https://doi.org/10.26599/AIR.2024.9150040 CAAI Artificial Intelligence Research

Artificial Intelligence in Emotion Quantification : A Prospective Overview

Feng Liu[1] ✉

ABSTRACT

The field of Artificial Intelligence (AI) is witnessing a rapid evolution in the field of emotion quantification. New possibilities for understanding and parsing human emotions are emerging from advances in this technology. Multi-modal data sources, including facial expressions, speech, text, gestures, and physiological signals, are combined with machine learning and deep learning methods in modern emotion recognition systems. These systems achieve accurate recognition of emotional states in a wide range of complex environments. This paper provides a comprehensive overview of research advances in multi-modal emotion recognition techniques. This serves as a foundation for an in-depth discussion combining the field of AI with the quantification of emotion, a focus of attention in the field of psychology. It also explores the privacy and ethical issues faced during the processing and analysis of emotion data, and the implications of these challenges for future research directions. In conclusion, the objective of this paper is to adopt a forward-looking perspective on the development trajectory of AI in the field of emotion quantification, and also point out the potential value of emotion quantification research in a number of areas, including emotion quantification platforms and tools, computational psychology, and computational psychiatry.

KEYWORDS

affect quantification; multi-modal emotion recognition; computational affection; computational psychiatry

Emotion recognition technology, especially multi-modal emotion recognition, has become an important branch of research in the field of Artificial Intelligence (AI). With the increase in computing power and the advancement of algorithms, emotion recognition technology continues to mature, and application areas are expanding from laboratory research to actual commercial and social applications[1]. The core challenge of emotion recognition is to accurately identify human emotional states from complex and diverse data, which requires not only strong technical support, but also a deep understanding of human emotional expression and change[2]. Machine perception and analysis of human emotional states can significantly improve the human-computer interaction capabilities of robots or computer systems. multi-modal emotion recognition has received widespread attention[3] because it fuses multiple perceptual channels such as visual, auditory, and haptic, providing a more comprehensive analysis of emotion than a single modality.

In this paper, we focus on emotion recognition techniques, specifically the analysis and recognition of emotional states from information collected from different data sources, including facial emotion recognition, voice emotion recognition, gesture recognition, multi-modal emotion recognition, and emotion recognition techniques based on other physiological signals. These different data sources are able to provide complementary information about an individual's emotional state, which greatly improves the accuracy of emotion recognition.

In addition, this paper takes an AI perspective on emotion quantification research and present the results of current research in emotion quantification and classification[4], emotion change feature extraction, quantitative analysis of personality traits[4], and quantitative interventions for mental health.

Meanwhile, this paper also explores emotion quantification research in non-AI fields such as automation, physiology, and psychiatry[4, 5], highlighting the urgency of developing efficient emotion quantification tools and methods. These studies have not only demonstrated the potential of AI technology in emotion quantification, but also validated its significant value in mental health-based emotion regulation and intervention.

Finally, based on the technological path of emotion quantification, this paper looks forward to the development direction of computational psychiatry in the future, exploring how to build accurate computational models and make practical progress in the field of quantitative identification of emotions such as depression, loneliness, anxiety, and stress, which will ultimately facilitate the monitoring, diagnosis, and intervention treatment of human mental health.

1 Advance in Emotion Recognition Technology Research

Emotion recognition technology is particularly important because of its central role in human-computer interaction[6]. This technology enables machines to understand and respond to human emotions, greatly improving the user experience, and has played a key role in education[7, 8], healthcare[9], care services[10], and other areas. In addition, emotion recognition has shown its potentially significant value in monitoring mental health and providing emotional support.

With the rapid development of artificial intelligence, emotion recognition technology has evolved from preliminary expression recognition to complex multi-modal recognition systems. These

1 School of Computer Science and Technology, East China Normal University, Shanghai 200062, China
Address correspondence to Feng Liu, lsttoy@163.com

systems comprehensively determine an individual's emotional state by integrating multiple data sources including speech, text, facial expressions and physiological signals. A variety of algorithms and models have been applied to emotion recognition, such as support vector machines, neural networks and deep learning, and these techniques have demonstrated high accuracy in laboratory environments, but still face challenges in real-world applications. In facial expression recognition, deep learning, particularly deep Convolutional Neural Networks (CNNs)[11], has been widely used to extract features from complex facial images to effectively detect and classify different emotional expressions. These models rely on large amounts of image data for training and are able to accurately predict emotional states. Natural Language Processing (NLP) techniques are equally important in textual sentiment analysis, which assesses the author's emotions by analysing the emotional colors in social media text[12], including emotion classification (positive, negative, neutral) and intensity measurement. In addition, machine learning techniques excel in physiological data analysis to identify physiological markers of emotional arousal by analysing physiological indicators such as heart rate variability[13], electrical skin response[14], etc., and to comprehensively assess the emotional state by integrating a wide range of physiological signals.

With the rapid development of emotion computing technology, emotion recognition technology has evolved from simple expression recognition to complex multi-modal recognition systems.

These systems comprehensively determine an individual's emotional state by integrating various data sources such as speech, text, facial expressions and physiological signals as shown in Fig. 1. A variety of algorithms and models have been applied to mainstream affective computing for emotion recognition, such as support vector machines, neural networks, deep learning, etc., and these techniques have demonstrated high accuracy in laboratory environments, but still face challenges in real-world applications.

1.1 Expression emotion recognition technology

In the field of expression emotion recognition techniques, the recognition and generation of facial expressions, micro-expressions and dynamic expressions are key ways of understanding and analysing human emotions. These techniques range from the simple capture of static expressions to the continuous monitoring of dynamic expression changes, demonstrating the wide range of applications of artificial intelligence in complex emotion analysis tasks.

Facial Expression Recognition (FER). In the field of FER techniques, the recognition and generation of facial expressions, micro-expressions, and dynamic expressions are key tools for understanding and analysing human emotions. These techniques not only range from simple capture of static expressions to continuous monitoring of dynamic expression changes, but also demonstrate the wide range of applications of artificial intelligence in complex emotion analysis tasks. Facial expression recognition techniques detect and classify changes in a person's expression by analysing facial images to infer their emotional state. Deep learning, in particular convolutional neural networks, is widely used in this field due to its excellent performance in image recognition tasks. Deep learning has been shown to be effective in recognising emotions from "neutral" to "fear" and from "happiness" to "pain". Deep learning has been shown to be effective in recognising emotions from "neutral" to "fear" and from "pleasure" to "pain" in order to accurately interpret complex emotional states[15].

Facial Micro-expression Recognition (FMiR) focuses on brief and minute facial expressions[16], which usually occur when a person is trying to hide his or her true emotions. Current research

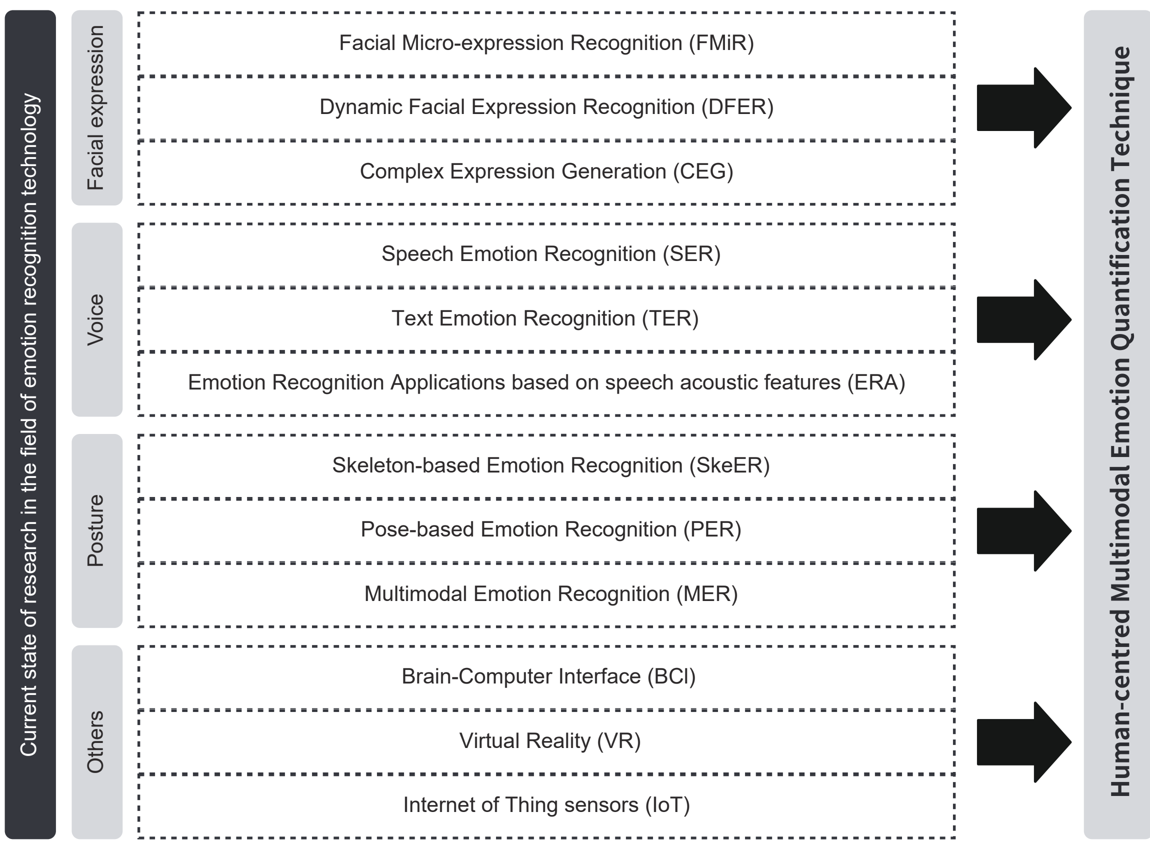

Fig. 1 Dominant technological directions in mainstream affective computing.

is exploring how to capture these brief and subtle facial movements[17] using high-speed cameras and fine image processing techniques, combined with deep learning techniques[18] to capture fine features that provide a more realistic analysis of emotions.

Dynamic Facial Expression Recognition (DFER) focuses on changes in facial expressions over time, which includes not only static facial information analysis, but also invalid frame removal during dynamic expressions[19], contextual relationships between intra- and inter-pair frames[20], static appearance of expressions and dynamic movements, information mining[21], etc. The potential value of dynamic expression recognition techniques is enormous. For example, the ability to extract facial emotional features from video frames in real time and effectively predict student engagement and emotional state using a single efficient neural network based method developed by Savchenko et al.[22] demonstrates the practical application of this technique in educational technology.

Complex Expression Generation (CEG) techniques are rapidly evolving through deep learning approaches and Large Language Models (LLMs). Deep learning frameworks such as Generative Adversarial Networks (GANs) and Variational Autocoders (VAEs) excel at recognising and generating dynamic and micro-expressions, e.g., the GAN model evoGAN[23], which incorporates optimization algorithms, exemplifies the potential of generating complex quantified expressions. In addition, LLMs such as SORA, which combines visual and textual data in multi-modal application scenarios, have been able to generate facial expressions that are closely linked to verbal commands, greatly enhancing the realism and interactivity of virtual characters[24]. These techniques are particularly important for areas such as video games, virtual reality, and emotional health diagnosis.

Despite considerable advances in the techniques used to recognize facial expressions, the accuracy of this process remains a challenge due to the presence of blurred expressions and poor image quality. To illustrate, the self-correcting network proposed by Wang et al.[25] enhances the quality of training samples through a self-attention mechanism and a relabelling mechanism, which effectively mitigates uncertainty in the recognition process. Furthermore, the adaptive correlation loss proposed by Fard and Mahoor[26] enhances the efficacy of expression recognition in field settings by optimizing the correlation between feature vectors of analogous samples. With regard to the implementation of technology, Zhang et al.[27] put forth a novel federated learning framework for unsupervised multi-domain face recognition after motion. This framework employs adversarial learning to enhance the recognition efficacy of the federated framework across diverse domains. The Medjden team's[28] adaptive user interface design through RGB-D sensor integration effectively detects and responds to the user's emotional state, thereby improving the personalization and usability of the user interface. This plays a key role in enhancing the user experience. These advances not only demonstrate technological innovation, but also provide new directions for the practical application of emotion recognition.

Overall, the development of expression recognition technology not only promotes in-depth research on AI in the field of emotion recognition, but also provides practical tools for a variety of application areas, such as mental health assessment, education, and customer service. With in-depth research and continuous improvement of the technology, future emotion recognition technology will be more accurate, intelligent, and sensitive, providing more powerful support for understanding and responding to human emotions.

1.2 Speech emotion recognition technology

Speech emotion recognition technology occupies an important position in the current field of artificial intelligence, and analysis technology based on speech acoustic features has become an important means of identifying human emotional states. Recent research results have highlighted the effective application of acoustic modeling, deep learning, and NLP in emotion recognition and analysis.

Audio Emotion Recognition (AER) technology has significantly improved the accuracy of emotion analysis by integrating fuzzy logic with NLP. This technology effectively detects and responds to the user's emotional state, enhancing the fluidity and naturalness of human-computer interaction[29]. The use of mixed emotion models for emotional speech conversion not only improves the naturalness and realism of emotion recognition[30], but also demonstrates the potential to simulate complex emotional responses in scenarios without direct human involvement[31]. In addition, the use of emotional feature detection techniques to identify deepfake speech enhances the security of the system[32]. The development of emotional speech databases has greatly enhanced the ability of speech generation systems to control emotional dimensions, which is crucial for improving the accuracy and effectiveness of speech systems in processing emotional data[33].

Text Emotion Recognition (TER). In the field of TER, this technology plays an irreplaceable role in Natural Language Understanding (NLU). By analysing user behavior on social media, researchers are able to identify and quantify emotional features[12]. In addition, a multi-modal sentiment analysis model based on interactive attention mechanisms can combine text, audio, and video data to significantly improve the accuracy of sentiment recognition[34]. By using a context-based sentiment recognition model, the researchers demonstrated the potential of applying NLU techniques in processing multilingual texts[35]. Meanwhile, a hybrid sentiment detection approach by combining contextual and semantic analysis aims to improve the accuracy and complexity of sentiment classification[36]. In addition, by comparing different machine learning algorithms and logistic regression models optimized by genetic algorithms, the researchers improved sentiment detection methods in text and searched for the most effective method[37].

Application of emotion recognition based on speech acoustic features. Emotion recognition based on speech acoustic features is an important research direction in the field of artificial intelligence, where these techniques use advanced acoustic modeling, deep learning, and natural language processing techniques to achieve accurate recognition of emotional states and generation of emotionally driven responses. A comprehensive review of this field[38] identifies the applications of acoustic features in emotion recognition and delineates future research directions and challenges.

In addition, a new method combining deep convolutional neural networks and acoustic features[39] has achieved more than 93% accuracy in emotion recognition, demonstrating the potential of these techniques for real-world applications. Ahmed et al.[40] used deep learning techniques for speech emotion recognition to improve the system's ability to be used in real-world applications, especially in human-computer interactions with naturalness and fluency. The review by Lope and Graña[41] highlights the critical role of acoustic features in understanding human emotions and explores the limitations of current techniques and possible future research directions. Daneshfar and Jamshidi[42] explored the potential and adaptability of these techniques through the use of nonlinear echo state networks in a meta-universe environment using an octonionic number base for speech emotion recognition, demonstrating the promise of emotion recognition techniques for complex environments. Carvalho et al.[43] also explored approaches to sentiment and emotion analysis using IBM's natural language processing technologies, highlighting the potential of these techniques for parsing sentiment expressions in social media and other textual data. In addition, Zhang's[44] research focuses on using deep learning to process sentiment analysis in Japanese, improving the accuracy and semantic understanding of sentiment recognition.

Speech emotion recognition technology is not only evolving towards more accurate, comprehensive, and richer semantic understanding, but also providing powerful technical support for mental health assessment and user experience optimization.

By combining acoustic models and advanced computational methods, researchers are able to extract key emotional information from speech, providing a strong technical foundation for real-time emotional interaction and emotion-driven human-machine interface development.

1.3 Gesture emotion recognition technology

In the development of gesture emotion recognition technology, significant technological advances in recent years are particularly reflected in the comprehensive use of multi-source data, which greatly improves the accuracy and efficiency of emotion analysis. The application of deep learning models has enabled emotion patterns from upper body movements to facial expressions to be effectively integrated and recognized[45], providing an effective means of emotion and gesture recognition through facial features and dynamic motion vectors.

Novel skeleton recognition frameworks, such as SAGN[46], using advanced graphical network structures, are able to deeply analyze human skeleton movements, which not only optimizes the recognition process of emotions and behaviors, but is also particularly suitable for processing dynamic visual data. The proposed attention-enhanced semantic-guided graph convolutional network AeS-GCN[47], further improves the efficiency of extracting emotion features from multi-source data and simplifies the complexity of data processing, thus significantly improving the adaptability and accuracy of emotion recognition systems.

The gesture emotion recognition technique also introduces a two-stage multi-modal approach[48], which effectively acquires face trajectories and character trajectories through face tracking and skeleton-based character tracking techniques, and then co-trains the multi-modal features in the same semantic space, which significantly improves the effect of emotion recognition.

By 2023, de Lope and Graña[41] discussed in detail the limitations of current technology and looked at future research directions, in particular how gesture data can be used more efficiently to improve the accuracy and naturalness of emotion recognition. They emphasized that human gesture data provides important clues about human emotional states, which is crucial for the development of more efficient emotion recognition systems.

Daneshfar and Jamshidi's research[42] explores the potential of the gesture-based emotion recognition technique for application in meta-universe environments, and their research demonstrates the adaptability of this technique in complex environments, particularly in simulating complex human-computer interaction situations, through a non-linear echo state network.

Starting from the traditional skeleton recognition, the gesture-based emotion recognition technique has gradually developed techniques such as hemi-recognition and multi-modal fusion recognition, which enhances the effect and also highlight its potential application value in improving the human-computer interaction experience, and provide a new direction and strong technical support for the future research and application of emotion recognition.

1.4 Multi-modal sentiment recognition techniques

In the field of multi-modal sentiment analysis, recent technological advances have demonstrated how innovative algorithms and model optimizations can effectively improve the accuracy and interpretability of sentiment recognition. Current state-of-the-art research has highlighted the use of pre-trained models and advanced neural network architectures in the integration of audio, visual and textual data to improve the performance of sentiment recognition systems. In particular, the Explanatory Multi-modal Emotion Reasoning (EMER) method[49] not only predicts emotions but also explains the motivations behind these predictions, providing a new perspective on open-set emotion recognition and significantly improving the reliability and explanatory power of the system.

In addition, the kernel-typical correlation analysis technique based on *K*-means clustering[50] improves the intermodal heterogeneity in human-computer interaction by accurately fusing multi-modal features, thus increasing the accuracy of emotion recognition. The combination of a multi-channel weight-sharing self-encoder and a cascaded multi-head attention mechanism[51] effectively solves the heterogeneity problem in multi-modal emotion recognition and improves the recognition performance on a public dataset. The feature decoupling method for multi-modal emotion recognition[52] enhances modal independence and reduces distributional differences by mapping features to modality-invariant and modality-specific subspaces.

In terms of fusing data and techniques from different modalities, the introduction of assistive tasks has been shown to more accurately capture and align emotionally relevant features, thus optimising the prediction of affective states[53]. In particular, significant progress has been made in emotion recognition by combining Wav2Vec2.0 and pre-trained models such as Bert on audio and text data. This approach not only improves the accuracy of sentiment analysis through a multi-layered multi-head attention mechanism, but also guides the network to capture key sentiment information through auxiliary tasks, as demonstrated in the results presented at ICASSP 2023[53].

Meanwhile, the application of learning algorithms in the analysis of multi-modal data is very promising, which not only improves the depth of the methodology, but also broadens the scope of applications for sentiment recognition[54]. The performance of multi-modal feature fusion in human-computer interaction has been effectively enhanced by the application of advanced techniques in image processing and behavioral modeling, such as *K*-means clustering and kernel-type correlation analysis[50]. The further development of these techniques, especially real-time emotion expression techniques in humanoid robots, provides new perspectives and tools for the design of future emotion recognition systems.

multi-modal emotion recognition technology has become one of the hottest areas in current affective computing, fully demonstrating the possibility of achieving accurate recognition of emotional states through advanced computational models and algorithms. It provides a technical foundation for further emotion quantification of multidimensional feature acquisition and analysis in future practical applications.

1.5 Other emotion recognition technologies

Brain-Computer Interface (BCI)-based emotion recognition techniques. The application of BCI technology to emotion recognition has made significant progress in recent years, particularly in the accurate and reliable detection of emotional states. Using advanced signal processing techniques and deep learning methods, researchers have been able to effectively integrate electronic brainwave (Electroencephalography (EEG)) signals and significantly enhance EEG-based emotion recognition through improved feature extraction and classification algorithms, such as discrete wavelet decomposition[55]. In addition, the hybrid BCI system combines P300 potentials and emotion patterns to improve the performance of emotion and conscious state detection in patients with disorders of consciousness[56].

In addition, Virtual Reality (VR) technology has emerged as an effective tool for inducing and detecting emotions in emotion research, especially when combined with neurophysiological measures such as EEG, providing new perspectives for a deeper understanding of emotional responses. Studies have shown that VR environments can effectively induce emotions by simulating immersive scenarios, while accurately monitoring the user's emotional state through physiological feedback mechanisms such as EEG[55]. For example, an adaptive VR scene generation system driven by EEG data can adjust the content according to the user's emotional response to optimize the effect of the psychological intervention. In addition, continuous physiological and behavioral emotion annotation datasets provide support for 360-degree VR video, improving the accuracy of emotion recognition and the evaluation of immersive experiences[57]. Specific examples include an emotionally adaptive VR experience for older adults, which enhances the user experience by dynamically adjusting environmental settings to alleviate stress and negative emotions[58]. Meanwhile, the flexible VR platform adjusts task difficulty and skill level to accurately induce and recognize specific emotions, providing a precisely controlled environment for emotion recognition. These applications not only improve the methodology of emotion research, but also extend the potential of VR technology in mental health interventions and user experience enhancement.

The development of these technologies not only provides new tools for understanding human emotion processing mechanisms, but also promotes the application of BCI in areas such as mental health management and psychotherapy. Meanwhile, VR as a technological tool has demonstrated its unique advantages and prospects for wide application in the field of emotion recognition and management, especially when combined with advanced signal processing and machine learning techniques that can provide more accurate data support for emotion recognition.

Finally, research into emotion recognition in multi-modal physiological measures continues to demonstrate its far-reaching impact, particularly in the ability to fuse different sensor data to accurately quantify human emotions. By combining electrical skin responses and electrocardiogram signals, researchers have successfully developed interpretative feature selection methods that have demonstrated high effectiveness in distinguishing between emotional states on the activation and pleasure dimensions, effectively improving the accuracy of emotion classification[14]. Meanwhile, the construction of the multi-modal emotion recognition platform further improves the data synchronization from EEG and eye movement technology to expression monitoring, and achieves the precise fusion of the three signal types based on EEG, thus optimizing the accuracy and efficiency of emotion recognition[59].

Privacy preserving stress detection systems have also made important advances in this area, using physiological data for effective emotion and stress detection through federated learning algorithms without violating user privacy[60]. In addition, the psychophysiological analysis approach has been widely applied to multi-modal systems, providing new tools and perspectives for emotion research by assessing the user's emotional state during task performance through psychophysiological feedback devices[61]. The hybrid multi-modal emotion recognition framework significantly improves the performance of emotion recognition by combining generalized hybrid functions and multi-view learning methods, effectively handling the heterogeneity of emotion data, further confirming the value of this technique for application in user experience evaluation[62].

Despite the considerable benefits that affective computing offers in terms of privacy protection, multi-modal analysis and hybrid framework applications, the field is confronted with four significant challenges. These relate to the psychological complexity of human beings, the diversity of cultures, and the necessity for further refinement in order to facilitate practical applications.

Firstly, current affective computing techniques lack the capacity to enhance the understanding of the interaction between emotions and personality traits. Given the intricate interconnections between emotions and personality characteristics, and the constraints of existing models in capturing the nuances of emotional responses, future AI emotion quantification techniques could integrate more comprehensive psychological theories. Secondly, the existing techniques for computing emotions are not yet adapted to cross-cultural characteristics. In order to address the cultural dependence of emotion models, it would be beneficial for researchers to conduct extensive cross-cultural data collection and analysis in order to construct more inclusive and universal emotion recognition models. Furthermore, existing affective computing technologies should also consider the fine-grained modeling of personalization and individual differences.

Finally, in consideration of the individual differences in emotional experience, future model development should focus on personalization. This can be achieved by utilising big data and reinforcement learning techniques to provide customized sentiment analysis services for each user.

2 Research in AI-Based Emotion Quantification

The importance of emotion quantification is particularly prominent in the field of psychology, as it relates to understanding and measuring the nuances of human emotions, which is crucial for mental health assessment, treatment and emotional understanding. By using AI techniques, particularly machine learning and deep learning, researchers are able to extract emotional information from large amounts of complex data to enable more accurate predictions and analysis of emotional states. To illustrate, AI algorithms are capable of accurately quantifying changes in emotional expression when analyzing facial movements. This enables the manipulation of emotions[63], for instance, from a neutral state to a positive or negative emotional state, with the resulting change being quantified from 0 to 100%. Furthermore, the application of quantum heuristics and variational quantum state fidelity estimation shows great potential for quantifying emotions in real-time scenarios[64].

Exploring fine-grained analyses of sentiment using LLMs not only helps in sentiment classification, but also quantifies the magnitude of each emotion in a sample, providing a new approach to understanding complex sentiment composition[65]. In addition, a nonlinear speech modeling approach through nonlinear multi-fractal analysis reveals the complexity of speech signals in different affective states, providing a new language-independent sentiment quantification strategy[66].

While these technological advances provide powerful tools for quantifying emotion, they still face challenges in terms of psychological interpretability. The "black box" nature of deep learning models limits their application to psychological interpretation, as these models often struggle to reveal their internal decision-making processes[67]. This is particularly problematic in psycho-therapeutic and clinical applications, as these fields rely on a deep understanding of the psychological reasons behind emotional states.

Furthermore, cultural differences are a challenge that cannot be ignored in the application of globalized sentiment analysis. Differences in sentiment expression across cultures can lead to algorithms misinterpreting or misclassiying sentiment expressions[68]. Therefore, the development of emotion quantification models and algorithms that can operate effectively across cultures is an important direction for future research.

To improve the effectiveness and psychological interpretability of AI applications in emotion quantification, researchers need to develop more transparent and interpretable models[69, 70]. These models should be able to provide detailed information about their decision-making processes and be adaptable to the needs of emotion recognition in different cultural contexts.

However, while significant progress has been made in quantifying emotions and personality traits in contemporary psychology as shown in Fig. 2, reflecting the fact that accurate description and measurement of personality and emotional states are crucial for both theory and application. These approaches have limitations in the following four areas.

- **Complex interactions between emotion and personality traits.** The subjective and dynamic nature of affective experience is difficult to capture in quantitative models of emotion. For example, Mehrabian[71] used Pleasure-Arousal-Dominance (PAD) models to examine the relationship between personality traits and affective states, but these models do not capture the complexity of affective responses. In addition, Plutchik and van Praag[72] suggested that the relationship between emotion and behavior is more complex, particularly in situations involving violent behavior, and that the explanatory power of quantitative models of emotion is questioned.
- **Cultural dependence of models of affect.** The cultural dependence of quantitative models of emotion limits their global applicability. The same expression of emotion may have different meanings and consequences in different cultural contexts, as shown in Mehrabian's study[73], but his study was mainly based on Western culture, which is particularly important in cross-cultural research, e.g. Mehrabian's study[74] showed that cultural context profoundly affects the interpretation and expression of emotion.
- **Methodological limitations of affective measures.** The results of self-report questionnaires, the primary data collection instrument, are susceptible to the psychological state of the respondent and social expectations. This method is not effective in capturing immediate and dynamic changes in emotions[75]. In addition, there is an inherent bias between the subjective experience of emotion and the objective measurement of the scale, such as the contradiction between the multidimensionality of emotional experience and the single dimensionality of the scale as pointed out by Mehrabian[74].
- **Quantitative models and individual differences.** Quantitative models of emotion emphasize group-level regularities and ignore significant differences in emotional experience between individuals. This is confirmed by Mehrabian's[73] study in which individual differences in affective responses were significant but often ignored by quantitative models.

Given the significance of emotion quantification in psychological research, this method offers a precise assessment of emotional state. Furthermore, it has the potential to inform mental health assessment, therapeutic interventions, and communication in everyday contexts. Therefore, research on AI emotion quantification that incorporates psychological interpretability is crucial.

3 AI Emotion Quantification Research Incorporating Psychological Interpretability

Incorporating psychological interpretability is crucial in research on AI applications for emotion quantification, as it strengthens the comprehension and interaction quality of the model, enabling it to interact more effectively with human users in real-world

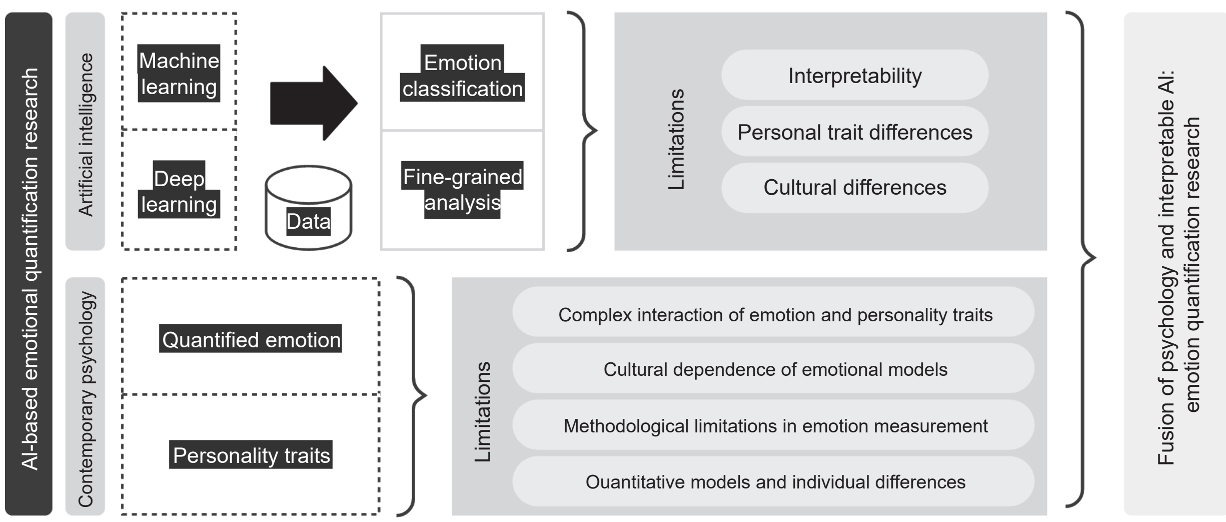

Fig. 2 A critical examination of the limitations of emotion recognition techniques and an overview of the latest developments in emotion quantification research.

applications. The OPO-FCM[76] model is an excellent example of this endeavor, which, by integrating deep learning techniques and psychological theories, not only improves the accuracy of emotion recognition, but also enhances the psychological interpretability of the model.

The model demonstrates the power of combining psychological theories and computational models by analyzing the expressive features in the video, mapping these features to the PAD emotion space, and then analyzing the individual's emotions and personality traits based on the character-emotion mapping relationship. The experimental results show that the effectiveness of the OPO-FCM model applied in a simulated environment reaches an accuracy rate of 79.56%, proving its feasibility and effectiveness in real-world situations[76].

The successful application of such a model suggests that the deep integration of psychological theories can significantly enhance the utility and depth of understanding of affective AI.

In addition, the development of methods such as EMER has further demonstrated that the transparency and explanatory power of affective AI can be improved by integrating different data sources and psychological theories[49]. By explaining in detail the process of inferring emotional states, this approach not only increases the user's trust in the AI system, but also facilitates the user's understanding of the system's decision logic.

Similarly, the proposed kernel-typical correlation analysis algorithm based on *K*-means clustering demonstrates how to improve the accuracy of emotion recognition in human-computer interaction by integrating the emotional features of different modalities[50]. This approach can effectively deal with non-linear relationships between complex data and increase the flexibility and accuracy of the model in dealing with different emotional states.

Overall, the integration of cognitive models with AI techniques is seen as an important trend in future emotion AI research. This integration not only improves the performance of quantitative emotion models, but also provides new tools for psychological research, transforming AI systems from simple emotion recognition tools to advanced platforms capable of deeply understanding and interacting with human emotions. As the technology develops, it is expected that this convergence will lead to the emergence of more sophisticated and human-like AI systems that understand and interact with human emotions at a deeper level based on psychological principles.

4 Future Direction

A modern view of the importance of emotion regulation and identification as essential to understanding and improving mental health treatment is provided in related articles from *Nature Reviews Psychology*. They emphazized the need for psychological interventions to be flexible and personalized to reflect the complexity and individuality of mental health problems. For example, a 2021 *Nature* editorial[77] discusses the complex needs of preventing and treating anxiety and depression in adolescents, highlighting the lack of a single effective treatment and the need for a combination of individualized interventions. In another 2023 *Nature* article[77], Carla Nardelli reviewed the importance of emotion regulation and its flexibility in mental processes, discussing the implications of being able to tailor emotional responses to situational demands, which is essential for maintaining mental health.

Recent research in AI affective computing has greatly enhanced the accuracy of emotion recognition compared to earlier approaches in the application of deep learning techniques, particularly models such as CNNs, Recurrent Neural Networks (RNNs), Long and Short-Term Memory Networks (LSTMs), and Transformer. Meanwhile, the fusion of multiple information sources, including vision, audio, and text, enhances the robustness and accuracy of emotion recognition by capturing cross-modal complementary information using deep learning models. Furthermore, significant progress has been made in the analysis of multidimensional information pertaining to the human body, particularly in the field of fine emotion expression and adaptive learning techniques.

In the field of sentiment analysis, the accuracy and depth of understanding of sentiment recognition have been greatly improved from the traditional model that relied on manual features at the beginning to automatic feature learning driven by deep learning, especially through multi-modal fusion technology. At the same time, the research expands from basic emotion classification to emotion continuum models (e.g., 2D valence-arousal model), which depicts the emotion dimensions more delicately. The combination of large-scale datasets with pre-trained models and the advancement of personalized emotion understanding through adaptive and reinforcement learning indicates that emotion recognition technology is undergoing a period of rapid evolution, becoming increasingly complex, subtle and efficient.

While AI has shown great potential in quantifying emotion, there are a number of challenges and limitations. These include cross-cultural adaptation of algorithms, the ability to process complex emotional data, and ensuring transparency and ethics in data processing. Future research needs to focus on the social and ethical implications of these techniques, while improving algorithmic accuracy and generalizability.

By combining psychological sentiment quantification with AI techniques, this study not only extends the methodology of sentiment quantification, but also improves the accuracy and efficiency of sentiment recognition and analysis.

Future research should continue to explore new biomarkers and algorithms to improve the accuracy and applicability of the model, especially in cross-cultural and multi-contextual settings as shown in Fig. 3.

In addition, future research should also focus on the ethical aspects of the technology in practical applications, especially when dealing with sensitive personal emotion data, to ensure that data security and personal privacy are adequately protected.

Such a review deepens the understanding of emotion quantification and demonstrates the potential of AI to advance psychological research and application areas.

4.1 Quantitative research towards computational psychiatry

With the development of computational psychiatry, the field has shown great promise in quantifying and treating emotional and cognitive disorders using machine learning and computational modeling. Studies, such as Nair et al.[78] applying computational frameworks to psychotherapy, have uncovered mechanisms of psychotherapy and improved personalization and effectiveness of treatment. In addition, Ribba[79] explored precision drug delivery through augmented learning, demonstrating the flexibility and adaptability of computational approaches in optimising therapeutic regimens. Ging-Jehli et al.'s[80] study validates the use of neurofeedback in ADHD treatment through computational modeling, highlighting the importance of computational psychiatry in personalized medicine. Finally, Ennis[81] discussed the role of behavioral quantification in neurobiology and digital

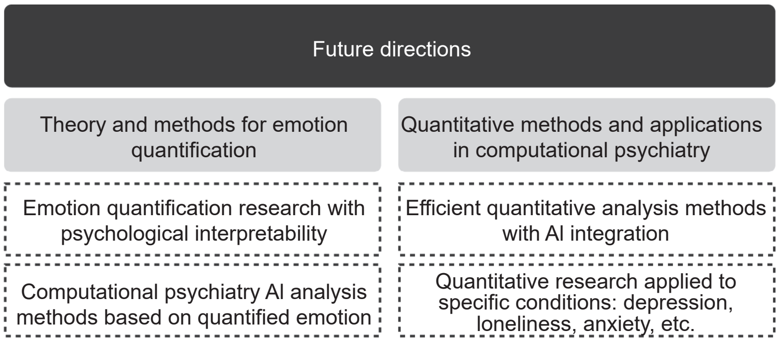

Fig. 3 Aims and objectives of future research in the field of emotion quantification based on AI technology.

psychiatry, highlighting the potential of modern technology in behavioral analysis.

Looking ahead, computational psychiatry is expected to continue to play a key role in mental health, particularly in improving the accuracy of treatment and disease prediction. Continued technological innovation and interdisciplinary collaboration will be the main drivers of the field moving forward. Researchers need to focus on how these advanced computational tools and theories can be applied to clinical practice, while also considering ethical and legal issues to ensure that scientific and technological advances are not only scientifically valid, but also ethically and socially acceptable.

4.2 Quantitative research on loneliness emotions

Techniques for quantifying emotions have played an important role in studies of loneliness and its impact on mental health. With continued advances in cognitive science and computational technology, future studies will be able to explore the underlying mechanisms of loneliness and its complex relationship with mental health problems in greater depth. For example, latent class analysis reveals the existence of a multidimensional structure of loneliness and points to its association with different dimensions of mental health[82]. This type of analysis not only helps us to identify subtypes of loneliness, but also provides a scientific basis for tailored mental health intervention strategies.

In addition, art therapy has shown unique utility in alleviating loneliness, particularly in promoting psychological rehabilitation and social integration through interaction with military historical objects among veterans, demonstrating the potential for applying emotional therapy to specific groups[83]. Community support projects, such as the Compassionate City programme in Vic, Barcelona, have been effective in reducing loneliness by strengthening social connections between community members, highlighting the role of social structure in maintaining emotional health[84].

Going forward, research should continue to explore how advanced computational modeling and machine learning techniques can be used to improve the accuracy and utility of emotion quantification, while addressing privacy and ethical issues to ensure that the psychological state of participants and the security of their personal data are respected in research and applications. In addition, interdisciplinary collaboration will be key to advancing the field by integrating expertise from multiple fields, including psychology, data science and artificial intelligence, to better understand and address loneliness and its complex psychosocial consequences. These efforts will enable us to provide more effective mental health support and social interventions to the public and improve overall well-being.

4.3 Emotion quantification studies in depression and anxiety

Emotion quantification techniques are rapidly becoming an indispensable tool in the identification and treatment of depression and anxiety. Research over the past five years has demonstrated the great potential of using advanced computational models and neural network techniques in this area. For example, using quantitative questionnaires and stress, anxiety and depression scales, studies have identified several factors that influence anxiety in care workers, such as lack of family and social support and lack of job autonomy, and these studies have provided targeted strategies to address them. In the field of depression identification and treatment, mood quantification techniques are rapidly becoming a key tool and have shown great potential for identifying depression using advanced computational modeling and neural network techniques. For example, a systematic evaluation by Wu and Xiao[85] highlighted the effectiveness of neuroimaging techniques in diagnosing psychiatric disorders including depression, anxiety disorders, etc. Huang[86] significantly improved the accuracy of identifying depression in older adults by combining facial microexpression analysis and deep learning techniques. Li[87] effectively detecteded mild depression using EEG and convolutional neural network techniques, further demonstrating the value of deep learning in mental health applications. Sun et al.[88] pioneered the application of multi-granularity graph convolutional networks to analyze major depressive disorder, demonstrating the progress of modern technology in capturing complex brain signals[88].

In the future, quantitative mood research in depression and anxiety will continue to rely on interdisciplinary approaches combining neuroscience, machine learning, and big data analytics to optimize diagnostic tools and treatment strategies. The research community could also focus on developing ethical application protocols to ensure that the use of these advanced technologies is not only scientifically valid, but also meets ethical standards, especially when dealing with sensitive mental health data. As technology advances and data accessibility increases, it is expected that more personalized and accurate diagnostic and treatment approaches for depression will be developed in the coming years, greatly improving the quality of life of patients.

5 Conclusion

This paper provides an in-depth analysis of the key developments of emotion quantification technology in the field of artificial intelligence, especially its application in emotion recognition technology and its revolutionary impact on areas such as health monitoring, educational experience optimization, and customer service. The utilization of multi-source data, encompassing facial expressions, speech, text, and physiological signals, has led to a notable enhancement in the precision of emotion recognition. Furthermore, it has fostered the acceptance and implementation of emotion quantification techniques in a multitude of real-world scenarios. In the domain of healthcare, for instance, this approach has been instrumental in expediting the diagnosis and personalization of treatment plans for various ailments, underscoring its vast potential for impact.

However, the development and application of emotion quantification technology still faces several challenges. Firstly, one of the key issues that the technology needs to address is to improve transparency and interpretability. Although current deep learning models for emotion recognition are highly accurate, their "black box" nature limits their use in application scenarios that require a high level of trust, such as medical and psychological counselling. In addition, the issue of cross-cultural adaptation needs to be addressed, as emotion expression and recognition vary significantly across cultures, and future model development could take this into account to ensure accuracy and adaptability for global applications.

Privacy and security are also issues that need to be taken seriously with sentiment quantification technologies. As the use of sentiment data increases, the question of how to make the best use of this data while ensuring that individual privacy is not violated is an issue that needs to be addressed in the development of the technology. This will require the research community and relevant stakeholders to develop more stringent data handling standards and privacy measures.

Furthermore, as a powerful tool, the prospective advancement of AI-based methodologies pertaining to emotional assessment is recommended to also accord greater attention to data privacy and ethical considerations, and strive to achieve a harmonious equilibrium between technological advancement and ethical and social responsibility. Further research should explore more transparent and interpretable methods of emotion recognition based on psychological and cognitive sciences that strengthen the model, while being accurately applicable in a global multicultural context. Through interdisciplinary collaboration and enhanced integration of theory and practical applications, emotion quantification techniques are expected to play a more central role in understanding and responding to human emotions, and to promote better cross-fertilization between human-computer interaction technologies, computational psychology, and computational psychiatry.

Acknowledgment

This work was supported by the Beijing Key Laboratory of Behavior and Mental Health, Peking University. We also thank the reviewers for giving us valuable comments to make the manuscript more valuable.

Article History

Received: 11 May 2024; Revised: 1 July 2024; Accepted: 23 July 2024

References

[1] R. W. Picard, Affective Computing, Cambridge, MA, USA: MIT Press, 2000.

[2] R. A. Calvo and S. D'Mello, Affect detection: An interdisciplinary review of models, methods, and their applications, *IEEE Trans. Affective Comput.*, vol. 1, no. 1, pp. 18–37, 2010.

[3] S. Poria, E. Cambria, R. Bajpai, and A. Hussain, A review of affective computing: From unimodal analysis to multimodal fusion, *Inf. Fusion*, vol. 37, pp. 98–125, 2017.

[4] J. J. Gross, Emotion regulation: Current status and future prospects, *Psychol. Inq.*, vol. 26, no. 1, pp. 1–26, 2015.

[5] G. M. Harari, N. D. Lane, R. Wang, B. S. Crosier, A. T. Campbell, and S. D. Gosling, Using smartphones to collect behavioral data in psychological science, *Perspect. Psychol. Sci.*, vol. 11, no. 6, pp. 838–854, 2016.

[6] S. George, From sex and therapy bots to virtual assistants and tutors: How emotional should artificially intelligent agents be? in *Proc. 1st Int. Conf. Conversational User Interfaces*, Dublin, Ireland, 2019, pp. 1–3.

[7] Y. Lim and M. Lee, Implications of emotional coaching and integrated art therapy teaching method on leadership education in the AI era, *J-Institute*, vol. 5, no. 2, pp. 42–49, 2020.

[8] P. P. Frank, M. X. E. Lu, and E. C. Sasse, Educational and emotional needs of patients with myelodysplastic syndromes: An AI analysis of multi-country social media, *Adv. Ther.*, vol. 40, no. 1, pp. 159–173, 2023.

[9] William Cho, https://globaljournals.org/GJMR_Volume22/E-Journal_GJMR_(A)_Vol_22_Issue_3.pdf, 2022.

[10] X. Jia, Research on the emotional impact of AI care robots on elderly living alone, *J. Artif. Intell. Pract.*, vol. 6, no. 6, pp. 50–55, 2023.

[11] A. Krizhevsky, I. Sutskever, and G. E. Hinton, ImageNet classification with deep convolutional neural networks, *Commun. ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[12] C. Liu, Q. Tian, and M. Chen, Distinguishing personality recognition and quantification of emotional features based on users' information behavior in social media, *J. Database Manag.*, vol. 32, no. 2, pp. 76–91, 2021.

[13] J. Luo, G. Zhang, Y. Su, Y. Lu, Y. Pang, Y. Wang, H. Wang, K. Cui, Y. Jiang, L. Zhong, et al., Quantitative analysis of heart rate variability parameter and mental stress index, *Front. Cardiovasc. Med.*, vol. 9, p. 930745, 2022.

[14] E. M. Polo, M. Mollura, M. Lenatti, M. Zanet, A. Paglialonga, and R. Barbieri, Emotion recognition from multimodal physiological measurements based on an interpretable feature selection method, in *Proc. 43rd Annu. Int. Conf. IEEE Engineering in Medicine & Biology Society* (*EMBC*), virtual, 2021, pp. 989–992.

[15] H. Prossinger, T. Hladky, J. Binter, S. Boschetti, and D. Riha, Visual analysis of emotions using AI image-processing software: Possible male/female differences between the emotion pairs "neutral"–"fear" and "pleasure"–"pain", in *Proc. 14th PErvasive Technologies Related to Assistive Environments Conf.* (*PETRA '21*), Corfu, Greece, 2021, pp. 342–346.

[16] Q. Li, S. Zhan, L. Xu, and C. Wu, Facial micro-expression recognition based on the fusion of deep learning and enhanced optical flow, *Multimed. Tools Appl.*, vol. 78, no. 20, pp. 29307–29322, 2019.

[17] P. Ekman and E. L. Rosenberg, What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS), Oxford, UK: Oxford University Press, 2005.

[18] J. Zhang, F. Liu, and A. Zhou, Off-TANet: A lightweight neural micro-expression recognizer with optical flow features and integrated attention mechanism, in *Proc. 18th Pacific Rim Int. Conf. Artificial Intelligence* (*PRICAI 2021*), Hanoi, Vietnam, 2021, pp. 266–279.

[19] H. Wang, B. Li, S. Wu, S. Shen, F. Liu, S. Ding, and A. Zhou, Rethinking the learning paradigm for dynamic facial expression recognition, in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition* (*CVPR*), Vancouver, Canada, 2023, pp. 17958–17968.

[20] F. Ma, B. Sun, and S. Li, Logo-former: Local-global spatio-temporal transformer for dynamic facial expression recognition, in *Proc. 2023 IEEE Int. Conf. Acoustics, Speech and Signal Processing* (*ICASSP*), Rhodes Island, Greece, 2023, pp. 1–5.

[21] L. Sun, Z. Lian, B. Liu, and J. Tao, MAE-DFER: Efficient masked

autoencoder for self-supervised dynamic facial expression recognition, in *Proc. 31st ACM Int. Conf. Multimedia*, Ottawa, Canada, 2023, pp. 6110–6121.

[22] A. V. Savchenko and I. A. Makarov, Neural network model for video-based analysis of student's emotions in E-learning, *Opt. Mem. Neural Networks*, vol. 31, no. 3, pp. 237–244, 2022.

[23] F. Liu, H. Wang, J. Zhang, Z. Fu, A. Zhou, J. Qi, and Z. Li, EvoGAN: An evolutionary computation assisted GAN, *Neurocomputing*, vol. 469, pp. 81–90, 2022.

[24] K. Chen, Z. Zhang, W. Zeng, R. Zhang, F. Zhu, and R. Zhao, Shikra: Unleashing multimodal llm's referential dialogue magic, arXiv preprint arXiv: 2306.15195, 2023.

[25] K. Wang, X. Peng, J. Yang, S. Lu, and Y. Qiao, Suppressing uncertainties for large-scale facial expression recognition, in *Proc. IEEE/CVF Conf. Computer Vision and Pattern Recognition* (*CVPR*), Seattle, WA, USA, 2020, pp. 6896–6905.

[26] A. P. Fard and M. H. Mahoor, Ad-corre: Adaptive correlation-based loss for facial expression recognition in the wild, *IEEE Access*, vol. 10, pp. 26756–26768, 2022.

[27] C. Zhang, M. Li, and D. Wu, Federated multidomain learning with graph ensemble autoencoder GMM for emotion recognition, *IEEE Trans. Intell. Transport. Syst.*, vol. 24, no. 7, pp. 7631–7641, 2023.

[28] S. Medjden, N. Ahmed, and M. Lataifeh, Adaptive user interface design and analysis using emotion recognition through facial expressions and body posture from an RGB-D sensor, *PLoS One*, vol. 15, no. 7, p. e0235908, 2020.

[29] M. Gopika Sri, G. Karthiga, K. Jayakarthika, N. Ilakkiya, V. Lavanyagayathri, and D. Uma Mageswari S, A fuzzy logic and NLP approach to emotion driven response generation for voice interaction, in *Proc. Int. Conf. Research Methodologies in Knowledge Management, Artificial Intelligence and Telecommunication Engineering* (*RMKMATE*), Chennai, India, 2023, pp. 1–5.

[30] K. Zhou, B. Sisman, C. Busso, and H. Li, Mixed emotion modeling for emotional voice conversion, arXiv preprint arXiv: 2210.13756, 2022.

[31] M. Prince, Real-time emotional expression generation by humanoid robot, *Int. J. Adv. Comput. Sci. Appl.*, vol. 12, no. 12, pp. 381–385, 2021.

[32] E. Conti, D. Salvi, C. Borrelli, B. Hosler, P. Bestagini, F. Antonacci, A. Sarti, M. C. Stamm, and S. Tubaro, Deepfake speech detection through emotion recognition: A semantic approach, in *Proc. 2022 IEEE Int. Conf. Acoustics, Speech and Signal Processing* (*ICASSP*), Singapore, 2022, pp. 8962–8966.

[33] A. Adigwe, N. Tits, K. El Haddad, S. Ostadabbas, and T. Dutoit, The emotional voices database: Towards controlling the emotion dimension in voice generation systems, arXiv preprint arXiv: 1806.09514, 2018.

[34] B. Zhou and X. Li, Multimodal emotion analysis model based on interactive attention mechanism, *Front. Comput. Intell. Syst.*, vol. 3, no. 2, pp. 67–73, 2023.

[35] L. Ahmed, I. K. Polok, M. A. Islam, M. Akhtaruzzaman, M. S. H. Mukta, and M. M. Rahman, Context based emotion recognition from Bengali text using transformers, in *Proc. 5th Int. Conf. Smart Systems and Inventive Technology* (*ICSSIT*), Tirunelveli, India, 2023, pp. 1478–1484.

[36] M. A. Mahima, N. C. Patel, S. Ravichandran, N. Aishwarya, and S. Maradithaya, A text-based hybrid approach for multiple emotion detection using contextual and semantic analysis, in *Proc. Int. Conf. Innovative Computing, Intelligent Communication and Smart Electrical Systems* (*ICSES*), Chennai, India, 2021, pp. 1–6.

[37] K. Hemakirthiga and J. Arunadevi, Improving emotion detection in text: A comparative analysis of machine learning algorithms and genetic algorithm-optimized logistic regression, in *Proc. Int. Conf. Data Science, Agents & Artificial Intelligence* (*ICDSAAI*), Chennai, India, 2023, pp. 1–6.

[38] M. J. Al-Dujaili and A. Ebrahimi-Moghadam, Speech emotion recognition: A comprehensive survey, *Wirel. Pers. Commun.*, vol. 129, no. 4, pp. 2525–2561, 2023.

[39] K. Bhangale and M. Kothandaraman, Speech emotion recognition based on multiple acoustic features and deep convolutional neural network, *Electronics*, vol. 12, no. 4, p. 839, 2023.

[40] W. Ahmed, S. Riaz, K. Iftikhar, and S. Konur, Speech emotion recognition using deep learning, in *Proc. 43rd SGAI Int. Conf. Artificial Intelligence*, Cambridge, UK, 2023, pp. 191–197.

[41] J. de Lope and M. Graña, An ongoing review of speech emotion recognition, *Neurocomputing*, vol. 528, pp. 1–11, 2023.

[42] F. Daneshfar and M. B. Jamshidi, An octonion-based nonlinear echo state network for speech emotion recognition in Metaverse, *Neural Netw.*, vol. 163, pp. 108–121, 2023.

[43] A. Carvalho, A. Levitt, S. Levitt, E. Khaddam, and J. Benamati, Off-the-shelf artificial intelligence technologies for sentiment and emotion analysis: A tutorial on using IBM natural language processing, *Commun. Assoc. Inf. Syst.*, vol. 44, pp. 918–943, 2019.

[44] X. Zhang, The application of natural language processing technology based on deep learning in Japanese sentiment analysis, in *Proc. Int. Conf. Ambient Intelligence, Knowledge Informatics and Industrial Electronics* (*AIKIIE*), Ballari, India, 2023, pp. 1–5.

[45] C. M. A. Ilyas, R. Nunes, K. Nasrollahi, M. Rehm, and T. B. Moeslund, Deep emotion recognition through upper body movements and facial expression, in *Proc. 16th Int. Joint Conf. Computer Vision, Imaging and Computer Graphics Theory and Applications* (*VISIGRAPP 2021*), virtual, 2021, pp. 669–679.

[46] Z. Fu, F. Liu, J. Zhang, H. Wang, C. Yang, Q. Xu, J. Qi, X. Fu, and A. Zhou, SAGN: Semantic adaptive graph network for skeleton-based human action recognition, in *Proc. 2021 Int. Conf. Multimedia Retrieval*, Taipei, China, 2021, pp. 110–117.

[47] Q. Xu, F. Liu, Z. Fu, A. Zhou, and J. Qi, AeS-GCN: Attention-enhanced semantic-guided graph convolutional networks for skeleton-based action recognition, *Comput. Animat. Virtual Worlds*, vol. 33, nos.3&4, pp. e2070, 2022.

[48] S. Sun, X. Xiong, and Y. Zheng, Two stage multi-modal modeling for video interaction analysis in deep video understanding challenge, in *Proc. 30th ACM Int. Conf. Multimedia*, Lisboa, Portugal, 2022, pp. 7040–7044.

[49] Z. Lian, H. Sun, L. Sun, H. Gu, Z. Wen, S. Zhang, S. Chen, M. Xu, K. Xu, K. Chen, et al., Explainable multimodal emotion reasoning: a promising way to open-set emotion recognition, arXiv preprint arXiv: 2306.15401, 2023.

[50] L. Chen, K. Wang, M. Li, M. Wu, W. Pedrycz, and K. Hirota, *K*-means clustering-based kernel canonical correlation analysis for multimodal emotion recognition in human–robot interaction, *IEEE Trans. Ind. Electron.*, vol. 70, no. 1, pp. 1016–1024, 2023.

[51] J. Zheng, S. Zhang, Z. Wang, X. Wang, and Z. Zeng, Multi-channel weight-sharing autoencoder based on cascade multi-head attention for multimodal emotion recognition, *IEEE Trans. Multimedia*, vol. 25, pp. 2213–2225, 2023.

[52] D. Yang, S. Huang, H. Kuang, Y. Du, and L. Zhang, Disentangled representation learning for multimodal emotion recognition, in *Proc. 30th ACM Int. Conf. Multimedia*, Lisboa, Portugal, 2022, pp. 1642–1651.

[53] D. Sun, Y. He, and J. Han, Using auxiliary tasks in multimodal fusion of Wav2vec 2.0 and bert for multimodal emotion recognition, in *Proc. 2023 IEEE Int. Conf. Acoustics, Speech and Signal Processing* (*ICASSP*), Rhodes Island, Greece, 2023, pp. 1–5.

[54] N. Ahmed, Z. Al Aghbari, and S. Girija, A systematic survey on multimodal emotion recognition using learning algorithms, *Intell. Syst. Appl.*, vol. 17, pp. 200171, 2023.

[55] R. S. Aparicio Garcia, G. Juarez Gracia, J. A. Alvarez Cedillo, J. Sandoval Gutierrez, and B. Tovar Corona, Evaluation of the design of a brain-computer interface for emotion detection, *Dyna*, vol. 93, no. 1, pp. 468, 2018.

[56] J. Pan, L. Wang, H. Huang, J. Xiao, F. Wang, Q. Liang, C. Xu, Y. Li, and Q. Xie, A hybrid brain–computer interface combining P300 potentials and emotion patterns for detecting awareness in patients with disorders of consciousness, *IEEE Trans. Cogn. Dev. Syst.*, vol.

15, no. 3, pp. 1386–1395, 2023.
[57] E. D. Floreani, S. Orlandi, and T. Chau, A pediatric near-infrared spectroscopy brain-computer interface based on the detection of emotional valence, *Front. Hum. Neurosci.*, vol. 16, pp. 938708, 2022.
[58] Y. Zhao, Wearable brain-computer interface technology and its application, *Theor. Nat. Sci.*, vol. 15, no. 1, pp. 137–145, 2023.
[59] W. Liu, H. Jiang, and Y. Lu, Research on multimodal emotion recognition platform construction, in *Proc. Information Science and Cloud Computing* (*ISCC 2017*), Guangzhou, China, 2018, pp. 1–9.
[60] Y. Lee, N. Lee, V. Pham, J. Lee, and T. M. Chung, Privacy preserving stress detection system using physiological data fromWearable device, in *Proc. 6th Int. Conf. Intelligent Human Systems Integration* (*IHSI 2023*): *Integrating People and Intelligent Systems*, Venice, Italy, 2023, pp. 340–347.
[61] M. Z. Baig and M. Kavakli, A survey on psycho-physiological analysis & measurement methods in multimodal systems, *Multimodal Technol. Interact.*, vol. 3, no. 2, pp. 37, 2019.
[62] M. A. Razzaq, J. Hussain, J. Bang, C. H. Hua, F. A. Satti, U. U. Rehman, H. S. M. Bilal, S. T. Kim, and S. Lee, A hybrid multimodal emotion recognition framework for UX evaluation using generalized mixture functions, *Sensors*, vol. 23, no. 9, p. 4373, 2023.
[63] J. Binter, S. Boschetti, T. Hladký, H. Prossinger, T. J. Wells, J. Jílková, and D. Říha, Quantifying the rating performance of ambiguous and unambiguous facial expression perceptions under conditions of stress by using wearable sensors, in *Proc. 24th Int. Conf. Human-Computer Interaction* (*HCII 2022*), virtual, 2022, pp. 519–529.
[64] J. Singh, F. Ali, B. Shah, K. S. Bhangu, and D. Kwak, Emotion quantification using variational quantum state fidelity estimation, *IEEE Access*, vol. 10, pp. 115108–115119, 2022.
[65] M. Sharma, I. Kandasamy, and W. B. Vasantha, Emotion quantification and classification using the neutrosophic approach to deep learning, *Appl. Soft Comput.*, vol. 148, pp. 110896, 2023.
[66] U. Sarkar, S. Nag, C. Bhattacharya, S. Sanyal, A. Banerjee, R. Sengupta, and D. Ghosh, Language independent emotion quantification using non linear modeling of speech, arXiv preprint arXiv: 2102.06003, 2021
[67] D. Castelvecchi, Can we open the black box of AI, *Nature*, vol. 538, no. 7623, pp. 20–23, 2016.
[68] S. Mohamed, M. T. Png, and W. Isaac, Decolonial AI: Decolonial theory as sociotechnical foresight in artificial intelligence, *Philos. Technol*, vol. 33, no. 4, pp. 659–684, 2020.
[69] M. Setzu, R. Guidotti, A. Monreale, F. Turini, D. Pedreschi, and F. Giannotti, GLocalX: from local to global explanations of black box AI models, arXiv preprint arXiv: 2101.07685, 2021.
[70] P. Verma, S. R. Marpally, and S. Srivastava, Discovering user-interpretable capabilities of black-box planning agents, in *Proc. 19th Int. Conf. Principles of Knowledge Representation and Reasoning*, Haifa, Israel, 2022, pp. 362–372.
[71] A. Mehrabian, Analysis of the big-five personality factors in terms of the pad temperament model, *Aust. J. Psychol.*, vol. 48, no. 2, pp. 86–92, 1996.
[72] R. Plutchik and H. Van Praag, The measurement of suicidality, aggressivity and impulsivity, *Prog. Neuro Psychopharmacol. Biol. Psychiatry*, vol. 13, pp. S23–S34, 1989.
[73] A. Mehrabian, Pleasure-arousal-dominance: A general framework for describing and measuring individual differences in Temperament, *Curr. Psychol.*, vol. 14, no. 4, pp. 261–292, 1996.
[74] A. Mehrabian, Relations among personality scales of aggression, violence, and empathy: Validational evidence bearing on the risk of eruptive violence scale, *Aggress. Behav.*, vol. 23, no. 6, pp. 433–445, 1997.
[75] A. Mehrabian, Beyond IQ: Broad-based measurement of individual success potential or “emotional intelligence”, *Genet. Soc. Gen. Psychol. Monogr.*, vol. 126, no. 2, pp. 133–239, 2000.
[76] F. Liu, H. Y. Wang, S. Y. Shen, X. Jia, J. Y. Hu, J. H. Zhang, X. Y. Wang, Y. Lei, A. M. Zhou, J. Y. Qi, et al., OPO-FCM: A computational affection based OCC-PAD-OCEAN federation cognitive modeling approach, *IEEE Trans. Comput. Soc. Syst.*, vol. 10, no. 4, pp. 1813–1825, 2023.
[77] C. Nardelli, From emotion regulation to emotion regulation flexibility, *Nat. Rev. Psychol.*, vol. 2, no. 11, pp. 660–660, 2023.
[78] A. Nair, R. B. Rutledge, and L. Mason, Under the hood: Using computational psychiatry to make psychological therapies more mechanism-focused, *Front. Psychiatry*, vol. 11, pp. 140, 2020.
[79] B. Ribba, Reinforcement learning as an innovative model-based approach: Examples from precision dosing, digital health and computational psychiatry, *Front. Pharmacol.*, vol. 13, pp. 1094281, 2023.
[80] N. R. Ging-Jehli, H. C. Kraemer, L. Eugene Arnold, M. E. Roley-Roberts, and R. de Beus, Cognitive markers for efficacy of neurofeedback for attention-deficit hyperactivity disorder–personalized medicine using computational psychiatry in a randomized clinical trial, *J. Clin. Exp. Neuropsychol.*, vol. 45, no. 2, pp. 118–131, 2023.
[81] M. Ennis, Behavior quantification as the missing link between fields: Tools for digital psychiatry and their role in the future of neurobiology, arXiv preprint ArXiv: 2305.15385, 2023.
[82] C. Ruiz, K. Ito, S. Wakamiya, and E. Aramaki, Loneliness in a connected world: Analyzing online activity and expressions on real life relationships of lonely users, in *Proc. AAAI 2017 Spring Symp.*, Stanford, CA, USA, 2017, pp. 726–733.
[83] J. Lobban and D. Murphy, Military museum collections and art therapy as mental health resources for veterans with PTSD, *Int. J. Art Ther.*, vol. 25, no. 4, pp. 172–182, 2020.
[84] X. Gómez-Batiste, S. Mateu, S. Serra-Jofre, M. Molas, S. MiR-Roca, J. Amblàs, X. Costa, C. Lasmarías, M. Serrarols, A. Solà-Serrabou, et al., Compassionate communities: Design and preliminary results of the experience of Vic (Barcelona, Spain) caring city, *Ann. Palliat. Med.*, vol. 7, no. S2, pp. S32–S41, 2018.
[85] J. Wu and J. Xiao, Effectiveness of the neuroimaging techniques in the recognition of psychiatric disorders: A systematic review and meta-analysis of RCTs, *Curr. Med. Imag. Rev.*, vol. 20, p. e260523217379, 2024.
[86] W. Huang, Elderly depression recognition based on facial micro-expression extraction, *Trait. du Signal*, vol. 38, no. 4, pp. 1123–1130, 2021.
[87] X. Li, R. La, Y. Wang, J. Niu, S. Zeng, S. Sun, and J. Zhu, EEG-based mild depression recognition using convolutional neural network, *Med. Biol. Eng. Comput.*, vol. 57, no. 6, pp. 1341–1352, 2019.
[88] X. Sun, Y. Xu, Y. Zhao, X. Zheng, Y. Zheng, and L. Cui, Multi-granularity graph convolution network for major depressive disorder recognition, *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 32, pp. 559–569, 2024.