NTU Journal of Engineering and Technology (ICSDT2025) special issue
DOI: https://doi.org/10.56286/ntujet.icsdt2025

IRAQI Academic Scientific Journals

P-ISSN: 2788-9971 E-ISSN: 2788-998X
NTU Journal of Engineering and Technology
Available online at: https://journals.ntu.edu.iq/index.php/NTU-JET/index

A Novel Face Emotion Recognition Based on Lite ResNet-50 for Embedded Systems

Elaf Abdulwahab [1], Ahmed Khazal Younis [1]
[1]Northern Technical University, Mosul, Iraq
elaf.salih@ntu.edu.iq , ahmedkhazal@ntu.edu.iq

Article Informations

Received: 15-03- 2025,
Accepted: 28-04-2025,
Published online: 01-03-2026

Corresponding author:
Name: E. Abdulwahab
Affiliation : Northern Technical University
Email: elaf.salih@ntu.edu.iq

Key Words:
Face emotion recognition,
Lite ResNet-50,
MobileNetV2,
Raspberry pi 4B.

A B S T R A C T

. Face Emotion Recognition (FER) is essential for improving human-computer interaction. Although deep learning has greatly enhanced the accuracy of FER, deploying these models on embedded devices is challenging due to limitations in computational power and memory. This study proposes a lite and effective FER system for real-time implementation in resource-constrained environments. Lite ResNet-50 and MobileNetV2 were compared for the classification of neutral and angry emotions using a carefully selected dataset. The focus was on measuring accuracy, inference speed, and resource efficiency. Real-time testing is also con-ducted to confirm their applicability in practice. The results show that Lite Res-Net-50 outperforms MobileNetV2 in all key areas, achieving an accuracy of 99.8, inference speeds of 229 ms and with 2 FPS. These findings establish Lite Res-Net-50 as the optimal choice for FER on embedded devices, bridging the gap be-tween deep learning advancements and real-world deployment to improve hu-man-computer interaction.

1. Introduction

Facial emotions have become an essential aspect of human interaction, as they naturally express different feelings that help in the interaction and understanding between humans. Face Emotion Recognition, therefore, is about automatically determining emotions based on facial expressions. Some of its applications include human-computer interaction, analysis of mental states, surveillance, and customer feedback [1].

Early FER systems relied heavily on handcrafted features and classical machine learning techniques, such as Local Binary Patterns (LBP), and Support Vector Machines (SVMs) [2]. While these methods achieved a certain degree of success, they were highly sensitive to variations in lighting conditions, facial poses, occlusions, and individual differences, which made them less robust in real-world scenarios. Besides, these approaches required extensive feature engineering and had no generalization in datasets of various diversity [3].

However, recent advancements in deep learning, especially Convolutional Neural Networks (CNN), have greatly enhanced FER accuracy by learning detailed, hierarchical facial representations [4]. Also, the OpenCV library has played a very crucial part in real-time detection [5][6], acting as a bridge between raw image data and deep models by providing precise face detection, preprocessing, and feature extraction efficiently [7].

Although deep learning has enhanced the accuracy of FER, deploying these models in real-time on embedded devices such as raspberry pi or FPGA [8] which is still challenging due to computational power and memory constraints. Many existing models are too large for practical use on systems with restricted resources [9]. Another critical challenge is the quality of the publicly available FER datasets, most of which contain low-resolution grayscale images lacking detailed color and texture information that may lower the recognition accuracy. Additionally, imbalanced class distributions, where certain emotions are overrepresented while others are underrepresented, lead to biased model predictions and hinder generalization in real-world scenarios [10].

This study addresses the main challenge of developing a lightweight yet accurate deep learning model that can effectively classify face emotions while ensuring low memory usage and high inference speed on embedded devices. It also addresses the careful selection of the dataset.

Our main research paper contributions are listed as follows:

1. Develop a user-friendly, real-time FER system capable of accurately classifying neutral and angry expressions.
2. Implement and evaluate two lightweight deep learning architectures, Lite ResNet-50 (a streamlined version of ResNet-50) and MobileNetV2, to enable efficient FER on embedded systems.
3. Applied various optimization techniques, including transfer learning and model architecture reduction, to improve computational efficiency while maintaining high accuracy.
4. The findings of the suggested system demonstrate that the Lite ResNet-50 outperforms both its original architecture and MobileNetV2 in terms of accuracy, inference speed, and real-time performance, making it the superior choice for embedded deployment.

The rest of the paper is organized as follows: Section 2 presents related studies, Section 3 demonstrates the methodology, Section 4 provides results and a discussion, and Section 5 draws the conclusion.

2. Research Method

S. A. Hussain and A. Salim [11] developed a system for real-time detection, recognition, and classification of human emotions using deep learning frameworks. The methodology uses the Viola-Jones and Haar Cascade for face detection, while KDEF dataset and VGG-16 model were used for face recognition and classification. From the analyzed data, it is observed that this system has classified faces into seven emotional expressions; the KDEF dataset consists of 4,900, from which the side postures have been removed to achieve an accuracy of 88 %. They have also created a new dataset consisting of 100 images. The implementation relies on the Python programming language, which utilizes computer vision libraries like dlib and OpenCV. The system can analyze the emotions of autistic children and measure the engagement of students in e-learning environments to improve in the area of understanding human emotion.

E. Pranav et al. [12] proposed a Deep Convolutional Neural Network (DCNN) model for the classification of five facial emotions. A self-collected image dataset was utilized to identify the exact emotions of real-world scenarios. The architecture consists of two convolution layers to process the images, followed by a dropout layer for preventing overfitting, an activation function, and pooling layers, which decrease the size of dimensionality while preserving important information. Finally, the model converts the 2-D array to a 1-D vector that is fed into a neural network. The model implemented in Keras, which runs on top of TensorFlow, efficiently enables building and compilation in an environment such as Jupyter Notebook. The model is trained utilizing a dataset of five classes, with a total of 2,550 images collected manually using a mobile phone camera.

This process achieved an accuracy of 78.04%. The researchers recommend an extension to the model to analyze changes in emotions over video sequences.

Li and Lima [13] suggested ResNet-50 for feature extraction combined with CNN. ResNet-50 introduces residual connections and identity mapping in deep neural networks to solve degradation problems, meaning that the network's performance worsens as depth increases so that the information is preserved in the deeper networks while simultaneously reducing gradient dissipation and explosion. Therefore, it can extract complex features without degradation in performance and achieve better accuracy in higher layers. This design is able to improve the ability of training with better convergence and achieve better performance in multiclassification tasks compared to traditional deep networks in facial emotion recognition systems. The researchers used a new facial expression dataset consisting of 700 images collected from 20 subjects representing seven different facial emotions to achieve an accuracy of 95.39 ± 1.41%. Experimental simulations show that the proposed model outperforms mainstream facial emotion recognition models in performance and accuracy.

Y. Nan et al. [14] proposed a deep learning A-MobileNet model with an attention module integrated into MobileNetV1 for face expression recognition from static photos or videos. An attention module will be used to extract fine-grained features from facial expressions to make emotion recognition more accurate with subtle expressions in complex scenarios. Moreover, it uses center loss and softmax loss to adjust the model parameters that decrease intraclass distance and increase inter-class distance for better identification accuracy. Experiments were conducted using the FERPlus dataset, which includes eight emotion categories and a total of 31,412 grayscale facial images, achieving an accuracy of 88.11%. Besides, the RAF-DB dataset includes 7 categories of emotion with a total number of 15,539 color facial images, which shows an accuracy of 84.49%. Moreover, it maintains efficiency without increasing model parameters, making it well-suited for resource-constrained applications. The researchers have pointed out that the unbalanced training datasets will make the performance suboptimal in some expressions with fewer samples.

K. Sarvakar et al. [15] presented the two-stage CNN model, Facial Emotion Recognition Convolutional (FERC), which removes the background from images for clearer facial recognition and extracts the facial vector for emotion recognition. This model contains a double-layered CNN structure that can update its weights and parameters incessantly for developing the performances. The Expression Vector (EV) enables the identification of five common facial expressions and reduces problems such as camera distance. On the other hand, the system may miss some emotions, but it will recognize the majority. using the FER-2013 dataset, which consists of 35,887 images. The developed CNN model will have the capability to classify emotions in real-time. This ability enables instant responses to emotional signals.

H. W. Ahmed and A. M. Murshid [16] addresses these issues of the imbalanced datasets, and suboptimal feature extraction by leveraging deep learning-based FER using seven pre-trained CNN models, including InceptionResNetV2, InceptionV3, ResNet101, ResNet50, Xception, MobileNetV2, and GoogleNet. The FER2013 dataset (35,886 grayscale images across seven emotions) is used, with experiments conducted on both imbalanced and balanced versions via undersampling. The models are trained using Adam and SGD optimizers with varying batch sizes. Results show that InceptionResNetV2 achieves the highest accuracy of 65.47% on the imbalanced dataset, while undersampling improves classification balance. The Adam optimizer outperforms SGD, demonstrating its effectiveness in handling imbalanced data. ROC curve analysis validates model performance. This research emphasizes the importance of balancing datasets and choosing the appropriate optimizer to improve the accuracy of FER and, in general, deep learning-based FER systems for more robust and real-world applications. Table 1 Summarizes previous studies in face emotion recognition.

Table 1. Comparative table of related studies in face emotions recognition

Year/ Ref	Function	Approach	Accuracy	Description
2020/ [11]	Facial expressions	VGG-16	88 %	Simplify and improve the process of face emotion recognition
2020/ [12]	Facial Emotion Recognition	DCNN	78.04 %	Explain the architecture of a DCNN and its role in identifying five facial emotions.
2021/ [13]	Facial expression	ResNet-50	95.39 ± 1.41 %	Using ResNet-50 for feature extraction in order to solve the problem of network degradation that occurs with deeper networks.
2022/ [14]	Facial expression	A-MobileNet	84.49% and 88.11% for different datasets	Improve the MobileNetV1 architecture by combining it with the built-in attention module to reduce the parameters for a lightweight model.
2023/ [15]	Facial emotion recognition	CNN	54 %	Discuss the role of the two-stage FERC as well as the EV in face emotion recognition.
2024/ [16]	Facial Expression Recognition	Seven pre-trained deep learning models	65.47% for imbalanced dataset	Discuss the role of data preprocessing, model selection, and hyperparameter tuning in enhancing FER systems.

3. Methodology

Face emotion recognition technology is foreseen to improve user-computer interactions by detecting emotions such as anger or neutrality. This is a very important application in smart environments where automatically detected emotions could adjust lighting or temperature to build up an atmosphere that corresponds to a user's emotional state. This will provide better communication between humans and machines, fill the gap between verbal and non-verbal cues, and make systems more empathetic and responsive towards human needs. Two modern CNN architectures, MobileNetV2 and lite ResNet-50, were tested to determine their performance in real-time scenarios. (Fig. 1) illustrates a face emotion recognition system outlining the main steps for building, training, and evaluating the models.

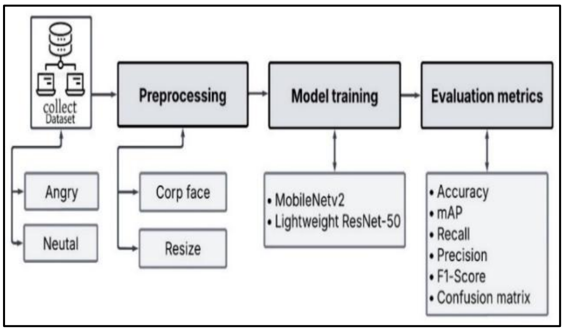

Fig. 1. The Main Steps for Building and Training the Model

3.1. Select dataset

The dataset consists of two emotion classes: angry and neutral. As an example, this study focuses on distinguishing between these two specific emotions for accurate classification. The images were carefully selected from the Getty Images website [17]. Each class includes 360 images of varying ages and forms. The dataset was selected in such a way that the age, gender, and cultural backgrounds were diverse to ensure inclusivity and generalization. The key factors included clear and distinguishable facial expressions, high image quality, and consistency in lighting and posture. Fig. 2 displays samples of the images that were utilized.

Fig. 2. Face Emotion Recognition Dataset Samples

3.2. Preparing dataset

The dataset was preprocessed for the improvement of the model by highlighting the features and ensuring uniformity across all images. First, face cropping was done to crop the face region, which would allow the model to focus entirely on face expressions while excluding irrelevant background details. Following this, data augmentation techniques were applied, including adjustments to lighting and random rotations, to increase the dataset's diversity to 1030 images in each class. After augmentation, all images were resized to a consistent resolution of 224 × 224 pixels to standardize input dimensions for the model. Finally, the preprocessed images were arranged in a structured folder, with each class of emotion placed in separate subfolders. The preprocessing steps prepare the dataset for training by enabling the model to pay attention to the relevant face features effectively. The preparation steps can be seen in Fig. 3.

Fig. 3. Dataset Preparation Steps

3.3. Training the model

The training of the face emotion recognition task was conducted in the Kaggle Notebook environment [18], taking advantage of the GPU resources for efficient computation. The whole idea was to balance model accuracy, computational efficiency, and resource utilization toward building lightweight models for deployment on embedded devices.

The dataset was first divided into a training set and a testing set. As shown in Fig. 4, this partitioning allocated 64% of the entire dataset for training, 16% for validation, and 20% for testing. Given the complexity of the task, this partitioning was determined to be the most suitable for the model.

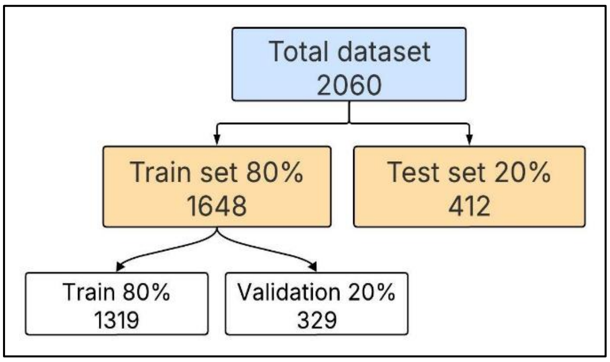

Fig. 4. Dataset Partitioning

Several models and training approaches were explored, including MobileNetV2 with transfer learning and variations of ResNet-50. MobileNetV2, renowned for its lightweight design and efficiency, was adapted for face emotion recognition by leveraging its pre-trained weights on ImageNet dataset [19]. Additional layers were appended to the pre-trained model, including a fully connected Dense layer with 256 neurons and ReLU activation, followed by a Dropout layer to mitigate overfitting. These adjustments enabled the model to effectively learn task-specific representations while benefiting from the robust feature extraction capabilities of MobileNetV2.

For ResNet-50 [20], multiple approaches were investigated to evaluate the relationship between model complexity and performance. The original ResNet-50 architecture, with its default 16 bottleneck blocks, was trained from scratch. To create lightweight variants, the number of bottleneck blocks was progressively reduced to 12, 10, 8, and even 4, significantly lowering computational costs, memory usage, and training time. Transfer learning was also applied to the original ResNet-50 by using pre-trained weights on ImageNet dataset and modifying the top layers to align with the number of emotion classes, retaining the model's pre-trained feature extraction capabilities.

Each approach was carefully designed to analyze the impact of model complexity on accuracy, training time, and resource usage. The lite ResNet-50 variants were much more efficient, with careful performance trade-offs to ensure minimum degradation in accuracy. Fig. 5 presents the architecture of the lite ResNet-50.

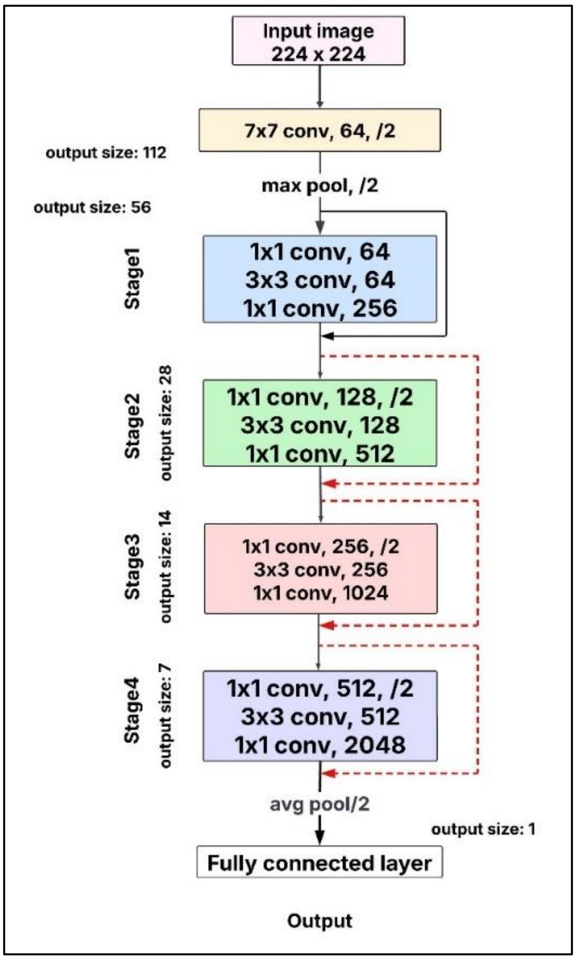

Fig. 5. Lite ResNet-50 Architecture

Table 2 shows the total parameters and used layers for the models. The parameters are the trainable weights and biases of a deep learning model that are learned during training. Parameters primarily include the weights of convolutional layers, which extract features from the input data, and the biases, which help refine the activation outputs. The number of parameters has a direct impact on the complexity, computational expense, and memory of a model. Additional parameters typically enable a model to learn more complicated patterns, but they can also lead to slower processing time and higher overfitting possibilities.
The original ResNet-50 contains 23.5 million parameters for trainable weights and biases. When reducing the size of the model, fewer parameters are used. Lite ResNet-50 and MobileNetV2 have 8 million and 2.5 million parameters, respectively. Both models are suitable for lower computational costs and are more appropriate for embedded applications like deployment on a Raspberry Pi.

Table 2. Model's parameters.

Model	Parameters	Layers
Original ResNet-50	23.5 M	50
Transfer learning	23.5 M	50
ResNet-50 reduced-12	19.9 M	37
ResNet-50 reduced-8	13.9 M	26
Lite ResNet-50	8 M	14
MobilenetV2	2.5 M	53

3.4. Model inference

Model inference is the process that occurs after training the model; it refers to using a trained model to make predictions based on new data. The goal of model inference is to extract valuable information from data that the model has not been trained on or seen, essentially allowing the model to predict outcomes based on its previous learning. where the main code for the real-time prediction is written and run using Visual Studio Code (VSC) based on the programming language for developing software, Python [21].

3.5. Evaluating metrics

In order to assess the face emotion recognition system's effectiveness, it's crucial to evaluate the model's performance to determine its reliability. The widely accepted metrics were employed such as accuracy, recall, precision, and confusion matrix to accomplish this.

- Accuracy: Accuracy is commonly considered to be the most explicit measure of performance, as it measures the fraction of data that were correctly predicted relative to the entire number predicted [22]. The equation is formulated below.

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (1)$$

Where:
- True Positives (TP): These are the cases where a model predicts the positive class correctly.
- True Negatives (TN): Those cases where the model correctly predicts the negative class.
- False Positives (FP): These are the cases where the model predicts the positive class when the actual class is negative.
- False Negatives (FN): Those cases when the model has predicted the negative class while it is actually positive.

- Recall: Also known as sensitivity or true positive rate, it is the ratio of correctly predicted positive observations to the total number of positive observations [23], as shown in the following equation.

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

- Precision: The model's precision is calculated based on the total number of guesses it makes. The next step is to divide the number of correct predictions by the total predictions [24], giving us the relationship between the total predictions TP+FP and the prediction percentage TP. An equation representing this information is provided below.

$$Precision = \frac{TP}{TP+FP} \quad (3)$$

- Confusion matrix: This matrix visualizes the performance of an algorithm to help the user understand where prediction errors have occurred. It shows actual class instances in rows and the predicted class instances in columns. In addition to presenting errors graphically, the confusion matrix can also incorporate other helpful metrics, including accuracy, F1-score, and recall. Fig. 6 presents the confusion matrix [25].

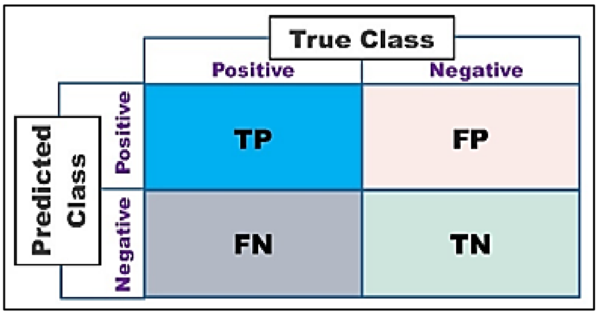

Fig. 6. Confusion Matrix [25]

4. Results and Discussion

This section will discuss two main categories: first, the results obtained from training the model, and second, the inference environment.

4.1. Training results

The training hardware environment utilizes a computer with Windows 10 Pro, Intel(R) Core (TM) i5-6300U CPU, and 8.00 GB RAM. Visual Studio Code (VSC) served as the integrated development environment (IDE) for coding depending on Python 3.11.5.

The key metrics considered to evaluate the performance of MobileNetV2 and variants of ResNet-50 are summarized in Table 3 to assess the feasibility and effectiveness of the proposed approach. These metrics reflect that the model is doing very well, hence providing reliable emotion recognition without sacrificing its lightweight architecture.

Table 3. Comparison of the performance of face emotion recognition based on the evaluation metrics.

Model	Train accuracy	Valid accuracy	Test accuracy	Epoch	Time	File size
Original ResNet-50	98.4	91.4	93.2	12	00:03:21	270 MB
Transfer learning	1	99	99.2	33	00:03:31	90.5 MB
ResNet50-12	98	76.9	98.5	17	00:03:47	228 MB
ResNet-50-8	99.5	78.7	97.3	12	00:02:11	160 MB
Lite ResNet-50	99.8	97.5	97	16	00:02:01	92.1 MB
MobileNetV2	99	95.4	94.4	25	00:01:57	12.8 MB

The most important key finding in Table III can be summarized as follows:

- The original ResNet-50 has a very large file size of 270 MB and relatively long training time, although it achieves high training accuracy and decent testing accuracy, making it less practical for deployment. While transfer learning improved the original ResNet-50 by reducing its file size to 90.5 MB and slightly enhancing testing accuracy, it still falls short of the lite ResNet-50 in terms of overall efficiency.
- The ResNet-50 models with 12 and 8 bottlenecks showed strong training accuracies, but their performance was less consistent. On the other hand, ResNet-50 with 8 bottlenecks could result in overfitting, showing lower testing accuracy while its training accuracy was higher. These variations bring into view how careful optimization is very important.
- The lite ResNet-50 has proven to be the most practical and efficient solution among all the models evaluated. It achieved the highest training accuracy, strong validation accuracy, and competitive testing accuracy with a compact file size of 92.1 MB and the fastest training time.
- The MobileNetV2 model offers an ultra-compact file size and a short training time, its slightly lower testing accuracy compared to the lite ResNet-50 makes it more suitable for situations where extreme resource constraints are prioritized over maximum accuracy.

For further analysis of the results, (Fig. 7) presents the confusion matrix that has been applied to the test dataset for all the used models.

- The original ResNet-50 model has the highest percentage of misclassifications for all the models tested in this approach, hence showing underperformance compared to other models in making class predictions.
- ResNet-50 with transfer learning reduces the misclassifications significantly, thus proving the value of using transfer learning in improving model performance.
- Comparing the lite ResNet-50 with the MobileNetV2, the lite version of ResNet-50 outperformed MobileNetV2 with lower misclassifications, proving to be superior again in terms of accuracy and also reliability

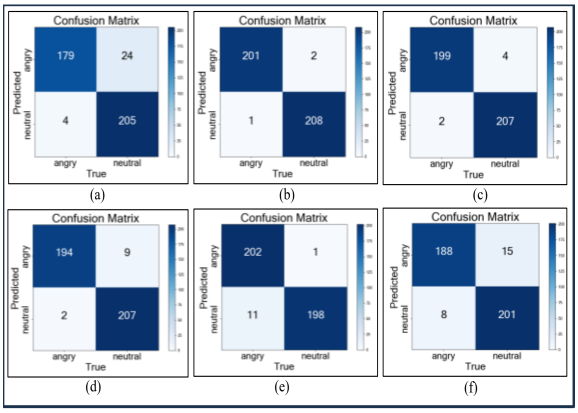

Fig. 7. The Confusion Matrix for: (a) Original ResNet-50, (b) ResNet-50 Transfer Learning, (c) ResNet-50 with 12 Bottlenecks, (d) ResNet-50 with 8 Bottlenecks, (e) Lightweight ResNet-50, (f) ModileNetV2

Finally, Table 4. presents a performance comparison of the utilized models with other commonly used algorithms in face emotion recognition.

Table 4. Comparing the Proposed System with Earlier Studies in Face Emotion Recognition

Year/ Study	Model	Description	Accuracy
2020 /[11]	VGG-16	Simplify and improve the process of face emotion recognition.	88 %
2020 /[12]	DCNN	Explain the role of DCNN in identifying facial emotions.	78.04 %
2021 /[13]	ResNet-50	Using ResNet-50 for feature extraction.	95.39 ± 1.41 %
2022 /[14]	A-MobileNet	Combining MobileNetV1 with the built-in attention module to reduce the parameters.	84.49% and 88.11%
2023 /[15]	CNN	Discuss the role of the two-stage FERC as well as the EV in face emotion recognition.	54 %
2024/ [16]	Seven pre-trained deep learning models	Discuss the role of data preprocessing, model selection, and hyperparameter tuning in enhancing FER systems.	65.47% for imbalanced dataset
Proposed model	Lite ResNet-50	Design a lite face emotion recognition for an embedded device in smart home automation.	99.8%

Ultimately, the lite ResNet-50 shows the best option that provides the perfect balance between accuracy, efficiency, and practicality. Its high performance across training, validation, and test sets, with its minimal computation cost, makes it perfect for real-world applications where precision and efficiency are reasonably required, especially in resource-constrained environments.

4.2. Inference environment

The hardware used to run the proposed system is the Raspberry Pi 4 Model B (RPi 4B), the RPi 4B is a small, affordable, and versatile single-board computer developed by the Raspberry Pi Foundation. It is powered by Raspberry Pi OS, which is a Debian-based operating system [26].

The RPi 4B is driven by a quad-core 64-bit ARM Cortex-A72 processor. Equipped with up to 8 GB of LPDDR4 RAM. For peripheral connectivity, it has two USB 2.0 ports and also two USB 3.0 ports on the board. It has dual micro-HDMI for display purposes, each supporting up to 4K resolution. The board is powered through a USB-C connector; however, a minimum 3A power supply is recommended for the best experience. All these features together make the RPi4B one of the most versatile and capable platforms for a number of computing projects. Fig. 8 presents the proposed system schematic.

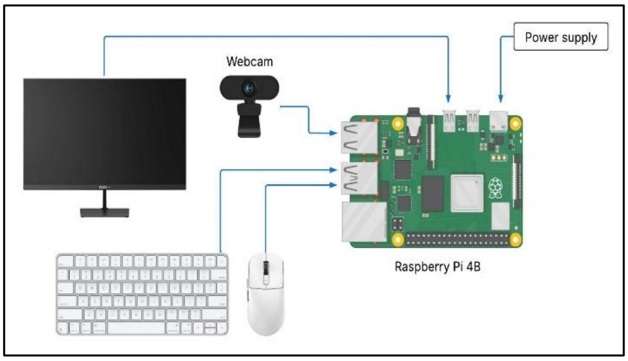

Fig. 8. Schematic Diagram of the FER System

Finally, the results for the real-time face emotion recognition system can be seen in Fig. 9. The proposed approach recognizes the emotion when the system captures their face as being either at a neutral stage or in anger. Once this emotion of the individual has been recognized to be that of anger, it sends impulses to instantly alter the mechanism of the cooling system, turning the fan on for maintaining the individual in a comfortable and composed condition.

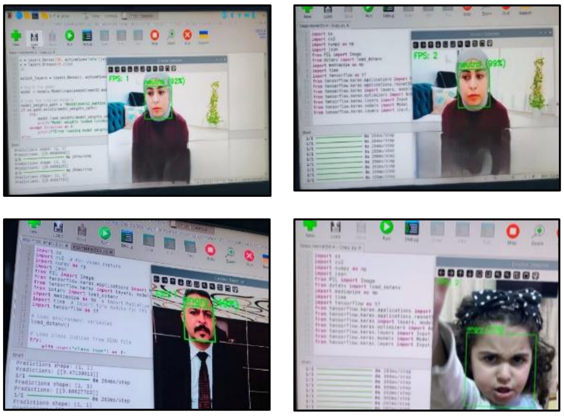

Fig. 9. Models Real-time Test Results: (a) MobileNetV2 for the Neutral Case, (b) Lite ResNet-50 for the Neutral Case, (c) MobileNetV2 for the Angry Case, and (d) Lite ResNet-50 for the Angry Case

5. Conclusion

In conclusion, this research introduces a lite and efficient FER system designed for real-time deployment on embedded devices, specifically the Raspberry Pi 4 Model B. We evaluated multiple deep learning architectures, and the experiments show that the Lite ResNet-50 model consistently outperforms both the original ResNet-50 variants and MobileNetV2 in classifying neutral and angry expressions. With a training accuracy of 99.8% and an inference speed of 229 ms, Lite ResNet-50 demonstrates strong performance with minimal misclassifications, as confirmed by confusion matrix analysis.

While MobileNetV2 provides an ultra-compact solution with a file size of only 12.8 MB, its slightly lower testing accuracy limits its practical use in applications where precision is critical. Overall, the Lite ResNet-50 model represents a great leap forward in FER, enabling better human-computer interaction with accurate real-time emotion recognition.

6. Future Works

This study successfully demonstrates the effectiveness of Lite ResNet-50 for real-time FER on embedded devices. However, there are several areas that require further research:

1. One of the most important areas of future research is to expand the model's ability to detect more emotions beyond the neutral and the angry to make the model more effective in most real-world applications.
2. Another key area of work is to explore other lightweight algorithms in order to achieve even lower inference time and higher frame rates.
3. Finally, the future works can involve comparisons of the model's performance on other embedded systems, such as the NVIDIA Jetson series, and FPGA-based platforms to decide on the optimum hardware for real-time FER systems.

References

[1] V. J, T. Kanwar, N. K. Rai, S. K. Nagesh, and Y. Nag, “A Multimodel Deep Learning Based Facial Emotion Recognition and Detection with Hyper Parameters,” *2024 Int. Conf. Adv. Comput. Res. Sci. Eng. Technol. ACROSET 2024*, pp. 1–6, 2024, doi: 10.1109/ACROSET62108.2024.10743754.

[2] A. K. Younis and T. Ercan, “Design of an FPGA-based Intelligent Gate,” *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 8, pp. 126–130, Jan. 2019.

[3] J. Li, K. Jin, D. Zhou, N. Kubota, and Z. Ju, “Attention mechanism-based CNN for facial expression recognition,” *Neurocomputing*, vol. 411, pp. 340–350, 2020, doi: 10.1016/j.neucom.2020.06.014.

[4] Y. Wang, Y. Li, Y. Song, and X. Rong, “The influence of the activation function in a convolution neural network model of facial expression recognition,” *Appl. Sci.*, vol. 10, no. 5, 2020, doi: 10.3390/app10051897.

[5] A. K. Younis, B. M. K. Younis, and M. S. Jarjees, “Hardware implementation of Sobel edge detection system for blood cells images-based field programmable gate array,” *Indones. J. Electr. Eng. Comput. Sci.*, vol. 26, no. 1, pp. 86–95, 2022, doi: 10.11591/ijeecs.v26.i1.pp86-95.

[6] D. Alhelal, A. K. Younis, and R. H. A. Al-Mallah, “Detection of brain stroke in the MRI image using FPGA,” *Telkomnika (Telecommunication Comput. Electron. Control.*, vol. 14, no. 5, pp. 1307–1315, 2021, doi: 10.12928/TELKOMNIKA.v19i4.18988.

[7] A. M. Riayth, M. S. Jarjees, and A. K. Younis, “A Review of Sickle Cells Classification System Based on Microscopic Images,” *AIP Conf. Proc.*, vol. 2862, no. 1, 2023, doi: 10.1063/5.0181048.

[8] B. M. K. Younis and A. K. Younis, “Hardware accelerator for anti-aliasing Wu’s line algorithm using FPGA,” *TELKOMNIKA (Telecommunication Comput. Electron. Control.*, vol. 19, no. 2, pp. 672–682, Apr. 2021, doi: 10.12928/TELKOMNIKA.V19I2.18158.

[9] S. Mikhaylevskiy, V. Chernyavskiy, V. Pavlishen, I. Romanova, and R. Solovyev, “Fast Emotion Recognition Neural Network for IoT Devices,” *Int. Semin. Electron Devices Des. Prod. SED 2021 - Proc.*, 2021, doi: 10.1109/SED51197.2021.9444517.

[10] R. A. Elsheikh, M. A. Mohamed, A. M. Abou-Taleb, and M. M. Ata, “Improved facial emotion recognition model based on a novel deep convolutional structure,” *Sci. Rep.*, vol. 14, no. 1, pp. 1–31, 2024, doi: 10.1038/s41598-024-79167-8.

[11] S. A. Hussain and A. Salim Abdallah Al Balushi, “A real time face emotion classification and recognition using deep learning model,” in *Journal of Physics: Conference Series*, 2020, pp. 0–13. doi: 10.1088/1742-6596/1432/1/012087.

[12] P. E., S. Kamal, S. Chandran C., and S. M. H., “Facial Emotion Recognition Using Deep Convolutional Neural Network,” in *2020 6th International conference on advanced computing and communication Systems (ICACCS)*, IEEE, 2020, pp. 317–320.

[13] B. Li and D. Lima, “Facial expression recognition via ResNet-50,” *Int. J. Cogn. Comput. Eng.*, vol. 2, no. January, pp. 57–64, 2021, doi: 10.1016/j.ijcce.2021.02.002.

[14] Y. Nan, J. Ju, Q. Hua, H. Zhang, and B. Wang, “A-MobileNet: An approach of facial expression recognition,” *Alexandria Eng. J.*, vol. 61, no. 6, pp. 4435–4444, 2022, doi: 10.1016/j.aej.2021.09.066.

[15] K. Sarvakar, R. Senkamalavalli, S. Raghavendra, J. S. Kumar, and R. Manjunath, “Facial emotion recognition using convolutional neural networks,” *Mater. Today Proc.*, vol. 80, pp. 3560–3564, 2023, doi: 10.1016/j.matpr.2021.07.297.

[16] H. W. Ahmed and A. M. Murshid, “Facial Expression Recognition using deep Convolutional Neural Network,” *J. Comput. Anal. Appl.*, vol. 33, no. 08, pp. 532–542, Sep. 2024, Accessed: Feb. 08, 2025. [Online]. Available: https://eudoxuspress.com/index.php/pub/article/view/1367

[17] “Royalty-Free Stock Photos, Creative Images & Vectors | News, Fashion, and Entertainment

imagery - Getty Images." Accessed: Jan. 21, 2025. [Online]. Available: https://www.gettyimages.com/

[18] "Kaggle: Your Home for Data Science." Accessed: Jan. 07, 2025. [Online]. Available: https://www.kaggle.com/

[19] "ImageNet." Accessed: Dec. 27, 2024. [Online]. Available: https://www.image-net.org/

[20] O. Alsaif and M. L. Muammer, "Pathological outcomes of covid-19 for lungs infections based on transfer learning technology," *3C Tecnol. innovación Apl. a la pyme*, vol. 12, no. 01, pp. 282–294, 2023, doi: 10.17993/3ctecno.2023.v12n1e43.282-294.

[21] E. Abdulwahab and A. K. Younis, "Surveillance System Based on Face Recognition Utilizing Deep Learning Technique," in *2024 1st International Conference on Emerging Technologies for Dependable Internet of Things (ICETI)*, IEEE, 2024, pp. 1–6. doi: 10.1109/ICETI63946.2024.10777082.

[22] R. H. M. Ameen, N. M. Basheer, and A. K. Younis, "Breast cancer diagnosis based on support vector machine techniques," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 32, no. 1, pp. 236–243, 2023, doi: 10.11591/ijeecs.v32.i1.pp236-243.

[23] L. A. Hussein and Z. S. Mohammed, "Static Arabic Sign Language Recognition in Real Time Using Machine Learning and MediaPipe," *1st Int. Conf. Emerg. Technol. Dependable Internet Things, ICETI 2024*, pp. 1–8, 2024, doi: 10.1109/ICETI63946.2024.10777193.

[24] R. M. Al-Alaf, F. Al-Azzo, and T. Thabet, "Building A Face Mask Recognition System Supported with Facial Landmark Features in Real-Time," *AIP Conf. Proc.*, vol. 2862, no. 1, 2023, doi: 10.1063/5.0172307.

[25] M. Sadiq, M. N. Kadhim, D. Al-Shammary, and M. Milanova, "Novel EEG Classification based on Hellinger Distance for Seizure Epilepsy Detection," *IEEE Access*, vol. PP, p. 1, 2024, doi: 10.1109/ACCESS.2024.3450449.

[26] Raspberry Pi Trading Ltd., "Raspberry Pi 4 Computer Model B," *Raspberrypi.Org*, no. May, pp. 1–6, 2019, [Online]. Available: www.raspberrypi.org