

# 多层特征融合的 PCANet 及其在人脸识别中的应用

陈飞玥<sup>1</sup>, 朱玉莲<sup>2</sup>, 陈晓红<sup>3</sup>

(1.南京航空航天大学计算机科学与技术学院, 江苏 南京 211106)

(2.南京航空航天大学公共实验教学部, 江苏 南京 211106)

(3.南京航空航天大学理学院, 江苏 南京 211106)

**[摘要]** 主成分分析网络(PCANet)是一种简单的深度学习网络模型,在图像识别领域具有很强的应用潜力.本文在 PCANet 的基础上,通过对 PCANet 结构进行分析,构造了一种基于多层特征融合的 PCANet(PCANet\_dense)网络模型.与单纯地只将前一层网络输出作为后一层网络输入的 PCANet 不同,PCANet\_dense 利用了不同层的特征信息.在 2 层网络结构中,它首先将原始图像特征和第 1 层网络的输出进行级联,然后将级联后的结果作为第 2 层网络的输入.而在 3 层网络结构中,它则将第 1 层和第 2 层网络的输出级联起来,作为第 3 层网络的输入.由于 PCANet\_dense 在训练每一层(除了第 1 层)时使用了更多信息,因此能够获得比原 PCANet 更好的效果.为了验证所提方法的有效性,本文使用 CMU PIE 数据集构建网络模型,并在 ORL、AR 和 Extended Yale B 3 个公开人脸数据集上对所提出方法的性能进行了测试,实验结果表明,本文提出的 PCANet\_dense 获得了比 PCANet 更好的性能.

**[关键词]** 人脸识别,主成分分析网络,密集网络,特征融合,多层特征融合的 PCANet

**[中图分类号]** TP391 **[文献标志码]** A **[文章编号]** 1001-4616(2021)02-0127-07

## Multi-stage Feature Fusion PCANet and Its Application to Face Recognition

Chen Feiyue<sup>1</sup>, Zhu Yulian<sup>2</sup>, Chen Xiaohong<sup>3</sup>

(1.College of Computer Science and Technology, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China)

(2.Fundamental Experimental Teaching Department, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China)

(3.College of Science, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, China)

**Abstract:** Principal component analysis network(PCANet) is a simple deep learning network model but shows strong application potential in many fields, especially in image recognition. In this paper, a new multi-stage feature fusion model based PCANet(PCANet\_dense) is constructed by analyzing the structure of PCANet. Different from PCANet which only takes the output of the former layer as the input of the latter layer, PCANet\_dense uses the feature information of different layers. In the two-layer network structure, it firstly cascades the original image features with the output of the first layer network, and then takes the fused features as the input of the second layer network. In the three-layer network structure, it cascades the output of the first and the second layer in series as the input of the third layer network. Due to the use of more information in each layer(except the first layer) in training process, the new model gets good performance. In order to verify the effectiveness of the proposed method, several comparative experiments on different face datasets, such as ORL, AR and Extended Yale B, are established. The network model of PCANet and PCANet\_dense are well constructed on CMU PIE dataset before all the experiments. The experimental results show that the proposed multi-stage feature fusion PCANet method achieves better performance than PCANet.

**Key words:** face recognition, PCANet, DenseNet, feature fusion, PCANet\_dense

人脸识别是目前应用最为广泛的生物特征识别技术之一.它在安全验证、银行系统、人机交互等领域均有广泛的应用.一般来说,人脸识别主要包括 3 个步骤:人脸检测、特征提取和人脸分类.人脸检测是将人脸区域从整个图像中分割出来,以便进行后续特征提取;特征提取是从每个人脸图像中提取最具鉴别能

收稿日期:2020-08-08.

基金项目:国家自然科学基金项目(61703206).

通讯作者:朱玉莲,副教授,研究方向:模式识别与人工智能. E-mail:lianyi\_1999@nuaa.edu.cn

力的特征;而人脸分类则是基于提取的特征来判别未知人脸图像是否属于某个人,或者属于哪个人. 本文的工作为特征提取和分类.

人脸识别也是计算机视觉中最具挑战性的问题之一. 因为人脸图像中不可避免地存在着不同的光照、姿态表情等变化,还可能出现不同程度的遮挡,从而出现大的类内变化和小的类间变化的情况,进而影响识别任务的性能. 为了克服这个问题,人们提出了许多基于局部特征的提取算法,典型的算法包括 Gabor 特征<sup>[1]</sup>、局部二值模式(LBP)<sup>[2]</sup>、尺度不变特征变换(SIFT)<sup>[3]</sup>和方向梯度直方图(HOG)特征<sup>[4]</sup>. 这些特征提取算法在一些特定的任务中取得了巨大的成功,但由于它们都是手工设计的特征,因此不能简单地应用于其他新任务. 根据“没有免费午餐定理”,一个高性能的算法必须利用数据的先验知识,因此,从数据本身来学习特征(基于学习的特征提取算法)成为弥补手动特征局限性的一种可行算法<sup>[5-6]</sup>.

在基于学习的特征提取算法中,卷积神经网络(CNN)<sup>[6]</sup>近年来受到了广泛的关注. CNN 通常由多个网络层组成,而每一网络层又由卷积滤波层、非线性处理层和特征池化层组成. 在过去的几年中,基于 CNN 体系结构,各种卷积神经网络的变体如 ResNet<sup>[7]</sup>、DeepFace<sup>[8]</sup>、FaceNet<sup>[9]</sup>等,相继被提出. 然而,基于 CNN 的网络结构通常非常复杂,如 ResNet<sup>[7]</sup>的层数高达 152,这导致了大量的结构参数(如滤波核参数)需要调整,这个过程非常耗时,而且复杂的网络结构通常需要一些额外的训练技巧才能获得良好的性能. 此外,复杂的网络结构通常需要大量的训练数据,才能保证通过深层网络获取对光照、表情等变化鲁棒的特征.

为了解决上述问题,一些研究者将注意力放在了简化训练过程上,典型的例子就是主成分分析网络(PCANet)<sup>[10]</sup>. PCANet 可视为 CNN 的简化版本,在 PCANet 中,卷积层的滤波核通过 PCA 算法来学习;非线性处理层使用最简单的二值化哈希编码,而最后的特征则通过直方图统计来表示. 尽管 PCANet 模型简单,但在某些图像识别领域取得了和 CNN 可比拟的性能<sup>[10]</sup>. 鉴于 PCANet 的成功,人们在 PCANet 的基础上做了各种改进<sup>[5,11-12]</sup>. 例如,Xi 等<sup>[11]</sup>提出了一种用于人脸识别的局部二值模式网络(LBPNet);Lei 等<sup>[5]</sup>提出了堆叠人脸图像描述符;Li 等<sup>[12]</sup>则提出了基于 RGBD(红、绿、蓝、深)图像的人体手势识别网络 SAE-PCA. 本文将研究如何在训练样本有限的情况下,使用 PCANet 提取到包含更丰富的有用的特征以实现性能的进一步提升.

PCANet 及上述 PCANet 的改进方法在构建卷积滤波核时只使用 PCA 的前  $k$  个最大的特征值对应的特征向量. 也就是说,只考虑了  $k$  个最大特征值对应的特征向量,而丢弃了剩余的特征向量. 实际上,被丢弃的特征向量中也存在一部分有用的辨别信息,丢弃它们会损失这部分辨别信息. 避免信息损失的一个简单方式是尽量多地使用 PCA 中的特征向量去构建卷积核,然而这样的操作将导致计算量及最后获得的特征维数大大增加. 最近,一种称为 DenseNet<sup>[13]</sup>的卷积神经网络模型被提出. DenseNet 的核心是每一层的输入均来自前面所有层的输出,并且通过特征的并行连接来实现特征重用,进而在参数更少的情况下获取比 CNN 更好的性能. 受 DenseNet 成功的启发,本文尝试将组合思想引入到 PCANet 中,对不同层次的特征信息进行融合,并相应地提出了基于多层特征融合的 PCANet(PCANet\_dense)模型. 与 PCANet 中只将第 1 层的输出反馈到第 2 层不同,PCANet\_dense 首先将原始图像特征和网络第 1 层的输出特征进行信道级的级联,然后将级联后的结果作为网络第 2 层的输入,因此 PCANet\_dense 比 PCANet 利用了更多的特征信息. 另外,本文也将 PCANet\_dense 扩展到 3 层网络结构中(在实验中将其命名为 PCANet\_dense3),即在保留 PCANet\_dense 前两层结构的基础上,将第 1 层与第 2 层的输出进行级联,然后将级联的结果作为第 3 层的输入. 为了验证所提方法的有效性,本文使用 CMU PIE 人脸数据库学习网络模型,并在 3 个人脸数据库(ORL、AR 和 Extended Yale B)上进行实验. 实验结果表明,相对于 PCANet,本文提出的 PCANet\_dense 和 PCANet\_dense3 不仅对光照、遮挡具有更强的鲁棒性,同时也适用于表情和姿态变化的测试图像.

## 1 基于多层特征融合的 PCANet

本节将详细描述基于多层特征融合的 PCANet 的网络层次结构和训练流程.

### 1.1 基于 2 层特征融合的 PCANet(PCANet\_dense)

和 PCANet 相似,PCANet\_dense 也包含 1 个输入层、2 个卷积层和 1 个输出层(如图 1 所示).

输入层:与 PCANet 的输入层相同. 首先对每个输入的训练图像进行块采样. 假设有  $N$  个大小为  $m \times n$

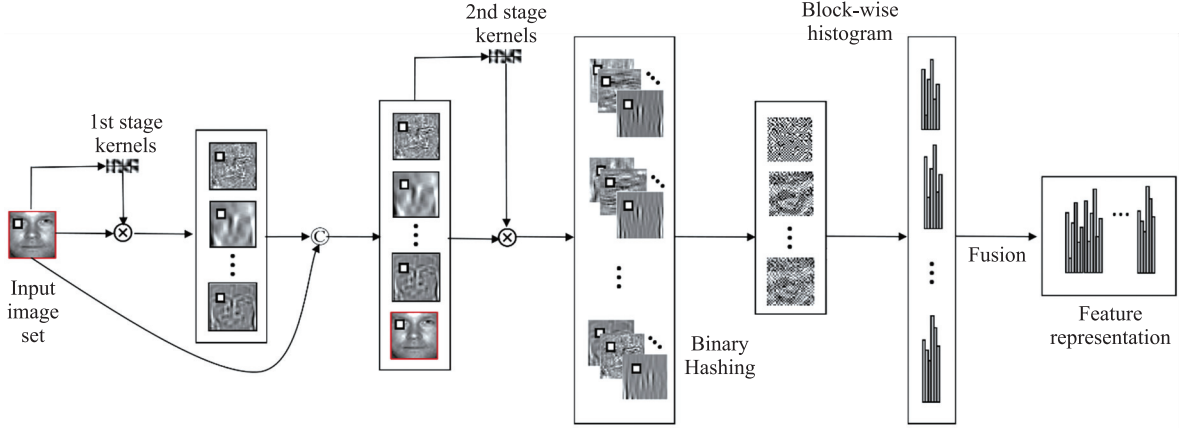


图1 PCANet\_dense 结构示意图 (其中⊕和⊗分别是信道级联和卷积运算)

Fig. 1 The illustration of structure of PCANet\_dense (⊕ and ⊗ is channel-wise concatenation and convolution operation, respectively)

的训练图像 (表示为  $\mathbf{I}^1 = \{\mathbf{I}_i^1\}_{i=1}^N$ ), 滤波核大小为  $k_1 \times k_2$ . 对于每幅图像  $\mathbf{I}_i^1$ , 首先以每个像素点为中心, 取 1 个  $k_1 \times k_2$  的块, 并对其向量化, 这样整个图像就可以构建成形如  $\mathbf{X}_i = [x_{i,1}, x_{i,2}, \dots, x_{i,\tilde{m}\tilde{n}}] \in R^{k_1 k_2 \times \tilde{m}\tilde{n}}$  的矩阵, 其中  $\tilde{m} = m - k_1 + 1, \tilde{n} = n - k_2 + 1$ . 然后从每个列中减去列平均值, 得到去均值化的矩阵  $\bar{\mathbf{X}}_i = [\bar{x}_{i,1}, \bar{x}_{i,2}, \dots, \bar{x}_{i,\tilde{m}\tilde{n}}]$ . 对所有  $N$  个训练图像进行此操作, 并将所有的  $\bar{\mathbf{X}}_i$  放在一起, 形成大矩阵  $\mathbf{X} = [\bar{\mathbf{X}}_1, \bar{\mathbf{X}}_2, \dots, \bar{\mathbf{X}}_N] \in R^{k_1 k_2 \times N\tilde{m}\tilde{n}}$ , 随后将  $\mathbf{X}$  作为 PCANet\_dense 的输入.

第 1 层卷积: 设第 1 层卷积层的滤波核个数为  $L_1$ . 首先对  $\mathbf{X}\mathbf{X}^T$  进行 SVD 分解, 并相应得到  $\mathbf{X}\mathbf{X}^T$  的前  $L_1$  个主特征值所对应的特征向量, 即  $\{\mathbf{V}_1^1, \mathbf{V}_2^1, \dots, \mathbf{V}_{L_1}^1\}$ . 随后将每个特征向量矩阵化表示为大小为  $k_1 \times k_2$  的矩阵, 并把每一个矩阵都看作是 PCANet\_dense 的第 1 层的滤波核. 也就是说, 第 1 层卷积的滤波核可以表示为如下形式:

$$\mathbf{W}_l^1 = \text{mat}_{k_1 k_2}(\mathbf{V}_l^1) \in R^{k_1 \times k_2}, l = 1, 2, \dots, L_1, \quad (1)$$

式中,  $\text{mat}$  是将向量  $\mathbf{V} \in R^{k_1 k_2}$  转换为矩阵  $\mathbf{W} \in R^{k_1 \times k_2}$  的函数. 因此, 对于图像  $\mathbf{I}_i^1$ , 第 1 层卷积后的输出就可以通过如下表达式得到:

$$\mathbf{I}_{i,l}^1 = \mathbf{I}_i^1 * \mathbf{W}_l^1, i = 1, 2, \dots, N; l = 1, 2, \dots, L_1, \quad (2)$$

式中,  $*$  为卷积符号. 很明显, 对于每一个输入图像, 经过第 1 层卷积后, 能得到  $L_1$  个输出映射.

第 2 层卷积: 设第 2 层卷积的 PCA 滤波核个数为  $L_2$ . 与原始 PCANet 不同, PCANet\_dense 将原始输入图像与第 1 卷积层的输出在信道上进行级联, 然后将级联后的结果作为第 2 层卷积的输入, 即 PCANet\_dense 的第 2 卷积层的输入为:  $\mathbf{I}^2 = [\mathbf{I}_1^1, \dots, \mathbf{I}_N^1, \mathbf{I}_{1,1}^1, \dots, \mathbf{I}_{1,L_1}^1, \dots, \mathbf{I}_{N,1}^1, \dots, \mathbf{I}_{N,L_1}^1] \in R^{k_1 k_2 \times (L_1+1)N\tilde{m}\tilde{n}}$ , 其中  $\mathbf{I}_i^1$  和  $\mathbf{I}_{i,l}^1$  ( $l = 1, 2, \dots, L_1$ ) 分别表示第  $i$  ( $i = 1, 2, \dots, N$ ) 个原始图像和其第 1 层卷积后的输出. 在  $\mathbf{I}^2$  的输入下, 重复几乎与第 1 层一样的操作过程, 就可以学习得到第 2 层的滤波核, 记为  $\{\mathbf{W}_1^2, \mathbf{W}_2^2, \dots, \mathbf{W}_{L_2}^2\}$ . 因此, 对于  $\mathbf{I}^2$  的每一个输入, 都有  $L_2$  个输出与之对应:

$$\mathbf{O}_i^l = \{\mathbf{I}_i^2 * \mathbf{W}_j^2\}_{j=1}^{L_2}, i = 1, 2, \dots, (L_1+1)N; l = 1, 2, \dots, L_1+1. \quad (3)$$

输出层: 输出层包括非线性处理、分块直方图和特征拼接. 在第 2 层卷积之后, 每个原始输入图像  $\mathbf{I}_i^1$  有  $L_1+1$  组输出映射, 而每个组包含  $L_2$  个特征映射, 也就是说, 共有  $(L_1+1)L_2$  个特征映射产生. 对于每个组, 首先使用 Heaviside 阶跃函数  $H$  对每个输出映射进行二值化, 然后使用二值哈希编码将每组的  $L_2$  个输出映射转换成 1 个元素值为  $0 \sim 2^{L_2}-1$  之间的输出矩阵, 即

$$\mathbf{\varpi}_i^l = \sum_{j=1}^{L_2} 2^{j-1} H(\mathbf{O}_i^l), l = 1, 2, \dots, L_1+1. \quad (4)$$

所以, 对于每个训练图像  $\mathbf{I}_i^1$ , 一共可以得到  $L_1+1$  个输出矩阵  $\mathbf{\varpi}_i^l$  ( $l = 1, 2, \dots, L_1+1$ ). 随后将每个  $\mathbf{\varpi}_i^l$  划分为  $B$  个块, 并统计每个块的直方图信息, 然后将这  $B$  个直方图特征进行串联, 表示为  $B\text{hist}(\mathbf{\varpi}_i^l)$ . 最后, 将所有的  $B\text{hist}(\mathbf{\varpi}_i^l)$  ( $l = 1, 2, \dots, L_1+1$ ) 连接成 1 个向量, 得到 1 个扩展的直方图特征向量, 而该向量将被

视为输入图像  $I_i^1$  的特征表示:

$$f_i = [Bhist(\varpi_i^1), Bhist(\varpi_i^2), \dots, Bhist(\varpi_i^{L_1+1})]^T \in R^{2^{L_1+1}B}.$$

(5)

1.2 基于 3 层特征融合的 PCANet(PCANet\_dense3)

PCANet\_dense 虽只是 2 层的网络模型,但同 PCANet 一样,它很容易扩展到 3 层或更多层.在此,通过在 PCANet\_dense 的基础上添加第 3 层卷积层,给出 PCANet\_dense 的 3 层版本(PCANet\_dense3).简单地说,PCANet\_dense3 在学习第 3 层卷积前,将第 1 层和第 2 层的输出进行信道级联,然后将级联后的结果作为第 3 层的输入. PCANet、PCANet\_dense 和 PCANet\_dense3 的流程简单比较如图 2 所示.

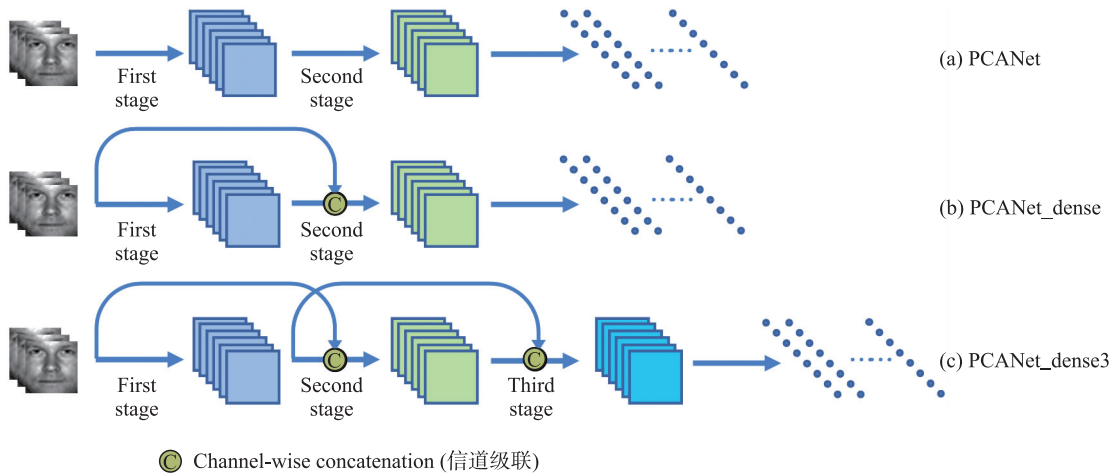


图 2 PCANet、PCANet\_dense 和 PCANet\_dense3 流程概述  
Fig. 2. Overview of PCANet, PCANet\_dense and PCANet\_dense3

2 实验结果与分析

为了评估文中所提方法的有效性,本文在人脸识别任务上进行了实验.实验中 PCANet、PCANet\_dense 和 PCANet\_dense3 使用 CMU PIE 人脸数据库学习网络模型,FaceNet 选用已在 VGGFace2 训练好的模型.实验分别在 ORL、AR 和 Extended Yale B 数据集上对 PCA、FaceNet、PCANet、PCANet\_dense 和 PCANet\_dense3 进行性能比较.

CMU PIE 数据库包含 68 个人不同姿势、不同照明条件和不同拍摄条件下共 41 368 张图像.本文使用的是 1 个子集,其仅包含 68 个人正面姿势的图像,并将所有图像分辨率调整为 100×82.在实验中,各项基本参数的设置如下:对于 PCA 算法,选取前 95%的主成分作为特征表示;对于 PCANet、PCANet\_dense 和 PCANet\_dense3,每层滤波核的大小均为 5×5;而滤波核的数目均为 8;所有实验用欧氏距离最近邻分类器进行分类.

2.1 在 ORL 数据集上的实验结果

ORL 数据库是剑桥大学 AT&T 实验室收集的人脸图像数据集,包括来自 40 个人的 400 张人脸图像,每个人有 10 张图像,所有图像分辨率为 112×92.对每个人来说,10 张照片在不同的时间拍摄并包含面部表情(睁眼、闭眼、微笑、中性)、面部细节(戴眼镜或不戴眼镜)和人脸尺度等变化(最多 10%左右).此外,所拍摄的图像允许脸部倾斜和旋转达 20°.在实验中,分别从每个人的图像中随机选取 5、6、7 和 8 幅人脸图像形成训练集,其余图像作为测试集,实验重复 10 次,取平均作为最后的结果.实验结果如表 1 所示.

由表 1 可以看出:(1) PCANet\_dense 比 PCANet 具有更好的性能,特别是当训练样本数足够多时(7 个或 8 个),PCANet\_dense 分别获得 97.5% 和 100% 的识别准确率,相对于 PCANet 的 95.53% 和 98.75%,PCANet\_dense 至少提高了 1%,这说明对不同层的特征进行组合是有效的;(2) PCANet\_dense3

表 1 ORL 数据集的识别率  
Table 1 Recognition rates on ORL dataset

模型	识别率/%			
	8	7	6	5
PCA	97.00	96.33	95.38	93.95
FaceNet	100	100	100	100
PCANet	98.75	95.53	95.63	95.50
PCANet_dense	100	97.50	95.63	95.50
PCANet_dense3	100	98.75	98.33	96.00



在所有情况下都优于 PCANet\_dense;在每类 6 个训练样本时,PCANet\_dense3 取得了 98.33% 的准确率,比 PCANet\_dense 高了 2.7%,这说明增加 PCANet\_dense 的层数,更有利于性能的提升;(3)FaceNet 在 ORL 数据集上表现得非常优秀,在所有情况下均获得了 100% 的识别率.这主要是因为:①FaceNet 是通过大量样本训练得到的真正的深度网络(深度可达 447);②ORL 中的图像受姿态、光线和遮挡等影响较少.虽然 PCANet\_dense3 在 ORL 上并没有获得与 FaceNet 可比拟的性能,但值得说明的是,PCANet\_dense3 仅仅是一个 3 层网络,如果增加层数将会进一步提高整体的性能.

## 2.2 在 AR 数据集上的实验结果

AR 人脸数据库是一个非常具有挑战性的数据库,它由 126 个人(70 名男性和 56 名女性),每人 26 张共计 3 000 多张正面图像组成.每人的照片在间隔 2 周的 2 个不同的时间段(session)拍摄.图 3 显示了 AR 数据库中的一些样本,每个 session 中的图像从左到右的详细信息为:第 1 张中性表情、第 2-4 张表情大幅变化、第 5-7 张照明变化、第 8-10 张为太阳镜遮挡、第 11-13 张为围巾遮挡.本文使用了包含 100 个人(50 名男性和 50 名女性)共 2 600 张图像的子集,并将原始图像的大小调整为  $66 \times 48$ .实验选择 session 1 的所有图像进行训练,并将 session 2 图像根据变化类别的不同分为 4 个测试子集(Exps:session 2 第 2-4 张;Illum:session 2 第 5-7 张;Sunglass:session 2 第 8-10 张;Scarf:session 2 第 11-13 张),以便测试在不同图像变化下的性能.



图 3 AR 数据库上某个人的样本

Fig. 3 Some samples of one person from AR database

图 4 列出了不同方法在 4 个子集上的识别准确率.从图 4 可以看出:(1)在 Exps 和 Illum 测试子集上,所有对比方法均获得了较好的性能.尤其是 FaceNet 达到了 99.67% 的识别性能;本文提出的 PCANet\_dense 和 PCANet\_dense3 也达到了 99% 的性能,高于 PCANet 的实验结果.(2)在 Sunglass 和 Scarf 测试子集上,PCA 和 FaceNet 模型的准确率相对于 Exps 和 Illum 明显下降.例如在 Scarf 子集上,由于人脸图像受到严重遮挡,PCA 的识别率只有 18.00%,FaceNet 也仅达到 91.67% 的识别性能;而 PCANet\_dense 和 PCANet\_dense3 分别取得了 97.00% 和 98.67% 的准确率,不仅高于 PCA 和 FaceNet,比 PCANet 也提升了至少 2.67%.这说明我们提出的方法对受遮挡的人脸图像数据表现得更为鲁棒.

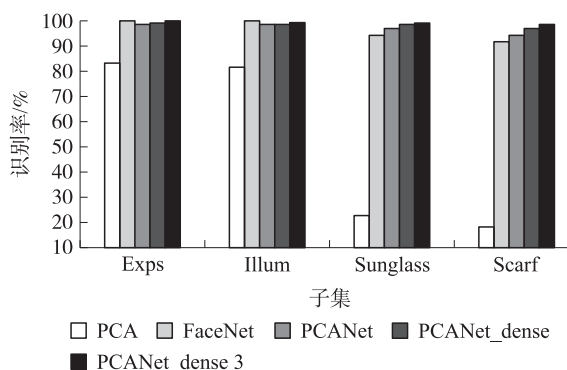


图 4 AR 数据集上各模型的识别率

Fig. 4 Recognition rates of each model on AR dataset

## 2.3 在 Extended Yale B 数据集上的实验结果

Extended Yale B 数据集包含了 38 个人的 2 414 张正面照片.所有的图像都是在实验室控制的各种光线下拍摄的.因此,该数据集主要用来测试算法对光照的鲁棒性.实验中,将原始图像大小从  $192 \times 168$  调整为  $48 \times 42$ ,并参考文献[14]中的策略,根据光源方向的角度(即光照方位角)大小将整个数据集分成 5 个不相交的子集(subset 1-subset 5).从图 5 所示的部分样本可看出从 subset 1 到 subset 5 图像的光照方位角逐渐增大.实验选择子集 subset 1 进行训练,其余 4 个子集(subset 2-subset 5)分别进行测试.实验的结果如表 2 所示.

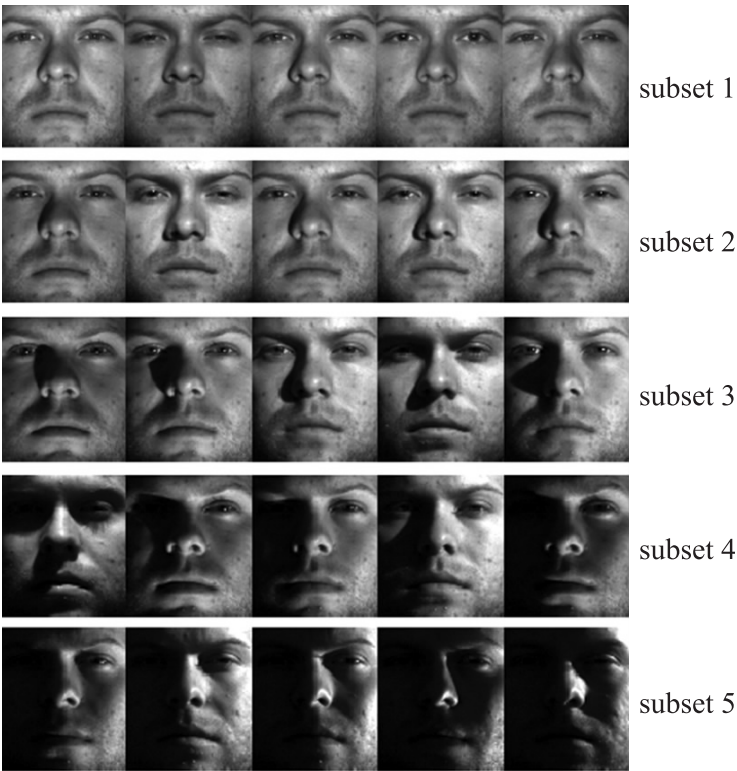


图 5 Extended Yale B 数据集的子集示例  
Fig. 5 Overview of subsets of Extended Yale B dataset

从表 2 可以看到:(1)同 PCA、FaceNet 和 PCANet 相比,PCANet\_dense 和 PCANet\_dense3 在性能上都有一定提升. 尤其是 PCANet\_dense3,在 subset 5 上达到了 98.03%的准确率,比 PCANet 多了 5%;而 PCA 和 FaceNet 在 subset 5 上只取得了 3.51%和 33.68%的准确率. (2)FaceNet 对光线变化的鲁棒性极不理想,这主要是因为在使用 VGGFace2 数据集学习 FaceNet 网络模型时并没有考虑光线等变化对图像干扰的问题. (3)当测试样本的光照方位角较小时,如 subset 2,所有对比方法都取得了比较好的效果,识别准确率都在 94%以上;随着光照方位角逐渐增大,所有方法的准确率都出现了不同程度地下降. 从 subset 2 到 subset 5,PCA、FaceNet 和 PCANet 的识别率分别下降了约 93%、60%和 7.2%;而文中提出的 PCANet\_dense3 下降幅度仅为 2%,远低于其他方法. 说明 PCANet\_dense3 对于识别受严重光照变化影响的人脸图像具有很好的鲁棒性.

2.4 讨论

由公式(5)可知,每张输入图像最终得到的特征表示向量,其维度与网络的卷积核个数和分块个数  $B$  有关. 在上述实验中,PCANet、PCANet\_dense 和 PCANet\_dense3 这 3 种方法的卷积层都采用了相同的卷积核大小与个数,输出层分块个数也相同. 然而,PCANet\_dense3 是 3 层结构,与 2 层结构的 PCANet 和 PCANet\_dense 相比在总体上产生了更多的卷积核,因此,PCANet\_dense3 所提取的特征维数肯定大于后两者. 本节将在 AR 数据集上进行实验,以讨论在最终特征维数尽量不变的情况下(相对于 PCANet\_dense)PCANet\_dense3 的性能. 在输出层分块个数相同的前提下,设 PCANet(或 PCANet\_dense)网络的 2 层滤波核个数分别为  $L_1$  和  $L_2$ ,PCANet\_dense3 中每层滤波核个数分别为  $L_3$ 、 $L_4$  和  $L_5$ ,若要保证 PCANet\_dense3 的特征维数尽量不变只需满足关系  $L_1 \cong (L_3+1)L_4+1$  且  $L_2=L_5$  即可. 设 PCANet 与 PCANet\_dense 的参数为:  $L_1=L_2=8$ ,因此 PCANet\_dense3 的参数可为:  $L_3=2$ , $L_4=2$ , $L_5=8$ . 实验均在 AR 数据集上进行建模和训练,相应的结果如表 3 所示.

表 2 Extended Yale B 数据集的识别率

Table 2 Recognition rates on Extended Yale B dataset

模型	识别率/%			
	subset 2	subset 3	subset 4	subset 5
PCA	97.22	70.56	12.14	3.51
FaceNet	94.44	76.94	60.24	33.68
PCANet	100	100	96.90	92.76
PCANet_dense	100	100	96.77	93.59
PCANet_dense3	100	100	99.75	98.03

从表 3 中的数据可知,降低每层的滤波核个数的确影响 PCANet\_dense3 的性能,但相对于 2 层的 PCANet 和 PCANet\_dense,整体性能仍然有提升,这也说明了在不增加特征维数的情况下增加 PCANet 的层数对性能的提高是有效的. 另外,根据 3 种不同网络的卷积核个数可以推算出,PCANet、PCANet\_dense 和 PCANet\_dense3 最终所产生的特征维数是 8:9:7,这说明 PCANet\_dense3 实际是以偏低的维度取得了表 3 中展现出的性能. 在实际应用中,可以考虑放宽特征维度限制,选择合适的卷积核数以提高 PCANet\_dense3 的性能.

表 3 维度限制下 AR 数据集的识别率

Table 3 Recognition rates under dimensional constraints on AR dataset

模型	识别率/%			
	Exps	Illum	Sunglass	Scarf
PCANet	98.67	98.67	97.00	94.33
PCANet_dense	98.67	99.00	97.00	95.67
PCANet_dense3 (no constraints)	98.33	99.33	98.00	96.33
PCANet_dense3 (dimensional constraints)	98.67	99.33	97.33	94.33

### 3 结论

本文在 PCANet 的基础上,结合 DenseNet 的思想,通过融合不同层的特征提出了 PCANet\_dense 模型. 该模型的核心是每个卷积层的输入为前两层输出的级联特征(当模型仅有 2 层时,第 2 层输入为第 1 层输出与输入层的级联特征),因此相对于 PCANet 而言,该模型使用了更多的特征信息. 本文分别构建了 2 层和 3 层 PCANet\_dense 模型,并在人脸识别任务上进行了实验. 实验结果表明,所提出的模型均取得了较好的效果,尤其当测试图像包含严重遮挡或受到严重光照干扰的时候,文中的方法获得了比 PCANet、FaceNet 等更为鲁棒和稳定的性能.

### [参考文献]

- [1] LIU C, WECHSLER H. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition[J]. IEEE transactions on image processing, 2002, 11(4): 467-476.
- [2] AHONEN T, HADID A, PIETIKAINEN M. Face description with local binary patterns: application to face recognition[J]. IEEE transactions on pattern analysis and machine intelligence, 2006, 28(12): 2037-2041.
- [3] LOWE D G. Object recognition from local scale-invariant features[C]//International Conference on Computer Vision (ICCV). Kerkyra, Greece: IEEE, 1999: 1150-1157.
- [4] DALAL N, TRIGGS B. Histograms of oriented gradients for human detection[C]//Computer Vision and Pattern Recognition (CVPR). San Diego, CA, USA: IEEE, 2005: 886-893.
- [5] LEI Z, YI D, LI S Z. Learning stacked image descriptor for face recognition[J]. IEEE transactions on circuits and systems for video technology, 2015, 26(9): 1685-1696.
- [6] LECUN Y, BOTTOU L. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11): 2278-2324.
- [7] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: IEEE, 2016: 770-778.
- [8] TAIGMAN Y, MING Y, RANZATO M, et al. DeepFace: closing the gap to human-level performance in face verification[C]//Computer Vision and Pattern Recognition (CVPR). Columbus, OH, USA: IEEE, 2014: 1701-1708.
- [9] SCHROFF F, KALENICHENKO D, PHILBIN J. FaceNet: a unified embedding for face recognition and clustering[C]//Computer Vision and Pattern Recognition (CVPR). Boston, MA, USA: IEEE, 2015: 815-823.
- [10] CHAN T, JIA K, GAO S, et al. PCANet: a simple deep learning baseline for image classification[J]. IEEE transactions on image processing, 2015, 24(12): 5017-5032.
- [11] XI M, CHEN L, POLAJNAR D, et al. Local binary pattern network: a deep learning approach for face recognition[C]//International Conference on Image Processing (ICIP). Phoenix, AZ, USA: IEEE, 2016: 3224-3228.
- [12] LI S, YU B, WU W, et al. Feature learning based on SAE-PCA network for human gesture recognition in RGBD images[J]. Neurocomputing, 2015, 151: 565-573.
- [13] HUANG G, LIU Z, LAURENS V, et al. Densely connected convolutional networks[C]//Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA: IEEE, 2017: 2261-2269.
- [14] LEE K, HO J, KRIEGMAN D J. Acquiring linear subspaces for face recognition under variable lighting[J]. IEEE transactions on pattern analysis and machine intelligence, 2005, 27(5): 684-698.

[责任编辑: 丁 蓉]