

doi:10.11835/j.issn.1000-582X.2018.05.012

中心损失与 Softmax 损失联合监督下的人脸识别

余成波, 田 桐, 熊递恩, 许琳英

(重庆理工大学 远程测控与信息处理研究所, 重庆 400054)

摘要:深度学习在人脸识别领域已经取得了巨大的成就,针对当前大多数卷积神经网络采用 Softmax 损失函数进行特征分类,增加新的类别样本会减小类间距离的增长趋势,影响网络对特征判别的问题,采用了一种基于中心损失与 Softmax 损失联合监督的人脸识别算法,来提高网络对特征的识别能力。在 Softmax 基础上,首先,分别对训练集每个类别在特征空间维护一个类中心,训练过程新增加样本时,网络会约束样本的分类中心距离,从而兼顾了类内聚合与类间分离。其次,引入动量概念,在分类中心更新的时候,通过保留之前的更新方向,同时利用当前批次的梯度微调最终的更新方向,该方法可以在一定程度上增加稳定性,提高网络的学习效率。最后,在人脸识别基准库 LFW 上的测试实验证明:所提的联合监督算法,在较小的网络训练集上,获得了 99.31% 的人脸识别精度。

关键词:深度学习;中心损失;Softmax 损失;动量;人脸识别

中图分类号:TP391.4

文献标志码:A

文章编号:1000-582X(2018)05-092-09

Joint supervision of center loss and softmax loss for face recognition

YU Chengbo, TIAN Tong, XIONG Di'en, XU Linying

(Institute of Remote Test and Control, Chongqing University of Technology,
Chongqing 400054, P.R.China)

Abstract: Nowadays, deep learning has made great achievements in face recognition. Most of the convolutional neural network uses the Softmax loss function to increase the distance between classes. However, adding samples of new classes will reduce the distance between classes and the performance of the network. In order to improve the recognition ability of the network characteristics, a face recognition approach based on joint supervision of center loss and Softmax loss is proposed. On the basis of Softmax, first of all, a class center is maintained in the feature space for each class of the training set. When a new sample is added to the training process, the network will constrain the distance of the classification center of the sample, and thus both intra-class aggregation and inter-class separation are considered. Secondly, the

收稿日期:2017-11-20

基金项目:国家自然科学基金资助项目(61402063);重庆市科技人才培养计划资助项目(CSJC2013KJRC-TDJS40012);重庆市高校优秀成果转化资助项目(KJZH14213)。

Supported by Natural Science Foundation of China (61402063), Chongqing Science and Technology Talent Training Program Funding Project (CSJC2013KJRC-TDJS40012) and Financed Projects for the Transformation of Excellent Achievement of University in Chongqing (KJZH14213).

作者简介:余成波(1965—),男,博士,重庆理工大学教授,主要从事远程测试与控制技术、信号与信息处理方向研究,(E-mail)yuchengbo@cqut.edu.cn。

田桐(联系人),女,硕士研究生,主要从事图像处理、信号与信息处理方向研究,(E-mail)669734574@qq.com。

concept of momentum is introduced. When the classification center is updated, by retaining the previous update direction and using the gradient of the current batch to fine-tune the final update direction, the method can increase the stability and improve the learning efficiency of the network. Finally, the test experiments on the face recognition benchmark library LFW (labeled faces in the wild) prove that the proposed joint supervision algorithm achieves 99.31% of face recognition accuracy on a small network training set.

Keywords: deep learning; central loss; Softmax loss; momentum; face recognition

近年来,随着深度学习的不断发展,人脸识别技术已经获得了飞跃式的发展^[1]。相较于传统的人脸识别技术,如:特征脸、局部特征分析法、基于几何特征法等,深度卷积神经网络(CNNs, convolutional neural networks)进行的人脸识别,能有效的克服诸如光照、多姿态、面部遮挡等干扰。深度卷积神经网络是通过不断的训练从图片中提取特征,克服了手动选择中人为的主观性^[2]。对于 CNNs 的常用人脸识别框架如图 1 所述。

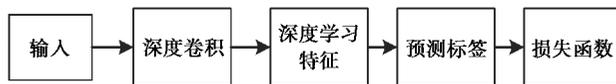


图 1 人脸识别框架图

Fig.1 Face recognition framework

早期的人脸识别通过不断的研究网络结构来提升人脸的识别精度,其中较著名的网络有 Lenet^[3]、Alexnet^[4]、Googlenet^[4]、VGG^[5]、Deep Residual Learning^[6]等,其研究结果表现在网络越深,获得的识别精度就越高。然而,这些网络结构需要大量的人脸数据库、较大的计算资源及消耗大量的计算时间,而让许多研究者望而却步,并成为人脸识别的研究瓶颈。为了解决这个问题,文献[7]提出了迁移学习方法,并得到迅速发展,被广大研究者广泛使用。迁移学习可以利用现有的训练好的网络模型,在样本数据不是特别充足的情况下获得较好的结果。但迁移学习的前提是训练数据具有相似性,否则性能将会降低。

由于相似性的训练数据不具有普遍性,而简单的网络结构不能满足需求,因此研究者将目光转向了对数据特征关系的研究。在 2005 年,文献[8]中就提出可通过计算 2 幅人脸图像的相似距离来识别人脸,孪生网络^[9]被广泛使用,孪生网络能够很好的提取 2 个输入获取到的特征的相似性。2014 年,Taigman 等人^[10]通过对人脸进行三维对称处理,使用 DeepFace 孪生网络结构进行训练,该方法在 LFW(labeled faces in the wild)人脸数据库上进行测试的精度达到了 97.25,接近人眼识别精度。Sun 等人^[11]在 DeepFace 的基础上提出了 DeepID(deep hidden identity features)网络,采用 Softmax 损失函数作为监督信号,通过在 CelebFaces 和 WDRef 结合的较大的人脸数据库上进行训练,最终在 LFW 上达到了 97.35%的精度。

Softmax 损失监督网络在面对未训练过的新标签样本时,由于网络对特征的泛化能力不足,其识别效果不佳。为了解决这个问题,文献[12]中提出使用识别信号和验证信号进行联合监督,其中识别信号能增大不同人脸的差异性,验证信号能减小人脸类内差异,使得在 LFW 上识别精度达到 99.15%。与选取某一层直接使用损失进行分类的方法不同,文献[13]提出了 triplet 损失函数作为监督信号,直接学习图像到欧式空间上点的映射,2 张图像所对应的特征的欧式空间上的点的距离直接对应着 2 个图像是否相似,来获得更好的分类效果,该方法在 LFW 上达到了 99.63%的精度。

文献[13]中的方法对训练样本数量要求很高,并且计算量增大,训练步骤也变得复杂。为了解决这个问题,研究采用了一种中心损失与 Softmax 损失进行联合监督的方法,该方法在增大不同人脸间距离的同时,减小同一个人脸内部的距离,从而使学习到的特征具有更好的泛化性和辨别能力,提高深度卷积网络对特征的识别能力。其优点是收敛速度更快、所需样本数量比其他方法少,最后,通过在基准 LFW 库的实验也证明了算法的有效性,获得了 99.31%的人脸识别精度。

1 损失函数及算法步骤

在卷积神经网络的前向传播中,通过网络最后一层预测结果与真实结果的比较得到损失,计算误差更新值,通过反向传播对网络的所有权值进行更新。其中,损失层是前向传播的终点,也是反向传播的起点。通过对预测值和真实值的计算,得到损失函数(loss function)的值,而分类中的优化过程实质就是对损失函数进行最小化的过程。可以采用梯度下降法等优化算法计算最小值。

1.1 Softmax 损失

Softmax 函数是 Logistic 函数^[14]的推广,主要用于解决多分类问题。Softmax 所获得的结果代表输入图像被分到每一类的概率,若是一个 K 分类器,则输出是一个 K 维的向量(向量中元素的和为 1)。

对于 m 个样本,为训练集及其所对应的标签为: $\{(x^{(1)}, y^{(1)}), (x^{(2)}, y^{(2)}), \dots, (x^{(m)}, y^{(m)})\}$, 其每一个样本的 k 个估计概率表示如式(1)所示

$$\mathbf{h}_\theta(x^{(i)}) = \begin{bmatrix} p(y^{(i)} = 1 | x^{(i)}; \mathbf{W}) \\ p(y^{(i)} = 2 | x^{(i)}; \mathbf{W}) \\ \vdots \\ p(y^{(i)} = k | x^{(i)}; \mathbf{W}) \end{bmatrix} = \frac{1}{\sum_{j=1}^k e^{\mathbf{w}_j^T x^{(i)}}} \begin{bmatrix} e^{\mathbf{w}_1^T x^{(i)}} \\ e^{\mathbf{w}_2^T x^{(i)}} \\ \vdots \\ e^{\mathbf{w}_k^T x^{(i)}} \end{bmatrix}, \quad (1)$$

其中: $\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \dots, \mathbf{w}_k$ 为网络模型参数; k 为分类个数; $\frac{1}{\sum_{j=1}^k e^{\mathbf{w}_j^T x^{(i)}}}$ 对结果进行归一化,同时所有概率之和为 1。

为了完成网络训练,需要对 Softmax 的损失函数采用梯度下降法,使其得到的结果达到最小值,其损失函数如公式(2)所示

$$L_s = -\frac{1}{m} \left[\sum_{i=1}^m \sum_{j=1}^k 1\{y^{(i)} = j\} \log \frac{e^{\mathbf{w}_j^T x^{(i)}}}{\sum_{l=1}^k e^{\mathbf{w}_l^T x^{(i)}}} \right], \quad (2)$$

其中, $\{y^{(i)} = j\}$ 是一个指示性函数,当其值为真时,该函数结果就为 1,否则结果就为 0。式(2)可简化为

$$L_s = -\frac{1}{m} \left[\sum_{i=1}^m \log \frac{e^{\mathbf{w}_{y^{(i)}}^T x^{(i)}}}{\sum_{l=1}^k e^{\mathbf{w}_l^T x^{(i)}}} \right]. \quad (3)$$

在实际应用中,为了不使任意的 $(\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3, \dots, \mathbf{w}_k)$ 参数为 0,通常会在给损失函数加入一权重衰减。损失函数越大,该分类器分类结果为真实标签的概率越小,通过逐次迭代计算损失函数的最小值,得到全局最优解,获得最好性能。

1.2 中心损失

文献[15]首次提出中心损失函数,该函数是一种典型的聚类算法。对于一个具有多个特征的 CNN 来说,在每个批次中可以计算出多个特征中心,计算特征中心的同时计算其损失函数。其损失函数根据该特征值与其对应中心的距离计算得来,因此损失函数可以通过公式(4)、(5)、(6)得出

$$L_c = \frac{1}{2m} \sum_{i=1}^m \|x_i - c_{y_i}\|_2^2, \quad (4)$$

$$\frac{\partial L_c}{\partial x_i} = x_i - c_{y_i}, \quad (5)$$

$$\Delta c_j = \frac{\sum_{i=1}^m \delta(y_i = j) \cdot (c_j - x_i)}{1 + \sum_{i=1}^m \delta(y_i = j)}, \quad (6)$$

其中: x_i 表示第 i 张图片的特征值; c_{y_i} 则表示该第 i 张图片所属分类的中心(该分类的特征值的中心); Δc_j 则表示分类中心的变化。式(5)描述了类内变化情况,其分类的特征中心 c_{y_i} 随着深度特征的变化而变化。式(6)中,当 $\delta(y_i = j)$ 取 0 或者 1,也就是当条件 $y_i = j$ 满足时, $\delta(y_i = j) = 1$,而当条件不满足时, $\delta(y_i = j) = 0$ 。

由于训练集较大,在训练过程中要一次性获取所有的中心是不现实,为了获得最优解,文献[15]在每一个小批量中对中心进行分别更新,如式(7)所述。

$$c_j^{t+1} = c_j^t + \Delta c_j^t, \quad (7)$$

其中, c_j^t 表示的是样本在 t 时刻的特征中心,在网络训练过程中需要对特征中心不断更新。受到梯度下降中动量^[16]概念的启发,为了解决文献[15]中心更新方法完全依赖当前的每一个批量,造成不稳定,在中心更新中引入动量来进行,该方法通过梯度和上一次的权重更新值的线性组合来更新特征中心。其基本思路如图 2 所示,首先,按照原来的更新方向更新一步(①线),然后在该位置计算梯度值(②线),然后用这个梯度值修正最终的更新方向(③线)。图 2 中描述了 2 步的更新示意图,其中④线是标准动量更新路径。

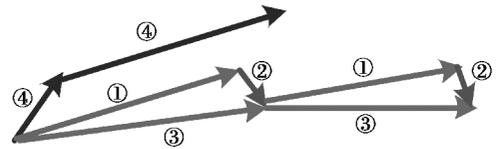


图 2 特征中心更新方法思路图

Fig.2 Loss joint supervision diagram

该特征中心更新方法的描述如式(8)所示

$$\Delta c_j^t = \eta \Delta c_j^{t-1} - \rho \nabla L_j^t(c_j^t + \eta \Delta c_j^{t-1}), \quad (8)$$

其中: ρ 表示学习率,取值为 $0-1$; ∇L_j^t 则表示 c_j 在批次的梯度; η 表示上一次更新的权重,这 2 个参数需要通过调整来得到最好的结果,一般是根据经验值进行设定的。

至此,经过上述 2 个步骤的修改,解决了原中心函数在对中心更新时,只考虑了当前批次问题。

1.3 Softmax 损失与中心损失联合监督函数

通过将 Softmax 损失与中心损失(center loss)进行联合监督,在增大类间距离的同时,减小类内距离,使其获得的特征具有更强的识别能力,如式(9)所示

$$L = L_S + \lambda L_C = -\frac{1}{m} \left[\sum_{i=1}^m \log \frac{e^{w_{y(i)}^T x^{(i)}}}{\sum_{l=1}^k e^{w_l^T x^{(i)}}} \right] + \frac{\lambda}{2} \sum_{i=1}^m \|x_i - c_{y_i}\|_2^2, \quad (9)$$

式中: λ 用来平衡 2 个损失,合适的 λ 选择有助于网络特征识别能力的增强; c_{y_i} 是 y_i 类的样本中心,其中,当 $\lambda=0$ 时,可以将该函数认为是仅有 Softmax 损失的情况。

通过联合监督方法不仅解决了对比损失中无法采集合适的样本对的问题,同时也解决了三重损失中复杂采样的问题。在该中心损失中,采用了对比损失中同类样本损失公式,用同类样本中的类中心替换了其中的一个同类样本,而异类时,是直接采用 Softmax 损失进行计算。

1.4 网络参数配置及算法步骤

为了有效的实现非限制条件下的人脸识别,将 Softmax 损失与中心损失相结合,其联合损失的监督示意图如图 3 所示,对于整个训练网络的具体参数的配置见实验部分。

通过对网络结构的研究,构建网络,对于联合监督算法部分,将算法的输入抽象为:训练数据集 $\{x_i\}$,分别初始化池化层和 loss 层的参数 θ_c, W 及 $\{c_j | j=1, 2, \dots, n\}$,同时输入参数: λ, α 和学习率 μ' ,其中迭代的次数 t 初始化为 0。输出则抽象为输出参数 θ_c ,因此归纳联合监督的人脸识别算法步骤如下所示。

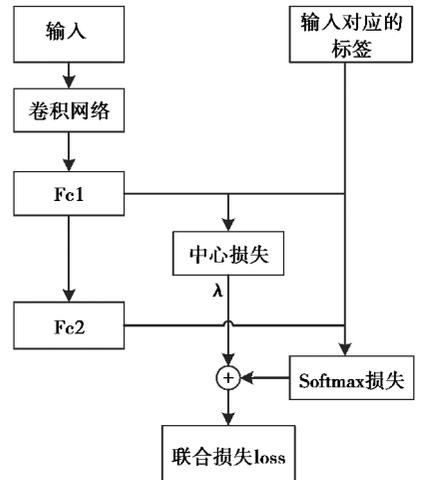


图 3 损失联合监督示意图

Fig.3 Loss joint supervision diagram

步骤 1:判断是否收敛,不收敛则执行步骤 2;

步骤 2:迭代次数 t 自增,即 $t \leftarrow t + 1$;

步骤 3:计算联合损失 L^t ; $L^t = L_S^t + L_C^t$

步骤 4:计算反向传播误差 $\frac{\partial L^t}{\partial x_i^t}$;通过公式 $\frac{\partial L^t}{\partial x_i^t} = \frac{\partial L_S^t}{\partial x_i^t} + \lambda \cdot \frac{\partial L_C^t}{\partial x_i^t}$

对每一张图片求解;

步骤 5:更新参数 W ;通过公式 $W^{t+1} = W^t - \mu' \cdot \frac{\partial L^t}{\partial W} = W^t - \mu' \frac{\partial L_S^t}{\partial W}$ 进行计算;

步骤 6:更新参数 c_j ,通过公式 $\Delta c_j^t = \eta \Delta c_j^{t-1} - \rho \nabla L_j^t(c_j^t + \eta \Delta c_j^{t-1})$ 公式进行详细计算;

步骤 7:更新参数 θ_c ,通过公式 $\theta_c^{t+1} = \theta_c^t - \mu' \sum_i \frac{\partial L^t}{\partial x_i^t} \cdot \frac{\partial x_i^t}{\partial \theta_c^t}$ 进行计算;

步骤 8:判断是否收敛,若收敛则迭代结束,否则,从步骤 1 开始重复执行。

2 实验及结果分析

为验证算法的有效性,选用人脸基准库 LFW 人脸库进行人脸识别实验。实验选用网络中单独使用 Softmax 损失监督的 Model A、Softmax 损失与对比损失联合监督的 Model B 几个使用最广泛的网络模型与算法的 Model C 进行对比分析。

2.1 网络模型及数据集

2.1.1 网络模型

VGGNet^[17]是 Andrew Zisserman 教授研究团队(Oxford)提出的一种 CNN,该网络相较于 AlexNet 具有更深的网络结构,通常有 16~19 层,且其所有卷积层都使用了同样的大小的卷积核,大小为 3×3 。VGGNet 在 2014 年的 ILSVRC 定位和分类 2 个问题上分别取得了第一名和第二名。因此,全部采用 3×3 的网络结构来组建网络。此外,受 Resnet^[18]的启发,在 VGGNet 基础上添加残差网络结构,其网络结构如图 4 所示。其卷积层由一个 3×3 的步长为 1 的卷积网络和一个 PReLU 激活函数组成。对于图 4 中的 LC 层,其结构如图 5 所示。其中,特征图大小分别为 128,256,521,2 个全连接层的输出维度都为 512。

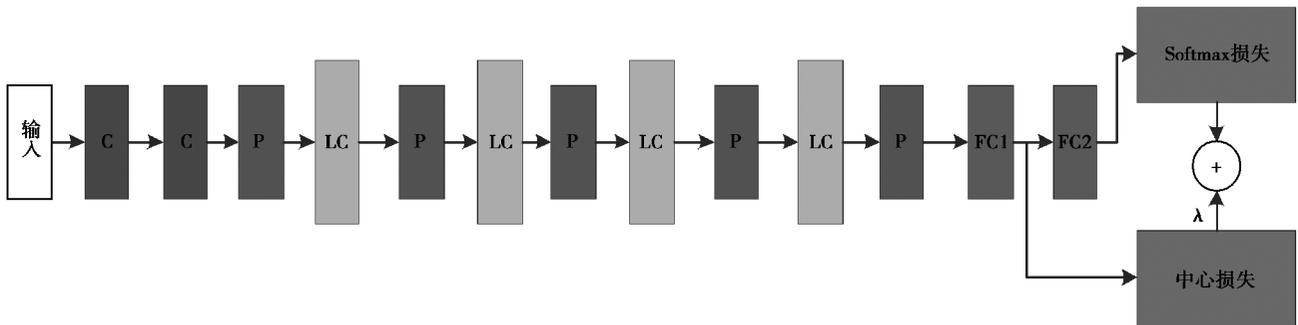


图 4 网络结构模型图

Fig.4 Network model

2.1.2 数据集

① 训练数据集

选取 CASIA-WebFace 作为训练数据集,其包含 10 575 人的约 500 K 张图片。由于该数据库与 LFW 有 3 个人脸是重复的,选取了 10 572 个人的人脸图像,对其进行剪切、对齐处理(图像剪切为 112×96),采用 4:1 的比例拆分训练数据和验证数据。在训练过程中,对训练和验证数据做镜像处理。其训练数据和验证数据如图 6 所示。

② 测试数据集

选取目前广泛使用的 LFW 标准人脸识别数据集,该数据集是专门用来研究人脸识别问题的数据集。该数据集包含 5 749 人,共 13 233 张人脸图像。LFW 定义了 3 个标准的协议(无监督、限制与非限制)去评估人脸识别的性能。这里选取了 6 000 组测试对比集,其中 3 000 组为相同人的对比集,3 000 组为不同人的对比集。经过对人脸进行矫正和剪切的预处理后^[19-20],其人脸如图 7 所示,其中绿色方框中的人脸表示 2 组中的人脸是同一个人,红色方框中的人脸表示 2 组中的人脸不是同一个人。

2.2 不同参数下人脸识别精度的影响

研究所提算法中,通过使用 ρ 来控制中心的学习率, λ 用来平衡 Softmax 及中心 2 个损失,从而加强网络对特征的识别能力。 η 则用来校正特征中心的更新,因此这 3 个参数对所述算法起着至关重要的作用,下面分别对其进行讨论。

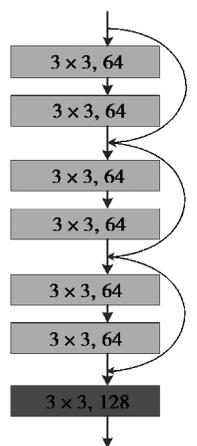


图 5 LC 层结构框图

Fig.5 Structure diagram of LC

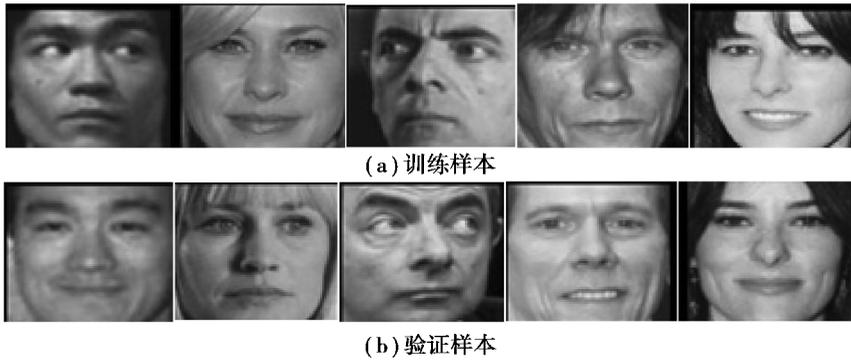


图 6 训练及验证数据集

Fig.6 Data sets of Training and validation



图 7 测试数据集

Fig.7 The test data set

1)参数 λ 对人脸识别精度的影响

本实验,通过固定 ρ 及动量 η ,改变 λ 参数的方式进行试验,从参考文献[15]中知道,当 $\rho=0.5$ 的时候,能够获得较高的中心学习率。因此,第一个试验,固定 $\rho=0.5, \eta=0.5$,而使 λ 从 0 变化到 1,从而学习到不同的网络模型。

在 LFW 数据集上的测试结果如图 8 所示,由图可知,所提 Softmax 损失及中心损失联合监督算法的优越性,其人脸识别测试精度,要高于仅 Softmax 损失监督下(当 $\lambda=0$ 时,训练网络仅有 Softmax 损失监督)的测试精度。同时,从图 8 中也可以得到,一个合适的 λ 值有助于提高人脸识别精度,参数 λ 的最佳结果为 0.003。

2)学习率 ρ 对人脸识别精度的影响

根据本章实验(1)中获得的最优结果,本节实验固定 $\lambda=0.003, \eta=0.5$,而让 ρ 从 0.01 变化到 1,从而学习到不同的模型。

分别采用这些模型在精准库 LFW 上做人脸识别测试,可以得到该实验的人脸识别精度如图 9 所示。由图 9 分析得出,选择合适的 ρ 值,同样有助于提高训练网络的人脸识别精度。根据实验分析,参数 ρ 的最佳结果为 0.6。

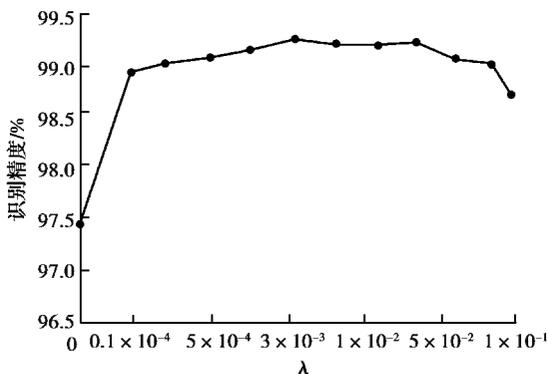


图 8 基于 LFW 数据集上的人脸识别精度($\alpha=0.5, \eta=0.5$)

Fig.8 Face recognition accuracy based on LFW data set

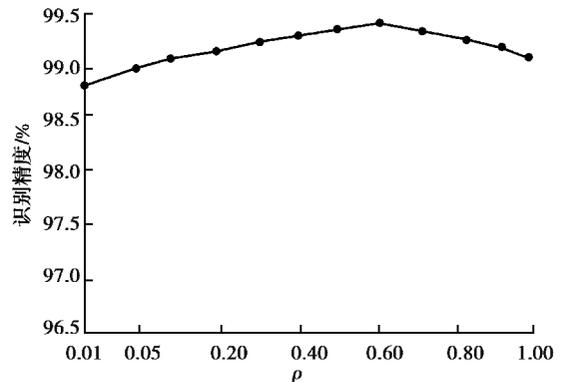


图 9 LFW 数据集上的人脸识别精度($\lambda=0.003, \eta=0.5$)

Fig.9 Face recognition accuracy on LFW data set

3) 参数 η 对人脸识别精度的影响

根据实验(1)、(2)获得的最优结果,本节实验通过固定 $\rho=0.6, \lambda=0.003$, 而让参数 η 在 0 到 1 之间 η 发生变化,从而训练到不同的网络模型。通过 LFW 数据集上的测试,得到该实验的人脸识别精度,其实验结果如图 10 所示。

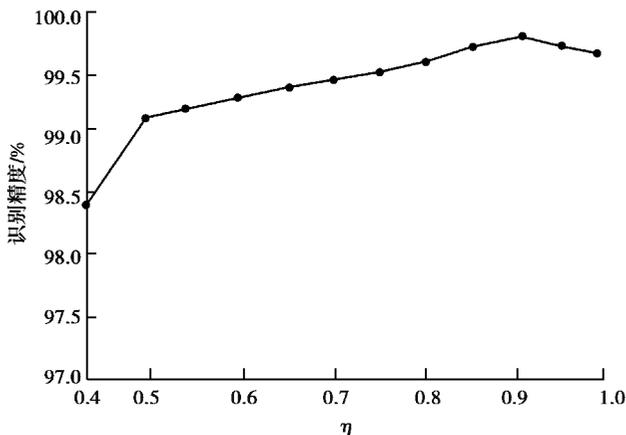


图 10 LFW 数据集上的人脸识别精度 ($\lambda=0.003, \rho=0.6$)

Fig.10 Face recognition accuracy on LFW data set

由图 10 分析得出,一个合适的动量选择,能够提高训练网络的人脸识别精度,从实验结果来看,当 η 取 0.9 时,人脸识别精度最高。同时,将其与图 9、图 10 对比发现,最佳的 λ, ρ 及 η 值,能够在总体上提高人脸识别精度。

2.3 不同网络模型下的人脸识别精度比较

ROC(receiver operating characteristic)曲线和 AUC(area under curve)常被用来评价一个二值分类器(binary classifier)的优劣^[21]。其中,AUC 被定义为 ROC 曲线下的面积。为了说明所提联合监督算法在人脸识别精度中的优势,将其与目前获得了最高人脸识别精度的几种经典网络模型比较分析,其中,网络模型是在 CAISA-webface 数据集上训练而来的,该数据集中的任何一张人脸都与 LFW 上的人脸不同,并且所采用的参数 $\rho=0.6, \lambda=0.003, \eta=0.9$ 。

在实验中,通过人脸验证试验绘制 ROC 曲线,从而评价算法训练下的分类特征的好坏。人脸验证试验是:对于 LFW 中给定的 6 000 组测试人脸,分别判断他们是否是同一个人。经过试验,在 LFW 测试集中达到了 99.31% 的平均精度,与表 1 中所列方法的平均精度和 ROC 曲线的比较分别如表 1 和图 11 所示,其中用 Model A 表示仅使用 Softmax 损失函数监视的模型,Model B 表示 Softmax 损失和对比损失联合监视的模型,而 Model C 则表示所提模型。

表 1 不同模型下的人脸识别性能比较

Table 1 Recognition of face recognition performance under different models

方法	图片/张	网络个数/个	Acc. On LFW
DeepFace	4 M	3	97.35% ± 1.13
DeepID-2+	—	1	98.70% ± 1.08
FaceNet	200 M	1	99.63% ± 0.26
Baidu	1.3M	1	99.13% ± 0.13
原始 center loss	0.7 M	1	99.28%
Model A	0.7 M	1	97.35% ± 0.08
Model B	0.7 M	1	99.24% ± 0.11
Model C(Proposed)	0.7 M	1	99.31% ± 0.10

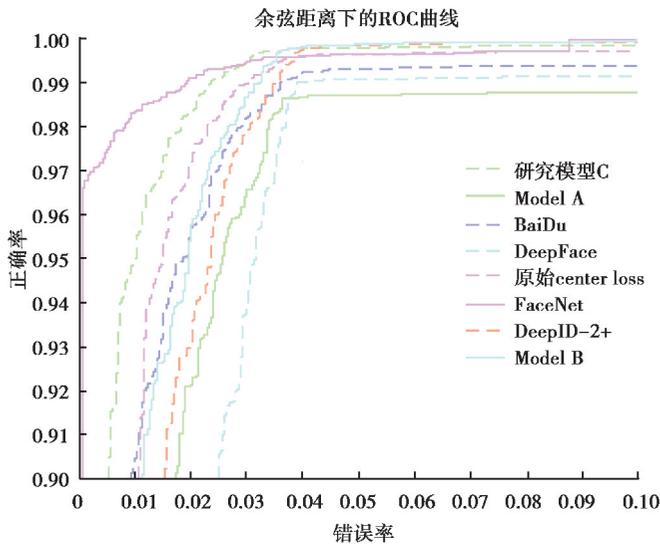


图 11 LFW 数据集上人脸识别的 ROC 曲线

Fig.11 ROC of face verification on LFW

表 2 不同 center loss 下的人脸识别性能比较

Table 2 The performance of face recognize under different center loss

方法	时间/h	训练轮数	训练样本(总数(张))
原始 center loss	14	28 K	17,189(0.7 M)
研究模型 C	12	24 K	10,572(0.7 M)

从表 1 及图 11 的实验数据分析来看,所提算法(Softmax 损失与 center 损失联合监督)比单一的损失监督 Model A,在相同训练数据图片的情况下获得了较高的识别精度,在 LFW 基准库上识别精度从 97.35% 提高到 99.31%。该对比结果可以证明,联合监督方法相比单一监督方法,在对特征的分类能力上有显著的提升。其次,研究还将算法与当前使用最为广泛的联合监督 Model B 进行比较,从实验结果来看,表 1 显示 Model C 的人脸识别精度从原来的 99.28% 提高到 99.31%,表 2 可见其训练时间已经缩短。实验表明网络收敛速度已经加快,在减小类内距离上有更加显著的优势。最后,表 1 还列出了当前表现最好的几种经典模型,结合图 11 的 ROC 曲线。可见所提算法模型的识别结果,能够在使用较小训练数据集的情况下,获得较高的人脸识别精度。

3 结 论

笔者提出了一种基于中心损失与 Softmax 损失联合监督的人脸识别算法,在 LFW 人脸库上的实验结果表明:算法能够提高网络对特征的识别能力,将中心损失与 Softmax 损失相结合,克服了 Softmax 损失在面对未训练过的新标签样本时,泛化能力不足的问题。同时,重点引入动量的概念,在分类中心更新的时候,通过保留之前的更新方向,同时利用当前批次的梯度微调最终的更新方向,该方法可以在一定程度上增加稳定性,提高网络的学习效率。

最后,实验表明:在较小的网络训练集的基础上,算法获得了 99.31% 的人脸识别精度,从而验证了算法在提高网络对特征的识别能力上的优越性。

参考文献:

- [1] Parkhi O M, Vedaldi A, Zisserman A, et al. Deep face recognition[C]// British Machine Vision Conference, Swansea, UK: British Machine Vision Association, 2015: 41-42.

- [2] Mufti H N, Abidi S, Abidi S R, et al. Predictors of post-operative delirium in cardiac surgery patients; a machine learning approach[J]. *Canadian Journal of Cardiology*, 2014, 30(10): 237-238.
- [3] LeCun L, Bottou Y, Bengio P, et al. Gradient-based learning applied to document recognition[J]. *Proceedings of the IEEE*, 1998, 86(11): 2278-2323.
- [4] Ballester P, Araujo R M. On the performance of GoogLeNet and alexNet applied to sketches[C]// *Thirtieth AAAI Conference on Artificial Intelligence*. Phoenix, USA: AAAI Press, 2016: 1124-1128.
- [5] Day M J, Horzinek M C, Schultz R D, et al. Guidelines for the vaccination of dogs and cats[J]. *Journal of Small Animal Practice*, 2007, 48(9): 528.
- [6] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition[G]// *Computer Vision and Pattern Recognition*, Las Vegas, USA: IEEE, 2015: 770-778.
- [7] Sinno J, Yang Q. A survey on transfer learning[J]. *IEEE Transactions on Knowledge & Data Engineering*, 2010, 22(10):1345-1359.
- [8] Hadsell R, Chopra S, Lecun Y, et al. Dimensionality reduction by learning an invariant mapping[C]// *Computer Vision and Pattern Recognition*, Conference on 2006 IEEE Computer Society. New York: IEEE, 2006: 1735-1742.
- [9] Chopra S, Hadsell R, Lecun Y, et al. Learning a similarity metric discriminatively, with application to face verification[C]// *Computer Vision and Pattern Recognition*, 2005. New York: IEEE, 2005: 539-546.
- [10] Taigman Y, Yang M. Deepface: closing the gap to human-level performance in face verification[C]// *Computer Vision and Pattern Recognition*. Columbus, USA: IEEE, 2014: 1701-1708.
- [11] Sun Y, Wang X, Tang X. Deep learning face representation from predicting 10,000 classes[C]// *IEEE Conference on Computer Vision and Pattern Recognition*. Columbus, USA: IEEE Computer Society, 2014: 1891-1898.
- [12] Sun Y, Wang X, Tang X, et al. Deep learning face representation by joint identification-verification[C]// *NIPS Montred*, CANADA: MIT Press, 2014, 27: 1988-1996.
- [13] Cheng D, Gong Y, Zhou S, et al. Person reidentification by multi-channel parts-based CNN with improved triplet loss function[C]// *Computer Vision and Pattern Recognition*, Las Vega S, USA: IEEE, 2016: 1335-1344.
- [14] King G, Zeng L. Logistic regression in rare events data[J]. *Political Analysis*, 2001, 9(2): 137-163.
- [15] Wen Y, Zhang K, Li Z, et al. A discriminative feature learning approach for deep face recognition[C]// *Computer Vision-ECCV 2016*. Amsterdam: Springer International Publishing, 2016: 11-26.
- [16] Qian N. On the momentum term in gradient descent learning algorithms[J]. *Neural Networks the Official Journal of the International Neural Network Society*, 1999, 12(1): 145-151.
- [17] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition[J]. *Computer Science*, 2014, 6(1): 1409-1556.
- [18] Szegedy C, Ioffe S, Vanhoucke V, et al. Inception-v4, Inception-ResNet and the impact of residual connections on learning[C]// *AAAI*. Palo Alto, California: AAAI, 2017: 4278-4284.
- [19] Kazemi V, Sullivan J. One millisecond face alignment with an ensemble of regression trees[C]// *Computer Vision and Pattern Recognition*. Columbus, USA: IEEE, 2014: 1867-1874.
- [20] Klawonn F, Höppner F, May S. An alternative to ROC and AUC analysis of classifiers[C]// *International Conference on Advances in Intelligent Data Analysis X*. Berlin: Springer, 2011: 210-221.