

特征匹配融合结合改进卷积神经网络的人脸识别

李佳妮, 张宝华*

内蒙古科技大学信息工程学院, 内蒙古 包头 014010

摘要 提出一种特征匹配融合结合改进卷积神经网络(CNN)的图像识别方法。针对局部二值模式(LBP)算子提取的特征信息局限以及不能准确描述图像轮廓信息的问题,使用梯度方向直方图(HOG)和LBP分层特征融合的方法在卷积神经网络中对训练集进行特征提取,再将匹配提取的特征图像输入改进的卷积神经网络进行训练、识别。以ReLU为激活函数,输出层用Softmax分类器,并通过TensorFlow框架进行训练,在ORL、YALE和CAS-PEAL人脸库上进行人脸识别仿真,该方法识别率分别达到了99.2%、98.7%、97.2%,高于其他对比算法。

关键词 机器视觉; 卷积神经网络; 特征匹配; 特征融合; 人脸识别

中图分类号 O436

文献标识码 A

doi: 10.3788/LOP55.101504

Face Recognition by Feature Matching Fusion Combined with Improved Convolutional Neural Network

Li Jiani, Zhang Baohua*

College of Information Engineering, Inner Mongolia University of Science and Technology,
Baotou, Inner Mongolia 014010, China

Abstract This paper presents an image recognition method based on feature matching fusion and improved convolutional neural network. Aiming at the problem that the texture features extracted by the local binary pattern (LBP) descriptors are limited and cannot describe the image edge and direction information effectively, the feature extraction of the training set is performed in the convolutional neural network by the histogram of oriented gradient (HOG) and LBP hierarchical feature fusion method. Then the extracted feature pictures are input into the improved convolutional neural network for training and recognition. The simulations are performed on ORL, YALE and CAS-PEAL face databases with ReLU as the activation function and the output layer with the Softmax classifier, and trained on the TensorFlow framework. The recognition rate of the proposed method reaches 99.2%, 98.7%, and 97.2% respectively, which is higher than other algorithms for comparison.

Key words machine vision; convolutional neural network; feature matching; feature fusion; face recognition

OCIS codes 150.1135; 100.2000; 100.4996; 330.5000

1 引言

当前,人工智能技术飞速发展,在很多领域中得以运用。机器快速检测并识别目标的问题,一直是机器视觉领域的热点^[1]。目前常用的生物识别技术主要有指纹识别、虹膜识别、语音识别以及人脸识别^[2]。具有唯一性的人脸是最自然的身份识别方

式。人脸识别相对于其他识别方法的优点是方便快捷,安全性高,易于实现和安装。目前已在门禁系统、摄像系统、考勤系统和金融系统等领域得到了广泛的应用。

特征提取是人脸识别最核心的步骤之一,准确并完整地描写人脸信息是提高识别精度的关键。全局特征提取和局部特征提取是特征提取的两种常用

收稿日期: 2018-04-25; 修回日期: 2018-05-03; 录用日期: 2018-05-14

基金项目: 国家自然科学基金(61663036,61261028)、国家海洋局海洋遥测工程技术研究中心创新青年基金(2014003)、内蒙古自治区高等学校“青年科技英才支持计划”青年科技骨干项目(NJYT-14-B11)、内蒙古自然科学基金(2014MS0610)、内蒙古科技大学创新基金(2014QNGG07)

* E-mail: zbh_wj2004@imust.cn

方式。全局特征提取方法有主成分分析(PCA)^[3]、线性判别分析(LDA)^[4]、独立成分分析(ICA)^[5]等。全局特征虽然可以表示完整的人脸内部信息,但在光照、姿态变化和遮挡等条件下表达信息的效果差。局部特征提取方法有局部二值模式(LBP)^[6]、尺度不变特征变换(SIFT)^[7]以及梯度方向直方图(HOG)^[8]等。因为局部特征对外界变化有很好的稳定性,所以局部特征提取方法在人脸识别中有很好的应用。

LBP算子是一个非参数的局部纹理描述运算符,可以有效获取图像中的细节特征^[9],具有很强的分类能力。但是,将整个图像的LBP的统计直方图作为特征,获取的特征较为模糊。其局限性在于 3×3 局部邻域不能捕获大量的结构特征,且相邻区域的LBP码与相邻像素有关,故对随机噪声的适应性差。与LBP算子不同,HOG算子根据块内统计提取特征。由于HOG描述子本身对光照和方向信息不敏感,所以可以有效获取人脸图片的轮廓和形状信息。Albiol等^[10]提出HOG特征与用于人脸识别的弹性图相结合,证实HOG特征对外部干扰有较强的稳健性^[11]。如何完整地提取LBP描述子所捕获的特征以及HOG描述子没有表达出的信息,且在人脸图像大数据的前提下,得到更高的人脸识别率?这一问题仍在探索中。

Ojala等^[12]采用LBP方法解决光照对人脸识别的影响,实现人脸识别对光照的高稳健性,但缺乏图像尺度、旋转等处理。Chen等^[13]采用高维LBP特征的方法,通过增加特征维度提高算法精度,但算法计算时间较长,实时性差。Pang等^[14]采用改进的HOG特征进行人脸识别,重复利用块内特征和重复单元特征进行特征提取。Yu等^[15]采用基于块和图像的HOG和LBP融合方法,将其与BoW模型结合实现图像分类检索识别功能。经过国内外科

研工作者多年的研究,人脸识别技术突飞猛进,但是之前的研究仅仅是利用LBP和HOG其中一种方式进行人脸特征提取,得到的人脸识别率较低;或者融合LBP和HOG提取特征后,利用其他模型实现图像分类及图像分类检索识别功能。研究成果没有充分利用两者特征融合结合卷积神经网络去处理人脸分类识别。

为了进一步提取更加完整的图像特征,本文在上述方法的启发下提出特征匹配融合结合改进卷积神经网络的人脸识别。首先,将分块后的人脸图像进行二进制编码,提取LBP统计直方图特征,之后将所有分块区域的直方图特征顺序串联起来,形成每一层特征获取图像的特征矩阵向量。对灰度图像再次进行HOG特征的提取,分层LBP特征向量与HOG特征向量串联级联,形成人脸图像的融合特征。在人脸图像大数据的前提下,为了提高人脸识别的准确率并减小时间复杂度,通过卷积神经网络(CNN)进行学习。CNN通过局部感受野和权值共享,减少训练时间,提高识别率。研究结果说明该算法能明显提高人脸识别率。

2 扩展卷积神经网络

深度学习技术中的卷积神经网络广泛应用于图像分类、目标检测、语义分割等领域^[16-19]。卷积神经网络本质上类似于一个多层感知机,具有良好的并行处理、自学习和泛化能力^[20]。卷积神经网络的经典模型结构是LeNet-5,如图1所示。原始图像首先在卷积层与滤波器进行卷积,得到若干特征图后,通过下采样层对特征进行模糊,逐层提取特征完毕后,通过全连接层输出用于识别图像的特征。卷积神经网络主要由三部分构成,分别为输入层、隐层和输出层。在隐层中包含多个卷积层和子采样层构成的重复结构。

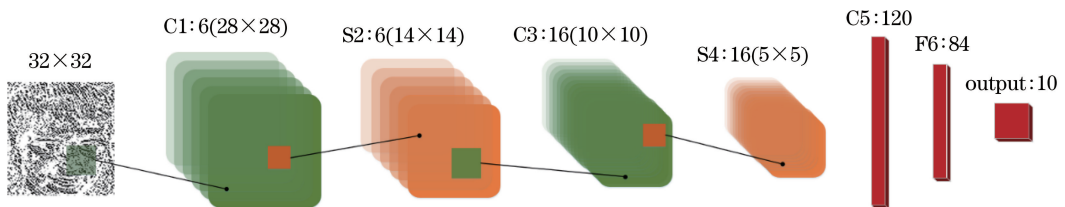


图1 LeNet-5 模式结构图

Fig. 1 LeNet-5 model structure

针对现有算法模式结构存在的计算量较大、时间复杂度高、图像纹理模糊等问题,在原有的模型结构上提出扩展模型^[21]。He等^[22]讨论了CNN中结

构深度、特征面数目以及卷积核大小之间的关系,实验证明卷积核小且较深的CNN结构比卷积核大且较浅的结构更好。然而在卷积核不变的前提下,网

络深度持续增加并且适当减少特征面的数目时,网络状态也会饱和。综上所述,在 CNN 中,深度变大、特征面数量增加,则学习能力越强,网络性能越好,但是缺点是计算复杂,容易出现过拟合现象。所

以,改进 LeNet-5 结构时,将深度加大,即增加一层卷积层和一层池化层;减小卷积核尺寸并增加学习率,使得训练能够获得好模型的同时减少时间,降低计算复杂度,如图 2 所示。

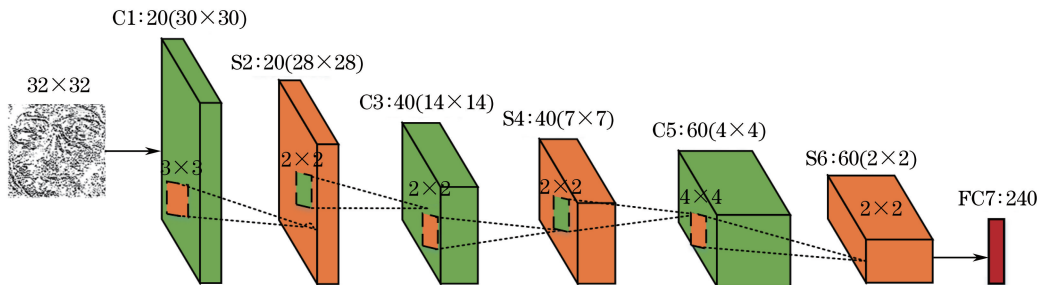


图 2 扩展卷积神经网络结构图

Fig. 2 Extended convolutional neural network structure

卷积操作对有形变的图片有很好的稳定性。假设卷积层的输入是 $p \times p$ 维矩阵 \mathbf{M} ,卷积核为 $q \times q$ 维矩阵 \mathbf{N} ,利用卷积核对矩阵 \mathbf{M} 中的感受野使用步长为 1 的卷积操作,得到 $(p-q+1) \times (p-q+1)$ 维特征图矩阵 \mathbf{F} 。卷积层的计算公式为

$$\mathbf{F} = \text{sigmoid}[\text{con}(\mathbf{M}, \mathbf{N}) + \mathbf{B}_1], \quad (1)$$

式中: sigmoid 为激活函数; con 为卷积函数; \mathbf{B}_1 为偏置矩阵。原始图像和进入池化层后得到的特征图像均可当作输入层的图像。卷积层进行优化的目的是提取特征准确度增加,并可以有效提高卷积神经网络的特征提取和分类能力。

池化的实质是池化函数 $\Phi_0(\cdot)$ 作用于池化域的过程。假设原始特征为矩阵 \mathbf{J} ,特征图的池化域为矩阵 \mathbf{K} ,池化得到的结果加上偏置 \mathbf{B}_2 进行计算,遍历特征图的池化域,获得自采样特征图 \mathbf{D} ,计算公式为

$$\mathbf{D} = \Phi_0(\mathbf{J}) + \mathbf{B}_2. \quad (2)$$

平均池化和最大池化是池化模型中最为普遍的模式。假设输入特征图矩阵 \mathbf{L} ,子采样池化域为 $c \times c$ 的矩阵 \mathbf{O} ,偏置为 b_2 ,得到的子采样特征图为

\mathbf{S} ,池化移动步长为 d 。平均池化和最大池化的计算公式为

$$S_{ij} = \frac{1}{d^2} \left(\sum_{i=1}^d \sum_{j=1}^d L_{ij} \right) + b_2, \quad (3)$$

$$S_{ij} = \max_{i=1, j=1}^d (L_{ij}) + b_2, \quad (4)$$

式中: $\max_{i=1, j=1}^d (L_{ij})$ 为从输入特征图 \mathbf{L} 的大小为 $c \times c$ 的池化域中取出的最大元素。

3 分层 LBP 和 HOG 特征融合

3.1 HOG 特征

Dalal 等^[23]提出了 HOG 的基本概念,HOG 算子计算局部图像梯度方向信息的统计值,可以描述局部目标的形状和边缘,对光照变化和梯度特征的少量偏置具有稳健性,可以很好地表示人脸轮廓特征信息。HOG 特征提取可描述为:确定一个 $n \times n$ 滑动窗口,将滑动窗口均匀分成若干个小块,让滑动窗口在整个图像上滑动,分别在每一个分块上运算每一个像素点的方向和幅度,计算公式为

$$\theta(x, y) = \arctan \left[\frac{I(x, y+1) - I(x, y-1)}{I(x+1, y) - I(x-1, y)} \right], \quad (5)$$

$$m(x, y) = \sqrt{[I(x, y+1) - I(x, y-1)]^2 + [I(x+1, y) - I(x-1, y)]^2}. \quad (6)$$

假设像素点为 x_0 ,周围有 8 个邻域 $x_i (i=1, 2, \dots, 8)$, $m_x(x_0)$, $m_y(x_0)$, $I(x_i)$ 分别表示该点的水平方向梯度、垂直方向梯度和像素大小,则该点的梯度为

$$m_x(x_0) = I(x_5) - I(x_4), \quad (7)$$

$$m_y(x_0) = I(x_7) - I(x_2). \quad (8)$$

将(5)式和(6)式代入(7)式和(8)式得到该像素点

x_0 的幅度和方向值。

对每一个分块进行 HOG 特征的统计,顺序串联每一个小块的直方图特征,从而得到完整图像的 HOG 特征。

3.2 LBP 特征

LBP 算子的作用在于提取人脸图像的特征值,

因本身对灰度和旋转具有很好的稳健性,所以 LBP 算子在提取特征方面有很大优势。

原始 LBP 算子定义在 3×3 邻域内,如图 3 所示,将中心像素点作为基础,分别与周围 8 个像素值

进行比较,若其像素值大于阈值,则该点记为 1,反之为 0。 3×3 邻域内的 8 个像素值可生成 8 位二进制数,转换为十进制后得到的值即为该窗口中心像素点的 LBP 值。

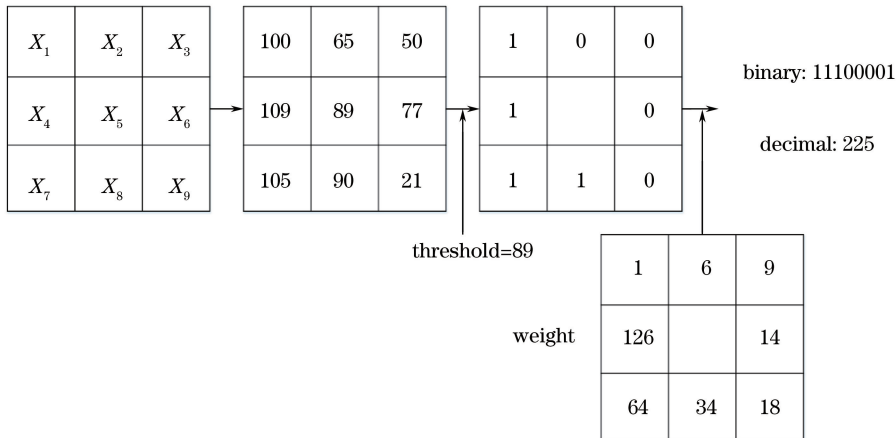


图 3 LBP 特征提取流程图

Fig. 3 Flow chart of LBP feature extraction

LBP 值计算公式为

$$f_{\text{LBP}}(x_c, y_c) = \sum_{i=0}^7 s(g_i - g_c) \cdot 2^i, \quad (9)$$

$$s(x) = \begin{cases} 1, & x > 0, \\ 0, & x < 0, \end{cases} \quad (10)$$

式中: $f_{\text{LBP}}(x_c, y_c)$ 为中心像素点 (x_c, y_c) 的 LBP 值; $g_i (i=1, 2, \dots, 8)$ 表示以 g_c 为中心的圆形区域的 8 个像素点。

3.3 特征融合

卷积神经网络本身具有特征提取能力,但在实验过程中,表情变化、遮挡情况、复杂光照环境等一系列因素对卷积神经网络训练有一定的影响,会降低识别精度,算法稳健性差。万源等^[24]提出了基于 LBP 和 HOG 的分层特征的人脸识别方法,能有效提取图像特征,说明了多特征提取能有效提高算法的识别精度。本文在此基础上,通过实验验证了融合特征和改进 CNN 结合对 CAS-PEAL 的人脸识别精度比基于 CNN 的识别算法提高了 4.7%。原因在于 HOG 特征侧重于提取形状信息, LBP 特征侧重于提取纹理信息,将融合特征进行 CNN 训练,不但能有效提高训练效率,还能降低误识别率,提升算法稳健性。因此,本文将 LBP 和 HOG 融合与改进 CNN 结合后识别精度得到了提高,识别性能得以改善。

所谓特征匹配融合是在特征融合的过程中,将三层 LBP 特征和三层 HOG 特征分别按如下步骤

进行匹配。

1) 对灰度人脸图像进行 n 次 HOG 特征的提取,获得人脸图像的 HOG 特征,即 $\text{HOG}(n), n=1, 2, 3$,如图 4 所示。

2) 对 $\text{LBP}(n), n=1, 2, 3$ 特征提取图像进行相同方法的分块形式,将图像分成大小相等、互不重叠的均匀子块图像。

3) 将所有子块的 LBP 直方图特征加以统计,并顺序串联为每层 $\text{LBP}(n), n=1, 2, 3$ 特征的特征向量。

4) 对灰度图像进行 n 次 LBP 特征的提取,即 $\text{LBP}(n), n=1, 2, 3$ 。

5) 第一层 HOG 特征提取图像能够清晰表达人脸的外形,遮挡边缘以及鼻子边缘信息。第二层和第三层虽清晰表达信息,但是边缘信息弱化。因此优先选择第一层 HOG 特征进行后续特征融合。

6) 将分层 $\text{LBP}(n), n=1, 2, 3$ 特征直方图与第一层 HOG 特征直方图进行顺序级联,得到 $\text{LBPHOG}(n, 0), n=1, 2, 3$ 。将三层 LBP 特征直方图和 HOG 级联得到直方图,此时表达信息最完善,特征完整,故优先选择最后一种特征融合方法,得到 $\text{LBPHOG}(3, 0)$ 。

通过上述步骤优选出特征完整、轮廓清晰的特征图作为 CNN 输入。图 4 中米黄色直线作为示例得出一组融合图,同理可得其他两组融合图,获得细节更加丰富的融合效果^[25-26]。

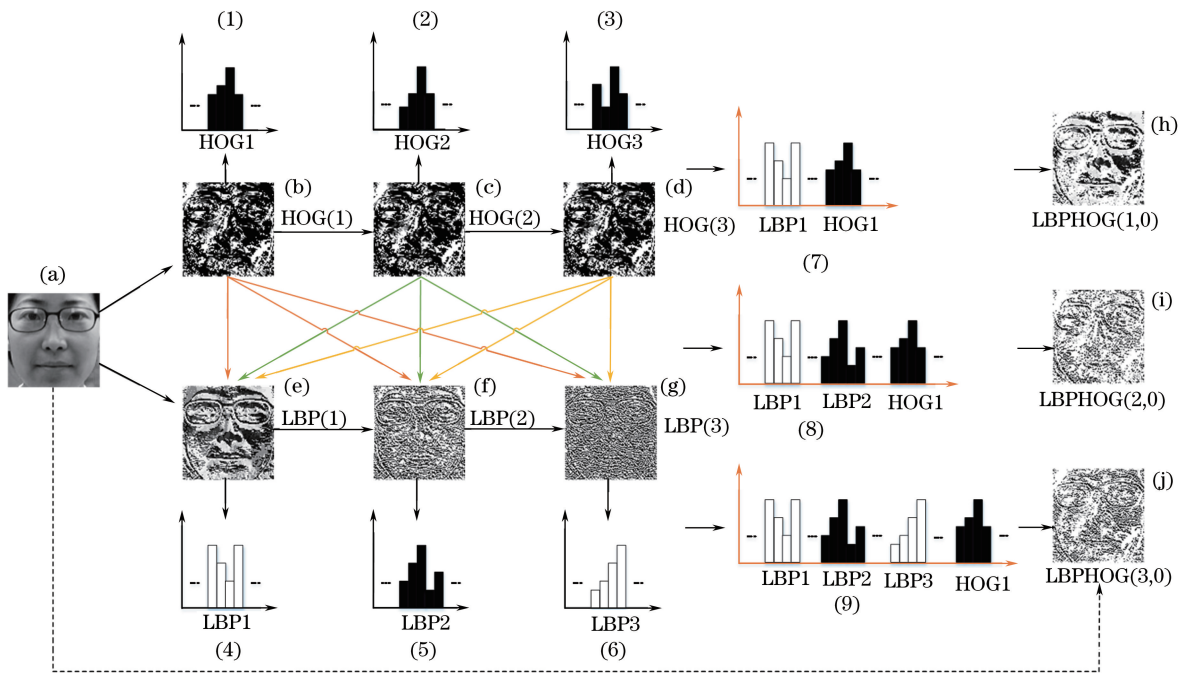


图 4 特征融合流程图

Fig. 4 Flow chart of feature fusion

4 本文算法

图 5 所示为本文算法结构图, 优选特征融合结合扩展卷积神经网络的图像识别主要由三部分组成: 图像预处理、优选特征融合以及 CNN 训练分类。图像样本首先通过直方图均衡化、尺度归一化和灰度归一化等预处理操作, 目的是缩短整体网络的计算时间, 同时可以删减原有图像信息中的无用

信息, 节省训练时间。经过预处理变换得到 64×64 、灰度值范围为 $[0, 1]$ 的图像集。优化特征融合完成对样本的提取, 用来训练卷积神经网络, 通过 Softmax 分类器对测试集的人脸进行判别输出。

具体步骤如下:

- 1) 样本通过预处理操作获得大小为 64×64 、灰度值范围为 $[0, 1]$ 的人脸图像集;
- 2) 提取图像的分块 LBP 特征, 得到各分块特

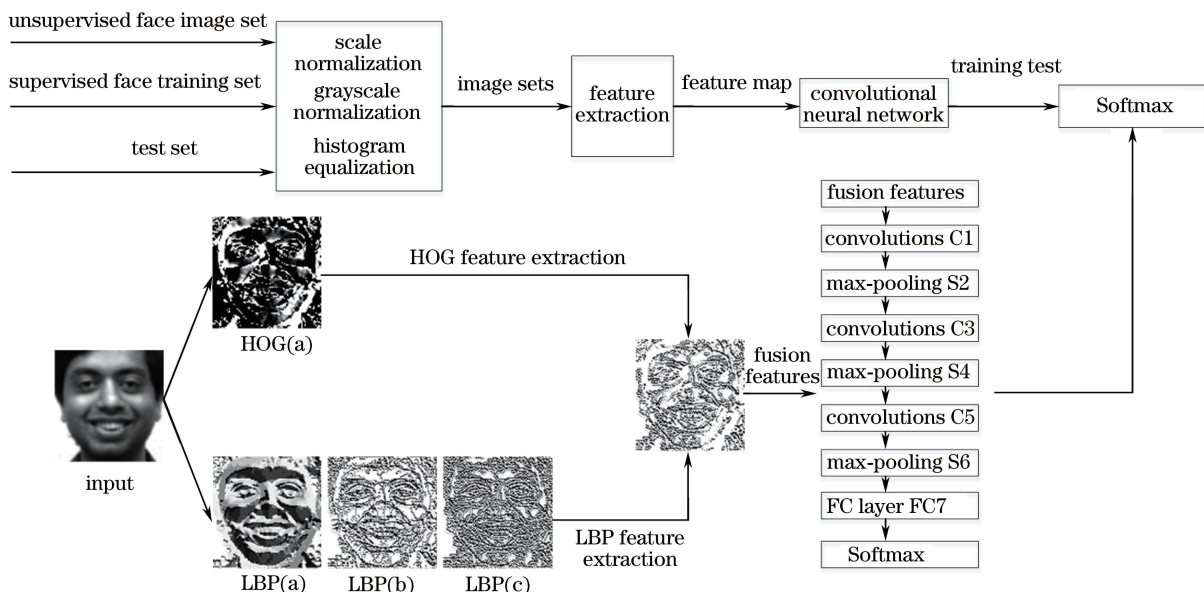


图 5 算法结构图

Fig. 5 Algorithm structure

征,级联形成整体 LBP 纹理特征,本文以 2×2 分块、半径 1、像素值 8 进行特征提取,共提取三次,提取得到的维度为 1024;

3) 提取原图像的 HOG 特征,同样进行三层特征提取,由于第一层提取的特征最为丰富,且第二层和第三层提取的 HOG 特征会弱化,表达信息不完整,故优先选择第一层特征;

4) 将三层 LBP 特征和 HOG 特征进行融合构成 32×32 维矩阵;

5) 提取的融合特征样本作为 CNN 的输入,进行卷积核池化,得到最终人脸特征 FC7,之后输出到 Softmax 分类器中,训练整个网络;

6) 将测试样本的融合特征输入训练好的网络

中,通过 Softmax 分类器进行分类识别。

5 仿真结果与分析

仿真模型基于 TensorFlow 深度学习框架搭建而成,仿真服务器硬件配置: Intel (TM) i7-3770 CPU, 8 G RAM, NVIDIA GeForce GTX 1060 GPU; 软件环境: Ubuntu 16.04 系统, MATLAB R2014a, TensorFlow 深度学习框架。

仿真使用的卷积神经网络共有 7 层,详细参数信息见表 1。为了验证提出的特征匹配融合方法在卷积神经网络训练中的有效性,分别在 ORL 人脸库、YALE 人脸库和 CAS-PEAL 人脸库上进行仿真。

表 1 卷积神经网络结构

Table 1 Architecture of CNN

Layer	Input layer	Convolution kernel	Pool	Number of convolution kernels	Number of input features	Output layer
Input						32×32
Convolution C1	32×32	3×3		20	20	30×30
Subsampling S2	30×30		2×2		20	28×28
Convolution C3	28×28	2×2		40	40	14×14
Subsampling S4	14×14		2×2		40	7×7
Convolution C5	7×7	4×4		60	60	4×4
Subsampling S6	4×4		2×2		60	2×2
FC layer FC7					1	240×1
Output	240×1				1	10×1

为了验证本文算法的有效性,选择 ORL、YALE 和 CAS-PEAL 人脸库对本文算法进行测试和比较。ORL 人脸库中的人脸图像共包含 400 张人脸正面图像,40 个人,每个人 10 幅图像,含有不同表情、不同拍摄角度,姿态变化较小,分辨率为 112×92 。分别取每个人的第 5~8 张图像进行训练,其他图像用于测试。YALE 人脸库中共包含 165 张人脸图像,15 个人,每个人 11 幅图像,主要区别在于光照、表情变化和饰品的有无,分辨率为

100×100 。分别取每个人的第 6~9 张图像进行训练,其他图像用于测试。CAS-PEAL 人脸库中共包含 99450 张人脸图像,1040 个人,仿真抽取 400 人,分辨率为 100×100 。每个人选 15 幅图像,共 6000 幅人脸图像作为实验对象,主要包括表情、遮挡物和背景的变化,分别取每个人的第 9~12 张图像进行训练,其余用于测试。可以看出,本文算法达到了较高的识别率。实验结果如图 6 所示。

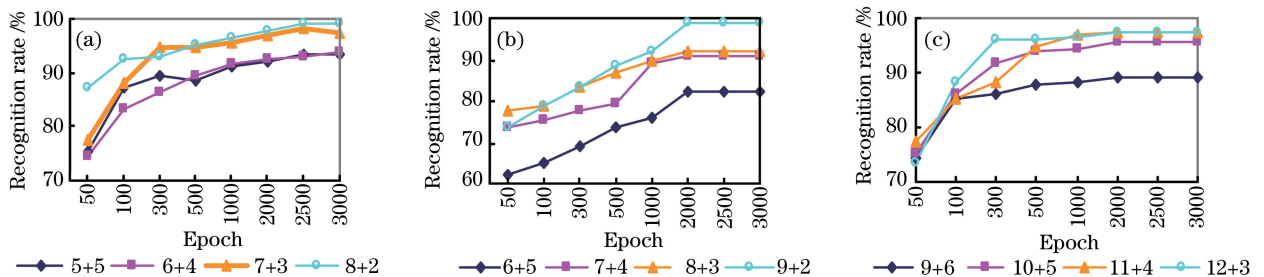


图 6 人脸库不同迭代次数对应的识别率。(a) ORL;(b) YALE;(c) CAS-PEAL

Fig. 6 Recognition rate corresponding to different iteration number in the database. (a) ORL; (b) YALE; (c) CAS-PEAL

分析结果可以得出,选择较多图像作为训练集且迭代次数达到 2000 次以上时,人脸识别率有较大的提高。这是由于训练次数增加,使网络更好地学习特征并降低了特征维度。

与其他算法相比,本文算法的确达到了较高的识别率,如表 2 所示。以基于 LBP 算子或者 HOG 算子和支持向量机(SVM)的方法作为对比。基本思路为采用 LBP 算子或 HOG 算子提取图像特征,并且利用 SVM 算法训练分类器实现分类识别。使用的 SVM 是台湾林智仁教授开发设计的 MATALAB 版 LIBSVM3.22。LibSVM^[27]大规模线性分类器包含了单类和多类支持向量机分类、利用网络实验计算得出最优的核函数参数、交叉验证等功能。

深度信念网络(DBN)与其他深度学习模型类似,通过输入低层特征描述进行学习,得到更抽象、有效的特征描述。DBN 的基本模型由若干个受限玻尔兹曼机(RBM)组成,在 DBN 算法中,首先利用对比度散度算法^[28]训练网络,为每个 RBM 优化参数,从而获得整个网络节点的权重参数;之后从标记数据的顶部调整网络;网络训练完成后,使用监督中的反向传播(BP)算法,可以获得训练有素的 DBN 网络。对比实验中,将融合测试样本的梯度特征输入到训练好的 DBN 网络中,学习更高层、完整的特征,并在网络顶层使用 Softmax 分类器进行测试样本的分类和识别。

结果表明,LBP 算子与 HOG 算子融合后提取的特征的鉴别能力比单独算子提取的特征更高,识别率高于单独使用这两个算子的情况。大样本人脸图像进行训练分类也保持较高的识别率,验证了本文算法的有效性和可行性。

表 2 六种算法的识别率对比(单位:%)

Table 2 Recognition rate comparison of six algorithms (unit: %)

Algorithm	ORL	YALE	CAS-PEAL
HOG+SVM	95.32	87.78	76.73
LBP+SVM	96.27	83.12	63.00
DBN+HOG	93.38	89.92	83.00
LBP+CNN	98.52	95.67	93.27
HOG+CNN	98.60	96.25	96.00
LBP+HOG+CNN	99.21	98.78	97.51

6 结 论

提出一种 LBP 和 HOG 分层特征融合的方法,输入到卷积神经网络中进行训练,并将其应用于人

脸识别。LBP 算子主要提取图像的纹理特征信息,HOG 算子主要提取图像边缘信息,两者特征匹配融合可提高识别率。进一步将融合信息输入到卷积神经网络中进行训练,实现人脸识别。在标准人脸库上比较分析不同数据库中识别率的变化,以及六种不同算法的识别率。结果表明,本文算法更为有效,分层融合方法可增强图像特征,显著提高了人脸识别率。提出的特征提取方案在神经网络中 SVM、DBN、CNN 等模型下有较好的效果。其他类型,如 K 最邻近分类算法(KNN)、深层神经网络(DNN)、BP、径向网络(RBF)、循环神经网络(RNN)等模型的结合,还没有通过实验作进一步分析。后续研究将考虑采用卷积稀疏字典等其他算法提取特征,或者采用显著性分析强化所需要的特征,利用更加简单、高效的多维图像进行分类识别,以期获得更好的效果。

参 考 文 献

- [1] Zuo D Y. A study of face detection based on Adaboost algorithm[D]. Hefei: University of Science and Technology of China, 2009: 1-2.
左登宇. 基于 Adaboost 算法的人脸检测研究[D]. 合肥: 中国科技大学, 2009: 1-2.
- [2] Chen H J. The comparison research of some algorithms of biological recognition[J]. Journal of the Hebei Academy of Sciences, 2007, 24(4): 33-37.
陈洪京. 几种生物识别方法的比较研究[J]. 河北省科学院学报, 2007, 24(4): 33-37.
- [3] Wold S, Esbensen K, Geladi P. Principal component analysis[J]. Chemometrics and Intelligent Laboratory Systems, 1987, 2: 37-52.
- [4] Lu J W, Plataniotis K N, Venetsanopoulos A N. Regularization studies of linear discriminant analysis in small sample size scenarios with application to face recognition[J]. Pattern Recognition Letters, 2005, 26(2): 181-191.
- [5] Bartlett M S, Movellan J R, Sejnowski T J. Face recognition by independent component analysis [J]. IEEE Transactions on Neural Networks, 2002, 13(6): 1450-1464.
- [6] Zhang Y, Hua C J. Driver fatigue recognition based on facial expression analysis using local binary patterns[J]. Optik, 2015, 126(23): 4501-4505.
- [7] Lowe D G. Distinctive image features from scale-invariant key points [J]. International Journal of Computer Vision, 2004, 60(2): 91-110.
- [8] Xu X M, Quan C Q, Ren F J. Facial expression

- recognition based on Gabor wavelet transform and histogram of oriented gradients [C] // IEEE International Conference on Mechatronics and Automation (ICMA), 2015: 2117-2122.
- [9] Yoo C H, Kim S W, Jung J Y, *et al.* High-dimensional feature extraction using bit-plane decomposition of local binary patterns for robust face recognition[J]. *Journal of Visual Communication and Image Representation*, 2017, 45: 11-19.
- [10] Albiol A, Monzo D, Martin A, *et al.* Face recognition using HOG-EBGM [J]. *Pattern Recognition Letters*, 2008, 29(10): 1537-1543.
- [11] Shan C F, Gong S G, McOwan P W. Facial expression recognition based on local binary patterns: a comprehensive study [J]. *Image and Vision Computing*, 2009, 27(6): 803-816.
- [12] Ojala T, Pietikäinen M, Harwood D. A comparative study of texture measures with classification based on featured distributions[J]. *Pattern Recognition*, 1996, 29(1): 51-59.
- [13] Chen D, Cao X D, Wen F, *et al.* Higher is better: high-dimensional feature and its efficient compression for face verification [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2013: 1-8.
- [14] Pang Y W, Yuan Y, Li X L, *et al.* Efficient HOG human detection [J]. *Signal Processing*, 2011, 91(4): 773-781.
- [15] Yu J, Qin Z C, Wan T, *et al.* Feature integration analysis of bag-of-features model for image retrieval [J]. *Neurocomputing*, 2013, 120: 355-364.
- [16] Girshick R, Donahue J, Darrell T, *et al.* Rich feature hierarchies for accurate object detection and semantic segmentation[C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2014: 580-587.
- [17] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolution neural networks [C] // Neural Information Processing Systems Conference, 2012: 1097-1105.
- [18] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition [C] // International Conference on Learning Representations (ICLR), 2014: 1-2.
- [19] Xu L, Zhao H T, Sun S Y. Monocular infrared image depth estimation based on deep convolutional neural networks[J]. *Acta Optica Sinica*, 2016, 36(7): 0715002.
许路, 赵海涛, 孙韶媛. 基于深层卷积神经网络的单目红外图像深度估计[J]. *光学学报*, 2016, 36(7): 0715002.
- [20] Li S M, Lei G Q, Fan R. Depth maps super-resolution reconstruction based on convolutional neural networks [J]. *Acta Optica Sinica*, 2017, 37(12): 1210002.
李素梅, 雷国庆, 范如. 基于卷积神经网络的深度图超分辨率重建 [J]. *光学学报*, 2017, 37(12): 1210002.
- [21] Long H Q, Tan T Z. Based on the depth of the convolution network and local binary pattern of face recognition[J]. *Computer Simulation*, 2017, 34(1): 322-325.
龙海强, 谭台哲. 基于深度卷积神经网络算法的人脸识别方法研究[J]. *计算机仿真*, 2017, 34(1): 322-325.
- [22] He K M, Sun J. Convolution neural networks at constrained time cost [C] // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2015: 5353-5360.
- [23] Dalal N, Triggs B. Histograms of oriented gradients for human detection [C] // Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005: 1886-1893.
- [24] Wan Y, Li H H, Wu K F, *et al.* Fusion with layered features of LBP and HOG for face recognition [J]. *Journal of Computer-Aided Design and Computer Graphics*, 2015, 27(4): 640-650.
万源, 李欢欢, 吴克凤, 等. LBP 和 HOG 的分层特征融合的人脸识别[J]. *计算机辅助设计与图形学学报*, 2015, 27(4): 640-650.
- [25] Chen M S, Cai Z S. Study on fusion of visual and infrared images based on NSCT [J]. *Laser & Optoelectronics Progress*, 2015, 52(6): 061002.
陈木生, 蔡植善. 基于 NSCT 的红外与可见光图像融合方法研究[J]. *激光与光电子学进展*, 2015, 52(6): 061002.
- [26] Zhang L B, Zhang J. A new adaptive fusion method based on saliency analysis for remote sensing images [J]. *Chinese Journal of Lasers*, 2015, 42(1): 0114001.
张立保, 章珏. 基于显著性分析的自适应遥感图像融合[J]. *中国激光*, 2015, 42(1): 0114001.
- [27] Fan R E, Chen P H, Lin C J. Working set selection using second order information for training SVM[J]. *Journal of Machine Learning Research*, 2005, 6(4): 1889-1918.
- [28] Hinton G E, Osindero S, Teh Y W. A fast learning algorithm for deep belief nets [J]. *Neural Computation*, 2006, 18(7): 1527-1554.