

Article

Robust and Lightweight System for Gait-Based Gender Classification toward Viewing Angle Variations

Jaychand Upadhyay¹ and Tad Gonsalves^{2,*} 

¹ Department of Information Technology, Xavier Institute of Engineering, Mumbai 400016, India; jaychand.u@xavier.ac.in

² Department of Information and Communication Sciences, Faculty of Science & Technology, Sophia University, Tokyo 102-8554, Japan

* Correspondence: t-gonsal@sophia.ac.jp

Abstract: In computer vision applications, gait-based gender classification is a challenging task as a person may walk at various angles with respect to the camera viewpoint. In some of the viewing angles, the person's limb movement can be occluded from the camera, preventing the perception of the gait-based features. To solve this problem, this study proposes a robust and lightweight system for gait-based gender classification. It uses a gait energy image (GEI) for representing the gait of an individual. A discrete cosine transform (DCT) is applied on GEI to generate a gait-based feature vector. Further, this DCT feature vector is applied to XGBoost classifier for performing gender classification. To improve the classification results, the XGBoost parameters are tuned. Finally, the results are compared with the other state-of-the-art approaches. The performance of the proposed system is evaluated on the OU-MVLP dataset. The experiment results show a mean CCR (correct classification rate) of 95.33% for the gender classification. The results obtained from various viewpoints of OU-MVLP illustrate the robustness of the proposed system for gait-based gender classification.



Citation: Upadhyay, J.; Gonsalves, T. Robust and Lightweight System for Gait-Based Gender Classification toward Viewing Angle Variations. *AI* **2022**, *3*, 538–553. <https://doi.org/10.3390/ai3020031>

Academic Editors: José Manuel Ferreira Machado and Kenji Suzuki

Received: 16 May 2022

Accepted: 10 June 2022

Published: 14 June 2022

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2022 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: gait; gender classification; biometrics; XGBoost; OU-MVLP

1. Introduction

Gait refers to the walking style of a person. Every person possesses a unique gait that can be utilized as a behavioral feature in recognition of that person. The system does not need any help from the person for gait acquisition, since it can be done from a distance. Hence, in recent years, gait recognition has gained popularity and it can be used in several fields, such as inspection of criminal activity, surveillance, access monitoring, and forensics. As CCTV is now installed in almost all public places [1–3], it captures the gait of the person unobtrusively.

In the recent research literature, the majority of the work has been done in gait recognition [4–14], whereas gender classification based on gait has a huge potential in real-time applications [15]. Gait-based gender classification can provide a cue in investigation and help in solving the problem of finding lost persons, especially children at railway stations, airports, bus stands, and other public places. The gait-based gender classification can also be used in commercial activities. For example, if the gender of the customer entering the shop is known, then related advertisements can be displayed on the digital screen. A talking robot installed in shopping malls can perceive the customers through cameras and help them with gender-related queries, such as finding toilets, beauty product shops, men's wear shops, and women's wear shops.

The majority of the studies done in gender classification [16–23] use gait energy image (GEI) [6]. GEI combines the silhouettes from one complete walking cycle. The brightness of every picture element in GEI reveals the gait dynamic of one complete walking cycle. Recent studies usually perform gait-based gender recognition using fixed

direction data [20,24], which means the recognition is only for a fixed viewpoint of the subject. But it is not always the same viewpoint in real-time. Unfortunately, among the recent research literature [13–18,25–29], very few have conducted gait-based gender classification experiments for viewing angle variations on a very large dataset, which can be considered as the main motivation of this study.

1.1. Research Gap and Motivation

- The performance of gender classification based on gait features mainly relies on the camera viewing angle. It can be observed in [27] that non-frontal viewing angles ($22.5^\circ / 45^\circ / 67.5^\circ / 90^\circ$) provide a straightforward way to distinguish the leg joint angles in the depth image XY-plane. However, in other viewing angles, the subjects' limb movement can be occluded from the camera, which hinders perceiving the gait features. Recent studies [18,28–30] have tried to investigate the problem of viewing angle variations in gait-based gender classification. However, they have used a relatively smaller dataset for performance evaluation of the results.
- In recent times, most studies [16,25,31–33] have adopted convolution neural network (CNN) based methods to perform gait-based gender classification. These methods demand an extensive amount of training, which can be accomplished with high-performance GPU machines. The end result is a high implementation cost. To overcome this drawback, this study makes a cost-effective attempt by proposing a lightweight system for gait-based gender classification.

1.2. Main Contributions

This study attempts to handle the research gap mentioned in Section 1.1 in the following manner:

- It adopts a lightweight approach for gait-based gender classification under viewing angle variations of subjects.
- It initially takes the input image, extracts silhouette, constructs GEI, applies discrete cosine transform (DCT) for feature extraction, and finally applies a XGBoost classifier for gender classification.
- It verifies the experimental results using the world's largest multi-view gait dataset (OU-MVLP) by confirming the efficiency and effectiveness of the proposed system against the results produced by the state-of-the-art models.

The contents of this paper are arranged as follows: Section 2: Related work, Section 3: Proposed system, Section 4: Experiments, and Section 5: Conclusions.

2. Related Work

Gait analysis methods are categorized into model-based approach [16,34–36], appearance-based approach [17,18,21,37], and deep learning-based approach [31,32,38,39]. Based on these approaches, gait-based systems can be categorized into model-based systems, appearance-based systems, and deep learning-based systems.

2.1. Model-Based Systems

The model-based approaches attempt to model the subject's body to perform gait-based gender classification and recognition tasks. These systems depend on the acquisition of a 3D skeletal model of subjects [40–42]. The benefit of 3D modeling in a gait-based gender classification and recognition system is that it can successfully tackle the viewing angle variation using the skeletal models. Since the essential parameters in the skeletal model are evaluated in 3D with the help of several calibrated 2D cameras or depth-sensing cameras, the outcome of the system shows robustness to viewing angle variations. Therefore, these systems attain a recognition accuracy of 92% for 5 various viewpoints [41]. Likewise, 3D dense models developed from several 2D cameras play a vital role in handling viewing angle variations [43]. These systems can transform the features retrieved from a model to a

particular viewpoint, attaining the accuracy of recognition of 75% over 12 different viewing angles of 20 subjects.

The disadvantages of the 3D model-based systems include things like camera calibration prior to usage, occlusion of gait, and limited range of depth-sensing cameras. Therefore, these systems become unproductive in the natural environment.

Yoo et al. [35] introduced a technique in which the gait outline is rearranged in the form of a two-dimensional link coordinate diagram and used as a distinct feature for recognition tasks. Lee et al. [34] partitioned the gait outline of the subject's body into 7 segments such that every segment is depicted in an elliptical shape. Further, the elliptical center, major axis, minor axis, and orientation of every ellipse are computed as distinct features. Isaac et al. [26] proposed a gender recognition system that removes the necessity of requiring a complete gait cycle by using the pose-based voting (PBV) method. The authors used linear discriminant analysis (LDA) in addition to the Bayes' rule for classification as an alternative to the popular support vector machine (SVM). Guffanti et al. [28] proposed a method that uses depth cameras to perceive most human gait features. They included 81 participants (40 females and 41 males) in their experiment. These participants were asked to walk at a self-selected speed across a 4.8-m walkway. A detailed analysis was done in the time domain. Further, the features with significant differences by gender were used to train a support vector machine (SVM) classifier. Lee et al. [29] proposed a gender recognition method that uses a support vector machine (SVM) and random forest (RF) based on recursive feature elimination to determine the best features. They investigated temporal, kinematics, and muscle activity to show the effects of gender-based differences on gait characteristics.

In general, it is observed that model-based systems fail to handle lower resolution input pictures; additionally, they end up increasing the computational complexity.

2.2. Appearance Based Systems

The appearance-based approaches depend on the spatio-temporal data captured from the gait dynamics of a person to perform gender classification and recognition. The appearance-based method mostly relies on a complete gait cycle of a subject to perform gender classification. The most famous appearance-based method is GEI. Lu et al. [44] introduced a gender classification system that is capable of handling the arbitrary movement of subjects. In this study, the gait sequences of subjects are compared and assigned to a particular cluster, depending on the comparison result. Finally, a cluster-based averaged gait image is computed, which is further used as a gait-based feature. Liu et al. [45] proposed a Fourier transform on gait energy image to extract the gait-based features.

In another study [46], the technique for gender classification is introduced in which the silhouette from every picture is extracted and gait dynamics over a certain time period are calculated to trace the person. Later, support vector machines (SVM) is applied for classification. The study [47] applied GEI and active energy image (AEI) in combination with the k-NN classification algorithm to perform gender classification. This study demonstrated the systems' performance on CASIA-B and SOTON-A public datasets. The study [48] utilized GEI and denoised energy images as distinct features and then applied SVM to perform gender classification. Bei et al. [16] introduced a method to calculate a subGEI using fewer frames rather than a complete gait cycle. They extracted the synthetic optical flow of multi-subGEIs and used them as temporal features. Further, they applied a two-stream CNN to combine the use of GEI and the optical flow information for further gait analysis. Hassan et al. [17] used a wavelet 5/3 lifting scheme for gait representation. They performed PCA to form dimensional distinctive vectors for each walk sequence.

The significance of appearance-based systems includes lower complexity for computation and reduced noise as it works directly on gait silhouette. Hence, this study uses an appearance-based approach.

2.3. Deep Learning-Based Systems

The deep learning approaches involving convolution neural networks (CNN) can be used for multiple tasks, such as gender classification, age estimation, and recognition, simultaneously. In [38], it can be observed that gender classification, age estimation, and recognition of subjects are carried out simultaneously. In [31], CNN is used with GEI; silhouette images are treated as input, and body mass index is computed as output. Zhang et al. [32] introduced a deep CNN, which performs multiple task learning for estimating the age along with gender classification. Sakata et al. [39] introduced CNN involving multiple stages to handle gait-based age and gender estimation. Liu et al. [33] used a VGGNet-16 deep convolution model in combination with SVM to perform gait-based gender recognition.

According to the available literature, it can be observed that deep learning-based systems for gait-based gender classification have produced superior results. However, these systems need a very high configuration of hardware systems [25]. The proposed system attempts to solve gender classification problems using a lightweight method by adapting an appearance-based approach.

3. Proposed System for Gender Classification

Figure 1 depicts the detailed functioning of gait-based gender classification. The following subsections explain the proposed system in detail.

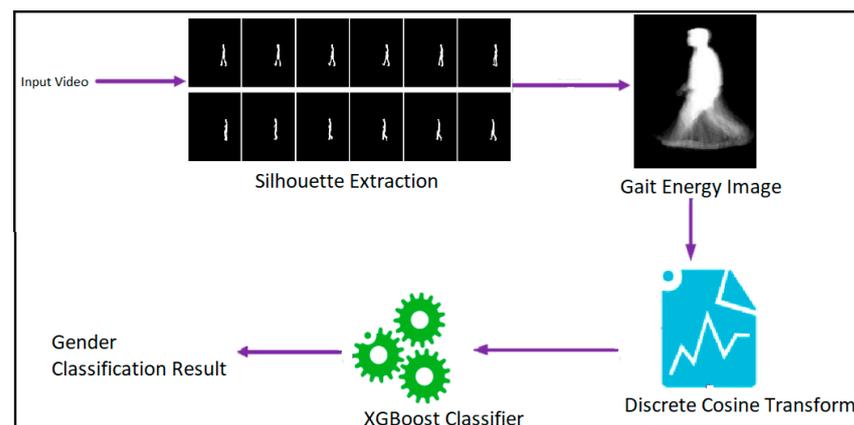


Figure 1. Proposed system for gait-based gender classification.

3.1. Silhouette Extraction

According to a study by Aslam et al. [49], the Gaussian mixture-based background/foreground segmentation algorithm is found suitable for moving object detection. Their experiment shows satisfactory results for background subtraction. It is also found that the Gaussian-based background/foreground segmentation algorithm is capable of detecting the moving object even though it is occluded. Therefore, this study performs silhouette extraction using the aforementioned algorithm. This algorithm partitions the pixels according to their intensity value into foreground and background.

Let M be the input video, (x_0, y_0) the pixel position, and t the time. The history of pixel position (x_0, y_0) can be represented with the following equation:

$$G_1, G_2, \dots, G_t = \{M(x_0, y_0, i) : 1 \leq i \leq t\} \quad (1)$$

where G_t represents video frames at instant t .

Let k = number of distributions, ω = associated weight with the i th Gaussian, Σ = standard deviation, μ = mean, σ = variance, and T = threshold value. The following equations illustrate the silhouette extraction process:

$$P(G_t) = \sum_{i=1}^k \omega_{i,t} R(G_t | \mu_{i,t}, \Sigma_{i,t}) \quad (2)$$

where $R(G_t | \mu_{i,t}, \Sigma_{i,t}) = \frac{1}{(2\pi)^{D/2}} \cdot \frac{1}{|\Sigma_{i,t}|^{1/2}} \exp(-\frac{1}{2}(G_t - \mu_{i,t})^T \Sigma_{i,t}^{-1} (G_t - \mu_{i,t}))$

$$C = \operatorname{argmin} \left(\sum_{i=1}^C \omega_{i,t} > T \right) \quad (3)$$

$$\left((G_{t+1} - \mu_{t+1})^T \sum_{i=1}^C (G_{t+1} - \mu_{t+1}) \right)^{0.5} > k \cdot \sigma_{i,t} \quad (4)$$

Once silhouette extraction is performed, a median filter of size 8×8 is applied to the silhouette image to remove the noise.

3.2. Gait Energy Image (GEI)

After silhouette extraction and noise removal, GEI is evaluated. The majority of appearance-based methods have adapted the GEI. In this experiment, we have fixed the GEI size as 88×128 pixels. Let $N_{G.F.}$ = number of frames in a gait cycle of an individual, i = frame sequence number, (x, y) = representation of image coordinate, and S_i = gait frame. GEI can be computed as follows:

$$GEI(x, y) = \frac{\sum_{i=1}^{N_{G.F.}} S_i(x, y)}{N_{G.F.}} \quad (5)$$

Figure 2 illustrates a pictorial representation of GEI.



Figure 2. Illustration of GEI.

Justification for GEI Representation

According to Equation (5), GEI is an average template; therefore, it is not affected by random noises in single silhouette frames. Further, the robustness may be enhanced by dropping the pixels whose energy values are below a threshold. In addition, with GEI templates, it is not required to split the sequence of silhouettes into cycles and carry out time normalization of the cycle length. Thus, the errors arising in this approach can be avoided. In comparison to binary silhouette, the GEI has an apparent information loss. The intensity value for a particular pixel in GEI indicates the frequency of silhouette taking place at that position over the whole sequence. We can partially rebuild the original silhouette sequence from the GEI with the knowledge of human walking. For instance, for the pixel closer to the leg contour, its GEI value indicates that the silhouette is taking place at this position in 30 frames out of 120 frames. These 30 frames are those frames in which persons are moving. Likewise, we can assign the GEI values to other limb movement regions for corresponding frames in the silhouette sequence. Generally, energy changes in the head and torso regions are regarded as noise. GEI can maintain the crucial contour of a person's walk, and it is also helpful to understand the changes in walking.

According to the study by J. Han et al. [6], GEI is more capable of saving both storage space and computation time for gait recognition than binary silhouette sequences. Their study also concluded that GEI is less sensitive to noise as compared to the individual frame. Therefore, this study uses GEI for gait representation.

3.3. Discrete Cosine Transform (DCT)

According to N. Ahmed et al. [50], DCT is preferred over Karhunen-Loeve Transform, Discrete Fourier Transform, Walsh-Hadamard Transform, and Haar Transform for pattern recognition application. In this study, comparing the performances of these transforms, DCT is found to be optimal. The study [51] compared the performance of Discrete Wavelet Transform (DWT) with DCT and concluded that DCT has excellent compaction for human image data. DCT provides an adequate trade-off between information packing ability and computational complexity, and is faster than DWT. Because of these advantages, this study applies DCT to the GEI for extracting the distinct gait features. Figure 3 illustrates the detailed process of DCT features extraction. The input GEI is divided into various blocks, with each block having a pixel size of 8×8 . We get 176 such blocks. Then, DCT is applied on every block from left to right in a top-down approach. These blocks are capable of representing the entire GEI with comparatively less memory requirement. The DCT matrix elements can be computed as:

$$DCT(x, y) = \frac{1}{\sqrt{2}}C(x)C(y) \sum_{m=0}^{N-1} \sum_{n=0}^{N-1} g(m, n) \cos \left[\frac{(2m + 1)x\pi}{2N} \right] \cos \left[\frac{(2n + 1)y\pi}{2N} \right] \quad (6)$$

where N represents the size of blocks, $C(u) = \begin{cases} \frac{1}{\sqrt{2}} & \text{if } u = 0 \\ 1 & \text{if } u > 0 \end{cases}$, and $g(m, n)$ represents the image matrix. Every DCT coefficient shows a specific spatial frequency. The first DCT coefficient, $X(0, 0)$, is the DC coefficient. The DC coefficient has zero frequency in both the horizontal and vertical directions. It specifies the brightness of the image block as it is computed by averaging the pixel values in the block. The other coefficients are termed the AC coefficients. The AC coefficients near the DC coefficient have lower spatial frequencies and the frequencies rise on moving away from the DC coefficient in all directions. AC coefficients react to gray level variations that are in the same direction as their spatial frequencies. AC coefficients values and signs are directly related to the strength, contour, and orientation of the movement in the image blocks. Once DCT is applied on GEI, a few coefficients are chosen for the feature vector, whereas others are discarded, resulting in the reduction of data dimensionality. The selected coefficient carries high energy components, inducing energy compaction in DCT.

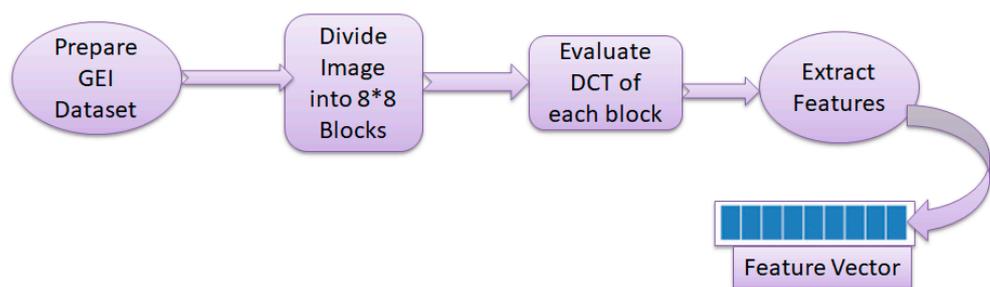


Figure 3. Flowchart for DCT feature extraction.

Figure 4 is a pictorial representation of the coefficients of DCT along with the GEI. The DCT coefficients matrix consists of three different frequency components—low, medium, and high. It is observed that low-frequency components are holding highly relevant and meaningful data. Since the low-frequency components are sufficient to regenerate the input image, these components have been utilized in DCT feature vector generation.

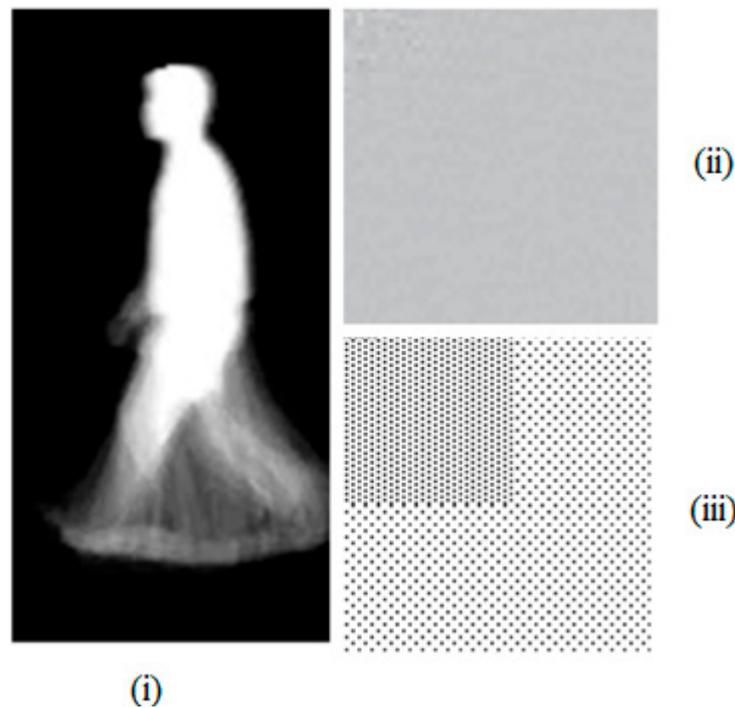


Figure 4. Illustration of DCT (i) GEI input, (ii) DCT Transform, and (iii) categorization of DCT coefficient.

3.4. XGBoost Classifier

According to [52], the performance of XGBoost is found to be superior to support vector regression and artificial neural network models. In addition, the performance of the XGBoost model has shown robustness to all input combinations, compared with the ANN and SVR models. Accordingly, this study used an XGBoost classifier for performing classification based on the features obtained from DCT. This classifier applies second-order gradients and advanced regularization for achieving higher accuracy in performance. The objective function of the classifier is computed as a summation of intermediate loss functions, which are observed in every iteration. The classifier also applies hessian to train the model by generating a tree. Hessian is described as a second-order derivative of the loss at any instant. The following steps explain the gender classification process.

Let X be the DCT features vector set obtained from the GEI of the individuals and Y a set of genders, $\alpha = \text{learning rate}$, and $M = \text{number of base learners}$.

$$X = [x_1, x_2, \dots, x_i]$$

$$Y = \{\text{male, female}\}$$

$$\text{Training Set} = \{(x_i, y_i)\}_{i=1}^N$$

Step (1) Preparing the model at the beginning

$$f_{(0)}(x) = \underset{\theta}{\operatorname{argmin}} \sum_{i=1}^N L(y_i, \theta) \tag{7}$$

Step (2) Performing iteration from $m = 1$ to M

(i) Evaluate hessian and gradient:

$$h_m(x) = \left[\frac{\partial^2 L(y_i, f(x_i))}{\partial f(x_i)^2} \right]_{f(x)=f_{(m-1)}(x)} \tag{8}$$

$$g_m(x) = \left[\frac{\partial L(y_i, f(x_i))}{\partial f(x_i)} \right]_{f(x)=f_{(m-1)}(x)} \tag{9}$$

where L is the log loss function.

(ii) Resolving optimization problem by fitting base learner with training pair

$$\left\{ (x_i, \frac{g_m(x_i)}{h_m(x_i)}) \right\}_{i=1}^N$$

$$\phi_m = \underset{\phi \in \Phi}{\operatorname{argmin}} \sum_{i=1}^N \frac{1}{2} h_m(x_i) \left[-\frac{g_m(x_i)}{h_m(x_i)} - \phi(x_i) \right]^2 \tag{10}$$

$$f_m(x) = \alpha \phi_m(x)$$

(iii) Upgrading the model

$$f_{(m)}(x) = f_{(m-1)}(x) + f_m(x) \tag{11}$$

Step (3) Estimating the result

$$f(x) = f_{(M)}(x) = \sum_{m=0}^M f_m(x) \tag{12}$$

Based on the outcome of step 3, the proposed system classifies the gender of the person under observation. Figure 5 illustrates the XGBoost classifier with the help of flowchart representation.

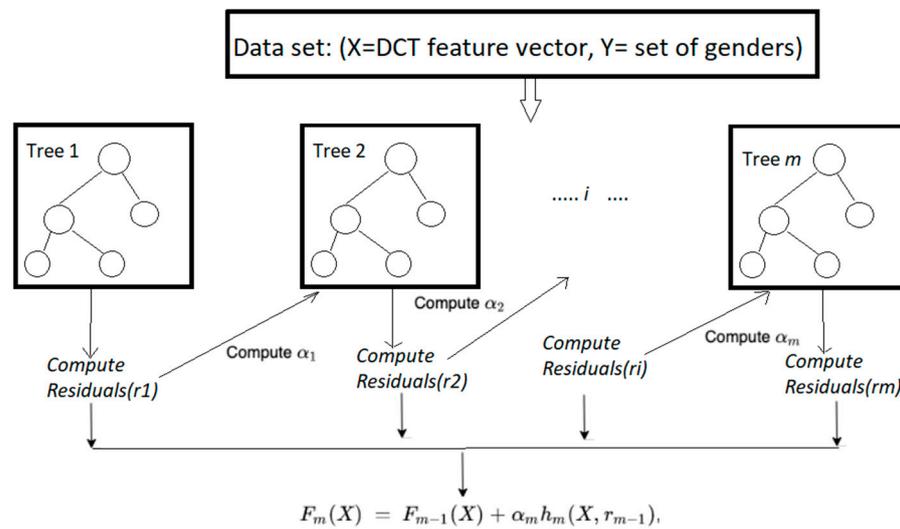


Figure 5. Illustration of XGBoost classifier with gradient tree.

4. Experiments

4.1. Dataset

This study demonstrates the performance of the proposed system for gait-based gender classification using the OU-MVLP dataset [53]. This dataset is developed by the Institute of Scientific and Industrial Research (ISIR), Osaka University (OU). It contains the gait of 10,307 individuals. Out of this, 5114 are male and 5193 are female. The gait recording of 10,307 individuals is done from 14 different viewing angles. These angles vary from 0° to 90° and 180° to 270° . The setup for gait recording includes seven network cameras, which are planted at the gap of 15° azimuth angles. These camera networks are planted at both ends of a walking track. The detailed setup (top view) for OU-MVLP is

illustrated in Figure 6. Here, orange-colored cameras are used for gait recording when a person walks from A to B, whereas blue-colored cameras are used for gait recording of a person moving from B to A. During the recording of gait, every individual is instructed to walk twice in forward (A to B) and backward (B to A) directions. Accordingly, 28 gait sequences of each individual are captured in this setup.

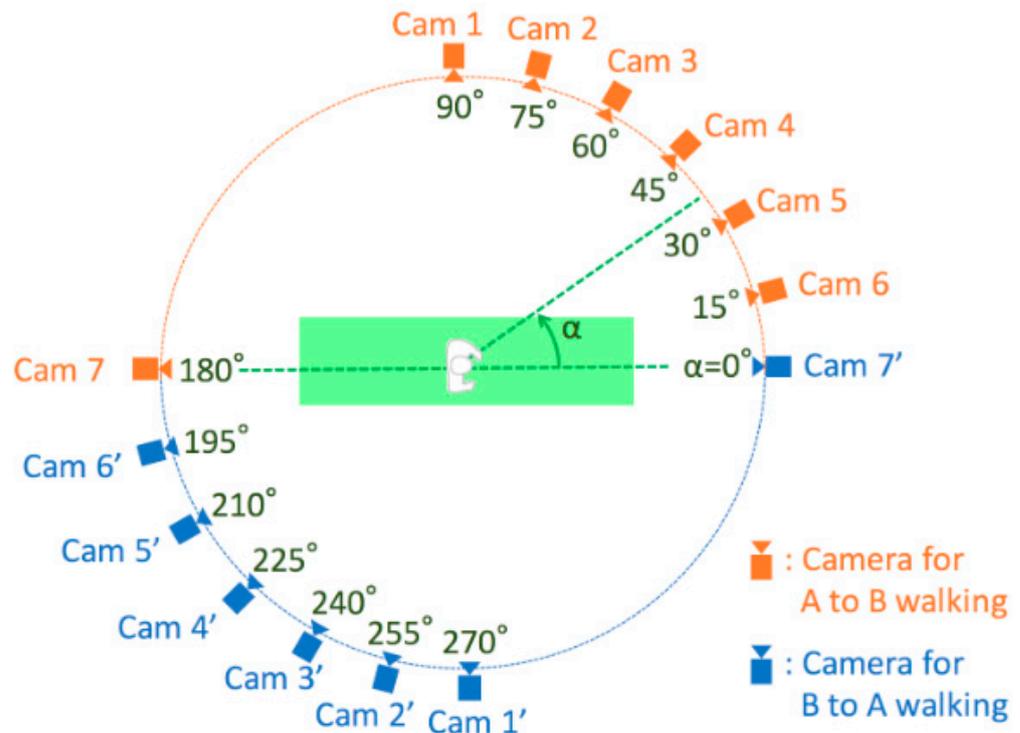


Figure 6. Illustration of gait capturing of OU-MVLP dataset (Adapted with permission from ref. [53] 2021, Yasushi Makihara).

The OU-MVLP gait dataset images are partitioned into two parts with similar image sizes. The first part contains gait images which are used for training, while the images in the other part are used for testing. This study adopts a similar train and test split for gait-based gender classification.

4.2. Classifier Tuning

This study tunes the XGBoost classifier by setting the value of boosting parameters like `max_depth`, `min_child_weight`, `gamma`, `subsample`, `colsample_bytree`, and `scale_pos_weight`. We have tuned the learning rate in the range from 0.2 to 0.4 according to the viewing angle variations. These learning rates have helped in the evaluation of the optimum number of trees required for classification. Once the learning rate and number of trees are evaluated, then the tree specific parameters were tuned as follows: `max_depth` = 3–5 (according to viewing angle variations), `min_child_weight` = 6, `gamma` = 0, `subsample` = 0.8, `colsample_bytree` = 0.8, `scale_pos_weight` = 1. To enhance the performance of the classifier, the regularization parameter `alpha` is tuned with a value of 0.005. Table 1 illustrates the tuning of `max_depth`, learning rate, and the number of estimators according to the viewing angle variations.

Table 1. Illustration of XGBoost parameters tuning with respect to viewing angle variations.

Viewing Angle	Max_Depth	Learning Rate	Number of Estimators
0°	5	0.2	1000
15°	4	0.3	900
30°	3	0.4	1000
45°	4	0.2	900
60°	5	0.2	900
75°	4	0.3	1000
90°	3	0.3	850
180°	4	0.2	1000
195°	5	0.2	850
210°	3	0.3	1000
225°	4	0.2	900
240°	5	0.3	900
255°	4	0.2	1000
270°	4	0.3	900

4.3. Performance Evaluation Criteria

The performance of the proposed system for gait-based gender classification is validated under the assumption that the gender information is known at the beginning. This study carried out the test for both correct and incorrect gender classification with respect to a specific individual to observe the effect of viewing angle variations on the gender classification. The correct classification rate (CCR) is evaluated by the following equation:

$$CCR = \frac{TP_g + TN_g}{N_g} \quad (13)$$

where, TP_g = true positive pointing out the scenario where the proposed system performs correct classification of positive samples (male samples), TN_g = true negative pointing out the scenario where the proposed system performs correct classification of negative samples (female samples), and N_g is the total number of samples. We have labeled the positive sample as 1 and the negative sample as -1 . The CCR depicts the correct classification of male and female samples.

4.4. Result Analysis

After tuning the classifier, the CCR for gender classification under every viewing angle is evaluated. Table 2 shows angle-wise CCR for gender classification. It is observed that the proposed system attains the highest CCR of 96.32% for gender classification under a viewing angle of 90°. In the side view of a person, the majority of the gait features are observed, which helps in classification. At 0° viewing angle, comparatively fewer features are observed; hence, the CCR declines to 92.85%. It can also be observed from the obtained results that when a person walks in the forward direction, the CCR obtained is comparatively better than when a person walks in the backward direction—refer to Figure 6 and Table 2.

Table 2. CCR for gender classification under each viewing angle.

Angle	CCR (in %)
0°	92.85
15°	94.7
30°	95.21
45°	95.35
60°	95.45
75°	96.09
90°	96.32
180°	94.45
195°	95.44
210°	95.92
225°	95.9
240°	95.41
255°	95.77
270°	95.7

4.5. Result Comparison and Discussion

In order to prove the robustness and efficiency of the proposed system, we have performed a statistical analysis of gait-based gender classification methods. Other studies excluding Gaitset [25] have evaluated the performance of their method for gender classification on a relatively smaller dataset. Nevertheless, the proposed system outperforms other methods. Table 3 enlists the state-of-art approaches for gait-based gender classification. The studies DGHEI [54], CNN+SVM [55], PBV-EFD (TUM-GAID) [26], and PBV-RCS (TUM-GAID) [26] have conducted experiments on a dataset containing the gait of 305 persons. Despite using a lower size dataset, the results obtained in the aforementioned studies do not match those of our study. The studies PBV-EFD (CASIA-B) [26], PBV-RCS (CASIA-B) [26], and SRML (CASIA-B) [45] have obtained better results, but these studies considered only 11 different viewing angles for calculation of mean CCR. Moreover, these studies conducted the experiments on a very small dataset of 62 persons, whereas our study conducted the experiment on the world's largest dataset, covering 14 different viewing angles. As can be seen in Table 3, the proposed system outperforms the studies by Hu et al. [56], SRML [45], PWC [30], and PWC+PF [30] in gender recognition accuracy. Although the studies Lifting scheme 5/3+PCA (OULP) [17], Lifting scheme 5/3+PCA (CASIA-B) [17], and SVM-RFE [29] show promising results, they have considered fewer viewing angle variations and used comparatively smaller datasets for gender classification. Thus, the Results section elucidates that DCT can successfully determine different frequency components and are insensitive to the changes in human appearances. Further, the GEI+DCT+XGBoost combination is capable of obtaining higher recognition rates than the methods using contour and texture-based features.

From the statistics presented in Table 3, it can be observed that mean CCR is inversely proportional to the number of viewing angles. As the number of viewing angles increases, the possibility of occlusion of gait features in several viewpoints also increases. This leads to the performance degradation of gender classification in those viewpoints. Thus, mean CCR reduces. However, the performance of gait-based gender classification depends on various factors, such as feature extraction technique, sample size, model training, and classification method, and the number of viewing angles, and so on.

Table 3. Statistical analysis of gait-based gender classification methods.

Gait-Based Gender Classification Method	No. of Samples	No. of Viewing Angles	Mean CCR (in %)
GaitSet [25]	10,307 (5114 are male and 5193 are female)	14	94.3
DGHEI [54]	305 (186 are male and 116 are female)	1	87.8
CNN+SVM [55]	305 (186 are male and 116 are female)	1	88.9
Hu et al. [56]	62 (31 are male and 31 are female)	11	94.36
PBV-EFD(CASIA-B) [26]	62 (31 are male and 31 are female)	11	95.34
PBV-EFD (TUM-GAID) [26]	305 (186 are male and 116 are female)	1	82.3
PBV-RCS(CASIA-B) [26]	62 (31 are male and 31 are female)	11	97.89
PBV-RCS(TUM-GAID) [26]	305 (186 are male and 116 are female)	1	78.2
SRML [45]	122 (85 are male and 37 are female)	2	93.7
SRML(CASIA-B) [45]	62 (31 are male and 31 are female)	11	98
PWC [30]	124 (93 are male and 31 are female)	11	73.26
PWC+PF [30]	124 (93 are male and 31 are female)	11	82.52
Lifting scheme 5/3+PCA(OULP) [17]	200 (100 are male and 100 are female)	4	97.50
Lifting scheme 5/3+PCA (CASIA-B) [17]	124 (93 are male and 31 are female)	11	97.98
SVM-RFE [29]	25 (12 are male and 13 are female)	1	99.11
Proposed Method	10,307 (5114 are male and 5193 are female)	14	95.33

To the best of our knowledge, very few studies have performed gender classification using the OU-MVLP gait dataset. This study attempts to demonstrate the performance of the proposed system under viewing angle variations by showing an angle-wise comparison. Table 4 illustrates angle-wise comparative analysis of the proposed study with GaitSet [25]. From Tables 3 and 4, it can be derived that the performance of the proposed system is superior to other methods for gait-based gender classification. The study [25] used a CNN-based method for feature extraction, whereas this study used DCT-based features for gender classification. This shows that the proposed method gives better results than the deep learning approach.

Table 4. Viewing angle-wise comparison of the proposed system with GaitSet [25].

Viewing Angle	GaitSet [25] CCR (in %)	Proposed System CCR (in %)
0°	91.5	92.85
15°	93.0	94.7
30°	94.5	95.21
45°	94.6	95.35
60°	95.0	95.45
75°	94.9	96.09
90°	95.7	96.32
180°	93.0	94.45
195°	93.6	95.44
210°	94.9	95.92

Table 4. *Cont.*

Viewing Angle	GaitSet [25] CCR (in %)	Proposed System CCR (in %)
225°	94.8	95.9
240°	94.4	95.41
255°	94.6	95.77
270°	94.9	95.7

In order to evaluate the statistical significance of Table 4, we assume μ_1 as the mean CCR of the study [25] and μ_2 as the mean CCR of the proposed study. Here, our null hypothesis is $H_0: \mu_1 = \mu_2$ and alternative hypothesis is $H_a: \mu_1 < \mu_2$. The test statistic is -2.887006 , and the p -value for the two-tailed test is 0.007729 . Since the p -value is *less than* the critical value, we can reject the null hypothesis at a 5% level of significance. Thus, we can accept the alternative hypothesis which states that the performance of the proposed method is superior to that of study [25].

4.6. Computational Efficiency

This study confirms the performance of a proposed system by showing an analysis of computational complexity for feature extraction and classification. The proposed system takes $O(N \log_2 N)$ time complexity [57] for feature extraction, $O(dn_{trees}x \log n)$ time complexity for training, and $O(dn_{trees})$ time complexity [58] for prediction of the classifier, where N is the size of a block, d is the depth of the tree, n_{trees} is the number of trees, x is the number of non-missing entries in training data, and n is the total input samples.

In order to measure the complexity of CNN-based architecture, the study [59] exploits the concept of Betti numbers and summarizes that Betti numbers can grow exponentially for deep network architecture (refer to Table 5). Therefore, in comparison to the complexity of several CNN-based (shallow and deep architectures) methods, the proposed system proves to be efficient in runtime complexity. The CNN-based systems represent exponential runtime complexity, whereas the proposed system represents quasilinear runtime complexity. Moreover, CNN-based systems need GPU [25] for implementation, whereas the proposed system is implemented on the following configuration: Intel(R) Core (TM) i5-3570K CPU @ 3.40 GHz, 16GB RAM, 64-bit Windows 10 operating system. Thus, the proposed system is computationally efficient and implementation-wise cost-effective.

Table 5. Runtime complexity analysis of shallow and deep CNN architectures (adapted from study [59]). Here l = hidden layer, h = hidden units, n = inputs, and r = degree of the polynomial.

Inputs	Layers	Activation Function	Upper Bound
n	3	threshold	$O(h^n)$
n	3	arctan	$O((n+h)^{n+2})$
n	3	Polynomial, degree r	$\frac{1}{2}(2+r)(1+r)^{n-1}$
1	3	arctan	h
n	any	arctan	$2^{h(2h-1)}O((nl+n)^{n+2h})$
n	any	tanh	$2^{h(h-1)/2}O((nl+n)^{n+h})$
n	any	Polynomial, degree r	$\frac{1}{2}(2+r^l)(1+r^l)^{n-1}$

5. Conclusions

A lightweight and robust system for gait-based gender classification is proposed. The proposed system uses GEI for human gait representation, DCT for extraction of the feature vector, and XGBoost for classification of gender. GEI is more capable of saving both storage

space and computation time for gait recognition than binary silhouette sequences. DCT provides an adequate trade-off between information packing ability and computational complexity; additionally, its performance time-wise is faster than DWT. The XGBoost model has shown robustness to all input combinations, and its performance is superior to that of support vector regression and artificial neural network models.

During this classification, the proposed system adopts an appearance-based approach for gait analysis. The performance of the proposed system is evaluated on the OU-MVLP dataset. Results obtained in the experiment are compared with conventional machine learning methods as well as deep learning methods to demonstrate the superior performance of the proposed system. The comparison results show superior CCR for each viewing angle and surpass the state-of-the-art models.

Most studies on gait-based gender recognition reported in the literature base their experiments on relatively smaller datasets, and the number of viewing angles considered is at most 11. The study targets the largest available data for classification and considers 14 different viewing angles. The proposed system also shows a mean CCR of 95.33% and runtime complexity of $O(N \log_2 N)$ for feature extraction, which is superior to CNN-based systems. Thus, it proves the robustness of the system against the viewing angle variations.

However, the proposed system does not consider other problems related to gait-based gender classification, such as occlusion of gait due to variations in clothing and bag carrying conditions. The future work of this study will focus on handling the problem of gait occlusion and making a system robust against partial access of gait.

Author Contributions: Conceptualization, J.U.; methodology, J.U. and T.G.; software, J.U.; validation, J.U. and T.G.; resources, J.U. and T.G.; data curation, J.U.; writing—original draft preparation, J.U.; writing—review and editing, T.G. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Bouchrika, I.; Goffredo, M.; Carter, J.; Nixon, M. On Using Gait in Forensic Biometrics. *J. Forensic Sci.* **2011**, *56*, 882–889. [[CrossRef](#)] [[PubMed](#)]
2. Iwama, H.; Muramatsu, D.; Makihara, Y.; Yagi, Y. Gait verification system for criminal investigation. *Inf. Media Technol.* **2013**, *8*, 1187–1199.
3. Lynnerup, N.; Peter, K.L. Gait as evidence. *IET Biom.* **2014**, *3*, 47–54. [[CrossRef](#)]
4. Upadhyay, J.L.; Gonsalves, T.; Katkar, V. A Lightweight System Towards Viewing Angle and Clothing Variation in Gait Recognition. *Int. J. Big Data Intell. Appl.* **2021**, *2*, 21–38. [[CrossRef](#)]
5. Gul, S.; Malik, M.I.; Khan, G.M.; Shafait, F. Multi-view gait recognition system using spatio-temporal features and deep learning. *Expert Syst. Appl.* **2021**, *179*, 115057. [[CrossRef](#)]
6. Han, J.; Bhanu, B. Individual recognition using gait energy image. *IEEE Trans. Pattern Anal. Mach. Intell.* **2006**, *28*, 316–322. [[CrossRef](#)]
7. Chtourou, I.; Fendri, E.; Hammami, M. Person re-identification based on gait via Part View Transformation Model under variable covariate conditions. *J. Vis. Commun. Image Represent.* **2021**, *77*, 103093. [[CrossRef](#)]
8. Wu, Z.; Huang, Y.; Wang, L.; Wang, X.; Tan, T. A Comprehensive Study on Cross-View Gait Based Human Identification with Deep CNNs. *IEEE Trans. Pattern Anal. Mach. Intell.* **2016**, *39*, 209–226. [[CrossRef](#)]
9. Makihara, Y.; Suzuki, A.; Muramatsu, D.; Li, X.; Yagi, Y. Joint intensity and spatial metric learning for robust gait recognition. In Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 21–26 July 2017; pp. 6786–6796.
10. Li, X.; Makihara, Y.; Xu, C.; Yagi, Y.; Ren, M. Joint Intensity Transformer Network for Gait Recognition Robust Against Clothing and Carrying Status. *IEEE Trans. Inf. Forensics Secur.* **2019**, *14*, 3102–3115. [[CrossRef](#)]

11. Takemura, N.; Makihara, Y.; Muramatsu, D.; Echigo, T.; Yagi, Y. On Input/Output Architectures for Convolutional Neural Network-Based Cross-View Gait Recognition. *IEEE Trans. Circuits Syst. Video Technol.* **2017**, *29*, 2708–2719. [[CrossRef](#)]
12. Chao, H.; He, Y.; Zhang, J.; Feng, J. Gaitset: Regarding gait as a set for cross-view gait recognition. In Proceedings of the 33th AAAI Conference on Artificial Intelligence (AAAI 2019), Honolulu, HI, USA, 27 January–1 February 2019.
13. Xu, C.; Makihara, Y.; Li, X.; Yagi, Y.; Lu, J. Cross-View Gait Recognition Using Pairwise Spatial Transformer Networks. *IEEE Trans. Circuits Syst. Video Technol.* **2020**, *31*, 260–274. [[CrossRef](#)]
14. Li, X.; Makihara, Y.; Xu, C.; Yagi, Y.; Ren, M. Gait recognition via semi-supervised disentangled representation learning to identity and covariate features. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 14–19 June 2020; pp. 13306–13316. [[CrossRef](#)]
15. Xu, C.; Makihara, Y.; Ogi, G.; Li, X.; Yagi, Y.; Lu, J. The OU-ISIR Gait Database comprising the Large Population Dataset with Age and performance evaluation of age estimation. *IPSJ Trans. Comput. Vis. Appl.* **2017**, *9*, 24. [[CrossRef](#)]
16. Bei, S.; Deng, J.; Zhen, Z.; Shaojing, S. Gender Recognition via Fused Silhouette Features Based on Visual Sensors. *IEEE Sens. J.* **2019**, *19*, 9496–9503. [[CrossRef](#)]
17. Hassan, O.M.S.; Abdulazeez, A.M.; TIRYAKI, V.M. Gait-based human gender classification using lifting 5/3 wavelet and principal component analysis. In Proceedings of the 2018 International Conference on Advanced Science and Engineering (ICOASE), Duhok, Iraq, 9–11 October 2018; pp. 173–178.
18. Kwon, B.; Lee, S. Joint Swing Energy for Skeleton-Based Gender Classification. *IEEE Access* **2021**, *9*, 28334–28348. [[CrossRef](#)]
19. Chen, L.; Wang, Y.; Wang, Y. Gender classification based on fusion of weighted multi-view gait component distance. In Proceedings of the 2009 Chinese Conference on Pattern Recognition, Nanjing, China, 4–6 November 2009; pp. 1–5.
20. Mannami, H.; Makihara, Y.; Yagi, Y. Gait analysis of gender and age using a large-scale multi-view gait database. In Proceedings of the 10th Asian Conference on Computer Vision, Queenstown, New Zealand, 8–12 November 2010; pp. 975–986.
21. Lu, J.; Wang, G.; Thomas, S.H. Gait-based gender classification in unconstrained environments. In Proceedings of the 21st International Conference on Pattern Recognition, Tsukuba, Japan, 11–15 November 2012.
22. Zhang, D.; Wang, Y. Using multiple views for gait-based gender classification. In Proceedings of the 26th Chinese Control and Decision Conference (2014 CCDC), Changsha, China, 31 May–2 June 2014; pp. 2194–2197.
23. Kalaiselvan, C.; SivananthaRaja, A. Robust gait-based gender classification for video surveillance applications. *Appl. Math. Inf. Sci.* **2017**, *11*, 1207–1215. [[CrossRef](#)]
24. Do, T.D.; Nguyen, V.H.; Kim, H. Real-time and robust multiple-view gender classification using gait features in video surveillance. *Pattern Anal. Appl.* **2019**, *23*, 399–413. [[CrossRef](#)]
25. Xu, C.; Makihara, Y.; Liao, R.; Niitsuma, H.; Li, X.; Yagi, Y.; Lu, J. Real-Time Gait-Based Age Estimation and Gender Classification from a Single Image. In Proceedings of the 2021 IEEE Winter Conference on Applications of Computer Vision (WACV), Waikoloa, HI, USA, 3–8 January 2021; pp. 3459–3469. [[CrossRef](#)]
26. Isaac, E.R.; Elias, S.; Rajagopalan, S.; Easwarakumar, K.S. Multiview gait-based gender classification through pose-based voting. *Pattern Recognit. Lett.* **2019**, *126*, 41–50. [[CrossRef](#)]
27. Yeung, L.-F.; Yang, Z.; Cheng, K.C.-C.; Du, D.; Tong, R.K.-Y. Effects of camera viewing angles on tracking kinematic gait patterns using Azure Kinect, Kinect v2 and Orbbec Astra Pro v2. *Gait Posture* **2021**, *87*, 19–26. [[CrossRef](#)]
28. Guffanti, D.; Brunete, A.; Hernando, M. Non-Invasive Multi-Camera Gait Analysis System and its Application to Gender Classification. *IEEE Access* **2020**, *8*, 95734–95746. [[CrossRef](#)]
29. Lee, M.; Lee, J.-H.; Kim, D.-H. Gender recognition using optimal gait feature based on recursive feature elimination in normal walking. *Expert Syst. Appl.* **2021**, *189*, 116040. [[CrossRef](#)]
30. More, S.A.; Deore, P.J. Gait-based human recognition using partial wavelet coherence and phase features. *J. King Saud Univ.-Comput. Inf. Sci.* **2020**, *32*, 375–383. [[CrossRef](#)]
31. Zhang, S.; Wang, Y.; Li, A. Gait-Based Age Estimation with Deep Convolutional Neural Network. In Proceedings of the 2019 International Conference on Biometrics (ICB), Crete, Greece, 4–7 June 2019; pp. 1–8. [[CrossRef](#)]
32. Zhang, Y.; Huang, Y.; Wang, L.; Yu, S. A comprehensive study on gait biometrics using a joint cnnbasedmethod. *Pattern Recognit.* **2019**, *93*, 228–236. [[CrossRef](#)]
33. Liu, T.; Ye, X.; Sun, B. Combining Convolutional Neural Network and Support Vector Machine for Gait-based Gender Recognition. In Proceedings of the 2018 Chinese Automation Congress (CAC), Xi'an, China, 30 November–2 December 2018; pp. 3477–3481. [[CrossRef](#)]
34. Lee, L.; Grimson, W.E.L. Gait analysis for recognition and classification. In Proceedings of the Fifth IEEE International Conference on Automatic Face Gesture Recognition, Washington, DC, USA, 21 May 2002; pp. 155–162. [[CrossRef](#)]
35. Yoo, J.-H.; Hwang, D.; Nixon, M.S. Gender classification in human gait using support vector machine. In Proceedings of the Advanced Concepts For Intelligent Vision Systems, Antwerp, Belgium, 18–21 September 2006; pp. 138–145.
36. Sudha, L.; Bhavani, R. Gait based Gender Identification using Statistical Pattern Classifiers. *Int. J. Comput. Appl.* **2012**, *40*, 30–35. [[CrossRef](#)]
37. Hu, M.; Wang, Y. A New Approach for Gender Classification Based on Gait Analysis. In Proceedings of the 2009 Fifth International Conference on Image and Graphics, Xi'an, China, 20–23 September 2009; pp. 869–874. [[CrossRef](#)]

38. Marn-Jimnez, M.J.; Castro, F.M.; Guil, N.; de la Torre, F.; Medina-Carnicer, R. Deep multi-task learning for gait-based biometrics. In Proceedings of the 2017 IEEE International Conference on Image Processing (ICIP), Beijing, China, 17–20 September 2017; pp. 106–110.
39. Sakata, A.; Takemura, N.; Yagi, Y. Gait-based age estimation using multi-stage convolutional neural network. *IPSJ Trans. Comput. Vis. Appl.* **2019**, *11*, 4. [[CrossRef](#)]
40. Sohrab, R.; Mahdi, B. Human gait recognition using body measures and joint angles. *Int. J. Sci. Knowl. Comput. Inf. Technol.* **2015**, *6*, 10–16.
41. Wang, Y.; Sun, J.; Li, J.; Zhao, D. Gait recognition based on 3D skeleton joints captured by kinect. In Proceedings of the IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, USA, 25–28 September 2016.
42. Zhao, G.; Liu, G.; Li, H.; Pietikainen, M. 3D gait recognition using multiple cameras. In Proceedings of the 7th International Conference on Automatic Face and Gesture Recognition, Southampton, UK, 10–12 April 2006.
43. Muramatsu, D.; Shiraiishi, A.; Makihara, Y.; Yagi, Y. Arbitrary view transformation model for gait person authentication. In Proceedings of the 2012 IEEE Fifth International Conference on Biometrics: Theory, Applications and Systems (BTAS), Arlington, VA, USA, 23–27 September 2012; pp. 85–90. [[CrossRef](#)]
44. Lu, J.; Wang, G.; Moulin, P. Human Identity and Gender Recognition From Gait Sequences with Arbitrary Walking Directions. *IEEE Trans. Inf. Forensics Secur.* **2013**, *9*, 51–61. [[CrossRef](#)]
45. Liu, Y.-Q.; Wang, X. Human Gait Recognition for Multiple Views. *Procedia Eng.* **2011**, *15*, 1832–1836. [[CrossRef](#)]
46. Stanley, A.W. Gait-Based Gender classification Using Zernike Moments. Support Vector Machines Linear Discriminant Classifier Nearest Neighbors. 2016. Available online: <https://www.semanticscholar.org/paper/Gait-Based-Gender-classification-Using-Zernike-Stanley/0c0aab94889dff7388a803c896a8f43a761d63e0> (accessed on 15 May 2022).
47. Martín Félez, R.; García Jiménez, V.; Sánchez Garreta, J.S. Gait-based Gender Classification Considering Resampling and Feature Selection. *J. Image Graph.* **2013**, *1*, 85–89. [[CrossRef](#)]
48. Chen, Y.; Yang, Y.; Lee, J. Gait based gender classification using Kinect sensor. In Proceedings of the 122nd ASEE Annual Conference & Exposition, Seattle, WA, USA, 14–17 June 2015.
49. Aslam, N.; Sharma, V. Foreground detection of moving object using Gaussian mixture model. In Proceedings of the 2017 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 6–8 April 2017; pp. 1071–1074. [[CrossRef](#)]
50. Ahmed, N.; Natarajan, T.; Rao, K. Discrete Cosine Transform. *IEEE Trans. Comput.* **1974**, *C-23*, 90–93. [[CrossRef](#)]
51. Hemachandran, K.; Justus Rabi, B. Performance Analysis of Discrete Cosine Transform and Discrete Wavelet Transform for Image Compression. *J. Eng. Appl. Sci.* **2018**, *13*, 436–440.
52. Osman, A.I.A.; Ahmed, A.N.; Chow, M.F.; Huang, Y.F.; El-Shafie, A. Extreme gradient boosting (Xgboost) model to predict the groundwater levels in Selangor Malaysia. *Ain Shams Eng. J.* **2021**, *12*, 1545–1556. [[CrossRef](#)]
53. Takemura, N.; Makihara, Y.; Muramatsu, D.; Echigo, T.; Yagi, Y. Multi-view large population gait dataset and its performance evaluation for cross-view gait recognition. *IPSJ Trans. Comput. Vis. Appl.* **2018**, *10*, 4. [[CrossRef](#)]
54. Hofmann, M.; Geiger, J.; Bachmann, S.; Schuller, B.; Rigoll, G. The TUM Gait from Audio, Image and Depth (GAID) database: Multimodal recognition of subjects and traits. *J. Vis. Commun. Image Represent.* **2014**, *25*, 195–206. [[CrossRef](#)]
55. Castro, F.M.; Marín-Jiménez, M.J.; Guil, N.; de la Blanca, N.P. Automatic learning of gait signatures for people identification. In Proceedings of the International Work-Conference on Artificial Neural Networks, Cadiz, Spain, 14–16 June 2017; Springer: Cham, Switzerland, 2017; pp. 257–270.
56. Hu, M.; Wang, Y.; Zhang, Z. Maximisation of mutual information for gait-based soft biometric classification using Gabor features. *IET Biom.* **2012**, *1*, 55–62. [[CrossRef](#)]
57. Pang, C.-Y.; Zhou, R.-G.; Hu, B.-Q.; Hu, W.; El-Rafei, A. Signal and image compression using quantum discrete cosine transform. *Inf. Sci.* **2018**, *473*, 121–141. [[CrossRef](#)]
58. Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In Proceedings of the KDD' 16: The 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; Association for Computing Machinery: New York, NY, USA, 2016; pp. 785–794. [[CrossRef](#)]
59. Bianchini, M.; Scarselli, F. On the Complexity of Neural Network Classifiers: A Comparison Between Shallow and Deep Architectures. *IEEE Trans. Neural Netw. Learn. Syst.* **2014**, *25*, 1553–1565. [[CrossRef](#)]