

Comparative Study of Recognition for Student Attention Analysis: YOLO-Based Face Detection in Classroom Environments

Thursday, 20 March 2025 15:15 (25 minutes)

The educational needs in the future classroom need to focus on a combination of student engagement in learning, inquiry-based approaches, curiosity, imagination, and design thinking. Smart classrooms leverage the advancements in Internet of Things to create intelligent, interconnected learning environments that enhance the quality of life and educational outcomes of students. With advancements in image recognition Face detection and recognition are pivotal in security and tracking, ensuring swift and effective identification and management of personal data. Currently, researchers are collecting data on student attention retention in learning environments to help facilitate teachers and students in a holistic learning environment. Furthermore, harnessing accurate visual tracking data to accurately monitor multiple students in real-time, with continued advancements in image recognition and deep learning frameworks. There is a challenge in choosing the most effective face-detection algorithms for specific use cases. We propose a system that first detects students in a classroom using object detection of a person with a custom-trained model. Once a student is identified, the video feed is cropped to the bounding box of the detected individual, and face detection is applied to this region. With the face detection step gaze estimation is performed to infer attention retention during lectures. Face detection is a critical technology in understanding and measuring student engagement through facial recognition and gaze analysis. This study focuses on evaluating the performance of three versions of the YOLO (You Only Look Once) face detection algorithm—YOLOv5, YOLOv8, and YOLOv11—in a real-world classroom setting. The research investigates their efficacy in face recognition among Japanese middle school students during lectures. The setup involved two cameras: a 3840 × 2160-pixel resolution at 30 frames per second camera positioned (from the student's point of view) at the front-top-left of the classroom and another mounted at the top-right-rear, facing the lecturer. These cameras provided complementary perspectives but introduced challenges, particularly in detecting and analyzing the faces of the farthest students due to reduced image sharpness and resolution constraints.

Given the constraints of the classroom environment, including limitations on adding additional cameras to avoid distractions for both students and the lecturer, the study emphasizes the need for robust face detection algorithms capable of delivering reliable results under suboptimal imaging conditions. The algorithms were evaluated across several dimensions: detection accuracy, speed, robustness to environmental variations (e.g., lighting and occlusion), and their ability to maintain performance on distant and partially visible faces.

The model to recognize the student face is trained on a 64 bit Windows operating system with 32 GB RAM and Intel(R) Core(TM) i9-9900K CPU@3.60GHz with a Nvidia GeForce RTX 3080 10 GB graphics processing unit (GPU). All YOLO variant will have the same dataset and key points used in the training.

The implication of this study extends to the broader application of face detection in education settings and to analyze gaze and student attention to aid facilitators and teachers in adjusting content delivery methods in real-time. By identifying the strengths and limitations of these algorithms, educators and scientists can make informed decisions about implementing AI tools to monitor student engagement. Future work will further add body position detection to provide a more comprehensive understanding of classroom attention dynamics, as well as refine the system for improved performance under similar constraints.

Primary authors: SANG-NGENCHAI, Apirak (Kanazawa Institute of Technology, Japan); OGAWA, Hayato (Kanazawa Institute of Technology); NAKAZAWA, Minoru (Kanazawa Institute of Technology)

Presenter: OGAWA, Hayato (Kanazawa Institute of Technology)

Session Classification: Artificial Intelligence (AI) - I

Track Classification: Track 10: Artificial Intelligence (AI)