

陈卫东,熊继平,李正浩,等. 人脸面部表情识别技术综述[J]. 智能计算机与应用,2024,14(11):16-23. DOI:10.20169/j.issn.2095-2163.241103

人脸面部表情识别技术综述

陈卫东¹,熊继平¹,李正浩¹,陈巍¹,李佳珂²

(1 浙江师范大学 物理与电子信息工程学院,浙江 金华 321004; 2 浙江师范大学 计算机科学与技术学院,浙江 金华 321004)

摘要: 面部表情识别是一项计算机视觉技术,该技术在医疗、行为分析、驾驶员疲劳监测、教育、心理健康分析以及其他人机交互系统等领域有广泛应用。本文首先对图像预处理、特征提取和特征分类的方法进行了简要总结,并介绍了面部表情识别的评估指标,总结了可用于面部表情识别的数据集。然后,对近三年的面部表情识别文献从基于对抗网络、卷积神经网络、注意力机制和滤波器这4个角度进行了分类和简要概述。最后,阐述了面部表情识别领域面临的挑战和未来的发展方向。

关键词: 面部表情识别; 深度学习; 对抗网络; 卷积神经网络; 注意力机制; 滤波器

中图分类号: TP399

文献标志码: A

文章编号: 2095-2163(2024)11-0016-08

An overview of facial expression recognition technology for the human face

CHEN Weidong¹, XIONG Jiping¹, LI Zhenghao¹, CHEN Wei¹, LI Jiako²

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

2 School of Computer Science and Technology, Zhejiang Normal University, Jinhua 321004, Zhejiang, China)

Abstract: Facial expression recognition is the computer vision technology that has wide applications in fields such as healthcare, behavior analysis, driver fatigue monitoring, education, psychological health analysis, and other human-computer interaction systems. This article provides a brief summary of methods for image preprocessing, feature extraction, and feature classification, as well as introduces evaluation metrics for facial expression recognition. Next, available datasets for facial expression recognition are summarized. Then, recent literature on facial expression recognition from the past three years is categorized and briefly outlined from four perspectives: adversarial networks, convolutional neural networks, attention mechanisms, and filters. Finally, the challenges and future directions in the field of facial expression recognition are discussed.

Key words: facial expression recognition; deep learning; adversarial networks; convolutional neural networks; attention mechanisms; filters

0 引言

人类在沟通中常常利用身体和面部信号、反应和语气来表达情感。面部表情在沟通中起着十分重要的作用,是由脸部肌肉的运动而形成的,反映了人们的情绪和心理变化。在日常生活中,面部表情不可或缺,因为人们的情感主要通过语言、语调、身体动作和面部表情来表达,而面部表情中蕴含着丰富的信息,能够传递出人们内心最真挚的感情。随着人工智能的快速发展,面部表情识别技术已成为计算机视觉领域的一个重要研究方向。面部表情识别在医疗、行为分析、驾驶员疲劳监测、教育、心理健康

分析以及许多其他人机交互系统等各种应用中起着至关重要的作用^[1-5]。为了正式确定面部表情识别的标准,Paul等学者^[6]提出将面部表情划分为6种,包括愤怒、厌恶、恐惧、快乐、悲伤和惊讶。无论文化和种族的差异如何,这6种基本表达方式大致反映了人类的情感状态。由于面部表情变化的高度非线性和身份偏差、头部姿势、光照和遮挡等较大的外观变化,因此人脸表情识别是一个很有挑战的问题。

1 面部表情识别技术

面部表情识别流程如图1所示。面部表情识别技

基金项目:浙江省金华市公益项目(2021-4-116)。

作者简介:陈卫东(1999—),男,硕士研究生,主要研究方向:深度学习;李正浩(2000—),男,硕士研究生,主要研究方向:深度学习;陈巍(1998—),男,硕士研究生,主要研究方向:深度学习;李佳珂(1999—),男,硕士研究生,主要研究方向:深度学习。

通信作者: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:2023-06-26

术包括 3 部分,分别是:人脸图像预处理、特征提取和 分类。

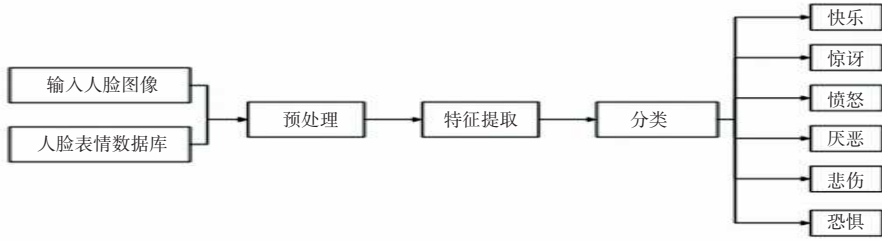


图 1 面部表情识别流程图

Fig. 1 Flow chart of facial expression recognition

1.1 图像预处理

在面部表情识别任务中,图像预处理是指在将图像输入模型之前对其进行处理的过程,旨在提高模型的准确性和稳定性。通过使用适当的预处理方法,可以提高模型的效率和准确性,降低模型的复杂性,使其更容易学习和理解面部表情的特征。在选择预处理方法时,应根据数据集和任务的特点来选择最适合实际情况的预处理方法。

人脸图像预处理的方法包括去除噪声、平滑边缘、增强对比度等操作。数据增强技术如旋转、缩放、平移和翻转能够增加数据集的多样性,提高模型的泛化能力。像素归一化、图像分辨率降低和图像裁剪等操作可以减少计算量和内存消耗,加速模型的训练和推理。中值滤波用于人脸图像归一化,减少光照变化的影响,提高稳健性和清晰度^[7]。裁剪和过滤操作可以提高模型的效率和准确性,去除与任务无关的信息并提取关键区域。颜色空间转换和直方图均衡化可以增强图像特征和对比度,帮助模型学习和理解图像信息。人脸对齐是重要的预处理方法,可以通过 SIFT 算法对图像进行对齐^[8]。ROI 分割是一种重要的预处理方法,通过分割脸部图像的颜色成分和关键区域来调节尺寸^[9]。

1.2 图像特征提取

图像特征提取是从图像中获取有意义信息的过程,在面部表情识别中起关键作用。成功提取代表分类信息的有效特征对于准确的表情识别至关重要。特征提取方法可分为基于纹理、统计、运动、几何和深度学习等 5 种类型。

1.2.1 基于纹理特征的方法

Gabor 滤波器是基于纹理的特征提取方法,用于描述图像的局部纹理特征,如边缘和纹理方向。并具有旋转不变性和适应不同尺度的能力,生成的特征图可解释性好,有助于分析图像的纹理特征和局部结构^[9-10]。局部二值算法(LBP)是另一种描述纹理特征的方法,通过比较像素与其邻域像素生

成局部二值模式。WPLBP 算法^[11]是对 LBP 算法的改进,引入权重和滤波器来加权提取特征,具有更好的纹理特征表达能力和鲁棒性。Weber Local Descriptor (WLD) 是一种基于局部灰度图像模式的特征描述符,通过引入 Weber 增量和 LBP 算法,从分割后的人脸图像中提取高判别纹理特征^[12]。

1.2.2 基于统计的方法

隐马尔科夫模型 (Hidden Markov Model, HMM),是在马尔科夫链的基础上研发出的概率模型,用于描述统计特征。文献[13]提出了改进的 HMM 训练算法,即类特定的 HMM 算法。该算法通过保留给定最大维度的信息,为每个类分配了独立的特征系统。类特定的 HMM 算法在识别率、检测率总体上优于传统 HMM 算法。

1.2.3 基于运动的方法

在面部表情识别中,光流法可以用于提取面部表情随时间变化的动态特征,从而获得面部表情的运动信息。光流法的基本假设是,在连续的 2 帧图像中,同一个像素点的灰度值是不变的,因此可以通过比较 2 帧图像中同一像素点的灰度值来计算其运动向量。当光流被用来描述面部表情时,用于克服运动不连续的运动近似方法有很大的影响^[14]。

1.2.4 基于几何特征的方法

在面部表情识别中,局部小曲线变换 (Local Curvelet Transform, LCT) 常被用于提取面部表情的形状和纹理特征。LCT 通过分解图像中的曲线信息,将图像分解为不同的局部频域中的子带,从而提取出图像的纹理和形状信息。提取的几何特征包括平均值、熵和标准差^[15]。

1.2.5 基于深度学习的方法

CNN 是一种神经网络,通过卷积层捕捉图像的局部特征,并通过池化层减少特征维度和计算量,最终通过全连接层将特征映射到目标类别空间。常见的 CNN 网络有 AlexNet、VGGNet、GoogleNet 和 ResNet 等。

RNN 是一种用于处理序列数据的神经网络,可在不同时间步中传递神经元信息,利用前一刻的状态作为后一刻的输入,可以提取长期依赖特征。

AE 是一种无监督学习方法,将输入数据映射到低维空间,然后通过解码器将低维空间的特征重新构建到原始空间中,以提取数据的关键特征。

GAN 是一种无监督学习方法,通过对生成器和判别器的对抗学习,得到与真实数据相似的样本。判别器学习将真实数据与生成器生成的数据区分开,从而生成高质量的数据样本并提取数据的潜在特征。

注意力机制能够自适应地对不同特征进行加权,以提取更重要的特征。一些常见的注意力机制包括 CBAM^[16] 和 SeNet^[17] 等。

1.3 图像表情分类

面部表情识别的最后阶段是分类,分类器用于对表情进行分类,例如愤怒、厌恶、恐惧、快乐、悲伤和惊讶。特征分类在此过程中起着关键作用。即能够自动从原始图像中提取最具区分性和代表性的特征,这些特征对于区分不同的面部表情至关重要。通过采用适当的特征分类算法,可以提高识别准确率,实现更好的面部表情识别。

1.3.1 支持向量机

支持向量机(SVM)是一种常见的分类和回归算法。通过将数据映射到高维空间,使原本线性不可分的数据在新的空间变得线性可分。能够有效处理高维数据,具有良好的泛化能力,适用于小样本情况下的分类,同时也可以处理非线性数据。

1.3.2 决策树

决策树(Decision Tree)是一种常见的用于分类和回归的算法。通过构建树形结构对数据进行分类或预测。决策树选择最优特征划分数据,节点表示特征,分支表示特征取值,叶子节点表示分类或回归结果。决策树易于理解和解释,适用于多类别问题,具有高效性和可扩展性。然而,决策树容易过拟合,需要剪枝操作来避免这种情况发生。处理连续性特征和缺失值比较困难,需对其进行数据预处理。

1.3.3 随机森林

随机森林(Random Forest)是一种基于决策树的集成学习算法,通过随机采样和特征选择来提高模型的泛化能力和准确性。适用于高维和非线性数据,具有较好的泛化性能和准确性。还可用于特征选择和缺失数据处理,减少过拟合风险,并对噪声和异常数据具有鲁棒性。此外,可估计变量重要性,指

导特征选择和模型优化。然而,随机森林也有缺点,如需要较多的内存和计算资源,对较少类别的分类处理可能欠佳,且模型解释性较差。

1.3.4 KNN

KNN(K-Nearest Neighbor)算法是一种基于实例的学习方法,用于分类问题。该算法的核心思想是通过计算待分类样本与训练集中各样本之间的距离来进行分类。KNN 算法具有简单易懂、容易实现和适用于多分类问题的优点,但其计算量相对较大,而且对于不平衡数据集的表现可能不佳。此外,距离度量 and K 值的选择也会影响算法的性能。

1.3.5 AdaBoost

AdaBoost 算法是一种集成学习方法,通常用于分类问题。该算法使用一组弱分类器来构建一个强分类器,其中每个弱分类器的输出通过加权投票来决定最终的分类结果。在每一轮迭代中,AdaBoost 会根据先前分类错误的点来调整训练数据集的权重,以便下一轮迭代时这些点能得到更多的关注。这种方法可以提高分类器集成的泛化能力和预测准确性。

1.4 面部表情识别评估指标

面部表情识别的性能评估可以从分类准确度和评价指标两方面进行评估。其中,分类准确度是指分类器正确分类的样本数与总样本数的比例,通常以百分比表示。评价指标包括准确率、精确率、召回率、 $F1$ 值和特异性。具体地,准确率表示分类器在所有测试样本中正确分类的比例;精确率表示在所有被分类器预测为某个类别的样本中,实际属于该类别的比例;召回率表示在所有属于某个类别的样本中,被分类器正确预测为该类别的比例; $F1$ 值综合考虑精确率和召回率,越高表示分类器性能越好;特异性表示在所有负样本中,模型正确识别为负样本的比例。这些指标用如下公式来表示:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Precision = \frac{TP}{TP + FP} \quad (2)$$

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

$$AccuracyB = \frac{TPR + TNR}{2} \quad (4)$$

$$Specificity = \frac{TN}{TP + FP} \quad (5)$$

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (6)$$

混淆矩阵是用来展示分类器预测结果和实际标签之间对比情况的矩阵,包括真阳性(TP)、真阴性(TN)、假阳性(FP)和假阴性(FN)。进一步地,则分别表示模型正确地将正类预测为正类、负类预测为负类,以及错误地将负类预测为正类、正类预测为负类。混淆矩阵有助于评估模型性能和选择最优模型。ROC 曲线和 AUC 值用于评估分类器在不同阈值下的分类效果,ROC 曲线越接近左上角, AUC 值越大,表示分类器性能越好。

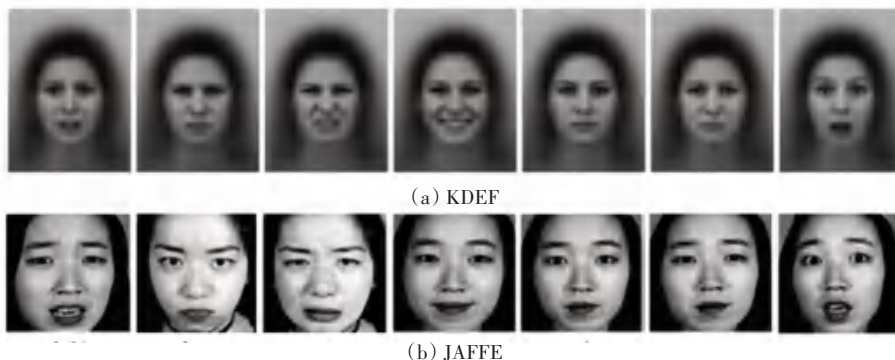


图 2 面部表情图像

Fig. 2 Facial expression images

2.1 KDEF 数据集^[18]

KDEF 数据集包含 4 900 张人类面部表情图像,每张图像的尺寸为 562×762 像素。涵盖了 7 种不同的表情,每种表情都有 5 个角度的图像。每个表情的图像都是以高分辨率拍摄的,并提供了多个不同的角度和照明条件。

2.2 CK+数据集^[19]

CK+数据集包含 123 名参与者的 593 张面部图像,包括 6 个基本表情,每张图像都有相应的面部标记,可以用于标识重要的面部特征。

2.3 MMI 数据集^[20]

MMI 数据集包含 2 900 个视频以及 740 张图片,分辨率 720×576 像素。MMI 数据集中出现的帧序列是根据起始点和偏移量来标记的。

2.4 JAFFE 数据集^[21]

该数据集包含大约 213 个面部表情样本。这些样本取自 10 名日本女性,每个人对应 3~4 张图像,分辨率为 256×256 像素。

2.5 FER 2013 数据集^[22]

FER-2013 数据集包含 35 887 张灰度图像,分为 7 个不同的情绪类别,每张图像都是 48×48 像素大小,由人工标注者对每张图像进行情绪类别标记。

2.6 RAF-DB 数据集^[23]

该数据集包含从互联网上的各种来源收集的大

2 人脸表情数据集

人脸表情数据集是一个用于研究和开发人脸表情识别和情感分析的数据集。数据集包含了大量的人脸图像和相应的表情标签,旨在帮助计算机系统理解和识别人类面部表情所传达的情感状态。在图 2 中展示了 KDEF 数据集和 JAFFE 数据集中的面部表情图像。图 2(a)~(b)中,从左到右依次为:害怕、生气、厌恶、开心、中性、伤心、惊讶。

量面部图像(29 672 张)。数据集中的图像涵盖了各种面部表情。该数据集包括手动注释的 7 个基本情感标签和 11 个复合情感标签。

2.7 Affect Net 数据集^[24]

Affect Net 是一个大规模的面部情感识别数据集,包含超过一百万张面部表情图像,并且每张图像都有相应的情感标签。覆盖了多种情感类别。数据集的标注采用了人工和机器标注相结合的方法,通过多个标注者的平均值提高了标注的准确性。

3 基于深度学习的表情识别方法

传统的面部表情识别技术依赖于人工设计的特征提取和分类器模型,需要大量的人工干预和调整,限制了识别效果。相比之下,深度学习通过自动学习特征和分类模型,显著提升了面部表情识别的准确率和鲁棒性。深度学习模型能够直接从原始图像中提取特征,无需手动设计特征提取器,减轻了特征提取的工作量,提高了模型的泛化能力。

3.1 基于对抗网络

对抗网络能提高面部表情识别模型的泛化能力和鲁棒性。通过训练生成器生成对抗噪声的图像,帮助模型更好地学习面部表情特征,提高识别性能。生成新的样本来平衡不同类别的数据,解决面部表情数据集中样本数量不平衡的问题,使得模型在各

个表情类别上都能得到充分训练。通过训练对抗网络,可以提高模型对抗攻击的能力,提升鲁棒性。

Sima 等学者^[25]将条件生成式对抗网络(cGAN)和用于生成2个关键帧的经过微调的三联距离模型结合,使用2个具有中性和最大表达强度的关键帧来减少受试者之间的差异。在CK+、MMI和北京航空航天大学数据集分别取得了98.05%、92.98%和93.06%的准确度。Yang等学者^[26]在GAN模型的基础上综合了二维和三维空间中任意强度的高强度表达式。从多模态增强的人脸中提取的面部特征被输入到融合子网中,从而获得更全面的面部表征用于表情识别。Ju等学者^[27]使用GAN生成不同角度的面部表情图像,使用深度回归网络进行关键点检测,然后将其发送到卷积神经网络进行分类和学习。在卷积神经网络中,通过引入阻力损失、识别损失、内容丢失和中心丢失来动态调整损失权重,从而提高识别精度。Cai等学者^[28]为了防止学习到的特征掺杂身份相关信息,提出了无身份条件生成对抗网络(IF-GAN),cGAN用来将给定的输入面部表情转化为与输入表情相同的“平均”身份脸。减少由身份相关的面部属性(如年龄、种族和性别)引起的主体间差异。Fan等学者^[29]使用对抗网络(TP-GAN)来补充头部姿势的面部表情信息。利用各种姿势和表情之间的联系来提高表情识别的准确性。Zhang等学者^[30]基于生成对抗网络(GAN),利用人脸图像的几何形状,实现了人脸图像合成和姿态不变人脸表情识别的同时进行。陈霖等学者^[31]将分类器嵌入到GAN网络中训练,采用GAN网络的变体生成具有不同姿态和表情的面部图像,以解决任意姿态下的面部表情识别问题。杨鼎康等学者^[32]结合生成对抗网络,利用GAN的对抗学习思想,通过博弈训练生成更多特定标签的面部表情图片,运用多个异质卷积神经网络提取不同深度、不同语义的特征向量,保留细微且表达能力强的表情特征,对不同尺度特征进行融合。陈亮等学者^[33]将生成对抗网络融合Softmax分类器构建出一个IC-GAN网络,在面部表情识别中建立对抗关系,提取面部表情特征从而进行面部表情识别。

3.2 基于卷积神经网络

卷积神经网络(CNN)能够学习判别性特征表示,并通过卷积层和池化层提取局部和全局的图像特征。适用于面部表情识别,因为CNN具有平移不变性和局部不变性。即使面部表情在图像中的位置发生变化,卷积神经网络仍能提取相同的特征,从而

实现稳定的表情识别。卷积神经网络由多层卷积层和全连接层组成,通过数据增强技术扩充训练数据集,提高网络的泛化能力,增加在真实世界面部表情识别任务中的鲁棒性。全连接层用于将特征映射到不同的面部表情类别,通过训练分类器,卷积神经网络能够学习到不同表情之间的差异,实现准确的面部表情分类和识别。

Zhou等学者^[34]在VGG-16网络结构进行改进,将平均采样层替换全连接层,并与长、短期存储网络相结合,将TAFMN激活函数应用于卷积神经网络,解决深度卷积网络中的梯度消失和梯度爆炸问题。使用SVM对特征进行分类以检测面部表情。Liu等学者^[35]在VGG-16的基础上引入ResNet的部分网络结构形成轻量级融合主干,采用姿势引导的面部对齐方法来减少类内差异并克服环境噪声的影响。同时使用混合特征表示方法,以获得高级判别性面部特征。在CK+、JAFPE、Oulu-CASIA和AR四个基准数据集上,模型的识别率分别为98.9%、96.8%、94.5%和98.7%。Khairuddin等学者^[36]基于VGGNet,对模型的超参数进行了微调,模型有4个阶段的卷积层和3个完全连接的层。4个卷积块提取图像的高级特征,而全连接层对图像进行分类。每个卷积阶段都有2个卷积块和1个最大池化层。卷积模块由卷积层、ReLU激活和批量归一化层组成。Zhang等学者^[37]通过简化卷积神经网络,只提取重要的面部特征。使用数据增强来扩展训练图像,并进行对齐、裁剪、直方图均衡化、Z-score归一化和下采样等预处理步骤来标准化图像数据。在训练过程中,采用随机旋转和水平翻转增加数据库规模。在CK+和JAFPE数据库上,平均识别准确率分别达到了97.38%和97.18%。Jia等学者^[38]将ALexNet、VGGNet、ResNet三个子网络组成面部表情识别模块,即将卷积神经网络进行集成,用SVM分类器整合3个网络的输出。郭玲玲等学者^[39]在保留原有卷积神经网络结构不变的前提下,设计了包含4层卷积层、2层全连接层的浅层网络结构结合LBP圆形算子进行面部表情特征的提取。

3.3 基于注意力机制

引入注意力机制可提升模型性能和准确度,卷积神经网络可聚焦于面部表情识别任务中的关键信息,如眼睛、嘴巴等。空间注意力机制和通道注意力机制可分别关注最具信息量和判别力的特征通道。引入注意力机制可使神经网络更加专注于面部表情的重要信息,减少对无关信息的干扰,提高表情识别

的准确性和鲁棒性。网络通过学习自动地分配和调整注意力,能够更好地适应不同面部表情的特征差异,并从中提取出更具判别性的特征表示。

Li 等学者^[40]引入了信道注意力机制模块 SeNet,用于自动学习张量特征中局部特征的权重,在卷积层中使用 3DPCA 算法从视频序列的空间和时间维度中提取动态表达特征。Lopes 等学者^[41]提出了多分支注意力卷积神经网络,网络结合了卷积块注意力模块、Depthwise 可分卷积和 3 种多分支结构方法,在 CBAM 的基础上提出了嵌入了具有深度可分离卷积的 DscCBAM 模块,使模型关注通道和空间中的重要特征。每个 DscCBAM 模块由 2 个深度可分离的卷积层、批处理归一化、Relu、最大池化和 CBAM 模块组成。Hu 等学者^[42]使用 Xception 模型和卷积方块注意力模块 (CBAM) 来学习关键的面部特征。采用轻量级卷积神经网络感知作为基本网络,嵌入卷积注意力模块,提高网络的表示能力。利用多尺度卷积核的深度可分卷积来扩展接受域,提取出更丰富的特征信息。Liu 等学者^[43]在模型中加入残差密集块中的残差和三重注意机制。将注意力、残差密集 (RRD) 块中的残差和注意力嵌入级联到 CNN 架构中进行面部表情。Wang 等学者^[44]利用空间变换网络对输入图像进行变换,然后在卷积网络中加入信道注意和空间注意。在卷积网络中使用了 GELU 激活函数,提高了面部表情的识别率。Xu 等学者^[45]以残差网络作为基础网络提取特征,用注意力模块定位有用信息,抑制无用特征的影响;然后将输出的等级特征分成一个阶段,构造 4 层特征金字塔网络并分别进行表达式预测,最后在决策层对预测值进行融合以获得最终的识别结果。郑伟^[46]改进了全局平均池化 ECANet,并提出了一种新的通道注意力模块 MECANet,对各通道特征信息进行全局池化操作。同时,引入多尺度注意力机制的密集连接网络 (DenseNet) 表情识别方法,通过减少网络层数和引入多尺度卷积层,在 CK+ 和 FER2013 数据集上实现了高识别率,分别为 96.2% 和 85.5%。李鹏威等学者^[47]将通道和空间注意力机制嵌入由 Propose Network (P-Net)、Refine Network (R-Net) 和 Output Network (O-Net) 三个阶段网络组成级联框架当中。在数据集 FER2013、JAFFE 上分别取得了 72.42% 和 95.71% 的识别准确率。

3.4 基于滤波器

卷积神经网络使用滤波器捕捉面部表情的局部

特征,如边缘、纹理和线条等。这些滤波器可识别面部表情中的细节和结构,提供重要信息。滤波器可通过学习进行调整和优化,以适应不同的样本。通过反向传播算法,网络可根据损失函数的反馈信号调整滤波器权重,更好地识别不同的表情类别。

Huang 等学者^[48]提出了基于肤色增强模型的人脸检测方法和基于分块主成分分析 (PCA) 的面部表情识别算法。使用同态滤波器扩大人脸图像的亮度范围并增强肤色对比度。通过建立肤色增强模型,利用 ycbCr 色彩空间组件定位面部区域,并使用面部差分水平积分投影提取特征。最后,使用带有深度神经网络的区块 PCA 完成面部表情识别。Revina 等学者^[49]提出了基于决策的规则导向中值滤波器 (DBROMF) 和多方向三角形模式 (FER-MDTP) 的人脸表情识别方法。DBROMF 用于去除面部图像中的脉冲噪声,FER-MDTP 提供了优越的图像描述。使用 MDTP-FES 算法定位人脸器官,提取特征地标,并与直方图导向的特征结合进行分类。最后,采用支持向量神经网络分类器 (SVNN) 进行表情分类。Revina 等学者^[50]提出基于增强修改决策的非对称修剪中位滤波器 (EMDBUTMF) 进行噪声消除的面部表情识别的方法,使用局部方向数 (LDN) 模式和显性梯度局部三元模式 (DGLTP) 描述符进行特征提取,以及使用 SVM 分类器进行面部表情分类。

4 结束语

4.1 面部表情识别的挑战

面部表情识别目前仍存在许多困难和挑战,包括以下几个方面:

- (1) 多样性:人脸表情多样,个人表情差异大,受环境、文化等因素影响,增加了面部表情识别难度。
- (2) 数据集不平衡:现有数据集中,面部表情类别不平衡,某些表情样本较少,导致对少数类别的识别效果较差。
- (3) 鲁棒性:面部表情识别模型容易受到噪声、光照、模糊等变化的影响,需要提高鲁棒性。
- (4) 隐私问题:面部表情识别技术广泛应用于安全、监控等领域,但引发隐私问题,需要进行技术和法律规范及保护。
- (5) 实时性:某些场景需要实时识别面部表情,如虚拟现实、视频会议,对算法速度和精度提出更高要求。

(6)跨数据集泛化能力:面部表情识别算法在特定数据集上训练良好,但在其他数据集上性能下降,需要提高泛化能力。

(7)识别面部微表情:面部微表情是短暂、微小的表情变化,在情感分析和谎言检测等领域具有重要应用价值,但难度更大。

4.2 面部表情识别的研究展望

未来的表情识别研究应将聚焦于跨模态表情识别,利用视频、声音和语言等多种信息源获取情感信息。同时,情感识别研究将深入探索人类情感的多样性和复杂性,包括不同文化和背景下的情感表达,并探索如何利用情感识别技术进行干预和治疗。在领域自适应的面部表情识别中,大规模真实场景下的面部表情视频数据库的建立成为一项迫切的研究任务,因为目前公开的数据库样本量有限且与真实场景存在差异,而样本量对于准确性至关重要。

参考文献

- [1] TRAN L, YIN Xin, LIU Xiaoming. Representation learning by rotating your faces[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 41(12): 3007–3021.
- [2] WANG Kai, PENG Xiaojiang, YANG Jiafei, et al. Suppressing uncertainties for large-scale facial expression recognition [C]// Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2020: 6897–6906.
- [3] WANG Kai, PENG Xiaojiang, YANG Jianfei, et al. Region attention networks for pose and occlusion robust facial expression recognition[J]. IEEE Transactions on Image Processing, 2020, 29: 4057–4069.
- [4] XU Ruyi, CHEN Jingying, HAN Jiaxu, et al. Towards emotion-sensitive learning cognitive state analysis of big data in education: deep learning-based facial expression analysis using ordinal information[J]. Computing, 2020, 102: 765–780.
- [5] AUNG N, TEWOGBOLA P. The impact of emotional labor on the health in the workplace: A narrative review of literature from 2013–2018[J]. AIMS Public Health, 2019, 6(3): 268.
- [6] PAUL E, FRIESEN W V. Constants across cultures in the face and emotion[J]. Journal of Personality and Social Psychology, 1971, 17(2): 124.
- [7] JI Yi, IDRISSE K. Automatic facial expression recognition based on spatiotemporal descriptors[J]. Pattern Recognition Letters, 2012, 33(10): 1373–1380.
- [8] DAHMANE M, MEUNIER J. Prototype-based modeling for facial expression analysis[J]. IEEE Transactions on Multimedia, 2014, 16(6): 1574–1584.
- [9] HERNANDEZ-MATAMOROS A, BONARINI A, ESCAMILLA-HEMANDEZ E, et al. Facial expression recognition with automatic segmentation of face regions using a fuzzy based classification approach[J]. Knowledge-Based Systems, 2016, 110: 1–14.
- [10] HEGDE G P, SEETHA M, HEGDE N. Kernel locality preserving symmetrical weighted fisher discriminant analysis based subspace approach for expression recognition [J]. Engineering Science and Technology, An International Journal, 2016, 19(3): 1321–1333.
- [11] KUMAR S, BHUYAN M K, CHAKRABORTY B K. Extraction of informative regions of a face for facial expression recognition [J]. IET Computer Vision, 2016, 10(6): 567–576.
- [12] COSSETIN M J, NIEVOLA J C, KOERICH A L. Facial expression recognition using a pairwise feature selection and classification approach [C]// International Joint Conference on Neural Networks. Piscataway, NJ: IEEE, 2016: 5149–5155.
- [13] LI Shaocong, BAI Yijun. Deep learning and improved HMM training algorithm and its analysis in facial expression recognition of sports athletes [J]. Computational Intelligence and Neuroscience, 2022, 2022: 1027735.
- [14] ALLAERT B, WARD I R, BILASCO I M, et al. A comparative study on optical flow for facial expression analysis [J]. Neurocomputing, 2022, 500: 434–448.
- [15] UÇAR A, DEMIR Y, GÜZELİŞ C. A new facial expression recognition based on curvelet transform and online sequential extreme learning machine initialized with spherical clustering [J]. Neural Computing and Applications, 2016, 27: 131–142.
- [16] WOO S, PARK J, LEE J Y, et al. Cbam: Convolutional block attention module [C]// Proceedings of the European Conference on Computer Vision (ECCV). Cham: Springer, 2018: 3–19.
- [17] HU Jie, SHEN Li, ALBANIE S, et al. Squeeze-and-Excitation networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 42(8): 2011–2023.
- [18] GOELEN E, RAEDT D R, LEYMAN L, et al. The Karolinska directed emotional faces: A validation study [J]. Cognition and Emotion, 2008, 22(6): 1094–1118.
- [19] TANUJAPATGAR T. Convolution neural network based emotion classification cognitive model for facial expression [J]. Turkish Journal of Computer and Mathematics Education (TURCOMAT), 2021, 12(13): 6718–6739.
- [20] VALSTAR M, PANTIC M. Induced disgust, happiness and surprise: an addition to the mmi facial expression database [C]// Proceedings of 3rd International Workshop on EMOTION (Satellite of LREC): Corpora for Research on Emotion and Affect. Piscataway, NJ: IEEE, 2010: 16083666.
- [21] LYONS M, AKAMATSU S, KAMACHI M, et al. Coding facial expressions with gabor wavelets [C]// Proceedings of Third IEEE International Conference on Automatic Face and Gesture Recognition. Piscataway, NJ: IEEE, 1998: 200–205.
- [22] GOODFELLOW I J, ERHAN D, CARRIER P L, et al. Challenges in representation learning: A report on three machine learning contests [C]// Proceedings of the 20th International Conference on the Neural Information Processing (ICONIP 2013), Part III 20. Cham: Springer, 2013: 117–124.
- [23] LI Shan, DENG Weihong, DU Junping. Reliable crowdsourcing and deep locality-preserving learning for expression recognition in the wild [C]// Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway, NJ: IEEE, 2017: 356–370.
- [24] MOLLAHOSSEINI A, HASANI B, MAHOOR M H. Affectnet: A database for facial expression, valence, and arousal computing in the wild [J]. IEEE Transactions on Affective Computing, 2017, 10(1): 18–31.
- [25] SIMA Y, YI J, CHEN A, et al. Automatic expression recognition

- of face image sequence based on key-frame generation and differential emotion feature[J]. *Applied Soft Computing*, 2021, 113: 108029.
- [26] YANG H, ZHU K, HUANG D, et al. Intensity enhancement via GAN for multimodal face expression recognition [J]. *Neurocomputing*, 2021, 454: 124-134.
- [27] JU L, ZHAO X. Mask-based attention parallel network for in-the-wild facial expression recognition[C]//*Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Piscataway, NJ:IEEE,2022: 2410-2414.
- [28] CAI J, MENG Z, KHAN A S, et al. Identity-free facial expression recognition using conditional generative adversarial network[C]//*Proceedings of IEEE International Conference on Image Processing (ICIP)*. Piscataway, NJ:IEEE, 2021: 1344-1348.
- [29] FAN Jiajun, WANG Shipu, YANG Po, et al. Multi-view facial expression recognition based on multitask learning and generative adversarial network[C]//*Proceedings of IEEE 18th International Conference on Industrial Informatics (INDIN)*. Piscataway, NJ: IEEE,2020: 573-578.
- [30] ZHANG F, ZHANG T, MAO Q, et al. Geometry guided pose-invariant facial expression recognition[J]. *IEEE Transactions on Image Processing*, 2020, 29: 4445-4460.
- [31] 陈霖, 段巍, 刘立志. 基于 GAN 网络的面部表情识别[J]. *电子技术与软件工程*, 2020, 171(1): 3-4.
- [32] 杨鼎康, 黄帅, 王顺利, 等. 基于生成对抗网络和网络集成的面部表情识别方法 EE-GAN[J]. *计算机应用*, 2022, 42(3): 750-756.
- [33] 陈亮, 吴攀, 刘韵婷. 面向类内差距表情的深度学习识别[J]. *中国图象图形学报*, 2020, 25(4): 679-687.
- [34] ZHOU Huiyu, TAJ M, CAVALLARO A. Target Detection and Tracking With Heterogeneous Sensors [J]. *IEEE Journal of Selected Topics in Signal Processing*, 2008, 2(4): 503-513.
- [35] LIU J, FENG Y, WANG H. Facial expression recognition using pose-guided face alignment and discriminative features based on deep learning[J]. *IEEE Access*, 2021, 9: 69267-69277.
- [36] KHAIREDDIN Y, CHEN Z. Facial emotion recognition: State of the art performance on FER 2013[J]. *arXiv preprint arXiv*, 2105. 03588, 2021.
- [37] ZHANG Xi, ZHANG Feifei, XU Changsheng. Joint expression synthesis and representation learning for facial expression recognition[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2021, 32(3): 1681-1695.
- [38] JIA Chen, CHU Lili, ZHANG Ying. Facial expression recognition based on the ensemble learning of CNNs [C]//*Proceedings of IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC)*. Piscataway, NJ:IEEE,2020: 1-5.
- [39] 郭玲玲, 苏冬娜, 胡绍彬. 结合 LBP 圆形算子的 CNN 面部表情识别研究[J]. *微型电脑应用*, 2023, 39(2): 1-4.
- [40] LI C, WEN C, QIU Y. A Video sequence face expression recognition method based on squeeze-and-excitation and 3DPCA network[J]. *Sensors*, 2023, 23(2): 823.
- [41] LOPES A T, AGULAR E D, SOUZA A F, et al. Facial expression recognition with Convolutional Neural Networks[J]. *Pattern Recognition*, 2017, 61(C): 610-628.
- [42] HU Zhibin, YAN Chunman. Lightweight multi-scale network with attention for facial expression recognition [C]//*Proceedings of the 4th International Conference on Advanced Electronic Materials, Computers and Software Engineering (AEMCSE)*. Piscataway, NJ:IEEE,2021: 695-698.
- [43] LIU K H, CHIU C H, LIU T J. Fusion of triple attention to residual in residual dense block to attention based CNN for facial expression recognition [C]//*Proceedings of IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. Piscataway, NJ:IEEE, 2022: 1045-1050.
- [44] WANG C, WANG Z, CUI D. Facial expression recognition with attention mechanism [C]//*Proceedings of the 14th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*. Piscataway, NJ: IEEE,2021: 1-6.
- [45] XU M, JI J, FANG C. Multi-feature facial expression recognition based on attention mechanism[C]//*Proceedings of the 14th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*. Piscataway, NJ: IEEE, 2022: 64-68.
- [46] 郑伟. 多尺度注意力机制 DenseNet 网络的表情识别方法[J]. *软件导刊*, 2023, 22(2): 81-86.
- [47] 李鹏威, 李剑. 结合级联框架与注意力机制的人脸表情识别[J]. *移动通信*, 2023, 47(4): 77-84.
- [48] HUANG Lei, XIE Fei, SHEN Shibin, et al. Human emotion recognition based on face and facial expression detection using deep belief network under complicated backgrounds [J]. *International Journal of Pattern Recognition and Artificial Intelligence*, 2020, 34(14): 2056010.
- [49] REVINA I M, EMMANUEL W R S. MDTP: A novel multi-directional triangles pattern for face expression recognition [J]. *Multimedia Tools and Applications*, 2019, 78: 26223-26238.
- [50] REVINA I M, EMMANUEL W R S. Face expression recognition using LDN and dominant gradient local ternary pattern descriptors [J]. *Journal of King Saud University-Computer and Information Sciences*, 2021, 33(4): 392-398.