Düzce University
Journal of Science & Technology

Research Article

Person Recognition from Gait Analysis for Smart Spaces by using MLP-based DNN model

Cüneyt YÜCELBAŞ [a,*]

[a] *Department of Electronics and Automation, Vocational School of Technical Sciences at Mersin Tarsus Organized Industrial Zone, Tarsus University, Mersin, TURKEY*
** Corresponding author's e-mail address: cuneytyucelbas@tarsus.edu.tr*
DOI: 10.29130/dubited.1187065

ABSTRACT

In smart fields, security measures are taken to protect people against threats that may arise by using technology and to provide crisis management, and the functions of measuring area security and ensuring its effectiveness are carried out. As an element of this measurement, it is thought that person recognition may be the most important factor in the future. It is seen that deep learning-based algorithms, which can provide fast and high-accuracy results with many data, will be an integral part of this sector in the future as they are today. However, when the literature is examined, it is understood that the number of research in which Deep learning algorithms are used in order to increase the success of the studies in this direction and the system practicality is insufficient. For this reason, in this study, deep learning was used to recognize people by using the walking data of 15 people obtained thanks to wearable sensors. Since the increase in the diversity of the data will positively affect the learning of the created model, data augmentation has been made and these data have been classified in the multilayer perceptron-based deep neural network model. The results were statistically analyzed and showed that this model exhibited excellent performance in person recognition from walking data. In addition, the accuracy rate was found to be 100%, and it proved that the method used to increase the data also produced successful results in walking data. It is thought that the success of the study can provide important perspective support to new studies for smart fields in the literature.

Keywords: *Artificial intelligence, Deep neural networks, Gait analysis, Person recognition, Smart fields*

MLP Tabanlı DNN Modeli Kullanılarak Akıllı Alanlar İçin Yürüyüş Analizinden Kişi Tanıma

ÖZET

Akıllı alanlarda, teknoloji kullanılarak oluşabilecek tehditlere karşı insanları korumak ve kriz yönetimi sağlamak için güvenlik önlemleri alınmakta, alan güvenliğinin ölçülmesi ve etkinliğinin sağlanması işlevleri yerine getirilmektedir. Bu ölçümün bir unsuru olarak gelecekte kişi tanımanın en önemli faktör olabileceği düşünülmektedir. Birçok veri ile hızlı ve yüksek doğrulukta sonuçlar verebilen derin öğrenme tabanlı algoritmaların bugün olduğu gibi gelecekte de bu sektörün ayrılmaz bir parçası olacağı görülmektedir. Ancak literatür incelendiğinde bu yöndeki çalışmaların başarısını artırmak için derin öğrenme algoritmalarının kullanıldığı araştırma sayısının ve sistem pratikliğinin yetersiz olduğu anlaşılmaktadır. Bu nedenle bu çalışmada giyilebilir sensörler sayesinde elde edilen 15 kişinin yürüme verileri kullanılarak insanları tanımak için derin öğrenme kullanılmıştır. Veri çeşitliliğindeki artış, oluşturulan modelin öğrenilmesini olumlu etkileyeceğinden veri artırması yapılmış ve bu veriler çok katmanlı algılayıcı tabanlı derin sinir ağ modelinde sınıflandırılmıştır. Sonuçlar istatistiksel olarak analiz edilmesi sonucunda önerilen modelin yürüme verilerinden kişi tanımada

Received: 10/10/2022, Revised: 28/01/2023, Accepted: 09/02/2023

mükemmel performans sergilediğini göstermiştir. Ayrıca doğruluk oranı %100 bulunmuş ve verileri artırmak için kullanılan yöntemin yürüme verilerinde de başarılı sonuçlar ürettiği kanıtlanmıştır. Çalışmanın başarısının literatürde akıllı alanlara yönelik yeni çalışmalara önemli bir perspektif desteği sağlayabileceği düşünülmektedir.

Anahtar Kelimeler: *Yapay zeka, Derin sinir ağları, Yürüyüş analizi, Kişi tanıma, Akıllı alanlar*

I. INTRODUCTION

Artificial Intelligence (AI) is a computer software that imitates human movements using a lot of data. Among the most important features of AI artificial intelligence are machine learning (ML), person identification through voice and speech, and virtual assistant inclusion. Integrating with ML, AI not only provides users with the necessary functionality, but also simplifies the business process. In addition, AI software is used to create and improve a smart implementation from scratch, thanks to its ML and deep learning (DL) skills [1]. Areas equipped with AI technology that connect all computing systems, various machines, animals, objects, and even people that can generate information and share it over the internet are called smart fields [2]. Today, companies of different sizes in almost all sectors use this technology to work more efficiently, improve business processes, make more effective decisions and increase the value of their business.

Creating smart systems with DL is an issue that has been on the agenda of large technology companies in recent years. DL can be used for prediction mechanisms in many smart fields. For example, results from genome changes can be predicted from the information of dozens of living cells. In other words, DL can be used to understand a particular disease of the person or to detect abnormal conditions [3]. In addition to creating a more secure system, through the applications developed in the security industry, the capacity to predict where the next attack will come from can be improved. Recognizing new threats that have never been detected before and thus being one step ahead of cybercrime organizations is possible with DL methods [4]. Another example is systems developed for vehicles using driver-assist services. Standard type vehicles may require the intervention of users in case of unexpected events. Recently, these vehicles have started to be replaced by vehicles that can travel without the need for a driver. The power of DL is used to simulate various driving conditions of the vehicle and to achieve an error-free autonomous driving [5]. Another is the developments in search engines. As it is known, text analysis has a very important place in the functioning mechanism of search engines. Indexing a site, generating its rank value, and presenting it to the user with logical sequences in the search mechanism are very important features for a search engine. DL algorithms are also used extensively in the development of these features [6]. As can be seen, the effectiveness of DL algorithms is very important in increasing the practicality and reliability of smart applications. In [7], the authors have conducted research on emotion-based gait classification. In this context, it has been tried to determine the types of emotions from gait characteristics and physical movements by using multiple DL methods. As a result, a success rate of 92% was achieved [7]. Researchers in [8] have worked from a different perspective to classify gait patterns with multi-scale learning. Briefly, neurophysiological signals such as electroencephalogram (EEG) and electromyogram (EMG) of subjects were used for gait detection. As a result of the analysis, 89.33% classification success was achieved [8]. In [9], various features were inferred from gait data obtained from many volunteers. Then, these features are presented as input to the support vector machine (SVM), k-nearest neighbors (kNN), and multilayer perceptron (MLP) classifiers. In conclusion, MLP showed the best performance in gait separation with approximately 94% rate [9]. One of the studies carried out in recent years was carried out by Lee et al. [10]. In this paper [10] stated that a new approach was presented to determine the best feature set for gender determination via gait. When this approach is used together with SVM and random forest, success rates of 99.11% and 98.89% were obtained, respectively [10].

As it is known, recently, analyses with biometric data have come to the fore in areas equipped with smart systems. Biometric data are biological measurements or physical features that can be utilized to characterize persons. Fingerprint matching, face recognition, and retinal scans are just the most well-known biometric identification technologies [11]. Physical features that are individual even in twins can be benefitted to replace or support the computer security logins, telephones, and staff-specific rooms and buildings. Unauthorized access to systems can be prevented at a higher level thanks to user-dependent security systems that have been previously defined. In order to make these systems more specific, physical characteristics such as weight and height of the authorized person can also be included in the system [12].

Researchers suggest that ears, veins, facial features, smells, sitting and walking styles may be individual differences and descriptive features. Based on this information, it is possible to reach the fact that it is possible to identify gender and/or person with an acceptable level of success from walking data of individuals, thanks to AI algorithms [13, 14]. However, when the literature is examined, it is obvious that it is necessary to carry out different researches using DL algorithms to increase the performance of the studies to a higher level in this direction and the practicality of the system. Because current technology-based smart systems are progressing within the framework of DL algorithms, as stated in the previous paragraphs. For this reason, it is thought that this gap in the literature should be filled with many studies with demonstrable statistical results. In this paper, which was carried out for the stated purpose, DL was used for person recognition by making use of walking data obtained thanks to wearable sensors. It is thought that the success of the study can provide an important perspective support to new studies for smart fields (in many aspects such as security, financial return, analysis, life, and DL) in the literature. The process steps of the study are shown in Figure 1.

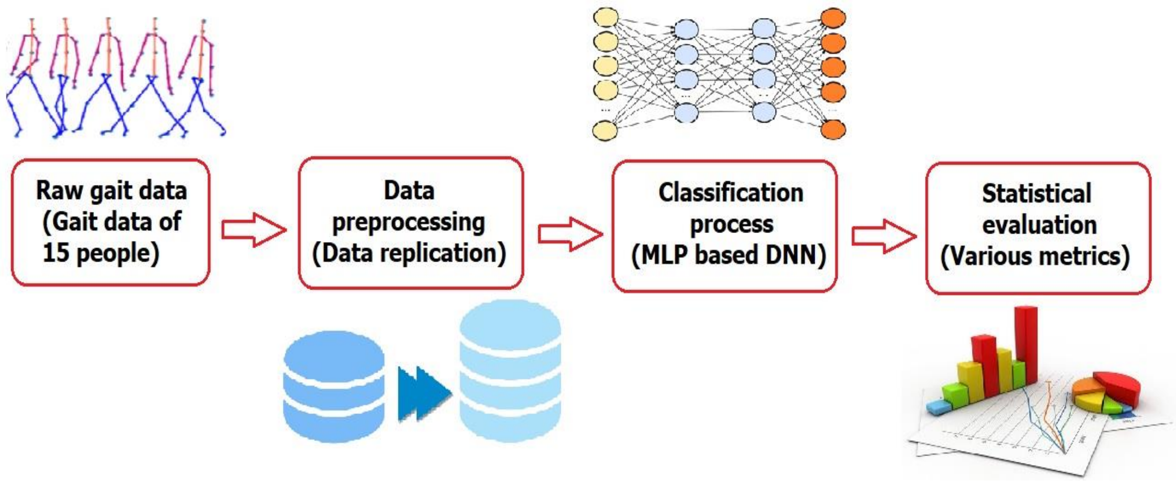

Figure 1. *The process steps*

II. MATERIALS AND METHODS

In order to ensure the forward movement of the trunk during walking, the completion of the movement of the heel to the ground for the first time and the second touch of the same heel to the ground and the continuation of this process is called the walking cycle. During the walking movement, one leg is thrown forward; after full contact of the thrown leg, the front movement of the second leg begins. The stance phase constitutes 60% of the walking cycle and the swing phase 40%. The stance phase begins with the contact of the heel to the ground and ends with the toes of the same foot leaving the ground. Figure 2 shows the phase distribution of a walking cycle.

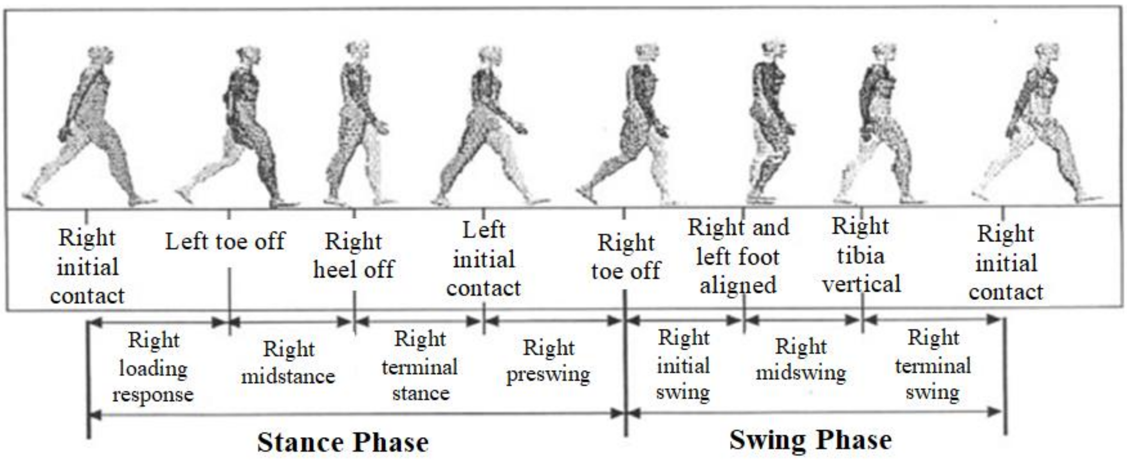

Figure 2. *Phase distribution of the walking cycle [15]*

Wearable technology industry is one of the most attacking and developing sectors. Technologies developed in this area are used in all areas of life such as sports, health, personal accessories and entertainment equipment. In wearable designs where sensors are actively used, wireless connection methods such as bluetooth and wi-fi are used in order to transfer data and infer from them. To give an example of wearable devices for this study, Figure 3 includes wearable sensor options for gait analysis.

Figure 3. *Wearable sensor options for gait analysis [16]*

In this study, the dataset created by calculating the gait parameters of a total of 15 different male and female volunteers was used by using the wearable gait analysis sensor (Physilog 5 sensor produced by Gaitup), which includes an accelerometer, gyroscope and barometric pressure sensor [13, 17]. The age range of the people is 20-34, and the weight range is 53-95. In order to calculate all gait parameters, subjects were carried out 3 times around a rectangular plane with an area of 10m x 5m. Baseline, temporal, spatial and elevation parameters were calculated for each gait of the individuals. Details of these parameters were given in [13, 17]. As a result, the data used in this study has 45 samples and 321 parameters [13].

A. PRODUCTION OF THE NEW DATA

As it is known, DL algorithms can obtain more accurate results with more data. Because the increase in the diversity of the data will make a positive contribution to the training of the model. Due to the insufficient number of data specified in chapter 2, the total number of records for each person in this study was increased by generating new data. This augmentation process was carried out by adding a random number between “0-1” to the existing 3 samples for each person. As a result, 16 new samples were derived from each sample of the individuals. Thus, 720 new samples belonging to 15 existing people were obtained. In total, person recognition process was started on 765 pieces of data. The matrix view of the first dataset and the new dataset obtained after data generation is shown in Figure 4.

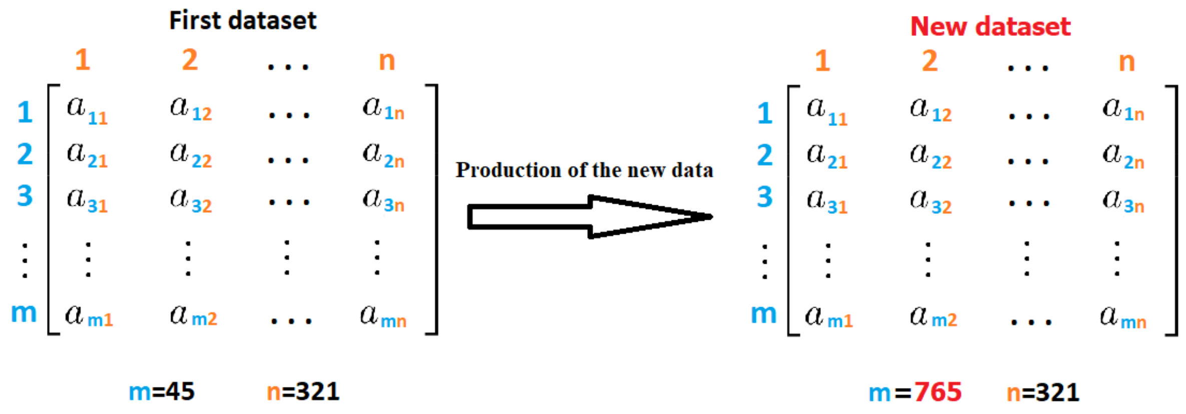

Figure 4. *Matrix views of the first dataset and the new dataset obtained after data generation*

B. CLASSIFICATION PROCESS: MLP-BASED DNN MODEL

Neural networks are the algorithms generated with the aim of automatically performing some abilities of the brain, such as deriving new implications, creating and discovering new data through learning, without any help [18]. The basis of this algorithm is the learning stages of the human brain. Flow diagrams were obtained by creating mathematical approaches of these stages. Thanks to the algorithm created as a result of all these stages, abilities such as the structure of the living brain, learning, and remembering are imitated [19]. There are many areas where this algorithm, which has emerged as an expert system, is used. The foremost of these is classification and estimation followed by purposes such as correlation, filtering, and control. In order to determine which mesh is more suitable for which problem, it is necessary to compare the properties of the networks with the properties of the problems.

In the MLP model, the data to be presented to the system is a one-way feed-forward neural network, which is taken from the input layer, processed in the hidden layers, and resulted in the output layer [20]. The neurons in the layers are interconnected and have their own weights. The same activation function is operated for the neurons in each layer in order to ensure that the processes are carried out smoothly as a whole. Generally, sigmoid has been used as a function of neurons in the hidden layer in many research and applications. This function is preferred as sigmoid or linear for the output layer [20] (Figure 5). The performance of the created MLP network depends on many factors. These are the layer, the number of neurons in the layers, the arrangement of data and variables; there are many other important factors such as learning rate, momentum constant, number of iterations, and stopping criteria from model parameters. Among these, learning rate and momentum constant parameters come to the fore as important internal factors affecting the learning process. An MLP model with only one hidden layer can complete the target function with high error. On the other hand, more hidden layers are more likely to complete the training phase with higher accuracy [21].

In its most basic definition, a deep neural network (DNN) is an algorithm that has more hidden layers and neurons than a traditional artificial neural network [22, 23]. There are many neural network models in the literature. All of them have the same common components such as layers, neurons, weights and functions [23]. These components, which imitate the work of our brain, have the feature of being trained with any network model. The most prominent and preferred advantage of the DNN algorithm is that it can model nonlinear problems that are difficult to solve [24]. Using more hidden layers than a traditional network offers the opportunity to combine features from previous layers. As a result, complex problems and data are solved with less error [22]. A DNN model, in its simplest definition, is a feed-forward network and is the processing of the presented data from the input to the output layer without feedback. In the first step in this algorithm, random numerical weight values are assigned to the connections between the generated neurons. A value between 0 and 1 is obtained by multiplying the inputs by the weights. If the created network model cannot make the desired definition, the weight values are revised within the framework of a certain rule [24]. The update operations are repeated in each iteration until these operations are detected with the desired/targeted accuracy of the data. In order for DNN to exhibit high performance, many training parameters such as the number of layers and neurons, and the learning rate should be precisely adjusted. Determining the optimum parameter values may not be possible due to time and computational cost. Instead, computational operations can be performed on many samples at the same time, instead of dealing with data one by one (grouping [25]) to speed up the network. Another way is to determine the optimum parameters by using the trial-and-error method, and as a result, the network can be accelerated. Because computers with multi-core processors can perform high-level processing and are suitable for multi-dimensional matrix calculations, these architectures increase the network training speed [26, 27]. Therefore, it is important that the architectural structures, processor capacities, and high level of the processors that perform the graphics operations of the computers where the DNN model will be used are important.

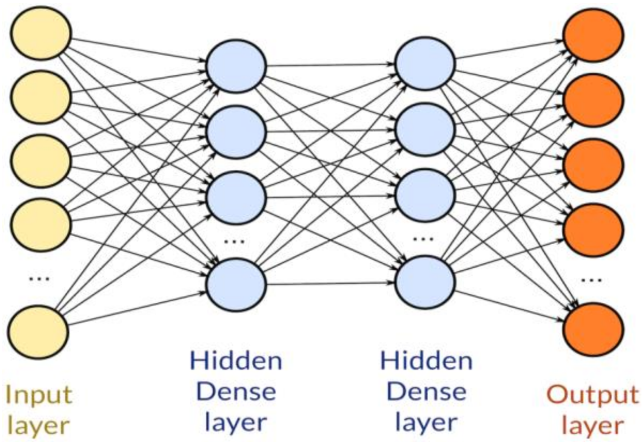

Figure 5. *The general architecture of MLP-based DNN model*

In deep learning applications, the number and size of data are important for the learning phase. The number of times that all training data is processed by the network during the learning phase is defined as the epoch. This parameter can be selected high, but the training process can be terminated before, according to certain criteria. The epoch number was taken as 100 for this study. If the update of the relevant model parameters for training is done after the entire dataset has been processed, both time loss and transaction cost will increase. In order to prevent this situation, the training set is divided into small parts, and learning is performed on them. The size of the small pieces that are separated is called batch_size. Selecting the batch_size value as 1 means that parameter updates are made after each data. This can lead to very different undesirable results. If the relevant parameter is selected as the whole number of data, the processing power will be insufficient since the process will take too long. When the literature is examined, it has been seen that the batch_size value is generally taken as a multiple of 2, such as 2, 4, 8, 16, 32, 64, ..., 512. In this research, the learning process was carried out successfully by taking the batch size value as 32. Another parameter is the learning rate. The step size to be taken on the path followed to reach the optimum point on the loss function is expressed as the learning rate. The learning rate is a hyperparameter that is determined before the training, and if it is selected as too small or too large, it may cause disadvantages such as longer training time and inability to reach the determined target. Therefore, choosing an ideal ratio value is important for a healthy learning phase. Generally, 0.1, 0.01, or 0.001 are used as default values in the literature for learning rate. In the learning phase of the network created in the study, this ratio was taken as 0.001 and successful results were obtained.

Sequential model of MLP-based DNN algorithm was used in this study. The sequential model allows you to create one or more layers. As a layer, "dense", which can work in many situations and is defined as a standard layer, was chosen. In this layer, all neurons in the previous layer are connected to those in the current layer. As a result of various trials, since the performance output was the highest, 64 nodes were used in the input layer, 128 and 64 were used in the 8 hidden layers, and finally 15 neurons were used for the output layer. The numbers of node and hidden layer were determined by examining deep learning applications used for different topics in the literature and by making necessary trials step by step through trial-and-error method. In this way, the best system performance was achieved. ReLu (Rectified Linear Unit) was preferred as the activation function. ReLu, which produces output in the range of $[0, +\infty)$ according to the input values, is faster than other functions because it has a less computational load. In addition, it contributes to the faster operation of the network created because it produces a zero value against negative inputs. As is known, loss functions aim to calculate how well the model training is. In this study, categorical crossentropy was chosen as the loss function. Briefly, this function detects the cross entropy loss between expected and predicted results. Since our dataset has more than two class labels as output, this function should be used for correct evaluation. In addition, the “adam” (adaptive moment) algorithm, which is generally the most used method in the literature, was chosen as the optimizer. This optimization is a stochastic gradient descent technique based on the adjustable estimation of first and second order moments [28].

C. STATISTICAL EVALUATION METRICS

In AI-based expert systems, a confusion matrix is needed to evaluate the performance over the true-false numbers as a result of the comparison between the predicted and the actual value. These are true positive (TP), false positive (FP), true negative (TN), and false negative (FN). The numerical values in this matrix are used to express the performance of the system through various statistical criteria. Accuracy rate (ACR) (1), Precision (2), Recall (3), Matthews Correlation Coefficient (MCC) (4), F1-score (5) [29-32], and Cohen's Kappa Coefficient (Kappa) [33], criteria were used to evaluate the performance of the DNN model used in this study. Kappa, one of these criteria, is used in statistical evaluation in many scientific fields. The main purpose of this evaluation criterion is to measure the degree of agreement or disagreement in identifying a particular target [34, 35]. In other words, Kappa is a statistical method that comparatively determines the reliability of agreement between the results of two observers [34, 35].

$$ACR = \left(\frac{CCI}{TNI}\right) \times 100 \tag{1}$$

$$Precision = \frac{TP}{TP + FP} \tag{2}$$

$$Recall = \frac{TP}{TP + FN} \tag{3}$$

$$MCC = \frac{(TP \times TN) - (FP \times FN)}{\sqrt{(TP + FP)(TP + FN)(TN + FP)(TN + FN)}} \tag{4}$$

$$F1\text{-}score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (5)$$

In addition to these statistical criteria, verification/validation - training achievements and verification/validation and training losses were calculated and curves were drawn.

III. EXPERIMENTAL RESULTS

In this paper, a dataset created by calculating the gait parameters of a total of 15 different male and female volunteers using a wearable gait analysis sensor, which includes an accelerometer, gyroscope and barometric pressure sensor [13, 14], was used. The sequential model of the MLP-based DNN algorithm was chosen as the classifier. Since DL algorithms can obtain more accurate results with more data, data augmentation has been made and the person recognition process has been started over a data matrix of [765x321] in total. The total data were randomly divided into 10 and the [77x321] data matrix was reserved for testing. Model training processes were carried out using the remaining [688x321] data. In this training phase, 10% of the training data was used for validation. After the model parameters were created, the dataset determined as a test was presented to the system and statistical performance metrics were calculated. Figure 6 shows the confusion matrix results obtained for each person in the test dataset.

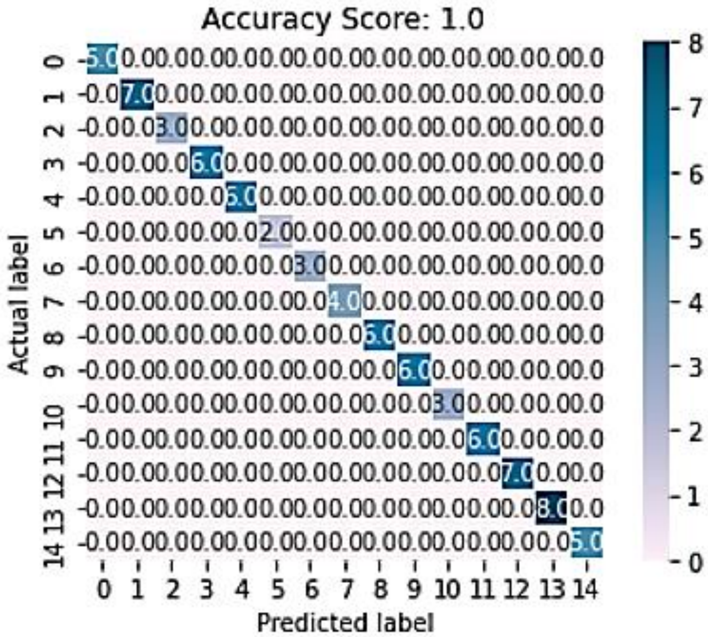

Figure 6. *Confusion matrix results of the test dataset for each person*

As can be seen in Figure 6, recognition was performed without any problems for each person in the test dataset and the ACR was calculated as 100%. This shows that the parameters of the MLP-based DNN model trained with the training-validation dataset are determined perfectly. The results of other statistical parameters calculated using this matrix are given in Table 1.

Table 1. *Obtained weighted average results for the test dataset with MLP based DNN model (TNI: Total Number of Instances, CCI: Correctly Classified Instances)*

Statistical Parameters	TNI	CCI	Precision	Recall	F1-score	MCC	Kappa	ACR %

Test Dataset	77	77	1	1	1	1	1	100

When Table 1 is examined, it is seen that all statistical parameters were calculated as 1 for the test dataset with 77 samples as a result of classification. This shows that the MLP-based DNN model has excellent performance in person recognition from walking data. In addition, the ACR was found to be 100%, and it proved that the method used to increase the data also produced a successful result in walking data. In Table 2, the ACRs and loss values obtained during the training of the MLP-based DNN model with the training and validation datasets are shown for each epoch.

Table 2. *ACRs and loss values for the training and validation datasets of the trained MLP based DNN model*

Number of epoch	**Training dataset**		**Validation dataset**	
	ACR	**Loss value**	**ACR**	**Loss value**
1	0.1793	2.7795	0.3043	2.0945
2	0.3457	2.0090	0.4783	1.6793
3	0.5137	1.5550	0.5217	1.1612
4	0.6511	1.0555	0.7681	0.8699
5	0.6333	1.0788	0.5072	1.2448
6	0.7851	0.7368	0.9420	0.3430
7	0.8045	0.8819	0.9710	0.2525
8	0.9661	0.1670	1	0.0394
9	0.9871	0.0496	1	0.0387
10	0.9952	0.0247	1	0.0096
11	1	0.0088	1	0.0033
12	1	0.0027	1	0.0017
13	1	0.0018	1	0.0014
14	1	0.0014	1	0.0011
15	1	0.0012	1	8.6304e-04

In Figure 7, the ACRs and loss values obtained during the training of the MLP-based DNN model with the training and validation datasets are shown graphically. When Table 2 and graphics are examined; in both datasets, it is seen that the network learns quickly from the first iterations (epoch), after 7 iterations, the network continues to learn, as can be seen from the ups and downs, and at the end of 11 iterations, the training is 100% successful.

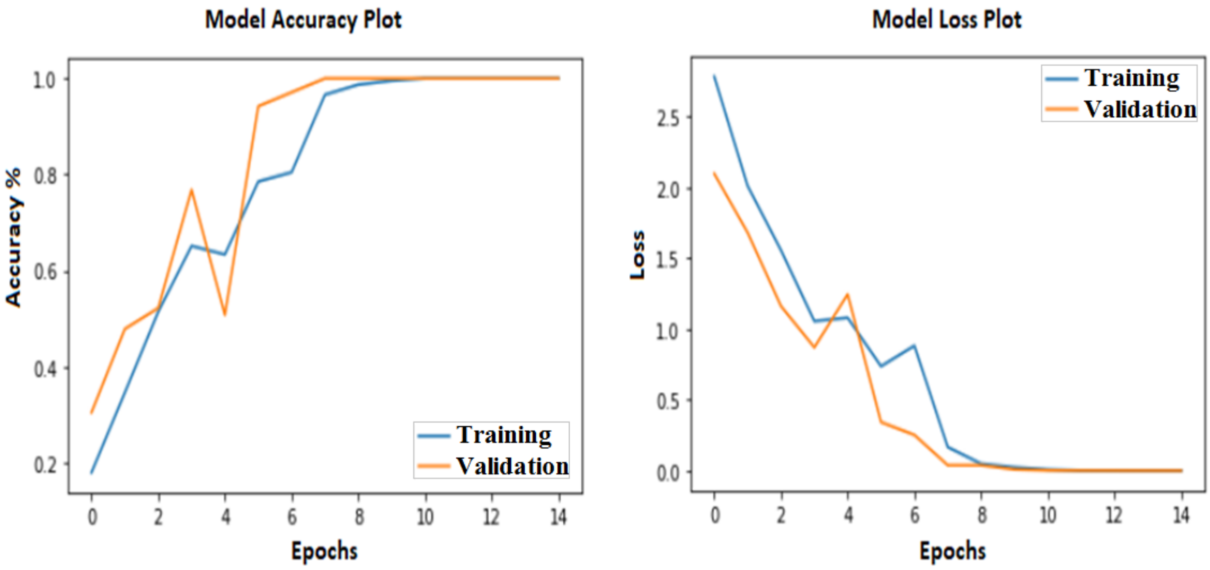

Figure 7. *Accuracy and loss graph of the trained MLP based DNN model*

The reason why the graphs in Figure 7 are cut off at epoch 14 is to avoid overfitting by making an early stopping in the analysis.

IV. DISCUSSION AND CONCLUSION

While there are applications that facilitate the daily lives of people in smart areas, applications that provide great benefits in corporate, production, and industrial areas and are also used intensively. Thanks to the many sensors and software included in wearable devices, it is possible to collect data, analyze it and communicate with other systems, making existing life easy, safe, and comfortable. It is also important that these wearable technologies can help ensure public and individual safety in smart spaces.

In this study, the dataset created by calculating the gait parameters of a total of 15 different volunteers, consisting of men and women, was used by using a wearable gait analysis sensor, which includes an accelerometer, gyroscope, and barometric pressure sensor [13]. Since the increase in the diversity of the data will positively affect the learning of the DL model, data augmentation has been made. The obtained data were classified in the MLP-based DNN model and the results were statistically analyzed. The results have proven that the MLP-based DNN model provides excellent performance in person recognition from walking data, and the method used for data augmentation also produces a successful result in walking data. With the MLP-based DNN used in this study, a 100% ACR was obtained without the need for feature selection. This shows that the proposed model is quite successful in recognizing person from gait data. In addition, as in every study, there are limitations for this study. Detailed analyzes could not be performed because the data content could not be reached completely. Also, the scarcity of data was another factor that limited the performance of the model.

In the future, it is thought that more information about people can be obtained with DNN-based software in order to further increase security in smart areas. Various studies can be carried out not only with motion sensors but also with different sensors such as pressure, temperature, EMG. Studies in which all sensors are used simultaneously and classified with DNN-based models can make a difference with detailed analysis of their results. In addition, it can be said that this study can shed light on other future studies.

Conflicts of interest: The author declares that he has no conflict of interest.

V. REFERENCES

[1] D. Rothman, *Artificial Intelligence by Example: Develop Machine Intelligence From Scratch Using Real Artificial Intelligence Use Cases*, 1st ed., Birmingham, UK: Packt Publishing, 2018, pp. 88-101.

[2] M. Molinara, A. Bria, S. De Vito, and C. Marrocco, "Artificial intelligence for distributed smart systems," *Pattern Recognition Letters*, vol. 142, pp. 48-50, 2021.

[3] C. Su, Z. Xu, J. Pathak, and F. Wang, "Deep learning in mental health outcome research: a scoping review," *Translational Psychiatry*, vol. 10, no. 1, pp. 1-26, 2020.

[4] M. A. Al-Garadi, A. Mohamed, A. K. Al-Ali, X. Du, I. Ali, and M. Guizani, "A survey of machine and deep learning methods for internet of things (IoT) security," *IEEE Communications Surveys & Tutorials*, vol. 22, no. 3, pp. 1646-1685, 2020.

[5] P. Ping, Y. Sheng, W. Qin, C. Miyajima, and K. Takeda, "Modeling driver risk perception on city roads using deep learning," *IEEE Access*, vol. 6, pp. 68850-68866, 2018.

[6] A. Makkar, and N. Kumar, "An efficient deep learning-based scheme for web spam detection in IoT environment," *Future Generation Computer Systems*, vol. 108, pp. 467-487, 2020.

[7] X. Sun, K. Su, and C. Fan, "VFL—A deep learning-based framework for classifying walking gaits into emotions," *Neurocomputing*, vol. 473, pp. 1-13, 2022.

[8] F. Duan, Y. Lv, Z. Sun, and J. Li, "Multi-Scale Learning for Multimodal Neurophysiological Signals: Gait Pattern Classification as An Example," *Neural Processing Letters*, vol. 54, no. 3, pp. 2455-2470, 2022.

[9] S. Krutaraniyom, K. Sengchuai, A. Booranawong, and J. Jaruenpunyasak, "Pilot study on gait classification using machine learning," in *International Electrical Engineering Congress (iEECON)*, Khon Kaen, Thailand, 2022, pp. 1-4.

[10] M. Lee, J.-H. Lee, and D.-H. Kim, "Gender recognition using optimal gait feature based on recursive feature elimination in normal walking," *Expert Systems with Applications*, vol. 189, pp. 1-14, 2022.

[11] K. Delac, and M. Grgic, "A survey of biometric recognition methods," in *46th International Symposium on Electronics in Marine*, Zadar, Croatia, 2004, pp. 184-193.

[12] J. Nelson, *The Handbook for School Safety and Security: Access Control, Access Badges, And Biometrics Characteristics For Schools*, 9th ed., Oxford, UK: Butterworth-Heinemann, 2014, pp. 241-253.

[13] A. Gümüşçü, "Improvement of wearable gait analysis sensor based human classification using feature selection algorithms," *Fırat Üniversitesi Mühendislik Bilimleri Dergisi,* vol. 31, pp. 463-471, 2019.

[14] A. Gümüşçü, "Wearable Sensor based Gait Recognition for Human Identification," in *International Conference on Multidisciplinary, Science, Engineering and Technology*, Dubai, United Arab Emirates, 2018, pp. 31-33.

[15] X. Li, H. Xu, and J. T. Cheung, "Gait-force model and inertial measurement unit-based measurements: A new approach for gait analysis and balance monitoring," *Journal of Exercise Science & Fitness*, vol. 14, no. 2, pp. 60-66, 2016.

[16] A. Saboor, T. Kask, A. Kuusik, M. M. Alam, Y. Moullec, I. K. Niazi, A. Zoha, and R. Ahmad, "Latest research trends in gait analysis using wearable sensors and machine learning: A systematic review," *IEEE Access*, vol. 8, pp. 167830-167864, 2020.

[17] D. Dua, and C. Graff, "UCI Machine Learning Repository [http://archive.ics.uci.edu/ml]. Irvine, CA: University of California," *School of Information and Computer Science*, vol. 25, pp. 27-27, 2019.

[18] S. E. Dreyfus, "Artificial neural networks, back propagation, and the Kelley-Bryson gradient procedure," *Journal of Guidance, Control, and Dynamics*, vol. 13, no. 5, pp. 926-928, 1990.

[19] J. L. Balcázar, R. Gavalda, and H. T. Siegelmann, "Computational power of neural networks: A characterization in terms of Kolmogorov complexity," *IEEE Transactions on Information Theory*, vol. 43, no. 4, pp. 1175-1183, 1997.

[20] C. M. Bishop, *Neural networks for pattern recognition*, 1st ed., Oxford, UK: Oxford University Press, 1995, pp. 116-191.

[21] H. Taud, and J. Mas, *Geomatic approaches for modeling land change scenarios*: *Multilayer perceptron (MLP)*, 1st ed., NewYork, USA: Springer, 2018, pp. 451-455.

[22] Y. Bengio, "Learning deep architectures for AI," *Foundations and Trends in Machine Learning*, vol. 2, no. 1, pp. 1-127, 2009.

[23] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Networks,* vol. 61, pp. 85-117, 2015.

[24] C. Szegedy, A. Toshev, and D. Erhan, "Deep neural networks for object detection," *Advances in Neural Information Processing Systems,* vol. 26, pp. 1-9, 2013.

[25] G. E. Hinton, *Neural networks Tricks of the trade: A practical guide to training restricted Boltzmann machines*, 2nd ed., NewYork, USA: Springer, 2012, pp. 599-619.

[26] Y. You, A. Buluç, and J. Demmel, "Scaling deep learning on gpu and knights landing clusters," in *International Conference for High Performance Computing, Networking, Storage and Analysis*, Dallas, USA, 2017, pp. 1-12.

[27] A. Viebke, S. Memeti, S. Pllana, and A. Abraham, "CHAOS: a parallelization scheme for training convolutional neural networks on Intel Xeon Phi," *The Journal of Supercomputing,* vol. 75, no. 1, pp. 197-227, 2019.

[28] D. P. Kingma, and J. Ba, "Adam: A method for stochastic optimization," in *3rd International Conference for Learning Representations*, San Diego, 2015, pp. 1-15.

[29] Y. Ma, L. Guo, and B. Cukic, *Advances in Machine Learning Applications in Software Engineering: A statistical framework for the prediction of fault-proneness*, 1st ed., USA: IGI Global, 2007, pp. 237-263.

[30] D. M. Powers, "Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation," *International Journal of Machine Learning Technology,* vol. 2, no. 1, pp.37-63, 2011.

[31] Ş. Yücelbaş, C. Yücelbaş, G. Tezel, S. Özşen, and Ş. Yosunkaya, "Automatic sleep staging based on SVD, VMD, HHT and morphological features of single-lead ECG signal," *Expert Systems with Applications,* vol. 102, pp. 193-206, 2018.

[32] Ş. Yücelbaş, C. Yücelbaş, G. Tezel, S. Özşen, S. Küççüktürk, and Ş. Yosunkaya, "Pre-determination of OSA degree using morphological features of the ECG signal," *Expert Systems with Applications,* vol. 81, pp. 79-87, 2017.

[33] J. Cohen, "A coefficient of agreement for nominal scales," *Educational and psychological measurement,* vol. 20, no. 1, pp. 37-46, 1960.

[34] F. Kayaalp, and M. Basarslan, "Performance analysis of filter based feature selection methods on diagnosis of breast cancer and orthopedics," in *6th International Congress on Fundamental and Applied Sciences 2019 (ICFAS2019),* Tirana, Albania, 2019, pp. 1-11.

[35] S. M. Vieira, U. Kaymak, and J. M. Sousa, "Cohen's kappa coefficient as a performance measure for feature selection," in *International conference on fuzzy systems*, Barcelona, Spain, 2010, pp. 1-8.