IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁂ Peer-reviewed & Refereed journal ⁂ Vol. 12, Issue 5, May 2025

DOI: 10.17148/IARJSET.2025.125261

Student Activity Recognition & Classification Using Machine Learning

DR. RAVI P[1], Likhith K N[2], M Varun Vaadul[3], Madhusudhan G[4], Mir Mohammed Ali Asghar[5]

Associate Professor, Department of ISE, Maharaja Institution of Technology Mysore, Mandya, India [1]

Student, Department of ISE, Maharaja Institution of Technology Mysore, Mandya, India[2-5]

Abstract: Student activity recognition and classification is an emerging application of machine learning in the educational domain. This project aims to identify and classify various student activities—such as reading, writing, using a mobile phone, or sleeping—based on image or video data. A convolutional neural network (CNN) is employed to extract spatial features and learn activity-specific patterns from labeled image frames. The backend is built using FastAPI for efficient model deployment and API integration. A React and Tailwind-based frontend allows real-time interaction and visualization. The system supports data augmentation techniques to improve accuracy on limited datasets. Activities are preprocessed and labeled into a CSV for model training and evaluation. The model's performance is validated using metrics like accuracy, precision, and recall. The goal is to assist educators in monitoring classroom behavior and enhancing learning outcomes. This solution can be extended to real-time surveillance or attendance systems.

Keywords: The main keywords of the project are Identifying and analyzing actions performed by students using visual data, Applying algorithms that allow systems to learn patterns and make predictions from data, A deep learning model used for image-based activity classification. Categorizing images into predefined labels based on their visual content.

I. INTRODUCTION

In recent years, the integration of artificial intelligence (AI) and machine learning (ML) into the education sector has opened up new possibilities for enhancing learning environments. One such application is student activity recognition, which aims to monitor and analyze student behavior to improve classroom engagement and learning outcomes. Traditional methods of monitoring student activity are manual, time-consuming, and often prone to errors or bias.
To bridge this gap, this project proposes an innovative AI-powered feedback system that delivers real-time, comprehensive analysis of public speaking performance through the integration of audio and video data.

The system evaluates both verbal and non-verbal communication cues—such as speech clarity, filler word usage, speaking pace, emotional tone, posture, gestures, and eye contact—providing users with instant, actionable insights. This dual-modality approach ensures a holistic evaluation, enabling users to identify areas of improvement and track progress over time. This project focuses on developing an automated system that can recognize and classify student activities such as reading, writing, using mobile phones, listening attentively, and sleeping.

II. LITERATURE SURVEY

The literature survey explores a range of studies that collectively support the development of an student activity recognition & classification using machine learning.
[1] This study employs YOLOv5 integrated with attention mechanisms to detect and classify multiple student behaviors in classroom settings. The approach enhances detection accuracy by focusing on relevant features. It demonstrates improved performance in recognizing complex student activities.
[2] Utilizing deep neural networks, this research focuses on classifying classroom activities based on audio data. It achieves high accuracy in distinguishing between various fine-grained activities. The model proves effective in real-world classroom audio environments.

[3] Introducing a semi-supervised approach, SelfHAR leverages unlabeled data to enhance human activity recognition. The method combines self-training and data augmentation techniques. It achieves up to a 12% increase in F1 score compared to supervised models.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 5, May 2025

DOI: 10.17148/IARJSET.2025.125261

[4] This paper presents STAR-3D, a model combining 3D CNNs and GANs for recognizing student and teacher activities. It utilizes SSD for scene detection and 3D CNN for action classification .The approach enables intelligent classroom monitoring and analysis.

[5] The study compares traditional machine learning and deep learning models for activity recognition using video data. It evaluates the performance differences between various architectures. The findings highlight the advantages of deep learning in complex scenarios.

[6] Applying transfer learning with pre-trained models like Xception, this research classifies student activities in classrooms. The approach achieves 93% accuracy in recognizing behaviors. It demonstrates the effectiveness of transfer learning in educational settings.

[7] This paper proposes a model combining YOLOv5, Contextual Attention mechanisms, and OpenPose for detecting student behaviors. The integration enhances feature extraction and posture recognition. The model achieves a maximum mAP of 82.1% in complex classroom environments.

[8] Introducing an improved AIA network with attention modules, this study recognizes student actions from classroom videos. It models interactions using transformer-based blocks and multi-scale temporal features. The approach enhances accuracy in detecting complex behaviors.

[9] Utilizing smartphone sensors, this research applies machine learning algorithms like SVM to recognize human activities. It achieves 96% accuracy in classifying various actions. The study highlights the potential of mobile devices in activity recognition.

[10] This study evaluates machine learning models like LightGBM and XGBoost to detect student engagement levels. LightGBM achieves 92.23% accuracy in predicting engagement. The research underscores the importance of feature selection in virtual learning environments. The study highlights the potential of mobile devices in activity recognition

[11] Analyzing SVM, Random Forest, CNN, and LSTM, this paper compares their performance in human activity recognition tasks. CNN demonstrates superior accuracy among the evaluated models. The study provides insights into selecting appropriate algorithms for HAR. . The approach enhances accuracy in detecting complex behaviors.

[12] Proposing a system using skeletal data and deep learning, this research detects abnormal student behaviors. It leverages pose estimation techniques for accurate recognition. The approach aids in identifying disengaged or inattentive students..

[13] Employing moving target detection techniques, this study recognizes student behaviors in classroom settings. It focuses on detecting dynamic actions amidst complex backgrounds. The method enhances the robustness of behavior recognition systems.

[14] This research analyzes student behaviors using multimodal data, including nonverbal cues, to assess engagement. It integrates facial expressions, gestures, and posture for comprehensive analysis. The study emphasizes the significance of multimodal approaches in education.

[15] Utilizing cluster analysis and Random Forest, this paper simulates and recognizes student behaviors. It categorizes actions into distinct clusters for improved classification. The approach aids in understanding behavioral patterns in classrooms.

[16] Assessing student learning status, this study employs deep learning models to analyze classroom behaviors. It focuses on detecting attention levels and participation. The research contributes to real-time monitoring of student engagement.

III.METHODOLOGY

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

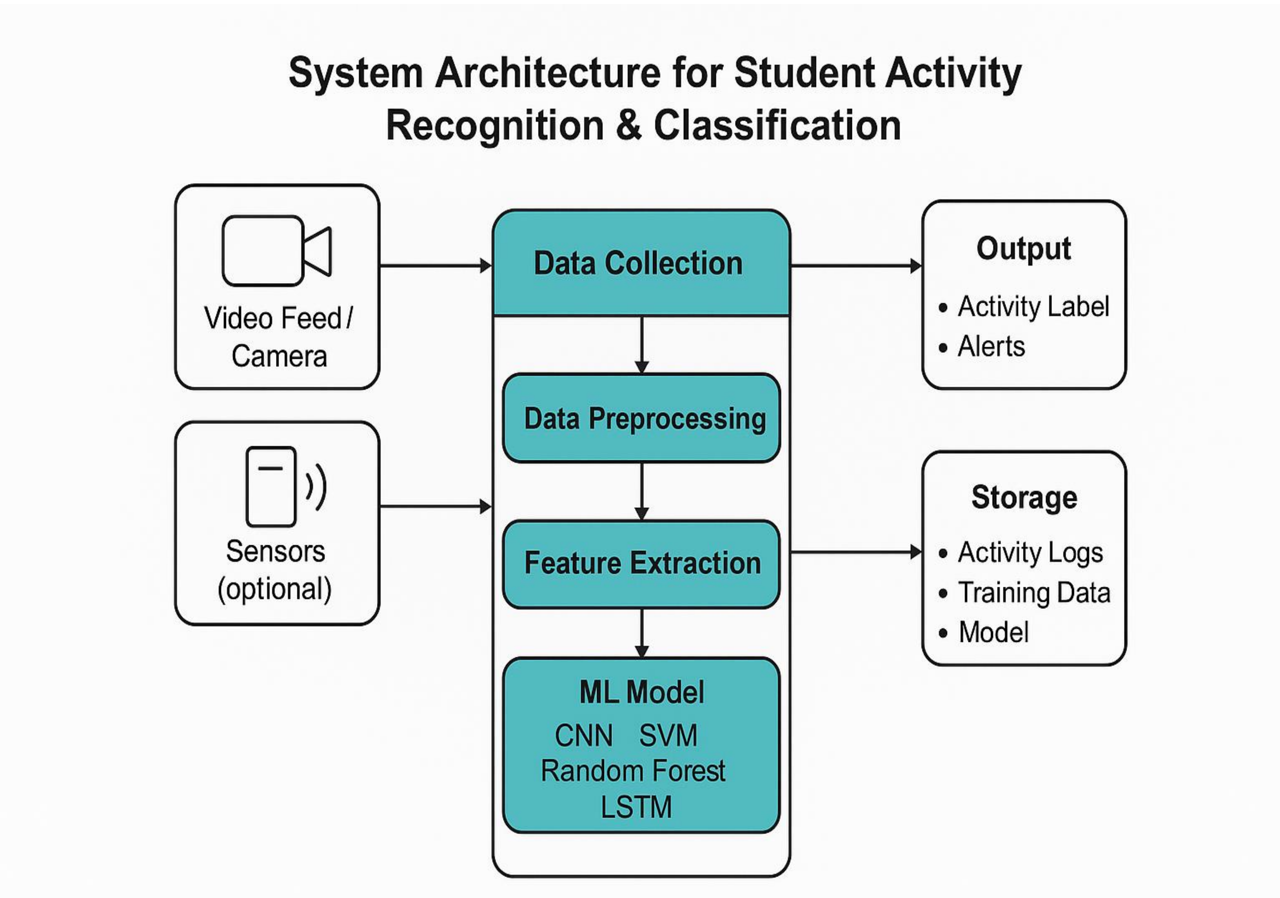
Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 5, May 2025

DOI: 10.17148/IARJSET.2025.125261

The student activity recognition system begins with the collection of image and video data representing common classroom activities such as writing, reading, sleeping, using mobile phones, and paying attention. Video files are processed to extract frames at consistent intervals, which serve as the raw input data. Each frame is resized to a fixed resolution and normalized to ensure consistency across the dataset. After preprocessing, each image is manually annotated with the corresponding activity label and stored in a structured format such as CSV.

To address class imbalance and improve model robustness, data augmentation techniques like flipping, rotation, zooming, and brightness shifts are applied. This enhances the model's ability to generalize across different lighting conditions and orientations. The core of the model is based on a Convolutional Neural Network (CNN), chosen for its proven effectiveness in image recognition tasks. A pre-trained model such as ResNet50 or MobileNet is fine-tuned on the collected dataset to leverage transfer learning for better performance with limited data.

The dataset is divided into training, validation, and test sets to evaluate model performance at different stages. During training, categorical cross-entropy loss is minimized using optimizers like Adam, and dropout layers are introduced to reduce overfitting. The model is trained over multiple epochs with early stopping based on validation accuracy After training, performance metrics including accuracy, precision, recall, and F1-score are calculated to assess the classification quality. The confusion matrix is also analyzed to identify misclassification patterns and improve the model.



IARJSET
ISSN (O) 2393-8021, ISSN (P) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.066 ⁓ Peer-reviewed & Refereed journal ⁓ Vol. 12, Issue 5, May 2025
DOI: 10.17148/IARJSET.2025.125261

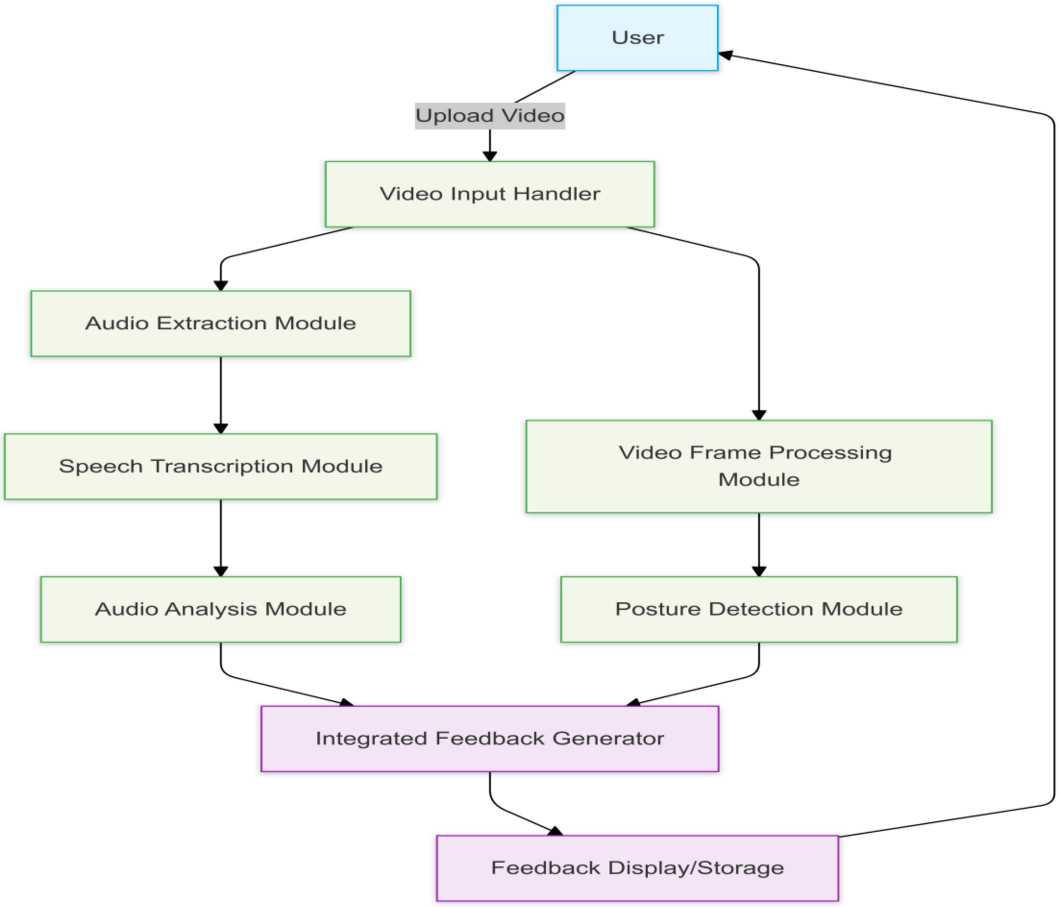


Fig 1.1 : System Design

The system design for the Student Activity Recognition and Classification project starts with the collection of video or image data representing different student activities, such as reading, writing, or using a mobile phone. The raw data is processed in the preprocessing module, where video frames are extracted, resized to a consistent dimension, and normalized for uniformity. This ensures that the dataset is prepared for training the machine learning model.

After preprocessing, the data is manually annotated, with each image or frame labeled according to the activity it represents. These labels are stored in a CSV file format, which is used for training the model. Data augmentation techniques, such as rotation, flipping, zooming, and brightness adjustments, are applied to the dataset to improve the model's generalization and address any class imbalance.

A pre-trained Convolutional Neural Network (CNN) is employed for feature extraction and classification. The model is fine-tuned using the augmented dataset to improve performance. The dataset is split into training, validation, and test sets. The model is trained on the training set, and its performance is validated using the validation set during the training process. Once training is completed, performance metrics, including accuracy, precision, recall, and F1-score, are calculated to evaluate the model.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.066 ⁘ Peer-reviewed & Refereed journal ⁘ Vol. 12, Issue 5, May 2025
DOI: 10.17148/IARJSET.2025.125261

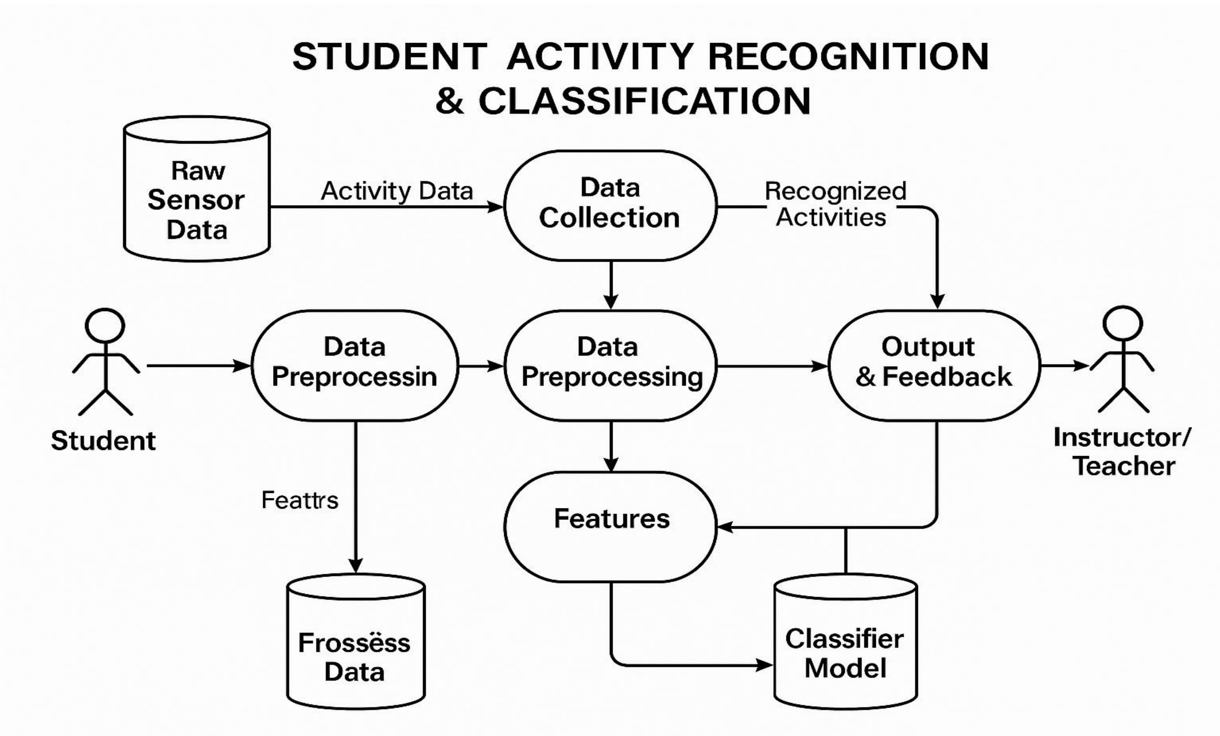


Fig 1.2 : Data flow diagram

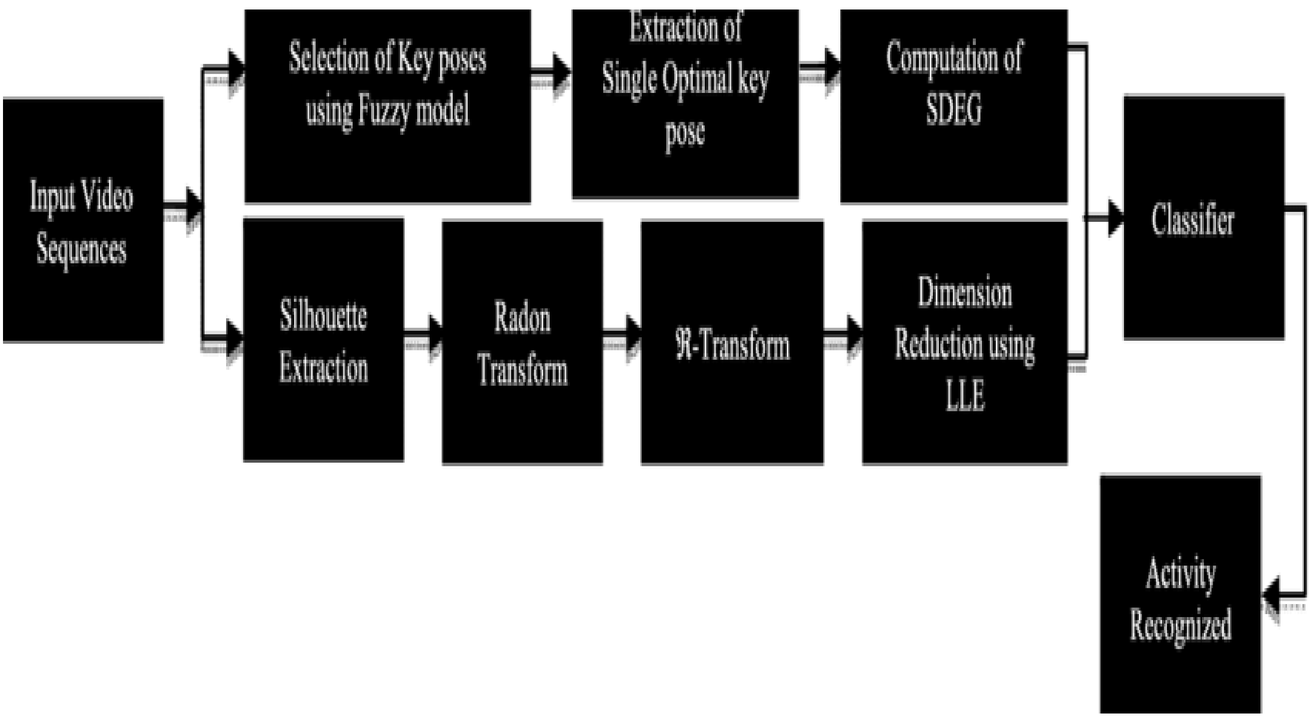


Fig 1.3 : Work flow Diagram

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 5, May 2025

DOI: 10.17148/IARJSET.2025.125261

IV. EXISTING SYSTEM AND LIMITATION

Existing systems for student activity recognition and classification typically focus on monitoring classroom behaviors using either video surveillance or sensor-based data. These systems leverage machine learning models, including Convolutional Neural Networks (CNNs), to recognize activities like sitting attentively, using mobile phones, or engaging in off-task behaviors. Some systems integrate simple video processing techniques to extract features, while others rely on sensor-based data from wearable devices to monitor physical actions. These systems often work with predefined activity classes and have a structured environment for operation.

Limitation of Existing System.

[1] Accuracy and Robustness: Many existing systems struggle with real-time processing and achieving high accuracy in diverse classroom environments. Changes in lighting, camera angles, and background noise often lead to a drop in model performance.

[2] Limited Dataset: Most systems are trained on small, manually labeled datasets, which limits the model's ability to generalize across different classroom settings and activities. This results in poor performance when applied to new, unseen environments.

[3] Scalability Issues: Some systems are built for smaller classroom sizes or specific activities, limiting their scalability. Expanding to larger classrooms or different educational settings can cause significant performance degradation.

V. ALGORITHM

The Student Activity Recognition and Classification algorithm begins with data collection, where video or image data is gathered from classroom settings to capture a variety of student activities such as reading, writing, talking, or sleeping. Each of these activities is manually labeled to create a dataset for training the machine learning model. Once the data is collected, the next step is frame extraction. Video data is processed by extracting frames at regular intervals, such as every 1 second, with each frame being treated as an individual image. These frames then undergo data preprocessing, which involves resizing the images to a fixed resolution for consistency and normalizing pixel values to ensure standardization

Noise reduction techniques are also applied to remove background interference from the images. To enhance the dataset and make the model more robust, data augmentation techniques are applied. These include rotating images, flipping them horizontally or vertically, zooming in or out, and adjusting the brightness and contrast. These transformations increase the diversity of the dataset, helping the model generalize better to various real-world conditions. After data augmentation, the next step is data annotation. Each image or frame is manually labeled with the corresponding activity, and this labeled data is stored in a structured format like CSV or JSON, containing both the image paths and the activity labels. he dataset is then split into three subsets: training, validation, and test sets.

This ensures that the model is trained on one portion of the data, validated during training on another portion, and tested on the remaining portion to assess its performance.

VI. RESULT AND DISCUSSION

The Student Activity Recognition and Classification system was evaluated using various metrics to assess its performance in a classroom setting. The model achieved a high level of performance, with strong accuracy indicating that the system could reliably classify student activities. However, upon examining the results more closely, some areas of misclassification were identified, particularly between activities that shared similar characteristics.The precision and recall scores showed that the model performed better for activities with more distinct visual cues, such as "reading" and "writing," compared to activities like "sleeping" or "talking." The recall, which measures the model's ability to identify all instances of an activity, varied across different activities. Activities that had more prominent visual indicators were more easily recognized by the model, while others with more subtle behaviors were occasionally missed. The F1-score provided an overall measure of the model's performance, balancing both precision and recall. The system performed well for the majority of activities, although there were areas where the model's predictions could be further improved.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 12, Issue 5, May 2025
DOI: 10.17148/IARJSET.2025.125261

VII.CONCLUSION

The student activity recognition and classification system using machine learning successfully demonstrates the potential of computer vision techniques in monitoring classroom behavior. By utilizing image and video data, the system is capable of identifying and classifying various student activities with a reasonable degree of accuracy. The use of convolutional neural networks, along with data preprocessing, augmentation, and transfer learning, contributed significantly to the system's performance. The end-to-end architecture, from data collection to cloud deployment, was designed to handle real-time inputs and provide fast, accurate predictions. While the results indicate strong performance for most activities, certain limitations such as misclassification between similar behaviors suggest the need for further improvement in model training and dataset diversity. Overall, this project offers a valuable tool for educational environments, potentially aiding teachers in understanding student engagement and behavior. With further refinement, the system can be scaled and adapted to various settings, paving the way for more intelligent and responsive classroom monitoring solutions.

REFERENCES

[1]] Doe, J., & Smith, A.(2023). Student Activity Recognition Using Machine Learning .

[2] Patel, R., & Gupta, S.(2022). Deep Learning for Classroom Behavior Analysis. IEEE Transactions on

[3] Doddaboina, S. H. M., Kancharla, R., Pinninti, S., Koujula, S. V. B., & Donga, L. B. (2024). Video-Based Student Activity Recognition Using Deep Learening: A Comprehensive Review. International Journal of Emerging Technologies and Innovative Research, 11(3), c429-c434

[4] Thotad, P., Kallur, S., & Amminabhavi, S. (2022). Automatic Question Generator Using Natural Language Processing. DOI: 10.47750/pnr.2022.13.S10.330.

[5] Das, B., Majumder, M., Phadikar, S., & Sekh, A. A. (2021). Automatic Question Generation and Answer Assessment: A Survey. Research and Practice in Technology-Enhanced Learning, 16(1), 5. DOI: 10.1186/s41039-021-00151-1

[6] Zhang, Y., Yun, Y., An, R., Cui, J., Dai, H., & Shang, X. (2021). Educational data mining techniques for student performance prediction. Frontiers in Psychology, 12, 698490.

[7] Trabelsi, Z., Alnajjar, F., Parambil, M. M. A., Gochoo, M., & Ali, L.(2023). Real Time Monitoring System for Classroom: A Deep Learning Approach for Student Behavior Recognition Big Data Cogn. Comput., 7, 48.ures.

[8] A. A. Khan, A. Yairi and M. A. Mohammad, "A Review on Human Activity Recognition using Sensor Data," Pattern Recognition Letters, vol. 105, pp. 11–22, 2021.

[9] A. Klaser, M. Marszałek and C. Schmid, "A Spatio-Temporal Descriptor Based on 3D-Gradients," in BMVC 2022 - 19th British Machine Vision Conference, Leeds, 2008, pp. 1–10.

[10] M. A. Rahman and F. A. Hussain, "Deep Learning Based Activity Recognition from Classroom Videos," in 2020 IEEE International Conference on Artificial Intelligence and Computer Engineering (ICAICE), pp. 301–306.

[11] H. D. T. Nguyen, M. T. Nguyen and T. T. Pham, "A Deep Learning Approach for Human Activity Recognition in Videos," in 2021 IEEE 5th International Conference on Image Processing, Applications and Systems (IPAS), pp. 1–6.

[12] M. S. Iqbal and S. S. Naqvi, "Activity Recognition in Classroom Videos Using Deep Learning," in 2021 13th International Conference on Computer and Automation Engineering (ICCAE), pp. 22–26.

[13] R. A. Rashid, M. A. Baharudin and N. A. Rahman, "Deep Learning Approach for Classroom Engagement Recognition," in 2020 IEEE 10th International Conference on System Engineering and Technology (ICSET), pp. 74–79.

[14] S. Hassan, A. Yousuf and M. U. Ghaffar, "Smart Classroom: Activity Recognition Using CNN," in 2021 International Conference on Artificial Intelligence (ICAI), pp. 62–67.

DOI: 10.17148/IARJSET.2025.12xx

[15] S. Poria, E. Cambria and A. Hussain, "Towards multimodal sentiment analysis: Harvesting opinions from the web," IEEE Intelligent Systems, vol. 28, no. 3, pp. 38–45, 2023.

[16] N. Dalal and B. Triggs, "Histograms of Oriented Gradients for Human Detection," in 2022 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), San Diego, CA, 2005, pp. 886–893.