2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Verification-Based Design of a Robust EMG Wake Word

Pradeep Kumar[1], Angkoon Phinyomark[1], *Member, IEEE*, and Erik Scheme[1], *Senior Member, IEEE*

Abstract— **Surface electromyography (sEMG) signals are now commonly used in continuous myoelectric control of prostheses. More recently, researchers have considered EMG-based gesture recognition systems for human computer interaction research. These systems instead focus on recognizing discrete gestures (like a finger snap). The majority of works, however, have focused on improving multi-class performance, with little consideration for false activations from "other" classes. Consequently, they lack the robustness needed for real-world applications which generally require a single motion class such as a mouse click or a wake word. Furthermore, many works have borrowed the windowed classification schemes from continuous control, and thus fail to leverage the temporal structure of the gesture. In this paper, we propose a verification-based approach to creating a robust EMG wake word using one-class classifiers (Support Vector Data Description, One Class-Support Vector Machine, Dynamic Time Warping (DTW) & Hidden Markov Models). The area under the ROC curve (AUC) is used as a feature optimization objective as it provides a better representation of the verification performance. Equal error rate (EER) and AUC are then used as evaluation metrics. The results are computed using both window-based and temporal classifiers on a dataset consisting of five different gestures, with a best EER of 0.04 and AUC of 0.98, recorded using a DTW scheme. These results demonstrate a design framework that may benefit the development of more robust solutions for EMG-based wake words or input commands for a variety of interactive applications.**

I. INTRODUCTION

Surface electromyography (sEMG) measures the electrical activity produced by skeletal muscles during contraction. The recorded signals contain a rich amount of information associated with human motion intent [1]. It is extensively used in continuous myoelectric control (for applications such as prosthetics), where sustained contractions are used for velocity or position control [2]. More recently, researchers have begun to consider EMG for human computer interaction (HCI) purposes with a focus on enabling multi-class recognition systems. The most commonly seen application for such an HCI is hand gesture recognition, which provides a natural method of hands-free interaction with emerging heads-up devices [3]. The majority of EMG-based HCI efforts, however, have focused on the recognition of multiple classes (such as for sign language recognition [4]), with little consideration for false activations from "other" classes. This lack of consideration for unknown contractions, which are inevitable in every day use, limits their applicability in real-world situations.

Some attempts have been made to limit the number of false activations by using techniques such as rejection thresholds [5] or majority voting [6]. For example, Robertson et al. [5] proposed a confidence-based rejection strategy to improve the usability of myoelectric control systems. They designed a Fitts' law-style virtual cursor control system governed using forearm EMG and a support vector machine (SVM) classifier and determined rejection thresholds heuristically to improve the overall performance of the system. A one-vs-one classification scheme was similarly proposed in [6] to reject unknown data patterns. The scheme was based on uncorrelated linear discriminant analysis (ULDA) projection and the distance threshold of the query data point to the class means followed by a majority vote. Their approach outperformed nine conventional classifiers; however, a fine tuning of different thresholds was required for best performance. Comparatively, there exists few works that have explored the usage of a 'wake' gesture (or 'click') to reduce the false activations [7], [8], although some have explored the inclusion of IMU information to reduce the likelihood of motion artifacts [9], however, EMG is still considered as the gold standard for retaining gesture specific information in comparison to IMU alone [10]. Tavakoli et al. [8] proposed the inclusion of a 'lock' gesture to reject unwanted movements. They used a fast double wrist flexion gesture (similar to double impulse mode switching in conventional EMG prosthesis control) to lock/unlock the system, in addition to a 15s timeout to lock the entire system. While possibly robust, such approaches detract from the usability and responsiveness of the interface. In contrast, the concept of gesture verification could be implemented to inherently limit the false activations by design. Verification systems are described as 1-to-1 matching systems, making them a good choice for applications targeting only one class. Furthermore, such a system would only require training of the target gesture class, and could be accomplished using one-class classifiers.

Most existing gesture recognition systems are also built using time-independent classifiers such as SVM, LDA or Naives Bayes [11]. These classifiers are trained on short windows of (presumably stationary) data, following the approach outlined by Englehart and Hudgins in 2003 [12]. This approach, when trained with a finite number of discrete classes, can lead to a large number of false activations during gesture transitions and other unwanted movements. The classification performance of such systems can be improved using voting strategies, but this imposes further constraints as the window

*This work was supported by MITACS, Canada and the New Brunswick Innovation Foundation

[1]P. Kumar, A. Phinyomark and E. Scheme are with the Institute of Biomedical Engineering, University of New Brunswick, Fredericton, NB, E3B 5A3, Canada pradeep.kumar@unb.ca, aphinyom@unb.ca, escheme@unb.ca

lengths and the number of decisions involved in voting remains fixed [13]. Rather, gestures may be represented as a sequence of features that vary temporally, and this temporal variation information can be used to improve classification. For example, in a typical hand gesture recognition system, looking at the hand's position at a single point in time during the transition may not robustly identify the gesture; this can only be achieved by analyzing changes over time [14]. Therefore, temporal classifiers could be used to classify specific gesture movements more robustly to make them robust to outlier movements that may contain similar features but follow dissimilar temporal paths.

In this paper, we propose a verification-based scheme for a robust EMG wake word in order to better reject the classification of outlier movements. We compare the performance of both window-based (Support Vector Data Description (SVDD), 1-class SVM (OC-SVM)) and temporal (Dynamic Time Warping (DTW) , Hidden Markov Model (HMM)) one class classifiers optimized using the area under an ROC curve as their optimization metric for feature selection.

II. Methods

A. Data Collection

Five healthy subjects participated in this pilot study. The data were recorded using a custom-built device [15] on the posterior side of their dominant wrist that consisted of two bipolar EMG channels placed laterally, 2.7 cm apart. The data were sampled at 600 Hz with an amplifier gain of 330x. Five repetitions each of five different gestures (thumb extension, index extension, pinky extension, hand open, and wrist extension) were recorded from each subject while in a seated position. All subjects were asked to perform the gesture naturally, initiating and releasing them at their own subjective rates, instead of sustaining a set contraction level for a fixed period of time. Subjects were also asked to elicit contractions at a comfortable and repeatable force level. All experiments were approved by the University of New Brunswick's Research Ethics Board under REB#2019-114. The recorded data were then high-pass filtered at 20 Hz using a 3rd-order Butterworth filter to remove any motion artifacts, and notch filtered at 60 Hz using a 2nd-order infinite impulse response (IIR) notch filter to remove power line interference. Next, a Hilbert transformation [16] based technique was applied to detect the onset of the gestures.

Classification results were computed by performing user-dependent training of the classifiers using a leave-one-trial-out cross validation scheme and the average results across users were reported. Feature selection was conducted across users using the area under the receiver operating curve (AUC) as the criterion function. The set of features corresponding to the maximum AUC across subjects for each classifier was used in the gesture verification process.

Gesture verification was performed for a given motion (e.g. index finger extension), and the remaining four gestures were treated as outlier movements. Based on the different classification schemes, different training protocols were followed. For example, the SVDD and OC-SVM classifiers were trained on a per-window basis and the confidence scores were averaged across windows for the entire gesture. In contrast, DTW and HMM were trained using the entire target gesture sequence. A Euclidean distance metric with a limited warping length constraint was used to get the similarity score between gesture sequences when using DTW [17]. Based on empirical testing, the HMM was trained using three states and 64 Gaussian mixture components.

B. Feature Selection and Classifiers

Feature Extraction & Selection: Before feature extraction, the data were segmented into 150 ms windows with an increment of 25 ms. Thirteen different features were explored: Mean Absolute Value (MAV), Slope Sign Changes (SSC), Waveform Length (WL), Zero Crossings (ZC), Mean Square Root (MSR), Maximum Fractal Length (MFL),Time-Domain Power Spectral Moments (TDPSD), Mean Absolute Values of the Second Difference (MAVSD), Difference Absolute Standard Deviation Value (DASDV), Willison Amplitude (WAMP), Sample Entropy (SAMPEN), Difference Absolute Mean Value (DAMV), and Difference Auto-Regressive Coefficients (DAR) [18].

To increase information density, a sequential feature selection (SFS) technique was applied instead of other projection based techniques (PCA, ULDA) which transform all features into the dimension-reduced feature space [19]. In contrast, SFS examined each feature individually and adds them iteratively until the objective function was satisfied. Although classification accuracy is commonly employed as the objective function, it does not consider the tradeoff between false-positive and false-negative samples. In this work, the objective was not so much to improve the classification performance of the system, but to optimize the verification of gestures in order to limit the false activations. For this reason, the area under the ROC curve was used as the criterion function in feature selection. The ROC curve is a graphical representation of the tradeoff between false-positive rate (FPR) and true-positive rate (TPR) and, as such, is considered to be an effective metric for assessing the performance of verification models [20], [21]. By extension, AUC measures the area underneath the entire ROC curve and provides an aggregate performance measure for the tradeoff. The higher the AUC, the better the model at distinguishing the classes with minimal false activations. Therefore, by optimizing the AUC during feature selection, the verification performance of the system is optimized more stringently than classification accuracy alone.

Classification: As with other verification problems (facial, signature, etc.), gesture verification is a one-to-one matching task, i.e., one must determine whether a particular gesture occurred or not. Therefore, in this work, we assessed a family of one-class classifiers including both window-based (OC-SVM, SVDD) and temporal (HMM, DTW). Neither one-class SVM nor SVDD classifiers are widely used in EMG pattern classification. However, Liu et al. [22] demonstrated the usability of an SVDD classifier to filter the non-target gestures when the authors trained an ensemble of SVDDs,

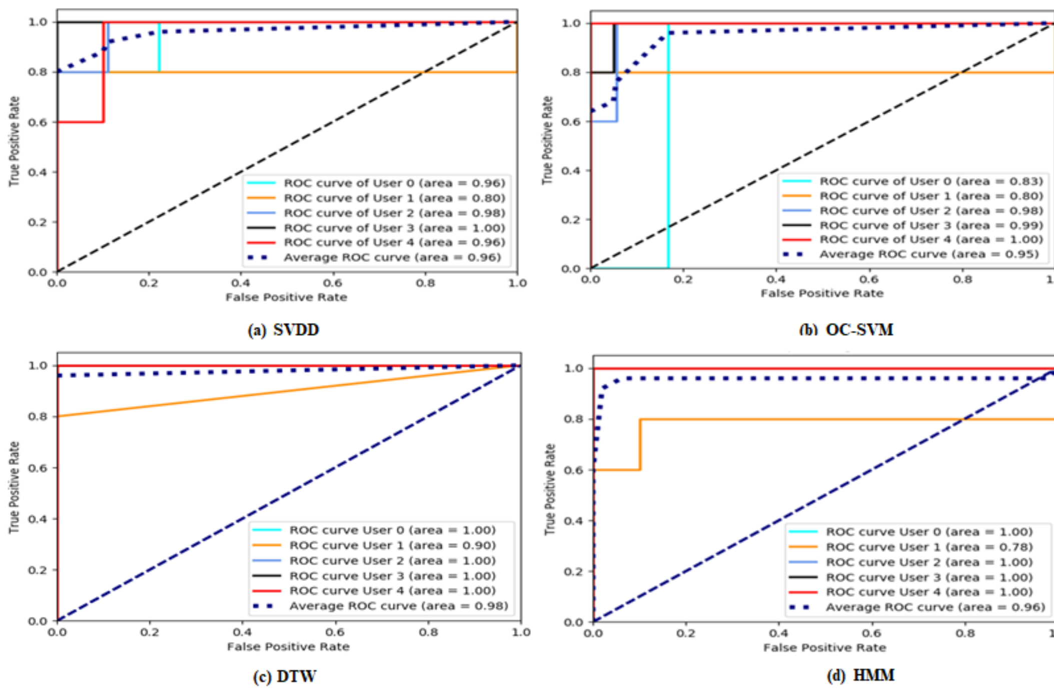

Fig. 1: ROC curves for gesture verification for Index extension across different classifiers.

one for each target class. For a given training set (T), the SVDD classifier aims to find the minimum-volume sphere in the feature space with center (c) and radius (r) such that all, or most of the training patterns are enclosed by the hypersphere. The hypersphere boundary is then used to distinguish between target and not-target data points [22]. Likewise, OC-SVM is an extension of SVMs and is commonly used in anomaly detection problems [23]. The classifier estimates a probability distribution that encompasses most of the observed training data and then labels those which lies far from the distribution as 'suspicious'.

Alternatively, sequential classifiers like HMM and DTW have been used in a handful of EMG gesture recognition problems for tasks such as transition-point detection [13] and sign-gesture recognition [24]. However, most of the existing works with these approaches have focused on the classification of multiple gestures, and not restricting outlier motions in a verification framework. Consequently, in this work, we use both HMM and DTW as one-class classifiers as part of a robust gesture verification system. HMM models are typically represented by three tuples (π, A, B), where π is the initial state probabilities and A denotes the state transition matrix from one state to other. Lastly, B is the observation probability that is modeled with the continuous probability density function for a given state. The model was trained using the Baum-Welch algorithm for initial output probability re-estimation and for maximizing the likelihood of the training set [25]. Likewise, DTW is popularly used for measuring the similarity between two time-series that may vary in time or speed. The algorithm recursively computes distances between two series and allows a non-linear mapping between them in the time-dimension by minimizing their distance. Thus, the similarity between the two time-series is generally represented by the DTW distance measure [26]. In this work, DTW distances were computed between the various instances of the target EMG gesture and a corresponding distance threshold was stored for the verification of test gestures.

C. Gesture Verification

For a given query gesture, the verification of its genuineness was computed based on a threshold value (th). Depending on the algorithm, the threshold value was a similarity measure or a distance value from the decision line defined for the genuine gesture. If the similarity measure was lower than the selected threshold for the classifier, the gesture was accepted as genuine, and it was otherwise rejected as an outlier motion. The verification decisions for all gestures were governed by Eq. (1), where 'X' and 'th' represent the query gesture and threshold value, respectively. If the similarity measure (S_m) of the query gesture was less than 'th', the gesture was treated as a target gesture.

$$Decision(X|S_m) = \begin{cases} Target\ gesture, & \text{if } S_m < th \\ Outlier, & \text{otherwise} \end{cases} \tag{1}$$

III. Results & Discussion

To evaluate their verification performance, ROC curves were generated for the different classifiers by varying their decision thresholds. The AUC was then recorded for each as shown in Fig. 1. This performance reflects the results for each classifier when using the features selected by the AUC-based SFS process, whose results are outlined before. The DTW-based gesture verification scheme outperformed the other classifiers with an average AUC score of 0.98. The DTW also yielded the minimum EER of 0.04, as shown in Table

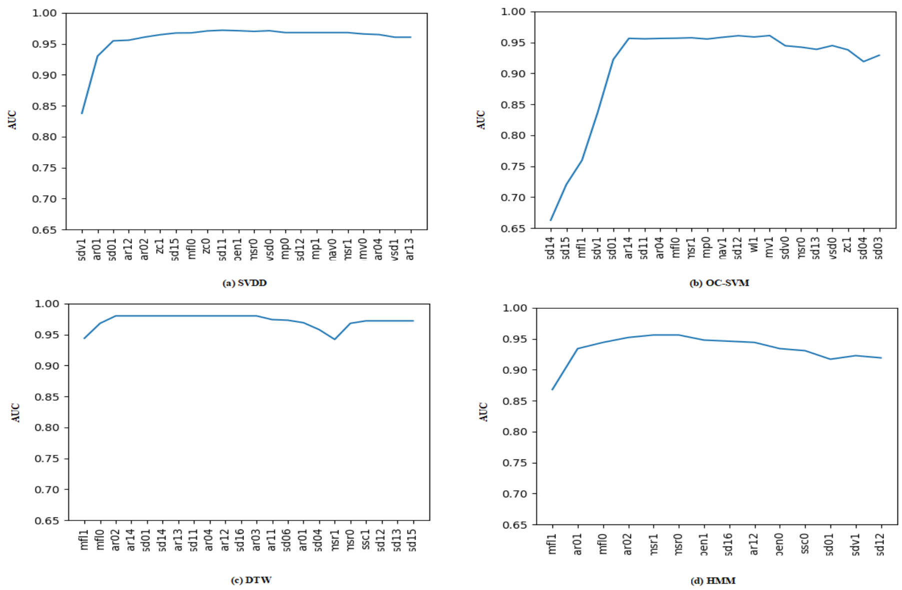

Fig. 2: AUC-based sequential feature selection across different classifiers: (a) SVDD, (b) OC-SVM, (c) DTW, (d) HMM.

I. This was observed to be due to the temporal modelling of the gesture using DTW, whereas the conventional classifiers were unable to learn the gesture template sequentially. Moreover, in the case of OC-SVM and SVDD, because the decision was based on average of the confidences across the windows of a gesture sequence, any misclassifications lowered the aggregate confidence score for target gestures. The HMM classifier yielded the second highest performance with an AUC of 0.96, along with SVDD. However, because training and tuning of the HMMs required more data, a frame increment of 2 ms was used during feature extraction for this scheme.

Feature Selection: During feature selection, both the sequential classifiers, HMM and DTW, converged quickly, with three (MFL, DAR & MSR) and two (MFL & DAR) features, respectively. By contrast, the time-independent classifiers, OC-SVM and, SVDD required a greater number of features to reach their best performance as seen in Fig. 2. For example, SVDD started with DASDV, DAR and TDPSD features but then added ZC & MFL features before converging. A similar trend can be seen with the OC-SVM classifier. This is likely because the sequential classifiers leveraged the temporal structure of the signals, and thus were more easily able to differentiate between target and non-target gestures. Also, the windowed classifiers naively assume that the class distributions are constant across the full gesture, requiring additional information to overcome the correspondingly reduced separability. This again demonstrates the power of the sequential classifiers for robust EMG-based gesture recognition applications.

To demonstrate the impact of the proposed AUC-based feature selection scheme, the AUC and EERs were compared with those obtained when selecting features based on the classical classification-error objective function. The results, presented in Table I, show that the AUC-based feature selection scheme outperforms the accuracy-based scheme for all classifiers. Although this is partially because the optimization criteria more closely aligns with the evaluation criteria, it also shows that the consideration of false positive and negatives in the feature selection process is able to reduce their impact during testing, and likely during real-world usage situations.

TABLE I: A comparison of gesture verification performance (AUC and EER) when comparing accuracy vs AUC-based feature selection. Note: CE is Classification Error

Classifier	SFS Criteria			
	CE		AUC	
	AUC	EER	AUC	EER
SVDD	0.80	0.24	0.96	0.11
OC-SVM	0.68	0.44	0.95	0.17
DTW	0.91	0.2	**0.98**	**0.04**
HMM	0.69	0.4	0.96	0.08

In considering the real-world application, a single-motion EMG-based gesture verification system (e.g. a mouse click or wake word), must employ a gesture that is robust to outlier motion artifacts and other gestures. The above results considered the use of index extension as the target motion, however, others could similarly be candidates. Consequently, a comparison of the different gestures was conducted, each time considering one gesture as the target, and all other gestures as extraneous motions. The results were computed using the DTW technique identified previously, and shown in Fig. 3. It can be seen that the index extension motion obtained the maximum AUC of 0.98. As in Fig. 3, it can also be noted that the index extension ROC starts with a TPR of 96% and 0% false positives, whereas the other gesture ROCs yield much lower TPRs for 0% FPR. Thus, a single-gesture application using a wrist-based EMG (with electrodes on the dorsal side) should consider index extension as the target class.

It should be noted that the proposed gesture verification system was implemented and tested on a limited number of gestures. These gestures represent common hand and finger movements, but do not cover the full range of pos-

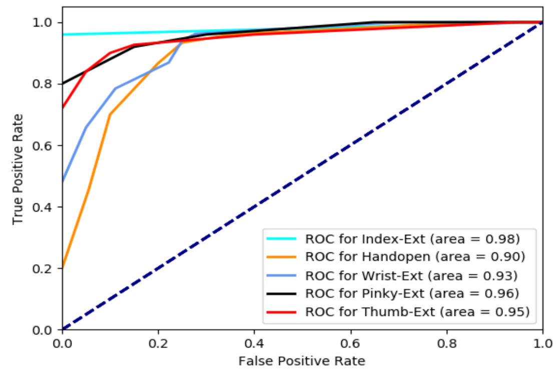

Fig. 3: ROC-AUC curves across different gestures using DTW.

sible movements. Thus, this study should be considered as preliminary work in this regard. Moreover, the performance of the system is compared among one-class classifiers, but future work should consider other classification techniques as well. Additionally, data from more subjects and more classes should be collected to further evaluate the robustness of the proposed system, and improve the resolution of the ROC curves.

IV. Conclusion

In this work, we proposed a robust gesture verification-based scheme for an EMG wake word. The verification process was carried out using both sequential and traditional classifiers and the performance was evaluated using ROC curves. Considering false positive and negative rates during the feature selection process improved the verification results. DTW was found to be more robust to outlier movements and yielded a high TPR, with a maximum AUC of 0.98 and an EER of 0.04. Index-extension was found to be most relevant gesture for single class based HCI applications. In the future, the performance of the proposed system could be improved by including IMU information along with EMG. Additionally, more repetitions of data, more subjects, and potentially leveraging deep learning techniques may further boost the system's performance.

References

[1] Q. Ding, Z. Li, X. Zhao, Y. Xiao, and J. Han, "Real-time myoelectric prosthetic-hand control to reject outlier motion interference using one-class classifier," in *32nd Youth Academic Annual Conference of Chinese Association of Automation*, pp. 96–101, IEEE, 2017.

[2] L. Bi, C. Guan, *et al.*, "A review on emg-based motor intention prediction of continuous human upper limb motion for human-robot collaboration," *Biomedical Signal Processing and Control*, vol. 51, pp. 113–127, 2019.

[3] C. Andronache, M. Negru, A. Neacsu, G. Cioroiu, A. Radoi, and C. Burileanu, "Towards extending real-time emg-based gesture recognition system," in *43rd International Conference on Telecommunications and Signal Processing*, pp. 301–304, IEEE, 2020.

[4] C. Savur and F. Sahin, "American sign language recognition system by using surface emg signal," in *International Conference on Systems, Man and Cybernetics (SMC)*, pp. 002872–002877, IEEE, 2016.

[5] J. Robertson, K. Englehart, and E. Scheme, "Effects of confidence-based rejection on usability and error in pattern recognition-based myoelectric control," *IEEE Journal of Biomedical and Health Informatics*, vol. 23, no. 5, pp. 2002–2008, 2018.

[6] E. Scheme, K. Englehart, and B. Hudgins, "Selective classification for improved robustness of myoelectric control under nonideal conditions," *IEEE Transactions on Biomedical Engineering*, vol. 58, no. 6, pp. 1698–1705, 2011.

[7] M. Tavakoli, C. Benussi, and J. L. Lourenco, "Single channel surface emg control of advanced prosthetic hands: A simple, low cost and efficient approach," *ESWA*, vol. 79, pp. 322–332, 2017.

[8] M. Tavakoli, C. Benussi, P. A. Lopes, L. B. Osorio, and A. T. de Almeida, "Robust hand gesture recognition with a double channel surface emg wearable armband and svm classifier," *Biomedical Signal Processing and Control*, vol. 46, pp. 121–130, 2018.

[9] T.-Y. Pan, W.-L. Tsai, C.-Y. Chang, C.-W. Yeh, and M.-C. Hu, "A hierarchical hand gesture recognition framework for sports referee training-based emg and accelerometer sensors," *IEEE Transactions on Cybernetics*, 2020.

[10] E. Campbell, A. Phinyomark, and E. Scheme, "Differences in perspective on inertial measurement unit sensor integration in myoelectric control," *arXiv preprint arXiv:2003.03424*, 2020.

[11] M. F. Wahid, R. Tafreshi, and R. Langari, "A multi-window majority voting strategy to improve hand gesture recognition accuracies using electromyography signal," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 28, no. 2, pp. 427–436, 2019.

[12] K. Englehart and B. Hudgins, "A robust, real-time control scheme for multifunction myoelectric control," *IEEE transactions on biomedical engineering*, vol. 50, no. 7, pp. 848–854, 2003.

[13] M. Rossi, S. Benatti, E. Farella, and L. Benini, "Hybrid emg classifier based on hmm and svm for hand gesture recognition in prosthetics," in *IEEE International Conference on Industrial Technology*, pp. 1700–1705, IEEE, 2015.

[14] M. W. Kadous *et al.*, *Temporal classification: Extending the classification paradigm to multivariate time series*. University of New South Wales Kensington, 2002.

[15] J. Chang, A. Phinyomark, S. Bateman, and E. Scheme, "Wearable emg-based gesture recognition systems during activities of daily living: An exploratory study," in *42nd Annual International Conference of the IEEE EMBC*, pp. 3448–3451, IEEE, 2020.

[16] S. Hooman, "Automatic activity detection in noisy signals with hilbert transfrom." (https://www.mathworks.com/matlabcentral/fileexchange/46139-automatic-activity-detection-in-noisy// -signals-with-hilbert-transfrom). MATLAB Central File Exchange, Accessed: 2020-09-10.

[17] S. S. Jambhale and A. Khaparde, "Gesture recognition using dtw & piecewise dtw," in *International Conference on Electronics and Communication Systems*, pp. 1–5, IEEE, 2014.

[18] A. Phinyomark, R. N. Khushaba, E. Ibáñez-Marcelo, A. Patania, E. Scheme, and G. Petri, "Navigating features: a topologically informed chart of electromyographic features space," *Journal of The Royal Society Interface*, vol. 14, no. 137, p. 20170734, 2017.

[19] J. Liu, "Feature dimensionality reduction for myoelectric pattern recognition: A comparison study of feature selection and feature projection methods," *Medical engineering & physics*, vol. 36, no. 12, pp. 1716–1720, 2014.

[20] R. Menon, G. Di Caterina, H. Lakany, L. Petropoulakis, B. A. Conway, and J. J. Soraghan, "Automatic misclassification rejection for lda classifier using roc curves," in *37th Annual International Conference of the IEEE EMBC*, pp. 482–485, IEEE, 2015.

[21] V. Mingote, A. Miguel, A. Ortega, and E. Lleida, "Optimization of the area under the roc curve using neural network supervectors for text-dependent speaker verification," *Computer Speech & Language*, vol. 63, p. 101078, 2020.

[22] Y.-H. Liu and H.-P. Huang, "Towards a high-stability emg recognition system for prosthesis control: A one-class classification based non-target emg pattern filtering scheme," in *IEEE SMC*, pp. 4752–4757, IEEE, 2009.

[23] S. M. Erfani, S. Rajasegarar, S. Karunasekera, and C. Leckie, "High-dimensional and large-scale anomaly detection using a linear one-class svm with deep learning," *Pattern Recognition*, vol. 58, pp. 121–134, 2016.

[24] X. Zhang, X. Chen, Y. Li, V. Lantz, K. Wang, and J. Yang, "A framework for hand gesture recognition based on accelerometer and emg sensors," *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, vol. 41, no. 6, pp. 1064–1076, 2011.

[25] P. Kumar, H. Gauba, P. P. Roy, and D. P. Dogra, "Coupled hmm-based multi-sensor data fusion for sign language recognition," *Pattern Recognition Letters*, vol. 86, pp. 1–8, 2017.

[26] M. Shah, J. Grabocka, N. Schilling, M. Wistuba, and L. Schmidt-Thieme, "Learning dtw-shapelets for time-series classification," in *Proceedings of the 3rd IKDD Conference on Data Science*, pp. 1–8, 2016.