Facial Expression Recognition Enhancement Using Convolutional Neural Network

Zahraa Fawzi Hassan*, M. N. Al-Turfi, Faiz Al Alawy*****

* Computer Engineering Department, College of Engineering, Al-Iraqia University, Iraq
zahraa.f.hassan@aliraqia.edu.iq
https://orcid.org/0009-0003-1365-5639

** Electrical Engineering Department, College of Engineering, Al-Iraqia University, Iraq
mohammed_alturfi@yahoo.com
https://orcid.org/0000-0001-5723-2660

*** Department of Computer Technology Engineering, AL Qalam University College, Kirkuk, Iraq
faiz.alawy@alqalam.edu.iq
https://orcid.org/0000-0002-6088-7315

Abstract

Facial expression recognition (FER) technology is quite popular in the fields of computer vision, security monitoring, image classification, and many other related applications. Enhancing the computer's ability to read facial expressions is important for human-computer interaction, as it enables machines to understand and interact with human emotions. In this paper, a modified approach using neural networks (CNNs) is presented to accurately identify unique facial expressions. using the changes to a commonly used 12-layer CNN model to improve its performance in facial expression recognition (FER). The model is trained using dataset images, which allows it to better infer people's emotions from their facial features. To enhance the accuracy of the system, a preprocessing stage is incorporated. This stage involves several operations for data augmentation, including changing the color in images, such as HSV (Hue, Saturation, Value), YCbCr (Luma, Blue-difference Chroma, Red-difference Chroma), etc., to facilitate better interpretation of color information. Additionally, the preprocessing stage refines facial expression recognition by manipulating facial features and adding extra RGB details, improving the visual information provided by the images. The study specifically focuses on evaluating the effectiveness of this approach using the KDEF database. The KDEF database contains standardized images of facial expressions, making it suitable for assessing the performance of the proposed system. By combining the modified CNN model, training with images, and the preprocessing operations, the system's performance was significantly enhanced. As a result, it achieved recognition rates of up to 95%, indicating a notable improvement in accurately identifying unique facial expressions compared to previous approaches.

Keywords- Facial Expression Recognition, CNN, KDEF, Data Augmentation.

I. INTRODUCTION

Facial emotions and expressions are mirrors of human thoughts, feelings and characteristics representation because it provides the viewer with an abundance of social signals, such as the focus of attention, intention, motivation, and emotion [1]. It is considered an effective method of nonverbal, efficient, effective communication tool. An analysis of these expressions yields a significantly deeper understanding [2]. Compared to verbal communication, facial expressions are faster and more effective in conveying messages. They play a crucial role in communication as they provide immediate visual cues about a person's emotional state [3]. Facial emotions are important for non-verbal communication as they enhance oral communication and aid in understanding concepts more effectively. Analyzing facial expressions allows human kind to interpret emotions, leading to more efficient and meaningful interactions [4] [5]. Predicting human emotion and using computers to diagnose psychiatric illnesses are the main goals of the significant branch of study known as emotion detection [6].

Understanding of human behavior in recent years, AI-based Facial Expression Recognition (FER) [7] has become one of the most important research topics, with applications in dynamic analysis, pattern recognition, interpersonal interaction, and monitoring of mental health, among others [1]. Knowing how a person feels from their face is hard for machines because they have to tell the difference between many feelings and understand what's happening around them. This opens up lots of ways that can be used, especially in alert systems about suspicious behavior or health matters which got very important not long ago [4] [8].

Intelligent FER techniques like CNNs based facial emotion recognition (FER) [9] systems utilized to classify images using deep learning (DL). CNN-based methods concentrate on static features and may be incapable of handling the dynamic character of facial expressions [6], [10]. CNNs are employed to extract features for FER systems, enabling intelligent systems to recognize and respond to human emotions. DL methods reduce the reliance on face-to-face physics-based techniques [11].
The advancement of deep learning (DL) algorithms throughout the 2010s significantly increased the accuracy and dependability of FER systems. Intelligent FER techniques like CNNs based facial emotion recognition (FER) systems utilized to classify images using deep learning (DL). The Convolutional Neural Networks (CNN) showed state-of-the-art performance in image identification, was one of the most notable efforts that popularized the use of CNNs for image recognition. researchers have developed new FER systems that are capable of detecting complex emotions and subtle nuances in facial expressions [12]. CNN-based methods concentrate on static features and may be incapable of handling the dynamic character of facial expressions [6], [10]. CNNs are employed to extract features for FER systems, enabling intelligent systems to recognize and respond to human emotions [11].
CNNs have demonstrated significant promise in this discipline, particularly when combined with data augmentation techniques. Researchers have investigated a variety of data augmentation techniques to enhance CNN performance on this dataset. The combination of neural network classifiers (CNNs) and data augmentation strategy has significantly improved the accuracy of FER [13]. Data augmentation is a crucial method for enhancing the efficacy of DL models by augmenting the dataset with additional data. It enables cutting-edge efficacy in numerous DL applications [14].
This paper is structured as follows: an overview of previous works on FER is provides in section 2. The research methodology is explained in section 3. Performance metric for classification algorithms in section 4. The qualitative and quantitative experimental results are discussed in section 5. The conclusion is discussed in section 6.

II. RELATED WORK

In this section, various related works dealing with human feelings, and particularly about human emotional recognition from the facial expression.
Eng et al. (2019) [15], utilized a support vector machine (SVM) as the classifier for facial emotion identification. They employed a histogram of oriented gradients (HOG) feature extraction technique to gather local information and orientation density distribution from the facial expression images. The preprocessing stage involved face identification and cropping, followed by extracting HOG features. These features were then organized into histogram bins to create a feature vector, which was input into the SVM classifier. Their approach achieved average recognition rates of 76.19% and 80.95% on the JAFFE and KDEF datasets, respectively, with varying performance across different expressions.
Sajjad et al. (2020) [14], focused on analyzing facial expressions in popular TV series footage to explore human behavior. Their approach involved several steps. Firstly, they employed the Viola-Jones algorithm to detect and recognize faces in the video frames. Then, the Kanade-Lucas-Tomasi method was used to track the detected faces throughout the video. HOG (Histogram of Oriented Gradients) features were extracted from the tracked faces and utilized with a support vector machine (SVM) classifier for facial recognition. Additionally, a lightweight convolutional neural network (CNN) was employed to identify facial emotions. Data augmentation techniques were applied to enhance face emotion recognition. The experimental results demonstrated improved performance in both face emotion recognition and human behavior comprehension.
Hussain & Al Balushi (2020) [16], presented three processes for facial recognition and detection using deep learning (DL) algorithms: face detection, face recognition, and face classification. In the face detection stage, a video camera captures a human face, and Haar cascade detection is employed to determine the face's position. Next, a CNN model, specifically VGG 16, is utilized to match the detected face with a database and categorize it based on expressions such as anger, fear, disgust, happiness, neutrality, and surprise. The system's effectiveness is demonstrated through experimental findings, and the performance of automatic face identification and recognition is evaluated using accuracy metrics.
Zhang et al. (2020) [17], proposed a multi-stream CNN fusion network along with an improved Local Binary Patterns (LBP) feature operator to extract facial features specifically from infant faces. Their approach involved combining multiple CNN streams to capture different aspects of facial information. The LBP feature operator was enhanced to better extract discriminative features from infant faces. By utilizing this multi-stream CNN fusion network and improved LBP feature operator, they achieved an improved accuracy rate of 91.67% in facial feature extraction for infants, surpassing the performance of traditional methods.
Liu et al. (2021) [18], propose a DL model called DML-Net for FER that specifically aims to handle position and identity fluctuations in facial expressions. DML-Net utilizes multiple simultaneous convolutional networks to extract global and local information from different face regions. These networks combine the extracted information in an embedding space to construct expression representations. The model is trained end-to-end using multiple loss functions, which helps improve recognition performance compared to previous techniques. The focus of their work is on building a robust FER model that can handle variations in facial position and identity.
Wafi et al. (2023) [19], compare different feature extraction methods, namely the gray level co-occurrence matrix, local binary pattern, and facial landmark (FL) techniques, for detecting human facial expressions. They evaluate the effectiveness of these methods using two datasets of facial expressions. Additionally, they propose an enhanced version of the extreme learning machine (ELM)

method called adaptive ELM (aELM). The aELM method dynamically selects the optimal number of hidden neurons, aiming to achieve improved performance. The results of their study demonstrate that utilizing the FL feature extraction method, their proposed approach achieves a maximum mean accuracy score of 88.07% on the CK+ dataset and 83.12% on the JFFE dataset. These results show some improvement over the standard ELM method.

Table 1 provides a summary of related works in the field of facial expression recognition. provide information about the method used, the dataset used for evaluation, the difference compared to the proposed work, and the results obtained.

Table 1: Summary of Related Works

Ref.	Method	Dataset	Difference to proposed work	Result
Eng et al. (2019) [15],	SVM as the classifier with HOG feature extraction technique	JAFFE and KDEF datasets	The main difference is the classification approach, with Eng et al. using SVM with HOG features, while the proposed work utilizes modified CNNs with preprocessing and data augmentation operations.	recognition rates are 76.19% and 80.95%
Sajjad et al. (2020) [14],	combination of Viola-Jones algorithm, Kanade-Lucas-Tomasi method, HOG features, SVM classifier, and a lightweight CNN	facial expressions in TV series	The study focused on analyzing facial expressions in TV series footage using a combination of Viola-Jones algorithm, Kanade-Lucas-Tomasi method, HOG features, SVM classifier, and a lightweight CNN, the proposed work focuses on modified CNNs with preprocessing and data augmentation operations for facial expression recognition.	Experiment results show enhanced performance in both face emotion recognition and human behavior
Hussain & Al Balushi (2020) [16],	Haar cascade detection, VGG 16 CNN model	-	the study focuses on reliable facial recognition and detection using processes like Haar cascade detection, VGG 16 CNN model, and expression classification, the proposed work focuses on modified CNNs with preprocessing operations and data augmentation operations for facial expression recognition.	-
Zhang et al. (2020) [17],	multi-stream CNN fusion network along with an improved LBP	infant faces	the study focuses on facial feature extraction for infant faces using a multi-stream CNN fusion network and an improved LBP feature operator, while the proposed work focuses on facial expression recognition using modified CNNs and preprocessing techniques and data augmentation.	accuracy rate of 91.67%
Liu et al. (2021) [18],	DML-Net DL model	-	the study proposed the DML-Net DL model for FER to handle position and identity fluctuations using multiple simultaneous convolutional networks and embedding spaces, the proposed work focuses on modifying CNNs for facial expression recognition and achieves higher recognition rates compared to previous techniques.	resulting in better recognition performance
Wafi et al. (2023) [19]	adaptive ELM (aELM)	CK+ and JFFE dataset	focus on comparing different feature extraction methods for facial expression detection and propose an enhanced version of the ELM method, the proposed work focuses on modifying CNNs for facial expression recognition and achieves higher recognition rates	mean accuracy score of 88.07% on the CK+ dataset and 83.12% on the JFFE dataset.

III. METHODOLOGY

This section explains the methodology that is used to recognize the different facial expressions utilizing intelligent systems with DL technique. The proposed method builds a FER system that can recognize and read each tested person's facial expression by using the KDEF dataset as a target example throughout the training phase. The model follows a multi-phase approach, including collecting/processing facial images, building a model based on the training dataset, as illustrates in Figure 1.

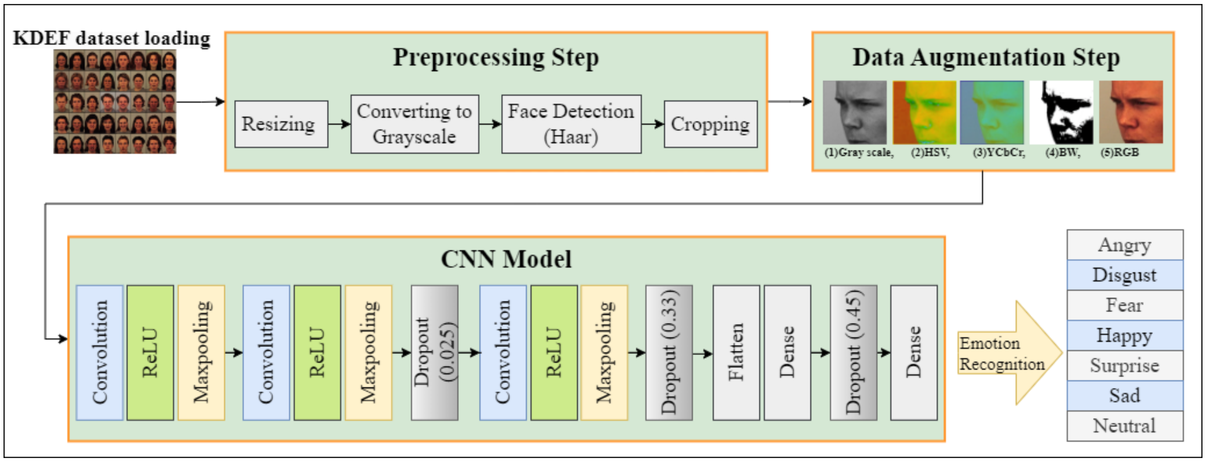


Figure 1: The Diagram of The Proposal Method

A. KDEF dataset

The "Karolinska Directed Emotional Faces" (KDEF) dataset is a collection of 4,900 photographs capturing human facial expressions. This dataset includes images of 70 individuals, with an equal distribution of 35 females and 35 males. It encompasses 6 primary emotions (such as "happiness, sadness, anger, surprise, fear, and disgust.") along with a neutral expression, making it an invaluable resource for studying and analyzing emotion identification [20].

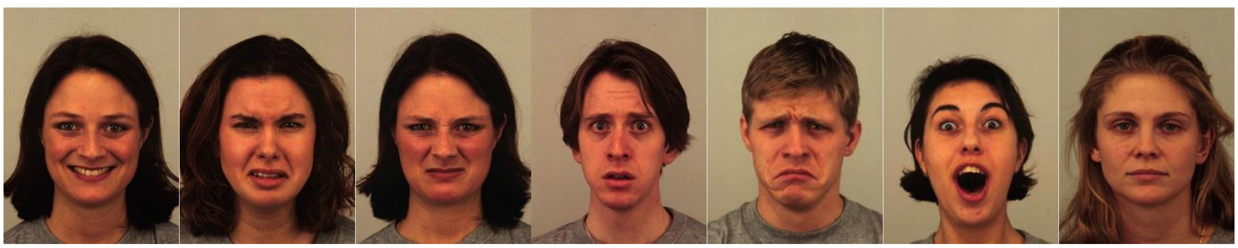


Figure 2: Primary KDEF Dataset image samples

B. Preprocessing

Preprocessing is a technique used to pressed images and normalized to make it ready for any image processing application including FER systems. The set of the preprocessing operations plays a crucial role in improving the quality of input images to the FER system by removing any unwanted noise or artifacts that could impact the accuracy of the system [21].

In this work, preprocessing involves the operations of resizing the images and enhancing facial features to ensure better image quality and more accurate recognition. Several other operations are carried out such as cropping the image to focus on the face, converting RGB images to grayscale, and resizing all the images to (64 × 64) pixels to ensure uniformity in size and shape of the extracted face region.

Face detection in images is accomplished using the Haar Cascade Classifier algorithm. It compares identified characteristics such as eyes, nose, mouth, and general face shape to a collection of trained classifiers by examining different scales and locations of a sliding window in the image. These preprocessing steps help in obtaining standardized facial images for accurate facial expression analysis.

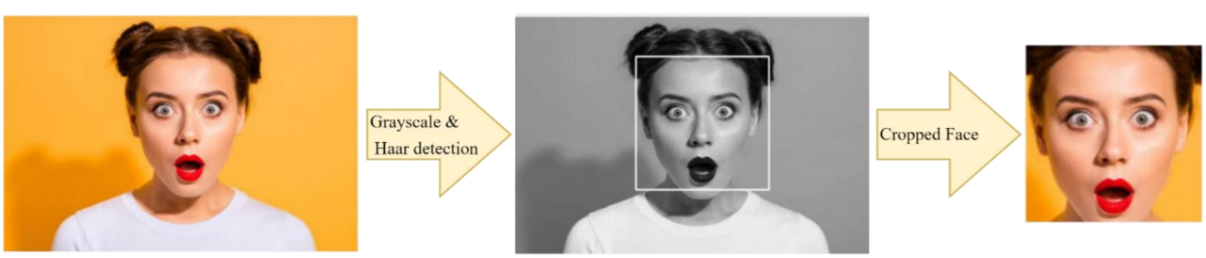


Figure 3: Preprocessing Steps

C. Data Augmentation Techniques

In machine learning (ML) and DL, data augmentation techniques are commonly employed to increase the size and diversity of training datasets. With the help of these techniques, current data can be transformed or modified to create new samples that are somewhat different but still similar to the original. Data augmentation improves the generalization and robustness of ML models, especially when there is limited training data available [22]. A several data augmentations were used in this work to enhance the quantity and diversity of the training datasets, as the following:

1- Grayscale image conversion: is a process where an RGB image is transformed into a simpler representation by combining the red, green, and blue color channels into a single grayscale channel. This results in a reduced image with only one channel instead of the original three, simplifying the image's representation for further analysis or processing.

2- Hue, Saturation, and Value (HSV): is a color space that represent colors in a more intuitive manner. It is visualized using a cone, where the phase angle of the color ranges from 0 to 360 degrees, allowing for a comprehensive representation of color variations.

3- YCbCr: it is one of the many color spaces used for image processing and video compression. Augmentation of image data that includes color transformation can be performed using the YCbCr model which is characterized by modification of Cb and Cr while leaving Y constant, thereby allowing changes in color information without changing brightness.

4- Binary (BW) – refers to Black and White images – data augmentation consists of the generation of binary images after subjecting a set of previous ones to various transformations and modifications. The use of such technique improves diversity within a given training set, that is important for computer vision problems.

5- RGB: one of the tools for computer vision, which aims at increasing the variety and quantity of images in a training dataset. It is a process that entails making various alterations to the RGB values of an image. Such a process helps avoid overfitting in deep neural networks and improves how AI models work.

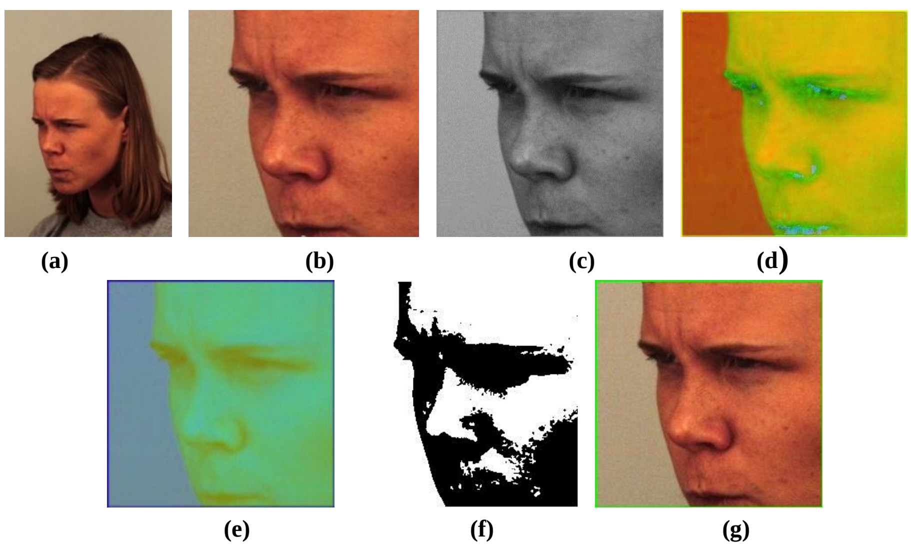


Figure 4: Data Augmentation (a)Original image, (b)Cropping, (c)Grayscale, (d)HSV, (e)YCbCr, (f)BW, (g)RGB

D. Convolutional Neural Network (CNN)

The proposed CNN model consists of three convolutional layers, three max-pooling layers, three dropout layers, and a flatten layer. The structure is offered as a full explanation of the layers that implement a CNN model, as follows:

- **Layer_1:** convolutional layer (Conv2D) with 128 filters of size 5x5. The activation function used is ReLU.
- **Layer_2:** max pooling layer (MaxPooling2D) with a pool size of 5x5. Max pooling takes the maximum value within each pooling window to minimize the spatial dimensions of the input.
- **Layer_3:** the second convolutional layer (Conv2D_1) with 64 filters of size 2x2 and ReLU activation.
- **Layer_4:** the second max pooling layer (MaxPooling2D_1) with a pool size of 2x2.
- **Layer_5:** dropout layer (Dropout) with a dropout rate of 0.025. During training, dropout randomly sets a portion of the input units to 0.

- **Layer_6:** the third convolutional layer (Conv2D_2) with 128 filters of size 2x2 and ReLU **activation.**
- **Layer_7:** the third max pooling layer (MaxPooling2D_2) with a pool size of 2x2.
- **Layer_8:** the second dropout layer (Dropout_1) with a dropout rate of 0.33.
- **Layer_9:** flatten layer (Flatten) to convert the multidimensional output from the previous layer into a one-dimensional vector.
- **Layer_10:** fully connected layer (Dense) with 64 neurons and ReLU activation.
- **Layer_11:** dropout layer (Dropout_2) with a dropout rate of 0.45.
- **Layer_12:** fully connected layer (Dense_1) with 7 neurons and ReLU activation.

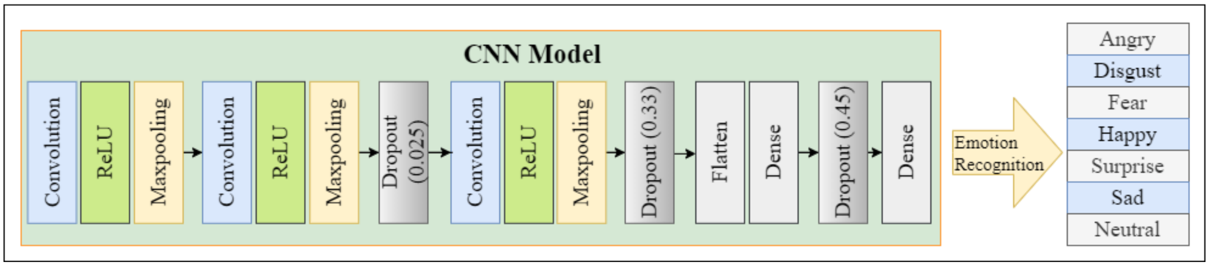


Figure 5: CNN Proposed Model

Table 2 shows the total parameters in the proposed model as well as some of the parameters (weights and biases) in each layer.

Table 2:CNN Parameters

Layer (Type)	Output Shape	Parameters
conv2d (Conv2D)	(None, 60, 60, 128)	9728
max_pooling2d (MaxPooling2D)	(None, 12, 12, 128)	0
conv2d_1 (Conv2D)	(None, 11, 11, 64)	32832
max_pooling2d_1 (MaxPooling2D)	(None, 5, 5, 64)	0
dropout (Dropout)	(None, 5, 5, 64)	0
conv2d_2 (Conv2D)	(None, 4, 4, 128)	32896
max_pooling2d_2 (MaxPooling2D)	(None, 2, 2, 128)	0
dropout_1 (Dropout)	(None, 2, 2, 128)	0
flatten (Flatten)	(None, 512)	0
dense (Dense)	(None, 64)	32832
dropout_2 (Dropout)	(None, 64)	0
dense_1 (Dense)	(None, 7)	455
Total params: 108,743 Trainable params: 108,743 Non-trainable params: 0		

IV. EXPERIMENTAL RESULTS

The proposed system for FER is developed using the publicly available dataset, KDEF. The images undergo data augmentation and preprocessing, the images are then used as input to the system to train an CNN model. Adam optimization algorithm is used in the CNN training operation. The available dataset is separated into 80:20 ratio to train the proposed CNN.

During the training phase, a batch size of 128 images per iteration and the training process was set to continue for a maximum of 100 epochs. 0.001 was the initial learning rate. and the size of training images is 64×64 pixels. Additionally, the training set is augmented, as mentioned earlier, to enhance the diversity and quantity of training data.

Table 3: Model training parameters

Parameters	The value
Epochs	100
Batch size	128
Learning rate	0.001

The performance of the proposed method evaluated utilizing confusion matrix, Figure 6. The CNN model performed well in terms of accurately predicting the test data, the test accuracy achieved of 0.9433% and the validation accuracy achieved of 0.9240%. As can be seen in Figure 7, test accuracy and loss are in line with training accuracy and loss. The model is not overfitting even if the validation loss and accuracy lines are not linear; the test loss is declining rather than rising, and the difference among training and testing accuracy is not very great.

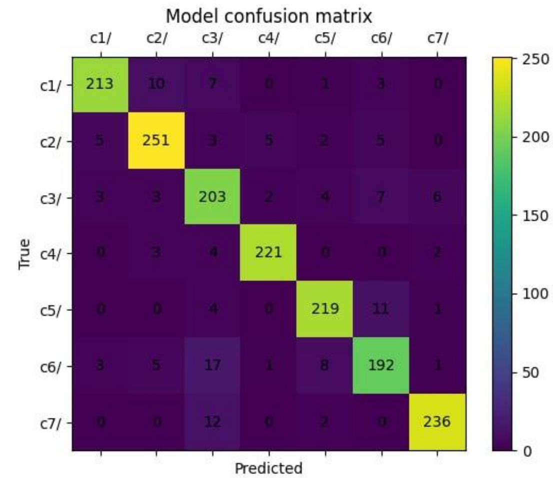
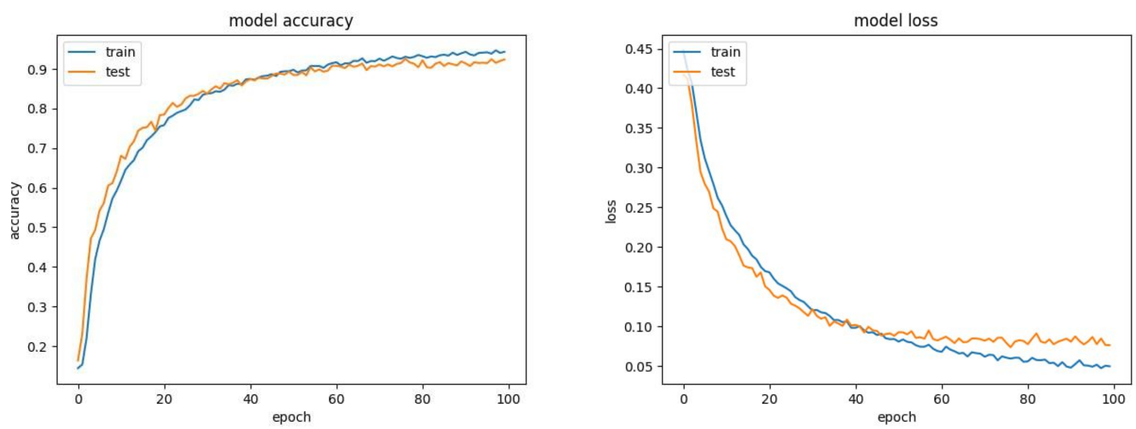


Figure 6: Model confusion Matrix.



(a) **(b)**

Figure 7: (a) Model Accuracy (b) Model Loss

In addition, Figure 8 displays the predictions accuracy that is obtained using the developed CNN.

Test image **Cropped face** **Output Emotion**

Figure 8: predictions result

Comparing the performance of this newly suggested system with other published techniques is giving a good indication that it is outperforming other techniques. Table 3 illustrates the results of the comparison. The results of two relevant researches using two different methodologies on the same KDEF dataset have been compared to the proposed model. The comparison's findings, show that our method is successfully outperformed the others.

Table 4: Comparison Between Our Proposed Model and The Current Models

Authors	Model	Accuracy
Eng, S. K.et al (2019) [15]	SVM with HOG	80.95%
Liu et al. (2021) [18],	DML-Net	88.2%
Our	**New CNN**	**94.33%**

V. Conclusion

The study highlights the difficulty of facial expression recognition (FER) as a computer vision task with practical applications in human-computer interaction and sentiment analysis. The paper proposes a novel approach that utilizes convolutional neural networks (CNNs) and data augmentation techniques to improve the accuracy of FER. The focus of the study is on the KDEF dataset, and the results demonstrate favorable performance compared to another research in the field. The proposed method achieves a high accuracy rate of 95% on the KDEF dataset. The study also suggests that incorporating traditional manual features into CNNs can further enhance the performance of FER models. By combining these manual features with the capabilities of deep neural networks, the model's generalization capabilities can be improved. Additionally, training the deep neural network using examples from multiple databases can contribute to better performance. Overall, the conclusion highlights the potential of the proposed approach in achieving high accuracy in FER and suggests avenues for further improvement by incorporating manual features and utilizing diverse training data.

In future works, we suggest Transfer learning techniques can be explored to leverage pre-trained models on large-scale image recognition tasks. By fine-tuning these models on FER datasets, it may be possible to achieve even higher accuracy rates and improve the generalization capabilities of the FER model.

Acknowledgment

I dedicate this research to my family, especially my martyr brother, and to the supervisors.

References

[1] Dalvi, C., Rathod, M., Patil, S., Gite, S., & Kotecha, K. , "A survey of ai-based facial emotion recognition: l techniques, age-wise datasets and future directions," *Ieee Access*, vol. 9, pp. 165806-165840, 2021.

[2] n, A., & Ferris, S. R., "Inference of attitudes from nonverbal communication in two channels," *Journal of consulting y*, vol. 31, no. 3, p. 248, 1967.

[3] Mohammadpour, M., Khaliliardali, H., Hashemi, S. M. R., & AlyanNezhadi, M. M., " Facial emotion recognition using deep convolutional networks.," *In 2017 IEEE 4th international conference on knowledge-based engineering and innovation (KBEI),IEEE.*, pp. 0017-0021, (2017, December).

[4] Khan, A. R., " Facial emotion recognition using conventional machine learning and deep learning methods: current achievements, analysis and remaining challenges," *Information*, vol. 13, no. 6, p. 268, 2022.

[5] Moret-Tatay, C., Wester, A. G., & Gamermann, D., " To Google or not: Differences on how online searches predict names and faces," *Mathematics*, vol. 8, no. 11, p. 1964, 2020.

[6] Ozdemir, M. A., Elagoz, B., Alaybeyoglu, A., Sadighzadeh, R., & Akan, A., "Real time emotion recognition from facial expressions using CNN architecture.," *In 2019 medical technologies congress (tiptekno),IEEE*, pp. 1-4, 2019.

[7] H. Mohammed, M. Nasser Hussain, and F. Al Alawy, "Feature Extraction Techniques for Facial Expression Recognition (FER)," *IJSER*, vol. 2, no. 3, p. 32–40, Sep. 2023..

[8] Elarbi-Boudihir, M., Rehman, A., & Saba, T. , "Video motion perception using optimized Gabor filter.," *International journal of physical sciences*, vol. 6, no. 12, pp. 2799-2806, 2011.

[9] H. Mohammed, M. Nasser Hussain, and F. Al Alawy,, "Facial Expression Recognition: Machine Learning Algorithms and Feature Extraction Techniques," *IJSER,* vol. 2, no. 2, p. 23–28, Jun. 2023..

[10] Liang, D., Liang, H., Yu, Z., & Zhang, Y. , "Deep convolutional BiLSTM fusion network for facial expression recognition," *The Visual Computer*, vol. 36, pp. 499-508., 2020.

[11] Ko, B. C. , " A brief review of facial emotion recognition based on visual information," *sensors,* vol. 18, no. 2, p. 401, 2018.

[12] Recht, B., Roelofs, R., Schmidt, L., & Shankar, V., "Do imagenet classifiers generalize to imagenet?," *In International conference on machine learning ,PMLR.,* pp. 5389-5400, (2019, May). .

[13] Ahmed, T. U., Hossain, S., Hossain, M. S., ul Islam, R., & Andersson, K. , "Facial expression recognition using convolutional neural network with data augmentation," *In 2019 Joint 8th International Conference on Informatics, Electronics & Vision (ICIEV) and 2019 3rd International Conference on Imaging, Vision & Pattern Recognition (icIVPR), IEEE.,* pp. 336-341, 2019.

[14] Sajjad, M., Zahir, S., Ullah, A., Akhtar, Z., & Muhammad, K., "Human behavior understanding in big multimedia data using CNN based facial expression recognition," *Mobile networks and applications,* vol. 25, pp. 1611-1621., 2020.

[15] Eng, S. K., Ali, H., Cheah, A. Y., & Chong, Y. F., " Facial expression recognition in JAFFE and KDEF Datasets using histogram of oriented gradients and support vector machine.," *In IOP Conference series: materials science and engineering, IOP Publishing.,* vol. 705, no. 1, p. 012031, 2019.

[16] Hussain, S. A., & Al Balushi, A. S. A. , "A real time face emotion classification and recognition using deep learning model.," *In Journal of physics: Conference series , IOP Publishing.,* vol. 1432, no. 1, p. 012087, 2020.

[17] Zhang, L., Xu, C., & Li, S., "Zhang, L., Xu, C., & Li, S. Facial expression recognition of infants based on multi-stream CNN fusion network.," *In 2020 IEEE 5th International Conference on Signal and Image Processing (ICSIP)IEEE.,* pp. 37-41, 2020.

[18] Liu, Yuanyuan, Wei Dai, Fang Fang, Yongquan Chen, Rui Huang, Run Wang, and Bo Wan. , "Dynamic multi-channel metric network for joint pose-aware and identity-invariant facial expression recognition," *Information Sciences,* vol. 578 , no. 2021, pp. 195-213, 2021.

[19] Wafi, M., Bachtiar, F. A., & Utaminingrum, F. , "Feature extraction comparison for facial expression recognition using adaptive extreme learning machine," *International Journal of Electrical and Computer Engineering,* vol. 13, no. 1, pp. 1113-1122, 2023.

[20] KDEF database, "kaggle," [Online]. Available: https://www.kaggle.com/datasets/tom99763/testtt.

[21] Huang, Y., Chen, F., Lv, S., & Wang, X., " Facial expression recognition: A survey," *Symmetry,* vol. 11, no. 10, p. 1189, 2019.

[22] Shorten, C., & Khoshgoftaar, T. M. , " A survey on image data augmentation for deep learning," *Journal of big data*, vol. 6, no. 1, pp. 1-48., 2019.