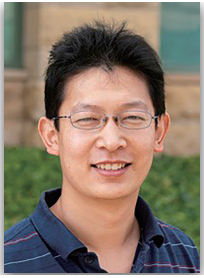


· 论 著 ·



李建强 北京工业大学软件学院科研副院长, 北京市特聘教授, 博士生导师, IEEE 高级会员 (Senior Member)。2004年1月在清华大学自动化系获博士学位; 2004—2005年在爱尔兰国立大学数字企业研究所 (DERI) 任高级研究助理; 2005—2013年在 NEC 中国研究院工作, 先后担任副研究员、研究员和高级研究员, 担任大规模语义计算研究组组长; 2009—2010年作为访问学者在斯坦福大学计算机系 InfoLab 实验室工作; 2014年作为北京市特聘教授加入北京工业大学软件学院, 担任北京市物联网工程中心物联网服务研究室主任, 并创建大数据技术实验室。主要从事语义计算、数据中心计算资源优化、数据挖掘、医疗大数据应用等研究。在医疗数据挖掘和信息处理领域发表论文 80 余篇, 其中 SCI 收录论文 30 余篇。申请发明专利 53 项, 其中 32 项已获中国、日本或美国专利授权。2014—2015年作为客座编委为 *Comput Ind* 期刊组织 *Emerging Information Technologies for Enhanced Healthcare* 学术专刊; 2016—2017年作为客座编委分别为 *IEEE Internet Things* 和 *Comput Ind* 期刊组织 *Emerging Social Internet of Things: Recent Advances and Applications* 和 *Big Data Technologies and Application* 学术专刊。

DOI: 10.16781/j.0258-879x.2018.08.0878

基于深度学习的白内障识别与分级

李建强^{1*}, 张苓琳¹, 张莉², 杨吉江³, 王青³

- 1. 北京工业大学软件学院, 北京 100124
- 2. 中国北京同仁首都医科大学附属北京同仁医院眼科中心, 北京 100730
- 3. 清华大学信息技术研究院, 北京 100084

[摘要] **目的** 利用深度学习自动提取眼底白内障特征, 构建白内障自动分类器, 并可视化分析深度网络中间层特征的逐层变换过程。**方法** 基于临床眼底图像, 使用深度卷积神经网络 (CNN) 从输入数据的原始表示直接学习有用的特征, 对比分析 CNN 自动提取的特征与预定义特征的性能表现。然后利用反卷积神经网络 (DN) 量化分析 CNN 各个中间层的特征, 进一步研究输入图像中对 CNN 的预测贡献最大的像素集, 探究 CNN 表征白内障的具体过程。**结果** 使用深度学习构建的分类器在四分类任务中达到 0.818 6 的平均准确率。与现有的预定义特征集相比, 利用深度 CNN 自动提取的特征集能提供更好的白内障特征表示。CNN 中间层特征呈现从低级抽象到高级抽象的分层变换, 如梯度变化到边缘, 然后到边缘状发散结构的组合, 最后到血管和视神经盘信息的高级抽象, 这种变换过程与临床检测白内障的诊断标准相吻合。**结论** 基于深度学习的分类器在性能表现上优于现有分类器。该方法对检测其他眼病也可能具有潜在的应用前景。

[关键词] 人工智能; 白内障; 深度学习; 深度卷积神经网络; 反卷积神经网络

[中图分类号] R 776.1 **[文献标志码]** A **[文章编号]** 0258-879X(2018)08-0878-08

Cataract recognition and grading based on deep learning

LI Jian-qiang^{1*}, ZHANG Ling-lin¹, ZHANG Li², YANG Ji-jiang³, WANG Qing³

- 1. School of Software Engineering, Beijing University of Technology, Beijing 100124, China
- 2. Beijing Tongren Eye Center, Beijing Tongren Hospital, Capital Medical University, Beijing 100730, China
- 3. Research Institute of Information Technology, Tsinghua University, Beijing 100084, China

[Abstract] **Objective** To automatically extract the characteristics of fundus cataract by deep learning, construct

[收稿日期] 2018-06-21 **[接受日期]** 2018-07-26

[基金项目] 国家科技重大专项(2017YFB1400803), 国家自然科学基金重点项目(71432004). Supported by National Science and Technology Major Project (2017YFB1400803) and Key Program of National Natural Science Foundation of China (71432004).

[作者简介] 李建强, 博士, 教授, 博士生导师.

*通信作者(Corresponding author). Tel: 010-67396842, E-mail: lijianqiang@bjut.edu.cn

a automatic classifier for cataract, and visualize the layer-by-layer feature transformation process of the intermediate layer of deep network. **Methods** Based on the clinical fundus image, a deep convolutional neural network (CNN) was used to directly learn useful features from the original representation of input data, and then the features extracted by the CNN were compared with pre-defined features. The deconvolution neural network (DN) method was used to quantitatively analyze the characteristics of each intermediate layer of CNN, analyze the pixel sets that have the most contribution to the prediction performance of CNN in the input image, and explore the process in characterizing cataract by CNN. **Results** The classifier constructed by deep learning achieved an average accuracy of 0.818 6 in four-category tasks. Compared with the existing pre-defined feature set, the feature set automatically extracted by the deep CNN performed better in representing characteristics of cataract. The features of the intermediate layer of CNN hierarchically transformed from low-level abstraction to high-level abstraction, including changed from gradient to edge, then to the combination of edge-like divergent structures, and finally to the high-level abstraction of blood vessel and optic disc information, and this transformation process coincided with the clinical diagnostic criteria of cataract. **Conclusion** The classifier based on deep learning is superior to the existing classifier in terms of performance. In addition, this method has potential application in detecting other eye diseases.

[Key words] artificial intelligence; cataract; deep learning; deep convolutional neural network; deconvolution neural network
[Acad J Sec Mil Med Univ, 2018, 39(8): 878-885]

白内障是视力受损的主要原因,也是导致失明的严重眼疾之一。据统计,世界上视力受损的人数约为 1.91 亿,其中 3 240 万人失明,18.4% 的视力损害病例和 33.4% 的失明病例是由白内障引起^[1]。预计由于白内障导致视力丧失的例数可能在 2025 年达到 4 000 万^[2]。

白内障筛查的主要目的是防盲工作,筛选出需要进行手术的白内障患者接受手术治疗;对于不需要手术治疗的轻度白内障患者,也可检查其眼底有无其他疾病,避免至白内障严重眼底看不清时盲目手术导致术后效果不佳。白内障检查耗时、昂贵,在低收入和中等收入国家和地区,由于健康投资较低,给白内障防盲工作的大范围普及带来了困难,导致白内障致盲率更高^[3]。构建白内障自动分类器对于降低白内障筛查的工作成本和和健康投资较低区域普及白内障筛查工作具有重要的实际意义,因此,基于医学影像的人工智能辅助白内障诊断技术越来越受到研究人员的关注。

在临床白内障筛查过程中,与裂隙灯检查法和虹膜投影法相比,非散瞳眼底照相机筛查白内障具有简单、安全、快捷、高效等优点,而且眼底数码照片的可存储和运输性使眼底图像成为一种更常用的白内障诊断依据。目前基于眼底图像的白内障自动分类研究方法主要包括预处理、特征提取、特征选择和分类器^[4-9]。在图像的获得和转移过程中可能会引入噪声,因此,预处理时需增强图像,如图像改善和噪声去除^[4]、视网膜结构

的分割和定位^[8]等。特征提取是白内障分类器构建过程中至关重要的一步,颜色、纹理、小波、轮辐、频谱等^[4-7]特征均可被用于构建分类器。特征选择主要用于防止特征维数爆炸,构建最优特征集。特征提取和特征选择合称为特征表示,良好的特征表示对于分类器的准确性有着关键的作用。最后是机器学习分类算法,不同的分类算法已被应用于白内障的检测和分级^[8-9]。

目前许多研究都集中于使用人工提取预先定义的特征集进行白内障分类^[4-9],即通过人工提取一组预定义的特征集构建白内障自动分类器,但这些预定义特征可能不具代表性、不完整或冗余,甚至可能会引入干扰特征。而且所有预定义的特征都是人为提取的,这是一种非常费力的启发式(需要专业知识)方法,在很大程度上取决于经验和运气。Gao 等^[10]提出了使用基于深度学习的卷积递归神经网络(convolutional-recursive neural network, C-RNN)自动提取的裂隙灯图像特征进行白内障分类,但缺乏对卷积神经网络(convolutional neural network, CNN)自动提取特征的可解释性分析,没有对自动提取的特征含义及特征的可用性进行说明。Zhang 等^[11]提出使用 CNN 自动提取的眼底图像特征进行白内障分类,可视化分析了最后一层池化层(pool5)的特征图,但 pool5 层的 6×6 特征图尺寸较小,且单凭 pool5 层的特征无法分析 CNN 中间层特征的转换过程,也无法详细分析特征集。

为解决以上问题,本研究使用深度 CNN 自

动提取眼底图像特征, 构建白内障自动分类器, 并利用反卷积神经网络 (deconvolutional neural networks, DN) 量化分析 CNN 如何表征眼底图像数据, 进一步研究 CNN 自动提取特征集的可解释性和可用性。

1 材料和方法

1.1 数据集与预处理方法

1.1.1 眼底图像数据集 白内障通常分为轻度、中

度和重度白内障^[7], 血管和视神经盘是其检测和分级的主要参考。图 1 所示为正常眼底和 3 个不同级别的白内障样本的眼底图像, 可见正常眼底图像中视神经盘、大小血管均清晰可见 (图 1A); 轻度白内障眼底图像中的血管信息少于正常眼底, 仅小血管不可见 (图 1B); 中度白内障的眼底图像中主要血管和视神经盘可见, 而大小血管不可见 (图 1C); 重度白内障的眼底图像中视神经盘模糊不清 (图 1D)。

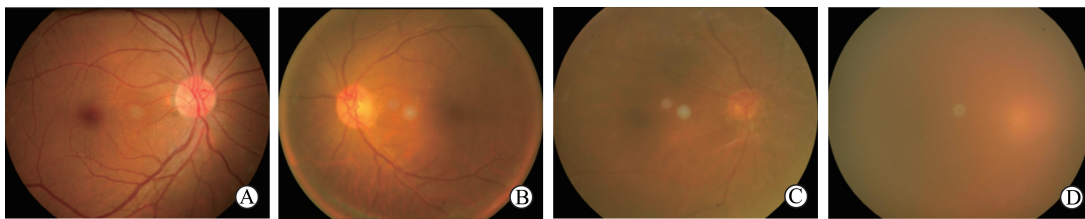


图 1 正常眼底与 3 个不同级别的白内障样本的眼底图像

Fig 1 Fundus images of normal fundus and three different levels of cataract diseases

A: Non-cataract; B: Mild cataract; C: Moderate cataract; D: Severe cataract

本研究试图找到开放的白内障眼底图像数据集用于训练分类模型, 但目前没有发现开放的数据集。本研究所使用图像的受试者来源于中国大陆地区, 年龄为 10~90 岁。为了使深度 CNN 能平等地学习各类眼底图像的特征, 使用随机过采样和欠采样的方法平衡各类别图像数据, 最终正常眼底与轻度、中度、重度 3 个不同级别白内障的图片数量均为 3 000。

1.1.2 去隐私处理 由于本研究采用的眼底图像来自不同规格的眼底摄像头, 图像尺寸大小不一, 首先需要将眼底图像尺寸统一为 3 048×2 432 像素大小。其次, 利用 MATLAB 编写设计椭圆截取和矩形切割算法切割图像, 在消除患者信息的同时尽可能保留眼底图片的原始信息。

1.1.3 消除眼底图像中不均匀色彩的影响 临床眼底照相时由于受局部不均匀照明和眼睛反射的影响, 难以精确检测和分级白内障。在前期研究中, 相较于其他两通道图像, 绿色通道图像最清楚且可保留最多的原始彩色图像基本信息, 能有效辅助特征识别。因此, 本研究在预处理阶段使用 G-filter 提取原始图像的绿色通道图像, 消除眼底图像中不均匀色彩的影响, 辅助白内障分类。

1.2 深度学习方法

1.2.1 基于深度 CNN 构建白内障自动分类器 深度

CNN 是一种人工神经网络, 它的权重共享网络结构使其更接近生物神经网络, 降低了网络模型的复杂度, 减少了权重的数量。当网络输入为多维图像时, 这种优势更加明显, 图像可以直接用作网络输入, 避免了传统识别算法的复杂特征提取和数据重构过程。卷积网络是专门设计用于识别二维形状的多层感知器, 这种网络结构对于平移、缩放、倾斜或其他形式的变形具有高度不变性。

与训练新的 CNN 结构不同, 本研究使用预先训练好的网络结构 AlexNet^[12]。原因如下: (1) 从大规模目标识别研究中以完全监督方式训练的 CNN 可以重新用于新的一般任务^[13]; (2) 本研究的训练集从数量和多样性上不如 ILSVRC2012 数据集^[14], CNN 模型的性能高度依赖训练集的多样性水平和数量大小; (3) 在许多备选的深度学习网络结构中, 本研究选择最轻量级和简单的网络结构来构建白内障分类器。最终, 本研究使用的白内障自动分类模型的网络结构如图 2 所示, 该网络结构共 8 层, 前 5 层是卷积层, 后 3 层是完全连接层。

在卷积层中, 其特征图是由前一层的特征图与一组权重 (也称过滤器) 进行卷积计算得到的。也就是说, 在 l 层的 i 通道的特征图 $Y_i^{(l)}$ 可定义为:

$$Y_i^{(l)} = \sum_{j=1}^{m^{(l-1)}} K_{j,i} \times Y_j^{(l-1)}$$

其中, K 代表过滤器, $i=1, 2, \dots, m^{(l)}$ 。

然后将所得到的特征图通过非线性单元 (ReLU) 输出。接下来在池化层中, 每个特征图将通过最大池化算法在 $q \times q$ 大小的邻接区域上进行子采样, 产生所谓的池化特征图。完成前 5 层的卷积和池化

操作后, 将其输出连接到完全连接层, 然后被馈送到四路 Softmax 分类算法, 在 4 个标签上产生概率分布, 输出最后的分类结果。

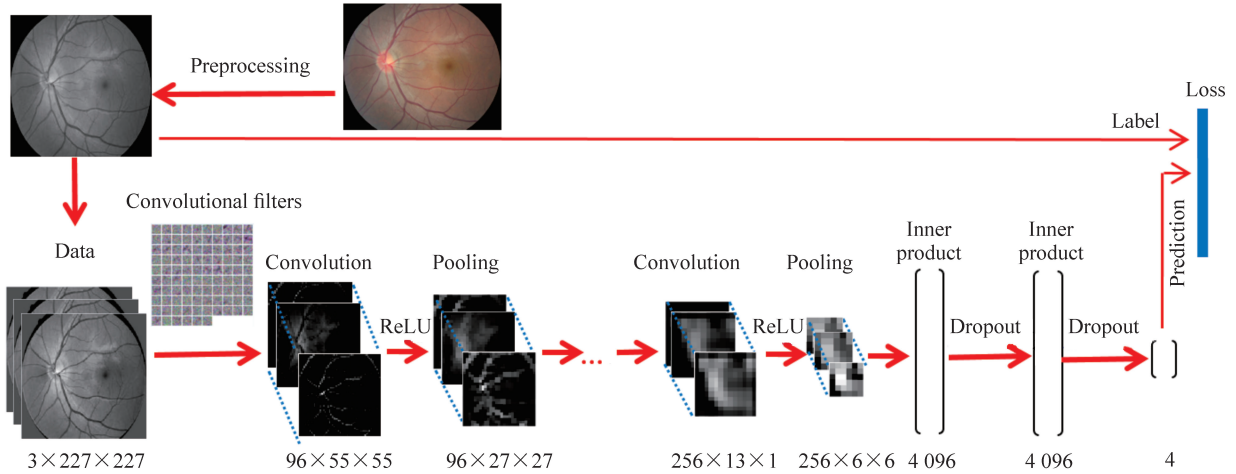


图 2 基于深度卷积神经网络 (CNN) 的白内障分类器结构图

Fig 2 Structure diagram of classifier for cataract grading based on deep convolutional neural network (CNN)

1.2.2 基于 DN 的特征可视化分析 深度 CNN 模型通过反向传播机制学习和优化每一层中的滤波器, 这些学习过滤器提取了唯一表示输入眼底图像的重要特征。因此, 为了理解 CNN 模型如何表征眼底图像, 需要通过 DN 方法可视化过滤器来观察特征的变换^[15]。本研究扩展了文献 [16] 中的方法, 对于层 l 中的所有绝对激活, 只考虑前 S 个最

大像素值, 其余部分设置为零, 并向下投影到像素空间, 构造定义如下的图像:

$$Y_i^{(l-1)} = \sum_{j=1}^{m^{(l)}} K_{j,i}^T \times Y_j^{(l)}$$

其中 $S=1, 2, \dots, \text{size}(Y_i^{(l)})$ 。基于此, 可以观察到该层输入图像中的高度活跃区域。图 3 描述了本研究以自下而上和自上而下的方式理解白内障自动分类的过程示意图。

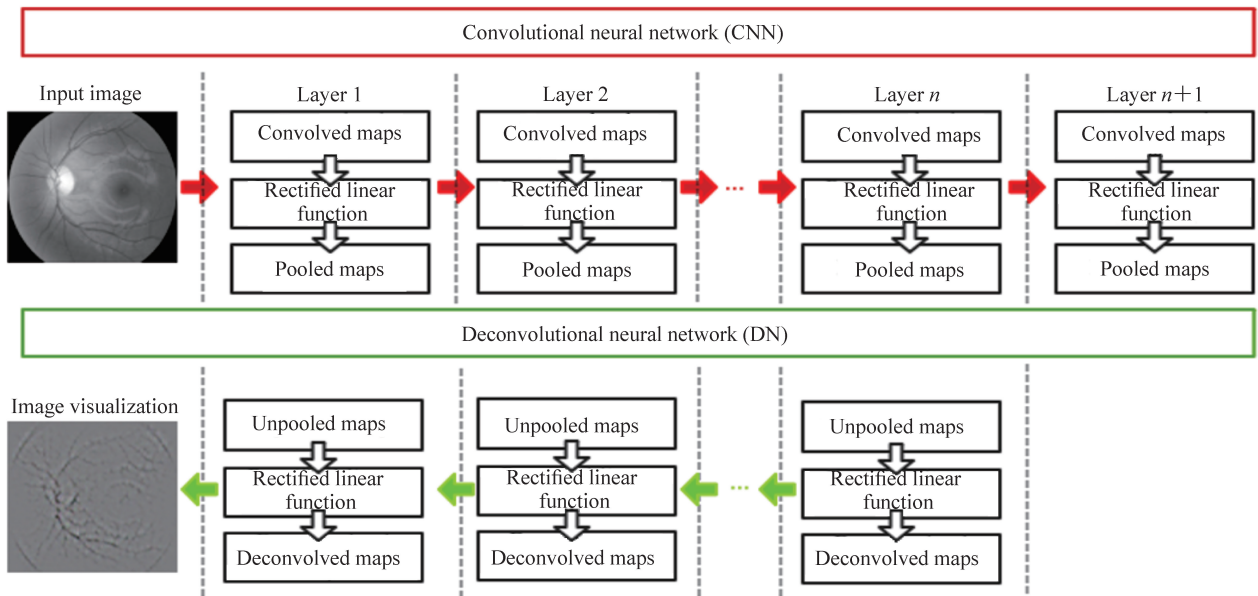


图 3 可视化分析卷积神经网络 (CNN) 中间层特征方法的总体架构

Fig 3 Overall architecture of visualized features of intermediate layer of convolutional neural network (CNN)

2 实验设计与结果

2.1 实验环境 本研究采用 Intel(R) Xeon(R) CPU E5-1620 v4 @ 3.50 GHz 处理器, 8 GB 内存, Quadro M2000 GPU 和 Ubuntu 16 操作系统。利用 Matlab 图像处理工具箱实现眼底图像的预处理。使用 Caffe (convolutional architecture for fast feature embedding) [17] 深度学习框架进行深度 CNN 的相关实验, 构建白内障自动分类器。对于训练过程中的参数设置, 本文采用逐步学习策略 Step, 初始学习率设为 10^{-3} , 每迭代 1×10^3 次, 学习率下降 0.1 倍, 设置 5×10^4 次迭代后停止。批样本数设置

为 10, 动量 (momentum) 设为 0.9, Dropout 比率设为 0.5。

2.2 基于深度 CNN 的白内障分类器性能验证

2.2.1 数据集大小对 CNN 分类器性能的影响 本研究分析了数据集大小对 CNN 分类器性能的影响。采用交叉验证实验, 每次分别使用各类随机抽样 1%、5%、10%、20%、30%、40%、50%、60%、70%、80%、90%、100% 的数据作为训练样本并分别进行 10 次独立交叉验证。实验结果如表 1 所示, 可见随着数据集规模的增大, 分类器的性能更好, 其表现的准确率也越稳定, 说明本研究采用的数据集大小具有提供参考性准确率的标准。

表 1 不同数量的眼底数据集对 CNN 在白内障分级任务中准确率的影响

Tab 1 Effect of different numbers of fundus datasets on accuracy of CNN in cataract grading task

Number	1	2	3	4	5	6	7	8	9	10	Acc	Var ($\times 10^{-5}$)
120	0.65	0.58	0.82	0.68	0.55	0.65	0.72	0.72	0.81	0.85	0.704 8	1 018.60
600	0.73	0.77	0.81	0.71	0.78	0.78	0.81	0.79	0.79	0.75	0.765 1	98.79
1 200	0.77	0.75	0.81	0.81	0.75	0.81	0.78	0.77	0.82	0.80	0.776 4	68.07
2 400	0.81	0.79	0.75	0.78	0.80	0.80	0.80	0.77	0.82	0.79	0.782 6	39.11
3 600	0.77	0.81	0.81	0.80	0.79	0.78	0.80	0.80	0.79	0.77	0.793 8	19.66
4 800	0.81	0.82	0.81	0.80	0.78	0.81	0.80	0.80	0.79	0.80	0.797 9	11.33
6 000	0.82	0.81	0.79	0.81	0.82	0.82	0.81	0.79	0.80	0.81	0.793 2	12.21
7 200	0.83	0.82	0.82	0.81	0.82	0.80	0.81	0.81	0.80	0.81	0.801 0	8.19
8 400	0.80	0.84	0.82	0.83	0.82	0.82	0.82	0.79	0.82	0.83	0.806 9	18.76
9 600	0.83	0.81	0.81	0.82	0.81	0.82	0.81	0.82	0.83	0.82	0.819 7	6.45
10 800	0.81	0.81	0.82	0.80	0.83	0.81	0.82	0.82	0.82	0.82	0.812 1	5.83
12 000	0.83	0.82	0.80	0.81	0.81	0.82	0.83	0.83	0.81	0.84	0.818 6	7.21

CNN: Convolutional neural network; Acc: Average accuracy; Var: Variance

2.2.2 CNN 分类器性能比较 为了验证深度学习自动提取的特征集在白内障检测和分类任务中的表现, 本研究与目前基于自定义特征集构建白内障分类器的最先进研究[7]中的方法进行了对比。为了使实验结果具有可比性, 实验过程中使用了相同的数

据集和相同的模型评估标准。结果如表 2 所示, G-filter 有助于克服局部不均匀色彩对分类模型的干扰, 利用 CNN 模型自动提取的特征比现有的预定义特征能提供更好的白内障特征表示。本研究 CNN 自动分类器的平均准确率为 0.818 6。

表 2 不同分类器表征白内障的平均准确率

Tab 2 Average accuracy of characterizing cataract by different classifiers

Method	Image color	Non-cataract	Mild cataract	Moderate cataract	Severe cataract	Average
Yang, et al[7]	G-channel	0.910 0	0.750 1	0.788 0	0.735 0	0.795 8
CNN	RGB	0.902 3	0.723 0	0.765 5	0.756 6	0.799 5
CNN	G-channel	0.941 2	0.752 1	0.799 4	0.781 8	0.818 6

CNN: Convolutional neural network

2.3 可视化分析 CNN 如何表征白内障的严重程度 本研究进一步分析了输入图像的哪些像素集对 CNN 预测的贡献最大。该方法在很大程度上取

决于探索直觉图, 即当深入网络隐藏层时, 特征图包含的不相关信息越来越少。本研究随机采样每层中最高激活的特征图, 并使用 DN 方法激活特定的

神经元, 然后将这些信号重新定位到原始图像中的相应像素, 研究每个卷积层中滤波器与输入图像之间的响应, 从而提高对 CNN 结构每层中学习到的特征的理解。

图 4 显示了输入图像经过第 1 层 (conv1 层)

卷积层计算后输出的特征图和其对应的反卷积结果。本研究随机选择了 conv1 层的最高激活的 3 个特征图, conv1 层的卷积计算过程类似于 1 组不同方向的 Gabor 滤波器组^[15], 从不同方向提取输入图像的边缘和线段等低级特征。

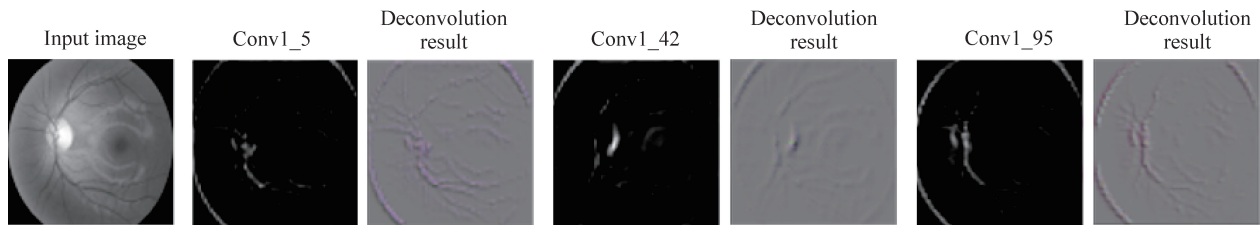


图 4 Conv1 层的最高激活特征图的反卷积结果示例

Fig 4 Deconvolution examples of the highest activation feature maps in conv1

Conv1: Convolution layer 1

图 5A 为 4 个类别的眼底图像样本, 从上至下依次为正常、轻度、中度和重度白内障样本, 图 5B~5D 为针对该样本在 conv2 层的最高激活的特征图的反卷积结果。可见样本图像中有明显像素变化的位置处的神经元被激活, 表现为在不同方向上的梯度变化的简单检测。因此, conv2 层的滤波器可以看作是一组梯度运算器, 用于提取眼底图像中的轮廓信息。但反卷积结果存在很多噪点, 即神经元在眼底图像整个区域被高度激活, 也说明该层特征是较低级特征, 不具有抽象意义。

Conv3 层可以观察到比 conv2 层更复杂的不变量, 如滤除了 conv2 层中大部分噪声信息, 且在血管、视神经盘区域和 4 个椭圆形边缘轮廓的神经元更易被激活。被激活的原因可能是滤波器学习了某种波边缘结构, 导致神经元激活与血管轮廓和视神经盘形状及眼底图像形状等基本特征相关联。见图 6。

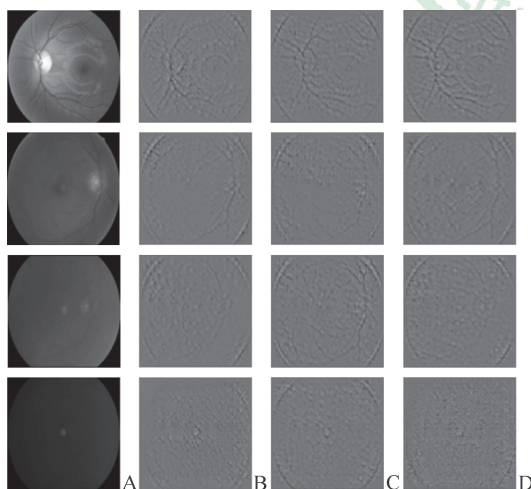


图 5 Conv2 层的最高激活特征图的反卷积结果示例

Fig 5 Deconvolution examples of the highest activation feature maps in conv2

From top to bottom, they are non-cataract, mild cataract, moderate cataract, and severe cataract samples, respectively. A: Input images; B-D: Examples of deconvolution results of the highly active feature maps from conv2. Conv2: Convolution layer 2

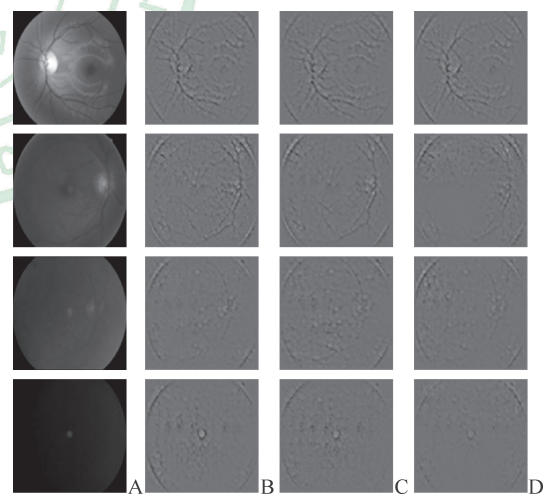


图 6 Conv3 层的最高激活特征图的反卷积结果示例

Fig 6 Deconvolution examples of the highest activation feature maps in conv3

From top to bottom, they are non-cataract, mild cataract, moderate cataract, and severe cataract samples, respectively. A: Input images; B-D: Examples of deconvolution results of the highly active feature maps from conv3. Conv3: Convolution layer 3

Conv4 层可观察到眼底图像结构的中级语义部分抽象。经 3 层卷积和最大池计算后, 特征图趋向

于保留更活跃的单元, 如视神经盘、大血管和完整眼底图像的四弧边, 这可以归因于 CNN 在特定方向上提取的曲率特征。然而, 血管信息的一些细节在这一过程中被丢弃。见图 7。

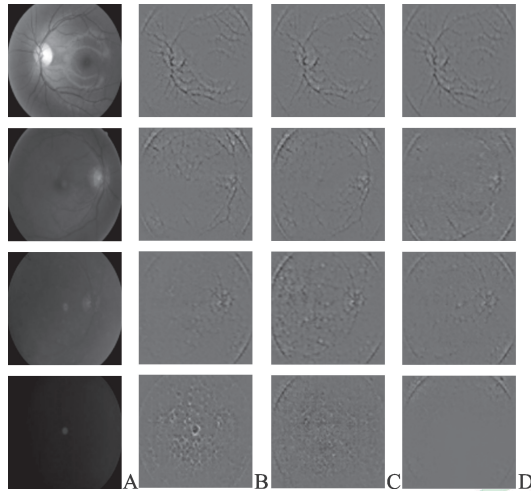


图 7 Conv4 层的最高激活特征图的反卷积结果示例
Fig 7 Deconvolution examples of the highest activation feature maps in conv4

From top to bottom, they are non-cataract, mild cataract, moderate cataract, and severe cataract samples, respectively. A: Input images; B-D: Examples of deconvolution results of the highly active feature maps from conv4. Conv4: Convolution layer 4

Conv5 层学习的特征与前几层相比变化显著, 且更具类别特异性, 可以看出对于正常样本, 由 CNN 自动提取的特征可以清晰显示血管和视神经盘的信息。轻度白内障图像中的血管细节比正常样本少; 中度白内障图像中, 高权重提取的特征主要落在视神经盘信息上, 部分大血管已被定位; 严重白内障图像中, 没有明确定位视神经盘信息的特征。见图 8。

除了用于分类的血管和视神经盘信息外, 本研究发现 CNN 模型自动提取的特征有一部分无用特征, 也可以称为干涉特征——围绕完整眼底的 4 个弧边缘特征, 且白内障严重程度越高, 干扰特征的权重越大 (电子形式的图片更容易观察)。

观察第 1 层到第 5 层的特征可视化结果, 可见从低层到中间层再到高层抽象的特征的分层变换。如梯度变化到边缘, 然后到边缘状发散结构的组合, 最后到血管和视神经盘信息的高级抽象。高层特征建立在中间层特征上, 而中级特征建立在低层特征上^[15], 彼此相关, 形成眼底图像的高级鲁棒特征表示。

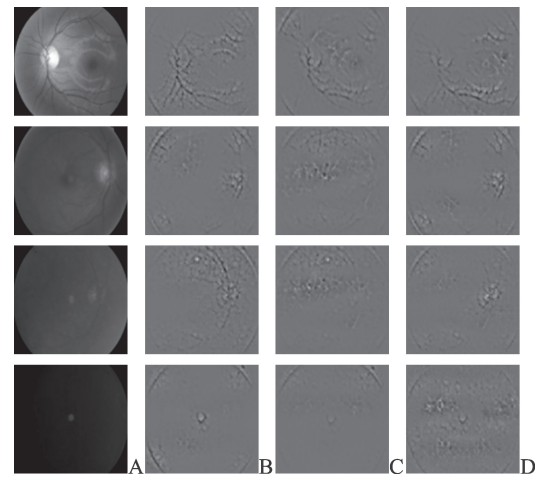


图 8 Conv5 层的最高激活特征图的反卷积结果示例
Fig 8 Deconvolution examples of the highest activation feature maps in conv5

From top to bottom, they are non-cataract, mild cataract, moderate cataract, and severe cataract samples, respectively. A: Input images; B-D: Examples of deconvolution results of the highly active feature maps from conv5. Conv5: Convolution layer 5

但在特征可视化过程中, CNN 自动提取的特征仍存在以下弊端: (1) 血管信息的丢失, CNN 的每层卷积层和池化层都着重提取轮廓特征, 经多层的卷积和池化计算后, 原始图像中的 2、3 级血管和毛细血管的细节信息容易丢失; (2) 无用特征的干扰, 原图像中眼底图像的 4 个椭圆形边缘轮廓的干扰特征占有很大权重, 从而影响整个白内障分类器的性能。

3 讨论

本文主要贡献表现在以下两个方面: (1) 使用深度 CNN 自动提取眼底图像特征, 并验证了利用 CNN 模型自动提取的特征集比现有的预定义特征集能提供更好的白内障特征表示; (2) 利用 DN 方法量化分析了 CNN 如何表征眼底图像数据, 进一步研究了输入图像的哪些像素集对 CNN 的预测贡献最大, 并观察了特征从低级到高级抽象的分层变换过程。本研究提出的基于深度 CNN 的白内障自动分类器在四分类任务中取得了 0.818 6 的平均准确率, 在相同的数据集中优于现有方法, 对于降低白内障筛查的工作成本和和健康投资较低区域普及白内障筛查工作具有重要的实际意义, 且该方法对于检测青光眼等其他眼病也可能具有潜在的应用前景。

CNN 模型自动提取眼底图像特征集在取得较好表现的同时, 也存在逐层细节血管信息丢失和干扰特征的问题, 这也为下一步继续完善白内障分类器提供了工作思路。

[参 考 文 献]

- [1] KHAIRALLAH M, KAHLOUN R, BOURNE R, LIMBURG H, FLAXMAN S, JONAS J, et al. Number of people blind or visually impaired by cataract worldwide and in world regions, 1990 to 2010[J]. *Invest Ophthalmol Vis Sci*, 2015, 56: 6762-6769.
- [2] MÜLLER-BREITENKAMP U, OHRLOFF C, HOCKWIN O. [Aspects of physiology, pathology and epidemiology of cataract][J]. *Ophthalmologet*, 1992, 89: 257-267.
- [3] HE K, ZHANG X, REN S, SUN J. Deep residual learning for image recognition[Z/OL]. arXiv: 1512.03385, 2015. <http://cn.arxiv.org/pdf/1512.03385v1>.
- [4] YANG M, YANG J J, ZHANG Q, NIU Y, LI J. Classification of retinal image for automatic cataract detection[C]//2013 IEEE 15th International Conference on E-Health Networking, Application & Services. IEEE, 2013: 674-679.
- [5] ZHENG J, GUO L, PENG L, LI J, YANG J, LIANG Q. Fundus image based cataract classification[C]//2014 IEEE International Conference on Imaging Systems and Techniques. IEEE, 2014: 90-94.
- [6] GUO L, YANG J J, PENG L, LI J, LIANG Q. A computer-aided healthcare system for cataract classification and grading based on fundus image analysis[J]. *Comput Indust*, 2015, 69: 72-80.
- [7] YANG J J, LI J, SHEN R, ZENG Y, HE J, BI J, et al. Exploiting ensemble learning for automatic cataract detection and grading[J]. *Comput Methods Programs Biomed*, 2016, 124: 45-57.
- [8] CAIXINHA M, VELTE E, SANTOS M, SANTOS J B. New approach for objective cataract classification based on ultrasound techniques using multiclass SVM classifiers[C]//IEEE International Ultrasonics Symposium. IEEE, 2014: 2402-2405.
- [9] FAN W, SHEN R, ZHANG Q, YANG J J, LI J. Principal component analysis based cataract grading and classification[C]//2015 17th International Conference on E-Health Networking, Application & Services. IEEE, 2016: 459-462.
- [10] GAO X, LIN S, WONG T Y. Automatic feature learning to grade nuclear cataracts based on deep learning[C]//Asian Conference on Computer Vision. New York: Springer International Publishing, 2014: 632-642.
- [11] ZHANG L, LI J, ZHANG I, HAN H, LIU B, YANG J, et al. Automatic cataract detection and grading using deep convolutional neural network[C]//2017 IEEE 14th International Conference on Networking, Sensing and Control. IEEE, 2017: 60-65.
- [12] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks[C]//International Conference on Neural Information Processing Systems. Curran Associates Inc., 2012: 1097-1105.
- [13] DONAHUE J, JIA Y, VINYALS O, HOFFMAN J, ZHANG N, TZENG E, et al. DeCAF: a deep convolutional activation feature for generic visual recognition[Z/OL]. arXiv: 1310.1531, 2013. <http://cn.arxiv.org/pdf/1310.1531v1>.
- [14] DONG C, CHEN C L, HE K, TANG X. Learning a deep convolutional network for image super-resolution[M]//Computer Vision-ECCV 2014. New York: Springer International Publishing, 2014: 184-199.
- [15] LEE S H, CHAN C S, MAYO S J, REMAGNINO P. How deep learning extracts and learns leaf features for plant classification[J/OL]. *Pattern Recogn*, 2017, 71: 1-13. doi: 10.1016/j.patcog.2017.05.015.
- [16] ZEILER M D, FERGUS R. Visualizing and understanding convolutional networks[J]. *Computer Vision-ECCV 2014*, 2013, 8689: 818-833.
- [17] JIA Y, SHELHAMER E, DONAHUE J, KARAYEV S, LONG J, GIRSHICK R, et al. Caffe: convolutional architecture for fast feature embedding[Z/OL]. arXiv: 1408.5093, 2014. <http://cn.arxiv.org/pdf/1408.5093v1>.

[本文编辑] 曾奇峰