

# Analyzing Classroom Behavior to Enhance Student Engagement and Learning Outcomes

Mr. Suresh Kumar S  
Professor of Artificial Intelligence and Data  
Science  
Rajalakshmi Engineering College  
Chennai, India  
[email]@rajalakshmi.edu.in

Atmakuru Siva Sandeep  
Artificial Intelligence and Data Science  
Rajalakshmi Engineering College  
Chennai, India  
221801501@rajalakshmi.edu.in

Buvankalyan P  
Artificial Intelligence and Data Science  
Rajalakshmi Engineering College  
Chennai, India  
221801506@rajalakshmi.edu.in

**Abstract—** This research presents a high-tech smart system that utilizes video records to monitor student behavior in classroom settings and identify specific activities like raising hands and dozing. The uneven frequency of certain behaviors along with variances in student posture, seating arrangements, and video quality make it extremely difficult to identify these behaviors in a live classroom. While some behaviors, like raising one's hand, may happen frequently and in a variety of ways, others, like napping, may happen less frequently but are just as crucial to comprehending student engagement. We use a sizable annotated dataset, in which bounding boxes are constructed around significant objects of interest in the video frames, to get over these difficulties. Our system's core is a convolutional neural network (CNN), which we have specially tuned for behavior identification. The CNN model is made to quickly evaluate the visual data and recognize characteristics and patterns connected to certain student behaviors in real time. We make sure the model can accurately identify actions like raising one's hand and falling asleep even when there are variations in lighting, seating configurations, and video quality by training it with a large dataset. Our model is specifically designed to achieve high accuracy in identifying minute cues that distinguish one movement from another, in contrast to conventional methods that might find it difficult to comprehend the intricacy of these actions.

Through extensive experimentation and validation, our system has demonstrated strong performance in detecting these classroom behaviors, offering valuable insights into student engagement. By providing real-time feedback, the system allows educators to monitor participation levels and identify students who may be disengaged. These insights empower teachers to make informed, data-driven decisions about their teaching strategies and classroom management, ultimately fostering a more interactive and productive learning environment. The implementation of this CNN-based behavior recognition system represents a significant step forward in the use of AI for enhancing classroom dynamics and improving student outcomes.

**Keywords—** Smart system, Video records, Student behavior, Convolutional neural network (CNN), Real-time feedback, Improving student outcomes.

## I. INTRODUCTION

Analyzing student behavior in the classroom is critical for assessing teaching quality and predicting long-term academic achievement. Teacher observations, which are the mainstay of traditional techniques of evaluating student behavior, can be laborious, subjective, and frequently unproductive when it comes to keeping an eye on big classes of kids.

This drawback emphasizes the urgent need for a more sophisticated system that can recognize and evaluate student behavior on its own, enabling thorough assessments and new perspectives on classroom dynamics. The goal of this study is to create an advanced intelligent system that can more effectively and accurately monitor classroom interactions by using video footage to examine student behavior. Using video analytics allows us to record a wider variety of student behaviors in real time, giving us a more comprehensive picture of student participation in the classroom. The current state of behavior detection algorithms can be broadly divided into three categories: pose estimation techniques that rely on the recognition of human body positions; object detection techniques that identify specific actions or objects within the video frames; and hand-crafted features-based algorithms that require manual design and tuning. Our suggested approach is made to increase the precision and efficacy of identifying student behaviors, giving teachers insightful data on student involvement and academic results.

By enhancing behavior detection accuracy, this study tackles the limitations of current approaches and advances automated classroom analysis. Our system's insights will enable educators to make data-driven, well-informed decisions that can improve learning outcomes and student engagement. The ultimate goal of this effort is to create a more active and effective learning environment in the classroom, which will eventually lead to better teaching strategies and increased student achievement. Through the deployment of this novel system, we intend to contribute to the burgeoning field of educational technology and promote the effective application of AI-driven solutions in the classroom. Automated behavior analysis reduces burdens for teachers and offers impartial evaluations of student participation. By using sophisticated algorithms to spot trends in behavior that more conventional techniques of observation might overlook, instructors can adjust their lesson plans in real time. Through this tracking, educational institutions may identify patterns and improve student performance and teaching effectiveness, laying the groundwork for a time when technology and education will be integrated to create productive learning environments.



Figure 1: Tech Stacks

## II. RELATED WORKS

In the paper "Teacher-Student Behavior Recognition in Classroom Teaching," Henghuai Chen describes a new approach to classroom behavior recognition that makes use of a modified YOLO-v4 model. The authors discuss the drawbacks of conventional classroom monitoring, which frequently depends on manual observation and makes it difficult to identify behavior in real time. The system recognizes student behaviors like hand-raising and inattention with over 90% accuracy by utilizing approaches like repulsion loss functions and cross-stage local networks. IoT technology integration improves real-time data collecting and makes it possible to continuously analyze behavior without the need for human interaction. However, environmental elements like lighting and occlusion might affect how well the system performs. Although this method greatly enhances behavior detection, there are several drawbacks, such as the dependence on accurate data and the requirement for real-time computational resources.

In the work "Learning Behavior Recognition in Smart Classroom with Multiple Students," Zhifeng Wang delves into an elaborate method of YOLOv5 improved with Squeeze-and-Excitation Networks-based classroom behavior detection. This study aims to precisely detect subtle and variable classroom actions in busy, dynamic classroom environments. The approach makes use of adaptive anchoring and numerous scales to recognize objects at different sizes efficiently, increasing overall accuracy in complicated situations. More efficient classroom management is made possible by this intelligent system, which gives teachers exact insights into the participation and engagement levels of their students. The project's ultimate goal is to aid in the creation of smart classrooms, where student engagement and teaching tactics can be greatly improved by real-time behavior recognition.

In the study "Real-Time Student Behavior Analysis Using Deep Learning," Xiaofei Wu offers a novel method for efficiently monitoring and analyzing student actions in real-time using convolutional neural networks (CNNs). After training on a large dataset, the system is capable of recognizing a variety of student behaviors, including raising their hands, being engaged, and being inattentive. The solution facilitates educators' ability to modify their teaching tactics in real-time to improve student engagement during lessons by giving them prompt feedback. The model's remarkable efficiency enables it to function at its best in a variety of classroom settings, such as ones with shifting lighting, seating configurations, and camera angles. The ability to provide real-time analytics is essential for maximizing student engagement in the classroom since it assists teachers in keeping students' attention and involvement throughout the lesson.

In the paper "Facial Expression Recognition in Classrooms," Rong Zhang describes a sophisticated system that uses the Facial Action Coding System (FACS) and convolutional neural networks (CNNs) to assess students' facial expressions while they are participating in class activities. Teachers can gain important insights into their students' emotional states by using this creative method to categorize a variety of emotions, including perplexity, impatience, and interest. This kind of information is essential for making real-time modifications to instructional tactics, allowing teachers to quickly address problems like student disengagement or frustration. By showing how facial expression recognition can result in more individualized and successful teaching strategies, Zhang's work emphasizes the value of emotional intelligence in education. The ultimate goal of this initiative is to use technology to better understand and promote a more responsive learning environment.

In the paper "Automated Classroom Monitoring System," Jianming Shi presents a novel approach that combines optical flow techniques with Support Vector Machines to Efficiently

detect and evaluate student movement in classroom environments. This method automates the identification of important actions, including hand-raising and fidgeting, which are critical for determining student involvement. With the help of optical flow to record motion patterns and SVM for behavior categorization, this research offers a productive way to keep an eye on classroom dynamics without requiring teachers to be there all the time. Teachers no longer have to spend as much time monitoring students, freeing them up to focus on teaching while still getting thorough feedback on their engagement levels.

In the article "Student Engagement Detection Using LSTM Networks," Wei Li explores the efficient recording and analysis of video frame sequences in classroom settings using Long Short-Term Memory (LSTM) networks. This novel approach, which stresses the significance of temporal patterns in contrast to conventional static behavior recognition models, allows it to monitor changes in student involvement over the course of a class. The LSTM model is trained on sequences that represent a variety of actions, including attentiveness, hand-raising, and periods of inattentiveness. This allows the system to provide educators with important insights into the reasons behind students' attention fluctuations. This feature helps teachers to detect important moments for intervention, promoting an environment that increases student engagement and learning results. Li's experiment demonstrates how utilizing cutting-edge machine learning techniques can result in a more

In the work "Gesture Recognition for Classroom Interaction," Yu Chen investigates the use of Hidden Markov Models (HMM) for the detection and identification of a range of classroom gestures, such as pointing, raising hands, and other non-verbal indicators. With this real-time gesture detection system, teachers can react to student motions instantly, which improves classroom engagement. The system's rapid gesture detection creates a more dynamic learning environment by motivating professors to communicate with their students in real time and enabling students to participate interactively. Chen's research promotes a more inclusive learning environment by emphasizing nonverbal communication, making sure that even kids who might not speak up are recognized and included in class discussions. This creative method highlights the importance of education while also enhancing the learning process.

In the paper "Smart Classroom Monitoring System Using Computer Vision," Amit Kumar presents a state-of-the-art system that makes use of computer vision methods to track and evaluate teacher actions in real time. Deep learning algorithms are being used in this research to detect and categorize student actions including involvement, distraction, and attentiveness. Through the processing of video feeds obtained from classroom cameras, the system provides teachers with actionable insights that allow them to dynamically assess student participation. The technology also has capabilities for identifying classroom dynamics, which enables teachers to modify their lesson plans in response to immediate feedback. The potential for incorporating technology into learning environments is highlighted by Kumar's research, which shows how Smart monitoring technologies can help teachers implement more successful pedagogies and raise student achievement levels.

In her work "Interactive Learning Environment Through Video Analysis," Sara Thompson delves into the creation of an interactive learning platform that makes use of video analysis to keep an eye on student behavior and classroom interactions. The research analyzes student involvement levels, such as participation in debates and attentiveness during courses, using cutting-edge image processing techniques and machine learning algorithms. The technology records and analyzes video footage to give teachers comprehensive analytics on how students interact, enabling them to spot patterns of engagement and disengagement. Thompson highlights how crucial real-time feedback is to modifying instruction to better suit the needs of pupils. The experiment also shows how using video analysis in the classroom can promote a more participatory and collaborative learning environment. This creative method not only improves teaching techniques but also

### III. PROPOSED SYSTEM

#### System Overview

The architecture of the Student Behavior Analysis System is designed to provide a comprehensive and automated solution for monitoring student engagement in real-time. There are three main steps to the system:

#### Student Behaviour Analysis system

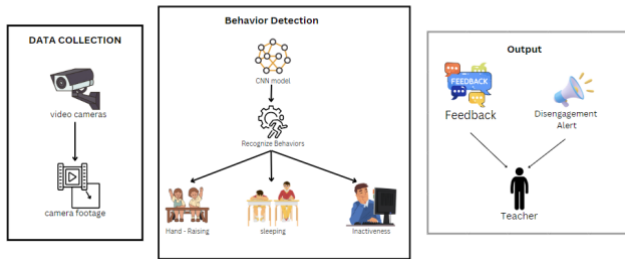


Figure 2: Overview of the System

Figure 2: Functions of the System

**Data collection:** The system's cornerstone, video cameras are placed in classrooms to continuously record pupils throughout instruction. The cameras are arranged such that they can view the classroom from a variety of perspectives, guaranteeing that no space is unobserved. Body posture, facial emotions, and hand movements are all captured by the video cameras. This video is the unprocessed data used in later phases. Accurate analysis depends on capturing even the smallest behavioral indicators, which is ensured by the deployment of high-definition cameras. After then, the camera footage is moved to the system's next phase for a more thorough analysis.

**Behavior Detection:** In the second stage, sophisticated machine learning models—more precisely, a Convolutional Neural Network (CNN)—are utilized. This CNN has been taught to identify particular behaviors in the classroom. Frame by frame, the CNN analyzes the video data to identify patterns associated with actions like raising one's hand, nodding off, and being inattentive. In real time, the CNN model recognizes important behavioral signs by concentrating on both little facial expressions and more extensive bodily actions. In order to split each video frame and emphasize regions of interest, the model uses bounding boxes, which enables the system to concentrate on certain pupils or groups. Additionally, the CNN ensures excellent accuracy by being resilient enough to manage obstacles like shifting seating arrangements, illumination fluctuations, and impediments in the classroom. This stage is essential for turning unprocessed video material into actionable data, where the system can classify and discriminate between activities like active participation and disengagement.

**Output and Feedback:** The last phase converts the identified behaviors into insightful knowledge. Teachers receive real-time feedback from the system informing them of the general levels of classroom participation. For example, the system will sound a disengagement alarm to alert the teacher to possible problems if it notices several children who are not paying attention or who are displaying indicators of disengagement, such as napping. The technology can also produce comprehensive reports on specific students or behavior patterns in the entire classroom. With the use of this data, teachers can make instant changes to their lesson plans or increase student participation. The teacher may manage the classroom dynamically and responsively thanks to the feedback loop that exists between the system and the teacher. This allows the teacher to step in at key times to enhance student engagement and learning objectives.

#### System Architecture

The Student Behavior Analysis System's architecture is painstakingly created to make it easier to monitor and analyze classroom video footage in order to successfully identify and interpret particular student actions. The video processing module, which serves as the first point of contact for handling visual data from uploaded video files, is essential to the operation of this system. With the help of sophisticated algorithms, this module can extract important elements from the video and identify important behaviors like raising the hand and other learning activities. This preprocessing stage is essential because it converts unprocessed video data into a format that later modules can readily understand and evaluate. Following the processing phase, the action detection module smoothly collaborates with the video processing module to classify and recognize distinct actions shown by students inside the classroom setting. This module uses machine learning approaches to determine student engagement levels with accuracy. This ensures that teachers receive precise and timely information on their students' participation and interactions. By giving deep insights into classroom dynamics, this module helps teachers recognize which students are actively engaged and which may be struggling, so supporting a more responsive teaching style.

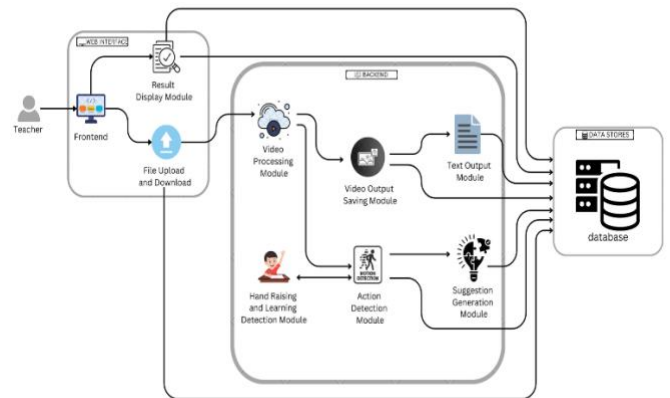


Figure 3: Architecture of the Project

The video output saving module is made to store the processed video data in an appropriate format, which improves system usability. With the help of this function, educators can watch the video again to gain a deeper comprehension of the dynamics of the classroom and the interactions between students. This kind of retrospective analysis can be quite helpful in seeing trends over time, which enables teachers to more easily modify their lesson plans to better meet the requirements of their pupils. Apart from its visual analysis features, the system also has a text output module that produces detailed reports according to the behaviors it detects. These reports are designed to give teachers practical insights into how their students participate in class and learn, which helps them make informed decisions about how best to teach. Additionally, the suggestion generating module greatly improves the general functioning of the system by providing tailored recommendations for instructional methods depending on the behaviors that are seen. This makes it possible for teachers to alter their methods in the moment, creating a more dynamic and productive learning environment.

Teachers can access historical data and monitor changes in student behavior over time because all processed data from these modules is safely kept in a central database. This architectural feature is essential for continuous evaluation and instructional technique.

advancement. With the user experience in mind, the frontend interface makes it simple for teachers to submit video files, evaluate analysis findings, and download reports. Because teachers frequently have limited time and need effective tools to improve their teaching efficacy in real-world classroom settings, this user-friendly design is essential for practical application. All things considered, this architecture creates a potent tool for improving educational results by fusing state-of-the-art video processing technology with a user-centric design approach. By mixing sophisticated detection methods with an easy interface, the system empowers educators to monitor and adjust to student requirements dynamically, ultimately leading to better engagement and learning achievement in classroom contexts.

## User Interface Design

The user interface (UI) of your student behavior analysis system is intended to be intuitive and user-friendly, allowing educators to easily upload video footage for analysis. The interface is simple to use and was created with Flask for the backend and HTML, CSS, and JavaScript for the frontend. Users can choose and submit video files for processing by simply navigating to the upload section.

After the video is uploaded, the system automatically analyzes the film to identify inattentiveness and hand-raising gestures. Following processing, consumers have access to feedback via a succinct and easy-to-read display module that presents the analysis's findings. Teachers can immediately analyze student engagement levels and modify their teaching tactics based on the actionable insights included in this feedback. The incorporation of By making the interface available on a range of devices, responsive design improves usability and gives teachers an effective tool to improve classroom dynamics.

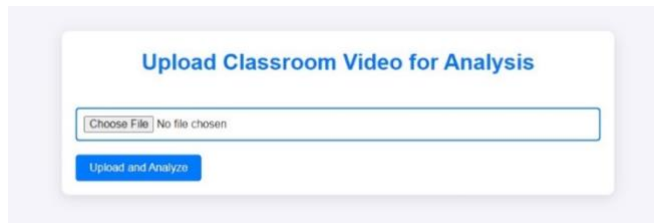
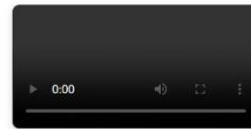


Figure 4: Video Uploading Section

The output portion of your user interface is critical for delivering practical insights to instructors based on classroom video analysis. As soon as the uploaded video is processed, the system produces feedback that draws attention to important behavioral observations. These observations include engagement levels and instances of particular behaviors like fighting, sleeping, reading, laughing, texting, and unknown activities. Teachers may easily evaluate the data and decide on their teaching tactics thanks to the feedback's clear and ordered presentation. A pie chart that shows the distribution of different behaviors seen during the video analysis is another aspect of the UI. Teachers can quickly spot patterns in student interactions by glancing at the pie chart, which is divided into segments that represent distinct behaviors. For example, teachers might identify possible areas of concern that may need quick attention if a sizable chunk of the chart shows behaviors like fighting or napping. By blending visual data representation with thorough feedback, the technology helps instructors to develop a more engaging and responsive classroom environment. All things considered, the output area improves the platform's usability by offering insightful information that can have a direct impact on student learning results.

## Classroom Video Analysis Results



### Classroom Engagement Suggestions

Classroom Engagement Suggestions: - Few hand-raising actions detected. Encourage more interaction. - Many students are leaning, indicating disengagement or tiredness. Consider a break or changing activities. - A significant number of students are texting. Encourage focus and participation. - More disengaged students detected. Try adding interactive elements.

Figure 5: Suggestion to user

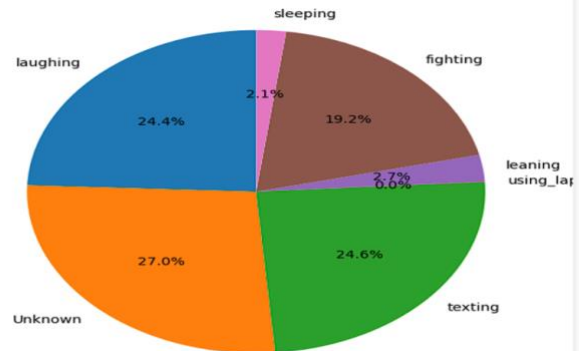


Figure 6: Overall Visualization

## System Workflow

The process of uploading a classroom video by a user, such as an administrator or instructor, initiates the data flow depicted in the diagram. After uploading, the video is routed to the server, which stores it in a special storage system (called D1: Video Storage). With the upload process is complete, the system may access and examine the video data. Pre-processing Video Frames is the initial step in accessing the stored video during the Processing and Analysis phase. In order to prepare the data for analysis, this phase involves removing frames from the video and using the appropriate pre-processing methods, such as shrinking, normalizing, or changing the frame quality. After being analyzed, these frames are fed into a CNN model, which uses them to identify different classroom behaviors like raising hands or nodding off. Following the detection of particular behaviors, the system classifies student activities using the CNN's output to determine each student's conduct in the classroom. The detected and classified actions are combined in the following step, Action Analysis. The system saves this activity data in a separate database (denoted as D2: Action Data Store). Using this information, a thorough Engagement Report reflecting the participation and engagement levels of the students is produced. Following this, the system provides suggestions for classroom improvement based on the students' behavior and engagement trends.

The technology presents the data in a Pie Chart during the last phase, Visualization and Feedback, giving a clear picture of student behavior and engagement. The teacher can monitor student participation and make well-informed decisions to improve classroom interaction and productivity by viewing this visual report. In order to enable teachers better understand student engagement and take appropriate action based on the data analysis, the complete system provides a real-time solution.

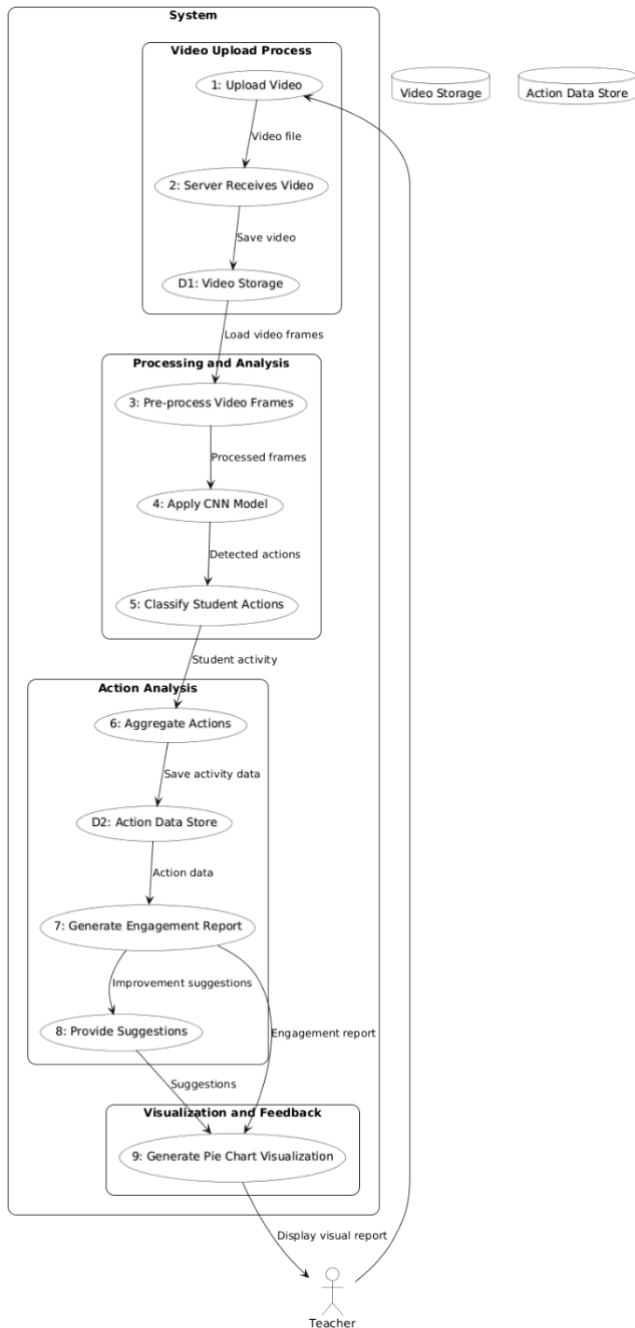


Figure 7: DFD of the Proposed System

#### IV. WORKING PRINCIPLE

##### Introduction to System Workflow

The workflow of the Student Behavior Analysis System begins with a user, generally a teacher or administrator, uploading a classroom video through an interface developed with Flask, HTML, CSS, and JavaScript. The video may be uploaded and delivered to the server with ease thanks to this interface, and it is kept there in a special Video Storage Database (D1). After that, the video is viewed and goes through a Pre-Processing stage in which each frame is taken out and ready for examination. Pre-processing tasks like as resizing, normalization, and frame quality correction are carried out in order to guarantee uniformity in a variety of classroom settings. After pre-processing, the frames are sent into a Convolutional Neural Network (CNN) model that has been trained to identify certain behaviors in the classroom, like raising hands, nodding off, and sleeping. The CNN examines every frame, identifying unique patterns of activity in real time.

The behavior detection process is followed by Action Detection and Classification, where the behaviors are categorized into pre-established groups by the system. Every student's activity is recorded with the associated timestamp, offering comprehensive data on student participation in the classroom. After that, this data is kept for later examination in an Action Data Store (D2). The system creates an extensive Engagement Report, which offers insights on student behavior during the class, based on this action data. Teachers can better grasp the dynamics of the classroom as a whole with the help of the report, which provides information on involvement frequency and disengagement incidents. In addition to this report, the system produces Suggestions for Classroom Improvement to assist the instructor in modifying their methods and boosting participation as needed. The Visualization and Feedback procedure happens at the last stage. Here, the system shows a Pie Chart that illustrates how various behaviors—like arguing, sleeping, reading, laughing, texting, or doing unknown things—were distributed during the session. Additionally, the technology gives teachers textual feedback that summarizes these behavioral tendencies, enabling them to keep a closer eye on student participation. The goal of this feedback is to provide instructors with timely, data-driven insights so they can make wise judgments. The entire system's workflow converts unprocessed video footage from classrooms into a wealth of information that teachers can use to enhance classroom management and instructional strategies and raise student engagement and learning objectives.

##### Algorithm

###### Step 1: Video Upload and Preprocessing

- Accept a user-submitted classroom video using the online interface.
- Save the video file that was uploaded to the server.
- To read and process the video frame by frame, use OpenCV.

###### Step 2: Preprocessing and Frame Extraction:

- Adjust the frame's dimensions to match the 224x224 pixel input size that the CNN model requires.
- Scale the pixel values to a range of 0 to 1 to normalize the pixel values.
- Transform the frame into a model-compatible format (such as RGB).

###### Step 3: Recognition and Categorization of Actions:

- Feed the trained CNN model with the preprocessed frame.
- Anticipating the action for the frame, the model will categorize it as one of the actions (e.g., listening, raising a hand, leaning, using a phone, etc.).
- Compile your predictions for every frame and note how frequently each action happens.

###### Step 4: Analysis and Postprocessing:

- Aggregate the predictions across all frames to estimate the frequency of each action.
- Over the course of the film, tally the instances of each classed activity.

###### Step 5: Generate Feedback and Suggestions:

- Determine the positive behaviors (such as raising your hands) based on the action numbers.
- Determine the detrimental behaviors (such as leaning and phone use).
- Based on the observed behaviors, make recommendations to the teacher on how to increase student engagement (e.g., encourage more hand-raising, limit distractions).



#### Step 6: Illustration:

- To see how the activities are distributed, create a pie chart.
- A detected action during the video is represented by each section of the pie chart.
- Put the analysis results and the pie chart on the webpage.

#### Step 7: Animation during Video Playback and Loading:

- Give the user the option to watch the analyzed video again.
- To demonstrate progress while the movie is being examined, use a loading animation.

#### Step 8: Results and Output Display:

- present the user with a pie chart, action counts, and feedback via the web interface.
- let the user download the analysis report.

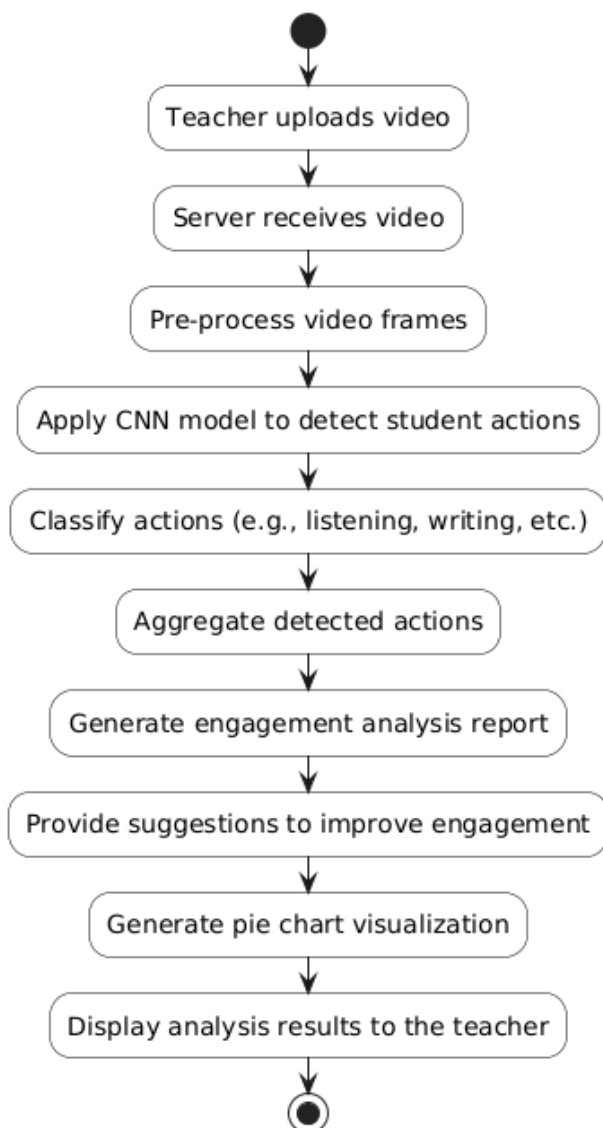


Figure 8: Algorithm of System

## V. RESULT AND CONCLUSION

### Result

The classroom footage is uploaded, processed by the system, which then produces a thorough analysis that highlights behaviors like raising hands, nodding off, and sleeping. A pie chart with the findings shows how the student actions that were seen in class are broken down clearly. It might indicate, for instance, that 25% of the time students raised their hands, 15% were not paying attention, and 50% were actively engaged in the lesson. In addition to the pie chart, the system provides comprehensive feedback that recommends ways to enhance instructional approaches by taking into account observed behaviors. Instructors can obtain a comprehensive report that includes behavior breakdowns and suggestions, enabling them to make data-driven modifications to improve student participation in the classroom.

### Conclusion

In conclusion, this project presents a significant advancement in the realm of classroom behavior analysis through the application of an optimized Fast R-CNN model. By effectively detecting critical student behaviors such as hand-raising and sleeping from video recordings, our proposed system empowers educators with valuable insights into student engagement and classroom dynamics. The utilization of advanced object detection techniques allows for accurate and real-time analysis, overcoming traditional limitations associated with manual observation methods. As a result, teachers can make informed decisions to enhance their instructional strategies and foster a more engaging learning environment. Furthermore, this system lays the groundwork for future research in automated classroom analysis, with potential applications extending beyond behavior detection to include broader aspects of student interaction and performance. Ultimately, this innovative approach not only benefits educators but also contributes to the overall improvement of educational outcomes, ensuring that every student has the opportunity to succeed.

## REFERENCES AND RESOURCES

- [1] Karpathy, A., & Fei-Fei, L. (2015). "Deep Visual-Semantic Alignments for Generating Image Descriptions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(4), 664-676.
- [2] Nguyen, T. D., & Kiyomoto, S. (2018). "Analyzing Student Behavior in Classrooms Using Convolutional Neural Networks," *Journal of Educational Technology & Society*, 21(2), 4-17.
- [3] Liu, W., et al. (2016). "SSD: Single Shot MultiBox Detector," *European Conference on Computer Vision (ECCV)*.
- [4] Donahue, J., et al. (2014). "Decaf: A Deep Convolutional Activation Feature for Generic Visual Recognition," *Proceedings of the 31st International Conference on Machine Learning (ICML)*.
- [5] Khan, A. F., et al. (2020). "A Comprehensive Review on Student Engagement Detection and Classification in Video Lectures: Current Trends and Future Directions," *Journal of Educational Computing Research*, 58(5), 1-30.
- [6] Li, Y., & Chen, Z. (2017). "Deep Learning-Based Object Detection for Visual Classroom Analysis," *Journal of Educational Technology & Society*, 20(4), 83-93.
- [7] Chen, T., et al. (2018). "Video-based Student Engagement Detection: A Review of the State-of-the-Art Techniques and Future Directions," *Education and Information Technologies*, 23(5), 2187-2208.
- [8] Xu, H., & Zhang, S. (2019). "Real-Time Video-Based Student Engagement Detection Using LSTM," *IEEE Transactions on Learning Technologies*, 12(2), 195-206.
- [9] Abdelhamed, A., & Kassem, R. (2020). "A Framework for Classroom Engagement Detection Using Neural Networks," *International Journal of Interactive Mobile Technologies*, 14(6), 150-160.
- [10] Bahl, S., & Gupta, V. (2020). "Analyzing Student Attention in Lecture Videos Using Deep Learning," *International Journal of Information and Education Technology*, 10(5), 415-421.
- [11] Cao, Y., et al. (2019). "Deep Learning for Classroom Behavior Analysis: A Survey," *Journal of Educational Technology & Society*, 22(3), 119-134.
- [12] Ma, L., et al. (2021). "Multi-Task Learning for Student Engagement Detection in Online Learning Environments," *Computers & Education*, 168, 104196.
- [13] Kaur, A., Mustafa, A., Mehta, L., & Dhall, A. (2018). "Prediction and Localization of Student Engagement in the Wild," *2018 Digital Image Computing: Techniques and Applications (DICTA)*, IEEE, 1-8.
- [14] Thomas, C., & Jayagopi, D.B. (2017). "Predicting Student Engagement in Classrooms Using Facial Behavioral Cues," *Proceedings of the 1st ACM SIGCHI International Workshop on Multimodal Interaction for Education*, 33-40.
- [15] Whitehill, J., Serpell, Z., Lin, Y.C., Foster, A., & Movellan, J.R. (2014). "The Faces of Engagement: Automatic Recognition of Student Engagement from Facial Expressions," *IEEE Transactions on Affective Computing*, 5(1), 86-98.