

# 一种特征加权融合人脸识别方法

孙劲光<sup>1,2</sup>, 孟凡宇<sup>2</sup>

(1. 辽宁工程技术大学 电子与信息工程学院, 辽宁 葫芦岛 125000; 2. 辽宁省数字化矿山装备工程技术研究中心, 辽宁 阜新 123000)

**摘 要:**针对传统人脸识别算法在非限制条件下识别准确率不高的问题,提出了一种特征加权融合人脸识别方法(DLWF<sup>+</sup>)。根据人脸面部左眼、右眼、鼻子、嘴、下巴等 5 个器官位置,将人脸图像划分成 5 个局部采样区域;将得到的 5 个局部采样区域和整幅人脸图像分别输入到对应的神经网络中进行网络权值调整,完成子网络的构建;利用 softmax 回归求出 6 个相似度向量并组成相似度矩阵与权向量相乘得出最终的识别结果。经 ORL 和 WFL 人脸库上进行实验验证,识别准确率分别达到 97% 和 91.63%。实验结果表明:该算法能够有效提高人脸识别能力,与传统识别算法相比在限制条件和非限制条件下都具有较高的识别准确率。

**关键词:**人脸识别;非限制条件;深度自信网络;局部特征;特征融合;全局特征  
**中图分类号:** TN911.73 **文献标志码:**A **文章编号:**1673-4785(2015)06-0912-09

中文引用格式:孙劲光,孟凡宇. 一种特征加权融合人脸识别方法[J]. 智能系统学报, 2015, 10(6): 912-920.  
英文引用格式:SUN Jinguang, MENG Fanyu. Face recognition by weighted fusion of facial features[J]. CAAI Transactions on Intelligent Systems, 2015, 10(6): 912-920.

## Face recognition by weighted fusion of facial features

SUN Jinguang<sup>1, 2</sup>, MENG Fanyu<sup>2</sup>

(1. School of Electronics and Information Engineering, Liaoning Technical University, Huludao 125000, China; 2. LiaoNing Digital Mining Equipment Engineering Technology Research Center, Fuxin 123000, China)

**Abstract:**The accuracy of face recognition is low under unconstrained conditions. To solve this problem, we propose a new method based on deep learning and the weighted fusion of facial features. First, we divide facial feature points into five regions using an active shape model and then sample different facial components corresponding to those facial feature points. A corresponding deep belief network (DBN) was then trained based on these regional samples to obtain optimal network parameters. The five regional sampling regions and entire facial image obtained were then inputted into a corresponding neural network to adjust the network weight and complete the construction of sub-networks. Finally, using softmax regression, we obtained six similarity vectors of different components. These six similarity vectors comprise a similarity matrix, which is then multiplied by the weight vector to derive the final recognition result. Recognition accuracy was 97% and 91.63% on the ORL and WFL face databases, respectively. Compared with traditional recognition algorithms such as SVM, DBN, PCA, and FIP+LDA, recognition rates for both databases were improved in both constrained and unconstrained conditions. On the basis of these experimental results, we conclude that the proposed algorithm demonstrates high efficiency in face recognition.

**Keywords:** face recognition; unconstrained condition; deep belief networks; local feature; feature fusion; overall feature

研究热点。人脸图像像素矩阵较大同时特征维数也较高,进行识别难度较大,因此设计出高效的特征提取器成为人脸识别的关键。其中最典型方法为主成分分析法(principal components analysis,PCA)<sup>[1]</sup>,其思想首先,把一副图像按行或列拉直成一个高维列向量,然后,求取整个训练样本集高维列向量的协方差矩阵,求此协方差矩阵的特征向量和特征值得到投影变换矩阵,最后,把人脸图像高维特征通过投影变换矩阵映射到低维空间中,得到人脸的低维表示特征;在人脸识别研究中,利用 PCA 方法对高维特征进行降维处理,获得了良好的效果,是最为成功的线性鉴别分析方法之一。但实际的应用中,PCA 方法进行线性鉴别分析通常会产生小样本问题(即样本个数远小于特征维数);被识别人脸图像是在非限制条件下提取时,应用 PCA 算法提取的特征并不适合作为分类的特征。

人们利用深度神经网络进行人脸识别主要分为 2 类:1)利用深度神经网络特征提取后再应用其他分类器进行识别<sup>[2-6]</sup>;2)直接在深度神经网络增加分类层进行识别<sup>[7-8]</sup>。汤晓鸥<sup>[9]</sup>等利用卷积神经网络进行人脸验证取得了良好的效果,在 LFW 人脸库上的识别率已经达到 99.15%;解决了人脸识别问题中的二分类问题。

为了解决传统人脸识别方法在复杂背景下和人脸多姿态条件下进行识别时,识别率不高的问题,本文提出了一种特征加权融合人脸识别方法。本文的主要贡献是通过将人脸面部进行区域划分提取得到了人脸面部的局部特征,并结合整幅人脸的全局特征进行相似度矩阵的求取,通过各部分的贡献度不同进行加权得出最终判别结果,提出了基于相似度矩阵加权的深度网络模型(deep learning and weighted fusion,DLWF+)人脸识别方法。

## 1 基于相似度加权评分深度网络模型

### 1.1 问题描述

利用传统深度信念网络<sup>[10-11]</sup>进行人脸识别,输入的是整张人脸的像素级特征,忽略了人脸图像局部结构特征,网络难以学习得到人脸局部特征<sup>[12]</sup>;在自然条件下,所获得的人脸图像通常受到背景、光照等外界因素的影响,很难确定哪些面部器官在识别过程中所起到的比重大,哪些器官起的比重小,使得在实际研究过程中对相似度加权的权值确定变得复杂。因此本文重点解决提取人脸局部特征的深度学习模型及多特征加权识别方法权值确定 2 个问题。

### 1.2 基本思想

为建立人脸面部局部特征深度学习模型,首先利用人脸面部的主要特征点将人脸分成若干子区域;然后将这些小的子区域和整幅人脸经归一化处理,输入到对应的 DBN 网络进行训练,得到各个子区域和整幅人脸的相似度概率向量;最后将左眼、右眼、鼻子、嘴、下巴等器官和整个面部的相似度向量加权后进行分类。DLWF+分为以下 4 个主要阶段:

#### 1) 人脸面部区域划分

利用 haar 方法进行人脸检测,提取出人脸面部主要区域,并根据人脸面部主要器官左眼、右眼、鼻子、嘴巴、下巴等进行子区域的划分。

#### 2) 网络模型的构建并确定权值

利用上一步求取得到的各个区域和整幅人脸图像分别输入到对应的子网络中,进行权值的调整,完成网络模型的构建。对各个子区域分别进行识别得出识别率作为各自的权重。

#### 3) 求取面部隶属度矩阵

通过将测试集不同区域的采样块分别输入到对应的子网络中,取出个区域的相似度向量,并利用求得的各个相似度向量组成相似度矩阵。

#### 4) 加权求和得出识别结果。

将相似度矩阵进行加权得出测试图像对应于每个类别的综合评分,评分最高的类别即作为最终的识别结果。

### 1.3 基本原理

#### 1.3.1 主动形状模型

主动形状模型是由 Cootes 等<sup>[13-15]</sup>在 1995 年提出。该模型主要由全局形状模型和局部纹理模型构成。

#### 1) 全局形状模型

ASM 利用点分布模型表示人脸形状  $S_i$ ,  $S_i = [x_1 \ y_1 \ x_2 \ y_2 \ \cdots \ x_n \ y_n]^T$  ( $n$  为特征点数目),其中具有  $N$  个训练样本的训练集表示为  $\Omega = \{S_1, S_2, \cdots, S_N\}$ 。

将训练样本对齐完成后,采用主成分分析(PCA)方法建立形状模型,过程如下:

①计算对齐完成后训练样本平均形状:

$$\bar{S} = \frac{1}{N} \sum_{i=1}^n S_i \tag{1}$$

②计算形状向量的协方差矩阵:

$$C = \frac{1}{N-1} \sum_{i=1}^N (S_i - \bar{S})(S_i - \bar{S})^T \tag{2}$$

③对  $C$  进行特征分解,得到特征值  $\lambda_k$  ( $\lambda_k \geq \lambda_{k+1}$ ,  $\lambda_k \neq 0, k=1, 2, \cdots, 2n$ ) 和特征向量  $P_k$ 。

④提取前  $t$  个特征值  $[\lambda_1 \ \lambda_2 \ \cdots \ \lambda_t]$ , 令

$$\frac{\sum_{i=1}^t \lambda_i}{\sum_{i=1}^{2n} \lambda_i} \geq \eta \quad (3)$$

式中:  $\eta$  为所选特征占总特征比例, 一般为 95% ~ 98%。求得特征向量  $\mathbf{P} = [p_1 \ p_2 \ \cdots \ p_t]$ , 最终求得全局形状模型:

$$\mathbf{S} = \bar{\mathbf{S}} + \mathbf{P}\mathbf{b} \quad (4)$$

式中:  $\mathbf{b}$  为主分量参数, 控制前  $t$  个模式系数, 且  $b_i$  取值:

$$-3\sqrt{\lambda_i} \leq b_i \leq 3\sqrt{\lambda_i} \quad (5)$$

全局形状模型的作用, 在搜索更新过程中, 通过改变形状参数, 找到最佳的匹配模型, 将形状变化约束在模型允许范围内。

## 2) 局部纹理模型

局部纹理模型构建过程如下:

①以第  $i$  个样本第  $j$  个标定点为中心, 与相邻 2 点连线垂直方向一定长度范围 (也称 Profile 邻域), 内外各取  $k$  个点的灰度值作为该点灰度信息:

$$\mathbf{g}_{ij} = [g_{ij,1} \ g_{ij,2} \ \cdots \ g_{ij,2k+1}] \quad (6)$$

②求  $\mathbf{g}_{ij}$  梯度  $\mathbf{dg}_{ij}$ :

$$\mathbf{dg}_{ij} = [g_{ij,2} - g_{ij,1} \ \cdots \ g_{ij,2k+1} - g_{ij,2k}] \quad (7)$$

对  $\mathbf{dg}_{ij}$  归一化处理得

$$G_{ij} = \frac{1}{\sum_{l=1}^{2k} |\mathbf{dg}_{ij,l}|} \mathbf{dg}_{ij} \quad (8)$$

式中:  $\mathbf{dg}_{ij,l} = g_{ij,(l+1)} - g_{ij,l}$ 。

对训练样本集所有图像的第  $j$  个特征点进行相同操作, 求得第  $j$  个特征点的  $N$  个局部纹理  $G_{1j}, G_{2j}, \cdots, G_{Nj}$ 。

③得到标记点  $j$  对应的平均纹理  $\bar{\mathbf{G}}_j$  和协方差矩阵  $\mathbf{C}_{gj}$ , 即为该点的局部纹理模型。

$$\bar{\mathbf{G}}_j = \frac{1}{N} \sum_{i=1}^N G_{ij} \quad (9)$$

$$\mathbf{C}_{gj} = \frac{1}{N} \sum_{i=1}^N (G_{ij} - \bar{\mathbf{G}}_j)(G_{ij} - \bar{\mathbf{G}}_j)^T \quad (10)$$

局部纹理模型作用是在搜索过程中, 确定每个特征点最佳候选点。其中, 马氏距离  $d(\mathbf{G}'_j)$  最小的点, 就是该特征点的最佳候选点。

$$d(\mathbf{G}'_j) = (\mathbf{G}'_j - \bar{\mathbf{G}}_j)^T \mathbf{C}_{gj}^{-1} (\mathbf{G}'_j - \bar{\mathbf{G}}_j) \quad (11)$$

式中:  $\mathbf{G}'_j$  是在  $j$  点附近采样得到, 经过归一化处理纹理向量。

利用上述模型方法提取到的 77 个面部特征点

如图 1 所示。

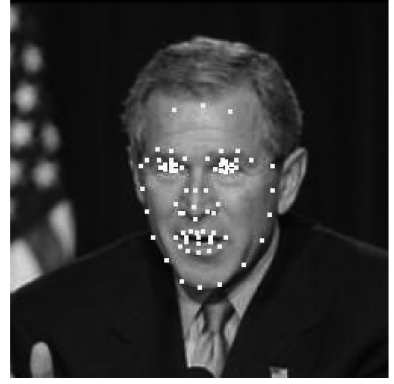


图 1 人脸面部特征点

Fig.1 Facial feature points

## 1.3.2 深度信念网络 (DBN)

DBN 模拟人脑的组织结构, 由低级到高级地提取数据的抽象特征<sup>[16]</sup>。DBN 由多个 RBM 堆叠而成, RBM 被当做是一个动力学系统在给定一组状态  $(v, h)$  下的能量定义为

$$E(v, h/\theta) = - \sum_{i=1}^{n_v} a_i v_i - \sum_{j=1}^{n_h} b_j h_j - \sum_{i=1}^{n_v} \sum_{j=1}^{n_h} h_j w_{ij} v_i \quad (12)$$

式中:  $n_v, n_h$  分别为可见层和隐含层中神经元数目。  $v_i$  为可见层第  $i$  个神经元状态,  $h_j$  为隐含层第  $j$  个神经元状态,  $a_i$  为可见层中第  $i$  个神经元偏置,  $b_j$  为隐含层第  $j$  个神经元偏置,  $w_{ij}$  为可见层第  $i$  个神经元与隐含层第  $j$  个神经元之间连接权重,  $\theta$  表示需要训练调整模型参数  $\theta = \{\mathbf{a}, \mathbf{b}, \mathbf{w}\}$ 。

用向量形式表示为

$$E(v, h/\theta) = -\mathbf{a}^T \mathbf{v} - \mathbf{b}^T \mathbf{h} - \mathbf{h}^T \mathbf{w} \mathbf{v} \quad (13)$$

当给定模型参数  $\theta = \{\mathbf{a}, \mathbf{b}, \mathbf{w}\}$  时, 利用式 (12) 定义的能量函数, 可以求出可视层  $v$  和隐含层  $h$  联合概率分布:

$$p(v, h/\theta) = \frac{1}{Z_\theta} e^{-E(v, h/\theta)} \quad (14)$$

$$Z_\theta = \sum_{v, h} e^{-E(v, h/\theta)} \quad (15)$$

式中:  $Z(\theta)$  为归一化因子。

当给定可视层节点时, 隐含层节点  $h_i$  为 0 或 1 的条件概率分布为

$$p(h_i = 1/v) = \sigma(a_i + \sum_j w_{ij} v_j) \quad (16)$$

$$p(h_i = 0/v) = 1 - p(h_j = 1/v) \quad (17)$$

当给定隐含层节点时, 可视层节点  $v_i$  为 0 或 1 的条件概率分布为

$$p(v_i = 1/h) = \sigma(b_i + \sum_j w_{ij} h_j) \quad (18)$$

$$p(v_i = 0/h) = 1 - p(v_i = 1/h) \tag{19}$$

第  $k$  和  $k+1$  层隐含单元满足:

$$p(h^{(k)} | h^{(k+1)}) = \prod_i p(h_i^{(k)} | h^{(k+1)}) \tag{20}$$

$$p(h_i^{(k)} = 1 | h^{(k+1)}) = \sigma(b_i^{(k)} + \sum_j W_{ij}^{(k)} h_j^{(k+1)}) \tag{21}$$

式中: $\sigma(x) = 1/(1+e^{-x})$ ,  $b_i^{(k)}$  为第  $k$  层的偏置,  $W_{ij}^{(k)}$  为第  $k$  和  $k+1$  层之间权值,在 DBN 中将  $p(h^{(l-1)}, h^{(l)})$  看作为为一个 RBM 模型。结构如图 2。

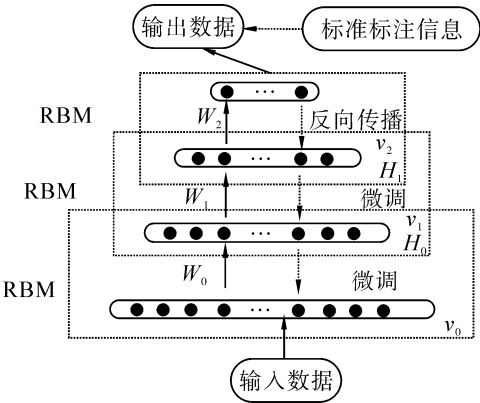


图 2 DBN 结构模型  
Fig.2 DBN structure model

1.3.3 Softmax 回归

Softmax 回归将只能够解决二分类问题的 Logistic 回归扩展至能够解决多分类问题。

假定 Softmax 回归模型训练样本来自  $k$  个不同的类别,共有  $m$  个,那么这  $m$  个训练样本组成的训练样本集为  $\{(x^{(1)}, y^{(1)}), \dots, (x^{(m)}, y^{(m)})\}$ , 则 Softmax 回归的假设函数为

$$h_{\theta}(x^{(i)}) = \begin{bmatrix} p(y^{(i)} = 1 | x^{(i)}; \theta) \\ p(y^{(i)} = 2 | x^{(i)}; \theta) \\ \vdots \\ p(y^{(i)} = k | x^{(i)}; \theta) \end{bmatrix} = \frac{1}{\sum_{j=1}^k e^{\theta_j^T x^{(i)}}} \begin{bmatrix} e^{\theta_1^T x^{(i)}} \\ e^{\theta_2^T x^{(i)}} \\ \vdots \\ e^{\theta_k^T x^{(i)}} \end{bmatrix} \tag{22}$$

式中:假设向量  $h_{\theta}(x^{(i)})$  的每个元素  $p(y^{(i)} = j | x^{(i)}; \theta)$  表示样本  $x^{(i)}$  属于类别  $j$  的概率。  $\theta_1, \theta_2, \dots, \theta_k$  为模型参数,其中  $\theta_i \in R^{n+1}$ , 将这些向量写成矩阵:

$$\theta = \begin{bmatrix} -\theta_1^T & - \\ -\theta_2^T & - \\ -\theta_3^T & - \end{bmatrix} \tag{23}$$

则模型代价函数定义:

$$J(\theta) = -\frac{1}{m} \left[ \sum_{i=1}^m \sum_{j=1}^k 1\{y^{(i)} = j\} \log \frac{e^{\theta_j^T x^{(i)}}}{\sum_{l=1}^k e^{\theta_l^T x^{(i)}}} \right] + \frac{\lambda}{2} \sum_{i=1}^k \sum_{j=0}^n \theta_{ij}^2 \tag{24}$$

式中: $1\{\cdot\}$  表示指示函数,花括号中表达式为真,那么函数值为 1,否则函数值为 0。加号后面部分为了防止模型过拟合所加的权衰减项。其中,  $\lambda$  为模型参数,在实验部分确定。分析可知无法直接求取可以使  $J(\theta)$  最小的  $\theta$  的解析解,通过迭代优化算法求解。计算代价函数梯度公式为

$$\nabla_{\theta_j} = -\frac{1}{m} \sum_{i=1}^m [x^{(i)} (1\{y^{(i)} = j\} - p(y^{(i)} = j | x^{(i)}; \theta))] + \lambda \theta_j \tag{25}$$

1.4 本文研究方法 (DLFW<sup>+</sup>)

本文研究重点内容如下:

1) 面部区域划分

利用主动形状模型确定人脸面部主要特征点,对于每张人脸面部图像能得到一个特征点位置向量:

$$x_i = (x_{i,0}, y_{i,0}, x_{i,1}, y_{i,1}, \dots, x_{i,(n-1)}, y_{i,(n-1)})$$

①左眼子区域划分

左眼区域采样框确定,根据左眼眼角确定采样框长,眉毛轮廓线最高点与眼睛轮廓线最低点确定采样框宽。为了获取左眼区域更多细节特征,长和宽同时向外延伸  $L_{\text{leye}}$  个像素。左眼区域长宽计算公式如下:

$$w_{\text{leye}} = x_{30} - x_{34} + L_{\text{leye}} \tag{27}$$

$$h_{\text{leye}} = y_{36} - y_{17} + L_{\text{leye}} \tag{28}$$

式中: $w_{\text{leye}}$  为左眼区域宽,  $h_{\text{leye}}$  为左眼区域长,  $L_{\text{leye}}$  为从原始图像矩阵  $I$  上截取的左眼区域矩阵。

②右眼子区域划分

右眼区域采样框确定,根据右眼眼角确定采样框长,眉毛轮廓线最高点及眼睛轮廓线最低点确定采样框宽。为获取右眼区域更多细节特征,长和宽同时向外延伸  $L_{\text{reye}}$  个像素。此区域长宽计算公式为

$$w_{\text{reye}} = x_{44} - x_{40} + L_{\text{reye}} \tag{29}$$

$$h_{\text{reye}} = y_{46} - y_{24} + L_{\text{reye}} \tag{30}$$

式中: $w_{\text{reye}}$  为右眼区域宽,  $h_{\text{reye}}$  为右眼区域长。

③鼻子子区域划分

鼻子区域采样框确定,根据鼻子轮廓线最低点及眉毛轮廓线最前端特征点确定鼻子区域采样框长。根据鼻子轮廓线最左边特征点及最右边特征点确定采样框的宽。长和宽同时向外延伸  $L_{\text{nose}}$  个像素。鼻子区域长宽计算公式如下:

$$w_{\text{nose}} = x_{54} - x_{58} + L_{\text{nose}} \tag{31}$$



$$h_{nose} = y_{56} - y_{22} + L_{nose} \tag{32}$$

式中: $w_{nose}$  鼻子区域宽, $h_{nose}$  为鼻子区域长。

④ 嘴部区域划分

嘴部区域采样框确定,根据嘴部特征点的左右嘴角确定采样框长。根据嘴部轮廓最上面特征点及最下面特征点位置确定采样框宽。并且长和宽同时向外延伸  $L_{mouth}$  个像素。嘴部区域长宽计算公式为:

$$w_{mouth} = x_{65} - x_{59} + L_{mouth} \tag{33}$$

$$h_{mouth} = y_{74} - y_{62} + L_{mouth} \tag{34}$$

式中: $w_{mouth}$  为嘴部区域宽, $h_{mouth}$  为嘴部区域长。

⑤ 下巴子区域划分

下巴区域采样框确定,根据与嘴角连线平齐的下巴轮廓线上 2 个特征点确定下巴采样框长,根据这 2 个特征点及下巴轮廓线最低点确定下巴采样框宽。长和宽同时向外延伸  $L_{chin}$  个像素。下巴区域长宽计算公式如下:

$$w_{chin} = x_9 - x_3 + L_{chin} \tag{35}$$

$$h_{chin} = y_6 - y_{74} + L_{chin} \tag{36}$$

式中: $w_{chin}$  为下巴区域宽, $h_{chin}$  为下巴区域长。

将一个  $m \times n$  的图像矩阵  $I$  按上述长宽截取 5 个子区域并和整幅人脸图像组成训练集即

$$\{I_{leye}, I_{reye}, I_{nose}, I_{mouth}, I_{chin}, I_{face}\} \in I \tag{37}$$

式中: $I_{leye}, I_{reye}, I_{nose}, I_{mouth}, I_{chin}$  表示截取得到的 5 个子区域像素矩阵, $I_{face}$  表示整幅人脸图像像素矩阵。

应用双线性内插法将上述区域归一化到固定大小,经过上述步骤截取得到的 5 个不同区域的采样块与整幅人脸图像一起构成了网络的训练样本集,如图 3 所示。



图 3 子区域划分  
Fig.3 Sub region division

2) 构造区域网络模型

由于截取得到的人脸面部 5 个子区域及整幅人脸图像大小不同,因此输入层结点个数也不相同,需要分别为 5 个人脸面部区域及整幅人脸图像构建相应的 DBN 网络,网络结构如图 4 所示。

隶属度求取网络是一个 5 层 DBN 网络,由 1 个输入层、1 个输出层和 3 个隐含层构成。如表 1 所示。

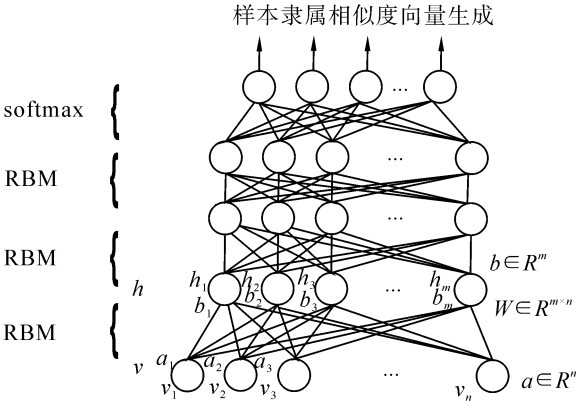


图 4 子区域网络结构  
Fig.4 Sub region network structure

表 1 不同区域网络结点数

Table1 Different area network node number						
结点数目	左眼	右眼	鼻子	嘴巴	下巴	人脸
输入层	$w_{leye} \times$	$w_{reye} \times$	$w_{nose} \times$	$w_{mouth} \times$	$w_{chin} \times$	$W_{face} \times$
	$h_{leye}$	$h_{reye}$	$h_{nose}$	$h_{mouth}$	$h_{chin}$	$h_{face}$
隐含层 1	$\beta_{leye}^1 \times$	$\beta_{reye}^1 \times$	$\beta_{nose}^1 \times$	$\beta_{mouth}^1 \times$	$\beta_{chin}^1 \times$	$\beta_{face}^1 \times$
	$L^1$	$L^1$	$L^1$	$L^1$	$L^1$	$L^1$
隐含层 2	$\beta_{leye}^2 \times$	$\beta_{reye}^2 \times$	$\beta_{nose}^2 \times$	$\beta_{mouth}^2 \times$	$\beta_{chin}^2 \times$	$\beta_{face}^2 \times$
	$L^2$	$L^2$	$L^2$	$L^2$	$L^2$	$L^2$
隐含层 3	$\beta_{leye}^3 \times$	$\beta_{reye}^3 \times$	$\beta_{nose}^3 \times$	$\beta_{mouth}^3 \times$	$\beta_{chin}^3 \times$	$\beta_{face}^3 \times$
	$L^3$	$L^3$	$L^3$	$L^3$	$L^3$	$L^3$

其中, $\beta$  为各层神经元的缩放系数,用来决定特征的降维尺度,同时也决定了特征损耗的多少。 $L^i$  为第  $i$  层的神经元个数。

输出层神经元个数为将要识别的类别数。

将提取得到的 5 个区域采样块及整个人脸分别输入到对应的 DBN 网络,经过自下而上逐层提取得到更抽象特征,输出层利用 Softmax 分类器求取各区域特征属于各类别的隶属度。求得的隶属度向量为  $P_i(P_{1,i} \ P_{2,i} \ \cdots \ P_{n,i})^T$ 。将各区域求出的隶属度向量进行组合便得到隶属度矩阵  $P$ 。

$$P = \begin{bmatrix} p_{1,1} & p_{1,2} & p_{1,3} & p_{1,4} & p_{1,5} & p_{1,6} \\ p_{2,1} & p_{2,2} & p_{2,3} & p_{2,4} & p_{2,5} & p_{2,6} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{n,1} & p_{n,2} & p_{n,3} & p_{n,4} & p_{n,5} & p_{n,6} \end{bmatrix} \tag{38}$$

式中:6 列代表 6 个区域,行数  $n$  代表  $n$  个类别。

3) 隶属度加权融合方法

利用在上一步骤中构建的 DBN 网络,求出各区域的正确识别率,识别率越高说明进行人脸识别时此区域起到的作用越大,识别率越低说明此区域起到的作用越小。将此识别率作为各区域权值系数确定依据。具体计算公式为

$$\omega_i = \frac{p_i}{p_1 + p_2 + p_3 + p_4 + p_5 + p_6} \quad (39)$$

使得  $\omega_1 + \omega_2 + \omega_3 + \omega_4 + \omega_5 + \omega_6 = 1$ , 对人脸图像进行识别时, 利用求得的权值矩阵  $P$  与隶属度向量  $\omega$  相乘, 计算公式如下:

$$\begin{bmatrix} p_{1,1} & p_{1,2} & p_{1,3} & p_{1,4} & p_{1,5} & p_{1,6} \\ p_{2,1} & p_{2,2} & p_{2,3} & p_{2,4} & p_{2,5} & p_{2,6} \\ p_{3,1} & p_{3,2} & p_{3,3} & p_{3,4} & p_{3,5} & p_{3,6} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ p_{n,1} & p_{n,2} & p_{n,3} & p_{n,4} & p_{n,5} & p_{n,6} \end{bmatrix} \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \\ \omega_4 \\ \omega_5 \\ \omega_6 \end{bmatrix} = \begin{bmatrix} q_1 \\ q_2 \\ q_3 \\ \vdots \\ q_n \end{bmatrix} \quad (40)$$

从而求得此待识别人脸图像隶属于各类别的综合评分向量  $Q = [q_1 \quad q_2 \quad \cdots \quad q_n]$ 。其中评分最高的类别, 即为最终识别结果。

1.5 算法描述

特征加权融合的人脸识别方法具体算法实现过程如下:

算法 1 训练算法。

1) 利用 ASM 模型标注 LFW 人脸库中全部 13233 幅图像, 人脸面部特征点向量  $x_i = (x_{i,0}, y_{i,0}, x_{i,1}, y_{i,1}, \cdots, x_{i,(n-1)}, y_{i,(n-1)})$ 。

2) 根据面部特征点向量  $x_i$ , 确定 5 个采样区域位置, 并计算 5 采样框的长和宽  $w_{leye}, h_{leye}, w_{reye}, h_{reye}, w_{nose}, h_{nose}, w_{mouth}, h_{mouth}, w_{chin}, h_{chin}$ 。

3) 在灰度化后原图像  $I$  中根据上步求出的采样框大小, 进行采样得到 5 个子区域  $I_{leye}, I_{reye}, I_{nose}, I_{mouth}, I_{chin} \in I$ , 以眼睛区域为例:

$$I_{leye} = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1,w_{leye}} \\ x_{21} & x_{22} & \cdots & x_{2,w_{leye}} \\ \vdots & \vdots & & \vdots \\ x_{h_{leye},1} & x_{h_{leye},2} & \cdots & x_{h_{leye},w_{leye}} \end{bmatrix} \quad (41)$$

式中:  $x_{ij}$  表示采样后子区域内像素值。

4) 将 5 个子区域及整幅人脸图像利用双线性内插法归一化到规定大小, 并拉直成列向量:

$$I_{leye} = [x_{11} \quad x_{12} \quad \cdots \quad x_{h_{leye},w_{leye}}]^T \quad (42)$$

5) 对 5 个子区域及整幅人脸图像进行灰度化、灰度拉伸等预处理操作, 形成初调训练样本集, 从将要用于测试的类别中选取一部分作为精调训练样本集。

6) 利用得到的 6 个初调训练样本集分别对 6 个 DBN 网络应用对比散度方法<sup>[17]</sup>进行预训练调整各层权值矩阵  $W_i$  及偏置矩阵  $B_i$ 。式中:  $i$  表示第  $i$  层。

7) 应用反向传播方法, 在精调训练样本集上对

DBN 网络各层权值矩阵  $W_i$  及偏置矩阵  $B_i$  精调。

8) 训练完成后, 将各层权值矩阵  $W_i$  及偏置矩阵  $B_i$  保存, 供识别时使用。

算法 2 分类算法

1) 对于一张给定的待识别人脸图像, 利用 ASM 模型提取出人脸面部特征点向量:

$$x = [x_0 \quad y_0 \quad x_1 \quad y_1 \quad \cdots \quad x_{(n-1)} \quad y_{(n-1)}]$$

2) 根据面部特征点向量  $x$ , 确定 5 个采样区域位置, 并计算 5 采样框的长和宽  $w_{leye}, h_{leye}, w_{reye}, h_{reye}, w_{nose}, h_{nose}, w_{mouth}, h_{mouth}, w_{chin}, h_{chin}$ 。

3) 在原图像  $I$  中根据上步求出的采样框大小, 进行采样得到 5 个子区域  $I_{leye}, I_{reye}, I_{nose}, I_{mouth}, I_{chin} \in I$ 。

4) 将 5 个子区域及整幅人脸图像利用双线性内插法归一化到规定大小。而后进行灰度化、灰度拉伸等预处理操作。

5) 将 4) 中得到的子区域拉直成列向量, 输入到由训练完成的权值矩阵  $W_i$  及偏置矩阵  $B_i$  构造的 DBN 网络中, 求出隶属度矩阵  $P$ 。

6) 将隶属度矩阵  $P$  与权值向量  $\omega(\omega_1, \omega_2, \omega_3, \omega_4, \omega_5, \omega_6)$  相乘求出此待识别人脸图像隶属于各个类别的综合评分向量  $Q = [q_1 \quad q_2 \quad \cdots \quad q_n]$ 。

7) 找出评分向量  $Q$  中获得评分最高的类, 作为分类结果。

2 实验结果及分析

本文算法分别在 LFW (Labeled Faces in the Wild Database) 人脸库和 ORL (Olivetti research laboratory) 人脸库上进行仿真实验, 证明了本文算法的有效性。在 LFW 人脸库实验目的是为了验证本文算法在非限制条件下识别的有效性, 同时, 由于 LFW 人脸库中有 13 233 幅人脸图像为 DBN 预训练提供了充足的训练样本, 保证 DBN 参数能够趋近全局最优<sup>[18]</sup>。为了进一步评估本文算法通用性, 在 ORL 人脸库上同样进行了比较和测试。进行本文实验硬件环境为: 3.50 GHz 的 Intel (R) Core (TM) i5-4690 CPU, 8.00 GB 内存。

2.1 LFW 人脸库上的实验

LFW 人脸库包含 5 749 人的 13 233 幅人脸图像。其中, 1 680 人包含 2 幅及以上人脸图像, 4 069 人只有一幅人脸图像。图像分辨率为  $250 \times 250$ , 以 JPEG 格式存储, 绝大多数为彩色图像, 少数为灰度图像。LFW 主要用于在非限制条件下人脸识别研究, 以成为学术界和工业界识别性能评价基准。LFW 人脸库能充分表现自然条件下人脸图像的变

化,如光照、姿态、表情、背景、性别等。

本文选取人脸库中包含图像数目大于或等于 25 幅的人作为实验对象,包括 29 人共 2458 幅图像。每人随机选取 5 幅人脸图像作为训练样本,其余图像作为测试样本。将这些图像首先根据人脸面部五官的位置进行区域的划分,然后归一化到固定大小(下巴区域 88×33、嘴部区域 51×34、左眼和右眼区域 44×31、鼻子区域 32×46),由此生成训练、测试样本。

1) 隐含层单元设置压缩比不同时对比实验

DBN 的分类能力主要取决于自下向上、多层提取得到特征的有效性。为了能够获得更多细节特征,5 个区域长和宽分别向外扩大了  $L_{\text{eye}}$ 、 $L_{\text{eye}}$ 、 $L_{\text{nose}}$ 、 $L_{\text{mouth}}$ 、 $L_{\text{chin}}$ ,它们的取值为各区域长的 0.2 倍。在构建 DBN 时,确定各隐含层神经元个数至关重要。因为 LFW 人脸库中图像都是在没有限制的自然条件下采集的,所以人脸图像特征较为复杂,并且含有较多噪声,为了能够有效提取人脸图像抽象特征,首先应该讨论网络构建问题中,不同隐含层神经元个数对网络识别性能的影响。实验结果如表 2 所示。

表 2 不同网络结构的正确识别率

Table 2 Recognition rate of different network structures			
隐含层 缩放系数	DBN 识别率/%	DLWF 识别率/%	DLWF <sup>+</sup> 识别率/%
0.70	55.75	56.25	58.45
0.75	64.50	66.75	69.35
0.80	68.25	70.25	73.85
0.85	71.10	74.50	79.25
0.90	82.50	85.25	88.55

由表 2 可知,随着隐含层神经元缩放系数增加,深度神经网络能够更好地提取得到适合于分类的人脸图像抽象特征,但计算量也随之增大。同时,隐含层神经元数目相同时,本文提出算法的识别率均高于只输入整张人脸图像像素级特征的传统深度神经网络算法。

2) 训练样本数不同时对比实验

实验 2 的目的是为了测试本文算法在不同数量训练样本的情况下,与传统人脸识别算法在识别率方面所表现的优劣程度。由于在实际应用环境中很难采集得到大量有标签的训练样本,所以在具有较低的训练样本个数条件下,能取得较好的识别效果就变得尤为重要。表 3 为深度神经网络隐含层结点数目的缩放系数为 0.9 时,在不同训练样本数目的条件下,与传统人脸识别方法进行比较的结果。

表 3 不同训练样本数正确识别率

Table 3 Recognition rate of different training samples			
方法	微调阶段训练样本数/%		
	5	10	15
PCA	23	27	29
SWM	35.5	44	55
DBN	82.5	83.5	86
FIP+LDA	83	85.5	86
DLWF	85.25	87.5	88
DLWF <sup>+</sup>	86.5	88.25	89.5

由表 3 可知,本文提出算法经过非监督的预训练,为 DBN 参数提供了较好的初始化值,再经过少量有标签的训练样本有监督的微调,便能达到较好的识别效果。随着训练样本数不断增加开始时识别准确率明显上升,而后趋于不变,表明本文算法在非限制条件下进行识别具有良好的鲁棒性。同时,表 3 也表明了本文算法比直接输入整张人脸图像像素级特征的 DBN 识别准确率更高。传统的人脸识别算法如 PCA、SVM 等在训练样本个数较少时,所提取的底维特征分类能力并不强。随着训练样本个数增加,提取的特征分类能力有所增强,但在非限制条件下算法的识别精度并不高。

3) 类别数不同时对比实验

实验 3 目的是为了验证本文算法在非限制条件下的普适性和有效性。从 LFW 中选择拥有人脸图像数目大于或等于 25 幅图像的 39 个人作为实验的候选样本,从其中随机挑选出 10、20、30 人作为实验对象,每人随机挑选出 5 幅人脸图像作为训练样本,其余图像作为测试样本。每次随机选择不同人脸图像进行实验,求取平均值作为最终结果。本次实验采用的 DBN 网络隐含层缩放系数与实验 2 相同,实验结果如表 4 所示。

表 4 类别数不同时不同算法正确识别率

Table 4 Recognition rate of different algorithms			
方法	类别数/%		
	10	20	30
PCA	27.32	25.74	24.63
SVM	41.65	38.23	36.46
DBN	83.71	84.05	85.21
FIP+LDA	84.61	86.72	87.51
DLWF	85.93	87.16	88.76
DLWF <sup>+</sup>	86.72	87.84	91.63

经过在 LFW 人脸库上实验可知,本文提出的人脸识别算法在非限制条件下的鲁棒性和准确率均高



于传统人脸识别算法。通过将人脸面部进行区域划分,分别提取局部抽象特征的方法,比传统仅输入整张人脸像素级特征方法更容易学习得到适合于分类的局部特征。并通过上述实验证明了本文算法在非限制条件下的有效性。

2.2 ORL 人脸库上的实验

ORL 人脸库又称 AT&T 人脸库,是剑桥大学历时 2 年拍摄完成的,包含 40 个人,每个人有 10 幅不同姿态、表情、光照的图像,共 400 幅,其中人脸的姿态和光照变化较小,图像分辨率为 112×92,图像背景为黑色,以 PGM 格式存储。实验过程中,神经网络首先利用 LFW 人脸库中所有人脸图像进行预训练,然后在 ORL 人脸库中每人随机选取 5 幅图像作为微调阶段训练样本集,其余的人脸图像作为测试样本集。

1) 隐含层单元数不同时对比实验

本实验目的是为了验证本文提出算法的通用性,以及在 ORL 人脸库上,不同隐含层缩放系数所表现出的识别性能的差异。表 5 为 DBN 网络隐含层缩放系数不同时的识别结果。

表 5 ORL 库上不同网络结构时正确识别率

Table 5 Recognition rate of different network structures in ORL			
隐含层 缩放系数	DBN 识别率/%	DLWF 识别率/%	DLWF <sup>+</sup> 识别率/%
0.70	44	57	61
0.75	52	64	67
0.80	71	77	82
0.85	82	86	89
0.90	94	97	97

从表 5 可知,当隐含层缩放系数逐渐增加时,网络逐层提取得到的人脸图像抽象特征的分类识别能力逐渐增强。当网络隐含层缩放系数为 0.9 时,获得了较高识别率 97%。

2) 与其他人脸识别方法对比实验

为了验证本文提出算法相对于传统人脸识别算法所具有的优越性,分别同 PCA、SVM 和传统 DBN 算法做了对比实验,表 6 为本文提出算法在 ORL 人脸库上与其他算法进行对比的实验结果。实验结果表明本文算法具有较好的通用性,与传统人脸识别方法进行比较,在识别率方面具有较为明显的优势。

表 6 ORL 库上不同算法识别率比较	
Table 6 Comparison of different algorithms on ORL	
方法	识别率/%
PCA	84.0
SVM	88.5
DBN	94.0
FIP+LDA	96.0
DLWF	97.0
DLWF <sup>+</sup>	97.0

3 结束语

本文提出了一种特征加权融合人脸识别方法,首先通过提取面部特征点将人脸图像划分成多个局部区域采样区域,然后将采样区域通过归一化处理后分别输入到对应 DBN 网络中,求取相似度矩阵,最后进行加权求和得出最终判别结果并计算识别准确率。经过在 LFW 人脸库上实验证明了本文算法在非限制条件下进行人脸识别的有效性,并且能够自下而上提取得到适合于分类的抽象人脸特征。识别率达到了 91.63%。同时在 ORL 人脸库上进行的实验充分证明了本文算法的通用性和有效性,识别率达到了较高的 97%,表现出了比传统人脸识别更高的优越性。

参考文献:

[1] TURK M, PENTLAND A. Eigenfaces for recognition[J]. Journal of Cognitive Neuroscience, 1991, 3(1): 71-86.

[2] SUN Yi, WANG Xiaogang, TANG Xiaou. Deep learning face representation from predicting 10, 000 classes[C]//Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA, 2014: 1891-1898.

[3] HU Junlin, LU Jiwen, TAN Y P. Discriminative deep metric learning for face verification in the wild[C]//Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA, 2014: 1875-1882.

[4] HUANG G B, LEE H, LEARNED-MILLER E. Learning hierarchical representations for face verification with convolutional deep belief networks[C]//Proceedings of the 2012 IEEE Conference on Computer Vision and Pattern Recognition. Providence, RI, USA, 2012: 2518-2525.

[5] ZHU Zhenyao, LUO Ping, WANG Xiaogang, et al. Deep learning identity-preserving face space[C]//Proceedings of



- the 2013 IEEE International Conference on Computer Vision. Sydney, NSW, Australia, 2013: 113-120.
- [6] SUN Yi, WANG Xiaogang, TANG Xiaoou. Hybrid deep learning for face verification[C]//Proceedings of the 2013 IEEE International Conference on Computer Vision. Sydney, NSW, Australia, 2013: 1489-1496.
- [7] TAIGMAN Y, YANG Ming, RANZATO M A, et al. Deep-face: Closing the gap to human-level performance in face verification[C]//Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, OH, USA, 2014: 1701-1708.
- [8] SUN Yi, CHEN Yuheng, WANG Xiaogang, et al. Deep learning face representation by joint identification-verification[J]. Advances in Neural Information Processing Systems. 2014.
- [9] HINTON G E, OSINDERO S, TEH Y W. A fast learning algorithm for deep belief nets[J]. Neural Computation, 2006, 18(7): 1527-1554.
- [10] HINTON G E, SALAKHUTDINOV R R. Reducing the dimensionality of data with neural networks[J]. Science, 2006, 313(5786): 504-507.
- [11] 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.
- [12] COOTES T F, TAYLOR C J, COOPER D H, et al. Active shape models-their training and application[J]. Computer Vision and Image Understanding, 1995, 61(1): 38-59.
- [13] MILBORROW S, NICOLLS F. Active shape models with SIFT descriptors and MARS[J]. VISAPP, 2014, 1(2): 5.
- [14] MILBORROW S, BISHOP T E, NICOLLS F. Multiview active shape models with SIFT descriptors for the 300-W face landmark challenge[C]//Proceedings of the 2013 IEEE International Conference on Computer Vision Workshops. Sydney, NSW, Australia, 2013: 378-385.
- [15] Bengio Y, Delalleau O. On the expressive power of deep architectures[C]//Proceedings of the 22nd International Conference. Espoo, Finland, 2011: 18-36.
- [16] HINTON G E. Training products of experts by minimizing contrastive divergence[J]. Neural Computation, 2002, 14(8): 1771-1800.
- [17] BENGIO Y. Learning deep architectures for AI[J]. Foundations and Trends in Machine Learning, 2009, 2(1): 1-127.

#### 作者简介:



孙劲光,女,1962年生,博士,教授,博士生导师,计算机学会(CCF)会员(21314S),主要研究方向为计算机图像处理、计算机图形学、知识工程。



孟凡宇,男,1991年生,硕士研究生,主要研究方向为计算机图像处理。