

陈琳, 邹远文. 人脸识别任务中角边距损失函数的对比研究[J]. 智能计算机与应用, 2024, 14(7): 160–164. DOI: 10.20169/j.issn.2095-2163.240725

人脸识别任务中角边距损失函数的对比研究

陈琳, 邹远文

(四川大学 生物医学工程学院, 成都 610065)

摘要: 深度卷积神经网络 (DCNNs) 和嵌入 (Embedding) 学习人脸特征是人脸识别任务的常用方法。在 DCNNs 网络中广泛使用 Softmax 损失函数, 将欧几里得边距强加到 DCNNs 学习到的特征矩阵中, 获得较好人脸识别正确率, 但是仍具有较大的提高空间。近年来提出的角边距损失函数进一步提升了人脸识别任务的正确率, 但是人脸识别任务中角边距损失函数的对比研究较少, 且缺乏在不同类别人脸数据集上对比角边距损失函数对人脸识别正确率的影响。本文使用 MS1MV2 数据集训练 ResNet50 模型, 以 Softmax 损失函数为基准, 对比研究 ArcFace、CosFace、Combined Margin (CM) 3 种角边距损失函数在 3 个不同类别人脸数据集 LFW、CFP-FP 和 AGEDB-30 上的识别准确率, 并调整角边距损失函数的参数值, 讨论其对人脸识别任务性能的影响。实验结果表明, 角边距损失函数较 Softmax 损失函数可以显著提升模型人脸识别正确率。在不同数据集和不同任务条件下, 选择合适的角边距损失函数和参数设置可以提高人脸识别性能。

关键词: 角边距损失函数; 人脸识别; 深度卷积神经网络

中图分类号: TP391.41

文献标志码: A

文章编号: 2095-2163(2024)07-0160-05

A comparative study of angular margin loss functions applied to face recognition

CHEN Lin, ZOU Yuanwen

(College of Biomedical Engineering, Sichuan University, Chengdu 610065, China)

Abstract: Using Deep Convolutional Neural Networks (DCNNs) and embedding learning to extract facial features is a common method in face recognition tasks. The Softmax loss function is widely used in DCNNs, where Euclidean distance is applied to the feature matrix learned by DCNNs, resulting in good face recognition accuracy, though there is still significant room for improvement. Recently proposed angular margin loss functions have further improved face recognition accuracy. However, comparative studies on angular margin loss functions in face recognition tasks are relatively scarce, and there is a lack of comparative research on the impact of angular margin loss functions on face recognition accuracy across different facial datasets. In this study, we trained a ResNet50 model using the MS1MV2 dataset and used the Softmax loss function as a baseline. We conducted a comparative study of three angular margin loss functions—ArcFace, CosFace, and Combined Margin (CM)—and evaluated their recognition accuracy on three different facial datasets: LFW, CFP-FP, and AGEDB-30. Additionally, we adjusted the parameter m of the angular margin loss functions and discussed its impact on the performance of face recognition tasks. Experimental results show that angular margin loss functions significantly improve the model's face recognition accuracy compared to the Softmax loss function. Choosing the appropriate angular margin loss function and parameter settings for different datasets and task conditions can enhance face recognition performance.

Key words: margin loss function; face recognition; Deep Convolutional Neural Networks

0 引言

使用深度卷积神经网络 (DCNNs) 和嵌入 (Embedding) 学习人脸特征是人脸识别任务的常用

方法^[1]。目前人脸识别研究方向分为人物分类和人脸验证。直接训练 DCNNs 可以得到对人脸进行多类别分类的网络, 例如使用 Softmax 损失函数训练得到的 DeepFace 网络^[2]。DCNNs 结合嵌入学习

作者简介: 陈琳 (1997-), 女, 硕士研究生, 主要研究方向: 信号处理, 图像处理。

通讯作者: 邹远文 (1967-), 男, 博士, 硕士生导师, 主要研究方向: 生物医学信息检测、处理和控制。Email: zyw@scu.edu.cn

收稿日期: 2023-12-14

(Embedding Learning)可以用于人脸验证,例如使用三元组损失函数(Triplet Loss)进行嵌入学习的Facenet模型^[3]。对于大规模训练数据驱动的DCNNs网络结构,使用Softmax损失函数的方法和三元组损失函数的方法都可以在人脸识别任务上获得较好的性能。然而,Softmax损失函数和三元组损失函数都是将欧几里得边距强加给模型学习到的特征矩阵中,在学习人脸特征时具有局限性^[4]。Softmax损失函数:线性变换矩阵 $W \in R^{d \times n}$ 的大小随恒等数 n 线性增加;对于闭集分类,学习到的特征是可分的,但对于开集人脸识别,学习到的特征没有足够的区分度。三元组损失函数:对于大规模数据集,人脸三元组的数量出现组合爆炸,导致迭代步数显著增加;对于样本差异较小的数据集,训练模型较困难。

有工作基于欧几里得边距的损失函数进行改进,Wen等^[5]提出了中心损失函数;Deng等^[6]提出了边距损失函数;Zhang等^[7]提出了范围损失函数。但是以上方法训练得到的模型对人脸识别正确率提升有限。

Liu等^[8]提出了SphereFace损失函数。SphereFace损失函数首次将Softmax损失函数与乘法角边距结合,将原始欧几里德特征空间投影到角度空间,同时加强类内紧凑性和类间差异性,从而使训练模型具有更好的判别能力。

尽管在SphereFace损失函数中引入了角边距(Angular Margin)这一重要概念,但其损失函数需要一系列近似值才能计算,而且Softmax损失函数在训练过程中占主导地位,基于整数的乘法角边距损失函数使得目标对数几率曲线非常陡峭,因此阻碍了收敛,导致网络训练不稳定。

为了稳定训练,Wang等^[9]提出了一个包含标准Softmax损失函数的混合损失函数CosFace。CosFace损失函数直接将余弦边距(Cosine Margin)惩罚添加到目标对数几率曲线,与SphereFace损失函数相比,CosFace损失函数使分类DCNNs在人脸识别任务上获得了更好的分类正确率,也更容易实现;Deng等^[10]提出了一种加性角边距损失函数ArcFace,进一步提高人脸识别模型的判别力并稳定训练过程。SphereFace、CosFace、ArcFace 3种角边距损失函数的角边距参数设置结合,得到CM(Combined Margin)损失函数。

综上所述,SphereFace损失函数由于其收敛性较差,模型难以稳定训练,在实际应用中使用较少。

在其基础上提出的CosFace、ArcFace和CM 3种角边距损失函数能够稳定地进行模型训练,具有较为广泛的应用,因此对比研究CosFace、ArcFace和CM 3种角边距损失函数在人脸识别任务中的表现和差异,有助于进一步探讨角边距损失函数对提高人脸识别性能的作用。本文使用MS1MV2数据集训练ResNet50模型,以Softmax损失函数为基准,对比研究ArcFace、CosFace、CM 3种角边距损失函数在3个不同类别人脸数据集LFW、CFP-FP和AGEDB-30上的识别准确率,并调整角边距损失函数的参数 m 值,讨论其对人脸识别任务性能的影响。

1 损失函数

1.1 Softmax损失函数和角边距损失函数

Softmax损失函数学习到的特征会按照一定角度分布^[8]。为了更好地加强类间的差异性和类内紧凑性,即增大不同类特征的角边距,引入了角边距概念。由于余弦与Softmax损失函数具有内在一致性,且余弦公式经常用于人脸识别中的相似度测量。因此在Softmax损失函数的基础上,直接引入余弦余量得到SphereFace损失函数,以增强余弦相关信息的判别能力。

从二元分类问题开始分析Softmax损失函数,其决策边界,式(1):

$$(W_1 - W_2)x + b_1 - b_2 = 0 \quad (1)$$

其中, W_i 、 b_i 是权重和偏置, x 为向量。

对 W_i 和 b_i 进行以下约束,式(2)和式(3):

$$\|W_1\| = \|W_2\| = 1 \quad (2)$$

$$b_1 = b_2 = 0 \quad (3)$$

则决策边界,式(4):

$$\|x\|(\cos(\theta_1) - \cos(\theta_2)) = 0 \quad (4)$$

其中, θ_i 是 W_i 和 x 之间的夹角。

新的决策边界就仅仅取决于 θ_1 和 θ_2 。优化后的Softmax损失函数能够直接优化角度,使得神经网络能够学习角度分布的特征。

虽然与传统的Softmax损失函数相比,修改后的损失函数学习到的特征是有角度分布的,但是不一定容易分类。因此,将修正后的损失函数进一步推广到A-Softmax损失函数,即SphereFace损失函数。

在优化后的Softmax损失函数基础上,SphereFace损失函数引入 $m(m \geq 1)$ 来定量控制决策边界。在二分类的情况下,第一类和第二类的决策边界如式(5)和式(6)所示:

$$\|\mathbf{x}\|(\cos(m\theta_1) - \cos(\theta_2)) = 0 \quad (5)$$

$$\|\mathbf{x}\|(\cos(\theta_1) - \cos(m\theta_2)) = 0 \quad (6)$$

其中, m 是超参数, 可以定量地控制角边距的大小。

由此可得 SphereFace 损失函数公式(7):

$$L = \frac{1}{N} \sum_i -\log\left(\frac{e^{\|\mathbf{x}_i\| \psi(\theta_{y_i,i})}}{e^{\|\mathbf{x}_i\| \psi(\theta_{y_i,i})} + \sum_{j \neq y_i} e^{\|\mathbf{x}_i\| \cos(\theta_{j,i})}}\right) \quad (7)$$

其中, $\psi(\theta_{y_i,i})$ 如公式(8)所示, $\theta_{y_i,i} \in$

$$\left[\frac{k\pi}{m}, \frac{(k+1)\pi}{m}\right], k \in [0, m-1], m \geq 1.$$

$$\psi(\theta_{y_i,i}) = (-1)^k \cos(m\theta_{y_i,i}) - 2k \quad (8)$$

通过优化 SphereFace 损失函数, 使判决区域更加分离, 同时扩大了类间间隔, 压缩了类内角度分布。SphereFace 损失函数具有清晰的几何解释。在 SphereFace 损失函数的监督下, 学习到的特征构造了一个判别角距离度量, 该度量等价于超球流形上的测地距离。SphereFace 损失可以被解释为将学习到的特征约束为可区分的超球面流形上, 其本质上与人脸图像位于流形上的先验相匹配^[11]。SphereFace 损失函数和超球流形之间的紧密联系使得学习到的特征在人脸识别中更加有效。

1.2 CosFace 损失函数

CosFace 损失函数在 SphereFace 损失函数的基础上进行改进。首先对特征和权重向量进行归一化消除径向变化, 并且将 Softmax 损失函数重新表述为余弦损失。与 SphereFace 损失函数不同的是, CosFace 损失函数引入的余弦余量项 m 是在余弦空间中定义的, 决策的边距不再取决于 θ , 使得决策边界如式(9)所示:

$$\cos(\theta_1) - m = \cos(\theta_2) \quad (9)$$

当满足条件(10)~(12)时:

$$W = \frac{W^*}{\|W^*\|} \quad (10)$$

$$x = \frac{x^*}{\|x^*\|} \quad (11)$$

$$\cos(\theta_j, i) = W_j^T x_i \quad (12)$$

CosFace 损失函数如式(13)所示:

$$L = \frac{1}{N} \sum_i -\log\left(\frac{e^{\|\mathbf{x}_i\| (\cos(\theta_{y_i,i}) - m)}}{e^{\|\mathbf{x}_i\| \psi(\theta_{y_i,i})} + \sum_{j \neq y_i} e^{\|\mathbf{x}_i\| \cos(\theta_{j,i})}}\right) \quad (13)$$

1.3 ArcFace 损失函数

ArcFace 在目标角度上添加一个加性的角度余量, 并通过余弦函数再次得到目标对数几率曲线, 公式(14):

$$L = \frac{1}{N} \sum_i -\log\left(\frac{e^{\|\mathbf{x}_i\| (\cos(\theta_{y_i,i} + m))}}{e^{\|\mathbf{x}_i\| \psi(\theta_{y_i,i})} + \sum_{j \neq y_i} e^{\|\mathbf{x}_i\| \cos(\theta_{j,i})}}\right) \quad (14)$$

与 CosFace 不同的是 CosFace 虽然实现了余弦量线性变化, 但是其角度变化仍然是非线性的^[5]。因此 ArcFace 具有更好的几何属性, 在整个区间内具有恒定的线性角边距, 而更好的边界设计可以进一步提高训练的速度。

1.4 Combined Margin 损失函数

在 SphereFace、ArcFace 和 CosFace 损失函数中, 提出了 3 种不同的角边距惩罚, 例如乘法角 m_1 , 加性角边距 m_2 和余弦加性角边距 m_3 。从数值分析的角度来看, 以上 3 种损失函数, 不同的角边距惩罚都通过惩罚目标对数几率曲线来加强类内紧致性和类间差异性。通过结合所有边距惩罚得到 Combined Margin 损失函数, 使得 m_1, m_2 和 m_3 作为超参数在统一的框架中实现了角边距损失函数, 公式(15):

$$L = \frac{1}{N} \sum_i -\log\left(\frac{e^{\|\mathbf{x}_i\| (\cos(m_1\theta_{y_i,i} + m_2) - m_3)}}{e^{\|\mathbf{x}_i\| \psi(\theta_{y_i,i})} + \sum_{j \neq y_i} e^{\|\mathbf{x}_i\| \cos(\theta_{j,i})}}\right) \quad (15)$$

2 实验

2.1 数据集

本文使用了 MS1MV2 作为训练数据, 在训练期间用 LFW, CFP-FP, AgeDB-30 等 3 种不同类别人脸验证数据集进行测试, 其中 LFW 数据集来源于真实生活场景, CFP-FP 是模拟口罩人脸数据集, AgeDB-30 是年龄差为 30 岁的数据集具体见表 1。

表 1 用于训练和测试的人脸数据集

Table 1 Dataset for training and testing

数据集	个体数量	图像总数
MS1MV2	85 000	5 800 000
LFW	5 749	13 233
CFP-FP	500	7 000
AgeDB-30	568	16 488

2.2 网络结构与参数设置

本文使用 ResNet50 模型作为骨干网络, 在 ResNet50 模型最后一个卷积层之后, 使用 BN-Dropout-FC-BN (BatchNorm-Dropout-Fully Connection-BatchNorm) 结构, 最终获得 512 维的嵌入特征。

本文实验使用 Pytorch 深度学习框架, 将批量大小设置为 512, 并在一个 NVIDIA 3090 GPU 上训练模型。用 MS1MV2 数据集训练模型时, 优化器选择 SGD, 动量设置为 0.9, 权重衰减设置为 $5e-4$, 学习

率从 0.1 开始衰减,如图 1 所示。

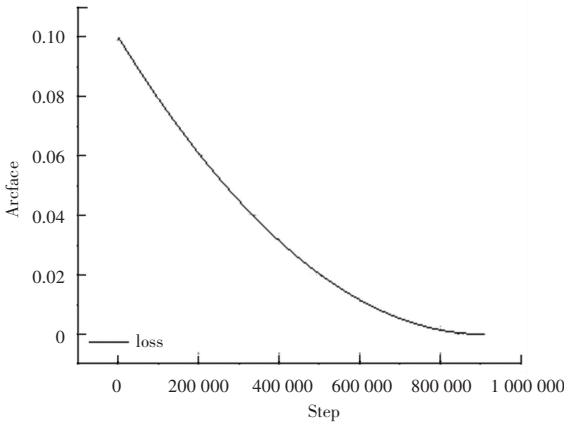


图 1 学习率的变化曲线

Fig. 1 Change curve of learning rate

2.3 对比实验

为深入探讨基于角边距的损失函数在加强类内紧凑性和类间差异性方面的作用,本文选择 Softmax 损失函数训练得到的模型作为基准模型。

使用 ArcFace 损失函数在 MS1MV2 上训练 ResNet50 模型,角边距超参数 m 设置为最佳值 0.5^[10]。随后引入 CosFace 损失函数,并将角边距 m 分别设置为 0.5、0.4 和 0.35,以评估不同 m 值对模型性能的影响。

此外,本文使用组合角边距进行了更广泛的实验,测试了两种 Combined Margin 损失函数,角边距超参数 m_1, m_2 和 m_3 分别设置为 (1.0, 0.3, 0.2) 和 (0.9, 0.4, 0.15)。

将上述角边距损失函数训练得到的模型在 LFW、CFP-FP 和 AgeDB-30 3 个数据集上进行测试,对比不同角边距损失函数在人脸识别任务中的表现,实验结果见表 2。

表 2 实验结果

Table 2 Results of loss function on datasets

方法	LFW	CFP-FP	AgeDB-30
Arcface(0.5)	0.998 33	0.976 29	0.981 00
Cosface(0.5)	0.998 17	0.975 43	0.980 17
Cosface(0.4)	0.997 83	0.978 14	0.980 50
Cosface(0.35)	0.998 17	0.978 43	0.979 67
CM1(1,0.3,0.2)	0.997 67	0.970 29	0.974 83
CM2(0.9,0.4,0.15)	0.996 67	0.965 86	0.970 17
Softmax	0.993 50	0.934 29	0.939 33

由表 2 可知,整体而言,基于角边距的损失函数训练得到的模型在 3 个测试数据集上的表现均优于基准 Softmax 损失函数训练所得模型的性能,表明

角边距损失函数可以增强判别深度学习特征的能力。Softmax 损失函数与角边距损失函数相比,角边距损失函数在 CFP-FP 和 AgeDB 上的表现均更加优秀。进一步说明在特征更混淆的情况下,角边距惩罚可以增强对嵌入特征的判别能力。

在人脸识别基准数据集 LFW 上和年龄分布跨度较大的 AgeDB-30 数据集上,加性角边距惩罚 ArcFace 损失函数训练模型性能较其他角边距损失函数稍显优势,但是 ArcFace 和 CosFace ($m = 0.5$) 损失函数之间的性能差异均小于 0.1%。各损失函数在 LFW 上的表现如图 2 所示。

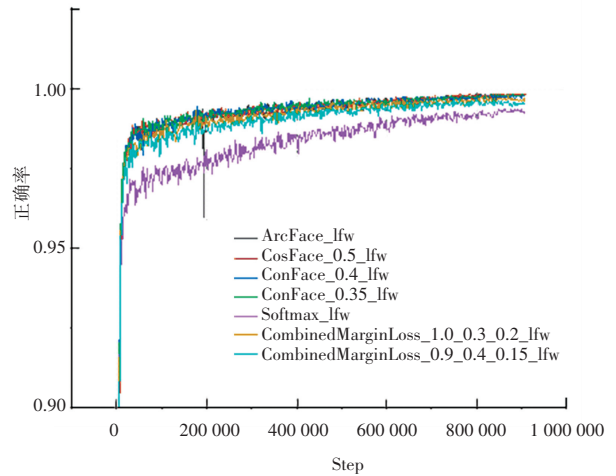


图 2 损失函数在 LFW 上的表现

Fig. 2 Loss functions performer on LFW

在局部遮掩的人脸数据集 CFP-FP 上,当 CosFace 损失函数的角边距 m 设置为 0.35 时,分类正确率达到最佳水平,不仅优于其它角边距损失函数,还超过了 CosFace 损失函数在其它 m 值下的表现,表明在无遮掩的人脸数据集上, ArcFace 损失函数在加强类内紧凑型和类间差异性上稍有优势。在人脸被遮掩时 CosFace ($m = 0.35$) 损失函数可以加强模型的人脸识别性能; Combined Margin 损失函数 (CM1 和 CM2) 同样可以获得较好的性能,虽然整体性能不如单独使用 CosFace 和 ArcFace 损失函数,但是相差无几,考虑到 CosFace 损失函数 3 种参数设置最后的性能差异,可以认为 Combined Margin 损失函数的性能差异是由于参数设置不是最优导致的。

3 结束语

本文使用 MS1MV2 数据集训练 ResNet50 模型,以 Softmax 损失函数为基准,对比研究 ArcFace、CosFace、Combined Margin (CM) 3 种角边距损失函数在 3 个不同类别人脸数据集 LFW、CFP-FP 和

AGEDB-30 上的识别准确率,验证了基于角边距损失函数在人脸识别问题上的优越性。实验结果表明,角边距损失函数较 Softmax 损失函数可以显著提升模型人脸识别正确率。在不同数据集和不同任务条件下,选择合适的角边距损失函数和参数设置可以提高人脸识别性能;在人脸局部遮掩的数据集, CosFace ($m = 0.35$) 损失函数最佳;日常场景和年龄分布方差较大的人脸数据集, ArcFace 损失函数表现最好,但与 CosFace ($m = 0.5$) 损失函数差距极小。

参考文献

- [1] SCHROFF F, KALENICHENKO D, PHILBIN J. Facenet: A unified embedding for face recognition and clustering [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2015; 815-823.
- [2] TAIGMAN Y, YANG M, RANZATO M A, et al. DeepFace: Closing the gap to human-level performance in face verification [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2014; 1701-1708.
- [3] SCHROFF F, KALENICHENKO D, PHILBIN J. Facenet: A unified embedding for face recognition and clustering [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2015; 815-823.
- [4] LEE J, WANG Y, CHO S. Angular margin-mining softmax loss for face recognition[J]. IEEE Access, 2022, 10; 43071-43080.
- [5] WEN Y, ZHANG K, LI Z, et al. A discriminative feature learning approach for deep face recognition [C]// Proceedings of Computer Vision - ECCV 2016; 14th European Conference. Amsterdam: IEEE, 2016; 499-515.
- [6] DENG J, ZHOU Y, ZAFEIRIOU S. Marginal loss for deep face recognition [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. IEEE, 2017; 60-68.
- [7] ZHANG X, FANG Z, WEN Y, et al. Range loss for deep face recognition with long-tail[J]. arXiv preprint arXiv:1611.08976, 2016.
- [8] LIU W, WEN Y, YU Z, et al. SphereFace: Deep hypersphere embedding for face recognition [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2017; 212-220.
- [9] WANG H, WANG Y, ZHOU Z, et al. CosFace: Large margin cosine loss for deep face recognition [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2018; 5265-5274.
- [10] DENG J, GUO J, XUE N, et al. ArcFace: Additive angular margin loss for deep face recognition [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. IEEE, 2019; 4690-4699.
- [11] HE X, YAN S, HU Y, et al. Face recognition using laplacianfaces [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2005, 27(3); 328-340.