

基于深度学习的指静脉识别研究

吴 超 邵 曦

(南京邮电大学 通信与信息工程学院 江苏 南京 210003)

摘 要: 提出了基于深度学习的指静脉识别算法。在指静脉图像采集过程中,由于受光照强度的影响,手指轮廓存在一定的模糊。为了获得良好的静脉区域图像,采用形态学算法对指静脉原始图像进行感兴趣区域提取,并进一步使用高斯高通滤波器来增强图像。指静脉采集过程中,手指存在不同程度的旋转,为了消除该影响,使用角度修正算法对指静脉图像进行矫正。由于深度学习在图像分类上表现优异,尤其是 AlexNet 在 ImageNet 大赛中的杰出表现,因此采用基于 AlexNet 的深度神经网络对指静脉图像进行分类。为了加快训练速度,在 AlexNet 深度神经网络的基础上提出改进方案,主要包括改变卷积核大小和卷积层的构造,从而减少网络参数,降低网络复杂度,加速网络的训练。实验结果表明,利用深度学习对指静脉图像进行分类具有较好的效果。

关键词: 指静脉识别; 形态学算法; 角度修正算法; 深度学习

中图分类号: TP301.6

文献标识码: A

文章编号: 1673-629X(2018)02-0200-05

doi: 10.3969/j.issn.1673-629X.2018.02.043

Research on Finger Vein Recognition Based on Deep Learning

WU Chao, SHAO Xi

(School of Communication and Information Engineering, Nanjing University of
Posts and Telecommunications, Nanjing 210003, China)

Abstract: In this paper, we put forward a finger vein recognition algorithm based on deep learning. In the collection of finger vein image, sometimes the finger contour is blur due to the influence of illumination intensity. In order to obtain a clear vein image, morphological algorithm is used to extract the region of interest for the original image, and the Gaussian high pass filter is adopted to enhance it. An angle correction algorithm is for eliminating the influence caused by the different degrees of finger rotation in the collection of finger vein image. Moreover, deep learning is performed well in image classification, especially the AlexNet in the ImageNet contest. Therefore, the neural network based on AlexNet is adopted to classify the finger vein image. In order to speed up the training, we modify the structure of the AlexNet network, including the size of the convolution kernel and the structure of the convolution layer, which lowers the network parameters, reduces network complexity, and accelerates the training of the network. The experiments show that using deep learning for classification of finger vein image has better effect.

Key words: finger vein recognition; morphological algorithm; angle correction algorithm; deep learning

1 概 述

生物识别技术是使用人体特征进行身份认证的一种方法,人体特征包括人脸、指纹、虹膜和静脉。与传统生物特征相比,人体的指静脉特征具有安全性、唯一性和长期不变性等优点。1992 年,北海道大学的研究人员第一次证明了世界上任意两个人即使是双胞胎都不可能拥有完全相同的静脉特征^[1-2]。由于指静脉图像是利用红外摄像头采集得到,与指纹采集等相比,指

静脉采集具有非接触性,并且很难被窃取和伪造。与指纹、声音和人脸等生物特征不同的是,即使某个手指受到损伤,依然可以利用其他手指静脉^[3-4]。

由于指静脉作为生物特征识别具有显著的优势,因此国内外都对此进行了广泛研究。国外对该领域研究较早,技术较为成熟。日本日立公司从 1997 年就开始这方面研究,2004 年推出静脉认证的自动存取款机^[5-6]。2010 年,该公司的 Naoto Miura 等在指静脉采

收稿日期: 2017-03-24

修回日期: 2017-07-31

网络出版时间: 2017-11-15

基金项目: 国家自然科学基金青年基金项目(61401227)

作者简介: 吴 超(1993-)男,硕士研究生,研究方向为图像分类和生物特征识别;邵 曦,博士,副教授,研究生导师,研究方向为多媒体信息处理系统、基于内容的音乐信息检索等。

网络出版地址: <http://kns.cnki.net/kcms/detail/61.1450.TP.20171115.1438.072.html>

集装置的设计、指静脉图像的获取和指静脉特征识别等多个方面进行了深入研究并发表多篇文献^[7]。韩国国立大学的 Joon Hwan 等针对指静脉图像的预处理进行研究,并在 2009 年发表的文献中指出通过梯度、主曲率与二值化等方法对指静脉图像进行预处理,可以有效提取静脉纹路^[8]。在国内,也有许多研究人员从事指静脉识别研究。比如,张忠波等针对手指静脉特征提取技术,提出了基于曲波和神经网络的多尺度特征提取方法,取得了较好的识别效果^[9]。

指静脉图像采集过程容易因光照强度、角度和环境温度等外界因素的影响而降低图像质量。因此需要对指静脉图像进行适当的预处理,以降低外界因素对指静脉识别的不利影响。指静脉采集过程中,由于手指的旋转、倾斜等会使同一手指静脉在不同时刻产生轻微的旋转角度,为了提高识别精度,需要对原始图像进行矫正。

近年来,由于计算机运算能力的不断提升,深度学习也有了巨大进步,获得了优异的成果。2012 年, Alex 等在 ImageNet 图像分类竞赛中利用 AlexNet 深度神经网络做出了非常好的效果,超过了其他所有浅层的方法^[10]。深度学习可以通过学习一个深度非线性网络,实现复杂函数逼近,展现了其从少量样本集中学习数据集本质特征的能力。因此文中提出基于深度学习的指静脉识别方法。

2 指静脉提取与预处理

2.1 指静脉采集装置原理

每个人的手指表皮下的静脉纹路都是独一无二的,即使是双胞胎的对应手指静脉分布也一定不相同。当进入成年以后,手指静脉图像不再改变,具有较好的稳定性和唯一性。这为指静脉识别提供了医学依据。

指静脉图像采集装置是由近红外光发生装置和红外摄像头采集装置组成。实验证明 760~850 nm 波长的近红外光对肌肉和骨骼具有较好的穿透能力,并且静脉中的血红蛋白对此范围的近红外光具有良好的吸收能力^[11-12]。近红外光从手指背部穿过,由于静脉血红蛋白可以吸收近红外光,因此红外摄像头采集到的手指图像会形成一些较暗的纹路,这些纹路便是手指静脉图像。

然而由于光照强度、光照角度、骨骼和肌肉厚度等因素的影响,使得红外摄像头采集到的图像除了包含指静脉纹路外,还包含一些不规则的暗影和噪声。这些噪声会对后期的指静脉识别产生干扰,需要对采集图像做进一步处理。

2.2 指静脉图像的矫正

同一手指在不同时间采集手指静脉时,不可避免

地会因为手指与设备角度的偏差而产生差异较大的手指静脉图像。因此需要对图像进行矫正。图像矫正前,需要计算手指倾斜角度,并根据倾斜角通过图像旋转矫正原始图像。

用 $F_1(x, y)$ 、 $F_2(x, y)$ 分别表示手指的左右边界曲线, N_1 表示曲线 F_1 上点的个数, N_2 表示曲线 F_2 上点的个数。 (x_{1i}, y_{1i}) 和 (x_{2j}, y_{2j}) 分别表示曲线 F_1 第 i 个点的坐标以及曲线 F_2 第 j 个点的坐标。图像旋转角度由以下公式给出:

$$X_{10} = \frac{\sum_{N_1} x_{1i}}{N_1} \quad (1)$$

$$Y_{10} = \frac{\sum_{N_1} y_{1i}}{N_1} \quad (2)$$

$$X_{20} = \frac{\sum_{N_2} x_{2i}}{N_2} \quad (3)$$

$$Y_{20} = \frac{\sum_{N_2} y_{2i}}{N_2} \quad (4)$$

$$X_1 = \sum_{N_1} |x_{1i} - x_{10}| \quad (5)$$

$$Y_1 = \sum_{N_1} |y_{1i} - y_{10}| \quad (6)$$

$$X_2 = \sum_{N_2} |x_{2i} - x_{20}| \quad (7)$$

$$Y_2 = \sum_{N_2} |y_{2i} - y_{20}| \quad (8)$$

$$\theta = \frac{1}{2} \left[\frac{Y_1}{X_1} \cdot \frac{180}{\pi} + \frac{Y_2}{X_2} \cdot \frac{180}{\pi} \right] \quad (9)$$

得到指静脉图像倾斜角后,通过反向旋转即可对图像进行矫正。

经过上述方法获取旋转角后进行反向旋转得到矫正图像。旋转矫正后的图像基本处于垂直角度,很好地解决了因为采集过程手指不同程度偏转而产生的图像差异问题。

2.3 指静脉图像的 ROI 提取

通过指静脉采集装置采集到的原始图像包含指静脉特征区域和背景区域,为了减少背景对后续指静脉特征提取的影响,需要去除背景区域,即感兴趣区域提取(region of interest, ROI)^[13-17]。文中使用形态学算法处理手指边缘,获得精确边缘检测。

使用 Sobel 算子进行原始图像的垂直边缘检测,但由于光线较暗,边缘检测效果并不是很好,会出现断断续续的点。因此下一步采用形态学算法进一步处理^[18-20]。

对上一步的边缘检测结果进行闭合操作,闭操作可使轮廓线更光滑,减少狭窄的间断和细长的鸿沟,消除小的空洞,并填补轮廓线中的断裂。接着进一步骨

髓化、细化边缘区域,由此可以得到清晰明确的两条边界线。根据图像左右的边界线切割图像得到指静脉区域图像。为了便于后续处理,把所有指静脉图像尺寸都放缩至 300×100 。

ROI 图像提取如图 1 所示。

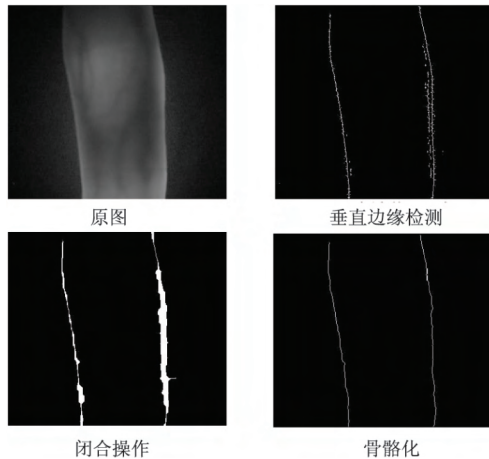


图 1 ROI 图像提取

2.4 指静脉图像增强

由于红外光易受到手指肌肉和骨骼的影响,在非静脉区域也会含有大量暗部。这使指静脉图像的对比度降低,对图像识别产生不利影响。这里引入一个高斯高通滤波器来增强图像的对比度^[21-22]。滤波器的公式如下:

$$H(x, y) = a(1 - e^{-D^2(x, y)/2D_0^2}) + b \quad (10)$$

$D(x, y)$ 定义如下:

$$D(x, y) = [(x - x_0)^2 + (y - y_0)^2]^{1/2} \quad (11)$$

其中, $D(x, y)$ 表示中心点与点 (x, y) 的距离,卷积模板的中心点为 $D_0(x_0, y_0)$ 。通过更改系数 a 、 b 的值来调节滤波器的滤波效果。文中卷积模板尺寸为 9×9 , a 和 b 的取值分别为 12.41 和 -4。

实验结果表明,滤波后的图像对比度明显高于原始图像,很好地去除了暗部的不利影响。

3 深度神经网络

3.1 深度学习

深度学习(也称深度结构化学习,分层学习或深层机器学习)是基于一组算法的机器学习的分支,深度学习通过组合低层特征形成更加抽象的高层表示属性类别或特征,以发现数据的分布式特征表示^[23]。

深度学习是机器学习研究中的一个新领域,其目的在于建立、模拟人脑进行分析学习的神经网络,它模仿人脑的机制来解释数据,例如图像、声音和文本。深度学习领域的研究主要致力于寻找对数据更好的表示以及构建模型来从大规模无标签的数据中学习这些特征表示。一些研究人员受到神经学科的启发,基于神

经系统中的信息处理和通信模式的解释构建出多种模型。诸如深度神经网络、卷积深度神经网络、深度信念网络和复现神经网络的各种深度学习架构已经应用于计算机视觉、自动语音识别、自然语言处理、音频识别和生物信息学等领域,并且显示出了一定的优势。

3.2 AlexNet 神经网络模型

2012 年,Geoffrey Hinton 和 Alex 在 ImageNet 的竞赛中使用一个 8 层的卷积神经网络(convolutional neural network, CNN)刷新了图像分类的记录,奠定了深度学习在计算机视觉中的地位,后来大家把这 8 层的 CNN 命名为 AlexNet。AlexNet 是一种典型的卷积神经网络,共有 8 层,其中前 5 层是卷积层,后 3 层全连接,最后的一个全连接层的输出是 softmax。

整体结构如图 2 所示。

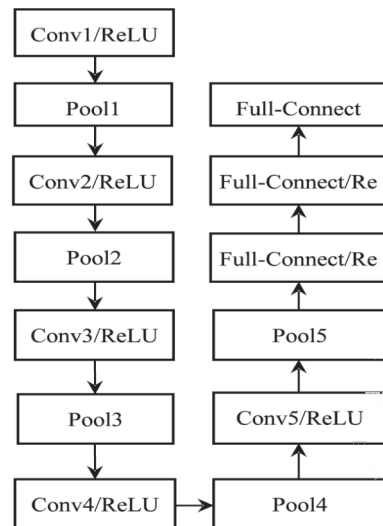


图 2 AlexNet 网络结构

该卷积神经网络涉及到以下几个部分:ReLU 激活函数、卷积和池化。

(1) ReLU 激活函数^[24]。

Sigmoid 是常用的非线性激活函数,它能够把输入的连续实值“压缩”到 0 和 1 之间。特别的,如果是非常大的负数,那么输出就是 0;如果是非常大的正数,输出就是 1。但是它有一些缺点:当输入非常大或者非常小时,会有饱和现象,这些神经元的梯度是接近于 0 的。如果初始值很大的话,梯度在反向传播时因为需要乘上一个 Sigmoid 的导数,使得梯度越来越小,导致网络变得很难学习。并且 Sigmoid 的输出不是 0 均值,这会导致后一层的神经元将得到上一层输出的非 0 均值的信号作为输入,会对训练产生不利影响。因此这里用 ReLU 激活函数来替代 Sigmoid 激活函数。

(2) 卷积。

AlexNet 的核心就是卷积层。卷积运算一个重要的特点就是增强原信号特征,并且降低噪音。不同的卷积核能够提取到图像中的不同特征。

(3) 池化。

池化是在卷积特征提取的基础上,对每个卷积特征进行取平均(或者取最大值),进一步缩小隐藏节点对应的卷积特征维数,减小分类器的设计负担。

3.3 改进的神经网络模型

AlexNet 神经网络结构有 6 000 万个参数,这给神经网络的学习带来了很大问题。较多的参数使得学习成本增高,并且可能带来过拟合的问题。因此提出改进方法,在不降低神经网络性能的前提下,进一步减少神经网络结构的参数数量。策略如下:

(1) 替换原来的 3×3 的卷积核为 1×1 的卷积核。

(2) 减少输入 3×3 卷积的特征图(feature map)的数量。

具体实现中,将 AlexNet 网络的每个卷积层分为两个自带 ReLU 函数的卷积层,第一层的卷积核大小为 1×1 ,第二层的卷积中包含 1×1 和 3×3 两种卷积核大小。在 ImageNet 上,参数数目可减少 50 倍。

改进后的 AlexNet 整体网络结构如图 3 所示。

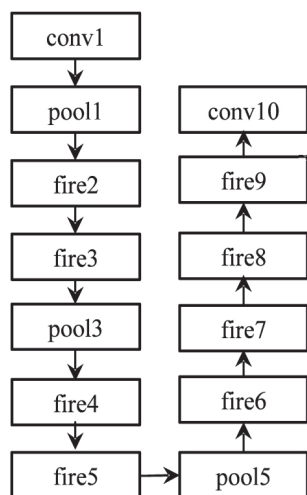


图 3 网络结构

3.4 Caffe 开源框架

Caffe 是一个清晰而高效的深度学习框架^[25]。由于其模型与相关的优化都是以文本的形式给出,因此使用 Caffe 框架可以方便地定义神经网络,并可以通过命令行或者接口的调用来方便地对网络进行训练。因此文中所有实验都是在 Caffe 框架下进行。

在 Caffe 中有两个核心配置文件 train_val.prototxt 和 solver.prototxt。train_val.prototxt 文件是网络配置文件,该文件通过 layer 参数来定义不同的网络层。solver.prototxt 文件是 Caffe 运行的关键,它协调着整个模型的运作。在深度学习中,loss 函数往往是非凸的,没有解析解,必须找到一个优化方法来求解。solver 的主要作用就是交替调用前向(forward)算法和后向(backward)算法来更新参数,从而最小化 loss,实际上就是一种迭代的优化算法。

在计算 loss 函数时,传统的梯度下降算法(gradient descent, GD)要通过所有样本计算梯度,而随机梯度下降算法(stochastic gradient descent, SGD)通过一个样本或部分样本计算一个近似的梯度,因此 SGD 计算量要更小。在训练集较大时,SGD 算法收敛速度更快。因此文中采用 SGD 算法。

4 实验

4.1 数据集的扩展

实验数据采集于 123 个人的四根手指,每根手指采集 12 张静脉图片,即数据集的大小为 492×12 。为了对网络模型进行更好的训练,减少数据集过拟合的问题,对该数据集进行扩展。扩展的方法是对原始图像进行随机抽取。抽取方式有两种,第一种方法是在尺寸为 300×100 的原始图像上随机抽取一个尺寸为 270×90 的子图像,并将原始图像和抽取图像统一压缩尺寸到 60×30 。第二种方法是对原始图像进行顺时针和逆时针旋转,旋转角度范围为 $[0, 2]$,并将原始图像和抽取图像统一压缩尺寸到 60×30 。利用上述两种抽取方式来扩展数据集,将原始数据集的大小由 492×12 扩大为 492×120 。

实验结果表明,通过这两种方式扩展得到的数据集具有良好的训练效果。

4.2 实验结果

通过上一节所提方法得到的数据集有 492 类手指图像,每类包含 120 张。利用该数据集进行模型训练。

该实验利用 Caffe 开源框架设计和实现深度神经网络。其中计算机 CPU 型号为 E7440,主频 2.4 GHz,8 核 16 G 内存。

利用 Caffe 框架对该神经网络进行迭代训练达到 80 000 次,实验结果如表 1、图 4 和图 5 所示。由表 1 可以看出,随着训练次数的增多,识别率逐渐提高,loss 值逐步降低。由图 4 和图 5 可知,当训练次数达到 70 000~80 000 次时,整个网络模型趋于稳定。

表 1 训练结果

训练次数	识别率	Loss 值
5 000	0.069	4.010
7 000	0.267	3.018
9 000	0.498	2.057
10 000	0.719	1.392
20 000	0.903	0.416
30 000	0.957	0.297
50 000	0.991	0.082
80 000	0.993	0.066

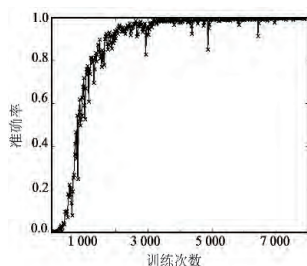


图 4 识别率变化曲线

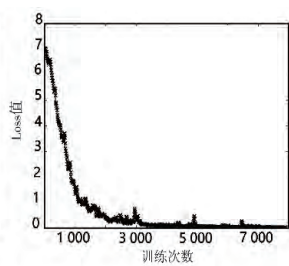


图 5 loss 值变化曲线

以上实验表明,改进过的 AlexNet 神经网络对指静脉图像进行分类能够得到较好的结果。

5 结束语

随着信息技术的高速发展,对生物特征识别技术的需求越来越高。由最初的指纹识别、语音识别到后来的虹膜识别和指静脉识别,每一次技术的突破都意味着信息安全的重要性不断得到重视。如今,电子产品与人类的交互越来越频繁,人们对设备的数据安全需求不断提高,因此研究指静脉识别具有很好的现实意义。近来发展火热的深度学习证明了其在图像分类等领域的优势。实验结果表明,利用深度学习构造深度神经网络模型,对指静脉图像的分类具有良好的效果。

参考文献:

- [1] XI X, YANG G, YIN Y, et al. Finger vein recognition with personalized feature selection [J]. *Sensors*, 2013, 13 (9) : 11243-11259.
- [2] KONO M, UEKI H, UMEMURA S. A new method for the identification of individuals by using of vein pattern matching of a finger [C]//Proceedings of the 5th symposium on pattern measurement. Yamaguchi, Japan [s.n.]: 2000: 9-12.
- [3] XIE S J, YANG J, YOON S, et al. Guided Gabor filter for finger vein pattern extraction [C]//Eighth international conference on signal image technology and internet based systems. [s.l.]: IEEE, 2012: 118-123.
- [4] LU Y, YOON S, XIE S J, et al. Finger vein recognition using histogram of competitive Gabor responses [C]//22nd international conference on pattern recognition. [s.l.]: IEEE, 2014: 1758-1763.
- [5] JAIN A K, ROSS A, PANKANTI S. Biometrics: a tool for information security [J]. *IEEE Transactions on Information Forensics and Security*, 2006, 1 (2) : 125-143.
- [6] KONO M, UEKI H, UMEMURA S. Near-infrared finger vein patterns for personal identification [J]. *Applied Optics*, 2002, 41 (35) : 7429-7436.
- [7] YAMADA J, NISHIKAWA Y. Personal identification device and method: U.S. 6907134 [P]. 2005-06-14.
- [8] CHOI J H. Finger vein extraction using gradient normalization and principal curvature [J]. *Proceedings of SPIE*, 2009, 7251: 725111.

- [9] ZHANG Z, MA S, HAN X. Multiscale feature extraction of finger-vein patterns based on curvelets and local interconnection structure neural network [C]//International conference on pattern recognition. [s.l.]: IEEE Computer Society, 2006: 145-148.
- [10] KRIZHEVSKY A, SUTSKEVER I, HINTON G E. ImageNet classification with deep convolutional neural networks [C]//International conference on neural information processing systems. [s.l.]: Curran Associates Inc, 2012: 1097-1105.
- [11] MIURA N, NAGASAKA A, MIYATAKE T. Feature extraction of finger-vein patterns based on repeated line tracking and its application to personal identification [J]. *Machine Vision and Applications*, 2004, 15 (4) : 194-203.
- [12] 杨数强, 王军强, 周涛. 手指静脉采集控制板的研制 [J]. *计算机工程与应用*, 2011, 47 (27) : 80-82.
- [13] YANG J, SHI Y. Finger-vein ROI localization and vein ridge enhancement [J]. *Pattern Recognition Letters*, 2012, 33 (12) : 1569-1579.
- [14] 杨金锋, 张海金. 手指静脉图像感兴趣区域提取方法研究 [J]. *山东大学学报: 工学版*, 2012, 42 (3) : 6-12.
- [15] 苑玮琦, 高洁睿. 基于局部灰度极小值的指静脉图像分割方法 [J]. *计算机技术与发展*, 2016, 26 (7) : 109-111.
- [16] 余成波, 秦华锋. 手指静脉图像特征提取算法的研究 [J]. *计算机工程与应用*, 2008, 44 (24) : 175-177.
- [17] 杨颖, 杨公平. 手指静脉识别技术 [J]. *计算机科学与探索*, 2012, 6 (4) : 343-354.
- [18] LIU Z, YIN Y, WANG H, et al. Finger vein recognition with manifold learning [J]. *Journal of Network & Computer Applications*, 2010, 33 (3) : 275-282.
- [19] WANG J X. Gabor filter based fingerprint image enhancement [C]//Fifth international conference on machine vision. [s.l.]: International Society for Optics and Photonics, 2013.
- [20] 冈萨雷斯, 伍兹, 埃丁斯, 等. 数字图像处理: MATLAB 版 [M]. 北京: 电子工业出版社, 2005.
- [21] ROSDI B A, SHING C W, SUANDI S A. Finger vein recognition using local line binary pattern [J]. *Sensors*, 2011, 11 (12) : 11357-11371.
- [22] 刘洋, 郭树旭, 张凤春, 等. 基于稀疏分解的指静脉图像去噪 [J]. *信号处理*, 2012, 28 (2) : 179-185.
- [23] AREL I, ROSE D C, KARNOWSKI T P. Deep machine learning - a new frontier in artificial intelligence research [research frontier] [J]. *IEEE Computational Intelligence Magazine*, 2010, 5 (4) : 13-18.
- [24] GLOROT X, BORDES A, BENGIO Y. Deep Sparse Rectifier Neural Networks [C]//International conference on artificial intelligence and statistics. [s.l.]: [s.n.], 2011: 315-323.
- [25] JIA Y, SHELHAMER E, DONAHUE J, et al. Caffe: convolutional architecture for fast feature embedding [C]//Proceedings of the 22nd ACM international conference on multimedia. [s.l.]: ACM, 2014: 675-678.