



Classification of Knee Abnormality Using sEMG Signals with Boosting Ensemble Approaches

Ayşenur Kapisız Altıntaş¹ , Derya Yılmaz^{*2} 

¹ Department of Electrical and Electronics Engineering, Institute of Science, Başkent University, Ankara, Turkey

² Department of Electrical and Electronics Engineering, Faculty of Engineering, Gazi University, Ankara, Turkey
(aysenur.kapisiz@tubitak.gov.tr, deryayilmaz@gazi.edu.tr)

Received: Sep.3, 2021

Accepted: Sep.16, 2021

Published: Oct.20, 2021

Abstract— Knee problems, although increasing in the elderly, are one of the most important orthopedic problems that occur at any age and reduce the person's standard of living by making it difficult to move. In recent years, increasing in the use of surface Electromyography (sEMG) signals from muscles has highlighted the use of these signals in the detection of movement and movement disorders. In this study, sEMG signals, from patients with different knee abnormalities and healthy individuals, the muscles responsible for the bending (flexion) and stretching/extension (extension) movements of the knee (rectus femoris (RF), biceps femoris (FB), semitendinosus (ST), vastus medialis (VM)), recorded during gait, sitting, and standing were evaluated with some statistical-based features. Unlike the literature, the classification processes were also performed for each muscle and each movement, and therefore the effect of the muscles on the classification performance was examined.

Keywords: knee disorder, knee muscle, surface EMG, decision trees, ensemble classification.

1. Introduction

The knee is one of the most frequently used and most frequently injured arthrosis in the human body, consisting of the thigh (femur), shin (tibia), kneecap (patella) and various ligaments, tendons and muscles that support them. Impacts, compelling movements, repetitive overloads and old age can cause damage to the anatomical structures in this arthrosis. Knee problems usually occur in the form of pain, difficulty in movement and walking, but affect daily life negatively and reduce the quality of life of the person. Since the knee arthrosis can affect other joints such as ankle and hip, it is important to start appropriate treatment, physiotherapy and rehabilitation programs with an easy and early diagnosis of knee problems. In recent years, EMG signals from muscles outside the arthrosis that provide knee movement have begun to be evaluated as an alternative to long and expensive diagnostic methods such as magnetic resonance imaging (MRI) in the detection of knee problems. EMG stands out as an examination method used to diagnose diseases affecting peripheral nerves, to confirm the diagnosis, to identify dysfunctions or structural damages in peripheral nerves, to monitor the disease process and to evaluate the effect of the treatment applied.

Recently, analyzes of myoelectric signals taken with a non-invasive approach by using surface electrodes, called surface EMG (sEMG), are preferred to evaluate the health of nerve cells that control muscles and muscle activities. sEMG signals which are used to understand muscle behavior in healthy and pathological conditions have a non-stationary structure that includes noises such as motion artifacts and electrode noise. There are some studies in the literature that attempt to classify the pathological knee abnormalities using lower extremity sEMG data, various signal processing and classification methods [1-5]. Some of the studies on the data set used in this study can be summarized as follows: Vijayvargiya et al. divided the lower extremity sEMG signals into 256ms segments with 25% overlap after denoising and extracted 11 features in the time domain [1]. In their study, five different machine learning classifiers

(k-nearest neighbor, support vector machine, decision tree, random forest and extra tree) were used for classification with ten cross validations, and extra tree classifier gave the highest accuracy (91%) in detecting knee abnormality. In another study using a deep neural network structure and time domain features combining 3D convolution layers and long short-term memory layer (LSTM), data augmentation methods and a transfer learning algorithm are proposed to improve performance [2]. With this proposed method, an accuracy of 94.12% was achieved in the sEMG-based recognition of lower extremity abnormality. Gautam et al. presented a transfer-learning based Long Term Recurrent Convolution Network (LRCN) called 'MyoNet' for classification of lower extremity movements with estimation of knee joint angle [3]. In this study, mean motion classification accuracies of 98.1% and 92.4% were obtained for healthy volunteers and the subjects with knee pathology, respectively. Janidarmian et al. proposed an automated diagnostic system with a bagged decision tree classifier and achieved an accuracy of 97.17% in distinguishing healthy ones from the patients with knee abnormalities using a set of time domain features extracted from sEMG data [4]. Uzun et al. evaluated the sEMG data for the detection of the knee problem with the statistical properties by using artificial neural network (ANN) and logistic regression (LR) models and discrete wavelet transform [5]. They showed that the classification success of the sitting signals with ANN was higher than the other conditions (gait and standing). From the results, it is seen that 100% classification performance is achieved by using only the sitting data.

In this study, determination of knee problem was studied by using few statistical features on lower extremity sEMG signals obtained from 11 healthy subjects with 11 knee problems, 3 different movement states (gait, sitting and standing) and 4 different muscles. In the study, unlike the literature, classification performances were evaluated separately for each muscle with each movement and also for each movement with all muscles together. The classification was made with ensemble approaches of boosted and RUSboosted (random under sampling) trees which have not been tried before for these signals. In addition, the number of features used in this study is 9 and Higuchi dimension feature have never been used for these signals in the literature. The results obtained showed that the sEMG signal from only one muscle (RF) and movement (standing) is sufficient to detect the knee problem with about 92% accuracy, and the highest classification accuracy of 98.8% was achieved by using all muscle with the movement type of sitting.

2. Materials and Methods

2.1. Data

The dataset from the UCI database includes 11 healthy individuals and 11 individuals with knee anomalies from 22 male subjects [7]. The dataset includes sEMG data from the vastus medialis (VM), semitendinosus (ST), biceps femoris (FB), and rectus femoris (RF) muscles. Signals were recorded from the relevant muscles while the subjects were in gait, standing and sitting positions. The sampling frequency of the signal is 1000 Hz. A four-level denoising was applied to the raw signal using dB8 wavelet and soft thresholding, thus noise removal was performed without disturbing the signal dynamics [1]. Figure 1 shows sEMG signals recorded from four different muscles in all motion types (gait, standing, sitting) from a healthy subject (N: normal) and a subject with knee problems (A: Abnormal).

2.2. Feature Extraction

The 20.000 samples of the sEMG signals were divided into 300ms segments using adjacent windowing technique and Hamming windowing was performed and then the features given in the Table 1 were calculated. In the mathematical expression, x_i and N represent to sEMG signal sample and the number of samples in the segment, respectively (Table 1).

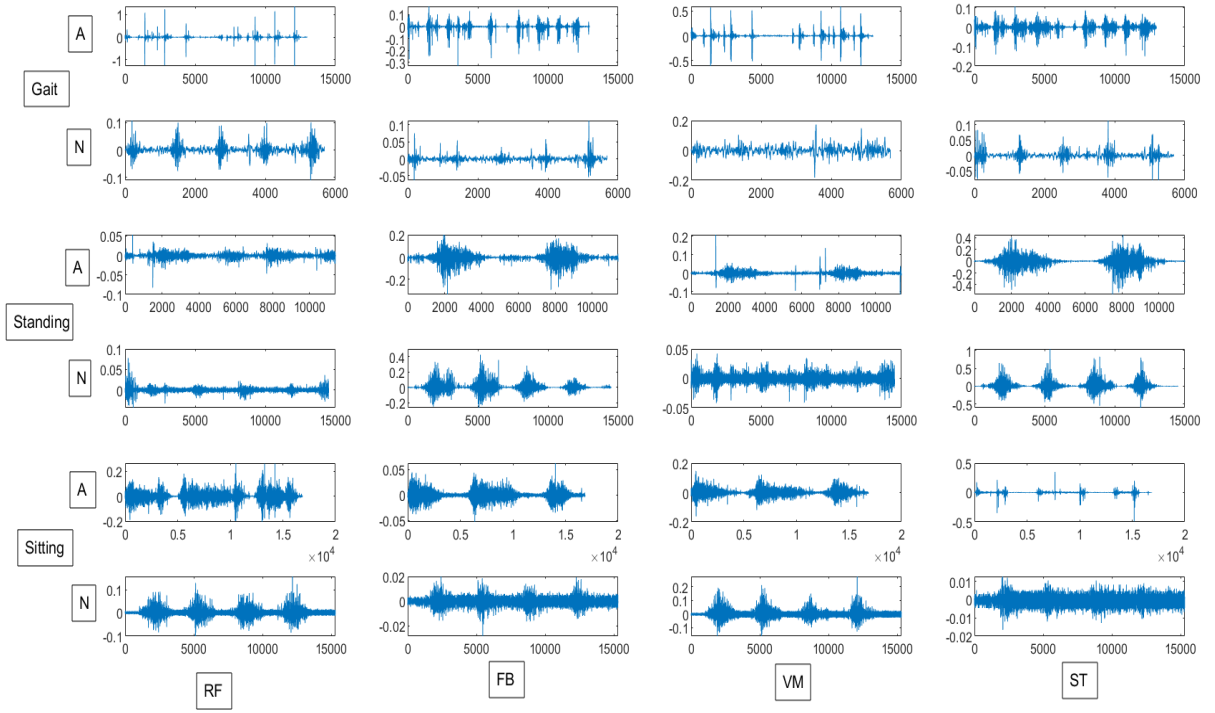


Figure 1. Signal samples from four different muscles in three different conditions from normal (N) (subject no: 1) and abnormal (A) (subject no: 1) subjects.

Table 1. Calculated features

No	Features and Mathematical Expression
1	Mean (μ) $\mu = \frac{1}{N} \sum_{i=1}^N x_i$
2	Standard Deviation (σ) $\sigma = \sqrt{\frac{1}{N} \sum_{i=1}^N (x_i - \mu)^2}$
3	Wilson Amplitude (WA) $WA = \sum_{i=1}^N u(x_{i+1} - x_i - T), T = 0.01$
4	Waveform Length (WL) $MA = \sum_{i=1}^{N-1} x_{i+1} - x_i $

5	<p>Slope Sign Change (SSC)</p> $SSC = \sum_{i=2}^{N-1} [f[(x_i - x_{i-1}) \times (x_i - x_{i+1})]]$ $f(x) = \begin{cases} 1, & \text{if } x \geq 0.01 \\ 0, & \text{otherwise} \end{cases}$
6	<p>Difference Absolute Standard Deviation Value (ASD)</p> $ASD = \sqrt{\frac{\sum_{i=1}^N (x_{i+1} - x_i)^2}{N-1}}$
7	<p>Maximum Fractal Length (MFL)</p> $MFL = \log_{10} \left(\sqrt{\sum_{i=1}^{N-1} (x_{i+1} - x_i)^2} \right)$
8	<p>Myopulse percentage rate (MPR)</p> $MPR = \frac{1}{N} \sum_{i=1}^N [f(x_i)], f(x) = \begin{cases} 1, & \text{if } x \geq 0.016 \\ 0, & \text{otherwise} \end{cases}$
9	<p>Higuchi Dimension (D_H)</p> <p>D_H is used as an irregularity measure and is calculated from signals directly without states pace constructions [8]. $L(d)$, of the signal are calculated for different dimension values (m), and in the double logarithmic axes, plotting the $\ln L(d)$ versus $\ln d$, and estimate the higuchi dimension value according to $L(d) \propto d^{D_H}$ assumption.</p> $L(d) = \frac{1}{d} \sum_{m=1}^d L_m(d)$

3. Classification Method

3.1. Ensemble Approaches

The ensemble tree classifier performs ensemble learning, that is, classification is done with multiple and different classifiers to improve performance. The results from these classifiers are average, voting, etc. evaluated in a variety of ways. That is, the decision is made on the basis of multiple classifiers instead of a single classifier model [10]. The most used algorithms in this approach are; it can be Naïve Bayes method, decision trees, neural networks etc. Bagging and Boosting methods are two of the most widely used techniques in machine learning. In the bagging method (bootstrap aggregating), the model is retrained by generating new data from the existing training set. The aim is to increase the classification accuracy by strengthening the model with the new datasets derived. An ensemble is created by applying estimators to the bootstrap samples obtained from the original data set. In this method, the bootstrapping application is used to randomly select and create subsamples. After creating multiple subsets from the original dataset, a base model (weak model) is created in each of these subsets. Models run in parallel and are independent of each other. Estimates are determined by combining estimates from all models. For regression trees, the average is taken when combining estimates, while in classification trees the results are determined by voting. Bagging makes it more convenient by using variables that have a low amount of bias but high variance. This method gives more effective results than single trees [11].

The basic idea in the boosting method is to make inferences from the collection of trees obtained as a result of giving different weights to the data set. Initially, all observations are equally weighted, and weighting is applied as the tree community begins to grow. A subset is created from the original dataset

and the weights are equal. A base model is created and used to make predictions for the entire dataset. Errors are calculated and the weight of misclassified observations is increased. Then another model is built and predictions are made. With this model, the errors in the previous model are tried to be corrected. Similarly, multiple models are created, each correcting the errors of the previous model. The final model (strong learner) is the weighted average of all models [12]. Some algorithms used by these methods are adaptive boosting, gradient boosting, extreme gradient boosting, random under sampling.

Thanks or Acknowledge

I would like to thank my Thesis advisor Derya Yılmaz and also my family and my husband who did not spare their support.

4. Conclusion

The ensemble trees methods of Boosted and RUSboosted trees were used in the classification. The results show that the knee problem can be identified by using single muscle sEMG (RF) and single movement, with a performance about 92% for the movement of standing. The highest accuracy rate is obtained as 98.8% with Boosted Trees classifier for sitting by using all muscles sEMG signals.

References

- [1] A. Vijayvargiya, N. Dey, R. Kumar, M. R. S. Tavares, “Comparative analysis of machine learning techniques for the classification of knee abnormality”, IEEE 5th International Conference on Computing Communication and Automation (ICCCA), Galgotias University, Greater Noida, UP, India. Oct 30-31, 2020
- [2] J. C. Huang, S. J. Shanq-JangRuan, W. C. Hsu, Y. T. Liu, C.H. Hsu, “3D-CLDNN: An effective architecture on deep neural network for sEMG-based lower limb abnormal recognition”, IEEE 8th Global Conference on Consumer Electronics (GCCE), 906-907, 2019.
- [3] A. Gautam, M. Panwar, D. Biswasand A. Acharyya, "MyoNet: A transfer-learning-based LRCN for lower limb movement recognition and knee joint angle prediction for remote monitoring of rehabilitation progress froms EMG", IEEE Journal of Translational Engineering in Healthand Medicine, 8, 1-10, 2020.
- [4] M. Janidarmian, K. Radecka, Z. Zilic, “Automated diagnosis of knee pathology using sensory data,” in Proc. 4th Int. Conf. Wireless Mobile Commun. Healthcare (Mobihealth), Nov. 2014, pp. 95–98, 2014.
- [5] R. Uzun, O. ErKaymaz, İ. Şenyay Yapıcı, “Comparison of artificial neural network and regression models to diagnose of knee disorder in different postures using surface Electromyography”, Gazi University Journal of Science, 31(1), 100-110, 2018.
- [6] O. Sanchez, J. Sotelo, M. Gonzales, G. Hernandez, “Emg dataset in lower limb data set”, UCI Machine Learning Repository, 2014, 2014-02.
- [7] A. M. Fraser, H. L. Swinney, “Independent coordinates for strange attractors from mutual information”, Phys. Rev. A, 33, 1134-1140, 1986.
- [8] T. Higuchi, “Approach to an irregular time-series on the basis of the fractal theory”, Physica D, 31, 277–283, 1988.
- [9] R. Agrawal, T. Imielinskiand A. Swami, “Database Mining: A Performance Perspective”, IEEE Transactions on Knowledge and Data Engineering, 5(6), 914-925, 1993.
- [10] Z.-H. Zhou, “Ensemble methods: foundations and algorithms”, New York: CRC Press, 2012.
- [11] Breiman, L., Bagging predictors. Machine Learning, 24 (2) (1996) 123-140.
- [12] Efron, B., Tibshirani, R., An Introduction to the Bootstrap. Chapman and Hall. London. (1993)