

# 局部遮挡的人脸表情识别研究进展

莫文彬<sup>1</sup> 宋向辉<sup>1</sup> 蒋 强<sup>1</sup>

MO Wenbin SONG Xianghui JIANG Qiang

## 摘 要

随着机器视觉和人工智能的发展,人脸表情识别的研究开始从实验室场景转变到自然场景。局部遮挡的人脸表情作为自然场景的重要研究方向,受到国内外研究人员的广泛关注。针对目前缺乏局部遮挡人脸表情数据样本、局部遮挡人脸表情数据样本有效特征提取不足的问题,对局部遮挡面部表情的识别算法进行综述探讨。首先,对局部遮挡的人脸表情的研究背景进行介绍;然后,对现有的常用数据集以及遮挡数据样本的处理方式进行介绍;接着,将局部遮挡的人脸表情识别算法分为传统方法、深度学习方法以及传统和深度学习相结合的方法,并对算法主要的技术手段进行综述;最后,对该领域未来的研究方向和挑战进行总结和展望。

## 关键词

图像处理;人脸表情识别;局部遮挡;机器学习;深度学习

doi: 10.3969/j.issn.1672-9528.2024.08.050

## 0 引言

有研究表明人脸表情在人类的情感交流中蕴含着丰富的信息。随着计算机视觉和人工智能的发展,人脸表情识别在许多领域也得到广泛的应用,比如人机交互、医疗服务、智慧课堂、疲劳驾驶等<sup>[1]</sup>。在人机交互中,智能手机可以根据用户的表情来调整音量、亮度等设置,智能音箱可以根据用户的情绪播放不同类型的音乐或者提供不同的娱乐内容;在医疗服务中,通过识别患者的表情来评估其情绪状态,辅助医生进行心理治疗和心理健康管理;在智慧课堂方面,通过识别学生的面部表情来判别学生的学习状态;在疲劳驾驶中,通过识别司机的面部表情,检测司机的精神状态,当司机表现出疲劳时及时提醒司机。人脸表情的识别受到很多因素的影响,比如种族、社交潮流等<sup>[2]</sup>。研究人员通过研究发现人类的面部有几种相同的基本表情,即高兴、悲伤、愤怒、吃惊、厌恶、恐惧<sup>[3]</sup>。由此,研究人员主要以基本表情为研究对象。

早期,数据集的采集以实验室场景为主,训练出的模型对理想环境条件下的人脸表情有较高的识别率,但是对于一些具有复杂场景的图像进行识别时识别率不高,泛化能力不强。因此,人脸表情的研究开始从实验室场景转变到自然场景,而局部遮挡的人脸表情就是复杂场景的主要类别之一。

局部遮挡的人脸表情识别以传统的机器学习、深度学习以及将两者相结合的方式为主,传统的机器学习提取的特征以依靠手工和浅层学习为主,深度学习使用卷积神经网络,通过前向传播、反向传播来提取图片的深层特征信息,传统和深度学习相结合的方法通过将两种特征融合来进行局部遮

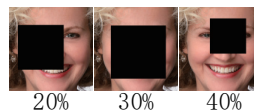
挡人脸表情的识别。尽管机器学习和深度学习对样本特征具有较好的分类效果,但在局部遮挡的人脸表情识别领域仍存在一些问题和挑战。在数据集方面,自然遮挡的人脸表情数据集少,不能提取到足够多的特征信息来提高算法的泛化能力。在特征提取方面,由于局部遮挡导致图片的部分信息丢失,不能提取到完整有效的特征,给识别带来挑战。

## 1 数据集

局部遮挡的数据集采集难度大,同时需要耗费大量的人力物力去对数据集进行人工标注。于是,有研究人员将现有的表情数据集采用黑色矩形遮挡的方法来代替实物遮挡<sup>[4]</sup>。为了更加逼真地模拟现实环境的遮挡物,研究人员除了遮挡与表情相关的眼睛、鼻子、嘴巴外,还对人脸表情按照图片面积的20%、30%、40%进行随机的遮挡<sup>[5]</sup>,处理后的图片如图1所示。还有一些学者通过人工合成遮挡来进行数据集的处理,如图2所示。常用的人脸表情数据集按照数据集名称、遮挡物、分类、大小以及参与人数总结如表1所示。



(a) 眼睛、鼻子、嘴巴的遮挡处理



(b) 随机遮挡

图1 遮挡不同面积的人脸表情图片

1. 陇南师范学院机电工程学院 甘肃陇南 742500

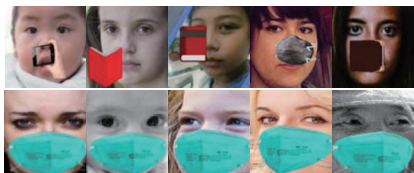


图 2 人工合成的遮挡人脸表情图片

表 1 常用人脸表情数据集

数据集名称	遮挡物	分类 / 类	大小	参与人数 / 人
JAFFE <sup>[6]</sup>	—	7	213 I	10
CK+ <sup>[7]</sup>	—	7	237 I	127
MMI <sup>[8]</sup>	自遮挡	7	2900 V, 740 I	75
AFEW/SFEW <sup>[9]</sup>	眼镜、眼罩、胡须、手、帽子等	7	957 V, 700 I	330
AffectNet <sup>[10]</sup>	头发、光照、手、眼镜、胡须、外界物体、自遮挡等	7	400 000 I	—
FERPlus <sup>[11]</sup>	手、眼镜、头发、侧面自遮挡等	7+ 轻蔑	35 887 I	—
RAF-DB <sup>[12]</sup>	头发、手、眼镜、化妆等	7+ 复合表情	30 000 I	—

注释：1.V 表示视频图像；2.I 表示图片；3. — 表示无或不详

2 局部遮挡人脸表情识别的研究现状

局部遮挡的人脸表情主要分为局部遮挡人脸表情数据的采集、数据集样本的预处理、样本的特征提取、人脸表情的分类识别几个阶段。流程图如图 3 所示。

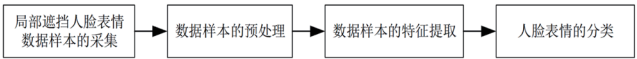


图 3 部遮挡人脸表情的流程

根据研究人员对局部遮挡人脸表情的识别研究，识别算法主要分为传统方法、深度学习方法、传统和深度学习相结合的方法三大类。以下内容将按照三类方法对局部遮挡的人脸表情识别展开综述。

2.1 传统方法的局部遮挡人脸表情识别

为了降低噪声、光照、背景等因素的影响来提取更有效的人脸表情特征，需要对采集到的人脸数据样本进行预处理操作，传统的预处理方法有：归一化<sup>[13]</sup>、人脸区域的框选<sup>[14]</sup>等。

特征提取中，有学者<sup>[15-16]</sup>通过传统的特征提取算法得到纹理特征或者 Gabor 特征，然后将其分割成相等的若干区域，分别对每个区域进行直方图特征的提取，从而得到局部细节的特征，最后通过特征融合进行局部遮挡人脸表情的识别。该方法的局限性在于数据集是通过黑色方块遮挡人脸代替实际的遮挡物，在实际应用中可能泛化能力差。王晓华等

人<sup>[17]</sup>提出一种基于非对称邻域韦伯局部描述子（Asymmetric Region Weber Local Descriptor, AR-WLD）和分块相似度加权的表情识别算法。该方法旨在关注人脸未遮挡区域的细节特征来提高遮挡人脸表情的识别率。有学者通过比对匹配的方法来进行遮挡人脸的表情识别，例如翁存福等人<sup>[18]</sup>通过将遮挡图像平滑降噪处理，然后将与表情识别息息相关的眼睛、鼻子、嘴巴作为参照标准，通过计算遮挡区域与参考标准之间的相似度进行表情识别；张芳艳等人<sup>[19]</sup>建立遮挡物字典和原始人脸的字典，通过结构化遮挡编码来去除遮挡物，然后对人脸表情进行识别，本质上还是通过与原始人脸进行比对来提高有效特征的提取，具体做法如图 4 所示。

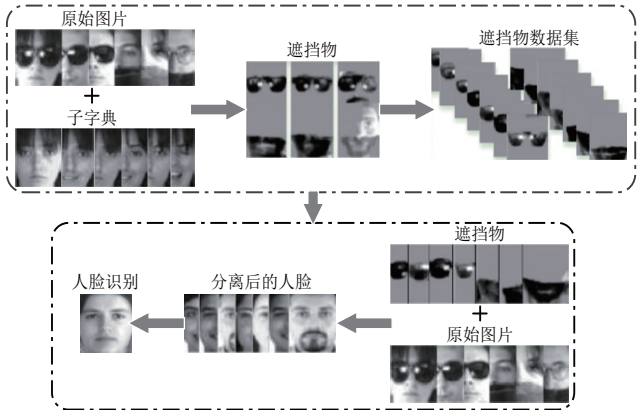


图 4 字典匹配修复的局部遮挡人脸识别模型

依靠原始人脸匹配的方式进行局部遮挡的人脸表情识别受原始字典池的影响，算法的泛化能力差。为了解决上述问题，Zhang 等人<sup>[20]</sup>介绍了一种基于随机 Gabor 模板的方法，用于在面部遮挡的图像中进行面部表情识别。该方法使用 Gabor 滤波器提取面部特征，并使用模板匹配和支持向量机进行分类。为了降低遮挡区域对于人脸表情识别的影响，薛雨丽等人<sup>[21]</sup>提出了一种对面部遮挡具有鲁棒性的表情识别方法，通过鲁棒主成分分析对待识别人脸进行重构，利用重构的区域将遮挡区域进行替换，然后对重构后的人脸表情进行识别。此外，李蕊等人<sup>[22]</sup>提出一种新的基于 Gabor 滤波和灰度共生矩阵的表情识别算法，该算法通过降低图片的维度来降低遮挡物对识别的影响。

通过阅读大量相关文献可以发现，人脸表情的识别主要以遮挡区域的重构或者修复，建立人脸字典和遮挡物字典来匹配或者比对，进行遮挡人脸的表情识别。利用图像修复和恢复技术，对被遮挡的部分进行补全，然后再进行表情识别。这种方法需要克服图像补全过程中可能引入的偏差和误差。匹配法同样在建立匹配的过程中会带入非相关区域的误差，给识别带来影响。总之，传统方法在处理复杂遮挡和非线性变换方面存在一定的局限性，因此近年来更多的研究集中在基于深度学习的方法上。

## 2.2 深度学习的局部遮挡人脸表情识别

深度学习的局部遮挡人脸表情识别在图像的预处理上除了采用归一化、人脸对齐<sup>[23]</sup>,还采用数据增强的方法(翻转、随机剪裁、旋转)来增加数据集样本的多样性<sup>[24]</sup>。在数据增强方面,还有学者使用编码-解码器生成各种表情图像来丰富人脸表情数据<sup>[25]</sup>。

在特征提取方面,为了解决网络训练过拟合、模型泛化能力差的问题,Pan 等人<sup>[26]</sup>使用网络微调的方式进行局部遮挡的人脸表情识别,具体做法是将未遮挡的人脸表情数据作为预训练的样本数据,然后利用预训练后的网络模型对遮挡的人脸表情数据进行训练,最后进行人脸表情的识别。这种方法的局限性在于提取到全局的特征,而没有考虑对人脸表情识别更重要的细节特征。Li 等人<sup>[27]</sup>提出使用关键点来定位面部的各个区域,通过使用注意力机制来给定位的区域分配权重,即给遮挡的区域分配小的权重,给未遮挡的区域分配大的权重,达到自动识别遮挡区域与未遮挡区域的目的。该方法仅仅考虑了局部的有效信息,而没有考虑全局的信息。考虑到这一点,该团队<sup>[28]</sup>又将全局特征引入到网络中,实验结果表明,局部遮挡的细节特征和全局特征相结合的方法能够取得相对较好的识别效果。受此启发,Chen 等人<sup>[29]</sup>提出了一种特征融合残余注意力网络,该网络采用了双分支网络,第一分支利用多尺度模块,而第二分支将提取的特征图分成多个不重叠的局部特征图,然后使用注意力机制进行处理,最终处理后的特征被融合进行分类,具体做法如图 5 所示。

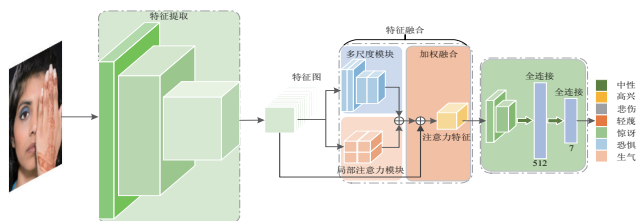


图 5 注意力机制的局部遮挡人脸表情识别模型

王海涌等人<sup>[30]</sup>提出一种基于改进生成式对抗网络的表情识别模型,先利用由自动编码器构成的生成器和两个鉴别器的对抗学习对遮挡人脸图像填补修复,再在全局鉴别器后面添加多分类层,利用全局鉴别器的部分卷积层并在后面添加多分类层构成表情分类器进行表情识别,具体做法如图 6 所示。

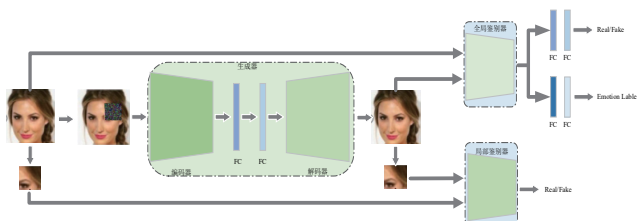


图 6 基于双鉴别网络的遮挡表情识别模型

姚乃明等人<sup>[31]</sup>提出了一种对人脸局部遮挡图像进行用户无关表情识别的方法。该方法包括一个基于 Wasserstein 生成式对抗网络(Wasserstein generative adversarial net, WGAN)的人脸图像生成网络和一个表情识别网络。前者能够为图像中的遮挡区域生成上下文一致的补全图像,后者能够通过表情识别任务和身份识别任务之间建立对抗关系来提取用户无关的表情特征并推断表情类别。奚琰<sup>[32]</sup>提出了一种结合对比学习和通道-空间注意力机制的人脸表情识别方法,学习各局部显著情感特征并关注局部特征与全局特征之间的关系。

深度学习的局部遮挡人脸表情识别主要围绕着人脸遮挡区域的修复和强调未遮挡的区域,修复的方式主要是采用对抗生成式网络或者自编码网络进行遮挡区域的修复;强调未遮挡区域的方法主要采用注意力机制的方式,通过训练网络给未遮挡的区域分配更多的权重来提高未遮挡区域的关注度。利用遮挡修复的方法忽略了局部特征对识别的影响,利用注意力机制的方法强调了局部特征而没有将全局特征考虑进去,因此,将全局特征和局部特征进行融合实现局部-全局的特征互补能够更加有效地提升网络的识别能力。

## 2.3 传统方法与深度学习方法相结合的局部遮挡人脸表情识别

随着遮挡人脸表情研究的深入,研究人员开始将传统方法提取的特征与深度学习提取的特征进行融合。例如,Poux 等人<sup>[33]</sup>通过获取遮挡人脸的光流特征图并与原始人脸相结合,利用自动编码器重建遮挡人脸的光流特征图,通过识别重建的光流特征图来识别人脸的表情。该方法过度依靠原始人脸来修复局部遮挡的光流特征图,遇到不确定身份人脸表情时识别率可能会降低。Dapogny 等人<sup>[34]</sup>提出了置信度加权的局部表达预测。该方法引入随机森林学习人脸局部子空间,并利用分层自编码器网络计算每个局部子空间的置信度。在 CK+ 数据集上,该方法在口腔遮挡下的准确率达到 72.7%,并且对面部遮挡表现出良好的鲁棒性。然而,目前对局部遮挡下面部表情识别的研究还很有限,在准确率上还有很大的提升空间。为了解决泛化能力差的问题,Zheng 等人<sup>[35]</sup>利用局部遮挡人脸来训练深度学习模型 Train-Swin-transformer,利用人脸表情的眼睛和眉毛来训练 k-最近邻(k-nearest neighbors, KNN)模型,最后通过将两者融合来进行遮挡人脸的表情识别。

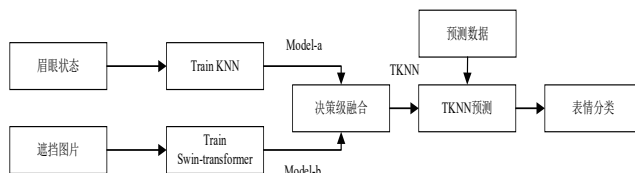


图 7 局部-全局的传统和深度学习结合的遮挡表情识别模型

传统方法与深度学习相结合的局部遮挡人脸表情识别方



法主要有两种：第一种方法是通过遮挡区域的重建来降低遮挡带来的识别误差，重建的主要手段是通过训练稳定的编码-解码器来填补遮挡区域；第二种方法是将提取的全局特征和局部特征进行融合，局部特征提取细节信息，全局特征提取轮廓信息降低遮挡带来的影响。相比于单纯依靠传统学习进行遮挡人脸表情的识别，将传统方法和深度学习相结合的方法能取得更好的识别效果。

局部遮挡人脸表情识别研究虽然已经取得了一定的成果，但仍然有一些亟待解决的问题。

(1) 局部遮挡人脸的数据集样本少，导致算法的泛化能力差。对此，可以采用预训练的方式，首先通过现有的未遮挡人脸数据集进行算法的训练，然后将训练好的模型在目标数据集上训练。

(2) 遮挡的人脸表情复杂，仅靠一种模态很难识别。对此，可以采用多模态进行局部遮挡人脸表情的识别。例如，通过将面部未遮挡区域与肢体动作结合起来进行遮挡人脸的表情识别。

(3) 人脸表情包含的信息复杂，遮挡人脸的微表情识别将更具有挑战性，可以采用改进损失函数的方式来提高算法的判别能力。

(4) 现有的网络模型注重提取人脸样本的特征完整性，忽略了模型在实际应用中的实时性。对此，可以在网络的小型化方面做研究，在保证网络模型有良好性能的同时，也有较好的识别速率。

### 3 结语

首先介绍了局部遮挡人脸表情识别的研究背景，然后介绍了常用的局部遮挡人脸表情数据集以及数据集的遮挡处理方式，接着按照传统方法、深度学习方法、传统方法与深度学习相结合的方法综述局部遮挡的人脸表情识别。通过分析研究可以发现，局部遮挡的人脸表情识别主要使用以下几种方法。(1) 修复遮挡区域。这种方法主要是通过填补人脸遮挡的区域来降低遮挡带来的干扰，修复以编码-解码器为主，其他方法次之。(2) 通过匹配的方式来进行遮挡人脸的识别。其主要内容是先建立原始人脸的字典，然后建立遮挡人脸的字典，最后通过匹配的方式进行识别。这种方式的缺点是对个人的身份要求高。(3) 通过强化未遮挡区域的特征提取，主要方法有两种：一是神经网络引入注意力机制的方式，通过给未遮挡区域分配更多的权重进行识别；二是划分人脸区域建立区域特征直方图，注重未遮挡区域的细节特征，降低遮挡的影响。对于人脸表情识别领域的研究，从数据集的角度，丰富数据的多样性和样本的类别可以有效防止网络训练过程的过拟合，提升网络模型的泛化能力；从算法的角度，

引入注意力机制、多种特征融合和多模态的遮挡人脸表情识别成为该领域研究的有效手段和趋势。

### 参考文献：

- [1] LI S, DENG W. Deep facial expression recognition: a survey[J]. IEEE transactions on affective computing, 2020, 13(3): 1195-215.
- [2] 李珊. 基于深度学习的真实世界人脸表情识别研究[D]. 北京: 北京邮电大学, 2021.
- [3] EKMAN P, MATSUMOTO D, FRIESEN W V. Facial expression in affective disorders [C]//What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS). New York: Oxford University Press, 1997: 331-342.
- [4] 刘帅师. 非均匀光照和局部遮挡情况下的鲁棒表情识别理论与方法研究[D]. 长春: 吉林大学, 2012.
- [5] 卢洋. 人脸表情图像识别关键技术的分析与研究[D]. 长春: 吉林大学, 2019.
- [6] LYONS M, AKAMATSU S, KAMACHI M, et al. Coding facial expressions with gabor wavelets[C]//Automatic Face and Gesture Recognition, 1998. [S.l.]: IEL, 1988: 200-205.
- [7] LUCEY P, COHN J F, KANADE T, et al. The extended cohn-kanade dataset (CK+): A complete dataset for action unit and emotion-specified expression[C]//2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition Workshops, [v.2]. Piscataway: IEEE, 2010: 1680-1687.
- [8] VALSTAR M, PANTIC M. Induced disgust, happiness and surprise: an addition to the mmi facial expression database[C]//Proceedings of the Proc 3rd Intern Workshop on EMOTION (satellite of LREC): Corpora for Research on Emotion and Affect, F, 2010. [S.l.]: [s.n.], 2010: 16083666.
- [9] DHALL A, GOECKE R, LUCEY S, et al. Collecting large, richly annotated facial-expression databases from movies[J]. IEEE multimedia, 2012, 19(3): 34-41.
- [10] 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.
- [11] BARSOUM E, ZHANG C, FERRER C C, et al. Training deep networks for facial expression recognition with crowd-sourced label distribution[C]//Proceedings of the Proceedings of the 18th ACM international conference on multimodal

- interaction. New York: ACM, 2016:279-283.
- [12] LI S, DENG W, DU J. Reliable crowdsourcing and deep locality-preserving learning for expression recognition in the wild[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition, [v.1]. Piscataway: IEEE, 2017:2584-2593.
- [13] 李蕊. 局部遮挡条件下的鲁棒表情识别方法研究与系统实现 [D]. 北京: 北京工业大学, 2015.
- [14] 李春虹, 卢宇. 基于深度可分离卷积的人脸表情识别 [J]. 计算机工程与设计, 2021, 42(5): 1448-1454.
- [15] LIU S, ZHANG Y, LIU K. Facial expression recognition under partial occlusion based on Weber Local Descriptor histogram and decision fusion[C]//2014 33rd Chinese Control Conference. Piscataway: IEEE, 2014: 4664-4668.
- [16] LIU S S, ZHANG Y, LIU K P, et al. Facial expression recognition under partial occlusion based on Gabor multi-orientation features fusion and local Gabor binary pattern histogram sequence[C]//2013 Ninth International Conference on Intelligent Information Hiding and Multimedia Signal Processing. Piscataway: IEEE, 2013: 218-222.
- [17] 王晓华, 陈影, 胡敏, 等. 基于 AR-WLD 和分块相似度加权的遮挡表情识别 [J]. 激光与光电子学进展, 2018, 55(4): 183-190.
- [18] 翁存福, 朱喜顺. 复杂背景下基于 SIFT 算法的局部遮挡人脸识别 [J]. 计算机仿真, 2024, 41(2): 232-236.
- [19] 张芳艳, 王新, 许新征. 基于结构化遮挡编码和极限学习机的局部遮挡人脸识别 [J]. 计算机应用, 2019, 39(10): 2893-2898.
- [20] ZHANG L, TJONDRONEGORO D, CHANDRAN V. Random Gabor based templates for facial expression recognition in images with facial occlusion[J]. Neurocomputing, 2014, 145: 451-464.
- [21] 薛雨丽, 毛峡, CATALIN-DANIEL C, 等. 遮挡条件下的鲁棒表情识别方法 [J]. 北京航空航天大学学报, 2010, 36(4): 429-433.
- [22] 李蕊, 刘鹏宇, 贾克斌. 局部遮挡条件下的人脸表情识别 [J]. 计算机应用与软件, 2016, 33(9): 147-150+175.
- [23] 澹台姝昱. 基于深度学习的遮挡人脸表情识别 [D]. 济南: 山东大学, 2022.
- [24] 王素琴, 高宇豆, 张加其. 基于生成对抗网络的遮挡表情识别 [J]. 计算机应用研究, 2019, 36(10): 3112-3115+3120.
- [25] ZHANG F, ZHANG T, MAO Q, et al. Joint pose and expression modeling for facial expression recognition[C]//2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, [Volume 5 of 13]. Piscataway, N.J.: IEEE, 2018: 3359-3368.
- [26] PAN B, WANG S, XIA B. Occluded facial expression recognition enhanced through privileged information[C]//Proceedings of the Proceedings of the 27th ACM International Conference on Multimedia. New York: ACM, 2019: 566-573.
- [27] LI Y, ZENG J, SHAN S, et al. Patch-gated CNN for occlusion-aware facial expression recognition[C]//2018 24th International Conference on Pattern Recognition, [v.1]. Piscataway: IEEE, 2018: 2209-2214.
- [28] LI Y, ZENG J, SHAN S, et al. Occlusion aware facial expression recognition using CNN with attention mechanism[J]. IEEE transactions on image processing, 2018, 28(5): 2439-2450.
- [29] CHEN Y, LIU S, ZHAO D, et al. Occlusion facial expression recognition based on feature fusion residual attention network[J]. Frontiers in neurorobotics, 2023, 17: 1250706.
- [30] 王海涌, 梁红珠. 基于改进的 GAN 的局部遮挡人脸表情识别 [J]. 计算机工程与应用, 2020, 56(5): 141-146.
- [31] 姚乃明, 郭清沛, 乔逢春, 等. 基于生成式对抗网络的鲁棒人脸表情识别 [J]. 自动化学报, 2018, 44(5): 865-877.
- [32] 奚琰. 基于对比学习的细粒度遮挡人脸表情识别 [J]. 计算机系统应用, 2022, 31(11): 175-183.
- [33] POUX D, ALLAERT B, IHADDADENE N, et al. Dynamic facial expression recognition under partial occlusion with optical flow reconstruction[J]. IEEE transactions on image processing, 2021, 31: 446-457.
- [34] DAPOGNY A, BAILLY K, DUBUISSON S. Confidence-weighted local expression predictions for occlusion handling in expression recognition and action unit detection[J]. International journal of computer vision, 2018, 126: 255-271.
- [35] ZHENG K, TIAN L, LI Z, et al. Incorporating eyebrow and eye state information for facial expression recognition in mask-obscured scenes[J]. Electronic research archive, 2024, 32(4): 2745-2771.

#### 【作者简介】

英文彬 (1991—), 男, 甘肃陇西人, 硕士, 助教, 研究方向: 智能控制与模式识别。

(收稿日期: 2024-05-23)