

DOI:10.3969/j.issn.1673-4785.201405060
网络出版地址: <http://www.cnki.net/kcms/detail/23.1538.TP.20150326.1015.003.html>

深度信念网络的二代身份证异构人脸核实算法

张媛媛,霍静,杨婉琪,高阳,史颖欢

(南京大学 计算机软件新技术国家重点实验室,江苏 南京 210023)

摘 要:二代身份证人脸核实问题是指判断二代身份证人像和身份证使用者当前头像是否为同一人。具体来说,即将二代身份证模糊人像和实际在光照、背景等因素不可控环境下拍摄的若干张二代证使用者的视频人像作匹配,判断是否为同一个人。由于低分辨率模糊图像和清晰视频图像属于 2 种不同的图像模态,因此该问题属于异构人脸识别问题。考虑到跨模态人脸图像的差异,传统的特征抽取方法很难抽取判别性特征来描述不同模态图像,使得传统方法难以达到精准辨别。针对这个问题,提出了一种新的基于深度学习的解决方法,其基本思想是通过深度信念网络(DBN)的非监督贪心逐层训练来提取人脸图像的高层特征,结合传统的图像预处理和相似性度量技术,达到人脸核实的目的。通过在 256 人的真实二代证数据集上和传统特征降维方法 PCA、LDA 进行比较,证实了所提出方法在准确率上相比 PCA 有约 12% 的提升,相比 LDA 有约 8% 的提升。实验同时表明,针对数据量增大的情况,基于深度学习的解决方法要优于传统的人脸识别方法。

关键词:人脸核实;多模态;深度学习;深度信念网络

中图分类号:TP391 **文献标志码:**A **文章编号:**1673-4785(2015)02-0193-08

中文引用格式:张媛媛,霍静,杨婉琪,等. 深度信念网络的二代身份证异构人脸核实算法[J]. 智能系统学报, 2015, 10(2): 193-200.

英文引用格式:ZHANG Yuanyuan, HUO Jing, YANG Wanqi, et al. A deep belief network-based heterogeneous face verification method for the second-generation identity card[J]. CAAI Transactions on Intelligent Systems, 2015, 10(2): 193-200.

A deep belief network-based heterogeneous face verification method for the second-generation identity card

ZHANG Yuanyuan, HUO Jing, YANG Wanqi, GAO Yang, SHI Yinghuan

(State Key Laboratory for Software Technology, Nanjing University, Nanjing 210023, China)

Abstract: The objective of the face verification method for the second-generation identity card is to determine whether the original head-photo stored in the corresponding identity card image and the currently captured head photo of the card-holder by using a video camera image actually belongs to the same person or not. To obtain a good verification result for the heterogeneous face verification method is a very challenging task because the two different types of images belong to two different modalities (e.g., different image resolutions, different illumination conditions). Considering the difference of trans-modal face images, it is hard to use traditional feature extraction methods to extract discriminative feature for description of images with different modes. Traditional feature extraction methods cannot distinguish images exactly. In this paper, a deep learning-based face verification method is proposed. The proposed deep learning-based face verification method integrates the deep belief network (DBN), which employs unsupervised greedy layer-by-layer training for high-level feature extraction of face photo and combines the popularly used image preprocessing and similarity measurement technologies to realize the purpose of face verification. The results were evaluated on a real dataset with two different modalities of 256 different people. This method outperforms the traditional principal component analysis (PCA) and linear discriminant analysis (LDA) methods with 12% and 8% improvements in terms of the verification accuracy, respectively. The results validated the advantage of the proposed method, especially when the amount of entries increases.

Keywords: face recognition; multimodes; deep learning; deep belief network

收稿日期:2014-05-28. 网络出版日期:2015-03-26.
基金项目:国家自然科学基金资助项目(61035003,61175042).
通信作者:张媛媛.E-mail:zhangyuanyuan2013nju@gmail.com.

目前,中国公民广泛使用第 2 代身份证作为身份识别的手段。然而,随着近年来社会对于治安监

控的逐步重视,基于计算机辅助的面向第 2 代身份证的人脸核实问题成为许多研究者关注的问题。

早期的人脸核实研究大都针对受控环境下的人脸识别问题,该问题目前已经可以被较好地解决。但在拍照环境不确定、用户不配合、年龄更替等不可控环境下,人脸核实的正确率急剧下降。因此,目前大部分的研究开始关注不受控环境下的人脸核实问题,该问题的主要难点在于环境、光照、姿态、年龄更替等因素引起的人脸外观急剧变化,从而导致识别困难^[1-2]。

1 相关工作

本文重点研究了二代身份证的人脸核实问题,该问题属于不受控环境下的人脸核实问题。如图 1,判断二代身份证的模糊人脸图像和二代身份证使用者的清晰视频人脸图像是否属于同一个人。在这个问题中,影响核实系统性能的因素主要有:

1) 模态差异。二代身份证上人像属于低分辨率图像模态,而现实情境下的视频人像属于清晰图像模态;

2) 人脸内部变化。年龄增长导致的面部变化、人脸的表情变化和装饰物等;

3) 外在环境因素。如拍摄地点不同而造成的光照问题、角度问题等。

不同于已有的不受控环境下的人脸核实问题,二代身份证的人脸核实问题由于涉及到匹配清晰和模糊 2 种不同模态的人像,该问题属于异构的人脸核实问题。由于模态不同导致的人脸图像之间的差异通常很大,因此对于异构人脸核实问题,想要达到精确判别十分困难。

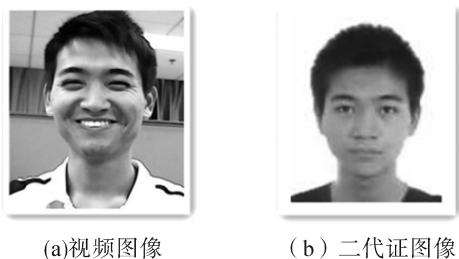


图 1 异构情境下的人脸核实

Fig.1 Heterogeneous face verification

对于传统的人脸核实算法,其基本假设是在进行人脸区域提取后,同一个个体内部的人像差异小于不同个体间的人像差异^[3],其解决手段主要为通

过降维的方法来提取人脸图像特征,包括主成分分析(PCA)^[4-5]、线性判别分析(LDA)^[6]、局部保持投影(locality preserving projection, LPP)^[7]等,然而上述方法皆属于浅层学习模型^[8],其局限性在于有限样本和计算单元情况下对复杂分类问题的函数表示能力有限,难以发现最具有判别能力的特征。此外,对于异构人脸核实问题,图像模态导致的差异通常伴随个体不同导致的人像差异,大部分传统方法无法考虑模态的差异进行特征提取,从而在此应用上会失效^[9]。

近年来,随着深度学习概念的提出,以深度信念网络(deep belief networks, DBN)为代表的深度神经网络模型,以其揭示数据中所隐藏的有用信息,而受到研究者们广泛的关注^[10-11]。

DBN 模型是一种逐层贪婪预训练的深层神经网络模型,它克服了传统神经网络在训练上的难度,通过多层来获得更加抽象的特征表达,挖掘隐藏在图像的像素特征之上的高维抽象特征。因此,针对二代身份证的人脸核实问题,在特征抽取部分首次使用了基于 DBN 的非监督贪心逐层预训练的方法进行权值初始化,结合传统的图像预处理和相似性度量技术,通过深层模型的特征抽取,学习到对数据有更本质的刻画的特征,继而提升人脸核实问题的准确性。

传统的人脸识别方法关注解决受控环境下的人脸识别,近年来,研究者们更多地关注不受控环境下的人脸识别问题^[12-14]。其中,异构人脸识别问题是一个研究热点。

一般来说,针对异构人脸图像的识别技术通常分为 2 类:1) 通过某种方法将由视图导致的人像特征差异减小,如针对近红外异构人脸数据集, Klare 等^[15]提出了对红外线人像(NIR)和可见光人像(VIS)进行随机子空间投影和稀疏表示来进行匹配的方法, Yi 等^[16]提出典型相关分析方法来学习 NIR 和 VIS 人像间的相似性;2) 通过将其中一种模态的人像转化为其他模态的人像,再进行匹配,如针对近红外异构人脸数据集, Wang 等^[17]提出了合成和分析的方法将 NIR 人像转变为合成的 VIS 人像, Chen 等^[18]使用局部线性嵌入算法将 NIR 人像转变为 VIS 人像。

深度学习是基于 Hubel-Wiesel 仿生学模型,通

过“逐层初始化”建立的多层人工神经网络模型。深度学习的实质是通过构建具有很多隐层的机器学习模型和海量的训练数据,来学习更有用的特征。区别于浅层学习,深度学习的不同在于:1)强调了模型结构的深度,通常有 4 层、5 层甚至 10 层的隐藏层节点;2)明确了特征学习的重要性,即通过逐层特征变换,将样本在原始空间的特征表示变换到一个新的特征空间,从而使分类或者预测更加容易。在人脸核实问题上,Sun 等^[19]提出将卷积神经网络(convolutional neural networks, CNN)和限制性玻尔兹曼机(restricted Boltzmann machines, RBM)相堆叠,利用 CNN 的二维特征提取和 RBM 的一维特征抽象的特性,在 LFW (labeled faces in the wild home)^[20]数据集上用以自然环境下的人脸核实问题,核实结果准确率为 86.88%的。本文相对于以往工作的主要创新是首次将深度学习引入到异构人脸核实问题上,解决了判别性特征难以提取的问题。

2 DBN 的基本原理

2.1 RBM 模型

RBM 是一个层内节点相互独立,层间节点的联合概率分布满足 Boltzmann 分布的二部图模型。如图 2,下层是可视层,即输入数据层,可视节点用 $v \in \mathbf{R}^l$ 表示,其偏置为 $b \in \mathbf{R}^l$,上层是隐藏层,隐藏节点用 $h \in \mathbf{R}^s$ 表示,其偏置为 $c \in \mathbf{R}^s$,可视层和隐藏层之间的连接权值用 $W \in \mathbf{R}^{l \times s}$ 表示。对于图像而言,像素层对应于可视层,特征描述子对应于隐藏层^[21]。

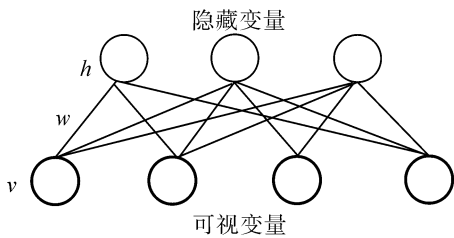


图 2 RBM 模型
Fig.2 RBM model

受统计学中能量泛函的启发,RBM 模型引入了能量函数,可视节点和隐藏节点的联合组态 (v, h) 的能量函数为

$$E(v, h) = - \sum_{j=1}^l b_j v_j - \sum_{i=1}^s c_i h_i - \sum_{j=1}^l \sum_{i=1}^s v_j h_i w_{ij}$$

依据 Boltzmann 分布,可视节点和隐藏节点的联合概率分布为

$$p(v, h) = \frac{1}{Z} e^{-E(v, h)}$$

式中: $Z = \sum_{v, h} e^{-E(v, h)}$, 为能量函数的指数函数的和。

有了联合概率,很容易得到

$$\begin{aligned} p(v) &= \frac{\sum_h e^{-E(v, h)}}{\sum_{v, h} e^{-E(v, h)}} \\ p(h) &= \frac{\sum_v e^{-E(v, h)}}{\sum_{v, h} e^{-E(v, h)}} \\ p(v|h) &= \frac{e^{-E(v, h)}}{\sum_v e^{-E(v, h)}} \\ p(h|v) &= \frac{e^{-E(v, h)}}{\sum_h e^{-E(v, h)}} \end{aligned} \tag{1}$$

此外,从统计学的概念出发,定义自由能量函数:

$$\text{FreeEnergy}(v) = - \ln \sum_h e^{-E(v, h)} \tag{2}$$

自由能量函数是描述整个系统状态的一种测度。系统越有序或者概率分布越集中,每个训练样本经过 RBM 网络编码到隐藏节点的取值概率也越集中,系统的能量就越小,此时 RBM 系统能够更好地拟合数据分布。

联合(1)、(2)可以得出

$$\ln p(v) = - \text{FreeEnergy}(v) - \ln Z \tag{3}$$

式(3)两边做个累加,可得

$$\sum_v \ln p(v) = - \sum_v \text{FreeEnergy}(v) - \sum_v \ln Z$$

可看出一个系统自由能量的总和最小的时候,正是 $\prod_v p(v)$ 最大的时候,也就是说,用极大似然估计去求得 $\prod_v p(v)$ 的参数能让 RBM 系统的自由能量总和达到最小,此时 RBM 系统能够最好地拟合数据分布。

定义 $\prod_v p(v)$ 的似然函数为

$$L(\theta) = \frac{1}{N} \sum_{n=1}^N \log p_{\theta}(v^{(n)})$$

式中: N 为节点的个数, $\theta = \{W, b, c\}$ 为 RBM 模型的参数。

用极大似然估计去求参数,可得到

$$\frac{\partial L(\theta)}{\partial W_{ij}} = E_{p_{\text{data}}}[v_j h_i] - E_{p_{\theta}}[v_j h_i]$$

式中: $E_{p_{\text{data}}}[v_j h_i]$ 为独立数据期望, $E_{p_{\theta}}[v_j h_i]$ 为模型期望。其中,模型期望的获得是十分困难的,其时

间与隐藏节点个数成指数关系。因此,Hinton 等^[22]提出了 CD_k (contrastive divergence)方法来模拟梯度的计算,整个算法过程如图3。

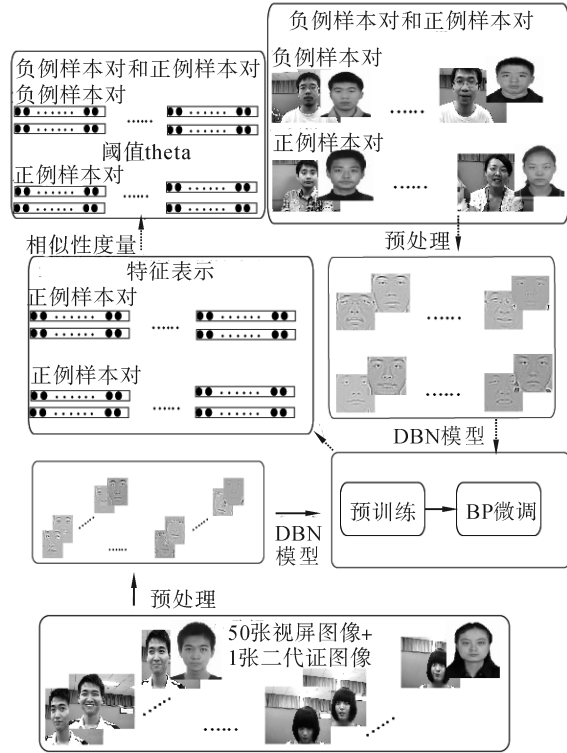


图3 系统框架

Fig.3 System framework

实验证明, CD_k 是一种很好的求解对数似然函数关于未知参数梯度的近似的方法。

CD_k 算法

输入: $RBM(V_1, 2, \dots, V_m, H_1, 2, \dots, H_n)$,

输出: 梯度估计 $\Delta w_{ij}, \Delta b_j, \Delta c_i$ 。

for $j = 1, 2, \dots, m, i = 1, 2, \dots, n$,

初始化 $\Delta w_{ij} = \Delta b_j = \Delta c_i = 0$ for $j = 1, 2, \dots, m$,

$i = 1, 2, \dots, n$ 。

for 所有的 v do

$v^{(0)} \leftarrow v$

for $t = 0, \dots, k - 1$ do

for $i = 1, 2, \dots, n$ do 采样 $h_i(t) \sim$

$p(h_i | v^{(t)})$

for $j = 1, 2, \dots, m$ do 采样 $v_j^{(t+1)} \sim p(v_j | h^{(t)})$

for $j = 1, 2, \dots, m, i = 1, 2, \dots, n$ do

$\Delta w_{ij} \leftarrow \Delta w_{ij} + p(h_i = 1 | v^{(0)})v_j^{(0)} - p(h_i = 1 | v^{(k)})v_j^{(k)}$

$\Delta b_j \leftarrow \Delta b_j + v_j^{(0)} - v_j^{(k)}$

$\Delta c_i \leftarrow \Delta c_i + p(h_i = 1 | v^{(0)}) - p(h_i = 1 | v^{(k)})$

2.2 RBM 学习算法

整个 RBM 的训练过程如下,这里用到了 CD_1 。

一层的 RBM 模型训练完毕后,固定其权值 w_{ij} 及偏置值 b_j, c_i ,隐藏层的输出 h_i 作为原输入信息的第 1 个表达,将隐藏层的输出 h_i 作为其上层 RBM 的输入,同样训练 RBM 模型,就会得到第 2 层的参数及其输出,其输出就作为原输入信息的第 2 个表达^[23]。如此不断往上叠加 RBM,就得到了 DBN 的初步模型。

以上在 DBN 模型的构建中统称为预训练阶段。预训练将网络参数训练到一组合适的初始值,从这组初始值出发会令代价函数达到一个更低的值^[24]。经过这种方式的训练后,再根据重构误差,使用传统的全局学习算法,比如 BP 算法对整个模型进行微调,从而使模型收敛到局部最优点。

RBM 训练算法符号说明:可视节点 j ,可视节点的偏置 b_j ,可视节点值为 1 的条件概率 $P(x_{2j} = 1 | h_1)$,sign 为 S 型函数, w_{ij} 为连接权值,隐藏节点 i ,隐藏节点的偏置 c_i ,隐藏节点值为 1 的条件概率 $Q(h_{1i} = 1 | x_1)$ 、 $Q(h_{2i} = 1 | x_2)$ 。

算法过程

输入: 可视节点值 x_1 ,

输出: 隐藏节点值 h_2 。

for 所有隐藏节点 i do

1) 计算 $Q(h_{1i} = 1 | x_1)$ (若为二值单元,则 Q 值为 $\text{sign}(c_i + \sum_j w_{ij}x_{1j})$)

2) 从 $Q(h_{1i} | x_1)$ 中采样 $h_{1i} \in \{0, 1\}$

end for

for 所有可视节点 j do

3) 计算 $P(x_{2j} = 1 | h_1)$ (若为二值单元,则 Q 值为 $\text{sign}(b_j + \sum_i w_{ij}h_{1i})$)

4) 从 $P(x_{2j} = 1 | h_1)$ 中采样 $x_{2j} \in \{0, 1\}$

end for

for 所有隐藏节点 i do

5) 计算 $Q(h_{2i} = 1 | x_2)$ (若为二值单元,则 Q 值为 $\text{sign}(c_i + \sum_j w_{ij}x_{2j})$)

end for

6) $W \leftarrow W + \varepsilon(h_{1i}x_{2j}' - Q(h_{2i} = 1 | x_2)x_{2j}')$

7) $b \leftarrow b + \varepsilon(x_1 - x_2)$

8) $c \leftarrow c + \varepsilon(h_1 - Q(h_2 = 1 | x_2))$

3 人脸核实问题的解决方案

人脸核实问题的解决算法流程如图 3,实线箭头部分为系统训练,虚线箭头部分为系统测试:

1)图像预处理
对收集到的数据集采用 Adaboost 算法^[25]和主动形状模型(active shape model, ASM)^[26]进行人脸检测和关键点定位,再按照瞳孔坐标位置进行归一化,如图 4,可以看到对齐后的效果。将归一化后的视频图像进行模糊化,最后对所有的图像进行 Quotient Image 方法去除光照干扰。

2)DBN 模型训练
将按照上述预处理后得到的图片进行高斯处理,使得均值为 1,方差为 0,如图 5,打乱训练样本图片,再将每张 40 × 40 的图片拉成一维行向量作为整个 DBN 模型的输入。

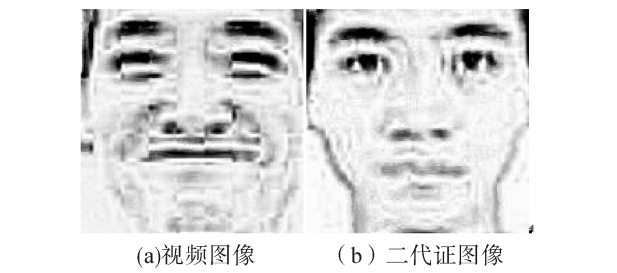


图 4 归一化效果图
Fig.4 Normalization



图 5 高斯处理后的效果图

Fig.5 Face pictures after Gaussian processing
用于二代证人脸核实问题上的 DBN 模型如图 6,第 1 层以 1 张分辨率为 40 × 40 的图像作为输入,即包含 1 600 个实值节点的可视层,第 2 层以及第 3 层均为二值隐藏层,分别包含 4 000,2 000 个节点,最后一层为实值隐含层,也为输出层,共包含 500 个节点。

整个 DBN 模型的训练过程为:
1)预训练。逐层叠加训练 RBM 模型,直至最高层。此时,整个模型网络参数的初始值在代价函数最优解附近,很大程度上避免了反向调节时 BP 算法陷入局部最优。

2)根据重构误差采用 BP 算法对整个模型进行反向调节。

整个深度学习模型可以看成是一个特征抽象的过程,即将原来 40 × 40 的图像经过一系列抽象,摒除无用信息,保留高维特征。

3)模型测试。将测试样本的数据均按照 1)所述进行预处理,进行高斯处理后再输入到 2)训练出的 DBN 模型中去,最高层的 500 个节点值作为图像最终的抽象特征向量。

4)相似性度量。将 50 个视频图像的特征向量和 1 个二代证图像的特征向量构成 50 个 (f_1, f_2) 对,其中 f_1 表示视频图像特征向量, f_2 表示二代证图像特征向量,计算向量的余弦距离,得到 50 个相似度数值,得到最大的相似度数值作为 50 张视频图像和一张二代证图像之间的相似度。对上述的相似度数值设置阈值,如果高于一个阈值则视频图像和二代证图像中是同一个人,否则不是同一个人。

DBN模型:预训练+反向微调

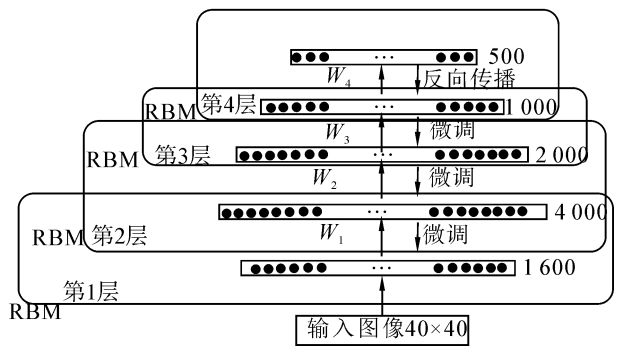


图 6 深度学习模型 DBN
Fig.6 DBN model

4 实验结果与分析

4.1 实验设置

实验用到的数据集为采集的二代身份证的人像数据集,具体包括:1)第 1 批采集的 98 人的二代身份证图像和视频图像(数据集 1);2)256 个人(第 1 批 98 个人+第 2 批 158 个人)的二代身份证图像和视频图像(数据集 2)。这 2 个数据集均是在特定的采集环境中采集,涵盖年龄、光照、姿态、表情等主要变化条件,每个人的数据为 51 张图像,包含二代身份证模糊照片 1 张和视频拍摄图像 50 张。

实验采用了 2 种方案进行:1)十折交叉验

证,以数据集 2 为例,即将 256 人的数据随机分成 10 份,训练模型使用其中的 9 份,测试使用其中的 1 份;2)非十折交叉,在数据集 1 上,训练模型使用 38 个人的数据,测试部分使用 60 人的数据;在数据集 2 上,训练模型使用 98 人的数据,测试部分使用 158 人的数据。

测试中使用的反例为十重交叉验证中一份数据中不同人的视频人像和二代证人像的组合。测试的准确率是正确判别的准确率(同一个人的图像分类为同一个人,不同的人的图像分类为不同的准确率;且阈值的选取为使得正确分类正例和正确分类反例的准确率相同时的值)。

4.2 深度学习核实结果

通过 DBN 深度学习模型的特征抽取,2 个数据集的人脸核实准确率结果如表 1。

表 1 2 个数据集上的正确率比较

Table 1 Accuracy rate on two datasets

数据集/人	十折交叉/%	非十折交叉/%
98	61.89 ± 0.7	51.20 ± 0.4
256	66.62 ± 1.1	62.41 ± 0.9

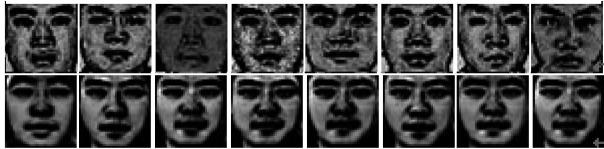
从表 1 可以看出,采用十折交叉的数据集 2 的效果最好,最高可以达到67.60%的准确率。从数据集 1 和数据集 2 的对比结果可知,训练数据越多,深度学习越能充分挖掘海量数据中隐藏的丰富信息。而数据集 1 的非十折交叉准确率仅有 51.21%,相当于一个弱分类器的效果。为了可视化地说明 RBM 用于特征描述的表现能力,给出数据集 1 和数据集 2 上的人脸图像及其经过 DBN 非监督贪心逐层训练得到的特征的重构图像示例,如图 7、8。由于训练样本过少,如图 9,测试样本图像与其重构结果相差较大,说明模型过拟合现象较为严重。将数据集 2 的十折交叉的测试结果也可可视化,如图 8,可看出重构结果比图 7(b)更具有区分能力。

4.3 和传统方法的对比结果

将深度学习模型和传统的特征降维 PCA、LDA 方法在二代证人脸核实上的准确率进行对比,以十折交叉的方案为例,如表 2 可看出,深度学习模型的解决方法在准确率上优于传统特征降维方法;应用问题数据量越大,深度模型越能充分挖掘更多有价值的信息和知识,效果越好,而传统方法则不然。



(a) 原图像



(b) 重构图像

图 7 数据集 1 中训练样本和测试样本的原图像和重构图像

Fig.7 The original pictures and constructed pictures of training set and test set in dataset_1



图 8 数据集 2 中训练样本和测试样本的原图像和重构图像

Fig.8 The original pictures and constructed pictures of training set and test set in dataset_2

表 2 与传统方法的正确率比较

Table 2 Compare with some traditional ways

方法	数据集 1/%	数据集 2/%
PCA	55.33	55.50
LDA	59.94	59.92
DBN 模型	61.89	67.60

4.4 层数对核实结果的影响

此外,还对 DBN 模型深度的设计做了相应的实验,如表 3 所示,以数据集 2 的非十折交叉为例,第 1 层到第 4 层逐层叠加之后的核实准确率是逐渐提升的,这表明此时的深度是有益的,而第 5 层时的准确率未有明显提升,表明模型的深度 4 已经足够。

表 3 DBN 模型对应的每层表现

Table 3 Accuracy rate for every level of DBN model

层数	准确率/%
1 层	58.23 ± 0.4
2 层	60.47 ± 0.7
3 层	61.39 ± 0.6
4 层	62.41 ± 0.9
5 层	61.16 ± 0.7

从以上结果和分析中可以看出,深度学习模型的实质是通过构建具有一定深度的机器学习模型和

海量的训练数据,来学习更有价值的特征,从而提升问题的准确性。所以,“深度模型”是手段,“特征学习”是目的^[27]。

当然,DBN模型也有其局限性,首先,其训练时间较长;其次,高分辨率图像必须经过压缩才能用于DBN模型;最后,DBN只关注图像的一维像素特征而忽略了图像本身固有的二维特性,诸如局部平移不变性等等,这些问题都有待进一步解决。

5 结束语

本文结合传统的图像预处理和相似性度量技术,把基于DBN的非监督贪心逐层训练的方法用于特征提取和降维,将深度学习运用在公安监控二代证核实问题上,提出的方法取得了相比传统方法更高的准确率。它为异构人脸核实和深度学习相结合的问题的研究提供了新的思路。而如何更加有效地深度挖掘图像的抽象特性,是一个值得深度研究的课题。一方面可以借鉴文献^[28],将RBM进行Convolution和Pooling操作,构成CDBN(convolutional deep belief networks)模型,另一方面也可借鉴文献[19],将CNN和RBM 2种基本模型层层堆叠成一个深度模型,此外,还可以结合二代证人脸核实问题的特殊性,即异构特征,针对不同模态下的人脸图像设计出不同的深度学习模型,用其他学习策略,如度量方法,学习不同模型之间的相似性等等,这方面的问题都有待于进一步研究。

参考文献:

- [1] CHAN C H, TAHIR M A, KITTLER J, et al. Multiscale local phase quantization for robust component-based face recognition using kernel fusion of multiple descriptors[J]. Pattern Analysis and Machine Intelligence, 2013, 35(5): 1164-1177.
- [2] LIAO S, JAIN A K, LI S Z, et al. Partial face recognition: alignment-free approach[J]. Pattern Analysis and Machine Intelligence, 2013, 35(5): 1193-1205.
- [3] CAO X, WIPF D, WEN F, et al. A practical transfer learning algorithm for face verification[C]//The IEEE International Conference on Computer Vision. Sydney, Australia, 2013: 3208-3215.
- [4] SIROVICH L, KIRBY M. Low-dimensional procedure for the characterization of human faces[J]. Journal of the Opti-

- cal Society of America, 1987, 4(3): 519-524.
- [5] TURK M A, PENTLAND A P. Face recognition using eigenfaces[C]//Computer Vision and Pattern Recognition. [S.l.], 1991: 586-591.
- [6] SCHOLKOPF B, MULLERT K R. Fisher discriminant analysis with kernels[C]//Proc of IEEE International Workshop on Neural Networks for Signal Processing. Madison, USA, 1999: 41-48.
- [7] HE X, NIYOGI P. Locality preserving projections[C]//Annual Conference on Neural Information Processing Systems. British Columbia, Canada, 2003, 16: 234-241.
- [8] BENGIO Y, LAMBLIN P, POPOVICI D, et al. Greedy layer-wise training of deep networks[C]//Neural Information Processing Systems. Vancouver, Canada, 2007, 19: 153-160.
- [9] LIN D, TANG X. Inter-modality face recognition[J]. European Conference on Computer Vision, 2006, 3954(4): 13-26.
- [10] HINTON G E, OSINDERO S, TEH Y W, et al. A fast learning algorithm for deep belief nets[J]. Neural Computation, 2006, 18(7): 1527-1554.
- [11] HINTON G E. Learning multiple layers of representation[J]. Trends in Cognitive Sciences, 2007, 11(10): 428-434.
- [12] BARKAN O, WEILL J, WOLF L, et al. Fast high dimensional vector multiplication face recognition[C]//The IEEE International Conference on Computer Vision. Sydney, Australia, 2013: 1-8.
- [13] SIMONYAN K, PARKHI O M, VEDALDI A, et al. Fisher vector faces in the wild[C]//Proc British Machine Vision Conference. Bristol, UK, 2013: 1-12.
- [14] CUI Z, LI W, XU D, et al. Fusing robust face region descriptors via multiple metric learning for face recognition in the wild[C]//Computer Vision and Pattern Recognition. Portland, Oregon, 2013: 3554-3561.
- [15] KLARE B, JAIN A K. Heterogeneous face recognition: matching NIR to visible light images[C]//International Conference on Pattern Recognition. Istanbul, Turkey, 2010: 1513-1516.
- [16] YI D, LIU R, CHU R, et al. Face matching between near infrared and visible light images[J]. Advances in Biometrics, 2007, 4642: 523-530.
- [17] WANG R, YANG J, YI D, et al. An analysis-by-synthesis method for heterogeneous face biometrics[J]. Advances in Biometrics, 2009, 5558: 319-326.
- [18] CHEN J, YI D, YANG J, et al. Learning mappings for face synthesis from near infrared to visual light images

[C]//Computer Vision and Pattern Recognition. Fla, USA, 2009: 156-163.

[19]SUN Y, WANG X, TANG X, et al. Hybrid deep learning for face verification [C]//IEEE International Conference on Computer Vision. Sydney, Australia, 2013: 1489-1496.

[20]HUANG G B, RAMESH M, BERG T, et al. Labeled faces in the wild: a database for studying face recognition in unconstrained environments [R]. Massachusetts, Amherst, 2007: 07-49.

[21]HINTON G. A practical guide to training restricted Boltzmann machines [J]. Momentum, 2010, 9(1): 599-619.

[22]HINTON G. Training products of experts by minimizing contrastive divergence [J]. Neural Computation, 2002, 14(8): 1771-1800.

[23]HINTON G. Learning multiple layers of representation [J]. Trends in Cognitive Sciences, 2007, 11(10): 428-434.

[24]HINTON G. Reducing the dimensionality of data with neural networks [J]. Science, 2006, 313(5786): 504-507.

[25]EDWARDS G J, COOTES T F, TAYLOR C J, et al. Face recognition using active appearance models [J]. Computer Vision—(ECCV). 1998, 1407: 581-595.

[26]VIOLA P, JONES M. Fast and robust classification using asymmetric adaboost and a detector cascade [C]//Neural Information Processing Systems. Vancouver, Canada, 2002: 1311-1318.

[27]余凯, 贾磊, 陈雨强, 等. 深度学习的昨天, 今天和明天 [J]. 计算机研究与发展 [J], 2013, 50(9): 1799-1804.

YU Kai, JIA Lei, CHEN Yuqiang, et al. Yesterday, today and tomorrow for deep learning [J]. Journal of Computer Research and Development, 2013, 50(9): 1799-1804.

[28]HUANG G B, LEE H, LEARNED-MILLER E, et al. Learning hierarchical representations for face verification with convolutional deep belief networks [C]//Computer Vision and Pattern Recognition. Newport, USA, 2012: 2518-2525.

作者简介:



张媛媛,女,1991 年生,硕士研究生,主要研究方向为机器视觉、机器学习等。



杨婉琪,女,1988 年生,博士研究生,主要研究方向为机器学习、机器视觉等。



高阳,男,1972 年生,教授,博士生导师,主要研究方向为强化学习、智能 agent、智能应用等。