

# 基于深度学习的视盘自动检测

张贵英 张先杰

(遵义医学院医学信息工程系, 贵州 遵义 563003)

**摘要:** 传统的视盘检测需要手动提取视盘几何、血管和自身属性等特征,该方法在一定程度上依靠经验,且耗时耗力。基于卷积神经网络(CNN),使用两种网络结构,在两种不同大小的数据集上,自动提取视网膜图像中视盘的特征,并依据CNN特征进行视网膜图像中视盘的自动检测。研究采用深度学习架构Caffe,在公开的眼底图像数据集上进行了实验验证。结果表明,此方法简单易行,准确率高达98.04%,超过了现有方法。同时,通过实验得出了一些有益的结论,为进一步的研究工作奠定了基础。

**关键词:** 视网膜图像; CNN 特征; 视盘; 自动检测

中图分类号: TP391 文献标识码: A 文章编号: 1674-7798(2017)03-0027-06

DOI:10.13391/j.cnki.issn.1674-7798.2017.03.007

## Deep learning based on optic disk automatic detection

ZHANG Gui-ying ZHANG Xian-jie

(Medical Information Engineering Department, Zunyi Medical University, Zunyi, Guizhou, 563003)

**Abstract:** Traditional Optic Disk (OD) detection is based on manual extracting features, which include OD's geometry, blood vessels and own properties. OD is detected by these features. Unfortunately, this method is based on experience and fortune to some degree, the worse is that it is time-consuming and complicated. To address this problem, a method based on Convolution Neural Network (CNN) is proposed in this paper, which can automatically obtain the features of OD in retinal image on two kinds of nets and two different size data sets, then to locate OD via CNN characteristics of retinal image. The experiments on public fundus image data show that the accuracy rate reach 98.04% by using deep learning architecture Caffe. At the same time, experimental results show that method proposed in this study is simpler and more effective than the previous ones. What's more, some useful conclusions are obtained through the analysis and explanation of the experimental data, which lay a solid foundation for further research.

**Key words:** Retinal image; CNN feature; Optic disk; Automatic detection

视网膜眼底图像中视盘(Optic Disk, OD)的自动检测在大范围病变甄别中有重要意义。视网膜图像中视盘自动检测是该类图像分析和计算机辅助诊断各类眼底病变的前提和重要步骤。在糖尿病视网膜病变、高血压视网膜病变、视网膜脱离

的裂孔定位、眼底肿瘤定位以及视网膜血管性等疾病的诊断中,均需要事先进行视盘的准确定位。视盘在通常为近似圆形的淡黄色或白色的亮斑。人工定位视盘时,由于个人差异、眼底疾病、光照以及眼底血管闭塞、噪声和模糊边缘等问题

收稿日期: 2017-01-05

基金项目: 贵州省科技厅联合基金(黔科合LH字[2014]7597号); 遵义医学院硕士启动基金(编号: F-641)。

作者简介: 1. 张贵英(1982-),女,甘肃文县人,在读博士,遵义医学院副教授,研究方向: 图像处理、模式识别。

2. 张先杰(1982-),男,贵州遵义人,硕士研究生,遵义医学院讲师,研究方向: 图像处理。

的存在,使得准确地定位视盘并非易事。

## 1 视盘定位研究现状

科研人员主要依据眼底视盘的几何、血管和自身属性对视盘进行自动检测,取得了较好的效果。G. A. Hoover<sup>[1]</sup>利用模糊收敛算法进行血管分割来定位视盘的位置,该算法在 STARE 数据库中准确率达 89%。M. Foracchia<sup>[2]</sup>首先检测出视网膜图像中的主要的血管,然后根据两个抛物线模型顶点的交汇来定位视盘,此方法依赖血管的提取,对于血管不明显或者模糊的图像,鲁棒性不强。H. Li<sup>[3]</sup>利用视盘外观灰度值较大的特性,得到视盘候选区域,然后再定位视盘,使用此特性的还有 T. Walter<sup>[4]</sup>。除此之外,科研人员综合利用视盘的各种特性来定位视盘,如 A. Youssif<sup>[5]</sup>、A. E. Mahfouz<sup>[6]</sup>、S. Ravishankar<sup>[7]</sup>、H. Yu<sup>[8]</sup> 及其 A. Rovira<sup>[9]</sup>等工作。

目前视盘定位主要基于手动提取视盘的几何、血管和自身属性特征。首先对眼底图像进行预处理,然后提取特征,再进行特征选择(即特征表达),最后依据人工提取的特征检测视盘。良好的特征表达,对最终视盘检测的准确性起着至关重要的作用。然而,手工完成视盘特征的提取在一定程度上靠经验和运气,且其调节也需要大量的时间,是一个耗时耗力的工作。深度学习是一种特征自动学习方法,能自动提取刻画视盘本质属性的特征。

20 世纪 60 年代,Hubel 和 Wiesel<sup>[10]</sup>在研究猫脑皮层中用于局部敏感和方向选择的神经元时发现其独特的网络结构可以有效地降低反馈神经网络的复杂性,继而提出了卷积神经网络(Convolution Neural Network, CNN)。K. Fukushima<sup>[11]</sup>提出的新识别机是 CNN 的第一个实现网络。基于 CNN 的检测算法大大提高了检测目标识别的效率。最具有代表性的是深度学习鼻祖 Hinton 的工作<sup>[12]</sup>,作者训练深度神经网络对 LSVRC-2010 和 LSVRC-2012 的数据集图像进行分类,获得当时最高的检测率。

此后,卷积神经网络在分类方面的精度和计算效率得到极大提升。2014 年 Girshick 和 Ross

等人提出了 R-CNN<sup>[13]</sup>,其思想为将每个图片分为约 2000 个区域输入 CNN 训练。该方法取得了很好的检测效果,不过其检测速度很慢。为了提高检测速度,何凯名等人提出 SPPNet<sup>[14]</sup>。之后, Fast-RCNN<sup>[15]</sup>在 SPPNet 的基础上,使用显著性检测方法在原始图像上提取出预选区域,大大提高了检测速度。2015 年 Ren 等提出 Faster-RCNN<sup>[16]</sup>,为 Fast-RCNN 的进一步加速。

综上所述,卷积神经网络处理图像和视频等方面带来了突破<sup>[17]</sup>。视网膜图像中视盘的检测主要借助人工提取传统特征,耗时耗力,结果依靠经验和运气。本研究针对这一情况,提出使用卷积神经网络自动学习视盘的本质特征,并依据 CNN 特征进行视网膜图像中视盘的自动检测。

## 2 基于卷积神经网络的视盘检测

卷积神经网络(CNN)直接以二维图像作为输入,具有自动学习特征的能力。图 1 为本研究设计的 CNN 网络,名为 Net1。该网络由 2 个卷积层、2 个池化层和两个全连接层组成。

输入图像为 256\*256 的彩色图像,第一个卷积层采用 20 个 5\*5 的卷积核,步长为 1,得到 20 个 252\*252 的特征图。第一个池化层采用卷积核为 2\*2 的最大值池化,步长为 2,池化后得到 20 个 126\*126 的特征图。第二个卷积层采用 50 个卷积核,大小和第一个卷积层相同,其后第二个池化操作和第一个类似,最后获得 50 个 61\*61 的特征图。连接两个全连接层,最后使用 softmax 分类。

为了验证层数对分类正确性的影响,本研究使用了基于 imageNet<sup>[12]</sup>网络,命名为 Net2。该网络由 5 个卷积和 5 个池化层组成,卷积层的核大小分别为 11、5、3、3 和 3,池化层的大小均为 3,步长均为 2。

## 3 实验及结果分析

为了验证视盘 CNN 特征,本文在常用的眼底图像数据上做了扩展实验,并和存在的算法进行了比较,验证了本文算法的有效性。

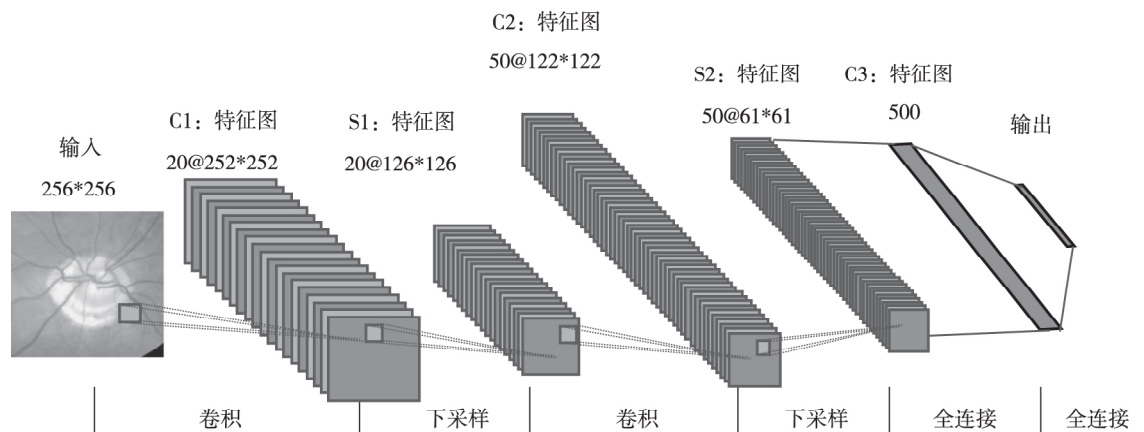


图1 Net1 网络结构

### 3.1 数据集

本文采用公开的 DRIVE<sup>[18]</sup> 数据集。该数据集共包括 40 张彩色眼底图片,大小为  $565 * 584 * 3$ ,其中 20 张作为训练集,20 张作为测试集,互不重叠。图 2 为 DRIVE 原始视网膜图像,图 2 (a) 为原始训练样本,图 2 (b) 为原始测试样本。

本实验将 DRIVE 训练集和测试集中的每副眼底图片剪裁成小块,并分别制作成正样本集和负样本集。具体操作为:对于训练集中的每张原视网膜眼底图像,以步长为 20,提取三张存在视盘的大小为  $256 * 256$  的小块作为正样本,提取 15 张无视盘的同样大小的小块作为负样本,如此操作,训练集中产生  $20 * 3 = 60$  个正样本和  $20 * 15 = 300$  个负样本。对于测试集中的每张原视网膜眼底图像,按同样的方法操作,得到 60 个正样本和

300 个负样本。称此种剪裁方法所获得数据集为 Data1。部分裁剪后的视网膜图像训练集和测试集中正负样本如图 3 和图 4 所示。

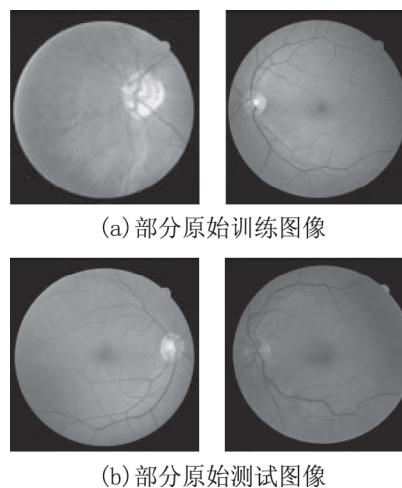


图2 部分原始视网膜图像

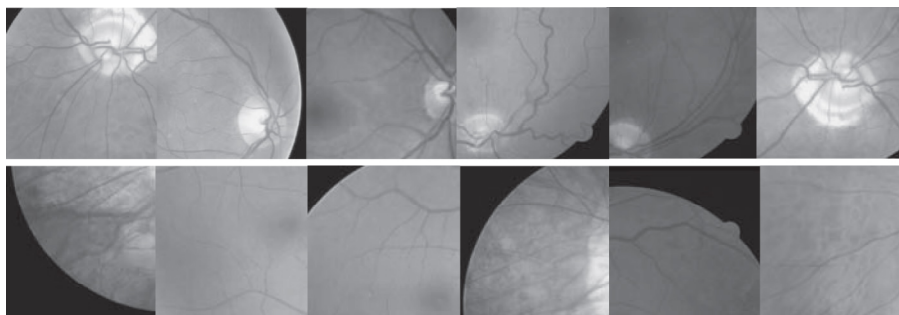


图3 Data1 中部分训练样本图例。第一行为正样本,第二行为负样本

为了验证样本数量对分类结果的影响,本研究在 Data1 的基础上,分别对训练集中的 60 个正样本旋转  $0^\circ$ 、 $90^\circ$ 、 $180^\circ$  和  $270^\circ$ ,以此获得

多个数据样本。按此方法,可获得  $60 * 4 = 240$  个正样本;对于测试集中 300 个负样本,按相同方法进行,得到  $300 * 4 = 1200$  个负样本,用此方法所

获得的数据集称之为 Data2。

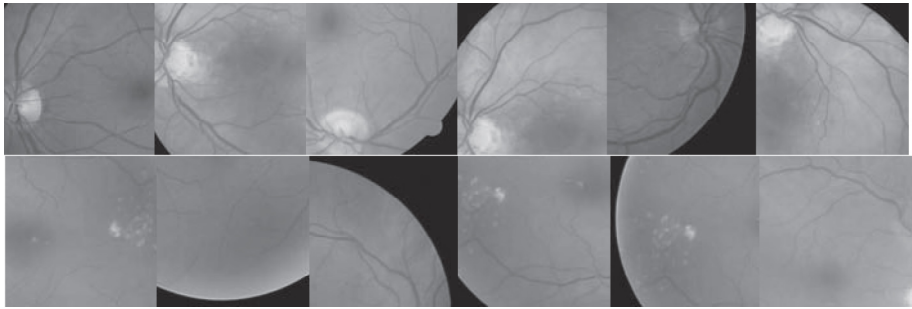


图 4 Data1 中部分测试样本图例。第一行为正样本,第二行为负样本

3.2 实验环境

硬件环境: 酷睿 i7 处理器,8G 内存,12G 显存的 GPU; 软件环境: Ubuntu 14.04, 深度学习框架 Caffe。

3.3 实验过程及结果分析

本研究所有实验均在深度学习框架 Caffe<sup>[19]</sup>上进行相应的实验。Caffe 是一个高效实用的计

算 CNN 相关算法的框架,由贾扬清博士编写和维护。在该框架中,所要实现的网络结构可在配置文件中指定,无需编码。该框架上手快,能够快速运行复杂的模型和海量的数据,且方便扩展,因此受到了科研工作者的广泛关注。本研究所采用 Solver 配置文件如表 1 所示。

表 1 Solver 配置

test_iter	test_interval	base_lr	lr_policy	gamma	max_iter	momentum	weight_decay
10	500	0.001	"step"	0.1	5000	0.9	0.0005

本研究分为四种情形进行实验。使用数据集 Data1 分别在卷积神经网络模型 Net1 和 Net2 上进行实验验证。使用数据集 Data2 分别在卷积神经网络模型 Net1 和 Net2 上再次进行实验验证。实验结果如表 2 所示。Iter 代表迭代次数,aver-

age 代表平均正确率,实验结果为数据集为 DataX,使用网络结构 NetX 下的正确识别率,X 代表 1、2。从表 2 中清楚的看出,在数据集为 Data2 和 CNN 网络为 Net1 时,本研究获得 98.04% 的准确率。

表 2 实验结果及分析

Model	iter = 1000	iter = 1500	iter = 2000	iter = 2500	iter = 3000	iter = 3500	iter = 4000	iter = 4500	iter = 5000	average
Data1 / Net1	97.80	98.40	97.20	97.00	98.2	97.00	98.40	97.6	97.8	97.71
Data2 / Net1	98.00	98.80	98.80	95.00	98.20	98.6	98.4	98.2	98.4	98.04
Data1 / Net2	96.60	94.8	96.80	95.00	95.4	96.4	97.00	96.2	96.2	96.04
Data2 / Net2	96.2	97.4	97.4	96.00	97.8	97.6	95.8	97.2	98.0	97.04

该方法无需提取视盘几何、血管和自身属性特征,更不需要分割血管,方法简单易行。该方法在训练数据为 Data2 和网络 Net1 下,最大正确率

为 98.04% (其它正确率分别为 97.71%、96.04% 和 97.04%)。文献 [20] 提出两种视盘检测算法,文献 [21] 就视盘检测算法提出一种算法。本



研究提出的算法和文献 [20 21] 在同一数据集 DRIVE 上的检测率进行了比较, 结果如表 3 所示。从表中清晰的看出, 本研究的视盘检测率均大大高于其他两种算法, 其中检测率最大提高了 18.04 个百分点。此为该论文的第二个贡献。

在网络为 Net1 时, 在数据集 Data2( 正确率为 98.04%) 上取得了比在数据集 Data1( 正确率为 97.71%) 上更好的准确率。相似地, 在网络为 Net2 时, 在数据集 Data2( 正确率为 97.04%) 上取得了比在数据集 Data1( 正确率为 96.04%) 上更好的准确率。说明样本的数量对准确率有一定的影响, 大样本对于准确率的提高起着有利作用, 此为该论文的第二个贡献。

表 3 DRIVE 数据集上的检测率

方法	检测正确率( % )
文献[20]中方法 1	80.00
文献[20]中方法 2	92.50
文献[21]	92.50
本文方法	98.04

据查文献可知, 本研究首次使用剪裁的方法产生视网膜图像数据集 Data1, 使用四种旋转的方法产生视网膜图像数据集 Data2。该方法扩充了数据集, 对于后续视盘检测工作起着积极作用。此为该论文的第三个贡献。

## 4 结束语

本研究在 DRIVE 数据集上的基础上, 通过两种大小不等的数据集 Data1 和 Data2, 在 Net1 和 Net2 两种网络模型上进行了实验, 得到了 98.04% 的正确率, 超过了已有算法。实验结果表明, 使用 CNN 自动提取特征, 无需分割血管, 不但大大减少了视盘检测所需要的人力物力, 还进一步提高了定位的准确性, 为眼科辅助系统提供了技术支持。本研究通过实验表明, 训练样本集越大, 检测正确率越高, 下步工作拟在更大的数据集上实验, 并设计复杂度相当的网络结构。传统的特征为视盘不可忽视的特征, 下一步工作准备融合传统特征和 CNN 特征, 筛选最能刻画视盘本质的特征, 以期进一步提高视盘检测的准确性。

## 参考文献:

- [1] A. Hoover, M. Goldbaum. Locating the optic nerve in a retinal image using the fuzzy convergence of the blood vessels [J]. IEEE Transactions on Medical Imaging, 2003, 22( 8): 951 – 958.
- [2] M. Foracchia, E. Grisan, A. Ruggeri. Detection of optic disc in retinal images by means of a geometrical model of vessel structure [J]. IEEE Transactions on Medical Imaging, 2004, 23( 10): 1189 – 1195.
- [3] H. LI, C. Opas. Automatic location of optic disk in retinal images [C]//Proceedings of the International Conference on Image Processing. Thessalonik: IEEE CS, 2001: 837 – 840.
- [4] Walter T., Klein J. C., Massin P., et al. A contribution of image processing to the diagnosis of diabetic retinopathy: detection of exudates in color fundus images of the human retina [J]. IEEE Transactions on Medical Imaging, 2002, 21( 10): 1236 – 1243.
- [5] Youssif A., Ghalwash A., Ghoneim A. Optic disc detection from normalized digital fundus images by means of a vessels' direction matched filter [J]. IEEE Transaction on Medical Imaging, 2008, 27( 1): 11 – 18.
- [6] Mahfouz A. E., Fahmy A. S. Ultrafast localization of the optic disc using dimensionality reduction of the search space [C]//Proceedings of the 12th International Conference on Medical Image Computing and Computer Assisted Intervention. London: IEEE, 2009: 985 – 992.
- [7] S. Ravishankar, Jain A., Mittal A. Automated feature extraction for early detection of diabetic retinopathy in fundus images [C]//Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on. IEEE, 2009: 210 – 217.
- [8] H. Yu, E. S. Barriga, C. Agurto, et al. Fast localization and segmentation of optic disk in retinal images using directional matched filtering and level sets [J]. Information Technology in Biomedicine, IEEE Transactions on, 2012, 16( 4): 644 – 657.
- [9] A. P. Rovira, E. Trucco. Robust optic disc location via combination of weak detectors [C]//Engineering in Medicine and Biology Society, 2008. EMBS 2008. 30th Annual International Conference of the IEEE. IEEE, 2008: 3542 – 3545.
- [10] Hubel, D. H. and T. N. Wiesel. Receptive fields, binocular interaction and functional architecture in the cat's

- visual cortex [J]. The Journal of physiology ,1962 ,160 ( 1) : 106 – 154.
- [11] Fukushima K. Neocognitron: A self – organizing neural network model for a mechanism of pattern recognition unaffected by shift in position [J]. Biological cybernetics ,1980 ,36( 4) : 193 – 202.
- [12] Krizhevsky A. ,I. Sutskever and G. E. Hinton ,ImageNet Classification with Deep Convolutional Neural Networks [J]. Advances in Neural Information Processing Systems 2012 ,25( 2) : 2012.
- [13] Girshick R. et al. Rich feature hierarchies for accurate object detection and semantic segmentation [A]. in Proceedings of the IEEE conference on computer vision and pattern recognition 2014.
- [14] Kaiming He ,Xiangyu Zhang ,ShaoqingRen ,JianSun. Spatial Pyramid Pooling in Deep Convolutional Networks for Visual Recognition [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence ,2014 ,37 ( 9) : 1904 – 1916.
- [15] Girshick R. Fast r – cnn [A]. in Proceedings of the IEEE International Conference on Computer Vision 2015.
- [16] Ren S. et al. Faster R – CNN: Towards real – time object detection with region proposal networks [J]. in Advances in Neural Information Processing Systems 2015.
- [17] LeCun Y ,Bengio Y ,Hinton G. Deep learning [J]. Nature 2015 ,521( 7553) : 436 – 444.
- [18] J. J. Staal ,M. D. Abramoff ,M. Niemeijer ,M. A. Viergever ,B. van Ginneken , " Ridge based vessel segmentation in color images of the retina " [J]. IEEE Transactions on Medical Imaging 2004( 23) : 501 – 509.
- [19] Y. Jia. Caffe: An open source convolutional architecture for fast feature embedding [DB/OL]. [http: //Caffe. berkeleyvision. org/](http://Caffe.berkeleyvision.org/). 2013.
- [20] Xiaolu Zhu et al. Detection of the optic disc in images of the retina using the Hough transform [A]. Engineering in Medicine and Biology Society ,2008. EMBS 2008. 30th Annual International Conference of the IEEE. IEEE 2008.
- [21] 罗汉源 ,戴培山. 一种快速的眼底图像视盘定位算法 [J]. 影像技术 2014( 05) : 20 – 21 ,17.

[责任编辑: 黄 梅]