

# 深度 PCANet 模型的人脸美丽预测

甘俊英 姜开永 曾军英 何国辉 谭海英

(五邑大学信息工程学院, 广东江门 529020)

**摘 要:** 深度卷积神经网络(Deep Convolution Neural Network, DCNN)在人脸识别、图像分类和目标检测领域已取得较好效果, 并得到广泛应用; 但是, 在人脸美丽预测中却存在拟合效果欠佳、网络训练难度大等问题。深度 PCANet 模型, 将深度主元分析网络(Principal Component Analysis Network, PCANet)作为特征提取器; 采用无监督预训练提取网络参数, 具有网络训练时间短、图像特征提取快等特点, 能有效避免 DCNN 存在的问题。为此, 本文将深度 PCANet 引入人脸美丽预测, 对训练集图像采用多尺度预处理, 训练深度 PCANet。该模型可提取人脸图像的结构性全局特征, 采用特征增强方法可生成更具表征能力的特征; 运用线性支持向量机(Support Vector Machine, SVM)和随机森林(Random Forest, RF)回归器进行训练和预测。基于 SCUT-FBP 人脸美丽数据库的实验结果表明, 深度 PCANet 模型具有结构简单、特征提取快和无需网络调参优化等特点; 选择合适的图像尺度与采用特征增强方法可提高人脸美丽评价结果, 证明了所提方法的有效性和可行性。

**关键词:** 深度主元分析网络模型; 线性支持向量机回归; 随机森林回归; 人脸美丽预测

**中图分类号:** TP391.4      **文献标识码:** A      **DOI:** 10.16798/j.issn.1003-0530.2018.12.014

## The Facial Beauty Prediction of the Model of Deep PCANet

GAN Jun-ying JIANG Kai-yong ZENG Jun-ying HE Guo-hui TAN Hai-ying

(School of Information Engineering, Wuyi University, Jiangmen, Guangdong 529020, China)

**Abstract:** Deep convolution neural network (DCNN) has achieved good results in face recognition, image classification and object detection, and has been widely used. However, DCNN has some problems in the facial beauty prediction, such as bad fitting effect, hard training etc. Deep PCANet model, using Principal Component Analysis Network (PCANet) as feature extractor. The model can get parameters through unsupervised pre-training with the advantages of less time in training and faster speed in feature extraction, which can effectively avoid the problems of DCNN. For this reason, the deep PCANet is introduced into the facial beauty prediction, and the training set image adopts multi-scale preprocessing to train the deep PCANet. The model can extract the structural global features of the face image, and the feature enhancement method can generate features with more representative capabilities. Finally, linear Support Vector Machine (SVM) and Random Forest (RF) regressors were used for training and prediction. Experimental results for facial beauty prediction on SCUT-FBP database show that deep PCANet has advantages of simpler structure, faster feature extraction, and few parameter adjustment; choosing the right image scale and using feature enhancement methods can improve the face aesthetic evaluation results, which proves that the network is valid and feasible.

**Key words:** deep principal component analysis network; linear support vector machine; random forest; facial beauty prediction

## 1 引言

人脸美丽预测是智能信息处理领域的一个重

要研究方向,与人脸识别、人脸表情识别一样,都是利用人脸特征进行判别的一种技术。在机器学习领域,早期的人脸美学研究注重对感性认识的验

证,如“近古典规则”、“黄金比例”和“三庭五眼”等,将人脸美总结为简单的几何关系,这种方法往往过于简单,而且缺乏严格的科学检验。目前,人脸美学研究从模式分析角度,采用数据驱动的统计学分析,挖掘表征人脸美丽特征,用于人脸美丽预测,如几何特征和表观特征等。其中,表观特征减少了主观干预,受到更多关注。

几何特征和纹理特征属于浅层特征,在人脸美丽预测中取得了大量的研究成果<sup>[1-7]</sup>。但是,由于人脸结构的复杂性,仍然丢失了许多有用的美丽特征,比如肌肉纹理、器官形状和皮肤颜色等。与之前手工设计的特征提取算法相比,深度学习能够提取抽象性、层次性的表观特征,具有更强的表达能力。2010年,Gray等<sup>[8]</sup>开发了一套基于深度卷积神经网络(Deep Convolution Neural Network, DCNN)的女性人脸美丽预测系统,以像素点作为系统输入,在没有人工干预的情况下实现了特征的全自动化提取。Xie等<sup>[9]</sup>在其开发的 SCUT-FBP 人脸美丽数据库上,设计了一个多层 DCNN,与人工标定值相比,取得了 0.8187 的相关系数。Gan 等<sup>[10]</sup>采用局部二值模式(Local Binary Pattern, LBP)、Gabor 和像素特征训练卷积深度信念网络(Convolutional Deep Belief Network, CDBN),在支持向量机(Support Vector Machine, SVM)分类和回归中取得了较好效果。Xu 等<sup>[11]</sup>设计了 3 个深度不等的卷积神经网络(Convolution Neural Network, CNN),采用 SCUT-FBP 人脸美丽数据库进行级联微调实验,在纹理图像上取得 0.82 的相关系数,在 RGB 彩色图像双级联上取得 0.88 的相关系数。因此,深度学习在人脸美丽预测中具有广泛的应用价值。

采用多尺度预处理,生成图像金字塔,能够获得相同结构、不同尺度和清晰度的图像集,有效表征人脸图像美的空间层次结构。同时,由于人脸美丽标签来自大量参评人员主观评分的平均值,在进行人脸美丽评价时,人脸图像的尺度大小、清晰度,成为人脸美丽主观评价的影响因素,会对参评人员的评分造成影响。因此,本文通过对人脸图像多尺度预处理,分解成不同尺度图像,再提取特征,进行训练和预测。例如,2010年,Gray等<sup>[8]</sup>构建了一个浅层 CNN 模型用于人脸美丽回归预测,在 HSV 颜色空间上对非受限人脸图像进行多尺度变换后,送

入网络进行训练,最终取得 0.458 的预测相关系数,高于单尺度图像 0.417 的预测相关系数。2016年,Gan 等<sup>[16]</sup>使用 Multi-scale K-means 方法在自建的人脸美丽数据库上,结合 SVM 分类器,在男性人脸美丽分类和女性人脸美丽分类中分别取得 57.49% 和 55.07% 的分类正确率,分别高于单尺度 k-means 男性人脸美丽分类和女性人脸美丽分类 55.36% 和 52.54% 的分类正确率。

DCNN<sup>[9,12]</sup>是一种有监督网络,在人脸识别和图像识别中被广泛应用,具有良好的分类判别性能。但是,DCNN 存在结构复杂、可调参数多、缺乏完善的选取准则和需要大量有标签数据集进行拟合等问题;尤其在小规模数据库上训练时,容易出现过拟合,训练难度大。因此,本文采用一种简化的深度主元分析网络(Principal Component Analysis Network, PCANet)<sup>[13]</sup>提取特征;该网络不仅具有 DCNN 的卷积功能和层次堆叠结构,而且还具有结构简单、系统资源开销小和对实验样本数量要求小的特点,能有效解决 DCNN 存在的问题;同时,采用线性支持向量机(Support Vector Machine, SVM)回归器<sup>[10]</sup>和随机森林(Random Forest, RF)回归器<sup>[14]</sup>,能够快速实现回归预测。为了提高回归预测效果,本文提出了深度 PCANet 的人脸美丽预测模型,可提取不同尺度图像的结构性全局特征,再用该特征训练线性回归器,最后进行模型性能评估。基于 SCUT-FBP 人脸美丽数据库的实验结果表明,深度 PCANet 模型具有较好的回归性能。

## 2 基于深度 PCANet 的人脸美丽预测模型

深度 PCANet 通过主元分析处理,提取图像块的最大主元分量作为深度 PCANet 卷积滤波器的参数,再通过卷积滤波提取图像深层次的结构性特征。深度 PCANet 提取的图像全局特征,经哈希编码和直方图处理后,转变成高维空间的向量特征,从而通过特征映射将图像转换到高维空间。为了获得更加完备和稳健的人脸美丽特征集合,本文将深度 PCANet 模型用于人脸美丽预测;并对图像采用多尺度预处理,得到不同尺度的人脸图像。如图 1 所示,包括多尺度图像预处理层、深度 PCANet 和回归器。

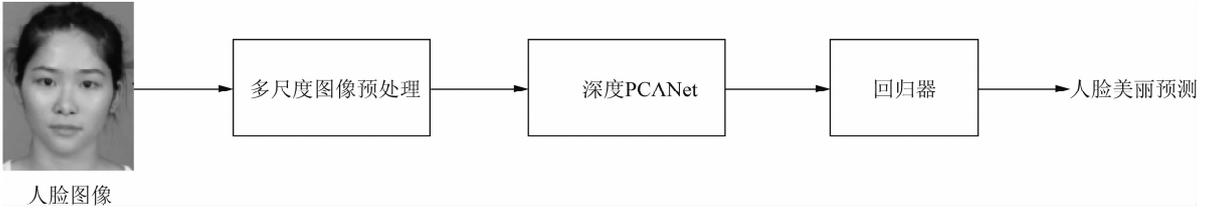


图 1 深度 PCANet 模型框图

Fig. 1 Deep PCANet model block diagram

### 2.1 多尺度图像预处理

图 1 中,多尺度图像预处理主要是对获取图像进行对齐、尺度归一化和下采样等尺度化处理,分别获得尺度为 90×90、60×60、30×30 的图像集,如图 2 所示。

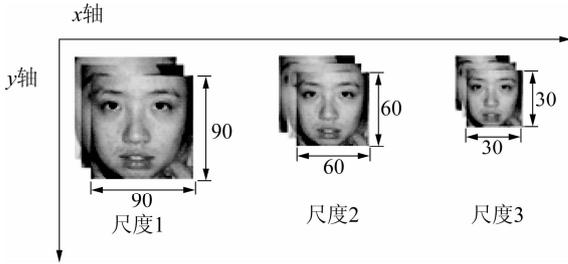


图 2 多尺度训练集

Fig. 2 Deep multi-scale training set

### 2.2 深度 PCANet

深度 PCANet<sup>[13]</sup> 是 Chan 等提出的一种简化的无监督深度卷积神经网络特征提取方法,整体采用层次堆叠结构,能够进行多层次级联,通过主元分析获取深度网络卷积滤波器参数,可提取更深层次更抽象的图像特征。与 DCNN 相比,深度 PCANet 通过无监督预训练提取卷积滤波器参数,无需调参优化,模型结构简单且运行速度更快。

深度 PCANet 通常由输入层、主元分析层、特征

输出层和训练分类器四个部分构成,如图 3 所示。其中,主元分析层由多级 PCA 层构成,用于提取模型的滤波器参数,通常使用 2 级 PCA 层;特征输出层由哈希编码和块直方图变换构成,用于特征图变换和输出特征;训练分类器使用线性回归器,用于评估模型性能。下面以图 3 为例,详细介绍其处理过程。

#### 2.2.1 PCA 层

第 1 级 PCA 主要是对输入图像分块并构建去均值化的图像块矩阵,通过主元分析方法提取滤波器参数。假设有  $N$  幅大小为  $h \times w$  的训练集图像  $\{I_i\}_{i=1}^N$ ,其中,PCA 层的分块大小为  $k_1 \times k_2$ 。从训练集图像提取每个像素周围的  $k_1 \times k_2$  像素块,采用覆盖提取方式,得到像素块矩阵,计算提取块的平均值,得到去均值化图像块矩阵,即

$$\bar{X}_i = [\bar{x}_{i,1}, \bar{x}_{i,2}, \dots, \bar{x}_{i,hw}] \in R^{k_1 k_2 \times hw} \quad (1)$$

对训练集中所有图像进行处理,得到每一幅图像的图像块矩阵,即  $\bar{X}_1, \bar{X}_2, \dots, \bar{X}_N$ ,则

$$X = [\bar{X}_1, \bar{X}_2, \dots, \bar{X}_N] \in R^{k_1 k_2 \times Nhw} \quad (2)$$

通过寻找标准正交矩阵来最小化重构误差,即

$$\min_{V \in R^{k_1 k_2 \times L_1}} \|X - VV^T X\|_F^2, V^T V = I_{L_1} \quad (3)$$

其中,  $V$  表示标准正交矩阵。再求解主元成分,即

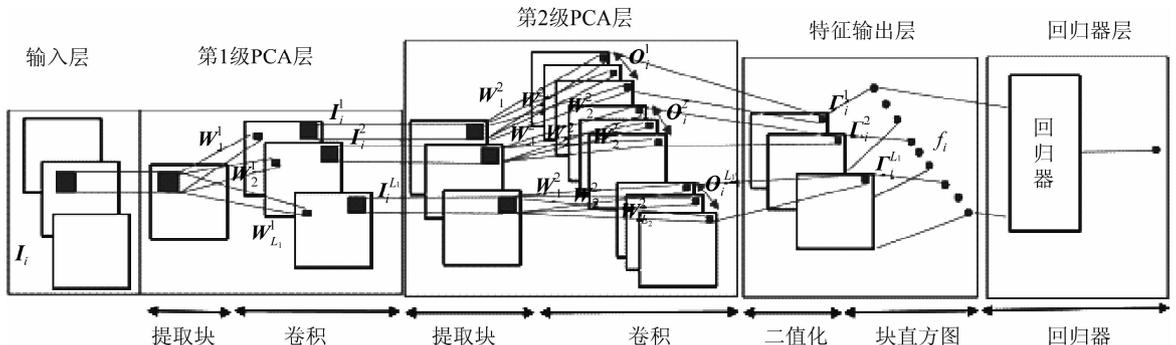


图 3 深度 PCANet 模型

Fig. 3 Deep PCANet model

$$W_l^i = \text{mat}_{k_1, k_2} [q_l(\mathbf{X}\mathbf{X}^T)] \in R^{k_1 \times k_2}, l = 1, 2, \dots, L_1 \quad (4)$$

其中,  $\text{mat}_{k_1, k_2}(\bullet)$  函数是  $k_1 k_2$  维向量到  $k_1 \times k_2$  矩阵的映射,  $q_l(\mathbf{X}\mathbf{X}^T)$  函数表示协方差矩阵的第  $l$  个主特征向量。卷积滤波后, 共有  $N \times L_1$  个映射输出, 即

$$I_l^i = \mathbf{I}E_i * W_l^i \quad i = 1, 2, \dots, N, l = 1, 2, \dots, L_1 \quad (5)$$

其中,  $*$  表示 2 维卷积,  $\mathbf{I}E_i$  表示训练集中第  $i$  幅扩充之后的图像。

第 2 级 PCA 以第 1 级 PCA 的映射输出作为输入, 进行卷积映射, 即

$$O_l^i = \{\mathbf{I}E_i^l * W_y^2\}_{y=1}^{L_2} \quad (6)$$

其中,  $i=1, 2, \dots, N, l=1, 2, \dots, L_1$ 。每一个训练样本经 2 级 PCA 处理后均会产生  $L_1 L_2$  个输出特征矩阵。

### 2.2.2 哈希和直方图处理

对第 2 级 PCA 输出的特征图像进行二值化, 得到只包括 1 和 0 的数值, 即

$$\{H(I_l^i * W_y^2)\}_{y=1}^{L_2} \quad (7)$$

其中,  $H(\cdot)$  函数表示二值处理, 相当于单位阶跃函数。然后, 进行二值化哈希编码, 编码位数与第 2 级 PCA 滤波器个数相同, 将  $O_l^i$  转换成一个整数图像, 即

$$I_l^i = \sum_{y=1}^{L_2} 2^{y-1} (I_l^i * W_y^2) \quad l = 1, 2, \dots, L_1 \quad (8)$$

图像的像素值编码为  $[0, 2^{L_2} - 1]$  区间内的整数。经哈希编程处理后, 得到  $L_1 N$  幅整数编码特征图像。将第  $i$  幅图像  $I_l^i$  的  $L_1$  幅图像分割成  $B$  块, 计算每块的直方图, 转成向量, 并用  $Bh(I_l^i)$  表示, 然后将直方图进行级联, 最终得到块直方图特征, 即

$$f_i = [Bh(I_l^1), Bh(I_l^2), \dots, Bh(I_l^{L_1})]^T \in R^{(2^{L_2})L_1 B} \quad (9)$$

### 2.2.3 回归算法

深度 PCANet 提取的图像块直方图特征高达上百万维, 使用加权主元分析 (Weighted Principal Component Analysis, WPCA) 进行降维后, 特征维数降低到几万维。由于特征维数很高, 以此训练的分类器模型在高维空间能够实现线性可分。

线性 SVM 分类器和随机森林分类器均属于高维数据分类模型。其中, 线性 SVM 分类器在样本数量有限且特征维数很高的情况下, 与 LibSVM 非线性分类器相比, 具有更高的分类效率和准确率。实验证实了线性 SVM 分类器在模型训练上用时更短、正确率更高。随机森林分类器是用随机方式建立的、包含多个决策树的分类器, 在高维数据训练时, 训练速度快且不易出现过拟合, 对数据噪声和错误具有鲁棒性。本文采用以上两种线性分类器的回归方法进行人脸美丽预测实验, 并使用皮尔逊相关系数 (Pearson Correlation Coefficient) 评价预测效果。

## 3 实验结果和分析

### 3.1 SCUT-FBP 数据库

SCUT-FBP 人脸美丽数据库是 Xie 等<sup>[9]</sup> 创建的用于人脸美丽预测的亚洲年青女性人脸美丽数据库, 包含 500 幅中性表情和少量遮挡的正面人脸图像, 如图 4 所示。

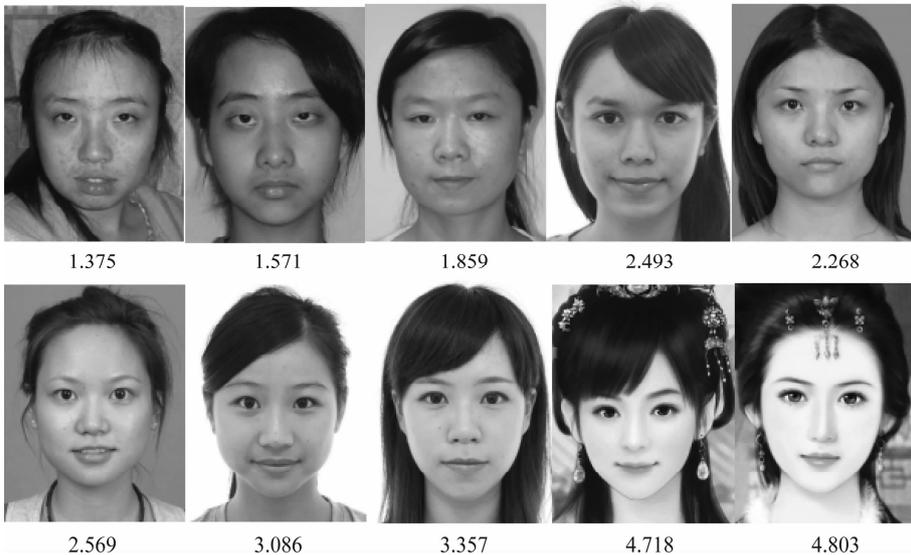


图 4 SCUT-FBP 人脸美丽数据库样本图像和标签

Fig. 4 Sample pictures and labels of SCUT-FBP

SCUT-FBP 人脸美丽数据库的标签分布符合高斯分布,如图 5 所示。人脸图像分布不均衡,极丑和极美的图像数量较少,大多为平均水平,如图 6 所示。

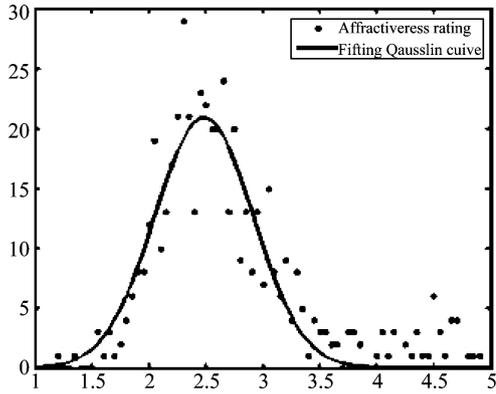


图 5 SCUT-FBP 标签分布图

Fig. 5 Histogram of rating distribution

计算人脸水平偏转角度,进行图像水平对齐;再以两眼中心和嘴巴中心相隔 48 个像素进行尺度归一化,对尺度归一化后的图像进行裁剪和灰度变换,最终得到尺度为 90×90 的图像,如图 8 所示。

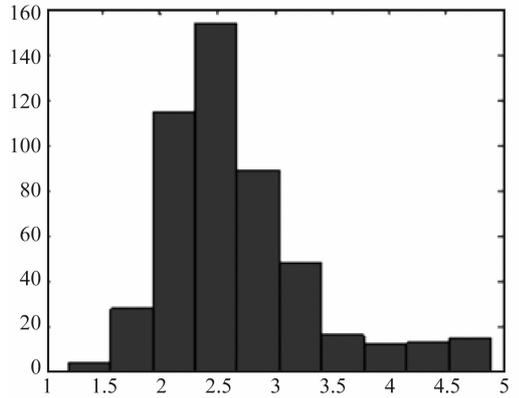


图 6 SCUT-FBP 标签分布直方图

Fig. 6 Grade histogram of SCUT-FBP

### 3.2 图像预处理

为了提高深度 PCANet 模型的训练效果和泛化能力,本文进行人脸检测和关键特征点检测<sup>[15]</sup>,如图 7 所示。以两眼中心关键点连线与水平线的夹角,

### 3.3 结果与分析

本文采用随机划分图像和平均划分图像两种方式,对 SCUT-FBP 人脸美丽数据库进行划分,并完成人脸美丽预测实验。



图 7 人脸检测和关键点检测

Fig. 7 Face detection and key points detection

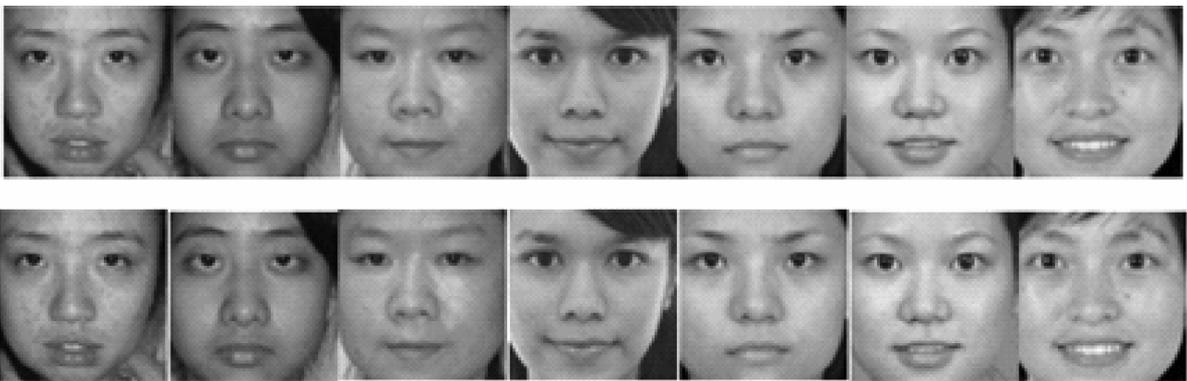


图 8 图像预处理后的彩色图像(上)、灰度图像(下)

Fig. 8 The preprocessed pictures, color(up) gray(down)

### 3.3.1 基于随机划分图像的实验结果

实验时,随机打乱 SCUT-FBP 人脸美丽数据库中的图像顺序,并平均分成 5 等份,然后进行 5 折交叉验证实验,用于实验的各尺度模型名称如表 1 所示。其中,图像尺度表示深度 PCANet 输入图像集的大小;DPCA 模型表示未对特征进行开平方增强处理的深度 PCANet;DPCAS 模型表示对特征进行开平方增强处理的深度 PCANet。例如,DPCA90 模型,未对提取的深度 PCANet 特征进行开平方处理,而是直接送入线性回归器进行模型训练和预测;DPCA90S 模型,对提取的深度 PCANet 特征进行了开平方处理,然后再送入线性回归器进行训练和预测。表 2 是深度 PCANet 模型的参数设置。其中,深度 PCANet 模型包括 2 级 PCA 滤波器,第 1 级有 12 个滤波器,第 2 级有 8 个滤波器。

表 1 深度 PCANet 模型名称  
Tab. 1 Names of deep PCANet model

图像尺度	DPCA 模型	DPCAS 模型
90×90	DPCA90	DPCA90S
60×60	DPCA60	DPCA60S
30×30	DPCA30	DPCA30S

表 2 深度 PCANet 模型参数  
Tab. 2 Parameters of deep PCANet model

参数	尺度 1	尺度 2	尺度 3
图像颜色	灰度	灰度	灰度
图像分辨率	90×90	60×60	30×30
分块大小	9×9	11×11	13×13
2 级滤波器	12×8	12×8	12×8
直方图分块	35×35	15×15	7×7
WPCA	是	是	是

在 3 个尺度 90×90、60×60、30×30 上训练深度 PCANet,并评估模型的回归预测性能。其中,60×60、30×30 的训练集由尺度为 90×90 的训练集降采样获得。训练结束后,提取 3 个尺度深度 PCANet 的滤波器参数,并进行可视化,如图 9 所示。其中,第一行是尺度为 90×90 图像集对应的深度 PCANet 的 2 级滤波器参数可视化图,从左侧起,前 12 个是第 1 级滤波器参数的可视化图,后 8 个是第 2 级滤波器参数的可视化图;第二行和第三行分别是尺度为 60×60、30×30 图像集所对应的深度 PCANet 的 2 级滤波器参数可视化图。

由表 3 可知,进行线性 SVM 回归训练和预测时,模型 DPCA90、DPCA60 和 DPCA30 取得的相关系数分别为 0.8403、0.8146、0.7440;使用特征增强方法后,模型 DPCA90S、DPCA60S 和 DPCA30S 取得的相关系数分别为 0.8558、0.8164、0.7718。DPCAS 模型的相关系数明显大于 DPCA 模型的平均相关系数,说明使用特征增强方法能有效提升回归性能,而且 DPCA90S 取得了最好的回归性能,表明选择合适的图像尺度可提高人脸美丽评价结果。

表 4 是深度 PCANet 模型在提取特征、训练和测试线性 SVM 与随机森林回归器上的时间。其中,第一行是 PCANet 模型训练和提取 400 幅训练集图像特征的总时间。在深度 PCANet 模型上,400 幅训练集图像的特征提取时间最长为 207.47 s,单幅图像的特征提取时间平均为 0.52 s;400 幅训练集图像的特征提取时间最短为 61.12 s,单幅图像的特征提取时间平均为 0.15 s。第二行、第三行分别是线性 SVM 回归器的训练时间和测试时间。其中,线性 SVM 回归器最长训练时间为 2.24 s,最短训练时间为 0.63 s;最长测试时间为 0.55 s,最短测试时间为 0.16 s。第四行、第五行分别是随机森林回归器的训

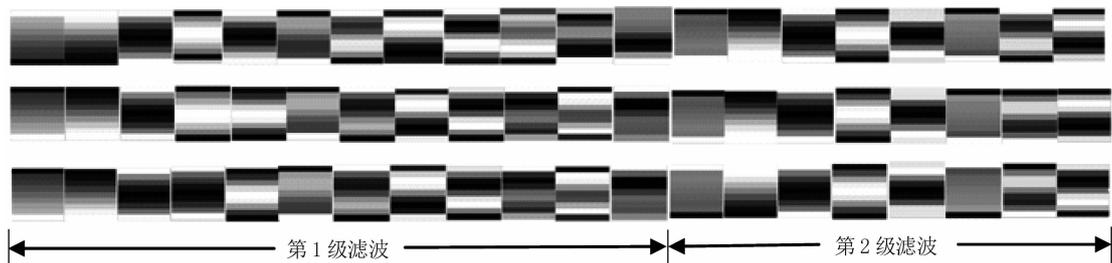


图 9 3 个尺度的滤波器可视化图

Fig. 9 Visualization of filters in 3 scales

表 3 回归预测结果  
Tab. 3 Results of regression

回归器	方法	DPCA90	DPCA90S	DPCA60	DPCA60S	DPCA30	DPCA30S
线性 SVM	1 折	0.8920	0.8965	0.8608	0.8573	0.7889	0.8128
	2 折	0.8813	0.9010	0.8714	0.8688	0.8122	0.7981
	3 折	0.7590	0.7734	0.6926	0.7114	0.5481	0.6420
	4 折	0.8576	0.8640	0.8125	0.8122	0.7771	0.7961
	5 折	0.8113	0.8441	0.8354	0.8322	0.7936	0.8098
	平均	<b>0.8403</b>	<b>0.8558</b>	<b>0.8146</b>	<b>0.8164</b>	<b>0.7440</b>	<b>0.7718</b>
随机 森林	1 折	0.8327	0.8328	0.7868	0.7789	0.7678	0.7670
	2 折	0.8816	0.8815	0.8405	0.8549	0.7292	0.7464
	3 折	0.7553	0.7555	0.6484	0.6503	0.6112	0.4995
	4 折	0.8338	0.8335	0.7519	0.7554	0.7016	0.7450
	5 折	0.7218	0.7219	0.7757	0.7573	0.7230	0.7159
	平均	0.8051	0.8050	0.7607	0.7594	0.7066	0.6948

练时间和测试时间。其中,随机森林回归器最长训练时间为 2000.60 s,最短训练时间为1388.50 s;最长测试时间为 37.44 s,最短测试时间为 25.71 s。显然,线性 SVM 回归器比随机森林回归器的运行速度更快。尽管 DCNN 可以通过增加网络深度,提升特征提取能力,改善回归性能,但是网络训练时间很长,一般需要几个小时或者几天,而本文提出的深度 PCA-Net 模型,使用 SVM 回归器从训练到测试总耗时不超过 10 分钟,而使用随机森林回归器总耗时不到 1 小时。因此,与现有算法相比,本文所构建的深度 PCA-

Net 模型具有网络训练时间短和图像特征提取速度快的特点,而且线性 SVM 回归器比随机森林回归器的训练时间更短,运行效率更高。

### 3.3.2 基于平均划分图像的实验结果

根据表 5,改变 SCUT-FBP 训练集和测试集的划分方式,将数据集分成 9 段,分别对每一段进行 5 折交叉分割,最后合并成 5 折交叉的训练集和测试集。表 5 中,第一行表示图像标签(Label)的 9 个划分区间,第二行表示每个区间内分割到的图像数量,9 个区间内的图像数量总和为 500。

表 4 不同尺度图像模型运行时间  
Tab. 4 The running time of models with different scale images

模型	DPCA90	DPCA90S	DPCA60	DPCA60S	DPCA30	DPCA30S
特征提取	199.32	<b>207.47</b>	127.36	141.43	<b>61.12</b>	80.80
SVM 回归/训练	<b>2.24</b>	2.13	1.53	1.48	<b>0.63</b>	0.84
SVM 回归/测试	<b>0.55</b>	0.51	0.39	0.37	<b>0.16</b>	0.21
随机森林/训练	<b>2000.60</b>	1948.30	1672.80	1896.20	<b>1388.50</b>	1519.70
随机森林/测试	<b>37.44</b>	36.21	31.30	35.75	<b>25.71</b>	28.32

表 5 标签分割  
Tab. 5 Label segmentation

标签区间	(0 1.9)	[1.9 2.3]	(2.3 2.4)	[2.4 2.5)	[2.5 2.7]	(2.7 3.1]	(3.1 3.5]	(3.5 4.4]	(4.4 5]
数量	24	123	39	41	82	94	44	31	22

模型名称和模型参数设置与3.3.1节相同。划分出训练集和测试集后,再次进行实验。由表6可知,进行线性SVM回归时,模型DPCA90、DPCA60和DPCA30取得的最高相关系数为0.8419,平均相关系数为0.8016;模型DPCA90S、DPCA60S和DPCA30S取得的最高相关系数为0.8548,平均相关系数为0.8078;使用随机森林回归时,最高相关系数为0.7831,低于线性SVM回归器的平均值0.8078。因此,深度PCANet模型在回归预测上,线性SVM回归器比随机森林回归器的预测性能好,且再次证明使用特征增强方法的DPCAS模型预测性能比不使用增强方法的DPCA模型好,而模型DPCA90S取得了最好的回归性能,再次说明选择合适的图像尺度可提高人脸美丽评价结果。表7是模型提取特征、

训练和测试线性回归器所消耗的时间。从表7与表4可知,无论是采用平均划分图像还是采用随机划分图像,深度PCANet模型具有网络训练时间短和图像特征提取速度快的特点。

### 3.4 回归测试效果对比

由3.3节的实验可知,在数据库规模比较小的情况下,深度PCANet仍具有较好的特征提取和泛化能力。由图6可知,SCUT-FBP人脸美丽数据库中极丑和极美的人脸图像数量较少,在3.3.1节的5折交叉验证实验中,极端美丽程度的人脸图像分配到训练集中的数量不固定,但结果显示,在深度PCANet回归预测上分别取得了最高的相关系数,进一步表明,深度PCANet对训练集中数量少的类别也具有较好的学习能力。

表6 回归预测结果  
Tab.6 Results of regression

回归器	方法	DPCA90	DPCA90S	DPCA60	DPCA60S	DPCA30	DPCA30S
线性 SVM	1折	0.8469	0.8578	0.8472	0.8454	0.8076	0.7987
	2折	0.8442	0.8538	0.7925	0.8038	0.6787	0.6906
	3折	0.8853	0.8895	0.8603	0.8521	0.7857	0.7892
	4折	0.8468	0.8514	0.8196	0.8266	0.7571	0.7552
	5折	0.7863	0.8213	0.7506	0.7512	0.7150	0.7302
	平均	0.8419	<b>0.8548</b>	0.8140	0.8158	0.7488	0.7528
随机 森林	1折	0.7678	0.7603	0.7812	0.7773	0.7730	0.7697
	2折	0.8192	0.8233	0.7647	0.7630	0.6619	0.6633
	3折	0.8327	0.8464	0.7905	0.8065	0.7604	0.7622
	4折	0.7688	0.7695	0.7053	0.7010	0.6426	0.6406
	5折	0.7072	0.7158	0.6879	0.6957	0.6848	0.6690
	平均	0.7791	<b>0.7831</b>	0.7459	0.7487	0.7045	0.7010

表7 不同尺度图像模型运行时间  
Tab.7 The running time of models with different scale images

模型	DPCA90	DPCA90S	DPCA60	DPCA60S	DPCA30	DPCA30S
特征提取	205.14	<b>206.24</b>	134.84	148.35	<b>64.55</b>	69.98
SVM回归/训练	2.36	<b>2.38</b>	1.20	1.51	<b>0.62</b>	0.67
SVM回归/测试	0.59	<b>0.60</b>	0.31	0.38	<b>0.16</b>	0.17
随机森林/训练	2076.60	<b>2114.20</b>	1618.30	1817.60	<b>1492.90</b>	1669.60
随机森林/测试	38.29	<b>39.35</b>	31.16	34.93	<b>27.64</b>	32.11

由表 8 可知,深度 PCANet 模型在测试性能上优于传统的浅层特征提取网络和 CNN 网络,取得的最大相关系数为 0.8558,高于深度卷积神经网络 GoogLeNet<sup>[17]</sup>、VGG-19<sup>[18]</sup>和 ResNet-101<sup>[18]</sup>,说明网络不是越深效果越好。本文提出的深度 PCANet 具有一定的优越性;虽然未达到深度卷积神经网络 CNN-3 的相关系数 0.88<sup>[11]</sup>,但是,使用深度卷积神经网络预训练时间长,且文献[11]所提方法需要多次级联微调,参数优化比较复杂。本文所构建的模型采用无监督方式进行预训练,无需调参优化,网络结构更简单和运行速度更快。

表 8 不同算法在 SCUT-FBP 数据库上的实验结果  
Tab.8 Results of different algorithms in SCUT-FBP

方法	相关系数
SVM Regression[10]	0.64
Gaussian Regression[9]	0.6482
PC[9]	0.8187
CNN-1[11]	0.76
CNN-2[11]	0.80
CNN-3 fine-tuning model[11]	0.82
CNN-3 cascaded fine-tuning model[11]	<b>0.88</b>
GoogLeNet[17]	0.81
VGG-19[18]	0.84
ResNet-101[18]	0.85
DPCA90S(本文)	<b>0.8558</b>

## 4 结论

本文构建了深度 PCANet 的人脸美丽预测模型,该模型能学习到图像的结构性特征,使其更稳健和更具表征能力。基于 SCUT-FBP 人脸美丽数据库的实验结果表明,深度 PCANet 模型不仅具有深度卷积神经网络卷积和层次堆叠的结构,而且还具有模型结构简单、运行速度快、特征提取时间短等特点,对深度学习在人脸美丽预测中的应用具有参考价值。同时,图像尺度为 90×90 的预测结果优于尺度为 60×60、30×30 的预测结果,对特征进行增强后的回归性能优于特征增强前的回归性能,说明选择合适的图像尺度与采用特征增强方法可提高人脸美丽评价结果。如何改进所构建的深度 PCANet 模型,提高其回归预测性能,值得进一步深入研究。

## 参考文献

- [1] Chang F, Chou C H. A Bi-prototype Theory of Facial Attractiveness[J]. *Neural Computation*, 2009, 21(3): 890-910.
- [2] Fan J, Chau K P, Wan X, et al. Prediction of facial attractiveness from facial proportions[J]. *Pattern Recognition*, 2012, 45(6): 2326-2334.
- [3] Chiang W C, Lin H H, Huang C S, et al. The cluster assessment of facial attractiveness using fuzzy neural network classifier based on 3D Moiré features[J]. *Pattern Recognition*, 2014, 47(3): 1249-1260.
- [4] Eishental Y, Dror G, Ruppim E. Facial attractiveness: beauty and the machine[J]. *Neural Computation*, 2006, 18(1): 119-142.
- [5] Mao H, Jin L, Du M. Automatic classification of Chinese female facial beauty using Support Vector Machine[C]// *IEEE International Conference on Systems*, 2009: 4842-4846.
- [6] Zhang D, Zhao Q, Chen F. Quantitative analysis of human facial beauty using geometric features[J]. *Pattern Recognition*, 2011, 44(4): 940-950.
- [7] Yan H. Cost-sensitive ordinal regression for fully automatic facial beauty assessment[J]. *Neurocomputing*, 2014, 129(5): 334-342.
- [8] Gray D, Yu K, Xu W, et al. Predicting Facial Beauty without Landmark[J]. *Computer Vision -ECCV 2010*, 6316: 434-447.
- [9] Xie D, Liang L, Jin L W, et al. SCUT-FBP: A Benchmark Dataset for Facial Beauty Perception[C]//*IEEE International Conference on Systems*, 2015, 9(6): 1821-1826.
- [10] Gan J, Li L, Zhai Y, et al. Deep self-taught learning for facial beauty prediction[J]. *Neurocomputing*, 2014(144): 295-303.
- [11] Xu J, Jin L W, Liang Z, et al. A new humanlike facial attractiveness predictor with cascaded fine-tuning deep learning model[J]. *Computer Science*, 2015, 70(1): 45-79.
- [12] Taigman Y, Yang M, Ranzato M, et al. DeepFace: Closing the gap to human-level performance in face verification[C]// *In Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition*, 2014: 1701-1708.
- [13] Chan T H, Jia K, Gao S, et al. PCANet: a simple deep learning baseline for image classification? [J]. *IEEE Transaction on Image Processing*, 2015, 24(12): 5017-5032.
- [14] Meguid M K, Levine M D. Fully automated recognition of spontaneous facial expressions in videos using random forest classifiers[J]. *IEEE Transactions on Affective Com-*

puting, 2014, 5(2): 141-154.

- [15] Sun Y, Wang X, Tang X. Deep convolutional network cascade for facial point detection[C]//In Proceedings of the 2013 IEEE Conference on Computer Vision and Pattern Recognition, 2013, 9(4): 3476-3483.
- [16] Gan J Y, Zhai Y K, Wang B. Unconstrained Facial Beauty Prediction Based on Multi-scale K-means [J]. Chinese Journal of Electronics, 2017, 26(3): 548-556.
- [17] 许杰. 基于深度学习的人脸美丽吸引力预测[D]. 广州: 华南理工大学, 2017.  
Xu Jie. Prediction of facial beauty attractiveness based on deep learning[D]. Guangzhou: South China University of Technology, 2017. (in Chinese)
- [18] Liu S, Li B, Fan Y, et al. Label DistributionBased Facial Attractiveness Computation by Deep Residual Learning[J]. IEEE Transactions on Multimedia, 2016, PP(9): 1-1.

#### 作者简介



**甘俊英** 女, 1964年生, 江西吉安人。五邑大学信息工程学院教授、硕士研究生导师。主要研究方向为生物特征识别。

E-mail: junyinggan@163.com



**姜开永** 男, 1984年生, 山东临沂人。五邑大学信息工程学院硕士研究生。主要研究方向为生物特征识别。

E-mail: yfdxgc@163.com



**曾军英** 男, 1977年生, 江西赣州人。五邑大学信息工程学院副教授、硕士研究生导师。主要研究方向为机器视觉和生物特征识别。

E-mail: zengjunying@126.com



**何国辉** 男, 1962年生, 江西萍乡人。五邑大学计算机学院教授、硕士研究生导师。主要研究方向为图像处理、虚拟现实和多媒体信息系统等。

E-mail: ghhe126@126.com



**谭海英** 女, 1994年生, 湖南衡阳人。五邑大学信息工程学院硕士研究生。主要研究方向为生物特征识别。

E-mail: haiyingtan@163.com