

基于局部二值模式与深度置信网络的人脸识别

满忠昂¹, 刘纪敏², 孙宗锜¹

(1.山东科技大学计算机科学与工程学院, 山东 青岛 266590;

2.山东科技大学智能装备学院, 山东 泰安 271000)

✉manzhongang@163.com; ThereliveAngel@163.com; skdljm@126.com



摘要: 针对现在的大多算法在提取人脸特征时直接提取整个人脸, 而忽略局部的细节特征, 提出一种将人脸图像进行分块局部运用LBP算子然后与深度置信网络结合的人脸识别算法(BPBN)。首先, 将人脸图像进行分块, 对分块后的图像提取LBP进行统计, 将生成的LBP直方图按照一定顺序组合连接成新的特征向量。其次, 将得到的LBP特征作为深度置信网络(DBN)的输入, 采用贪婪算法逐层进行训练, 然后用反向传播(BP)算法对训练得到的深度置信网络进行优化。最后, 用训练好的深度置信网络对人脸进行识别。在ORL人脸数据库上进行实验, 识别率达到96.0%, 然后与传统的PCA算法集成支持向量机(SVM)的方法进行相比, 识别率有较为显著的提升, 说明该方法具有更好的人脸识别效果。

关键词: 局部二值模式; 人脸识别; 受限波尔兹曼机; 深度置信网络

中图分类号: TP183 **文献标识码:** A

Face Recognition Based on Local Two Value Model and Deep Belief Network

MAN Zhongang¹, LIU Jimin², SUN Zongkun¹

(1.College of Computer Science and Engineering, Shandong University of Science and Technology, Qingdao 266590, China;

2.College of Intelligent Equipment, Shandong University of Science and Technology, Taian 271000, China)

✉manzhongang@163.com; ThereliveAngel@163.com; skdljm@126.com

Abstract: In view of the fact that most current algorithms directly extract the entire face when extracting face features, and ignore local detailed features, this paper proposes BPBN, a face recognition algorithm that uses the LBP operator to divide the face image into blocks and then combines it with deep belief networks algorithm. First, the face image is divided into blocks, and the LBP is extracted from the divided images for statistics, and the generated LBP histograms are combined in a certain order to form a new feature vector. Secondly, the obtained LBP features are used as the input of Deep Belief Network (DBN), and the greedy algorithm is used to train layer by layer, and then the back propagation (BP) algorithm is used to optimize the trained deep confidence network. Finally, the trained deep belief network is used to recognize faces. Experiments on the ORL face database show that the recognition rate reaches 96.0%, significantly improved compared with the traditional principal component analysis (PCA) algorithm integrated support vector machine (SVM) method, which indicates that the method has better face recognition results.

Keywords: local two value model; face recognition; restricted Boltzmann machine; Deep Belief Network (DBF)

1 引言(Introduction)

人脸识别是计算机视觉领域的重点研究课题, 人脸识别的实际应用: 由人脸识别衍生出来的人脸检索、人脸校验等应用被广泛用于家居安防、人机交互、公安部门等重要行业。随着计算机计算速度的增加和人工智能的发展, 人脸识别算法发展迅速^[1-6], 对各行各业的影响都非常大, 能够积极

提高人们生活各方面的方便性和安全性等。

人脸识别就是对人脸图像的特征信息进行识别的技术, 在这个过程中, 对人脸图像进行信息的提取和检测, 进而来进行身份识别。人脸识别的传统方法比较常见的是基于几何结构方法, 该方法通过对人脸面部的几何信息进行提取为相应的几何结构, 对结构进行详细分析和处理, 通过几何特征

矢量来反映面部几何信息，从而进行相应的信息匹配。此传统方法难点在于如何精确地对特征信息进行匹配，要求比较高也使得该方法适应性不足^[7-9]。因此，Sirovich等人提出了PCA应用和基于主成分分析的特征脸方法^[10,11]，该方法保留了人脸图像信息之间的几何关系和部分局部特征，在整体人脸特征信息提取的情况下，还具有计算量小、描述能力强、可分性好等优点。基于局部特征方法：该方法主要包括局部二值模式(LBP)、Gabor变换^[12]。这两种局部特征方法各有优劣，局部二值模式算法能够快速计算，对光照强度的改变等感应不敏感，不足之处在于位置方向的改变以及声音方面的敏感性都不是很突出。而Gabor变换算法与局部二值模式算法有很大不同，该算法具有对光照改变的不敏感，以及更全面提取特征信息的特点，但是计算量大且耗时，产生的冗余度比较大。

本文根据上述不足，提出BPBN模型，该模型通过分块提取人脸图像的LBP特征统计生成LBP直方图，然后将得到的LBP特征作为深度置信网络(DBN)的输入，采用贪婪算法逐层进行训练，然后用反向传播(BP)算法对训练得到的深度置信网络进行优化；最后，用训练好的深度置信网络对人脸进行识别。在ORL人脸数据库上进行实验，通过实验表明该方法比PCA+SVM的方法具有更好的人脸识别效果。

2 LBP算子和深度学习理论(LBP operator and deep learning theory)

2.1 LBP算子

LBP是一种纹理描述算子。在LBP算子定义中，是通过选取一个 3×3 的区域，在这个区域中包含一个中心像素，将阈值用中心像素来进行设置，这个区域内的其他像素与阈值进行对比，进一步得到二进制码，得到的二进制码可以描述图像的局部特征。二进制码的具体步骤为：大于阈值的，用1表示，反之用0表示。按照这种步骤得出的结果，通过顺时针的方式进行放置，得出一个8位的二进制码，为了得到中心像素的LBP值，通过换为十进制的方式实现。

在大小为 3×3 的区域上的LBP模式在实际应用中受到了一定程度的限制，主要原因在于：

(1)对于 3×3 这个区域大小，在实际应用的图像中显得太小，不能更好地描述局部特征；

(2)在选取的那个区域系统上，不容易旋转不变特征的处理。

为了解决上述问题(1)，通过将 3×3 的区域扩展为任意邻域的方式，得到可变区域LBP。如果用P表示邻域点的个数，R表示圆形邻域的半径，新的LBP算子表示为 LBP_{P-R} 。

LBP的编码计算公式为：

$$LBP(x_c, y_c) = \sum_{i=0}^8 f(g_i - g_c) 2^i \quad (1)$$

$$f(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (2)$$

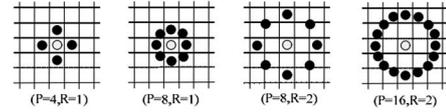


图1 不同P、R值对应的LBP算子

Fig.1 LBP operators corresponding to different P and R values

从LBP算子的定义能够得出，通过比较得出二进制模式的个数与邻域集中的采样个数密切相关，例如 LBP_{P-R} 会产生 2^P 种模式。通过增加邻域采样点个数，二进制模式的个数是几何增长的。针对这个问题，对二进制模式进行改进，提出等价模式，也称为均匀LBP。等价模式的原理就是将LBP算子所得的二进制序列进行首尾连接，当存在序列从0到1或是从1到0的变化不超过2次的情况。例如：00000000(0位转变)、01110000(2位转变)、11001111(2位转变)是等价模式，而11001001(4位转变)、01010010(6位转变)这种就不能叫作等价模式。通过优化的二进制模式，种数变为 $P(P-1)+2$ ，在一定程度上降低了特征信息的维数，经过统计，均匀LBP在整个的LBP特征中占85%—90%，所以不会对数据信息的遗漏，还会减少高频噪声的影响。

提取整个图像的LBP信息形成直方图来描述人脸特征时，只能得到图像的整体纹理和轮廓特征，而图像的具体特征细节得不到充分地描述。针对这个问题，分块思想被提出，将图像分块的基本思想是，在图像的局部区域运用LBP算子，然后把在每一个局部区域产生的直方图进行组合形成一个新的特征向量。通过这种方式，整个图像就由很多个统计直方图来进行表示。分块LBP直方图的提取过程如下：

(1)对实验数据进行分块处理，将实验图像分为 $(N \times N)$ ，如图2所示。

(2)分块处理之后对每一部分进行LBP提取，从而生成LBP直方图，如图3所示。

(3)通过(1)(2)得到多个直方图，按照得到的顺序进行排列，形成一个新的特征向量。



图2 原始人脸 2×2 分块

Fig.2 Original face 2×2 blocks

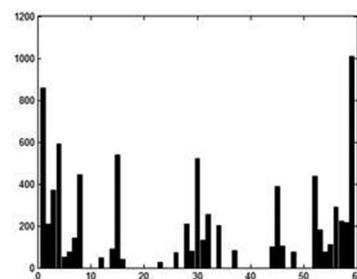


图3 分块LBP统计直方图

Fig.3 Block LBP statistical histogram

2.2 受限玻尔兹曼机

受限玻尔兹曼机(RBM)是一种特殊形式的玻尔兹曼机，从本质上说，它是一个随机的二值化对称连接的神经网络。能够把无向图的所有节点划分成两层，在一层中的节点之间是不相连的，层与层之间的节点互相连接。其结构如图4所示。

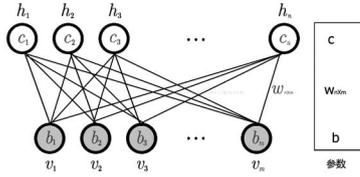


图4 RBM 结构图

Fig.4 RBM structure diagram

RBM网络结构是由两层构成的，通过向量表示层。第一层通过可见单元(显元)构成的可视层，记为 v ，这可以成为数据的输入；第二层通过可见隐单元(隐元)构成的隐藏层，记为 h ，这可以在数据的特征提取中发挥作用。第一层和第二层的节点之间的连接权重矩阵为 W ，权重为对称连接，即 $W_{ij}=W_{ji}$ 。

然后介绍训练受限玻尔兹曼机的模型：

对于 $v_i \in \{0,1\}$ 和 $h_j \in \{0,1\}$ 的RBM，它们的能量函数可以表示为：

$$E(v, h | \theta) = -\sum_{i=1}^M c_i h_i - \sum_{j=1}^N b_j v_j - \sum_{i=1}^M \sum_{j=1}^N v_j W_{ij} h_i \quad (3)$$

$$= -(v^T W + c^T) h - b^T v$$

其中， $\theta = \{W, b, c\}$ 是RBM的参数，这些参数均为实数。其中 W_{ij} 表示第 i 个显元与第 j 个隐元之间的连接权重， c_i 表示第 i 个单元的偏置， b_j 表示第 j 个单元的偏置。通过无向图能够看出， (v, h) 的联合概率密度用式子可表示为：

$$P(v, h | \theta) = \frac{1}{\sum_{v, h} e^{-E(v, h)}} e^{-E(v, h)} \quad (4)$$

对于一个实际问题，最关心的是基于RBM的观测数据 v 的分布 $P(v | \theta)$ ，也称为似然函数可表示为：

$$P(v | \theta) = \frac{\sum_h e^{-E(v, h)}}{\sum_{v, h} e^{-E(v, h)}} \quad (5)$$

RBM的特殊结构即层内无连接，当给定显元的状态时，各隐元之间是条件独立的，故第 j 个隐元的激活状态为：

$$P(h_j = 1 | v) = \frac{1}{1 + e^{-b_j - \sum_i v_i w_{ij}}} = \text{Sigmoid}(w_j^T v + b_j) \quad (6)$$

其中，Sigmoid为激活函数。

因为RBM的对称性，隐层 h 的分布 $P(h | \theta)$ 及第 i 个观测数

据的激活函数分别表示成：

$$P(h | \theta) = \frac{\sum_v e^{-E(v, h)}}{\sum_{v, h} e^{-E(v, h)}} \quad (7)$$

$$P(v_i = 1 | h) = \frac{1}{1 + e^{-c_i - \sum_j w_{ij} h_j}} = \text{Sigmoid}(w_j^T h + c_i) \quad (8)$$

假设有 N 个训练样本 $\{v_1, v_2, \dots, v_N\}$ ，则通过这些样本，采用对数似然函数得到RBM的参数 θ ，即

$$\theta^* = \arg \max_{\theta} L(\theta) = \arg \max_{\theta} \sum_{i=1}^N \log p(v_i | \theta) \quad (9)$$

用随机梯度下降法来求解 θ ，首先求得目标函数关于参数 θ 的偏导数，然后通过迭代得到目标函数的最优值：

$$L(\theta) = \sum_{i=1}^N \log p(v_i | \theta) = \sum_{i=1}^N \sum_h \log p(v_i, h | \theta)$$

$$= \sum_{i=1}^N \log \frac{\sum_h e^{-E(v_i, h; \theta)}}{\sum_v \sum_h e^{-E(v, h; \theta)}} \quad (10)$$

对 $L(\theta)$ 求偏导有：

$$\frac{\partial L(\theta)}{\partial \theta} = \sum_{i=1}^N \frac{\partial}{\partial \theta} \left(\log \left(\sum_h e^{-E(v_i, h; \theta)} \right) - \log \left(\sum_v \sum_h e^{-E(v, h; \theta)} \right) \right)$$

$$= \sum_{i=1}^N \left(\left\langle \frac{\partial (-E(v_i, h | \theta))}{\partial \theta} \right\rangle_{p(h|v_i, \theta)} - \left\langle \frac{\partial (-E(v, h | \theta))}{\partial \theta} \right\rangle_{p(h|v, \theta)} \right) \quad (11)$$

其中， $\langle \cdot \rangle_p$ 表示关于概率分布 p 的数学期望。

2.3 深度置信网络

深度置信网络(Deep Belief Network, DBN)是一种深度神经网络模型，这个模型是用概率来对数据进行输出结果的。模型在训练初期，要得到每个神经元的初始权重，运用无监督学习方法来进行计算，得到初始值之后，在这基础上运用监督学习方法来对数据进行学习，从而得到整个神经网络的权值。

DBN由多层的受限玻尔兹曼机(RBM)累叠而成，每个隐藏层都RBM的一个可视层和一个隐层构成。层跟层的神经元是全部连接的，在每层中的神经元是不联系的。在网络中，最上面两层的神经元连接是无向的，剩下的层中连接是有向的，在最下面一层对数据向量进行表示，有多少个神经元就代表该数据向量有多少维度。用联合概率分布表示输入向量 v 和隐含向量 h 之间的关系：

$$P(v, h^2, \dots, h^l) = \left(\prod_{k=2}^{l-2} p(h^k | h^{k+2}) \right) P(h^{l-2}, h^l) \quad (12)$$

在深度置信网络工作时，第一步进行预训练，得到初始权重；第二步在网络后期学习中，进行参数的调整。

网络工作的具体流程为：首先在可见层输入数据，产生一个向量 v 然后传递给隐层。隐层由多个RBM组成，首先由最

底层的RBM的隐藏单元对可视单元重新构造，然后对下一层的RBM的隐藏单元进行传递，通过无监督的贪婪训练方法逐层训练获得整个DBN的模型，然后对模型进行参数的调整，从而优化完成，得到最终的神经网络模型。可以把具体流程分为两个部分：预训练和参数调整。

3 实验结果及分析(Experimental results and analysis)

3.1 BPBN模型

BPBN模型主要功能包括了局部二值模式和置信网络，模型先将图片分块，之后对分块进行特征提取，统计分块直方图，之后组成全局特征向量，并将其作为模型输入数据放入置信网络进行参数训练，之后根据后向传播算法对参数进行误差调整，后通过输出层分类，模型框架如图5所示。

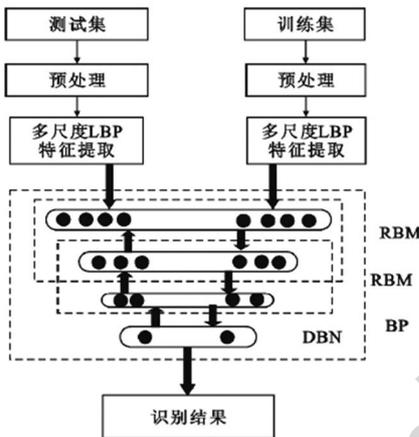


图5 人脸识别模型图

Fig.5 Face recognition model

算法步骤如下：

第一步：将训练图片按照4×4规格分块，后对分块图片进行特征提取并统计分块直方图，然后将组成最终LBP纹理特征。

第二步：将最终提取特征作为置信网络的输入数据，训练模型，在采用BP算法调整模型参数，调节方式采用fmincg函数，直到模型收敛。

第三步：将测试图片进行第一步和第二步的操作。

第四步：使用softmax对测试图片分类，并计算模型测试准确率。

3.2 实验环境和结果分析

本文采用的实验数据为ORL Faces Database，其包含400张图片，图片分别来自40人的人脸图，每人10张，分别包含不同表情。本文将数据集以7:3的比例随机划分训练集和测试集，即一个人的10张图片中，有七张会被作为训练数据，另外三张会被作为测试数据。实验环境为Win10，采用的处理器AMD A10-5750M APU with Radeon(tm) HD Graphics

2.50 GHz，内存为12.00 GB。

首先使用BPBN对数据集进行特征提取，将数据分块后对分块进行提取，并最终获得完整人脸特征，其结果如图6所示。

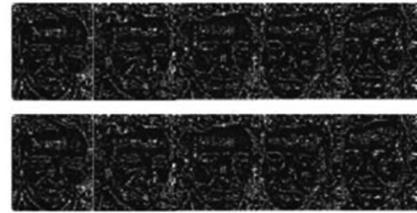


图6 特征提取

Fig.6 Feature extraction

本文实验中，数据分块采用4×4标准，BPBN模型中隐层数量为2，并采用不同规模的模型分别进行三次实验，分别对应隐层规模为150—150、200—150和250—150，实验总迭代次数设为3000，模型学习率设为 1×10^{-2} 。

BPBN模型在不同模型规模下的准确率和时间损耗由表1呈现，为了对比BPBN方法的实用性和有效性，将传统的主成分分析方法(PCA)和支持向量机(SVM)结合的模型与本文提出的BPBN方法进行实验对比，对比结果由表2呈现。通过实验结果和对比结果可以看出：BPBN模型的准确率随着隐层规模复杂性的变大而增高，相应模型训练时间会随着隐层规模复杂性的变大而增加，同时在隐层模型复杂度达到一定程度后，模型准确率不会再随着隐层规模复杂性的变大而增高，甚至由于过拟合的原因有轻微下降趋势。

PCA-SVM会将图片光照、表情等因素聚集在PCA的主向量上，从而使得准确率降低，而BPBN的特征提取更加精确，所以BPBN模型较PCA-SVM方法拥有更高的准确性和更好的鲁棒性。

表1 BPBN实验结果

Tab.1 The experimental results of BPBN

使用方法	隐层规模	准确率(%)	时间损耗(s)
BPBN	150-150	92.1	145.46
	200-150	94.2	294.97
	250-150	96.0	341.19
	300-200	95.7	633.51

表2 与其他方法比较

Tab.2 Comparison with other methods

使用方法	准确率(%)
BPBN	96.0
PCA-SVM	88.5

4 结论(Conclusion)

本文提出一种将人脸图像进行分块局部运用LBP算子，

(下转第12页)