

# 一种卷积神经网络和极限学习机相结合的人脸识别方法

余丹 吴小俊

(江南大学物联网工程学院, 无锡, 214122)

**摘要:** 卷积神经网络是一种较好的特征提取器, 但却不是最佳的分类器, 而极限学习机能够很好地进行分类, 却不能学习复杂的特征, 根据这两者的优点和缺点, 将它们结合起来, 提出一种新的人脸识别方法。卷积神经网络提取人脸特征, 极限学习机根据这些特征进行识别。本文还提出固定卷积神经网络的部分卷积核以减少训练参数, 从而提高识别精度的方法。在人脸库 ORL 和 XM2VTS 上进行测试的结果表明, 本文的结合方法能有效提高人脸识别的识别率, 而且固定部分卷积核的方式在训练样本少时具有较大优势。

**关键词:** 卷积神经网络; 极限学习机; 特征提取; 人脸识别

**中图分类号:** TP391.4      **文献标志码:** A

## Face Recognition Algorithm Based on Combination of Convolutional Neural Networks and Extreme Learning Machine

Yu Dan, Wu Xiaojun

(School of Internet of Things Engineering, Jiangnan University, Wuxi, 214122, China)

**Abstract:** Convolutional neural networks are good at learning features, but not always optimal for classification, while extreme learning machines are good at producing decision surfaces from well-behaved feature vector, but cannot learn complicated invariances. Based on the advantages and disadvantages of convolutional neural networks and extreme learning machine, we present a hybrid system where a convolutional neural network is trained to extract features and an extreme learning machine is trained from the features learned by the convolutional neural networks to recognize faces. We also propose prefix part of the filters in the convolutional layers to reduce parameters for improving the recognition accuracy. The experimental results obtained on the ORL and XM2VTS databases show that the proposed method can effectively improve the performance of face recognition, and the method of prefixing part of the filters is better than the method of stochastic filters in small training data.

**Key words:** convolutional neural networks; extreme learning machine; feature extraction; face recognition

## 引 言

人脸识别一直都是生物识别领域的研究热点, 在计算机、(应用)数学、图像处理和模式识别等学科

都有广泛的研究。近年来,卷积神经网络(Convolutional neural networks, CNNs)已成为语音分析和图像识别领域的研究热点<sup>[1-3]</sup>,尤其在图像识别领域取得了不俗的成果:Krizhevsky 等运用卷积神经网络在 ILSVRC-2012(Imagenet 2012 large-scale visual recognition challenge)取得了前所未有的识别率<sup>[4]</sup>;在 MNIST 手写数字数据库上,卷积神经网络也取得了很好的识别率(错误率小于 0.3%)<sup>[5]</sup>。卷积神经网络的局部感受野和权值共享使之更类似于生物神经网络,减少了网络模型参数数量,降低了模型复杂度。卷积神经网络将图像直接作为网络的输入,信息再通过网络前向传播依次传递到不同层,每层获得图像的不同特征,如轮廓、纹理特征等,从而避免了传统识别算法中复杂的特征提取过程。由于图像的局部感受区域允许神经元或处理单元可以访问到最基础的特征,如定向边缘或角点等,所以卷积神经网络对平移、缩放和旋转及其他形式的变形具有高度不变性<sup>[6]</sup>。卷积神经网络的最后一层可以看成是一个线性分类器,所以卷积神经网络虽然能够提取最显著的特征,但却不是一个最优的分类器。支持向量机(Support vector machine, SVM)及它的一些改进算法是应用最广泛的分类器之一<sup>[7-10]</sup>,因此 Yann 等<sup>[11]</sup>在 2006 年提出了卷积神经网络与支持向量机相结合的识别算法,卷积神经网络提取特征,然后支持向量机根据提取的特征进行识别,从而避免了卷积神经网络不是最优分类器的缺点。Huang 等<sup>[12]</sup>在 2004 年提出了极限学习机(Extreme learning machine, ELM),并在此基础上做了许多改进<sup>[13,14]</sup>。与支持向量机等传统的分类算法相比,极限学习机的参数不需要通过迭代训练,而是通过最小二乘法直接计算得到,所以在理论上具有最小误差以及非常快的训练速度,而且不需要人为地选择学习率等训练参数,避免了复杂的调优过程,将其应用于人脸识别取得了较好的效果<sup>[15]</sup>。但是 SVM,ELM 等都需要事先人为地挑选特征,在好的特征下分类效果很好,而一旦特征挑选不好,分类效果也就会很差。

考虑到卷积神经网络能够提取最优分类特征,而极限学习机训练速度快、训练精度高,本文提出将两者相结合的算法,并应用于人脸识别。该方法的特点是:先利用训练样本训练卷积神经网络,训练好后,卷积神经网络的其他层参数保持不变,并将最后一层替换成极限学习机,然后训练得到极限学习机的参数,相当于将卷积神经网络提取的特征作为极限学习机的输入,综合了卷积神经网络和极限学习机的优势。虽然卷积神经网络的权值共享结构减少了网络参数,但由于其复杂的多层结构,仍有大量的参数需要训练。在训练样本数较少的情况下,识别效果往往不太好。本文在此基础上提出固定卷积神经网络的部分卷积核(即部分卷积核用特定的滤波器代替)来减少训练参数,从而提高识别精度。

## 1 卷积神经网络和极限学习机

### 1.1 卷积神经网络

卷积神经网络以图像作为输入,紧跟着的几层为卷积层和子采样层的交替实现,在最后靠近输出的一些层是全连接的一维网络。若第  $l$  层为卷积层,则其第  $j$  个特征映射图  $x_j^l$  为

$$x_j^l = f\left(\sum_{x_i^{l-1} \in M_j} x_i^{l-1} * k_{ij}^l + b_j^l\right) \quad (1)$$

式中: $M_j$  表示选择的输入映射图的集合,即  $l-1$  层中与第  $j$  个映射图相连的映射图的集合; $f$  为非线性函数(本文选用 sigmod 函数); $k_{ij}^l$  为连接  $l-1$  层中第  $i$  个映射图和第  $l$  层中第  $j$  个映射图的卷积核; $b_j^l$  为偏置。

若第  $l$  层为子采样层,则其第  $j$  个特征映射图为

$$x_j^l = f(\omega_j^l \text{down}(x_j^{l-1}) + b_j^l) \quad (2)$$

式中: $\text{down}(\cdot)$  为下采样函数,有许多不同的方法,如求均值、最大值、随机采样及  $L_p$  采样<sup>[16]</sup>等; $\omega_j^l$  为权值; $b_j^l$  为偏置。

若第  $i$  层为全连接层, 则其第  $j$  个特征映射图为

$$x^i = f(\beta^i \cdot \mathbf{v} + b^i) \quad (3)$$

式中:  $\mathbf{v}$  为  $l-1$  层的子采样层或卷积层的特征图拉成的向量或全连接层输出的向量;  $\beta^i$  为权值;  $b^i$  为偏置量。

## 1.2 极限学习机

极限学习机最初为解决单隐层前馈神经网络 (Single-hidden layer feed-forward neural networks, SLFNs) 而提出。假设有  $N$  个任意不同样本  $(x_i, t_i)$ ,  $i=1, 2, \dots, N$ , 其中  $\mathbf{x}_i = [x_{i1}, x_{i2}, \dots, x_{in}] \in \mathbf{R}^n$ ,  $t_i = [t_{i1}, t_{i2}, \dots, t_{im}] \in \mathbf{R}^m$ , 则极限学习机的目标函数为

$$\mathbf{H}\boldsymbol{\beta} = \mathbf{T} \quad (4)$$

式中:  $\mathbf{H}$  为神经网络隐层的输出矩阵;  $\mathbf{T}$  为期望输出。

$$\mathbf{H} = \begin{bmatrix} h(x_1) \\ h(x_2) \\ \vdots \\ h(x_N) \end{bmatrix}, \quad \mathbf{T} = [t_1, t_2, \dots, t_N]^T \quad (5)$$

$$h(x) = g(\mathbf{W} \cdot \mathbf{x} + \mathbf{b}) \quad (6)$$

式中:  $\mathbf{W} \in \mathbf{R}^{L \times n}$  为输入层到隐含层的权值;  $\mathbf{b} \in \mathbf{R}^L$  为隐含层的偏置量;  $g(\cdot)$  为激励函数。

极限学习机的基本思想是: 随机产生  $\mathbf{W}$  和  $\mathbf{b}$ , 然后据此计算  $\mathbf{H}$  和  $\boldsymbol{\beta}$ 。  $\boldsymbol{\beta}$  的确定有很多方法<sup>[17,18]</sup>, 其中一种简单的方法为

$$\boldsymbol{\beta} = \mathbf{H}^+ \mathbf{T} \quad (7)$$

式中:  $\mathbf{H}^+$  为  $\mathbf{H}$  的 Moore-Penrose 广义逆。

## 1.3 卷积神经网络和极限学习机的结合

卷积神经网络是一个卷积层和子采样层交替出现的多层网络, 并逐层提取分类特征, 最后一层为一个线性分类器, 然后采用梯度下降法来调整各层的权值。由于卷积网络通过所有层同时训练来最小化全局误差, 因此, 特征提取是整个分类系统的一部分, 而不单独存在, 且完全由原始数据训练得到, 不需要提前设计, 所以通过卷积神经网络提取的特征最有利于分类, 但其最后一层的分类器却不是最好的分类器。

极限学习机是单隐层的神经网络, 它的输入层到隐含层的权值可以随机产生, 然后用广义逆的方法求得隐含层到输出层的权值, 并且该权值使得实际输出与期望输出的误差最小。由于极限学习机的权值直接通过计算, 而不是通过迭代训练得到, 所以它的训练速度非常快, 且训练精度很高。但是极限学习机只有在好的分类特征的前提下才能获得高的识别率, 如果特征选取不好, 识别效果就可能很不好。

将卷积神经网络和极限学习机结合起来, 可以克服两者的缺点, 并充分利用其优点。图 1 显示了卷积神经网络和极限学习机的结合方式。如图 1 所示, 卷积神经网络和极限学习机相结合的模式包括 3 个部分: (1) 特征学习, 为卷积神经网络学习部分, 本文实验时卷积神经网络采用 6 层结构, 包括输入层、两层卷积层 ( $C_1, C_3$ )、两层子采样层 ( $S_2, S_1$ ) 及最后一层分类器层, 并通过训练样本训练得到卷积网络的各层参数。(2) 参数转换部分, 即将卷积神经网络的最后一层替换成极限学习机, 而保持其他参数不变, 相当于将卷积神经网络作为一个训练好的特征提取器。(3) 分类器学习部分, 即将卷积神经网络提取的特征作为极限学习机的输入, 然后训练得到极限学习机的参数。

该网络的具体过程如下: 先采用梯度下降法调整卷积神经网络参数, 训练完成后, 将通过卷积神经

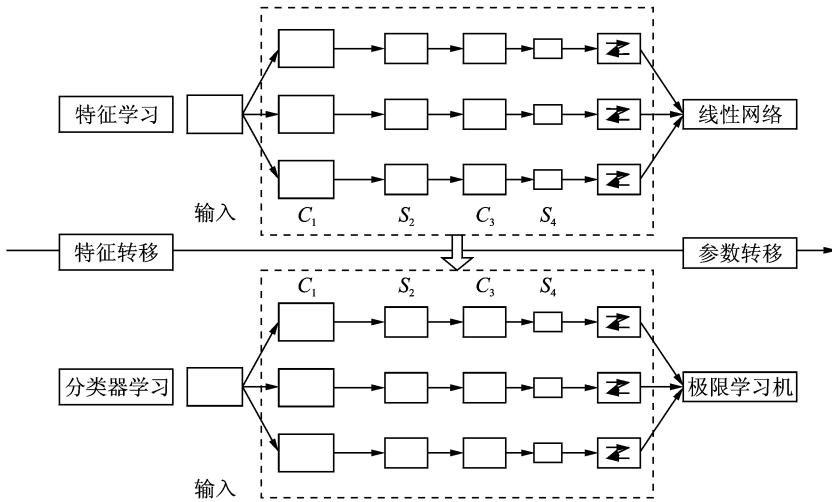


图1 卷积神经网络和极限学习机结合结构示意图

Fig.1 Schematic representation of the combination of convolutional neural network and extreme learning machine

网络提取的特征作为极限学习机的输入特征,根据极限学习机训练公式计算其参数,最终形成整个分类网络。

卷积神经网络和极限学习机结合算法如下:

输入:训练样本及对应的期望输出  $(x_i, t_i), i=1, 2, \dots, N$ 。

输出:训练好的分类模型。

(1) 初始化卷积神经网络参数、目标精度  $\epsilon$  及最大迭代次数 epochs。

(2) While ( $i < \text{epochs} \ \& \ \text{err} > \epsilon$ )

(a) 根据式(1~3)计算实际输出;

(b) 采用梯度下降法调整参数。

Endwhile

(3) 随机产生极限学习机参数  $\mathbf{W}$  和  $\mathbf{b}$ 。

(4) 根据式(1,2)计算卷积神经网络倒数第2层输出并其作为极限学习机输入。

(5) 根据式(6,7)计算  $\mathbf{H}$  和  $\beta$ 。

## 2 固定卷积网络部分卷积核的方式

卷积神经网络虽然采用权值共享的结构减少了训练参数,但其需要训练的参数还是很多,在训练样本少时,训练效果往往不是太好。考虑到卷积层的每一个卷积核都相当于提取一种图像特征,如某个方向的边缘,多个不同参数的卷积核就提取了图像的多种特征,通常这些参数通过学习得到,本文提出将部分卷积核用固定的滤波器代替,这些滤波器的参数将不再学习,从而使得网络在调节更少参数的同时提取了这些特征。本文分别用快速傅里叶变换(Fast Fourier transform, FFT)、小波变换(Wavelet transform, WT)、Contourlet 变换及 Sobel 算子和 Laplacian 算子作为固定的卷积核,提取某一特定的图像特征。本文卷积神经网络的第1层卷积层采用以上固定的卷积核,在训练过程将不调整这些卷积核的参数,其他层则随机产生并通过迭代调整参数。

### 3 实验结果与分析

#### 3.1 ORL 上的人脸识别实验

ORL 人脸图像库由 AT&T Laboratories Cambridge 提供, 是人脸识别中常用的数据库, 数据库中 共有 400 幅图像, 40 个人, 每人 10 幅图像, 照片为  $112 \times 92$  像素的灰度图像。照片拍摄于不同的时间, 有一定的面部表情变化、光照变化等, 能够较好地反映一个人不同人脸的变化和差异。

为了进行人脸识别, 将 ORL 人脸库分成两部分, 每人前 3~7 幅作为训练样本, 其余作为测试样本, 分别进行不同方法的实验。在本文中, 输入数据为  $112 \times 92$  的像素点构成的矩阵, 第 1 个特征图层( $C_1$  层)采用  $9 \times 9$  的窗口对输入图像进行卷积操作, 然后第 1 个子采样层( $S_2$  层)针对  $C_1$  进行子采样操作, 窗口大小为  $4 \times 4$ 。 $C_3$  层为一个卷积层, 卷积核大小同样为  $9 \times 9$ ,  $C_3$  与  $S_2$  中的所有特征图都相连。 $S_4$  则在  $C_3$  的基础上进行子采样, 大小同样为  $4 \times 4$ , 然后将  $S_4$  中的所有特征图拉成一个向量, 直接与输出层相连, 得到最终的结果。为了验证本文方法的有效性, 分别对极限学习机、卷积神经网络及本文方法进行了实验, 其中极限学习机的分类特征通过方向梯度直方图(Histogram of oriented gradient, HOG)<sup>[19]</sup>提取; 考虑到卷积核的影响, 分别对随机产生卷积核和固定部分卷积核的方式进行了实验:(1) 随机产生卷积核, 然后通过样本训练调整参数;(2) 第 1 层卷积层采用固定的卷积方式(FFT, Wavelet, Contourlet, Sobel 算子及 Laplacian 算子), 其他卷积层随机采用随机产生的卷积核; 另一个影响网络性能的因素是子采样方法的选取, 所以本文也分别进行了不同子采样方法的实验, 包括 Max-pooling, Average-pooling, Stochastic-pooling 及 Lp-pooling。实验还与当前一些流行算法, 如基于稀疏表示的分类(Sparse representation-based classification, SRC)<sup>[20]</sup>和 FV(Fisher vector)<sup>[21]</sup>等算法进行了比较。不同算法的结果如表 1 和图 2 所示。

表 1 不同方法在 ORL 人脸库上实验结果比较

Tab. 1 Performance comparison of different methods on ORL

方法	卷积核	子采样方法	训练样本数					
			3	4	5	6	7	
ELM			0.753 6	0.833 3	0.860 0	0.912 5	0.933 3	
SRC			0.793 6	0.875 0	0.885 0	0.937 5	0.941 7	
FV			0.767 9	0.850 0	0.845 0	0.887 5	0.908 3	
CNNs		Max-pooling	0.692 9	0.787 5	0.840 0	0.900 0	0.916 7	
CNNs		Average-pooling	0.757 1	0.829 2	0.865 0	0.931 3	0.933 3	
CNNs		Stochastic-pooling	0.753 6	0.833 3	0.860 0	0.918 7	0.941 7	
CNNs	随机	Lp-pooling	0.757 1	0.837 5	0.870 0	0.937 5	0.958 3	
CNNs+ELM		Max-pooling	0.757 1	0.837 5	0.880 0	0.943 7	0.966 7	
CNNs+ELM		Average-pooling	0.775 0	0.841 7	0.920 0	0.962 5	0.975 0	
CNNs+ELM		Stochastic-pooling	0.757 1	0.837 5	0.895 0	0.950 0	0.966 7	
CNNs+ELM		Lp-pooling	0.789 3	0.845 8	0.925 0	0.968 8	0.983 3	
CNNs		固定	Max-pooling	0.700 0	0.795 8	0.845 0	0.900 0	0.916 7
CNNs			Average-pooling	0.760 7	0.837 5	0.880 0	0.925 0	0.925 0
CNNs			Stochastic-pooling	0.753 6	0.833 3	0.870 0	0.918 7	0.933 3
CNNs	Lp-pooling		0.775 0	0.841 7	0.885 0	0.937 5	0.958 3	
CNNs+ELM	Max-pooling		0.792 9	0.841 7	0.890 0	0.943 7	0.966 7	
CNNs+ELM	Average-pooling		0.814 3	0.866 7	0.925 0	0.962 5	0.975 0	
CNNs+ELM	Stochastic-pooling		0.803 6	0.858 3	0.895 0	0.956 3	0.966 7	
CNNs+ELM	Lp-pooling		0.817 9	0.875 0	0.930 0	0.968 8	0.983 3	

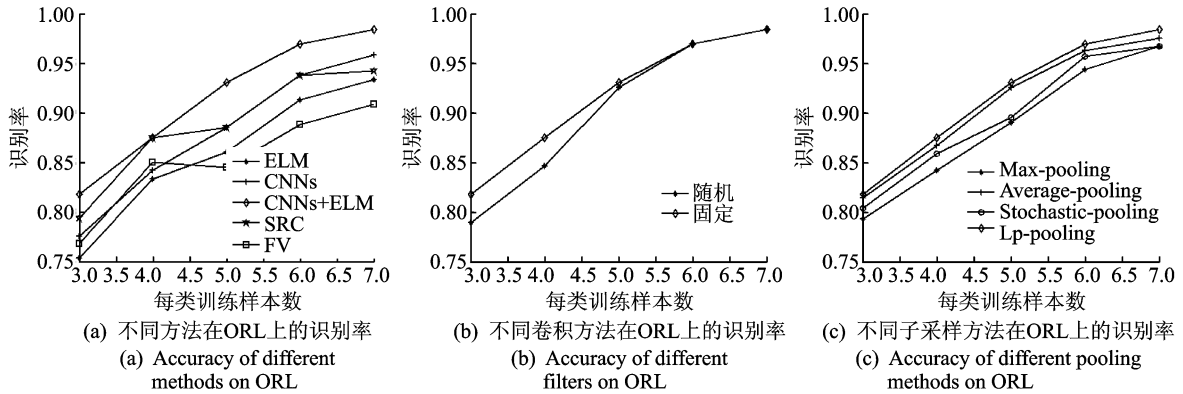


图2 不同方法在 ORL 人脸库上实验结果比较

Fig.2 Performance comparison of different methods on ORL

从表 1 数据和图 2(a)可以看出,本文方法比单独使用卷积神经网络或极限学习机的识别效果有显著的提高,这是因为极限学习机比卷积神经网络的最后一层的线性分类器具有更好的分类效果,而卷积神经网络比一般的特征提取方法更能够提取更优的分类特征。随机产生卷积核的方式在训练样本少时,效果不如 SRC, FV 等方法,但随着样本数的增加,识别率逐渐超过了这些方法,而固定卷积核的方式在训练样本少时,比 SRC 和 FV 效果略好,而且随着样本数增加,优势越来越明显。从图 2(b)可以看出,采用固定的卷积核的方式在训练样本较少时,识别率高于随机产生卷积核的方式,这是因为固定卷积核的方式在提取了固定滤波器产生的特征的同时,减少了需要训练的参数,使得在样本少的情况下,能够更好地训练这些参数,从而提取出更好的特征。图 2(c)显示了不同子采样方法对识别率的影响,其中 Lp-pooling 子采样的效果最好。

### 3.2 XM2VTS 上的人脸识别实验

XM2VTS 人脸库包括 295 人在 4 个月时间内 4 次录制的人脸和语音数据。每次采集都包括 2 个头部旋转视频和 6 个不同语音视频片断。从每个时期中提取两幅图像构成这个用于人脸识别的人脸库,分辨率为  $55 \times 51$ 。为了进行人脸识别,将 XM2VTS 人脸库分成两部分,每人前 3~7 幅作为训练样本,其余作为测试样本,分别进行不同方法的实验。在本文中,输入数据为  $55 \times 51$  的像素点构成的矩阵,第 1 个特征图层( $C_1$  层),采用的窗口对输入图像进行卷积操作,然后第 1 个子采样层( $S_2$  层)层针对  $C_1$  进行子采样操作,窗口大小为  $2 \times 2$ 。 $C_3$  层为一个卷积层,卷积核大小同样为  $5 \times 5$ , $C_3$  与  $S_2$  中的所有特征图都相连。 $S_4$  则在  $C_3$  的基础上进行子采样,大小同样为  $2 \times 2$ ,然后将  $S_4$  中的所有特征图拉成一个向量,直接与输出层相连,得到最终的结果。

在 XM2VTS 人脸库上也如 ORL 人脸库上讨论了不同的卷积和子采样方法并与 SRC 和 FV 进行了对比实验。不同算法的识别率如表 2 和图 3 所示。从表 2 和图 3 可以看出,在 XM2VTS 人脸库上的实验和 ORL 上的一样,卷积神经网络和极限学习机相结合的方法优于单独使用极限学习机或卷积神经网络,并且优于 SRC 和 FV,而且固定部分卷积核的方式在样本少时具有优势。

从以上在 ORL 和 XM2VTS 两个数据库上的实验结果可以看出,本文方法在识别效果上总体优于单独的卷积神经网络和极限学习机方法以及 SRC 和 FV 方法,当训练样本少时,采用固定部分卷积核的方式可以提高识别率。

表 2 不同方法在 XM2VTS 人脸库上实验的识别率  
Tab. 2 Performance comparison among different methods on XM2VTS

方法	卷积核	子采样方法	训练样本数				
			3	4	5	6	7
ELM			0.745 8	0.830 5	0.894 9	0.889 8	0.935 6
SRC			0.753 2	0.848 3	0.897 2	0.906 8	0.922 0
FV			0.755 9	0.847 5	0.898 3	0.905 1	0.928 8
CNNs	随机	Max-pooling	0.678 0	0.834 7	0.888 1	0.906 9	0.925 4
CNNs		Average-pooling	0.748 5	0.837 3	0.894 9	0.911 8	0.935 6
CNNs		Stochastic-pooling	0.747 1	0.835 6	0.892 7	0.908 5	0.932 2
CNNs		Lp-pooling	0.750 5	0.841 5	0.896 0	0.913 6	0.935 6
CNNs+ELM		Max-pooling	0.750 5	0.842 4	0.897 2	0.913 6	0.939 0
CNNs+ELM		Average-pooling	0.755 3	0.847 5	0.906 2	0.916 9	0.949 2
CNNs+ELM		Stochastic-pooling	0.753 2	0.844 9	0.904 0	0.915 3	0.939 0
CNNs+ELM		Lp-pooling	0.755 9	0.851 7	0.910 7	0.922 2	0.959 3
CNNs	固定	Max-pooling	0.711 8	0.848 3	0.896 0	0.906 9	0.928 8
CNNs		Average-pooling	0.750 5	0.856 8	0.904 0	0.911 8	0.939 0
CNNs		Stochastic-pooling	0.754 8	0.856 8	0.902 8	0.910 2	0.932 2
CNNs		Lp-pooling	0.752 5	0.858 5	0.905 1	0.913 6	0.909 0
CNNs+ELM		Max-pooling	0.752 5	0.858 5	0.905 1	0.916 9	0.935 6
CNNs+ELM		Average-pooling	0.767 5	0.861 0	0.909 6	0.920 3	0.949 2
CNNs+ELM		Stochastic-pooling	0.756 8	0.859 3	0.908 5	0.918 6	0.945 8
CNNs+ELM		Lp-pooling	0.779 7	0.866 1	0.915 2	0.922 0	0.952 5

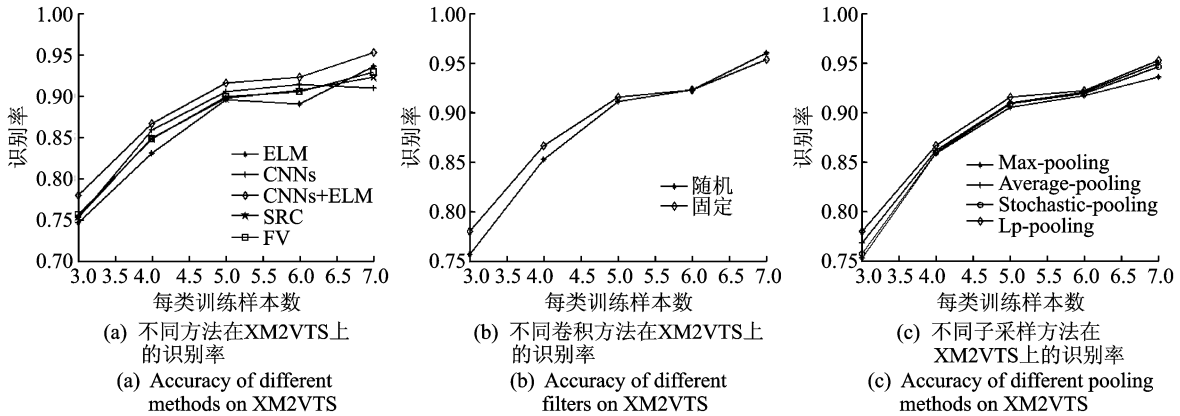


图 3 不同方法在 XM2VTS 人脸库上实验结果比较

Fig. 3 Performance comparison of different methods on XM2VTS

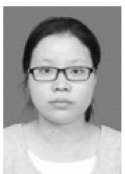
### 4 结束语

本文提出将卷积神经网络和极限学习机相结合的人脸识别方法,该方法充分利用了卷积神经网络的特征提取能力和极限学习机的分类能力,综合了两者的优势,并提出了固定部分卷积核的方式来提高训练样本少时的识别率,同时讨论了不同子采样方法对识别率的影响。实验结果证明了结合方法的优越性,以及当训练样本少时采用固定卷积核的方式比采用随机卷积核的方式识别效果好。此外,实验发现不同的子采样方法得到的结果不同,因此采用什么子采样方法和卷积核才能得到更高的识别率这个问题有待于进一步研究。

## 参考文献:

- [1] 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.
- [2] Ossama A H, Mohamed A R, Jiang H, et al. Applying convolutional neural networks concepts to hybrid NN-HMM model for speech recognition[C]//2012 IEEE International Conference on Acoustics, Speech and Signal Processing. Kyoto, Japan: IEEE Computer Society Press, 2012:4277-4280.
- [3] Turaga S C, Murray J F, Jain V, et al. Convolutional networks can learn to generate affinity graphs for image segmentation [J]. Neural Computation, 2010, 22(2): 511-538.
- [4] Krizhevsky A, Sutskever I, Hinton G E. Image net classification with deep convolutional neural networks[J]. Advances in Neural Information Processing Systems, 2012,25(2): 1097-1105.
- [5] Ciresan D, Meier U, Schmidhuber J. Multi-column deep neural networks for image classification[C]//2012 IEEE Conference on Computer Vision and Pattern Recognition. [S. l.]: IEEE Computer Society Press, 2012: 3642-3649.
- [6] Arel I, Rose D C, Karnowski T P. Deep machine learning—A new frontier in artificial intelligence research[J]. Computational Intelligence Magazine, 2010, 5(4): 13-18.
- [7] Burges C J C. A tutorial on support vector machines for pattern recognition[J]. Data Mining and Knowledge Discovery, 1998, 2(2): 121-167.
- [8] 宋晓宁,吴小俊,杨静宇,等.支持向量机的新分类器算法研究[J].系统仿真学报,2009,21(12):3617-3620.  
Song Xiaoning, Wu Xiaojun, Yang Jingyu, et al. Research on new classifier algorithms for support vector machine[J]. Journal of System Simulation, 2009,21(12):3617-3620.
- [9] 邹永祥,吴宗亮.一种广义不可分的支持向量机算法[J].数据采集与处理,2015,30(2):434-440.  
Zou Yongxiang, Wu Zongliang. Generalized C-support vector machine algorithm[J]. Journal of Data Acquisition and Processing, 2015,30(2):434-440.
- [10] 奚吉,赵力,左加阔.基于改进多核学习的语音情感识别算法[J].数据采集与处理,2014,29(5):730-734.  
Xi Ji, Zhao Li, Zuo Jiakuo. Speech emotion recognition based on modified multiple kernel learning algorithm[J]. Journal of Data Acquisition and Processing,2014,29(5):730-734.
- [11] Huang F J, Yann L C. Large-scale learning with SVM and convolutional networks for generic object recognition[C]//2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. [S. l.]: IEEE Computer Society Press, 2006:284-291.
- [12] Huang G B, Zhu Q Y, Siew C K. Extreme learning machine: A new learning scheme of feedforward neural networks[C]//2004 IEEE International Joint Conference on Neural Networks. [S. l.]: IEEE Computer Society Press, 2004: 985-990.
- [13] Huang G B, Zhu Q Y, Siew C K. Extreme learning machine: Theory and applications[J]. Neurocomputing, 2006, 70(1): 489-501.
- [14] Huang G B, Ding X J, Zhou H M. Optimization method based extreme learning machine for classification[J]. Neurocomputing, 2010, 74(1): 155-163.
- [15] Wang Chao, Yang Jucheng, Chen Yarui, et al. Image latent semantic analysis based face recognition with ensemble extreme learning machine[C]//2014 IEEE Computing, Communications and IT Applications Conference. Beijing, China: IEEE Computer Society Press, 2014: 300-304.
- [16] Sermanet P, Chintala S, LeCun Y. Convolutional neural networks applied to house numbers digit classification[C]//21st International Conference on Pattern Recognition. Sukuba, Japan: IEEE Computer Society Press, 2012: 3288-3291.
- [17] Huang G B, Chen L, Siew C K. Universal approximation using incremental constructive feedforward networks with random hidden nodes[J]. IEEE Transactions on Neural Networks, 2006, 17(4): 879-892.
- [18] Widrow B, Greenblatt A, Kim Y, et al. The no-prop algorithm: A new learning algorithm for multilayer neural networks [J]. IEEE Transactions Neural Networks, 2013, 37: 182-188.
- [19] Dalal N, Triggs B. Histograms of oriented gradients for human detection[C]//2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. San Diego, CA, USA: IEEE Computer Society Press, 2005: 886-893.
- [20] Wright J, Yang A Y, Ganesh A, et al. Robust face recognition via sparse representation[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence,2009,31(2):210-227.
- [21] Csurka G, Perronnin F. Fisher vectors: Beyond bag-of-visual-words image representations[M]. Berlin: Springer Berlin Heidelberg, 2011: 28-42.

## 作者简介:



余丹(1990-),女,硕士研究生,研究方向:人工智能与模式识别,E-mail: yudan\_jiangsu@163.com.



吴小俊(1967-),男,教授,研究方向:计算机视觉、图像处理、模式识别及人工智能等。



