Human-Computer Interaction on Facial Recognition and Emotional Feedback

Jinze Han

Brunel London School, North China University of Technology, Beijing, China

2161299@brunel.ac.uk

Abstract. The advancement of science and technology has resulted in the growing prevalence of facial recognition technology in human-computer interaction, particularly in the domain of emotion recognition, where it shows significant potential. By identifying the user's facial expressions and emotional responses, the system is capable of conducting further analysis and prediction of the user's needs, thereby optimizing the emotion recognition experience of human-computer interaction. This article will elucidate the principles of facial recognition technology, with a particular focus on its practical applications and technical realization in emotion recognition. Furthermore, this article will examine the current limitations of the technology, discuss potential avenues for improvement, and speculate on the future development of facial recognition technology in the field of human-computer interaction. The technology is currently being used in a variety of areas, including smart home technology, medical rehabilitation and public services. The technology has the potential to improve the user experience by increasing the intelligence of devices and providing more personalised services to users through sentiment analysis.

Keywords: Human-computer interaction, facial recognition, emotional identification.

1. Introduction

With the rapid development of modern science and technology, human-computer interaction technology has gradually penetrated into all aspects of life, becoming a bridge of communication between humans and machines and systems. With the advancement of artificial intelligence, the Internet of Things and big data technology, the application forms of human-computer interaction have become more diversified and intelligent. For example, voice assistants, virtual reality, gesture control and facial recognition technologies have been widely used in people's daily life, work, education, medical care and many other fields. Especially in the field of smart devices and robots, how to interact with users in a more natural and efficient way to improve the user experience has become a key direction of current research.

Facial recognition technology, as an important part of human-computer interaction, mainly relies on biometric identification technology to achieve personal identity verification or emotional state recognition through data analysis and comparison of facial feature points. Currently, most HCI-based systems often rely on user-initiated feedback, such as questionnaires or direct feedback from users, and developers then debug and improve based on this feedback. However, facial recognition systems offer a new way of thinking about this process: by capturing and analysing the user's facial expressions, the

system is able to automatically identify the user's emotional state and respond immediately, without relying on the user's active feedback.

At this stage, the application of facial recognition technology at home and abroad has begun to bear fruit. Domestically, for example, the technology has been widely used in education and public services, such as KDDI's Xiaobing robot (2020), public safety and patrol robots (since 2018), as well as facial recognition retail robots (since 2020) and medical rehabilitation robots (2021). At the international level, Amazon's Echo Show and Google's Nest Hub Max will be widely used in home services from 2019, while some European research teams have developed facial recognition robots to detect emotions in the elderly. While these technologies are pushing human-computer interaction towards a more intelligent and humanised form, they have also shown great potential for applications in different fields.

In this paper, based on the recognition of facial feature points by facial recognition technology, we will focus on how to identify the user's emotional state by comparing the digital feature vectors of the feature points with the data corresponding to emotions in the database, and apply the results of this analysis to the decision-making system of service robots to provide a better user experience. Furthermore, this paper will examine the potential implications of this technology at this stage and delineate the pertinent measures for ensuring the continued advancement of this field.

2. The technologies utilised in the domain of facial recognition

2.1. The human face functions as a conduit for the expression of information
Human facial muscles or organ movements are an important way to convey feelings and states, which lies mainly in the three key areas of eyes, nose and mouth, including the surrounding muscle groups, the combination of the three will form a unique biometric feature belonging to the individual. For example, facial expression is an important way to convey emotions, the muscles in these areas are composed of smiling, frowning, surprise and other expressions are an important reflection of the mood of the person who responded to the expression, and by analysing the user's facial expression, eye movement and gaze point, the system can determine the user's concentration level. Expressions reflected after an event can also convey social signals such as agreement, disagreement or confusion, reflecting the user's level of satisfaction with the experience.

2.2. Data mining of facial features
Recognising facial expressions is not as simple and easy as it sounds in theory, because everyone's facial muscles are different and there is no absolute standard to speak of, so it would be quite complex and difficult to data-materialise the whole face, so we focus on exploring the data of the key areas, the different parts of the face are automatically recognised by the camera module, and then the movement trajectories or subtle changes in movement of the key areas are Then the movement trajectory or subtle changes in movement data of the focus area is collected.

2.3. Database of algorithms for a wide range of applications
In the field of facial expression recognition, popular datasets include FER2013, MMI, Japanese Female Facial Expressions (JAFFE), and Extended Cohn-Kanade (CK+). Japanese Female Facial Expressions (JAFFE) was released earlier and is smaller in size [1]. MMI Face The database contains static still images and videos. It contains sequences of consecutive facial expressions. Additionally, the Facial Action Coding System (FACS) categorizes the action units associated with each expression. This database is distinguished by a relatively larger sample size and higher standard requirements. The most significant advantage of this dataset is that it encompasses both induced and spontaneous facial expressions as two distinct components [2]. This approach is optimal for the extension and debugging of subsequent programs, which are designed to facilitate the acquisition of knowledge regarding the models in question [3]. CK+ is an extension of CK [4]. The dataset consisted of 210 participants and 2105 images from individuals aged between 18 and 40 years. Approximately 65% of the subjects were female and approximately 35% were male [5]. The FER2013 dataset was created using keywords related

to emoticons in a Google image search. The dataset contains 35,886 dimensions of facial expressions. Each image is constituted by a grayscale image with a fixed width and height of 48 pixels. It is important to note, however, that the data set is not balanced [6].

2.4. The Facial Feature Recognition Algorithm

It is found that this data can be efficiently processed by constructing a Convolutional Neural Network (CNN) model, while this deep learning model reduces or eliminates the dependency on other models or existing models. The model shows good performance in a series of experiments, with an accuracy and average f1 score of 96% and 0.95 on the test data, which is more accurate than Recurrent Neural Networks (RNN) with 92% and 0.9, Support Vector Machines (SVMs) with 88% and 0.87 [7], and Random Forests (RFs) with 85% and 0.83, and the K-Nearest Neighbour (KNN) algorithm with 80% and 0.78 are much more accurate [8]. However, when the face is wearing glasses or too many accessories, and in poor lighting conditions, the accuracy of the recognition can be affected and the recognition points can be distorted in critical areas [9].

The process is conducted in a series of stages, each of which is designed to achieve a specific objective. In the initial stage of the process, the camera module acquires the video image and processes it frame by frame [7]. In this phase, the face region is removed for the purposes of data screening and classification. The video data is transformed into image data (frame sequence) through the application of deep learning methods. Subsequently, the facial regions are identified and analysed for each frame, and the explicit expression frames corresponding to emotions are filtered from the database. Images that are of insufficient clarity or lack the requisite detail are excluded from subsequent processing. The management of data is enhanced through the utilisation of data frames, which facilitate a more balanced distribution of data through reordering prior to inputting into the model. Subsequently, the data may be employed for the purposes of training the model and validating the accuracy of the database in recognising user emotions. This may also be employed as a means of providing the robot with a foundation for improvement, with the objective of enhancing the accuracy and user experience of human-robot interaction. The training and validation data are directly implicated in the model training process, while the images within the frame are scaled and resized in order to align with the input requirements of the convolutional neural network (CNN) model.

The accuracy of FER is directly impacted by the quality of facial expression characteristics [10]. Consequently, another area of interest for FER research is accurately identifying face characteristics to convey facial emotions. Geometric and statistical feature extraction are the two basic categories into which feature extraction methods fall. Using the key characteristics of face expressions in photos, geometric feature extraction primarily extracts rich spatial geometric information. Statistical feature extraction is usually performed using the methods of algebra such as Principal Component Analysis (PCA), Local Binary Pattern (LBP) [11], etc. In addition to capturing the most varied data components in the sample and choosing a number of significant people from each feature component, PCA may also be used for feature extraction [12]. The technique for extracting certain feature points of the human face using PCA is 99.67% accurate, according to the experimental results of Sharma et al [13]. While PCA is easy to use and has good resolution in low-dimensional space, the global features it extracts are highly dependent on the surrounding area and the full greyscale image. If specific local texture features can be obtained to support the global greyscale features, the recognition impact will be more pronounced and effective. It can be effectively improved by utilizing a hybrid PCA and LBP strategy. Based on the experimental observation, LBP is utilized to extract the mouth region's local characteristics. The global characteristics of the entire image are extracted using PCA. For classification, Support Vector Machines (SVM) are employed. Ultimately, 93.75% classification accuracy was attained using the suggested approach.

Furthermore, the two most crucial characteristics of LBP features are their computational simplicity and their ability to withstand changes in light [14]. We can see this in the experiment of extracting facial expression image features using this feature by Shan et al [15]. The LBP feature is a key issue in low-resolution facial expression recognition, but it is rarely addressed in existing research. Based to their

tests, the LBP feature works reliably and consistently over the effective range of low-resolution facial photos. It also does well in compressed low-resolution video sequences that were taken in actual settings. A popular feature descriptor in domains where an object's form is crucial is the Histogram of Oriented Gradients (HOG) [16].

In Yadav and Singha's paper, LBP and HOG features are used for feature extraction and then these two features are combined called hybrid features [17]. While LBP depicts the image's local texture features, HOG features are very sensitive to the object's form. After extracting and storing the HOG and LBP features for every facial component, they concatenated all of the collected features to create the feature matrix. Four databases—JAFFE, CK+, MMI, and a bespoke dataset—were used for training and testing, and their respective accuracy rates were 97.65%, 98.56%, 97.5%, and 99.77%.

In addition, Haar-like features and HOG features were extracted for each facial region by Kim et al [18]. The extracted features were then used for face recognition using SVM. The findings from the experiment demonstrate that the suggested approach provides an 87.59% classification accuracy.

2.5. Factors and Conditions Affecting the Precision of Recognising Emotional Feedback
Factors affecting the accuracy of the whole system can be analysed from different perspectives: firstly, from the user's point of view, if the recognised person wears glasses with thick lines or complex shapes and other decorations, this may interfere with and obscure the recognition points in key areas of the face, resulting in the possibility that only a small number of frames of the image are retained during the data screening phase, and therefore too few samples are available for a single category in the data classification phase, which may lead to erroneous analysis results [9]. In addition, if the user has large facial scars or suffers from a medical condition that prevents the facial nerves from producing the appropriate expression, the accuracy of the emotional feedback will also be affected.

Looking at objective factors, since the technology's recognition is done by using a camera to capture images of the human face, the lighting conditions in the environment can significantly affect the visibility and clarity of facial features, thus affecting the accuracy of emotion recognition. Different angles of the face can also affect accuracy. Cluttered backgrounds can lead to an inability to accurately locate faces, the front and back of the face can affect recognition results, and there are large differences in facial features between different countries and nationalities, and these differences can also affect recognition accuracy [19]. Not only is each person's facial expression almost unique in existence, preventing the system from using predetermined criteria to filter and categorize the received data, but also the differences in facial features, complexion and facial structure of different people can affect the generalisation ability of the emotion recognition system.

In addition to these factors, the most important are the precision of the system and the accuracy of the algorithms, as well as the adequacy of the samples in the database and their availability for matching with newly scanned data.

2.6. For testing of influencing factors, model training methods
The choice of dataset plays a very important role in dealing with the influencing factors of facial emotion recognition techniques. Different datasets serve different research purposes.

The Yale Face Database B (YALE-B), the CMU PIE Face Database, the CMU Multi-PIE Database and the Oulu CASIA datasets are made to investigate the effect of lighting circumstances on facial expression recognition [20]. The CMU PIE database is primarily used to study the effect of lighting conditions on FER. To obtain variations in light, the room was equipped with a flash system. It is divided into two main parts. The first part contains the variations of gestures and expressions and the second part contains the variations of poses and light [21]. CMU Multi-PIE was developed at Carnegie Mellon University and is based on the CMU PIE database [6].The Oulu-CASIA dataset was experimented with both near-infrared and visible light imaging systems and was photographed under three different lighting conditions: normal room light, low light, and dark light [22].

The Radboud Faces Database (RaFD), Karolinska Directed Expressional Face (KDEF) and Celebrities in Frontal-Profile (CFP) datasets have been developed with the objective of investigating the

impact of multi-angle faces on identifying emotions. Furthermore, the RaFD dataset can be utilised to evaluate the impact of cultural dissimilarities on the recognition of expressions [23]. As illustrated in Figure 1.

Figure 1. Exemplar photos from the RAF-DB collection [19]

Images of seven distinct human facial expressions, each seen from five different perspectives, taken from 70 individuals (50 percent male and 50 percent female) render the basis of the KDEF dataset [24]. The CFP dataset is designed for multi-angle face verification. Each participant produces 10 frontal and 4 lateral images, with the angle of change between each image less than 10° and the angle of separation between frontal and lateral close to 90° [25].

The purpose of the BAUM-1, FACES, and MUG datasets is to examine emotional intensity through the use of video sequences. The BAUM-1 database encompasses not only the six fundamental facial expressions (happy, angry, sad, disgusted, fearful, and surprised), but also boredom and contempt [26]. The FACES database contains six facial expressions: neutral, sad, disgust, fear, anger and happiness [27]. The data photos in the MUG database are divided into two categories: those created by users imitating facial expressions and those captured as they watched a film intended to provoke certain facial emotions [28].

The goal of the AffectNet dataset, the Acting Facial Expressions in the Wild (AFEW), the Static Facial Expressions in the Wild (SFEW), and the Real-world emotion Database (RAF-DB) is to examine how context affects emotion detection. A database of photos of people's expressions taken from actual settings is called RAF-DB [29]. The dynamic dataset AFEW contains facial expressions from movies that were filmed in settings similar to real life [30]. The SFEW database's photos come from movies and are more representative of reality [31]. As illustrated in Figure 2.

Figure 2. Exemplar photos from the SFEW collection [19]

More than a million photos make up the AffectNet dataset, which was gathered by utilizing expression-related metadata in search engine queries across many platforms, and not only that, almost half of these images have been manually labelled. The model is divided into two models: a classification model and a dimensional model [32].

The Binghamton University 4-Dimensional Facial Expression Database (BU-4DFE) and the Binghamton-Pittsburgh 3-D Dynamic Spontaneous Facial Expression Database (BP4D-Spontaneous) are relatively new databases of high-resolution 3D dynamic facial expressions. The most recent of them is the high-resolution 3D dynamic facial expression database, BU-4DFE [33]. The BP4D-Spontaneous database is constituted by data obtained from the Di3D dynamic face capture system, which encompasses both three-dimensional model sequences and two-dimensional texture videos. The latter are captured using two stereo cameras and one texture camera. The images of facial expressions were obtained through spontaneous expression in response to external stimuli, including video footage, actions, and odours [34].

3. Applications and examples of facial emotion data

The advancement of technology has enabled the intelligent system to interact with the user through facial recognition technology, thereby enhancing the robot's responsiveness to the user through emotion recognition. This is a fantastic example of how emotion recognition technology might advance further.

3.1. Detection of Fatigue-Hazardous Driving

The system utilises facial recognition technology to scan and track the driver's face, thereby extracting information pertaining to the state of the eyes and mouth. This is achieved through the utilisation of metrics such as PERCLOS, EAR and MAR, which are subsequently processed through a visual library. Furthermore, image processing and fatigue value calculations are conducted via the aforementioned visual library. In the event that the driver is identified as being in a state of fatigue or engaging in dangerous driving behaviour, an auditory alert will be issued by the system, prompting the driver to resume safe driving practices [35].The aforementioned facial recognition technique employs an estimate of the eye aspect ratio (EAR) and a specific parameter to provide real-time response results, thereby analysing the blinking phenomenon [9].

3.2. The application of intelligent hotel management techniques

The system employs facial recognition technology to facilitate intelligent hotel management. It utilises the OpenCV open-source computer vision library for face recognition and integrates a range of hardware components, including the STC15W4K58S4 microcontroller, Wi-Fi module, Bluetooth module, voice broadcasting module, and infrared sensors, to enable the functions of identity authentication, intelligent unlocking, voice control, and burglary alarm for hotel guests [36].

3.3. The Case for Facial Recognition-Based Welcome Robots

Luo Peng's thesis for his Master of Engineering degree at Harbin Institute of Technology states that this research improves the accuracy and variety of facial emotional responses by using a facial recognition system to identify the user's data, analyse the user's expression [37], and then determine and select the simulated 'emotional expressions' that the robot should display. MobileNet [38], a convolutional neural network which is lightweight, was considered for this research. When facial expression recognition systems perform expression recognition, they typically use geometric and textural features extracted from feature points to represent human emotions, including changes in the geometric structure of the face as well as changes in the texture of the facial skin [39]. Automated learning and classification of these feature points using deep learning models, especially deep neural networks, improves the accuracy and efficiency of emotion recognition by learning more complex representations of emotional features from large amounts of data[40].After receiving the user's facial expression information, the robot should autonomously judge the next action, for example, the hotel reception robot should receive the documents to be given to the receptionist for check-in after welcoming the guests, or the dissatisfaction to be sent to the data centre after delivering the goods in the guest room, and calculate the route to return to the standby area, and so on. Since this type of human-robot interaction environment requires high-precision route calculation, the exact optimal search methods in SLAM path planning are depth-first and breadth-first methods, which select the optimal solution after the robot has traversed all paths, and are optimised from the priority-defined breadth-first method and the Dijkstra algorithm. Although time-consuming, the robot needs to select the optimal path according to different scenarios to enhance the human-robot interaction experience in environments where interior decoration and furniture are everywhere.

4. Conclusion

In the context of human-computer interaction, facial recognition technology has a great deal of potential for identifying human emotions. By accurately capturing the user's facial features and combining them with deep learning algorithms, the system is able to effectively identify the emotional state and adjust the device's behaviour accordingly, thereby providing a more intelligent and personalised service experience. However, the implementation of this technology is not without its challenges. Factors such

as light, occlusion and facial angle have the potential to impact the accuracy of facial recognition, while the complexity of emotion recognition represents a further obstacle. Moreover, the high cost of the technology, response delays and limitations in the capacity of the emoji library present additional challenges that must be addressed. To this end, LED screens can be employed in lieu of mimetic materials to reduce costs, algorithms can be enhanced to improve recognition accuracy, and a cloud database can be created using 5G technology to reduce system latency. As technology continues to advance, the applications of facial recognition and emotion analysis will become increasingly prevalent in a diverse range of contexts, influencing the way smart devices interact and offering substantial benefits in domains such as smart homes, mental health monitoring, and public safety.

References
[1]   Lyons MJ, Akamatsu S, Kamachi M & Gyoba J (1998) Japanese Female Facial Expression (JAFFE) *Database. In Proceedings of the Third IEEE International Conference on Automatic Face and Gesture Recognition*, pp. 200-205.
[2]   Cohn JF, Ambadar Z, Ekman P (2007) Observer-based measurement of facial expression with the facial action coding system. *Neurosci Letters*, pp.203-221
[3]   Valstar M & Pantic M(2010). Induced disgust, happiness and surprise: An addition to the MMI facial expression database. In *Proceedings of the International Conference on Language Resources and Evaluation, Workshop on Emotion*, pp.65–70
[4]   Kanade T, Tian Y & Cohn JF (2002) Comprehensive database for facial expression analysis. In *Proceedings of the IEEE* International *Conference on Automatic Face & Gesture Recognition* (Grenoble, France), pp.46-53
[5]   Lucey P , Cohn JF, Kanade T et al. (2010) The extended Cohn-Kanade dataset (CK+): A complete dataset for action unit and emotion-specified expression. *In Computer Vision and Pattern Recognition Workshops* (San Francisco, CA, USA), pp.94-101
[6]   Kuang L, Zhang M & Pan Z (2016) *Facial expression recognition with CNN ensemble. In International Conference on* Cyberworlds (IEEE, Chongqing, China), pp.259-262
[7]   SatrioAgung E, Rifai AP & Wijayanto T (2024) Image-based facial emotion recognition using convolutional neural network on Emognition Dataset. *Scientific Reports*, 14:(14429)
[8]   Sultana S, Mustafa R & Chowdhury MS (2023) Human emotion recognition from facial images using convolutional neural network. In *Machine Intelligence and Emerging Technologies. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering* (Springer, Cham). Vol. 491, pp.141–153
[9]   Bisen D, Shukla R, Rajpoot N, et al. (2022) Responsive human-computer interaction model based on recognition of facial landmarks using machine learning algorithms. *Multimedia Tools and Applications*, 81:(14577-14599)
[10] Liu J, Wang HX & Feng YJ (2021) An end-to-end deep model with discriminative facial features for facial expression recognition. *IEEE Access*, 9:(12158–12166)
[11] Wang M, Hong J & Ying L (2010) Face recognition based on DWT/DCT and SVM. In *2010 International Conference on Computer Application and System Modeling* (IEEE, Taiyuan)
[12] Li C., Diao Y, Ma H & Li Y (2009) A statistical PCA method for face recognition. In *2008 Second International Symposium on Intelligent Information Technology Application* (IEEE, Shanghai, China)
[13] Sharma S, Verma R, Sharma MK, Kumar V (2016) Facial expression recognition using PCA. *International Journal of Scientific Research and Development,* 4(2):1905–1910
[14] Luo Y, Wu CM, Zhang Y (2013) Facial expression recognition based on fusion feature of PCA and LBP with SVM. *Optik - International Journal of Light and Electron Optics*, 124(17): 2767–2770
[15] Shan C, Gong S, Mcowan PW (2009) Facial expression recognition based on local binary patterns: a comprehensive study. *Image and* Vision *Computing*, 27(6):803–816

[16]   Mishra N, Bhatt A (2021) Feature extraction techniques in facial expression recognition. In *2021 5th International Conference on Intelligent Computing and Control Systems* (Madurai, India)

[17]   Yadav KS, Singha J (2020) Facial expression recognition using modified Viola-John's algorithm and KNN classifier. *Multimedia Tools and Applications*, 79(19):13089–13107

[18]   Kim S, An GH, Kang SJ (2017) Facial expression recognition system using machine learning. In *International SoC Design Conference* (Seoul, Korea)

[19]   Guo X, Zhang Y, Lu S, Lu Z (2023) Facial expression recognition: a review. *Multimedia Tools and Applications*, 83(8):23689–23735

[20]   Georghiades AS, Belhumeur PN, Kriegman DJ (2002) From few to many: Illumination cone models for face recognition under variable lighting and pose. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 23(6):643–660

[21]   Saito JH, Carvalho TVD, Hirakuri M et al. (April 27–29, 2005) Using CMU PIE human face database to a convolutional neural network - Neocognitron. In *ESANN 2005, 13th European Symposium on Artificial Neural Networks* (Bruges, Belgium) , Proceedings (pp.491–496)

[22]   Zhao G, Huang X, Taini M, Li SZ, Pietikäinen M (2011) Facial expression recognition from near infrared videos. *Image and Vision Computing*, 29(9):607–619.

[23]   Langner O, Dotsch R, Bijlstra G, Wigboldus DH, Hawk ST et al. (2010) Presentation and validation of the Radboud faces database. *Cognition and Emotion*, 24(8):1377–1388.

[24]   Lundqvist D, Flykt A, hman A (1998) The Karolinska directed emotional faces – KDEF [CD-ROM]. Department of Clinical Neuroscience, Psychology Section, Karolinska Institute.

[25]   Sengupta S, Chen JC, Castillo C, Patel VM, Jacobs DW (2016) Frontal to profile face verification in the wild. In *2016 IEEE Winter Conference on Applications of Computer Vision*. (Lake Placid, NY, USA), pp.1–9

[26]   Zhalehpour S, Onder O, Akhtar Z, Erdem CE (2017) BAUM-1: a spontaneous audio-visual face database of affective and mental states. *IEEE Transactions on Affective Computing*, 8(3):300–313

[27]   Ebner NC, Riediger M, Lindenberger U (2010) FACES—a database of facial expressions in young, middle-aged, and older women and men: Development and validation. *Behavior Research Methods*, 42(1):351–362.

[28]   Aifanti N, Papachristou C, Delopoulos A (2010) The MUG facial expression database. In *2010 11th International Workshop on Image Analysis for Multimedia Interactive Services* (IEEE, Desenzano del Garda, Italy), pp.1–4

[29]   Li S, Deng W, Du JP (2017) Reliable crowdsourcing and deep locality-preserving learning for expression recognition in the wild. In *2017 IEEE Conference on Computer Vision and Pattern Recognition* (IEEE, Honolulu, HI, USA), pp.2584–2593

[30]   Dhall A, Goecke R, Lucey S, Gedeon T (2011) Acted facial expressions in the wild database. Technical Report TR-CS-11-02. Australian National University.

[31]   Dhall A, Goecke R, Lucey S, Gedeon T (2011) Static facial expression analysis in tough conditions: Data, evaluation protocol and benchmark. In *IEEE International Conference on Computer Vision Workshops* (IEEE, Barcelona, Spain), pp.2106–2112

[32]   Mollahosseini A, Hasani B, Mahoor MH (1949) AffectNet: A database for facial expression, valence, and arousal computing in the wild. *IEEE Transactions on Affective Computing*, 10(1): 18–31

[33]   Yin L, Wei X, Yi S, Wang J, Rosato MJ (2006) A 3D facial expression database for facial behavior research. In *7th International Conference on Automatic Face and Gesture Recognition* (IEEE, Southampton, UK.), pp. 211–216

[34]   Xing Z, Yin L, Cohn JF, Canavan S, et al. (2013) A high-resolution spontaneous 3D dynamic facial expression database. Image *and Vision Computing*, 32(10):692–706

[35]   Huang CC, Zhao HM, Chen YC, et al. (2021) A system for the detection of hazardous driving due to driver fatigue based on facial recognition.*Electromechanical Engineering Technology*, 50(12):143–146

[36] Yu Yanxiu, Yu Rui, Yan Feijie (2021) Intelligent hotel management system based on facial recognition. *Journal of Intelligent Systems*, , 50(1):116

[37] Luo Peng (2020, April) A study of facial recognition-based reproduction control of anthropomorphic expressions in a welcoming robot. Master of Engineering Thesis, Harbin Institute of Technology.

[38] Kalpana Chowdary M, Nguyen TN, Hemanth DJ (2021) Deep learning-based facial emotion recognition for human–computer interaction applications, *Neural Computing and Applications*, 35(10):692–706

[39] Palanichamy N (2023) Occlusion-aware facial expression recognition: A deep learning approach. *Multimedia Tools and Applications*, 82(28):32895–32921

[40] Ujjwal MK, Parameswaran S, Chowdary VG, et al.(2024) Comparative analysis of facial expression recognition algorithms. In Data Science and Security (Lecture Notes in Networks and Systems, vol. 922) (pp. 419–431)