

基于 LBP 和数据扩充的 CNN 人脸识别研究

任飞凯, 邱晓晖

(南京邮电大学 通信与信息工程学院, 江苏 南京 210003)

摘要:针对卷积神经网络在人脸识别存在的数据集比较少,容易发生拟合的问题,提出对人脸进行局部二值模式处理,提升图像特征,再引入深度卷积生成对抗网络对局部二值化的人脸进行生成,有效扩充数据集,提升卷积神经网络的泛化能力。该人脸识别卷积神经网络模型包括3层卷积层,3层池化层,1个全连接层,1个 Softmax 分类回归层。仿真实验中,选取 ORL 人脸数据库中40人每人10张的人脸图像按8:1:1比例设置为训练集、验证集和测试集,并选取 Yale 人脸数据库中15人每人11张的人脸图像按9:1:1的比例设置训练集、验证集和测试集,通过 LBP 算法提取人脸纹理特征对其进行生成,分别扩充数据集至990张和2 200张。结果表明,该算法的人脸识别率不仅高于未扩充数据 PCA 和 LBP 等传统人脸识别方法的识别率,而且也将卷积神经网络的识别率提升了约2%,有效提高了泛化能力。

关键词:CNN; LBP; 数据集扩充; 人脸识别

中图分类号:TP391

文献标识码:A

文章编号:1673-629X(2020)03-0062-05

doi:10.3969/j.issn.1673-629X.2020.03.012

Research on Face Recognition of CNN Based on LBP and Data Expansion

REN Fei-kai, QIU Xiao-hui

(School of Telecommunications & Information Engineering, Nanjing University of Posts and Telecommunications, Nanjing 210003, China)

Abstract:In view of the problem that convolutional neural network has few data sets in face recognition and is prone to over-fitting, the local binary mode processing is carried out to enhance the image features, and the deep convolution generation is introduced to generate the anti-network to generate the local binarized face, which effectively expands the data set and improves the generalization of the convolutional neural network. This convolutional neural network model of face recognition consists of 3-layer convolutional layer, 3-layer pooling layer, a fully connected layer and a Softmax classification regression layer. In the simulation experiment, the 10 face images of each of 40 people in the ORL face database are selected as the training set, verification set and test set according to the ratio of 8:1:1, and 11 face images of each of 15 people in the Yale face database are selected as the training set, verification set and test set according to the ratio of 9:1:1. The face texture features are extracted by LBP algorithm to generate them, and the data set is expanded to 990 sheets and 2 200 sheets respectively. The results show that the face recognition rate of the proposed algorithm is not only higher than that of the traditional face recognition methods such as unexpanded data PCA and LBP, but also the recognition rate of the convolutional neural network is increased by about 2%, which effectively improves its generalization.

Key words:CNN; LBP; dataset expansion; face recognition

0 引言

近年来,人脸识别技术是人工智能领域研究的热点之一。人脸识别因其非接触性、非强制性、成本低等特点,使得其在身份认证领域有着很高的应用价值。经典的人脸识别算法有主成分分析法(PCA)^[1]、线性判别分析(LDA)^[2]、局部保真映射(LPP)^[3]、稀疏表示^[4]、小波分析^[5]等。

卷积神经网络是神经科学原理影响深度学习的典

型代表,其最早可追溯到八十年代^[6],之后发展一度停滞,直到反向传播算法的提出^[7],才有了突破性的进展。现在卷积神经网络的主要结构有 LeNet-5^[8], AlexNet^[9], GoogleNet, VGGNet, ResNet^[10]等。这些网络结构都比较深,网络参数比较多,训练不充分很容易造成网络的过拟合,而解决过拟合的方法一般是扩充数据集或者增强数据集。由于网络结构的原因,训练样本越多,网络的泛化能力也会越好。因此,文中提出

收稿日期:2019-03-27

修回日期:2019-07-30

网络出版时间:2019-12-05

基金项目:江苏省自然科学基金(BK2011789)

作者简介:任飞凯(1991-),男,硕士研究生,研究方向为信号处理;邱晓晖,博士,教授,通信作者,研究方向为信号与信息处理。

网络出版地址: <http://kns.cnki.net/kcms/detail/61.1450.TP.20191205.1113.028.html>

了将人脸进行局部二值模式处理,在进行训练之前提取人脸面部纹理特征,然后将通过 LBP 算法提取出的人脸特征纹理放入深度卷积生成对抗网络中进行训练生成新的人脸图像来扩充数据集。

1 局部二值模式和数据集扩充

1.1 局部二值模式

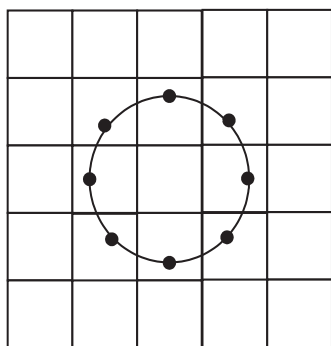
局部二值模式(local binary patterns, LBP)是一种用于描述局部纹理特征的算子,具对光照不变性和灰度不变性等显著优点^[11],因此广泛应用于人脸识别、表情识别等图像处理领域。

基本 LBP 算子可以理解为一个 3×3 的矩阵,其中心像素的灰度值为一个阈值,周围相邻的八个像素的灰度值分别与中心阈值比较,如果像素灰度值高于中心阈值则标为 1,低于则标为 0,如图 1 所示。最终得到一组包含 0 和 1 的序列,转化成十进制后便是中心像素的 LBP 值, LBP 值可以用于获得面部特征。



图 1 LBP 算子

但是这种基本 LBP 有一定的局限性,如果图像旋转, LBP 值也会随着改变^[12],所以文中采用圆形 LBP 算子,如图 2 所示。圆形算子的邻域不限于 LBP 算子的 8 个,且可以自由调节采样半径。圆形算子符号表示为 $LBP_{P,R}$ 其中 P 为邻域采样点数, R 为采样半径。图 2 中采样点数 P 为 8, 采样半径 R 为 1。



$LBP_{8,1}$

图 2 圆形 LBP 算子

圆形 LBP 算子计算公式如下:

$$LBP_{P,R} = \sum_{p=0}^{P-1} S(g_p - g_c) * 2^p \quad (1)$$

其中, g_p 表示邻域像素灰度值, g_c 表示中心像素灰度值, $S(x)$ 表示中心像素与邻域像素比较之后的布尔值,公式如下:

$$S(x) = \begin{cases} 1, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (2)$$

1.2 深度卷积生成对抗网络

深度卷积生成对抗网络(deep convolutional generative adversarial networks, DCGAN)^[13]是一种改进的生成对抗网络(generative adversarial networks, GAN)^[14]。GAN 中包含两个模型,一个生成模型,一个判别模型。可以用微分方程 D 和 G 来表示判别模型和生成模型。生成器 G 接受随机噪声 z , 通过这个噪声生成图片,记为 $G(z)$, 判别模型 D 输入参数为 x , $D(x)$ 代表 x 为真实图片的概率,如果为 1,则表示 x 为真实图片,为 0 则表示为假图片。在训练整个模型的过程中,生成模型 G 的目的就是为了生成尽量真实的图片去欺骗判别模型 D ,而 D 的目的是为了将真实图片和 G 生成的图片区分开来,这样便形成了一个动态博弈。最终的博弈结果是 $D(G(z)) = 0.5$,这意味着判别模型 D 无法区分生成模型 G 生成的图片是否为真实图片,这样就可以用训练好的生成模型 G 来生成图片。

DCGAN 是在 GAN 的基础上进行了改进。DCGAN 的判别模型 D 是一个卷积神经网络,生成模型 G 的输入是一个 100 维的噪声向量 z , G 的第一层就是一个全连接层,将 100 维向量变成一个 $4 \times 4 \times 1024$ 的向量,然后使用转置卷积做上采样,逐层减少通道,最后得到的输入为 $64 \times 64 \times 3$,即一个三通道宽高都是 64 的图像。

文中将通过 LBP 算法提取出的人脸特征纹理当作判别模型 D 的输入,训练判别模型的判别能力,然后通过生成模型 G 来生成图片,放入判别模型 D 中进行判别,当 $D(G(z)) = 0.5$ 时,就可以用 DCGAN 中的生成模型 G 来生成图像,达到扩充数据集的目的,提高卷积神经网络的泛化性。

2 局部二值模式扩充数据集的 CNN 人脸识别

卷积神经网络(convolutional neural networks, CNN)用于人脸识别是通过卷积操作来提取人脸面部特征,通过不同的卷积核提取到特征所具有的性质和表现力也不一样。其结构一般分为卷积层、池化层、全连接层和 Softmax 分类。文中设计的网络结构如图 3 所示。首先将人脸图进行局部二值化处理,提取面部纹理特征,然后将提取出的面部纹理特征当作判别模式 D 的输入,当 $D(G(z)) = 0.5$ 时,用生成模式 G 的生成图像来扩充数据集,提高卷积神经网络的泛化能力,然后将真实的局部二值化人脸和生成的局部二值化人脸输入至卷积神经网络中进行训练学习参数。卷

积神经网络由3个卷积层 C_1 、 C_2 、 C_3 , 3个池化层 S_1 、 S_2 、 S_3 , 一个全连接层, 一个 Softmax 分类器组成。输入人脸纹理特征图像经处理统一为 $64 * 64$, C_1 有 5 个

$3 * 3$ 卷积核, C_2 有 10 个 $3 * 3$ 卷积核, C_3 有 15 个 $3 * 3$ 卷积核, 池化层选择 $2 * 2$ 的最大池化, 全连接层为 $1 * 100$, Softmax 分类器根据数据库人数另外设置。

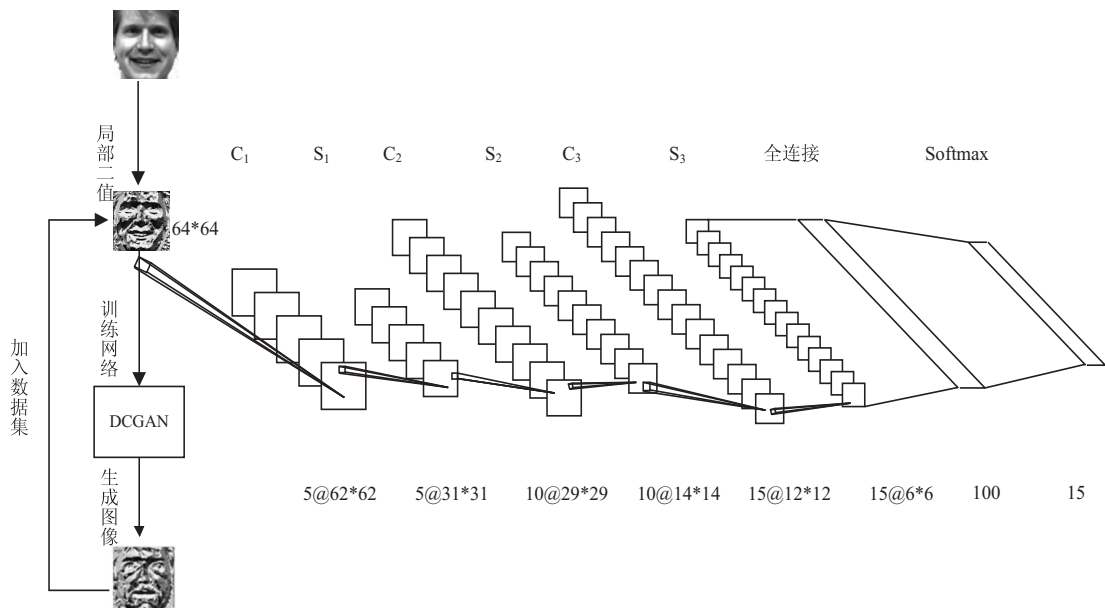


图3 人脸识别的结构设计

2.1 网络结构

(1) 卷积层。

卷积神经网络核心便是卷积运算。卷积运算具有稀疏交互、参数共享等特性,通过卷积可以有效地提取出图像中的特征数据,并且减少训练参数^[15]。卷积神经网络中的卷积可以理解为用一个滤波器扫描图像像素,而这个滤波器就是卷积核,通过卷积核与图像的卷积,可以得到包含图像特征信息的特征图,具体计算公式如下:

$$y_i^j = f(\sum_{i=1}^{N_j^{l-1}} \omega_j \otimes X_j^{l-1} + b_j^l), j = 1, 2, \dots, M \quad (3)$$

其中, y_i^j 表示第 j 个特征图, ω_j 表示卷积核, X_j^{l-1} 表示上一层第 i 个特征图作为当前输入, b_j^l 表示偏置, N_j^{l-1} 表示每个特征图的特征数量, M 表示当前卷积层的特征图数量, $f(\cdot)$ 为激活函数,常见的激活函数有 Sigmoid, Tanh 和 ReLU,文中选用 ReLU,其公式如下:

$$\text{ReLU}(x) = \begin{cases} x, & x > 0 \\ 0, & x \leq 0 \end{cases} \quad (4)$$

通过卷积核与图像的卷积,带有原始图像特征的特征图就会被提取出来,经过多层卷积之后,低层特征图会被提取出更具表征能力的特征图。因此为了能够得到更多的特征,更深层次的特征图也会更多。

卷积层 C_1 采用 5 个 $3 * 3$ 的卷积核,对 $64 * 64$ 的图像进行卷积,得到 5 个 $(64 - 3 + 1) * (64 - 3 + 1)$ 即 $62 * 62$ 的特征图,即提取了 5 个人脸纹理的特征图; C_2 采用 10 个 $3 * 3$ 的卷积核对 $31 * 31$ 的 S_1 进行卷积,得到 10 个 $(31 - 3 + 1) * (31 - 3 + 1)$ 即 $29 * 29$ 的特征

图; C_3 采用 15 个 $3 * 3$ 的卷积核对 $14 * 14$ 的 S_2 进行卷积,得到 15 个 $(14 - 3 + 1) * (14 - 3 + 1)$ 即 $12 * 12$ 的特征图。

(2) 池化层。

池化层也称作下采样层,图像经过多次卷积之后,特征图的数量会越来越多,网络结构也越来越复杂,同时网络的参数也越多越容易引发过拟合。所以在卷积层后会加入一个池化层,来对特征图进行特征聚合,降低维度。常见的池化方法有最大池化、平均池化和随机池化。最大池化对于纹理特征提取明显,所以文中选用最大池化。

池化层 S_1 对卷积层 C_1 进行 $2 * 2$ 的最大池化,得到了 5 个大小为 $31 * 31$ 的特征图。同理 S_2 对 C_2 进行下采样,得到 10 个 $14 * 14$ 的特征图, S_3 对 C_3 进行下采样,得到 15 个 $6 * 6$ 的特征图。

(3) 全连接层。

为了增强网络的非线性能力,同时限制网络规模,一般会在卷积池化层之后加入一个全连接层。全连接层相当于传统神经网络中的多层感知机的隐藏层,即全连接层的每一个神经元与前一层的所有神经元互相连接,同层神经元之间不连接。其公式如下。

$$y_{w,b}(x) = f(\mathbf{W}^T \mathbf{x} + b) \quad (5)$$

其中, $y_{w,b}(x)$ 表示神经元的输出, \mathbf{W} 表示权值向量, \mathbf{x} 表示神经元输入向量, b 表示偏置, $f(\cdot)$ 表示激活函数。实验中为了降低网络复杂度,避免过拟合, b 设置为 0。

(4) Softmax 分类器。

卷积神经网络输出层选择 Softmax 分类器,其公式如下:

$$S_i = \frac{e^{V_i}}{\sum_i e^{V_i}} \quad (6)$$

其中, V_i 表示分类器上一级输入单元的输出, i 表示类别索引, C 表示总类别个数, S_i 表示当前元素的指数与所有元素指数和的比值。

当输入了给定样本时,每一个神经元会输出一个 0 到 1 的值,这个值便是给定样本属于每个分类的概率,从而可以判定给定样本是哪一分类,达到识别人脸的效果。

2.2 网络训练

卷积神经网络主要分为两个部分^[16],第一部分为前向传播,第二部分为反向传播。前向传播时,选取一个样本 x 和它的标签 y ,将样本 x 输入至网络中,通过整个卷积池化全连接之后,Softmax 输出一个 \tilde{y} , \tilde{y} 是一个 n 维向量, n 取决于 Softmax 选定了多少分类,也就是系统可以识别的人数,如 Yale 数据库有 15 人, n 则为 15。 \tilde{y} 中的每个元素代表了输入样本 x 是其中每个人的概率。反向传播时,计算 \tilde{y} 与 y 的误差,按照极小化误差的方法调整权重矩阵。

2.3 算法步骤

文中算法的步骤如下:

(1)将数据库人脸进行局部二值化处理,提取人脸特征纹理。

(2)将局部二值化的人脸放入 DCGAN 中训练,训练完成后用生成模式 D 来生成局部二值化的人脸,以扩充用于下一步中的 CNN 人脸识别的数据集,提高网络的泛化性。

(3)将真实局部二值化的人脸和生成的局部二值化的人脸共同训练 CNN 人脸识别模型。

3 仿真实验结果与分析

3.1 实验选用的人脸数据集

(1)ORL 人脸数据库。

由英国剑桥大学 AT&T 实验室创建,包含了 40 个人,每人 10 张共 400 张面部图像,其中部分图像包含了姿态、表情和面部饰物的变化。采集对象的面部表情和细节均有变化,例如笑与不笑,眼睛睁着或闭着,戴或者不戴眼镜。

(2)Yale 人脸数据库。

由耶鲁大学计算视觉与控制中心创建,包含了 15 个人,每人 11 张共 165 张图片,数据库里的人脸图像包含了光照、表情和姿态的变化。相较于 ORL 人脸数据库,Yale 人脸数据库中每个对象采集的样本包含更

加明显的光照、表情和姿态以及遮挡的变化。

3.2 实验结果及分析

(1)人脸图像局部二值模型处理。

Yale 人脸数据库中选取了一张,如图 4(a)所示,经过局部二值模型处理后的结果如图 4(b)所示,面部纹理特征得到增强,轮廓和边缘特性更为明显。

(2)生成的局部二值图像。

将局部二值模式处理过的图像放入 DCGAN 训练,训练完成后,通过 DCGAN 生成的局部二值模式处理过的人脸如图 4(c)所示,人脸轮廓和纹理都十分清晰。根据 DCGAN 生成对抗的原理,DCGAN 中的判别模式无法区分生成图是真是伪。



(a) 原图 (b) LBP处理图 (c) 生成图

图4 LBP处理和生成的图像

通过 DCGAN,Yale 人脸数据库每个人生成了 55 张,整个数据集扩充为 990 张。ORL 人脸数据库每个人生成了 45 张,数据库扩充为 2 200 张。数据集得到显著扩充,为比较原始数据集和文中方法扩充数据的识别率,保持训练集、验证集和测试集按照整个数据集划分为 8 : 1 : 1 的比例不变。

(3)与其他方法比较。

图 5 给出了在 Yale 人脸数据库中使用原始数据集和文中方法数据集中验证集和测试集的识别率对比情况,其中 Epochs 表示迭代次数,图例 CNN Val Acc 为原始数据集的验证集识别率,CNN Test Acc 是原始数据集测试集识别率,This article Val Acc 和 This article Test Acc 为文中方法扩充数据集的验证集和测试集识别率。表 1 前四个给出了在 ORL 人脸数据库中各个算法中的识别率,后四个给出了在 Yale 人脸数据库中各个算法的识别率,其中 CNN 是使用原始数据集进行训练的卷积神经网络。

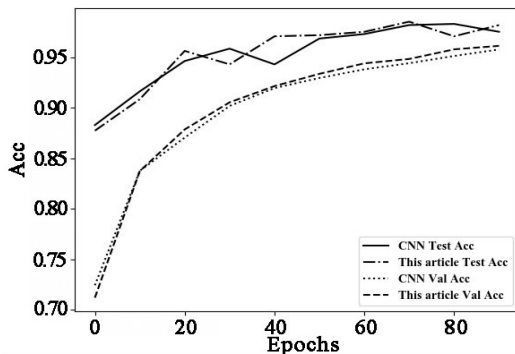


图5 识别率随迭代次数变化曲线

文中将数据集按照 8 : 1 : 1 的比例划分成训练

集、验证集和测试集,其中训练集用于训练网络参数,验证集不参与网络参数训练,用于调整迭代数和学习率,可以认为训练集和验证集都参与了网络的训练过程,所有完全没有参与过训练的测试集用于网络训练完之后识别率测试。测试集完全不参与网络训练,所以识别率较验证集会有所下降。测试集可以体现出网络泛化能力。

表 1 ORL 和 Yale 人脸数据库识别率 %

算法	识别率
ORL-PCA	83.5
ORL-LBP	89.7
ORL-CNN	95.4
ORL-文中算法	97.5
Yale-PCA	87.8
Yale-LBP	85.2
Yale-CNN	90.1
Yale-文中算法	91.9

由图 5 可以得出,文中提出的算法经过 LBP 算法提取人脸特征纹理并进行数据集扩充后,在测试集的识别率高于使用原始数据集的测试集识别率,泛化能力得到了提升。由表 1 得出,文中提出的识别率同时也高于传统的人脸识别算法。实验结果表明,经过局部二值模式处理提取面部纹理特征,扩充数据集,能够有效地提升测试集识别率和卷积神经网络的泛化能力。

4 结束语

针对卷积神经网络人脸识别普遍存在的训练数据少,容易造成训练特征不足以及网络的过拟合,泛化能力不足等问题,引入局部二值模型和深度卷积生成对抗网络。在图像预处理方面,首先通过局部二值模型提取出人脸面部纹理特征,再通过 DCGAN 训练生成图像,扩充数据集,提高卷积神经网络的泛化能力。在 ORL 人脸数据库和 Yale 人脸数据库地测试结果表明,经过局部二值模式处理且扩充数据能提升卷积神经网络的泛化能力,同时识别率也得到了提升。

参考文献:

[1] KWON O W, LEE T W. Phoneme recognition using ICA-

based feature extraction and transformation[J]. Signal Processing,2004,84(6):1005-1019.

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

[3] 刘 君,黄燕琪,熊邦书.融合核主成分分析和最小距离鉴别投影的人脸识别算法[J]. 计算机工程,2016,42(4):221-225.

[4] 徐静妹,李 雷.基于稀疏表示和支持向量机的人脸识别算法[J]. 计算机技术与发展,2018,28(2):59-63.

[5] 闫 河,王 朴,刘 婕,等.自适应加权多尺度 LTP 的人脸识别[J]. 计算机工程与设计,2016,37(4):1027-1031.

[6] 李彦冬,郝宗波,雷 航.卷积神经网络研究综述[J]. 计算机应用,2016,36(9):2508-2515.

[7] 周飞燕,金林鹏,董 军.卷积神经网络研究综述[J]. 计算机学报,2017,40(6):1229-1251.

[8] LECUN Y L,BOTTOU L,BENGIO Y,et al. Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE,1998,86(11):2278-2324.

[9] KRIZHEVSKY A,SUTSKEVER I,HINTON G. ImageNet classification with deep convolutional neural networks[J]. Advances in Neural Information Processing Systems,2012,25(2):1097-1105.

[10] HE K,ZHANG X,REN S,et al. Deep residual learning for image recognition[C]//IEEE conference on computer vision and pattern recognition. [s.l.]:IEEE,2015:770-778.

[11] 张洁玉,赵鸿萍,陈 曙.自适应阈值及加权局部二值模式的人脸识别[J]. 电子与信息学报,2014,36(6):1327-1333.

[12] 刘中华,史恒亮,张兰萍,等.基于多尺度局部二值模式的人脸识别[J]. 计算机科学,2009,36(11):293-295.

[13] RADFORD A,METZ L,CHINTALA S. Unsupervised representation learning with deep convolutional generative adversarial networks[C]//International conference on learning representations (ICLR). [s.l.]: Computational and Biological Learning Society,2016:1151-1167.

[14] GOODFELLOW I,POUGET-ABADIE J,MIRZA M,et al. Generative adversarial nets[C]//Advances in neural information processing systems. Palais des Congr s de Montr al, Montr al, Canada:[s.n.],2014:2672-2680.

[15] 叶 浪.基于卷积神经网络的人脸识别研究[D]. 南京:东南大学,2015.

[16] 卢官明,何嘉利,闫静杰,等.一种用于人脸表情识别的卷积神经网络[J]. 南京邮电大学学报:自然科学版,2016,36(1):16-22.