Ben McAteer

0029592670

BME 511 Final Project

Hand Gesture Prediction Using Forearm EMG Data

Abstract

Approximately 5000-10000 individuals require a hand amputation every year. The significant loss in functionally often causes the patient to feel like a nuisance to society, hurting their emotional well-being. The goal of this project is to create a gesture prediction model using electromyography (EMG) data that could be implemented into a low-cost hand prosthetic. The EMG data consisted of eight unmarked sensors and data was recorded for six trials of twenty seconds at a 200Hz sampling rate. The key innovation of this project involved finding the main differential characteristic to restructure the data before applying the model. The key differential characteristic was identified as the mean of the absolute value of each sensor. The data was defined by this characteristic in 1.00, 0.5, 0.25, 0.05, and 0.005 second time windows that would be used for the model. After restructuring the data, five models are applied in a 70% train and 30% Test split. A peripheral objective became determining which type of model best predicts EMG data. The five models are: Multi-Layer Perceptron (MLP), Linear, Naïve Bayes Gaussian (NBG), K-Nearest Neighbor (KNN), and Principal Component Analysis (PCA). After applying each model, the MLP, Linear, and KNN models showed accuracies above 95% with only 1/20th of a second of data. These findings show that the four hand gestures can be predicted in a near instantaneous 1/20th of a second and therefore could allow for further gesture control devices. Future work would entail adding more gestures and testing these same models.

Introduction

Every year, 5000-10000 individuals require their hand or wrist to be amputated [1]. The functional impact losing a hand has on a patient is immense, and usually severely affects their emotional quality of life too. They lose power to perform simple tasks like getting dressed or cooking causing them to feel like a nuisance to people around them. The problem is low-cost hand prosthetics do not offer an acceptable level of function, and other hand prosthetics may use implantable technologies that are difficult to maintain, semi-permanent, and cost an average of $18,703 without insurance [2]. Therefore, there is a need to create a hand prosthetic model that can provide function to the user while remaining non-invasive and low cost. The solution involves using electromyography (EMG) to read forearm muscle activation to predict how the missing hand would move. The scope for this project focuses on the "gesture" prediction software that would be used to predict hand movement. The data used to create a gesture prediction model was recorded from Kirill Yashuk using a non-invasive eight-sensor EMG [3]. The exact "MYO Armband" EMG device is seen in Figure 1.

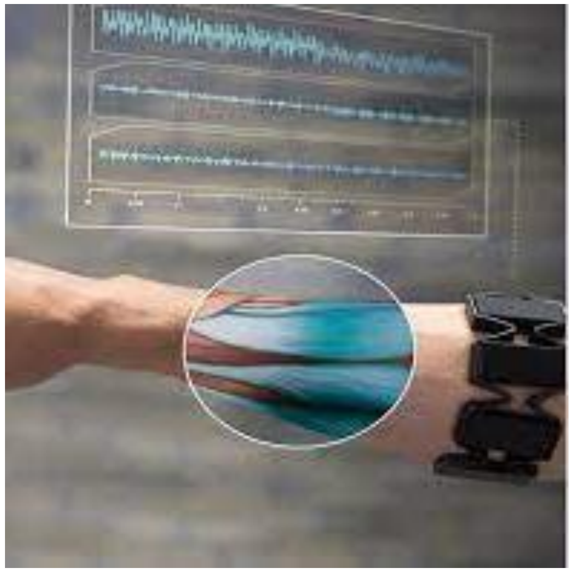

Figure 1: 8-Sensor Forearm EMG [4]

There are four recorded gestures: “Rock”, “Paper”, “Scissors”, and an “OK” gesture where the user touches their pointer finger to their thumb and spreads the other three fingers out. Each gesture was flexed for approximately 15 seconds for eight trials, totaling to about 120 seconds of EMG data per gesture. The specific goal set for this project was to create a model that predicted the gesture with 99% accuracy with only one second of data input. This was chosen as it could be translated to the low-cost prosthetic. A user would flex the gesture they’d want to make, and with a one second delay, the prosthetic would perform the gesture. As the project continued, a secondary goal was created to determine the best type of model to predict EMG data.

Methods

As mentioned previously, the dataset came in with approximately 120 seconds of raw EMG data for four hand gestures. Given a 200Hz sampling rate, this led to about 24,000 data points per gesture. Each gesture had its own CSV file organized by readings from each of the eight sensors, for eight trials, with the final column classifying what gesture was truly being performed. This created a 65-column matrix with each row representing EMG amplitude in mV per sample depicted in Table 1. EMG data tends to have a mean value of zero, a frequency between 50-150Hz, and shows muscle activation via increased peak to peak values and variance [5]. These characteristics correspond to biological processes within the muscles and are the same on all humans. Although the mean value stays near zero, if the absolute value of the data is taken, then the resulting mean value increases tremendously when muscles become active. These characteristics that define muscle activity are used to shape and define the models. Data was collected only on one person and in a single session, eliminating variance from outside sources such as switching users or the armband electrodes shifting locations on the forearm. It should be noted that although reducing the number of variables during testing side decreases variance, it creates a large bias that would need to be addressed before prosthetic implementation can occur. The EMG sensors used were very high-end and yielded almost no movement artifacts or other noise in the collected data. This experiment created raw data containing very low variance with itself between each test and gesture, leading to an optimal situation to create a gesture predicting model. Unfortunately, the sensor location on the subject’s arm was not recorded which shunted any attempt at creating reference electrodes from expected inactive muscles per gesture.

Table 1: Raw Data Layout for One Gesture

	Trial 1	Trial 2	Trial 3	Trial 4	Trial 5	Trial 6	Trial 7	Trial 8	
Voltage Reading 200Hz	[8sensors]	[8sensors]	[8sensors]	[8sensors]	[8sensors]	[8sensors]	[8sensors]	[8sensors]	[GESTURE_CLASS]
1									
2									
3									
...									

Table 2: Method Steps and Descriptions

Description of the Signal Processing Step	**Rationale**	**Inputs and Outputs of the Steps**
Data Restructuring	The data needed to be restructured to begin comparing each gesture. The most successful method was to concatenate each trial together. This was done by selecting each eight sensors per trial, then flattening the data. Although the variations of each trial were lost in the process, the restructuring allowed for the comparison of each gesture’s characteristics on a sensor-to-sensor basis as seen in Figure 2. Without this step, there is would be no way of predicting hand gestures using only eight sensors. The model would be taking all eight trials into account, leading to false results.	The data began with each gesture having 64 columns split into eight trial sets of eight sensors and their EMG data as the rows. After restructuring, each gesture had eight columns total representing the entire raw 120 seconds of EMG data per sensor. The reduction in dimensions allowed the observing of sensor-to-sensor changes per gesture as seen in Figure 2
Key Characteristic Selection & Time-Windowing	After restructuring, the key characteristic for the model were chosen from the background research and then tested. The key characteristics were variance per sensor, variance between each sensor, the absolute value of the mean sensor value, and the mean top 10% peak values per sensor. These characteristics were calculated given an explicit amount of datapoints, creating precise time windowing. One second of data was used to calculate each of these characteristics and applied to the model. This process was repeated for shorter time windows to visualize the time-accuracy tradeoff of each model. Without this step, there	Using the restructured data and a sampling rate of 200Hz, each characteristic was calculated from the previous 200 data points-creating one value for one second of data. The values were different as each characteristic is calculated differently. These values are ultimately what is input into the model to predict gestures. This process was repeat with 100, 50, 10, and 1 datapoint(s) to create a model that works with only 0.50, 0.25, 0.05, and 0.005 seconds of data. These repeats show the visualization of speed-to-accuracy tradeoff in order

	would be no way to test a certain time-point of data, such as one second of data. Instead, the model would be predicting the gesture off a single datapoint representing 1/200th of a second each time. This step allowed for a reduction of variance by using a characteristic instead of all the raw data, an increase in bias, and a smoothing effect of the data to help the gesture prediction.	to help pick the fastest data acquisition possible for the model. Working with minimal data while remaining accurate is the final goal.
Model Application	All different gestures' data were concatenated with its gesture classifier and then randomly sorted into a 70% train 30% test split. The five models tested were a Multi-layer Perceptron (MLP), Linear, Naïve Bayes Gaussian (NBG), PCA on a Naïve Bayes Gaussian (referred to just as PCA), and a K-Nearest Neighbors. These were picked given relevant background research [6][7][8]. These models are the whole premise of the project and predict the hand gestures using the key characteristics.	The four key characteristics from the previous step were z-scaled and applied to each model for each of the five time-windows for a total of 100 tests (4characters x 5models x 5time-windows). The training group is used to fit the model and the model is tested on the test group. Both the accuracy of the training and testing groups were calculated. The very best combination of characteristic and model type were recorded.
Statistical Representation	After the best model was chosen, a confusion matrix was created to show where the model excels and where the model struggles to predict gestures. This helps with future work in improving the model and applying more preprocessing steps.	The very best model and key characteristic combination is applied to a confusion matrix. This matrix spreads the predicted gestures by the actual gestures. It shows where false predictions occur and gives insight as to how the model can be improved.

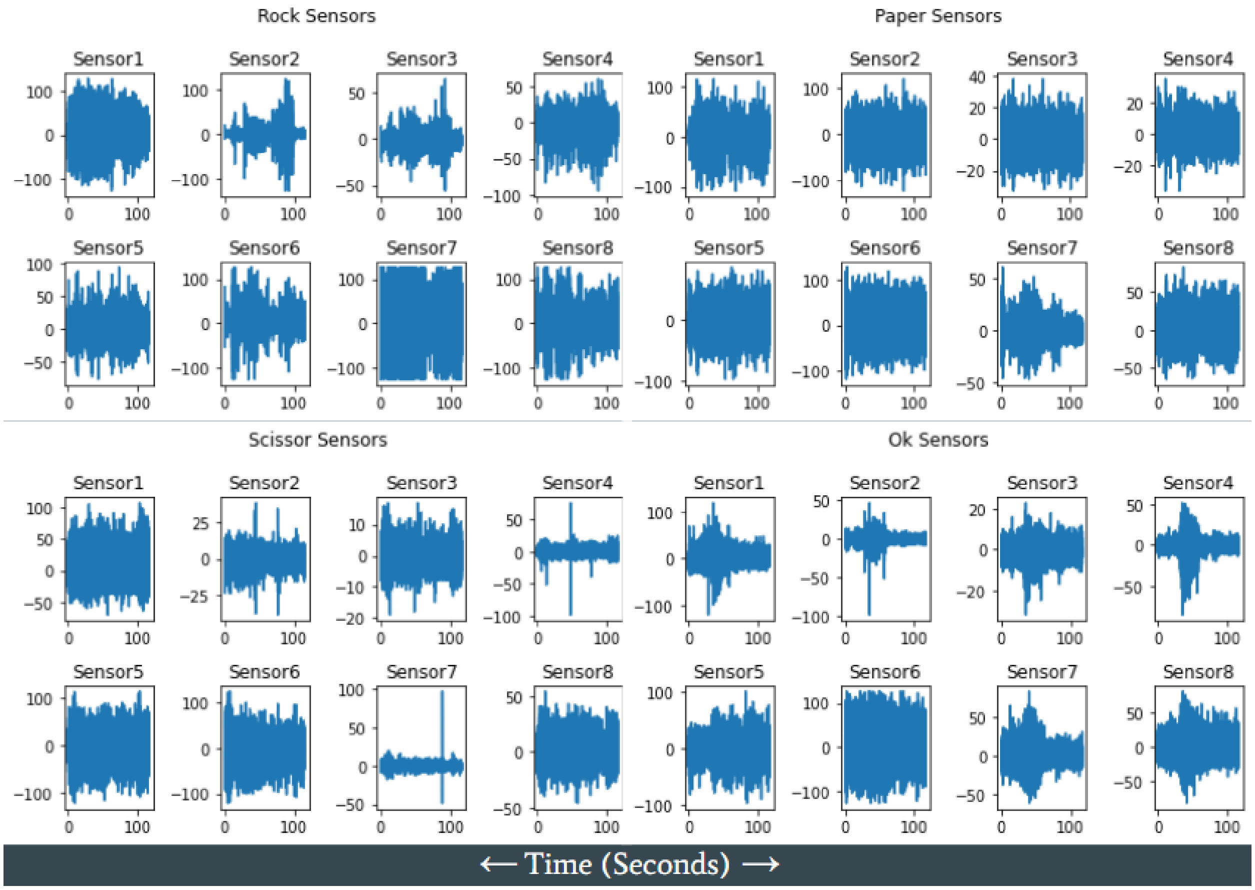

Figure 2: Restructured EMG Data Allowing for Deeper Observations

To create a model predicting hand gestures, data restructuring, key characteristic selection, and model application were applied as seen in Table 2. Each step fed into the next and ultimately allowed for 100 simulations to be ran and the best model to be chosen.

Results

As seen in Figure 2, there are some major differences in sensors depending on the gesture held. This alone serves as a proof of concept that predicting hand gestures is possible. All four key characteristics were applied to the models and the absolute value of the mean of each sensor yielded the highest accuracy by around 5%, regardless of the model. This characteristic was applied to all five models and five time-windows, and the accuracies were calculated seen in Table 3 below.

Table 3: All Five Model's Accuracy for Each Time-Window

	1.00 Seconds		0.50 Seconds		0.25 Seconds		0.05 Seconds		0.005 Seconds	
Model	Train	Test	Train	Test	Train	Test	Train	Test	Train	Test
MLP	100.0%	99.7%	100.0%	99.8%	99.8%	99.6%	97.9%	98.4%	75.1%	76.2%
Linear	100.0%	100.0%	99.8%	100.0%	99.2%	99.8%	96.9%	97.1%	31.9%	31.7%
Naïve Bayes Gaussian	97.3%	99.3%	97.7%	98.6%	96.9%	97.7%	91.0%	91.5%	-	-
Optimized KNN	100.0%	99.1%	99.6%	99.5%	99.1%	99.5%	96.5%	97.5%	67.7%	73.6%
PCA	97.9%	99.3%	98.5%	98.6%	96.6%	97.3%	91.0%	91.5%	-	-

All percentages will vary from run to run by around 0.1%

All models are very accurate given one whole second of data to differentiate each gesture. At half a second and a quarter of a second of data, the MLP, Linear, and KNN models all predicted gestures with over 99% in the test group. With 1/20th of a second, MLP, Linear, and an Optimized KNN still predict gestures with over 95% accuracy. Given only one datapoint and a time window of 1/200th of a second, no model was able to predict hand gestures with over 95% accuracy. This highlights the importance of determining a key characteristic as they simply cannot be calculated with only one datapoint. The models at 1/200th of a second are simply modeling off the raw data as the mean of a single number is just that number etc. It should be noted that on every repeat of the code, MLP operates around 1% more accurately than any other model at 1/20th of a second. For this reason, it was chosen as the model to perform a Confusion Matrix on. A classification report was also printed for all time windows to give insight per time window for the MLP model, but the Confusion matrix shown in Figure 3 was only made for the very shortest time window as it gives the most visual insight into the struggles and strengths of the model. A confusion matrix on a quarter second of data with 99.6% test accuracy will simply show all predicted values as true values other than one or two outliers, thus providing little insight.

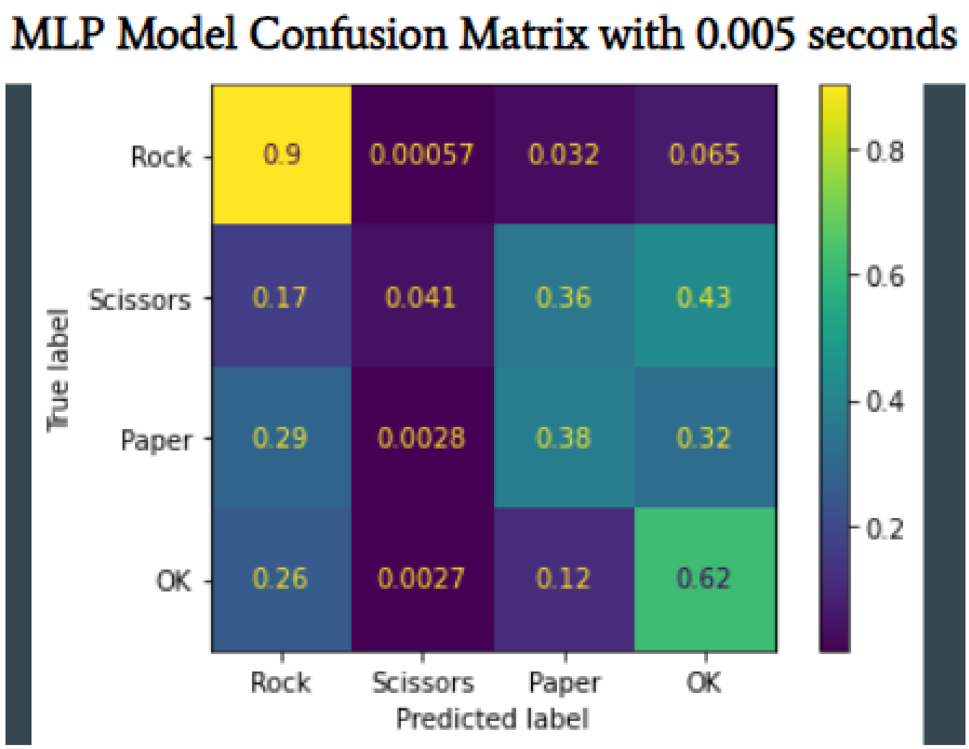

Figure 3: Confusion Matrix of an MLP Model given 1/200th of a Second of Data

In the confusion matrix of 1/200th of a second in Figure 3, the scale is from 0-1 with 1 representing a relationship of 100% pick frequency and 0 being 0% selection frequency. 90% of the time rock was predicted, it was rock. This would be considered a strength of the model. Unfortunately, the model never really predicted scissors and would often misplace a scissors gesture for Paper or OK. Overall, this can be expected with this low of a time window and the model's accuracy greatly increases with all larger time windows. This is because at 1/200th of a second, the key characteristic is not even calculated correctly as it is taking the mean on a single value as mentioned previously. Therefore, this confusion matrix more so highlights the importance of picking a key characteristic and windowing before applying the model. Even with 10 datapoints, the key characteristic allowed for the model to predict all testing data with close to 99% accuracy.

Conclusion

Each mechanic of this project was valuable in achieving the end results. The original goal was to determine the hand gesture with 99% accuracy given only 1 second of data. In the end, the MLP model can predict the hand gesture with 99% accuracy in only ¼ of a second. This is a four times improvement compared to the original goal and is partially expected given the extraordinary clean dataset as mentioned in the introduction. Future work is still needed to optimize and pick the best model in the scope of run-time and accuracy. MLP gives directional evidence that it performs the best, but it may also operate the slowest, causing unwanted lag in a real-world scenario and does not show a statistically significant performance increase over KNN or Linear models with the limited testing. Otherwise, this software just needs to be adjusted for constant streaming of data and then could easily control a hand prosthetic with a quarter second lag plus computing lag. The main learnings from this project were the importance of preprocessing as well as the rigor of real-world dataset restructuring as it took nearly a week of coding before a model could even be applied. Data may not always be in a format that is conducive to the goal objective and manipulating it can sometimes be very abstract and difficult.

References

1. Fahrenkopf, M. P., Adams, N. S., Kelpin, J. P., & Do, V. H. (2018, September 28). *Hand amputations*. Eplasty. Retrieved December 17, 2021, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6173827/
2. *How much does a prosthetic arm cost?* CostHelper. (n.d.). Retrieved December 17, 2021, from https://health.costhelper.com/prosthetic-arms.html
3. Yashuk, K. (2018, December 9). *Classify gestures by reading muscle activity.* Kaggle. Retrieved December 17, 2021, from https://www.kaggle.com/kyr7plus/emg-4?select=0.csv
4. *Forearm surface EMG Palm Movement classification*. Signal and Image Processing Lab. (n.d.). Retrieved December 17, 2021, from http://sipl.eelabs.technion.ac.il/projects/forearm-surface-emg-palm-movement-classification/
5. Shaw, L., & Bhaga, S. (2012, October). *Characteristics of EMG Signal*. Retrieved December 17, 2021, from https://www.researchgate.net/figure/Characteristics-of-EMG-signal_fig1_232905752
6. Dellacasa Bellingegni, A., Gruppioni, E., Colazzo, G., Davalli, A., Sacchetti, R., Guglielmelli, E., & Zollo, L. (2017, August 14). *NLR, MLP, SVM, and Lda: A comparative analysis on EMG data from people with Trans-radial amputation - journal of Neuroengineering and Rehabilitation*. BioMed Central. Retrieved December 17, 2021, from https://jneuroengrehab.biomedcentral.com/articles/10.1186/s12984-017-0290-6
7. Furui, A., Hayashi, H., Nakamura, G., Chin, T., & Tsuji, T. (2017, June 22). *An artificial EMG generation model based on signal-dependent noise and related application to motion classification*. PLOS ONE. Retrieved December 17, 2021, from https://journals.plos.org/plosone/article?id=10.1371%2Fjournal.pone.0180112
8. Bosco, G. (2010). *Principal Component Analysis of Electromyographic Signals: An Overview* . Retrieved December 17, 2021, from https://benthamopen.com/contents/pdf/TOREHJ/TOREHJ-3-127.pdf