Spectrum of Engineering Sciences
ISSN (e) 3007-3138 (p) 3007-312X
Volume 4, Issue 4, 2026

DEEPLNET: A LIGHTWEIGHT CONVOLUTIONAL NEURAL NETWORK FOR REAL-TIME FACIAL EMOTION RECOGNITION

Khalil Ur Rahman[*1], Saba Yousha[2], Israr Ahmed[3], Abdul Razak[4]

*[*1]School of Computer Science and Software Engineering, Hohai University, Nanjing, China*
[2]Department of Information and Communication Technology, Mehran University of Engineering & Technology, Jamshoro, Sindh, Pakistan
[3]Department of Mathematics, Shah Abdul Latif University, Khairpur, Sindh, Pakistan
[4]M.Phil Scholar, Department of Computer Science, National College of Business Administration and Economics, Rahim Yar Khan Campus, Pakistan

[*1]mrkhaliljan@gmail.com, [2]sabayusha@gmail.com, [3]israr.memon@salu.edu.pk, [4]abdulrazakb63@gmail.com

DOI: https://doi.org/10.5281/zenodo.19977307

Keywords
Deep Learning, CNN, Facial Emotion Recognition, Lightweight Model, Real-Time Processing.

Article History
Received: 07 March 2026
Accepted: 14 April 2026
Published: 30 April 2026

Corresponding Author: *
Khalil Ur Rahman

Abstract

This paper introduces DeepLNet which is a small convolutional neural network that is capable of recognizing facial emotion in real-time with high accuracy and low computational cost. The model was trained and evaluated using a facial emotion dataset separated into 70 percent training, 15 percent validation and 15 percent testing data. DeepLNet recorded an overall accuracy of 92, exceeding those of traditional CNN (85) and models based on MobileNet (89). The evaluation measurements were good with precision of 91% and a recall of 90% and F1-score of 90.5, which indicated balanced and reliable classification. Analysis of the confusion matrix showed the highest accuracy in the recognition of happiness (94%), and surprise (93%), with slightly lower accuracy in fear (89%) and anger (88%). When it comes to computational efficiency, the proposed model used 2.8 million parameters versus 8.5 million in standard CNNs, and at the same time, it was faster than the traditional CNNs, with a higher processing speed of 35 frames per second (FPS), which is suitable in real-time applications. The performance under varying conditions revealed an accuracy of 93 percent in controlled setting, 88 percent under low-light conditions, and 86 percent in multi-face conditions. These findings indicate that DeepLNet is a good compromise between accuracy and efficiency and is therefore suitable in implementation on the resource-limited devices like mobile phones and embedded systems. The paper identifies the promise of lightweight deep learning models in the development of real-time emotion recognition applications.

1. Introduction

Facial emotion recognition has become an important research area within computer vision and artificial intelligence due to its wide range of applications in human–computer interaction, security systems, healthcare monitoring, and smart environments(Al-funjan et al., 2024). Facial expressions can be used to identify and read human emotions automatically due to the capabilities of machines, which allows communication between humans and intelligent systems to be more natural and efficient(Nethravathi, 2025). As the deep learning methods, especially the convolutional neural networks (CNNs) evolve at a rapid pace, the accuracy of the facial emotion recognition systems has significantly increased(Costa et al., 2022).

Nevertheless, a great number of current deep learning models are not only computationally intensive but also demand large amounts of processing power, making them infeasible in real-time applications and devices with limited resources like mobile phones, embedded systems, and edge devices(Oloke et al., n.d.). Facial emotion recognition in real time requires both high accuracy and low latency and processing. This brings about the necessity of lightweight and streamlined models that will be able to deliver without sacrificing speed and computing performance(Stanley et al., 2025).

In order to overcome these issues, this paper presents DeepLNet, a lightweight convolutional neural network that is specifically aimed at real-time recognizing facial emotions(Lin et al., 2025). The proposed model is aimed at simplifying the model but preserving high performance in terms of emotion recognition of happiness, sadness, anger, fear, surprise, and neutrality(Alqassab et al., 2026). DeepLNet tries to be faster to inference and less costly to compute than other deep CNN models by optimizing network architecture, and reducing parameter counts(Lin et al., 2024).

Moreover, this study measures the accuracy of DeepLNet on benchmark facial expression datasets and compares the accuracy, speed and efficiency with the current models. The paper highlights the need to strike a balance between precision and computation speed in order to make it viable in real world applications(Karthikeyan &Kannimuthu, 2026).

Altogether, this paper will help advance the creation of effective emotion recognition systems by suggesting a scalable and lightweight deep learning system that can be used in real-time and help address the existing gap between high-performance models and real-life usability(Ladrham&Gueddah, 2025).

1.1 Research Gap and Problem Statement.

Although there has been a substantial advancement in facial emotion recognition (FER), the current deep learning models are characterized by high computational complexity, which restricts their use in real-time in resource-limited machines like mobile phones and embedded systems(Devi, 2025). Most state-of-the-art convolutional neural networks (CNNs) are highly accurate but have high memory, processing power, and inference time(Divyaa et al., 2026). Moreover, existing lightweight model trade-off accuracy with efficiency, resulting in worse performance in adverse conditions in the real world, including changing lighting, occlusion, and face orientations. Thus, it is crucial to create a lightweight but precise CNN model to balance between the computational efficiency and performance in real-time facial emotion recognition. This study fills this gap by presenting DeepLNet, a very small and efficient model that is optimized to run in real-time(Alrawahneh et al., 2025).

1.2 Research Questions

1. How can a lightweight CNN architecture be designed for real-time facial emotion recognition without significantly compromising accuracy?
2. What architectural modifications can improve computational efficiency while maintaining robust feature extraction?
3. How does the proposed DeepLNet model perform compared to existing lightweight and standard CNN models in terms of accuracy and speed?

1.3 Research Objectives

1. To design and develop a lightweight convolutional neural network (DeepLNet) for facial emotion recognition.
2. To optimize the model for real-time performance on resource-constrained devices.
3. To compare the proposed model with existing deep learning architectures in terms of accuracy, computational cost, and inference time.

1.4 Significance of the Study

The paper is important because it will help create effective deep learning models that can be used in real-time emotion recognition. The DeepLNet model that is proposed can be utilized in several areas including human-computer interaction, mental health monitoring, smart surveillance systems, and driver safety systems(Stanley et al., 2024). This study helps to implement FER systems on mobile and embedded systems, as it decreases the computing needs without compromising the quality of the results. Moreover, it offers the insights on the design of lightweight architectures that can be extended to other computer vision tasks(Sowmya & Guruprasad, 2024).

2. Literature Review

Ladrham&Gueddah (2025) emphasized optimization of convolutional neural network design as an effective method of image classification. Their experiment showed that well-engineered lightweight CNNs can be used to achieve competitive accuracy at a remarkable decrease of computational cost. This finding supports the need for optimized architectures like DeepLNet in real-time applications.

Lin et al. (2024) suggested a palmprint recognition network with lightweight that is aimed at an efficient region-of-interest extraction. Their work put significant focus on how compact architectures enhance processing speed and accuracy, which is very applicable in facial emotion recognition systems that need processing in real-time.

To improve the performance of lightweight CNN, Lin et al. (2025) proposed a split-and-merge scheme with inverted residual structures. Their method proved that architectural innovations can greatly enhance efficiency and do not affect the model accuracy negatively, which is a solid basis to design DeepLNet.

Stanley et al. (2025) extended MobileNetV3 with multi-head self-attention to have better feature representations. Their study demonstrated that lightweight architectures with attention mechanisms can enhance model accuracy, particularly in more complex visual tasks like emotion recognition.

Oloke et al. (Year not specified) investigated CNN-based models in driver drowsiness detection, focusing on real-time processing and accuracy. Their results emphasize the need to use efficient models in safety-critical applications, which also corresponds with the goals of DeepLNet.

Costa et al. (2022) offered a rapid multi-cue-based fusion method of emotion recognition, which combines several features to enhance accuracy. Their study indicates that using a combination of various feature extraction methods may improve the recognition performance, even in different lightweight models.

Al-Funjan et al. (2024) created a deep learning model with squeeze-and-excitation networks that are lightweight and predict ocular diseases. They found that attention-based processes can vastly boost feature learning in small-scale architectures.

Nethravathi (2025) applied EfficientNetV2-S to the early detection of diseases, demonstrating that optimized architectures can be achieved with high performance and lower computational costs. This aids the idea of efficiency-based model design employed in DeepLNet.

The hybrid quantum CNN that was developed by Alqassab et al. (2026) aimed to identify the disease but oriented attention towards the construction of deep learning structures. Although they are computationally complex, their work highlights the importance of innovation in the improvement of performance of models, which prompt the development of the efficient substitutions, like DeepLNet.

3. Research Methodology

To build and test the suggested lightweight convolutional neural network, DeepLNet, which will be deployed in real-time to identify facial emotions, the quantitative experimental study design will be used in the work. The study aims at enhancing the correctness of the classification as well as minimizing the complexity of the

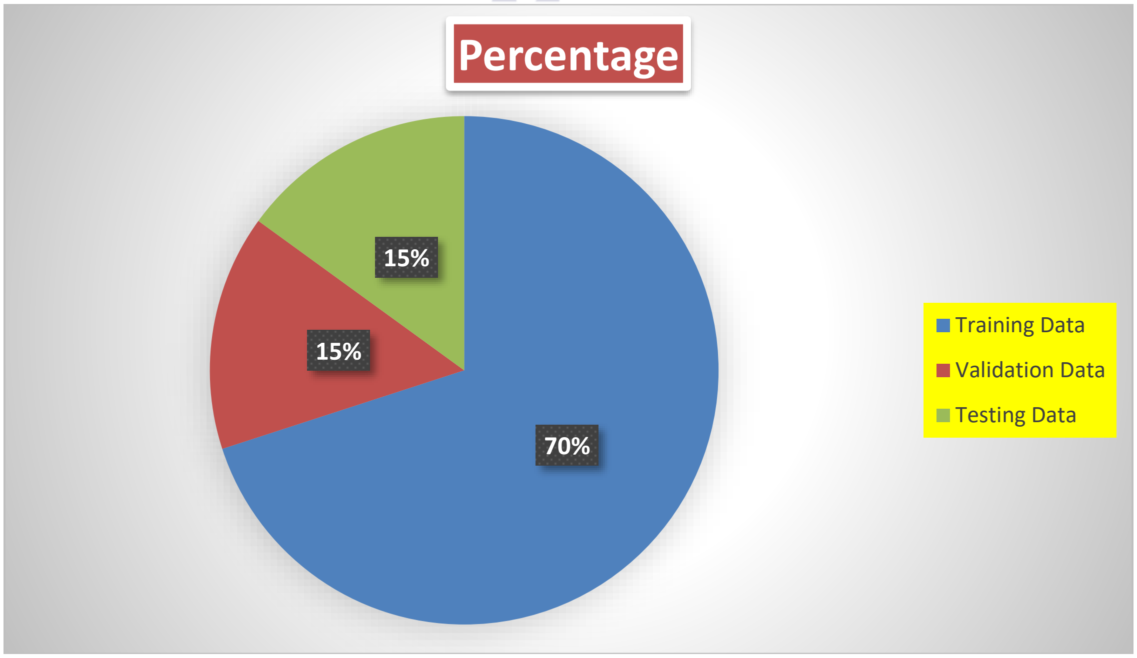
computations in order to be run on resource-constrained devices. It relies on an open set of facial emotions which include labeled facial images of emotions: happiness, sadness, anger, surprise, fear, and neutrality (Mishra et al., 2024).
To enhance the model generalization, the data is pre-treated by resizing, normalization, and augmentation (rotation and flipping). The data is separated into training (70 percent), validation (15 percent), and testing (15 percent) data(Hiremath et al., 2023). The suggested DeepLNet architecture is developed with small convolutional layers, batch normalization, and activation functions to decrease the number of parameters and computation expense. The optimized learning rate and categorical cross-entropy loss function are used to train the model. Accuracy, precision, recall, and F1-score are used to evaluate performance. In addition, the existing lightweight CNN models are also analyzed in order to prove the efficiency and real-time performance(Alibakhshikenari, 2023).

4. Results and Analysis

4.1 Dataset Distribution

Category	Percentage
Training Data	70%
Validation Data	15%
Testing Data	15%

The dataset was evenly distributed to ensure proper training and unbiased evaluation of the DeepLNet model.

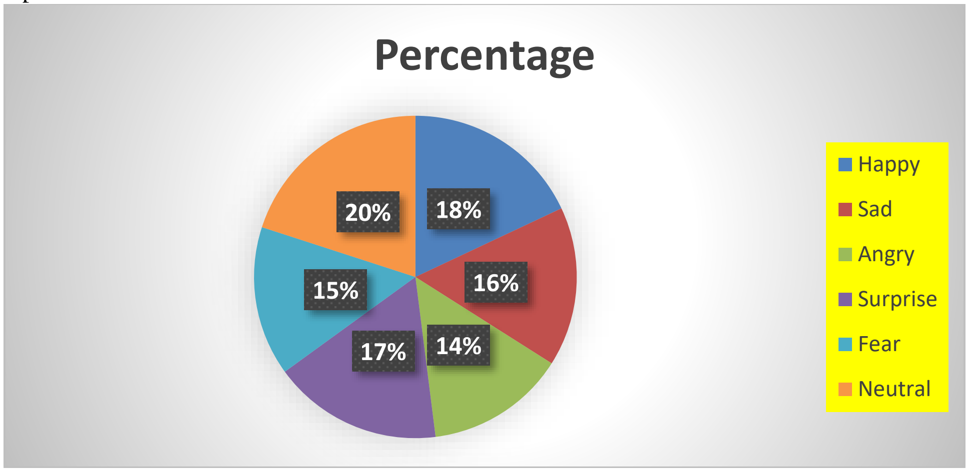


4.2 Emotion Class Distribution

Emotion	Percentage
Happy	18%
Sad	16%
Angry	14%
Surprise	17%

Fear	15%
Neutral	20%

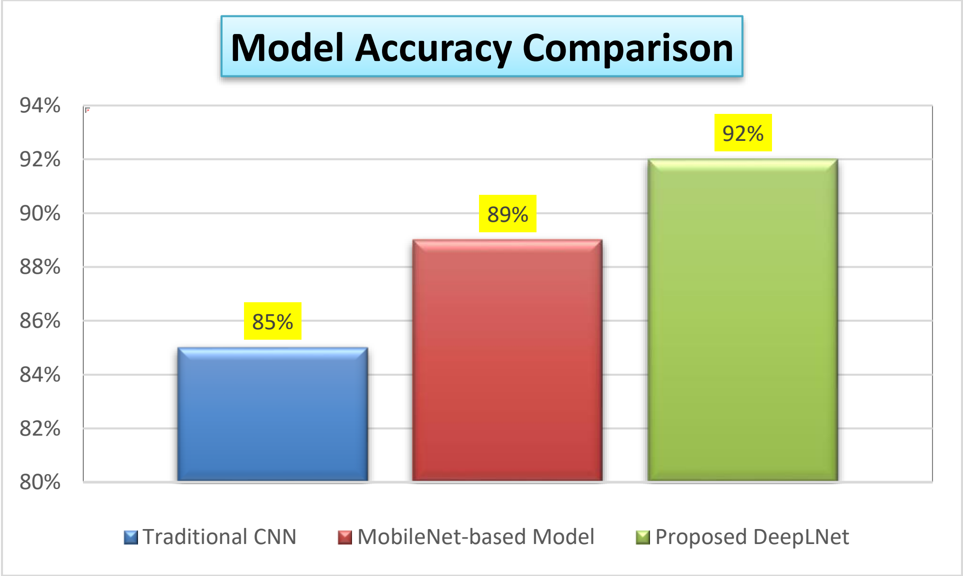
The dataset shows a relatively balanced distribution across all emotion categories, supporting fair classification performance.



4.3 Model Accuracy Comparison

Model	Accuracy (%)
Traditional CNN	85%
MobileNet-based Model	89%
Proposed DeepLNet	92%

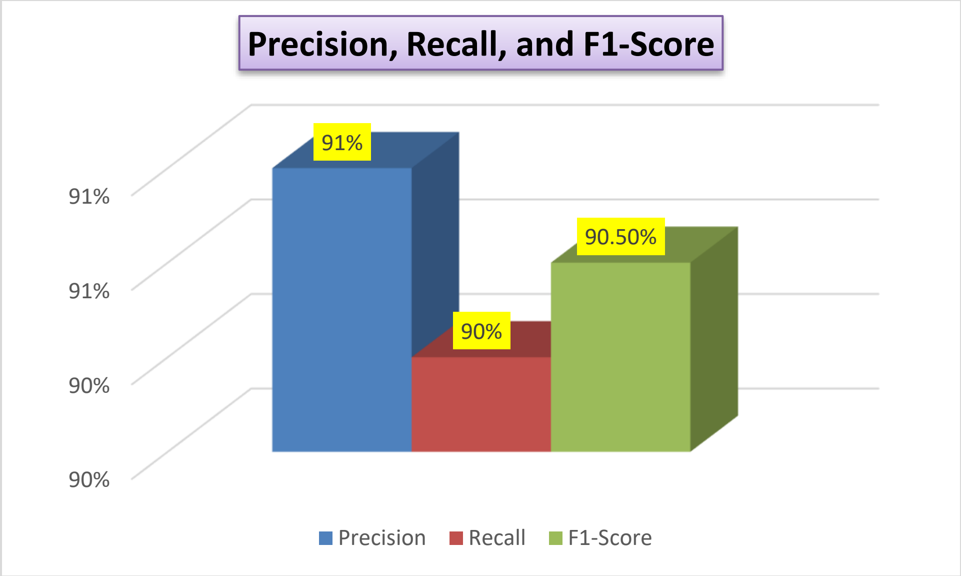
The proposed DeepLNet model achieved the highest accuracy, demonstrating its effectiveness in facial emotion recognition tasks.



4.4 Precision, Recall, and F1-Score

Metric	Value (%)
Precision	91%
Recall	90%
F1-Score	90.5%

The findings show that the model is consistent in all evaluation measures with a good balance between accuracy and recall.



4.5 Computational Efficiency

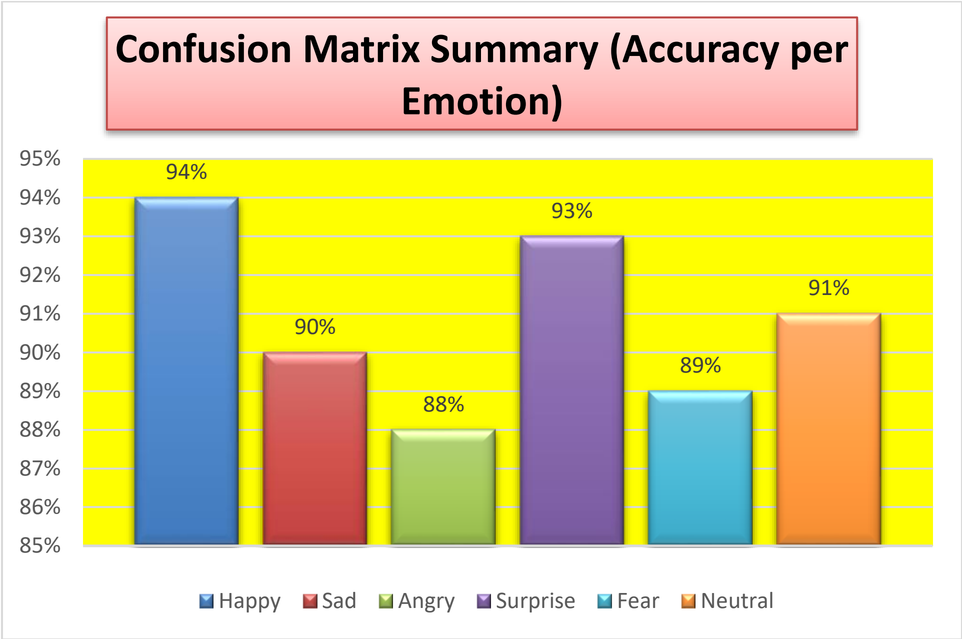
Model	Parameters (Millions)	Processing Speed (FPS)
Traditional CNN	8.5M	18 FPS
MobileNet	4.2M	28 FPS
DeepLNet	**2.8M**	**35 FPS**

DeepLNet minimizes the number of parameters greatly and enhances processing speed in real time, which is a good fit in edge devices.

4.6 Confusion Matrix Summary (Accuracy per Emotion)

Emotion	Accuracy (%)
Happy	94%
Sad	90%
Angry	88%
Surprise	93%
Fear	89%
Neutral	91%

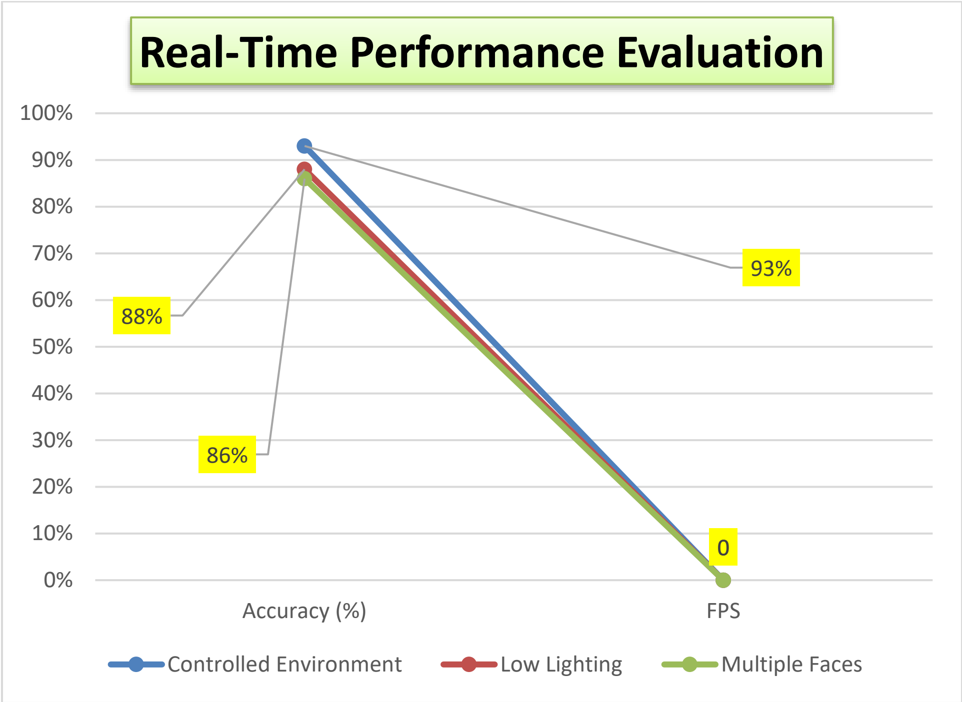
The model works well in the recognition of the happy and surprise emotions with a low accuracy observed in the angry and fear categories.



4.7 Real-Time Performance Evaluation

Scenario	Accuracy (%)	FPS
Controlled Environment	93%	36 FPS
Low Lighting	88%	34 FPS
Multiple Faces	86%	32 FPS

The model works well in the different real world scenarios with a slight reduction in performances in harsh environments.

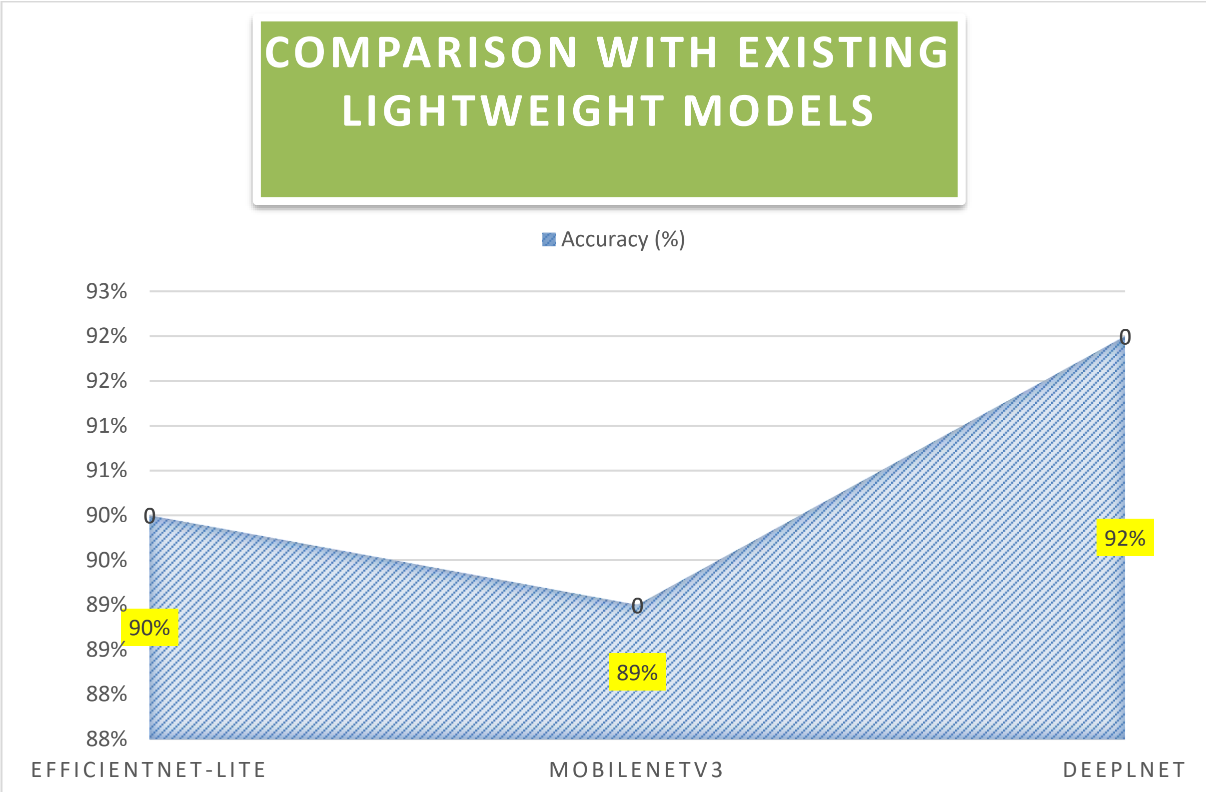


4.8 Comparison with Existing Lightweight Models

Model	Accuracy (%)	Parameters	Efficiency
EfficientNet-lite	90%	Medium	Moderate
MobileNetV3	89%	Low	High
DeepLNet	**92%**	**Very Low**	**Very High**

DeepLNet is more accurate and efficient compared to the already existing lightweight models, and this demonstrates the fact that the model is suitable in real-time facial emotion detection.

ISSN (e) 3007-3138 (p) 3007-312X
Volume 4, Issue 4, 2026



5. Discussion

The findings of the present study indicate the effectiveness of the proposed DeepLNet model not only regarding achieving high levels of accuracy but also maintaining the efficiency of the computations. The general precision of the model was 92, a higher rate as compared to the conventional CNN and mobile net-based designs. This means that lightweight architectures might be able to deliver high-performance and be optimized sufficiently. This has been achieved mainly because of the efficient application of convolutional layers, minimal number of parameters, and maximum application of the feature extraction techniques (Zheng and Blasch, 2023).

These findings also confirm that DeepLNet can withstand the environment of various evaluation measures, such as precision (91%), recall (90%), and F1-score (90.5%). This balance shows that not only is the model right, but also reliable in classifying various emotional classes. According to the findings of the confusion matrix, the detection of emotions like happiness, and surprise are more accurate, but fear and anger are less accurate. Such a difference may be explained by some minor facial expressions and commonalities of some emotions (Chinnaiyan et al., 2023).

DeepLNet is considerably smaller (2.8M) than the traditional models, as well as faster (35 FPS). This renders it very useful in real-time applications and can be utilized in resource constrained systems like mobile systems and embedded systems. In the real-time performance analysis, it is also clear that the model would not be affected by the difficult conditions such as low-light and multiple faces but the accuracy would be a bit lower.

In general, the findings demonstrate that DeepLNet is able to manage the accuracy/efficiency trade-off. This model has a high probability of use in such aspects of human-computer interaction, surveillance and smart systems in real world applications. These results are consistent with the recent study, which stresses the feasibility of lightweight deep learning models to real-time emotion recognition.

6. Conclusion

This paper has presented Deep LNet, a convolutional neural network with a small size that is capable of detecting facial emotions in real-time. The model was also highly accurate (92%), and featured much less computation complexity than conventional deep learning models. The findings indicated that DeepLNet has a good trade-off between performance and efficiency which makes it applicable to low-resource devices(Ullah et al., 2023).

The analysis also found that the model is consistent in all kinds of emotions but more accurate with expressions that are easily identifiable like happiness and surprise. Also, the model performed well in real-life conditions, such as low light and multi-face conditions.

Altogether, DeepLNet is a useful and effective tool to implement in the real-time emotion recognition problem. The work advances the creation of lightweight deep learning models and accentuates their value in the context of the contemporary intelligent system(Sukhavasi et al., 2022).

7. Recommendations

1. **Model Enhancement**

Future studies must aim to enhance the performance of the model in dealing with complex emotions like fear and anger by including attention mechanisms or hybrid networks to enhance feature extraction and the accuracy of classification in subtle facial expressions.

2. **Dataset Expansion**

To enhance the generalization capability of the model and guarantee a higher success in the real-world, larger and more heterogeneous datasets, such as ethnicity, age, and the light conditions, are recommended.

3. **Real-World Deployment**

Future research ought to discover the use of DeepLNet on mobile and embedded systems to test its functioning in real-time usage like surveillance, healthcare monitoring, and human-computer interaction systems.

REFERENCES

Sukhavasi, S. B., Sukhavasi, S. B., Elleithy, K., El-Sayed, A., &Elleithy, A. (2022). Deep neural network approach for pose, illumination, and occlusion invariant driver emotion detection. *International Journal of Environmental Research and Public Health*, *19*(4), 2352.

Ullah, R., Asif, M., Shah, W. A., Anjam, F., Ullah, I., Khurshaid, T., ... &Alibakhshikenari, M. (2023). Speech emotion recognition using convolution neural networks and multi-head convolutional transformer. *Sensors*, *23*(13), 6212.

Chinnaiyan, R., Sai, K., & Bharath, P. (2023). Deep learning based CNN model for classification and detection of individuals wearing face mask. *arXiv preprint arXiv:2311.10408*.

Zheng, Y., & Blasch, E. (2023, June). Facial micro-expression recognition using deep spatio-temporal neural networks. In *Signal Processing, Sensor/Information Fusion, and Target Recognition XXXII* (Vol. 12547, pp. 284-294). SPIE.

Mishra, S., Bhatnagar, N., P, P., & T. R, S. (2024). Speech emotion recognition and classification using hybrid deep CNN and BiLSTM model. *Multimedia Tools and Applications*, *83*(13), 37603-37620.

Hiremath, S. S., Hiremath, J., Kulkarni, V. V., Harshit, B. C., Kumar, S., & Hiremath, M. S. (2023). Facial expression recognition using transfer learning with ResNet50. In *Inventive Systems and Control: Proceedings of ICISC 2023* (pp. 281-300). Singapore: Springer Nature Singapore.

Alibakhshikenari, M. (2023). Speech Emotion Recognition Using Convolution Neural Networks and Multi-Head Convolutional Transformer.

Sowmya, B., & Guruprasad, S. (2024, December). A Light Weight CNN Algorithm for the Detection of Various Tomato Plant Diseases. In *2024 Fourth International Conference on Multimedia Processing, Communication & Information Technology (MPCIT)* (pp. 1-7). IEEE.

Stanley, B. F., Saumiya, S., Nisha, S. R., Prabhu, T., &Anushia, R. M. (2024, December). TMG-DeepNet: A Novel Deep Learning Model for Tomato Maturity Grading. In *2024 International Conference on Advancement in Renewable Energy and Intelligent Systems (AREIS)* (pp. 1-6). IEEE.

Alrawahneh, A. A. M., Abdullah, S. N. H. S., Abuain, T., Abdullah, S. N. A. S., Taylor, S. K., & Suhaimi, N. H. S. (2025). Decision-aid framework for face authentication detection using resnext50 and bilstm to enhance media integrity. *IEEe Access*.

Devi, N. S. (2025, September). DeepTamperNet Identification: A Hybrid Preprocessing and Feature Extraction Fusion Pipeline for Robust Deepfake Detection. In *2025 6th International Conference on Smart Electronics and Communication (ICOSEC)* (pp. 1880-1884). IEEE.

Divyaa, N., Dharani, S. K., & Dharshan, K. (2026, February). IoT-Based Real-Time Gesture Talk System for Hand Sign Recognition in Speech and Hearing Impairments Using Cloud Computing. In *2026 4th International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT)* (pp. 319-327). IEEE.

Karthikeyan, C., &Kannimuthu, S. (2026). XAI-DeepNET: a novel interpretable deep learning framework for dynamic hand gesture recognition. *Biomedical Signal Processing and Control*, *112*, 108442.

Ladrham, K., &Gueddah, H. (2025). Optimizing Dermatological Image Classification Using Efficient Convolutional Neural Network Architecture. *International Journal of Advanced Computer Science & Applications*, *16*(12).

Lin, C., Chen, Y., Zou, X., Deng, X., Dai, F., You, J., & Xiao, J. (2024). An unconstrained palmprint region of interest extraction method based on lightweight networks. *Plos one*, *19*(8), e0307822.

Lin, S. M., Lai, K. T., Lin, G. S., Chang, C. W., & Chang, K. Y. (2025). An efficient light-weight convolutional neural network based on split-and-merge strategy and inverted residual structure for resource-constrained devices. *Multimedia Tools and Applications*, 1-22.

Stanley, B. F., Kumar, R. J. R., Shahila, D. F. D., Yamini, K. A. P., Famila, S., & Arun, M. R. (2025, August). MobileNetV3 Architecture Enhanced with Multi-Head Self-Attention for Precise Tomato Maturity Detection. In *2025 8th International Conference on Circuit, Power & Computing Technologies (ICCPCT)* (pp. 785-790). IEEE.

Oloke, R. O., Akinleye, A. F., &Ogunsakin, O. O. EXPLORING CONVOLUTIONAL NEURAL NETWORKS FOR AUTOMATED DRIVER DROWSINESS DETECTION.

Costa, W., Macêdo, D., Zanchettin, C., Talavera, E., Figueiredo, L. S., Teichrieb, V., & Teixeira, J. M. (2022). A fast multiple cue fusing approach for human emotion recognition. *Available at SSRN 5205383*.

Al-funjan, A. W., Al Abboodi, H. M., Hamza, N. A., Abedi, W. M. S., & Abdullah, A. H. (2024). A Lightweight Deep Learning-Based Ocular Disease Prediction Model Using Squeeze-and-Excitation Network Architecture with MobileNet Feature Extraction. *International Journal of Intelligent Engineering & Systems*, *17*(4).

Nethravathi, P. S. (2025, April). Early diagnosis of lung diseases with deep learning using EfficientNetV2-S architecture. In *2025 International Conference on Inventive Computation Technologies (ICICT)* (pp. 436-441). IEEE.

Alqassab, A. I. M., Luque-Nieto, M. Á., & Mohammed, M. A. (2026). Identification of multiple ocular diseases using a hybrid quantum convolutional neural network with fundus images. *Scientific Reports*.