Mathematical Formulation and Experimentation for Multi-View GEI-Based Gait Identification Using an Ensemble Learning and Optical Flow

Babita Sonare
Department of Computer Science and Engineering, Poornima University, Jaipur, Rajasthan, India
sonare.babita@gmail.com (corresponding author)

Deepika Saxena
Department of Computer Science and Engineering, Poornima University, Jaipur, Rajasthan, India
deepika.saxena@poornima.edu.in

Vijay Katkar
School of Artificial Intelligence, Bennett University, Greater Noida, India
katkarvijayd@gmail.com

Received: 24 January 2026 | Revised: 15 February 2026 and 3 March 2026 | Accepted: 10 March 2026

 DOI: https://doi.org/10.48084/etasr.17731

ABSTRACT

The challenge of identifying individuals in various situations, such as walking orientation and clothing conditions, from their gait patterns remains challenging in biometric recognition. To mitigate this challenge, this work presents a unified mathematical formulation that incorporates optical flow-based viewing angle estimation to reduce computational complexity and a Kernel-PCA (KPCA) ensemble framework, combining the strengths of multiple pre-trained Convolutional Neural Networks (CNNs), including EfficientB0 and B1, MobileNet and MobileNetV3 variants, as well as several RegNetX and RegNetY models. The most suitable lightweight model among pretrained CNNs is used to extract the features based on the estimated angle and Gait Energy Image (GEI) generated from the CASIA B dataset. The extracted features are refined using KPCA to improve class separability and recognition. These transformed features are utilized to train Machine Learning (ML) classifiers such as K-Nearest Neighbors (KNN), Support Vector Machine (SVM), Random Forest, and XGBoost. From these classifiers, KNN using transformed features produced the best result. The proposed method achieved recognition accuracies between 95.6% and 99.2% across viewing angles from 0° to 180° in the training and 92.07% to 95.57% in testing. By integrating angle estimation with selective model usage, the framework significantly reduces computational overhead while maintaining high recognition performance. Collectively, the proposed approach delivers a robust and efficient solution for GEI-based person identification under a varied range of real-world conditions and viewing angles.

Keywords-gait analysis; machine learning; convolutional neural network; ensemble; optical flow

I. INTRODUCTION

In the security-related domain, biometric methodologies, such as the physiological or behavioral properties of a person, are used to recognize identity. Although conventional methods, such as fingerprint analysis, facial recognition, and iris scanning, are significantly effective, their use is typically restricted to consent and cooperative scenarios, restricting their applicability. To overcome these limitations, the Gait Energy Images (GEIs) are developed based on the gait appearance of the subject's silhouettes, showing great potential as a non-invasive method to identify a person. Nevertheless, there are several challenges associated with this innovative strategy. When subjects walk at varying cross angles, the identification of individuals based on GEIs becomes a significant difficulty. Load carrying capacity can change the length of the stride, ultimately changing the gait [1]. Both traditional and modern models have struggled to achieve high levels of accuracy in situations when people modify their appearance by wearing different garments or by carrying objects such as bags and coats, as people's appearance can be altered by the things they are carrying.

Persons can be recognized by their silhouette using model-based [2] and model-free [3] techniques. The former method examines the silhouette's joint location, joint angle, height, and stride length, and the latter considers the subject's appearance. Identifying humans based on gait uses one complete gait cycle.

A GEI is a model-free/appearance-based method to construct features. The problem of identifying people by analyzing their GEIs is approached from several different angles, including the application of a wide range of ML methods and neural network topologies. However, these efforts typically run into the problem of not adequately resolving variances in walking angles or clothing situations. This constraint draws attention to a substantial gap in the already available literature, establishing favorable conditions for the development of more advanced methods.

In [4], the MPGR-CF architecture was used to extract dynamic and spatiotemporal features, trained at the decision level using SVM and HMM classifiers on the OU-ISIR and CASIA B datasets, demonstrating a high correct recognition rate. In [5], GEI features were used with an MSVM classifier to evaluate performance in both supportive and adversarial scenarios, but the impact of training data size and the degree of Carrying Status (CS) label led to recognition difficulties. The shape invariance in [6] suggested features in grid frameworks, horizontal and vertical silhouettes, on the OU-ISIR and CASIA B datasets based on statistical moments, offering a higher degree of precision. Several classifiers, including RF, NB, KNN, and DT (C4.5), were used to assess the resilience of the system. In [7], a cross-view gait recognition method used adversarial domain adaptation. In [8], a part-based deep learning method was proposed for cross-view silhouette gait recognition.

In the GEINet model [9], a GEI is fed to two triplet layers of convolution, pooling, and normalization, and two fully connected layers with a softmax classifier. In [10], activity was monitored using gait analysis with the help of different ML algorithms. In a two-stage network [11], a 3D CNN was used for identifying the point angle for the CASIA B dataset, which worked well for multi-angle. In [12], a Siamese network used contrastive loss for verification and a triplet network with triplet ranking loss for identification. In [13], three raw features, gray pixels, optical flow channels, and depth maps, were input to a CNN, gait features were extracted, and fused features were given to a softmax classifier. In VGGNet-SVM [14], the GEI-supplied image descriptor was retrieved from the fully connected layer of the pre-trained VGGNet-16 model. The results showed that for gait-based gender detection, SVM outperformed softmax with an accuracy of 87.94%. CAMNet [15] is a deep gait CNN for feature extraction on the view angle, performing classification using KNN and SVM.

GaitPart [16] was tested on the CASIA B and OU-MVLP datasets, performing well compared to GEINet and GaitSet. In [17], multi-view gait recognition was performed using spatiotemporal features. In [18], pairwise GEIs with similar or dissimilar labels were fed to two parallel CNNs. In [19], robust deep learning features, produced by two fully connected pre-trained networks (VGG16 and AlexNet), were combined in a Kernel Extreme Learning Machine (KELM). In [20], a 3D CNN for gait recognition in multiview with spatio-temporal features was presented. In [21], a CNN was supplied with clothing-invariant GEIs, performing better than cutting-edge approaches on the OU-ISIR B dataset. In [22], motion information/descriptors were extracted using deep convolutional models, such as VGG-19 and ResNet, by providing five cropped optical flow patches of human gait images. In VGRNet [23], intra-gait cycle segments were fed to a pre-trained 3D CNN to extract spatiotemporal features to train an LSTM classifier, performing well on the CASIA B dataset. In [24], a deep CNN called ITCNet was proposed for person identification. In [25], activity recognition was performed using a BiLSTM and a ResDLCNN-GRU attention network, respectively, to be used for person recognition in a security-concerned scenario.

This study aimed to improve performance and reduce computational cost through the introduction of a unique ensemble-based methodology. Different types of pretrained CNNs, such as EfficientNetB0, EfficientNetB1, MobileNetV3Large, MobileNetV3Small, MobileNet, RegNetX002, RegNetX004, RegNetX008, RegNetY002, RegNetY004, and RegNetY008, were used to extract robust features that exhibit greater invariance. To obtain a better distinction between classes, GEI features were preprocessed using Kernel Principal Component Analysis (KPCA). Subsequently, a diverse range of ML classifiers, including KNN, SVM, XGBoost, and Random Forest, were trained and evaluated. Among these options, an ensemble consisting of a 5-neighbor KNN classifier trained on features extracted using MobileNetV3Small, RegNetY002, and RegNetY004 exhibited the best performance.

This work advances person identification based on GEIs, providing a strong and reliable solution that can operate effectively in many settings. The novelty of the proposed work lies in formulating an appropriate ensemble of TL and ML models on selective feature fusion strategies to reduce computational complexity, offering a light-weight, high-performance design, by proposing an optical flow technique for the calculation of the subject's viewing angle.

II. PROPOSED METHODOLOGY

This section outlines the suggested approach, which leverages pretrained deep neural networks and Machine Learning (ML) in covariant situations to identify humans based on their gait. Figure 1 outlines the method, followed by a mathematical representation. The proposed solution consists of two tasks, video processing and classification.

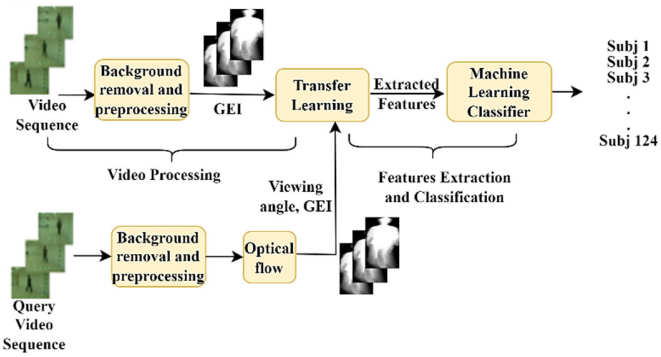

Fig. 1. Proposed method for human identification using TL and ML.

A. Mathematical Representation for Gait-Based Person Identification

The model assumptions are:

- Silhouettes are accurately extracted from video sequences
- Walking sequences represent steady gait cycles.

1) Step 1: Data Acquisition

Let V_n be the sequence of gait video frames for person n, where V_n^T represents normal walking, V_n^B represents walking with a backpack, and V_n^L represents walking with a long coat. Thus, the input dataset is:

$$D = \{V_n^T, V_n^B, V_n^L\}, \quad n \in \{1,2,\dots,N\}$$

where N is the total number of persons.

2) Step 2: Video/Data Preprocessing

Each video V_n consists of frames F_t, where t is the time index: $Fn = \{F_n^t \mid t = 1,2,\dots,T\}$

3) Step 3: Gait Energy Image (GEI) Generation

The GEI for each person n is computed as: $GEI_n = \frac{1}{T}\sum_{\{t=1\}}^{\{T\}} F_{\mathrm{n}}$, where GEI_n is the averaged silhouette across frames.

4) Step 4: Feature Extraction Using a TL and ML Ensemble

Define a $f_{\{\theta\}}$ representing the extracted features using a TL model with parameter θ, $f_{\{\theta\}}$: $GEI_n^T \rightarrow y_n$, where θ are the trainable parameters of the model.

Apply KPCA with an RBF kernel to map original high-dimensional data to a nonlinear subspace, increasing inter-class operability, $\tilde{F}_{\mathrm{n}}^{\mathrm{t}} = KPCA(F_{\mathrm{n}}^{\mathrm{t}}), \forall\, t \in \{1,\dots,T\}$, where $\tilde{F}_{\mathrm{n}}^{\mathrm{t}}$ represents the reduced feature representation.

- Use Label Binarization to encode identities:

  $y_n = LabelBinarizer(n) \in \mathbb{R}^{\mathrm{N}}$,

  where y_n is a label encoded identity label of vector length N.

- Train the model using ML classifiers: In the transformed-based space, similarity between features can be computed using distance-based metrics, which can be achieved using basic ML classifiers. The reduced features $\tilde{F}_n^t$ are subsequently used to train ML classifiers.

5) Step 5: Model Testing & Evaluation

- Use the trained model to predict on testing data for carrying a backpack and wearing a long coat:

  $\hat{y}_n^B = f_{\{\theta\}}\, GEI_n^B, \quad \hat{y}_n^L = f_{\{\theta\}}\, GEI_n^L$

  where $\hat{y}_n^B$ and $\hat{y}_n^L$ is the predicted identity.

- Evaluate performance using metrics such as accuracy

6) Step 6: Obtain the Viewing Angle

Apply an optical flow-based technique to get the viewing angle (illustrated in part B).

7) Step 7: Identify Identity

Apply the pretrained model ensemble with an ML classifier, evaluated from steps 4 and 5, to obtain the person's identity based on gait, which has been optimized as lightweight to function best at specified angles.

B. Minimization of Computational Cost for Query Video

Optical flow is a key idea in computer vision, used to estimate how an item moves between successive frames in an image sequence [26]. The brightness constancy constraint, which asserts that a pixel's intensity stays constant as it moves across time, is the fundamental premise of optical flow. Formally, this is expressed as:

$$I(x,y,t) = I(x+\delta x, y+\delta y, t+\delta t)$$

where $I(x,y,t)$ is the image intensity at point (x,y) at time t and $(x+\delta x, y+\delta y)$ is the location of the same point at time $(t+\delta t)$.

An optical flow-based technique is used to estimate the subject's viewing angle from video frames to reduce the computational complexity of gait detection. Since various angles result in different flow signatures, optical flow, which records motion patterns like direction and magnitude, is a useful indicator for determining the camera perspective. Following the estimation of the viewing angle, the system routes the input to one of several lightweight convolutional neural networks, namely MobileNetV3Small, RegNetY002, or RegNetY004, which have been optimized to function best at specified angles. Compared to processing all inputs with a single broad or generalized model, this angle-specific model selection ensures that only the most pertinent and effective model is utilized for feature extraction, thus reducing needless computation.

GEIs are produced from silhouette sequences connected to the input video after the angle has been estimated. To extract discriminative gait features, GEIs, which offer a condensed spatiotemporal representation of an individual's walking pattern, are subsequently run through the chosen lightweight model. In addition to increasing computational efficiency, this pipeline's modular design enables scalable deployment across many perspectives by adding more lightweight models as required. Additionally, this method maintains minimal resource consumption, appropriate for real-time or edge-based applications, while supporting strong performance in both same-view and cross-view gait detection scenarios.

The mathematical model for minimizing computational cost can be described as shown the following algorithm.

Algorithm 1
Let:
$\mathcal{V} = \{v_1, v_2, \dots, v_{\mathrm{n}}\}$ be the set of viewing angles,
$\mathcal{M} = \{M_1, M_2, \dots, M_{\mathrm{n}}\}$ be pretrained models for each angle.
$x \in \mathcal{X}$ be the input query video.
$\mathcal{F}(x) \rightarrow v_{\mathrm{i}} \in \mathcal{V}$ is the viewing angle estimator (optical flow).

$C(M_i)$: computational cost of model M_i.
$A(M_i, x)$: accuracy of M_i on input x.
Objective – Minimize expected cost:
$min_i \ \mathbb{E}_x \in \mathcal{X} \ [\ C(M_i)\ ] \ \ subject\ to\colon \ A(M_i, x) \geq \delta$
with probabilistic angle estimation.
If $\mathcal{F}(x)$ (the optical flow method) gives a probability distribution over angles, i.e.
$\mathcal{F}(x) = \{(v_1, p_1), (v_2, p_2), \ldots, (v_n, p_n)\}$
Then the expected cost can be written as:
$min_i \sum_j p_j \cdot C(M_i)$
subject to: $\sum_j p_j \cdot A(M_j, x) \geq \delta$
With the generalization model M_0:
$min_i \in \{0, \ldots, n\} \ [\ C(M_i) + \lambda \cdot (1 - A(M_i, x))]$
where:
M_0 is a general-purpose model (e.g., a multitask or ensemble model),
λ is a regularization factor balancing cost and accuracy.

This mathematical model helps in formalising the proposed system as a process for making decisions:

- Uses optical flow to estimate the viewing angle
- Chooses the pretrained model tailored to that particular angle.
- Reduces expenses while maintaining a high level of precision.

This mathematical model can ensure a decrease in computational complexity.

III. EXPERIMENTAL RESULTS

The CASIA-B dataset [27, 28] was used for experimentation on the Jupyter platform on an Intel Core i7 processor, 8 GB RAM, and an NVIDIA RTX 1080 Ti GPU. The dataset consists of gait-corresponding silhouettes of 124 subjects in three walking variations, Normal Walking (6 variations, NM), walking with a Carrying Bag (2 variations, BG), and walking while wearing a Coat (2 variations, CL), in 11 different viewing angles, 0°, 18°, 36°, ...,180°. Each silhouette is cropped to a size of 80×120.

The features are extracted by using T1-EfficientNetB0, T2-EfficientNetB1, T3-MobileNetV3Large, T4-MobileNetV3Small, T5-MobileNet, T6-RegNetX002, T7-RegNetX004, T8-RegNetX008, T9-RegNetY002, T10-RegNetY004, and T11-RegNetY008 (all models renamed as TN-Modelname, where N is from 1 to 11). The extracted features are preprocessed through the use of KPCA to reduce dimensionality.

ML classifiers, such as KNN (5-neighbor), SVM, XGBoost, and Random Forest, are used to classify the subject. The dataset was partitioned into training and testing (NM, BG, CL). The training part consists of a Normal walk (Nm-01, Nm-02, Nm-03, Nm-04), Bg-01, Cl-01 of all subjects, and the testing contains Nm-05, Nm-06, Bg-02, and Cl-02 silhouettes.

Figure 2 shows representative silhouettes of a normal walking of subject 001 at 36° and its GEI. This GEI underwent dimensionality reduction using KPCA and was then fed to the pre-trained networks. Human recognition with ML classifiers was performed for 11 different viewing angles. Accuracy was considered for correct identification:

$$Accuracy = \frac{(TP+TN)}{(TP+FP+TN+FP)}$$

where TP is True Positives, i.e., cases that the model correctly predicted, TN is True Negatives, i.e., cases that the model correctly predicted other classes, and FP and FN are False Positives and False Negatives, i.e., cases that the image belongs to one class but the model predicted another class and vice versa.

(a) (b)

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

Table I shows the accuracy (%) of the pre-trained networks with the KNN classifier for all angles. MobileNetV3Small with KNN with 5 neighbors offered satisfactory accuracies at 94.90%, 92.07%, 94.47%, and 95.53% for the 18°, 36°, 144°, and 162° angles, respectively. MobileNetV3Small, RegNetY002, and RegNetY004 also worked well.

TABLE I. ACCURACY OF PRE-TRAINED NETWORKS WITH KNN CLASSIFIER FOR ALL ANGLES

Angle	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11
0°	92.6	91.1	95.0	95.6	85.9	91.1	89.0	90.0	95.7	**95.8**	94.1
18°	89.8	85.9	92.5	**94.9**	72.8	89.7	83.9	82.7	92.6	91.9	89.7
36°	86.5	86.8	89.5	**92.1**	69.4	86.0	79.6	82.5	91.8	88.7	89.3
54°	88.7	86.4	91.7	92.9	64.1	88.3	82.2	83.4	**93.7**	92.8	91.2
72°	88.6	91.3	91.9	94.2	78.4	86.8	85.8	81.6	91.3	**94.5**	91.4
90°	89.6	88.0	91.1	92.6	82.8	86.4	83.3	82.3	**92.7**	91.9	87.8
108°	87.3	86.3	88.2	92.3	80.4	86.5	80.3	75.9	**92.7**	90.7	87.9
126°	86.9	87.6	87.0	92.7	80.0	88.5	82.3	82.9	**92.8**	91.2	89.9
144°	89.9	87.6	89.0	**94.5**	81.0	86.5	82.8	84.6	93.7	91.1	90.7
162°	89.8	90.0	91.7	**95.5**	83.4	87.8	83.8	84.9	93.7	94.4	92.1
180°	92.5	90.2	94.6	94.2	88.6	90.3	88.2	89.1	**94.7**	94.1	92.7
Avg	89.3	88.3	91.1	**93.8**	78.8	88.0	83.7	83.6	93.2	92.5	90.6

Table II shows the accuracy (%) of the pre-trained networks with the SVM classifier. MobileNetV3Small with KNN with 5 neighbors achieved satisfactory accuracies at 93.83%, 90.60%, 90.47%, 92.40%, 93.83%, 91.90%, 90.13%, and 95.43% for 0°, 18°, 36°, 54°, 90°, 126°, 144°, and 180°, respectively. MobileNetV3Small, RegNetY004, and RegNetY002 also achieved good performance.

TABLE II. ACCURACY OF PRE-TRAINED NETWORK WITH SVM CLASSIFIER FOR ALL ANGLES

Angle	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11
0°	91.0	86.8	92.1	**93.8**	77.1	85.6	84.8	83.8	93.4	93.6	91.8
18°	87.9	81.5	86.8	**90.6**	65.7	79.2	79.8	75.7	90.1	87.9	85.1
36°	84.0	84.8	88.1	**90.5**	66.5	80.4	74.3	78.6	90.2	89.5	86.8
54°	84.4	86.9	87.9	**92.4**	61.2	81.7	79.3	79.3	92.2	92.3	88.5
72°	84.8	88.2	94.2	93.4	79.8	80.2	80.6	78.3	92.6	**95.6**	92.3
90°	87.5	86.7	90.2	**93.8**	75.5	80.5	81.0	79.4	92.9	92.2	88.6
108°	84.6	84.8	88.2	91.8	75.4	78.8	76.3	74.4	**92.5**	90.5	87.0
126°	80.6	82.5	85.9	**91.9**	73.2	80.0	75.2	75.6	91.1	90.2	85.0
144°	85.8	83.7	84.5	**90.1**	75.5	78.3	75.5	77.3	89.9	89.3	84.1
162°	86.8	82.5	87.9	89.9	73.6	79.1	75.4	76.3	90.6	**91.1**	87.3
180°	91.0	91.3	92.5	**95.4**	85.9	81.3	81.5	85.0	93.6	91.6	91.8
Avg	86.2	85.4	88.9	**92.2**	73.6	80.5	78.5	78.5	91.7	91.3	88.0

Table III shows the accuracy (%) of the pre-trained networks with the Random Forest classifier for all angles. RegNetY002 with a Random Forest classifier achieved higher accuracies at 79.40%, 71.10%, 65.17%, 64.83%, 64.27%, 66.23%, 78.20% for the 0°, 18°, 36°, 54°, 72°, 90°, and 180° angles, respectively.

TABLE III. ACCURACY OF PRE-TRAINED NETWORK WITH RANDOM FOREST CLASSIFIER FOR ALL ANGLES

Angle	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11
0°	75.0	72.7	72.2	76.9	46.8	70.0	65.6	60.8	**79.4**	74.6	72.4
18°	64.5	61.5	64.9	69.5	37.2	55.1	45.7	53.5	**71.1**	69.3	61.0
36°	58.6	60.7	55.1	60.7	35.2	47.1	40.7	39.1	**65.2**	59.4	51.6
54°	62.7	53.9	55.0	59.4	31.1	48.0	43.2	45.1	**64.8**	58.9	52.0
72°	60.9	61.3	58.6	63.3	38.1	46.8	43.8	37.2	**64.3**	59.5	57.8
90°	59.5	55.7	55.2	62.6	36.7	47.8	45.7	35.6	**66.2**	56.1	54.8
108°	53.2	57.8	57.0	**64.8**	36.6	48.4	39.6	33.9	63.0	55.7	52.8
126°	57.5	57.7	53.3	**66.0**	41.0	47.2	42.9	38.2	62.1	62.4	49.8
144°	62.0	55.9	57.6	**65.0**	43.5	47.1	46.3	37.5	63.4	62.5	54.5
162°	65.5	64.3	67.2	**75.9**	44.9	52.3	51.8	51.1	72.1	65.5	60.1
180°	71.9	75.1	76.5	77.3	53.2	63.2	60.6	61.0	**78.2**	75.7	67.2
Avg.	62.9	61.5	61.1	67.4	40.4	52.1	47.8	44.8	**68.2**	63.6	57.6

Table IV shows the accuracy (%) of the pre-trained networks with the XGBoost classifier for all angles. MobileNetV3Small with XGBoost achieved accuracies of 68.77%, 65.87%, 56.17%, 57.53%, 58.57%, 67.07% for the 0°, 18°, 36°, 54°, 144°, and 180° angles, respectively. RegNetY002 also achieved close performance. Figure 3 shows the average accuracy of pre-trained networks with an ML classifier.

TABLE IV. ACCURACY OF PRE-TRAINED NETWORK WITH XGBOOST CLASSIFIER FOR ALL ANGLES

Angle	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10	T11
0°	48.3	58.6	56.7	**68.8**	39.8	52.6	55.4	46.7	60.5	63.3	55.3
18°	61.1	54.4	57.4	**65.9**	35.6	50.0	47.0	46.9	61.2	59.8	52.4
36°	50.4	50.5	46.9	**56.2**	34.4	45.3	40.9	39.8	53.0	50.3	50.3
54°	54.9	49.5	50.3	**57.5**	30.5	41.8	40.2	44.0	54.4	52.4	48.0
72°	53.4	53.8	52.9	52.7	39.8	43.6	40.5	37.6	**57.0**	53.8	52.7
90°	53.5	51.6	**54.2**	53.3	34.1	45.2	40.6	39.4	53.9	51.2	50.0
108°	49.2	50.7	51.6	53.1	37.9	41.0	37.9	35.5	**55.4**	50.8	46.1
126°	54.3	52.4	51.3	51.7	36.7	45.5	40.5	37.6	52.8	**55.2**	47.1
144°	55.9	55.6	51.3	**58.6**	39.7	44.3	41.1	36.4	54.4	53.2	48.6
162°	57.1	58.9	56.2	59.8	39.6	50.5	44.4	48.4	**60.2**	57.6	51.9
180°	59.3	64.4	61.4	**67.1**	50.0	58.3	53.2	52.8	65.9	65.9	57.5
Avg	54.3	54.6	53.7	**58.6**	38.0	47.1	43.8	42.3	57.1	55.8	50.9

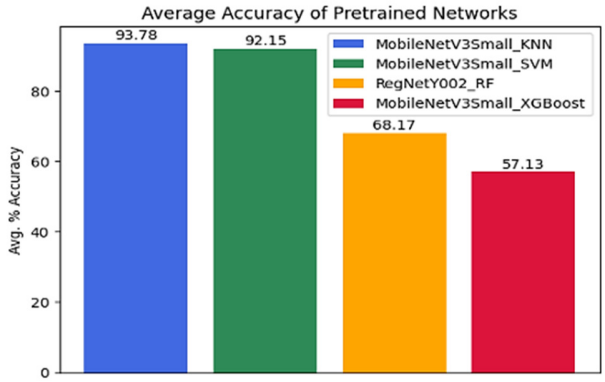

Fig. 3. Average accuracy (%) of pre-trained networks with ML classifier for all angles.

These results show that the MobileNetV3Small pretrained network with KNN (5 neighbors) has achieved the best performance considering all viewing angles. KNN is a nonparametric, instance-based classifier, capable of handling complex class distributions. A one-way ANOVA was used to test whether the means of three or more groups are significantly different from each other. The null hypothesis (H_0) is $\mu_1 = \mu_2 = \mu_3 = \cdots = \mu_k$, where $\mu_1, \mu_2, \mu_3, \cdots, \mu_k$ are the means of the k groups. In this hypothesis, the means of all groups are equal, and any differences observed are due to random chance, not actual group effects. The alternative hypothesis (H_a) is that at least one $\mu_i \neq \mu_j$ for some $i \neq j$. This means at least one group's mean is statistically significantly different, suggesting that an effect or difference exists. Table V shows a statistically significant difference in mean probe performance across the four classifiers. KNN achieved the best performance, followed closely by SVM, while RF and XGB lag significantly behind.

TABLE V. STATISTICAL VALIDATION

Classifier	KNN	SVM	RF	XGB
EfficientNetB0 (T1)	89.29	86.22	62.85	54.29
EfficientNetB1 (T2)	88.30	85.43	61.51	54.58
MobileNetV3Large (T3)	91.12	88.93	61.14	53.66
MobileNetV3Small (T4)	93.78	92.15	67.39	58.60
MobileNet (T5)	78.78	73.59	40.40	38.00
RegNetX002 (T6)	88.00	80.46	52.09	47.10
RegNetX004 (T7)	83.73	78.53	47.82	43.78
RegNetX008 (T8)	83.64	78.53	44.82	42.28
RegNetY002 (T9)	93.24	91.73	68.17	57.13
RegNetY004 (T10)	92.48	91.25	63.60	55.76
RegNetY008 (T11)	90.61	88.03	57.64	50.91

TABLE VI. ANOVA: SINGLE FACTOR SUMMARY

Groups	Count	Sum	Average	Variance		
KNN	11	972.94	88.45	21.92		
SVM	11	934.86	84.99	39.63		
RF	11	627.42	57.04	88.06		
XGB	11	556.11	50.56	45.89		
Source of variation	SS	df	MS	F	P-value	F crit
Between groups	12218.82	3	4072.941	83.33456	2.98E-17	2.838745
Within groups	1954.983	40	48.87457			
Total	14173.81	43				

Table VII compares the recognition rate/accuracy of the proposed system with the existing ones.

TABLE VII. COMPARISON WITH THE EXISTING SYSTEMS

Reference	Methodology	Accuracy
[22]	VGG + blocks	65.10%
[29]	WideResNet	89.40%
[30]	ResNet50	93%
	InceptionV3	85%
	NASNet Mobile	80.10%
	NASNet Large	75.80%
	InceptionResNet	89.20%
	VGG16	92.50%
	VGG19	91.50%
	ResNet101	91%
	ResNet152	87.20%
Proposed	MobileNetV3Small + KNN (GEI)	93.78%

These results show that for GEI features, MobileNetV3Small with a KNN classifier performs well, with 93.78% accuracy compared to existing systems. The performance of the KNN-based ensemble can be attributed to the nature of the extracted gait features and the classifier's ability to exploit local similarity structures in the feature space. The deep features obtained from lightweight pretrained CNNs, followed by KPCA preprocessing, result in compact and well-separated class distributions.

The computational complexity can be calculated in terms of the expected cost. The optical flow technique gives the probability distribution over the subject's viewing angle. After testing, the TL with ML ensemble that gives the best accuracy at each viewing angle was selected, as shown in Tables I, II, III, and IV. The mathematical model proposed to minimize computational cost ensures a decrease in computational complexity by selecting a specified model only for a particular viewing angle.

IV. CONCLUSION

Recognizing persons through GEIs presents considerable difficulties, especially when subjects exhibit different walking angles or wear different types of clothing. This paper presented a unified mathematical formulation for strong, cost-effective optical flow and an ensemble-based strategy to effectively tackle these issues. Through the utilization of a hybrid approach using pretrained CNNs to extract features and KPCA to preprocess them, this approach successfully improved the distinguishability between different classes. A detailed assessment was conducted using different ML classifiers, showing that an ensemble of KNN classifier trained using MobileNetV3Small, RegNetY002, and RegNetY004 models achieved the highest level of effectiveness.

In training, the accuracy of recognition achieved a peak of 99.2% and rarely fell below 95.6%, despite the subject's angle ranging from 0° to 180°. In testing, the recognition accuracy exhibited a range of 92.07% to 95.57% across the same angles, demonstrating the resilience of the proposed method. Overall, the obtained results reflect a considerable advancement in the state of GEI-based recognition and meet the demand for accurate and adaptable to real-world applications.

However, the system's performance may change due to severe or prolonged occlusion of silhouettes, significant clothing variations influenced by demographic diversity, and real-world deployment scenarios, as the proposed approach was evaluated under a controlled environment. In a real-world scenario, challenges such as background clutter, shadow effects, etc, need to be addressed.

DECLARATION OF COMPETING INTERESTS

The authors declare no competing interests.

ACKNOWLEDGMENT

Not applicable to this work.

DATA AVAILABILITY

The dataset used in this work is publicly available at [27]. Codes are available from the corresponding author upon request.

FUNDING SOURCES

This research did not receive any specific grant from any of the funding agencies.

REFERENCES

[1] E. Y. Abdullah and L. I. Jalil, "Effective stride length on load carrying capacity for ankle joint," *Journal of Interdisciplinary Mathematics*, vol. 24, no. 8, pp. 2345–2353, Nov. 2021, https://doi.org/10.1080/09720502.2021.1984565.

[2] G. Ariyanto and M. S. Nixon, "Model-based 3D gait biometrics," in *2011 International Joint Conference on Biometrics (IJCB)*, Oct. 2011, pp. 1–7, https://doi.org/10.1109/IJCB.2011.6117582.

[3] S. Shirke, S.S.Pawar, and K. Shah, "Literature Review: Model Free Human Gait Recognition," in *2014 Fourth International Conference on Communication Systems and Network Technologies*, Apr. 2014, pp. 891–895, https://doi.org/10.1109/CSNT.2014.252.

[4] X. Wang and S. Feng, "Multi-perspective gait recognition based on classifier fusion," *IET Image Processing*, vol. 13, no. 11, pp. 1885–1891, Sept. 2019, https://doi.org/10.1049/iet-ipr.2018.6566.

[5] Md. Z. Uddin *et al.*, "The OU-ISIR Large Population Gait Database with real-life carried object and its performance evaluation," *IPSJ Transactions on Computer Vision and Applications*, vol. 10, no. 1, Dec. 2018, Art. no. 5, https://doi.org/10.1186/s41074-018-0041-z.

[6] A. Nandy, R. Chakraborty, and P. Chakraborty, "Cloth invariant gait recognition using pooled segmented statistical features," *Neurocomputing*, vol. 191, pp. 117–140, May 2016, https://doi.org/10.1016/j.neucom.2016.01.002.

[7] T. Huang, X. Ben, C. Gong, W. Xu, Q. Wu, and H. Zhou, "GaitDAN: Cross-View Gait Recognition via Adversarial Domain Adaptation," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 34, no. 9, pp. 8026–8040, Sept. 2024, https://doi.org/10.1109/TCSVT.2024.3384308.

[8] Md. Z. Uddin, K. Hasan, M. A. R. Ahad, and F. Alnajjar, "Horizontal and Vertical Part-wise Feature Extraction for Coss-view Gait Recognition," *IEEE Access*, pp. 1–1, 2024, https://doi.org/10.1109/ACCESS.2024.3513541.

[9] K. Shiraga, Y. Makihara, D. Muramatsu, T. Echigo, and Y. Yagi, "GEINet: View-invariant gait recognition using a convolutional neural network," in *2016 International Conference on Biometrics (ICB)*, June 2016, pp. 1–8, https://doi.org/10.1109/ICB.2016.7550060.

[10] Y. L. Ng, X. Jiang, Y. Zhang, S. B. Shin, and R. Ning, "Automated Activity Recognition with Gait Positions Using Machine Learning Algorithms," *Engineering, Technology & Applied Science Research*, vol. 9, no. 4, pp. 4554–4560, Aug. 2019, https://doi.org/10.48084/etasr.2952.

[11] D. Thapar, A. Nigam, D. Aggarwal, and P. Agarwal, "VGR-net: A view invariant gait recognition network," in *2018 IEEE 4th International Conference on Identity, Security, and Behavior Analysis (ISBA)*, Jan. 2018, pp. 1–8, https://doi.org/10.1109/ISBA.2018.8311475.

[12] N. Takemura, Y. Makihara, D. Muramatsu, T. Echigo, and Y. Yagi, "On Input/Output Architectures for Convolutional Neural Network-Based Cross-View Gait Recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 29, no. 9, pp. 2708–2719, Sept. 2019, https://doi.org/10.1109/TCSVT.2017.2760835.

[13] F. M. Castro, M. J. Marín-Jiménez, N. Guil, and N. Pérez De La Blanca, "Multimodal feature fusion for CNN-based gait recognition: an empirical comparison," *Neural Computing and Applications*, vol. 32, no. 17, pp. 14173–14193, Sept. 2020, https://doi.org/10.1007/s00521-020-04811-z.

[14] T. Liu, X. Ye, and B. Sun, "Combining Convolutional Neural Network and Support Vector Machine for Gait-based Gender Recognition," in *2018 Chinese Automation Congress (CAC)*, Nov. 2018, pp. 3477–3481, https://doi.org/10.1109/CAC.2018.8623118.

[15] S. Muhammed and E. Çelebi, "CAMNet: DeepGait Feature Extraction via Maximum Activated Channel Localization," *Intelligent Automation & Soft Computing*, vol. 28, no. 2, pp. 397–416, 2021, https://doi.org/10.32604/iasc.2021.016574.

[16] C. Fan *et al.*, "GaitPart: Temporal Part-Based Model for Gait Recognition," in *2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, June 2020, pp. 14213–14221, https://doi.org/10.1109/CVPR42600.2020.01423.

[17] S. Gul, M. I. Malik, G. M. Khan, and F. Shafait, "Multi-view gait recognition system using spatio-temporal features and deep learning," *Expert Systems with Applications*, vol. 179, Oct. 2021, Art. no. 115057, https://doi.org/10.1016/j.eswa.2021.115057.

[18] C. Zhang, W. Liu, H. Ma, and H. Fu, "Siamese neural network based gait recognition for human identification," in *2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Mar. 2016, pp. 2832–2836, https://doi.org/10.1109/ICASSP.2016.7472194.

[19] M. A. Khan *et al.*, "Human gait analysis for osteoarthritis prediction: a framework of deep learning and kernel extreme learning machine," *Complex & Intelligent Systems*, vol. 9, no. 3, pp. 2665–2683, June 2023, https://doi.org/10.1007/s40747-020-00244-2.

[20] T. Wolf, M. Babaee, and G. Rigoll, "Multi-view gait recognition using 3D convolutional neural networks," in *2016 IEEE International Conference on Image Processing (ICIP)*, Sept. 2016, pp. 4165–4169, https://doi.org/10.1109/ICIP.2016.7533144.

[21] T. Yeoh, H. E. Aguirre, and K. Tanaka, "Clothing-invariant gait recognition using convolutional neural network," in *2016 International Symposium on Intelligent Signal Processing and Communication Systems (ISPACS)*, Oct. 2016, pp. 1–5, https://doi.org/10.1109/ISPACS.2016.7824728.

[22] A. Sokolova and A. Konushin, "Pose-based deep gait recognition," *IET Biometrics*, vol. 8, no. 2, pp. 134–143, Mar. 2019, https://doi.org/10.1049/iet-bmt.2018.5046.

[23] D. Thapar, G. Jaswal, A. Nigam, and C. Arora, "Gait metric learning siamese network exploiting dual of spatio-temporal 3D-CNN intra and LSTM based inter gait-cycle-segment features," *Pattern Recognition Letters*, vol. 125, pp. 646–653, July 2019, https://doi.org/10.1016/j.patrec.2019.07.008.

[24] M. Babaee, L. Li, and G. Rigoll, "Person identification from partial gait cycle using fully convolutional neural networks," *Neurocomputing*, vol. 338, pp. 116–125, Apr. 2019, https://doi.org/10.1016/j.neucom.2019.01.091.

[25] A. Dey, S. Biswas, and L. Abualigah, "Efficient Violence Recognition in Video Streams using ResDLCNN-GRU Attention Network," *ECTI Transactions on Computer and Information Technology (ECTI-CIT)*, vol. 18, no. 3, pp. 329–341, July 2024, https://doi.org/10.37936/ecti-cit.2024183.255679.

[26] B. K. P. Horn and B. G. Schunck, "Determining optical flow," *Artificial Intelligence*, vol. 17, no. 1–3, pp. 185–203, Aug. 1981, https://doi.org/10.1016/0004-3702(81)90024-2.

[27] Y. Fu, Shibei Meng, Saihui Hou, Xuecai Hu, and Yongzhen Huang, "CASIA-B Pose." Science Data Bank, Aug. 17, 2023, https://doi.org/10.57760/SCIENCEDB.10304.

[28] Y. Fu, S. Meng, S. Hou, X. Hu, and Y. Huang, "GPGait: Generalized Pose-based Gait Recognition," in *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*, Oct. 2023, pp. 19538–19547, https://doi.org/10.1109/ICCV51070.2023.01795.

[29] H. Chao, Y. He, J. Zhang, and J. Feng, "GaitSet: Regarding Gait as a Set for Cross-View Gait Recognition," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 33, no. 01, pp. 8126–8133, July 2019, https://doi.org/10.1609/aaai.v33i01.33018126.

[30] K. Apostolidis, P. Amanatidis, and G. Papakostas, "Performance Evaluation of Convolutional Neural Networks for Gait Recognition," in *24th Pan-Hellenic Conference on Informatics*, Nov. 2020, pp. 61–63, https://doi.org/10.1145/3437120.3437276.