Identification of Neurodegenerative Diseases Based on Multivariate Gate Signal Self-Regression Model and Fusion in Intelligent Classifiers

Pouya Gholian[1], Mohammad Reza Yousefi[2*], Khoshnam Shojaei[3]
1- Department of Electrical Engineering, Najafabad Branch, Islamic Azad University, Najafabad, Iran
Email: puyagholian96@gmail.com
2- Digital Processing and Machine Vision Research Center, Najafabad Branch, Islamic Azad University, Najafabad, Iran.
Email: mr-yousefi@iaun.ac.ir (Corresponding author)
3- Digital Processing and Machine Vision Research Center, Najafabad Branch, Islamic Azad University, Najafabad, Iran.
Email: khoshnam.shojaee@gmail.com

Received: October 2021 Revised: December 2021 Accepted: February 2022

ABSTRACT:
Diagnosis of diseases with the help of new methods has received much attention. One of these diseases is amyotrophic lateral sclerosis. In this disease, neurons cause progressive and irreparable damage to the central nervous system (brain and spinal cord) and peripheral nervous system. Symptoms of upper motor neurons as well as symptoms of lower motor neurons are seen. It is possible to diagnose this disease from kinematic dynamics analysis data. Clinical methods in diagnosing this disease face significant errors. Machine learning methods are an effective way to diagnose these diseases. The proposed method of this research consists of five steps. Pre-processing, feature extraction, dimension reduction, classification and evaluation. The novelty of this article is in using classifier in the diagnosis of this disease. In classification fusion, the types of linear and non-linear classifications in a fusion method with each other will diagnose the disease with higher accuracy.

KEYWORDS: Neurodegenerative Diseases, Dynamic Gait Analysis, Classification, Principal Component Analysis, Classification Fusion.

1. INTRODUCTION

With the change in people's lifestyles, The methods of diagnosing degenerative diseases in clinical settings are more qualitative, so that the diagnosis of these diseases is limited to drawing circular geometric shapes (such as circles or ovals), breaking walnuts or performing rhythmic operations. Studies in the field of identifying patients with neurodegeneration, although growing; But they are still in the early stages of processing. It seems that no specific methods of preprocessing and noise reduction for this type of signal in the studies are in the early stages due to the unknown frequency range and noise sources [1, 2].

Obtaining Anshan motions is one of the most important researches in biomechanics. In ancient times it was used to detect diseases related to skeletal and muscular as well as rehabilitation. Despite the analysis of movement in hospitals and medical centers is very useful [2]. The use of gait analysis in hospitals and treatment centers is very limited and has not grown significantly. It should be noted that the use of motion analysis methods in these centers is very valuable [3]. On the other hand, the cost and time of data collection, processing and interpretation will be saved. In order to diagnose gait-related diseases, it is necessary to first obtain the necessary parameters to examine the movement of the teeth and also to acquire knowledge about the problems and issues of walking [4].If we pay attention to people's gait, we will see that despite the fact that the movement and gait seem automatic. But it is complicated. For this purpose, the hands are pulled back and forth and they walk with almost a rhythm. There are also a limited number of people who walk differently and are not actually able to walk like

Paper type: Research paper
DOI: https://doi.org/10.52547/mjms.11.1.37
How to cite this paper: P. Gholian, M. R. Yousefi and Kh. Shojaei, **"Identification of Neurodegenerative Diseases Based on Multivariate Gate Signal Self-Regression Model and Fusion in Intelligent Classifiers"**, *Majlesi Journal of Mechatronic Systems*, Vol. 11, No. 1, pp. 37-46, 2022.

normal people. This is called the problem of walking. In fact, among the various models of walking, only one is the correct model. In order to organize the models, although some of them are included in a special group [5].

Neurodegenerative diseases are one of the most complex and common neurological diseases that have been studied in various motor and cognitive aspects in the last decade. Early diagnosis of these complications can help in faster treatment, so it is necessary to propose a method that is firstly appropriate and optimal in terms of screening cost, and secondly, the method's segregation is different for different classes of diseases compared to the control group. In this study, we try to present an optimal technique based on the extraction of temporal, frequency, statistical and nonlinear properties and the application of linear and nonlinear classifications for dynamic walking signal. This study examines and determines the appropriate properties for linear and nonlinear classifiers, and yet tries to meet the requirements of other steps in a way that achieves the desired result [6].

Also, the analyzes performed in the field of signal processing are limited to time processes that are sensitive to noise, and the analysis of this signal in the time-frequency and nonlinear domains has been less. Also, the methods of reducing the dimension of features have been considered as a tool to increase the speed and accuracy of classifications and increase less resolution [7, 8]. Finally, linear and nonlinear classification structures in a set or a council structure have not been comparatively evaluated [9].

The dynamics of gait are regulated by the nervous system. It is believed that the feedback loop requires feedback from a physiological system at temporal and spatial scales to be able to adapt to the environment [10]. Therefore, the time series of stepping intervals can be a demonstration of this power of adaptation in different conditions [11]. In recent studies, computer-based diagnostic tools have been used to measure gait distance parameters in healthy adults as well as to describe the distinctive features of gait in patients with neurodegeneration.

In [12] the time series showed the steps taken from the gait signal as a sequence of symbols (ie converting gait distances into a set of symbols and signs). They then used a threshold dependent on the application of the symbolic entropy method to analyze the complexity of walking. In [13] used the Super Central Pattern Generator model (SCPG) to simulate the dynamics of human gait, as well as to evaluate the random and chaotic features of gait in patients with amyotrophic lateral sclerosis, Huntington and Parkinson's. In [14] presented a statistical analysis method to classify the gait of amyotrophic lateral sclerosis and healthy individuals. In his proposed method, the probability density function of the walking pace was estimated using the Parzen window method, and then the Kullback – Leibler divergence was extracted. In [15] It has been hypothesized that the number of oscillations of walking intervals in patients with amyotrophic lateral sclerosis differs from that in healthy individuals. In [16] the accuracy of classification was 89.66% for classifying healthy people and people with amyotrophic lateral sclerosis. In [17] Gait asymmetry was examined in patients with neurodegeneration using multi-rate entropy analysis of step time fluctuations. In [18] For pattern analysis and classification, support vector machine (SVM) has been used to differentiate healthy individuals from patients with Parkinson's disease, Huntington's disease and amyotrophic lateral sclerosis. The results of this study showed that only by using step intervals, this mapping can be used to classify the classes of neurodegenerative diseases.

From the conclusion of the researchers' studies, it can be concluded that the main gap in the studies is in the way of data preprocessing, lack of optimal signal processing and processing in the field of conversion and lack of combination methods at the level of features and classifiers. The novelties of this article are:

1. Optimal preprocessing of walking signal in such a way that time and frequency patterns and improvement of signal-to-noise ratio index.
2. Using principal component analysis in reducing the number of features based on principal component analysis, which has previously been based more on statistical tests.
3. Combining at the level of intelligent classifications in the form of a council machine system that has previously been a single decision maker.

The structure of this article is as follows. Section 2 will present fusion methods. In section 3, the proposed method will be examined. Section 4 will evaluate the classifications. Finally, in Section 5, the conclusion of the article is presented.

2. FUSION CLASSIFIER

In order to improve the accuracy of learning, the results of different categories can be combined and a council system can be created. In fact, using multidisciplinary results called group learning is an effective way to improve the performance of identification systems. First, the appropriate baseline categories (type and number) must be selected. Then the output of the classifiers are combined in such a way that the best result is obtained for classifying the patterns. A noteworthy point in creating base classifiers is the variation in error of one classifier over another in order to improve the overall system recognition rate. Also, the characteristics of each classifier should be selected in such a way that the recognition rate of that classifier alone is acceptable. According to the application and requirements of the problem, the type of classifiers, their number and how to combine the output is selected. The main goal is to maximize the recognition rate of the hybrid classification system. In the problem of pattern recognition, the combination of classifiers and features is of particular importance. Due to the variety of rules for combining features or output results of categories, several views can be mentioned.

A: The power to recognize patterns is limited in each category; Therefore, it is not possible to find a classification that is able to correctly identify all the patterns in all conditions. Hence, due to the inability of a classifier to meet the required recognition rate, a combination of several classifications is required for a particular application.

B: Each pattern can be represented differently depending on the different characteristics it creates in the signal. To do this, recognizing each pattern usually requires extracting different features.

A: The extraction of some properties leads to the emergence of a large property vector. The analysis of large property vectors by a classifier greatly prolongs the processing time. This is a problem in applications such as biometric systems. Combining classifiers allows feature vectors with large dimensions to be subdivided into smaller vectors and processed simultaneously with smaller, simpler classifiers. The final classification is done by combining the results of these classifications [13]. In fact, the efficiency of the pattern recognition system, especially for complex patterns, is improved by combining the results of the classifiers

2.1. Ensemble methods

A combination is the supervised learning algorithm itself, because it can be taught and then used to predict. Compositions can be shown to have more flexibility in the functions they can provide. This flexibility in theoretical science enables them to make a model too appropriate for the educational data used. In practice, however, some combination methods (especially bagging) tend to reduce the problems associated with over-fitting educational data.

Empirically, combinations tend to provide better results when there is considerable variability in the models [12]. Therefore, many combination methods seek to promote diversity among their models. The use of different types of robust learning algorithms has been shown to be more effective than the use of techniques that attempt to simplify models to promote diversity [12]. Some common methods of combinations include: optimal Bayesian classification, Bootstrap aggregating (bagging), Boosting, Adaboost, etc .; In this study, bagging and boosting of their collections have been used.

2.2. Bagging

The term stacking comes from bootstrap aggregation. As the name implies, the two most important elements of bagging are bootstrap and aggregation. We know that the combination of independent basic learners leads to a significant reduction in errors; Therefore, we want basic learners to be as independent as possible. Sampling a number of non-overlapping data subsets and then training the basic learner from each subset is a solution in this regard. But since we do not have infinite educational information, such a process leads to very small samples, which leads to low performance of basic learners.

Bagging Exactly, by giving the training data set that contains m training instance, a subset of the main data set is given to each of the classifiers. That is, each classifier observes a portion of the data set and must build its model based on the same portion of the data provided (ie, the entire database is not given to each of the classifiers). In fact, for each of the classifiers. , A subset of the original data is selected, which will be selected by replacement. That is, a sample can be selected multiple times. So some original samples appear more than once, while some original samples do not appear in the sample. This process is performed T times and T samples are obtained from M training samples. Then, one basic learner from each instance can be trained using the basic learning algorithm.

Bagging uses the most popular strategies for aggregating classifier outputs, namely voting for classification and averaging for regression. Each of the categories expresses its prediction, which is used as a vote in the final vote based on the majority vote. In the case of forecasting and using the bagging method, the final forecast value is obtained from the average forecast of each of the combined forecasting systems [13]. The bagging algorithm is as follows:

- D={(x1, y1), (x2, y2), … , (xm, ym)} is the data set
- $\mathcal{L}$ is Basic learning algorithm
- T is Number of basic learners
- **for**t=1,…,T:
- $h_t = \mathcal{L}(D, D_{bs})$($D_{bs}$) is Bootstrap distribution
- **end**
- Output: $H(x) = argmax \sum_{t=1}^{T} \mathbb{1}\,(h_t(x) = y)\;\; y \in Y$

2.3. Boosting

The term boosting is a hybrid meta-algorithm in the field of machine learning that is used to reduce imbalances as well as variance. This method is used in supervised learning and is a family of machine learning algorithms. This technique is a way to transform weak learning systems into strong ones based on the combination of the results of different classifications [19] .

In boosting, a weight is assigned to each training instance. A series of k classifiers are taught repeatedly. Once the Mi classifier is trained, the weights are updated for the next Mi + 1 classifier so that more attention is paid to data from the training data field that has not been properly categorized by the previous classifier Mi. The final boosted classifier combines the votes of the individual classifiers, in which the vote of each classifier is weighted. The weight applied to the output of each classifier is a function of the classifier accuracy [23]. The algorithm can also be developed to predict continuous values. The general procedure for the boosting method is as follows:

- Input
- Distribution sample
- $\mathcal{L}$ is Basic learning algorithm
- T is Number of basic learners
- $h_t=\mathcal{L}(D_t)$;

1-$D_1=D_0$: Start the initial distribution

2- **for** t=1,…,T:

3Teach D_t a poor learner of distribution

4- $\epsilon_t = p_x \sim D_t(h_t(x) \neq f(x))$; calculate the error of h_t

5- D_{t+1}=Adjust_Distribution(D_t,ϵ_t)

6- end

The weighting of the samples is such that at each stage, the weight of the samples that are correctly classified decreases and the weight of the samples that are not properly classified increases, so that in the next steps (by new learners) more attention is paid and more accurately classified. To turn; Therefore, the focus of new weak learners will be more on data that the system has not been able to classify correctly in previous stages [20] The efficiency of accuracy and sensitivity will increase significantly.

3. PROPOSED METHOD

In order to diagnose the complication and the disease, a machine-based method has been used in which the fusion accuracy of the classification and the diagnosis of the disease have been provided by merging the categories. The proposed method block diagram is shown in Fig. 1.

Firststep: data collection

In this study, the physiotherapy database for neurodegenerative diseases has been used. This database contains a set of data recorded by force-sensitive resistance sensors used to analyze the walking dynamics that are placed underfoot. The data is in the Physion database and includes 15 records of Parkinson's patients, 20 records of Huntington's patients, 13 records of amyotrophic lateral sclerosis and 16 records for the control group. After downloading each person's data, a zip file is provided for each person, which contains a mat file and an Excel file that contains information about the sample from which it was recorded. The dimensions of each matrix after entering MATLAB are 2 rows and 90,000 columns, two rows are related to left and right foot data and 90,000 are related to samples recorded per unit time. The sampling frequency of the device was 300 Hz, so it can be concluded that the duration of each recording of 300 seconds was equal to 5 minutes.

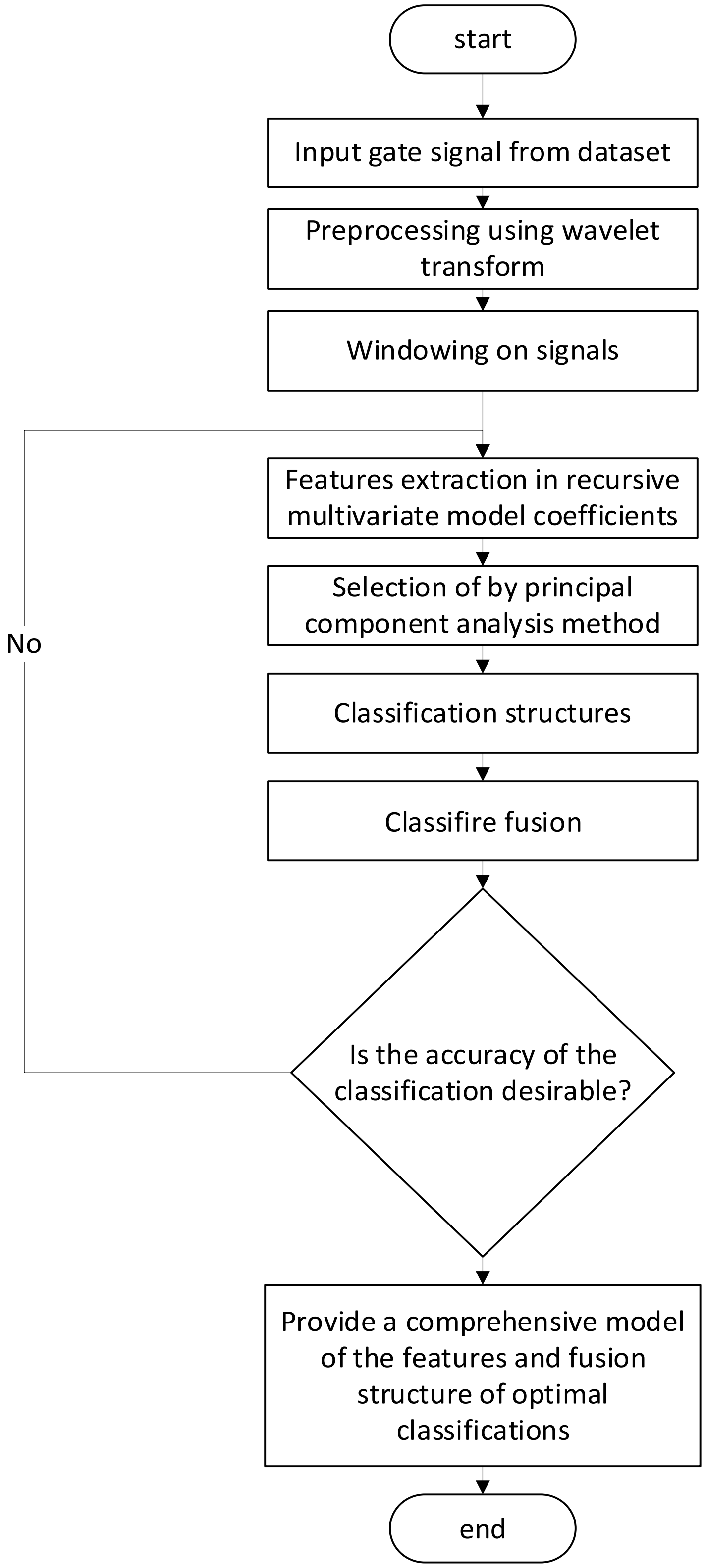

Fig. 1. General algorithm of the proposed method.

Step 2: Preprocessing

In order to reduce the noise from the data, the MATLAB software wavelet transform toolbox has been used. As mentioned before, in order to reduce noise, the default wavelet transform coefficients have been used to eliminate noise with unspecified sources. Researchers' studies have shown that the Dabigiz 4 wavelet with eight levels of decomposition retains the original information of the signal after holding the approximation coefficients. Noise is done. Of course, the set of noise removal operations has been accomplished using the ddencmp command on the editor page. Noise removal for each movement, each repetition and each electrode is performed separately and again the noise removal matrix is used to perform the processing process. Fig.s (2) and (3) summarize this process using the graphical interface. MATLAB software depicts. In order to evaluate the noise reduction process, trial and error method was used.

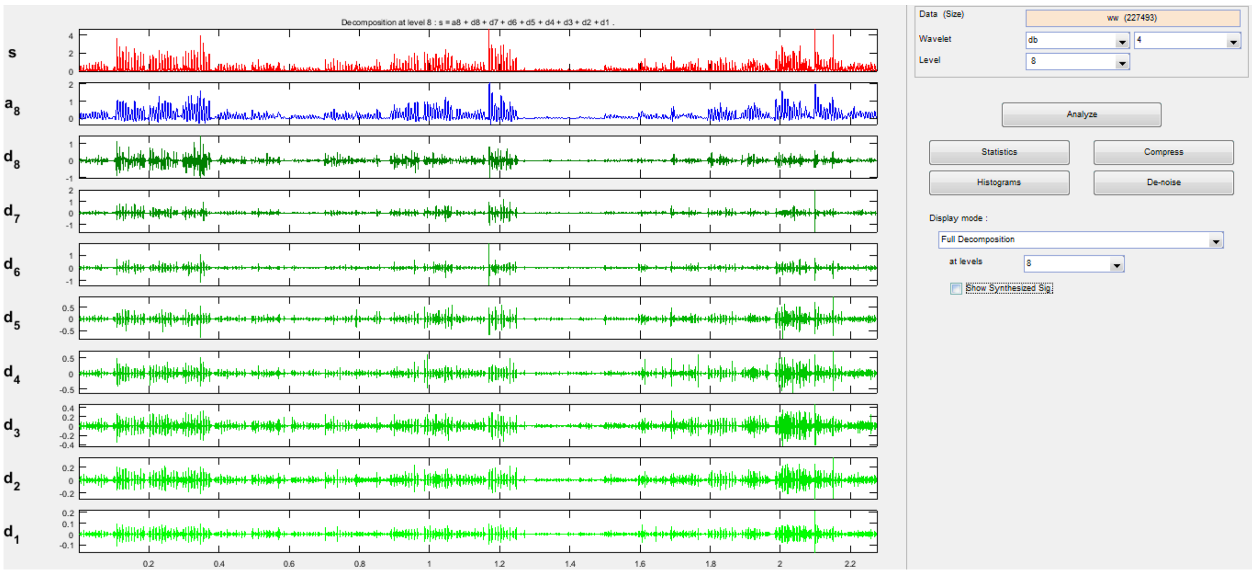

Fig. 2. Decomposition of the first signal electrode into eight levels with Dabigiz mother wavelet.

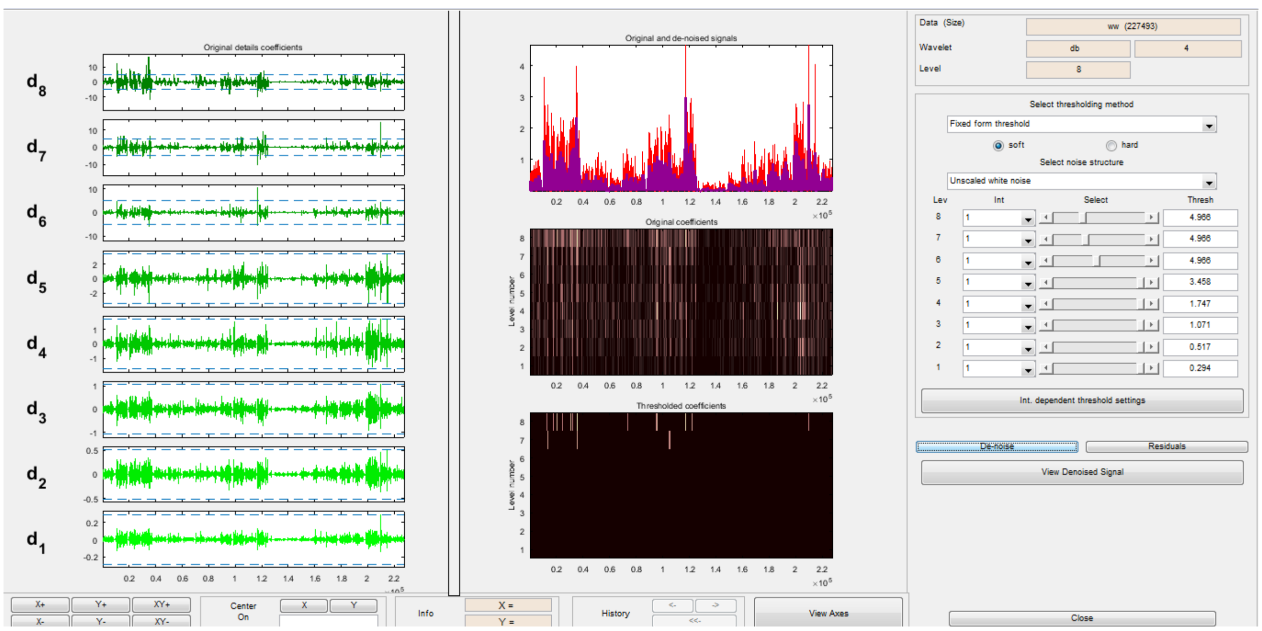

Fig. 3. Noise removal from the first signal electrode after eight levels of Dabigiz mother wavelet analysis.

Step 3: Feature Extraction

Properties are extracted from the output matrix of the processing stage. A total of 11 properties are extracted from each electrode, so if the number of columns of the property matrix expressing the number of properties is extracted, 11 properties are calculated for each electrode, which produces a total of 22 columns for the property matrix. Each property is normalized at each electrode and their box diagrams are plotted, which are reported for movements to examine the characteristic changes for each electrode and each movement in the next chapter.

Step 4: Reduce feature space

In order to reduce the property space, principal component analysis has been used. In the principal component analysis method, either the number of remaining properties is determined by the order of eigenvalues determining the volume of the property space, or it must be done to determine the percentage of property space matrix information. That is, reducing the feature space. The machine learning toolbox shown in Fig. (4) was used to perform the principal component analysis.

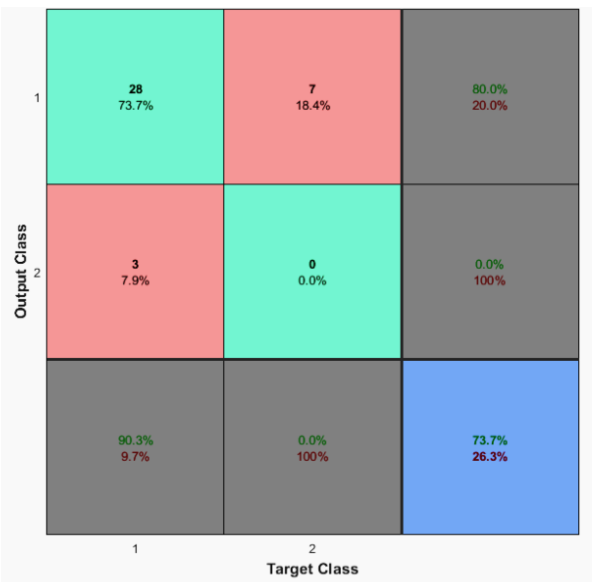

Fig. 4. An example of a clutter matrix.

Step 5: Classification and diagnosis

In order to classify the extracted features and the reduced reduction features, the four classification structures mentioned in the previous sections were used.

Linear and nonlinear kernels were used to classify the support vector machine, the training stopping criterion was used based on cross-validation 5, and the data were randomly assigned to the classifier input at each move. Percentage of training data was 75% of the total data and 25% for testing. Similar settings were applied to the nearest neighbor classifier, and the k-criterion was used to classify the nearest neighbor 10.

For perceptron multilayer neural network, a lattice with a latent layer with different number of neurons in the latent layer was used. According to the study of Man-chung et al. (2000), the number of hidden layer neurons should be considered as half the sum of inputs and outputs. Also, the number of latent layer neurons according to Freisleben (1992) study is equal to the square root of the sum of inputs and outputs and according to Gencay et al. (1999) study is equal to the logarithm of the natural basis of the number of inputs and the results were compared. 70% of data For training, 25% were used for testing and 5% for validation and stopping the training process. The training algorithm was also a gradient descentfeatures will be used to improve accuracy.

4. RESULTS

In the next part of the proposed method, the results of the combination at the level of the categories after applying the reduction in the feature space for the category of composition movements are presented. Only three other types of classifications were used in the amalgamation process. Fig. (5) shows the results obtained by combining at the classifier level for k-fold with a coefficient of 5. The results show that the combination at the level of classifiers has improved the results, and among these, the decision template method and the Dempster-Schaefer method have been the best way to combine the classifiers and have improved the classification results.

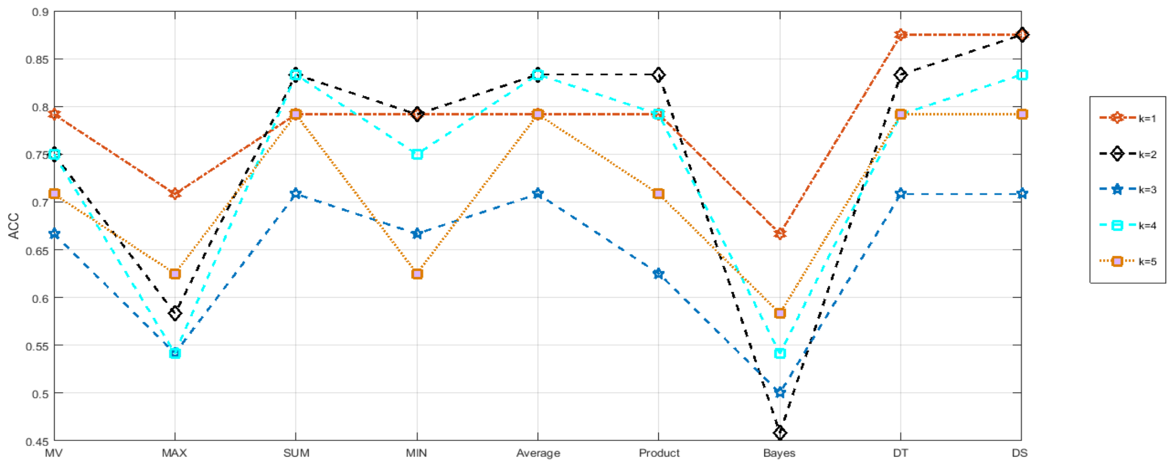

Fig. 5. Accuracy of classification after combination at the level of classifiers for 5 tests and 9 combination methods.

To compare the reproducibility of the composition results at the classifier level, Fig. (6) is presented. In this Fig., the results of classification accuracy are presented for the 9 methods of combination. This diagram shows that the maximum repeatability is obtained for the majority voting method and the minimum repeatability is obtained for the Bayesian method. Because the diagram of the implementation of the algorithm three times for Bayesian method has caused the separation of accuracy diagrams, so it does not have good reproducibility. This is while the correctness diagrams in the majority voting method are matched in all three times of the combined algorithm and have better reproducibility.

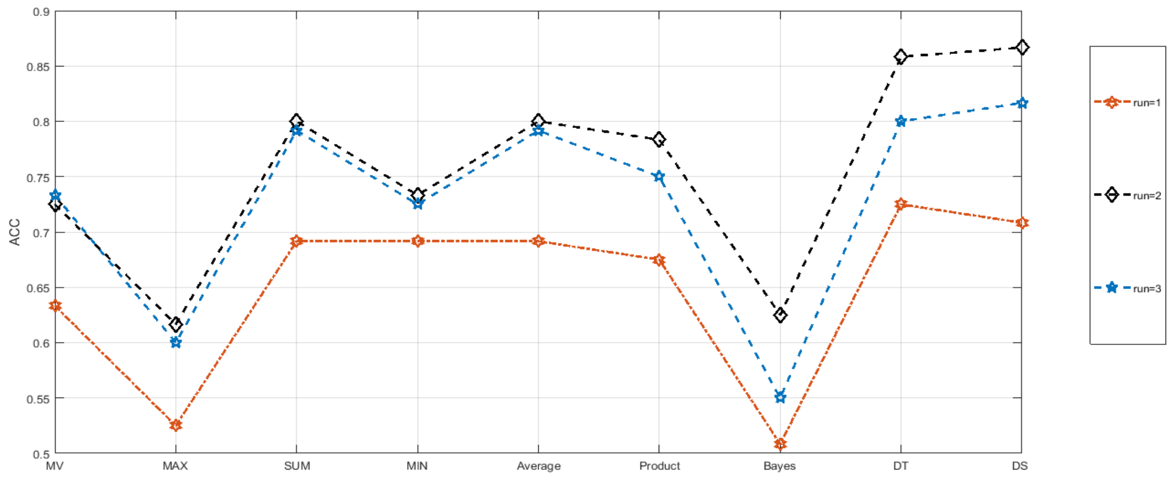

Fig. 6. Accuracy of classification after combination at the level of classifiers for 3 times performed and 9 methods of combination.

Also compare the results of optimal classifications, ie support vector machine (with Gaussian nonlinear kernel), nearest Gaussian micro-scale neighbor and multilayer perceptron neural networks with one hidden layer and number of neurons in the hidden layer equal to the square of the total inputs and output for k-fold with coefficient 5 is drawn in Fig.7

.

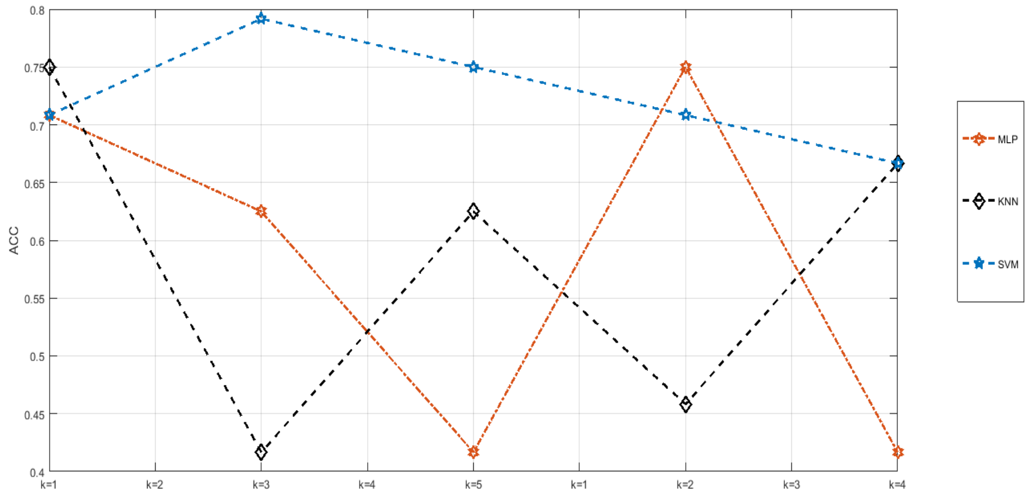

Fig. 7. Reproducibility of classification accuracy for three types of classification.

5. CONCLUSION

The simulation results showed that the use of wavelet transform bank filter has the necessary efficiency in improving the signal and reducing noise. The classification results show that it can be concluded that the backup vector machine with nonlinear kernel and the nearest microscale nearest neighbor classifier in the classification of features with high efficiency. This can be explained by the fact that because the properties of the first moves are not linearly separable, the linear classifiers do not have the necessary efficiency in classification.

REFERENCES

[1] A. Raftarai, R. R. Mahounaki, M. Harouni, M. Karimi, and S. K. Olghoran, **"Predictive Models of Hospital Readmission Rate Using the Improved AdaBoost in COVID-19,"** *in Intelligent Computing Applications for COVID-19: CRC Press*, 2021, pp. 67-86.

[2] M. Karimi, M. Harouni, A. Nasr, and N. Tavakoli, **"Automatic Lung Infection Segmentation of Covid-19 in CT Scan Images,"** *in Intelligent Computing Applications for COVID-19: CRC Press*, 2021, pp. 235-253.

[3] .Karimi, M. Harouni, and S. Rafieipour, **"AutomatedMedical Image Analysis in Digital Mammography,"** *in Artificial Intelligence and Internet of Things: CRC Press*, 2021, pp. 85-116.

[4] G.-M. Jeong, P. H. Truong, and S.-I. Choi, **"Classificationof three types of walking activities regarding stairs using plantar pressure sensors,"** *IEEE Sensors Journal*, Vol. 17, No. 9, pp. 2638-2639, 2017.

[5] M. Batool, A. Jalal, and K. Kim, **"Sensors Technologies for Human Activity Analysis Based on SVM Optimized by PSO Algorithm,"** *in 2019 International Conference on Applied and Engineering Mathematics (ICAEM)*, 2019, pp. 145-150: IEEE.

[6] M. Harouni, M. Karimi, and S. Rafieipour, **"Precise Segmentation Techniques in Various Medical Images,"** *Artificial Intelligence and Internet of Things: Applications in Smart Healthcare*, p. 117, 2021.

[7] A. R. Khan et al., **"Authentication through gender classification from iris images using support vector machine,"** *Microscopy research and technique*, 2021.

[8] S. Aryanmehr, M. Karimi, and F. Z. Boroujeni, **"CVBL IRIS Gender Classification Database Image Processing and Biometric Research, Computer Vision and Biometric Laboratory (CVBL),"** *in 2018 IEEE 3rd International Conference on Image, Vision and Computing (ICIVC)*, 2018, pp. 433-438: IEEE.

[9] L. Sun, Y.-X. Yuan, Q. Zhang, and Y.-C. Wu, **"Human Gait Classification Using Micro-Motion and Ensemble Learning,"** *in IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium*, 2018, pp. 6971-6974: IEEE.

[10] J. M. Hausdorff, **"Gait dynamics, fractals and falls: finding meaning in the stride-to-stride fluctuations of human walking,"** *Human movement science*, Vol. 26, No. 4, pp. 555-589, 2007.

[11] J. Hollman, M. K. Watkins, A. C. Imhoff, C. E. Braun, K. A. Akervik, and D. K. Ness, **"Complexity and Fractal Dynamics of Gait in Treadmill Ambulation: Implications for Rehabilitation Providers,"** *Archives of Physical Medicine and Rehabilitation*, Vol. 97, No. 10, p. e82, 2016.

[12] Y. Wu and S. Krishnan, **"Computer-aided analysis of gait rhythm fluctuations in amyotrophic lateral sclerosis,"** *Medical & biological engineering & computing*, Vol. 47, No. 11, pp. 1165-1171, 2009.

[13] M. R. Daliri, **"Automatic diagnosis of neuro-degenerative diseases using gait dynamics,"** *Measurement*, Vol. 45, No. 7, pp. 1729-1734, 2012.

[14] D. Blokh and I. Stambler, **"The application of information theory for the research of aging and aging-related diseases,"** *Progress in neurobiology*, Vol. 157, pp. 158-173, 2017.

[15] N. Scafetta, R. E. Moon, and B. J. West, **"Fractal response of physiological signals to stress conditions, environmental changes, and neurodegenerative diseases,"** *Complexity*, Vol. 12, No. 5, pp. 12-17, 2007.

[16] A. F. Costa, G. Humpire-Mamani, and A. J. M. Traina, **"An efficient algorithm for fractal analysis of textures,"** *in 2012 25th SIBGRAPI Conference on Graphics, Patterns and Images*, 2012, pp. 39-46: IEEE.

[17] O. C. Yurdakul, M. Subathra, and S. T. George, **"Detection of Parkinson's Disease from gait using Neighborhood Representation Local Binary Patterns,"** *Biomedical Signal Processing and Control*, Vol. 62, p. 102070, 2020.

[18] A. Alharbi, **"A genetic-ELM neural network computational method for diagnosis of the Parkinson disease gait dataset,"** *International Journal of Computer Mathematics*, Vol. 97, No. 5, pp. 1087-1099, 2020.

[19] M. Woźniak, M. Grana, and E. Corchado, **"A survey of multiple classifier systems as hybrid systems,"** *Information Fusion*, Vol. 16, pp. 3-17, 2014.

[20] R. E. Singh, K. Iqbal, G. White, and J. K. Holtz, **"A review of EMG techniques for detection of gait disorders,"** *Artificial Intelligence-Applications in Medicine and Biology*, 2019.