

Research Article

Optimization of Real-Time Student Face Recognition Attendance Using the YOLO v10 Algorithm

Mesra Betty Yel^{1*}, Elviwani², Nandang Sutisna³, Ziyad Fernanda Syams⁴

- ¹ Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika Jakarta, Indonesia; email: mesrabettyyel@stikomcki.ac.id
 - ² Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika Jakarta, Indonesia; email: elviwani@stikomcki.ac.id
 - ³ Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika Jakarta, Indonesia; email: nandangsutisna@stikomcki.ac.id
 - ⁴ Sekolah Tinggi Ilmu Komputer Cipta Karya Informatika Jakarta, Indonesia
- * Corresponding Author: mesrabettyyel@stikomcki.ac.id

Abstract: This research is motivated by the problems in manual attendance systems at schools, which remain vulnerable to fraud, time-consuming, and inefficient. The expected solution is to develop an automated attendance system based on face recognition that can operate in realtime with high accuracy. The research object is vocational high school students, with the applied method implementing the YOLO v10 algorithm for face detection, followed by the face_recognition library for identification. The instruments used include an Imou CCTV camera as the input device, a mid-range laptop as the hardware platform, and Python with SQLite as the software environment for data processing and attendance storage. The results show that the developed system achieved an average face detection accuracy of 96% under normal lighting and 91% under low lighting, with an average processing speed of 27 FPS. The implementation of an anti-duplication feature also ensured data validity by allowing each student to be recorded only once per day. In conclusion, the use of YOLO v10 in face-based attendance proved to be effective, efficient, and capable of reducing fraud. The implication of this study is that the system can be applied in both Islamic boarding schools and general schools as a modernization of attendance systems, with a recommendation for further development through web-based application and cloud database integration.

Keywords: Automated Attendance; Computer Vision; Face Detection; Realtime; YOLO v10.

Received: 20 August 2025
Revised: 12 September 2025
Accepted: 14 October 2025
Published: 30 October 2025
Curr. Ver.: 30 October 2025



Copyright: © 2025 by the authors.
Submitted for possible open
access publication under the
terms and conditions of the
Creative Commons Attribution
(CC BY SA) license
(<https://creativecommons.org/licenses/by-sa/4.0/>)

1. Introduction

Student attendance is an important indicator in supporting learning achievement and student discipline. Ideally, an attendance system within a school environment should provide accurate, real-time, and tamper-resistant data as a foundation for managerial and academic decision-making (Hidayat et al., 2025). In the context of a boarding-based school such as Skill Village Islamic School Vocational High School (SMK Skill Village Islamic School), Jonggol, Bogor, student attendance not only reflects participation in learning activities but also serves as a benchmark for monitoring students' daily lives in the dormitory.

However, the current attendance system is still largely manual, relying on paper-based attendance sheets or signature records. This approach creates opportunities for manipulation, delays in data recapitulation, and inefficiencies in staff workload (Budiman, 2021; Hidayat et al., 2025). Furthermore, such conditions do not support the school's vision of becoming an educational institution that integrates technology into its operational and learning processes.

A notable social phenomenon is the relatively low utilization of digital technology in educational administration, despite the fact that today's students are highly familiar with the digital environment. This disparity creates an urgent need to develop an automated, intelligent, and real-time attendance system that can enhance the professionalism and effectiveness of school management (Benedict, 2022; Hidayat et al., 2025).

This research is significant because it not only contributes to the application of advanced technologies, such as the YOLOv10 algorithm in computer vision, but also addresses the need for administrative efficiency and accurate attendance data management in modern technology-based educational environments (Bellout & others, 2024; Nurmaini & others, 2021).

This study offers several novel contributions compared to previous face-recognition-based attendance system research, including:

a. Implementation of the YOLOv10 Algorithm

Most previous studies have utilized YOLOv5, YOLOv7, or YOLOv8 as the foundation for face detection and attendance systems (Nurhaliza, 2023; Ramadhan, 2021; Syarifuddin, 2022). This research employs YOLOv10, the latest version of the YOLO family, which provides significant improvements in inference speed and detection accuracy, making it more suitable for real-time student attendance applications (Bellout & others, 2024; A. Wang & others, 2024).

b. Integration with Classroom CCTV Cameras

While previous face-recognition attendance systems have predominantly relied on laptop cameras or standard webcams (Ramadhan, 2021; Syarifuddin, 2022), this study integrates the system with Imou CCTV cameras. This innovation provides greater practical relevance, as CCTV systems are better suited to classroom environments that require wider coverage, centralized monitoring, and permanent installation (Tawakkal, 2023).

c. Optimization through an Anti-Duplication Attendance Feature

The proposed system not only performs face detection and recognition but also incorporates an anti-duplication mechanism that ensures each student is recorded only once per day. This feature is rarely found in previous studies and significantly enhances the validity and reliability of attendance records (Budiman, 2021; Nurhaliza, 2023).

d. Testing under Dynamic Classroom Conditions

Another novel aspect of this research is the evaluation of the system in real classroom environments, including variations in lighting conditions, student movements, and changes in facial appearance. This approach provides a more comprehensive assessment of system performance in real-world implementations rather than relying solely on static datasets (Aulia, 2023; Hayati et al., 2023).

Through these innovations, this research focuses not only on the development of an automated attendance system but also on improving its reliability and practical relevance for direct implementation in school environments (Hidayat et al., 2025; Tawakkal, 2023).

2. Literature Review

Attendance Optimization

Attendance optimization is the process of improving the performance of attendance systems to make them more efficient, accurate, and adaptive to user needs. The primary objectives of this optimization are to minimize recording errors, reduce dependence on manual methods, and accelerate the attendance verification process. Optimization may involve various approaches, ranging from algorithm improvements and facial recognition technology integration to the simplification of system interfaces (Hidayat et al., 2025; Nurhaliza, 2023).

In the context of information systems, attendance optimization plays a crucial role in establishing transparent attendance management that can be monitored in real time and generate valid data to support decision-making processes (Hidayat et al., 2025; Santosa, 2022).

Attendance systems optimized through real-time digitalization can reduce operational time and minimize fraudulent activities while enabling direct attendance monitoring through web- and mobile-based interfaces. This demonstrates that optimization not only improves operational speed but also strengthens the overall accountability of attendance management systems (Cyganek, 2013; Hidayat et al., 2025).

Face Identification

Face identification is a technology used to recognize individuals by utilizing the unique characteristics of the human face. This technology operates through several main stages, including image acquisition, image preprocessing, feature extraction, comparison with stored data, and final classification (Fetik et al., 2022). Various approaches have been applied in face identification systems, ranging from geometric-based methods and statistical techniques to machine learning algorithms designed to improve identification accuracy (Li & Jain, 2005; Prasetya, 2023; Zhao et al., 2003).

Other studies describe facial recognition technology as an advanced analytical system capable of reading and distinguishing facial characteristics such as the shape of the eyes, nose, and mouth, and then matching them with reference data stored in a database. This technology has been widely applied in security systems, user authentication, and law enforcement applications (Lestari, 2023).

Nevertheless, the implementation of facial recognition systems still faces technical challenges, including varying lighting conditions, differences in facial viewing angles, and concerns related to privacy and ethical issues (Li & Jain, 2005; Zhao et al., 2003).

In the context of this study, facial recognition technology is implemented using a real-time face recognition approach based on the YOLOv10 model. The system is designed to perform image acquisition, facial feature extraction, and data matching quickly and simultaneously, even when multiple individuals appear within a single frame. The model is also optimized to detect faces from various positions and viewing angles, thereby maintaining high accuracy under real-world classroom conditions (Ardiansyah, 2024; Maulana, 2024).

Real-Time Systems

A real-time system is a system capable of providing immediate responses to events within a predetermined and predictable time frame (Manapa & others, 2023; A. Wang & others, 2024). In other words, the system must be able to respond to every event within a predictable period. Failure to respond within the required timeframe may cause serious disruptions to its primary functions; therefore, timeliness is a critical aspect of real-time systems (J. Wang, 2017).

In this study, the student attendance detection system was designed to operate in real time. Each video frame captured by the CCTV camera is directly transmitted to the automated attendance application. The frame is then processed to detect and recognize the faces of students captured by the camera. After processing, the recognition results are displayed immediately on a web-based monitoring interface, allowing teachers or administrators to monitor attendance in real time (Santosa, 2022; Tawakkal, 2023)

If the processing time is too slow, the system will be unable to display video streams in real time, thereby hindering the attendance monitoring process. Therefore, data processing speed is a crucial factor in determining the successful implementation of automatic attendance technology based on facial recognition in school environments (A. Wang & others, 2024).

YOLO (You Only Look Once)

YOLO (You Only Look Once) is one of the leading paradigms in real-time object detection, effectively balancing computational cost and detection performance (Zhao et al., 2003). YOLO is a CNN (Convolutional Neural Network)-based object detector widely recognized for its ability to perform object detection quickly and accurately through a one-stage detection approach, unlike two-stage approaches such as R-CNN (He et al., 2017; Nurmaini & others, 2021).

YOLO operates by dividing an image into grids and directly predicting bounding boxes and object classes in a single inference process. However, previous versions of YOLO relied on Non-Maximum Suppression (NMS) for post-processing, which could increase inference latency and hinder end-to-end deployment (A. Wang & others, 2024). To address this issue, YOLOv10 introduces an NMS-free training strategy known as "consistent dual assignments," which combines the advantages of one-to-many assignments for training and one-to-one assignments for inference, thereby eliminating the need for NMS.

In addition, YOLOv10 optimizes the model architecture holistically through an "efficiency-accuracy driven model design" approach. These optimizations include reducing computational redundancy, utilizing large-kernel convolution, and implementing partial self-attention modules to improve both performance and efficiency (Zhao et al., 2003).

Case Study

SMK Skill Village Islamic School is an Islamic boarding-based vocational high school located in Sukasirna Village, Jonggol District, Bogor Regency, West Java, Indonesia. The school adopts a vocational education model integrated with an Islamic boarding school

system, combining the national curriculum, thematic Qur'anic learning, and the development of digital and entrepreneurial skills. The primary areas of competency development include information technology, construction, and digital agriculture. As a vocational educational institution that upholds integrity and professionalism based on Islamic values, SMK Skill Village Islamic School aims to produce outstanding graduates who are prepared to face global challenges while maintaining their identity as Muslims.

Learning activities are conducted intensively within a boarding school environment, supported by technology-based monitoring through CCTV cameras installed at various strategic locations, including classrooms and dormitory areas. During the 2024/2025 academic year, the school had 14 active students, all of whom resided in the dormitory and participated in full-time educational activities.

In this context, the need for a facial recognition-based digital attendance system becomes highly relevant, considering that students' daily activities and learning processes are continuously supervised by the institution. Therefore, the development of an automated attendance system using YOLOv10 technology and CCTV implementation represents an important step in the digital transformation of school administration, aligning with the characteristics and operational needs of SMK Skill Village Islamic School (Ardiansyah, 2024; Nurhaliza, 2023; Zhao et al., 2003).

3. Materials and Method

Research Design

This study employed a computer vision-based approach to develop a real-time student attendance system using the YOLOv10 algorithm for face detection. The research was conducted at Skill Village Islamic School Vocational High School (SMK Skill Village Islamic School), Jonggol, Bogor, utilizing classroom CCTV recordings as the primary data source. The proposed system was designed to automatically detect student faces and record attendance in real time.

Dataset Collection

The dataset used in this study was obtained directly from the school environment through classroom CCTV recordings and face image acquisition of students. A total of 14 students participated in the data collection process. Each student was photographed from four different facial orientations: frontal, upward, left-facing, and right-facing positions. Consequently, 56 facial images were collected and used as the dataset for model development.

In addition, classroom CCTV videos were recorded during active school hours for ten consecutive school days. The CCTV videos had a resolution of 1280×720 pixels and were sampled into image frames at five-second intervals. The dataset included variations in illumination, facial expressions, student movements, and camera perspectives to simulate real classroom conditions.

Data Preprocessing

Prior to model training, all images underwent preprocessing. Face regions were manually annotated using Roboflow by drawing bounding boxes around each detected face. The annotated images were then resized to 640×640 pixels to ensure input consistency during the training process.

To improve model generalization and robustness, data augmentation techniques were applied, including horizontal flipping, image rotation, and brightness adjustment. Subsequently, the dataset was divided into training, validation, and testing subsets with a ratio of 70%, 20%, and 10%, respectively.

YOLOv10 Model Training

The face detection model was developed using the YOLOv10 architecture implemented through the Ultralytics framework in Python with PyTorch support. YOLOv10 was selected due to its improved inference speed and detection accuracy, enabling efficient real-time object detection without requiring a Non-Maximum Suppression (NMS) stage.

The training process utilized the prepared dataset to learn facial features from the annotated images. Model optimization was performed through iterative training until satisfactory detection performance was achieved on the validation dataset.

System Implementation

The developed attendance system was implemented using Python programming language version 3.12.11 within the Anaconda environment. Several supporting libraries, including PyTorch, OpenCV, and Ultralytics YOLO, were utilized during system development.

The implementation workflow consisted of the following stages:

- 1) Dataset acquisition from student face images and CCTV recordings.
- 2) Data annotation and preprocessing.
- 3) YOLOv10 model training and validation.
- 4) Real-time face detection using classroom CCTV cameras.
- 5) Automatic attendance recording into the digital attendance system.
- 6) Prevention of duplicate attendance records through an attendance validation mechanism.

The final system was integrated with an Imou CCTV camera installed in the classroom to support continuous and centralized monitoring.

Model Evaluation

The performance of the proposed model was evaluated using the testing dataset and real-world classroom scenarios. Evaluation metrics included Precision, Recall, F1-Score, Accuracy, and Mean Average Precision (mAP). These metrics were calculated from the confusion matrix generated during the testing phase.

Precision measures the proportion of correctly detected faces among all detected faces, while Recall evaluates the model's ability to detect all actual faces. The F1-Score provides a balanced assessment of Precision and Recall, whereas mAP represents the overall detection performance of the model.

Furthermore, real-time system performance was assessed by measuring Frames Per Second (FPS) and detection latency during classroom implementation. Field testing was conducted by simulating student attendance activities in front of the CCTV camera to evaluate system effectiveness under practical conditions.

4. Results and Discussion

Implementation and Testing

Data Processing

This study utilized a facial image dataset collected from 14 students of the Business Boarding School at Skill Village Islamic School Vocational High School. The dataset consisted of 56 facial images captured under four pose variations: frontal, upward, left-facing, and right-facing. The dataset was divided into three subsets: 39 images for training, 11 images for validation, and 6 images for testing.

Dataset Annotation

The facial image annotation process was conducted using the Roboflow platform to enable the system to automatically recognize and distinguish student faces under various conditions. During this stage, each face in the image was annotated using a bounding box, which served as a reference for the model to learn essential facial features. This step is crucial in developing an automated attendance system based on facial recognition, as it helps the model understand specific visual patterns required for identification tasks.

Data Splitting

After the preprocessing stage was completed, the dataset was divided within Roboflow into three categories: 70% training data, 20% validation data, and 10% testing data, as illustrated in Figure 1.



Figure 1. Dataset Split Distribution.

The experimental results demonstrated that the YOLO v10-based face detection stage achieved an average accuracy of 96%. Furthermore, the integration of YOLO v10 with the face_recognition library yielded an average face recognition accuracy of 98%. In terms of processing speed, the system maintained an average performance of 27 FPS under normal lighting conditions and 23 FPS under dim lighting conditions. Distance-based testing indicated that the system could reliably detect faces up to a distance of 1.5 meters, while performance began to decline at a distance of 2 meters.

Face Detection Accuracy Test

Table 1. Face Detection Accuracy Results.

No	Student Name	Number of Tests	Correct Detections	Accuracy
1	Andini	5	5	100%
2	Aulia	5	5	100%
3	Abdul	5	5	100%
4	Almun	5	5	100%
5	Rafi	5	4	80%
6	Syahla	5	5	100%
7	Fachri	5	5	100%
8	Jojo	5	5	100%
9	Ziyad	5	4	80%
10	Jundi	5	5	100%
11	Raka	5	5	100%
12	Eka	5	5	100%
13	Zuhdan	5	4	80%
14	Aufa	5	5	100%
Average				96%

Face Recognition Accuracy Test

Table 2. Face Recognition Accuracy Results.

No	Student Name	Number of Tests	Correct Recognitions	Accuracy
1	Andini	5	5	100%
2	Aulia	5	5	100%
3	Abdul	5	5	100%
4	Almun	5	4	80%
5	Rafi	5	5	100%
6	Syahla	5	5	100%
7	Fachri	5	5	100%
8	Jojo	5	5	100%
9	Ziyad	5	4	80%
10	Jundi	5	5	100%
11	Raka	5	5	100%
12	Eka	5	5	100%
13	Zuhdan	5	5	100%
14	Aufa	5	5	100%
Average				98%

Processing Speed Test

Table 3. Processing Speed Results.

Condition	Average FPS
Normal Lighting	27 FPS
Dim Lighting	23 FPS

Detection Distance Test

Table 4. Detection Distance Results.

Distance	Detection Result
0.5 m	Detected
1.0 m	Detected
1.5 m	Detected
2.0 m	Not Detected

Optimization Analysis

The optimization of the real-time facial attendance system was achieved by combining object detection using the YOLO v10 algorithm with facial matching through the face_recognition library. Experiments were conducted on an MSI GF75 Thin laptop equipped with an Intel Core i7-10750H processor, 16 GB RAM, NVIDIA GTX 1650 Ti GPU, and a camera resolution of 640×480 pixels.

The experimental results showed that YOLO v10 maintained an average processing speed of 27 FPS, outperforming YOLOv8, which achieved 22 FPS, and the Haar Cascade method, which achieved 19 FPS. In terms of face detection accuracy, YOLO v10 achieved 96% accuracy, while YOLOv8 reached 92% and Haar Cascade achieved 85% under identical testing conditions.

Under low-light conditions, YOLO v10 maintained an accuracy rate of 91%, demonstrating greater robustness compared to YOLOv8 (86%) and Haar Cascade (72%). The improvement was particularly evident when processing tilted faces, partially occluded faces, or faces located more than one meter from the camera. Furthermore, the integration of the face_recognition fallback mechanism increased face recognition accuracy from 94% to 98%, especially in cases where YOLO v10 detected only partial facial regions.

Additionally, the implementation of an anti-duplication attendance mechanism ensured that each student was recorded only once per day, thereby improving attendance data validity and preventing duplicate records. These combined optimizations enhanced the reliability of the system and demonstrated its suitability for real-time attendance applications in educational environments.

Comparison of Face Detection Methods

Table 5. Performance Comparison of Face Detection Methods.

Method	Average FPS	Normal Lighting Accuracy (%)	Low-Light Accuracy (%)
Haar Cascade	19 FPS	85%	72%
YOLOv8	22 FPS	92%	86%
YOLO v10	27 FPS	96%	91%
YOLO v10 + Face Recognition	25 FPS	98%	94%

5. Conclusion

Based on the results of the research and implementation of the real-time student facial recognition attendance system using the YOLO v10 algorithm, it can be concluded that: a) The automated attendance system based on facial recognition was successfully developed by implementing the YOLO v10 algorithm for face detection and the face_recognition library for the identification process, which was subsequently integrated with an Imou CCTV camera as the classroom input device. The system is equipped with an SQLite database for attendance data storage and an automatic logging mechanism, enabling attendance recording to be conducted quickly, systematically, and with minimal manual intervention. b) The testing results indicate that the developed system possesses a high level of effectiveness and accuracy in detecting student attendance in real time. The average face detection accuracy reached 96% under normal lighting conditions and 91% under low-light conditions, with an average processing speed of 27 FPS. These findings demonstrate that the system is more efficient than conventional manual attendance methods, which require more time and are prone to fraudulent practices, as the system records attendance only once per student per day. c) Under dynamic classroom conditions, such as student movement, variations in lighting, or changes in facial appearance (e.g., wearing glasses or displaying different facial expressions), the system was still able to operate with relatively stable performance. However, under dim lighting conditions or suboptimal camera viewing angles, the accuracy tended to decrease, particularly for students whose enrollment data were less representative. Overall, the system demonstrated reliable performance and can be considered a dependable solution for supporting the modernization of technology-based attendance systems in educational environments.

To further improve and optimize the attendance system for future implementation, several recommendations can be proposed: a) Developing system integration with web-based or mobile applications so that attendance data can be accessed in real time by teachers, homeroom teachers, and parents. b) Utilizing cameras with higher resolution and better image

quality to enhance detection accuracy, particularly at longer distances or in environments with poor lighting conditions. c) Adding support for face recognition when users wear full-face masks or other accessories that partially obstruct facial features. d) Integrating the system with a cloud-based database for centralized data storage, thereby facilitating data management on a larger school scale. e) Implementing YOLO v10 optimization techniques, such as model pruning or quantization, to accelerate inference performance on devices with limited hardware specifications.

References

- Ardiansyah, M. D. (2024). *Penerapan model YOLO v10-S dalam sistem presensi berbasis face recognition secara real-time*.
- Aulia, C. (2023). *Evaluasi akurasi deteksi wajah menggunakan YOLOv6 dalam sistem presensi*.
- Bellout, A., & others. (2024). YOLO v10-enabled IoT robot car for accurate disease detection in strawberry cultivation. *ITM Web of Conferences*, 69, 1011. <https://doi.org/10.1051/itmconf/20246901011>
- Benedict, S. (2022). *Deep learning technologies for social impact*. IOP Publishing. <https://doi.org/10.1088/978-0-7503-4024-3>
- Budiman, A. (2021). *Sistem deteksi kehadiran otomatis menggunakan kamera dan YOLO*.
- Cyganek, B. (2013). *Object detection and recognition in digital images: Theory and practice*. Wiley.
- Ferik, Y., Octavianto, H., & Wahyu, H. (2022). *Deteksi wajah menggunakan algoritma Viola-Jones*.
- Hayati, N. J., Singasatia, D., & Muttaqin, M. R. (2023). Object tracking menggunakan algoritma YOLOv8 untuk menghitung kendaraan. *KOMPUTA Journal*, 12(2).
- He, R., Wu, X., Sun, Z., & Tan, T. (2017). Wasserstein CNN: Learning invariant features for NIR-VIS face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*. <https://doi.org/10.1109/JPROC.2017.170802412>
- Hidayat, W., Ba'a, F. A., Prasetio, O., & Haryono, W. (2025). Perancangan sistem aplikasi absensi real time untuk meningkatkan efisiensi manajemen kehadiran PT. Asia Sinergi Solusindo. *Switch: Jurnal Sains Dan Teknologi Informatika*, 3(1), 37–48. <https://doi.org/10.62951/switch.v3i1.322>
- Lestari, A. (2023). *Penerapan YOLOv8 untuk identifikasi wajah siswa secara real-time*.
- Li, S. Z., & Jain, A. K. (Eds.). (2005). *Handbook of face recognition*. Springer.
- Manapa, E. S., & others. (2023). *Review of real-time operating system research in Indonesia*.
- Maulana, F. (2024). *Face recognition real-time dalam sistem keamanan kelas menggunakan YOLO v10*.
- Nurhaliza, G. (2023). *Sistem presensi digital siswa berbasis deep learning dengan YOLOv8*.
- Nurmaini, S., & others. (2021). *Pengenalan deep learning dan implementasinya*. Unsri Press.
- Prasetya, E. (2023). *Deteksi wajah otomatis untuk sistem absensi mahasiswa berbasis YOLOv8*.
- Ramadhan, D. (2021). *Penerapan algoritma YOLOv7 untuk pencatatan kehadiran siswa di sekolah dasar*.
- Santosa, B. (2022). Face recognition dalam sistem absensi real-time berbasis web. *Jurnal Teknologi Dan Sistem Komputer*, 10(2), 150–158. <https://doi.org/10.14710/jtsiskom.2022.150-158>
- Syarifuddin, A. (2022). *Implementasi sistem presensi mahasiswa menggunakan face recognition dengan YOLOv5*.
- Tawakkal, M. A. (2023). *Rancang bangun aplikasi penghitungan okupansi manusia dengan CCTV menggunakan object detection model YOLOv8*.

Wang, A., & others. (2024). *YOLO v10: Real-time end-to-end object detection*.

Wang, J. (2017). *Real-time embedded systems*. John Wiley & Sons.

Zhao, W., Chellappa, R., Phillips, P. J., & Rosenfeld, A. (2003). Face recognition: A literature survey. *ACM Computing Surveys*, 35(4), 399–458. <https://doi.org/10.1145/954339.954342>