

# 基于特征融合的轻量级新残差人脸识别方法

惠康华, 闫建青, 高思华\*, 贺怀清

(中国民航大学计算机科学与技术学院, 天津 300300)

**摘要:** 针对现有轻量级模型在嵌入式设备的人脸识别应用中存在识别精度难以提升的问题, 提出一种融合人脸对齐关键特征点信息的轻量级新残差网络模型(Lightweight New Residual Network, LNRN). LNRN利用深度残差网络结构能够解决网络退化且避免干扰因素影响的优势, 结合人脸对齐环节产生的关键特征点信息, 对深度残差网络结构进行简化和合理设计, 实现对关键特征信息和全局信息的提取. 为避免特征提取过程中丢失重要特征信息, 该模型在新残差网络中加入结合空间和通道的注意力机制进行辅助. 在公开的四个标准人脸数据集上的仿真实验表明, 该模型识别速度在接近主流轻量级人脸识别方法的同时, 平均识别精度比MobiFace提高了0.6%.

**关键词:** 轻量级新残差网络模型; 人脸识别; 关键特征信息; 注意力机制

**基金项目:** 国家重点研发计划项目(No.2020YFB1600101); 天津市教委科研项目(No.2020KJ024)

**中图分类号:** TP391.41

**文献标识码:** A

**文章编号:** 0372-2112(2024)03-0937-08

**电子学报 URL:** <http://www.ejournal.org.cn>

**DOI:** 10.12263/DZXB.20221024

## Lightweight New Residual Face Recognition Method Based on Feature Fusion

HUI Kang-hua, YAN Jian-qing, GAO Si-hua\*, HE Huai-qing

(College of Computer Science and Technology, Civil Aviation University of China, Tianjin 300300, China)

**Abstract:** Aiming at the problem that the existing lightweight models are difficult to improve the recognition accuracy in the face recognition applications of embedded devices, a new lightweight residual network model (Lightweight New Residual Network, LNRN) that integrated the key feature point information of face alignment is proposed. The advantage of deep residual network structure that can solve the network degradation and avoid the influence of interference factors are absorbed by LNRN. In order to realize the extraction of key feature information and global information after combining the key point information generated by the face alignment, the deep residual network structure is simplified and reasonably designed. In order to avoid losing important feature information in the process of feature extraction, an attention mechanism combining space and channel is added to the new residual network for assistance. Simulation experiments on the four standard face datasets showed that the recognition speed of the proposed model was close to the mainstream lightweight face methods, and the average recognition accuracy of the proposed model is 0.6% higher than that of MobiFace.

**Key words:** lightweight new residual network model; face recognition; key feature information; attention mechanism

**Foundation Item(s):** National Key Research and Development Project of China (No.2020YFB1600101); Scientific Research Project of Tianjin Education Commission (No.2020KJ024)

### 1 引言

人脸识别作为一种生物特征识别技术, 可快速准确地判别出视频或图像中人脸的身份, 具有非侵扰、安全快捷等特点, 被广泛应用于金融、考勤门禁和安防等领域<sup>[1]</sup>. 传统的人脸识别算法, 大部分都是基于降维思想或人工设计的特征, 例如: Sharma 等人<sup>[2]</sup>利用主成分

分析法(Principal Component Analysis, PCA)获得人脸图像的特征表示, 有效地提升了计算效率, 但在受到光照、遮挡等因素干扰时, 难以得到人脸数据的真实子空间结构, 识别性能下降. Tao 等人<sup>[3]</sup>提出了一种使用内核组合的局部敏感支持向量机(Support Vector Machine, SVM)算法, 对单因素造成的人脸变化具有较强的适应性, 但不适用于大规模数据集训练场景. 以上这

些浅层算法的局限性在于,当数据集规模和特征学习单元有限时,无法拟合出有效的复杂函数。

随着深度学习技术的发展<sup>[4]</sup>和大规模人脸数据集的公开,深度学习在人脸识别领域的研究主要集中在网络结构<sup>[5]</sup>和损失函数的设计<sup>[6]</sup>,以达到进一步提高人脸识别准确度的目的。Chan 等人<sup>[7]</sup>提出的 PCANet 将浅层算法和深度学习相结合,对存在光照和遮挡影响的人脸进行识别时具有很强的鲁棒性<sup>[8]</sup>。Schroff 等人<sup>[9]</sup>提出的 FaceNet 和 Parkhi 等人<sup>[10]</sup>提出的 VGGNet 均是利用 Triplet 损失函数在欧式空间上学习图像的特征表示,在公开人脸数据集上识别效果均很好。后续研究者们分别提出的 SphereFace<sup>[11]</sup>、CosFace<sup>[12]</sup>、ArcFace<sup>[13]</sup> 三种损失函数在应用于 ResNet<sup>[14]</sup> 的度量学习时,在公开人脸数据集上的识别性能都有相应的提升。此外,汪等人<sup>[15]</sup>提出一种基于迹变换和旋转增量调制编码的特征提取方法,提高了在不同特殊场景下的模糊人脸识别的准确性和鲁棒性。Kim 等人<sup>[16]</sup>提出了一种新的损失函数 AdaFace,即错误分类样本的策略应根据其图像质量进行调整,提高了低质量人脸数据集中人脸识别性能。Wang 等人<sup>[17]</sup>提出一种更快的人脸分类训练方法,采用动态类池对身份特征进行动态存储和更新,有效减少了训练的时间和成本。上述基于深度学习的人脸识别算法从大量的训练数据集中不断提高人脸特征的判别性,在受到年龄、姿态等因素影响的人脸数据集上具有很好的识别效果<sup>[18]</sup>。但这些方法无论在数据训练实验还是实际应用中都需要强大的硬件资源和大规模的数据集作支撑。在硬件资源不充分的实际场景中,识别往往采用便携的小型设备,其算力难以满足上述人脸识别算法对计算资源的要求。因此,为了适应实际情况所需,一批可以部署到移动或嵌入式设备中的轻量级网络开始逐渐得到应用,如 SENet<sup>[19]</sup>、MobileNetV2<sup>[20]</sup>、ShuffleNet<sup>[21]</sup> 等。尽管借助上述文献中的 SE (Squeeze-and-Excitation Networks) 和 CBAM (Convolutional Block Attention Module, CBAM) 等注意力机制<sup>[22]</sup> 能够使轻量级网络快速准确地完成识别任务,但传统轻量的网络设计导致其提取的特征信息不够丰富,无法接近或达到复杂网络的高识别精度需要。

针对上述问题,本文提出一种融合人脸对齐关键特征点信息的轻量级新残差网络模型 LNRN。相比传统人脸识别所采用的轻量级模型方法,在满足实时性的基础上,提高其在小型移动设备中的识别性能。首先,LNRN 借鉴深度残差网络的恒等映射策略,将新残差网络的模块设计为两个分支,分别用于提取图像关键信息和全局信息,使学习到的特征更具有判别性。其次,获取人脸关键点的位置信息,并将其融入对应分支完成对关键特征信息的提取,增强特征信息的丰富性。再

次,在注意力机制辅助下,将特征图在通道和空间两个维度进行信息交互,增强重要信息的权重。最后,利用新残差网络模块提取的特征信息,训练出轻量级的分类器实现对面脸身份的判别。

## 2 LNRN 模型

为解决轻量级模型在小型移动设备中识别性能难以提升的问题,本文提出了 LNRN 模型。该模型的设计思路是在残差网络模块的基础上,通过引入人脸对齐的关键特征点位置快速提取出对面脸身份判别有效的关键特征,之后使用注意力机制持续学习获取到的人脸关键特征,从而使得模型学到的人脸特征表示更具判别力。不同于以往的轻量级模型,LNRN 模型为新残差结构引入人脸对齐的关键特征点信息,从而提取更为重要的特征。与此同时,使用注意力机制辅助重要特征的提取。以下将从算法流程、人脸对齐关键特征和轻量级新残差网络三个方面详细介绍 LNRN 模型。

### 2.1 LNRN 算法

人脸识别技术在实际应用中往往需要同时满足实时性和高精度两个要求,以 ResNet 为代表的复杂网络可以实现高精度的需求,但复杂网络本身的训练和使用都需要强大的计算资源开销和充分的训练时间。因此大型的重量级网络无法直接应用到硬件资源受限的实际场景中,但可以通过学习重量级网络框架中的优势,以此获得具有其优点的轻量级网络。ResNet 网络对非限制场景下的人脸数据集表现出了很强的鲁棒性,而且其自身的残差结构可以解决网络退化现象<sup>[11]</sup>。本文通过保留 ResNet 网络的优势,然后结合人脸对齐获得的关键特征点信息,设计出一种能提取具有判别性特征的轻量级新残差网络。LNRN 模型可以在满足实时性的前提下,进一步提高识别精度。具体算法流程如下。

**步骤 1 数据预处理:**归一化处理 and 标准化处理减少无效的信息。

**步骤 2 人脸数据处理:**对预处理完的数据进行人脸对齐,裁剪为指定大小并保留人脸对齐得到的关键点位置。

**步骤 3 模型训练:**将处理好的人脸数据送入 LNRN 模型,使用 ArcFace 损失监督模型分类器的训练。

**步骤 4 模型测试:**将训练好的分类器在测试集上进行相关测试,验证模型的泛化性能。

步骤 1 是图像数据普遍进行的标准化操作,可降低后续处理的计算复杂度,并减少冗余的无效信息。步骤 2 中的人脸对齐关键特征提取是人脸识别环节的关键,在 2.2 小节会阐述其在本模型中起到的作用。步骤 3 的重点在于保持模型轻量化,并联合人脸对齐的关键特征点信息提高识别效果,这一部分会在 2.3 小节进行相

关介绍. 步骤4模型的测试验证将在第3节详细展开.

### 2.2 人脸对齐关键特征

人脸识别分为检测、对齐和识别三个环节,检测是识别的基础,对齐是识别的关键. 人脸对齐<sup>[23]</sup>可在检测好的人脸区域内定位出面部关键特征点,对齐的效率和准确性直接影响识别算法的性能.

针对识别环节而言,现有的大多深度学习人脸识别算法,都是把人脸对齐后的图像作为直接输入,本质上送入识别算法的都是普通的特征,并没有充分利用人脸对齐定位出的关键特征点信息. 如果能在送入人脸图像的同时,将人脸对齐中的关键特征点信息也同时引入进来,不但能使得识别算法更快地聚焦于判别人脸身份的关键特征,而且可以减少输入普通/无用特征的干扰. 在2.3小节新残差模块中,通过引入人脸对齐中的关键点信息,可以在特征提取的过程中获取更多对判别人脸身份有效的关键特征,从而增强网络的学习能力.

由于人脸对齐的关键特征点信息无法直接融入到识别算法中,本文采取一种间接方式. 首先通过K-Means聚类算法<sup>[24]</sup>计算出关键点位置边界,之后将关键点位置边界送入LNRN模型当中. 关键点坐标信息的获取步骤分为四步:第一步,进行数据预处理;第二步,对训练集进行人脸对齐,保存每一张人脸的关键点坐标信息;第三步,将保存好的关键点坐标信息和簇的个数送入聚类算法中进行聚类;第四步,根据聚类算法得到的簇中心坐标和所有的关键点坐标,确定关键点的坐标边界.

### 2.3 轻量级新残差网络

ResNet网络对光照、年龄、姿态和遮挡等因素影响下的人脸识别具有良好的鲁棒性,并且还有避免过拟合等优点<sup>[12,13]</sup>. 为获得更高的精度,本文在吸收ResNet网络残差结构优点的基础上,联合人脸对齐标出的关键特征点信息,设计一种能提取更具判别性特征的新残差模块. 轻量级新残差网络如图1所示.

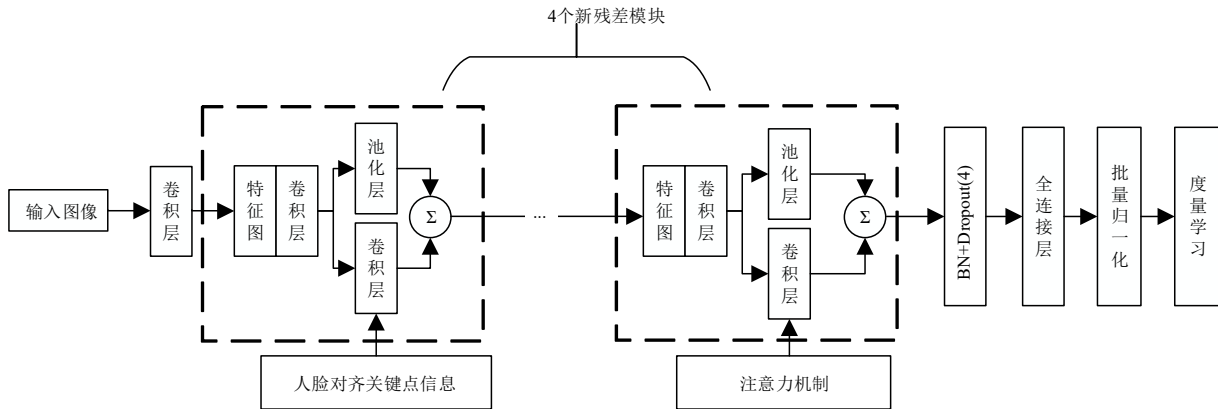


图1 轻量级新残差网络基本模型

#### 2.3.1 新残差模块

为了在提高识别精度的前提下满足实时性,本文的轻量级新残差网络选取结构简洁的四个新残差模块用来完成特征的提取. 图1所示的四个新残差模块在分支处都有共享特征图的操作,可以减少特征图的重复计算从而降低时间开销. 其中带池化层的分支用来提取全局信息,保证上下文语义丰富性的同时可以减少特征提取网络的参数量;另一个分支负责对人脸对齐定位出的位置完成关键判别性特征信息的提取.

新残差模块分为两种,分别如图2(a)和图2(b)所示. 图1中第一个新残差模块采用的是图2(a)所示的第一种新残差模块,图1中剩余三个模块是图2(b)所示的第二种新残差模块. 第一种新残差模块通过人脸对齐定位的关键点信息聚焦关键特征所处的位置,一方面可从分支处共享的特征图减少大量的非关键位置特征,从而降低整个网络的计算复杂度;另一方面可以使得关键特征区域的权重系数变大,在后续特征提取时

提取到更多对判别人脸起决定性作用的重要特征. LNRN可在特征提取网络的开始部分就对特征的重要性做出分类,从而有效减少后续参与计算的特征数量,实现对重要特征的提取和时间开销的降低. 第二种新残差模块通过引入注意力机制来增强特征学习,提高特征提取网络的表达能力,关于注意力机制在模块中起到的作用分析详见2.3.2. 两种新残差模块都吸收了深度残差网络的优点,既可以解决网络退化的问题,又能降低外界因素对人脸识别的影响.

#### 2.3.2 CBAM注意力机制

轻量级新残差网络在联合人脸对齐的关键点信息后,在特征提取网络的刚开始不仅可以对特征的重要性做出分类,而且能有效减少后续参与计算的特征数量,从而实现对重要特征的提取和降低时间开销. 联合人脸对齐的目的在于关注重要特征的位置,但次要特征并不能直接丢弃. 重要特征的提取可作为判别人脸身份的决定性特征,而次要特征可作为辅助信息防止

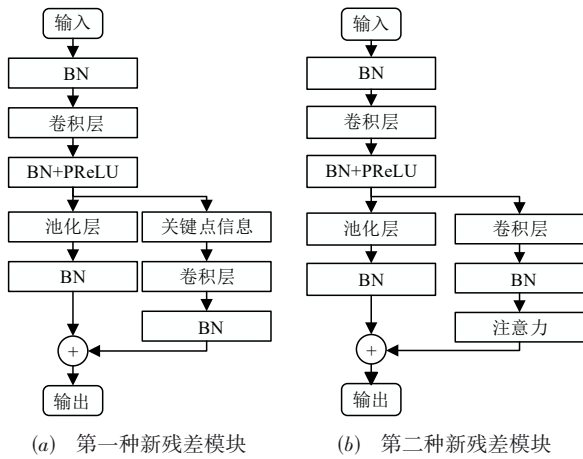


图2 两种新残差模块

识别出现错误. 两种特征信息相互补充, 才能达到较好的识别效果.

本文模型的主体是一种轻量级网络, 在层数较少的深度学习中难以充分提取数据中隐藏的特征信息. 因此, 本文使用CBAM注意力机制作为辅助, 在空间维度和通道维度上进行信息交互, 进一步提高轻量级网络的表示能力, 并且注意力机制的加入只会增加很小的时间开销<sup>[22]</sup>. 从人脸对齐关键特征点位置信息间接提取到的人脸关键特征与全局特征进行融合, 送入由第二种新残差模块构成的网络中进行后续特征学习. 而第二种新残差模块中引入的注意力机制, 是用来保持对融合特征当中人脸关键特征的关注度, 进而实现人脸关键特征的持续表示学习, 从而有效提高人脸特征的判别性.

### 2.3.3 ArcFace 损失函数

在人脸识别领域的研究中, 前期主要是网络结构研究占据主导地位, 但网络结构的改进难以持续提升识别性能. 后期为了解决该问题, 损失函数的研究取得了很大进展, 在相关研究中占据了主导地位. 损失函数是在分类正确的基础上, 通过尽可能最大化类间距离的同时最小化类内距离, 提高特征的判别性. 因此, 损失函数可以帮助卷积神经网络在训练时以度量学习的方式获得泛化性更强的分类器. 目前, ArcFace 损失是人脸识别算法中普遍使用的一种损失, 图1中的度量学习即是使用该损失监督模型的训练, 可以将原始特征分布空间中判别困难的样本进行维度转换, 降低外界干扰造成的影响, 从而提升识别性能. LNRN 模型的 ArcFace 损失如式(1)所示.

$$L_{\text{ArcFace}} = -\frac{1}{N} \sum_{i=1}^N \ln \frac{e^{s(\cos(\theta_i + m))}}{e^{s(\cos(\theta_i + m))} + \sum_{j=1, j \neq y_i}^n e^{s(\cos \theta_j)}} \quad (1)$$

其中, 与样本相关的权重向量和特征向量皆进行归一

化,  $y_i$  代表第  $i$  个样本的深度特征属于第  $y_i$  个类别,  $s$  为缩放因子,  $m$  为加性角间隔,  $\theta$  为样本特征和权重之间的角度.

## 3 实验结果与评估

本节涉及的算法均在 CASIA-WebFace<sup>[25]</sup> 上先训练至收敛, 然后分别在野外标记人脸数据集<sup>[26]</sup> (Labeled Faces in the Wild, LFW)、数据集 VGG-Face<sup>[10]</sup>、年龄数据集<sup>[27]</sup> (Age DataBase, AgeDB) 和名人的正脸侧脸数据集<sup>[28]</sup> (Celebrities in Frontal Profile -Frontal-Profile, CFP-FP) 上进行实验, 以验证 LNRN 模型的有效性. 首先, 通过聚类方法确定人脸对齐关键特征点位置范围; 其次, 通过关键点信息引入前后的实验结果验证其作用; 再次, 通过注意力机制添加前后的实验结果验证其作用; 最后, 将 LNRN 模型与近年来的主流人脸识别算法进行比较.

### 3.1 实验数据集

CASIA-WebFace 作为实验的训练集, 有 10 575 类, 总计 494 414 张图像, 其中人脸图像存在面部遮挡、姿态夸张、光照暗和拍摄角度欠佳等干扰因素. 实验的测试数据集分别是 LFW、VGG-Face、AgeDB 和 CFP-FP 四个标准数据集. LFW 是在自然生活场景下采集的基准数据集, 其图像受姿态变化、年龄变化和面部遮挡等因素影响, 共有 5 749 人的 13 233 张人脸图像. VGG-Face 是一个大规模人脸数据库, 主要具有姿态变化、年龄变化及种族等特点, 且每类的样本数量充足, 实验选取 500 人的共计 160 483 张图像. AgeDB 数据集的主要特点是年龄变化差异大, 包含 440 人的 12 240 张人脸图像. 实验使用其子数据集 AgeDB-30, 由 300 个正样本对和 300 个负样本对组成. CFP-FP 是一个姿态变化影响占比较大的数据集, 有 500 人的图像信息, 其中每人有 10 张正脸和 4 张侧脸图像.

首先, 对训练集和测试集进行相同的图像预处理操作; 其次, 对数据集中人脸图像进行 MTCNN (Multi-task Cascaded Convolutional Networks)<sup>[29]</sup> 网络检测并标识出面部的五个关键点位置: 眼睛 (两个眼球中心)、嘴巴 (两个嘴角) 和鼻子 (一个鼻尖); 再次, 对图像完成相似度转换对齐后输出  $112 \times 112$  大小的 RGB 彩色图像; 最后, 在送入模型前对 RGB 彩色图像的像素值做归一化操作, 像素值先减去 127.5, 再除以 128.

### 3.2 实验环境及模型训练

本文实验的硬件配置如下: 操作系统是 Ubuntu 18.04.5, 处理器为 Intel Xeon (R) Silver 4242, 内存是 32 GB, 显卡为 Quadro RTX 5000. 在此硬件基础上, 通过在 JetBrains PyCharm Community Edition 2020.2.1 软件上配置 python 3.8.3 和 PyTorch 1.6.0 深度学习框架完

成本文模型的训练和测试。

从 CASIA-WebFace 中选取 453 580 张数据输入 LNRN 模型,选取方法和文献[13]相同,而后进行数据增强,最后在 ArcFace 损失的监督下完成分类器训练.模型训练中涉及的参数根据经验设置,如损失中的缩放因子  $s$  为 32,间隔  $m$  为 0.5.训练使用随机梯度下降进行参数的更新,模型的初始学习率为 0.1,权重衰减系数为 0.000 5,动量为 0.9.训练共 18 轮,每轮的批量大小为 128,同时,在第 6、11、16 轮依次将学习率按步长为 0.1 衰减,LNRN 模型迭代训练 63 700 次结束.

### 3.3 实验过程和结果评估

为确保实验的结果具有可比性,本文涉及的训练集和测试集均采用相同的预处理方式.参与对比的算法均使用相同的训练集完成训练,并采用十折交叉验证对测试集进行测试.本部分所涉及的 LNRN-1 网络主体结构在图 2 两种新残差模块的基础上,保留注意力机制,舍弃引入人脸对齐的关键点信息.LNRN-2 网络主体结构在图 2 两种新残差模块的基础上,保留引入人脸对齐关键点信息,舍弃注意力机制.

#### 3.3.1 人脸对齐关键点引入前后的结果对比

人脸对齐关键特征结果如图 3 所示,在检测到的人脸区域可标识出面部关键特征点位置,有利于后续识别.为了能够将人脸对齐的关键点信息融入到识别算法中,先通过 K-means 聚类算法对关键点的位置信息进行训练,然后找到五个关键点对应的五个簇中心,得到关键点位置的界限,最后再将其加入到识别算法中完成模型的训练.聚类出的五个簇中心和所有定位的关键点坐标可视化结果如图 4 所示.散落在图 4 四周的异常点是侧脸、表情等因素影响下定位的关键点.从图 4 的分布趋势显示,采用聚类获得的簇中心表示所有关键点是可行的.通过五个簇中心坐标确定的关键点位置范围是 [39:73, 51:88],但这个边界并没有包含眼睛上方的眉毛等信息.为符合模型的计算方便以及能提取到眉毛等人脸关键信息,本文最后选择的边界是 [28:84, 34:90].



图 3 人脸对齐的关键点定位

表 1 所示的 LNRN-1 方法没有引入人脸对齐关键点信息,LNRN(CBAM)方法中引入了人脸对齐关键点信息.从实验结果可知,LNRN(CBAM)的结果相较除

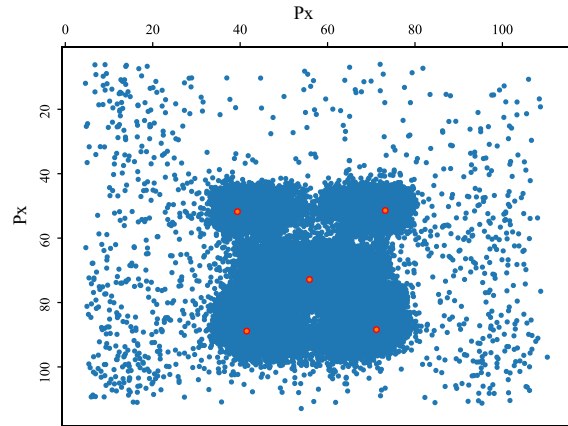


图 4 人脸关键点坐标聚类可视化

CFP-FP 以外其他三个数据集均有所提升,同时其识别速度降低 1 ms.其原因在于引入人脸对齐关键点信息的 LNRN(CBAM),在特征提取阶段能够指导提取出更多对判别人脸起决定作用的关键正脸区域特征,相同的层数下获取更为丰富的关键特征,有利于正脸姿态识别精度的提高.与此同时,关键点信息在第一个模块中以坐标边界的形式引入,在新残差模块的分支处直接缩小了共享特征图的大小,从而减少了参与计算的特征个数,使得识别速度有所提升.

表 1 关键点信息引入前后的识别精度对比

方法	识别精度/%				速度/ms
	LFW	VGG-Face	AgeDB	CFP-FP	
LNRN-1	99.12	95.62	92.16	<b>93.89</b>	18
LNRN(CBAM)	<b>99.12</b>	<b>95.95</b>	<b>93.30</b>	93.41	17

LNRN(CBAM)的结果在 CFP-FP 数据集中略低于 LNRN-1 的原因在于,引入人脸对齐关键点信息后,正脸与侧脸的特征区域重叠部分较少.在正脸的特征区域权重系数变大的同时,侧脸特征区域权重系数则会相应地变小.CFP-FP 数据集中的侧脸所占的比例为 28%,是四个数据集中占比最高的,又因为模型中侧脸特征区域权重系数变小,导致提取出来的特征对侧脸姿态的识别效果略差.

#### 3.3.2 注意力机制引入前后的结果对比

为验证注意力机制引入可有效提高网络的表示能力,本文分别在四个基准数据集上做了对比试验.如表 2 所示,引入关键点信息的 LNRN-2 方法的新残差模块中,分别加入 SE、CAM、SAM 和 CBAM 四种注意力机制<sup>[19,22]</sup>来验证其作用.SE 是一种通道注意力机制,而 CBAM 是由 CAM 和 SAM 两种注意力机制组成.轻量级新残差网络分别加入 SE 和 CAM 两种注意力模块的实验结果显示,加入 CAM 后在除 AgeDB 以外的三个数据

集上,识别精度均比加入SE提升的更多. SE和CAM都是一种通道注意力且结构相似,区别在于CAM比SE多了一个并行的池化层,可以多一条提取信息的分支,从而获得的信息比SE更全更丰富. 四种注意力的加入方式绝大多数都在LNRN-2模型的基础上提升了识别的精度,是因为轻量级网络在注意力机制的辅助下,可以提取更多的特征信息,进而提升网络的表达能力. 加入CBAM效果最好的原因在于,CBAM可以从通道和空间两个维度上对来自人脸关键点位置的关键特征图进行自适应的特征优化,保留更多关键的特征信息进行特征学习.

表2 注意力机制引入前后的识别精度对比

方法	识别精度/%				速度/ms
	LFW	VGG-Face	AgeDB	CFP-FP	
LNRN-2	98.98	95.86	92.65	92.54	16
LNRN(SE)	99.02	95.78	92.65	92.68	16
LNRN(CAM)	99.08	95.86	92.60	93.21	16
LNRN(SAM)	99.03	95.70	92.62	92.74	16
LNRN(CBAM)	<b>99.12</b>	<b>95.95</b>	<b>93.30</b>	<b>93.41</b>	17

### 3.3.3 LNRN与主流算法的结果对比

为进一步证明LNRN算法可以提高轻量级网络的识别准确率,本文将其与近年主流的ResNet、MobiFace<sup>[30]</sup>、GhostNet<sup>[31]</sup>、HRNet<sup>[32]</sup>等多个识别算法进行了比较. 本部分的LNRN均采用实验效果较好的CBAM注意力机制作为辅助. 实验结果如表3所示.

表3 不同算法的识别精度对比

方法	识别精度/%				速度/ms
	LFW	VGG-Face	AgeDB	CFP-FP	
ResNet	<b>99.46</b>	<b>95.95</b>	<b>94.15</b>	<b>94.78</b>	23
MobiFace	98.60	95.70	92.32	92.83	15
HRNet	99.40	95.98	93.10	93.65	20
GhostNet	99.17	95.81	91.97	92.67	18
LNRN(CBAM)	99.12	95.95	93.30	93.41	17

ResNet是作为一种大型复杂网络,在四个数据集上的识别性能均高于表3中其他算法,但其识别速度最慢. 本文采用的ResNet为50层,50层的残差网络可以在测试数据集上表现出很好的性能和抗干扰性,但受限于网络的深度导致其识别速度最慢.

LNRN算法在四个不同数据集上的平均识别准确率均比MobiFace高0.6%,单张识别速度比MobiFace慢2 ms. LNRN算法在4个测试集上识别效果与GhostNet相比,一个基本持平,三个有所提升,且单张识别速度比GhostNet快1 ms. LNRN能够提高识别精度,一是吸

收了残差网络可以有效减低外界干扰的特点,二是归功于人脸对齐关键点信息的引入和注意力机制的辅助,通过这两种方式可以使得提取的特征信息更丰富且更具判别性. 与此同时,LNRN的这两种操作也在一定程度上增加了时间开销. MobiFace和GhostNet算法作为轻量级的网络框架,轻量的设计以及未充分利用人脸对齐关键点信息的劣势,使得其无法在尽可能快的速度下获得更高的识别精度. HRNet算法在四个数据集上的识别效果均较好,但识别速度比LNRN慢17.6%. 造成这种现象的原因在于,HRNet网络提取的特征信息更全面,需要通过多分支的多次并行卷积执行不同尺度的特征融合实现的,从而带来更多的时间开销. 从以上分析来看,LNRN算法的识别速度居于第二位,但综合速度和识别精度两方面来考虑,LNRN是一个可以在满足实时性的基础上提高识别精度的轻量级新残差网络.

## 4 结束语

本文在ResNet网络的基础上提出了一种结合人脸对齐关键特征点信息的轻量级新残差网络,用于解决轻量级网络在小型设备中识别精度难以提高的问题. 多组实验结果表明,和MobiFace等主流的轻量级网络相比,LNRN模型能在满足实时性的前提下进一步提高识别性能,适合在小型便携设备中移植使用. 本文将人脸对齐关键特征点信息融入识别环节,使得算法具有更好的泛化能力,可为轻量级网络在小型设备上的使用提供重要的参考.

## 参考文献

- [1] TAKALKAR M, XU M, WU Q, et al. A survey: Facial micro-expression recognition[J]. Multimedia Tools and Applications, 2018, 77(15): 19301-19325.
- [2] SHARMA R, PATTERH M S. A new hybrid approach using PCA for pose invariant face recognition[J]. Wireless Personal Communications, 2015, 85(3): 1561-1571.
- [3] TAO Q Q, ZHAN S, LI X H, et al. Robust face detection using local CNN and SVM based on kernel combination [J]. Neurocomputing, 2016, 211: 98-105.
- [4] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521: 436-444.
- [5] 李倩玉, 蒋建国, 齐美彬. 基于改进深层网络的人脸识别算法[J]. 电子学报, 2017, 45(3): 619-625.  
LI Q Y, JIANG J G, QI M B. Face recognition algorithm based on improved deep networks[J]. Acta Electronica Sinica, 2017, 45(3): 619-625. (in Chinese)
- [6] WANG M, DENG W H. Deep face recognition: A survey

- [J]. *Neurocomputing*, 2021, 429: 215-244.
- [7] CHAN T H, JIA K, GAO S H, et al. PCANet: A simple deep learning baseline for image classification? [J]. *IEEE Transactions on Image Processing: a Publication of the IEEE Signal Processing Society*, 2015, 24(12): 5017-5032.
- [8] 李小薪, 梁荣华. 有遮挡人脸识别综述: 从子空间回归到深度学习[J]. *计算机学报*, 2018, 41(1): 177-207.
- LI X X, LIANG R H. A review for face recognition with occlusion: From subspace regression to deep learning[J]. *Chinese Journal of Computers*, 2018, 41(1): 177-207. (in Chinese)
- [9] SCHROFF F, KALENICHENKO D, PHILBIN J. FaceNet: A unified embedding for face recognition and clustering[C]//2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2015: 815-823.
- [10] PARKHI O M, VEDALDI A, ZISSERMAN A. Deep face recognition[ED/OL]. (2015-09-10) [2022-08-05]. [https://www.robots.ox.ac.uk/~vgg/publications/2015/Parkhi15/parkhi15.pdf?utm\\_source=top.caibaojian.com/31386&from=weibo.com/kujian](https://www.robots.ox.ac.uk/~vgg/publications/2015/Parkhi15/parkhi15.pdf?utm_source=top.caibaojian.com/31386&from=weibo.com/kujian).
- [11] LIU W Y, WEN Y D, YU Z D, et al. SphereFace: Deep hypersphere embedding for face recognition[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2017: 212-220.
- [12] WANG H, WANG Y T, ZHOU Z, et al. CosFace: Large margin cosine loss for deep face recognition[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 5265-5274.
- [13] DENG J K, GUO J, XUE N N, et al. ArcFace: Additive angular margin loss for deep face recognition[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2019: 4690-4699.
- [14] HE K M, ZHANG X Y, REN S Q, et al. Deep residual learning for image recognition[C]//2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2016: 770-778.
- [15] 汪宇玲, 陈立, 黎明, 等. 基于迹变换和旋转增量调制特征的模糊人脸识别[J]. *电子学报*, 2021, 49(12): 2437-2448.
- WANG Y L, CHEN L, LI M, et al. Rotational delta modulation feature and its application in blurry face recognition based on trace transform[J]. *Acta Electronica Sinica*, 2021, 49(12): 2437-2448. (in Chinese)
- [16] KIM M, JAIN A K, LIU X M. AdaFace: Quality adaptive margin for face recognition[C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2022: 18750-18759.
- [17] WANG K, WANG S, ZHANG P P, et al. An efficient training approach for very large scale face recognition [C]//2022 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2022: 4083-4092.
- [18] 吴长虹, 苏剑波, 陈叶飞. 抗年龄干扰的人脸识别[J]. *电子学报*, 2018, 46(7): 1593-1600.
- WU C H, SU J B, CHEN Y F. Age invariant face recognition[J]. *Acta Electronica Sinica*, 2018, 46(7): 1593-1600. (in Chinese)
- [19] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 7132-7141.
- [20] SANDLER M, HOWARD A, ZHU M L, et al. MobileNetV2: Inverted residuals and linear bottlenecks[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 4510-4520.
- [21] ZHANG X Y, ZHOU X Y, LIN M X, et al. ShuffleNet: An extremely efficient convolutional neural network for mobile devices[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway: IEEE, 2018: 6848-6856.
- [22] WOO S, PARK J, LEE J Y, et al. CBAM: Convolutional block attention module[C]//Computer Vision—ECCV 2018. Cham: Springer International Publishing, 2018: 3-19.
- [23] 万俊, 李晶, 常军, 等. 基于局部形状组合模型的人脸对齐[J]. *计算机学报*, 2018, 41(9): 2162-2174.
- WAN J, LI J, CHANG J, et al. Face alignment on local-shape-based combined model[J]. *Chinese Journal of Computers*, 2018, 41(9): 2162-2174. (in Chinese)
- [24] 李骛, 王卓, 于晓洋, 等. 多核低冗余表示学习的稳健多视图子空间聚类方法[J]. *通信学报*, 2021, 42(11): 193-204.
- LI A, WANG Z, YU X Y, et al. Robust multiview subspace clustering method based on multi-kernel low-redundancy representation learning[J]. *Journal on Communications*, 2021, 42(11): 193-204. (in Chinese)
- [25] YI D, LEI Z, LIAO S C, et al. Learning face representation from Scratch[EB/OL]. (2014-11-28) [2022-08-05]. <https://arxiv.org/pdf/1411.7923.pdf>.
- [26] HUANG G B, MATTAR M, BERG T, et al. Labeled faces in the wild: A database for Studying face recognition in unconstrained environments[EB/OL]. (2008-07-

01) [2022-08-05]. <https://vis-www.cs.umass.edu/lfw/lfw.pdf>.

- [27] MOSCHOGLOU S, PAPAIOANNOU A, SAGONAS C, et al. AgeDB: The first manually collected, In-the-wild age database[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW). Piscataway: IEEE, 2017: 1997-2005.
- [28] SENGUPTA S, CHEN J C, CASTILLO C, et al. Frontal to profile face verification in the wild[C]//2016 IEEE Winter Conference on Applications of Computer Vision (WACV). Piscataway: IEEE, 2016: 1-9.
- [29] ZHANG K P, ZHANG Z P, LI Z F, et al. Joint face detection and alignment using multitask cascaded convolutional networks[J]. IEEE Signal Processing Letters, 2016, 23 (10): 1499-1503.
- [30] DUONG C N, QUACH K G, JALATA I, et al. MobiFace: A lightweight deep learning face recognition on mobile devices[C]//2019 IEEE 10th International Conference on Biometrics Theory, Applications and Systems (BTAS). New York: ACM, 2019: 1-6.
- [31] HAN K, WANG Y H, TIAN Q, et al. GhostNet: More features from cheap operations[C]//2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2020: 1580-1589.
- [32] SUN K, XIAO B, LIU D, et al. Deep high-resolution representation learning for human pose estimation[C]//2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2019: 5693-5703.



高思华 男, 1987年1月出生于承德市. 现为中国民航大学计算机科学与技术学院讲师. 主要研究方向为强化学习理论及应用.

E-mail: shgao@cauc.edu.cn



贺怀清 女, 1969年11月出生于吉林白山市. 现为中国民航大学计算机科学与技术学院教授. 主要研究方向为图形图像与可视分析.

E-mail: hqhe@cauc.edu.cn

#### 作者简介



惠康华 男, 1982年9月出生于江苏连云港市. 现为中国民航大学计算机科学与技术学院副教授. 主要研究方向为计算机视觉.

E-mail: khhui@cauc.edu.cn



闫建青 男, 1998年9月出生于山西忻州市. 现为中国民航大学计算机科学与技术学院硕士研究生. 主要研究方向为图像处理.

E-mail: 2019051017@cauc.edu.cn