

文章编号: 2095-2163(2024)03-0061-06

中图分类号: TP391.41

文献标志码: A

# 融合 MobileNet 与 Contextual Transformer 的人脸识别研究

陈经纬, 熊继平, 程汉权

(浙江师范大学 物理与电子信息工程学院, 浙江 金华 321004)

**摘要:** FaceNet 作为人脸识别的一大跨越,以其高精度、低硬件配置等优势被广泛应用于各个人脸识别相关领域。本文开源了首个餐厅支付场景下的中国人脸数据集 CN-Face,该数据集拥有 13 000 人的脸图像,总计 100 000 张。此外,本文以 CASIA-WebFace 作为训练集,利用改进后的 MobileNet 主干网络,采取不同的注意力机制添加方法,改变激活函数并且融入 Contextual Transformer 模块,大大降低了参数量和识别速度,显著提升了人脸识别精度。相较于原版 FaceNet,在 LFW 测试集下,准确率达到 98.79%,提升了 2.74%,在 CN-Face 数据集中准确率达到 95.22%,提升了 1.35%。

**关键词:** ECA 注意力机制; 人脸识别; FaceNet; 深度学习; Contextual Transformer

## Research on face recognition combining MobileNet and Contextual Transformer

CHEN Jingwei, XIONG Jiping, CHENG Hanquan

(College of Physics and Electronic Information Engineering, Zhejiang Normal University, Jinhua 321004, Zhejiang, China)

**Abstract:** FaceNet, as the major progress in face recognition, has been widely used in related face recognition fields due to its advantages of high precision and low hardware configuration requirements. The laboratory collects face data from multiple restaurants and combines them to form a Chinese face dataset (CN-Face) with more than 13 000 face IDs and 100 000 face images. Explore the use of CASIA-WebFace as the training set, use the improved MobileNet backbone network, adopt different attention mechanism addition methods, change the activation function and integrate the Contextual Transformer module, which greatly reduces the number of parameters and recognition speed, and significantly improves face recognition precision. Compared with the original FaceNet, under the LFW test set, the accuracy rate reaches 98.79%, an increase of 2.74%. At the same time, the accuracy rate reaches 95.22% under the Chinese face dataset (CN-Face) collected in the laboratory, an increase of 1.35%.

**Key words:** ECA attention mechanism; face recognition; FaceNet; deep learning; Contextual Transformer

## 0 引言

人脸识别是个人身份认定中的重要一环,相较于指纹识别和虹膜识别等身份认定手段,人脸识别具有无接触性、非强制性、并发性等优势,在金融安全、智能监控领域具有广泛的应用。人脸识别系统的基础是提取面部特征,并使用这些特征进行比较。其中,深度学习(DL)模型提供了一种成功提取分析数据特征的方法<sup>[1]</sup>。许多应用都需要先进的人脸识别技术,例如获取人脸信息、人机界面应用、多媒体通信、创建人造人脸和管理基于内容的人脸数据库<sup>[2]</sup>。每个比较过程的接近值,是通过将识别过程中被质疑图像的独特特征与人脸数据库中的其他特征集进行比较来计算,通过考虑所有计算出的接近

度值来获得结果。人脸识别应用程序的质量与人脸识别数据库的创建,以及人脸识别过程中获得的人脸姿势的质量成正比<sup>[3]</sup>。人脸匹配最关键的是确定 2 个不同人脸的特征表示,以确保正确匹配<sup>[4]</sup>。

人脸识别自上世纪 90 年代被 Turk 团队实现并开始应用迄今为止已经历了 30 余年<sup>[5]</sup>,各种人脸识别算法也陆续问世,一直作为身份认定的重要方法。21 世纪,随着深度学习的兴起,运用深度学习挖掘人脸特征成为人脸识别的主要途径。2006 年, Hinton 团队利用深度学习,提出了深度信念网络,挖掘图像之间的非线性特征,获取图像的高维特征<sup>[6-7]</sup>。卷积神经网络(CNN)的提出,加快了人脸图像识别的发展。2014 年, Taigman 团队提出 Deep Face 深度学习模型<sup>[8]</sup>,准确率达到 97.35%,突破

基金项目: 金华市公益项目(2021-4-116)。

作者简介: 陈经纬(1996-),男,硕士研究生,主要研究方向:深度学习。

通讯作者: 熊继平(1982-),男,博士,副教授,主要研究方向:深度学习、图像处理。Email: xjping@zjnu.cn

收稿日期: 2023-08-07

当时人脸识别准确率的最高记录。同年, Sun 团队在人脸识别任务中引入了度量学习, 开发出 DeepID<sup>[4]</sup>。2015 年, Google 提出 FaceNet 人脸识别模型, 采用了卷积神经网络和 *Triplet loss* 作为损失函数, 大大提高了人脸识别的精度<sup>[9]</sup>。利用网络将人脸图像映射入一个可测的向量空间, 通过计算人脸向量的欧氏距离, 判定人脸相似度。近年, 对损失函数的修改取得有效的进展, AdaCos<sup>[10]</sup>、ArcFace<sup>[11]</sup>、Adaptive Face<sup>[12]</sup>、Circle Loss<sup>[13]</sup> 等损失函数的提出, 对人脸识别的识别率、推断速度等方面有着不同程度的改进。人脸数据集的发展是人脸识别的重要途径, 主要分为两大方向。其一, 对数据集进行提纯清洗, 例如通过清理 MS-Celeb-1M 数据集(包括 10 万人的 1 亿张人脸图像), iBUG 得到 MS1M-iBUG 数据集(包括 8.5 万人的 380 万张人脸图像), 同样是 MS-Celeb-1M 数据集, ArcFace 得到 MS1M-ArcFace 数据集(包括 8.5 万人的 580 万张人脸图像); 其二, 通过收集更多的人脸图像, 当前最大的人脸数据集为文献[13]提出的数据集 WebFace260M(包括 400 万人的 2.6 亿张人脸图像), 并且设计了可以自动利用且自训练的自动清洗 pipeline, 所以 WebFace260M 人脸数据集虽然巨大, 但也具有相当的清洁度。

## 1 模型介绍

### 1.1 FaceNet

FaceNet 是谷歌研究人员开发的用于人脸识别的深度神经网络模型。该模型使用卷积神经网络(CNN)架构, 从面部图像中提取特征, 然后将这些特征映射到称为“面部嵌入”的高维空间。该模型使用大量人脸数据集进行训练, 即使在光照、姿势和表情变化等条件下也能准确识别人脸。FaceNet 网络结构如图 1 所示。

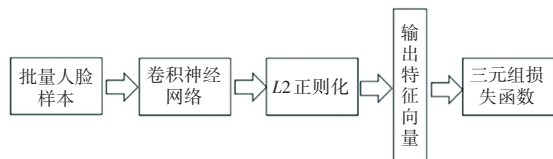


图 1 FaceNet 网络结构

Fig. 1 FaceNet network structure

FaceNet 的关键创新是在训练期间使用三元组损失函数 (*Triplet loss*), 其作用是使模型将同一个人的图像映射到相似的人脸嵌入, 将不同人的图像映射到不同的嵌入, 有助于模型学习, 可以区分不同个体的判别特征。

*Triplet loss* 三元组损失是深度学习中用于度量学习的一种损失函数, 其目的是学习将输入映射到度量空间的函数, 使得相似的输入在度量空间中比不同的输入彼此更接近。度量空间可用于图像检索、人脸识别和行人再识别等任务。

三元组损失函数接受 3 个输入: 锚样本、正样本和负样本。其中, 锚样本和正样本是同一类/类别的 2 个示例, 而负样本是不同类/类别的示例。三元组损失函数计算锚点和正样本间的距离, 以及度量空间中锚点和负样本间的距离。目标是最小化锚点和正样本间的距离, 同时最大化锚点和负样本间的距离。*Triplet loss* 可以表示为:

$$\mathcal{L} = \max(d(a, p) - d(a, n) + \text{margin}, 0) \quad (1)$$

其中,  $d(a, p)$  是 anchor 样本  $a$  (取一张人脸图像) 和正样本  $p$  (同一人的不同人脸图像) 的距离;  $d(a, n)$  是 anchor 样本  $a$  (取一张人脸图像) 和负样本  $n$  (不同人的脸像) 的距离;  $\text{margin}$  作为一个超参数, 指定为零锚样本同负样本间最小期望距离。若锚点同正样本间的距离较之锚点同负样本间的距离至少小  $\text{margin}$ , 则损失为零, 否则损失为正, 并且与距离间的差异成正比。

三元组损失函数通常与连体或三元组网络架构结合使用, 该架构由共享权重的同一网络的多个副本组成, 并采用多个输入, 其输入将用于三元组中的每个样本。通过使用随机梯度下降或类似的优化算法优化三元组损失函数来训练网络。*Triplet loss* 训练阶段流程如图 2 所示。

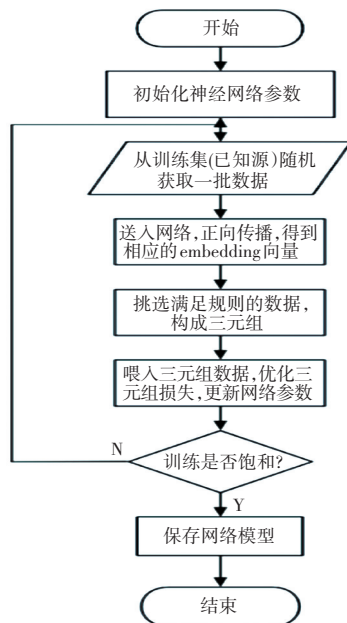


图 2 *Triplet loss* 训练阶段流程图

Fig. 2 Flow chart of the training phase of *Triplet loss*

## 1.2 主干网络

MobileNet 是一种深度神经网络架构<sup>[14]</sup>,专为高效移动和嵌入式视觉应用而设计,已广泛用于移动设备的计算机视觉应用程序,以及其他资源受限的环境。

MobileNet 的关键思想是使用深度可分离卷积代替传统卷积。通过将卷积操作分成 2 个单独的层:depthwise convolution 层对每个输入通道应用一个过滤器;pointwise convolution 层应用了  $1 \times 1$  卷积来组合跨通道的 depthwise 卷积结果。以此减少网络的计算成本和参数数量,同时保持视觉识别任务的高精度。

MobileNet 由 3 个阶段和自适应平均池化层与全连接层构成,其中 3 个阶段分别为 *stage1*、*stage2*、*stage3*。*stage1* 和 *stage2* 包含 6 个卷积层,*stage3* 包含 2 个卷积层。

MobileNet 还使用了一种“瓶颈”技术,用来进一步降低计算成本。在瓶颈中,输入通道首先使用  $1 \times 1$  卷积扩展到更多通道,然后使用深度可分离卷积进行处理,不仅有助于提高网络的表示能力,同时还保持了较低参数数量。

## 2 实验方法

### 2.1 数据集

#### 2.1.1 训练集

CASIA-WebFace 是中国科学院自动化研究所(CASIA)开发的大规模人脸识别数据集<sup>[15]</sup>,其中包含从互联网上收集的超过 10 000 个人的 500 000 多张图像。在这个数据集中涵盖了每一个人不同姿势、表情、光照条件和年龄的图片,并标记了每个人的身份,便于数据集后续的扩展。

CASIA-WebFace 数据集广泛用于人脸识别算法的开发和评估,尤其是基于深度学习的方法。数据集的庞大规模和多样性,使其可用于训练稳健且准确的人脸识别系统。

#### 2.1.2 测试集

##### 2.1.2.1 LFW

LFW(Labeled Faces in the Wild)是一种流行的人脸识别数据集<sup>[16]</sup>,由马萨诸塞大学的研究人员创建。该数据集包含从互联网上收集的 13 000 多张人脸图像,每张图像都标有图像中人物的身份。

LFW 数据集广泛用于评估人脸识别算法的性能,特别是用于人脸验证任务(即确定 2 张图像是

否显示同一个人)。这是一个具有挑战性的数据集,因为图像通常是低质量的,在不同的光照条件下拍摄的,并以各种姿势和表情显示面部。

本文实验利用 pairs 文件随机读取图像对,并通过计算人脸图像特征向量的欧氏距离,将人脸阈值确定为 0.90,测试所得准确率为 96.05%。

##### 2.1.2.2 CN-Face

实验室通过与中国某公司合作,经过人脸截取、图像聚类、图片清理等操作,共采集得到 134 00 个人的 100 000 张左右的人脸图像。该数据集仅包含中国人脸,为当前最大的中国人脸数据集。

该数据集可用于 FaceNet 人脸训练和人脸图像准确率测试。本实验采用的测试集为部分上述数据集,共包含 857 个人的 8 000 余张人脸图像。相较于 LFW 测试集利用 pairs 文件随机读取图像,本实验通过遍历整个人脸数据集,共产生 320 万人脸图像对。人脸阈值定为 0.90 时,测试所得准确率为 93.87%。

## 2.2 FaceNet 融合 ECA 注意力机制模块

ECA-Net 是香港中文大学研究人员于 2020 年推出的一种深度神经网络架构<sup>[17]</sup>。ECA-Net 的主要思想是通过引入一种更有效的新型通道注意机制,用以提高卷积神经网络(CNN)的性能。

ECA-Net 中使用的通道注意机制基于一种“有效通道注意”的新技术。该技术沿通道维度对每个特征图应用一维卷积核,并对输出应用 *sigmoid* 激活函数,以生成一组注意力权重,然后使用这些注意力权重来缩放特征图,有效地放大或抑制结果响应。

ECA-Net 中使用的有效通道注意机制与以往的通道注意方法相比,需要的参数和计算量少很多,同时在 ImageNet 分类和 COCO 对象检测等各种基准上实现了相当、甚至更好的性能。

ECA 注意力机制侧重于通过对每个卷积层内通道之间的交互进行建模,来学习通道方面的特征依赖性。通过将具有固定内核大小的一维卷积运算应用于每个卷积层的输出,接着使用逐元素乘法,将此操作的输出与原始特征图组合。

ECA 注意力机制旨在计算高效且易于实施,同时仍能在图像识别基准测试中取得最先进的结果,已被证明在准确性和计算效率方面优于其他注意力机制,如挤压和激发(SE)和全局上下文(GC)注意力。ECA 注意力机制模块如图 3 所示。

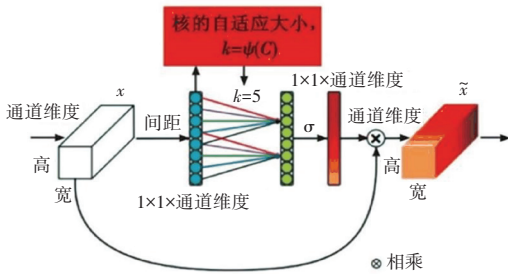


图3 ECA 注意力机制模块

Fig. 3 ECA attention mechanism module

本文实验利用2种方式将FaceNet与ECA注意力机制融合:

(1)用ECA模块代替一层DW卷积。在 $stage_2$ 中,利用ECA模块取代一层DW卷积层。采用通道注意力机制,相当于对通道做减法,能有效减少参数量和推断所需时间成本。

在LFW测试集下得到了95.32%的准确率,虽然准确率降低了0.73个百分点,但参数量较之原版减少了27%,推断所需时间减少30%。有利于在低配置硬件条件下达到实时检测效果。

(2)直接加入ECA模块。在 $stage_2$ 中,间隔dw卷积层,加入3个ECA注意力机制模块。在ECA注意力机制模块中,一维卷积核大小 $K$ 默认为 $K=5$ 。实验旨在探求人脸识别任务中,ECA模块添加最合适位置及一维卷积核 $K$ 的最佳大小。

ECA-Layer不同 $K$ 值测试结果见表1。由表1可知,当 $K=7$ 时,在LFW测试集上取得的准确率最高,为98.31%。

表1 ECA-Layer不同 $K$ 值测试结果Table 1 Different  $K$  value test results of ECA-Layer %

一维卷积核 $K$	准确率
3	92.87
5	97.81
7	98.31
9	97.40
11	87.63

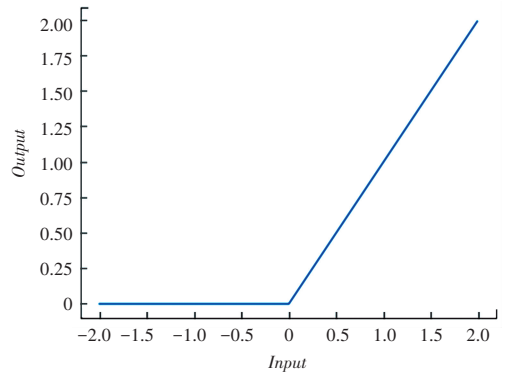
### 2.3 SiLU 激活函数

$ReLU$  (Rectified Linear Unit)<sup>[18]</sup>是神经网络中使用的激活函数,函数的数学公式为:

$$f(x) = \max(0, x) \quad (1)$$

其中, $x$ 表示函数的输入, $f(x)$ 表示输出。

激活函数 $ReLU$ 如图4所示。由图4可知,当输入为正时,函数返回与输入相同的值;当输入为负时,函数返回0。

图4 激活函数 $ReLU$ Fig. 4 Activation function  $ReLU$ 

MobileNet原版使用 $ReLU$ ,为适配ECA模块,现将激励函数 $ReLU$ 改为 $SiLU$ 。

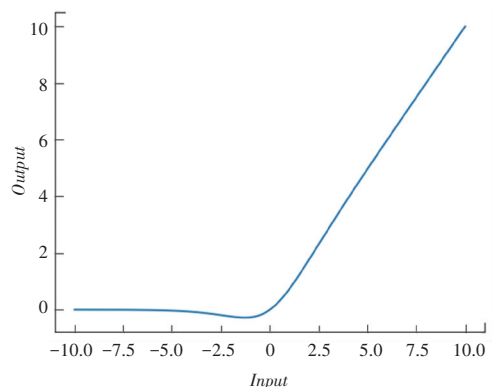
$SiLU$  (Sigmoid 线性单元)是神经网络中使用的激活函数<sup>[19]</sup>,也被称为 $Swish$ ,是“具有线性输出的自门控激活函数”的缩写。 $SiLU$ 函数的数学公式为:

$$f(x) = \frac{x}{1 + e^{-x}} \quad (2)$$

其中, $x$ 表示函数的输入, $f(x)$ 表示输出。

$SiLU$ 是一个平滑且连续的函数,如图5所示。由图5可看到,其形状与Sigmoid函数相似,但具有正范围内的线性分量。

实验证明,在MobileNet作为主干网络的情况下,以 $SiLU$ 函数代替 $ReLU$ 函数可以在LFW数据集下取得97.23%的准确率。

图5 激活函数 $SiLU$ Fig. 5 Activation function  $SiLU$

## 2.4 融合 Contextual Transformer 模块

Contextual Transformer 是一种神经网络架构,通常用于自然语言处理(NLP)任务,例如语言翻译、文本分类和问答等。这是 Transformer 架构的一个变体,在 2017 年 Vaswani 等学者的论文“Attention is All You Need”中被引入<sup>[20]</sup>。

传统的 Transformer 架构<sup>[21]</sup>的关键创新是使用自注意力,允许模型一次性捕获输入序列中所有位置对之间的依赖关系,使其非常适合需要理解远程依赖关系的任务(如:机器翻译)。

Contextual Transformer 网络的核心在于通过上下文转换器来提取全局特征,这也是其相异于传统 Transformer 的最大区别。在上下文转换器中,输入序列首先使用嵌入层转换为向量序列;这些向量由一系列自注意层处理,其中每个向量关注序列中的所有其他向量,以计算序列的加权表示;然后将得到的加权向量序列通过前馈网络产生最终输出,这也是 CNN 网络的优点。

Contextual Transformer 模块能有效结合 CNN 和 Transformer 的优点,利用 Transformer 的空间自注意力机制改善 CNN 所固有的平移不变性和局限性,有效提高人脸图像特征提取的有效性和完整性。

实验证明,Contextual Transformer 模块能明显提高人脸识别准确率,在 LFW 数据集下得到 98.16% 的准确率。

## 2.5 实验结果分析

在 MobileNet 同时融合 ECA ( $K = 7$ ) 和  $SiLU$  激活函数与 Contextual Transformer 模块后,人脸识别准确率得到了显著提升,在 LFW 数据集下达到了 98.79%,超过原版 2.74 个百分点,同时在实验室自采集数据集下准确率达到 95.22%,提升了 1.35 个百分点,表明该网络具有较好的泛化能力,能更好地提高模型提取人脸特征能力。测试结果见表 2。

表 2 不同网络结构在 LFW 数据集上测试结果

改进策略	准确率 %
MobileNet+ECA ( $K = 5$ )	97.81
MobileNet+ECA ( $K = 7$ )	98.31
MobileNet+ $SiLU$	97.23
MobileNet+Contextual Transformer	98.16
MobileNet+ $SiLU$ + ECA ( $K = 7$ )+Contextual Transformer	98.79

同样,在 CASIA-WebFace 数据集下,不同的 ECA 注意力机制添加方法对实验结果具有较大偏差。当人脸识别系统对实时性要求较高时,可考虑利用 ECA 模块代替一层 dw 卷积层,虽然降低了 0.73 个百分点,但是可以大大提高推断速度和参数量,在特定场合具有一定的参考意义;在  $stage2$  中间隔 DW 卷积层添加 3 个 ECA 模块为最优方式,并且通过对一维向量大小的实验证明,在  $k = 7$  时,ECA 模块最为适合人脸识别任务。

为适配 ECA 模块,将 DW 卷积层的激活函数由  $ReLU$  改为  $SiLU$ ,同样在准确率方面得到了提升。Contextual Transformer 模块能结合 CNN 和 Transformer,改善 CNN 固有的平移不变性和局限性,充分发掘 Transformer 的自注意力机制,加强全局性、有效性方面特征的挖掘。综合以上的改进,最终在 LFW 测试集上取得了 98.79% 的准确率,提升了 2.74 个百分点,在 CN-Face 数据集上取得了 95.22% 的准确率,提升了 1.35 个百分点。图 6 为  $loss$  曲线。由图 6 可见,最终  $loss$  已趋近于 0。

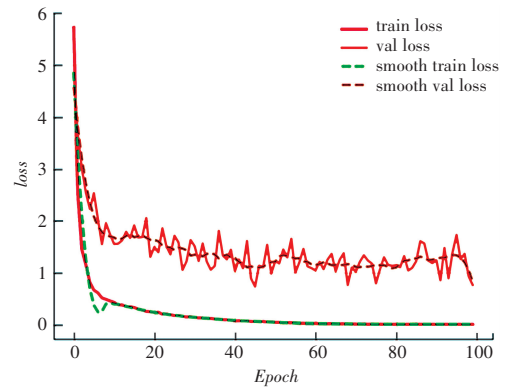


图 6 损失函数  $loss$  曲线

Fig. 6 The curve of function  $loss$

## 3 结束语

本文首次将 ECA 注意力机制模块与 FaceNet 相结合,并将 Contextual Transformer 模块应用到人脸识别任务中。一方面从提高人脸识别判断速率进行实验,另一方面从提高人脸准确率入手,解决人脸识别现实工程中面临的两大问题。分别在 LFW 数据集和自采集的中国人脸数据集上进行测试,实验结果表明均存在明显提升,说明该改进方法具有一定泛化能力。CN-Face 数据集已经在 github 中进行了开源 (<https://github.com/Planktonwithcrab/face-dataset-of-CN>)。

当前开源数据集中,中国人脸乃至亚洲人脸极

其匮乏,现有的亚洲人脸数据集通过爬虫方法获得,大多属于明星人脸,由于妆容特效等原因,不能很好地应用到人脸识别研究中。本实验室将进一步加大中国人脸数据集的收集及清理,为广大学者提供更优质的数据集。

## 参考文献

- [1] BROWNEEE J. Deep learning for natural language processing: Develop deep learning models for your natural language problems [EB/OL]. [2023-08-01]. <https://ozturkibrahim.com>.
- [2] KOESTINGER M, WOHLHART P, ROTH P M, et al. Annotated facial landmarks in the wild: A large-scale, real-world database for facial landmark localization [C]//2011 IEEE International Conference on Computer Vision Workshops. Barcelona, Spain; IEEE, 2011:2144-2151.
- [3] QI Xuan, LIU Chenu, SCHUCKERS S. CNN based key frame extraction for face in video recognition [C]// 2018 IEEE 4<sup>th</sup> International Conference on Identity Security and Behavior Analysis. Australia; IEEE, 2018:1-8.
- [4] SÜTCÜLER E. Real-time face localization and recognition system by using video sequences [D]. Turkey: Yıldız Technical University, 2006.
- [5] GAO Yongbin, LEE H J. Pose-invariant features and personalized correspondence learning for face recognition [J]. Neural Computing and Applications, 2019,31 (1) : 607-616.
- [6] ATIK M, DURAN Z. 3D face recognition and accuracy analysis using local featurebased methods[J]. Journal of the Faculty of Engineering and Architecture of Gazi University, 2020,36 (1) : 359-372.
- [7] YANIV T, YANG Ming, AURELIO R M, et al. DeepFace: Closing the gap to human level performance in face verification [C]// 2014 IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA;IEEE, 2014; 1701-1708.
- [8] SUN Y, WANG X, TANG X. Deep learning face representation from predicting 10,000 classes[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Columbus, USA;IEEE, 2014; 1891-1898.
- [9] FLORIAN S, DMITRY K, JAMES P. FaceNet: A unified embedding for face recognition and clustering[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Boston, USA;IEEE, 2015; 815-823.
- [10] ZHANG Xiao, ZHAO Rui, QIAO Yu, et al. Adacos: Adaptively scaling cosine logits for effectively learning deep face representations [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Long Beach, USA; IEEE, 2019; 10823-10832.
- [11] DENG Jiankang, GUO Jia, YANG Jing, et al. Arcface: Additive angular margin loss for deep face recognition [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Long Beach, USA;IEEE, 2019; 4690-4699.
- [12] LIU Hao, ZHU Xiangyu, LEI Zhen, et al. Adaptiveface: Adaptive margin and sampling for face recognition [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Long Beach, USA; IEEE, 2019; 11947-11956.
- [13] SUN Yifan, CHENG Changmao, ZHANG Yuhan, et al. Circle loss: A unified perspective of pair similarity optimization [C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2020; 6398-6407.
- [14] ZHU Zheng, HUANG Guan, DENG Jiankang, et al. WebFace260M: A benchmark unveiling the power of million-scale deep face recognition [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2021; 10492-10505.
- [15] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [C]//Advances in Neural Information Processing Systems. Long Beach, USA;NIPS Foundation, 2017; 5998-6008.
- [16] MENSAH J A, APPATI J K, BOATENG E K A, et al. FaceNet recognition algorithm subject to multiple constraints: Assessment of the performance [J]. Scientific African, 2024, 23; e02007.
- [17] BÜYÜKTAŞ B, EROĞLU E Ç, ERDEM T. More learning with less labeling for face recognition [J]. Digital Signal Processing, 2023, 136; 103915.
- [18] CHANG H C, YU L W, LIU B Y, et al. Classification of the implant-ridge relationship utilizing the MobileNet architecture [J]. Journal of Dental Sciences, 2024, 19(1): 411-418.
- [19] LIU Lichao, LIANG Jing, WANG Jianqing, et al. An improved YOLOv5-based approach to soybean phenotype information perception [J]. Computers and Electrical Engineering, 2023, 106; 108582.
- [20] KROŽAJ J, PLESH R O, BANAVAR M, et al. Deep face decoder: Towards understanding the embedding space of convolutional networks through visual reconstruction of deep face templates [J]. Engineering Applications of Artificial Intelligence, 2024, 132; 107941.
- [21] AGARWAL M, GUPTA S, BISWAS K K. A new Conv2D model with modified ReLU activation function for identification of disease type and severity in cucumber plant [J]. Sustainable Computing: Informatics and Systems, 2021, 30; 100473.