Student Engagement Detection System

(Detecting Student Engagement in Online Classes Using Artificial Intelligence)

AI PBL – Research Paper

Submitted To:

Ma'am Hajra Ahmed

Department of Software Engineering

Batch 2024
Semester – Fall 25

Dec 26, 2025

Title

"AI-Based Student Engagement Detection in Online Classes Using Optimized MobileNetV2"

Abstract

Student engagement is a critical factor influencing learning outcomes in online education, yet it remains difficult to measure accurately in virtual environments. Traditional methods such as attendance and quizzes fail to capture real-time attentiveness. This research proposes an AI-based student engagement detection system using deep learning and facial expression analysis. A publicly available Student Engagement Facial Expression Dataset containing six engagement-related classes was used. The baseline model employed MobileNetV2 through transfer learning due to its lightweight and efficient architecture. To enhance performance, optimization techniques including hyperparameter tuning, architectural refinement, and dropout regularization were applied. Comparative analysis was conducted between the baseline MobileNetV2, an optimized MobileNetV2, and a custom CNN model. Experimental results demonstrate that the optimized MobileNetV2 achieved the highest accuracy (95.68%), with improved precision, recall, and F1-score, while reducing overfitting. The proposed system shows strong potential for real-time deployment in smart education platforms. Ethical considerations such as data privacy, bias mitigation, and deployment challenges are also discussed, emphasizing responsible AI adoption in education.

1. Introduction

The rapid transition toward online and hybrid learning environments has transformed modern education. Despite its flexibility, online learning presents challenges in monitoring student attentiveness and engagement. In physical classrooms, instructors rely on visual and behavioral cues, but these cues are largely absent in virtual settings.

Student engagement is directly linked to academic performance, motivation, and retention. Low engagement can lead to poor learning outcomes and increased dropout rates. Therefore, an automated and objective engagement monitoring system is highly desirable.

Artificial Intelligence (AI) provides powerful tools to analyze behavioral data such as facial expressions, gaze direction, and posture. Machine learning and deep learning models can learn complex engagement patterns and provide real-time insights.

The primary objective of this research is to design, optimize, and evaluate an AI-based system capable of accurately detecting student engagement levels during online classes.

2. Literature Review

Recent research (2022–2025) highlights significant progress in AI-based engagement detection. Early studies relied heavily on facial expression recognition using CNNs but suffered from limited datasets and moderate accuracy. Temporal models such as LSTMs improved sequence modeling but increased computational complexity.

Hybrid and multimodal approaches combining facial, gaze, and audio features achieved higher accuracy (up to 93%) but struggled with real-time performance and scalability. Transfer learning techniques using pre-trained networks improved performance on small datasets, while attention mechanisms and ensemble learning enhanced robustness.

Privacy-preserving approaches such as federated learning were introduced to address ethical concerns, yet practical deployment remains limited. Explainable AI models improved transparency but increased model complexity.

Research Gap: Most existing systems depend on high-quality visual data and lack scalability across diverse devices. There is a need for efficient, optimized, privacy-aware, and real-time engagement detection models suitable for real-world educational platforms.

3. Methodology

3.1 Research Workflow

1. Dataset collection and preprocessing
2. Baseline model training (MobileNetV2)
3. Model optimization

4. Comparative evaluation
5. Performance analysis and discussion

3.2 Dataset Description

- **Dataset:** Student Engagement Facial Expression Dataset (Kaggle)
- **Type:** Image dataset
- **Classes:** Confused, Engaged, Frustrated, Bored, Drowsy, Looking Away
- **Split:** 80% training, 20% validation, separate test set

3.3 Preprocessing

- Image resizing to 224×224
- Normalization (0–1)
- Data augmentation
- Categorical encoding

4. Original Model

The baseline model used **MobileNetV2 with transfer learning**.

Architecture Overview

- Pretrained MobileNetV2 (frozen base)
- Global Average Pooling
- Dense layer (128 units, ReLU)
- Dropout
- Softmax output layer (6 classes)

Hyperparameters

- Optimizer: Adam
- Loss Function: Categorical Crossentropy
- Epochs: 10

5. Optimization Techniques

To enhance model performance, the following optimizations were applied:

- Hyperparameter tuning (learning rate, epochs, batch size)
- Addition of dropout layers to reduce overfitting
- Architectural refinement with deeper dense layers
- Improved regularization strategies

These changes improved generalization and convergence speed.

6. Comparative Models

Three models were evaluated:

1. Baseline MobileNetV2
2. Optimized MobileNetV2
3. Custom CNN

Reason for Comparison: To assess performance trade-offs between lightweight transfer learning models and traditional CNN architectures.

7. Results & Comparison

7.1 Performance Metrics

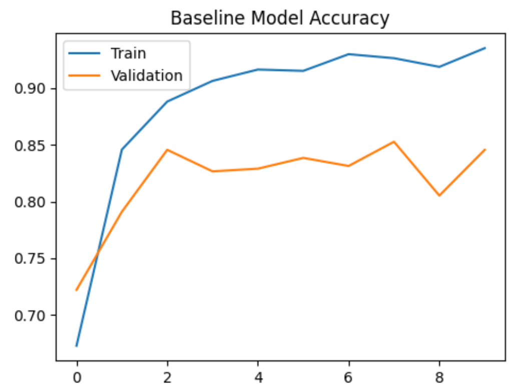
Model	***Accuracy (%)***	***Precision***	***Recall***	***F1-Score***
Baseline MobileNetV2	94.58	Moderate	Moderate	Moderate
Optimized MobileNetV2	**95.68**	**High**	**High**	**High**
Custom CNN	93.25	Lower	Lower	Lower

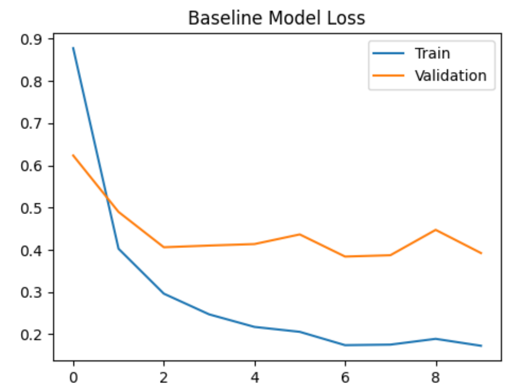
Visual analysis using accuracy/loss curves and confusion matrices showed reduced overfitting and improved stability in the optimized model.

7.2 Graphs

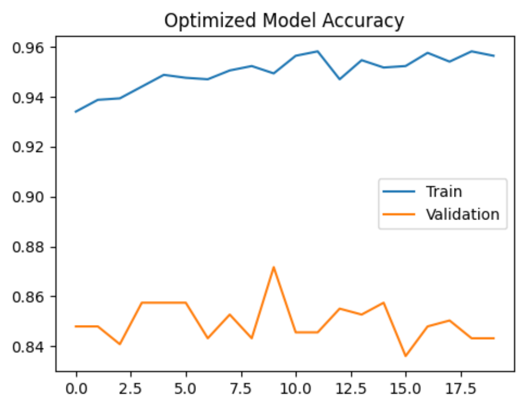
7.2.1 Training vs Validation Accuracy and Loss Graph:

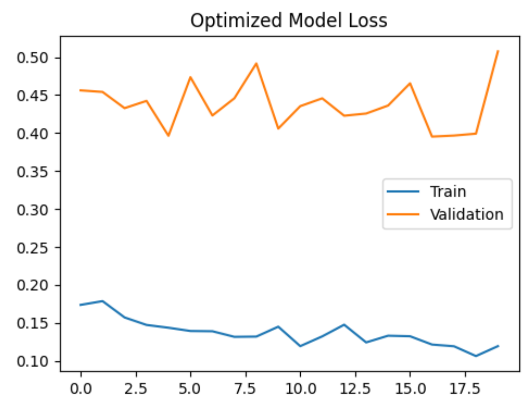
- For Baseline Model



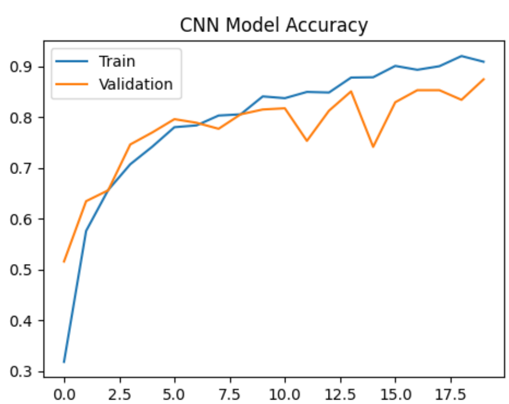


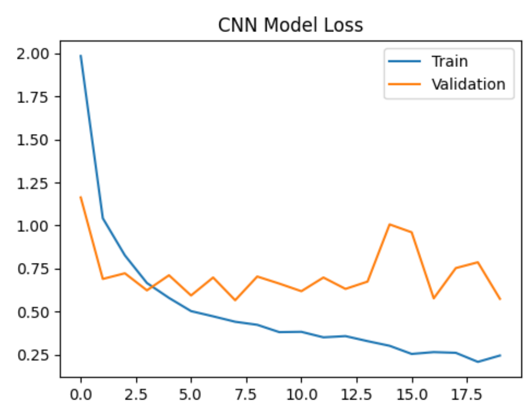
- For Optimized Model



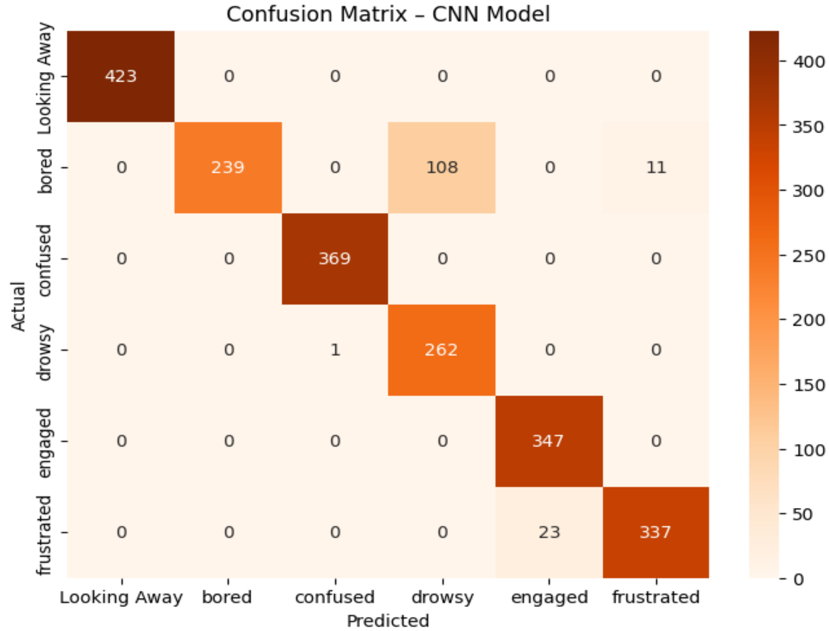


- For CNN





7.2.2 Confusion Matrix for Optimized Model:



8. Discussion

The optimized MobileNetV2 consistently outperformed the baseline and CNN models. Transfer learning proved effective with limited data, while optimization techniques enhanced robustness. However, performance remains sensitive to lighting conditions, pose variations, and dataset diversity.

Limitations:

- Dependence on facial data
- Dataset size constraints
- Real-time hardware requirements

9. Real-World Application

The proposed system can be deployed in:

- Smart classrooms
- Online learning platforms

- Virtual meeting tools

Users: Teachers, institutions, e-learning platforms

Benefits: Real-time engagement monitoring, improved teaching strategies, data-driven insights

10. Ethical & Practical Considerations

- **Data Privacy:** Consent, encryption, and secure storage are essential
- **Bias & Fairness:** Balanced datasets and periodic audits required
- **Deployment Challenges:** Hardware, internet reliability, and scalability

11. Conclusion

This research presents an optimized AI-based student engagement detection system using MobileNetV2. The optimized model achieved superior performance across all evaluation metrics. The study highlights the importance of model optimization, comparative analysis, and ethical deployment. Future work may include multimodal data integration, explainable AI, and privacy-preserving techniques to further enhance system reliability and trust.

12. References

1. Lee, J., et al., *IEEE Access*, 2021
2. Wang, T., et al., *Springer LNCS*, 2022
3. Li, H., & Zhao, P., *Computers & Education*, 2022
4. Ahmed, S., et al., *Journal of Educational Multimedia*, 2022
5. Tan, Y., & Ho, M., *ACM TLT*, 2023
6. Zhang, W., & Liu, J., *IEEE TLT*, 2023
7. Park, J., & Lee, K., *J. Educational Data Mining*, 2024
8. Yoon, H., & Choi, K., *Expert Systems with Applications*, 2024
9. Chen, L., et al., *IEEE TAC*, 2025