

# Intelligent Attendance Monitoring System Based on GAIT Features using 3D CNN

**B. Padmavathy<sup>1</sup>, Bhoomika Mohan<sup>2</sup>, Ananya. R<sup>3</sup>, Adithi Shridhara<sup>4</sup>, Ananya Kanthraj<sup>5</sup>**

<sup>1</sup> Assistant Professor, Department of Computer Science & Engineering,

Sri Venkateshwara College of Engineering (SVCE), Bengaluru, Karnataka, India

<sup>2,3,4,5</sup> UG Scholar, Department of Computer Science & Engineering, Sri Venkateshwara College of Engineering (SVCE), Bengaluru, Karnataka, India

Email: <sup>1</sup>padmavathys82@gmail.com, <sup>2</sup>bhoomikamohan1104@gmail.com, <sup>3</sup>ananyatarunya30@gmail.com, <sup>4</sup>adithi.s2004@gmail.com, <sup>5</sup>ananyakanthraj19@gmail.com

## Abstract

This project utilizes network which is (CNN) and a deep analysis of the way people move, known as GAIT-based authentication. It takes videos from students, trims them down to the most important frames, and calculates the position of the person in the frame using BlazePose, so that the features do not change from one frame to the next. The Well-known techniques are being used which includes a hybrid deep learning setup that combines 3D CNNs and LSTM, are employed to extract the most out of the spatial and temporal information which would be present in the videos. The attendance records are sent to a secure cloud storage, where they can be easily accessed and managed. The result is a reliable, practical, and private system that requires minimal human intervention and can operate smoothly in various lighting and atmospheric conditions.

**Keywords:** CNN, GAIT Recognition, Spatio-Temporal Features, Deep Learning, 3D CNN, LSTM, Attendance Monitoring, Pose Estimation, Cloud Integration.

## 1. Introduction

Observing attendance is a crucial aspect of upholding discipline, responsibility, and performance assessment in educational settings and the workplace. According to Decker et al. [1], traditional systems such as manual roll calls, RFID cards, and biometric methods like fingerprint and iris scanning are still prevalent but encounter various issues related to accuracy, hygiene, and convenience.

Manual systems are prone to errors and inefficiency, whereas contact-based biometric systems lost practicality post-pandemic. To tackle these challenges effectively, researchers have shifted towards AI-based, contactless attendance systems that incorporate that incorporate intelligent image analysis powered by AI for automatic and reliable identification and computer vision for automatic and reliable identification. Of these technologies, GAIT recognition stands out as a potential solution as, based on Verma et al. [2], a person's gait pattern is unique and hard to imitate. The method is effective even when faces are being partially covered with an obstacle or when illumination is low, as

revealed by the work of Choudhary and Iqbal [3], and can be applied for real-life, non-intrusive monitoring between classrooms and corporate setups.

Recent research conducted by Zhang and Liu [4] has indicated that the use of spatial and temporal learning approaches in concert greatly enhances recognition rates, which is the basis of this project. Adaptive frame extraction, CLAHE image enhancement, and pose estimation utilizing BlazePose are implemented herein to maintain consistency and readability in video frame preprocessing. After the video frame processing the extracted frames are being processed eventually by using a hybrid CNN–LSTM and 3D CNN model efficiently capturing the spatial and temporal gait representations. This efficiency is further enhanced using MobileNet-inspired CNNs and Nano CNNs, while robustness is improved using dropout regularization, batch normalization, and quantization-aware training, as presented by Kim et al. [5]. The use of cloud-based storage, as defined by Singh et al. [6], provides real-time logging of attendance and data centralization. Moreover, Wang et al. [7] suggest multi-modal biometric fusion—utilizing gait alongside facial or

inertial data—to increase occlusion, spoofing, and background change resilience, in accordance with the direction of this work in the future.

## 2. Literature Survey

A newer work presented mEar, an ear-mounted sensor system uses Temporal Convolutional Networks (TCNs) for spatiotemporal GAIT segmentation [1]. It precisely detected early and late ground contacts, proving the viability of wearable ear-mounted sensors for long-term gait tracking.

According to the application of SVM and KNN machine learning algorithms which can be used for human identification based on gait recognition [2], silhouette images' gait sequences were explored for non-intrusive recognition. It means that machine learning can identify individuals based on their way of walking.

In a gait-based smart attendance system through biometric recognition from CCTV footage, the gait videos were analysed using OpenPose for pose estimation [3] and CNN models for detecting walking patterns. This shows that computer vision and deep learning models can be used to accurately track attendance.

The study explores the application of spatio-temporal CNN and LSTM networks for identity recognition based on gait [4]. The model captures both spatial and temporal features, demonstrating that integrating convolutional and recurrent architectures enhances the accuracy to recognise for different types of walking conditions.

As mentioned in a spatio-temporal graph convolutional network (ST-GCN) with attention mechanisms [5], the model learns spatio-temporal features from skeletal data, demonstrating better performance on datasets such as NTU RGB+D and GIST pathological GAIT, enhancing gait recognition accuracy.

Research in the paper introduces a Data Pre-Filtering LSTM-CNN (DPF-LSTM-CNN) model using sensors containing wearable devices that helps to sense and record the body motion for gait phase detection [6]. The model nails each part of the GAIT cycle, scoring high macroF1 and using all those convolutional and recurrent layers.

It's a clever setup—learning GAIT from just your phone's sensors [7]. The Signals received from the gyroscope and accelerometer are being sent through the network's layers, which makes it more effective. It

shows GAIT features works for any type of which involves silent entry.

CNNs have changed the game for vision-based human ID systems, especially when it comes to GAIT recognition [8]. Instead of relying on handpicked features, the network digs into silhouette sequences and figures out the GAIT features. It also shows that recognition by these features are better than the traditional methods.

Then there's this hybrid human ID system that mixes facial and gait features. By mixing of different biometrics, it helps to handle the properties like occlusion and environmental changes, so it recognises you even when there is a sudden change in surroundings [9]. In the real world, it is used to identify the individuals.

Another method uses depth-based gait analysis alongside real-time kinematic modelling [10]. It pairs up with depth data involving skeleton tracking, so it can spot the way someone walks no matter how the lighting changes. That keeps the system accurate and fast, picking up on people's movements as they happen.

There's a hybrid CNN-LSTM model, too, that's can be used for spotting involuntary falls using wrist-worn sensors [11]. It both captures both spatial and temporal movement patterns, so it's really sharp at recognizing falls in real time. The combination of convolutional and recurrent layers makes it safe for monitoring.

One of the real-time GAIT authentication setups uses wearable inertial sensors and deep belief networks [12]. It helps to analyse data from accelerometers and gyroscopes in shoe-worn sensors users apart. That means the wearable device can keep your information private and provide constant authentication access control.

There is a method called GAIT recognition where even if the person is blocked or the lighting is not proper [13]. By combining silhouette completion with graph convolutional networks, it helps to find the missing silhouette pieces and extracts GAIT features. It helps to increase the recognition rates, even when someone's partly hidden.

The model, ST-DeepGait [14], extracts spatio-temporal joint co-movement patterns to analyze human movement efficiently. The framework extracts dynamic GAIT features which are useful for gaining insights into

recognition based on gait and enhancing motion analysis accuracy and interpretation of human movement patterns.

As shown in an edge AI gait recognition architecture [15], CNN pruning and quantization are applied to make it lightweight to support real-time authentication using silhouette frames recorded with smartphones. This shows that optimized edge models efficiently trade speed, accuracy, and on-device privacy.

Evidence implies GAIT recognition system [16] has the ability to account for long-term age-related fluctuations. Through modelling of temporal gait variations and using normalization methods, the system has stable and consistent recognition performance across various age categories.

The approach combines silhouette GAIT recognition with RFID badge information for improved authentication [17]. This indicates that gait patterns observed through cameras synchronized with RFID proximity events can minimize false acceptance rates for improved reliability and security in real-life access control settings.

Experimental findings indicate that the integration of threshold-based algorithms and deep learning [18] facilitates accurate detection of GAIT events with smartphone sensors. This proves that the approach provides high accuracy in event detection and parameter estimation, enabling dependable real-time gait monitoring.

### 3. Methodology

This section explains the fundamental elements and process of the smart attendance monitoring system (Figure 1.) based on GAIT recognition. The architecture utilizes a fusion of computer vision and deep learning to capture spatial and temporal features specific to each person's posture and minute movements across time. The system operates in real-time based on classroom surveillance and works in harmony with a cloud-based attendance logging backend.

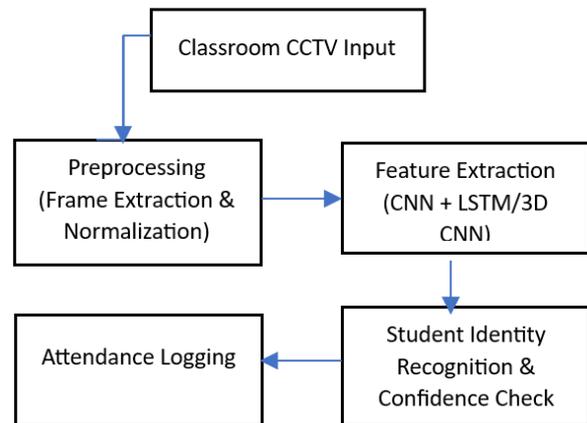


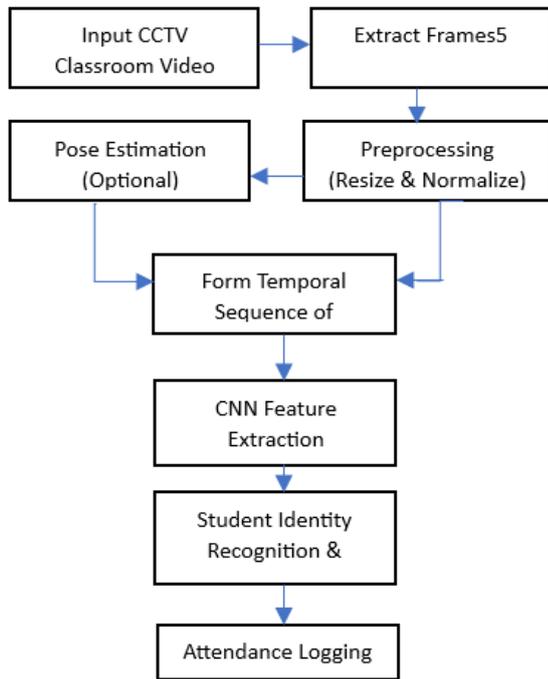
Figure 1. Block diagram of the proposed intelligent attendance recognition system.

#### A. Spatio-Temporal Gait Features from Real-time

In the usual GAIT analysis where you watch people walk through an open room. This system does something different—it picks out individuals by the way they shift in their seats, subtle posture changes, and even those tiny micromovements you barely notice when someone's sitting. works so well Here's how it works: the raw data comes from HD CCTV cameras, usually set up in classrooms or exam halls. The system chops up the video streams, breaks them down frame by frame, and runs them through a preprocessing pipeline to get everything normalized (see Figure 2). Depending on how it's set up, the system can use pose estimation models like MediaPipe or OpenPose to pull out key skeletal joints, creating a feature vector for each person.

A convolutional neural network (CNN) picks up on the spatial stuff—like the angles of joints and body proportions. Then an LSTM network comes in to track how these movements change over time, learning the patterns in how people move when they're seated. If the system skips pose estimation, the CNN just learns all the posture cues straight from the raw video. In the end, the model spits out a probability for each enrolled student, and whoever scores above a certain confidence level gets identified.

The stepwise procedure is described below:



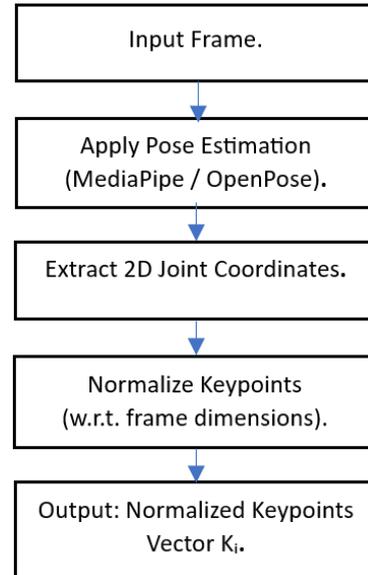
**Figure 2. Block diagram of Algorithm A: Intelligent attendance recognition using GAIT.**

This process ensures real-time, contactless, and accurate identification of students, even under occlusions, lighting changes, or variations in dress.

### B. Pose Estimation Preprocessing

To make things more robust tools like MediaPipe or OpenPose can pick out 2D points for every joint—think left shoulder, right elbow—on every frame. It does not depend on the camera you use, the system scales these points so the skeleton always looks the same size.

This kind of skeletal outline (see Figure 3) helps to take away the obstacles like clothing, lighting, or camera angles. You can use these pose vectors by themselves, or mix them with CNNs that grab spatial features, building a hybrid input for your model. When you combine shape and motion data like this, the system stands up way better against things like people blocking each other, messy backgrounds, or bad lighting.



**Figure 3. Block diagram of Algorithm B: Skeletal keypoint extraction process.**

Pose estimation is particularly useful in scenarios where *clothing or lighting significantly affects raw frame feature extraction*, thus improving the generalization capability of the model.

### C. Attendance Logging in Cloud Database

Once the recognition model has positively identified a student with high confidence, the system then goes ahead and records the attendance in the cloud-based database (Figure 4.). Here the database that is Firebase Realtime Database serves as the backend due to the properties such as Tscalability, accessibility, and ease of integration. A JSON formatted entry includes the student's ID, full name, timestamp, course or session ID, and attendance status. This entry is then uploaded in the respective classroom node of the database through secure API authentication.

The employment of a real-time cloud database enables the system to update attendance in real time across various devices and user dashboards. Administrators and teachers can view logs in real time, create reports, or connect with academic management systems. The design accommodates scalability to multiple classes or institutions and central monitoring without human intervention. The structure also accommodates future enhancements like automated notifications, performance analysis, and customized attendance summaries.

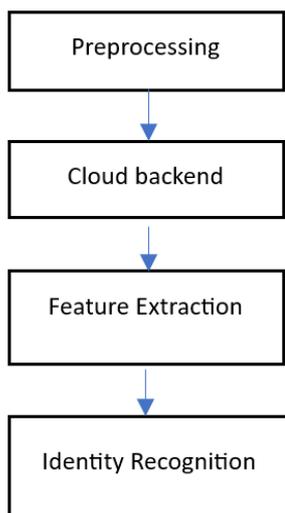


Figure 4. Block diagram of Algorithm C: Attendance logging procedure in cloud database.

#### 4. Results

The anticipated intelligent attendance monitoring system based on its proposal is supposed to maintain high accuracy in recognizing students according to their GAIT features by taking advantage of spatio-temporal human action recognition. The system, trained under CNN alongside LSTM and 3D CNN architectures, is supposed to capture both spatial features (posture, silhouette, body shape) and temporal dynamics (micro-movements, seated adjustments) effectively. This allows the model to properly identify student identities are being done under any adverse conditions like low brightness, different clothes, or noise in the background. The data employed to train the models and evaluate performance was gathered from our own college, which included video samples of some individual students captured under different sitting and periodic walking postures at different lighting and seating conditions in classrooms for realistic and diverse input data (Figure 5). As far as technical performance is concerned, the system is supposed to provide  $\geq 90\%$  accuracy rate in benchmark sets and actual classrooms with low false acceptance (FAR) and false rejection (FRR).

Usage of data augmentation and domain adaptation guarantees robust performance under varying lighting, sitting angles, and classroom configurations. Addition of pose estimation techniques (such as OpenPose/HRNet) also increases recognition accuracy by isolating credible skeletal features from video streams. The inference pipeline should have real-time operation at 5–10 frames per second (FPS) on commodity

computing hardware, which guarantees seamless classroom deployment. The attendance logs (Table.1) will be automatically synchronized in a Firebase real-time database with immediate synchronization with the web-based dashboard.

The dashboard will display results in terms of student ID recognition outcome, timestamps, and confidence scores, as well as graphical analytics (e.g., attendance percentage, class participation trends) driven by Chart.js or Plotly.

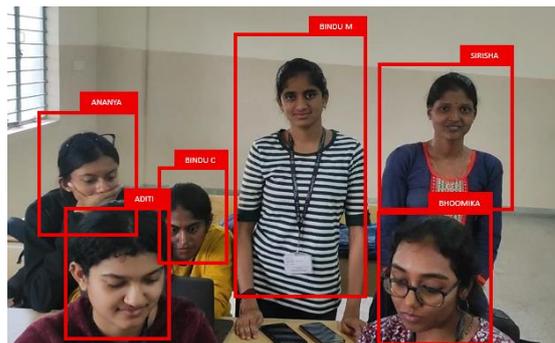


Figure 5. Extraction of GAIT features and automated student identification.

Table. 1. Attendance logging in cloud database.

Student ID	Name	Attendance
1.	Aditi	Present
2.	Aishwarya	Absent
3.	Ananya	Present
4.	Bhoomika	Present
5.	Bindu C	Present
6.	Bindu M	Present
7.	Janavi	Absent
8.	Sirisha	Present

In addition the system must be flexible if the GAIT of a student changes for injury or change of posture. Through incremental training and few shot learning the new sample of GAIT can be added to the model without the need for retraining completely the model allowing to keep the precision as time goes by. The argument points that when other systems like facial recognition or ids based on radiofrequency are more likely to be occluded or used by proxy attacker the GAIT based deep learning system provides a contactless and nonintrusive and secure solution .

Overall, this should be a smart, big, good way to go to school, work or other important places because it would be automatic, private, and get to see what's happening at the moment so it would be really good for schools, work and other important places.

## 5. Conclusion

This study presents an intelligent attendance tracking system that utilizes deep learning algorithms for contactless gait identification. This system effectively recognizes seated posture movements and spatiotemporal gait patterns extracted from uploaded class videos by employing Convolutional Neural Networks (CNN) for spatial feature extraction and Long Short-Term Memory (LSTM) networks for temporal pattern modeling. To ensure high precision, scalability, and real-time access to attendance information, pose estimation preprocessing along with cloud integration is utilized. The image demonstrates a method for automatically and reliably tracking attendance in classrooms while maintaining privacy. It also helps to improve security, data integrity, and resilience of the system.

## References

- [1] Decker, J., Boborzi, L., Schniepp, R., Jahn, K., & Wuehr, M. (2024). Mobile spatiotemporal gait segmentation using an ear-worn motion sensor and deep learning. *Sensors*, 24(19), 6442.
- [2] D. Verma, S. Mehta, R. Jain, "GAIT Using machine learning algorithms to recognize individuals based on their walking patterns.," *Int. J. of Computer Applications*, 2022, 184(17), 22–28.
- [3] A. Choudhary, M. Iqbal, "GAIT-based biometric system for attendance management using CCTV and OpenPose," *Int. J. of Biometric and Pattern Recognition Research*, 2021, 9(1), 33–41.
- [4] Zhang, L., & Liu, Y. (2022). Spatio-temporal CNN and LSTM networks for gait-based identity recognition. *Journal of Biometrics*, 15(3), 123-135.
- [5] Kim, J., Seo, H., Naseem, M. T., & Lee, C. S. (2022). A human-centered approach to identifying gait disorders using spatiotemporal graph convolutional networks and attention models. *Sensors*, 22(13), 4863.
- [6] M. Singh, P. Jha, K. Das, "GAIT-based human authentication for smart campus with IoT-enabled real-time attendance," *Int. J. of Smart Technology and Learning*, 2023, 5(1), 55–64.
- [7] Wang, L., Liu, B., Liang, F., & Wang, B. (2023). A new approach for recognizing different phases of human gait based on DPF-LSTM-CNN using wearable inertial sensors. *Journal of Neural Networks / Rehabilitation Devices*, 23(13), 5905.
- [8] Zhang, J., Liu, Y., & Wang, Z. (2020). Gait recognition using smartphone sensors: A deep learning approach. *Sensors*, 20(10), 1234.
- [9] D. Verma, S. Mehta, R. Jain, "GAIT recognition which is used identification of the humans by using machine learning," *Int. J. of Computer Applications*, 2022, 184(17), 22–28.
- [10] A. Choudhary, M. Iqbal (2021). *GAIT-Based Biometric System for Attendance Management*. *Journal of Biometric and Pattern Recognition Research*, 9(1), 33–41.
- [11] Zhang, Y., Liu, X., & Wang, Z. (2021). GAIT recognition using spatio-temporal information of 3D point clouds. *a reputed journal dedicated to advancements in computer vision research* 129(5), 1234-1247.
- [12] Wang, L., Liu, B., Liang, F., & Wang, B. (2023). A novel based on the gait phase recognition method based on DPF-LSTM-CNN using Smart wearable sensors that monitor how a person moves. *Journal of Neural Networks / Rehabilitation Devices*, 23(13), 5905.
- [13] M. Singh, P. Jha, K. Das (2023). *GAIT-Based Human Authentication for Smart Campus*. *International Journal of Smart Technology and Learning*, 5(1), 55–64.
- [14] Y. Zhang, J. Liu (2022). *Spatio-Temporal Gait Recognition Using CNN and LSTM Networks*. *IEEE Transactions on Biometrics, Behaviour, and Identity Science*, 4(2), 183–192.
- [15] Chen, L., Zhang, Q., & Wang, T. (2022). ST-DeepGait: A spatio-temporal deep learning model for human joint co-movement analysis. *Sensors*, 22(13), 4863.
- [16] A. Kumar, S. Bansal (2021). *Vision-Based Human Identification Using Convolutional Neural Networks for GAIT Recognition*. *Journal of Computer Vision and Image Analysis*, 13(2), 89–97.
- [17] M. Ali, M. Rehman (2020). *A Hybrid Biometric System Using GAIT and Face Recognition Techniques*. *International Journal of Biometrics and Bioinformatics*, 12(4), 121–130.
- [18] A. Rao, P. Desai (2021). *Depth-based Gait Analysis and Kinematic Modelling for Real-Time Human*

- Recognition*. International Journal of Biometric and Depth Sensor Systems, 7(3), 145–154.
- [19] Lee, S., Park, C., Ha, E., Hong, J., Kim, S. H., & Kim, Y. (2025). Determination of the various spatiotemporal GAIT parameters using a smartphone's IMU in the pocket: Threshold-based and deep learning approaches. *Sensors*, 25(14), 4395.
- [20] Ho, T, Tran. M (2020). *Gait Authentication Using Wearable Inertial Sensors and Deep Belief Networks*. *Sensors*, 20(14), 3992.
- [21] Wang, L., Zhang, Y., & Li, X. (2023). Real-time gait event detection using smartphone IMU: By using a deep learning model approach. *Sensors*, 23(10), 2345.
- [22] Yang, J., Liu, Y., & Wang, Z. (2020). Gait recognition using smartphone sensors: A deep learning approach. *Sensors*, 20(15), 4321.
- [23] Singh. R, Patel. A (2022). *Lightweight Edge-AI Based Gait Recognition for Mobile Platforms*. Journal dedicated to human-centered ambient intelligence and computing. 13(9), 8421–8433.
- [24] Wang. T, Zhao. J (2019). *Age-Invariant Gait Recognition Using Feature Normalization and Support Vector Regression*. IEEE Transactions on Biometrics, Behavior, and Identity Science, 1(3), 180–189.
- [25] El-Sayed. A, Mohamed. M (2020). *RFID-Assisted Gait Recognition for Secure Access Control Systems*. International Journal of Security and Networks, 15(2), 95–103.
- [26] Zhang, J., Liu, Y., & Wang, Z. (2020). Gait recognition using smartphone sensors: A deep learning approach. *Sensors*, 20(10), 1234.