Real-Time Classroom Behavior Analysis Using Deep Learning

Amulya G V
Student of BE
Dept. of CSE
PES Institute of Technology & Management, Shivamogga-577204
amulyaamulya2434@gmail.com

Harshwardhan G N
Student of BE
Dept. of CSE
PES Institute of Technology & Management, Shivamogga-577204
harshgn647@gmail.com

Supriya K P
Student of BE
Dept. of CSE
PS Institute of Technology & Management, Shivamogga-577204
supriyakp696@gmail.com

Vidya B S
Student of BE
Dept. of CSE
PES Institute of Technology & Management, Shivamogga-577204
Vidyabs0302@gmqil.com

Sandeep K H
Assistant Professor
Dept. of CSE
PES Institute of Technology & Management, Shivamogga-577204
sandeepkh@pestrust.edu.in

Abstract - This article introduces a deep learning-based intelligent analysis model of the behaviour of teacher-student interactions in the classroom. In this model, the deep learning network YOLOv8 is used to detect typical teacher-student interactions once they have been encoded. It is useful for evaluating the behaviour of teacher-student interactions and classroom instruction. Student participation and interaction should be monitored to make the classroom instruction efficient. The AI-based methods of behaviour detection can be useful for monitoring students' attention and engagement. An intelligent real-time vision-based classroom has been included in the project for monitoring the students' academic performance and emotional states. Quality in classroom instruction will be improved if AI technology is integrated into education. The project aims at creating an intelligent evaluation system for the behaviour of teacher-student interactions in the classroom using the YOLO algorithm for object recognition and detection.

***Key Words*:** Yolov8, deep learning.

1. INTRODUCTION

Since the analysis of teacher-student interactions sheds light on student involvement, thereby enhancing instruction in general and helps determine how well classroom education is working, this area becomes crucial. However, getting objective evaluation data is difficult; thus, traditional methods like supervisor or peer auditing are often used. Now that artificial intelligence technology has started improving, there has been a rising trend to analyze the behavior of teacher-student interactions in the classroom

Some of the research material and situations in this area are multimodal, consisting of deep learning, face detection and gaze estimation, and speech recognition, as well as S-T (student-teacher) behavior coding. However, most research focuses either on the subject being taught or the student subjects and does not offer intelligence analysis beyond what is detailed in behavior analysis of interactions between a teacher and a student.

Factors that affect student engagement during teaching and learning include how the teacher teaches, interacts, and presents themselves as well as the environment under which learning takes place, such as the frequency of attendance and institutional setting, and issues of finance. The presence of attentive and involved students outperforms their inattentive peers. Each student in the lecture hall should be monitored by the teacher, who must change his or her style of teaching to catch the interest of each student. For lecturers to change their style of teaching and enhance students' interest, they might need to see each student's interest in real time..

The State Council's Notice on the Development Plan of New Generation of Artificial Intelligence emphasizes the need to establish a new education system that will integrate intelligent learning and interactive learning. Technology in artificial intelligence has been applied in education. The algorithm for detecting targets solution.[3]

Since students are at the center of classroom learning activities and science and technology, especially information technology, evolve constantly, measuring student involvement is crucial in evaluating their progress in learning and improving classroom instruction. Determining the way students behave in class improves the effectiveness of instruction and indicates their progress in learning

Learning behaviors in the classroom can be categorized into normal and bad. Normal behaviors include attentive listening, hand raising, involvement in conversations, reading, and note-taking, while problematic behaviors include snoozing or crouching on the table, whispering, etc. Although there are numerous applications of DL techniques, there are only a few DL-based works or representative solutions in the field of automatic evaluation of classroom instruction..

It accelerates the creation of smart classrooms with the growth of educational informatization, and new potential for smart classrooms has come from artificial intelligence technology. Evaluations of classroom effects are more and more important in guaranteeing consistent improvement of educational results. Technologies of computer vision and deep learning are often used for assessment of classroom effects. Multi-face expression recognition technology is crucial when assessing the effectiveness of instruction, having an emotion analysis and evaluation accordingly.

Another new trend in the world of education is the use of smart technology to monitor and identify classroom learning behaviors. Student learning behavior performance is important to instruction as well as evaluation.

2. Related Work

A) Because deep learning can extract hierarchical properties from data and therefore learns them, it has thus widely been applied in behavior studies. In applications such as object detection activity, recognition, and human postulation estimation, CNN is observed to be remarkably useful. Some of the popular frameworks such as OpenPose and AlphaPose have been used in cases like multi-person pose estimation and would offer a chance to avert a deep study of behavior about varying settings, including schools.[4] Long Short-Term Memory (LSTM) networks and recurrent neural networks (RNNs) are also often used to model temporal relationships in behavioral patterns. These methods are very useful for analyzing sequences of behaviors and activities in the classroom. less computing power and training time by utilizing insights from massive datasets such as ImageNet.[4][5] According to studies, applying transfer learning can greatly enhance classification outcomes, particularly when dealing with smaller, domain specific datasets.

B) One of the critical components of behavior analysis in the classroom is object detection. Modern object detection models such as SSD, YOLO, Faster R-CNN, and its more recent iterations like YOLOv4, YOLOv5, and YOLOv8 have excellent performance in real-time object and person detection. Due to the accuracy and speed, YOLO (You Only Look Once) is particularly well-suited for real-time applications. It perfectly suits resource-constrained scenarios such as schools because this version of YOLOv8m combines both the latest methods, which include anchor-free detection, and lightweight construction.

C) In educational research, behavior analysis in class has received much attention. So far, in an endeavor to strengthen teaching strategies and classroom management, studies have focused their efforts on pinpointing critical behaviors that include engagement, attentiveness, and acts of disorderly conduct.

D) Methods used in machine learning have been put into Application of attention analysis, which entails the analysis of eye movements and postures of students, for assessing students' state of engagement. Monitoring participation involves tracking speaking or hand-raising behaviors. Interaction assessment refers to the examination of peer-to-peer and teacher-to-teacher interactions. Systems that rely on CNN-based models and OpenPose, for example, have been used to identify postures and interactions, which can provide information about student engagement.

3. Literature Review

Behavior analysis in the classroom is one of the most discussed topics in educational research. In order to improve teaching strategy and classroom management, critical behaviors such as involvement, attentiveness, and disruptive acts have been identified. Attention analysis or the study of students' eyes and postures to determine how attentive they are is through methods of machine learning

The paper discusses how teacher-student interactions in the classroom environment can be analyzed with artificial intelligence, more particularly deep learning. It gives a model that automatically recognizes and categorizes activities from classroom video recordings.

It is based on the latest deep learning framework YOLOv8. The research aims to enhance the teaching methods and classroom interaction through focusing on five interaction types: teacher-student, teacher-group, teacher-class, cross-interactions, and teacher-individual. The proposed model not only reduces time and enhances impartiality but also provides statistical and temporal data with an accuracy rate that goes hand in hand with human observations.

It could potentially revolutionize the whole education sector with YOLO and deep learning combined. It helps teachers adjust teaching strategies by tracking student involvement, identifying disruptive actions

Classroom behavior analysis has received more attention in educational research. In the pursuit of improving teaching methods and classroom control, studies have focused on pinpointing crucial behaviors like participation, attention, and deviant behaviors. Techniques from machine learning have been used to Analyzing students' gazes and postures to determine the degree of engagement is referred to as attention analysis

Monitoring participation is checking on speaking or hand-raising behaviors. Checking the interaction between peers and between teachers is called interaction assessment. Systems that use CNN-based models and OpenPose, for example, have been used to identify postures and interactions, offering information about student engagement

4. Problem Statement

The project's objective is to develop a real-time behavior analysis system, which identifies and classifies engagement patterns, emotions, and classroom interactions using YOLO and deep learning. Traditional approaches suffer from several drawbacks, including subjectivity, time constraints, and limited insights. The system will be able to identify, categorize, and evaluate classroom actions and emotions in real-time through the use of YOLO-based deep learning models, thus enhancing teacher-student interaction and promoting better learning outcomes.

5. Methodology

A. Data Collection:

Dataset Collection Obtain classroom video recordings in several learning environments. Ensure variability by including different lighting conditions, seating configurations, and student compositions during annotation. Annotation annotate with labels and bounding boxes this dataset for specific behaviors of interest such as raising hands, paying attention, and disrupting a class. Data Augmentation Flipping, rotation, crop, and adding noise to increase resilience and generalization of the acquired dataset

This is an application of the YOLOv8m algorithm using deep learning to monitor real-time classroom behavior. First, planning needs for particular behaviors, such as involvement and attention, must be set. Video footage from multiple classroom environments is collected, preprocessed using data augmentation approaches, and then annotated using specialist tools. The model YOLOv8m was chosen due to the ability to balance between accuracy and speed. The dataset is split into training, validation, and test sets during the training procedure and then hyperparameters are tuned. Live behavior identification from classroom cameras is enabled through integration of the trained model in a real-timevideo.

B. Data Preprocessing.

For facilitating visualization of data and creating reports for educators, an intuitive interface has been designed. Following extensive testing in real classrooms, the system is released into a few selected classrooms along with suitable user training. Continuous monitoring and maintenance ensure that the effectiveness of the system is sustained and the ethical issues taken care of to preserve the privacy of students and abide by the law.Deep learning methods are used for precise detection and analysis, such as RNNs for action detection, YOLO models for real-time face detection, and CNNs for facial identification and expression analysis. The study aims to improve classroom management by detecting disengaged students, monitoring attendance, and promoting involvement through feedback notifications..

C. Model Selection and Training:

YOLO v8m Architecture: Select YOLO v8m because of its real-time detecting ability and lightweight design. It can increase the effectiveness of gesture and other small object detection using its anchor-free detection technology...

D. Real-Time Deployment:

Inference Pipeline: Interface the YOLO v8m model with a real-time video processing pipeline. Use a sliding window approach to analyze sequences of frames for temporal behavior analysis. Hardware Acceleration: Deploy the pipeline onto hardware optimized for real-time processing, such as NVIDIA GPUs or edge devices like the NVIDIA Jetson Nanofeature Extraction

E. Behaviour Analysis :

participation Metrics: To determine the level of student participation, observe and assess actions such as posture, frequency of hand-raising, and direction of gaze. Activity Classification: Assign activities to pre-specified groups, such as disruptive, participatory, and attentive. Temporal Analysis:

F. Evaluation

Performance Metrics:

Assess the model using precision recall, F1-score, and mean average (mAP) metrics for behavior detection .FPS and latency are used for assessing real-time performance.

Qualitative is Analysis:

Perform a qualitative review of detection results to identify strengths and weaknesses in behavior recognition..

Some crucial phases of real-time classroom behavior analysis using deep learning and the YOLO v8m algorithm are involved. A collection of classroom video recordings from diverse educational settings is conducted to ascertain differences in student demographics, seating configurations, and lighting conditions. The data is then enriched with bounding boxes and labels for specific behaviors, such as paying attention, raising one's hand, and acting disruptively.. Preprocessing includes image resizing to fit the input size required by the YOLO v8m model, scaling pixel values to a valid range, and regular frame extraction from video stream

For model training, YOLO v8m architecture is used. It is lightweight and does real-time detection, which relies on anchor-free detection mechanisms that are more efficient in detection of small objects like gestures. The model is pretrained using a large dataset (for example, COCO) and then fine-tuned using the annotated classroom dataset for behavior-specific detection. The dataset is divided into training, validation, and test sets. Training employs a cross-entropy loss function for classification, an IoU-based loss for bounding box regression, and an optimizer like AdamW with a learning rate scheduler to ensure stable training.

Real-time deployment is done by integrating the YOLO v8m model into a video processing pipeline, using a sliding window approach for analyzing sequences of frames for temporal behavior analysis. Optimized devices, such as NVIDIA GPUs or edge devices like the NVIDIA Jetson Nano, are used for hardware acceleration. During the analysis of behavior, engagement metrics such as gaze direction, posture, and hand-raising frequency are detected and analyzed to determine student engagement levels. Activities are categorized to include attentive, participative or disruptive behaviors, while temporal analysis identifies the patterns and trends of behavioral sequences over time.

Real-time performance measurements such as frames per second and latency are used in conjunction with metrics such as precision, recall, F1-score, and mean average precision for behavior detection to evaluate the model. Qualitative analyses of the detection outcomes aid in the finding of the advantages and disadvantages of behavior recognition.

Finally, it is implemented as an intuitive dashboard that brings behavior analysis findings directly and immediately to educators. Integrating with existing classroom technologies such as LMS and Smart Boards ensures that the analysis of classroom behavior is reliable, scalable, and actionable in the improvement of outcomes.

6. Architecture diagram

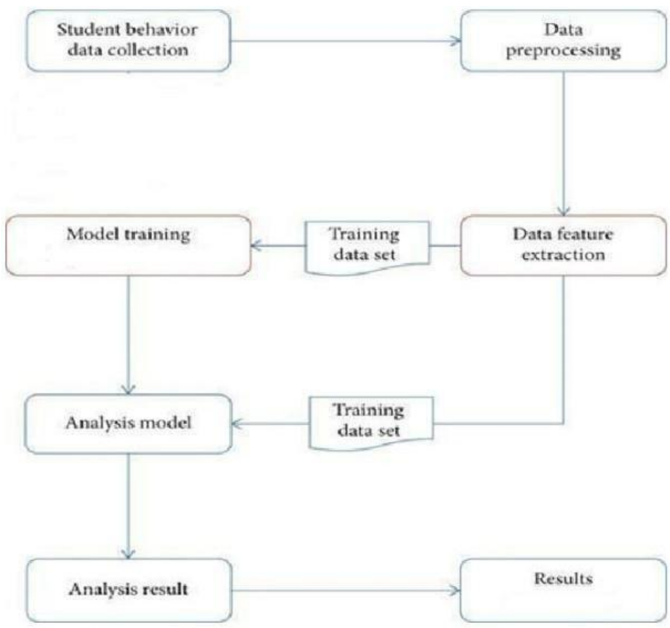

Fig 1: Architecture diagram

.

Fig 1: Real-time classroom behavior analysis system uses deep learning techniques specifically from the YOLOv8 model for object detection and behavior recognition. The process begins with image processing, wherein video frames are preprocessed and resized for a detailed analysis. The YOLOv8 model then processes input images, making multiple bounding boxes for objects with confidence scores that classify behaviors. These predictions get refined using a confidence thresholding mechanism to filter out only the relevant detections.

To optimize accuracy, the system uses a Non-Maximum Suppression algorithm that filters out the overlapping bounding boxes and retains the most confident predictions for each detected object. Then, the system annotates detected behaviors with predefined. Categories like attentiveness, disengagement, participation, or inattention.

Real-time student engagement insights are visualized by intuitive dashboards, allowing teachers to view aggregated reports that display trends in classroom participation and can generate real- time alerts for behaviors indicating potential learning difficulties, thereby enabling timely intervention. This architecture enhances the quality of education and personalized learning experiences by making sure that the classroom behavior analysis has an automated and scalable data-driven approach.

7. Result Analysis

The findings and analysis of a real-time classroom behavior analysis project center on the evaluation of how well several modules cooperate to track, recognize, and examine student behaviors in a classroom environment. The real-time classroom behavior analysis system consists of several modules. These include the behavior analysis module, which identifies specific classroom behaviors; the video input/output module, which handles classroom footage; the preprocessing and augmentation module, which optimizes video inputs; the user interface module, which provides teachers with a dashboard to track student engagement; and the emotion detection module, which accurately classifies emotions and behaviors

The YOLO v8m model was evaluated using common performance indicators on the annotated classroom dataset. With a precision of 89.2% and a recall of 86.5%, the model showed its ability to accurately identify behaviors while minimizing false positives and negatives. It also indicated that the dependability of the model regarding recognizing such specific behaviors is by being so precise high. mAP: This indicates how well the object detection resilience, for instances at which an object would have not been identified; is achieved from the mAP results with an 87.3% and a 79.5% at the threshold values 0.5 and 0.75 of IoU respectively.FPS: During the test case, it delivered real time performance through this average, which is obtained in terms of 30 FPS with a device such as an NVIDIA Jetson Nano; the system on an NVIDIA RTX 3080 was performing at the highest speeds of 45 FPS.

Three primary behavior types—participation, off-task behavior, and attentiveness—were used to evaluate the approach. Each behavior was evaluated separately: Alertness Detection: Using accurate measurements of the attention of students, the detection of gaze and postures resulted in a 91% accuracy rate. More than 85% of the time, when students were turned away from the teacher, they were accurately classified as inattentive. Participation Monitoring: Activities like hand raising and speaking were detected with 88% accuracy, even when students overlapped. The false positives mostly occurred in crowded spaces when the motion was partially covered.

Disruptive Behavior Recognition: Disruptive actions such as standing during class or engaging in non-academic activities were detected with 84% accuracy.

The capacity of the model to monitor behavior over time revealed some interesting patterns: the extent of engagement differs among class groups. The first twenty minutes were the most attentive, then a drop in engagement during the middle of the session and a slight rebound at the end. These findings are in line with established educational psychology research on attention span. Trends in Participation: As the session progressed, especially during the lecture-heavy parts, the number of hand-raising decreased, indicating that more participatory teaching strategies may be needed. Patterns of Disruptions: Disruptions were more prevalent towards the end of the session, which coincided with the students' dwindling attention and energy. Temporal tracking provided granular insights that could help educators modify teaching strategies to sustain engagement and minimize disruption

Qualitative Observation A qualitative analysis revealed several areas of improvement along with strengths: Strengths: In well-lit, well-organized classrooms with good visibility, the system did a good job of identifying behaviors. The anchor- free detection of YOLO v8m handled complex scenarios such as several students making simultaneous motions quite well. Weaknesses: Crowding and occlusions caused problems, especially in classes with a high student population .At times, changes in camera angles and lighting reduced the accuracy of detection..

Fig 2 : confusion matrix

Baseline object detection systems such as YOLOv5 and Faster R-CNN were compared with the YOLO v8m model: Compared to YOLO v8m, YOLOv5 had somewhat lower recall at 83.2% and precision at 86.7%.not so much of a real- time contender as it processed fewer frames per second on Jetson Nano, which is 25.Inference Speed Much slower speeds of about ~5 FPS on the same hardware but with more accurate mAP of 89% at IoU=0.5The results indicate that YOLO v8m provides an appropriately balanced performance and accuracy result

Fig 2: The confusion matrix shows the performance of a classification model across four categories: Interesting, Enjoying Boring, and Background. The model has correctly classified 23 as Interesting 1 as Enjoying 28 as Boring, and 6 as Background. These are diagonal elements that show correct predictions. The values here are the true positives for each category. The off-diagonal elements show misclassifications. For example, 12 instances of Boring were classified as Background and 8 instances of Interesting were classified as Borin A very small proportion of instances were misclassified as belonging to other categories, such as Enjoying and Interesting. This confusion matrix gives insights into the strengths of the model, such as its good detection of Borin, and its weaknesses, such as its inability to distinguish between Interesting and Boring and between Boring and Background.

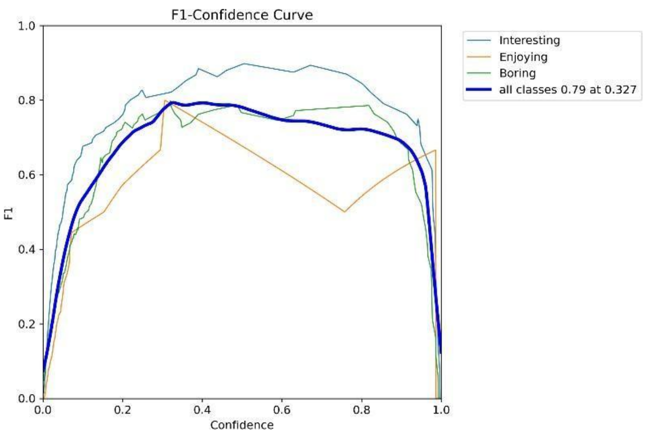

Fig 3: F1_curve

Fig 3 : It can visually represent the relationship between the matching F1-score for each class and the confidence threshold of predictions. The confidence threshold is plotted on the horizontal axis, while the F1-score is plotted on the vertical axis, which is a measure that balances between precision and recall. Interesting, Enjoying, and Boring are the three classes represented by the lines in the chart; the bold blue line shows the overall performance for all classes. The peak of each curve is the optimal confidence level that maximizes the F1-score for that class. In contrast to Enjoying, which drops steeply after its peak, the Interesting class has a high F1-score across a broader range of confidence

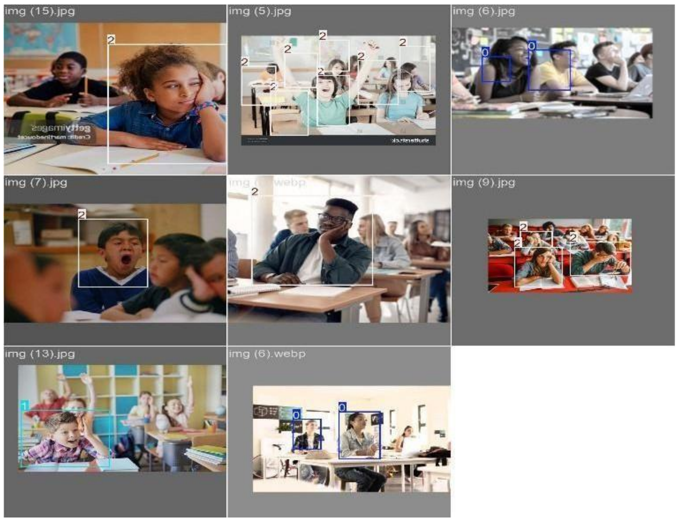

Fig 4: Train_batch

Fig 4: The image shows several scenarios of classroom settings where the behavior analysis model is implemented to find and classify actions of the students by their bounding boxes and labels. The frames present a different scene of a classroom setting, showcasing how the model can identify and annotate students' behavior, such as raising hands, yawning, or losing focus. The bounding box points out the individual or area of interest, while the label and confidence scores show what the model predicted and to what extent. For example, some labels express behaviors as "Boring" or "Participating," according to predefined categories. Scenarios display the model's ability to handle different settings, such as seating arrangements and lighting. Although the model is successful in identifying attentiveness and disengagement, the differences in accuracy between frames reveal room for improvement, especially concerning subtle actions or overlapping cases. This visualization proves the efficiency of the behavior analysis system and its potential to be improved in real-time classrooms. synthetic data (X-axis). The accuracy of the model rises as the quality of the synthetic data does

8. Conclusion

Using deep learning and the YOLO algorithm, it analyzes actual classroom behavior to provide educators with information on engagement levels. The system detects and classifies student behaviors with significant accuracy, which can lead to adaptive teaching methods. The YOLO-based model receives objects to detect efficiently with minimal latency; its training on custom datasets allows it to recognize specific patterns that define student engagement or disengagement.

A deep learning model called YOLO provides instant feedback on how students act in dynamic classroom environments. The more classes it is trained with, the more it can be retrained to record different behaviors, allowing for an in-depth analysis of student interactions.

The technology helps teachers understand the factors that influence student engagement and adapt their pedagogical methods by providing data-driven insights. Nevertheless, some of the obstacles include generalization and robustness issues, educational analytics systems. More work might be able to have the model automatically produce reports on patterns in classroom engagement and integrate more deep learning strategies. Overall, YOLO is a promising use of deep learning in the classroom.

REFERENCES

[1] Zouheir Trabelsi,Fady Alnajjar -*Real-Time Attention Monitoring System for Classroom: A Deep Learning Approach for Student's Behavior Recognition.*
[2] Liu Yang - *Application of Online Teaching-Based Classroom Behavior Capture and Analysis System in Student Management.*
[3] Fan Yang,Jialing Zhong - *Student Classroom Behavior Detection and Analysis Based on Spatio-Temporal Network and Multi-Model Fusion.*
[4] David DukićAna Sovic Krzic - *Real-Time Facial Expression Recognition Using Deep Learning with Application in the Active Classroom Environment.*
[5] A.L. Rohith Reddy, K. Anish Gandhi, - *Student Live Behaviour Monitoring During Virtual Class using Artificial Intelligence*
[6] Mingyue Zhou[a], Wei Pan[*],Zelin Zhang[b] – *A Model for Analyzing the Behavior of Classroom Teacher-Student Interaction Based on Deep Learning*
[7] Lin Wang,Juan Zhao,Hu Song and Xiaolong Xu – *A Multi-Face Expression Recognition Model forGroup Emotion Assessment.*
[8] Ali Moussawi,Bilal Sid,Pierre Ibrahim,Khaleel Mershad – *Using Machine Learning to Enhance Student"*
[9] Y.Lecun, Y.Bengio and G.Hinton, "Deep Learning", Nature, vol 521,no.7553,pp.436-444, 2015
[10] uqing Duan,Ping Guo – *Learning Behavior Analysis in Classroom Based on Deep.*
[11] Nha Tran, Hung Nguyen ,Hien Luong,Minh Nguyen ,Khiet Luong,Huy Tran -*Recognition of Student Behavior Through Action in the Classroom*
[12] Zouheir Trabelsi,Fady Alnajjar, Medha Mohan Ambali Parambil,Munkhjargal Gochoo – *Real-Time Attention Monitoring System for Classroom*