Sarcopenia Evaluation by using surface EMG-Based Platform with Supervised Classification

Andrea Manni[1], Gabriele Rescio[1], Andrea Caroppo [1], Pietro Siciliano[1] and Alessandro Leone[1]

[1] *National Research Council of Italy, Institute for Microelectronics and Microsystems, Via per Monteroni c/o Campus Universitario Palazzina A3, Lecce, Italy*

Abstract

Loss of strength and muscle mass, known as sarcopenia, has a high impact on the health status, quality of life and independence of the older population. The use of smart wearable devices is becoming increasingly relevant for the diagnosis and prevention of the disease. For this purpose, surface electromyography is becoming more popular thanks to its minimally invasive characteristics. A hardware/software platform was designed and implemented, based on the processing of the electromyographic signal derived from the Gastrocnemius Lateralis and Tibialis Anterior muscles. These signals are used to analyze the strength of the muscles in order to identify three different confidence levels of sarcopenia. Subsequently, the effectiveness of three state of the art supervised classifiers in the evaluation of sarcopenia was compared. To validate the proposed approach, a series of experiments were performed to verify the effectiveness and its operation in real-time. A total of 32 patients were recruited from Casa Sollievo della Sofferenza Hospital in San Giovanni Rotondo (Foggia, Italy). All patients were considered at risk or suffering from sarcopenia. The obtained results demonstrated the ability of the proposed platform to classify the three confidence levels of sarcopenia, with the Support Vector Machine classifier outperforming the other classifiers in terms of accuracy.

Keywords

Sarcopenia, Surface EMG, Machine Learning.

1. Introduction

Ageing is a process involved in numerous physical changes, including muscle mass. In particular, muscle shows a progressive reduction in size and number of muscle fibers (up to 30%) beginning from 25 to 80 years of human life [1]. This loss of muscle mass and strength as a consequence can cause "sarcopenia". This term is derived from the Greek words, "sarx" meaning flesh and "penia" meaning loss.

Sarcopenia is not easily diagnosed and, moreover, its treatment is still laborious because it is not easily estimated the time trend of its fundamental components, i.e. muscle mass, physical performance (such as walking speed) and muscle strength. Currently, muscle mass and strength are assessed by various gold standard techniques such as magnetic resonance imaging, computed tomography or dual energy X-ray absorptiometry [2]. These clinical tests seem to be rarely used in this environment due to the high equipment costs, absence of portability and the need for highly qualified medical staff. To bypass this limitation, smart sensors are becoming increasingly popular in recent years.

Electromyography (EMG) is a robust method to monitor muscle fatigue and estimate the efficiency of muscles, performed using their electrical potentials. There are two types of electromyography

Italian Workshop on Artificial Intelligence for an Ageing Society (AIxAS 2022)
EMAIL: andrea.manni@le.imm.cnr.it (A. 1); gabriele.rescio@cnr.it (A. 2); andrea.caroppo@cnr.it (A. 3); pietro.siciliano@le.imm.cnr.it (A. 4); alessandro.leone@cnr.it (A. 5)
ORCID: 0000-0001-5716-5824 (A. 1); 0000-0003-3374-2433 (A. 2); 0000-0003-0318-8347 (A. 3); 0000-0002-1312-4593 (A. 4); 0000-0002-8970-3313 (A. 5)

analysis: the superficial measurement method (sEMG) and the intramuscular technique. The first one is less invasive since it uses non-invasive surface electrodes on the skin [3].

Several papers have demonstrated the use of EMG signals in a medical context [4-5]. In particular, in [4] the potential clinical value of sEMG-based techniques in neurorehabilitation was addressed. In order to detect fatigue in multiple sclerosis patients, a wireless medical measurement system was developed in [5], including EMG, motion detection and muscle strength.

To detect muscle fatigue in real time during exercise, an EMG patch to be worn on the lower leg (gastrocnemius muscle) was designed and developed in [6]. An application was also designed and developed to display muscle fatigue levels and running information for the end user.

In this context, good performance was achieved using supervised Machine Learning (ML) schemes. Specifically, in [7] the atrophy disorder was analyzed using recorded sEMG signals from the biceps and evaluating three different classifiers (Linear Discriminant Analysis, Quadratic Discriminant Analysis and Support Vector Machine-SVM) to separate the samples into two classes atrophic and normal using different extracted features. The results showed that Quadratic Discriminant Analysis performed best to detect this specific disorder. In [8], the feasibility of the SVM classifier to identify upper limb intention from sEMG signals was investigated, developing a new human-machine interface for self-rehabilitation training of stroke patients.

sEMG has been widely used for the analysis of specific diseases, but very few papers in the literature have focused on the application of sEMG for the assessment of sarcopenia or its monitoring over time. Specifically, in [9], the muscle strength of seventy-one hip fracture patients was compared according to the presence of sarcopenia after surgery, and the measured values of the sEMG and dynamometer were correlated in the post-operative measurement of muscle strength. It was shown that the dynamometer and sEMG values were highly correlated, although no statistically difference in muscle strength with or without sarcopenia was detected. In contrast, [10] examined the possibility of observing age-specific effects in the sEMG representation of healthy individuals engaged in cyclic back extension exercises, in order to identify new biomarkers to be used for the screening of very early forms of sarcopenia.

With the aim of assessing sarcopenia, the focus of this research is the design and implementation of a platform combining a smart sEMG technology and a software module as a decision support system (DSS) for healthcare staff. This system is able to provide useful information to evaluate the user's muscle condition during physical performance assessment in a non-intrusive manner. Furthermore, an easy-to-use and low-invasive system can also be used at home allowing more frequent monitoring. ML techniques were used in this study to evaluate the user’s sarcopenia level. Specifically, since there is no predefined and validated model guaranteeing good performance with any type of test data, the efficiency of three ML classifiers was compared. In addition, the integrated platform was included in an initial validation in a hospital ward to demonstrate its effectiveness.

The remainder of this paper is organized as follows. Section 2 explains our proposed algorithmic pipeline and provides an overview of the methodology by detailing the implemented algorithmic step. The results obtained are reported in Section 3. Finally, Section 4 shows our conclusions and discussions on some ideas for future work.

2. Method

The proposed platform is composed of two main components: a hardware system for collecting EMG signals and a software tool for processing raw EMG data and extracting sarcopenia-related features. A commercial product was employed in the platform. The employed probes are part of the FREEEMG1000 system, produced by the BTS Bioengineering. The system is entirely based on wireless technology, and it can use up to ten lightweight, minimally invasive wireless EMG probes (dimensions are 41.5 × 24.8 × 14 mm and the weight is about 13 g). The probes are clipped to the pre-gelled Silver/Silver Chloride (Ag/AgCl) electrodes, providing for a rapid, simple and stable mounting for the user’s movements at the highest level of usability. It is possible to stream and record raw signal up to 6 h thanks to the rechargeable batteries. Probes were placed on two muscles (Gastrocnemius Lateralis and Tibialis Anterior muscles, as shown in Figure 1: Positioning of the electrodes on Gastrocnemius Lateralis and Tibialis Anterior muscles and used platform.) of the lower limbs that were involved

throughout the execution of the activities to obtain raw EMG data. One probe was placed on each muscle on both legs; thus, a total of 4 probes and channels were considered. The probes are placed along the approximated direction of muscle fibres, with the inter-electrode distance of about 20mm to obtain the maximal surface EMG amplitude.

Figure 1: Positioning of the electrodes on Gastrocnemius Lateralis and Tibialis Anterior muscles and used platform.

The designed algorithmic pipeline consists of four main blocks. In the first block, signal pre-processing techniques are included to format the data for the next steps. The second block implements the feature extraction procedures, and is followed by a logic block designed to select the most effective set of EMG features for the assessment of the considered pathology. At last, a module for the classification of the confidence level of sarcopenia has been implemented. An overview of the pipeline with a block diagram representation is shown in Figure 2.

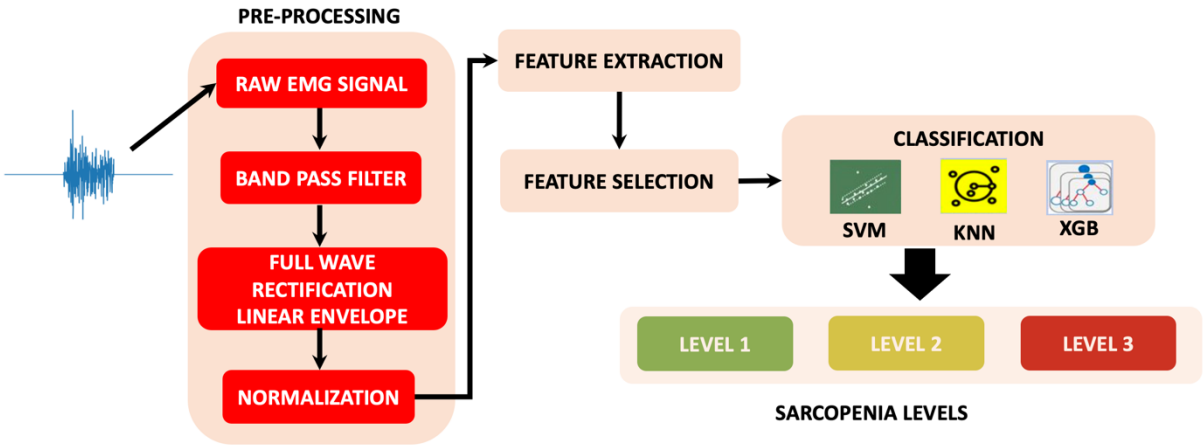

Figure 2: Overview of the proposed algorithmic pipeline designed and implemented for the distinction of different confidence levels of sarcopenia.

2.1. Data acquisition

In order to evaluate the performance of the proposed system, a total of 32 patients (19 males, with an average age of 63.95 ± 5.54 years and 13 females, with an average age of 65.62 ± 7.30 years) were recruited from the Casa Sollievo della Sofferenza Hospital in San Giovanni Rotondo (Foggia, Italy). All patients were considered at risk or suffering from sarcopenia. Due to the COVID-19 emergency, a larger number of patients could not be recruited. Tests were performed considering: (a) the SARC-F

questionnaire, (b) muscle strength analysis by means of the hand grip strength test and (c) functional performance evaluation by means of sit-to-stand and gait speed tests.

The SARC-F can be considered a powerful tool for screening sarcopenia [11]. The questionnaire examines typical symptoms of sarcopenia, such as weakness, difficulty getting up from a chair or climbing stairs, need for assistance when walking, and falls. Each of the parameters is assigned a score between 0 and 2 for minimum and maximum values. SARC-F values ≥ 4 are associated with limitation of physical activities and risk of sarcopenia.

The hand grip strength test was performed using a hand-held dynamometer with 2 tests for each hand and alternating sides during the test, considering the maximum values measured during all tests. The cut-offs suggested by the reports of the European Working Group on Sarcopenia in Older People (EWGSOP2) [2] were considered: 27 kg for men and 16 kg for women.

Sit-to-stand and gait speed tests were performed as illustrated in Figure 3 and detailed below. Without using their arms, the participants stood up as fast as possible from a sitting to a standing position; their arms were either extended to their sides or held along their chest. They then walked for 5 m and their average speed was monitored. Each participant wore sEMG sensors as detailed in the previous section. The first phase was used to acquire electromyographic data in order to assess lower body strength, while the second phase analyzed the users' physical performance. A gait speed lower than 0.8 m/s was defined as slow according to [2].

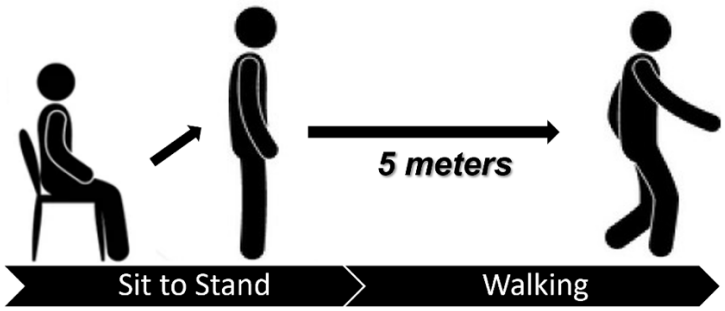

Figure 3: Sit-to-stand and gait speed tests.

In order to reduce inter-individual variability of EMG signals between different users, maximum voluntary contraction (MVC) values were calculated and averaged under the following three conditions:

1. the subject is at rest for a period of 5 s to obtain a baseline signal;
2. the subject performs a plantar flexion of the ankle against a fixed resistance and keeps it constant for 5 s to obtain the highest possible sEMG signal resulting from the contraction of the Gastrocnemius Lateralis muscle;
3. the subject performs a plantar flexion of the ankle against a fixed resistance and keeps it constant for 5 s to obtain the highest possible sEMG signal resulting from the contraction of the Tibialis Anterior muscle.

These values were used to normalize the processed data.

In order to define the confidence level of sarcopenia in the examined patients and to label the acquired electromyographic signals, the following criteria were adopted:

- confidence level 1: if SARC-F ≥ 4;
- confidence level 2: if SARC-F ≥ 4 and hand grip-strength < cut-off values;
- confidence level 3: if SARC-F ≥ 4 and hand grip-strength < cut-off values and gait speed < 0.8 m/s.

Accordingly, the patients were divided into three groups: three with confidence level 1, twenty-two with confidence level 2 and seven with confidence level 3. In order to avoid unbalanced data, an oversampling strategy was used. Specifically, our dataset was pre-processed using Synthesizing Minority Over-sampling Technology (SMOTE) in combination with Edited Nearest Neighbours (ENN). SMOTE+ENN [12] is a sampling method that combines SMOTE [13] and Wilson's ENN [14] reducing the potential overfitting in synthetic data.

2.2. Pre-Processing

Pre-processing consists of: (a) reduction of noise and signal artefacts, (b) EMG enveloping and (c) data normalization. In the first step, the raw signals were filtered with a 4th-order Butterworth bandpass filter with a frequency between 20 Hz and 450 Hz.

In the second step, in order to make the signals comparable and suitable for further processing, the linear envelope of the signal was determined using full rectification and a Butterworth low-pass filter (with a cutoff frequency of 10 Hz). Lastly, the normalization step was performed as described in Section 2.1.

2.3. Feature Extraction & Selection

To extract relevant information from the surface EMG signal to identify muscle problems, several time-domain and time-frequency domain characteristics used in medical applications for monitoring lower limb muscles were analyzed [15, 16].

In this work, the size of the sliding window was set to 200 ms, with an incremental window of 50 ms. Twelve EMG features (integrated EMG, mean absolute value, modified mean absolute value type 1, modified mean absolute value type 2, root mean square, variance, average amplitude change, zero crossing, simple square integral, slope sign change, Willison amplitude, average instantaneous frequency) were extracted from each EMG channel, resulting in a feature size vector of 48. To reduce the complexity of signal processing and increase the performance of the system, the Modified Binary Tree Growth Algorithm (MBTGA) feature selection technique was applied, showing good performance for EMG signal analysis [17]. By applying this algorithm, the best subset of features was obtained: integrated EMG, root mean square and averaged instantaneous frequency for all channels; consequently, the dimension of the feature vector is 12.

2.4. Classification

After previous blocks, the data were labelled as different classes according to the sarcopenia confidence level. Three ML algorithms were trained on this data for comparison: SVM, K-Nearest Neighbors (KNN) and Extreme Gradient Boosting (XGB).

SVM [18] is a classification method developed in statistical learning. It has been shown to perform better in terms of accuracy than other classifiers and, in addition, to be scalable for large problems. SVM attempts to find a hyperplane in the N-dimensional space (where N is the number of features) in order to maximize the margin or distance between each dataset and the baseline for data classification. It identifies a set of support vectors, which represent the most representative observations for each baseline class. A kernel is used, which can be linear, polynomial or radial. In our approach, a linear kernel was applied.

KNN is another popular classification method due to its ease of implementation and high classification performance. The basic idea of the algorithm is to assign a sample to a category if most of the k neighboring samples of the considered sample belong to the same category. Usually, k is no greater than 20 [19]. K must be neither too small or too large because in the former case, the approach is sensitive to noise, while in the latter case, the neighborhood may include samples from other classes. The selected neighbors are those that have been classified correctly.

XGB [20] was designed to compensate the drawbacks of gradient boosting, enabling fast classification and excellent prediction results. It also avoids overfitting due to internal cross-validation in each iterative step.

To select the optimal parameters of each ML classifier, a grid search technique was applied [21]. These parameters are shown in Table 1.

Table 1
Parameters used for classification models.

Model	Parameters
SVM	decision_function_shape=ovo, max_iter=100, kernel=linear, C=0.1
KNN	n_neighbors = 5, metric = minkowski, algorithms = auto, weights = distance
XGB	learning_rate = 0.001, max_depth = 2, n_estimators = 214

3. Results and discussion

In order to validate the proposed approach, a series of experiments were carried out to verify its effectiveness and its operation in real time.

The EMG signal acquisition and processing steps are as follows: (1) acquisition of the raw signal through software routines implemented in C#, (2) data set balancing, data processing, feature extraction and classification using Python language (3.7.1). The experiments were performed on an embedded PC with an Intel Core i5 processor and 8 GB RAM. The classifiers' performance was evaluated using four different metrics: accuracy (Acc), precision (Pr), recall (Re) and F1-score. Some useful terms for these metrics are introduced in Table 2.

Table 2
Definition of terms used in metrics.

Predicted Label	Actual Label	Definition
Positive	Positive	True Positive (TP)
Positive	Negative	False Positive (FP)
Negative	Positive	False Negative (FN)
Negative	Negative	True Negative (TN)

Metrics can be defined by the following equations:

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \tag{1}$$

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

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

$$F1 - score = \frac{2 * TP}{2 * TN + FP + FN} \tag{4}$$

where accuracy is the ratio between all correctly classified samples and all samples. Precision indicates how accurately the model predicts positive occurrences. Recall shows how the model is able to detect positive cases using all positive cases. F1-score has a greater impact on true positive cases than precision.

The data set analyzed is unbalanced, particularly between the class with confidence level 2 of sarcopenia and the other two, in fact the ratio of confidence level 1 to confidence level 2 is 1:7.33. To avoid this imbalance and generate a more robust dataset, an augmentation strategy with all classes as

sampling strategy was used, as described at the end of section 2.1, resulting in a dataset of 150 patients with balanced sarcopenia confidence levels (47-52-51).

In this work, each ML model was trained by first considering all available features and then considering only the features obtained with the previously described feature selection technique. A significant improvement in model performance of approximately 10% was observed by applying feature selection.

Subsequently, the performance of these models was compared on the basis of test sets by applying a 10-fold cross-validation [22]. This procedure is used to perturb the training set of each classifier by randomizing the original dataset. In this way, each classifier is trained using 90% of the data, while the remaining 10% is used for testing. To avoid over-fitting of the training set, 10% of the training data is used to create a validation set. The procedure is repeated 10 times, training the classifier with a different training set and testing it with a separate test set making sure that the same samples do not appear in the training and test sets at the same time.

The performance of each ML model with dataset augmentation are shown in Table 3. SVM showed the best performance in terms of accuracy, precision, recall and F1 score considering the features obtained by feature selection with a percentage deviation in accuracy of approximately 3%.

Table 3
Classifier results with dataset augmentation and feature selection.

Model	Accuracy	Precision	Recall	F1
SVM	0.967	0.969	0.953	0.956
KNN	0.933	0.941	0.926	0.932
XGB	0.936	0.944	0.927	0.931

The confusion matrices of the average accuracy obtained for each considered classifier are shown in Figure 4. From the analysis of the confusion matrices, it can be seen that for the classifier with the best performance, errors occur in the classification of adjacent classes.

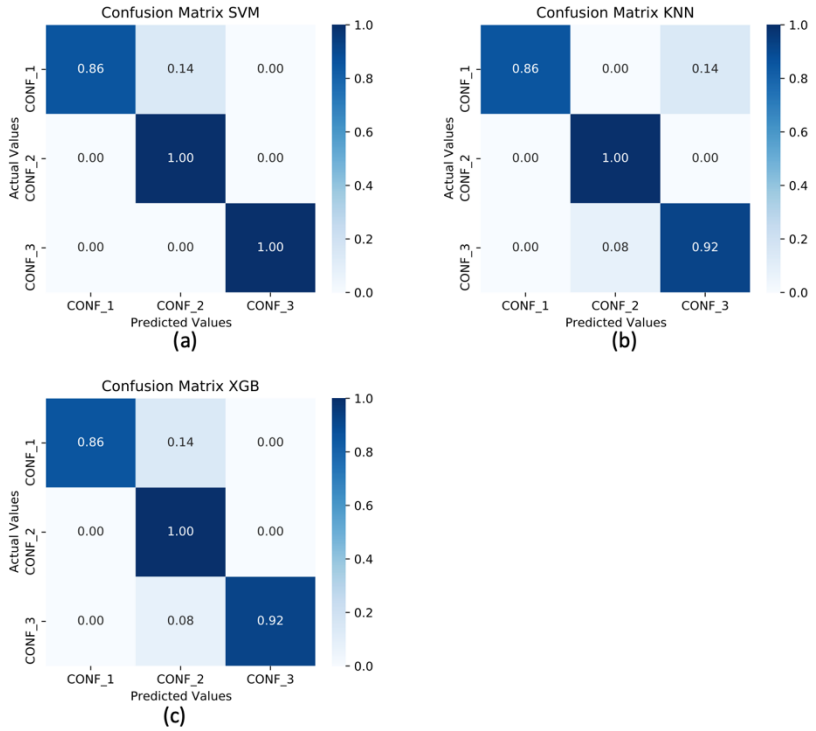

Figure 4: Confusion matrices for three classes of sarcopenia confidence levels using (a) SVM, (b) KNN, (c) XGB as classifiers with feature selection.

4. Conclusion

The present work has two main objectives. Firstly, an algorithmic pipeline was designed and implemented with the aim of classifying three different confidence levels of sarcopenia using electromyographic signals acquired via a commercial device (sEMG). Secondly, the performance of three supervised ML algorithms was evaluated to assess the level of disease severity. Due to the pandemic period and the consequent difficulty in obtaining real data, synthetic data from electromyographic signals were used from the real data of 32 subjects.

The results obtained confirmed the algorithmic choices made; in fact, all measured metrics showed a numerical improvement using the 'feature selection' logic block. The SVM classifier outperformed the other 2 supervised classifiers. However, an important limitation of the presented study must be emphasized. In order to compare the performance of the classification algorithms, only the synthesized data of a specific technique were considered.

A future development of this work will consist of evaluating a new test protocol to be able to assess muscular behavior even in subjects who are unable to perform sitting-to-standing speed and gait tests. In addition, other muscles (not only those of the lower limbs) will be analyzed, and long-term monitoring of electromyographic signals will be carried out to develop an intelligent tool for the early diagnosis of the considered disease.

5. Acknowledgements

This work has been carried out within the project PON 4Frailty funded by MUR-Italian Ministry for University and Research.

6. References

[1] J. Lexell, C. C. Taylor, and M. Sjöström. “What is the cause of the ageing atrophy?: Total number, size and proportion of different fiber types studied in whole vastus lateralis muscle from 15-to 83-year-old men”, J. Neurol. Sci. 84 (1988), 275-294.

[2] A. J. Cruz-Jentoft et al., “Sarcopenia: Revised European consensus on definition and diagnosis”, Age Ageing 48 (2019), 16-31.

[3] H. Ghasemzadeh, R. Jafari, and B. Prabhakaran, “A body sensor network with electromyogram and inertial sensors: Multimodal interpretation of muscular activities”, IEEE Trans. Inf. Technol. Biomed. 14 (2009), 198-206.

[4] I. Campanini, C. Disselhorst-Klug, W. Z. Rymer, and R. Merletti, “Surface EMG in clinical assessment and neurorehabilitation: Barriers limiting its use”, Frontiers in neurology 11 (2020), 934.

[5] F. Yu, A. Bilberg, E. Stenager, Wireless medical sensor measurements of fatigue in patients with multiple sclerosis, in: Proceedings of the 2010 Annual International Conference of the IEEE Engineering in Medicine and Biology, Buenos Aires, Argentina, 31 August-4 September 2010; pp. 3763-3767.

[6] S. H. Liu, C. B. Lin, Y. Chen, W. Chen, T. S. Huang, and C. Y. Hsu, “An EMG patch for the real-time monitoring of muscle-fatigue conditions during exercise”, Sensors 19 (2019), 3108.

[7] S. E. Askarinejad, M. A. Nazari, S. R. Borachalou, Experimental detection of muscle atrophy initiation Using sEMG signals, in: Proceedings of the 2018 IEEE 4th Middle East Conference on Biomedical Engineering (MECBME), Tunis, Tunisia, 28-30 March 2018; pp. 34-38.

[8] S. Cai, Y. Chen, S. Huang, Y. Wu, H. Zheng, X. Li, and L. Xie, “SVM-based classification of sEMG signals for upper-limb self-rehabilitation training”, Front. Neurorobot. 13 (2019), 31.

[9] J. I. Yoo, H. Byun, H. S. Kim, Y. J. Jang, and C. H. Lee, “Evaluating Postoperative Muscle Strength Using Surface Electromyography in Hip Fracture Patient”, J. Bone Metab. 27, (2020), 125.
[10] R. Habenicht, G. Ebenbichler, P. Bonato, J. Kollmitzer, S. Ziegelbecker, L. Unterlerchner, P. Mair, and T. Kienbacher, “Age-specific differences in the time-frequency representation of surface electromyographic data recorded during a submaximal cyclic back extension exercise: A promising biomarker to detect early signs of sarcopenia”, J. Neuroeng. Rehabil. 17, (2020), 8.
[11] J. Woo, J. Leung and J. E. Morley, “Validating the SARC-F: A suitable community screening tool for sarcopenia?”, J. Am. Med. Dir. Assoc. 15, (2014), 630-634.
[12] G. E. Batista, R. C. Prati and M.-C. Monard, “A study of the behavior of several methods for balancing machine learning training data”, ACM SIGKDD Explor. Newsl. 6, (2004), 20-29.
[13] Y. Mi, “Imbalanced classification based on active learning SMOTE”, Res. J. Appl. Sci. Eng. Technol. 5, (2013), 944-949.
[14] D. L. Wilson, “Asymptotic properties of nearest neighbor rules using edited data”, IEEE Trans. Syst. Man Cybern. 2, (1972), 408-421.
[15] R. Barioul, S. Fakhfakh, H. Derbel and O. Kanoun, Evaluation of EMG signal time domain features for hand gesture distinction, in: Proceedings of the 2019 16th International Multi-Conference on Systems, Signals & Devices (SSD), Istanbul, Turkey, 21-24 March 2019; pp. 489–493.
[16] A. Phinyomark, G. Chujit, P. Phukpattaranont, C. Limsakul and H. Hu, A preliminary study assessing time-domain EMG features of classifying exercises in preventing falls in the elderly, in: Proceedings of the 2012 9th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), Phetchaburi, Thailand, 16-18 May 2012.
[17] J. Too, A. R. Abdullah, N. Mohd Saad and N. Mohd Ali, “Feature selection based on binary tree growth algorithm for the classification of myoelectric signals”, Machines, 6, (2018), 65.
[18] J. A. Suykens and J. Vandewalle, “Least squares support vector machine classifiers”, Neural Process. Lett., 9 (1999), 293-300.
[19] S. Zhang, “Efficient kNN classification with different numbers of nearest neighbors”, IEEE Transac Neural Networks Learning Systems, 5 (2018), 1774-1785.
[20] T. Chen and C. Guestrin, Xgboost: A scalable tree boosting system, in: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, San Francisco, CA, USA, 13–17 August 2016; pp. 785-794.
[21] P. C. Bhat, H. B. Prosper, S. Sekmen and C. Stewart, “Optimizing event selection with the random grid search”, Comput. Phys. Commun., 228 (2018), 245-257.
[22] T. Hastie, R. Tibshirani and J. Friedman, The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Springer Science & Business Media: Berlin/Heidelberg, Germany, 2009.