

光学相干断层扫描视网膜图像的迁移学习分类

连超铭¹, 钟舜聪^{1*}, 张添福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 350005

摘要 光学相干断层扫描是目前检测糖尿病视网膜黄斑病变较为灵敏的方法之一,但病变的人工判断易产生主观失误,且比较耗时。为此,本文在迁移学习的基础上提出了一种改进的深度神经网络,用于视网膜图像的自动分类。先基于自适应阈值联合高斯滤波算法对图像进行预处理;然后以预训练模型为基础,通过微调解决样本差异的问题,并以全局平均池化方法替代传统的全连接层来提取深层特征,以降低网络的过拟合现象。基于实验数据对该网络进行验证,所提网络对视网膜病变图像的分类准确率可达 97.3%,说明了所提网络对视网膜黄斑病变图像自动分类的有效性。

关键词 医用光学; 光学相干断层扫描; 迁移学习; 高斯滤波; 微调; 全局平均池化

中图分类号 TP391.4

文献标识码 A

doi: 10.3788/LOP202158.0117002

Transfer Learning-Based Classification of Optical Coherence Tomography Retinal Images

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

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

²The First Affiliated Hospital of Fujian Medical University, Fuzhou, Fujian 350005, China

Abstract Currently, optical coherence tomography is one of the most sensitive methods for detecting diabetic retinopathy. However, the artificial detection of diabetic retinopathy is time consuming and prone to subjective errors. Accordingly, this paper proposed an improved deep learning network based on transfer learning for automatic classification of retinal images. First, the image was preprocessed via adaptive threshold combined with the Gaussian filter algorithm. Then, on the basis of the pretraining model, the problem of sample difference was solved through fine-tuning, and the traditional fully connected layer was replaced by the global average pooling method for extracting deep features and reducing overfitting. The network was validated based on the experimental data, with the accuracy of the retinal image classification being 97.3%. Results reveal that the proposed network is effective for the automatic classification of retinal macular lesions.

Key words medical optics; optical coherence tomography; transfer learning; Gaussian filter; fine-tuning; global average pooling

OCIS codes 170.4500; 170.4470; 100.3008

1 引言

糖尿病视网膜黄斑病变是糖尿病的常见并发症之一,也是导致工作年龄人群致盲的主要疾病。国

际糖尿病联盟 2017 年的报告指出,我国糖尿病患者已超 1.14 亿^[1]。脉络膜新生血管(CNV)、糖尿病性黄斑水肿(DME)和眼底玻璃膜疣(drusen)是最常见的几种黄斑病变并发症,这几种病症会伴随着

收稿日期: 2020-05-09; 修回日期: 2020-05-18; 录用日期: 2020-06-04

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

年龄的增长而严重^[2]。因此,及时的视网膜眼底检查对于糖尿病患者来说是必不可少的。

视网膜眼底检查的传统方法是眼底相机与血管荧光造影检查方法。近年来,光学相干断层扫描(OCT)技术凭借其快速、无创的微观视网膜结构成像优势,逐渐成为各大医院进行视网膜眼底成像的主流工具^[3-5]。但临床上对于病变的观察判断依旧依靠医生的个人经验,这种判断不仅耗费时间成本,同时也容易出现主观错误。因此,在深度学习的基础上对糖尿病视网膜 OCT 图像进行快速准确的自动分类,对眼科疾病的治疗发展具有重要意义。

深度学习^[6]自问世以来,就在机器视觉方面展现出了广阔的前景,其在医学影像的识别上也具有突出优势。目前,已有多种基于深度卷积神经网络(DCNN)的视网膜眼底图像识别方法被提出。Gargeya 等^[7]对 75137 张 OCT 视网膜眼底图像进行了卷积神经网络训练,达到了 0.97 的 AUC(area under curve)值。丁蓬莉等^[8]设计了继承于 AlexNet 神经网络的 CompactNet 神经网络,其分类指标达到了 0.87。庞浩等^[9]提出了利用弱监督学习的深度卷积神经网络,其训练样本为 Kaggle 数据集上的 35124 张视网膜眼底图像,该网络在验证集上得到了 0.85 的分类指标。传统的分类方法需要设计合适的神经网络对样本进行训练,同时,为了提取更多的图像特征,还需要增加大量的网络层和全连接层,因此会导致数据量大、过拟合现象发生。本文在微调(fine-tuning)迁移学习的基础上,采用全局平均池化方法代替全连接层对 OCT 视网膜眼底图像进行识别和分类,并结合数据图像的去噪处理来增强特征提取效率,减少过拟合,提高识别的准确率。

2 基本原理

2.1 迁移学习

近年来,深度学习技术在图像识别、机器视觉、语音识别等领域应用广泛,得到了各行业的重视。但深度学习面临着一个严峻的问题,即数据依赖严重。深度学习在训练、生成神经网络模型过程中往往需要海量数据,训练数据量和模型完整度存在线性关系^[10],待解决的问题越复杂,所需要的数据量越大。迁移学习(TL)技术的出现缩短了构建模型和训练样本所需的时间,同时迁移学习所需的样本数据量更小,模型的泛化性更好^[11]。近年来,迁移学习已被广泛应用于各领域的机器学习中,它允许采用已训练好的模型进行新数据的训练,不必再针对一个问题设

计一个全新的解决方案。同时,迁移学习适用于样本数量较少的 OCT 视网膜眼底图像的分类。

2.1.1 迁移学习原理

迁移学习就是在已有知识的基础上学习目标知识,关键在于寻找已有知识和目标知识之间的相似性。这里有两个重要的概念:域和任务。域(D)和任务(T)的定义为

$$D = \{\sigma, P(\sigma)\}, \quad (1)$$

$$T = \{y, f(\cdot)\}, \quad (2)$$

式中: σ 为特征空间,且 $\sigma = \{\sigma_1, \sigma_2, \dots, \sigma_n\}$; $P(\sigma)$ 为特征空间的边缘分布; y 为标签空间; $f(\cdot)$ 为目标预测函数^[12]。已有的知识被称作源域(source domain),需要迁移的目标知识被称为目标域(target domain),两者相互联系。通常的做法是减小源域和目标域之间数据分布的差异,从源任务(source tasks)中提取特征,并将特征应用于目标域中。

2.1.2 预训练模型

迁移学习按照学习方式可以分为四类,即基于样本的迁移、基于特征的迁移、基于模型的迁移和基于关系的迁移。本研究采用的是基于模型的迁移,且采用了 VGG19、ResNet50、InceptionV3 三个预训练网络模型。

VGG19 网络包含 16 个卷积层和 3 个全连接层,主要使用 3×3 的卷积层和 2×2 的池化层。通过加深网络结构,VGG19 模型能够提取出更多的图像内层特征。ResNet50 是基于 CNN 框架开发出的 50 层网络模型,具有更少的参数量和更快的训练速度。InceptionV3 网络具有良好的局部扩扑结构,能够并行地对图像进行若干的卷积层或池化层操作。

2.1.3 微调

当实际应用视网膜的数据集进行迁移学习时,直接使用预训练模型进行训练很难得到理想的识别准确率。微调是在预训练模型的基础上对网络进行适当的调整,然后再将网络应用于新数据集的过程。迁移学习网络中前几层网络的特征更加泛型,后几层网络具有特定数据集的特征,因此,对于图像特征相似度与预训练模型差异较大的数据集采用冻结前 k 层网络,并训练剩余 $(n - k)$ 层网络的微调方法。微调采用较低的学习率训练网络。预训练模型具有良好的权重比例,较低的学习率不会过快地改变模型的权重,可使模型保持优秀的特征提取性能。

2.2 全局平均池化

在迁移学习过程中,卷积层提取图像的局部特征,然后通过若干个全连接层对提取到的特征进行

再处理,使获得的特征图(feature map)维度减少,便于分类器分类。但是全连接层参数过大会使模型数据量变得臃肿,容易丢失特征图的空间信息。

本研究使用全局平均池化^[13](GAP)函数代替迁移学习网络中的全连接层。GAP 函数把特征图进行全局平均得到一个值,以直观地反映特征图之间的联系(图 1),实现特征提取过程中特征图的降维和卷积层向量的展开,使特征图的空间信息保持完整,并大幅减少网络模型的参数量,减少过拟合的发生。

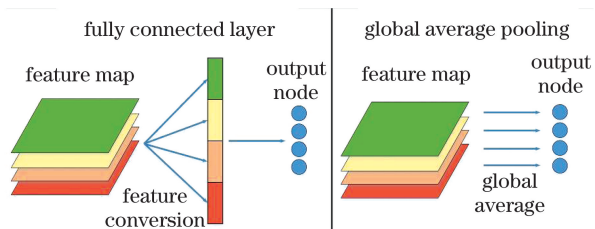


图 1 全局平均池化函数对图像特征的提取方式

Fig. 1 Global average pooling function for image feature extraction

2.3 评价指标

深度学习中对二分类图像分类的指标通常有以下四种: True Positive(真正类,记为 α_{TP}),即将正类样本预测为正类的数量; True Negative(真负类,记为 α_{TN}),即将负类样本预测为负类的数量; False Positive(假正类,记为 α_{FP}),即将负类样本预测为正类的数量; False Negative(假负类,记为 α_{FN}),即将正类样本预测为负类的数量。

本文中糖尿病视网膜黄斑病变的识别属于多分类问题,对测试图片的病变判断正确代表正类,判断错误代表负类。分类结果的评估将从训练过程的准确率和受试者工作特征(ROC)曲线两个方面进行。

模型训练结果的最直观的评价标准就是训练集和验证集的准确率曲线以及测试准确率。

在 ROC 曲线^[14]图中,纵轴是“真正例率”(记为 r_{TPR}),也称为敏感度;横轴是“假正例率”(记为 r_{FPR}),也称特异性。 r_{TPR} 和 r_{FPR} 的计算公式分别为

$$r_{TPR} = \frac{\alpha_{TP}}{\alpha_{TP} + \alpha_{FN}}, \quad (3)$$

$$r_{FPR} = \frac{\alpha_{FP}}{\alpha_{TN} + \alpha_{FP}}. \quad (4)$$

ROC 曲线包裹下的面积称为 AUC(area under curve),反映该类别图像测试的准确程度,表示测试中正例排在负例前面的概率。AUC 面积越大,模型的预测准确率就越高。理想状态下的 AUC 面积值为 1。假设 ROC 曲线是由若干个坐标分别为

$(x_1, y_1), (x_2, y_2), \dots, (x_n, y_n)$ 的点连接而成的,则 AUC 的面积为

$$S_{AUC} = \frac{1}{2} \sum_{i=1}^{n-1} (x_{i+1} - x_i) \cdot (y_i + y_{i+1}). \quad (5)$$

3 分类实验

3.1 实验数据集

3.1.1 数据集详情

本研究选用的数据集是 Kaggle 平台提供的开源 OCT 视网膜病变图像。数据集中的图像数量为 84495 张,包括四种类别的图像,即脉络膜新生血管、糖尿病性黄斑水肿、眼底玻璃膜疣和正常视网膜图像,如图 2 所示。

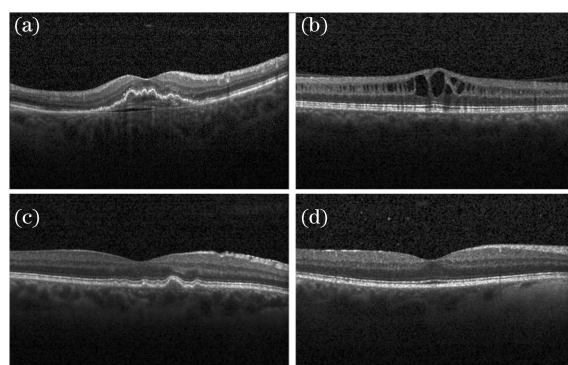


图 2 糖尿病视网膜黄斑病变 OCT 图像。(a)脉络膜新生血管;(b)糖尿病性黄斑水肿;(c)眼底玻璃膜疣;(d)正常视网膜

Fig. 2 OCT images of diabetic retinopathy. (a) Choroidal neovascularization; (b) diabetic macular edema; (c) drusen; (d) normal retina

数据集的收集标记机构为加利福尼亚大学圣地亚哥分校 Shiley 眼科研究所、加利福尼亚视网膜研究基金会、上海眼科医学中心和北京同仁眼科中心。数据集以 8.5 : 1 : 0.5 的比例随机划分为训练集、验证集和测试集,如表 1 所示。

表 1 视网膜数据集的划分

Table 1 Retina data set partition

Category	Number of images in data set			
	Training	Validation	Test	All
CNV	32477	3820	1912	38209
DME	9647	1136	567	11350
Drusen	7323	862	431	8616
Normal	22372	2632	1316	26320
All	71819	8450	4226	84495

3.1.2 图像去噪

OCT 技术在采集视网膜眼底图像时容易受到外界光线的干扰,导致图像出现畸变和散斑噪

声^[15]。因此,在进行模型训练之前需要对数据集中的散斑噪声进行去噪处理,同时增强图像的对比度,突出病变部位的特征。去噪处理结果如图 3 所示。

具体的去噪方法如下:1)将图片转化为灰度图像。2)在图像的不同区域采用不同的阈值进行分

割,因为当图像各局部的亮度不同时,全局阈值很难将背景和目标进行分割。阈值的取值为相邻区域的加权和。3)采用高斯滤波算法进行去噪,在保留图像边缘细节的同时减小了噪声的影响,对消除散斑噪声和椒盐噪声具有显著效果。

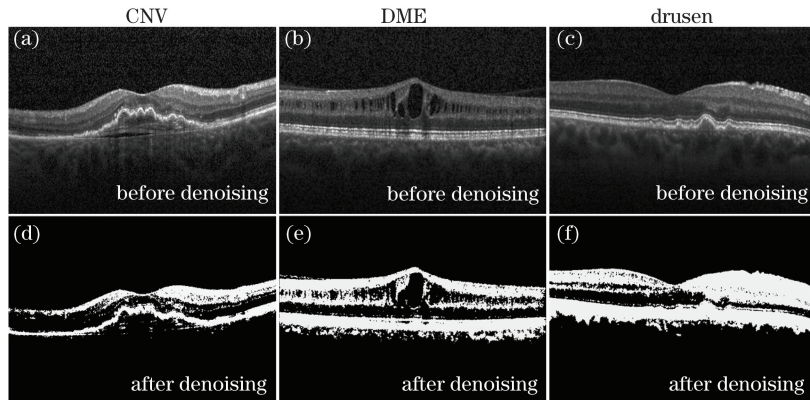


图 3 去噪前后的图像

Fig. 3 Images before and after denoising

3.2 实验设置

本研究基于迁移学习网络进行视网膜数据集的分类,以 VGG19、ResNet50 和 InceptionV3 三个预训练模型为基础,通过微调的方法获得最佳网络权重,然后用全局平均池化层替代全连接层。具体方法如下:

1) 数据的预处理。首先将输入的图像设置为 224 pixel×224 pixel,并以 1/255 的尺度因子进行缩放;然后使用多种数据增强变换来增加训练数据量,以提取到更深病变部位的特征。增强变换包括旋转(rotation)、翻转(flip)、缩放(zoom)和平移(shift)。

2) 在迁移学习中,预训练模型采用 ImageNet 中的 1000 类图像进行训练。而视网膜图像和预训练模型的图像特征差异较大,若直接使用模型的分类器对图像进行分类,准确率会很低。为解决这一问题,本文对预训练模型进行微调,即:在直接迁移学习的基础上将预训练模型的部分网络层进行冻

结,开放剩余的层进行训练,以得到更好的权重参数;在网络的池化层后接入 Dropout 正则化函数,并随机丢弃一定比例的神经元(这是因为过拟合问题会导致训练过程中测试数据的准确率远低于训练数据的准确率,没有考虑模型的泛化能力)。

3) 为验证 GAP 函数对视网膜图像分类的作用,设置两组对照实验。

实验一:①直接迁移学习组,即使用预训练模型直接进行迁移学习训练;②微调迁移学习组,即对预训练模型进行微调后的神经网络采用维度分别为 2048 和 1024 的全连接层进行特征的融合。

实验二:①微调迁移学习组;②GAP 迁移学习组,即在微调的基础上,用全局平均池化层替代全连接层,以减少过拟合与数据冗余,提高识别的准确率。

实验二中使用的全局平均池化层的神经网络模型结构如图 4 所示。网络中的 GAP 层对预训练模

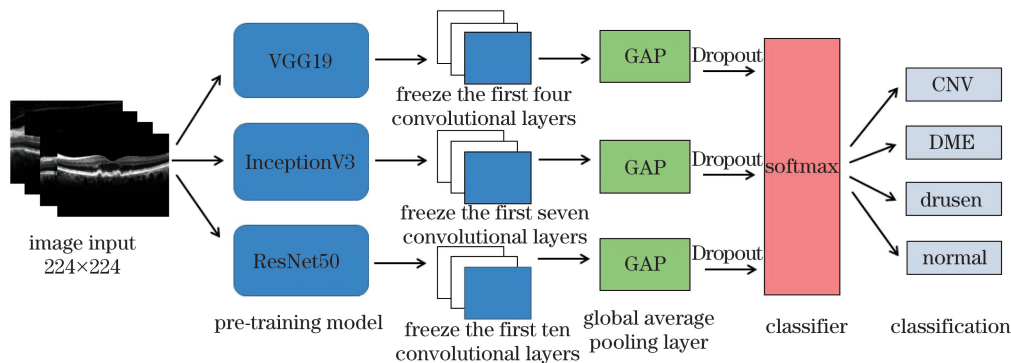


图 4 GAP 迁移学习网络模型结构

Fig. 4 GAP transfer learning network model structure

型输出的特征图进行降维平均化、正则化,然后再将得到的结果直接输入 softmax 分类器中对图像的特征进行分类。

4 实验结果分析

4.1 对照实验分析

4.1.1 实验一的结果分析

实验一针对的是微调方法对视网膜图像分类的影响。图 5(a)、(d)分别是直接对三个预训练模型进行迁移学习的训练集曲线和验证集曲线;图 5(b)、(e)分别为微调后三个模型的训练集曲线和验证集曲线。可见,微调后的迁移学习网络的分类准确率显著提高,其中 InceptionV3 网络的验证准确率达到 96%,VGG19 网络的验证准确率达到 95%,ResNet50 网络的验证准确率达到 92%。

深度卷积神经网络^[16]的一个特征是靠前的网络层的特征更为一般化(适用于各种不同的数据集),而靠后的网络层包含的特征更加特殊。所以微调方法通过冻结预训练模型的初始部分网络层,训练靠后的网络层,既可以优化各网络层的权重之比,也可以加快网络的收敛速度。从图 5 所示的曲线可知,微调后的迁移学习网络迭代 30 次后,训练集曲线已经接近饱和,而直接迁移学习的网络训练集曲线迭代 60 次左右才达到饱和。

4.1.2 实验二的结果分析

实验二针对的是 GAP 函数对视网膜图片分类的影响。图 5(c)、(f)分别为采用 GAP 函数的迁移学习网络的训练集曲线和验证集曲线,将其与图 5(b)、(e)所示的微调网络进行对比后可知,采用 GAP 函数后,模型在训练准确率和验证准确率上均得到了提高。InceptionV3、VGG19、ResNet50 模型验证集曲线的准确率分别在 98%、97%、93%左右波动,且相比于微调后的模型准确率稍有提升。模型训练结束后采用测试集进行测试,由表 2 所示的准确率可知,在 GAP 迁移学习组中,InceptionV3、VGG19、ResNet50 模型的测试准确率分别为 97.3%、96.9%和 94.4%,优于 Bhowmik 等^[17]使用迁移学习方法得到的 94%的测试准确率,以及王翀等^[18]基于联合决策的神经网络方法得到的 94.5%的平均识别准确率。

表 2 对照实验中不同预训练模型的测试准确率

Table 2 Test accuracy of different pre-training models in control experiment

Group	Test accuracy / %		
	InceptionV3	VGG19	ResNet50
Fine-tuning transfer learning group	94.9	94.8	92.9
GAP transfer learning group	97.3	96.9	94.4

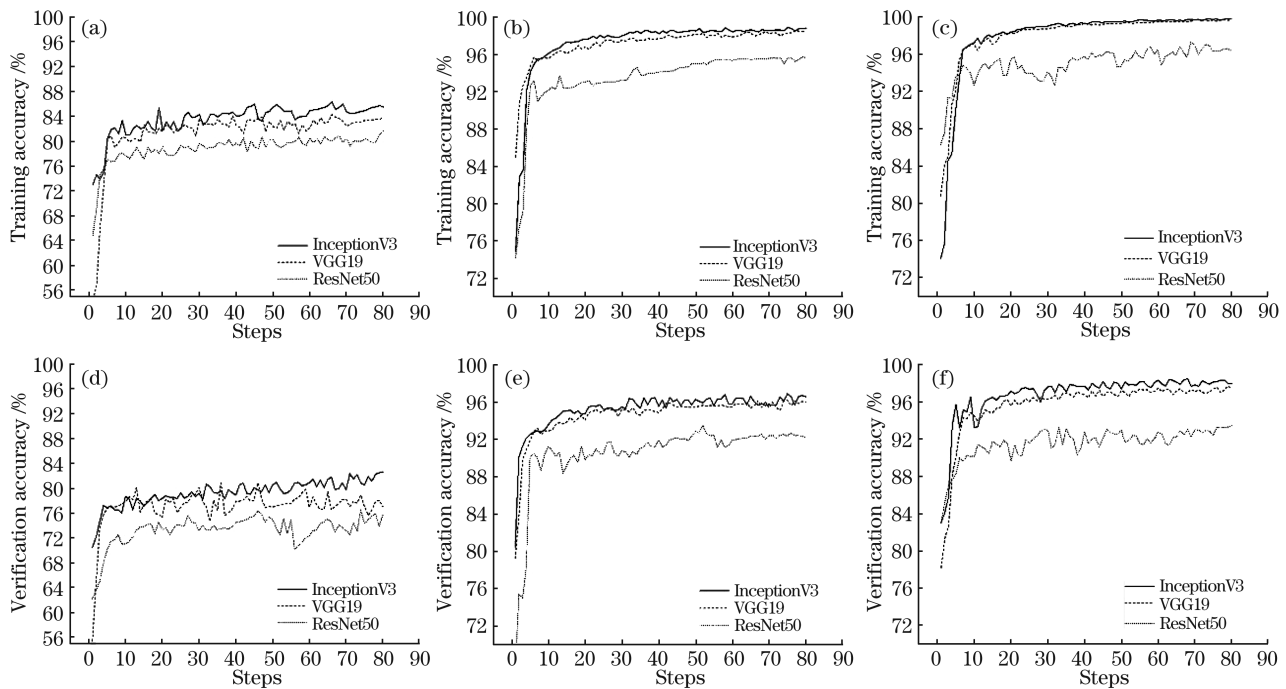


图 5 对照实验网络的训练集曲线。(a)(d)直接迁移学习组的训练集曲线;(b)(e)微调迁移学习组的训练集曲线;(c)(f)GAP 迁移学习组的训练集曲线

Fig. 5 Training set curves of control experimental networks. (a)(d) Training set curves of direct transfer learning group; (b)(e) training set curves of fine-tuning transfer learning group; (c)(f) training set curves of GAP transfer learning group

迁移学习网络在卷积层特征输出到分类器层过程中存在一个缺点,即参数量过大,容易造成过拟合,图像深层特征的提取量下降。GAP 函数跳过了将特征图转化为向量拼接然后再进行全连接的过程,直接对特征图进行特征维度的提取;另一方面,GAP 函数在结构上对网络进行正则化,可以减少训练过程发生的过拟合现象,有效提升图像分类的准确率。

4.2 ROC 分析

为了进一步探究 GAP 函数对糖尿病视网膜图像分类网络的作用,采用 ROC 曲线绘制 GAP 迁移学习组网络对四种 OCT 视网膜病变图像的分类结果。

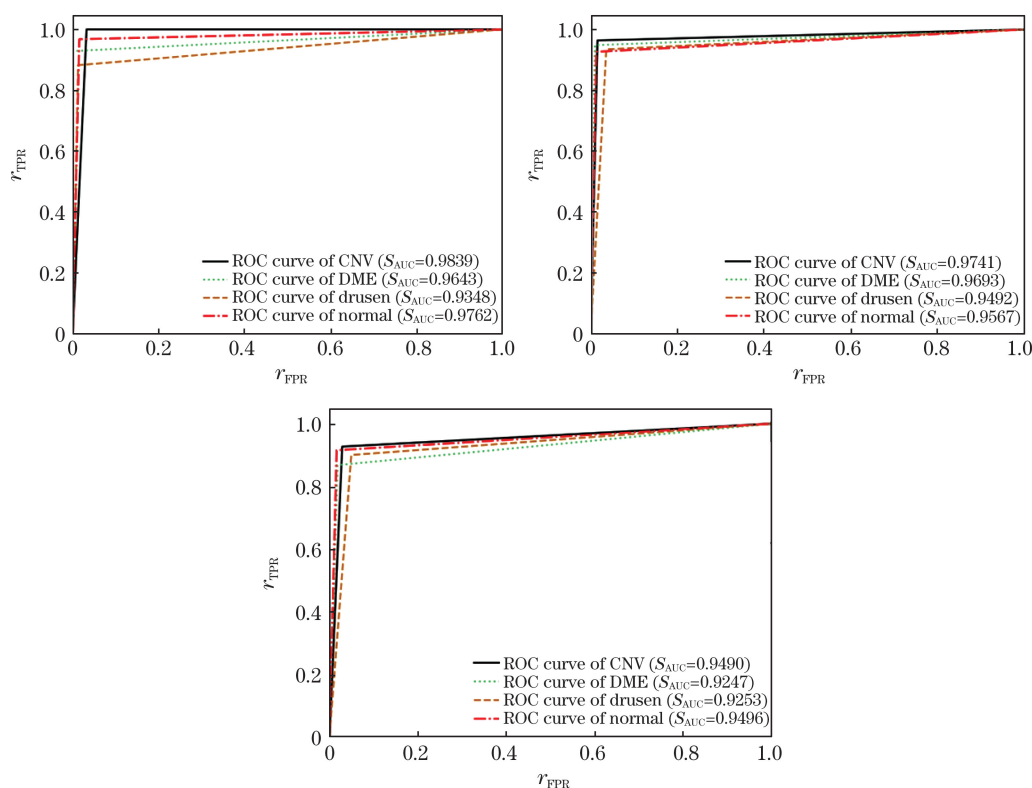


图 6 GAP 迁移学习组网络分类结果的 ROC 曲线。(a) InceptionV3 网络;(b) VGG19 网络;(c) ResNet50 网络

Fig. 6 ROC curve of GAP transfer learning group network classification result. (a) InceptionV3 network; (b) VGG19 network; (c) ResNet50 network

5 结 论

本文采用改进的迁移学习网络对 OCT 视网膜病变图像进行自动分类。首先采用高斯滤波算法去除散斑噪声,同时保留图像的边缘特征;然后采用微调开放相对应的模型网络权重,以解决数据量不平衡的问题,并用 GAP 函数替代传统的全连接层,跳过了将特征图转化为向量拼接然后再全连接的过程,直接对特征图进行特征维度的提取。此外,在结

构上对网络进行正则化,减少训练过程发生的过拟合现象,增强特征的空间映射鲁棒性。模型的分类准确率可达 97.3%,分类 AUC 值可达 0.9839,模型的有效性得到了证明。

参 考 文 献

- [1] Chen D W, Ran X W. New drugs for the treatment of diabetes mellitus[J]. Chinese Journal of Diabetes Mellitus, 2018, 10(2): 103-106.
陈大伟,冉兴无. 糖尿病新药热点[J]. 中华糖尿病

- 杂志, 2018, 10(2): 103-106.
- [2] Ophthalmology Group of Ophthalmology Society of Chinese Medical Association. Guidelines for clinical diagnosis and treatment of diabetic retinopathy in Ophthalmology (2014) [J]. Chinese Journal of Ophthalmology, 2014, 50(11): 851-865.
中华医学会眼科学会眼底病学组. 我国糖尿病视网膜病变临床诊疗指南(2014 年)[J]. 中华眼科杂志, 2014, 50(11): 851-865.
- [3] Huang D, Swanson E A, Lin C P, et al. Optical coherence tomography [J]. Science, 1991, 254 (5035): 1178-1181.
- [4] Barros Garcia J M B, Isaac D L C, Avila M. Diabetic retinopathy and OCT angiography: clinical findings and future perspectives [J]. International Journal of Retina and Vitreous, 2017, 3(1): 1-10.
- [5] Gao Y Z, Yuan Y, Ma Z H. High-resolution cortical blood flow imaging based on optical coherence tomography [J]. Laser & Optoelectronics Progress, 2019, 56(11): 111101.
高英哲, 袁毅, 马振鹤. 基于光学相干层析成像的皮层血流高分辨率成像 [J]. 激光与光电子学进展, 2019, 56(11): 111101.
- [6] Xu L X, Chen X J, Ban Y, et al. Method for intelligent detection of parking spaces based on deep learning [J]. Chinese Journal of Lasers, 2019, 46(4): 0404013.
徐乐先, 陈西江, 班亚, 等. 基于深度学习的车位智能检测方法 [J]. 中国激光, 2019, 46(4): 0404013.
- [7] Gargeya R, Leng T. Automated identification of diabetic retinopathy using deep learning [J]. Ophthalmology, 2017, 124(7): 962-969.
- [8] Ding P L, Li Q Y, Zhang Z, et al. Diabetic retinal image classification method based on deep neural network [J]. Journal of Computer Applications, 2017, 37(3): 699-704.
丁蓬莉, 李清勇, 张振, 等. 糖尿病性视网膜图像的深度学习分类方法 [J]. 计算机应用, 2017, 37 (3): 699-704.
- [9] Pang H, Wang C. Deep learning model for diabetic retinopathy detection [J]. Journal of Software, 2017, 28(11): 3018-3029.
庞浩, 王枫. 用于糖尿病视网膜病变检测的深度学习模型 [J]. 软件学报, 2017, 28(11): 3018-3029.
- [10] Zhuang F Z, Luo P, He Q, et al. Survey on transfer learning research [J]. Journal of Software, 2015, 26 (1): 26-39.
庄福振, 罗平, 何清, 等. 迁移学习研究进展 [J]. 软件学报, 2015, 26(1): 26-39.
- [11] Pan S J, Yang Q. A survey on transfer learning [J]. IEEE Transactions on Knowledge and Data Engineering, 2010, 22(10): 1345-1359.
- [12] Venkateswara H, Chakraborty S, Panchanathan S. Deep-learning systems for domain adaptation in computer vision: learning transferable feature representations [J]. IEEE Signal Processing Magazine, 2017, 34(6): 117-129.
- [13] Hsiao T Y, Chang Y C, Chou H H, et al. Filter-based deep-compression with global average pooling for convolutional networks [J]. Journal of Systems Architecture, 2019, 95: 9-18.
- [14] Lin D Y, Sun L, Toh K A, et al. Twin SVM with a reject option through ROC curve [J]. Journal of the Franklin Institute, 2018, 355(4): 1710-1732.
- [15] Zhang C X, Chen M H, Wang F, et al. Optical coherence tomography image denoising algorithm based on wavelet transform and fractional integral [J]. Laser & Optoelectronics Progress, 2019, 56 (18): 181008.
张晨曦, 陈明惠, 王帆, 等. 小波变换和分数阶积分结合的 OCT 图像去噪算法 [J]. 激光与光电子学进展, 2019, 56(18): 181008.
- [16] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks [J]. Communications of the ACM, 2017, 60(6): 84-90.
- [17] Bhowmik A, Kumar S, Bhat N. Eye disease prediction from optical coherence tomography images with transfer learning [C]//Engineering Applications of Neural Networks. [S.l.:s.n.], 2019: 104-114.
- [18] 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.