Applied Computer Systems
ISSN 2255-8691 (online)
ISSN 2255-8683 (print)
December 2021, vol. 26, no. 2, pp. 164–172
https://doi.org/10.2478/acss-2021-0020
https://content.sciendo.com

Real-Time Identification from Gait Features Using Cascade Voting Method

Berk Ercin[1], Abdulkadir Karacı[2*]
[1,2] *Kastamonu University, Kastamonu, Turkey*

***Abstract* – There are several biometric methods for identification. These are generally classified under two main groups as physiological and behavioural biometric methods. Recently, methods using behavioural biometric features have gained popularity. Identification made using gait pattern is also one of these methods. The present study proposes a machine learning based system performing identification in real time via gait features using a Kinect device. The data set is composed of 23 individuals' skeleton model data obtained by the authors. From these data, 147 handcrafted features have been extracted. Deep Neural Network (DNN), Random Forest (RF), Gradient Boosting (GB), XG-Boost (XGB) and *K*-Nearest Neighbour (KNN) classifiers have been trained with these features. Furthermore, the output of these five machine learning models has been combined with a voting approach. The highest classification has been obtained with 97.5 % accuracy via a voting approach. The classification accuracies of the RF, DNN, XGB, GB and KNN classifiers are 95 %, 87.5 %, 85 %, 80 % and 65 %, respectively. The classification accuracy obtained via a voting approach is higher than in the previous studies. The developed system successfully performs real-time identification.**

***Keywords* – Deep neural network, identification, Kinect, machine learning, voting approach.**

I. Introduction

Development of accurate and efficient identity authentication methods via biometrics and computer-assisted vision systems is an important research area. Biometrics is the measurement and analysis of the distinguishable biological properties [1]. While biometrics has been used in various commercial applications, from access control against potential frauds to preventing voters from voting twice, it has also attracted a great deal of attention from researchers [2]. Biometric recognition, or simply biometrics, means the recognition of individuals automatically according to their physiological or behavioural features. While features such as iris image, face and fingerprint are used in physiological biometrics, behavioural characteristics of individuals such as voice, signature and gait information are used in behavioural biometrics [3]. The importance of the automatic identification based on biometrics increases with each passing day in crowded environments frequently used by societies. Although the use of physiological biometrics is quite common in such environments, it already requires authentication of the person before they can come into physical contact with the sensor to initiate the identification process. These sensors, which are in common use of individuals in social settings, may cause various contagious diseases to spread [4]. Therefore, the importance of automatic remote identification or authentication using gait information increases in such social settings.

The identification of people based on their gait can be performed using low resolution images unlike other biometrics such as iris and fingerprint and requires no physical interaction from the subjects [5]. Identification based on gait style of people aims at practicability of identification process. The gait style of individuals is a unique feature that is difficult for others to imitate [6]. Gait information used in identification is based on recording the gait of individuals with the help of a camera. The extraction of the gait information after this recording is generally classified under two categories as model-based and model-free.

In the model-free method, the gait information of the individuals is based on the features extractable from the silhouettes obtained from the image [7]. For the gait information of individuals to be extracted, silhouettes are obtained by separating the background and the individuals. Model-free approaches can be applied easily but they are not fixed-scaled and require recording from various angles [8]. Model-based approaches, on the other hand, aim at identifying by matching a skeleton model to the individual whose image is taken [9]. In model-based approaches, the information obtained from the gait features is generally called static and dynamic features. Static features consist of the person-specific measurements such as bone lengths, height of the individual, dimensions of the body parts and the locations of the joints. Dynamic features, on the other hand, consist of features such as the walking pace that can change during the walk, length of the step, the orbits the body parts follow during the walk and the angles of the joints relative to one another. Model-based approaches have slightly higher computational costs; however, new generation devices make it possible to overcome this limitation.

One of these devices, Microsoft's Kinect, has become a device often preferred by the researchers as it is highly affordable due to its low cost and can easily extract the characteristic gait information of individuals thanks to its depth

* Corresponding author's e-mail: akaraci@kastamonu.edu.tr

©2021 Berk Ercin, Abdulkadir Karacı.
This is an open access article licensed under the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0), in the manner agreed with Sciendo.

sensors. It provides the researchers with the opportunity to collect detailed three-dimensional information using the Time-of-Flight technology [10]. Various data of 25 joints of individuals whose images are recorded with Kinect SDK can easily be obtained. Kinect provides the user with the three-dimensional real world coordinate information of these joints. The gait features of the individuals are extracted by analysing their movements with this coordinate information. Identification applications can be developed by using the obtained gait features.

Some researchers have developed feature extraction techniques that allow extracting both static and dynamic features from the gait information. Bobick and Johnson [11] have measured static features such as height, body width and leg length, and the distance between the feet at moments when both feet contact the ground during the walk for each walking period. Singh and Jain [12] have obtained dynamic features by creating two triangles between three points to represent the human body. Jianwattanapaisarn et al. [13] have achieved a high level of identification performance by using some joint coordinate points as a gait feature. Also, Sahak et al. [5] have stated that while low identification performance is observed in studies using only dynamic features, in the studies using the combination of static and dynamic features high identification performance is observed.

In this study, real-time identification is performed by using all joint coordinate points of human skeleton model and refracting angles of some joint points relative to joint points, to which they are associated provided by Kinect device. The data set has been obtained from 23 volunteers, 13 men and 10 women, within the scope of the study. For classification, Deep Neural Network (DNN), Random Forest (RF), Gradient Boosting (GB), XG-Boost (XGB) and *K*-Nearest Neighbour (KNN) classifiers have been used. Furthermore, these methods have been combined via a voting approach and the classification performance of all classifiers has been compared.

The main contribution of the study can be summarised as follows:

1. Original data set obtained by the authors has been used.
2. The classification performance has been enhanced by combining five different machine learning models via a cascade voting method.
3. Higher classification performance compared to other real-time identification studies in the literature has been obtained.

The paper is organised as follows. In the second section of the study, theoretical background of the biometrics is presented. The third section describes methods used, feature extraction and creation of the data set. In the fourth section, real-time classification performance of the classifiers is presented.

II. Theoretical Background

A. Biometrics

Biometrics is a discipline that applies statistical methods to biological problems. It includes the use of physical or biological properties to identify individuals. Biometric identification is known as a discipline interested in the identification and measurement of the physiological and behavioural features. By means of biometrics, the features to be used in identification are measured and person-specific characteristic features are collected. Then, these features are analysed and biometric systems that perform identification are designed. One of the problems encountered in the design of such systems is that there can be noise in these samples taken from individuals. This noise results from the measurements of two samples taken from the same individual at different times being different [14]. Biometric identification processes are generally separated into two processes, such as recording and identification. In the recording process, distinguishing biometric data are taken from a group of individuals, these data are analysed, and features are extracted for each individual. Using these extracted features, a template is created for each individual. In the identification process, the obtained template is compared with the templates obtained during the recording process using certain algorithms and the identification is achieved. In scenarios where more than one algorithm and template are used, a single output is tried to be produced by a voting method [1]. Biometric features are divided into two groups: physiological and behavioural biometrics.

B. Physiological Biometrics

Physiological biometrics are obtained from the physical measurements of the human body. The accuracy rates of the systems based on physiological characteristics in identification are high [15]. However, in some circumstances, these biometric features may change. For instance, fingerprints of individuals from professions in contact with chemical substances may vary from day to day. Similarly, the eyes of individuals with diabetes may also vary [16]. Examples of application regarding physiological biometrics include identification with iris, identification with face and identification with fingerprint.

In identification with iris, the iris pattern which is a distinguishing feature for individuals is used. As in fingerprints, the irises of the twins are also different. The iris image taken for identification is subjected to data analysis for the distinguishing features to be extracted. This analysis is generally made by scanners specifically designed for iris scanning [17]. Even though the use of these scanners is simple and quick, they are expensive and pose threats to human health as different individuals at close range use them.

The face recognition systems ensure the easy, rapid and comfortable acquisition of the data required without the need for cooperation of the individuals or interfering with their social lives. They also demonstrate high classification performance. For these reasons, these are one of the most used biometric identifications in our daily lives. One major problem in automatic face recognition is that they need to perform well despite the changes a face experiences through time, different poses, illumination, cosmetics such as make-up products and glasses individuals use and various facial expressions. [18].

Fingerprint identification is one of the most common and reliable biometrics due to its inter-person uniqueness, consistency and authenticity [19]. Each finger of each person

has a different pattern, and even the fingerprints of twins are different [20]. The patterns of ridges and valleys on the fingertips are distinguishing features for individuals, and the locations and directions of these patterns on the finger are important. Collection of fingerprints is done by sensors designed for this purpose [21]. These sensors are generally preferred as they are cheap and decrease the application costs [22].

C. Behavioural Biometrics

Behavioural biometrics are the behavioural characteristic features individuals respond to the situation they encounter. While a part of the individual's body is used for identification or authentication in physiological biometrics, in behavioural biometrics, unique responses individuals present are used for identification. In addition, while physiological biometrics methods identify the individual once and at only one point, behavioural biometrics methods perform identification or authentication continuously as long as the individual is in proximity to the related sensor. Behavioural biometrics methods include identification via gait, signature and voice.

The way individuals sign is their characteristic feature. In identification via signature, the digital form of the signatures of individuals is used. Typically, there are two methods: static method and dynamic method [23]. In static authentication, similarities between a signature, the image of which is recorded with the help of a scanning device, and signatures of the individual stored in the database are detected using various pattern detection methods [24]. In the dynamic identification, on the other hand, the signatures of individuals are recorded in real time by means of electronic tablets or computer screens sensitive to pens; features of the signature such as speed, acceleration, pressure and the angle of the pen in relation to the surface are extracted. The authentication is made in accordance with the analysis of these features [25]. The voice biometrics performs authentication by using the characteristic features of the voice of an individual. The features of the voice consist of physiological and behavioural biometric features [1]. Identification by voice is used to identify or authenticate the identity of an individual by decoding the speaking structure of the said individual.

Another behavioural biometrics used in identification is identification via gait. The gait pattern is person-specific and is a behavioural biometric consisting of complex data. It is quite difficult for one's gait to be imitated by others [6]. People constantly walk in a gait pattern unique to themselves. Identification or authentication based on gait information is a pattern identification in its essence. The information of gait patterns of individuals is obtained using a variety of methods and devices, and various machine learning models are trained by the obtained information, and systems that can perform real-time identification can be developed. In addition, gait information is among one of the few biometric features that can be obtained from individuals without interfering with their daily lives and from a certain distance with the help of a camera. Therefore, it is quite appropriate for the surveillance and security scenarios in shared areas often used by societies without harming the privacy of individuals while identifying an individual within the system. The combination of biometric-based static features and the features obtained from motion analysis of some joints may create an effective data set to identify an individual [8]. Many systems providing identification based on a gait pattern extract the silhouette of the individual from the image and perform the identification either by obtaining gait information from this silhouette or by the gait information obtained from the various joint points provided by the emerging model as a result of the matching of these silhouettes to a human skeleton model.

III. METHOD

The section of the study presents a model-based approach used for real-time identification and machine learning methods used for classification by obtaining the data set via Microsoft Kinect device.

The success of the identification process depends on getting the gait information in a healthy way. For this reason, the methods to obtain gait information should be chosen carefully. As mentioned before, the methods to obtain the gait information generally consist of model-based and model-free methods. Model-free methods extract the gait information from silhouettes obtained from image [26]. Static features are more important for these methods [27]. The application of these methods is easier, and their computational complexities are smaller compared to model-based methods. However, model-free methods depend on the angle and scale of the image taken and require recording with multiple cameras [8]. In addition, clothing of individuals also has an important effect on the gait information obtained. In the phase of feature extraction, these limitations can be overcome by using model-based methods [28]. Model-based methods use matching of a human skeleton model to the individual whose silhouette is extracted [11]. Gait information is obtained by using the skeleton model that emerged as a result of the matching. Furthermore, model-based methods are independent of the angle and scale of the image taken [29]. They are not sensitive to the background and the noise in the image.

Model-based approaches have high computational costs due to complex algorithms such as extracting the background, detecting the silhouette and matching this detected silhouette to human skeleton model [30]. These disadvantages of the model-based approach can be tackled using new generation image recording devices emerging as the camera systems develop. Kinect sensor is one of these devices. Kinect sensor provides high-quality human skeleton models containing three-dimensional real-world coordinate information of 25 joint points belonging to at most 6 individuals and orientation data of the joint points relative to the associated joint points at the same time. However, because some of these joint points are leaf nodes, Kinect does not provide any orientation data for these joint points. These joint points are head, left foot, right foot, left hand tip, right hand tip, left thumb and right thumb. Fig. 1 shows a human skeleton model and joint points provided by Kinect device. In Table I, the labels of these joint points are presented [31]–[33].

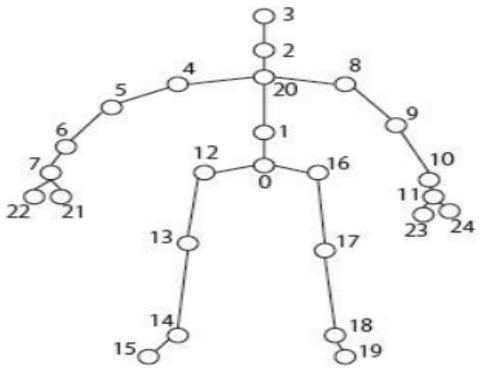


Fig. 1. Kinect human skeleton model and joint points.

TABLE I
LABELS CORRESPONDING TO INDEX INFORMATION OF THE JOINT POINTS OF THE KINECT SKELETON MODEL

Joint Points	Index Numbers	Joint Points	Index Numbers
Lower Spine	0	Left Knee	13
Middle Spine	1	Left Ankle	14
Neck	2	Left Foot	15
Head	3	Right Hip	16
Left Shoulder	4	Right Knee	17
Left Elbow	5	Right Ankle	18
Left Wrist	6	Right Foot	19
Left Hand	7	Shoulder Spine	20
Right Shoulder	8	Left Hand Tip	21
Right Elbow	9	Left Thumb	22
Right Wrist	10	Right Hand Tip	23
Right Hand	11	Right Thumb	24
Left Hip	12		

A. Obtaining the Data Set

There is no readily available data set acknowledged by everyone for identification with Kinect device. Therefore, in this study, an authentic data set is created by the authors. The data set is collected from the students and lecturers of Kastamonu University, Faculty of Engineering. Data collection processes are carried out after the purpose of the research is explained to the people from whom the gait data are to be collected and their consent is obtained. Data have been collected from 23 people of different age in total, 13 being men and 10 being women. Fig. 2 shows the data collection and data collection setting.

To collect data from Kinect sensor, PyKinect2 wrapper (https://github.com/Kinect/PyKinect2) is used on Phyton programming language. During the data collection process, Kinect device is placed on a platform of 70 cm height. The gait information of individuals is obtained by recording their normal walk on a straight line towards Kinect device five times. Each walk takes about 2 sec, and 30 rows of data are collected for each individual per walk. Since Kinect assigns a negative infinite value to the coordinate information of the joint points that it cannot detect immediately, there are undefined data in the recorded phases of individuals' gait data. After obtaining the data of five gait phases of 23 individuals, virtual values are assigned to these undefined values within each gait phase with a simple algorithm. In this algorithm, if the undefined value is the first or the last value, the next or the previous value is assigned as the virtual value. In case the previous or next values are also undefined, the previous or next values are kept being checked until a defined value is found. The data manipulation is achieved by assigning a virtual value to undefined data between any two defined data values by taking the average of the previous and next data values.

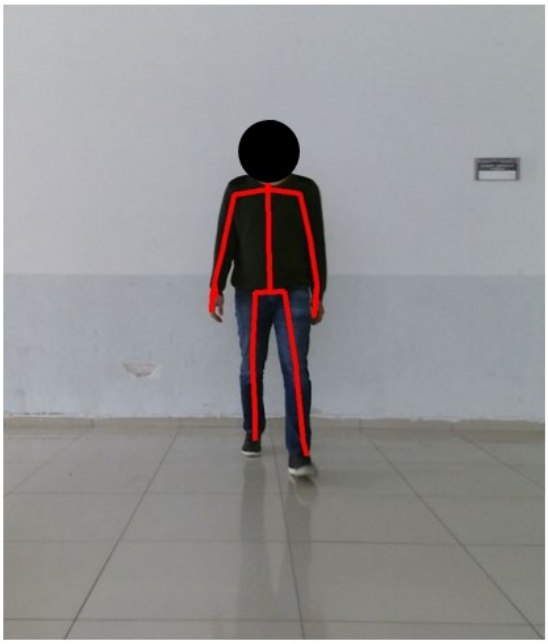


Fig. 2. Data collection setting and an exemplary moment of data collection.

The following features are also extracted based on the study of Jianwattanapaisarn et al. [13] during the data collection process: (i) the real-time three dimensional *X*, *Y* and *Z* coordinate information of 25 joint points, (ii) the *X*, *Y*, and *Z* vectors of the joint points (except for the leaf node joint points representing the head, left foot, right foot, left hand tip, right hand tip, left thumb and right thumb) calculated from the orientation data in accordance with the joint points they are associated to, and (iii) the angles of refraction of the joint points with orientation data calculated in accordance with the joint points they are associated to. The orientation data herein are the rotation motion of the joint points they make with the subsequent joint points towards leaf points and provided by Kinect device in quaternion form. Total of 147 features have been obtained from each frame. 75 of these 147 features are real world *X*, *Y*, and *Z* real-time coordinate information of the 25 joint points relative to Kinect camera. This information is provided by Kinect device. The remaining 72 features are the angles of refraction θ of 18 joint points (apart from 7 joint points index numbers of which are 3, 15, 19, 21, 23, 22 and 24, whose quaternion information cannot be provided by Kinect device as they are leaf joint points and therefore joint vectors and angles of refraction cannot be calculated) relative to

subsequent joint points they are related to and their *X*, *Y*, and *Z* vector values. These values are calculated through quaternion data provided by Kinect device. Said calculations are explained in the next section.

B. Obtaining Joint Vectors and Angles of Refraction

Kinect provides the orientation information of the joint points to the user in quaternion forms. In mathematics, quaternions are the number systems expanding the complex numbers to four dimensions, one being real and three being virtual. Quaternions have been defined and applied to mathematics in the three-dimensional space by the Irish mathematician Sir William Rowan Hamilton [34]. Quaternions are generally used to calculate the rotational movements in the three-dimensional space. Quaternary numbers are stated as in (1).

$$q = w + xi + yj + zk\ , \tag{1}$$

where q is a quaternary number; w, x, y and z are each real numbers and i, j and k are basic quaternion units. Basic quaternion units are imaginary unit numbers and are shown as in (2).

$$i^2 = j^2 = k^2 = ijk = -1. \tag{2}$$

Quaternions are defined as in (3) and quaternions consist of scaler and vector parts. If the equivalent of w is to be shown with qa as the scaler part (real dimension); then the equaivalent of $x + y + z$ is shown with $qb + qc + qd$ as the vector part (virtual dimensions).

$$Qo = qa + iqb + jqc + kqd. \tag{3}$$

Kinect provides the orientation quaternion of the joint points of users in $Qo(qa, qb, qc, qd)$ form. When represented in this form, orientation quaternions can be expressed as (4) for θ angle around one unit of axis considering similar to the axis-angle [35].

$$Qo = \cos\left(\frac{\theta}{2}\right) + i\left(x * \sin\left(\frac{\theta}{2}\right)\right) + j\left(y * \sin\left(\frac{\theta}{2}\right)\right) + k\left(z * \sin\left(\frac{\theta}{2}\right)\right), \tag{4}$$

where θ is the angle of refraction made with the associated joint point, and *x*, *y*, and *z* are vectors representing the rotation axis. After calculating the θ angle of refraction of this Qo orientation quaternion provided by Kinect as in (5), *x*, *y* and *z* magnitudes of the vector, part of orientation quaternion could be calculated as in (6).

$$\theta = \arccos(qa) \times 2. \tag{5}$$

$$x = qb/\sin\left(\frac{\theta}{2}\right), y = qc/\sin\left(\frac{\theta}{2}\right), z = qd/\sin\left(\frac{\theta}{2}\right). \tag{6}$$

C. Classification

The section explains the details of the RF, GB, XGB, KNN and DNN classifiers used for identification and the cascade voting methods used in this study. Among these classifiers, while Keras library of TensorFlow is used for the DNN, for all other algorithms, Scikit-learn library is used. In the training of classifiers, the entirety of the data set is used because the testing is performed in real time with the data obtained online from the Kinect device. In the training of RF, GB, and XGB classifiers, default parameters provided by Scikit-learn library are used. KNN classifier is trained with $k = 1$ parameter. The input layer of DNN classifier consists of 147 neurons, the output consists of 23. It has 3 hidden layers with 200 neurons. While the activation function is selected as ReLU for hidden layers, in the last layer, activation function is selected as Softmax. Adamax is chosen as the optimization algorithm, Sparse Categorical Crossentropy – as the loss function, 80 – as the batch size and 120 epoch training is carried out.

As mentioned earlier, the testing has been made in real time using a cascade voting method. The cascade voting method approach applied for identification is shown in Fig. 3. For testing, the data have been taken for 2 sec at 15 frames per second, 30 frames in total. If there is an undefined value in the instant data without any data manipulation being done on the data, that data are skipped and given as input to the machine learning models. The reason is that data manipulation processes that will be carried out where necessary by controlling all the received data increase the cost of and decrease the efficiency of the system.

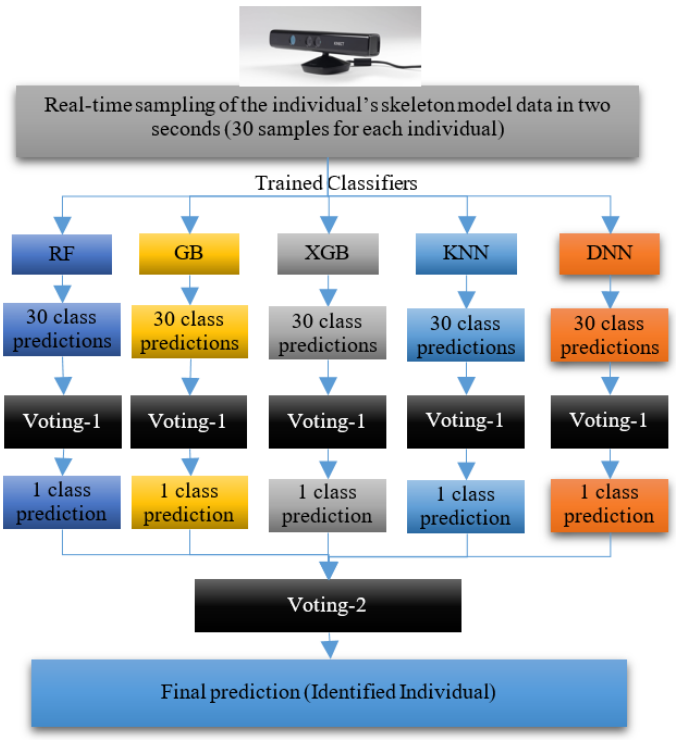


Fig. 3. The architecture of cascade voting approach.

In fact, one frame is sufficient for identification. However, since a real-time classification has been performed relying on a single frame of data, it can create security breaches. To avoid this situation, 30 frames are taken, and the classification results of each classifier are obtained for each frame. In other words, one classifier produces 30 class predictions in testing. Here, voting approach (Voting-1) is used to reduce 30 class predictions of a classifier to one and the class receiving most votes is accepted as the output of that classifier. By this way, individual classification predictions of 5 different classifiers are obtained. However, here the classification predictions of 5 different classifiers still creates an uncertainty. To eliminate this uncertainty, the voting approach is applied to the predictions of these 5 classifiers and the class receiving the most vote becomes the final class prediction of the system. By this way, the identified individual is clearly revealed.

IV. Results and Discussion

In this section, real-time classification performance of the machine learning algorithms and voting approach is given and is compared to previous studies. Real-time tests are carried out on four individuals. The first two are a middle-aged woman and a young woman, and the other two are young men. The 147 features to be used as input to classification algorithms for testing have been collected from each frame during video stream and used as a single row. Data have been collected from 30 frames with no undefined data in them and the collected data have been given to five machine learning algorithms. The reason for choosing the number of frames as 30 is because the depth sensor used to receive the three-dimensional data in the Kinect device can only be extracted from the skeleton model of the person detected between 1.5 and 4.5 meters away from the Kinect device. It has been observed that to cover the start and end stages of walking, 3-meter distance for the individuals who need to walk during the test, can be crossed over in the period of obtaining approximately 30 frames.

TABLE II
Real-Time Identification Performance of Machine Learning Algorithms

Person Code	RF (acc %)	GB (acc %)	XGB (acc %)	KNN (acc %)	DNN (acc %)	Voting (acc %)
20	100	100	100	40 (3, 6, 21)	90 (21)	100
21	100	90 (20)	100	80 (2,6)	100	100
22	80 (8, 10)	30 (4, 8, 10, 11)	40 (4, 8, 10)	90(8)	80 (8, 10)	90 (8)
23	100	100	100	50 (14, 19)	80 (14)	100
Mean Accuracy	95	80	85	65	87.5	97.5

In the experimental study conducted, 10 experiments have been made for each individual. Individual classification performance of the classification algorithms obtained by voting-2 approach is shown in Table II. There are 23 people within the data set used in training of the machine learning models. In real-time testing, tests have been carried out on 4 of these 23 individuals. These four individuals are represented with person codes as 20 for young male adolescent, 21 for young male, 22 for young female, and 23 for middle-aged female. The values given in parenthesis next to classification performance in Table II is the person code information that machine learning algorithm predicted wrongly.

The individual with person code 20 is misclassified with individuals with person codes 3, 6 and 21 by KNN for 6 times and is predicted wrongly; and is misclassified with individual with person code 21 by DNN algorithm and is predicted wrongly. For the individual with person code 21, GB algorithm has predicted wrongly once as individual with person code 20; and KNN has predicted wrongly as individuals with person codes 2 and 6. It should be stated that individuals with person codes 20 and 21 are siblings and are misclassified with each other in wrong predictions. For the individual with person code 22, RF, GB, XGB, and DNN algorithms have made wrong predictions as individuals with person codes 8 and 10. In addition, KNN and voting methods have predicted wrongly as individual with person code 8. Furthermore, GB algorithm has predicted wrongly as individuals with person code 4 and 11, and XGB algorithm has predicted wrongly as individual with person code 4. For the individual with person code 23, KNN and DNN algorithms have made wrong predictions as individual with person code 14. Also, KNN algorithm has predicted wrongly as individual with person code 19. When individuals predicted wrongly for one another are considered, it is seen that they are of the same sex generally.

When the classification performance of the machine learning methods and voting approach (Voting-2) are compared in terms of their accuracy values, the highest rate of accuracy (97.5 %) is obtained with a voting approach as can also be seen in Fig. 4. The voting approach has predicted individuals correctly apart from individual with person code 22. It has misclassified individual with person code 22 once and twice with individual with person code 8. The classification performance of other machine learning algorithms is as follows: RF (accuracy = 95 %), DNN (accuracy = 87.5 %), XGB (accuracy = 85 %), GB (accuracy = 80 %) and KNN (accuracy = 65 %), respectively. The lowest classification performance belongs to KNN.

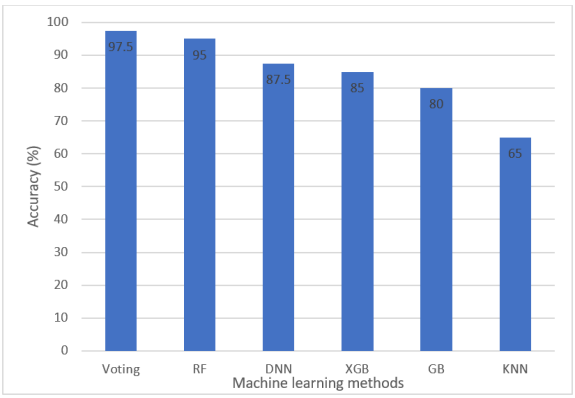


Fig. 4. The comparison of the classification performance of the classifiers in terms of their accuracy values.

The comparison of results obtained in this study with the previous studies is shown in Table III. In the literature, a few studies making offline identification from the data obtained from Kinect device stand out. Sahak et al. [5] have prepared a data set composed of the three-dimensional joint point information received from 30 individuals walking towards Kinect device. Then, they have performed an offline identification procedure with the support vector machine algorithm. The highest classification accuracy they obtained was 98.67 %. Wang et al. [36] have obtained a data set from Kinect device comprising the three-dimensional information of the joint points of the human skeleton model and two-dimensional information obtained from the silhouettes corresponding to these joint points of 52 individuals. Using this data set, they performed offline identification with KNN classification algorithm. The highest classification accuracy they obtained was 94.23 %.

In the literature, there also are studies performing real-time identification. Jiang et al. [37] have created their own data set of 10 individuals and performed real-time identification with KNN algorithm using the static and dynamic features they obtained by means of the three-dimensional information of the joint points of the human skeleton model from Kinect device. For the static features, the distances of some joint points from one another were calculated. The dynamic features consisted of the swinging angles of arms and legs which change during the gait. They obtained the highest identification performance with 82 % accuracy using static and dynamic features together. Jianwattanapaisarn et al. [13] have created their own data set of 90 individuals and performed identification using the gait features that they have obtained from individual's walking freely at separate times using the Kinect device. Detecting the gait subsequences with fixed length at random numbers and giving these gait subsequences to a group of extra-tree classifiers, they obtained predictions for an entire gait. As a result of the real-time performed tests, they obtained the highest identification performance with 92.22 % accuracy. Choi et al. [38] have created their own data set to perform real-time identification with a gait pattern of 12 individuals, using a frame level matching method to minimise the effects of the noisy gait patterns and to preserve the distinguishing power of each frame. In their study, they obtained the highest real-time identification performance with 95.94 % accuracy.

TABLE III
COMPARISON OF STUDY PERFORMANCE

Test Environment	Studies	Identification Performance (Accuracy (%))
Offline	Sahak et al. [5]	98.67
	Wang et al. [36]	94.23
Real-Time	Jiang et al. [37]	82
	Jianwattanapaisarn et al. [13]	92.22
	Choi et al. [38]	95.94
	Our study	**97.5**

The classification accuracy obtained in this study is higher than in the previous studies. It has a slightly lower classification performance from the study of Sahak et al. [5]. However, this study performs offline classification. For this reason, it is normal that the classification performance of our study is slightly lower. In real-time classification, classification performance usually decreases due to occasionally undefined data received from the depth sensor of Kinect device.

V. CONCLUSION

In this study, a system performing identification in real time via gait features using a Kinect device has been developed. The features within the data set have been obtained using a model-based approach. These features consist of real-world three-dimensional coordinates of 25 joint points of human skeleton model and the angles of refraction of some joint points with the joint points they are associated to. RF, DNN, XGB, GB and KNN classifiers have demonstrated the accuracy rate of 95 %, 87.5 %, 85 %, 80 % and 65 %, respectively. The highest classification performance has been obtained via a voting method (Voting-2) with 97.5 % accuracy rate. Obtaining a common classification result through a comparison of the results of the five machine learning algorithms, the voting method increases the classification performance. When the performance of the developed system is reviewed, it is seen that it is a system that can successfully perform real-time identification. As the distance of the depth sensor of the used Kinect device is limited, a system more tolerant to changes can be developed by using devices with larger depth sensor distance to observe the longer distance walks. Also, a limitation of this study is that, even though data have been collected from 23 individuals during the data collection phase, only four individuals have been reached in the testing phase due to Covid-19 pandemic.

REFERENCES

[1] A. K. Jain, A. Ross, and S. Prabhakar, "An introduction to biometric recognition," *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 1, p. 4–20, Jan. 2004. https://doi.org/10.1109/TCSVT.2003.818349

[2] A. Drosou, D. Loannidis, D. Tzovaras, K. Moustakas, and M. Petrou, "Activity related authentication using prehension biometrics," *Pattern Recognition*, vol. 48, no. 5, p. 1743–1759, May 2015. https://doi.org/10.1016/j.patcog.2014.12.008

[3] B. Miller, "Vital signs of identity [biometrics]," *IEEE Spectrum*, vol. 31, no. 2, p. 22–30, 1994. https://doi.org/10.1109/6.259484

[4] K. A. Reynolds, P. M. Watt, S. A. Boone, and C. P. Gerba, "Occurrence of bacteria and biochemical markers on public surfaces," *International Journal of Environmental Health Research*, vol. 15, no. 3, p. 225–234, 2005. https://doi.org/10.1080/09603120500115298

[5] R. Sahak, N. Tahir, A. I. M. Yassin, F. H. Kamaruzaman, and A. A. Misreb, "Human gait recognition using skeleton joint coordinates with orthogonal least square and locally linear embedded techniques," *International Journal of Simulation: Systems, Science and Technology*, vol. 19, no. 5, p. 25.1–25.9, 2019. https://doi.org/10.5013/IJSSST.a.19.05.25

[6] P. C. Cattin, "Biometric Authentication System Using Human Gait", Doctoral Thesis, ETH Zurich, Zurich, Switzerland, 2002.

[7] A. Kale, A. Sundaresan, A. N. Rajagopalan, N. P. Cuntoor, A. K. R. Roy-Chowdhury, V. Kruger, and R. Chellappa "Identification of humans using gait", *IEEE Transaction on Image Processing*, vol. 13, no. 9, p. 1163–1173, Aug. 2004. https://doi.org/10.1109/TIP.2004.832865

[8] N. Sadeghzadehyazdi, T. Batabyal, N. K. Dhar, B. O. Farmiloni, K. M. Iftekharuddin, and S. T. Acton, "GlidarCo: gait recognition by 3D skeleton estimation and biometric feature correction of flash lidar data", *arXiv preprint arXiv:1905.07058,* 2019.
[9] H. Fujiyoshi, and A. J. Lipton, "Real-time human motion analysis by image skeletonization," in *IEEE Workshop on Applications of Computer Vision*, Princeton, NU, USA, Oct. 1998, pp. 15–21. https://doi.org/10.1109/ACV.1998.732852
[10] H. Sarbolandi, D. Lefloch, and A. Kolb, "Kinect range sensing: Structured-light versus Time-of-Flight Kinect," *Computer Vision and Image Understanding*, vol. 139, pp. 1–20, Oct. 2015. https://doi.org/10.1016/j.cviu.2015.05.006
[11] A. Bobick and A. Y. Johnson, "Gait recognition using static activity-specific parameters," in *IEEE 2001 Computer Society Conference on Computer Vision and Pattern Recognition*, Kauai, USA, Dec. 2001. https://doi.org/10.1109/CVPR.2001.990506
[12] J. P. Singh and S. Jain, "Person identification based on gait using dynamic body parameters," in *Trendz in Information Sciences & Computing,* Chennai, India, Dec. 2010, pp. 248–252. https://doi.org/10.1109/TISC.2010.5714649
[13] N. Jianwattanapaisarn, N. Khamsemanan, and C. Nattee, "Gait recognition from freestyle walks using relative coordinates and random subsequence-based Sum-Rule classification," in *Joint 8th International Conference on Soft Computing and Intelligent Systems (SCIS) and 17th International Symposium on Advanced Intelligent Systems (ISIS)*, Sapporo, Japan, Aug. 2016, pp. 5–9. https://doi.org/10.1109/SCIS-ISIS.2016.0016
[14] Y. Wang, S. Rane, S. C. Draper, and P. Ishwar "A theoretical analysis of authentication, privacy and reusability across secure biometric systems," *IEEE Transactions on Information Forensics and Security,* vol. 7, no. 6, pp. 1825–1840, Dec. 2012. https://doi.org/10.1109/TIFS.2012.2210215
[15] A. El-Sallam, F. Sohe, and M. Bennamoun, "Robust pose invariant shape-based hand recognition," in *IEEE 2011 6th Conference on Industrial Electronics and Applications*, Beijing, China, June 2011, pp. 1–10. https://doi.org/10.1109/ICIEA.2011.5975595
[16] K. Simoens, P. Tuyls, and B. Preneel, "Privacy weaknesses in biometric sketches," in *IEEE 2009 30th Symposium on Security and Privacy*, Oakland, USA, May 2009, pp. 188–203, doi: 10.1109/SP.2009.24.
[17] A. Eng and L. A. Wahsheh, "Look into my eyes: A survey of biometric security," in *10th International Conference on Information Technology: New Generations*, Las Vegas, USA, Apr. 2013, pp. 422–427. https://doi.org/10.1109/ITNG.2013.65
[18] J. Lu, X. Yua, and T. Yahagi, "A method of face recognition based on fuzzy c-Means clustering and associated Sub-NNs," *IEEE Transactions on Neural Networks*, vol. 18, no. 1, p. 150–160, Jan. 2007. https://doi.org/10.1109/TNN.2006.884678
[19] D. Souza, A. Burlamaqui, and G. Souza Filho, "Improving biometrics authentication with a multi-factor approach based on optical interference and chaotic maps," *Multimedia Tools Applications*, vol. 77, p. 2013–2032, 2018. https://doi.org/10.1007/s11042-017-4374-x
[20] K. Dharavath, F. A. Talukdar, and R. H. Laskar, "Study on biometric authentication systems, challenges and future trends: A review," in *IEEE 2013 International Conference on Computational Intelligence and Computing Research,* Enathi, India, Dec. 2013, pp. 1–7. https://doi.org/10.1109/ICCIC.2013.6724278
[21] A. N. Kataria, D. M. Adhyaru, A. K. Sharma, and T. H. Zaveri, "A survey of automated biometric authentication techniques," in *IEEE 2013 Nirma University International Conference on Engineering*, Ahmedabad, India, Nov. 2013, pp. 1–6. https://doi.org/10.1109/NUiCONE.2013.6780190
[22] W. Meng, D. S. Wong, S. Furnell, and J. Zhou, "Surveying the development of biometric user authentication on mobile phones," *IEEE Communications Surveys and Tutorials,* vol. 17, no. 3, pp. 1268–1293, 2015. https://doi.org/10.1109/COMST.2014.2386915
[23] S. Pal, U. Pal, and M. Blumenstein, "Signature-based biometric authentication," in *Computational Intelligence in Digital Forensics: Forensic Investigation and Applications. Studies in Computational Intelligence*, A. Muda, Y. H. Choo, A. Abraham, N. S. Srihari, Eds., vol 555. Springer, Cham, pp. 285–314, 2014. https://doi.org/10.1007/978-3-319-05885-6_13
[24] D. Impedovo, G. Pirlo, L. Sarcinella, E. Stasolla, and C. A. Trull, "Analysis of stability in static signatures using Cosine similarity," in *International Conference on Frontiers in Handwriting Recognition,* Bari, Italy, Sep. 2012, p. 231–235. https://doi.org/10.1109/ICFHR.2012.180
[25] R. Plamondon and M. A. Djioua, "Multi-level representation paradigm for handwriting stroke generation," *Human Movement Science*, vol. 25, no. 4–5, p. 586–607, 2006. https://doi.org/10.1016/j.humov.2006.07.004
[26] N. K. Hosseini and M. J. Nordin, "Human gait recognition: A silhouette based approach," *Journal of Automation and Control Engineering*, vol. 1, no. 2, p. 103–105, 2013. https://doi.org/10.12720/joace.1.2.103-105
[27] A. Veeraraghavan, A.K.R. Chowdhury, and R. Chellappa, "Role of shape and kinematics in human movement analysis," in *IEEE 2004 Computer Society Conference on Computer Vision and Pattern Recognition*, Washington, USA, June-July 2004, pp. I–I. https://doi.org/10.1109/CVPR.2004.1315104
[28] D. Cunado, M. S. Nixon, and J. N. Carter, "Automatic extraction and description of human gait models for recognition purposes", *Computer Vision and Image Understanding*, vol. 90, no. 1, pp. 1–41, 2003. https://doi.org/10.1016/S1077-3142(03)00008-0
[29] J. P. Singh, S. Jain, S. Arora, and U. P. Singh, "A survey of behavioural biometric gait recognition: Current success and future perspectives," *Archives of Computational Methods in Engineering*, vol. 28, pp. 107–148, Nov. 2021. https://doi.org/10.1007/s11831-019-09375-3
[30] C. Yam, M. S. Nixon, and J. N. Carter, "Automated person recognition by walking and running via model-based approaches," *Pattern Recognition*, vol. 37, no. 5, pp. 1057–1072, 2004. https://doi.org/10.1016/j.patcog.2003.09.012
[31] J. Y., Kang, R. Kim, H. Kim, Y. Kang, S. Hahn, Z. Fu, M. I. Khalid, E. Schenck, and T. Thesen, "Automated tracking and quantification of autistic behavioural symptoms using Microsoft Kinect," *Studies in Health Technology and Informatics,* vol. 220, pp. 167–170, Jan. 2016.
[32] E. Gianaria and M. Grangetto, "Robust gait identification using Kinect dynamic skeleton data", *Multimedia Tools and Application,* vol. 78, pp. 13925–13948, 2019. https://doi.org/10.1007/s11042-018-6865-9
[33] M. W. Rahman and M. L. Gavrilova, "Kinect gait skeletal joint feature-based person identification," in *IEEE 2017 16th International Conference on Cognitive Informatics and Cognitive Computing*, Oxford, UK, July 2017, pp. 423–430. https://doi.org/10.1109/ICCI-CC.2017.8109783
[34] W. R. Hamilton, "II. On quaternions, or on a new system of imaginaries in algebra," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 25, no. 163, pp. 10–13, 1844. https://doi.org/10.1080/14786444408644923
[35] W. R. Hamilton, "XXXIX. On quaternions, or on a new system of imaginaries in algebra," *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science*, vol. 25, no. 166, pp. 241–246, 1844. https://doi.org/10.1080/14786444408644984
[36] Y. Wang, J. Sun, J. Li, and D. Zhao, "Gait recognition based on 3D skeleton joints captured by Kinect," in *IEEE 2016 International Conference on Image Processing*, Phoenix, USA, Sep. 2016, pp. 3151–3155. https://doi.org/10.1109/ICIP.2016.7532940
[37] S. Jiang, Y. Wang, Y. Zhang, and J. Sun, "Real time gait recognition system based on Kinect skeleton feature," in *Asian Conference on Computer Vision,* Singapore, Nov. 2014, pp. 1–11. https://doi.org/10.1007/978-3-319-16628-5_4
[38] S. Choi, J. Kim, W. Kim, and C. Kim, "Skeleton-based gait recognition via robust frame-level matching," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 10, pp. 2577–2592, 2019. https://doi.org/10.1109/TIFS.2019.2901823



Berk Ercin received B. sc. ing. Degree in Computer Engineering from Kastamonu University, Turkey in 2020. At the moment, he is a Master student at Akdeniz University, Turkey. His current research interests include machine learning, data mining and game development.
E-mail: berkerciks@gmail.com
ORCID iD: https://orcid.org/0000-0002-0440-1528



Abdulkadir Karacı received the Doctoral degree in Electronic and Computer Education (intelligent tutoring system, machine learning and fuzzy logic) from Gazi University, Turkey, in 2013. He is currently an Associate Professor at the Department of Computer Engineering, Kastamonu University. His current research interests include machine learning, convolution neural network, intelligent tutoring system and programming languages.
E-mail: akaraci@kastamonu.edu.tr
ORCID iD: https://orcid.org/0000-0002-2430-1372