Journal of Engineering Science and Technology
Special Issue on ICIT2022 Vol. 19, No. 2 (2024) 163 - 172

DESIGNING A RESEARCH MODEL FOR DEPRESSION TENDENCY DETECTION USING FACIAL EXPRESSION RECOGNITION

NG KWAN LUNG, CHONG PUI LIN*, DESHINTA ARROVA DEWI

Faculty of Data Science and Information Technology,
INTI International University, Malaysia
*Corresponding Author: puilin.chong@newinti.edu.my

Abstract

Most people across the world experience stress and depression at various points in their lives, and this is due to a variety of factors. There are many demands on people today, ranging from young educational activities to competitive and challenging duties for young or middle-aged people, employment stress, family commitments, different types of interpersonal management, health issues, and the elderly. University students, however, are under a lot of strain because they must strike a balance between the demands to excel in their studies and the pressure to land a specific vocation in the future. To assess the levels of depression and mental wellness among university students, the authors aim to design a preliminary study by developing a research model for depression tendency detection using machine learning and deep learning. To identify university students who are at risk for depression, the authors suggest Face Expression Recognition (FER) technology in this model. Python, OpenCV, Machine Learning, Deep Learning - Convolutional Neural Network (CNN), and JAFEE and CK+ datasets are the suggested tools for the model. In addition, the authors recommend several facial expression recognition processes and pipelines, such as input image, pre-processing, facial component detection, feature extraction, classifier, and output, as these are better suited and more effective for predicting the tendency for depression in university students.

Keywords: Computer vision; Convolutional neural network, Depression tendency, Facial expression recognition; JAFEE; CK+; Machine learning; Public health.

1. Introduction

The World Health Organization (WHO) estimates that 350 million people globally, across all age groups, suffer from depression. One of the most significant but frequent mental illnesses in the world is depression, often known as depressive disorder. Depression can seriously affect a person's capacity to perform daily activities and interact with or restrict that ability [1]. If the symptoms are minor, the person will feel down and lose faith in everything; if the symptoms are severe, the person may have trouble paying attention, sleeping, or even have suicidal thoughts. Depression can affect anyone at any age, and due to its widespread prevalence and severe impact, it is receiving increasing attention from society. Although depression is a serious condition, it can be treated with medication, psychotherapy, and other professional procedures. However, the majority of traditional processes for diagnosing depression are based on the clinical interview reports provided by the patients themselves and the behaviour questionnaires completed by their loved ones and friends [2]. However, the traditional techniques for diagnosing depression rely on subjective assessment, and the results vary depending on the circumstances and the time of day. Additionally, several psychologists or clinical research experts must be included in the diagnosis process to acquire an objective evaluation, which is very time-consuming and constrained for monitoring the diagnosis and evaluating the effectiveness of treatment [3].

Students of all ages, whether in high school, college, or graduate school, experience depression continuously. One in fifteen college students has experienced some form of depression, and the average age at which they encounter it ranges from teens to 25 years old. Despite the fact that depression is frequent among university students, it is crucial to treat it properly. Nearly every element of university life, including classes, social life, and studies, is impacted by depression. Depression can also result in thoughts of self-harm and suicide if it progresses to its last stage. In order to minimize the possible harm in actual society, facial expression recognition based on machine learning and deep learning can detect the mental condition of university students more swiftly and objectively [4].

The Facial Expression Recognition technique is used to examine facial expressions and emotions in various media, including photos and videos. The ability to recognize and introduce human emotion and emotional states for study and calculation is known as emotional computing, and facial expression recognition is typically the computing technique used. In recent years, the interpretation of human emotional expression is no longer a research interest in the field of psychology, which also involves the field of human-computer interaction technology. Facial expression recognition technology has advanced due to the advent and development of camera technology, biometric analysis, machine learning, deep learning, and pattern recognition in recent years. A facial expression recognition system can more quickly detect a person's emotion and intention by using facial expression, which is the most important information clue. Basic emotions, such as happiness, sadness, anger, surprise, fear, and disgust, are recognized through facial expression recognition. Each expression is connected to a specific facial expression, even in various settings and at various times [5].

Artificial intelligence technology is also the foundation of this technology. It is hard to discuss artificial intelligence without mentioning its two main aspects, machine learning, and deep learning. One would argue that machine learning and deep learning are paving the way for the automation industry's next technological advancement. Convolution Neural Network (CNN) and Deep Neural Network (DNN) are new and improved techniques used by facial expression recognition technology to evaluate human expressions and emotions. Both methods use computer vision to identify faces and calculate a person's stress and despair in order to determine their mood and state of mind [6].

2. Methodology

The methodology or research process employed in this study starts with a definition of the problem statement, which addresses the depressive tendency now present in young and middle-aged people, particularly university students. The next phase in the research process is to specify the research scope, which determines the breadth of the study area's investigation. Research questions and objectives are then developed after that. The next step in the study technique is to identify applicable datasets, instruments, and techniques, with a focus on the recognition of facial expressions and depression tendencies in young and middle-aged persons. Performance measurements for the proposed technique are produced before the paper's primary goal, which is the development of a research model. Figure 1 offers a synopsis of this explanation.

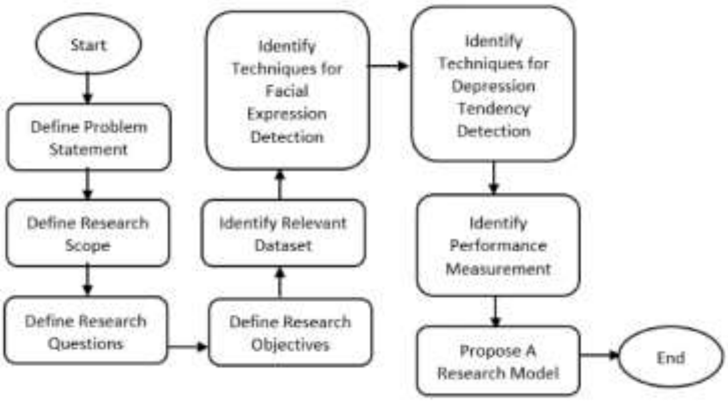

Fig. 1. The research methodology used in this study.

The next crucial step after identifying the research objectives is locating the pertinent dataset that will be used to train the facial expression identification module. It takes a lot of data collection and data processing to recognize facial expressions. Facial expression detection systems use techniques like machine learning and deep learning, both of which need a lot of data to create a database and make precise calculations. It is essential to have sufficient labelled training data with as many variations of the populations and environments as is practical for the development of a deep expression recognition system [7]. More explanation of the dataset used in this study is outlined in the next section below.

3. Results and Discussion

This section highlights the result summary of the methodology used in this paper. According to literature that has been highlighted by previous researchers, we decided to focus on the problem of assessing the levels of depression and mental wellness among university students [8].

Accordingly, we define research questions and objectives as below:

RQ1 How a research model for detecting depressive tendencies be created?
RQ2 Which datasets are useful for research on detecting depression tendencies?
RQ3 What methods can be used for facial expression recognition?

RO1 To determine the steps necessary to create a research model focusing on depression tendency identification.
RO2 To select the relevant datasets for detecting depression tendency.
RO3 To determine the most effective methods for facial expression recognition.

In terms of datasets, we recommend three datasets that have been utilized and proposed by earlier scholars [8]. They are JAFEE, CK+, and FER2013, and Table 1 shows further information about each.

Table 1. Datasets used in this paper.

Database	Samples	Subject	Expression distribution
JAFEE	213 images	10	6 basic expressions + neural
CK+	593 image sequences	123	6 basic expressions + contempt and neural
FER2013	35,887 images	N/A	6 basic expressions + neural

3.1. JAFEE

The JAFFE database is frequently used for facial expression analysis. It consists of 213 still images of neutral faces positioned in controlled environments with six universal facial expressions (happy, sad, fear, furious, disgust, and surprise) as depicted in Fig. 2. Each statement has three to four examples presented to each subject. The image size is 256×256 pixels, and the ground truth semantic judgments of the expressions were made on seven different emotion categories by 60 female volunteers. The annotations for the expression images are created based on an image's dominant expression. Because the eyelids and overall shape of our test subject's face differ slightly from those in the JAFEE dataset. Therefore, just a portion of the data collection - such as eyebrows - is used in this paper [9].

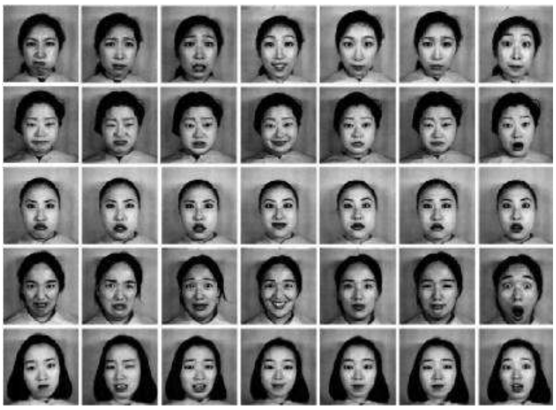

Fig. 2. Example of JAFEE dataset.

3.2. CK+

For evaluating FER systems, the Extended CohnKanade (CK+) database is the most extensively used laboratory-controlled database. 593 video clips from 123 different topics may be found on CK+. The segments show a shift in facial expression from neutral to peak and range in duration from 10 to 60 frames. The seven fundamental expressions of anger, contempt, disgust, fear, happiness, sorrow, and surprise are assigned to 327 sequences from 118 individuals based on the Facial Action Coding System (FACS). Because CK+ does not provide particular training, validation, or test sets, the algorithms tested on this database are not uniform. The most often used method of data extraction for static-based techniques is to record the first and last one to three frames of each sequence that contain peak formation (neutral face) [10].

3.3. FER2013

The FER2013 database as depicted in Fig. 3 was made public during the ICML 2013 Challenges in Representation Learning. The Google image search API automatically created a massive, unrestricted database known as FER2013. After being rejected from frames with improper labels, all photographs were registered, cropped, and reduced to 48*48 pixels. FER2013 includes 28,709 training images, 3,589 validation images, and test images with seven expression labels (anger, disgust, fear, happiness, sadness, surprise, and neutral) [11].

Fig. 3. Example of FER 2013 dataset.

The six action units that were determined using Jupyter Notebook were used as characteristics in the classification process after the authors trained the dataset. Figure 4 shows typical AU values for a facial expression example. Numerous emotions are depicted in the images, including neutral, happy, surprised, angry, sad, afraid, and disgusted [7].

ES	neutral	joy	surprise	anger	sadness	fear	disgust
AU0	0.21	0.77	-0.10	0.30	0.17	-0.11	0.91
AU1	-0.06	0.09	0.60	-0.07	-0.04	0.20	0.13
AU2	-0.25	1.00	-0.49	0.06	-0.37	-0.60	0.88
AU3	-0.21	0.00	-0.13	0.04	-0.09	-0.17	0.00
AU4	-0.04	-0.47	0.58	-0.19	-0.02	0.28	-0.32
AU5	-0.23	-0.30	0.10	-0.34	-0.27	-0.02	-0.39

Fig. 4. The facial expression and corresponding AU.

Figure 5 shows an illustration of the AU2 and AU5 measurement distribution for Subject #3. Even with a simple analysis of the distribution of the traits, we can see that some emotions are detectable. Our research decided to investigate the possibility of using AU to automatically detect emotions. The tests made use of neural network classifiers and K-nearest neighbours.

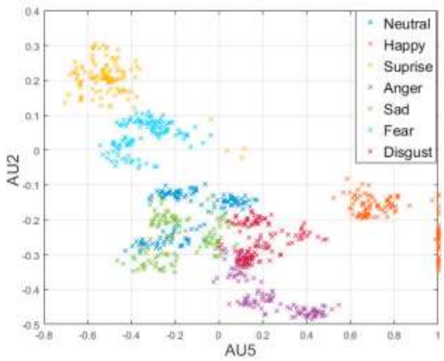

Fig. 5. Example of K-nearest neighbours.

For Facial Expression Recognition (FER), we recommend the following steps illustrated in Fig. 6. To compute and recognize different facial emotions more quickly, a variety of processing stages must be done to the image.

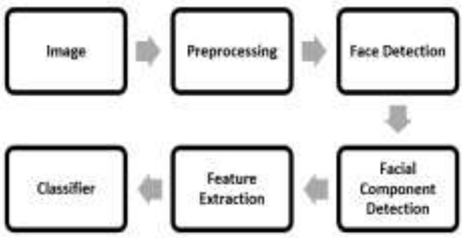

Fig. 6. Pipeline of facial expression recognition.

After a series of images arrive in the model, pre-processing is the first step in the procedure. The initial action in this phase is to ensure consistency in the input image for pre-processing. The proposed pre-processing of the FER system is to use face alignment and face augmentation. Face alignment is a frequent pre-processing step in many tasks involving the recognition of faces. A collection of popular deep FER algorithms publicly available for implementation. After employing a set of training data to identify the human face, background and non-facial features are then removed. One kind of data (face) augmentation is on-the-fly augmentation, which makes it a crucial step in the deep FER procedure. The process of creating additional data points from current data in order to artificially increase the amount of data is known as data augmentation. This includes enhancing the dataset by making small changes to the data or by creating new data points using machine learning models to explore the latent space of the original data. On-the-fly data augmentation is typically included in deep learning toolkits to lessen overfitting [12].

The pre-processing stage of the standardized image is followed by the face detection stage, which extracts the single face and completely removes any unfavourable backdrops. Histogram of Oriented Gradient is a method of face detection that is often used (HOG). Within a particular region of an image, the algorithm counts instances of gradient orientation. The HOG description emphasizes an object's composition or form. It is superior to other edge descriptors because it computes the features using both the magnitude and the angle of the gradient [6].

This stage involves tracking and identifying various facial features using a detector known as the Action Unit, which can pinpoint the precise site of facial muscle contraction (AUs) as shown in Fig. 7. The examinations concentrate on the eyes, brows, lips, nose, and ears. The basic actions of a single muscle or a group of muscles are referred to as action units (AUs). The metrics in the Emotion Facial Action Coding System are AUs rather than muscles for two reasons. First, two or more muscles are combined into a single AU for various appearances because the changes in appearance they produce cannot be distinguished. Second, the appearance changes brought about by a single muscle are occasionally separated into two or more AU to represent the somewhat autonomous activity of different muscle components [11].

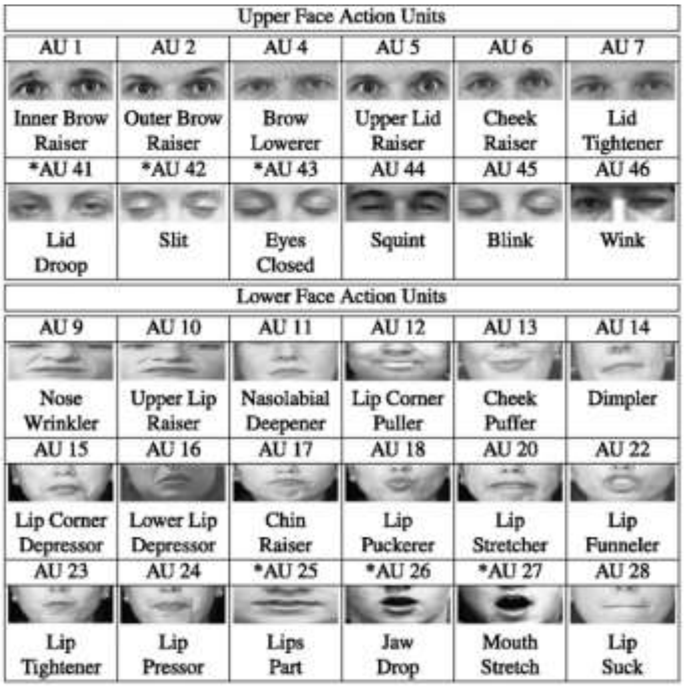

Fig. 7. Upper face action units and lower face action units.

Subsequently, feature extraction is a crucial phase of the Facial Expression Recognition system. The basic objective of facial feature extraction is to produce an accurate and useful representation of facial features while preserving facial information. Currently, feature extraction makes extensive use of Principal Component Analysis (PCA) and geometric-based methods. The authors reduce dimensionality by using PCA. To characterize the shape and placement of facial components, the authors also employ geometric-based extraction algorithms, which rely on predetermined geometric landmark positions [13].

The most distinctive aspects of face images are extracted and prepared, and then the images are categorized using a specific classification technique to the correct class label of facial emotions. The authors implement the Support Vector Machine to establish the best line or decision boundary that can divide n-dimensional space into classes, allowing us to quickly classify fresh data points in the future. Other than that, the authors also implement neural networks into the FER system. Different types of neural networks exist depending on their performance and properties. Convolutional Neural Networks are most frequently utilized in FER systems. In this study, we proposed an emotion recognition system that uses a single network with a single hidden layer neural network classifier to categorize facial expressions with selected dynamic variables as suggested by previous research [12].

As FER has been increasingly popular in recent years, over the past ten years, numerous new FER algorithms have been created. This essay provides a comprehensive examination of recent FER technology developments. The author covers the following aspects, in addition to defining depression and discussing FER's study. The authors then incorporated the existing in-use FER approaches into deep learning and machine learning-based methodologies. The authors specifically divide the six primary processes of the standard techniques into picture pre-processing, face detection, facial component detection, feature extraction, and expression categorization. Each phase introduces and discusses many potential strategies. It analyses facial photos and categorizes them into one of the six facial expressions: surprised, furious, disgusted, joyful, neutral, or happy [12].

In addition, we proposed Chi-square feature selection that has been demonstrated in previous research by [14] involving performance measurements that consist of TP Rate, FP Rate, Precision, Recall, F-Measure, and Accuracy. This measurement has successfully classified the expression accuracy of KNN, RF, J48, and RBF as depicted in the following Fig. 8.

Fig. 8. Results of Chi-Square feature selection performance towards four algorithms [12].

Apart from that, we also suggest incorporating the JAFEE, FER2013, and Ck+ databases to measure the algorithm performance. We believe that when university students' facial expressions can be detected and recognized using FER, then the depression tendency index can also be measured and displayed as the output. In this view, FER can benefit university counsellors in identifying and locating university students' depression tendencies quickly and accurately.

Altogether, the previous deliberation has suggested the following research model aim for Depression Tendency Detection using Facial Expression Recognition. Our proposed model is illustrated in Fig. 9.

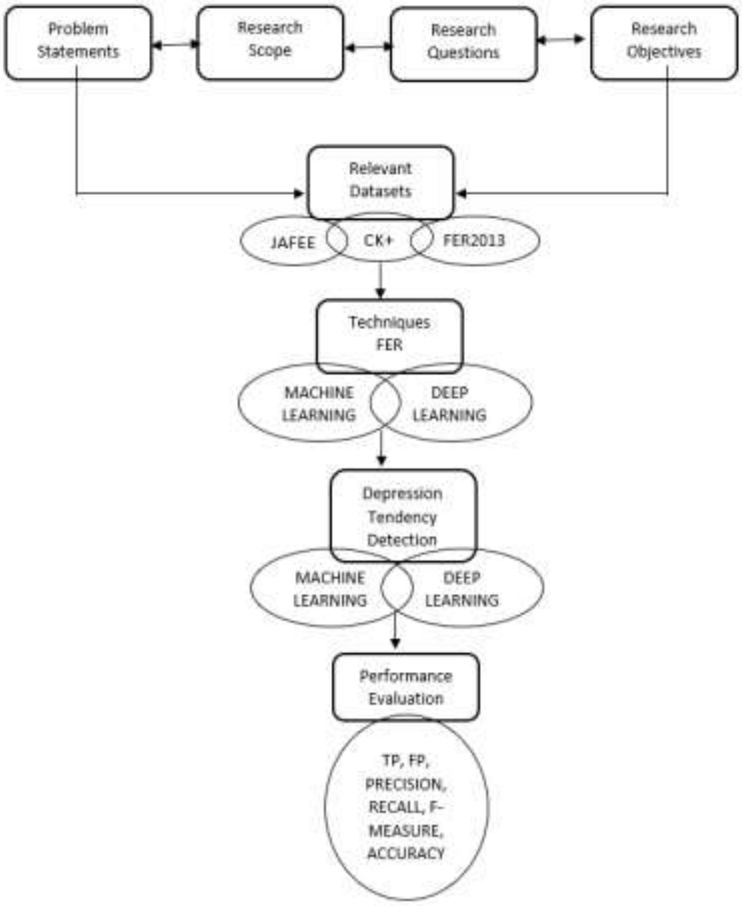

Fig. 9. Proposed research model.

4. Conclusion

This paper discusses a research model for detecting depression tendencies through facial expression recognition. To move forward with the study's execution, several strategies have been found in the literature, along with the necessary datasets. The research on this subject that comes after will mainly concentrate on two tactics. In the implementation, the authors focus on applying the research model with a variety of algorithms, specifically by exploiting the probability values for each class that the SVM algorithm gives, as increasing recognition rates is one of our primary objectives. The other is to automate the process of identifying veiled areas in obscured images.

References

1. Mahmood, M.R.; Abdulrazzaq, M.B.; Zeebaree, S.; Ibrahim, A.K.; Zebari, R.R.; and Dino, H.I. (2021). Classification techniques' performance evaluation for facial expression recognition. *Indonesian Journal of Electrical Engineering and Computer Science*, 21(2), 176-1184.
2. Higuera, V. (2020). Everything you need to know about depression (Major depressive disorder). Healthline. Retrieved October 5, 2023, from https://www.healthline.com/health/depression
3. Kim, J.-H.; Kim, B.-G.; Roy, P.P.; and Jeong, D.-M. (2019). Efficient facial expression recognition algorithm based on hierarchical deep neural network structure. *IEEE Access*, 7, 41273-41285.
4. Wang, W.; Xu, K.; Niu, H.; and Miao, X. (2020). Retracted: Emotion recognition of students based on facial expressions in online education based on the perspective of computer simulation. *Complexity*, Volume 2020, Article ID 4065207.
5. Khan, A.R. (2022). Facial emotion recognition using conventional machine learning and deep learning methods: current achievements, analysis and remaining challenges. *Information*, 13(6), 268.
6. Chen, M.; Cheng, J.; Zhang, Z.; Li, Y.; and Zhang, Y. (2021). Facial expression recognition method combined with attention mechanism. *Mobile Information Systems*, Volume 2021, Article ID 5608340.
7. VanDerBill, B. (2021). *What to know about depression in students*. Psych Central. Retrieved October 5, 2023, from https://psychcentral.com/depression/depression-in-students
8. Li, S.; and Deng, W. (2020). Deep facial expression recognition: A survey. *IEEE Transactions on Affective Computing*, 13(3), 1195-1215.
9. Rajan, S.; Chenniappan, P.; Devaraj, S.; and Madian, N. (2019). Facial expression recognition techniques: a comprehensive survey. *IET Image Processing*, 13(7), 1031-1040.
10. Lucey, P.; Cohn, J.F.; Kanade, T.; Saragih, J.; Ambadar, Z.; and Matthews, I. (2010). The extended cohn-kanade dataset (ck+): A complete dataset for action unit and emotion-specified expression. *Proceedings of the 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition – Workshops.* San Francisco, CA, USA, 94-101.
11. Fathallah, A.; Abdi, L.; and Douik, A. (2017). Facial expression recognition via deep learning. *Proceedings of the 2017 IEEE/ACS 14th International Conference on Computer Systems and Applications* (*AICCSA*), Hammamet, Tunisia, 745-750.
12. Tarnowski, P.; Kołodziej, M.; Majkowski, A.; and Rak, R.J. (2017). Emotion recognition using facial expressions. *Procedia Computer Science*, 108, 1175-1184.
13. World Health Organisation (2021). Depressive disorder (depression). World Health Organization; Retrieved October 5, 2023, from https://www.who.int/news-room/fact-sheets/detail/depression
14. Sikander, G.; Anwar, S.; Husnain, G.; Thinakaran, R.; and Lim, S. (2023). An adaptive snake based shadow segmentation for robust driver fatigue detection: A 3D Facial Feature based Photometric Stereo Perspective. *IEEE Access*, 11, 99178-99188.