

ELBOW FLEXION AND EXTENSION IDENTIFICATION USING SURFACE ELECTROMYOGRAPHY SIGNALS

A. Rubiano^{1,3}, J. L. Ramírez^{1,3}, M. N. El Korso^{1,3}, L. Gallimard^{1,3}, N. Jouandeau^{2,3}, O. Polit^{1,3}

¹LEME, Université Paris Ouest Nanterre La Défense,

²LIASD, Université Paris 8,

³Université Paris Lumière,

50, rue de Sèvres, 92410 Ville d'Avray, France.

ABSTRACT

In this paper, a new approach is presented for the analysis and the identification of the surface electromyography (EMG) signals of biceps and triceps muscles. The objective of this study is the accurate classification of elbow flexion and extension movements. We propose a cropping method based on the agreement of the movement changes and the EMG signal using the upper limb kinematic. Then, we perform the extraction and selection of several well known features in time and frequency domain. The selected features are used as inputs for our support vector machine classifier which is designed using an optimal weight vector criterion. Afterward, the training and test steps are performed in the proposed scheme. Finally, numerical simulation assess the accuracy of the classification, as well as the robustness of the proposed approach considering noisy measurements.

Index Terms— Electromyography, feature extraction, upper limb kinematic, support vector machine, quaternions.

1. INTRODUCTION

The surface electromyography (EMG) signal contains relevant information about the electrical activity produced by the neuromuscular process during contraction or relaxation [1]. Due to the nature of the EMG, it is possible to extract some characteristics, which are useful to identify the produced movements, as function of the measured electrical activity. Nevertheless, the complexity of the measured EMG signals makes the detection and/or classification process a difficult task [2].

In the literature, different studies had been carried out, with the aim to identify movements from EMG signals. In [3], the authors focused on the identification of five hand movements in which the obtained accuracy range was 94% to 99%. In the same way, [4] proposed the detection of the wrist and ring finger movements, and achieved an accuracy of 87.3%.

More recently, in [5], the authors put forward the recognition of eight grasping gestures, and earned an accuracy between 96.9% and 99.65%.

Although these studies have reached good accuracies, none of them have considered the synchronization of the upper limb kinematic with the EMG signal, in order to extract the signal fragment that match with each movement. One of the achievement of this communication, is the design of a real time synchronized process w.r.t. both kinematic and EMG signal, which guarantees the correct extraction of signals. Likewise, we consider variations in kinematical parameters, such as speed and angle range of the performed movements. These variations affect considerably the EMG signals, and so that, the classification accuracy. Even that, we achieve and maintain an accuracy of 99.79% using noisy measurements as input (i.e., measurements corrupted by an additive noise in order to assess the robustness of the proposed scheme) unlike the previous studies which used noise-free measurements (i.e., without adding an artificial additive noise).

Furthermore, we note that the obtained accuracy is obtained by testing the EMG signal from a new subject. This is of importance, since it means that the proposed scheme has both accurate accuracy and robustness to new subjects which could be not included in the learning process. Consequently, the proposed scheme earn generalization and applicability in real applications.

The main contributions of this paper are as follows:

1. The development of an open source software, which is able to: a) show in real time the synchronization of upper limb kinematic and EMG signals, b) fill automatically data gaps in kinematics using quaternion's approach, c) extract EMG signals in an accurate way through the calculation of angle and speed of movements, during the flexion and extension.
2. Design a methodology for analyzing and choosing the bests features to identify and classify movements, avoiding black boxes, which improve the control of the classification and generalization of the method.
3. Propose a nonlinear support vector machine (SVM) classifier which is feeded with an experimental data set, that con-

This work has been supported by the Université Paris Lumière through the ProMain project. A. Rubiano acknowledges Universidad Militar Nueva Granada, where she is assistant professor.

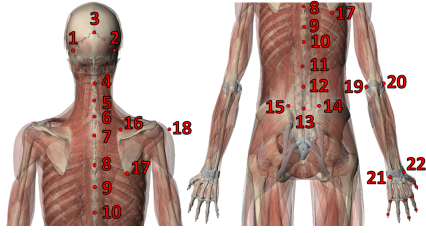


Fig. 1. Distribution of markers in the upper limb kinematic chain.

sider variations in kinematical parameters. Furthermore, we perform a proper exploration of the SVM settings, which give us an identification of the parameters influences in the learning process. Finally, the robustness of the proposed scheme is test by artificially adding zero mean white gaussian noise.

2. EXPERIMENT AND MATERIALS

Three healthy subjects have participated in the proposed experiment. Each subject performed flexion and extension movements, changing the covered angle and speed of movement, while the upper limb motion tracking (ULMT) and EMG signals for biceps and triceps are recorded. Subject one, two, and three, performed one, four, and six trials respectively, thus making a total of 44 recorded EMG signals.

The EMG signals have been recorded using the ZEROWIRE wireless EMG system with a sampling frequency of 1 kHz . All subjects were equipped with two EMG sensors, which were located over the biceps brachii and triceps. We selected these muscles, because they are the main muscles involved in the elbow flexion and extension [6], and consequently, they produce high electrical activity during the movements.

The ULMT, is performed using 7 3D VICON cameras and Nexus software. The system has a sampling frequency of 200 Hz . The subjects were equipped with 30 retro-reflective markers, positioned all over the upper limb kinematic chain including the spinal column, see Fig. 1, distributed as follows: i) Occipital bone (markers 1 to 3). ii) Cervical and lumbar vertebra (markers 4 to 13). iii) Sacrum (markers 13 to 15). iv) Scapula (markers 16 and 17) v) Medial epicondyle of humerus (marker 18). vi) Coronoid process of ulna and radius (markers 19 and 20). vii) Styloid process of ulna and radius (markers 21 and 22). viii) Metacarpophalangeal joint of fingers 1, 2, 3 and 5 (markers 23 to 26). ix) Proximal interphalangeal joint of the fingers 2, 3 and 5 (markers 27 to 29). x) Interphalangeal joints of finger 1 (marker 30).

3. PRE-PROCESSING OF DATA

The pre-processing of data consist in the visualization and extraction of the EMG signals, that corresponds to the flexion and extension movement. For this issue, we develop a open

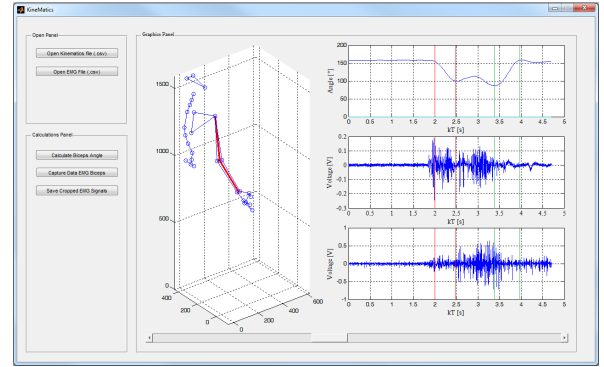


Fig. 2. Software's interface.

source software¹ (developed in Matlab, see Fig. 2) which: i) Displays ULMT data, ii) displays the corresponding EMG signals, iii) calculates the elbow angle and the corresponding angular velocity, iv) synchronizes in real time the EMG signals and the ULMT data, with the aim of extracting signals for flexion and extension movements from the raw EMG data taking into account the different sampling frequencies. For the purpose of plotting the ULMT data, the markers are clustered by regions as follows: i) Head markers 1 to 3, ii) spinal column markers 4 to 13, iii) sacrum markers 13 to 15, iv) scapula markers 16 to 18, v) arm markers 18 to 20, vi) forearm markers 19 to 22 and vii) hand markers 21 to 30. These regions are used to draw lines between markers to allow an easy interpretation of the 3D visualization, linking the scattered points in the graph (this explains that some markers are repeated in different clusters).

3.1. Automatic correction of missing data and outliers

Frequently, some markers of the ULMT disappear within several periods of times. Vicon Nexus Software try to fill the gaps of these missing markers. Nevertheless, sometimes, it appears that the software fails to fill it and let empty the values of the missing markers (which are set to 0 in Matlab, see Fig. 3). Nevertheless, each zero has a dramatic effect on the accuracy of the following classification steps.

Consequently, to overcome this disadvantage, we develop a mathematic model, which is able to determine the values of the markers which disappear. The model use quaternions, since they are used to formulate spatial rotation in many fields, including computer graphics and robotics [7]. Due to space limitation, we briefly describe the model, more details are included in the open source package software.

Our proposed model consists of a coordinate system fixed to each marker, and a set of vectors which links the markers between them, maintaining the order of the markers cluster (e.g. marker 1 is linked to marker 2 by a vector, and marker

¹The open source software will be available soon at <http://leme.u-paris10.fr/promain-565200.kjsp?RH=1415865400449>

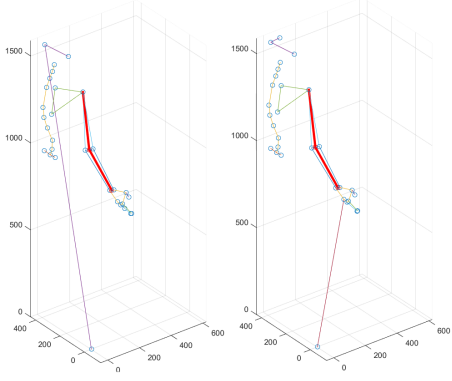


Fig. 3. Examples of missing ULMT Markers.

2 is linked to marker 3 by another vector, and so on). Then, we define a group of quaternions, which express the rotations between the created vectors, in order to verify the position of each marker. When the position of a marker exceed a certain threshold (in our case 3mm) w.r.t. the calculated position with the quaternions, the marker is replaced by the calculated point, so that, the 3D model is automatically corrected.

3.2. Elbow angle evaluation and analysis

Once the 3D model is corrected, we can determine correctly the elbow angle φ . For this, we use two vectors: one from the elbow to the wrist and the other from the elbow to the shoulder. Since there are two markers in the elbow, it is required to find the position of the middle point p_1 between the coranoid process ulna and radius (markers 19 and 20), and also the point p_2 in the wrist, between the markers placed over the styloid process of ulna and radius (markers 21 to 22).

The arm vector denoted by $\mathbf{v}_a \in \mathbb{R}^3$ is the vector that links the point p_3 (marker placed over the epicondyle of humerus), to the point p_1 . The forearm vector denoted by $\mathbf{v}_{fa} \in \mathbb{R}^3$ links points p_1 and p_2 . With these two vectors, the elbow angle φ is calculated as $\varphi = \arccos\left(\frac{\langle \mathbf{v}_a, \mathbf{v}_{fa} \rangle}{\|\mathbf{v}_a\| \|\mathbf{v}_{fa}\|}\right)$. Since the EMG signal and the UMLT are synchronized, our software calculate the angular velocity, defined as $\omega = d\varphi/dt$, to identify if the angle φ is increasing or decreasing. The behavior of φ allows the exact identification of elbow flexion and extension movements, which is important for the correct and accurate extraction of EMG signals. As result of the pre-processing, we obtain a set of EMG signals which corresponds to electromyography of biceps and triceps for flexion and extension at different angles and speeds (see Table. 1) for all subjects and all trials.

4. FEATURE EXTRACTION

Feature extraction is a technique to obtain relevant information of a signal, in time or frequency domain, which is given

Subj.	Trial	Range Angle [°]		Speeds [rad/s]	
		Flex	Ext	Flex	Ext
1	1	057,92	-071,29	1,80	-2,04
	1	097,37	-102,28	1,24	-1,30
2	2	108,69	-117,79	1,93	-1,76
	3	117,27	-117,87	8,15	-6,22
	4	117,27	-117,87	1,49	-1,32
3	1	048,54	-038,74	2,72	-1,77
	2	043,18	-039,88	1,71	-1,22
	3	036,80	-041,40	2,56	-2,18
	4	046,60	-044,44	2,62	-1,52
	5	046,76	-046,56	2,54	-1,84
	6	045,63	-043,73	2,41	-1,49

Table 1. Angles and speeds of movements.

by a single vector value [3]. There are several features often used for analyzing EMG signals. Specifically, we consider the following features:

- Entropy : $\text{Ent}(\mathbf{s}) = -\sum_i o_i^2 \log_2(o_i^2)$ where o_i represents the projection coefficients of the signal \mathbf{s} in an orthonormal basis [8].
- Mean Absolute Value : $\text{MAV}(\mathbf{s}) = n^{-1} \sum_{i=1}^n |[\mathbf{s}]_i|$, where $[\mathbf{s}]_i$ represents the i -th sample of the signal \mathbf{s} , and n is the number of samples.
- Mean value: $\text{M}(\mathbf{s}) = n^{-1} \sum_{i=1}^n [\mathbf{s}]_i$.
- Harmonic mean: $\text{HM}(\mathbf{s}) = n(\sum_{i=1}^n |[\mathbf{s}]_i|^{-1})^{-1}$.
- Mean frequency: $\text{MF}(\mathbf{s}) = (\sum_{j=1}^N I_j)^{-1} \sum_{j=1}^N I_j f_j$ where N denotes the number of harmonics in the spectrum, I_j represents the magnitude of the j -th harmonic, and f_j is the frequency of the j -th harmonic.

As previously mentioned, the aim of this study is to identify two movements (elbow flexion and extension) from EMG signals. Therefore, the proposed approach is to extract the aforementioned features of electromyography signals, to create two sets of features, one for flexion, and one for extension. These sets are used to characterize the difference between the features which is evaluated using an Euclidean distance. It is important to note that the generated sets of features are affected by the variations in kinematical parameters, such as speed and angle range of the performed movements.

In order to chose the best set of feature for our classification scheme, the maximum distance between the nearest values of each couple of feature is calculated. For this, we propose the following criterion which is applied to all trials in the two muscles (biceps and triceps):

$$(b, u) = \arg \max_{b, u} (\min (|F_b^{\text{ext}} - F_b^{\text{fle}}| + |F_u^{\text{ext}} - F_u^{\text{fle}}|))$$

$b \neq u$ in which F_b^{ext} and F_b^{fle} are the b -th features for extension and flexion, F_u^{ext} and F_u^{fle} are the u -th features for extension and flexion. Based on the the proposed criterion, the selected features are found to be the entropy and the mean frequency features. Therefore, with these selected features, we build a matrix \mathbf{X} , in which each row is the value of

Flexion					Extension				
row	Biceps		Triceps		row	Biceps		Triceps	
	Ent	MF	Ent	MF		Ent	MF	Ent	MF
1	3,3	269,4	3,4	270,2	12	3,7	267,1	2,3	262,6
2	3,4	270,3	2,0	259,0	13	2,0	259,3	2,2	261,9
3	3,5	271,0	2,0	259,7	14	2,0	259,3	2,2	262,0
4	3,2	272,6	2,5	262,7	15	2,5	262,6	2,5	263,8
5	3,5	269,8	2,1	258,0	16	1,8	257,6	2,4	264,6
6	4,4	280,6	2,8	266,7	17	2,6	263,4	2,9	269,8
7	3,0	269,8	2,2	262,1	18	2,3	265,6	2,1	267,1
8	3,2	272,6	2,5	262,7	19	2,5	262,6	2,5	263,8
9	3,4	269,5	2,1	260,0	20	2,3	263,8	1,7	260,7
10	4,3	273,7	2,7	264,4	21	2,6	263,7	3,0	271,2
11	3,3	269,1	2,3	262,6	22	2,1	261,5	2,2	265,9

Table 2. The feature matrix \mathbf{X} in which rows one and twelve are trial extracted from subject one, rows two to five and rows thirteen to sixteen are trials extracted from subject two, and the other rows are trials extracted from subject three.

$[\text{Ent}(\mathbf{s}_{bi}), \text{MF}(\mathbf{s}_{bi}), \text{Ent}(\mathbf{s}_{tr}), \text{MF}(\mathbf{s}_{tr})]$ for one trial, where \mathbf{s}_{bi} and \mathbf{s}_{tr} are the biceps and triceps recorded EMG signals respectively (note that the extracted features are influenced by variations in kinematical parameters). Table 2 shows the extracted values of the selected features stored in the matrix \mathbf{X} for different subject and different trials.

5. CLASSIFICATION SCHEME

The matrix \mathbf{X} , can be considered as a set of points in a 4-th dimensional features space, wherein each l -th point $\mathbf{x}_l \in \mathbb{R}^4$, has an etiquette $y_l \in \{1, -1\}$. The etiquette classifies the points into two categories: flexion or extension. The points of the matrix \mathbf{X} can be separated, using a learning strategy based on statistical theory, called support vector machine (SVM). The SVM calculate an optimal hyperplane between such categories, using a matrix \mathbf{X}' composed by q rows, where each row \mathbf{x}'_q is a support vector. In our case, the support vectors are the rows three, four, six, fourteen, fifteen and seventeen of the matrix \mathbf{X} , i.e. three vectors for flexion and three for extension.

In our study, a linear separation is not possible, as can be seen from Fig. 4. Consequently, we propose to use a non-linear SVM classifier, where the optimal hyperplane is obtained by solving a quadratic programming problem (QPP) [9], given by $\min_{\mathbf{w}, b, \xi} \frac{1}{2} \mathbf{w}^T \mathbf{w} + C \sum_{q=1}^m [\xi]_q$ s.t. $y_q (\mathbf{w}^T \phi(\mathbf{x}'_q) + b) \geq 1 - [\xi]_q$ and $[\xi]_q \geq 0$ for $q = 1, \dots, m$ where \mathbf{w} and b determine the hyperplane in feature space, $[\xi]_q$ represents the error soft margin, C is a term to control the overfitting, m denotes the amount of support vectors inside \mathbf{X}' , and ϕ maps \mathbf{x}'_q into high-dimensional space. The solution of this problem, is the following decision function:

$\Gamma(\mathbf{x}_l) = \text{sign} \left(\sum_{q=1}^m \alpha_q y_q K(\mathbf{x}'_q, \mathbf{x}_l) + b \right)$ in which α_q are the lagrange coefficients used to solve the QPP problem [8], and $K(\mathbf{x}'_q, \mathbf{x}_l)$ denotes a kernel function, that is, in

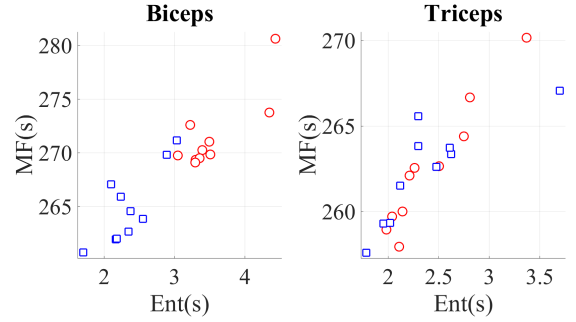


Fig. 4. Distribution of features for flexion and extension. Red circles and blue squares represent, respectively, flexion and extension features.

our case, given by the following radial basis function (RBF):

$K(\mathbf{x}'_q, \mathbf{x}_l) = \exp^{-\frac{\|\mathbf{x}'_q - \mathbf{x}_l\|^2}{2\sigma^2}}$ where σ denotes a positive parameter for controlling the radius. This RBF function is selected due to its capacity of maximizing the distance between the hyperplane and the points [10].

Since we use six support vectors, the decision function $\Gamma(\mathbf{x}_l)$ is applied to the remaining points \mathbf{x}_l of the matrix \mathbf{X} . Then, the result of $\Gamma(\mathbf{x}_l)$, is compared with y_l . If the values match, it is considered that the SVM has correctly classified the point \mathbf{x}_l . As result, using a non optimal setting of σ and C , the algorithm may achieve a classification percentage of 90.91%. This result shows that only two movements of all the trials (flexion of second subject with trial three, and flexion of the third subject with trial three) are not correctly classified. It is important to remark that the SVM is trained only with 13.63% of the whole trials (considering trials of subject two and three), but the algorithm is able to classify correctly the EMG signals produced during flexion and extension of the first subject whose patterns are new for the classifier and was not used in the training step.

In order to improve this result, we analyse the effect of the parameters σ and C in order to identify their influence in the final percentage of classification. This numerical analysis consists of a double iteration changing the values of the enounced parameters. As result, we find that there is an optimal area for values of σ near to 1.1 and C close to 2.5, where the percentage of classification reaches 100% as is shown in Fig. 5. It is important to note that the 100% of successful classification is due to the fixed (finite) number of testing data. To test the method in a more realistic environment, the EMG signal may be corrupted by some noisy data. Such scenario is discussed in the following section.

6. CLASSIFICATION ROBUSTNESS IN THE PRESENCE OF NOISY MEASUREMENTS

Even if we obtained an accurate result of EMG signals identification with the nonlinear SVM classifier, it is necessary to consider that there are several conditions that could influence

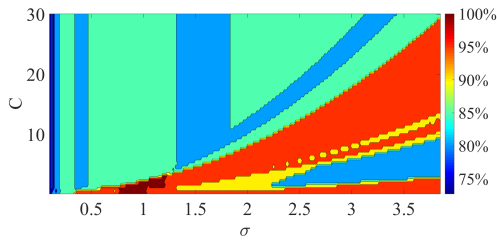


Fig. 5. Percentage of classification as function of σ and C .

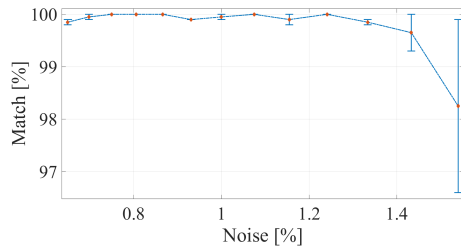


Fig. 6. Classification results after adding noise for the first subject and trial one.

the results (e.g. the sweat and fatigue of the subject and displacement of the recording electrodes). The effect of some non idealistic conditions can be modeled as an additive white Gaussian noise (AWGN) [1]. In our simulations, we added AWGN to the biceps and triceps EMG signal measurements of: (a) the first subject for trial 1, (b) the second subject for trial 1, (c) the third subject for trial 5. Note that the contaminated trials are only used to test robustness and thus, they were not in the training set X' of SVM. The noise is generated for several values of signal to noise ratio (SNR), using 1000 Monte Carlo trials for each value of SNR. The corresponding percentages of the additive power of noise, over the power of EMG signal, lies between 0.64% and 1.43%. Subsequently of adding noise to these signals, the same process described in section 4 is applied in order to obtain features of the noisy signals. The SVM achieves a successful classification of 99.79%. The results of percentage of classification with respect to the percentage of noise added to the signal, for the subject one, is presented in Fig. 6 (for the other subjects the same behavior is noticed but due to space limitation it is not presented here).

7. CONCLUSION

In this paper, a new elbow flexion and extension identification scheme is proposed. Quaternion based method was used to correct the ULMT data analysis, which allow us to calculate the position of missing kinematic markers. Accordingly, with the ULMT, we calculated, in an accurate way, the angles and speeds of the movement. These values are important, in order to select the exact time window, that matches with each movement, avoiding to extract features with wrong information. Furthermore, we designed a novel criterion to select the

best couple of feature for the classification. Then, we set up a non linear SVM classifier which has as input the selected features. Finally, numerical simulation showed that the proposed scheme performs almost perfectly (99,79%) with a moderate contaminated measurement.

REFERENCES

- [1] M. B. I. Reaz, M. S. Hussain, and F. Mohd-Yasin, "Techniques of EMG signal analysis: detection, processing, classification and applications," *Biological Procedures Online*, vol. 8, no. 1, pp. 11–35, Dec. 2006.
- [2] A. Alkan and M. Gnay, "Identification of EMG signals using discriminant analysis and SVM classifier," *Expert Systems with Applications*, vol. 39, no. 1, pp. 44–47, Jan. 2012.
- [3] M.A. Oskoei and H. Huosheng, "Support vector machine-based classification scheme for myoelectric control applied to upper limb," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 8, pp. 1956–1965, Aug. 2008.
- [4] G. R. Naik, D.K. Kumar, and Jayadeva, "Twin SVM for gesture classification using the surface electromyogram," *IEEE Transactions on Information Technology in Biomedicine*, vol. 14, no. 2, pp. 301–308, Mar. 2010.
- [5] N. Wang, Y. Chen, and X. Zhang, "The recognition of multi-finger prehensile postures using LDA," *Biomedical Signal Processing and Control*, vol. 8, no. 6, pp. 706–712, Nov. 2013.
- [6] A. Ali, K. Sundaraj, R. B. Ahmad, N. U. Ahamed, and A. Islam, "Surface electromyography for assessing triceps brachii muscle activities: A literature review," *Bio-cybernetics and Biomedical Engineering*, vol. 33, no. 4, pp. 187 – 195, 2013.
- [7] P. Bell and I. Potapov, "Reachability problems in quaternion matrix and rotation semigroups," *Information and Computation*, vol. 206, no. 11, pp. 1353 – 1361, 2008.
- [8] Q. Liu, X. Zhao, and Z. Hou, "Metric learning for event-related potential component classification in EEG signals," in *Proceedings of the 22nd European Signal Processing Conference (EUSIPCO)*, Lisbon, Portugal, Sept. 2014, pp. 2005–2009, .
- [9] F. F. Chamasemani and Y. P. Singh, "Multi-class support vector machine (SVM) classifiers - an application in hypothyroid detection and classification," in *Sixth conference on Bio-Inspired Computing: Theories and applications, IEEE*, Penang, Malaysia, Sept. 2011, pp. 351–356, .
- [10] X. Peng, "TPMSVM: A novel twin parametric-margin support vector machine for pattern recognition," *Pattern Recognition*, vol. 44, no. 10-11, pp. 2678–2692, Oct. 2011.