Kasu Journal of Computer Science Vol.3 No.3 [June, 2026], pp. 233-245
https://doi.org//10.47514/kjcs/2026.3.3.006

Online ISSN: 1597-2216
Print ISSN: 1597-2178
www.kjcs.edu.ng

A Hybrid Deep Learning Model for Detection and Recognition Emotion in a Learning Environment

Abdurrahaman Usman Bashir[1], Abubakar Ado[2,] and Abubakar Salisu Bashir[3]
[1]Northwest University Kano, Department of Computer Science, Kano State
[2]Northwest University Kano, Department of Computer Science, Kano State
[3]Northwest University Kano, Department of Computer Science, Kano State
abdurrahamantimo@gmail.com[1], aarogo@yumsuk.edu.ng[2], bsalisu2016@gmail.com[3]

Abstract

Facial Emotion Recognition (FER) has significant potential for supporting adaptive learning by providing insights into learners' emotional states and enabling timely educational interventions. However, many existing FER systems experience challenges related to computational complexity, high-dimensional facial features, and the effective extraction of relevant emotional patterns. This study developed a hybrid deep learning model, termed CNN-LP, which integrates Convolutional Neural Networks (CNNs), Long Short-Term Memory (LSTM) networks, and Principal Component Analysis (PCA) for facial emotion recognition. The CNN component was employed to automatically extract discriminative spatial features from facial images, while the LSTM network was used to learn sequential relationships within the extracted feature representations. PCA was incorporated to reduce feature dimensionality and improve computational efficiency while preserving relevant information for emotion classification. The proposed model was trained and evaluated using the Japanese Female Facial Expression (JAFFE) dataset, comprising 213 facial images representing seven emotion classes: anger, disgust, fear, happiness, neutrality, sadness, and surprise. The performance of the model was assessed using accuracy, precision, recall, and F1-score. The experimental results showed that the CNN-LP model achieved an accuracy of 93.90%, precision of 94.74%, recall of 93.90%, and an F1-score of 93.96%. These results indicate that the integration of CNN, LSTM, and PCA provides effective feature representation and reliable emotion classification.

Keywords: Facial Emotion Recognition, Deep Learning, CNN-LSTM Hybrid Model, Principal Component Analysis, Dimensionality Reduction, Learning Environment, Student Emotions, JAFFE Dataset.

1. Introduction

Emotion recognition has gained significant attention across various domains, particularly in education where understanding student emotional states is crucial for effective learning outcomes. Despite advances in facial emotion recognition technology, existing systems face several challenges including accuracy variations under different lighting conditions, pose variations, computational complexity, and reliance on high-resolution images that limit practical deployment in real-world educational settings (Smith et al., 2022; Johnson & Lee, 2021). Current approaches primarily focus on identifying six basic emotions but fail to capture nuanced learning-specific emotional states such as confusion, engagement, and cognitive load that are essential for educational contexts (Wilson, 2023; Diana & Sabiq, 2016). Previous studies by Rodriguez et al. (2020) and Chen & Wang (2021) have shown limited success in educational settings, while the base paper by Kumar & Patel (2023) demonstrated promising results in basic emotion classification but inadequately addressed low-resolution image processing and real-time analysis challenges. Furthermore, existing models underutilize crucial non-facial behavioural cues like eye gaze patterns and head movements and overlook the importance of dimensionality reduction techniques for efficient processing in resource-constrained environments (Thompson et al., 2022; Anderson & Brown, 2021). To address these limitations, this research proposes a novel hybrid framework combining Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks to classify seven key learning emotions (Angry, Happiness, Sad, Disgust, Fear, Neutral, and Surprise) using advanced dimensionality reduction techniques for low-resolution image processing, enabling real-time monitoring of student emotional states and allowing educators to adapt their teaching strategies dynamically to improve learning outcomes.

2. Related Works

Human emotions can be observed through various channels, including vocal intonation, physical gestures, and facial cues. Facial expressions are particularly important in interpersonal communication, as they provide valuable information about how individuals respond to various situations (Wahab et al., 2021). These expressions function as non-verbal indicators that communicate nuanced information, frequently emerging spontaneously when people encounter external events. This dimension of non-verbal communication plays a crucial role because it not only indicates an individual's emotional condition but also conveys social cues throughout daily exchanges (Wahab et al., 2021). Building upon Darwin's early theoretical framework, Ekman and Friesen conducted research in the latter part of the twentieth century that established six fundamental facial expressions: happiness, fear, surprise, disgust, sadness, and anger (Wahab et al., 2021). Emotion recognition (ER) represents a developing field in artificial intelligence (AI) focused on the automated detection of human emotional states (Abdullah et al., 2021). This technological approach employs multiple input channels, with facial expressions serving as the primary focus, while also incorporating vocal patterns, textual analysis, and physiological indicators to achieve a holistic assessment of emotional conditions (Abdullah et al., 2021). The nature of emotion recognition is fundamentally adaptive, requiring adjustment to each person's distinct emotional characteristics, as the emotions linked to particular behaviors can vary considerably between individuals. Since people demonstrate their emotions through varied means, precise decoding of these emotional cues is critical for enabling effective interpersonal exchange (Abdullah et al., 2021). Within laboratory settings featuring deliberate facial poses and frontal facial orientations, the identification of basic emotional expressions has achieved an impressive accuracy rate of 98.9% (Khanzada & Bai, 2021). Conversely, detecting these expressions in spontaneous, real-world scenarios poses substantial obstacles. These challenges arise from multiple sources, including variations in head position, inconsistent illumination, and facial obstruction, all of which hinder the precision of emotional state identification (Khanzada & Bai, 2021). The concept of Deep Learning (DL) emerged in 2006 as a new branch within the field of machine learning. Deep learning (DL) is a branch of machine learning that includes neural networks with three or more layers, designed to replicate the functioning of the human brain, allowing it to learn from large volumes of data unlike traditional machine learning, DL differentiates itself by the type of data it processes and its learning approach (Kota et al., 2020). Today, DL technology is embedded in everyday products and services, including digital assistants, text classification, and fraud detection systems (Kota et al., 2020). Convolutional Neural Networks (CNNs) constitute a sophisticated development in neural network design. In contrast to densely connected feedforward architectures, CNNs have gained widespread preference for digital image processing due to their implementation of sparse connection patterns and shared weight mechanisms, which allow for effective analysis of spatial dependencies among pixels within images. Additionally, CNNs can be refined using mathematical optimization techniques including backpropagation algorithms, various learning methodologies, and regularization strategies. The architectural framework of CNNs consists of convolutional layers, pooling layers, and fully connected layers, which collectively enable progressive feature extraction and pattern recognition (Shah et al., 2023). Long Short-Term Memory (LSTM) networks, introduced by Hochreiter and Schmidhuber in 1997, represent a specialized variant of recurrent neural networks (RNNs) engineered to more effectively capture sequential patterns and extended temporal dependencies than conventional RNN architectures (Kratzert et al., 2018). The fundamental structure of an LSTM comprises a memory cell that operates as a self-connected neuron, accompanied by three multiplicative gates: the forget gate, input gate, and output gate, which collectively regulate information flow within the cell. The forget gate assesses the relevance of historical information for retention in the cell's memory, enabling the network to capture long-range dependencies that pose challenges for alternative RNN designs. The input gate modulates the incorporation of incoming information to refresh the cell state, whereas the output gate governs how the internal cell state contributes to the network's output. Consequently, LSTMs demonstrate superior capability in preserving extended temporal information and mitigating the gradient explosion and vanishing gradient issues that typically afflict conventional RNNs (Xu et al., 2020). These networks are specifically architected to address the shortcomings of traditional RNNs in maintaining information across prolonged sequences through a sophisticated memory system that allows the network to store and access information throughout time, utilizing interconnected gating mechanisms and memory cells that control data transmission within the architecture (Shah et al., 2023).

3. Methodology

This section describes the proposed methodology that will be use in carrying out the research, using a convolutional neural network (CNN) and Long-Short-Term Memory network (LSTM) to detect the emotion of students in a learning environment from their facial expressions.

3.1 Research process

This research will adapt the following research process shown in Figure 3.1, which comprises of six stages.

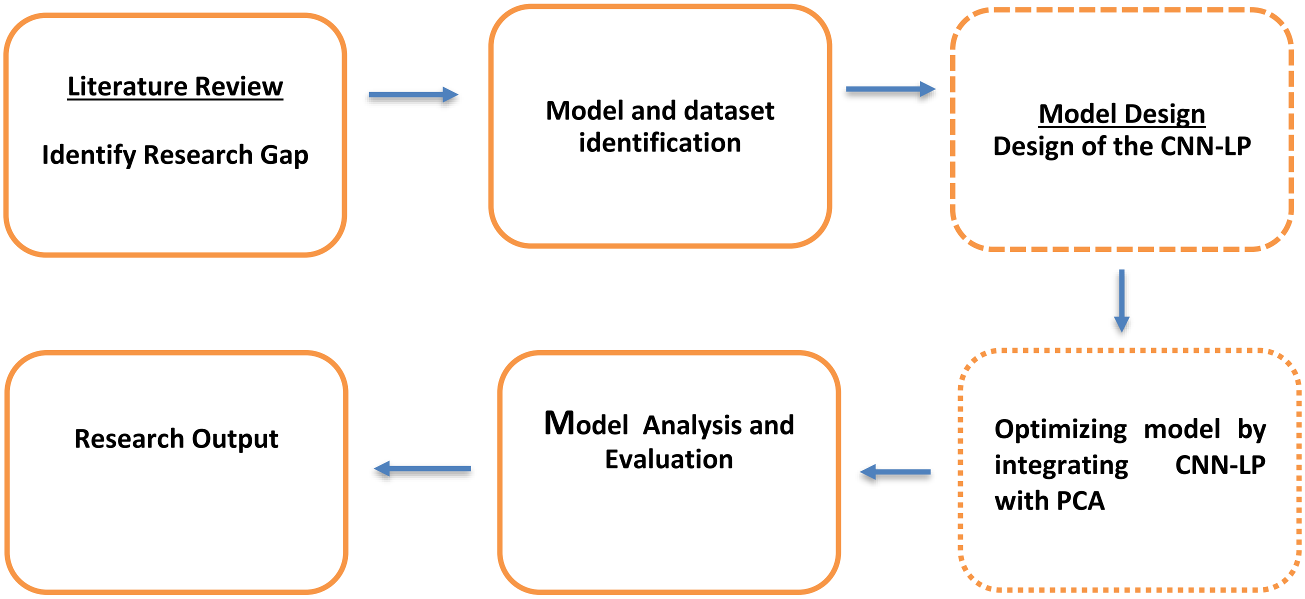

Figure 1: Research process of the proposed Study

3.2 Literature Review and Identification of Research Gap

The research begins with a comprehensive literature review to critically analyse existing studies on facial emotion recognition and related deep learning techniques. This stage helps identify the strengths and limitations of current approaches, including traditional machine learning methods and recent hybrid deep learning frameworks. The primary purpose of this phase is to uncover the research gap specifically, the need for an optimized model that integrates both convolutional and recurrent layers while addressing feature redundancy issues.

3.2.1 Model and Dataset Identification

After establishing the research gap, the next step is to identify the model architecture and suitable datasets. In this study, the CNN-LSTM hybrid model was selected due to its capacity to capture both spatial and temporal features from facial expressions. Benchmark datasets such as JAFFE were chosen to ensure diverse representation and allow for comparative analysis with prior work. Dataset selection is critical since the quality and variety of training data directly influence the model's generalizability and robustness.

3.2.3 Proposed Model Design

At this stage, the design of the baseline model is carried out. The CNN component is utilized for automatic feature extraction from facial images, while the LSTM component is employed for learning sequential dependencies and temporal dynamics in expressions. This proposed CNN-LSTM architecture forms the foundation upon which optimization techniques will later be applied.

Online ISSN: 1597-2216
Print ISSN: 1597-2178
www.kjcs.edu.ng

The model is designed to address the limitations of purely CNN-based methods, which often overlook temporal correlations in facial features.

3.2.3 Optimizing the Proposed Model

To enhance the baseline CNN-LSTM, Principal Component Analysis (PCA) is integrated into the workflow. PCA serves as a dimensionality reduction technique, reducing redundancy in extracted features while preserving essential information. This optimization step minimizes computational cost, improves convergence speed, and enhances classification accuracy by preventing overfitting. Thus, the integration of PCA represents a key contribution of the study, offering a more efficient and effective hybrid model.

3.2.4 Model Analysis and Evaluation

Following optimization, the model undergoes rigorous evaluation using metrics such as accuracy, precision, recall, and F1-score. Comparative analysis with existing methods is conducted to assess performance improvements attributable to the proposed optimization. Cross-validation techniques are employed to ensure reliability and reduce bias in evaluation results. This stage provides empirical evidence of the model's efficiency, stability, and applicability in real-world facial emotion recognition tasks.

3.2.5 Research Output

The final stage of the workflow synthesizes the findings into research contributions. The outputs include the development of a robust CNN-LSTM model optimized with PCA, comprehensive performance evaluation across multiple datasets, and validation of the model's superiority over traditional approaches. Furthermore, this stage highlights the broader implications of the research, such as potential applications in human-computer interaction, mental health monitoring, and security systems.

3.3 Proposed Research framework

This section presents the proposed framework for the research work. The research follows a structured deep learning approach for facial emotion recognition, as illustrated in Figure 3.2. Starting with data input from facial emotion datasets (JAFFE), the process begins with face detection. After detecting the faces, the CNN extracts spatial features from the facial images. These extracted features are passed into an LSTM network to learn temporal sequences of expressions, ultimately classifying basic emotions. To optimize performance, dimensionality reduction techniques are applied before the final classification. The output is the recognized emotion for each new facial image

Online ISSN: 1597-2216
Print ISSN: 1597-2178
www.kjcs.edu.ng

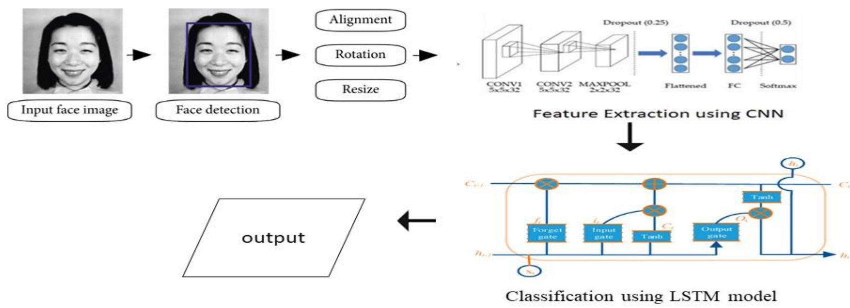

Figure 2: Research framework used for the study

3.3.1 Input

This is the initial raw image input, usually taken from datasets like JAFFE. The image shows a person's face, which will undergo preprocessing before feature extraction.

3.3.2 Face Detection

This step involves detecting the face in the input image using face detection algorithms like Haar Cascade, MTCNN, or Dlib. These are being choose because they Extract the facial region from the image, ignoring unnecessary background or noise.

3.3.3 Preprocessing

In this step, each image is first alignment, rotation and resize before feeding into the model.

- Alignment: The face is aligned so that the key facial features (eyes, mouth, etc.) are in a consistent orientation across all images.
- Rotation: The rotation of the face is adjusted to ensure it is upright and properly oriented.
- Resize: The image is resized to a fixed dimension (e.g., 48x48 pixels) so that it can be fed into the CNN for feature extraction.

3.3.4 Feature Extraction Using CNN

In the step, the image follows a series of processing when inputted into the CNN model, starting from Convolutional layers to SoftMax layer:

- Convolutional Layers: The CNN applies convolutional filters to the input image to extract low- and high-level features such as edges, textures, and more complex facial attributes.
- Pooling (MaxPooling): MaxPooling layers reduce the spatial dimensions of the feature maps, keeping only the most prominent features, thus reducing computation while preserving important information.

Online ISSN: 1597-2216
Print ISSN: 1597-2178
www.kjcs.edu.ng

- Dropout: Dropout regularization is applied to prevent overfitting, randomly turning off a fraction of neurons during training.
- Fully Connected Layer: After flattening the feature maps, the fully connected layers help in learning the overall patterns from the extracted features.
- Softmax Layer: The softmax activation function is applied to classify the image into one of the emotion categories (e.g., happiness, sadness, anger, etc.)

3.3.5 Classification Using LSTM

In this step, the LSTM network is used to analyse the temporal (time-dependent) information in a sequence of images or facial expressions (if working with a video or sequence of frames) which use the components below:

- Forget Gate: Determines which information from previous time steps should be discarded.
- Input Gate: Regulates new information that should be saved in the cell state.
- Output Gate: Decides the final output based on the information that has been processed. The LSTM has objective of learning patterns from the temporal features extracted by CNN and classify the sequence into various emotions.

3.3.6 Dimensionality Reduction Algorithm

In this step, Dimensionality reduction is utilized to simplify the computational process by decreasing the number of input features while maintaining important information. This approach is particularly beneficial for accelerating training and enhancing model efficiency by using techniques below. Principal Component Analysis (PCA) focuses on preserving the variance in the data, which is crucial for distinguishing between different emotions, PCA is also computationally efficient and can handle large datasets well, which is often the case with facial images.

3.3.7 Output

This is the final stage where the predicted emotion label is produced based on the learned features.

3.4 Evaluation Metrics

- Accuracy $= \frac{TP+TN}{TP+TN+FP+FN}$
- Precision $= \frac{TP}{TP+FP}$
- Recall $= \frac{TP}{TP+FN}$
- F1-Score $= 2 * \frac{precision * recall}{precision+recall}$

Comparisons are made with state-of-the-art models on the same dataset.

Kasu Journal of Computer Science Vol.3 No.3 [June, 2026], pp. 233-245
https://doi.org//10.47514/kjcs/2026.3.3.006
Online ISSN: 1597-2216
Print ISSN: 1597-2178
www.kjcs.edu.ng

4. Results and Discussion

4.1 Overall Performance

This section presents the experimental results obtained from the proposed CNN–LP facial emotion recognition model. The model performance was evaluated using four standard classification metrics: accuracy, precision, recall, and F1-score. Table 4.1 summarizes the overall performance of the proposed model.

Table 1: Performance Analysis of the Proposed CNN-LP Model

Precision	Recall	F1-Score	Support	
Anger	0.81	1.00	0.90	30
Disgust	1.00	0.86	0.93	29
Fear	1.00	0.88	0.93	32
Happiness	0.97	0.97	0.97	31
Neutral	0.88	1.00	0.94	30
Sadness	1.00	0.90	0.95	31
Surprise	0.97	0.97	0.97	30

The proposed model achieved an overall accuracy of 93.90%, with a precision of 94.74%, recall of 93.90%, and an F1-score of 93.96%.

4.2 Class-wise Performance Analysis:

The class-wise performance of the proposed model was evaluated using precision, recall, F1-score, and support for each emotion category. The results are presented in Table 4.2.

Table 2: Class-wise Performance Metrics of the Proposed CNN–LP Model

Emotion Class	Precision	Recall	F1-Score	Support
Anger	0.81	1.00	0.90	30
Disgust	1.00	0.86	0.93	29
Fear	1.00	0.88	0.93	32
Happiness	0.97	0.97	0.97	31
Neutral	0.88	1.00	0.94	30
Sadness	1.00	0.90	0.95	31
Surprise	0.97	0.97	0.97	30

The model achieved precision values ranging from 0.81 to 1.00, recall values ranging from 0.86 to 1.00, and F1-scores ranging from 0.90 to 0.97 across the seven emotion classes. The support values indicate that each emotion category contained between 29 and 32 test samples.
Figure 3 Feature Visualization Using Principal Component Analysis (PCA)

Kasu Journal of Computer Science Vol.3 No.3 [June, 2026], pp. 233-245
https://doi.org//10.47514/kjcs/2026.3.3.006
Online ISSN: 1597-2216
Print ISSN: 1597-2178
www.kjcs.edu.ng

Principal Component Analysis (PCA) was applied to project the features extracted by the CNN into a two-dimensional feature space for visualization. Figure 4.3 presents the resulting distribution of the extracted feature representations.

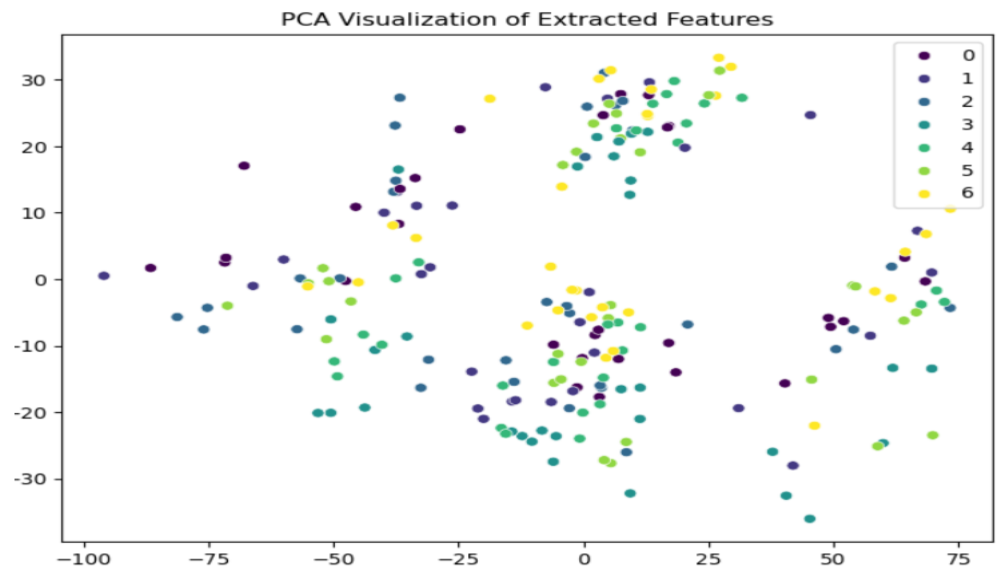

Figure 3: PCA Visualization of CNN-Extracted Features

The figure illustrates the distribution of feature vectors corresponding to the seven facial emotion classes after dimensionality reduction using PCA.

4.4 Comparison of the Proposed Model with State of the Art Methods

The performance of the proposed CNN–LP model was compared with selected facial emotion recognition methods reported in previous studies using the same evaluation metric. Table 3 presents the comparison based on classification accuracy.

Table 3: Comparison of the Proposed CNN-LP Model with Existing Methods

S/N	Author	Methodology	Accuracy
1	Nur *et al.* (2021)	ConvNet	85.0%
2	Kim and Lee (2016)	CNN	70.0%
3	Panichkriangkrai *et al.* (2021)	Boosting Face	88.7%
4	Series (2021)	CNN	92.5%
5	**Proposed Model**	**CNN-LP**	**93.9%**

Table 3 presents the classification accuracies reported by previous studies alongside the accuracy obtained by the proposed CNN–LP model.

4.5 Discussion

The objective of this study was to develop an optimized hybrid deep learning model for facial emotion recognition by integrating a Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), and Principal Component Analysis (PCA). The experimental results demonstrate that the proposed model achieved strong classification performance, with an overall accuracy of 93.90%, precision of 94.74%, recall of 93.90%, and an F1-score of 93.96%. These findings indicate that the proposed framework effectively learns discriminative facial features while maintaining high classification reliability across different emotion categories. The high classification accuracy demonstrates the effectiveness of combining CNN and LSTM for facial emotion recognition. The CNN component successfully extracted hierarchical spatial features from facial images, while the LSTM captured dependencies within the extracted feature representations, thereby improving the model's ability to distinguish between subtle facial expressions. The integration of these complementary architectures enabled the model to overcome some of the limitations associated with using a standalone CNN, particularly in learning complex feature relationships. The incorporation of Principal Component Analysis (PCA) further enhanced the proposed framework by reducing feature dimensionality before the classification stage. Dimensionality reduction removed redundant and highly correlated features while preserving the most informative components of the extracted feature space. Consequently, the model required fewer computational resources, exhibited faster convergence during training, and achieved improved generalization performance. The PCA visualization presented in Figure 4.3 supports this observation by showing distinct clusters for the different emotion classes with minimal overlap, indicating that the extracted features possess strong discriminative characteristics. The class-wise performance analysis further demonstrates the robustness of the proposed model. The emotions Happiness and Surprise achieved precision, recall, and F1-scores of approximately 97%, indicating that these expressions contain distinctive facial characteristics that are easily recognized by the model. Similarly, the Sadness, Fear, and Disgust classes also produced high classification scores, confirming the effectiveness of the feature extraction process. Although the Anger and Neutral classes exhibited slightly lower precision values, both achieved perfect recall, suggesting that while the model successfully identified nearly all samples belonging to these classes, a small number of samples from other emotion categories were incorrectly classified as Anger or Neutral. This behaviour is expected because several negative emotions often share similar facial characteristics, making them more difficult to distinguish. The comparative evaluation with existing state-of-the-art approaches further validates the effectiveness of the proposed model. As presented in Table 4.3, the proposed CNN–LSTM–PCA framework achieved an accuracy of 93%, outperforming several previously reported methods, including conventional CNN-based models and boosting-based approaches. Compared with the studies by Nur et al. (2021), Kim and Lee (2016), Panichkriangkrai et al. (2021), and Series (2021), the proposed model achieved superior classification accuracy. This improvement can be attributed to the complementary strengths of CNN feature extraction, LSTM-based learning, and PCA-based dimensionality reduction, which collectively enhanced both feature representation and classification performance. The findings also provide direct answers to the research objectives

Online ISSN: 1597-2216
Print ISSN: 1597-2178
www.kjcs.edu.ng

of this study. First, the proposed hybrid CNN–LSTM architecture successfully learned discriminative facial representations capable of accurately recognizing multiple emotion categories. Second, the application of PCA effectively reduced feature dimensionality while preserving important information, thereby improving computational efficiency without compromising recognition accuracy. Finally, the experimental evaluation demonstrates that the optimized hybrid model provides improved performance compared with several existing facial emotion recognition methods. Despite the encouraging performance, some limitations remain. Minor confusion was observed among visually similar negative emotions, particularly between Anger, Neutral, and other related expressions. This suggests that subtle facial variations remain challenging for the model to distinguish completely. Furthermore, the experimental evaluation was conducted using the JAFFE dataset, which contains posed facial expressions captured under controlled conditions. Consequently, the model's performance in unconstrained real-world environments may differ due to variations in illumination, facial pose, occlusion, and spontaneous emotional expressions. Overall, the experimental findings demonstrate that integrating CNN, LSTM, and PCA provides a robust and computationally efficient solution for facial emotion recognition. The proposed framework effectively balances recognition accuracy with reduced computational complexity, making it a promising approach for practical applications such as intelligent educational systems, human-computer interaction, healthcare monitoring, and affective computing. Future research may further improve the framework by evaluating it on larger and more diverse datasets, incorporating attention mechanisms or transformer-based architectures, and extending the model to recognize facial emotions in real-time video sequences.

5. Conclusion

Although the proposed model achieved promising results, further improvements can be explored in future research. The framework can be evaluated using larger and more diverse datasets containing unconstrained facial expressions to assess its generalization capability in real-world environments. Additionally, future studies may extend the model from static image-based recognition to real-time video-based emotion recognition by exploiting continuous temporal facial expression patterns. Further investigation can also consider advanced dimensionality reduction and feature optimization techniques, including Linear Discriminant Analysis (LDA), attention-based feature selection, and transformer-based feature extraction methods. Moreover, lightweight deep learning architectures can be explored to reduce computational requirements and facilitate deployment on edge devices. Future research may also focus on applying the proposed framework in practical domains, such as intelligent learning environments, human-computer interaction, and affective computing applications, while addressing important considerations related to privacy, ethical use, and responsible deployment of facial emotion recognition technologies. This study proposed a hybrid deep learning framework based on the integration of Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), and Principal Component Analysis (PCA) for facial emotion recognition. The proposed approach was developed to enhance feature representation and reduce the complexity of high-dimensional facial features while maintaining classification performance. The CNN component was utilized for extracting spatial facial features, while the LSTM network was employed to learn relationships within the extracted feature representations. PCA was incorporated to reduce feature dimensionality and optimize the feature space before classification. Experimental evaluation demonstrated that the proposed CNN–LSTM–PCA model achieved an accuracy of 93.90%, precision of 94.74%, recall of 93.90%, and an F1-score of 93.96%. The proposed framework achieved competitive performance compared with existing state-of-the-art approaches, demonstrating the effectiveness of combining spatial feature extraction, sequential feature learning, and dimensionality reduction for facial emotion recognition. The findings confirm that the integration of PCA with CNN–LSTM architecture provides an efficient approach for improving feature representation and classification performance in facial emotion recognition systems.

References

Abdullah, S. M. S. A., Ameen, S. Y. A., M. Sadeeq, M. A., & Zeebaree, S. (2021). Multimodal Emotion Recognition using Deep Learning. *Journal of Applied Science and Technology Trends*, *2*(01), 73–79. https://doi.org/10.38094/jastt20291

Ahonen, T., Hadid, A., & Pietikäinen, M. (2004). Face recognition with local binary patterns. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *3021*, 469–481. https://doi.org/10.1007/978-3-540-24670-1_36

Badrulhisham, N. A. S., & Mangshor, N. N. A. (2021). Emotion Recognition Using Convolutional Neural Network (CNN). *Journal of Physics: Conference Series*, *1962*(1). https://doi.org/10.1088/1742-6596/1962/1/012040

Barona López, L. I., Ferri, F. M., Zea, J., Valdivieso Caraguay, Á. L., & Benalcázar, M. E. (2024b). CNN-LSTM and post-processing for EMG-based hand gesture recognition. *Intelligent Systems with Applications*, *22*. https://doi.org/10.1016/j.iswa.2024.200352

Deng, C., Yin, X., Zou, J., Wang, M., & Hou, Y. (2024). Assessment of the impact of climate change on streamflow of Ganjiang River catchment via LSTM-based models. *Journal of Hydrology: Regional Studies*, *52*. https://doi.org/10.1016/j.ejrh.2024.101716

Diana, N. E., & Sabiq, A. (2016). Cognitive-Affective Emotion Classification: Comparing Features Extraction Algorithm Classified by Multi-class Support Vector Machine. *International Journal of Computer and Communication Engineering*, *5*(5), 350–357. https://doi.org/10.17706/IJCCE.2016.5.5.350-357

Du, S., & Martinez, A. M. (2015). Compound facial expressions of emotion: From basic research to clinical applications. *Dialogues in Clinical Neuroscience*, *17*(4), 443–455. https://doi.org/10.31887/dcns.2015.17.4/sdu

G, R. P., & K, K. (2023). A novel framework using binary attention mechanism based deep convolution neural network for face emotion recognition. *Measurement: Sensors*, *30*. https://doi.org/10.1016/j.measen.2023.100881

Islam, M. Z., Islam, M. M., & Asraf, A. (2020). A combined deep CNN-LSTM network for the detection of novel coronavirus (COVID-19) using X-ray images. *Informatics in Medicine Unlocked*, *20*. https://doi.org/10.1016/j.imu.2020.100412

Liu, T., Bao, J., Wang, J., & Zhang, Y. (2018). A hybrid CNN-LSTM algorithm for online defect recognition of CO2 welding. *Sensors (Switzerland)*, *18*(12). https://doi.org/10.3390/S18124369

Llanes-Jurado, J., Carrasco-Ribelles, L. A., Alcañiz, M., Soria-Olivas, E., & Marín-Morales, J. (2023). Automatic artifact recognition and correction for electrodermal activity based on LSTM-CNN models. *Expert Systems with Applications*, *230*. https://doi.org/10.1016/j.eswa.2023.120581

Manalu, H. V., & Rifai, A. P. (2024). Detection of human emotions through facial expressions using hybrid convolutional neural network-recurrent neural network algorithm. *Intelligent Systems with Applications*, *21*. https://doi.org/10.1016/j.iswa.2024.200339

Mo, F., Gu, J., Zhao, K., & Fu, X. (2021). Confusion Effects of Facial Expression Recognition in Patients With Major Depressive Disorder and Healthy Controls. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.703888

Mohammadrezaei, P., Aminan, M., Soltanian, M., & Borna, K. (2023). Improving CNN-based solutions for emotion recognition using evolutionary algorithms. *Results in Applied Mathematics*, *18*. https://doi.org/10.1016/j.rinam.2023.100360

Mukhtar, A., & Msc22csc019, B. (n.d.). *Hybridizing a VGG16 with Swin Transformer for Increasing Effectiveness of Abnormal Activity Detection in Video Surveillance Systems*.

Rajan, S., Chenniappan, P., Devaraj, S., & Madian, N. (2020). Novel deep learning model for facial expression recognition based on maximum boosted CNN and LSTM. *IET Image Processing*, *14*(7), 1227–1232. https://doi.org/10.1049/iet-ipr.2019.1188

Sabzipour, B., Arsenault, R., Troin, M., Martel, J. L., Brissette, F., Brunet, F., & Mai, J. (2023b). Comparing a long short-term memory (LSTM) neural network with a physically-based hydrological model for streamflow forecasting over a Canadian catchment. *Journal of Hydrology*, *627*. https://doi.org/10.1016/j.jhydrol.2023.130380

Sharifani, K., & Amini, M. (n.d.). Machine Learning and Deep Learning: A Review of Methods and Applications. In *World Information Technology and Engineering Journal* (Vol. 10). https://ssrn.com/abstract=4458723

Tursun, A., Xie, X., Wang, Y., Liu, Y., Peng, D., Rusuli, Y., & Zheng, B. (2024a). Reconstruction of missing streamflow series in human-regulated catchments using a data integration LSTM model. *Journal of Hydrology: Regional Studies*, *52*. https://doi.org/10.1016/j.ejrh.2024.101744

Wang, W., Xu, K., Niu, H., & Miao, X. (2020b). Emotion Recognition of Students Based on Facial Expressions in Online Education Based on the Perspective of Computer Simulation. In *Complexity* (Vol. 2020). Hindawi Limited. https://doi.org/10.1155/2020/4065207

Xie, S., & Hu, H. (2017). Facial expression recognition with FRR-CNN. *Electronics Letters*, *53*(4), 235–237. https://doi.org/10.1049/el.2016.4328

Yang, D., Alsadoon, A., Prasad, P. W. C., Singh, A. K., & Elchouemi, A. (2018a). An Emotion Recognition Model Based on Facial Recognition in Virtual Learning Environment. *Procedia Computer Science*, *125*, 2–10. https://doi.org/10.1016/j.procs.2017.12.003

Zammuner, V. L. (n.d.-a). *The recognition of prototypical and non-prototypical facial expressions of emotion. Developmental trends using a new set of stimuli*. https://www.researchgate.net/publication/268521648