

面部表情识别技术在智能家居中的应用研究

周 勉¹, 刘静超¹, 杨婧敏¹, 谢宜辰², 左欣垚^{1*}

¹西京学院, 计算机学院, 陕西 西安

²西京学院, 体育与美育教育中心, 陕西 西安

收稿日期: 2024年12月24日; 录用日期: 2025年1月23日; 发布日期: 2025年1月29日

摘 要

随着智能家居技术的发展, 情绪识别(FER)技术的应用日益广泛, 成为提升人机交互体验和个性化服务的重要工具。本文通过综述现有的FER技术及其在智能家居中的应用, 分析了该领域的主要研究进展, 并指出了现有研究中的局限性。文章提出了对FER技术的应用进行系统性归纳和分析的新视角, 特别是在情绪调节与个性化环境适配方面的潜力。通过总结不同研究的成果, 本文明确了当前技术面临的挑战, 并展望了未来的研究方向, 包括多模态数据融合、深度学习优化以及更智能的情绪调节机制。本文的创新性在于通过对现有技术的全面分析, 提出了一些尚未被充分研究的应用方向, 并为未来智能家居领域的研究提供了新的视角和思路。

关键词

面部情绪识别, 智能家居, 深度学习, 人机交互

The Application of Review of Facial Expression Recognition Technology in Smart Homes

Mian Zhou¹, Jinchao Liu¹, Jingmin Yan¹, Yichen Xie², Xinyao Zuo^{1*}

¹School of Computer Science, Xijing University, Xi'an Shaanxi

²Center of Physical and Aesthetic Education, Xijing University, Xi'an Shaanxi

Received: Dec. 24th, 2024; accepted: Jan. 23rd, 2025; published: Jan. 29th, 2025

Abstract

With the development of smart home technologies, emotion recognition (FER) has become a key

*通讯作者。

文章引用: 周勉, 刘静超, 杨婧敏, 谢宜辰, 左欣垚. 面部表情识别技术在智能家居中的应用研究[J]. 计算机科学与应用, 2025, 15(1): 269-278. DOI: 10.12677/csa.2025.151027

tool for enhancing human-computer interaction and providing personalized services. This paper reviews the existing FER technologies and their applications in smart homes, analyzing the major research progress in the field and identifying the limitations in current studies. The paper offers a new perspective on the application of FER, particularly its potential in emotional regulation and personalized environmental adaptation. By synthesizing findings from different studies, this paper highlights the challenges facing current technologies and anticipates future research directions, including multi-modal data fusion, deep learning optimization, and more intelligent emotional regulation mechanisms. The innovation of this paper lies in its comprehensive analysis of existing technologies, proposing underexplored application directions and providing new perspectives and ideas for future research in the smart home domain.

Keywords

Facial Emotion Recognition, Smart Home, Deep Learning, Human-Computer Interaction

Copyright © 2025 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



1. 引言

人工智能(Artificial Intelligence, AI)和物联网(Internet of Things, IoT)的快速发展,推动了智能家居系统在现代生活中的广泛应用。这些系统通过集成传感器、通信网络和自动化控制技术,能够实时调节居家环境,并提供个性化服务。情绪感知技术作为人机交互的核心部分,对于提升用户体验至关重要,是智能家居研究中的一个重要方向。

面部表情识别(Facial Expression Recognition, FER)技术,作为一种基于计算机视觉的情绪识别方法,通过分析面部图像或视频来检测和分类人类的情绪状态。近年来,随着深度学习技术的进步,面部表情识别的准确性和实时性得到了显著提升,其在情绪监测、心理健康评估、娱乐互动等领域的应用潜力逐渐得到重视。将面部表情识别与智能家居系统相结合,不仅可以增强家居设备的情感交互能力,还能根据用户的情绪状态动态调节环境参数(如灯光、温度、音乐等),为用户提供更为智能化和个性化的居住体验。本文将综述面部表情识别技术的基础原理与应用场景,并重点探讨其在智能家居中的潜在应用和未来发展方向。

2. 面部表情识别技术概述

2.1. 技术原理

面部表情识别的技术可分为三个主要环节:人脸检测、特征提取和表情分类。

人脸检测是面部表情识别的前置步骤,旨在从输入图像或视频中定位人脸区域。传统方法包括基于Haar特征的级联分类器和Dlib库,这些方法效率较高,但对光照变化和复杂场景的鲁棒性不足。近年来,基于深度学习的模型(如MTCNN和RetinaFace)通过卷积神经网络实现了更高的精度,尤其在多角度和动态场景下表现优异。

特征提取是将检测到的人脸区域转化为可用于情绪分类的数值特征的过程。常用的传统特征提取方法包括局部二值模式(Local Binary Pattern, LBP) [1],该方法通过计算像素与其邻域像素的灰度差异,生成局部纹理模式。LBP对光照变化具有一定的鲁棒性,广泛应用于低分辨率图像的表情识别。然而,由于

其仅捕捉局部特征, 在处理复杂表情和多样化场景时表现受限; 以及梯度方向直方图(Histogram of Oriented Gradients, HOG) [2], HOG 通过统计图像中梯度方向的分布来捕捉全局形状特征, 适合用于轮廓信息显著的表情识别任务。虽然 HOG 在早期表情识别任务中取得了一定成效, 但其对光照和遮挡的敏感性限制了实际应用效果。特征提取也可以通过深度学习的方法来完成, 神经网络的卷积层能自动提取多层次的特征, 有效捕捉面部细微变化和复杂结构, 从而提高情绪识别的准确性与鲁棒性。

最后是表情分类, 表情分类的目标是根据特征提取的结果将情绪归类为某种情感类别。传统分类器如支持向量机(Support Vector Machine, SVM) [3]和随机森林(Random Forest)因灵活性和效率曾被广泛应用。近年来, 基于深度学习的端到端分类方法取得了显著进展, 尤其在多情绪类别识别任务中表现突出。

2.2. 基于深度学习的表情识别技术

传统的方法(如上述的 Haar 特征、LBP、HOG 和 Gabor 小波[4]等)对图像的低层特征进行分析, 能够处理简单的情绪识别任务, 被广泛使用。然而, 它们对复杂情境(如光照变化、背景干扰等)的适应性较差, 鲁棒性有限[5]。这一阶段的技术为后来的深度学习方法奠定了基础。深度学习技术的快速发展也推动了表情识别的研究进入新阶段: 基于深度学习的表情识别方法具有自动提取高维特征的能力, 显著提升了识别精度与鲁棒性。

(1) 卷积神经网络(Convolutional Neural Network, CNN)

随着卷积神经网络(CNN)在图像分类领域的重大进展, 众多以 CNN 为基础的表情识别技术应运而生, 有效提升了传统方法在鲁棒性方面的不足[6] [7]。针对图像中存在的局部遮挡问题, 一部分研究者巧妙地利用 CNN 构建了无监督学习模型。他们通过编解码网络架构, 从被遮挡的图像中抽取并学习特征编码。在这一基础上, 进一步实现图像合成或变换, 确保在过程中保留了个体的身份信息和表情特征, 最终再进行精准的表情识别[8] [9]。此外, 近年来提出的轻量化模型, 如 MobileNetV3 [10]和 EfficientNet [11], 探索了神经网络在资源受限场景的可能性。通过多层卷积, 神经网络能够自动从图像中提取层次化的特征, 具备更强的鲁棒性和准确性, 能够处理更为复杂的情绪变化、不同个体的面部特征等。然而, 深度神经网络也带来了较高的计算复杂度和对数据的依赖, 这推动了更高效模型的出现。

(2) 多模态融合方法

为了进一步提升情绪识别的准确性和鲁棒性, 研究者开始尝试融合面部表情、语音语调和生理信号(如心率、皮肤电活动)等多种数据源。Stamos 等人[12]建立了基于脑电图和心电图的情绪识别数据库, Gamage 等人[13]提出了使用音素序列来编码与情绪表达相关的语言线索, 将文本信息与语音特征相结合, 从而提高情绪识别的准确率。在此基础上, Xu 等人[14]建立 ESIM (Enhanced Sequential Inference Model) 交互注意力机制, 获取语音和文本模态之间的相似特征, 实现语音和文本模态在时域上的对齐, 得到更加准确的多模态特征表示。融合语音与视觉特征的语音-视觉情绪识别框架(Speech-Vision Emotion Recognition Framework) [15]和融合面部手势和言语的情感强度分析模型[16], 都对复杂情感状态的识别做了一定的探索。多模态融合能够弥补单一模态识别中的缺陷, 尤其在面对遮挡、复杂背景等情况时, 能有效提升识别效果。尽管这种方法的计算和存储开销较大, 但它极大地丰富了情绪识别的维度, 成为近年来的研究热点。

(3) 生成对抗网络(Generative Adversarial Network, GAN)

GAN [17] [18]作为一种生成模型, 最初应用于图像生成领域, 但逐渐在情绪识别中得到了应用[19]。Luo 等[20]提出了一种基于生成对抗网络的域适应方法, 该网络由四个核心组件构成: 源域与目标域的生成器、判别器以及分类器, 其训练流程分为预训练与对抗训练两个阶段。在预训练阶段, 借助两个域生成器的作用, 源域与目标域的数据被映射到一个共同的高维特征空间。进入对抗训练阶段, 则通过对抗

机制来缩小这两个映射之间的距离。这一方法有效解决了跨被试情绪识别模型面临的域偏移难题。此外, Ma 等[21]人基于域对抗网络进一步提出了域残差网络(DResNet), 其结构设计类似于残差网络, 但优势在于它构建了一个域泛化框架, 使得在情感模型训练过程中无需依赖目标域被试的任何信息。Yao 等[22]提出了一种基于 WGAN (Wasserstein generative adversarial net, WGAN)的人脸图像补全算法, 能够以生成方式近似还原被遮挡的人脸图像, 缓解因局部表情信息缺失带来的影响, 提高识别算法的鲁棒性。尽管 GAN 在训练过程中可能出现不稳定性和生成图像质量的限制, 但其在数据增强方面的应用已取得了显著成果。

(4) Transformer 及自注意力机制

Transformer 模型的提出为情绪识别带来了新的发展契机。与传统的 CNN 不同, Transformer 能够更好地处理图像中的长距离依赖, 提升对复杂情绪表达和微表情变化的捕捉能力, 如 Vision Transformer (ViT) [23]和其改进模型 Swin Transformer [24], 通过局部特征的学习在小样本训练上也能达到较高的情绪分类准确率和较快的测试速度。自注意力机制允许模型在处理面部表情时关注图像中的重要区域, 从而提高识别精度。虽然 Transformer 模型具有较高的计算需求, 但其在复杂情绪和微妙变化的识别方面展现了巨大的潜力, 尤其是在与 LSTM、CNN 结合的多模态框架[25]中得到了广泛应用。

(5) 强化学习

强化学习在情绪识别中的应用较为新颖, 主要通过训练模型自我调整识别策略, 优化情绪识别的效果。尽管强化学习的训练时间较长, 并且可能面临收敛性问题, 但实验[26]表明它在动态情境下(如实时情绪识别与用户交互)展现了极大的潜力, 这也在情绪反馈和策略优化方面提供了新的研究方向。

在情绪识别算法的发展过程中, 各种算法类型逐渐成熟, 并在不同应用场景中表现出各自的优势与局限性。为了更好地理解这些算法的特点和发展趋势, 表 1 对主流的情绪识别算法进行了总结与对比。

Table 1. Comparison and analysis of mainstream emotion recognition algorithms

表 1. 主流情绪识别算法对比分析

算法类型	优点	缺点	发展方向
特征提取方法	计算效率高, 实时性强	鲁棒性差, 对环境变化敏感	逐步被深度学习方法取代
CNN	高效的特征自动提取、鲁棒性更强	计算复杂度高, 训练难度大	深度卷积网络和残差网络发展, 性能提升
多模态融合方法	提高准确性, 更真实的情绪识别	计算复杂度高, 数据收集难度大	深度学习结合多模态信息, 提高情绪识别精度
GAN	改善数据不平衡问题, 增强训练数据	训练不稳定, 生成的图像质量有限	在数据增强方面发挥重要作用
Transformer	捕捉长距离依赖, 全局信息建模能力强	计算复杂度高, 数据需求大	应用于情绪识别中的微表情和细节变化识别
强化学习	动态调整策略, 自我学习能力强	训练时间长, 收敛性差	强化学习逐渐应用于实时情绪识别和用户交互优化

此外, 表情识别技术在轻量化、自监督等方面取得了显著进展。针对资源受限的设备, 研究人员通过模型剪枝、量化和知识蒸馏等方法开发了高效的轻量化模型, 使得表情识别技术能够在边缘设备上实现实时性和高精度的平衡。自监督学习和少样本学习的引入则有效缓解了数据稀缺问题, 自监督学习利

用未标注数据生成训练信号，而少样本学习使模型在极少样本条件下快速适应新任务。这些技术进步为表情识别在智能家居中的实际应用奠定了基础，同时也为未来实现更智能化和个性化的情绪感知系统提供了更多可能性。

2.3. 常用公开数据集

表情识别研究离不开高质量的公开数据集，它们为算法训练和评估提供了统一的标准。以下是一些常用的公开数据集及其特点：

(1) FER2013

FER2013 是一个用于面部表情识别的基准数据集，包含 35,887 张灰度图像，分为七种基本表情类别（如快乐、愤怒、悲伤等）。该数据集以其大规模和通用性广泛用于表情识别任务的模型训练和性能比较。

(2) CK+ (Cohn-Kanade Plus)

CK+ 是表情识别领域经典的数据集，包含 593 个面部表情视频序列，涵盖了 123 名被试者的表情变化过程。从中提取的关键帧被标注为七种表情类别，主要用于动态表情分析和序列识别研究。

(3) AffectNet

AffectNet 是目前最大规模的情感表情识别数据集之一，包含超过一百万张带有情感标签的面部图像，标签包括七种基本表情以及中性表情，并额外标注了情绪的情感值(Arousal-Valence)。其数据来源于互联网，具有多样性强、噪声大的特点，是研究深度学习模型鲁棒性的良好资源。

(4) RAF-DB (Real-world Affective Faces Database)

RAF-DB 数据集包含 30,000 多张静态面部图像，专注于现实场景中的表情识别，标签分为七种基本表情和 11 种复合表情。其真实感和多样性为表情识别技术在实际场景的应用提供了支持。

(5) EmotiW (Emotion Recognition in the Wild)

EmotiW 数据集是由 “In the Wild” 系列比赛提供的资源，涵盖视频和音频的多模态情绪识别任务。其场景多样且光照复杂，特别适合于智能家居中非受控环境下表情识别研究。

(6) Oulu-CASIA

Oulu-CASIA 数据集包含 2880 个视频片段，记录了在不同光照条件下的六种基本表情(如快乐、惊讶)。其多样化的光照场景为研究表情识别在不同环境中的表现提供了宝贵的数据支持。

这些数据集各具特色，涵盖了从实验室环境到现实场景的多种应用需求，为表情识别技术的发展提供了重要的基石。在智能家居应用中，根据实际需求选择适合的数据集是构建高效模型的关键步骤。

3. 智能家居中的情绪调节需求

智能家居正在从功能性设备集成向用户情感需求适配的方向迈进的过程中，情绪调节作为改善居住体验、提升生活质量的重要环节，是智能家居研究与应用中的关键目标之一。本部分从情绪调节的目标功能出发，探讨其在智能家居中的具体表现与重要性。

3.1. 情绪调节需求的背景

在现代快节奏社会，工作与生活边界的模糊使得情绪压力不断积累，个体面临的情绪压力和心理负担日益加重。根据《2022 年国民抑郁症蓝皮书》，我国抑郁症患者已超过 9500 万，其中年轻人群体尤为严重，随着情绪健康问题的日益严重，情绪调节已成为社会关注的重要议题。越来越多的人开始寻求在舒适的家庭环境中释放和调节情绪的方式。

智能家居作为提升生活质量的重要手段，正逐步渗透到日常生活中，通过高效地捕捉和理解用户情

绪,实现个性化的情绪关怀。比如,在用户感到疲惫时,智能家居系统可以自动调节室内灯光、播放轻音乐、调节室温等方式,帮助用户缓解压力,恢复情绪平衡。

情绪调节功能的引入不仅能够满足用户对情感需求的高期待,也为智能家居市场带来了新的发展机遇,成为家庭物联网技术发展的重要方向。

3.2. 情绪调节的目标及场景

情绪调节在智能家居中的目标涵盖多个层面,主要包括舒缓压力、提升生活幸福感以及促进家庭情感交流,这些目标可通过具体的调节场景实现。

(1) 光环境调节

灯光与情绪之间有密切的关联。通过情绪感知技术,智能家居动态调整灯光的颜色、亮度和冷暖色调,营造适宜的环境。例如,当用户感到紧张或焦虑时,灯光可以自动切换为柔和的暖色调,以舒缓情绪;当用户情绪愉悦时,则可增加灯光亮度,营造明快的氛围。

(2) 温湿度调节

环境温湿度对情绪的影响显著,智能家居通过监测用户情绪状态,调整室内温湿度。例如,在炎热的夏季,当用户表现出易怒情绪时,智能空调可适当降低温度,以提升用户的舒适度;在寒冷的冬季,则可通过增加暖气或调节湿度,提供更为宜人的居住环境。

(3) 背景音乐与个性化服务

通过识别用户的情绪状态,提供个性化的音乐和服务推荐。例如,当检测到用户感到压力时,可播放舒缓的背景音乐;在情绪低落时,系统可以选择振奋人心的曲目。此外,智能家居还可联合冥想应用、启动按摩椅,甚至触发香薰设备,为用户提供多样化的情绪调节手段。

(4) 促进家庭成员的互动与关怀

智能家居不仅关注个体情绪,也重视家庭成员之间的情感联结。通过监测每位家庭成员的情绪状态,系统可以在检测到某人情绪低落时触发提醒,鼓励其他家庭成员给予关心,从而促进家庭情感的交流与融洽。

(5) 情绪数据的长期记录与反馈

智能家居也需要能够长期记录用户的情绪数据,分析情绪变化模式,为用户提供健康报告或情绪管理建议。例如,当检测到用户长期处于高压状态时,系统可建议其参与放松训练或寻求心理健康支持。

情绪调节需求是智能家居发展的重要目标之一,通过整合面部表情识别技术、物联网设备和个性化服务,智能家居能够更好地满足用户的情感需求。情绪调节不仅提升了家居环境的智能化水平,还为用户提供了更加人性化和关怀的居住体验,这标志着智能家居从“功能性”向“情感适配性”的重要转型。

4. 表情识别在智能家居中的典型研究案例分析

现有的研究主要集中在人机交互、智能监控和情感计算等方面。例如,研究者们探讨了基于面部表情的情绪识别在课堂监控、超市精准营销、刑事审讯等情境下的应用[27]。在这些场景中,通过面部表情的分析,能够评估用户或顾客的情绪变化,从而优化服务或做出适当的应对。然而,情绪驱动的环境调节和个性化服务在智能家居中的研究相对较少,仍处于初步探索阶段。

4.1. 情绪感知与环境调节

一项研究[28]提出了一种基于情绪感知的智能家居自动化系统,通过结合计算机视觉、环境感知与数据处理技术,实现了面部表情的识别与情绪检测。该系统采用了卷积神经网络(CNN)和支持向量机(SVM)

算法,不仅能够实时识别用户面部的情绪变化,还能根据用户的情绪状态调节环境参数从而改善家庭自动化系统的安全性、效率和用户体验。这一研究强调了情绪感知在智能家居中的应用潜力,特别是在提升用户体验方面的优势。

4.2. 多模态情绪识别的集成应用

另一项相关研究则将焦点放在了语音情绪识别技术的集成应用[29]上,提出了 TF-Mix 情绪识别方法。该方法通过音频数据增强和多特征融合策略,提升了情绪识别的准确性。研究者通过结合卷积神经网络(CNN)、长短期记忆网络(LSTM)和 Transformer 架构,开发了多个情绪预测模型,并通过加权平均法生成了一个集成模型(D),其在多个数据集上的表现都具有较高的准确率。该研究展示了语音情绪识别在智能家居系统中的潜力,尤其是在结合语音交互和情绪感知方面的应用。

4.3. 绿色家居与情绪感知

“智能家居 2.0” [30]尝试通过整合植物物联网与情绪感知技术,提升家庭环境的舒适性和健康性。该研究认为,传统的智能家居系统主要关注与家电的互动,而忽略了环境中的绿色植物对人类情绪的影响。研究表明,室内植物有助于改善环境质量,并对身心健康产生积极影响。因此,结合情绪感知技术和绿色家居环境,能够为用户提供更加个性化、健康的生活空间,从而增强智能家居系统的情感交互能力。

尽管已经开展了多项相关研究,并在情绪识别与智能家居结合方面取得了一定的进展,但总体而言,情绪驱动智能家居系统仍处于初步研究阶段。未来的研究应进一步探索如何通过情绪感知技术实现更加精准的环境调节和个性化服务,以提升智能家居系统的用户体验。

5. 面临的挑战与未来展望

5.1. 挑战

尽管面部表情识别技术在智能家居中展现了广阔的应用前景,但在实际应用过程中,仍然面临诸多挑战。要使该技术在智能家居中得到更加广泛和深入的应用,亟需解决以下关键问题:

(1) 实时性与资源优化的平衡

面部表情识别技术的应用对系统的实时性和资源优化提出了双重要求[31]。一方面,系统需要在复杂的居住环境中实时处理和分析用户的表情,以迅速响应情绪变化并及时调整环境配置,这对硬件计算能力和软件处理效率形成了巨大挑战;另一方面,智能家居设备的硬件成本和能耗问题也限制了技术的普及,尤其是在低功耗设备中,如何在有限计算资源下平衡算法性能与硬件需求成为关键难题。因此,提升识别精度与效率,同时优化资源消耗,是未来技术发展的重要方向。

(2) 环境适应性与情境感知的多重挑战

面部表情识别技术在智能家居中的应用面临环境适应性和情境感知的多重挑战。一方面,外部因素如环境光线、摄像头角度,以及用户表情的个体差异(包括文化背景、性别、年龄等)会显著影响识别的准确性。这要求模型具备更强的泛化能力,可以考虑通过数据增强和自适应学习机制提升对复杂环境的鲁棒性。另一方面,情绪识别需结合用户的情境进行动态分析,例如在特殊情况下(如生病或忙碌),模型不仅要准确识别情绪,还需提供适配的个性化响应[32]。为实现这一目标,未来研究需重点发展上下文感知算法,提升系统在不同环境和情境中的适应能力。

(3) 隐私保护与用户接受度

首先,用户的面部数据属于高度敏感信息,这对数据的安全存储与传输机制提出了严苛要求[33]。系

统需确保数据不被滥用或泄露，并遵守相关法律法规。其次，技术的普及还面临用户对隐私、安全性以及技术可靠性的顾虑。在推广过程中，应关注用户心理与社会文化因素，通过提高技术透明度、提供用户控制权，提升社会接受度。这些工作是确保技术普及与信任建立的关键环节。

5.2. 未来展望

面部表情识别技术在智能家居中的应用仍有巨大的发展空间，未来的研究方向可以从以下几个方面进行探索：

(1) 多模态融合

未来，情绪识别将从单一模态转向多模态融合，将面部表情与语音、手势、心率等生理信号相结合。通过引入 Transformer 等新型网络架构以及跨模态的特征提取方法，进一步提升情绪识别的精确度和适应性。这将赋予智能家居更全面的用户情绪感知能力，增强系统的智能化与交互性。

(2) 云边协同与分布式架构

随着物联网技术的发展，智能家居系统正从单一设备模式向多设备互联的分布式架构演化。未来，面部表情识别技术将深度集成于智能家居中的其他设备(如智能音响、智能温控器等)，通过云平台实现数据共享，构建更加智能的情绪识别与调节网络。在此基础上，边缘计算与云平台协同的处理架构需要逐步完善：本地设备可快速完成初步情绪识别，而云端则负责分析用户长期情绪变化趋势，为个性化服务提供数据支持。这种协作模式不仅能显著提升系统的性能与响应速度，还将极大增强其可扩展性和服务质量。

(3) 个性化模型与情境适应

个性化情绪识别模型也是情绪识别的一个发展趋势。通过引入联邦学习等隐私友好的个性化训练方法，智能家居系统可以针对不同用户的表情特点进行模型自适应优化。此外，研究如何根据用户所在情境(如工作、休息、娱乐)动态调整识别逻辑，也是提升用户体验的重要方向。

(4) 隐私优先的技术规范探索

在推动智能家居技术发展的过程中，隐私保护和技术规范的构建应始终以用户需求为核心。未来，透明化算法设计和用户可控的数据管理机制将成为增强用户信任的关键方向。同时，通过深入研究隐私保护算法(如差分隐私)和可信数据交互机制，可以更有效地保障用户数据在采集、存储与传输过程中的安全性。

此外，智能家居行业需要技术、法律和社会多方力量的联合推动，以制定统一的技术标准和法规政策，实现全球范围内伦理与技术的协同发展。通过构建可信赖的技术环境，面部表情识别技术将不仅服务于更广泛的用户群体，还将推动智能家居的持续健康发展。

面部表情识别技术的未来发展将深刻影响智能家居的智能化与个性化进程。尽管仍面临隐私保护、伦理规范和技术稳定性等多方面的挑战，但通过多模态融合、云边协同架构的优化、个性化模型的创新，以及全球范围内标准与规范的协同发展，这一技术将不仅为用户带来更加智能的生活体验，也将推动智能家居行业迈向新的高度。

未来，智能家居或许能够真正成为用户的“情绪管家”，通过更加精准的情绪识别和实时调节，为每一位用户提供温暖、智能、安全的家居环境。

基金项目

本研究得到西安市科技计划软科学研究一般项目(项目编号：24RKYJ0065)和陕西省大学生创新创业训练计划项目(项目编号：202412715013)的支持。

参考文献

- [1] Ghimire, D., Park, S.H. and Kim, M.J. (2017) Facial Expression Recognition Using Local Region Specific Dense Optical Flow and LBP Features. *Advanced Concepts for Intelligent Vision Systems*, Antwerp, 18-21 September 2017, 28-38. https://doi.org/10.1007/978-3-319-70353-4_3
- [2] Zhang, Y. and Ji, Q. (2005) Active and Dynamic Information Fusion for Facial Expression Understanding from Image Sequences. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **27**, 699-714. <https://doi.org/10.1109/tpami.2005.93>
- [3] Schuldt, C., Laptev, I. and Caputo, B. (2004) Recognizing Human Actions: A Local SVM Approach. *Proceedings of the 17th International Conference on Pattern Recognition*, 2004. *ICPR 2004*, Cambridge, 26 August 2004, 32-36. <https://doi.org/10.1109/icpr.2004.1334462>
- [4] Zhang, Z., Lyons, M., Schuster, M., et al. (1998) Comparison between Geometry-Based and Gabor Wavelets-Based Facial Expression Recognition Using Multi-Layer Perceptron. *Proceedings Third IEEE International Conference on Automatic Face and Gesture Recognition*, Nara, 14-16 April 1998, 454-459. <https://doi.org/10.1109/AFGR.1998.670990>
- [5] 徐峰, 张军平. 人脸微表情识别综述[J]. *自动化学报*, 2017, 43(3): 333-348.
- [6] Mollahosseini, A., Chan, D. and Mahoor, H.M. (2015) Going Deeper in Facial Expression Recognition Using Deep Neural Networks. 2016 *IEEE Winter Conference on Applications of Computer Vision (WACV)*, Lake Placid, 7-10 March 2016, 1-10. <https://doi.org/10.1109/WACV.2016.7477450>
- [7] 孙晓, 潘汀, 任福继. 基于 ROI-KNN 卷积神经网络的面部表情识别[J]. *自动化学报*, 2016, 42(6): 883-891.
- [8] Yim, J., Jung, H., Yoo, B., Choi, C., Park, D. and Kim, J. (2015) Rotating Your Face Using Multi-Task Deep Neural Network. 2015 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, 7-12 June 2015, 676-684. <https://doi.org/10.1109/cvpr.2015.7298667>
- [9] Zhang, J., Kan, M., Shan, S. and Chen, X. (2016) Occlusion-Free Face Alignment: Deep Regression Networks Coupled with De-Corrupt Autoencoders. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, 27-30 June 2016, 3428-3437. <https://doi.org/10.1109/cvpr.2016.373>
- [10] Howard, A., Sandler, M., Chen, B., Wang, W., Chen, L., Tan, M., et al. (2019) Searching for MobileNetV3. 2019 *IEEE/CVF International Conference on Computer Vision (ICCV)*, Seoul, 27 October-2 November 2019, 1314-1324. <https://doi.org/10.1109/iccv.2019.00140>
- [11] Tan, M. and Quoc V. Le. (2019) EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *Proceedings of the 36th International Conference on Machine Learning*, Long Beach, 9-15 June 2019, 6105-6114.
- [12] Katsigiannis, S. and Ramzan, N. (2018) DREAMER: A Database for Emotion Recognition through EEG and ECG Signals from Wireless Low-Cost off-the-Shelf Devices. *IEEE Journal of Biomedical and Health Informatics*, **22**, 98-107. <https://doi.org/10.1109/jbhi.2017.2688239>
- [13] Gamage, K.W., Sethu, V. and Ambikairajah, E. (2017) Saliency Based Lexical Features for Emotion Recognition. 2017 *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, New Orleans, 5-9 March 2017, 5830-5834. <https://doi.org/10.1109/icassp.2017.7953274>
- [14] 徐志京, 高姗. 基于 Transformer-ESIM 注意力机制的多模态情绪识别[J]. *计算机工程与应用*, 2022, 58(10): 132-138.
- [15] Akinpelu, S., Viriri, S. and Adegun, A. (2024) An Enhanced Speech Emotion Recognition Using Vision Transformer. *Scientific Reports*, **14**, Article No. 13126. <https://doi.org/10.1038/s41598-024-63776-4>
- [16] Zadeh, A., Zellers, R., Pincus, E. and Morency, L. (2016) Multimodal Sentiment Intensity Analysis in Videos: Facial Gestures and Verbal Messages. *IEEE Intelligent Systems*, **31**, 82-88. <https://doi.org/10.1109/mis.2016.94>
- [17] Wang, C. (2022) Improved Generative Adversarial Networks for Student Classroom Facial Expression Recognition. *Scientific Programming*, **2022**, Article ID: 9471334. <https://doi.org/10.1155/2022/9471334>
- [18] Yang, H., Ciftci, U. and Yin, L. (2018) Facial Expression Recognition by De-Expression Residue Learning. 2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, 18-23 June 2018, 2168-2177. <https://doi.org/10.1109/cvpr.2018.00231>
- [19] 权学良, 曾志刚, 蒋建华, 等. 基于生理信号的情感计算研究综述[J]. *自动化学报*, 2021, 47(8): 1769-1784.
- [20] Luo, Y., Zhang, S., Zheng, W. and Lu, B. (2018) WGAN Domain Adaptation for EEG-Based Emotion Recognition. *Neural Information Processing*, Siem Reap, 13-16 December 2018, 275-286. https://doi.org/10.1007/978-3-030-04221-9_25
- [21] Ma, B.Q., Li, H., Zheng, W.L. and Lv, B.L. (2019) Reducing the Subject Variability of EEG Signals with Adversarial Domain Generalization. *Neural Information Processing*, Sydney, 12-15 December 2019, 30-42. https://doi.org/10.1007/978-3-030-36708-4_3

-
- [22] 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.
- [23] Dosovitskiy, A., Beyer, L., Kolesnikov, A., *et al.* (2020) An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. arXiv: 2010.11929.
- [24] 陈宗楠, 金家瑞, 潘家辉. 基于 Swin Transformer 的四维脑电情绪识别[J]. 计算机技术与发展, 2023, 33(12): 178-184.
- [25] Wani, T.M., Gunawan, T.S., Qadri, S.A.A., Kartiwi, M. and Ambikairajah, E. (2021) A Comprehensive Review of Speech Emotion Recognition Systems. *IEEE Access*, **9**, 47795-47814. <https://doi.org/10.1109/access.2021.3068045>
- [26] 杨亚飞. 基于深度学习的自然场景下人脸表情识别方法研究[D]: [硕士学位论文]. 成都: 电子科技大学, 2023.
- [27] Khare, S.K., Blanes-Vidal, V., Nadimi, E.S. and Acharya, U.R. (2024) Emotion Recognition and Artificial Intelligence: A Systematic Review (2014-2023) and Research Recommendations. *Information Fusion*, **102**, Article 102019. <https://doi.org/10.1016/j.inffus.2023.102019>
- [28] Patil, V.K., Hadawale, O., Pawar, V.R. and Gijre, M. (2021) Emotion Linked AioT Based Cognitive Home Automation System with Sensovisual Method. 2021 *IEEE Pune Section International Conference (PuneCon)*, Pune, 16-19 December 2021, 1-7. <https://doi.org/10.1109/punecon52575.2021.9686498>
- [29] Wang, M., Ma, H., Wang, Y. and Sun, X. (2024) Design of Smart Home System Speech Emotion Recognition Model Based on Ensemble Deep Learning and Feature Fusion. *Applied Acoustics*, **218**, Article 109886. <https://doi.org/10.1016/j.apacoust.2024.109886>
- [30] Chen, M., Yang, J., Zhu, X., Wang, X., Liu, M. and Song, J. (2017) Smart Home 2.0: Innovative Smart Home System Powered by Botanical IoT and Emotion Detection. *Mobile Networks and Applications*, **22**, 1159-1169. <https://doi.org/10.1007/s11036-017-0866-1>
- [31] Yang, J., Qian, T., Zhang, F. and Khan, S.U. (2021) Real-Time Facial Expression Recognition Based on Edge Computing. *IEEE Access*, **9**, 76178-76190. <https://doi.org/10.1109/access.2021.3082641>
- [32] Zhao, S., Gholaminejad, A., Ding, G., Gao, Y., Han, J. and Keutzer, K. (2019) Personalized Emotion Recognition by Personality-Aware High-Order Learning of Physiological Signals. *ACM Transactions on Multimedia Computing, Communications, and Applications*, **15**, 1-18. <https://doi.org/10.1145/3233184>
- [33] Beltrán, M. and Calvo, M. (2023) A Privacy Threat Model for Identity Verification Based on Facial Recognition. *Computers & Security*, **132**, Article 103324. <https://doi.org/10.1016/j.cose.2023.103324>