

DOI:10.13364/j.issn.1672-6510.20210165

基于细粒度特征的人脸属性识别

韩 菲, 周卫斌, 杨永刚, 王 阳

(天津科技大学电子信息与自动化学院, 天津 300222)

摘要: 人脸属性识别通常受光照、亮度、头部姿态等客观因素的影响而导致人脸属性识别的准确度降低. 为了提高复杂场景下人脸属性识别的准确度, 本文以卷积神经网络(convolutional neural network, CNN)技术为基础提出了一种基于细粒度特征的人脸属性识别网络模型(multiple granularity face attribute net, MGFA Net). 在 MTFL 数据集上, 通过对人脸图像提取细粒度特征, 增加总体特征对各个人脸属性任务的表达能力, 4 种人脸属性的平均识别准确度达到 88%, 与 ResNet18 网络模型相比, 有效提高了 2.79%, 并通过对比实验证明了细粒度特征网络结构的有效性.

关键词: 人脸属性识别; 细粒度特征; 卷积神经网络

中图分类号: TP389.1 **文献标志码:** A **文章编号:** 1672-6510(2022)03-0073-08

Face Attribute Recognition Based on Multiple Granularity Features

HAN Fei, ZHOU Weibin, YANG Yonggang, WANG Yang

(College of Electronic Information and Automation, Tianjin University of Science & Technology,
Tianjin 300222, China)

Abstract: Face attribute recognition is usually affected by such objective factors as illumination, brightness and head posture, which reduces the accuracy of face attribute recognition. In order to improve the accuracy of face attribute recognition in complex scenes, using convolutional neural network (CNN) technology, in this article we propose a face attribute recognition network model based on multiple granularity face attribute net (MGFA Net). On the MTFL datasets, by extracting multiple granularity features from face images, the overall feature's ability to express each face attribute task was increased. The average recognition accuracy of the four types of face attributes reached 88%, which was significantly improved by 2.79% compared with the ResNet18, and the effectiveness of the multiple granularity feature network structure was proved through comparative experiments.

Key words: face attribute recognition; multiple granularity features; convolutional neural network

人脸属性识别对人脸识别、身份识别、改善人机交互体验、提供定制化服务等相关人工智能技术或社会服务具有重要的辅助作用. 随着人工智能技术的不断发展, 基于卷积神经网络的人脸属性识别技术已经取得了较好的识别效果, 但在实际场景中, 人脸图像数据的有效信息易受光照、亮度、头部姿态等外部客观因素的影响而导致人脸属性识别的准确度低或可靠性不佳. 因此, 如何提高复杂场景下人脸属性识别的准确度逐步成为这一课题的重要研究方向, 同时对推动人脸属性识别技术的生产应用具有重要意义^[1]. 目前, 人脸属性识别的研究通常集中在正脸和

较高质量的人脸图像上, 且通常使用带有提取共享特征的网络结构模型按分类任务进行实现. Zhang 等^[2]提出了 TCDCN 网络进行人脸关键点检测并同时进行了多任务人脸属性识别任务, 该模型较好地处理了属性任务间的相关性, 但并未重点关注细粒度的人脸特征. Hand 等^[3]为了解决人脸属性识别中特征表达能力不强的问题, 同时提高模型的训练效率, 采用先对属于同一范畴的人脸属性进行分组, 然后再根据分组结果设计卷积神经网络结构的方法对人脸属性识别任务进行实现; 但这种方法具有较强的主观性, 在复杂场景下容易发生错误识别. 郑欣^[4]通过使用双向残

收稿日期: 2021-08-10; 修回日期: 2021-12-07

作者简介: 韩 菲 (1996—), 女, 天津人, 硕士研究生; 通信作者: 周卫斌, 副教授, zhouweibin@tust.edu.cn

差注意力机制获取更加准确的局部特征和全局特征,用以提高语义信息的丰富度,尽管这种方法与细粒度方法不同,但进一步验证了融合局部信息的判别特征,可以较好地提高人脸属性识别的精度.胡鼎昌^[5]利用属性之间可能具有相关性这种客观因素,使用属性内排序和各个属性间排序约束的方法提高识别准确度.曹穆赞^[6]使用对训练样本多尺度提取的方法,先对训练样本使用双线性差值进行扩大或缩小,再对获得的多尺度特征进行整合处理,以提高人脸属性识别的准确度.这种方法的思想与基于细粒度方法网络结构的思想相似.刘婧怡^[7]结合使用头肩检测和64点人脸关键点检测的方法对人脸面部进行裁剪,把裁剪后的人脸部件图像作为各个子网络的输入分别进行训练,以提升对疲劳驾驶的检测效果.但是,人脸在头部姿态大角度变化时,人脸关键点的预测精度也会随之下降,因此这种方法在处理复杂场景的人脸多类属性识别问题时可能准确度较低.

本文以卷积神经网络(convolutional neural network, CNN)为基础提出了一种基于细粒度特征的人脸属性识别网络(multiple granularity face attribute net, MGFA Net)结构,对复杂场景下人脸属性识别任务进行实现.该网络结构首先通过主干网络提取共享特征;然后对共享特征分别设置细粒度特征提取和全局特征提取的分支结构,把人脸图像区域拆分为4个子区域,并在每个子区域内提取更加细化的特征增强;最后用于判别人脸属性任务总体特征的表达能,有效提高了在复杂场景下人脸属性识别任务的准确度.

1 细粒度卷积神经网络

1.1 细粒度模型基本理论

在细粒度网络模型的设计中,用列向量表征子区域(block)中图像语义信息的方法是构建细粒度网络模型的核心思想之一.假设存在3维图像数据的张量 T ,然后把 T 经过激活函数后的特征图沿着通道维度(C)上的向量定义为列向量集合,以 h_i 表示^[8];假设存在某细粒度网络模型 M ,且在模型 M 中划分 g_i 个分支用于提取不同子区域的细粒度特征. g_i 对应的权重矩阵记为 W_i ,以集合 h_{dist} 表示列向量集合 h_i 之间的相似性,则细粒度网络模型的主要目标和约束条件可以表述为^[9]

$$\text{target} = \arg \min_{W_i} \sum_i \text{loss}_{g_i} + \arg \min_{h_i} \sum_i h_{\text{dist}} \quad (1)$$

式(1)表明在细粒度网络模型中,通常需要优化两部分损失函数使模型最优,即:需要使各个 g_i 分支上各损失函数值最小,以实现最佳的分类效果;需要优化列向量集合 h_i (包括影响 h_i 的主干网络中的权重参数和第 i 个分支上的权重参数),使列向量的相似性集合 h_{dist} 的求和值最小,以实现 h_i 中各列向量间的相似性最低.因此,细粒度网络的总体损失函数的作用可以概括为:通过模型的迭代训练使每个 g_i 分支具备提取该分支特征的能力,同时通过降低各组间列向量相似性的方法使每个分支结构都只对所在子区域具有针对性的特征提取能力,以防止各分支间提取相似的特征^[10-12].

对于式(1)中的 g_i 损失函数部分,通常采用交叉熵损失函数(softmax loss),且文献[8]指出分支数 g_i 对细粒度模型的准确度具有重要影响,过多划分子区域会增加误差值,导致模型精度降低.在构建细粒度模型结构时,通常使用全局平均池化(GAP)方法对 g_i 个分支区域进行下采样,且下采样率与 T 的分辨率要适配,以满足在划分block时特征图的完整性.通过GAP下采样可以获取 $C \times 1 \times 1$ 的特征图,然后在通道维度上切分可以获取 C 个 1×1 的特征图.对这些特征图使用降维操作后得到列向量集合 h_i ,且 h_i 中每个列向量都是一个特征描述符.对于式(1)中的 h_{dist} 损失函数部分,通常采用元组损失函数(triplet loss),文献[9]证明了在细粒度特征提取分支结构中,权重间不使用权值共享的有效性.因此,要尽量降低 h_{dist} 中各列向量的相似性,即优化权重参数,使 h_i 都独立关注其所在子区域内的局部特征.

1.2 人脸图像的细粒度特征

卷积神经网络模型在对图像提取特征时,不同模型会在不同损失函数和超参数的影响下,对同一图像数据产生不同的特征关注区域.对人脸属性识别而言,当人脸图像在光照亮度过强或过弱、人脸侧脸斜度较大的情况下,只使用提取全局特征的方法通常难以获得具有鲁棒性的特征以及对所有属性任务进行正确分类.为了解决这一问题,本课题组提出把人脸区域划分为4个子区域分别提取细粒度特征的方法,如图1所示.提取人脸细粒度特征,是指在划分出的每个区域上(如图1中的1~4子区域)均设定针对该区域提取特征的网络结构,通过模型的迭代训练,以实现对该区域提取更加细化的局部特征.在划分子区域时,不需要考虑区域内图像的语义信息,只需要按人脸图像宽、高的二分之一比例均匀切分.在构建网络结构时,使用 2×2 特征图进行实现,但本节需要说明的是,按图1所述的子区域划分方法,MGFA

Net 网络模型中 4 个细粒度分支上以卷积核为 4×4 的均值池化层代替全局平均池化层,且在 2×2 的特征图上同时进行切分,以获取图 1 中 1~4 区域上的特征表示. 这是由于人脸图像通常使用正方形数据,且标准正面人脸图像上人脸部件通常均匀分布在 1~4 区域中. 由于每个人脸子区域内包括具有不同表达能力的图像语义信息,因此模型中对应的每个细粒度分支结构只关注区域内的有效信息^[8],这对提高复杂场景下人脸属性识别的准确度和训练效率具有重要作用^[13].

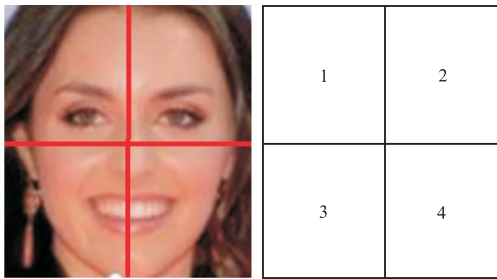


图 1 人脸子区域划分示意图

Fig. 1 Schematic diagram of face subregion division

1.3 损失函数

对于本文提出的人脸细粒度网络,结合细粒度模型基本理论,MGFA Net 网络模型构建的主要思想可以表述为:假设存在人脸图像的张量 T_{attr} ,设定 $g_i = 2$,即 2 个分支用于提取局部特征和全局特征,定义总列向量集合为 h ,定义分支的列向量集合为 h_{mn} ,其中 m, n 分别表示在 2×2 特征图上的第 m 行和第 n 列,则 $h \in \{h_1, h_2, h_{mn}\}$; h_1 用于提取第 1 个分支的全局特征, h_2 用于提取第 2 个分支上的全局特征, h_{mn} 用于提取图 1 中 1~4 区域对应的局部特征,并将每

个区域对应的行列向量集合分别记为 $h_{11}, h_{12}, h_{21}, h_{22}$. MGFA Net 网络模型的主要目标和约束条件可以定义为

$$\text{target}_{\text{MGFA}} = \arg \min_{W_{\text{task}}} \sum_i \text{loss}_{g_i} + \arg \min_{h_1, h_2, h_{mn}} \sum_i h \quad (2)$$

其中 $\text{loss}_{g_i} = \text{loss}_{\text{binary}} + \text{loss}_{\text{multi}}$ (3)

$$\text{loss}_{\text{binary}} = -[y \times \log(p) + (1-y) \times \log(1-p)] \quad (4)$$

$$\text{loss}_{\text{multi}} = -\sum_{i=1}^K y_i \log(p_i) \quad (5)$$

在式(2)中,由于 MGFA Net 网络模型使用了图 1 所示的分支划分方式,因此对于 h_{mn} 仍然在 $m \times n$ 的矩阵上分割获取列向量. $W_{\text{task}} \in \{\text{gender, smile, wear glass, head pose}\}$, W_{task} 为包含 MGFA Net 网络模型中共享网络部分的全部权重参数.

按照细粒度模型基本理论,对于本文训练使用的性别、是否微笑、是否佩戴眼镜这 3 个属性,每个人脸属性分类任务均使用式(4)的二分类交叉熵损失函数^[14]. 在式(4)中: y 为样本标签值, $y \in \{-1, 1\}$; p 为模型预测为正样本的概率, $p \in [0, 1]$. 对于头部姿态属性,使用式(5)的多分类交叉熵损失函数. 在式(5)中: K 为类别数, $K \in \{1, 2, 3, 4, 5\}$; y_i 为第 K 类标签值的判断值, $y_i \in \{1, 0\}$; p_i 为模型预测为第 K 类的概率.

1.4 MGFA Net 网络模型的结构

细粒度特征提取(MGN)网络^[9]在行人重识别研究领域已经取得了较好的效果,本文在 MGN 网络基础上,提出一种基于细粒度特征的人脸属性识别网络结构(MGFA Net). MGFA Net 网络以 ResNet18 网络模型为基础主干网络,其模型结构如图 2 所示.

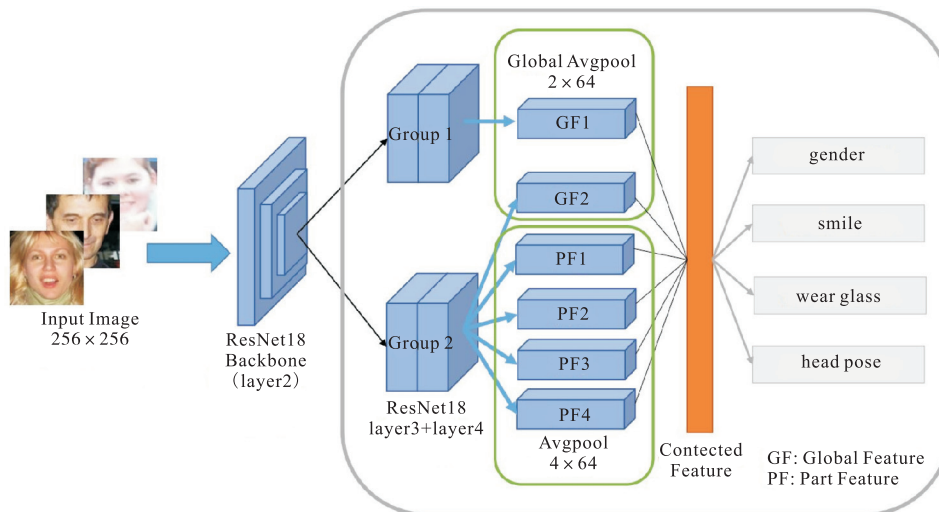


图 2 MGFA Net 网络模型的结构图

Fig. 2 Structure diagram of MGFA Net network model

在 MGFA Net 网络模型的结构中,使用 ResNet 18 网络模型的第 1 层至 layer2 部分作为提取共享特征的主干网络,然后设置 2 个特征提取分支,在每个分支中都使用 ResNet18 网络模型的 layer3 和 layer4 部分进一步提取特征.第 1 个分支(GF1)不对人脸区域划分,即提取人脸的全局特征;第 2 个分支部分包含 5 个分支结构,其中第 1 个分支结构(GF2)提取第 2 个分支的全局特征,然后第 2~5 个分支结构(PF1—PF4)提取图 1 中 1~4 对应的人脸子区域的细粒度特征.上述 6 个分支结构均提取 64 维度的特征,并把 6 部分的特征组合为最终判别各个属性任务的总特征.

由于 GF1 和 GF2 为全局特征,这 2 个分支在特征降维前均使用全局平均池化,但 PF1—PF4 细粒度分支结构在特征降维前使用卷积核为 4×4 的均值池化,详细特征图尺寸和特征图通道数见表 1.

表 1 MGFA Net 网络模型的特征图与通道数

Tab.1 Feature map and channel number of MGFA Net network model

层名称	特征图尺寸	特征图通道数
Input Image	256×256	3
Backbone	16×16	256
Group 1、Group 2	8×8	256
GF1、GF2	1×1	64
Avgpool_out	2×2	512
PF1、PF2、PF3、PF4	1×1	64

注: Avgpool_out 为切分 PF1—PF4 前的中间变量

对于 MGFA Net 网络模型的第 2 个特征提取分支中的细粒度特征提取网络(PF1—PF4),由于主干网络输出的特征图为 $256 \times 8 \times 8$,为加强特征提取能力和降低特征图分辨率,首先对该特征图进行卷积核

为 4×4 且增加卷积核数量的均值池化操作,得到 $512 \times 2 \times 2$ 的特征图;然后对于 2×2 的特征图在宽、高维度上按图 1 所示的方式分别切分出 PF1—PF4 结构的特征图,经降维操作后可以获取每个区域的细粒度特征.

2 实验

2.1 MTFL 数据集

MTFL 数据集包含 12 995 张采集自网络的人脸图像,包括性别(gender)、是否微笑(smile)、是否佩戴眼镜(wear glass)、头部姿态(head pose)4 种人脸属性类别. MTFL 数据集中的人脸图像数据如图 3^[15] 所示.

该数据集集中的大部分人脸图像具有光照亮度变化,且具有相对丰富的头部姿态变化.这些光照变化和头部姿态变化较好地反映了复杂场景下人脸图像的情况,同时每张人脸图像都标记有人脸关键点以方便数据预处理.实验使用 MTFL 数据集中性别、是否微笑、是否佩戴眼镜、头部姿态 4 种人脸属性,其类别标签格式见表 2.

表 2 MTFL 数据集中 4 种人脸属性的标签格式

Tab.2 Label format for 4 face attributes in the MTFL datasets

属性名称	标签值
性别	1(男性), -1(女性)
是否微笑	1(微笑), -1(不微笑)
是否佩戴眼镜	1(戴眼镜), -1(不戴眼镜)
头部姿态	1(左侧脸), 2(轻微左侧脸), 3(正脸), 4(轻微右侧脸), 5(右侧脸)

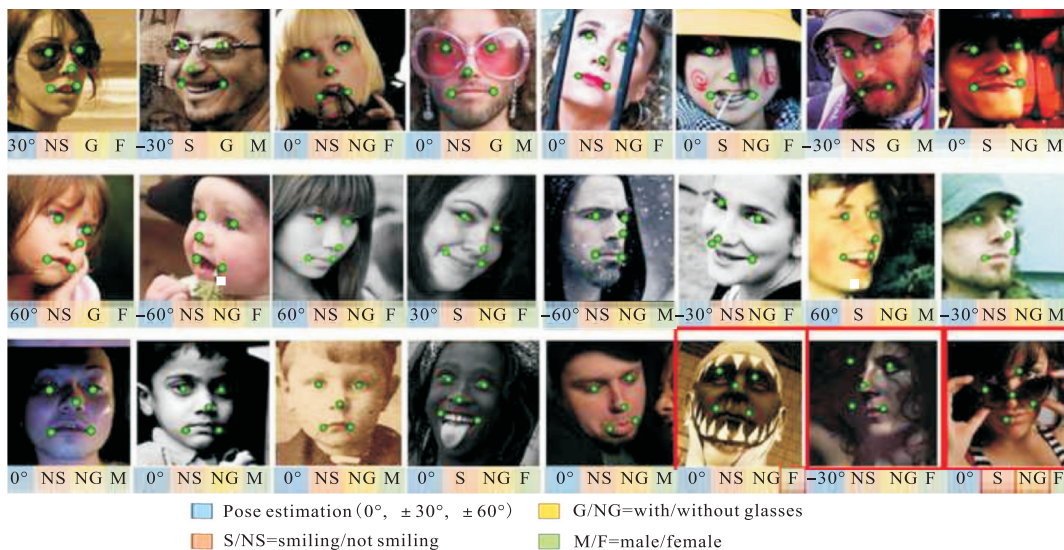


图 3 MTFL 数据集中的人脸图像数据

Fig.3 Human face images in MTFL datasets

2.2 硬件环境与训练参数

实验中使用 PyTorch 深度学习框架搭建 MGFA Net 网络模型结构,使用 2 块 NVIDIA GeForce GTX 1080 Ti 显卡进行训练,训练中 batch size 设为 50,起始学习率设为 0.01,权重衰减系数设为 5×10^{-3} ,模型训练使用的人脸图像分辨率为 256×256 . 为了将模型训练至最佳效果,ResNet18 网络模型的训练次数 epoch 设为 100,学习率衰减策略为在 epoch = (20, 40, 80) 时线性衰减,学习率衰减系数设为 0.1. MGFA Net 网络模型的训练次数 epoch 设为 400,学习率衰减策略为在 epoch = (320, 380) 时线性衰减,学习率衰减系数设为 0.1. 为了提高每个分支的特征提取能力,GF1—GF2 和 PF1—PF4 的权重值均不共享^[8-9]. 训练时均使用 ResNet18 网络模型的预训练权重.

2.3 实验结果与实验分析

2.3.1 对比实验

实验中设定一组对比实验证明 MGFA Net 网络模型提取细粒度特征的有效性,结果见表 3. 对比 MGFA Net 网络模型和 ResNet18 网络模型的实验结果可知:MGFA Net 网络模型提取的细粒度特征在 MTFL 数据集的性别、是否微笑、是否佩戴眼镜、头部姿态属性上具有更强的特征表达能力,分别实现了

准确度 2.50%、3.81%、0.50% 和 4.36% 的提升,平均识别准确度实现了 2.79% 的提升. 由于是否佩戴眼镜这一外在人脸属性相对于其他人脸属性的难度较低,因此识别准确度的提升较小. 对比 Ehrlich 等^[16]使用相对浅层的卷积神经网络和主成分分析方法可知,端到端的深层卷积神经网络对处理 MTFL 数据集中的属性任务具有明显优势,MGFA Net 网络模型相比 MTM-RBMs 方法提升明显,平均识别准确度上升 6.27%.

表 3 在 MTFL 数据集上的对比实验结果

Tab. 3 Comparative experiment results on the MTFL datasets

网络模型名称	识别准确度/%				平均识别准确度/%
	性别	是否微笑	是否佩戴眼镜	头部姿态	
MTM-RBMs ^[16]	79.00	79.00	93.30	75.70	81.75
ResNet18	85.44	81.03	98.73	75.70	85.23
MGFA Net	87.94	84.84	99.23	80.06	88.02

为了进一步验证细粒度特征的有效性,把 MGFA Net 网络模型的细粒度分支结构输出的特征图与 ResNet18 网络模型输出的特征图进行可视化对比,如图 4 和图 5 所示.

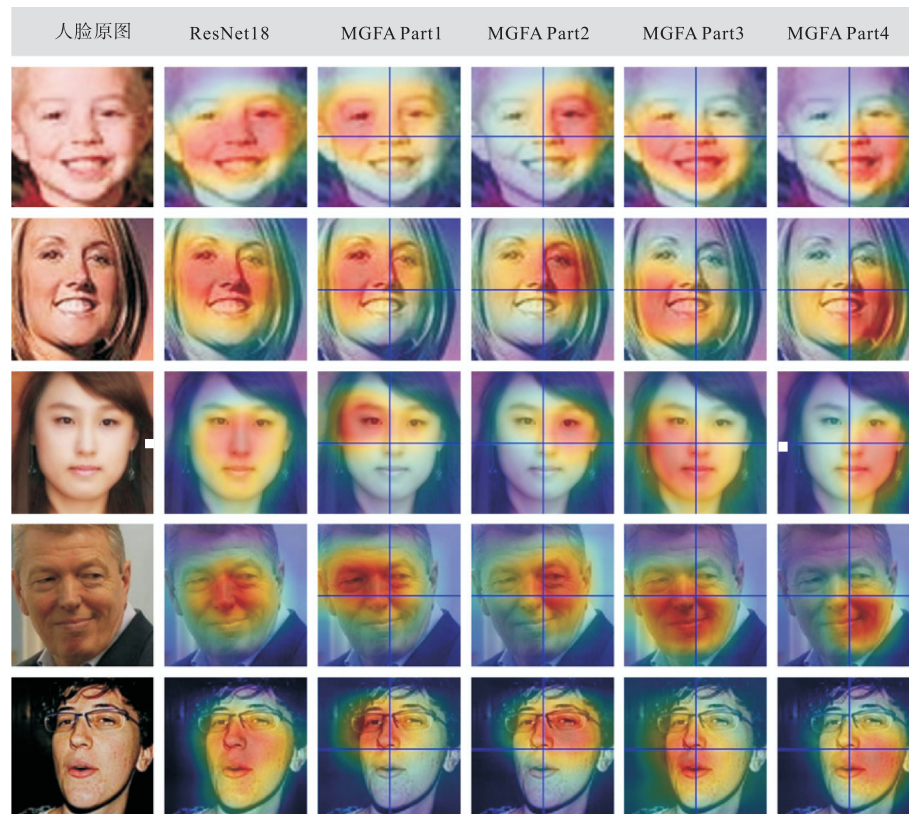


图 4 小姿态人脸图像特征图可视化对比结果

Fig. 4 Visualized comparison results of feature maps of small pose face images

从小姿态人脸图像的预测结果中可以观察到: ResNet18 网络模型在关注全局特征时可能会丢失必要的人脸区域或减弱对必要人脸区域的关注度,尤其在人脸边界位置; MGFA Net 网络模型的细粒度特征

提取分支有效地加强了局部特征的提取能力,例如图 4 中第 1 行第 2 列缺少对左侧眼部的关注,而第 1 行第 3 列的 MGFA Part1 则较好地关注了左侧眼部区域内的特征信息.

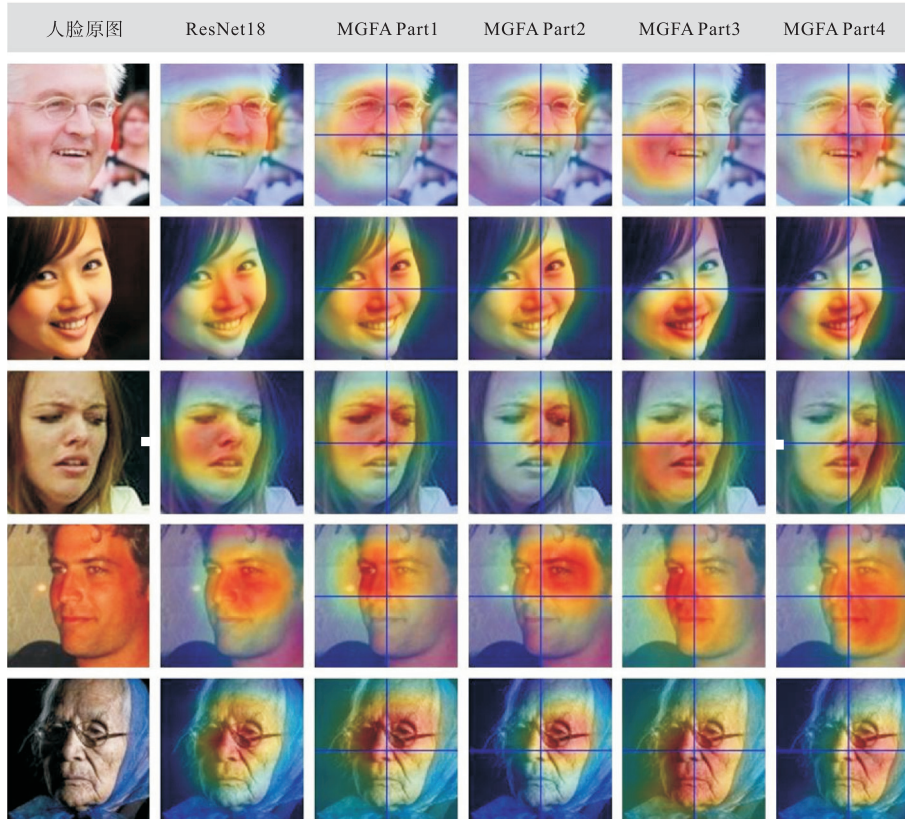


图 5 大姿态人脸图像特征图可视化对比结果

Fig. 5 Visualized comparison results of feature maps of large pose facial images

在大姿态人脸图像的预测结果中, ResNet18 网络模型在关注全局特征时丢失必要的人脸区域或减弱对必要人脸区域的关注度的现象更加严重^[17]. 当全局特征的关注区域不能包含所有属性任务必要的关注区域时,对应的属性任务的识别精度将会降低. 如图 5 中第 1、4、5 行第 2 列, ResNet18 网络模型对人脸图像中嘴部的关注度较弱,而第 2、3 行第 2 列的人脸图像分别对左侧眼部和左脸、右侧眼部和右脸的关注度较弱.

在 MGFA Net 网络模型大姿态人脸图像的预测结果中,由于过大的姿态可能会导致均匀分割的某些区域内所包含的有效人脸信息很少(或不存在),此时某些细粒度分支结构的关注区域也可能会扩大到相邻的区域上. 实验结果表明这些关注邻域信息分支结构提取的特征是有效的,并且这些特征在大姿态人脸图像的推理过程中加强了对于某些人脸属性识别任务的特征表达能力^[18-19]. 与 ResNet18 网络模型相

比,图 5 中第 1 行的第 3、6 列分别加强了对眼部区域和右侧脸部区域的关注度,第 4 行中第 5、6 列分别加强了对左侧脸部和右侧脸部下方区域的关注度. 这些强化的局部特征对于是否微笑和头部姿态的识别准确度的提升更加明显.

2.3.2 特征长度与分支区域实验

对于 MGFA Net 网络模型结构的构建,实验中分别设定一组不同特征长度实验和一组部分区域划分实验,结果见表 4 和表 5.

由表 4 可知:不同的特征维度对 MGFA Net 网络模型的检测精度有一定的影响,其中过高和过低的特征维度会引起模型对 4 种属性的识别准确度下降. 这也是由于输入数据的人脸图像通常分辨率较低,过高的特征维度会使提取特征冗余并容易导致模型难以收敛,过低的特征维度则难以充分表达各区域的信息. 由表 5 可知:当划分 2 个和 8 个分支区域时,模型在 4 种属性识别上的准确度下降,且分支区域越多

模型的准确度下降越明显.实际上,对于正方形的人脸图像以及人脸面部器官分布而言,按照图1所示的

分支区域划分方法更近似于人眼观察人脸面部信息的情况^[20].

表4 不同特征长度对比实验结果

Tab. 4 Comparison of experimental results with different feature lengths

特征长度/维	识别准确度/%				平均识别准确度/%
	性别	是否微笑	是否佩戴眼镜	头部姿态	
32	87.47	84.00	99.23	79.66	87.59
64	87.94	84.84	99.23	80.06	88.02
128	87.31	84.54	99.43	78.79	87.52
256	86.87	83.23	99.16	77.92	86.80

表5 不同分支区域个数对比实验结果

Tab. 5 Comparison of experimental results of the number of different branch areas

分支区域/个	识别准确度/%				平均识别准确度/%
	性别	是否微笑	是否佩戴眼镜	头部姿态	
2	86.97	84.54	99.39	78.43	87.33
8	86.07	83.90	99.23	76.69	86.47

3 结论

本文以卷积神经网络技术为基础提出了一种基于细粒度特征的人脸属性识别网络结构 MGFA Net 网络模型,用以解决实际场景中复杂人脸图像的人脸属性识别准确度低的问题.通过对比实验,证明了细粒度特征可以有效提高模型对于局部特征的提取能力,在 MTFL 数据集上对性别、是否微笑、是否佩戴眼镜、头部姿态4种人脸属性的平均识别准确度提高了2.79%.但是,受数据集中人脸数据复杂程度和数量的限制,对基于细粒度卷积神经网络的人脸属性识别的研究,还需要进一步在实际复杂场景的数据集上进行对比实验;同时,对于如何构建有效的细粒度网络分支结构,例如是使用只有列向量的条状分块,还是按照模拟人脸区域使用行列向量进行分块,也是接下来主要的探索方向.

参考文献:

- [1] 葛宏孔,罗恒利,董佳媛.基于深度学习的非实验室场景人脸属性识别[J].计算机学报,2019,46(S2):246-250.
- [2] ZHANG Z, LUO P, LOY C C, et al. Learning deep representation for face alignment with auxiliary attributes[J]. IEEE Transactions on pattern analysis and machine intelligence, 2016, 38(5):918-930.
- [3] HAND E M, CHELLAPPA R. Attributes for improved attributes: a multi-task network for attribute classification[EB/OL]. [2021-08-09]. <https://arxiv.org/abs/1604.07360>.
- [4] 郑欣.深度双向阶梯式人脸属性识别网络研究[D].大连:大连理工大学,2020.
- [5] 胡鼎昌.人脸属性分析关键技术[D].北京:北京邮电大学,2020.
- [6] 曹穆赞.基于深度学习的人脸属性识别研究与实现[D].南京:南京邮电大学,2020.
- [7] 刘婧怡.基于机器视觉的人脸属性识别和状态检测技术研究[D].北京:北方工业大学,2021.
- [8] SUN Y, ZHENG L, YANG Y, et al. Beyond part models: person retrieval with refined part pooling[EB/OL]. [2021-08-09]. <http://arxiv.org/pdf/1711.09349>.
- [9] WANG G, YUAN Y, CHEN X, et al. Learning discriminative features with multiple granularities for person re-identification[EB/OL]. [2021-08-09]. <https://arxiv.org/abs/1804.01438v1>.
- [10] 邓绍伟,张伯泉.基于 Grad-CAM 与 B-CNN 的细粒度图像分类方法研究[J].计算机科学与应用,2020,10(5):841-850.
- [11] 邹承明,罗莹,徐晓龙.基于多特征组合的细粒度图像分类方法[J].计算机应用,2018,38(7):29-32.
- [12] 贾军伟.基于深度学习的细粒度图像识别算法研究及应用[D].杭州:浙江工商大学,2019.
- [13] 曹猛,田青,马廷淮,等.人类面部属性估计研究:综述[J].软件学报,2019,30(7):2188-2207.
- [14] 周志华.机器学习[M].北京:清华大学出版社,2016:425.
- [15] ZHANG Z, LUO P, LOY C C, et al. Facial landmark detection by deep multi-task learning[C]//European

- Conference on Computer Vision(2014). Cham : Springer, 2014: 94–108.
- [16] EHRlich M, SHIELDS T J, ALMAEV T, et al. Facial attributes classification using multi-task representation learning[C]//IEEE. 2016 IEEE Conference on Computer Vision and Pattern Recognition Workshops(CVPRW). New York: IEEE, 2016: 7789589.
- [17] 靳一凡. 基于级联卷积神经网络的人脸关键点检测算法[D]. 杭州: 浙江大学, 2015.
- [18] 王子衡. 基于 CNN 多模型集成和难三元组损失的细粒度家居商品图像识别[J]. 中国新通信, 2018, 20(24): 218–219.
- [19] 赵浩如, 张永, 刘国柱. 基于 RPN 与 B-CNN 的细粒度图像分类算法研究[J]. 计算机应用与软件, 2019, 36(3): 210–213.
- [20] FU S, HE H, HOU Z G. Learning race from face: a survey[J]. IEEE Transactions on pattern analysis and machine intelligence, 2014, 36(12): 2483–2509.

责任编辑: 郎婧

(上接第 53 页)

- 240.
- [2] MANOHARAN P, SUBRAMANIAM U, BABU T S, et al. Improved perturb and observation maximum power point tracking technique for solar photovoltaic power generation systems[J]. IEEE Systems journal, 2020, 15(2): 3024–3035.
- [3] LI T Y, ZHU H. Research of solar energy generate system maximum power point tracking algorithm based on Matlab[C]//Proceedings of 2019 Chinese Automation Congress(CAC). Hangzhou: IEEE, 2019: 1861–1865.
- [4] DZIWIŃSKI P, BARTCZUK Ł. A new hybrid particle swarm optimization and genetic algorithm method controlled by fuzzy logic[J]. IEEE Transactions on fuzzy systems, 2020, 28(6): 1140–1154.
- [5] DURSUN E H, KOYUNCU H, KULAKSIZ A A. A novel unified maximum power extraction framework for PMSG based WECS using chaotic particle swarm optimization derivatives[J]. Engineering science and technology, an international journal, 2021, 24(1): 158–170.
- [6] ZHANG J H, WEI X Y, HU L, et al. A MPPT method based on improved fibonacci search photovoltaic array[J]. Tehnički vjesnik, 2019, 26(1): 163–170.
- [7] IMTIAZ T, KHAN B H, KHANAM N. Fast and improved PSO(FIPSO)-based deterministic and adaptive MPPT technique under partial shading conditions[J]. IET Renewable power generation, 2020, 14(16): 3164–3171.
- [8] WU J C, JOU H L, HUANG P H. Seven-level power conversion system for solar power generation system[J]. IET Renewable power generation, 2020, 14(8): 1387–1394.
- [9] FAKHOURI H N, HUDAIB A, SLEIT A. Multivector particle swarm optimization algorithm[J]. Soft computing, 2020, 24(15): 11695–11713.
- [10] ZHU Y L, YOUSEFI N. Optimal parameter identification of PEMFC stacks using Adaptive Sparrow Search Algorithm[J]. International journal of hydrogen energy, 2021, 46(14): 9541–9552.
- [11] TOO J, ABDULLAH A R. Chaotic atom search optimization for feature selection[J]. Arabian journal for science and engineering, 2020, 45(8): 6063–6079.
- [12] WANG Y Z, WANG D, ZHANG X, et al. McDPC: multi-center density peak clustering[J]. Neural computing and applications, 2020, 32(17): 13465–13478.
- [13] ZHOU J, DONG S B. Hybrid glowworm swarm optimization for task scheduling in the cloud environment[J]. Engineering optimization, 2018, 50(6): 949–964.
- [14] LIANG Y L, HE Y Y. Simulation research on the grid connected generation system of solar thermal power generation[J]. Thermal science, 2020, 24(5): 3239–3248.
- [15] KANO F, KASAI Y, KIMURA H, et al. Buck-boost-type MPPT circuit suitable for vehicle-mounted photovoltaic power generation[J]. IEEE Transactions on electrical and electronic engineering, 2021, 16(9): 1229–1238.

责任编辑: 周建军