

Chapter 9

Enhanced Real-Time Raw sEMG Signal Classification Through Bypass of Manual Feature Engineering and Extraction



Shriniket Subramanian, Narayanan Aravind, and Kane Ng Zheng Kang

Abstract The use of surface electromyographic (sEMG) signals obtained from the residual stump of an arm has been embraced for the purpose of developing prosthetics and bionic limbs. These sEMG signals are processed and classified into gestures. As the signals are stochastic and non-stationary, classification of these signals is a complex task. Conventional methods of feature recognition require handcrafted, manual feature extraction followed by classification. This paper explores the bypass of this preliminary step of feature extraction in order to reduce lag and computing complexity. Through data collected with the aid of an 8-channel Myoband developed by Thalmic Labs, our group trained and implemented three main raw data classifier models for real-time sEMG signal classification of nine gesture classes. The three methods presented are a raw naïve Bayes model implementation (81.55%), a novel raw sEMG ConvNet (96.88%), and finally, a novel ConvXGB model implementation (90.62%). The ConvXGB implemented is an adaptation of the new deep learning model developed in 2020. The significance of these novel models developed is the reduction in computational resources required for real-time classification of signals, leading to a reduction in lag time in real-time prosthetic devices, and the increased viability of embedded-systems classifications. The findings have the potential to also pioneer a new generation of more responsive prosthetic devices that are also easy to control.

Keywords Multi-channel time series classification · sEMG signals · Prosthetic devices · Bypassing feature extraction · ConvXGB model · GRU-FCN model · Naïve Bayes model

S. Subramanian (✉) · N. Aravind · K. N. Z. Kang
NUS High School of Math and Science, Singapore, Singapore
e-mail: H1610129@nushigh.edu.sg

9.1 Introduction

The loss of a forearm is a traumatic experience and severely impairs an individual's quality of life. Every year, around 50,000 people undergo amputation in the USA alone, and it is unimaginable as to how disrupted their life will become post-procedure. As such, biomedical research into prosthetics has gained traction over the past few years, and many prosthetics have been developed to attempt to achieve comparable degrees of freedom to the actual human hand. However, it has also increasingly come into question as to how natural control of the prosthetics is to be achieved. On top of degrees of freedom, main design goals of prosthetics would be to be lag free and be lightweight.

Many state-of-the-art prosthetics achieve control by detecting impulses in the muscles in the user's residual limb using mounted differential electrodes. These differential electrodes are invasive and need to be inserted under the skin of the user. While these electrodes provide more precise signals of greater amplitude and frequency, they are also inconvenient, intrusive, and introduce a chance of complications. The Thalmic Labs Myoband has thus been seen as a potential game changer by researchers for this purpose. The bands are worn on the subject, and collect data using non-invasive surface electrodes, that do not require prior preparation of the subject. The bands detect 8 channel surface electromyographic (sEMG) data from the muscles and wirelessly transmit the data to any device. These bands simplify the data collection process and are hence embraced by researchers. We have adopted this technology for acquiring data in this project. Electromyographic signals are complicated signals influenced by a range of physiological and anatomical properties such as muscle fatigue, muscle size, fiber type, orientation, electrode placement, impedance of skin, and more. While it is more convenient, sEMG signals are less precise and may carry more crosstalk than EMG signals from differential electrodes, the signal processing is further complicated. Being non-stationary stochastic signals with multiple channels of input [1, 2] and over 100 timesteps per gesture [3], sEMG signals tended to be too complicated for models to make sense of directly [4].

Sequence-based models such as (K-nearest neighbors) KNNs that can work with one input channel become too computationally expensive and impractical to implement in real time as the lag has to be kept within 200–300 ms [5] for the system to feel responsive. As such, manual feature extraction is often carried out to reduce the dimensionality of the data. These features are then fed into a classifier. Over time, accurate feature extraction from EMG signals has become the focus for accurate classification in both real-time and offline systems and is essential to the motion command identification. The performance of traditional feature-based methods depends on the quality of handcrafted features. However, it is difficult to design good features to capture all intrinsic properties of EMG to classify the various gestures accurately due to the signal's non-stationary nature. Thus, models are additionally limited by the feature extraction capabilities available. With manual feature extraction, the classification accuracy of nine classes in real time as of 2020 is currently 92.1% [6].

With current models having to extract over 50 features [6] in order to reliably classify nine classes of data, the lag time due to feature extraction is simply too much, and greatly exceeds the required 200–300 ms for a generally responsive prosthetic device. Inspired by recent advances in image classification, where automatic feature extraction occurs, this project investigates the feasibility of implementing a real-time classification system without manual feature extraction. Bypassing the computationally expensive feature extraction step allows for the lag time to be smaller. Automatic feature extraction might allow the model to be more accurate as well as the optimal feature set for each gesture is not fixed and can be learnt.

We thus developed several different artificial intelligence (AI) models capable of classifying this data and relaying the results to a mechanical hand.

9.2 Methodology (Data Collection)

One of the authors for the study was recruited for surface electromyography (sEMG) data collection.

Muscular activity was gathered using 2 Thalmic Myobands on the right forearm, each having eight non-invasive sEMG electrodes. This paper will refer to the band closer to the wrist as the lower band and the one further from the wrist as the upper band. Data was collected wearing two bands at the same time to reduce time spent collecting data, while ensuring both upper and lower placement data are not affected by the small inconsistencies between every repetition of each gesture. As a result, 8-channel data was collected for every movement at the maximal 200 Hz allowed by the Myobands. Nine classes of gestures were collected, with 30 repetitions in each class (see Fig. 9.1) [7, 8].

The Thalmic Myoband collects 200 data points every second, each expressed as an array of eight elements corresponding to the signal detected at each of the eight electrodes of the band. The data will then be stored as a text file with eight data points separated by spaces in each line, for each timestamp.

Two separate datasets were collected on different dates, one collecting 1 s of data per gesture (200 lines) and another collecting 2 s of data per gesture (400 lines). We will refer to the two datasets later on as the 200-line dataset and the 400-line dataset for clearer understanding [8].

The timer begins when the subject begins to form the gesture, from a neutral rest position. This method of data acquisition was practiced in place of recording 1 s of holding the gesture, because the nine gestures involve closely related muscles to hold the gesture, but obviously activate different muscles to perform the gesture from rest. For example, the index and middle gestures have similar sEMG oscilloscope traces and have roughly the same muscles activated, but the difference lies in the potentials fired at the start of the gesture. Rest was given as requested by the subject before the recording of each action to avoid fatigue.

The nine gestures were chosen primarily due to their prevalence in usage in both common literature and in the real world. Note that some of the gestures that we

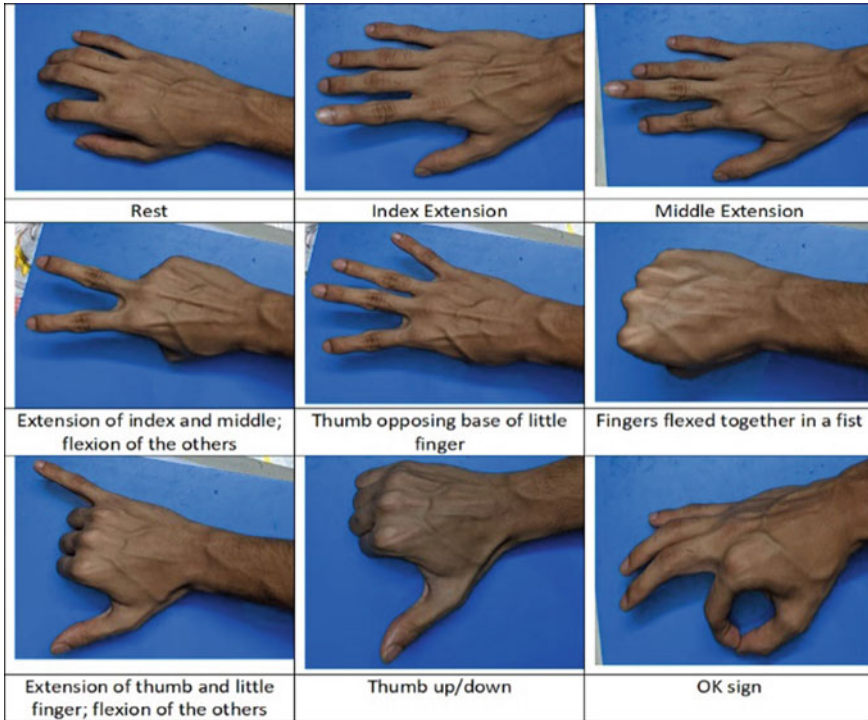


Fig. 9.1 Images of nine performed gestures

trained the classifier with are named after numbers, where the gesture represents the one-handed representation of a one-digit number. For example, to show the number four, all fingers bar the thumb are extended.

Four other plausible gestures such as large diameter grip, four gesture, three gesture, and seven gesture were trialed to form a 13-class dataset but were omitted due to lack of distinct data obtained from the gestures. Instead, a net 9-class dataset was used to train all the classifiers instead. In later parts, we will refer mainly to this 9-class dataset, and in certain parts will refer to the 13-class dataset as the **trial** classification dataset.

The data collected from each gesture was written into a text file by the Python program and saved separately. Following which, data was read from all the text files to save in arrays and NumPy objects for manipulation and training. Using the data collected, multiple datasets with variable number and combination of classes were created to test the efficacy of certain elements in the data that can contribute to the classification performance.

The data collected is represented as an (1, 8, 200) array for the 200-line dataset and (1, 8, 400) for the 400-line dataset. This arrangement of data is manipulated into various data formats for loading into the various models. Some strategies used by

the team in manipulating the data are elaborated below. A form of data manipulation we trialed was filtering.

When passing through various tissues, EMG signals acquire noise. Noise is the unwanted electrical signal in an EMG signal. Noise is primarily from the ambient electromagnetic radiation, inherent noise in electrical devices and electrode impedance. Motion artifacts are introduced into the signal when the electrodes move relative to each other when a muscle flexes prominently. This noise and artifacts in the signal are serious issues to be considered, as this will adversely affect the quality of the signal. As these factors cannot be remedied easily without upgrading to a more complex electrode set up, we have decided to implement noise removal systems. As such, the team designed and adapted a Butterworth bandpass filter for this purpose to filter out noise that is not in the dominant frequency range of the band. We identified that the dominant frequency spectrum of the signals from the band was between 10 and 500 Hz through literature data [9]. A graph showing the difference between the noisy signal and the filtered signal is shown (Fig. 9.2). The impact of this designed filter will be discussed later.

For the naïve Bayes model which could not take the raw data in its multi-dimensional state as an input, the team carried out a flattening of the 3D (1 * 8 * 200) array into a 2D (1 * 1600) array.

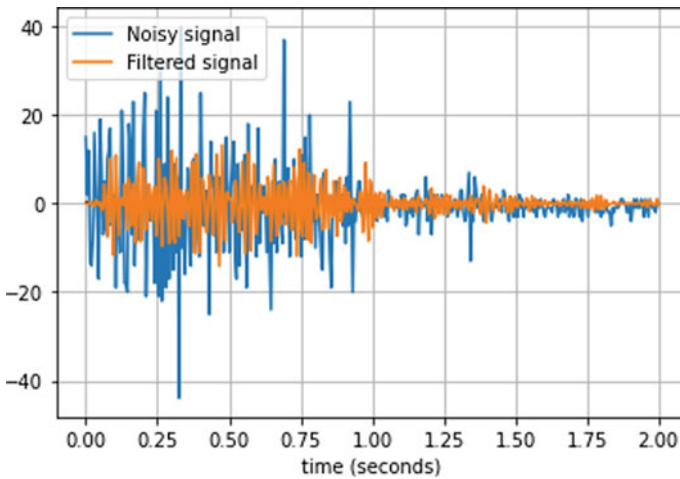


Fig. 9.2 sEMG trace of noisy versus filtered signal

9.3 Models Developed

9.3.1 *Naïve-Bayes Classifier*

The naïve Bayes model was trained using the 400-line dataset because it performed better with this dataset as compared to the other 200-line dataset.

We use this model, as it is a very commonly used model, albeit with extracted features. We can thus compare other solutions with this approach to check how much better other approaches fare when pitched against this model.

Raw data in the shape of $(x, 8, 400)$ was flattened to form an $x * 3200$ array and input into a naïve Bayes model developed with the aid of Python's scikit-learn package. The naïve Bayes model was also optimized using grid search to find the optimal smoothing variable for best observed accuracy.

It is worth noting that such flattening of data, reduces the interdependency between the channels therefore potentially resulting in frequent misclassifications as the models will not be learning any properties specific to the interaction between the eight distinct channels of the Myoband.

This model developed is not state of the art; it was simply an elementary approach by which we can gauge the performance of the models in the absence of feature extraction. While such an approach compromises accuracy, it is fast.

Regardless, accuracy is still an important factor in determining the efficacy of the model, and it is not desirable to have the prosthetic device perform some other movement when given a gesture by the user.

Hence, we describe in later parts, other solutions to match, if not exceed the accuracies that are achieved by the current line of state-of-the-art models which practice feature extraction.

9.3.2 *Gated-Recurrent-Unit, Fully-Connected-Network*

A typical GRU cell consists of two gates: reset gate and update gate. The GRU FCN model was chosen due to GRUs being able to combat vanishing or exploding gradient problems due to its reset and update gates [10] and being easier to implement than LSTMs.

For the GRU, raw data was simply loaded as is with a data shape of $(x, 8, 200)$ where x is the number of samples available. This model turned out to be competitive with state-of-the-art models, while showing minimal and consistent processing times. The model was also tested with 400-line data and 200-line data and was found to perform better with the 200-line dataset. This finding is important and is analyzed in later parts.

9.3.3 Modified ConvXGB Classifier

As an outstanding classifier and feature extractor, CNNs have achieved great success, especially in the field of image recognition [2]. Inspired by this advantage, we also exploit the GRU-FCN above as the feature extractor, and a new classifier XGBoost to replace the dense soft-max classifier. XGBoost is an efficient implementation of gradient boosted decision trees (GBDT) due to its block structure to support the parallelization of tree construction. In GBDT, gradient boosting refers to a kind of ensemble technique creating new models to predict the residuals or errors of prior models and making the final decision by summing up the predictions from all models. This model was chosen to be implemented due to its promise of strong feature extraction and accurate classification due to the blend of CNN and XGBoost. As such, we have implemented a novel GRU-CNN-XGB. In 2020, the general ConvXGB model was introduced [9]. Our implementation also includes a GRU for feature extraction and is the first of such models to be applied in this field, to our best knowledge [11].

9.4 Results and Discussion

Our team obtained four main datasets as a result of the manipulations described in data collection segment. There is a 200-line dataset and a 400-line dataset which are raw datasets derived directly from the Myoband. Secondly, through the designed Butterworth filter, we derived two filtered datasets for the 200-line and the 400-line, respectively.

We tried to optimize the model performance by identifying the dataset that led to the best accuracy for each of the models through trial and error.

It was found that the models performed better with the unfiltered datasets. It was further observed that the naïve Bayes model performed better with the 400-line dataset while the GRU and ConvXGB performed better with the 200-line dataset. As such, models were ultimately trained with the dataset they performed best with, and we obtained the following results.

The graph (Fig. 9.3) shows the highest accuracies that each of the three models described in our project has managed to yield. The orange line shows the current state-of-the-art accuracy, and it is worth noting that the GRU-FCN surpassed this standard while the ConvXGB is close to matching it. Unsurprisingly, the naïve Bayes did not perform as well.

The accuracy of the classical naïve Bayes model was 81.55%. The GRU-FCN and ConvXGB models were able to achieve accuracies that are well above that of the naïve Bayes model and very competitive with the conventional accuracy of 92.1% for nine classes of gestures. The results show that bypassing feature extraction is possible without compromising the accuracy classification. Further, the GRU-FCN and ConvXGB models are competitive alternatives that can be considered when

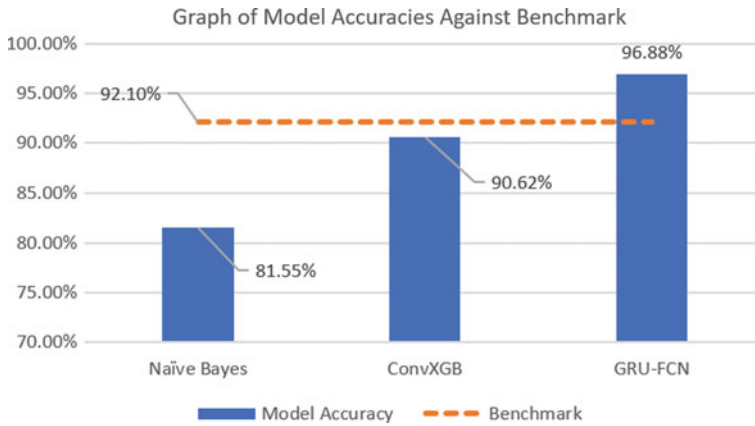


Fig. 9.3 Accuracies of models against testing data

designing prosthetic devices seeing as they are both faster and as accurate as state-of-the-art alternatives.

In addition to the above results, the data supports the conclusion that the lower placement of the band allows for more distinct data to be collected per class, leading to better classification performance. This result is consistently reflected in both the 9 class and **trial** 13 class classification (see Fig. 9.4) using the GRU-FCN model. This result can be explained biologically, with greater number of finer muscles fibers present as we progress down the forearm toward the wrist, which contribute to more distinct signals for each gesture when the band is placed lower along the forearm.

We also mentioned briefly earlier that the GRU-FCN performs better when the 200-line dataset is used. (see Fig. 9.5) a plausible explanation involves data density.

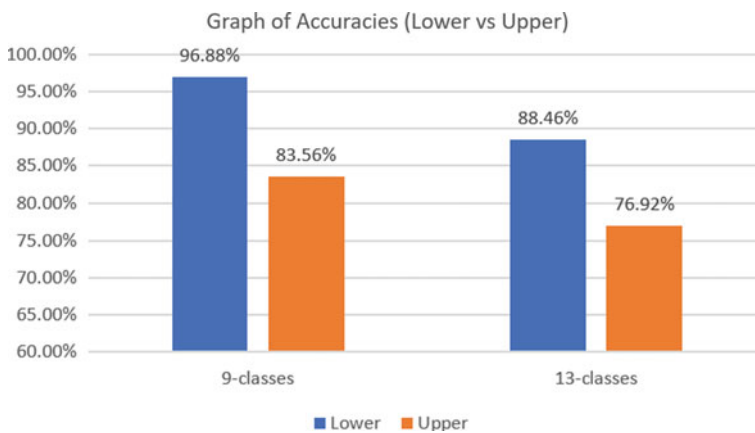


Fig. 9.4 Comparison of model performance between lower and upper band data for different training classes

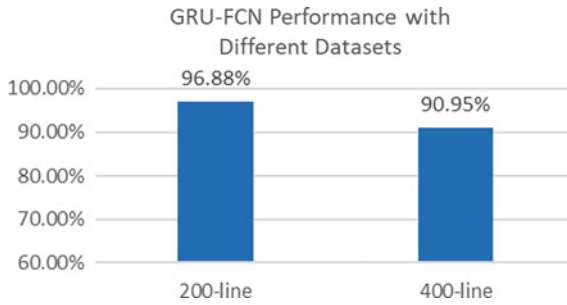


Fig. 9.5 GRU-FCN performance versus dataset types

The 400-line dataset involves a longer sequence to consider which results in more points being stored in short-term memory, leading to a greater chance of false associations and patterns being discovered, which reduces classification accuracy. This finding is useful and can be further explored in an attempt to boost model accuracies.

It is worth noting that the models developed did not perform too well with the filtered dataset. This was surprising as the filter that was developed with the goal of improving model accuracies; however, this was not observed and instead the models performed rather sub optimally instead in real time. After careful analysis of other studies involving sEMG signals and the physiological aspect of the electrodes, we have a plausible explanation. EMG electrodes are placed in the skin, closer to the muscles unlike sEMG electrodes which are on the surface. This means that the motion artifacts for EMG signals generally remain the same across gestures, while they differ between gestures sEMG data. That is to say, the noise generated is unique to each gesture and can enhance classification. Filtering reduces these motion artifacts and thus leading to a loss of important data. This results in a lowered model accuracy when fed filtered data. On the flip side, this means that using sEMG technology requires even lower computation, as filtering is not required.

Finally, yet most importantly, in terms of lag time, our models performed exceptionally well with the classifiers showing a remarkable 172.499 ms (± 35.36977 ms) average lag time. Our models thus successfully fall under 300 ms of lag standard and thus fulfill the original objective of the implementation and bypass of the feature extraction step in order to improve lag times. Laboratory studies on pattern recognition usually report 300 ms of response lag [12]. Hence, the models developed are not only of comparable accuracies to existing alternatives but are also significantly faster. This fulfills the engineering goal of designing faster, yet accurate classifier models such that the responsiveness of prosthetic devices in real time can be improved [13].

9.5 Implications and Conclusions

This paper thus contributes three plausible novel implementations of raw sEMG signal classification models without feature extraction through exploration of models such as naïve Bayes, GRU-FCN, and GRU-ConvXGB. The GRU-ConvXGB and GRU-FCN models developed are also of comparable accuracy to state-of-the-art models that use feature extraction, and thus are equally as competitive and lucrative for prosthetics developers to consider.

The benefits of the models within this research paper include but are not limited to lightened computing complexity and reduced signal processing times when used to classify signals in real time. We highlight that reduced processing times are integral when it comes to mimicking the real hand with prosthetics. It would allow faster reaction of the prosthetic itself and thus would make it more cohesive and authentic. The novel approach explored by this paper, to the best of our knowledge, is one of the first of such endeavors in reducing lag times and can be further explored in future.

9.6 Limitations of Findings

However, there are certain limitations to the methods proposed and implemented. The data collected in this report is from an able-bodied participant. However, prosthetics largely concern users without a physical hand. As such, data derived from the residual stump of a hand—needs to be considered for a more realistic case scenario. Further, these gestures will need to be tried among a few human subjects to determine whether the EMG trace observed is user-specific or generally observed for any subject.

Further data can be collected from various subjects for training of classifiers to increase the generalizing abilities of the models as well.

Moreover, it was observed that there was some deviance in terms of data collected between different bands. This is something worth noting especially because most would assume that the data would be consistent across the same band product. As such, models may have to be retrained depending on the band being used for the prosthetic devices.

9.7 Future Work

Further optimization can be carried out to render the models to be more accurate as accuracy is highly important in real-time prosthetic devices. We believe that dataset sizes could be boosted to achieve this purpose such that the model can learn more detailed features of the data collected.

Further, we also recommend exploring variable timesteps of data recording. As mentioned earlier, we found deviations in the GRU-FCN's performance with 2 and

1-s timestep data. As such, this deviation can be explored further, and we could thus possibly improve the accuracy of the model further, by considering other possible timesteps of data collection.

Appendix

	Accuracies		
	GRU-FCN (%)	Naïve Bayes (%)	ConvXGB (%)
200 lines (1 s of data) (9 classes)	96.88	81.55	90.62

GRU-FCN	Lower		Upper	
	Loss	Accuracy (%)	Loss	Accuracy (%)
9 classes	0.4029	96.88	0.7225	83.56
13 classes	0.3854	88.46	1.1370	76.92

References

1. Lin, H. (2019, November 1). A multivariate time series classification method based on self-attention. https://doi.org/10.1007/978-981-15-3308-2_54
2. Zheng, Y., Liu, Q., Chen, E., Ge, Y., & Zhao, J. L. (2014, June). Time series classification using multi-channels deep convolutional neural networks. In *International conference on web-age information management* (pp. 298–310). Springer, Cham. http://staff.ustc.edu.cn/~cheneh/paper_pdf/2014/Yi-Zheng-WAIM2014.pdf
3. Reaz, M. B. I., Hussain, M. S., & Mohd-Yasin, F. (2006). Techniques of EMG signal analysis: Detection, processing, classification and applications. *Biological Procedures Online*, 8(1), 11–35.
4. Hatami, N., Gavet, Y., & Debayle, J. (2018, April). Classification of time-series images using deep convolutional neural networks. In *Tenth international conference on machine vision (ICMV 2017)* (Vol. 10696, p. 106960Y). International Society for Optics and Photonics.
5. Singh, R. E., Iqbal, K., Rasool, G., White, G., Wingate, S., & Hutchinson, H. Muscle synergy extraction from EMG (electro myogram) data: A comparison of algorithms (PCA, NNMF, pICA).
6. Scheme, K. E., Nurhazimah Nazmi, M. A., Tenore, A. R. F. V., Zecca, S. M., Guo, M. P. S., Phinyomark, P. P. A., & Kuiken, G. L. T. (2019, November 21). Evaluation of surface EMG-based recognition algorithms for decoding hand movements. Retrieved May, 2020, from <https://doi.org/10.1007/s11517-019-02073-z>
7. Balandinodidonato (n.d.). Balandinodidonato/MyoToolkit. Retrieved September, 2020, from <https://github.com/balandinodidonato/MyoToolkit/blob/master/SoftwareforThalnic'sMyoarmband.md>
8. Eliotxu (n.d.). Eliotxu/pyMyo. Retrieved August, 2020, from <https://github.com/eliotxu/pyMyo>

9. Englehart, K., & Hudgins, B. (2003). A robust, real-time control scheme for multifunction myoelectric control. *IEEE Transaction on Biomedical Engineering*, *50*, 848–854.
10. Chung, J., Gulcehre, C., Cho, K., & Bengio, Y. (2014). Empirical evaluation of gated recurrent neural networks on sequence modeling. [arXiv:1412.3555](https://arxiv.org/abs/1412.3555)
11. Thongsuwan, S., Jaiyen, S., Padcharoen, A., & Agarwal, P. (2020, August 02). ConvXGB: A new deep learning model for classification problems based on CNN and XGBoost. Retrieved December, 2020, from <https://www.sciencedirect.com/science/article/pii/S1738573319308587>
12. Huang, Y., Englehart, K. B., Hudgins, B., & Chan, A. D. C. (2005). A Gaussian mixture model-based classification scheme for myoelectric control of powered upper limb prostheses. *IEEE Transactions on Biomedical Engineering*, *52*, 1801–1811.
13. Chan, A. D. C., & Englehart, K. B. (2005). Continuous myoelectric control for powered prostheses using hidden Markov models. *IEEE Transactions on Biomedical Engineering*, *52*, 121–124.