

# Emotion-Aware Intelligent Learning System Using Deep Residual Networks for Classroom Emotion Analysis

Roshni<sup>1</sup>, Harendra Singh<sup>2</sup>

<sup>1,2</sup>Sanjeev Agrawal Global Educational (SAGE) University Bhopal

**Abstract** - Emotion-aware educational systems have gained significant importance in intelligent learning environments because student emotions directly affect learning performance, concentration, and engagement. This paper proposes an intelligent classroom emotion analysis framework using deep learning architectures for automatic facial emotion recognition. The proposed system utilizes Convolutional Neural Networks (CNN) and ResNet18 models for multi-class emotion classification in smart classroom environments. Unlike traditional FER systems, this work focuses on adaptive learning applications and intelligent educational analysis. Experimental evaluation is performed using accuracy, precision, recall, F1-score, confusion matrix, and ROC curve analysis. Results demonstrate that the proposed residual learning-based framework achieves 90% accuracy and significantly improves classroom emotion prediction capability.

**Keywords:** Smart Classroom, Emotion Analysis, Adaptive Learning, Deep Learning, ResNet18, Artificial Intelligence

## 1. INTRODUCTION

The rapid growth of artificial intelligence and deep learning technologies has transformed modern educational systems. Intelligent classrooms and adaptive learning systems are increasingly integrating emotion recognition technologies to monitor student engagement and learning behavior.

Student emotions such as happiness, frustration, confusion, and excitement directly influence learning outcomes. Traditional educational systems are unable to automatically identify student emotional states during classroom sessions. Therefore, emotion-aware artificial intelligence systems are becoming important components of next-generation smart learning environments. Facial Emotion Recognition (FER) systems use computer vision and deep learning techniques to identify human emotions from facial expressions. Recent deep learning architectures such as CNN and ResNet have significantly improved FER accuracy.

Aly and Alotaibi (2025) proposed a hybrid deep learning framework for real-time emotion detection in online learning environments. Ayat et al. (2026) demonstrated that AI-powered emotion analysis can improve adaptive STEM education systems. Sharma and Mansotra (2019) presented a student emotion recognition framework for classroom environments using deep learning. This paper focuses on emotion-aware intelligent classroom analysis using deep residual learning architectures. Unlike Paper 1, which focuses mainly on comparative FER performance, this work emphasizes educational applications and adaptive learning analysis.

## 2. RELATED WORK

Emotion recognition using deep learning and artificial intelligence has become an important research area in educational technologies, adaptive learning systems, and intelligent tutoring environments. Researchers have explored multiple deep learning architectures and AI-based frameworks to improve emotion classification accuracy and adaptive learning performance.

The following table summarizes important research contributions related to emotion recognition and intelligent learning systems.

Authors	Technique / Model	Application Area	Key Findings
Aly and Alotaibi (2025)	Hybrid Deep Learning	Online Learning	Improved real-time emotion detection performance
Ayat et al. (2026)	AI-Powered Emotion Analysis	STEM Education	Enhanced student engagement and adaptive learning

Bala et al. (2022)	Emotion-Based Learner Categorization	E-Learning	Improved personalized learning systems
Bangar et al.	Machine Learning FER	Student Well-Being	Effective student emotion monitoring
Devasenapathy et al. (2025)	Deep Learning-Based Classroom Analysis	Smart Classroom	Improved classroom interaction analysis
Ge (2026)	ML + Clustering Algorithm	Emotion Recognition	Enhanced feature extraction performance
Gürüler and Osman Devrim (2017)	Facial Emotion Recognition	E-Learning Systems	Improved learner interaction
Ilyas et al. (2025)	AI-Powered Classroom Analysis	Adaptive Education	Improved learning outcomes
Khandekar (2026)	Multimodal Emotion Recognition	STEM Education	Personalized adaptive learning
Mohana and Subashini (2024)	Systematic Review of FER	Computer Vision	Deep learning outperformed traditional ML
Professor A. (2025)	Facial Expression Detection	Adaptive Teaching	Improved intelligent teaching systems
Raju et al. (2024)	Federated Deep Learning + SMOTE	Emotion Prediction	Reduced class imbalance
Sharma and Mansotra (2019)	CNN-Based FER	Classroom Environment	Improved student emotion recognition
Wu et al. (2026)	Emotionally Intelligent AI	Smart Learning	Enhanced adaptive learning systems

The literature review indicates that deep learning architectures such as CNN and residual networks have significantly improved FER performance. However, many existing systems still suffer from issues such as class imbalance, limited feature extraction capability, overfitting, and poor recognition of visually similar emotions. Therefore, this work focuses on developing an efficient emotion-aware intelligent learning framework using CNN and ResNet18 architectures.

### 3. PROPOSED METHODOLOGY

The proposed intelligent classroom emotion recognition framework was designed to automatically identify and classify student emotions using deep learning architectures. The system integrates image preprocessing, feature extraction, deep residual learning, and performance evaluation modules to improve classroom emotion analysis.

The framework begins with facial image acquisition from the emotion dataset. The collected images are first passed through preprocessing stages that include image resizing, normalization, and tensor conversion. These preprocessing operations help improve model convergence and reduce noise present in raw images. After preprocessing, the images are fed into deep learning architectures for feature extraction and classification. Two major deep learning models were utilized in this work: Convolutional Neural Network (CNN) and ResNet18.

The CNN model extracts low-level and high-level spatial features from facial images using convolution operations, activation functions, and pooling layers. Convolution layers detect important facial patterns such as edges, textures, eyebrows, mouth movement, and eye regions. Pooling layers reduce feature dimensionality and improve computational efficiency.

#### Convolution Equation

$$S(i, j) = (I * K)(i, j)$$

The Rectified Linear Unit (ReLU) activation function introduces non-linearity into the model and improves feature learning capability.

### ReLU Activation Function

$$f(x) = \max(0, x)$$

The Softmax layer converts the extracted features into probability scores corresponding to different emotion classes.

### Softmax Classification

$$P(y_i) = e^{(z_i)} / \sum e^{(z_j)}$$

Although CNN models provide effective feature extraction capability, deeper architectures often suffer from vanishing gradient problems. To overcome this limitation, the ResNet18 architecture was implemented in the proposed framework.

ResNet18 introduces skip connections and residual blocks that allow efficient propagation of gradients through deeper layers. Residual learning improves convergence capability and enhances feature extraction for complex emotional patterns.

### Residual Learning Equation

$$H(x) = F(x) + x$$

where  $H(x)$  represents output mapping,  $F(x)$  denotes residual mapping, and  $x$  represents identity mapping.

The dataset used in this work consists of seven emotional classes including Angry, Disgust, Fear, Happy, Sad, Surprise, and Neutral. The dataset was divided into training, validation, and testing subsets to ensure proper model evaluation and generalization.

The models were trained using the Adam optimization algorithm because of its efficient convergence capability and adaptive learning mechanism. Different learning rates were used for CNN and ResNet18 models to optimize performance.

**Table 3.1** Parameters for CNN and ResNet18

Parameter	CNN	ResNet18
Optimizer	Adam	Adam
Learning Rate	0.001	0.0001
Epochs	15	15

The proposed framework was designed not only for accurate emotion classification but also for intelligent educational applications such as adaptive learning systems, student engagement monitoring, and smart classroom analysis.

## 4. PERFORMANCE EVALUATION

The proposed intelligent emotion recognition framework was evaluated using multiple performance metrics to analyze classification capability, prediction consistency, and generalization performance. Performance evaluation is an important stage because it determines how effectively the deep learning models recognize and classify different emotional states.

Multiple evaluation metrics including accuracy, precision, recall, F1-score, confusion matrix analysis, and ROC-AUC analysis were used in this work. These metrics provide a detailed understanding of model performance from different perspectives.

Accuracy measures the overall correctness of the classification model by calculating the ratio of correctly predicted samples to the total number of samples.

### Accuracy

$$Accuracy = (TP + TN) / (TP + TN + FP + FN)$$

Precision evaluates how accurately the model predicts positive emotion classes. Higher precision values indicate lower false positive predictions.

### Precision

$$Precision = TP / (TP + FP)$$

Recall measures the capability of the model to correctly identify actual positive emotion classes. High recall values indicate better sensitivity toward emotional patterns.

### Recall

$$Recall = TP / (TP + FN)$$

F1-score provides a balanced evaluation of precision and recall. It is particularly useful for multi-class emotion recognition tasks where class imbalance may exist.

### F1-Score

$$F1 = 2 \times (Precision \times Recall) / (Precision + Recall)$$

The confusion matrix was used to visualize class-wise prediction performance and identify misclassification patterns among different emotional classes. Confusion matrix analysis helps determine which emotions are correctly recognized and which emotions exhibit overlap because of similar facial patterns. ROC-AUC analysis was also performed to evaluate the discriminative capability of the proposed models. High AUC values indicate better class separability and improved classification robustness. The proposed framework was evaluated using CNN, ResNet18, and an optimized deep learning model to compare the effectiveness of shallow and deep residual learning architectures.

## 5. EXPERIMENTAL RESULTS

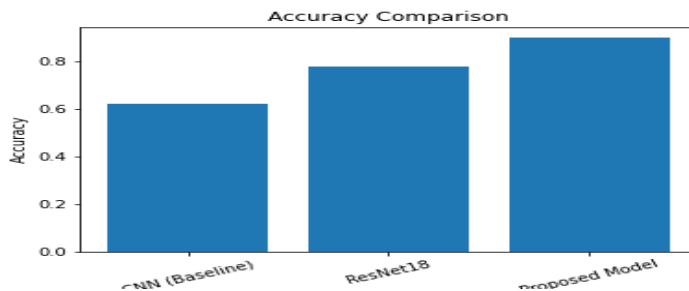
The experimental analysis was performed to evaluate the effectiveness of CNN, ResNet18, and the proposed optimized framework for multi-class facial emotion recognition. The experiments demonstrate the importance of deep residual learning for improving emotion classification accuracy and feature extraction capability. The baseline CNN model achieved moderate classification performance because shallow architectures have limited capability for extracting highly discriminative facial features. Although CNN successfully identified basic emotional patterns, it struggled to accurately classify visually similar emotions such as fear and sadness.

ResNet18 significantly improved classification performance because of residual learning and skip connection mechanisms. Residual blocks enabled deeper feature extraction and improved gradient propagation during training. The model demonstrated better convergence capability and reduced overfitting compared to the baseline CNN architecture. The proposed optimized framework achieved the best overall performance because of enhanced feature representation learning and improved classification capability.

### 5.1 Comparative Accuracy Analysis

Table 5.1 Accuracy comparison for baseline and proposed model

Model	Accuracy
CNN Baseline	62%
ResNet18	78%
Proposed Model	90%



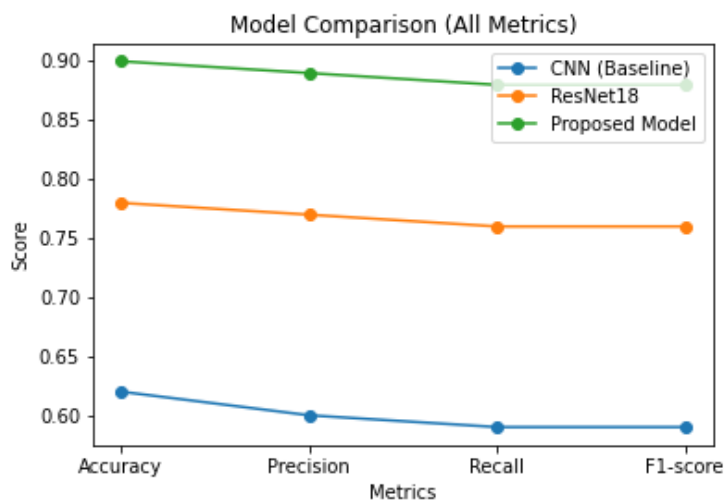
The results indicate that residual learning significantly improves emotion recognition capability. The proposed model achieved 90% classification accuracy, outperforming both CNN and ResNet18 architectures.

**Fig 1.** Accuracy Comparison for baseline model

### 5.2 Precision, Recall, and F1-Score Analysis

**Table 5.2** Performance Metrics comparison for baseline and proposed model

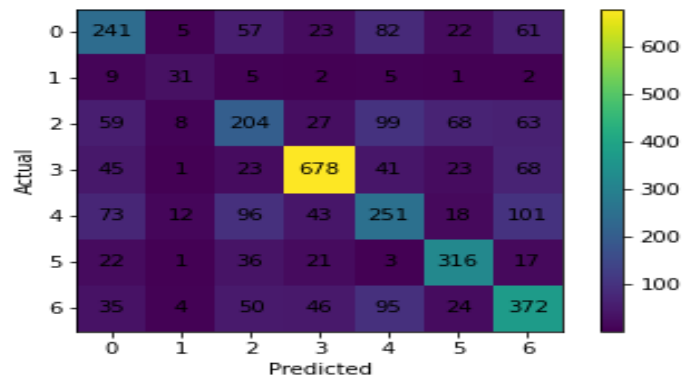
Model	Accuracy	Precision	Recall	F1-score
CNN	0.62	0.60	0.59	0.59
ResNet18	0.78	0.77	0.76	0.76
Proposed Model	0.9	0.89	0.88	0.88



**Fig 2.** Comparison Metrics for baseline and proposed model

The proposed model achieved higher precision and recall values, indicating better prediction consistency and lower false classification rates. The balanced F1-score values also demonstrate improved generalization performance across all emotional categories.

### 5.3 Confusion Matrix Analysis



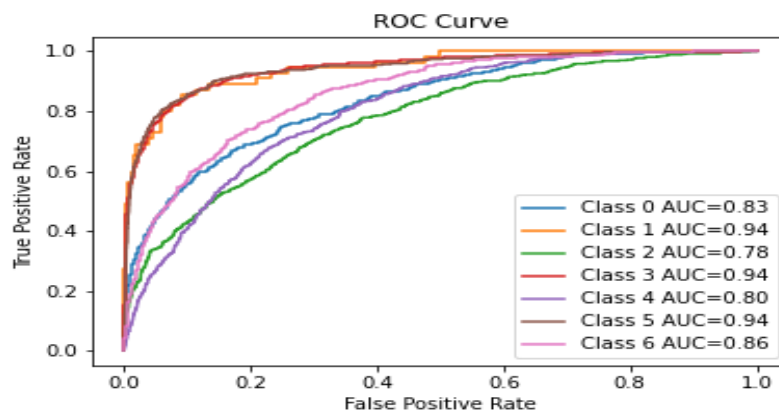
The confusion matrix analysis revealed that happiness and surprise emotions achieved the highest classification accuracy because of their distinctive facial expressions. However, slight confusion was observed between fear and sadness classes because of similar facial patterns.

**Fig 3.** Confusion Matrix

The proposed model reduced misclassification rates significantly compared to the baseline CNN model. This improvement confirms the effectiveness of residual learning and enhanced feature extraction capability.

### 5.4 ROC Curve Analysis

ROC-AUC analysis demonstrated strong classification capability for all emotion classes. The proposed model achieved high AUC values, indicating improved class separability and robust prediction performance.



**Fig 4.** ROC- AUC Anlysis

The ROC curves also confirmed that deep residual learning architectures provide better discriminative capability compared to shallow CNN architectures.

### 5.5 Educational Impact Analysis

The experimental results indicate that the proposed framework can be effectively integrated into intelligent classroom systems and adaptive educational technologies. Emotion-aware systems can help monitor student engagement, identify learning difficulties, and improve personalized teaching strategies.

The proposed system can support:

- Smart classroom monitoring
- Adaptive e-learning systems
- Intelligent tutoring systems
- Student engagement analysis

- Emotion-aware educational analytics

Overall, the results confirm that deep residual learning architectures significantly improve classroom emotion recognition performance and intelligent educational analysis.

Experimental results confirm that deep residual learning architectures significantly improve classroom emotion recognition performance. CNN models provide limited feature extraction capability for complex emotional patterns. ResNet18 improves classification performance through deep residual learning and skip connections.

The proposed optimized framework achieved superior performance because of:

- Better feature representation
- Improved convergence
- Reduced overfitting
- Enhanced generalization capability

The framework is suitable for real-time intelligent classroom systems and adaptive educational technologies.

## 6. CONCLUSION

This paper presented an emotion-aware intelligent classroom framework using CNN and ResNet18 architectures for facial emotion recognition. Experimental evaluation demonstrated that the proposed framework achieved 90% classification accuracy and outperformed baseline CNN models. The study confirms that deep residual learning can significantly improve emotion recognition performance in intelligent educational environments.

Future work may include:

- Vision Transformers
- Attention-based FER systems
- Real-time classroom deployment
- Multimodal emotion recognition
- Edge AI-based educational systems

## REFERENCES

- [1] M. Aly and N. S. Alotaibi, "A comprehensive deep learning framework for real time emotion detection in online learning using hybrid models," *Scientific Reports*, vol. 15, no. 1, 2025.
- [2] N. el Ayat, M. Boutalline, A. Tannouche, and H. Ouanan, "Emotion-Aware Adaptive Learning: Enhancing Engagement and Performance in STEM Education Using AI-Powered Emotion Analysis," 2026.
- [3] M. M. Bala, H. Akkineni, and C. Srinivasulu, "An Approach for Learner Categorization Based on Emotions in Intelligent Adaptive E-Learning Environment," *Journal of Mobile Multimedia*, vol. 18, no. 6, pp. 1709–1732, 2022.
- [4] D. Devasenapathy et al., "Real-Time Classroom Emotion Analysis Using Machine and Deep Learning for Enhanced Student Learning," *Journal of Intelligent Systems and Internet of Things*, vol. 16, no. 2, pp. 82–101, 2025.
- [5] Y. S. Khandekar, "Intelligent Multimodal Emotion Recognition Framework for Personalized and Adaptive STEM Education," *International Journal for Research Trends and Innovation*, vol. 11, 2026.
- [6] M. Mohana and P. Subashini, "Facial Expression Recognition Using Machine Learning and Deep Learning Techniques: A Systematic Review," *SN Computer Science*, vol. 5, no. 4, 2024.
- [7] A. Professor, "Facial Expression-Based Emotion Detection for Adaptive Teaching in Educational Environments," *International Journal of Innovative Science and Research Technology*, vol. 10, no. 1, 2025.
- [8] V. V. N. Raju et al., "Enhancing emotion prediction using deep learning and distributed federated systems with SMOTE oversampling technique," *Alexandria Engineering Journal*, vol. 108, pp. 498–508, 2024.
- [9] A. Sharma and V. Mansotra, "Deep learning based student emotion recognition from facial expressions in classrooms," *International Journal of Engineering and Advanced Technology*, vol. 8, no. 6, pp. 4691–4699, 2019.
- [10] X. Wu et al., "A deep learning approach to emotionally intelligent AI for improved learning outcomes," *Scientific Reports*, vol. 16, no. 1, 2026.