

应用视觉注意力的糖尿病视网膜病变眼底影像筛查及分级

万加龙¹ 胡建斌² 金炜东¹ 唐鹏¹

¹西南交通大学电气工程学院,成都 611756;²成都东区爱尔眼科医院眼底专科 610056

通信作者:胡建斌,Email:1809707278@qq.com

【摘要】 目的 构建基于视觉注意力的糖尿病视网膜病变(DR)辅助诊断的智能分析系统,实现 DR 眼底影像的自动筛查及分级。**方法** 从数据建模及数据分析竞赛平台(Kaggle)中的 Diabetic Retinopathy Detection 竞赛上下载得到 35 126 张 DR 眼底图片,并从 Messidor 网站上下载得到 1 200 张彩色眼底照片。首先,针对现有的 DR 眼底图像的特征,对视网膜图像进行一系列预处理;然后,在 VGG16 网络的基础上引入视觉注意力 SENet 模块,以提高病灶特征的显著性,生成一个网络结构较为复杂的深度卷积神经网络(CNN)SEVGG,该网络基本上继承了 VGG16 的一些结构参数,而 SENet 模块参数则根据基本网络及训练数据集进行相应的调整;最后,应用 SEVGG 网络模型对 DR 眼底图像进行筛查,并根据不同时期 DR 的病变程度把眼底图像分成不同等级。配置训练平台及环境并进行算法性能验证实验。**结果** 将本研究中提出的方法在不同的公开标准数据集上进行检验,最终在基于图像的分类上实现了较高的准确率,其中 Kaggle 数据集中 5 分类准确率可达 83%,病变检测的敏感性为 99.86%,特异性为 99.63%,Messidor 数据集中 4 分类准确率可达 88%,病变检测的敏感性为 98.17%,特异性为 96.39%。引入视觉注意力对于病灶点的关注更加显著,有助于 DR 的精准检测。**结论** 应用视觉注意力的 DR 眼底影像筛查及分级方法有效避免了传统人工特征提取和眼底图像分类的一些缺点,且对于病灶点的识别更加精确,不仅优于之前的方法,而且具有较好的鲁棒性及泛化性。

【关键词】 糖尿病视网膜病变;视觉注意力;卷积神经网络;注意力模块;分级

DOI:10.3760/cma.j.issn.2095-0160.2019.08.008

Screening and grading of fundus images of diabetic retinopathy based on visual attention

Wan Jialong¹, Hu Jianbin², Jin Weidong¹, Tang Peng¹

¹School of Electrical Engineering, Southwest Jiaotong University, Chengdu 611756, China; ²Fundus Specialist, Eastern Chengdu Aier Eye Hospital, Chengdu 610056, China

Corresponding author: Hu Jianbin, Email: 1809707278@qq.com

【Abstract】 Objective To construct an intelligent analysis system based on visual attention for diabetic retinopathy (DR) assistant diagnosis and to realize the automatic screening and grading of fundus images of DR.

Methods Total of 35 126 DR fundus images were downloaded from the Diabetic Retinopathy Detection competition in the Data Modeling and Data Analysis Competition Platform (Kaggle), and 1 200 fundus images were downloaded from the Messidor website. Firstly, according to the characteristics of DR fundus images, a series of preprocessing was carried out for retina images. Then, on the basis of VGG16 network, visual attention SENet module was introduced to improve the saliency of lesion features, and a deep convolution neural network SEVGG with complex network structure was generated. The network basically inherited some structural parameters of VGG16, and the parameters of SENet module were adjusted according to the basic network and training data set. Finally, the SEVGG network model was used to screen the DR fundus image, and the fundus image was divided into different levels according to the degree of lesions of DR in different periods. Configure the training platform and environment and perform algorithm performance verification experiments. **Results** The method proposed in this study was tested on different open standard datasets, and finally achieved high accuracy in image-based classification. The accuracy of 5 classification in Kaggle dataset was 83%, the sensitivity of lesion detection was 99.86% and the specificity was 99.63%. The accuracy rate of the 4 classification in the Messidor data set was up to 88%, the sensitivity of the lesion detection was 98.17%, and

the specificity was 96.39%. The introduction of visual attention was more significant for the focus of the lesion, which helped the accurate detection of DR. **Conclusions** This method effectively avoids some shortcomings of traditional artificial feature extraction and fundus image classification, and is more accurate for lesion recognition. It is not only superior to the previous method, but also shows better robustness and generalization.

[Key words] Diabetic retinopathy; Visual attention; Convolutional neural network; Attention module; Grading

DOI:10.3760/cma.j.issn.2095-0160.2019.08.008

糖尿病视网膜病变(diabetic retinopathy, DR)是糖尿病常见和严重的微血管并发症,是40~60岁人群视力损害的首要原因^[1-2],其早期诊断和治疗十分重要,眼底影像的精准分类是眼科医生对DR采取有效诊治方案的重要前提。随着深度学习技术在医学图像模式识别领域受到广泛关注^[3-5],已有研究采用基于迁移学习的BNnet网络对视网膜图像进行分类,利用CompactNet网络对眼底图像进行自动分级,利用增强的半监督生成对抗网络对DR等级和程度进行识别,以及通过构建两级深度卷积神经网络(convolutional neural networks, CNN)来完成对眼底图像的特征提取、特征组合和结果分类任务^[6-9]。然而,已有算法多存在检测步骤繁琐、识别精度不高等问题,且关于引入视觉注意力的DR检测目前研究尚少见。本研究中构建一种结合视觉注意力机制的CNN模型,以期对DR不同分期的眼底图像进行自动筛查及分级。

1 材料与方法

1.1 DR 图像收集及来源

从数据建模及数据分析竞赛平台(Kaggle)中的Diabetic Retinopathy Detection竞赛上下载得到35126张不同分期的DR图像。该数据集中的眼底影像属于高分辨率的红、绿、蓝(red-green-blue, RGB)三色图像,分辨率约为3500×3000,并依据DR病变程度将图片分为5个等级。不同时期的DR图像样本中,正常眼底图像为0级,具有轻微、中度、重度和增生性病变的眼底图像分别为1、2、3和4级。训练的数据集存在着严重的不平衡现象,其中0级图片数量约为4级的40倍(表1)。从Messidor官网上下载得到1200张彩色眼底照片,该数据集中图像的分辨率有1440×960、2240×1488和2304×1536像素3种。数据集中800张是瞳孔扩张图像,400张是瞳孔未扩张图像。根据微血管瘤、出血点、硬性渗出物和新血管的个数将每张图像病变等级分为0~3级(表2),且该数据也存在着不平衡现象。

表1 Kaggle 数据 DR 图片数量分布

病变等级	病变程度	图片数量
0	正常	25 810
1	轻微病变	2 443
2	中度病变	5 292
3	严重病变	873
4	增生性病变	708

注:DR:糖尿病视网膜病变

表2 Messidor 数据 DR 图片数量分布

病变等级	病变描述	图片数量
0	($\mu A = 0$) and ($H = 0$)	546
1	($0 < \mu A \leq 5$) and ($H = 0$)	153
2	(($5 < \mu A < 15$) or ($0 < H < 5$)) or ($NV = 0$)	247
3	($\mu A \geq 15$) or ($H \geq 5$) or ($NV = 1$)	254

注:DR:糖尿病视网膜病变; μA 表示微动脉瘤的个数;H表示出血点的个数;NV=0/1分别表示无/有新生血管

1.2 算法流程

算法流程主要分为数据预处理过程和神经网络训练过程^[7]。预处理包括去除背景及噪声影像、数据归一化和数据增强。(1)将下载的眼底图像进行挑选,并切掉图像周围的黑色边框,剔除噪声严重污染的图像。(2)将偏暗的眼底图像进行白平衡及直方图均衡化处理,对所有图片进行归一化处理,使所有眼底图像的色调统一,避免由于光线问题而产生的噪声干扰。(3)由于所得到的眼底图像样本量较少,且存在着数据的严重不平衡问题,故采用镜像、逆时针旋转90°/180°/270°、拉伸、翻转、平移等方法进行数据扩充。(4)在CNN的基础上引入视觉注意力机制,设计针对眼底影像特征的CNN模型,利用前馈式神经网络完成对DR眼底影像的筛查及分级任务(图1)。

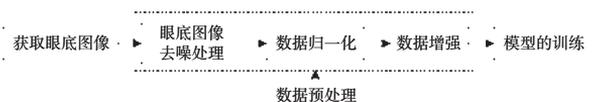


图1 算法流程图^[7]

1.2.1 数据集筛选及归一化处理 首先剔除无意义及噪声严重污染的图像(图 2A),裁掉图像中黑色边框(图 2B)部分,对光线偏暗的图像进行白平衡和直方图均衡化处理,解决由光线问题引起的噪声干扰(图 2C,D);所有处理过的眼底图像做归一化操作(图

2E),采用线性函数转化方式转换: $b = (a - C_{min}) / (C_{max} - C_{min})$,其中 a 是转化前的值, b 是转换后的值, C_{max} 和 C_{min} 分别对应眼底图像中的最大值和最小值^[6]。

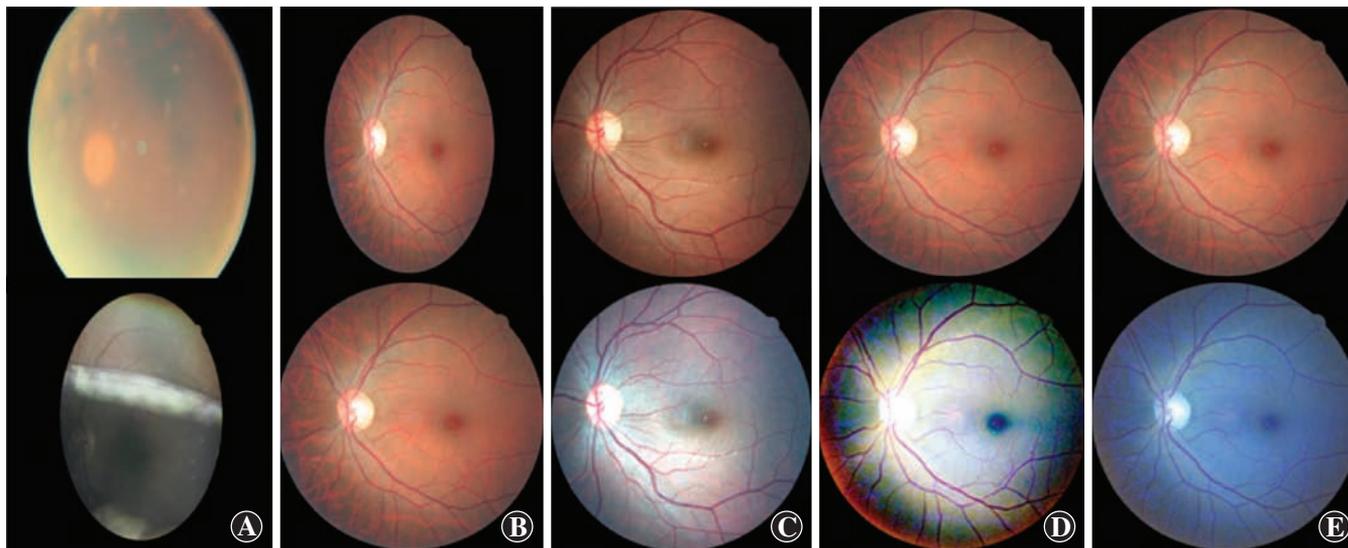


图 2 眼底图像预处理示例 A:需要删除的图像 B:切除黑色边框 C:白平衡处理 D:直方图均衡化 E:图像归一化 除了 A 组表示噪声图像外,其他 4 个组第 1 行均表示预处理之前图像,第 2 行均表示预处理之后图像

1.2.2 数据增强 为了消除网络模型的验证结果偏向某一种类别,对视网膜数据集进行数据扩充。在 DR 眼底影像原数据的基础上对眼底图像进行镜像、逆时针旋转 90°/180°/270°、拉伸、翻转、平移及组合操作进行数据扩增。Kaggle 数据中 0 级数量较多则挑选质量好的 2 万张,其他类别数量均扩增到 2 万张,使得实验数据集数量由原来的 3 万多张增加到 10 万张,这样将实验样本数量较少的类别向样本数目较多的类别靠拢,解决数据不平衡和数据量不足的问题。Messidor 数据集中每类数据量均较少,所以每类均扩增到 1 万张。

于通道域的视觉注意力模块 SENet^[13]。

1.2.3 SEVGG 网络模型 由于神经网络技术在图像识别领域中产生了很多非常经典的网络模型,如在 ImageNet 国际计算机视觉挑战赛(ILSVRC)中拔得头筹的 AlexNet(2012 年冠军)、GoogleNet(IlsVRC 2014 年冠军)及 VGG(IlsVRC 2014 年亚军)网络^[10-11]。为了在一定程度上降低网络模型的复杂度,本研究中选择 VGG16 作为基本的 CNN。而在计算机视觉领域,注意力机制被引入以进行视觉信息处理。随之产生了空间域的注意力机制、通道域的注意力机制、时间域的注意力机制以及注意力机制与深度残差网络的结合^[12-15]。本研究中根据 DR 的特点,采用基

本研究在 VGG16 网络的基础上引入视觉注意力 SENet 模块,构建成 SEVGG 网络模型(图 3)。该模型基本继承了大部分 VGG16 的结构参数,在 VGG16 网络中每个子部分后面都加入视觉注意力 SENet 模块(图 4)。此外,在每个 SENet 模块之前需加入 Ioffe 等^[16]提出的批归一化层 BN,全连接层的后面添加了 Dropout 层^[17],网络中的每个卷积层及全连接层后面都带有 ReLU 激活函数层以及网络模型的最后使用 Softmax 函数输出,图中输出类别可根据不同任务调整。

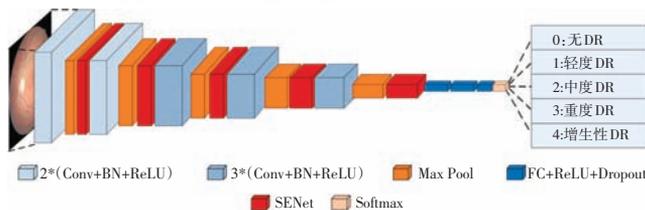


图 3 SEVGG 网络结构图 DR:糖尿病视网膜病变

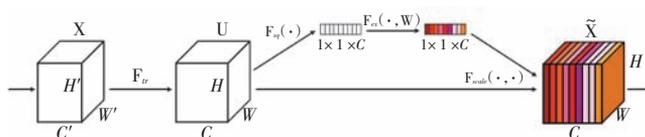


图 4 视觉注意力 SENet 模块结构图^[13]

1.3 实验平台设置及算法性能验证实验

本研究的训练实验平台及环境配置如表 3 所示。为解决数据格式不一致所带来的训练结果的偏差,本研究中的训练实验均采用统一的图片格式。此外,根据网络结构特点输入图片的尺寸均设置成 224×224 像素;并从数据扩增后的眼底图像中随机取 75% 作为训练集,用于网络模型的训练实验;剩下的 25% 作为测试集用于网络模型的测试实验。为验证本研究算法的优越性,在 2 种标准数据集上且均进行数据增强的条件下,分别就分类准确率、敏感性及特异性的结果进行对照。同时,为验证不同网络结构对实验结果产生的影响,本研究中还进行了自身对比实验。

表 3 训练平台及环境配置		
编号	配置	具体信息
1	深度学习框架	PyTorch 1.0
2	操作系统	Windows 10
3	GPU	NVIDIA GeForce GTX 1080
4	CPU	Intel E5-2630 v2@2.60 GHz
5	编程语言	Python 3.6

注:GPU:图形处理器;CPU:中央处理器

1.3.1 视觉注意力的增强效果实验 为验证 CNN 中引入视觉注意力 SENet 模块对 DR 病变区域的关注程度以及对实验结果的影响,本研究中采用增强后的眼底图像对不同结构的神经网络模型进行 BNvGG16 和 SEvGG 网络模型训练实验。为便于直接有效地观察视觉注意力机制的引入对网络训练损失的变化以及测试准确率的影响,分别研究 BNvGG16 和 SEvGG 的训练实验损失和测试准确率曲线。

1.3.2 批归一化层的提升效果实验 为验证在 CNN 中加入批归一化层 BN 对实验产生的影响,也是作为本研究的自身对比实验,本研究中采用数据扩充后的眼底图像对 VGG16 和 BNvGG16 网络模型进行实验。为便于直接观察批归一化层的加入对于网络模型收敛速度的影响,分别研究 VGG16 和 BNvGG16 网络模型的训练损失和测试准确率曲线。

1.3.3 视觉注意力增强效果的可视化实验 为进一步验证 CNN 中加入视觉注意力 SENet 模块对 DR 病

变区域的关注程度,根据卷积层学习到的特征可以很容易地被理解和可视化^[18]及提出的残差注意力网络中的可视化思想^[15],本研究中针对不同深度的网络层及引入视觉注意力模块进行 VGG16 和 SEvGG 网络模型可视化实验。

1.3.4 视觉注意力的通用性及有效性验证实验 为进一步验证视觉注意力 SENet 模块的通用性及有效性,本研究中以对 AlexNet 为基本网络^[14]及引入视觉注意力的 AlexNet 网络进行验证实验。为便于检验视觉注意力机制的引入对测试准确率的影响及其通用性,分别研究 AlexNet 和加入 SENet 模块的网络训练实验误差和测试准确率曲线。

1.4 评价指标

针对医疗图像的分类,可根据图像的分类准确率来评判该分类模型的优越性,例如将 Ntotal 代表测试集中眼底图像的总数量,Nrec 代表其中被正确分类的图像个数,则分类准确率 = Nrec/Ntotal^[6,19]。

此外,当前检验是否患病的自动检测方法有效性的评判指标主要有基于图像水平和基于病灶区域两类。在基于图像水平的评价指标中其单位为整张视网膜图像的个数,而在基于病灶区域水平的评价指标中其单位为区域的个数。在医学领域中常用的评价自动检测算法的指标主要包括敏感性、特异性、准确率和预测率,其中敏感性代表不被漏判的可能性;特异性代表正常视网膜图像不被误判的可能性;准确率表示评判对应算法的准确率;预测率一般作为基于病变区域的判断指标,表示视网膜图像中病变区域被检测到的可能性。由于眼底影像的分级是属于图像级的分类,所以本研究中采用敏感性和特异性作为评价指标,即根据识别整张眼底图像中是否患病来判定^[20]。

2 结果

2.1 不同等级的 DR 图像样本表现

DR 第 1 期与第 2 期间、第 2 期与第 3 期间眼底图像差异不明显(图 5,6),传统方法中人为地进行病灶特征提取,然后进行眼底图像的分类很有可能会产生

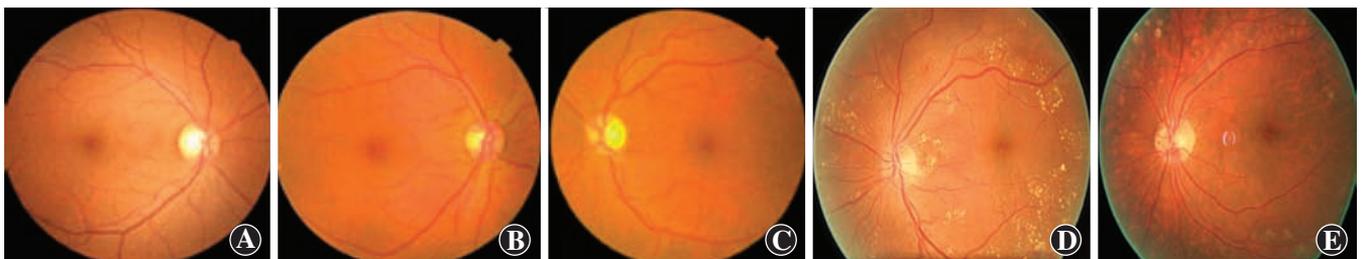


图 5 Kaggel 数据不同程度 DR 眼底影像 A:正常眼底图 B:轻微病变眼底 C:中度病变眼底 D:重度病变眼底 E:增生性病变眼底

误判,并且工作效率低下。此外,Kaggle 数据集中眼底图像采集时由于镜头、背景等其他客观因素的干扰,不同等级的眼底图像都不同程度地受到噪声污染(图 7)。

2.2 不同网络模型的实验结果对比

不同网络结构的分类准确率及其他评价指标结果如表 4,5 所示。其中,CompactNet 是丁蓬莉等^[7]采用的继承 AlexNet 前 2 层参数的网络;BNnet 是李琼等^[6]对 AlexNet 引进批归一化层的网络;增强的 SSGAN 是张文勇等^[8]针对 DR 识别的网络模型,表中选用各文献中与本研究相同条件下的最优实验结果。

2.3 视觉注意力 SENet 模块的增强效果对比

在 Kaggle 数据集下 BNVGG16 和 SEVGG 网络分别进行训练时,SEVGG 网络模型在测试集上的分类准确率、敏感性及特异性较高(表 4)。BNVGG16 网络在训练集上的损失下降速度和准确率的上升速度均比较缓慢,当训练迭代次数在 100 000 时,损失不再变化(图

8A),网络模型趋于收敛状态,分类准确率约为 76% (图 8B)。SEVGG 网络训练实验中,误差曲线的下降速度和准确率的上升速度较为平稳,当训练次数同样到达 100 000 时,SEVGG 网络模型趋于收敛状态(图 8C),分类准确率可达 83% (图 8D)。由实验结果可以看出 SEVGG 网络分类准确率明显高于其他模型,具有较好的分类效果。

在 Messidor 数据集下 BNVGG16 和 SEVGG 网络分别进行训练时,SEVGG 网络模型在测试集上的分类准确率、敏感性及特异性较高(表 5)。当训练迭代次数在 100 000 时,损失不再变化(图 9A),网络模型趋于收敛状态,BNVGG16 网络分类准确率约为 79% (图 9B)。SEVGG 网络训练次数在 80 000 时,损失不再变化(图 9C),网络模型处于收敛状态,SEVGG 网络分类准确率约为 88% (图 9D)。由实验结果可以看出 SEVGG 网络分类准确率明显高于其他模型,有较好的分类效果。

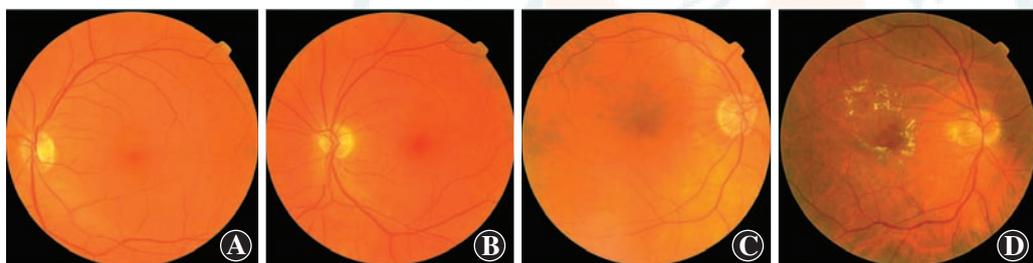


图 6 Messidor 数据不同程度 DR 眼底影像 A:正常眼底 B:轻微病变眼底 C:中度病变眼底 D:重度病变眼底

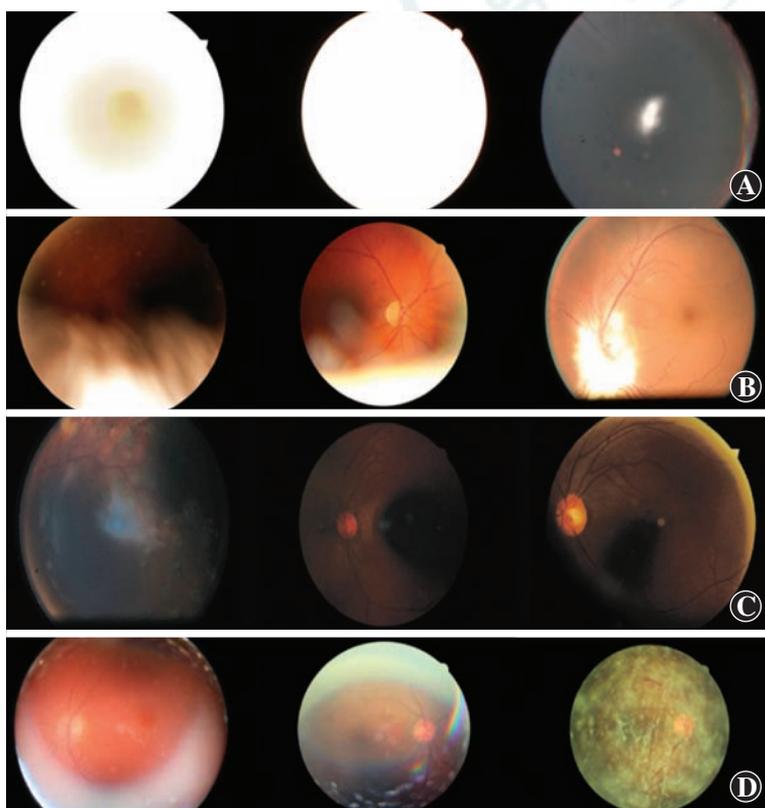


图 7 噪声图像 A:不存在任何纹理结构特点的第 1 类噪声图 B:光照引起的第 2 类噪声图 C:光线太暗引起的第 3 类噪声图 D:镜头引起的第 4 类噪声图图中列举每类中 3 组典型噪声图像,每一行为同类型噪声的图像

表 4 Kaggle 数据下不同网络模型训练的各类评价指标对比

方法	网络模型	分类准确率 (%)	敏感性 (%)	特异性 (%)
文献[7]	CompactNet	69	-	-
文献[6]	BNnet	71	-	-
本研究	AlexNet	63	88.17	86.06
本研究	SE+AlexNet	67	91.31	89.19
本研究	VGG16	65	96.60	95.36
本研究	BNVGG16	76	98.00	96.70
本研究	SEVGG	83	99.86	99.63

注:-:无数据

表 5 Messidor 数据下不同网络模型训练的各类评价指标对比

方法	网络模型	分类准确率 (%)	敏感性 (%)	特异性 (%)
文献[8]	增强的 SSGAN	77.2	92.1	86.1
本研究	AlexNet	64	89.65	88.53
本研究	SE+AlexNet	69	90.31	88.74
本研究	VGG16	70	93.09	91.54
本研究	BNVGG16	79	95.00	92.77
本研究	SEVGG	88	98.17	96.39

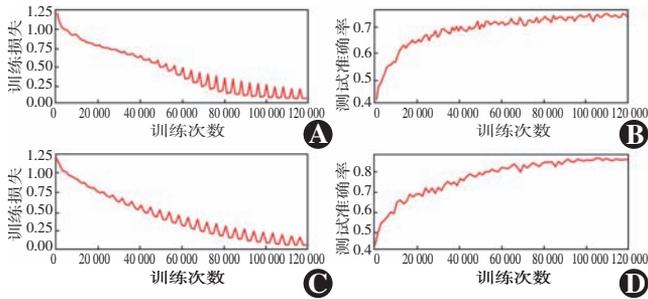


图 8 增强后的 Kaggle 眼底图像对不同结构的神经网络训练损失和测试准确率曲线图 A:BNVGG16 训练损失随训练次数的变化曲线 B:BNVGG16 测试准确率随训练次数的变化曲线 C:SEVGG 训练损失随训练次数的变化曲线 D:SEVGG 测试准确率随训练次数的变化曲线

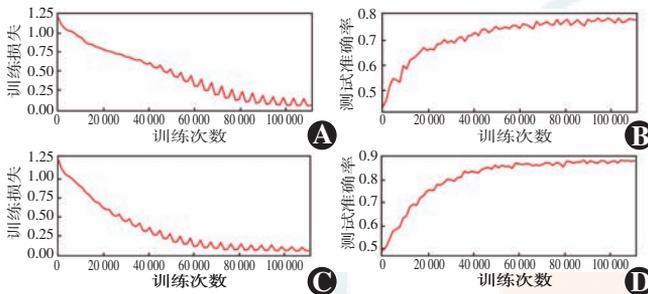


图 9 增强后的 Messidor 眼底图像对不同结构的神经网络训练损失和测试准确率曲线图 A:BNVGG16 训练损失随训练次数的变化曲线 B:BNVGG16 测试准确率随训练次数的变化曲线 C:SEVGG 训练损失随训练次数的变化曲线 D:SEVGG 测试准确率随训练次数的变化曲线

2.4 批归一化层对于网络模型实验结果的影响

VGG16 网络在训练集上的损失下降速度和准确率的上升速度均比较平缓,且均会出现一定程度的波动现象(图 10A);当训练次数达到 110 000 次时,网络模型趋于收敛状态,网络的分类准确率约为 65%(图 10B)。在 BNVGG16 训练实验中,误差曲线的下降速度和准确率的上升速度均较 VGG16 明显加快,当训练的迭代约 100 000 次时,BNVGG16 误差值已降至最低值,约为 0.001(图 10C)且准确率在 75% 以上,并且网络已经收敛。BNVGG16 网络的收敛速度相对较快,但在整个训练过程中网络模型并没有出现过拟合的现象(图 10D)。

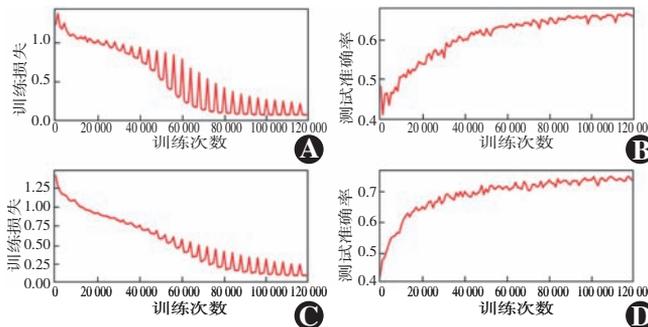


图 10 Kaggle 眼底图像对不同结构的神经网络训练损失和测试准确率曲线图 A:VGG16 训练损失随训练次数的变化曲线 B:VGG16 测试准确率随训练次数的变化曲线 C:BNVGG 训练损失随训练次数的变化曲线 D:BNVGG 测试准确率随训练次数的变化曲线

2.5 不同深度的网络层及引入视觉注意力的可视化实验结果对比

在相同网络模型中,不同层特征图响应的视觉注意力效果不同,在浅层结构中,网络的视觉注意力集中于眼底图中背景等区域;而在深层结构中,网络的视觉注意力特征图聚焦于待分类的 DR 病变区域。这些与之前的很多实验结论相类似,即深层次的特征图具有更高的抽象性和更强的语义表达能力,对于物体分类较浅层特征有较大的作用(图 11,12)。

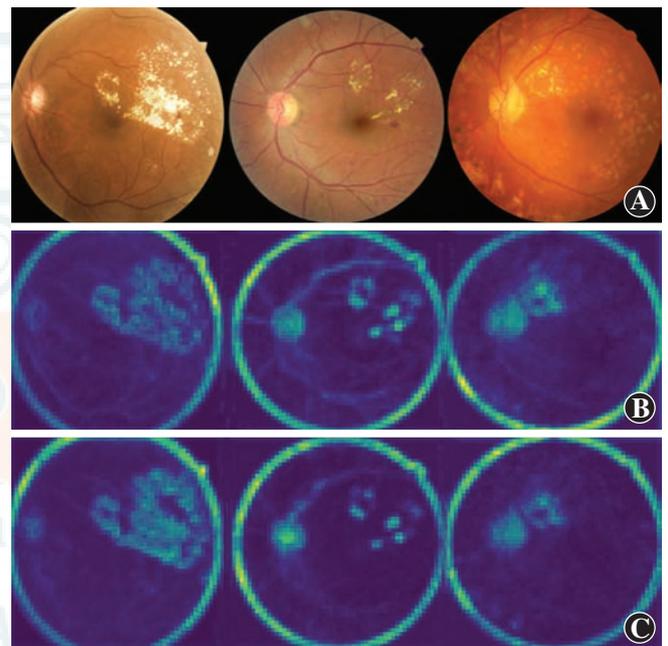


图 11 网络浅层可视化效果图 网络浅层特征融合眼底图像的可视化实验 A:原图 B:无视觉注意力图 C:视觉注意力图

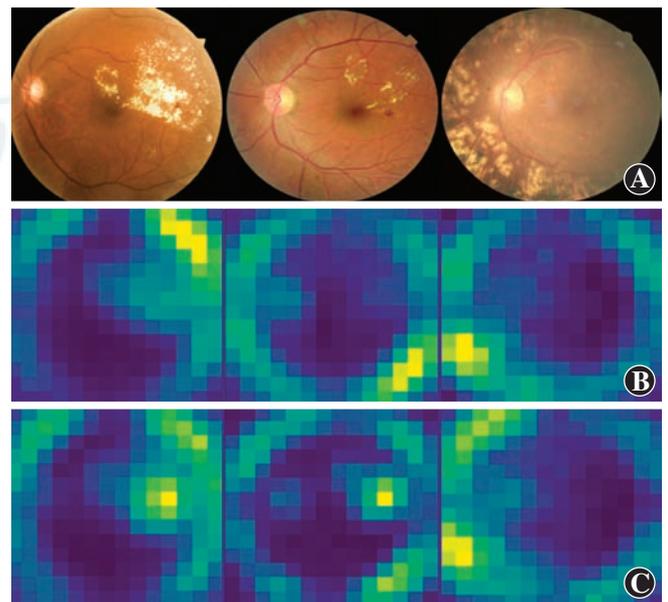


图 12 网络深层可视化效果图 网络深层特征融合眼底图像的可视化实验 A:原图 B:无视觉注意力图 C:视觉注意力图

相比较于网络浅层,网络中层引入视觉注意力 SENet 模块虽不如网络深层那样效果明显,但也在一定程度上关注了 DR 病变区域的特征,使其朝着网络深层更显著特征区域的方向发展(图 11,13)。

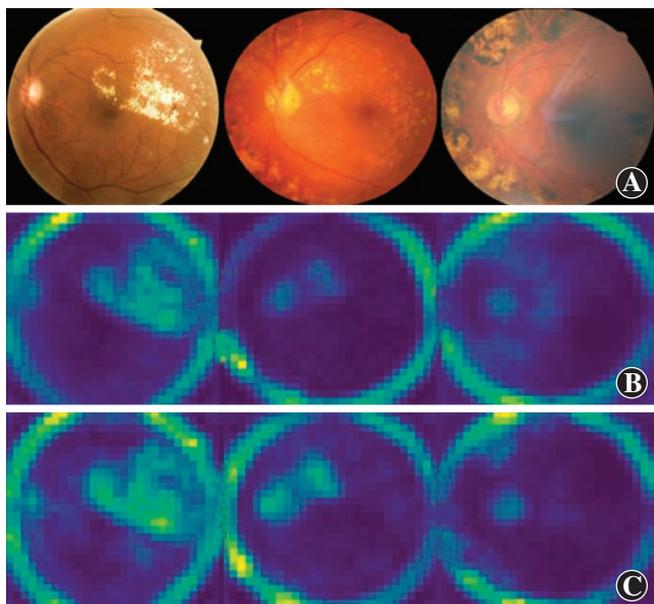


图 13 网络中层可视化效果图 网络中层特征融合眼底图像的可视化实验 A:原图 B:无视觉注意力图 C:视觉注意力图

2.6 视觉注意力通用性实验

本研究以 AlexNet 为基本网络在数据集上进行训练实验,当训练次数达到 110 000 次时,网络模型基本趋于收敛状态,AlexNet 网络的分类准确率约为 64% (图 14A, B)。引入视觉注意力 SENet 模块后,同样地,当训练迭代约 110 000 次时,网络模型也基本收敛,最后的分类准确率约为 69% (图 14C, D)。

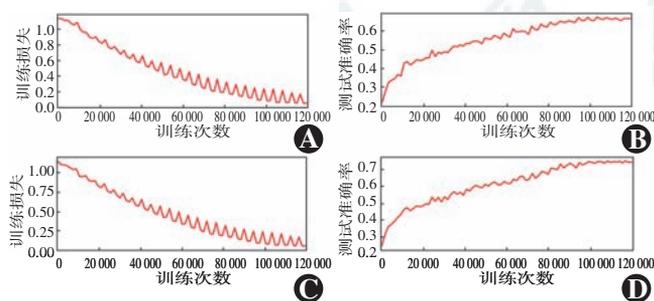


图 14 增强后的 Messidor 眼底图像对 AlexNet 网络训练损失和准确率曲线图 A: AlexNet 网络训练损失随训练次数变化曲线 B: AlexNet 网络测试准确率随训练次数变化曲线 C:加入视觉注意力的 AlexNet 网络训练损失随训练次数变化曲线 D:加入视觉注意力的 AlexNet 网络测试准确率随训练次数变化曲线

3 讨论

本研究利用深度学习结合视觉注意力的方法对不

同病变程度 DR 的眼底影像进行筛查及分级。因为眼底图像中存在噪声干扰大、数据量小、数据严重不均衡等问题,所以在进行网络模型训练前必须先对眼底图像进行一系列的预处理操作。在网络模型的构造方面,本研究基于 VGG16 网络结构及视觉注意力特点,结合 DR 眼底图像的特征,引入视觉注意力 SENet 模块以提高 DR 病灶点的显著性,构建了 SEVGG 网络模型。

为验证在相同的增强数据和基本网络的条件下,加入视觉注意力 SENet 模块对实验结果产生的影响以及引入批归一化层的必要性,进行了 3 个组的实验,结果表明引入批归一化层的网络比原始网络训练的效果好很多;加入视觉注意力模块能够提高 DR 病灶点的显著性,且分级效果好。为验证视觉注意力的通用性及有效性,以 AlexNet 为基本网络分别进行训练实验,结果表明视觉注意力 SENet 模块针对不同网络均能起到一定的加强作用,因此具有一定的普适性和有效性。

3.1 数据归一化及 BN 的作用

数据归一化的目的是加速网络模型收敛,从而达到更好的训练效果并且能够对 DR 眼底图像的特性点进行统一。传统的深层神经网络中,较高的学习率可能会导致梯度爆炸或者梯度消失,也有可能陷入局部最优化,BN 层可以很好地解决这个问题。在每层中归一化激活后,能防止因参数小幅度的改变而造成数据在网络传播过程中变动不断放大的问题。

本研究中批归一化层的加入不仅能加快网络模型的收敛速度,而且能够显著提高网络模型对 DR 病变分级的准确率。本研究中训练集较多且各类别均衡,保证了训练集和测试集特征分布一致,Dropout 的加入也在一定程度上避免了神经网络的过拟合。

3.2 视觉注意力作用及其通用性

在神经网络中,注意力模块通常是一个额外的神经网络,能够硬性选择输入的某些部分,或者给输入的不同部分分配不同的权重。本研究中的视觉注意力网络 SENet 是一种通过精确的建模卷积特征各个通道之间的作用关系来改善网络模型的表达能力的模块,即能够让网络模型对特征进行校准的机制,使网络从全局信息出发来选择性地放大有价值的特征通道并且抑制无用的特征通道。视觉注意力模块 SENet 的核心思想在于通过网络根据误差变化去学习特征权重,使得有效的特征图权重重大,无效或效果小的特征图权重小的方式训练模型达到更好的结果。

本研究中进行 DR 病灶点的检测过程,即通过 Squeeze 和 Excitation 2 个关键性操作方式,显式地建模特征通道之间的相互依赖关系,然后通过学习的方

式来自动获取到每个病灶点特征通道的重要程度,再根据这个重要程度去加强有用的 DR 病变特征,并抑制对当前分级任务中用处不大的病变特征。

本研究中结果表明,对于 AlexNet 网络是否加入视觉注意力模块同样影响着分类准确率,视觉注意力的加入在一定程度上能够提高模型分类准确率。由 2 种不同结构基本网络得到的实验结果均表明,视觉注意力 SENet 模块的加入能够提高模型分类精度。因此,视觉注意力具有一定程度上的通用性和有效性。

3.3 应用视觉注意力的 DR 检测

DR 病变在眼底图像中分布位置具有随机性,且形状、大小不一,增加了神经网络的检测难度。在基本网络 VGG16 中引入视觉注意力模块是本研究的创新点之一。为节约资源,且为降低网络模型的复杂度,所以这也是本研究采用 VGG16 网络作为基本网络的创新点之二。此外,对于以其他基本网络的训练实验是作为视觉注意力通用性及有效性的验证。

本研究中注意力机制模块的加入不仅能加快网络模型的收敛速度,而且能够提高网络模型分类准确率。在引入视觉注意力 SENet 模块后,在网络深层中对于 DR 病变区域特征的关注程度更加显著,使其不仅关注于最具代表性的病变特征,也在一定程度上关注更多重要特征,使 DR 的整体特征轮廓趋于明显。因此,视觉注意力的引入对于提高 DR 眼底影像的分类准确率起到了关键作用。

综上所述,本研究结果表明应用视觉注意力能够增加对 DR 检测有用病变特征的权重,使 DR 最具代表性的病变特征更加显著,对于提高 DR 检测识别效果具有重要意义。由于 DR 病变特征的特殊性以及网络结构的局限性,本研究结果对于临床要求还有差距。在后续的研究工作中,将会尝试调整网络结构,结合其他的视觉注意力机制模块,例如混合域的注意力模块,进一步提高 DR 检测识别率。

利益冲突 所有作者均声明不存在利益冲突

参考文献

[1] 美丽巴努·玉素甫,陈雪艺. 视力损害的流行病学研究[J]. 国际眼科杂志, 2010, 10(2): 304-307. DOI: 10. 3969/j. issn. 1672-5123. 2010. 02. 033.
Mehriban · Yusup, Chen XY. Epidemiology survey of visual loss[J]. Int J Ophthalmol, 2010, 10(2): 304-307. DOI: 10. 3969/j. issn. 1672-5123. 2010. 02. 033.

[2] 王芳,王春芳,闫建林. 45 岁以上糖尿病患者中糖尿病视网膜病变的患病率调查及相关危险因素分析[J]. 中华实验眼科杂志, 2013, 31(8): 783-787. DOI: 10. 3760/cma. j. issn. 2095-0160. 2013. 08. 018.
Wang F, Wang CF, Yan JL. Survey of diabetic retinopathy of the diabetic population over 45 years old[J]. Chin J Exp Ophthalmol, 2013, 31(8): 783-787. DOI: 10. 3760/cma. j. issn. 2095-0160. 2013. 08. 018.

[3] Shen D, Wu G, Suk HI. Deep learning in medical image analysis[J]. Annu Rev Biomed Eng, 2017, 19: 221-248. DOI: 10. 1146/annurev-bioeng-071516-044442.

[4] Litjens G, Kooi T, Bejnordi BE, et al. A survey on deep learning in medical image analysis[J]. Med Image Anal, 2017, 42: 60-88. DOI: 10. 1016/j. media. 2017. 07. 005.

[5] Shen W, Zhou M, Yang F, et al. Multi-scale convolutional neural networks for lung nodule classification[J]. Inf Process Med Imaging, 2015, 24: 588-599.

[6] 李琼,柏正尧,刘莹芳. 糖尿病性视网膜图像的深度学习分类方法[J]. 中国图象图形学报, 2018, 23(10): 1594-1603.
Li Q, Bai ZY, Liu YF. Automated classification of diabetic retinal images by using deep learning method[J]. J Image Graphics, 2018, 23(10): 1594-1603.

[7] 丁蓬莉,李清勇,张振,等. 糖尿病性视网膜图像的深度学习神经网络分类方法[J]. 计算机应用, 2017, 37(3): 699-704. DOI: 10. 11772/j. issn. 1001-9081. 2017. 03. 699.
Ding PL, Li QY, Zhang Z, et al. Diabetic retinal image classification method based on deep neural network[J]. J Comput Appl, 2017, 37(3): 699-704. DOI: 10. 11772/j. issn. 1001-9081. 2017. 03. 699.

[8] 张文勇,申妍燕,王书强. 基于增强型半监督生成对抗网络的糖尿病视网膜病变识别[J]. 图像与信号处理, 2019, 8(1): 1-8. DOI: 10. 12677/jisp. 2019. 81001.
Zhang WY, Shen YY, Wang SQ. Identification of diabetic retinopathy based on enhanced semi-supervised generation of anti-network[J]. Image Signal Process, 2019, 8(1): 1-8. DOI: 10. 12677/jisp. 2019. 81001.

[9] 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. DOI: 10. 13328/j. cnki. jos. 005332.
Pang H, Wang C. Deep learning model for diabetic retinopathy detection[J]. J Software, 2017, 28(11): 3018-3029. DOI: 10. 13328/j. cnki. jos. 005332.

[10] Krizhevsky A, Sutskever I, Hinton G. ImageNet classification with deep convolutional neural networks[J]. Adv Neural Inform Process Sys, 2012: 1097-1105.

[11] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J/OL]. Computer Sci, 2014, arXiv preprint arXiv: 1409. 1556. [#.XTUo4cSGWg.](http://www. oalib. com/paper/4068791)

[12] Jaderberg M, Simonyan K, Zisserman A. Spatial transformer networks[C]//Advances in Neural Information Processing Systems, 2015: 2017-2025.

[13] Hu J, Shen L, Albanie S, et al. Squeeze-and-excitation networks[J]. IEEE Trans Pattern Anal Mach Intell, 2019, DOI: 10. 1109/TPAMI. 2019. 2913372.

[14] Mnih V, Heess N, Graves A, et al. Recurrent models of visual attention[J]. Adv Neural Inform Process Sys, 2014: 2204-2212.

[15] Wang F, Jiang M, Qian C, et al. Residual attention network for image classification[J/OL]. Comput Vis Pattern Recogn, 2017: arXiv: 1704. 06904[cs. CV][2019-02-10]. <https://arxiv.org/abs/1704.06904>.

[16] Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift[C]//Proceedings of the 32nd International Conference on International Conference on Machine Learning, Lille, France: PMLR, 2015: 448-456.

[17] Srivastava N, Hinton G, Krizhevsky A, et al. Dropout: a simple way to prevent neural networks from overfitting[J]. J Machine Learn Res, 2014, 15(1): 1929-1958.

[18] Zeiler MD, Fergus R. Visualizing and understanding convolutional networks[C]//European Conference on Computer Vision, Cham: Springer, 2014: 818-833.

[19] Spanhol FA, Oliveira LS, Petitjean C, et al. A dataset for breast cancer histopathological image classification[J]. IEEE Trans Biomed Eng, 2016, 63(7): 1455-1462. DOI: 10. 1109/TBME. 2015. 2496264.

[20] 蔡震震,唐鹏,胡建斌,等. 深度卷积神经网络实现硬性渗出的自动检测[J]. 计算机科学, 2018, 45(z2): 203-207.
Cai ZZ, Tang P, Hu JB, et al. Auto-detection of hard exudates based on deep convolutional neural network[J]. Comput Sci, 2018, 45(z2): 203-207.

(收稿日期:2019-02-25 修回日期:2019-07-04)

(本文编辑:刘艳)