International Journal of Engineering Trends and Technology
ISSN: 2231-5381/ https://doi.org/10.14445/22315381/IJETT-V74I5P113

Volume 74 Issue 5, 195-204, May 2026

Original Article

Improving Gait-Based Human Identification Under Variability Using Deep and Machine Learning Ensembles

Babita Sonare[1], Deepika Saxena[2], Vijay Katkar[3]

[1,2] Department of CSE, Poornima University, Jaipur, Rajasthan, India.
[3]School of Artificial Intelligence, Bennett University, Greater Noida, India.

[1]Corresponding Author : sonare.babita@gmail.com

Received: 26 August 2025 Revised: 12 March 2026 Accepted: 28 March 2026 Published: 30 May 2026

Abstract - *This paper investigates the convolutional neural network ensembles with machine learning classifiers for individual identification based on Gait. Gait recognition is a biometric modality wherein a person's unique Gait is used to identify. However, when a person wears a different pair of shoes, carries a load on their back, or walks differently, their gait changes, making it more difficult to identify them. In light of this difficulty, this work suggests an ensemble method to optimize the precision and resilience of gait-based human identification systems. It does this by combining the CNN and LSTM with the machine learning classifier independently by providing features like Gait Energy Image (GEI), Accumulative Frame Difference Energy Image (AFDEI), and their fusion. The proposed system's performance is meticulously checked with the existing system on the "CASIA B" dataset and shows remarkable outcomes, as proved by statistical validation. The proposed methodology is experimented on a newly generated dataset, PCCOE-GAIT, which has given comparable results.*

Keywords - *Biometrics, Gait, ML Classifiers, CNN, LSTM.*

1. Introduction

Person identification is a critical aspect of security systems and surveillance applications. The conventional biometric methods, like fingerprint or facial recognition, have limitations in certain scenarios, such as being affected by physical conditions, wear and tear, injuries, spoofing, etc. It makes the false positive rate higher and results in incorrect person recognition. Another biometric method, Gait, can be an alternative approach for person identification. With its many benefits, it is a useful method for person identification, especially in situations requiring non-intrusive, distance-based identification. It is a promising biometric technology with a wide range of applications due to its resilience to appearance changes and difficulty in imitation, which improves its security capabilities.

Although useful for non-intrusive identification, gait recognition has several challenges, especially when it comes to looking/appearance. For example, differences in footwear, carrying objects, clothing variability, and viewpoint variability. A person's walking pattern changes, making it difficult to identify accurately. These covariates introduce intra-class variability, making it difficult to train the model under controlled conditions.

Deep learning, specifically CNNs, has shown outstanding performance in image recognition tasks. Applying these techniques to gait recognition can potentially improve accuracy and reliability. Ensembling multiple CNN models further enhances the overall performance by leveraging diverse representations.

The ML algorithms that are Support Vector Machine (SVM), Naïve Bayes (NB), K-nearest Neighbors (KNN), Random Forest (RF), Decision Tree (DT), Multilayer Perceptron (MLP), Adaboost, Gradient Boost, and Extreme Gradient Boost are used for classification.

The following are the main contributions to the field of gait biometrics through this research.

- Hybrid features generation: By averaging the two features, GEI and AFDEI, hybrid features are generated. From the gait silhouettes, spatiotemporal and dynamic contents can be regarded. The combined feature properties of GEI with AFDEI in the literature are not found.
- Experimentation with the combined properties of GEI with AFDEI using CNN and LSTM. It is found that GEI-CNN with SVM Classifier works better than hybrid features and AFDEI alone.
- New Dataset generation in the context of an Indian person walking, and its experimentation is done, which shows the generalizability of the trained model.

2. Literature Review

In the field of biometric authentication, the unique and non-intrusive identification approach of gait recognition has garnered a lot of interest. The main advancements, approaches, and research contributions in gait recognition are highlighted in this survey of the literature, which is focused on the feature extraction techniques using handcrafted and deep learning [1].

In the literature study, it has been seen that authors have extracted features from gait silhouettes manually, called handcrafted features, and then applied supervised machine learning algorithms to recognize the person. Conversely, automatic extraction using deep learning. In [2], it was found that KPCA works well with the KNN classifier on gait silhouettes.

Feature extraction using PCA with MLP performs well on large-scale human identification presented in [3] on the CASIA dataset. Authors [4] considered 3 joint angles as features from the CASIA dataset silhouettes and used the dynamic time-warping method for recognizing the person. Soft biometric, which is Gait, has been used to identify humans by using an SVM classifier on the SOTON dataset in [5]. In [6], the use of features such as GEI and AFDEI with PCA and HMM, respectively, for human identification using KNN, was demonstrated on the CASIA -B dataset.

A shape analysis method used in [7], wherein GEI is divided into horizontal, vertical, and grid structures and applied KNN, NB, DT, and RF classifiers. The gait depth data from the TUM GAID dataset in [8] was used, where GEI with the height feature is fused and applied to the KNN, SVC, RFC, and NB. KNN has given better recognition. [9] Experimented with dynamic gait features from various clothing variant datasets such as CASIA B, OUISIR, SACV, and TUM-IITKGP by using SVM, achieving optimized accuracy. [10] Extracted the shape, geometric, and texture features from enhanced frames, subtracted from the background fed to SVM, and experimented on CASIA datasets.

In the advancement of techniques, deep learning algorithms are also used in extensive applications. CNN is helpful to extract features automatically from low-level to high-level, which can be further used for binary or multi-class classification. Optical flow features from the TUM-GAID dataset in [11] are fed to a CNN consisting of 4 layers of convolutional, 2 pooling, and 2 fully connected layers, using the SVM, which gives a good recognition rate. Multi-attribute gait identification is shown using a CNN using the CASIA B dataset in [12] with GEI features and logistic regression classifiers. A human stick model proposed in [13], consisting of knee joints and angles, and classified using an ANN, was demonstrated on the TUM GAID dataset. One uniform model has been proposed in [14], useful for converting from one view to another view angle, which makes cross-view convenient. A two-branch CNN that uses middle and last-fusion strategies is used in [15] to make use of inherent features, experimented on the CASIA and OU-ISIR LP datasets. Different modalities of gait images are used with a CNN in [16], and it was observed that the selection of the architecture of the CNN alters state-of-the-art or poor results. A CNN and LSTM combined approach for gait recognition is presented [17]. A demonstration of GEI with 3D CNN by using CASIA and OULP datasets [18] has been presented.

[19] employed GEI and MSI fusion and fed the results to CNN using a softmax classifier to obtain excellent accuracy on the CASIA B dataset. GEI to CNN in [20] fed and demonstrated the relu activation function on the CASIA B dataset. [21] done the gait recognition using an 8-layer CNN with relu activation function, experimented on the CASIA B dataset. [22] utilizes DenseNet-201, VGGNet-16, and VT feeding with an average silhouette that GEI from three different datasets. Predictions from each network have been averaged and fed to MLP. Authors in [23] introduced a DPF-LSTM-CNN model using wearable inertial sensors, demonstrating the value of combining CNNs and LSTMs across both vision- and sensor-based gait analysis. [24] presented an ensemble approach processing GEI segment through part-wise CNNs and aggregating their features, significantly boosting rank-1 accuracy under covariate conditions.

Further advances include a study proposing a Global-Local Feature Extractor based on SENet (GLFES) with attention to fuse local and global gait cues, achieving strong performance on CASIA-B and OUMVLP [25]. Meanwhile, an EfficientNet-based ensemble with CGAN-generated GEIs achieved good accuracy on CASIA-B [26]. Together, these works show that combining spatio-temporal modeling, feature fusion, and ensemble strategies significantly enhances gait recognition robustness. Authors in [27] proposed a multimodal approach based on a pretrained network and attention-based models; later, decision-level, feature-level, and fusion-based strategies were employed.

While prior work has explored spatial and temporal features independently, the potential Hybrid of GEI and AFDEI into hybrid feature generation has not been investigated; furthermore, a comprehensive comparison of CNN-LSTM has not been experimented on hybrid features. Existing work is more based on a standard dataset, which raises the generalizability of the trained model on a population with different physical characteristics, clothing style, and walking habits.

3. Proposed Methodology

The proposed methodology for human identification using ML and DL is presented in this section.

3.1. Input data

The CASIA B [28] dataset includes 124 people's gait-corresponding silhouettes in three different walking configurations. The subjects walked normally in six variations, carried a bag in two variations, and wore a coat in two variations while moving in each of the 11 view angles, which are '000°, 018°, 036°, ..., 180°. There are roughly 8000 frames available for a single person for every angle variation, in addition to 10 variations for walking and carrying. Thus, approximately 1118373 frames are accessible for 124 subjects. For each silhouette, the size is trimmed to 120 by 80.

To provide the input to the network, Gait Energy Images (GEI) [29], Accumulative Frame Difference Images (AFDEI) [6], and hybrid images are generated. GEIs are the average of silhouettes over one gait cycle. GEIs are condensed spatiotemporal representations of gait silhouettes. GEI calculation is shown in Equation (1).

$$G(x, y) = \frac{1}{N} \sum_{t=1}^{N} I_t\,(x, y) \qquad (1)$$

Where $I_t(x, y)$is the input silhouette at the time?

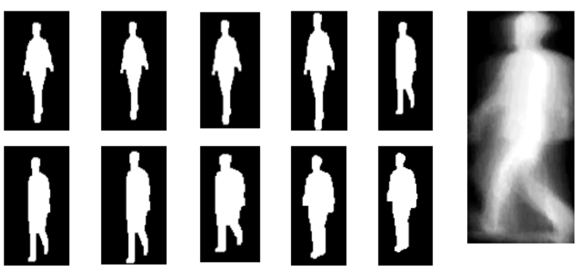

Fig. 1(a) Silhouettes of normal walking of (b) GEI subject 001 at 36°

Figures 1(a) and 1(b) show representative silhouettes of normal walking of subject 001 at 36° and its GEI.

AFDEI images are calculated by using forward and backward differences of images over one gait cycle, which helps retain the dynamic part of walking images. AFDEI calculation is shown in Equation (2).

Let $I_t(x, y)$ be the binary image frame at time t, where (x, y) denotes the pixel coordinates. A full gait cycle consists of T frames. The AFDEI can be defined as:

$$AFDEI(x, y) = \frac{1}{2(T-1)} \sum_{t=1}^{T-1} [|I_{t=1}\,(x, y) - I_t(x, y)| + |I_t(x, y) - I_{t-1}(x, y)|] \qquad (2)$$

This equation highlights motion dynamics at each pixel location during a gait cycle by calculating the average of the absolute changes between consecutive forward and backwards frames.

Fig. 2 (a) Silhouettes of normal walking of (b) AFDEI subject 001 at 36°

Figures 2(a) and 2(b) show representative silhouettes of normal walking of subject 001 at 36° and its AFDEI.

3.1.1. Hybrid image generation from GEI and AFDEI

The video sequence of the Gait has been taken into consideration to identify the person. Extracted frames from a video sequence are subjected to background subtraction [30] and binarization. There has been frame resizing. Afterward, AFDEI and GEI images are produced. The average silhouette of a person's stride is represented by GEI, while the dynamic variations between consecutive frames are captured by AFDEI, which may help identify action-specific features.

Averaging in their respective viewing angles is a blending approach used to merge the GEI and AFDEI to get hybrid images. Figure 3 shows the hybrid image generation from GEI and AFDEI in the respective angle hybrid images.

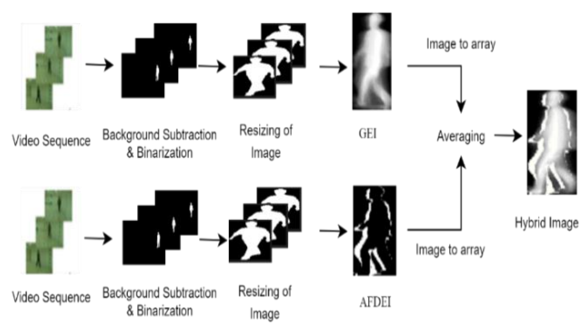

Fig. 3 Generation of a hybrid image from GEI and AFDEI

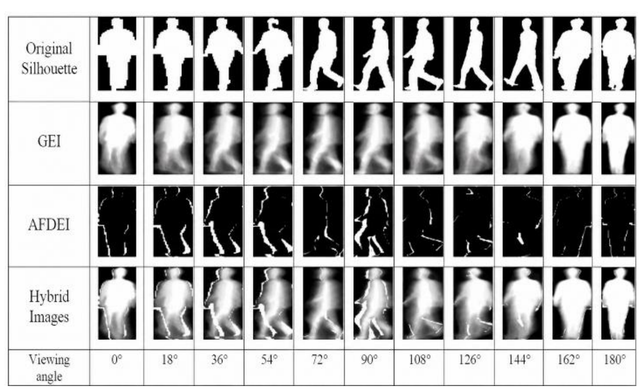

Fig. 4 Transformed images for subject 001 at different angles

Figure 4 shows the original silhouette, GEI, AFDEI, and its hybrid image in respective viewing angles. Original images are representative images from the respective viewing angle. By considering frames from one gait cycle, GEI, AFDEI, and Hybrid images are generated.

3.2. Network architecture

In this paper, person identification based on gait silhouettes using Deep Learning (DL) is proposed. DL techniques such as CNN [31] and LSTM [32] are used. After experimentation, an optimized convolutional neural network is proposed.

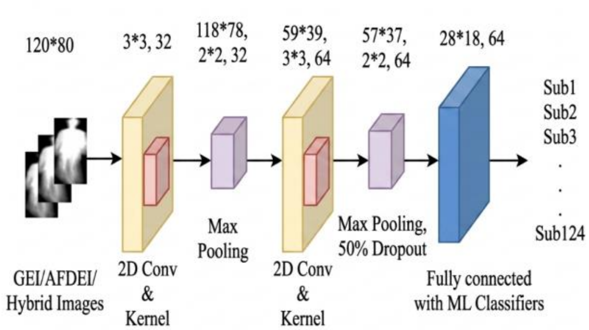

Fig. 5 CNN-based human gait identification system

The above Figure 5 shows the proposed CNN-based human gait identification system. The system consists of only 2 convolutional layers along with a kernel, rectilinear unit, max-pooling with a dropout layer, and a fully connected layer with different machine learning classifiers. 120*80 Resized GEIs are fed to 2D convolutional layer along with 3*3, 32 filters without padding. Its output is 118*78, 32 feature maps have undergone max pooling and fed to 59*39, 32 to a 2D convolutional layer along with 3*3, 64 filters. The output of the previous layers, 57*37, 64, is passed through a 2*2 max pooling layer and 50% dropout. Its output is 28*18, 64 given to the fully-connected layer after flattening. Adam optimizer is used with a learning rate of 0.001. Instead of using the Softmax classifier, different machine-learning classifiers are used for classification and comparison.

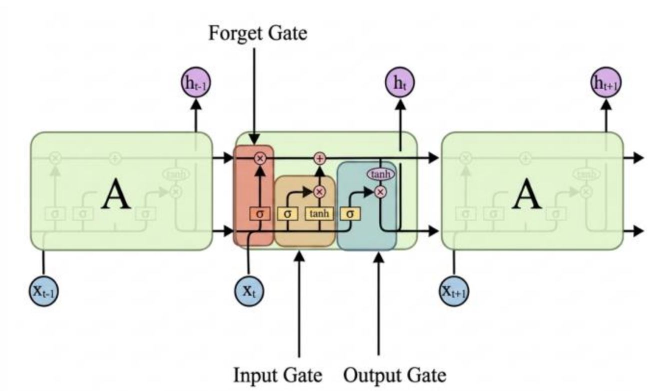

Fig. 6 Human Gait Identification System using LSTM

The variant of the RNN, Long Short-Term Memory (LSTM) [32] architecture, is also used for experimentation. Figure 6 shows a Human gait identification system using LSTM. The forget gate, input gate, and output gate are the three gates that are used to control the flow of information. The forget cell helps identify data that can be deleted from the network by identifying which of it is no longer required. Input data is what the network uses to obtain new input. The third gate, the output gate, decides which data should be displayed as the design's output at any particular moment. These components make it possible for LSTMs to acquire and retain knowledge over long periods, which makes person recognition highly beneficial.

4. Results and Discussion

The Python Jupyter environment on an Intel Core i7 processor, 8 GB RAM, and an NVIDIA RTX 1080 Ti GPU has been utilized for testing with the CASIA B dataset. The dataset consists of 124 subjects' walking videos. Dataset separation is done as all normal walk images, some samples of carrying a backpack and wearing a coat (normal-01, normal-02, normal-03, normal-04, bag-01, cloth-01) for training, (normal-05, normal-06, bag-02, cloth-02) for testing. Exhaustive experimentation with CNN is performed to get an optimized network by using the Keras Tuner library for hyperparameter tuning. The variant of the RNN, Long Short-Term Memory (LSTM) [32], is also used for experimentation. Experimented for 2 to 5 layers of convolutional layers, 32 to 128 filters, optimizers like 'adam', 'sgd', 'rmsprop', and 'adadelta', and learning rate from 0.1, 0.01, 0.001. With the help of a random search, the best parameter was identified.

Assumptions for experimentation

- Silhouettes are precisely extracted from video sequences
- Walking sequences represent stable gait cycles with respect to the number of frames per second.

Performance of all the techniques is measured with the help of accuracy, considering it as the recognition rate. Accuracy is the ratio of correctly classified images to all provided images. 〖RecognitionRate

$$RecognitionRate = (TP + TN)/(TP + FP + TN + FP) \quad (3)$$

Where TP is correctly predicted class P, TN is correctly predicted class N, FP is wrongly predicted class P, and FN is wrongly predicted class N. Categorical cross-entropy loss is used to calculate the loss, which gives a probabilistic prediction for all class categories.

$$CE = -\sum_{i=1}^{n} y\_true * \log(y\_pred) \quad (4)$$

Equation (4) shows categorical cross-entropy loss.

4.1. Ablation Study

An ablation study has been conducted to systematically examine the architecture of CNN and LSTM to prevent loss or overfitting of the data. A variety of convolutional, pooling, dropout, and normalization configurations for CNN architecture are tested at various learning rates and optimizers. Two LSTM layers, one dropout, a dense layer, several optimizers, learning rates, activation functions, and cost functions were also tested with the LSTM network. An attempt has been made to combine CNN and LSTM with many machine learning classifiers, such as SVM, LR, RF, MLP, NB, KNN, Adaboost, GB, and XGB. Tables 1 and 2 shows the hyperparameter tuning for CNN and LSTM networks.

Table 1. CNN Parameter tuning

Layer (type)	No. of Units
Conv_01_filter	48
Conv_01_kernel	3
Conv_02_filter	64
Conv_02_kernel	5
Dense_01_units	112
Learning_rate	0.001
Optimizer	Adam
Dropout	0.5
loss	Categorical_Crossentropy

Table 2. LSTM Parameter Tuning

Layer (type)	No. of Units
LSTM_1_units	128
LSTM_2_units	48
optimizer	rmsprop
Dropout	0.5
loss	categorical_crossentropy

4.2. Recognition Rate/Accuracy using GEI_CNN and Hybrid_CNN

A Convolutional Neural Network with several machine learning classifiers, including SVM, RF, MLP, LR, KNN, NB, Adaboost, GradientBoost, and Extreme Gradient Boost, is implemented by feeding GEI and Hybrid images to the network. Additionally, the recognition rate was confirmed.

The following are the results achieved, shown in the following tables. It is the average result of test probes that are normal, coat-wearing, and bag-carrying variants. Using multiple classifiers for experimentation helps to decide which viewing angle is significant. above Table 3 indicates that Angles 0°, 54°, and 180° are significant. In comparison to other classifiers, with an average accuracy of 83.6%, the CNN with the SVM classifier outperforms.

Table 3. Recognition Rate/Accuracy using GEI_CNN

Viewing Angle	SVM	RF	MLP	LR	KNN	NB	Adaboost	GB	XGB
0°	88.6	79.5	89.3	83.2	81.4	43.8	48.1	82.8	67.6
18°	85.6	91.4	83.9	86.0	80.2	74.5	44.8	79.4	64.4
36°	80.0	81.9	78.8	71.4	74.5	66.3	42.3	75.9	64.1
54°	85.5	76.5	85.6	89.5	86.8	84.7	20.2	83.2	71.3
72°	88.3	84.5	88.8	83.3	81.5	79.7	53.4	82.1	62.2
90°	85.9	69.6	82.7	85.2	81.1	77.6	43.2	85.6	47.0
108°	84.5	81.2	80.2	85.4	81.7	77.2	49.2	79.2	55.7
126°	73.1	83.8	71.6	82.7	81.5	77.1	35.8	82.4	53.5
144°	80.1	83.4	75.2	75.6	80.1	68.1	42.0	83.0	62.3
162°	82.6	79.3	77.9	81.9	83.5	55.9	40.1	75.2	69.9
180°	85.4	81.5	89.0	91.0	85.0	75.9	45.2	86.2	53.9
Avg.	83.6	81.1	82.1	83.2	81.6	71.0	42.2	81.4	61.1

Table 4. Recognition Rate/ Accuracy using Hybrid_CNN

Viewing Angle	SVM	RF	MLP	LR	KNN	NB	Adaboost	GB	XGB
0°	50.0	23.7	44.9	43.4	49.9	28.4	27.6	45.8	17.9
18°	36.4	17.1	39.1	35.5	37.4	37.2	6.3	36.8	24.5
36°	23.8	12.9	26.1	29.8	29.0	29.6	14.5	28.9	9.3
54°	29.0	10.8	34.8	32.7	27.8	6.3	11.4	21.2	8.6
72°	29.6	15.6	35.8	33.2	33.6	25.1	17.9	23.9	16.5
90°	24.7	12.4	27.7	24.6	25.1	21.6	16.4	25.4	14.7
108°	23.7	12.5	20.0	20.2	22.2	20.6	12.4	24.2	15.2
126°	17.3	5.7	19.9	18.3	17.6	11.3	7.3	15.6	5.1
144°	23.4	4.8	19.9	25.8	21.6	7.0	12.0	17.6	11.8
162°	29.6	5.2	34.4	42.5	35.6	11.8	17.3	29.4	20.2
180°	38.3	8.5	37.0	46.0	36.0	19.5	15.5	33.5	19.4
Avg.	29.6	11.7	30.9	32.0	30.5	19.9	14.4	27.5	14.8

Compared with many classifiers, Table 4 indicates that viewing Angle 0° is significant, and CNN with SVM gives an accuracy of 50% for hybrid features, followed by KNN with 49.9% accuracy.

4.3. Recognition Rate/Accuracy using GEI_LSTM and Hybrid_LSTM

To recognize the individual, the LSTM network with two LSTM layers, with different Machine learning classifiers that are SVM, RF, MLP, LR, KNN, NB, Adaboost, GradientBoost, and Extreme Gradient Boost, is applied by feeding GEI and Hybrid images to the network.

The viewing angles of 72° and 90° are significant and give a recognition rate of up to 76% in the case of GEI and LSTM combination with a logistic regression classifier. Angle 0 is significant, but the performance is lower compared to GEI_CNN in the case of hybrid features fed to the LSTM using the KNN classifier. The recognition rate is up to 65%. This recognition rate can be increased by tuning the hyperparameter more rigorously.

4.4. Statistical Analysis

A paired t-test on the GEI_CNN and GEI_Hybrid groups is applied to prove which group performs well. Here, the mean is the average value for all viewing angles under each group.

Step 1: Set the Hypothesis

Null Hypothesis (H_0): There is no difference in means between GEI_CNN and Hybrid_CNN.

Alternative Hypothesis (H_1): There is a significant difference in means between GEI_CNN and Hybrid_CNN.

Step 2: Calculation of p-value after applying the paired t-test for degrees of freedom 8, as 9 classifiers are considered in each group.

Table 5. Independent t-Test Results Comparing GEI_CNN and Hybrid_CNN Performance

Mean (GEI_CNN)	Mean (Hybrid_CNN)	t Stat	Degrees of Freedom	P(T<=t) two-tail	t Critical two-tail
74.13	23.48	14.20	8	5.9E-07	2.31

Step 3: Conclusion

Since |t Stat| (14.20) > t Critical (2.31) and p-value < 0.05, reject the null hypothesis.

There is a highly significant difference between GEI_CNN and Hybrid_CNN means. The difference is not due to random chance but is statistically significant at a very high confidence level (>99.9999%). Based on the mean value, GEI_CNN is performing much better than Hybrid_CNN.

4.5. Generation of a new gait dataset (PCCOE-Gait) and experimentation

The above methods are verified on the new dataset also. The new real-time gait dataset of 53 subjects is generated from 11 viewing angles (0°, 18°, 36° …, 180°) by considering different variants/challenges at 6 normal walking, 2 carrying a backpack, and 2 wearing a coat.

Figure 7 shows the direction of viewing angle, and Figure 8 shows the subject's Gait. For 53 subjects from 11 viewing angles (0°, 18°, 36° …, 180°), by considering different variants/challenges at 6 normal walking, 2 carrying a backpack, and 2 wearing a coat, videos are captured.

These videos are processed and generate silhouettes by using a masked image and morphological operators (erosion and dilation). From the silhouette, gait energy image, accumulative frame difference images, and hybrid images (Average image of gait energy image and accumulative frame difference image) are generated. These images are fed to the Convolutional Neural Network and its ensemble, and the recognition rate is calculated.

Figure 9 shows the recognition rate analysis of CNN along with its ensemble using GEI and Hybrid features. It shows that GEI-based methods consistently outperform Hybrid-based methods.

Among all combinations, the best performance is achieved with GEI + CNN MLP, reaching a recognition rate of 69.9% for CL and 77.3% for BG, followed closely by GEI + CNN RF and GEI + CNN GB, which also deliver relatively high accuracies.

In contrast, the lowest performance is observed with GEI + CNN XGB, yielding only 25.5% for CL and 34.9% for BG. On the other hand, Hybrid features demonstrate a significant drop in recognition accuracy across all classifiers.

The highest result among Hybrid models is obtained with Hybrid + CNN LR (40.1% for CL and 42.8% for BG), while the weakest performance is again observed with Hybrid + CNN XGB, giving around 10% for both conditions.

Overall, BG values are consistently higher than CL values across methods, but the results clearly highlight that GEI features provide a stronger representation for CNN-based recognition compared to Hybrid features, which show limited effectiveness.

Fig. 7 Viewing Angles

Fig. 8 Sample of Subject walking

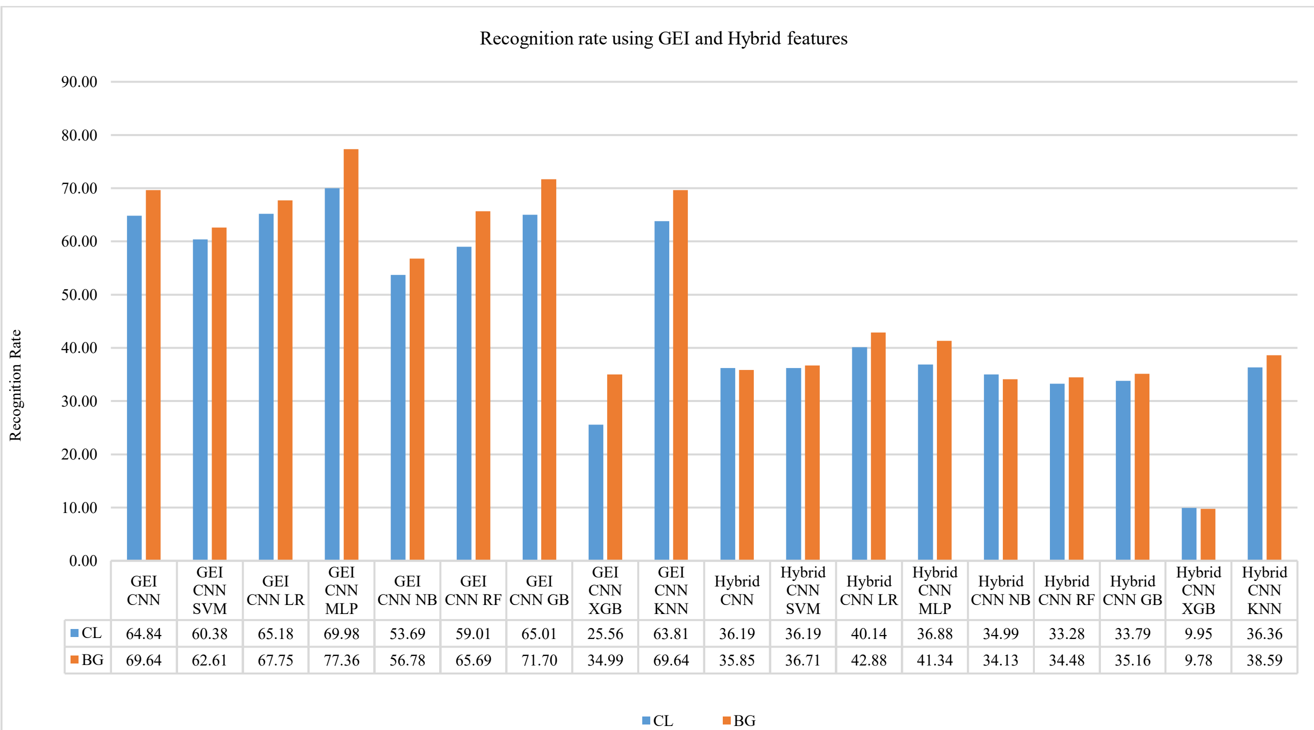

Fig. 9 Recognition rate of CNN and its ensemble using GEI and Hybrid features for PCCOE-Gait dataset

4.6. Result Comparison with Existing Work

The aforementioned experimental findings are further contrasted with previous research that used a Convolutional Neural Network on the CASIA-B gait dataset.

An evaluation of the relative performance gains made by the suggested strategy versus well-known CNN-based techniques is made possible by this comparison. Table 6 shows the comparative results of the proposed system with existing work. It has been experimentally shown that the suggested 2-layered CNN architecture with SVM performs better for different covariates when it comes to person recognition based on Gait, as it is able to extract robust features. Ensemble learning techniques further strengthen the classification performance.

By adjusting the hyperparameters, it can rise more. The highlighted value shows the proposed system recognition rate. The methodology offers a fair and consistent foundation for assessing accuracy, robustness, and generalization capabilities by comparing the two models using the same benchmark dataset. The results validate the suggested method's contribution to the gait recognition field by highlighting whether it provides a statistically significant advantage in terms of recognition rates.

From an engineering and deployment perspective, the proposed system has a computationally low cost as it works offline and the architecture has fewer layers, scalable in access control and surveillance systems.

Table 6. Comparative results of the proposed systems with the existing work

Authors, Reference	Models	Different Challenging Conditions in Train/Test	Performance (%)
Zhang et al.[33]	VT-GAN	Yes	48.53
Min et al.[20]	Deep CNN	Yes	48.23
Liu & Liu[34]	TS-Net	-	56.233
Wang & Yan [35]	GCF-CNN	-	62.36
Yeo et al.[36]	triplet CNN	Yes	72.987
Wu et al.[37]	FWCN	-	74.507
Xu [38]	DLMNN	-	80.67
Wei et al.[39]	Cycle-GAN	Yes	81.30
Deng et al.[40]	Cycle-GAN-SiaNet	Yes	81.50
Chao et al.[41]	GaitSet	-	85.733
Proposed method	2-layered CNN-SVM	Yes	85.767

5. Conclusion

The human identification based on Gait using a CNN and LSTM network with machine learning classifiers such as SVM, LR, RF, MLP, NB, KNN, Adaboost, GB, and XGB is used on GEI, AFDEI, and a hybrid of them for the CASIA B dataset, 124 subjects' silhouettes in 11 viewing angle experiments individually. GEI, AFDEI, and their hybrid combinations are fed to the CNN and LSTM with machine learning classifiers to achieve an optimum recognition rate with GEI and hybrid features. The CNN with a support vector machine classifier works well with 85.767% recognition accuracy for the gait energy image features, followed by its hybrid features, giving better results. Gait AFDEI features are not suitable for CNN architecture because of their poor performance. One important benefit of this finding is that optimized accuracy is achieved with fewer layers. The performance is comparable on the new real-time dataset as well, which yielded a recognition rate on CL-BG 71.5% for Gait Energy Images. The future scope for this system is to increase the recognition rate by using pre-trained network techniques and to further work on different uncontrolled challenges in Gait.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Acknowledgements

All authors acknowledge Pimpri Chinchwad Education Trust's Pimpri Chinchwad College of Engineering, Pune, Maharashtra, India, for granting permission to generate the dataset of 53 students used in the study.

References

[1] Jashila Nair Mogan, Chin Poo Lee, and Kian Ming Lim, "Advances in Vision-based Gait Recognition: From Handcrafted to Deep Learning," *Sensors*, vol. 22, no. 15, pp. 1-24, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[2] Ning Suo, Xu Qian, and Jinhui Zhao, "Gait Recognition based on KPCA and KNN," *2010 The 2nd Conference on Environmental Science and Information Application Technology*, Wuhan, China, pp. 432-435, 2010. [CrossRef] [Google Scholar] [Publisher Link]

[3] Emdad Hossain, and Girija Chetty, "A Multi-Modal Gait based Human Identity Recognition System based on Surveillance Videos," *2012 6th International Conference on Signal Processing and Communication Systems*, Gold Coast, QLD, Australia, pp. 1-4, 2012. [CrossRef] [Google Scholar] [Publisher Link]

[4] Soumia Benbakreti, Samir Benbakreti, and Mohammed Benyettou, "Recognizing Human Gait in Video Sequences," *2012 International Conference on Multimedia Computing and Systems*, Tangiers, Morocco, pp. 323-327, 2012. [CrossRef] [Google Scholar] [Publisher Link]

[5] Daniel A. Reid, Mark S. Nixon, and Sarah V. Stevenage, "Soft Biometrics; Human Identification using Comparative Descriptions," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 36, no. 6, pp. 1216-1228, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[6] Jing Luo et al., "Gait Recognition using GEI and AFDEI," *International Journal of Optics*, vol. 2025, no. 1, pp. 1-5, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[7] Anup Nandy, Rupak Chakraborty, and Pavan Chakraborty, "Cloth Invariant Gait Recognition using Pooled Segmented Statistical Features," *Neurocomputing*, vol. 191, pp. 117-140, 2016. [CrossRef] [Google Scholar] [Publisher Link]

[8] Kristijan Lenac et al., "Extending Appearance based Gait Recognition with Depth Data," *Applied Science*, vol. 9, no. 24, pp. 1-18, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[9] Hajra Masood, and Humera Farooq, "An Appearance Invariant Gait Recognition Technique using Dynamic Gait Features," *International Journal of Optics*, vol. 2021, no. 1, pp. 1-15, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[10] Muhammad Sharif et al., "A Machine Learning Method with Threshold based Parallel Feature Fusion and Feature Selection for Automated Gait Recognition," *Journal of Organizational and End User Computing (JOEUC)*, vol. 32, no. 2, pp. 67-92, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[11] Francisco Manuel Castro et al., "Automatic Learning of Gait Signatures for People Identification," *Advances in Computational Intelligence: 14th International Work-Conference on Artificial Neural Networks, IWANN 2017*, Cadiz, Spain, pp. 257-270, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[12] Anubha Parashar et al., "Data Preprocessing and Feature Selection Techniques in Gait Recognition: A Comparative Study of Machine Learning and Deep Learning Approaches," *Pattern Recognition Letters*, vol. 172, pp. 65-73, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[13] Win Kong et al., "Human Gait State Classification using Artificial Neural Network," *2014 IEEE Symposium on Computational Intelligence for Multimedia, Signal and Vision Processing (CIMSIVP)*, Orlando, FL, USA, pp. 1-5, 2014. [CrossRef] [Google Scholar] [Publisher Link]

[14] Shiqi Yu et al., "Invariant Feature Extraction for Gait Recognition using only one Uniform Model," *Neurocomputing*, vol. 239, pp. 81-93, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[15] Xiuhui Wang, and Jiajia Zhang, "Gait Feature Extraction and Gait Classification using Two-Branch CNN," *Multimedia Tools and Applications*, vol. 79, no. 3-4, pp. 2917-2930, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[16] Francisco M. Castro et al., "Multimodal Feature Fusion for CNN-based Gait Recognition: An Empirical Comparison," *Neural Computing and Applications*, vol. 32, no. 17, pp. 14173-14193, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[17] David Kreuzer, and Michael Munz, "Deep Convolutional and LSTM Networks on Multi-Channel time Series Data for Gait Phase Recognition," *Sensors*, vol. 21, no. 3, pp. 1-15, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[18] Saba Gul et al., "Multi-View Gait Recognition System using Spatio-Temporal Features and Deep Learning," *Expert Systems with Applications*, vol. 179, pp. 1-9, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[19] Md. Iman Junaid, and Samit Ari, "Gait Identification using Ensemble Deep Learning Technique," *2022 IEEE Silchar Subsection Conference (SILCON)*, Silchar, India, pp. 1-5, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[20] Pa Pa Min, Shohel Sayeed, and Thian Song Ong, "Gait Recognition using Deep Convolutional Features," *2019 7th International Conference on Information and Communication Technology (ICoICT)*, Kuala Lumpur, Malaysia, pp. 1-5, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[21] Abhishek Sheth et al., "Gait Recognition using Convolutional Neural Network," *International Journal of Online and Biomedical Engineering*, vol. 19, no. 1, pp. 107-118, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[22] Jashila Nair Mogan et al., "Gait-CNN-ViT: Multi-Model Gait Recognition with Convolutional Neural Networks and Vision Transformer," *Sensors*, vol. 23, no. 8, pp. 1-15, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[23] Kun Liu et al., "A Novel Gait Phase Recognition Method based on DPF-LSTM-CNN using Wearable Inertial Sensors," *Sensors*, vol. 23, no. 13, pp. 1-17, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[24] Büşranur Yaprak, and Eyüp Gedikli, "Enhancing Part-based Gait Recognition via Ensemble Learning and Feature Fusion," *Pattern Analysis and Applications*, vol. 28, no. 2, pp. 1-19, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[25] Qizhen Chen et al., "Gait Recognition based on Temporal Gait Information Enhancing," *MultiMedia Modeling: 30th International Conference, MMM 2024*, Amsterdam, The Netherlands, pp. 451-463, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[26] Entesar T. Burges, Zakariya A. Oraibi, and Ali Wali, "Gait Recognition using CGAN and EfficientNet Deep Neural Networks," *Proceedings of the 20th International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISIGRAPP 2025)*, vol. 3, pp. 339-346, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[27] Kamrul Hasan et al., "MMF-Gait: A Multi-Model Fusion-Enhanced Gait Recognition Framework Integrating Convolutional and Attention Networks," *Symmetry*, vol. 17, no. 7, pp. 1-24, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[28] Shuai Zheng et al., "Robust View Transformation Model for Gait Recognition," *2011 18th IEEE International Conference on Image Processing*, Brussels, Belgium, pp. 2073-2076, 2011. [CrossRef] [Google Scholar] [Publisher Link]

[29] Changsheng Wan, Li Wang, and Vir V. Phoha, "A Survey on Gait Recognition," *ACM Computing Surveys (CSUR)*, vol. 51, no. 5, pp. 1-35, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[30] Helly M. Desai, and Vaibhav Gandhi, "A Survey: Background Subtraction Techniques," *International Journal of Scientific and Engineering Research*, vol. 5, no. 12, 2014. [Google Scholar]

[31] Laith Alzubaidi et al., "Review of Deep Learning: Concepts, CNN Architectures, Challenges, Applications, Future Directions," *Journal of Big Data*, vol. 8, no. 1, pp. 1-74, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[32] Benjamin Lindemann et al., "A Survey on Long Short-Term Memory Networks for Time Series Prediction," *Procedia CIRP*, vol. 99, pp. 650-655, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[33] Peng Zhang, Qiang Wu, and Jingsong Xu, "VT-GAN: View Transformation GAN for Gait Recognition Across Views," *2019 International Joint Conference on Neural Networks (IJCNN)*, Budapest, Hungary, pp. 1-8, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[34] Xiaoyang Liu, and Jinqiang Liu, "Gait Recognition Method of Underground Coal Mine Personnel based on Densely Connected Convolution Network and Stacked Convolutional Autoencoder," *Entropy*, vol. 22, no. 6, pp. 1-17, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[35] Xiuhui Wang, and Ke Yan, "Gait Classification through CNN-based Ensemble Learning," *Multimedia Tools and Applications*, vol. 80, no. 1, pp. 1565-1581, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[36] Sang-Soo Yeo et al., "A Triplet-Branch Convolutional Neural Network for Part-based Gait Recognition," *Computer Systems Science and Engineering*, vol. 47, no. 2, pp. 2027-2047, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[37] Yalian Wu et al., "Gait Recognition based on Feedback Weight Capsule Network," *2020 IEEE 4th Information Technology, Networking, Electronic and Automation Control Conference (ITNEC)*, Chongqing, China, pp. 155-160, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[38] Wanjiang Xu, "Deep Large Margin Nearest Neighbor for Gait Recognition," *Journal of Intelligent Systems*, vol. 30, no. 1, pp. 604-619, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[39] Longhui Wei et al., "Person Transfer GAN to Bridge Domain Gap for Person Re-identification," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, pp. 79-88, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[40] Weijian Deng et al., "Image-Image Domain Adaptation with Preserved Self-Similarity and Domain-Dissimilarity for Person Re-identification," *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT, USA, pp. 994-1003, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[41] Xinnan Ding et al., "Sequential Convolutional Network for Behavioral Pattern Extraction in Gait Recognition," *Neurocomputing*, vol. 463, pp. 411-421, 2021. [CrossRef] [Google Scholar] [Publisher Link]