Applied Acoustics 173 (2021) 107701

Contents lists available at ScienceDirect

Applied Acoustics

journal homepage: www.elsevier.com/locate/apacoust

# A novel biometric recognition method based on multi kernelled bijection octal pattern using gait sound

Emrah Aydemir [a], Turker Tuncer [b,*], Sengul Dogan [b], Musa Unsal [a]

[a] Department of Computer Engineering, Faculty of Engineering and Architecture, Kirsehir Ahi Evran University, Kirsehir, Turkey
[b] Department of Digital Forensics Engineering, Technology Faculty, Firat University, Elazig, Turkey

ARTICLE INFO

*Article history:*
Received 11 March 2020
Received in revised form 14 September 2020
Accepted 25 September 2020
Available online 28 October 2020

*Keywords:*
Gait recognition
Biometrics
Multi kernelled bijection octal pattern
Information fusion
Sound recognition

ABSTRACT

*Background:* Many gait based methods have been presented about biometric identification in the literature. Gait recognition methods have generally used images and sensors signals. In this work, a novel gait based biometric recognition method is presented. A novel Multi Kernelled Bijection Octal Pattern (MK-BOP) is presented in this study.
*Object:* The main aim of the proposed MK-BOP is to extract distinctive and comprehensive features from a signal (gait sound). By using the proposed MK-BOP, a novel biometric recognition method is proposed. Gait sounds are collected, and two novel datasets are collected. The first dataset is a noisy and heterogeneous dataset. The second dataset is a clear and homogenous dataset. A multileveled method is presented to authenticate subjects from these datasets. One dimensional discrete wavelet transform (1D-DWT) is applied to sound signal with Symlet 6 (sym6) filter, and levels are calculated.
*Conclusion:* The proposed MK-BOP generates features from each level signals, and the generated features are concatenated. A hybrid feature selector (RFNCA) selects the most discriminative feature, and selected most discriminative features are forwarded to classifiers. 0.980 and 0.949 success rates were achieved for clear and noisy datasets, respectively.



## 1. Introduction

Gait recognition is an essential behavioral biometric feature used to identify people [1,2]. Behavioral biometry identifies people by evaluating the characteristics of people such as writing style, speech style, signature style with appropriate algorithms [3]. For example, every person has a typical walking style. Distinguishing this style with computerized systems is among the important research topics of today [4,5]. Especially considering that very different diagnosis systems are developed with computer systems [6–11], gait recognition is easy to sample compared to other biometric features. Every person's gait has a characteristic, and it is difficult to imitate. In this way, recognition can be realized even without the knowledge of people [12]. Thus, people can be automatically identified from the camera images in places such as airports and shopping centers. These advantages are an essential way of identifying the people wanted, especially in crowded places [13–15]. However, gait recognition systems also have some difficulties. One of the most important of these difficulties is the correct differentiation of the body part. At the same time, viewing angle, different clothes, lighting, gait speed, crowded places significantly affect gait recognition performance [16–19]. Many methods have been developed to prevent this situation from turning into a disadvantage. These methods can be divided into three as follows [20].

- Obtaining salient gait features
- Evaluating gait features according to the view angle
- Creating a three-dimensional model

The primary purpose of these methods is to ensure that gait is evaluated under different conditions [21]. Gait recognition methods can also be used in many different situations such as identification, disease, gender, and human mood recognition [22,23]. The studies in these different situations in the literature are presented in Table 1.

In this method, we collected a novel two gait sound datasets and proposed a novel signal classification method. The presented signal classification method is multileveled, and these levels are calculated by using *four* leveled 1D-DWT with a sym6 filter [48]. MK-BOP extracts features from raw gait signal and DWT coefficients of it. ReliefF [49] and Neighborhood Component Analysis

* Corresponding author.
*E-mail address:* turkertuncer@firat.edu.tr (T. Tuncer).

https://doi.org/10.1016/j.apacoust.2020.107701
0003-682X/

**Table 1**
Related work.

| | Purpose | Method | Dataset | Evaluation Criteria |
|---|---|---|---|---|
| Liao et al. method (2020) [24] | Gait recognition based on human prior knowledge and body pose | PoseGait, Spatio-temporal feature | CASIA B Gait Dataset [19], CASIA E Dataset [25] | Recognition rate, the computational cost |
| Yao et al. method (2019) [26] | Gait recognition based on model-based features and model-free features | Hybrid descriptors | CASIA GaitDatabase B [19] | Accuracy |
| Tong et al. method (2019) [27] | Cross view gait recognition | Restrictive triplet network | USF [28], OU-ISIR [19,28] and CASIA-B [19] dataset | Recognition score |
| Verlekaret et al. method (2018) [29] | Gait recognition based on silhouettes | Segmentation, identification, and rectification method | KU IR database [30] | Recognition rate |
| Wu et al. method (2018) [31] | Gait recognition | Convolutional neural network | OU-ISIR [19,28] and CASIA-B [19] dataset | Accuracy, average time |
| Alotaibi and Mahmood method (2017) [32] | Improved gait recognition | Deep convolutional neural network | CASIA-B [19] dataset | Computation time, correct classification rate, accuracy |
| Sun et al. method (2019) [33] | Gait analysis | Transfer learning and neural network | HMDB [34] and UCF [35] dataset | Time |
| Rubio-Solis et al. method (2019) [36] | Walking activities | Interval Type-2 Fuzzy Logic and Multilayer Neural Networks | Collected data | Box-and-whisker diagrams, accuracy, confusion matrix |
| Sun et al. method (2020) [37] | Gait-based identification | Multiplematching algorithm | OU-ISIR [19,28] dataset | Recognition rate |
| Yu et al. method (2017) [38] | Gait recognition | Stacked Progressive Auto-Encoders | SZU RGB-D [39] Gait Dataset and CASIA Gait Dataset B [19] | Recognition rate |
| Lempereur et al. method (2020) [40] | Gait disorders in children | Recurrent neural network | Collected data | Velocity and position |
| Pepa et al. method (2020) [41] | Gait detection for Parkinson's Disease | Fuzzy logic | Collected data | ROC curve, sensitivity, specificity |
| Battistone and Petrosino method (2019) [42] | Gait recognition | Time-based Graph Long Short-Term Memory | TUM-GAID [43], CASIA Gait B [19], MSR Action3D [44] and CAD-60 [45] dataset | Accuracy |
| Deng et al. method (2016) [46] | Human gait recognition | Deterministic learning | CASIA-B gait dataset [19] and CMU MoBo gait dataset [47] | Recognition rate |

(NCA) [50] are used together to propose a 2-layered NCARF feature selector. The used NCARF selects 128 most discriminative features, and these features are utilized as the input of the classifiers. The used classifiers are Decision Tree (DT) [51], Support Vector Machine (SVM) [51], Linear Discriminant (LD) [52], K Nearest Neighbor (KNN) [53] and Subspace Discriminant (SD) [54] classifiers. Key contributions of the proposed MK-BOP based gait sounds authentication method are;

- Gait based biometric recognition systems generally use sensors (gyroscope, accelerometer) data or images. In this study, gait sounds are used for biometric recognition. Two novel datasets are collected to perform gait sounds based biometric authentication.
- A novel feature extractor is proposed, and this feature extractor is called MK-BOP to generate effective and discriminative features from sound signals.
- A hybrid feature selector is presented to select the most distinctive features. This feature selector uses ReliefF and NCA together. Therefore, this method is called RFNCA, and the success of them is shown in results.
- The achieved best classification rates are calculated as 98% and 94.90% for clean and noisy datasets. These results show that the proposed MK-BOP based multileveled method is robust and highly accurate.

The innovative elements of this research are;

- A new biometric feature for gait is investigated and this biometric feature is named gait sound-based human identification. Two sound datasets were collected and published publicly to show the results of these biometric features.
- A new textural feature generation/extraction function is presented (MK-BOP).
- A hand-crafted features based accurate sound classification model is presented.

Our primary motivation is to show the human identification capability of the gait sounds.

## 2. Materials

The gait sounds of 55 subject (15 women, and 40 men) were recorded. Table 2 shows the status of data for individuals.

The subjects were asked to walk as they walk in normal life and not to exhibit a different walk. Then, each person in the subject group was fitted with a sound recording apparatus developed from the waist. They were told to go on a tour in the corridor and return to the same starting point, and the pitch sounds were recorded. Fig. 1 below shows a subject ready for gait.

The gait distance is defined as 56 m. First, the floor was cleaned to prevent the occurrence of noise when crushing objects such as dust and stones on the floor with shoes. The corridor used for the walkway is shown in Fig. 2 below. The floor has a 60x60 cm hard stone coating. There was no interference with the shoes of

**Table 2**
Attributes and numerical properties.

| Attribute | Min-Max | Mean ± Std |
|---|---|---|
| Age | 18–60 | 20.79 ± 8.08 |
| Height (cm) | 158–191 | 175.38 ± 8.25 |
| Weight (kg) | 47–104 | 71.46 ± 12.20 |

**Fig. 1.** Preparing the Subject for Voice Recording.

**Fig. 2.** Operating Environment.

the subjects. Each subject used the shoes he used in normal life. The shoe of each subject was found to be different in both shape and feature.

Three microphones were used to acquire gait sounds. These are attached to the shoe, ankle, and waist, respectively. It was observed that too much wind noise occurred in the microphone connected to the shoe, and too much rubbing sound of the garment fabric occurred in the microphone connected to the ankle. Therefore, it was finally decided to obtain recordings with a microphone attached to the waist. When each subject completed the registration process, the sound recording file was named as first and last name and recorded on the mobile device. The same process was repeated for each subject. If the preparation process for voice recording is also considered, the average recording time for a subject was 4 min. Ambient silence was maintained during recording. Care was taken not to create any other sound than gait. However, despite all these, some of the audio files of the pending subjects were observed to be mixed up in the speech sounds. In addition, it was seen that gait sounds were difficult to distinguish because some of the subjects were wearing clothes. Considering this situation, the audio files were separated into two separate groups. In the first group, all the sound files were used, including the difficult to distinguish gait sounds. In the second group, recordings containing speech sounds from outside or too much noise were removed. Thus, the first group consisted of 55 people, and the second group consisted of 30 people. The group containing all the audio files was called as noisy, and the group containing the noise was called clean. Noise reduction was also applied to the clean group with noise reduction software. The data of each person is separated into *five* steps. It is ensured that each person has approximately *ten* records. The presence of some people was found when the number of steps did not reach 10. Table 3 below provides information for clean and noisy data. The sampling rate of these sounds is 44.1 KHz.

## 3. Proposed MK-BOP based gait sounds classification method

The main aim of the MK-BOP based classification method is to achieve high success rates by using two collected datasets. The proposed MK-BOP based gait sound classification method consists of channel concatenation, multilevel 1D-DWT, MK-BOP based feature generation, feature selection with RFNCA, and classification. Schematic demonstration of the proposed method is shown in Fig. 3.

Fig. 3 indicate that the proposed MK-BOP based classification method is a hand-extracted multileveled method. The acquired sounds are two channeled. In the first phase, channel concatenation is performed. Then, four leveled 1D-DWT applied to the concatenated signal, and low sub-bands are calculated. MK-BOP generates 3072 features from each level. These features are combined, and 15,360 features are obtained. To select 128 most discriminative features from 15,360 features, RFNCA is used. The selected 128 features are forwarded to classifiers. All operations were done using Matlab 2019b version. Intel i7 (Turbo 3.60 Ghz) processor, 16 GB ram, GTX 1050TI 4 GB display card, and a computer with 240 GB SSD hard disk were used for the study. The phases are detailed explained in the subsection.

### 3.1. Channel concatenation

The collected gait sounds are two channeled. In the first step of the proposed MK-BOP based gait sound classification method, these channels are concatenated.

Step 1: Concatenate channels of the acquired gait sounds.

$$Voice = Ch_1|Ch_2 \tag{1}$$

where *Voice* is concatenated gait sound, $Ch_1$ and $Ch_2$ represent channel *one* and channel *two*, respectively. | defines concatenation operator.

**Table 3**
Clean and Noisy data set features.

| Dataset | Number of People | Number of Records | Download link |
|---|---|---|---|
| Clean | 30 | 300 | https://websiteyonetimi.ahievran.edu.tr/_Dosyalar/Genel/ffea88e5-d2c3-4e9c-8183-d8b6ca39834a-98467bf3-f26a-4365-bc7a-79b92137ebe3.rar |
| Noisy | 55 | 529 | https://websiteyonetimi.ahievran.edu.tr/_Dosyalar/Genel/0c793668-a8cc-4cab-8e3a-6af04c0b7a02-b74a758b-4396-4f21-9df0-58239d526711.rar |

The PSNR rate of the noisy signals was calculated as approximately 20 dB.

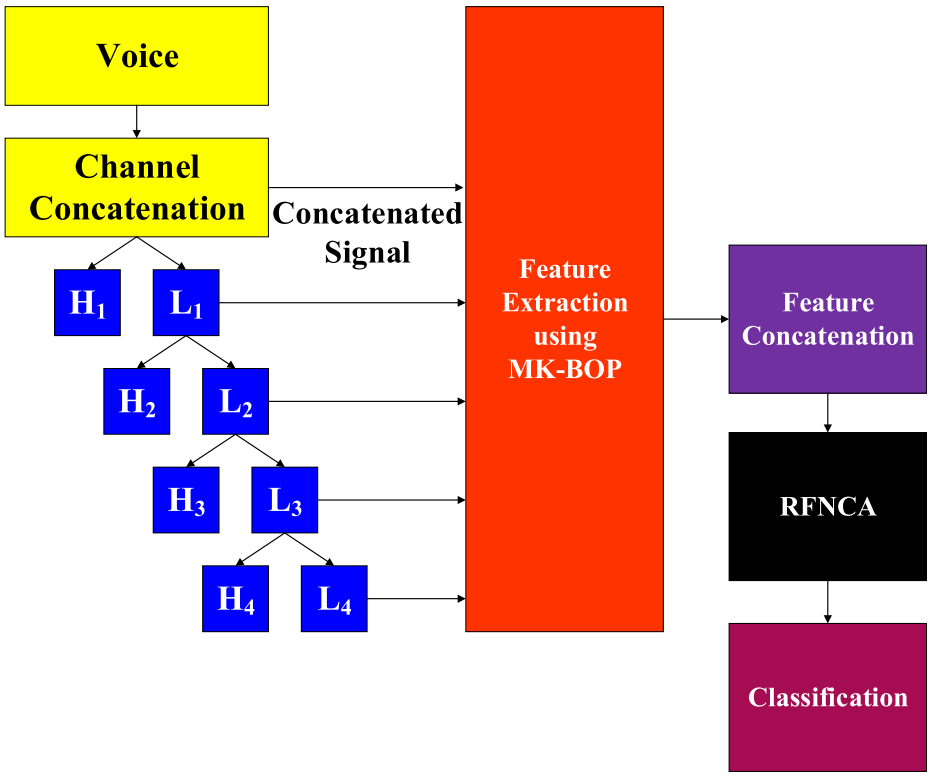


**Fig. 3.** Graphical illustration of the proposed MK-BOP based gait sound classification method.

### 3.2. Multilevel one-dimensional discrete wavelet transform

In this work, we proposed a multileveled feature generation method. 1D-DWT [48] with a sym6 filter is used to construct levels. As we know from the literature, sym6 filtered 1D-DWT are very effective transformation for sounds. The symlets filter have been generally used for signal de-noising and sound decomposition. Here, sym4 and sym6 filters were tested and the best results were calculated by implementing sym6 filter. Therefore, we used 1D-DWT with sym6.

Step 2: Apply 1D-DWT to *Voice* and calculate first, second, third, and fourth low bands. The mathematical equations of these processes are given below.

$$[L_1, H_1] = 1D - DWT(Voice, sym6) \tag{2}$$

$$[L_2, H_2] = 1D - DWT(L_1, sym6) \tag{3}$$

$$[L_3, H_3] = 1D - DWT(L_2, sym6) \tag{4}$$

$$[L_4, H_4] = 1D - DWT(L_3, sym6) \tag{5}$$

where $L_i$ and $H_i$are $i^{th}$ leveled low and high wavelet coefficients respectively.

### 3.3. Feature generation using the proposed MK-BOP

Sections 3.1 and 3.2 defines the preprocessing phase of the proposed MK-BOP based gait sounds recognition method. MK-BOP extracts feature from *Voice*, $L_1$, $L_2$, $L_3$ and $L_4$.

Step 3: Generate features from *Voice*, $L_1$, $L_2$, $L_3$ and $L_4$ using the proposed MK-BOP.

$$feat^1 = MK - BOP(Voice) \tag{6}$$

$$feat^2 = MK - BOP(L_1) \tag{7}$$

$$feat^3 = MK - BOP(L_2) \tag{8}$$

$$feat^4 = MK - BOP(L_3) \tag{9}$$

$$feat^5 = MK - BOP(L_4) \tag{10}$$

where $feat^i$ is $i^{th}$ level features with size of 3072. As seen from Eqs. (6)–(10), the primary extraction function is the proposed MK-BOP. Details of the proposed MK-BOP are explained in Section 3.3.1.

#### 3.3.1. Multi kernelled bijection octal pattern: MK-BOP

In the literature, there are variable one-dimensional local feature generators. Most of them use either a determined pattern or a determined kernel for feature generation. Therefore, they have limited feature generation abilities. A multi-pattern and multi kernel-based pattern is presented to improve feature generation ability, and it is named MK-BOP. The MK-BOP uses *four* variable patterns and two kernels. The used kernels are signum and ternary. The proposed MK-BOP is a textural feature extractor. It uses non-overlapping blocks with a size of 17. Each block divided into three parts, and these parts are called block 1, center, and block 2. Signum and ternary functions [51] are utilized as a binary feature extraction function (kernel). Therefore, this method is called a multi kernelled feature extraction method. Steps of the proposed MK-BOP are given as below.

Step 3.0: Load concatenated gait sound or one of the low pass-band of it.

Step 3.1: Divide gait sound into a 17 size overlapping block as seen from Fig. 4.

As seen in Fig. 4, the used 17 sized overlapping block divided into three parts. These parts are called as block 1, block 2, and center. Blue cells describe block 1, the green cell is the center value, and yellow cells represent values of block 2. Signum and ternary functions are also used for binary feature generation by using these

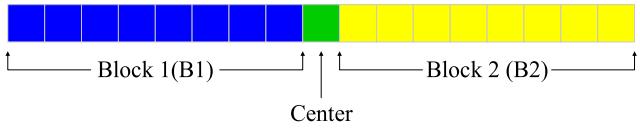


**Fig. 4.** Used 17 sized overlapping block.

parts. Eqs. (11)–(14) mathematically explain signum and ternary functions.

$$sgnm(a,b) = \begin{cases} 0, a-b<0 \\ 1, a-b \geq 0 \end{cases} \tag{11}$$

$$trnr(a,b) = \begin{cases} -1, a-b<-thr \\ 1, a-b>thr \\ 0, -thr \leq a-b \leq thr \end{cases} \tag{12}$$

$$bit_{trnr}^{lower} = \begin{cases} 0, trnr(a,b) \geq 0 \\ 1, trnr(a,b) = -1 \end{cases} \tag{13}$$

$$bit_{trnr}^{upper} = \begin{cases} 0, trnr(a,b) \leq 0 \\ 1, trnr(a,b) = 1 \end{cases} \tag{14}$$

where *sgnm* and *trnr* are signum and ternary functions, respectively. $bit_{trnr}^{upper}$ and $bit_{trnr}^{lower}$ refine upper and lower bits of the ternary functions.*a* and *b* represent the input value of the functions. *thr* is the threshold value. To automated calculate threshold function, a standard deviation based function, and this function is shown in Eq. (15).

$$thr = 0.5\sqrt{\frac{\sum_{k=1}^{L}(Voice_k - (\frac{1}{L}[\sum_{k=1}^{L} Voice_k]))}{L}} \tag{15}$$

In Eq. (15), *L* defines the length of the sound. We used half of the standard deviation of the used gait sound as the threshold value.

As seen Eqs. (11)–(15), MK-BOP uses ternary and signum functions together.

Step 3.2: Generate MK-BOP values. These steps consist of four levels. In the first level, the Center value and block 1 are used to extract 1st level MK-BOP value, and three decimal values are generated.

$$dv_t^1 = \sum_{i=1}^{8}\left(sgnm(block1_i, center) * 2^{8-i}\right), i = \{1,2,\cdots,8\} \tag{16}$$

$$ter_i = trnr(block1_i, center) \tag{17}$$

By using Eq. (17), ternary values of the first level are generated. Eqs. (13) and (14) are applied to generated ternary values, and lower and upper bits of each ternary value are calculated. Second and third decimal values are calculated by using these bits.

$$dv_t^2 = \sum_{i=1}^{8}(bit_{trnr}^{upper}{}_i * 2^{8-i}) \tag{18}$$

$$dv_t^3 = \sum_{i=1}^{8}(bit_{trnr}^{lower}{}_i * 2^{8-i}) \tag{19}$$

where $dv_t^h$ is h$^{th}$ decimal value of the MK-BOP, and *t* defines index number and $t = \{1,2,\cdots,L-16\}$.

A numerical explanation of the first level is shown in Fig. 5. In the examples, the signum function is used.

Fig. 5 schematically explains the first level. In purple cells, extracted binary features are shown. These binary features are converted to the decimal value to obtain decimal value.

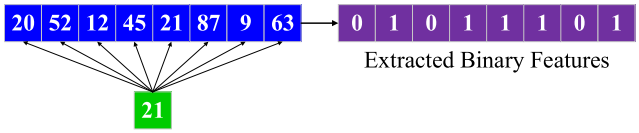


**Fig. 5.** A numerical example of the first level by using the signum function.

In the second level, block2 and center value are used to generate 4th, 5th, and 6th decimal values. The mathematical explanations of this level are given below.

$$dv_t^4 = \sum_{i=1}^{8}\left(sgnm(block2_i, center) * 2^{8-i}\right) \tag{20}$$

$$ter_i = trnr(block2_i, center) \tag{21}$$

$$dv_t^5 = \sum_{i=1}^{8}(bit_{trnr}^{upper}{}_i * 2^{8-i}) \tag{22}$$

$$dv_t^6 = \sum_{i=1}^{8}(bit_{trnr}^{lower}{}_i * 2^{8-i}) \tag{23}$$

Here, block 2 and center value are used together, and Eqs. (20)–(24) clearly indicate that three values are generated at this level. A graphical explanation about this level is shown in Fig. 6.

In Fig. 6, yellow block defines block 1, the green cell is the center value. The extracted bits using the signum function are shown with purple cells.

In the third level, block 1 and block 2 are used to generate MK-BOP decimal values. In this level, features are bipartitely generated. The generation of the decimal values of the third level is mathematically shown in Eqs. (24)–(27).

$$dv_t^7 = \sum_{i=1}^{8}\left(sgnm(block1_i, block2_i) * 2^{8-i}\right) \tag{24}$$

$$ter_i = trnr(block1_i, block2_i) \tag{25}$$

$$dv_t^8 = \sum_{i=1}^{8}(bit_{trnr}^{upper}{}_i * 2^{8-i}) \tag{26}$$

$$dv_t^9 = \sum_{i=1}^{8}(bit_{trnr}^{lower}{}_i * 2^{8-i}) \tag{27}$$

In Eqs. (25)–(28), 7th, 8th, and 9th decimal values generation are given. The numerical example of this stage for 7th decimal value is shown in Fig. 7.

The last level is the fourth level, and in this level block 1 and block 2 are used center symmetric. 10th, 11th, and 12th values are generated.

$$dv_t^{10} = \sum_{i=1}^{8}\left(sgnm(block1_i, block2_{9-i}) * 2^{8-i}\right) \tag{28}$$

$$ter_i = trnr(block1_i, block2_{9-i}) \tag{29}$$

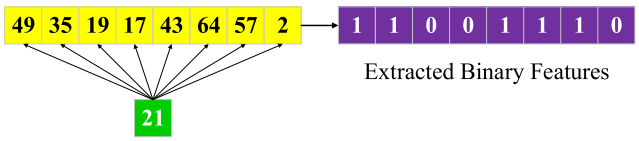


**Fig. 6.** Bit generation example of the second level of the proposed MK-BOP.

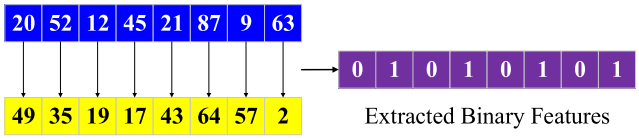


**Fig. 7.** A numerical example of the third level MK-BOP values using signum function.

$$dv_t^{11} = \sum_{i=1}^{8} (bit_{trnr\ i}^{upper} * 2^{8-i}) \tag{30}$$

$$dv_t^{12} = \sum_{i=1}^{8} (bit_{trnr\ i}^{lower} * 2^{8-i}) \tag{31}$$

Graphical and numerical explanation of the 4th level is shown in Fig. 8.

Fig. 8 shows that this level is center symmetric.

Step 3.3: Create 12 signals by using the calculated 12 decimal values.

Step 3.4: Extract histograms of the created 12 signals. Each signal is coded in 8 bits. Therefore, the extracted histogram has $2^8 = 256$ values.

Step 3.5: Concatenate histograms of the calculated signals to obtain a feature set with a size of 3072.

$$feat = H^1 | H^2 | H^3 | H^4 | H^5 | H^6 | H^7 | H^8 | H^9 | H^{10} | H^{11} | H^{12} \tag{32}$$

where *feat* is a feature vector with the size of 3072.

Steps 3.1–3.5 are defined as the MK-BOP feature extraction procedure.

### 3.4. Feature concatenation

In the MK-BOP based feature extraction phase, 3072 features are extracted in five levels. In this phase, 15,360 sized feature vector is created by concatenating the extracted features, and these features are normalized.

Step 4: Concatenate features to obtain the final feature vector.

$$ff = feat^1 | feat^2 | feat^3 | feat^4 | feat^5 \tag{33}$$

where *ff* is the final feature with a size of 15360.

Step 5: Normalize the final feature by using Eq. (35).

$$ff = \frac{ff - ff_{min}}{ff_{max} - ff_{min}} \tag{34}$$

$ff_{min}$ and$ff_{max}$ are minimum and maximum values of the final features, respectively.

### 3.5. Feature selection

Here, we employed both ReliefF and NCA feature selectors. These feature selectors are weight-based selectors. Both of them used a distance-based fitness function (L1-norm) to generate weights, and they can select the most informative features by using one-dimensional feature generators. Therefore, both of them are utilized as a feature selector (RFNCA) to use the advantages of both of them. The main objective of this phase is to select 128 most discriminative features from 15,360 normalized features. According to tests, the best results are obtained from 128 features. Therefore, we select 128 most discriminative features. In order to select these features, ReliefF [49] and NCA [50] are used together. These two feature selectors generate weights of the features. The main difference between them is defined as follows. NCA extracts non-negative weights, but ReliefF generates both positive and negative weights. According to ReliefF, negative weights define redundant features. By using them, the RFNCA feature selector is used in this phase, and steps are given below.

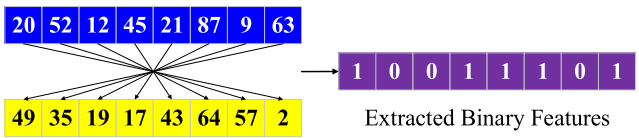


**Fig. 8.** Fourth level bit extraction.

Step 6: Extract ReliefF weights to final features.

$$w^{RF} = ReliefF(ff, target) \tag{35}$$

where $w^{RF}$ describes ReliefF weights with a size of 15360.

Step 7: Eliminate redundant features using Eq. (37).

$$fr(cnt) = \begin{cases} ff(i) and cnt = cnt + 1, w^{RF}(i) > 0 \\ continue, w^{RF}(i) \leq 0 \end{cases} \tag{36}$$

As seen from Eq. (36), positive weighted features are selected in the ReliefF phase. *fr* is selected features using ReliefF, and *cnt* is the length of the *fr*.

Step 8: Generate features of NCA using *fr* and *target*. Mathematical notation of this step is given in Eq. (37).

$$w^{NCA} = ReliefF(fr, target) \tag{37}$$

where $w^{NCA}$ is weights of the NCA with a size of *cnt*.

Step 9: Select 128 most distinctive features of the *fr*.

$$[sortedweights, endex] = sort(w^{NCA}, descending) \tag{38}$$

$$fs(i) = fr(endex(i)); i = \{1, 2, \cdots, 128\} \tag{39}$$

In this step, generated weights are sorted by descending, and indices of the sorted weights are calculated in the Eq. (39). 128 most distinctive features are selected, and these 128 features are called as *fs*. The selection process is shown in Eq. (39).

### 3.6. Classification

RFNCA selects 128 most discriminative features. These features are utilized as input in the classification phase. To comprehensively evaluate the performance of the MK-BOP based feature extraction method, five conventional classifiers are selected. These classifiers are Decision Tree (DT), Linear Discriminant (LD), Support Vector Machine (SVM), K Nearest Neighbor (KNN), and Subspace Discriminant (SD). Attributes of these five classifiers are given in Table 4. 10-folds cross-validation is used to obtain test results.

**Table 4**
Attributes of the used five classifiers.

| Name | Explanation |
|---|---|
| DT [51] | The decision tree is one of the most known and simple classifiers in the literature. DT uses entropy to calculate predicted values. In this study, Gini's diversity index is chosen for splitting, and 100 is selected as the maximum number of the split. |
| LD [52] | LD is a nonparametric and conventional classifier. Here, full is selected as a covariance structure. |
| SVM [51] | SVM has variable kernels and is an optimization-based classifier. Kernel function of the used SVM selected as Gaussian and 11 is selected as a kernel scale. The one-vs-one classification method is used. |
| KNN [53] | KNN is a distance-based classifier, and it uses variable distance metrics and K values. K is selected as 1, and the City Block distance metric is utilized in this work. |
| SD [54] | SD is an ensemble classifier. 64 and 30 are selected as subspace dimensions and learner numbers, respectively. |

MATLAB classification learner toolbox was used to calculate the results of the presented MK-BOP and RFNCA based identification model. The best accurate five classifiers are selected with their default parameters.

## 4. Results and discussions

In this section, the performance (recognition rate) of the proposed MK-BOP based gait recognition method is evaluated by using accuracy, precision, recall, geometric mean, and F1 score [55]. We gathered two datasets, and these datasets are called as noisy and clear. Therefore, we define tow cases by using these datasets. Five classifiers that are described in Section 3.6 are used to obtain a comprehensive benchmark. The defined cases are explained below.

Case 1: The proposed MK-BOP based gait recognition method is tested on a clear dataset in this case.

Case 2: The proposed MK-BOP based gait recognition method is tested on a noisy dataset in this case.

To calculate numerical results from the used five conventional classifiers, accuracy, precision, recall, F1 score, and geometric mean are used. Mathematical explanations of these performance metrics are given in Eqs. (41)–(45) [55].

$$accuracy = \frac{nTP + nTN}{nTP + nFP + nTN + nFN} \tag{40}$$

$$precision = \frac{nTP}{nTP + nFP} \tag{41}$$

$$recall = \frac{nTP}{nTP + nFN} \tag{42}$$

$$F1 = 2\frac{precision * recall}{precision + recall} \tag{43}$$

$$geometricmean = \sqrt{recall * \frac{nTN}{nTN + nFP}} \tag{44}$$

where *nTP* is the number of true positives, *nTN* defines the number of true negatives, *nFP* represents the number of false-positive and *nFN* is the number of false negatives.

The used tests of the used five classifiers were repeated 1000 times to get comprehensive results, and average, standard deviation, and maximum value of the used performance metrics according to the used case are given in Tables 5 and 6.

**Table 5**
Results of Case 1.

| Classifier | Performance metric | Statistical moment | | |
|---|---|---|---|---|
| | | Mean | Std | Max |
| DT | Accuracy | 70.93 | 1.86 | 76.33 |
| | F1-score | 71.46 | 1.87 | 76.84 |
| | Recall | 70.93 | 1.86 | 76.33 |
| | Precision | 72.01 | 1.92 | 76.33 |
| | G-mean | 68.93 | 3.01 | 75.14 |
| LD | Accuracy | 93.12 | 0.80 | 95.33 |
| | F1-score | 93.50 | 0.78 | 95.54 |
| | Recall | 93.12 | 0.80 | 95.33 |
| | Precision | 93.89 | 0.77 | 95.80 |
| | G-mean | 92.59 | 0.90 | 95.08 |
| SVM | Accuracy | 93.77 | 0.32 | 94.67 |
| | F1-score | 94.22 | 0.28 | 94.99 |
| | Recall | 93.77 | 0.32 | 94.67 |
| | Precision | 94.68 | 0.24 | 95.32 |
| | G-mean | 93.56 | 0.34 | 94.50 |
| KNN | Accuracy | 95.57 | 0.26 | 96.67 |
| | F1-score | 95.87 | 0.24 | 96.90 |
| | Recall | 95.57 | 0.26 | 96.67 |
| | Precision | 96.18 | 0.23 | 97.13 |
| | G-mean | 95.28 | 0.28 | 96.51 |
| SD | Accuracy | 95.88 | 0.76 | 98.00 |
| | F1-score | 96.15 | 0.68 | 98.09 |
| | Recall | 95.88 | 0.76 | 98.00 |
| | Precision | 96.43 | 0.61 | 98.18 |
| | G-mean | 95.62 | 0.81 | 97.87 |

**Table 6**
Results of the Case 2.

| Classifier | Performance metric | Statistical moment | | |
|---|---|---|---|---|
| | | Mean | Std | Max |
| DT | Accuracy | 57.07 | 1.59 | 62.00 |
| | F1-score | 57.58 | 1.65 | 63.37 |
| | Recall | 56.84 | 1.64 | 62.19 |
| | Precision | 58.34 | 1.74 | 64.59 |
| | G-mean | 40.29 | 1.84 | 58.56 |
| LD | Accuracy | 91.82 | 0.47 | 93.38 |
| | F1-score | 91.80 | 0.48 | 93.25 |
| | Recall | 91.15 | 0.50 | 92.73 |
| | Precision | 92.45 | 0.50 | 92.73 |
| | G-mean | 90.05 | 0.69 | 92.23 |
| SVM | Accuracy | 93.22 | 0.59 | 94.90 |
| | F1-score | 93.48 | 0.58 | 95.01 |
| | Recall | 93.09 | 0.65 | 94.82 |
| | Precision | 93.87 | 0.53 | 95.22 |
| | G-mean | 92.45 | 0.73 | 94.39 |
| KNN | Accuracy | 90.75 | 0.19 | 91.49 |
| | F1-score | 91.26 | 0.20 | 91.95 |
| | Recall | 90.70 | 0.18 | 91.41 |
| | Precision | 91.83 | 0.25 | 92.58 |
| | G-mean | 88.99 | 0.26 | 90.01 |
| SD | Accuracy | 93.74 | 0.37 | 94.90 |
| | F1-score | 93.92 | 0.38 | 95.13 |
| | Recall | 93.61 | 0.45 | 95.00 |
| | Precision | 94.24 | 0.33 | 95.27 |
| | G-mean | 93.05 | 0.50 | 94.55 |

As seen from the results above, the best classifier is SD for both datasets. The best classification accuracies are calculated as 94.90% and 98.0% for Case 1 and Case 2 by using the SD classifier. These results clearly showed that the proposed MK-BOP based gait recognition method is a robust recognition method. KNN and LDA achieved higher 90% classification accuracies for both datasets. The worst classifier is DT because it resulted in the worst in both clean and noisy datasets. The advantages of the proposed MK-BOP based gait recognition method are given below.

- The proposed method is a robust method because it achieved high classification rates for both noisy and clear datasets.
- Here, a novel textural feature extractor is presented, and the success of the proposed MK-BOP is shown. (See Tables 5 and 6)
- The presented MK-BOP based method is a hand-crafted classification method, and there is no need to set millions parameters as deep learning.
- The application of the proposed MK-BOP based gait recognition method is simple.

The main hypothesis of the MK-BOP based gait recognition method is to show biometric identification can be performed using gait sound. Limitations of our method are to implement biometric authentication with gait sound, extra hardware should be used, and it is depended on the used shoe.

## 5. Conclusions

In this work, a novel MK-BOP based gait sound classification method is presented. The objective of the proposed MK-BOP based is to show biometric identification can be performed by using gait sounds. Therefore, a multileveled signal processing method is presented to recognize subjects using gait sounds. We collected two datasets. These datasets are called as a noisy and clean dataset. The robustness of the proposed MK-BOP was tested on these datasets. *Five* conventional classifiers were used to obtain a comprehensive benchmark, and the results of them were listed in Tables

4 and 5. The best-resulted classifier is SD according to these Tables. SD achieved 94.9% and 98.0% classification accuracies for noisy and clear datasets, respectively. This work is crucial for crime investigators. Among the advantages of the study is its high success rate and its ability to recognize in both noisy environments and environments with clear sound. The disadvantages of the study include the effect of the dress, walking environment, and walking speed on the success rate. In the future, novel datasets can be created for gait sound recognition, and novel intelligence digital forensics applications can be developed by using this method or novel deep networks.

## CRediT authorship contribution statement

**Emrah Aydemir:** Conceptualization, Methodology, Formal analysis, Investigation, Writing - original draft, Supervision, Project administration, Funding acquisition. **Turker Tuncer:** Conceptualization, Methodology, Software, Validation, Resources, Writing - review & editing, Funding acquisition. **Sengul Dogan:** Software, Validation, Writing - original draft, Writing - review & editing, Visualization, Funding acquisition. **Musa Unsal:** Data curation, Funding acquisition.

## Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References


[1] Wang L, Tan T, Ning H, Hu W. Silhouette analysis-based gait recognition for human identification. IEEE Trans Pattern Anal Mach Intell 2003;25:1505–18.
[2] Tao D, Li X, Wu X, Maybank SJ. General tensor discriminant analysis and gabor features for gait recognition. IEEE Trans Pattern Anal Mach Intell 2007;29:1700–15.
[3] A.W. Mitas, M.D. Bugdol, An idea of human voice reaction measurement system under the aspect of behavioral biometric, Information Technologies in Biomedicine, Springer2010, pp. 219-228.
[4] Frank M, Biedert R, Ma E, Martinovic I, Song D. Touchalytics: on the applicability of touchscreen input as a behavioral biometric for continuous authentication. IEEE Trans Inf Forensics Secur 2012;8:136–48.
[5] Sitová Z, Šeděnka J, Yang Q, Peng G, Zhou G, Gasti P, et al. HMOG: New behavioral biometric features for continuous authentication of smartphone users. IEEE Trans Inf Forensics Secur 2015;11:877–92.
[6] Pławiak P, Abdar M, Pławiak J, Makarenkov V, Acharya UR. DGHNL: a new deep genetic hierarchical network of learners for prediction of credit scoring. Inf Sci 2020;516:401–18.
[7] Pławiak P, Abdar M, Acharya UR. Application of new deep genetic cascade ensemble of SVM classifiers to predict the Australian credit scoring. Appl Soft Comput 2019;84:105740.
[8] Rzecki K, Sośnicki T, Baran M, Niedźwiecki M, Król M, Łojewski T, et al. Application of computational intelligence methods for the automated identification of paper-ink samples based on LIBS. Sensors 2018;18:3670.
[9] Abdar M, Książek W, Acharya UR, Tan R-S, Makarenkov V, Pławiak P. A new machine learning technique for an accurate diagnosis of coronary artery disease. Comput Methods Programs Biomed 2019;179:104992.
[10] Pławiak P, Acharya UR. Novel deep genetic ensemble of classifiers for arrhythmia detection using ECG signals. Neural Comput Appl 2020;32:11137–61.
[11] Abdar M, Wijayaningrum VN, Hussain S, Alizadehsani R, Plawiak P, Acharya UR, et al. IAPSO-AIRS: a novel improved machine learning-based system for wart disease treatment. J Med Syst 2019;43:220.
[12] Hayfron-Acquah JB, Nixon MS, Carter JN. Automatic gait recognition by symmetry analysis. Pattern Recogn Lett 2003;24:2175–83.
[13] Boulgouris NV, Hatzinakos D, Plataniotis KN. Gait recognition: a challenging signal processing technology for biometric identification. IEEE Signal Process Mag 2005;22:78–90.
[14] Iwama H, Muramatsu D, Makihara Y, Yagi Y. Gait verification system for criminal investigation. Inform Media Technol 2013;8:1187–99.
[15] D. Muramatsu, Y. Makihara, H. Iwama, T. Tanoue, Y. Yagi, Gait verification system for supporting criminal investigation, 2013 2nd IAPR Asian Conference on Pattern Recognition, IEEE, 2013, pp. 747-748.
[16] Gafurov D. A survey of biometric gait recognition: Approaches, security and challenges, Annual Norwegian computer science conference. Annual Norwegian Computer Science Conference Norway 2007:19–21.
[17] Liu Z, Sarkar S. Effect of silhouette quality on hard problems in gait recognition. IEEE Trans Systems Man Cybern Part B (Cybernetics) 2005;35:170–83.
[18] Jain AK, Flynn P, Ross AA. Handbook of biometrics. Springer Science & Business Media; 2007.
[19] Yu S, Tan D, Tan T. A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition. In: 18th International Conference on Pattern Recognition (ICPR'06). p. 441–4.
[20] Deng M, Fan T, Cao J, Fung S-Y, Zhang J. Human gait recognition based on deterministic learning and knowledge fusion through multiple walking views. J Franklin Inst 2019.
[21] L. Lee, W.E.L. Grimson, Gait analysis for recognition and classification, Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition, IEEE, 2002, pp. 155-162.
[22] Yu S, Tan T, Huang K, Jia K, Wu X. A study on gait-based gender classification. IEEE Trans Image Process 2009;18:1905–10.
[23] Cavanagh P, Derr J, Ulbrecht J, Maser R, Orchard T. Problems with gait and posture in neuropathic patients with insulin-dependent diabetes mellitus. Diabet Med 1992;9:469–74.
[24] Liao R, Yu S, An W, Huang Y. A model-based gait recognition method with body pose and human prior knowledge. Pattern Recogn 2020;98:107069.
[25] O.S.B. Database, http://www.am.sanken.osaka-u.ac.jp/BiometricDB/index.html.
[26] Yao L, Kusakunniran W, Wu Q, Zhang J, Tang Z, Yang W. Robust gait recognition using hybrid descriptors based on Skeleton Gait Energy Image. Pattern Recogn Lett 2019.
[27] Tong S-B, Fu Y-Z, Ling H-F. Cross-view gait recognition based on a restrictive triplet network. Pattern Recogn Lett 2019;125:212–9.
[28] Iwama H, Okumura M, Makihara Y, Yagi Y. The ou-isir gait database comprising the large population dataset and performance evaluation of gait recognition. IEEE Trans Inf Forensics Secur 2012;7:1511–21.
[29] Verlekar TT, Soares LD, Correia PL. Gait recognition in the wild using shadow silhouettes. Image Vis Comput 2018;76:1–13.
[30] Iwashita Y, Kurazume R, Stoica A. Gait identification using invisible shadows: robustness to appearance changes. In: 2014 Fifth International Conference on Emerging Security Technologies. p. 34–9.
[31] Wu H, Weng J, Chen X, Lu W. Feedback weight convolutional neural network for gait recognition. J Vis Commun Image Represent 2018;55:424–32.
[32] Alotaibi M, Mahmood A. Improved gait recognition based on specialized deep convolutional neural network. Comput Vis Image Underst 2017;164:103–10.
[33] Sun C, Wang C, Lai W. Gait analysis and recognition prediction of the human skeleton based on migration learning. Physica A 2019;532:121812.
[34] Kuehne H, Jhuang H, Garrote E, Poggio T, Serre T. HMDB: a large video database for human motion recognition. 2011 International Conference on Computer Vision, IEEE 2011:2556–63.
[35] M.D. Rodriguez, J. Ahmed, M. Shah, Action mach a spatio-temporal maximum average correlation height filter for action recognition, 2008 IEEE conference on computer vision and pattern recognition, IEEE, 2008, pp. 1-8.
[36] Rubio-Solis A, Panoutsos G, Beltran-Perez C, Martinez-Hernandez U. A multilayer interval type-2 fuzzy extreme learning machine for the recognition of walking activities and gait events using wearable sensors. Neurocomputing 2020.
[37] Sun F, Zang W, Gravina R, Fortino G, Li Y. Gait-based identification for elderly users in wearable healthcare systems. Inform Fusion 2020;53:134–44.
[38] Yu S, Chen H, Wang Q, Shen L, Huang Y. Invariant feature extraction for gait recognition using only one uniform model. Neurocomputing 2017;239:81–93.
[39] S. Yu, Q. Wang, Y. Huang, A large RGB-D gait dataset and the baseline algorithm, Chinese Conference on Biometric Recognition, Springer, 2013, pp. 417-424.
[40] Lempereur M, Rousseau F, Rémy-Néris O, Pons C, Houx L, Quellec G, et al. A new deep learning-based method for the detection of gait events in children with gait disorders: Proof-of-concept and concurrent validity. J Biomech 2020;98:109490.
[41] Pepa L, Capecci M, Andrenelli E, Ciabattoni L, Spalazzi L, Ceravolo MG. A fuzzy logic system for the home assessment of freezing of gait in subjects with Parkinsons disease. Expert Syst Appl 2020;113197.
[42] Battistone F, Petrosino A. TGLSTM: a time based graph deep learning approach to gait recognition. Pattern Recogn Lett 2019;126:132–8.
[43] 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.
[44] W. Li, Z. Zhang, Z. Liu, Action recognition based on a bag of 3d points, 2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition-Workshops, IEEE, 2010, pp. 9-14.
[45] J. Sung, C. Ponce, B. Selman, A. Saxena, Human activity detection from RGBD images, Workshops at the twenty-fifth AAAI conference on artificial intelligence, 2011.
[46] Deng M, Wang C, Chen Q. Human gait recognition based on deterministic learning through multiple views fusion. Pattern Recogn Lett 2016;78:56–63.
[47] R. Gross, J. Shi, The cmu motion of body (mobo) database, (2001).
[48] Vidya V, Farheen N, Manikantan K, Ramachandran S. Face recognition using threshold based DWT feature extraction and selective illumination enhancement technique. Proc Technol 2012;6:334–43.
[49] Robnik-Šikonja M, Kononenko I. Theoretical and empirical analysis of ReliefF and RReliefF. Mach Learn 2003;53:23–69.
[50] Goldberger J, Hinton GE, Roweis ST, Salakhutdinov RR. Neighbourhood components analysis. Adv Neural Inform Process Syst 2005:513–20.

[51] Tuncer T, Dogan S. Novel dynamic center based binary and ternary pattern network using M4 pooling for real world voice recognition. Appl Acoust 2019;156:176–85.
[52] Ye J, Janardan R, Li Q. Two-dimensional linear discriminant analysis. Adv Neural Inform Process Syst 2005:1569–76.
[53] M.-L. Zhang, Z.-H. Zhou, A k-nearest neighbor based algorithm for multi-label classification, 2005 IEEE international conference on granular computing, IEEE, 2005, pp. 718-721.
[54] Y. Zhu, S. Schwartz, M. Orchard, Fast face detection using subspace discriminant wavelet features, Proceedings IEEE Conference on Computer Vision and Pattern Recognition. CVPR 2000 (Cat. No. PR00662), IEEE, 2000, pp. 636-642.
[55] Tuncer T, Dogan S, Pławiak P, Acharya UR. Automated arrhythmia detection using novel hexadecimal local pattern and multilevel wavelet transform with ECG signals. Knowl-Based Syst 2019;186:104923.