Face Expressions Recognition by using Deep Learning

Salwa Almoshity
College of Electrical and Electronics Technology (CEET), Benghazi, Benghazi, Libya

Salema Younus
College of Electrical and Electronics Technology (CEET), Benghazi, Benghazi, Libya

Sarah Amer Al-asbaily
Faculty of information technology University of Benghazi, Benghazi, Libya

ABSTRACT
Facial expression recognition is a technology that uses biometric features to classify expressions in human faces. This technology plays a significant role in social communication since it conveys a lot of information about people, is considered a sentiment analysis tool, and is able to automatically recognize the seven basic or universal expressions: anger, contempt, disgust, fear, happiness, sadness, and surprise. Deep learning methods boost the learning process and facilitate the data creation task. In this work, the proposed approach used a non-classical technique, Inception-Resnet-v2, to pre-trained deep neural networks (DNNs) on more than a million images from the ImageNet and tested utilizing the face expression database from the Cohn-Kanade (CK+). The system had a loss validation of 0.014668% and attained 100% accuracy.

Keywords
Face Expressions, DNNs, InceptionResnet-V2.

1. INTRODUCTION
Technology continues to increase and improve on a daily basis in different fields of our lives. At the same time, the need for security with technology is also growing.

Modern and efficient methods of achieving technological security objectives are biometric approaches. For identifying persons based on their biological traits, these technologies are substitutes for personal identification numbers (PIN), passwords, and smart cards, which rely on remembering a specific code [1] [2] [3] [4] [5]. Biometric is used to recognize or verify human identity based on behavioral or physical characteristics. Biometric features such as fingerprint, palm print, face, iris, signature, hand geometry, and have been used for human identification and recognition [5] [6].

Human beings communicate with each other in the form of speech, gestures and emotions. As such systems that can recognize the same are in great demand in many fields [7]. Emotion is the most important factor that can be expressed in form of the psychological state of the human being, Where the user's ability to recognize their emotions is crucial for maximizing user happiness. Face expressions are a significant means of expressing emotions, whereas methods for classification facial expressions are crucial [8].

Facial expressions are the vital identifiers for human feelings, because it corresponds to the emotions. Most of the times (roughly in 55% cases), A person can tell if they are telling the truth or not by looking at their facial expression, which is a nonverbal form of emotional expression [9].

The classification of facial expressions is beneficial in terms of the capability of social interactions and communicating with machines. Classification of human facial expressions can be done by the Face Action Coding System (FACS) structure, which uses Action Units (AU) as its building blocks and specifies the action units, can be used to classify human facial expressions. The numerous types of action units are Main action units, Head movement action units, Eye movement action units, and Emotion action units [10].

The primary goal of facial expression recognition (FER) is to map different facial expressions to their corresponding emotional states. The two major components of the traditional FER are feature extraction and emotion recognition. Additionally, image preprocessing is required. This includes face detection, cropping, scaling, and normalizing. Face detection first crops the face region before removing the background and non-facial areas. In a traditional FER system, feature extraction from the processed image is the most crucial operation. Current methods employ prestigious techniques like discrete wavelet transform (DWT), linear discriminant analysis, etc. Finally, by classifying the retrieved data using a neural network (NN) and other machine learning techniques, it is possible to interpret emotions [11].

2. RELATED WORK
Due to the importance of recognition and classification facial expressions, many researchers have devoted time to designing algorithms for it.

Suneeta et al [12] presented a comparative analysis of facial expression recognition technologies. In general, facial expression recognition involves three steps: face detection, feature extraction, and classification of expressions. A face detection and extraction method based on the Haar cascade features is described in the suggested study. Through the use of numerous positive and negative images, the classifier is trained. It is used to extract the features. Convolutional kernel pictures are utilized for conventional Haar feature images. The CK+ database is then used to use Fisher face classifier, a supervised learning technique, to construct an eight-class system for categorizing facial expressions (neutral and the seven fundamental emotions). In the CK+ database, a recognition rate of 65% is achievable.

Singh et al [13] In this work, facial expressions are automatically recognized using DCNN features. The project will be finished using the dataset of Google search images.

One of the aforementioned datasets is used for facial recognition during the pre-processing stage. Using OpenCV21, frontal faces are reduced. Utilizing the Caffe on the graphics processing unit (GPU), feature extraction is carried out. Image Net 15 is used to gather facial data in order to construct the convolution neural network architecture used for object detection. In Image Net, there are eight learned layers, just the first five properties are collected using the first five layers.

Santra et al [14] In this approach been applied, the kaggle facial expression dataset with seven facial expression labels—happy, sad, surprise, fear, anger, disgust, and neutral—is used to classify emotions based on specific regions of interest on the face, including the lips, lower jaw, eyebrows, cheeks, and many

more. On the testing dataset, the system achieved a precision of 0.57 and an accuracy of 56.77%.

Facial [15] proposed a novel technique called facial emotion recognition (FERC) convolutional neural networks, is a unique method that has been developed based on a two-part convolutional neural network (CNN), the FERC: The backdrop is first removed from the image, and then the facial feature vector extraction is the focus of the second half. Expressional vector (EV) is utilized in the FERC model to identify the four different categories of typical face expression. 154 people's supervisory data were taken from the 10,000-image database that was saved. Using an EV of length 24 values, the emotion may be accurately highlighted with 96% accuracy. Before the development of EV, an innovative backdrop removal technique was employed to avoid dealing with a variety of potential issues (such as distance from the camera). Caltech faces, CMU, NIST, expanded Cohn-Kanade expression, and more than 750K photos were used to comprehensively test FERC.

Altaher et al [16] In this study, method for identifying facial emotions is provided using the CMU face data collection of four different emotions. The four distinct emotion categories are joyful, sad, furious, and neutral. Some of the various techniques employed in this study include the restricted Boltzmann machine (RBM), deep belief networks (DBN), convolutional neural networks (CNN), and multi-inception ensemble convolution neural networks. These approaches' accuracy has been contrasted. The latter is evidently considerably more accurate than the others. The results of using the procedures indicated Results for the Deep Belief Networks (DBN) model are almost i, while those for the Restricted Boltzmann Machine (RBM) model are 26.1 percent. identical and slightly higher than 26 percent, and the results for the standard CNN model are 55 percent.

Saravanan et al [17] The purpose of this study is to classify images of human faces into one of seven basic moods. A number of models, including decision trees and neural networks, were explored before choosing a final Convolutional Neural Network (CNN) model. CNNs are more efficient for picture identification jobs since they can input spatial data thanks to their many filters. The proposed model consists of six convolutional layers, two max pooling layers, and two fully linked layers. The final accuracy of this model after modifying the various hyperparameters was 0.60.

In this work, pre-trained DNNs from InceptionResnet-v2 are used, and trials and enhancements are carried out on the CK+ face expressions database, demonstrating the efficiency of the deep learning method in face expressions recognition.

The rest of the paper is organized as follows Section 3 is the proposed methodology, which includes introduces the face expressions preprocessing method and the face expressions recognition method based on deep learning Section 4 describes the experimental results obtained by verifying the algorithm on face expressions database; Section 5 conclusion and future work the paper.

3. PROPOSED APPROACH

The structure of the proposed methodology for the main stages that occur after face image acquisition: input face expressions database, preprocessing, classification using inception ResNet V2, result, the main methodology of this work has illustrated in Fig.1

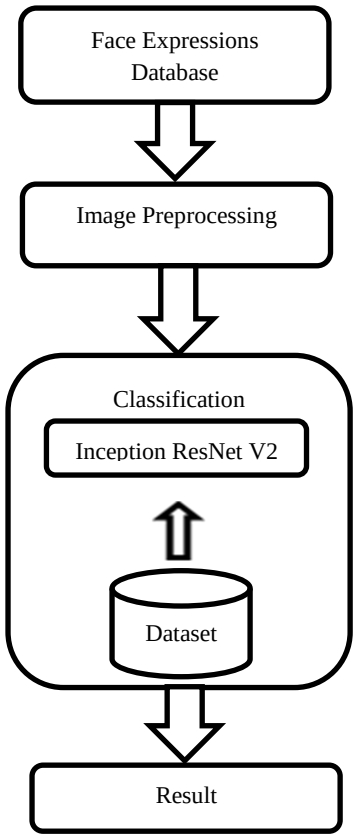

Fig. 1 Stages of proposed face expressions recognition system

3.1 Face expressions database

In the Extended Cohn-Kanade database (CK+), 97 people with ages ranging from 18 to 30 were represented by facial photographs. There were 65 percent female subjects in the database. Three percent of the participants were Asian or Latino, and 15% of the subjects were African-American. The subject was immediately in front of the camera. The individuals made a variety of facial expressions (single action units and combinations of action units), starting with a neutral expression and ending with the intended emotion. The displays (i.e., neutral, glad, surprised, angry, fearful, contemptuous, disgusted, and sad) were based on descriptions of typical emotions [18]. as shown in Fig. 2.

Fig. 2 CK+ face expressions image

3.2 Image pre-processing

The image has been resized (224×224) and rescale (1, 0) pre-processing procedures, Image augmentation (raising the amount of Image to better train utilizing rotation and zoom in zoom out) was utilized.

3.3 Inception ResNet V2

The InceptionResNetV2 architecture is a hybrid deep learning model that combines contemporary deep learning models with convolutional neural networks. It has the benefits of a residual network while retaining the distinctive characteristics of the multi convolutional core of the Inception network [19].

Inception ResNet V2 has used and improved in this study. Our earlier Inception models were trained with each replica being partitioned into numerous sub-networks in order to fit the complete model in memory. Contrarily, the Inception design is extremely adaptable and lets the number of filters in the various layers be changed without affecting how well the fully trained network performs. Manage the layer sizes carefully to distribute the computation among the numerous model sub-networks in order to speed up training. [20].

The development of Tensor Flow, on the other hand, has maintained stability and made it possible to train without separating the replicas. This has been possible thanks to recent memory advancements for back propagation that were acquired by carefully identifying the tensors needed for gradient computation and structuring the computation to restrict the number of such tensors. In the past, we have been

cautious when making architectural changes, limiting our testing to particular network components while the rest of the network remained stable [20], as shown in Fig.3

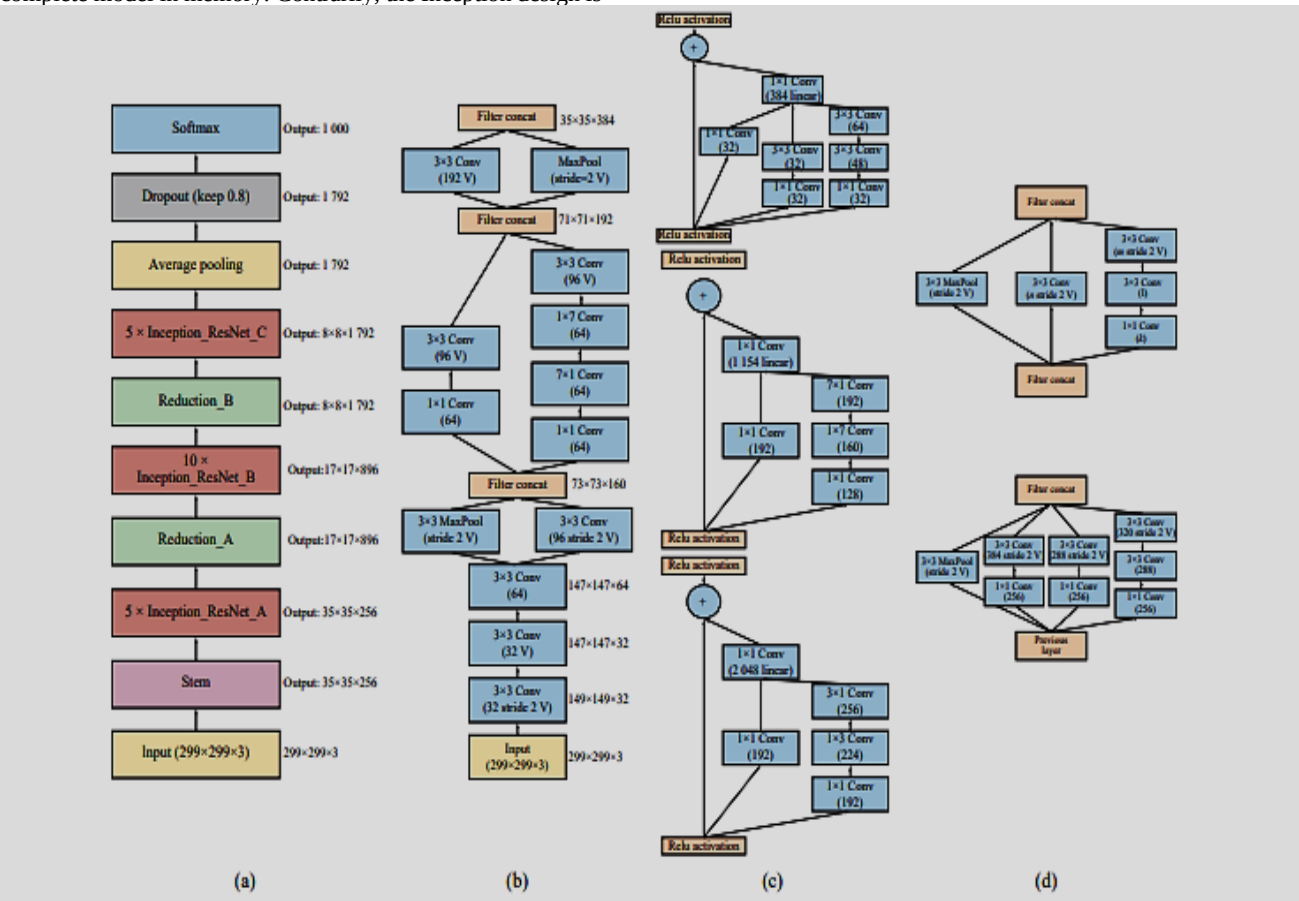

Fig. 3 Inception_ResNet_v2 overall structure and module structure: (a) Overall structure; (b) Stem module; (c) Up: Inception_ResNet_A, Middle: Inception_ResNet_B, Bottom: Inception_ResNet _C; (d) Up: Reduction_A, Bottom: Reduction_B [21]

4. RESULT AND DISCUSSION

The proposed method's efficacy and robustness were evaluated using CK+ face emotion database. For face expressions recognition, the best results are obtained by utilizing Inception ResNet-v2 pre-trained DNNs.

In the experiment, use seven expression classes: anger, contempt, disgust, fear, happiness, sadness, and surprise, we utilized 70% images for training and 15% for validation and 15% for testing. The correct recognition rates achieved after multiple rounds of training were 100% and loss validation were 0.014668%, as shows in Fig. 4

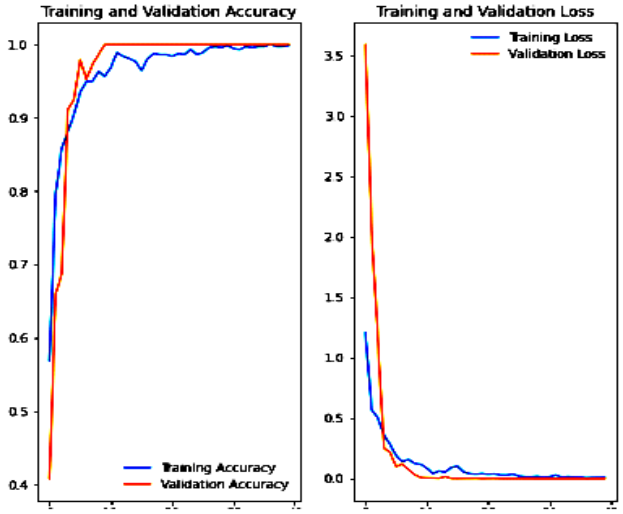

Fig. 4: (a) Training Accuracy and Validation Accuracy; (b) Training Loss and Validation Loss.

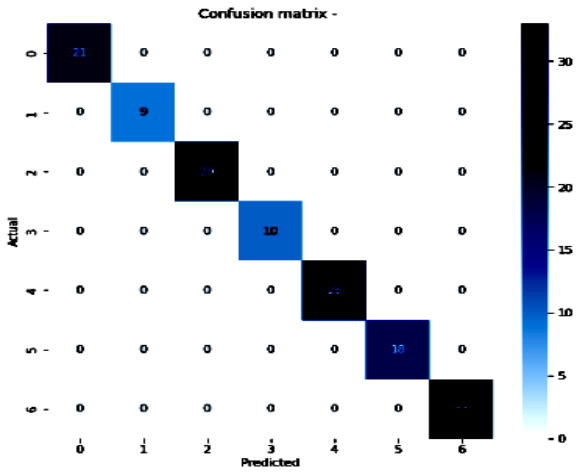

Fig. 5: confusion matrix

The precision, recall, f1 score, and accuracy are the main metrics utilized to assess the performance of the model in this research based on the performance of biometrics-based verification systems. The four categories utilized in these two metrics—True Positive (TP), False Positive (FP), False Negative (FN), and True Negative (TN)—are among them. It is TP if the test results are positive and are categorized as positive. It is FP if the test is both positive and labeled as negative. It is FN if the test is negative but nevertheless considered positive. It is TN if the test results are negative and classed as negative. The following formulas can be used to calculate the precision, recall, f1 score, and accuracy of our performance measures used in this paper:

$$precision = \frac{TP}{TP + FP}$$

$$recall = \frac{TP}{TP + FN}$$

$$F1\ score = \frac{TP}{TP + \frac{1}{2}(FP + FN)}$$

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

In Table I by using Python, we note that in all cases in the database we get accuracy = 1.00, macro avg = 1.00, weighted avg = 1.00. Since palm prints are abundant features, they may be affected by accuracy, which can also lead to recognition errors.

Table 1. The result of proposed system

No. of cases	precision	recall	f1 score	support
0	1.00	1.00	1.00	1
1	1.00	1.00	1.00	1
.				
.				
30	1.00	1.00	1.00	1
accuracy			1.00	30
macro avg	1.00	1.00	1.00	30
weighted avg	1.00	1.00	1.00	30

Comparisons with other methods are shown in Table II. All authors used (CK+) face expressions database

Table 2. Comparison with other methods

Year	Authors	Methods	Recognition Rate
2020	Shaees et al. [22]	Transfer Learning (CNNs, AlexNet + SVM)	98.3%
2021	Minaee et al. [23]	Attentional Convolutional Network	98.0%
2021	Debnath et al. [24]	ConvNet (f LBP, ORB + CNN)	98.13%
2023	Proposed Approach	Inception_ResNet _v2	100%

5. CONCLUSIONS

In this study, an efficient DNN using Inception Resnet-v2 pre-trained and the technique was assessed using the CK+ face expressions database. the proposed method shows very high recognition accuracy. Additionally, a result of 100% and a loss validation of 0.014668% are both achievable. In future work, it will be done multiple databases are used with the same methodology, the concept of facial emotion recognition could be expanded to include expressions recognition from speech or body language.

6. REFERENCES

[1] Mehraj, H., & Mir, A. H. (2021). A Survey of Biometric Recognition Using Deep Learning. *EAI Endorsed Transactions on Energy Web*, *8*(33), e6-e6.

[2] Mustapha, M. F., Mohamad, N. M., & AB, S. H. (2022). A Survey On Video Face Recognition Using Deep Learning. *Journal of Quality Measurement and Analysis JQMA*, *18*(1), 49-62.

[3] Minaee, S., Abdolrashidi, A., Su, H., Bennamoun, M., & Zhang, D. (2023). Biometrics recognition using deep learning: A survey. *Artificial Intelligence Review*, 1-49.

[4] Bogdanov, M. R., Dumchikova, I. N., Dokuchaev, I. S., & Dumchikov, A. (2018). Deep learning based person biometric identification. *Industry 4.0*, *3*(4), 219-222.

[5] Younus, S., Abdulali, E. O., Bozed, K. A., & Boudjellal, S. E. (2022, May). PalmPrint Recognition using Deep Convolutional Neural Networks. In *2022 IEEE 2nd International Maghreb Meeting of the Conference on Sciences and Techniques of Automatic Control and Computer Engineering (MI-STA)* (pp. 539-543). IEEE..

[6] Shepley, A. J. (2019). Deep learning for face recognition: a critical analysis. *arXiv preprint arXiv:1907.12739*.

[7] Sarvakar, K., Senkamalavalli, R., Raghavendra, S., Kumar, J. S., Manjunath, R., & Jaiswal, S. (2023). Facial emotion recognition using convolutional neural networks. *Materials Today: Proceedings*, *80*, 3560-3564.

[8] Choi I., Ahn H., & Yoo J.,(2018) "*Facial Expression Classification Using Deep Convolutional Neural Network*", Journal of Electrical Engineering and Technology,vol.13,no.1,pp.1,pp. 485-492.,

[9] Mehendale N.,(2020)"*Facial,emotion recognition using convolutional neural networks*", SN Applied Sciences,vol.2,no.3,pp.1-8.

[10] Wibowo, H., Firdausi, F., Suharso, W., Kusuma, W. A., & Harmanto, D. (2019). Facial expression recognition of 3D image using facial action coding system (FACS). *Telkomnika*, *17*(2), 628-636.

[11] Akhand, M. A. H., Roy, S., Siddique, N., Kamal, M. A. S., & Shimamura, T. (2021). Facial emotion recognition using transfer learning in the deep CNN. *Electronics*, *10*(9), 1036.

[12] Suneeta, V. B., Purushottam, P., Prashantkumar, K., Sachin, S., & Supreet, M. (2019, September). Facial expression recognition using supervised learning. In *International Conference On Computational Vision and Bio Inspired Computing* (pp. 275-285). Cham: Springer International Publishing.

[13] Singh, A., Srivastav, A. P., Choudhary, P., & Raj, S. (2021, April). Facial emotion recognition using convolutional neural network. In *2021 2nd International Conference on Intelligent Engineering and Management (ICIEM)* (pp. 486-490). IEEE.

[14] Santra,A., Rai, V. , Das , D. Kundu, S.(2022) "Facial Expression Recognition Using Convolutional Neural Network "vol.10,pp. 2321-9653,2022.

[15] Mehendale, N. (2020). Facial emotion recognition using convolutional neural networks (FERC). *SN Applied Sciences*, *2*(3), 446.

[16] Altaher, A., Salekshahrezaee, Z., Abdollah Zadeh, A., Rafieipour, H., & Altaher, A. (2020). Using multi-inception CNN for face emotion recognition. *Journal of Bioengineering Research*, *3*(1), 1-12.

[17] Sarvakar, K., Senkamalavalli, R., Raghavendra, S., Kumar, J. S., Manjunath, R., & Jaiswal, S. (2023). Facial emotion recognition using convolutional neural networks. *Materials Today: Proceedings*, *80*, 3560-3564.

[18] Tripathi , A., Thakurdesai, N.,(2018)"Implementation and Comparison of Facial Expression Detection and Classification Techniques" International Journal of Computer Applications,vol.182 ,no. 18,pp. 0975 – 8887.

[19] Wable , S., Shitole , S., Thorat , A., Sarde , R.& Zope, M. ,(2022) "Classification and Identification of Multiple Leaf Diseases Using Inception-ResNet V2 (CNN Architecture)", ijariie,vol.8,no.2,pp. 2395-4396.

[20] Szegedy, C., Ioffe, S., Vanhoucke, V., & Alemi, A. (2017, February). Inception-v4, inception-resnet and the impact of residual connections on learning. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 31, No. 1).

[21] Jia, W., Gao, J., Xia, W., Zhao, Y., Min, H., & Lu, J. T. (2021). A performance evaluation of classic convolutional neural networks for 2D and 3D palmprint and palm vein recognition. *International Journal of Automation and Computing*, *18*(1), 18-44.

[22] Shaees, S., Naeem, H., Arslan, M., Naeem, M. R., Ali, S. H., & Aldabbas, H. (2020, September). Facial emotion recognition using transfer learning. In *2020 International Conference on Computing and Information Technology (ICCIT-1441)* (pp. 1-5). IEEE.

[23] Minaee, S., Minaei, M., & Abdolrashidi, A. (2021). Deep-emotion: Facial expression recognition using attentional convolutional network. *Sensors*, *21*(9), 3046.

[24] Debnath, T., Reza, M. M., Rahman, A., Beheshti, A., Band, S. S., & Alinejad-Rokny, H. (2022). Four-layer ConvNet to facial emotion recognition with minimal epochs and the significance of data diversity. *Scientific Reports*, *12*(1), 6991.