

基于深度反卷积网络的虹膜定位方法研究

徐 霄^{1,2} 陈 阳^{1,3} 张飞云^{1,4} 乔 宇¹

¹(中国科学院深圳先进技术研究院 深圳 518055)

²(中国科学院大学深圳先进技术学院 深圳 518055)

³(中国科学技术大学 合肥 230026)

⁴(江苏大学 镇江 212013)

摘 要 精确而快速地进行虹膜定位是有效地进行虹膜识别的重要前提。传统的虹膜定位方法有 Daugman 定位法、Hough 变换定位法等,但其对睫毛比较浓密、虹膜被遮挡较多等情况下的图片处理效果不是很好。文章在总结前人工作的基础上,实现了一套基于深度学习的虹膜定位系统。该系统利用深度学习的方法,根据虹膜图像区域的特点,对图像进行像素级分类。根据像素分类结果,可以很好地标识虹膜区域和非虹膜区域,达到定位识别虹膜区域的目的,并在中国科学院自动化所公布的虹膜数据集 CASIA-IrisV3-Interval 上验证了文章工作的有效性,像素分类精度达到约 98.4%,达到了较高的鲁棒性。

关键词 虹膜定位;深度学习;像素级分类

中图分类号 TG 156 **文献标志码** A

Iris Location Based on Deep Deconvolution Network

XU Xiao^{1,2} CHEN Yang^{1,3} ZHANG Feiyun^{1,4} QIAO Yu¹

¹(Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, Shenzhen 518055, China)

²(Shenzhen Institutes of Advanced Technology, University of Chinese Academy of Sciences, Shenzhen 518055, China)

³(University of Science and Technology of China, Hefei 230026, China)

⁴(Jiangsu University, Zhenjiang 212013, China)

Abstract Obtaining the iris localization precisely and fast is the prelude of effective iris recognition. Traditional iris localization methods, including Daugman localization and Hough transformation localization, are weak in processing images with thick eyelashes and severely shlder. In this paper, combined with previous work of other researchers, deep learning method was employed to classify iris images based on the characteristics of iris region. We carried our experiment on the CASIA-IrisV3-Interval dataset, to verify the effectiveness of our work. The accuracy of pixel-wise classification is around 98.4%, with a higher robustness.

Keywords iris location; deep learning; pixel-wise classification

收稿日期: 2015-08-28 修回日期: 2015-09-14

基金项目: 深圳市科技计划(KQCX2015033117354153, CYJ20130402113127496); 广东省科技计划(201001D0104648280, 2014B050505017, 2015B010129013)

作者简介: 徐霄, 硕士研究生, 研究方向为计算机视觉、图像处理; 陈阳, 硕士研究生, 研究方向为计算机视觉、图像处理; 张飞云, 硕士研究生, 研究方向为计算机视觉、图像处理; 乔宇(通讯作者), 研究员, 研究方向为计算机视觉、图像处理, E-mail: yu.qiao@siat.ac.cn.

1 引言

1.1 虹膜识别简介

虹膜识别是一种新颖的基于生物特征的身份鉴别方法,也是所有生物识别技术中,应用最方便和最精确的方法,在信息安全领域有着重要的应用价值^[1,2]。

虹膜定位是虹膜识别的基础,虹膜定位的时间占整个虹膜识别过程的一半以上^[3],是影响整个系统性能的主要因素。如何迅速准确地定位出虹膜一直是虹膜图像处理的一个研究热点和难点。一般来讲,基于虹膜的特征识别主要由四部分组成^[4]:

(1) 虹膜图像采集

不同人种的虹膜颜色有很大的差别。白种人的虹膜颜色浅,纹理显著;而黄种人的虹膜则多为深褐色,纹理非常不明显,必须采用专门的虹膜图像装置才可拍摄出纹理丰富的虹膜图像。采集时在遮蔽杂光的前提下,用户需站在离采集设备 10~50 cm 的范围内,采用稳定的近红外光源,调整光源的入射角度,以获得高质量的 8 位灰度图象。

(2) 虹膜图像的预处理

虹膜图像的预处理包括对虹膜图像的定位,归一化和增强等步骤。

虹膜定位:是整个虹膜识别过程中最重要的环节,准确地确定虹膜的内边界和外边界,保证每次进行特征提取的虹膜区域不存在较大偏差,定位的速度和准确性决定了整个虹膜系统是否实用可行。

虹膜归一化:一般虹膜识别系统都要对虹膜进行归一化处理,其目的是将每幅原始图像调整到相同的尺寸和对应的位置,从而消除平移、缩放和旋转对虹膜识别的影响。

虹膜图像增强:采集设备自身的原因使虹膜图像光照不均匀,一般通过直方图均衡化处理。

采集过程中还存在各种噪声的干扰,通常通过同态滤波去除噪声干扰。图像增强的目的是减小光照不均、各种噪声等因素对虹膜识别系统的识别性能的影响。

(3) 虹膜特征提取及编码

采用适当的算法提取虹膜图像细节特征并进行编码,最后形成特征模板模型。从特征提取角度,现有方法可分为以下三类:基于相位分析的方法,如 Daugman 的相位编码方法^[5];基于过零检测的方法,如 Boles 的一维小波零交叉编码方法^[6],以及基于纹理分析的方法,如 Wildes 的拉普拉斯金字塔算法^[7]。

(4) 匹配与识别

将采集图像的特征与数据库中的虹膜图像特征模板进行比对,判断两个虹膜是否属于同一类。随着虹膜识别技术研究和应用的进一步发展,虹膜识别系统的自动化程度越来越高,神经网络算法、模糊识别算法也逐步应用到虹膜识别之中。

传统的基于非监督学习的虹膜定位方法,其定位质量受到图像质量与外界因素的影响较大,且计算量和占用内存均较大,提取的参数受参数空间的量化间隔制约,导致虹膜定位效果并不好。基于此,受深度神经网络的启发,本文提出一种基于深度反卷积网络的基于监督学习的虹膜定位方法,旨在有限的时间和存储开销下,既能解决非监督学习方法的不足,又能准确高效地定位出虹膜区域。

1.2 几种常见的虹膜定位算法

目前,具有代表性的 4 种虹膜定位方法分别为: Daugman 定位法、Hough 变换定位法^[9]、几何特征定位法^[10]和主动轮廓线定位法^[11]。

1.2.1 Daugman 定位法

该方法由英国剑桥大学的 Daugman 教授提出。在拍摄获取的虹膜图像中,灰度分布存在着一定的差异:一般而言虹膜比瞳孔亮,巩膜又

比虹膜亮。根据虹膜的形状类似圆环形的状况, Daugman 利用圆形检测匹配器的方法分割虹膜。Daugman 定位法的优点是定位比较准确, 但是该算法较强地依赖图像亮度的梯度, 因此只有在被处理的图像照度比较均匀而且没有大面积的亮度反常区域时有效。

1.2.2 Hough 变换定位法

该方法具体过程是: 先通过边缘检测方法得到瞳孔和虹膜以及虹膜和巩膜的边缘, 再利用虹膜是圆环状的特点, 应用 Hough 变换得到圆环的内外半径, 从而分割出虹膜。由于拍摄的图片不可能消除眼皮和睫毛, 所以用 Hough 变换法去分割虹膜往往会包含眼皮和睫毛。一个改进方法是以抛物线来拟合眼睑, 这样就可以消去眼皮、睫毛的影响, 保留其他虹膜信息。

Hough 变换的优点主要在于: 它对于图像中的噪声点不敏感, 利用它得到的效果可以有效地消除噪声的影响。缺点是计算量大、占用内存大、提取的参数受参数空间的量化间隔制约。

1.2.3 几何特征定位法

几何特征定位法是利用虹膜边缘图像的几何特征以及圆相交弦的性质进行虹膜定位的方法。具体过程是: 首先根据虹膜图像的整体灰度分布信息, 利用边缘检测算子提取虹膜的内边缘; 然后利用圆的相交弦性质提取虹膜的中心; 最后利用 Hough 变换提取虹膜外半径。

几何特征法综合了 Hough 变换定位法, 提高了虹膜的定位速度。但是定位速度受到图像质量的影响很大, 当图像质量不好时, 算法性能急剧下降。

1.2.4 主动轮廓线定位法

主动轮廓线是一项复杂的轮廓提取及图像解释技术, 它通过不断地极小化自身的能量函数来达到物体的边界。定位的具体过程为: 先用灰度检测的方法检测出瞳孔内一点作为瞳孔的伪圆心; 然后以该点为中心, 在其周围取几个点作为

初始的蛇形模型, 按照蛇形模型的运行机制不断进化, 找到虹膜的内边界, 接着计算进化后的蛇形模型的形心和控制点, 定位出虹膜内边界位置; 最后按照 Daugman 方法定位外边界。

主动轮廓线定位法的优势是: 不用边缘检测和二值化, 对瞳孔初始的圆心要求不高, 算法的鲁棒性更强。其缺点是定位质量受到图像质量与外界因素的影响较大。

2 基于深度学习的虹膜定位识别

2.1 卷积神经网络

卷积神经网络是一种可训练的多级前馈神经网络^[12], 它的人工神经元可以响应一部分覆盖范围内的周围单元^[13], 对于大型图像处理有出色表现。

卷积神经网络是人工神经网络的一种, 是当下图像分析、识别、检测等领域内的研究热点。其网络结构与生物神经网络非常相似, 降低了网络模型的复杂度, 减少了权值的数目。其优点在于: 当网络的输入是多维图像时, 图像可以直接作为网络的输入, 避免了传统识别算法中复杂的特征提取和数据重建过程。同时, 网络结构对平移、比例缩放、倾斜或者其他形式的变形具有高度不变性。

典型的卷积神经网络^[12,14]通常由 4 种类型的层组成: 卷积层、池化层、非线性层和全连接层。每一层的输入和输出叫做特征图。

(1) 卷积层将输入特征图和三维滤波器做卷积处理以产生输出特征图。每个滤波器在输入特征图的所有位置抽取相同类型的局部特征。一般地, 卷积层的步长为 1。但是在最近的一些卷积神经网络模型中, 大于 1 的步长也同样在卷积层被使用, 用以下采样输出特征图, 降低计算维度。

位于位置 (i, j) 处的数据向量 x_{ij} 经过卷积层以后的输出 x_{ij} 的计算公式为:

$$y_{ij} = f_{ks} \left(\left\{ X_{s_i + \delta_i, s_j + \delta_j} \right\} \right) \quad 0 \leq \delta_i, \delta_j \leq k \quad (1)$$

其中, k 表示卷积核大小; s 为步长或降采样因子, f_{ks} 决定了网络层的类型: 卷积或平均池化的矩阵乘操作、激活函数的元素非线性操作等。

(2) 池化层的作用是降低特征图的分辨率, 使得输出特征图对输入偏移和畸变不敏感。最大池化和平均池化是最常用的池化方法。通常情况下, 池化层的步长和它的卷积核的大小相等。但是在最近的一些卷积神经网络模型中, 和卷积核大小不相等的步长也被使用。

(3) 非线性层是一个对特征图的每一项进行逐点非线性函数计算的过程, 这一过程也被叫做非线性激活, 使用的非线性函数叫激活函数。常用的激活函数有 sigmoid 函数、双曲正切函数、修正线性单元 (ReLU) 等。本文采用 ReLU 作为激活函数, 其公式如下:

$$g(x) = \max(0, x) \quad (2)$$

其好处在于训练非常快速, 节省了大量时间。

(4) 全连接层是把输入转换成一个一维的向量, 然后对这个一维向量进行点乘, 得到最后的输出概率的过程。

上述特征使得卷积神经网络能够利用输入数据的二维结构, 产生对应的输出。与其他深度学习结构相比, 卷积神经网络在图像和语音识别方面能够给出更优的结果。这一模型也可以使用反

向传播算法进行训练。相比其他深度前馈神经网络, 卷积神经网络需要估计的参数更少^[15], 使之成为一种颇具吸引力的深度学习结构。

2.2 网络结构

在卷积神经网络的基础上, 来构建本论文中的网络。图 1 描述了整个深度反卷积神经网络的详细配置, 采用传统的 caffe^[16] 框架, 网络由两部分组成——卷积和反卷积网络。我们采用自己定义的网络结构, 网络总共有 6 层, 3 层卷积层和 3 层反卷积层, 每一层后面紧跟着一个归一化线性层, 整个网络没有池化操作和全连接层, 这也是我们的网络有别于其他网络的地方。池化层往往在卷积层后面, 通过池化来降低卷积层输出特征向量的维度, 同时改善结果, 不易出现过拟合。本文没有采用池化操作, 仍然达到了理想的效果。众所周知, 神经网络极大部分的参数计算过程都在全连接层, 随着网络结构的复杂化, 大量的数据计算将会影响网络性能。同时, 如果使用全连接层, 那么网络必须使用固定大小的输入, 这是非常不方便以及不合理的。2015 年 3 月以来, 全卷积网络越来越流行。这种网络的好处就是不包含全连接层, 网络可以使用任意大小的输入, 受此方法的影响, 我们去掉了全连接层, 使得网络更加灵活。

在反卷积层的最后一层后面, 我们接了一个 Softmax 层, 用以输出二分类概率图。网络的每

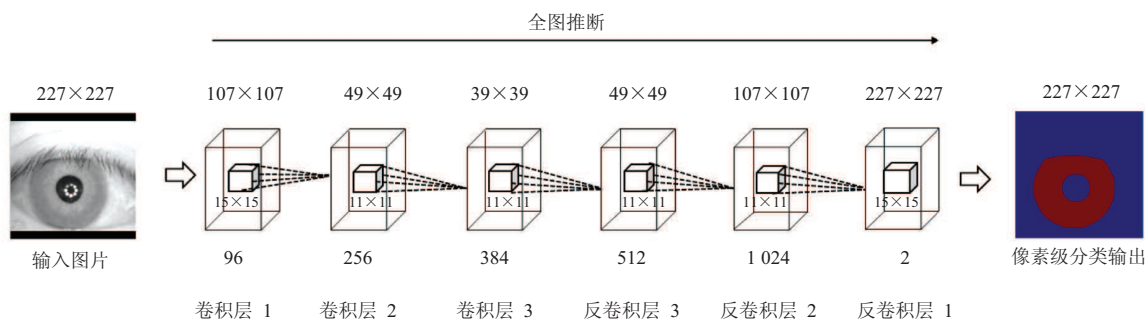


图 1 深度反卷积网络结构图

Fig. 1 Structure of Deep Deconvolution Network

一层是一个大小为 $w \times h \times d$ 的三维矩阵, 其中 h 和 w 代表图像的高度和宽度, d 是特征或者信道维数。利用上述三维矩阵可得出网络每一层的神经元个数。网络的输入为 $51\ 529 (227 \times 227 \times 1)$, 以下计算方式相同) 维, 网络剩下层的神经元个数依次为 $1\ 099\ 104$ 、 $614\ 656$ 、 $584\ 064$ 、 $1\ 229\ 312$ 、 $11\ 723\ 776$ 、 $103\ 058$, 卷积核的大小依次为 15×15 、 11×11 、 11×11 、 11×11 、 11×11 、 15×15 。从卷积核大小的角度也可以看出, 反卷积是卷积过程的镜像。网络最后一层的输出个数为 2, 分别对应虹膜区域-非虹膜区域二分类的两个概率图。网络每一层的输出计算公式为:

$$h_o = \frac{h_i + 2 \times \text{pad} - \text{kernel_size}}{\text{stride}} + 1 \quad (3)$$

其中, h_o 为输出高度; h_i 为输入高度; pad 为填充大小; kernel_size 为卷积核大小; stride 为步长。输出宽度 w_o 采用同样的方法计算。

卷积网络对应于特征提取, 将输入图像转换为多维特征表达。而反卷积网络则是一个形状产生器, 将从卷积网络中提取的特征转换为最终的像素级分类结果。最后的网络输出是一个和输入图像等大的概率分布图, 指示每个像素属于虹膜的概率。反卷积网络是卷积网络的一个镜像版本, 包含 3 层反卷积层和归一化线性层。卷积网络减小了前向传播中激活单元的大小, 反卷积网络增大了激活单元的大小。下面将详细描述如何利用反卷积网络进行本实验。

2.3 反卷积网络像素级分类

2.3.1 反卷积

反卷积是指通过测量输出和已知输入重构未知输入的过程, 广泛应用于语音识别、图像恢复等领域。从数学角度来讲, 反卷积是卷积操作的逆运算。不同于卷积层, 用一个滤波窗口连接多个输入激活单元, 以产生一个单一的激活值, 而反卷积层的一个单一的输入激活单元对应多个输出, 如图 2 所示。

由图 2 可知, 反卷积层的输出是一个扩大且密集的激活图。我们对增大的激活图的边界进行裁剪, 以保证输出图的大小与前一层的输入大小保持一致。

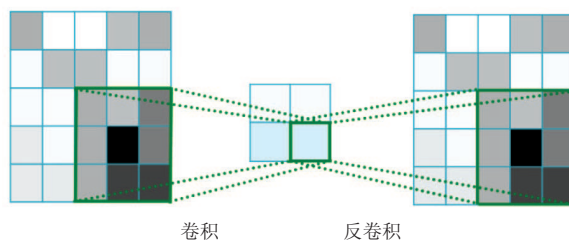


图 2 反卷积运算图解

Fig. 2 Diagram of deconvolution operation

反卷积层学习到的滤波器是输入目标形状重构的基础。因此, 与卷积网络类似, 反卷积层的层次结构用于捕获形状细节的不同级别。较低层的滤波器倾向于捕捉物体的整体形状, 而类相关的细节则在高层编码。通过这种方式, 网络直接考虑与类相关的形状信息用于像素级分类, 这是其他仅基于卷积层的方法经常忽略的^[17,18]。

2.3.2 反卷积网络分析

在该算法中, 反卷积网络是精确到像素级分类的关键部分。我们也尝试过不使用反卷积操作, 而直接使用传统的卷积神经网络, 如 AlexNet^[19] 网络等以及我们自己设计的一些卷积神经网络: 将一幅 227×227 的输入图像, 通过卷积、池化、全连接等操作, 输出一张 20×20 左右的特征图, 然而实验效果并不好。与传统方法相比, 我们的方法使用反卷积网络生成虹膜区域像素级分类掩膜, 通过连续的卷积、归一化线性限制和反卷积操作, 生成一个和输入图像等大的、密集的像素级类概率图。由粗到细的物体结构重建过程由反卷积层的传播进行; 低层主要获取物体整体的粗略结构(如位置、形状和区域), 而更复杂的模式主要在高层被发现。因此, 它能够在更精细的分辨率下有效重构物体的详细结构。另一方面, 在反卷积层学习到的滤波器能够

捕获与类相关的特有形状信息。通过反卷积，与目标类密切相关的激活单元的作用被增强，而来自其他区域的噪声激活单元则被有效地抑制。

3 实验

3.1 数据集

本论文中，我们采用中国科学院自动化研究所公布的虹膜数据集 CASIA-IrisV3-Interval，其包含了 249 位受试者，在不同实验条件下采集的图片总数为 2 639 张。为了节省存储空间，提高数据传输速度，所有的虹膜图像均被压缩成 8 位的 JPEG 文件，在靠近红外光的条件下采集。CASIA-IrisV3 包括 3 个子集：CASIA-IrisV3-Interval、CASIA-IrisV3-Lamp、CASIA-IrisV3-Twins。3 个子集各具特色，包含了反映时间跨度变化、光照变化以及双胞胎特殊人群的虹膜图像。本文仅仅使用第一个子数据集，一些统计数据 and 特征如表 1 所示。除了有少量外国人外，本数据集中的受试者几乎都是中国人。

数据采集过程示例如图 3 所示。

图 4 展示了 CASIA-IrisVe-Interval 数据集中的部分示例图片，从图中可以看到如下特点：

- (1) 虹膜图片相对比较简单，虹膜位置总体比较固定，虹膜细节特征比较清晰；
- (2) 虹膜图片光照强度不同；
- (3) 上下眼睑对虹膜的遮挡程度不同：有的

毫无遮挡，有的遮挡比较严重；

(4) 睫毛对虹膜区域有不同程度的遮挡，有些图片中的遮挡甚至很严重；

(5) 有些图片中虹膜和巩膜区域的区分度并不高。

表 1 CASIA-IrisV3-Interval 数据统计

Table 1 CASIA-IrisV3-Interval data statistics

数据集属性	CASIA-IrisV3-Interval
采集装置	CASIA-close-up iris camera
采集环境	室内
采集时间跨度	大部分图像分多次采集，每次采集的时间间隔超过 1 个月
人数	249
类别数	395
图像数	2 639
分辨率	320×280
特点	高质量的虹膜图像，虹膜细节特征清晰可见

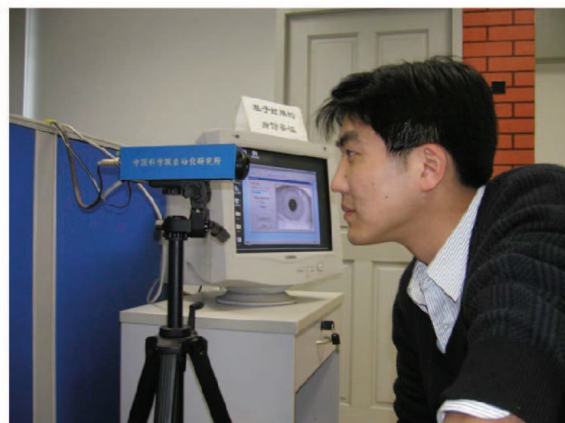


图 3 CASIA 研发的虹膜摄像机

Fig. 3 The iris camera developed by CASIA

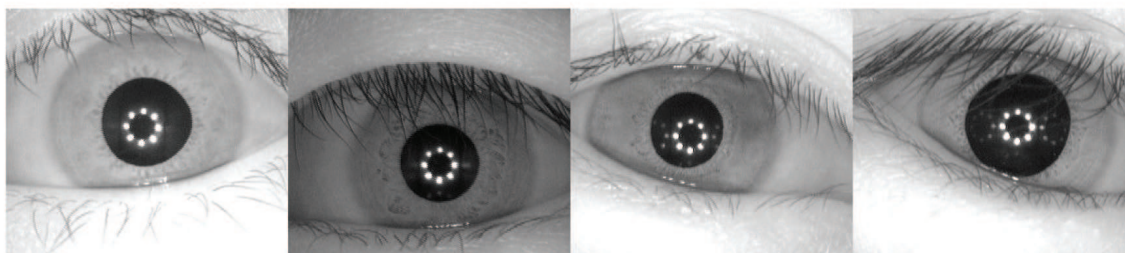


图 4 CASIA-IrisVe-Interval 数据集示例图片

Fig. 4 Sample images of the CASIA-IrisVe-Interval datasets

针对上述虹膜图片相对比较简单的特点, 我们构造 3 层的反卷积网络结构理论上是合理的。眼睑和睫毛遮挡的情况一直是各种虹膜定位算法的难点。

3.2 数据标定

对虹膜区域进行像素级分类的问题, 本质上是一个二类分类问题, 即判断一个像素点是属于虹膜区域还是非虹膜区域。因此, 我们以 Michael Boyd^[20]的虹膜识别软件为基础, 开发了一款符合我们任务的标定工具, 标定时虹膜内外边缘采用圆拟合, 上下眼睑遮挡部分曲线采用贝塞尔曲线拟合。标定完成后, 软件自动将虹膜区域标注为 1, 非虹膜区域标注为 0, 生成一个二值化掩膜作为训练数据的标签。生成训练数据时, 首先将 8 位灰度图像转换为 double 格式, 然后和二值化掩膜标签一起存为 mat 文件, 最后用 python 脚本将其转换为 caffe 可接受的 hdf5 数据格式。标定过程示意图如图 5 所示。

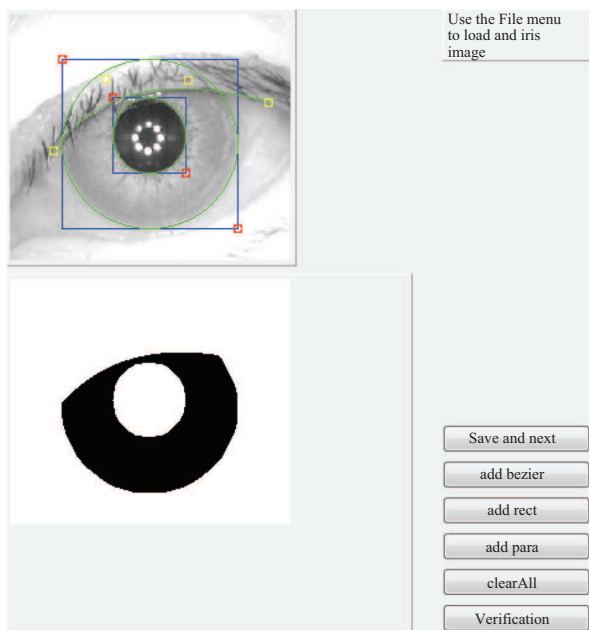


图 5 标定过程示意图

Fig. 5 Labeling process diagram

3.3 数据扩充

通常情况下, 训练一个神经网络, 如果没有

充分的训练数据, 模型将极大可能陷入过拟合, 而增加训练集的数量则会让一个过拟合的网络取得更好的训练结果。数据扩充是人为通过一些手段(旋转、平移、缩放、添加噪声等)增加训练数据的个数, 经过数据扩充后, 使得训练样本更加丰富, 可以更好地学习特征。显然, 这要比手工收集更多的样本经济许多。所以, 数据扩充是深度学习过程中必不可少的一项操作。在很多经典的深度学习论文中, 均采用了数据扩充, 这样可以使得网络具有更强的泛化能力和鲁棒性。需要说明的是, 数据扩充并不表示重复样本, 重复样本是放入多张完全相同的图片进行训练, 而数据扩充是改变图片的某个条件, 以考虑同一图片可能在不同情况下的状态, 扩充后的图片和原图片的类别是一致的, 但是并不完全和原图片相同。

由于原始的 CASIA-IrisV3-Interval 数据集只有 2 639 张虹膜图像, 对于神经网络来说, 这些数据是远远不够的。因此, 人为地进行了一些数据扩充, 包括裁剪、缩放、旋转等操作, 具体操作如下:

(1) 裁剪

首先将图片缩放到 256×256 , 然后从图片的左上、左下、右上、右下和中间五个位置裁剪出 227×227 的大小, 示例如图 6(a) 所示。

(2) 缩放

将图片分别缩放到 128×128 、 160×160 、 192×192 、 224×224 大小, 为了保证缩放前后长宽比不变, 将缩放后的小图分别放在左上、左下、右上、右下四个位置, 然后将其填充到 227×227 大小, 示例如图 6(b) 所示。

(3) 旋转

原始数据集中包含少量稍微倾斜的虹膜图片, 但比例很少, 为了使这部分训练数据仍然能学到比较好的特征, 人为地挑选出一些图片, 将其进行旋转, 旋转角度分别为 5° 、 15° 、 30° 、 45° , 示例如图 6(c) 所示。

经过上述数据扩充操作后，训练图片总数达到了 38 225 个，约为原来的 15 倍。

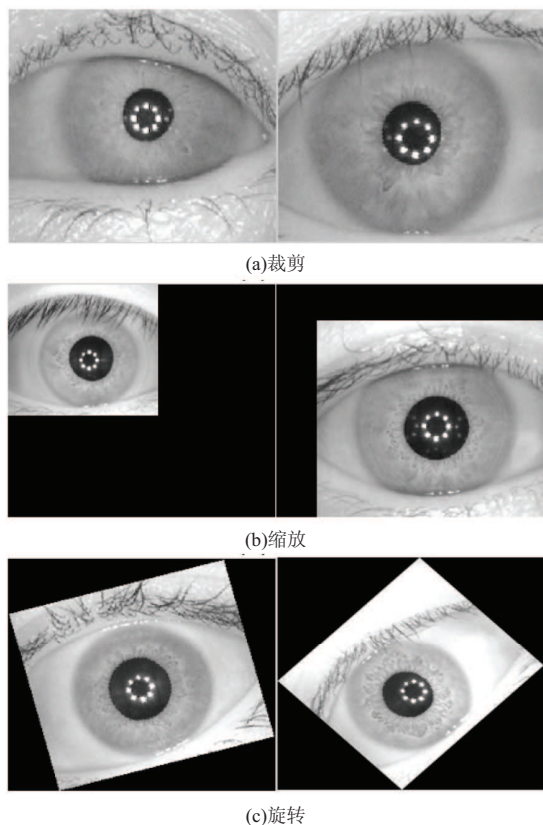


图 6 数据扩充后部分数据示例

Fig. 6 Sample images after data enhancement

4 实验结果

4.1 与 DeconvNet 比较

本文主要对 DeconvNet^[21]、我们的方法以及传统方法进行了比较。DeconvNet 主要采用 VGGNet-16^[22] 网络作为初始化卷积训练网络，但其移除了最后的分类层。其卷积网络有 13 个卷积层，卷积层之间有归一化线性操作和池化层，最后还有两个全连接层。反卷积网络部分是卷积网络的逆操作，包含一系列的反池化，归一化线性限制和反卷积操作。加上反卷积网络部分，网络层数接近 30 层，输入图片大小为 224×224 ，输出结果为和输入图片等大的概率图，每批次训

练图片个数为 10，测试图片个数为 1。本文仅用了 3 层卷积层，反卷积过程是卷积过程的镜像，网络总共 6 层，比起 DeconvNet，网络深度大大减小，训练时间和代价也随之降低。我们使用 227×227 的输入图片大小，每批次训练图片个数为 50，测试图片个数为 10。测试精度和训练误差随着迭代次数的变化情况分别如图 7 和图 8 所示。

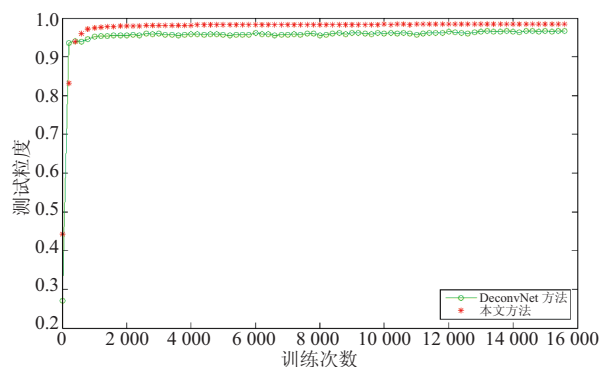


图 7 测试精度随训练次数的变化图

Fig. 7 Testing accuracy versus training times

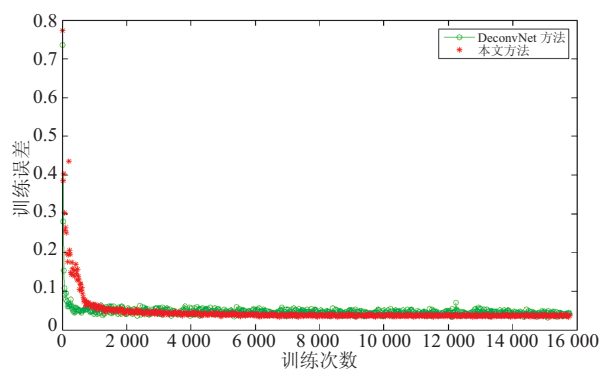


图 8 训练精度随训练次数的变化图

Fig. 8 Training accuracy versus training times

图 9 列出了一些像素分类结果图。因为是二值分类(0-1分类)，中间列表示预测为 0 的概率图，右边列表示预测为 1 的概率图，即虹膜区域。红色代表高响应值，蓝色代表低响应值。由图可以看到，右边列在虹膜区域有很高的响应值，在非虹膜区域有很低的响应值；在虹膜和巩膜的交界处，响应值出现了渐变，这也从实验结果的角度证明了某些情况下虹膜和巩膜之间确实

有不高的区分度。中间列预测出背景有很高的响应值, 与右边列虹膜区域的预测结果正好相反。

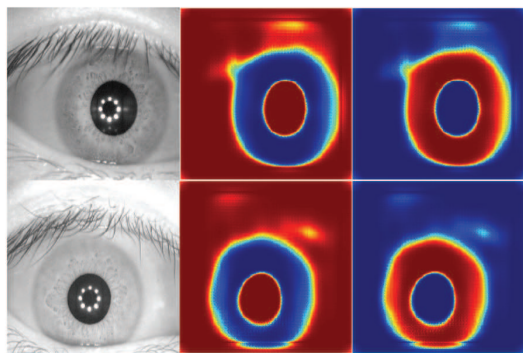


图 9 像素分类结果图

Fig. 9 Result of the pixel-wise classification

其他一些方法对比实验结果如表 2 所示。

表 2 各实验方法结果对比

Table 2 Comparison of various experimental methods

方法	是否进行数据扩充	误差	像素分类精确度
微调 DeconvNet	是	0.032	0.974 3
3 层卷积-反卷积网络	否	0.082	0.955 7
3 层卷积-反卷积网络	是	0.034	0.983 9
5 层卷积-反卷积网络	是	0.034	0.983 6

由表 2 可以看出, 进行数据扩充后, 精确度提高了近 3 个百分点, 这证明进行数据扩充, 增加样本丰富性对本实验是有显著意义的; 而增加卷积-反卷积层的层数似乎对实验结果没有明显的提高。虽然进行数据扩充后微调 DeconvNet, 也能达到和本实验差不多的精度, 但 DeconvNet 使用了 14 层卷积层, 每隔几层使用池化操作, 随后的反卷积网络是卷积网络的镜像操作, 总共有 28 层, 这种太复杂的网络结构运算量大、资源消耗大, 且精度没有提升, 对我们的数据而言显然不是必需的。相比之下, 我们的方法具有更高的精确度和鲁棒性。对虹膜图片而言, 其图片信息是比较简单的, 包含的内容也不如其他一些复杂场景的图片多, 虹膜的整体位置也比较集中, 太复杂的网络涉及的参数计算更多, 反而不太适合虹膜识别任务。其他一些基于本文网络结

构所使用的方法, 如使用不同的学习率、改变网络结构卷积核的大小、改变网络每层输出个数等, 因其结果均明显不如表 2 所述几种方法, 故实验结果不列出。

4.2 与传统方法比较

与传统方法相比, 我们的网络也具有比较明显的优势。传统方法主要采用 Daugman 和 Hough 变换相结合的方法: 首先对图像二值化, 找到大致的瞳孔位置, 为了避免睫毛的干扰, 对角度进行了限制; 然后用 Daugman 微积分算子定位虹膜内外边缘。眼睑检测使用 Hough 变换, 这里假定眼睑曲线为抛物线。睫毛检测根据睫毛区域两个特性——灰度值较低以及灰度变化大来确定。由图 10 可以看出, 传统方法对于定位虹膜的内外边缘已经比较准确, 但对于上下眼睑遮挡部分的定位仍然存在较大问题。

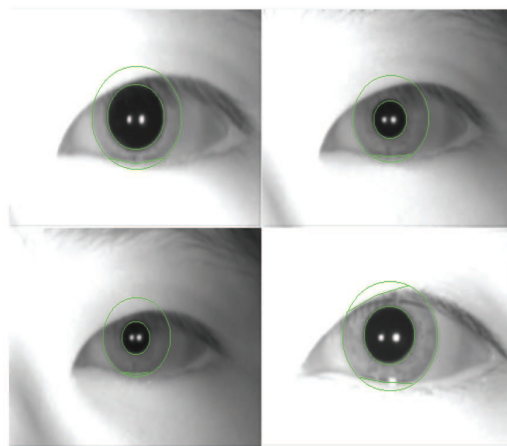


图 10 Daugman 与 Hough 变换结合的虹膜定位结果图

Fig. 10 Iris locating results of Daugman combined with Hough transformation method

4.3 限制

与很多方法相似, 对于上下眼睑遮挡虹膜、睫毛与虹膜区域有重叠的情况, 本方法的定位效果并不好(图 11)。一个可能的原因是, 在标定的时候, 没有对眼睫毛进行标定, 导致在虹膜区域的眼睫毛随着虹膜一起被标注为正样本, 而其他地方的眼睫毛被标注为负样本; 在训练的时候,

没有对睫毛进行预处理,对于部分睫毛为正、部分睫毛为负的情况,神经网络没有学习到很好的分类模型,导致最后的像素分类结果很差。对此,提出了两个可能的解决方案:

(1)标定的时候将眼睫毛标定出来,但这种方法既费时又费力,且没有一定的标准;

(2)在预处理阶段,将眼睫毛先去除,再用处理过的图片进行训练。

在以后的工作中,将对上述两种方案分别进行验证。

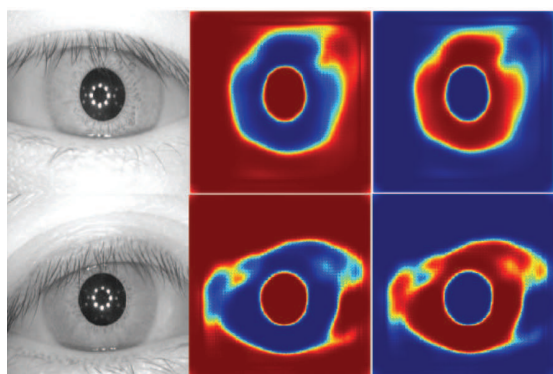


图 11 定位不好的图片示例

Fig. 11 Examples of bad locating result

5 讨论

在本工作中,解决了虹膜区域的定位识别问题,保证了识别的准确性和高效性。针对不同类型的虹膜图片,提出了一套完整的分析方案。对于所使用的方法,文章摒弃了传统的定位识别方法,采用基于深度学习的方法,使用反卷积神经网络对虹膜图片进行像素级分类,在 CASIA-IrisV3-Interval 数据库上取得了很好的效果。

当前的解决方案也有一些限制,首先是上下眼睑对虹膜区域遮挡比较严重的情况,识别方法精度会下降。因为在标注中,挡住虹膜区域的眼睫毛部分仍然被标注为正样本。一个可能的解决方法是采用一种合理的方法,将眼睫毛也单独进行标注,这样在训练样本的时候就能对睫毛和虹

膜部分有很好的区分。但由于人的眼睫毛太细而且比较多,人为地将每根眼睫毛标注出来既费时又费力。

其次,我们使用的数据集中数据有限,使用的实验方法主要是针对 CASIA-IrisV3-Interval 这个数据集进行的,对于其他复杂的数据集,如虹膜图像来自于室外场景采集、人眼不同程度地关闭状态下采集的虹膜数据,我们也将进一步尝试。

此外,虹膜识别也有一些固有的限制:

- (1)很难将图像获取设备的尺寸小型化;
- (2)镜头可能产生图像畸变而使可靠性降低;
- (3)黑眼镜极难读取;
- (4)需要较好的光源。

如何减小或避免这些限制带来的影响,也是一项长期而值得研究的课题,本文不做过多描述。

6 结论

本文在总结前人工作的基础上,实现了一套基于深度学习的虹膜定位系统:利用深度反卷积网络方法,根据虹膜图像区域的特点,对图像进行像素级分类。根据像素分类结果,可以很好地标识虹膜区域和非虹膜区域,以达到定位识别虹膜区域的目的。

文章所提方法还有待改进之处,由于时间限制,本实验对影响神经网络实验效果的其他因素,如学习率、网络初始化参数、每层网络接收域的大小、网络深度等因素没有做过多地探讨,这些均有待尝试。在将来的研究工作中,将考虑这些因素,同时,也将采用更复杂、种类更丰富的虹膜数据。

参考文献

- [1] 孙冬梅,裘正定.生物特征识别技术综述[J].电子学报,2001,29(12A):1744-1748.

- [2] 张敏贵, 周德龙, 潘泉, 等. 生物特征识别及研究现状 [J]. 生物物理学报, 2002, 18 (2): 156-161.
- [3] 王春, 叶虎年. 虹膜识别算法的研究 [J]. 贵州工业大学学报 (自然科学版), 2000, 29 (3): 48-52.
- [4] 虹膜识别技术流程 [EB/OL]. [2015-08-03]. <http://www.techshino.com/tech/detail.html?id=3>.
- [5] Daugman J. How iris recognition works [J]. IEEE Transactions on Circuits and Systems for Video Technology, 2004, 14 (1): 21-30.
- [6] Boles WW, Boashash B. A human identification technique using images of the iris and wavelet transform [J]. IEEE Transactions on Signal Processing, 1998, 46 (4): 1185-1188.
- [7] Wildes RP. Iris recognition: an emerging biometric technology [J]. Proceedings of the IEEE, 1997, 85 (9): 1348-1363.
- [8] 5.1.185 资源: 虹膜数据库 [EB/OL]. [2015-08-04]. http://www.ecas.cn/xxzy/dep/dwz/kxsjk/201101/t20110112_3061414.html.
- [9] Cherabit N, Chelali FZ, Djeradi A. Circular hough transform for iris localization [J]. Science and Technology, 2012, 2 (5): 114-121.
- [10] 周俊, 罗挺, 王友斌, 等. 一种基于几何特征的虹膜定位算法 [J]. 后勤工程学院学报, 2010, 26 (3): 81-85.
- [11] 苑玮琦, 马军防, 狄文彬. 基于主动轮廓线的虹膜定位方法 [J]. 计算机工程与应用, 2004, 39 (34): 104-107.
- [12] LeCun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition [J]. Proceedings of the IEEE, 1998, 86 (11): 2278-2324.
- [13] LISA lab. Convolutional Neural Networks [EB/OL]. (2015-11-09) [2015-08-03]. <http://deeplearning.net/tutorial/lenet.html>.
- [14] LeCun Y, Boser B, Denker JS, et al. Backpropagation applied to handwritten zip code recognition [J]. Neural Computation, 1989, 1 (4): 541-551.
- [15] LeCun Y. LeNet-5, convolutional neural networks [EB/OL]. [2015-08-12]. <http://yann.lecun.com/exdb/lenet/>.
- [16] Jia Y, Shelhamer E, Donahue J, et al. Caffe: convolutional architecture for fast feature embedding [C] // Proceedings of the ACM International Conference on Multimedia, 2014: 675-678.
- [17] Chen LC, Papandreou G, Kokkinos I, et al. Semantic image segmentation with deep convolutional nets and fully connected CRFs [EB/OL]. [2015-09-03]. <http://arxiv.org/abs/1412.7062>.
- [18] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation [EB/OL]. [2015-09-03]. <http://arxiv.org/abs/1411.4038>.
- [19] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks [C] // Advances in Neural Information Processing Systems, 2012: 1097-1105.
- [20] Boyd M, Carmaciu D, Giannaros F, et al. MSc Computing Science Group Project: Iris Recogniton [R/OL]. (2010-03-19) [2015-09-03]. <http://projectiris.co.uk/final.pdf>.
- [21] Noh H, Hong S, Han B. Learning Deconvolution Network for Semantic Segmentation [EB/OL]. [2015-09-03]. <http://arxiv.org/abs/1505.04366>.
- [22] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition [EB/OL]. [2015-09-03]. <http://arxiv.org/abs/1409.1556>.