [J]. *光学学报*, 2019, 39(6): 0615006.
- [14] Yang M J, Tang Y Q, Jiang X J. Novel shoe type recognition method based on convolutional neural network [J]. *Laser & Optoelectronics Progress*, 2019, 56(19): 191505.
杨孟京, 唐云祁, 姜晓佳. 基于卷积神经网络的鞋型识别方法 [J]. *激光与光电子学进展*, 2019, 56(19): 191505.
- [15] Wu C Y, Yi B S, Zhang Y G, et al. Retinal vessel image segmentation based on improved convolutional neural network [J]. *Acta Optica Sinica*, 2018, 38(11): 1111004.
吴晨玥, 易本顺, 章云港, 等. 基于改进卷积神经网络的视网膜血管图像分割 [J]. *光学学报*, 2018, 38(11): 1111004.
- [16] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift [EB/OL]. (2015-03-02) [2020-05-25]. <https://arxiv.org/abs/1502.03167>.
- [17] Li K, Zou C Q, Bu S H, et al. Multi-modal feature fusion for geographic image annotation [J]. *Pattern Recognition*, 2018, 73: 1-14.
- [18] Lecun Y, Bengio Y, Hinton G. Deep learning [J]. *Nature*, 2015, 521(7553): 436.
- [19] Kermay 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.
- [20] Zhang K, Wang X, Guo Y, et al. Competing ratio loss for multi-class image classification [C] // 2019 IEEE Visual Communications and Image Processing (VCIP), December 1-4, 2019, Sydney, Australia. New York: IEEE: 19297873.