IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429

Volume 6, Issue 2, May 2026

Impact Factor: 8.2

# Facial Emotion Recognition using AI

**Mrs. Swati Y. Kale[1], Mr. More Sarthak[2], Mr. Marale Yash[3],**
**Miss. Nhavakar Priya[4], Miss. More Vedshri[5]**

Prof. Computer Engineering Department, Adsul's Technical Campus, Ahilyanagar, India [1]

Students, Computer Engineering Department, Adsul's Technical Campus, Ahilyanagar, India [2,3,4]

Students, Information & Technology Engineering Department, Adsul's Technical Campus, Ahilyanagar, India[5]

**Abstract:** *This paper introduces a study employing artificial intelligence (AI) to utilize computer vision algorithms for detecting human emotions in video content during user interactions with diverse visual stimuli. The research aims to unveil the creation of software capable of emotion detection by leveraging AI algorithms and image processing pipelines to identify users' facial expressions. The process involves assessing users through images and facilitating the implementation of computer vision algorithms aligned with psychological theories defining emotions and their recognizable features. The study demonstrates the feasibility of emotion recognition through convolutional neural networks (CNN) and software development and training based on facial expressions. The results highlight successful emotion identification; however, precision improvement necessitates further training for contexts with more diverse images and additional algorithms to distinguish closely related emotional patterns. The discussion and conclusions emphasize the potential of A.I. and computer vision algorithms in emotion detection, providing insights into software development, ongoing training, and the evolving landscape of emotion recognition technology. Further training is necessary for contexts with more diverse images, alongside additional algorithms that can effetely distinguish between facial expressions depicting closely related emotional patterns, enhancing certainty and accuracy.*

**Keywords:** facial emotion, recognition, A.I., convolutional neural network (CNN), Transfer Learning

## I. INTRODUCTION

A facial expression is exhibited by the movement of muscles underneath the face skin. Facial emotion recognition is the process of detecting human emotions from facial expressions. The facial expression for emotion detection has always been a challenging task in achieving through computer algorithms. In the field of Artificial Intelligence, Facial Emotion Recognition (FER) is an active research area, with several recent studies. With the recent advancement in computer vision and machine learning, it is possible to detect emotion from images. The automatic emotion recognition based on facial expression is an interesting research field, which has presented and applied in several areas such as safety, health and in human machine interfaces. The human brain recognizes emotions automatically, and software has now been developed that can recognize emotions as well. This technology is becoming more accurate all the time, and will eventually be able to read emotions as well as our brains do. In this project we propose a technique called facial emotion recognition using convolutional neural networks and haar cascade classifier.

Human facial expressions that people see visually are all around them. They are natural signals that help them understand emotions from any person in front of them or via images or videos. These emotions are highly complex and challenging to understand for machines but easily understandable by humans. To understand how humans could understand such emotions, Mehrabian, a famous psychologist, found from his research that the emotional data that humans classify as emotions are distributed in sections. He found that only 7% of the emotional data total is passed by language, and 38% is transported by our language auxiliary, which differs from culture to culture, such as the rhythm of speech, tone, pitch, etc. So far, the highest percentage of emotional data shown by facial expression is 55% [1]. This indicates that many sensible emotional data can be obtained by recognizing facial emotions that effectively understand any human's state of mind and actions directly associated with emotions [2]. So, it is essential to explore this research

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, May 2026
Impact Factor: 8.2

domain in more detail as less accurate systems plague its commercial implementation. Human facial emotion recognition has been broadly used in numerous human-computer interactions such as smartphones, affective computing, intelligent control systems, psychological, behavioral study, pattern searching, defense, social sites, robotics, and other fields [3]–[5]. By evaluating these emotions, one could deliver maximum user satisfaction and feedback to improve current technologies. This can only be done in the domains of computer vision and deep learning. To create several Facial Emotion Recognition (FER) systems that have been evaluated for encoding and transmitting Information from facial representations. In the twentieth century, Ekman and Friesen identified six fundamental emotions based on cross-cultural research that revealed that humans convey these fundamental emotions in the same way regardless of culture [6]. Face expressions include anger, disgust, fear, happiness, sadness, and surprise. Contempt was later added to this list of feelings. To do Facial Emotion Recognition, there are basic initial steps which are divided into three essential stages. The facial features of the face are detected from the entire frame of a video at the first stage, which is a pre-processing stage. The eyebrows, brows, nose, mouth, and chin are among the facial features. More descriptive features from different areas of the Face are removed in the second level. Likewise, more descriptive features from different areas of the Face are removed in the second level.

## II. METHODOLOGY

Face expressions include anger, disgust, fear, happiness, sadness, and surprise. Contempt was later added to this list of feelings. To do Facial Emotion Recognition, there are basic initial steps which are divided into three essential stages. The facial features of the face are detected from the entire frame of a video at the first stage, which is a pre-processing stage. The eyebrows, brows, nose, mouth, and chin are among the facial features. More descriptive features from different areas of the Face are removed in the second level. Likewise, more descriptive features from different areas of the Face are removed in the second level. Finally, a classifier is trained using the training data before generating labels for the Emotions, illustrated in Figure 1.

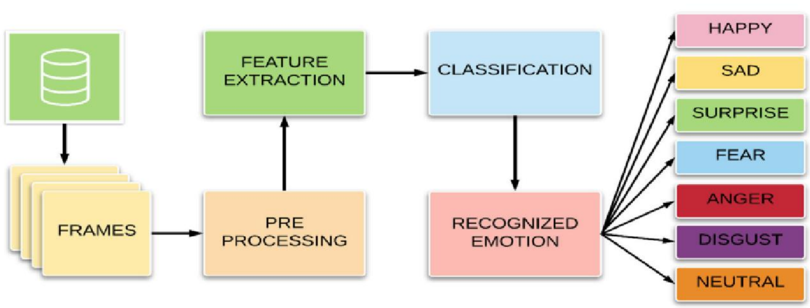


Fig.1 Facial Emotion Classification Process

Recent advancement in neuroscience and psychology research has sparked a debate that Ekman's model of six basic emotions claimed to be universal is culture-specific and not universal. Because of this, it has raised questions about whether emotions differ based on gender, age, and culture. Today the need for emotion classification has surpassed the barrier of age because of the global shift towards online platforms such as online education to teach or gain knowledge virtually globally to all remote areas, IoT enabled health monitoring systems and temperature setters in cars and households, robotics, psychiatric evaluation based on violent behaviors of criminals or those mentally disturbed, mood swings study on adolescents to help guide them mentally, deepfake detection, gaming, and many such applications are currently being innovated using state of the art technologies. Also, numerous studies have been conducted on Facial Emotion Recognition by using Computer vision because of its practicality in intelligent robotics, health-related treatment, IoT, Security surveillance, criminal psychological analysis, observation of driver exhaustion, and other humancomputer interfaces mechanisms [7]–[9]. With more virtual connectivity through videos and images, the need to adopt the latest technology based on people's emotions is now a critical factor in driving user-friendliness and

IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429 Volume 6, Issue 2, May 2026 Impact Factor: 8.2

maximum user satisfaction. Emotions are nothing but a cognitive state or phase perceived by a human and associated with moods. Usually, these emotions are often twisted with attitude, temper, character, disposition, and motivation. They can also be defined into binary sentiments such as positive(pleasure) or negative (displeasure) under different circumstantial psychological tasks or events. Such emotions bend a person's mind psychologically that the behavior of humans changes over time. Humans handle these emotions by either behavioral response, psychological states triggered by any events or by a person in front of themselves, subjective experience of the situation, and cognitive processes. Humans understand that emotions are not easy to quantify or replicate artificially from this complex set of actions. Many researchers use their version of emotion definitions and assumptions. This makes research in human facial emotions troublesome because all the studies that have been done have significant variance in them and do not draw a generalized conclusion. Although all humans have naturally occurring sets of emotions that can be perceived even cross-culturally, this is also mentioned in the Discrete Emotion Theory, which says that such emotions are distinguishable by an individual's features [10]. Ekman claimed that these emotions are perceived by humans not only culturally but also universally. His proposed model suggested that emotions are categorized into Fear, Happiness, Sad, Surprise, Disgust, and Anger. These categorical emotions are classified using facial and vocal data, which allows them.
Along with the category, the number of samples for each emotion also varies from each other. Hence, it is necessary to use a balanced dataset for a good result. Figure 4 and Figure 5, respectively, below show the difference between adult (RML) and kids' (LIRIS) datasets. There is a difference between the category of emotions as well as the recording conditions. RML dataset consists of 8 posed emotions recorded in a controlled environment whereas the LIRIS dataset has 6 spontaneous emotions recorded from a webcam. Recording emotions in a controlled environment gives RML and edge over LIRIS dataset in terms of quality. Apart from dataset quality and emotion category, the facial features also differ from each other. Apart from dataset quality and emotion category, the facial features also differ from each other. Apart from Plutchik's model, which depicts the well-known wheel of emotions that classifies emotions, the wellknown Circumplex Model of Affects is also illustrated in Figure 3, proposed in a study [12] comparable to Plutchik's model. It is divided into four portions: arousal (activation/deactivation) and valence (pleasant/unpleasant) axes. Every emotion depicted directly results from linear combinations of these two parts of varying degrees of valence and arousal. Four quadrants are created by combining

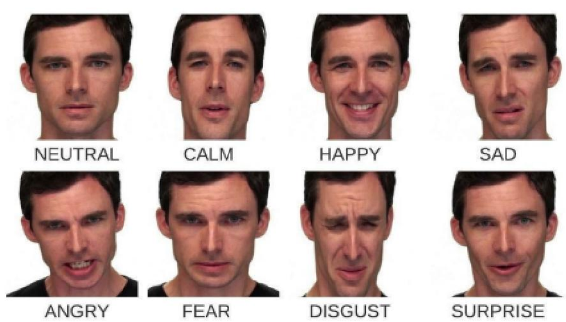


Fig.2 Adult emotions (RML dataset).

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, May 2026
Impact Factor: 8.2


Fig.3 . Kids' emotions (LIRIS dataset) ( the emotion labels font vary based on the datasets).

## III. TECHNOLOGY USE

Our system represents a comprehensive integration of cutting-edge technologies, primarily focusing on Image Processing and Machine Learning, to develop an advanced emotion classification system. At the project's inception, we employed Haar cascade classifiers, a highly precise face detection technique. These classifiers form the foundation of our system, ensuring accurate identification of facial regions for subsequent analysis. Moving forward, Image Processing techniques were strategically applied to translate facial images into a digital format, setting the stage for further computational analysis.

A critical aspect of our methodology involved the extraction of features from key facial components such as eyes, brows, nose, and mouth. This meticulous process aims to capture the intricate details of facial expressions, facilitating a nuanced understanding of the emotional content conveyed through these expressions. By dissecting facial features, our system gains the capability to discern subtle cues and variations, contributing to the overall accuracy of emotion classification.

The core Machine Learning component of our system is embodied by Convolutional Neural Networks (CNN). These sophisticated neural networks serve as the cornerstone of our emotion classification model. Trained on a wealth of data, the CNNs learn to associate extracted facial features with specific emotional states, enabling them to make informed predictions during the classification phase. The adaptability and learning capacity of CNNs make them particularly well-suited for this task, as they can discern complex patterns and relationships within the facial features dataset.

To ensure a well-organized and efficient project structure, we adopted a modular approach. This approach encompasses distinct modules dedicated to key stages of the emotion classification pipeline, including face detection, feature extraction, classification, and evaluation. Each module contributes to the seamless functioning of the overall system, emphasizing both effectiveness and real-time processing capabilities.

our project represents a harmonious integration of Image Processing and Machine Learning technologies, culminating in an advanced emotion classification system. The utilization of Haar cascade classifiers, Image Processing techniques, and Convolutional Neural Networks collectively form a robust framework for real-time emotion recognition. The comprehensive and modular nature of our approach positions the project at the forefront of advancements in the burgeoning field of Facial Emotion Recognition.

### A. CNN Algorithm:

A CNN is a Deep Learning algorithm which takes an input image, assigns importance to various aspects/objects in the image and is able to differentiate between images. The preprocessing required in a CNN is much lower than other classification algorithms. Figure shows the CNN operations. The architecture of a CNN is analogous to that of the connectivity pattern of neurons in the human brain. One role of a CNN is to reduce images into a form which is easier to process without losing features that are critical for good prediction. This is important when desi FER Dataset consists of 48x48 pixel grayscale images of faces. The faces have been automatically registered so that the face is more or less centered and occupies about the same amount of space in each image. The task is to categorize each face based on the emotion shown in the facial expression in to one of seven categories. These are: 0=Angry 1-Disgust 2=Fear 3-Happy 4-Sad 5=Surprise 6-Neutral. When using CNN algorithm the train.csv contains two columns, "emotion" and "pixels". The "emotion" column contains a numeric code ranging from 0 to 6, inclusive, for the emotion that is present in the image. The "pixels" column contains a string surrounded in quotes for each image. The contents of this string a space-separated pixel values in row major order. test.csv contains only the "pixels" column and your task is to predict the emotion column

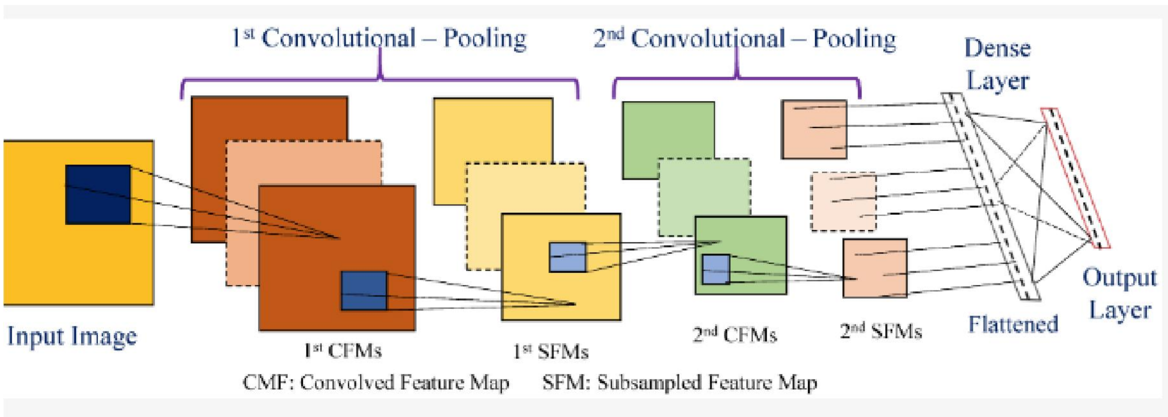


Fig. 2 The generic architecture of a convolutional neural network with two convolutional-pooling layers.

**B. Facial Emotion Recognition using TL in Deep CNNs:**

The available VGG-16 model is trained with the ImageNet dataset to classify 1000 image objects. The pre-trained model is modified for emotion recognition redefining the dense layers, and then fine-tuning is performed with emotion data. In defining the architecture, the last dense layers of the pre-trained model are replaced with the new dense layer(s) to recognize a facial image into one of seven emotion classes (i.e., afraid, angry, disgusted, sad, happy, surprised, and neutral). A dense layer is a regular, fully connected, linear layer of a NN that takes some dimension as input and outputs a vector of the desired dimension. Therefore, the output layer contains only seven neurons. The fine-tuning is performed on the architecture having the convolution base of the pre-trained model plus the added dense layer(s). A cleaned emotion dataset prepared through preprocessing (i.e., resizing, cropping, and other tasks) is used to train in fine-tuning. In the case of testing, a cropped image is placed to the input of the system, and the highest output probability of emotion is considered to be the decision. VGG-16 may be replaced with any other DCNN models, e.g., ResNet, DenseNet, Inception. Therefore, the size of the proposed model depends on the size of the pre-trained model used and the architecture of the added dense layer(s).

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, May 2026
Impact Factor: 8.2

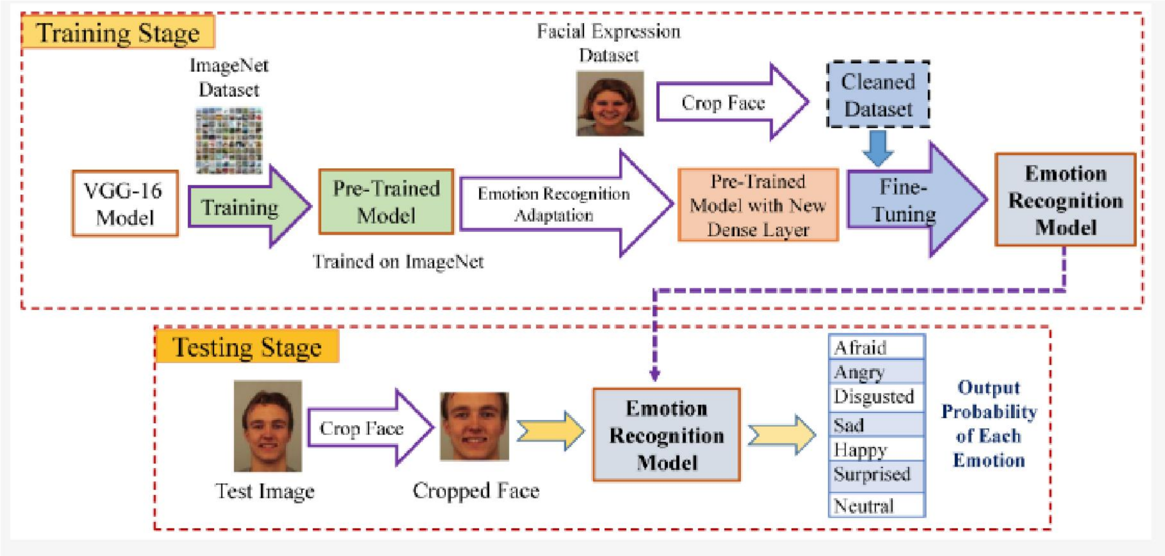


Fig. 3. Illustration of the proposed FER system based on transfer learning in deep CNN.

**C. Experimental Studies:**

This section investigates the efficiency of the proposed FER system using TL on DCNN on two benchmark datasets. Firstly, a description of benchmark datasets and experimental setup are presented. Finally, the outcome of the proposed model on the benchmark datasets is compared with some existing methods to verify the effectiveness of the proposed model.

There are few datasets available for the emotion recognition problem; among those, Karolinska Directed Emotional Faces (KDEF) [75] and Japanese Female Facial Expression (JAFFE) [76] datasets are well-known and considered in this study. Images of the datasets are categorized into seven different emotion classes: Afraid (AF), Angry (AN), Disgusted (DI), Sad (SA), Happy (HA), Surprised (SU), and Neutral (NE). The brief description and selection motivation of the datasets are given below.

The KDEF [75] dataset (also refer as KDEF for simplicity, henceforth) was developed by Karolinska Institute, Department of Clinical Neuroscience, Section of Psychology, Stockholm, Sweden. The dataset images were collected in a lab environment, so the emotion of the participants was artificially created. Specifically, the purpose of the dataset was to use for perception memory emotion attention, and backward masking experiment. Although the primary goal of the material was not emotion classification, it is popular for such a task because medical and psychological issues sometimes related to emotion. The dataset contains 4900 images of 70 individuals, each expressing seven emotional states. Photos of an individual were taken from five different angles, which resemble frontal (i.e., strait) view and four different profile views (full left, half left, full right, and half right). In the angular value variation point of view, images are from −90° (full left) to +90° (full right). In a full left or full right profile view, one side of the face with only one eye and ear is visible and makes FER more challenging. Some sample images from the KDEF dataset are shown in Figure 5. FER from the dataset is challenging due to the diversity in images with different profile views along with the frontal view. Profile views mimic the expectation of FER from different angular positions, and therefore, the complete dataset is considered in this study to evaluate the efficiency of the proposed method for such critical cases, which is necessary for industrial applications. Moreover, a few studies are available with the dataset, but they are mostly based on 980 frontal images .


Fig.5. Sample images from KDEF dataset.

## IV. CONCLUSION

The purpose of this work was not to get conclusive results but to bear out the main challenges and difficulties involved in emotion detection from facial expressions. The project helps in identifying the different emotions of a person which are angry, disgust, happy, sad, fear, surprise and neutral. These facial emotions can be mainly used in E-learning, criminal investigation, online gaming, psychology and many other applications. Using a public dataset of train and test samples for seven different classes of human emotions, we build and train a convolutional neural network model and use it along with haar cascade classifier to produce the final predicted emotion of an uploaded image.

Future Scope:

As emotion recognition technology becomes more sophisticated and more deeply embedded in our array of devices, it will become expected that our computers and phones provide us with a continual progression of customized triggers and messaging. The technology will be found even in future car dashboards, refrigerator doors, and conference room walls etc. The future enhancements of this project are:

- To increase the accuracy
- To predict the compound emotions

This work can be carried out by others in order to convert it into a model for predicting the emotions of a person by using the real time capturing of the face. Similarly, the other advancement of this project may include the extraction of images from the videos and predicting the emotions from those images.

## ACKNOWLEDGMENT

It gives us great pleasure in presenting the paper on “Facial Emotion Recognition using AI”. We would like to take this opportunity to thank our guide, Prof. Swati Y. Kale, Professor, Computer Department, Adsul’s technical Campus, Ahlyanagar, for giving us all the help and guidance we needed. We are grateful to her for hers kind support, and valuable suggestions were very helpful.

## REFERENCES

[1]. Albaladejo, X., Díaz, J. R., Quesada, A. X., and Iglesias, J. (2021). Proyectos agiles.org. Available online at: https://proyectosagiles.org/pm-partners (accessed July 12, 2023).

[2]. Banafa, A. (2016). Qué es la computación afectiva? OpenMind BBVA. Available online at: https://www.bbvaopenmind.com/tecnologia/mundo-digital/que-es-lacomputacion-afectiva/ (accessed September 14, 2023).

[3]. Bledsoe, W. W. (1966). Man-Machine Facial Recognition: Report on a Large-Scale Experiment. Technical Report PRI 22. Palo Alto, CA: Panoramic Research. Centeno, I. D. P. (2021). MTCNN Face Detection

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 2, May 2026
Impact Factor: 8.2

Implementation for TensorFlow, as a Pip Package. Available online at: https://github.com/ipazc/mtcnn (accessed September 14, 2023)

[4]. Chollet, F. (2017). "Xception: deep learning with depthwise separable convolutions," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (Honolulu, HI), 1251–1258.

[5]. Darwin, C., and Prodger, P. (1996). The Expression of the Emotions in Man and Animals. Oxford: Oxford University Press. Ekman, P. (1994). Strong evidence for universals in facial expressions: a reply to Russell's mistaken critique. Psychol. Bull. 115, 268–287. doi: 10.1037/0033-2909.115.2.268

[6]. Ekman, P. (1999). Basic emotions. Handb. Cogn. Emot. 3, 45–60. doi: 10.1002/0470013494.ch3

[7]. Ekman, P., Sorenson, E., and Friesen, W. (1969).

[8]. Pan-cultural elements in facial displays of emotion. Science 164, 86–88. doi: 10.1126/science.164.3875.86 Frijda, N. H. (2017). The Laws of Emotion. London: Psychology Press.

[9]. García, A. R. (2013). La educación emocional, el autoconcepto, la autoestima y su importancia en la infancia. Estudios y propuestas socioeducativas. 44, 241–257.

[10]. Ghotbi, N. (2023). The ethics of emotional artificial intelligence: a mixed method analysis. Asian Bioethics Rev. 15, 417–430. doi: 10.1007/s41649-022-00237-y

[11]. Hernández Sampieri, R., Fernández, C., and Baptista, L. C. (2003). Metodología de la Investigación. Chile: McGraw Hill.

[12]. Kaggle (2019). FER−2013. Available online at: https://www.kaggle.com/ (accessed October 5, 2023).

[13]. Krizhevsky, A., Sutskever, I., and Hinton, G. E. (2017). ImageNet classification with deep convolutional neural networks. Commun. ACM 60, 84–90. doi: 10.1145/3065386

[14]. Lee, Y. S., and Park, W. H. (2022). Diagnosis of depressive disorder model on facial expression based on fast R-CNN. Diagnostics 12:317. doi: 10.3390/diagnostics12020317

[15]. Lu, X. (2022). Deep learning based emotion recognition and visualization of figural representation. Front. Psychol. 12:818833. doi: 10.3389/fpsyg.2021.818833