EMG-Based Classification of Hand Movement in Virtual and Physical Reality

James Wong
Carnegie Mellon University
Pittsburgh, PA
jtw2@andrew.cmu.edu

Luigi Burda
Carnegie Mellon University
Pittsburgh, PA
lborda@andrew.cmu.edu

Dailyn Despradel
Carnegie Mellon University
Pittsburgh, PA
ddespradel@andrew.cmu.edu

Abstract

Daily hand movements are used when training upper limb individuals whether through physical or virtual environments. Electromyogram signals are electrical signals acquired from muscles that allow for generated movement classification due to the distinct signature of each movement. The classification of electromyography (EMG) while a participant performs basic hand movements allows for personalized rehabilitation training approaches that could enhance motor performance. Thus, this study proposes the use of machine learning (ML) to select a classifier that best discriminates EMG patterns of a set of muscles while performing grasping movements within a virtual and physical environment. This work concluded that there is a substantial difference between EMG signals in real and virtual environments. All the tested classifiers showed an accuracy on the test set higher than 97% when classifying in two environments. Given its computational efficiency, this study selected logistic regression as the classifier that best discriminated EMG patterns during object manipulation within the physical and virtual environment.

1. Introduction

According to the literature, many people in the USA suffer from upper limb motor dysfunction due to a neurological deficit every year [3]. Many techniques have therefore been considered to facilitate motor recovery during assigned rehabilitation processes [7] [9]. Although promising results have been shown, rehabilitation is a long-lasting process, during which every subject, based on the impairment level and the engagement in the rehabilitation program, makes progress differently. Subject-specific training techniques are therefore required to attain motor recovery [14] [5] [17]. In recent years, Virtual Reality (VR) is one of the most promising tools to support personalized rehabilitation training [18]. VR technologies allow for the simulation of Activities of Daily Living (ADL) and provide a range of personalized training to facilitate the engagement of patients [5] [10] [19]. At the same time, feedback related to muscle activity while performing a task plays an important role in motor rehabilitation. Researchers over the years have extensively used the electromyography signal (EMG) to gather muscle activity information [11]. EMG signals are small electric currents that are generated when a muscle is contracted [1]. These signals provide information related to movement and are therefore helpful to understand the functions and dysfunctions of the muscular system [8]. The feature analysis of EMG signals can therefore offer muscle activity information, such as fitness, fatigue, endurance level, and gestures [22]. Thus, this project aims to examine the EMG signal of a participant while manipulating virtual vs physical objects in their respective environments. To evaluate the classification of such interactions, and eventually allow for progress updates while the participant interacts with the virtual object in the VE.

2. Related Work

Among rehabilitation treatments dedicated to upper limb rehabilitation, VR has been explored extensively in the last decades, with positive results as a potential intervention [2]. Studies have pointed out three key factors that have influenced the integration of VR into rehabilitation: a) enjoyable repetition process, b) personalized feedback to reproduce intense and massive stimuli of the interaction, and c) motivation and/or presenting the therapy in a pleasant and attractive way [12].

In the literature, a few studies can be found that aimed at investigating VR as a rehabilitation tool for personalized training [15] [6] [20]. Studies like Yoo et al. [21], investigated the effects of VR games and EMG for pediatric rehabilitation aimed at kids with cerebral palsy. Where they showed that the integration of VR and EMG biofeedback allowed for improved activation of muscles, augmented visual feedback of this activation, and overall engagement of the participant. Also, Muri et al. [13], designed a VR upper-limb model that allowed for user control via the use of self-generated EMG signals. Ultimately proposing a wearable robot simulator device, for upper limb rehabilitation train-

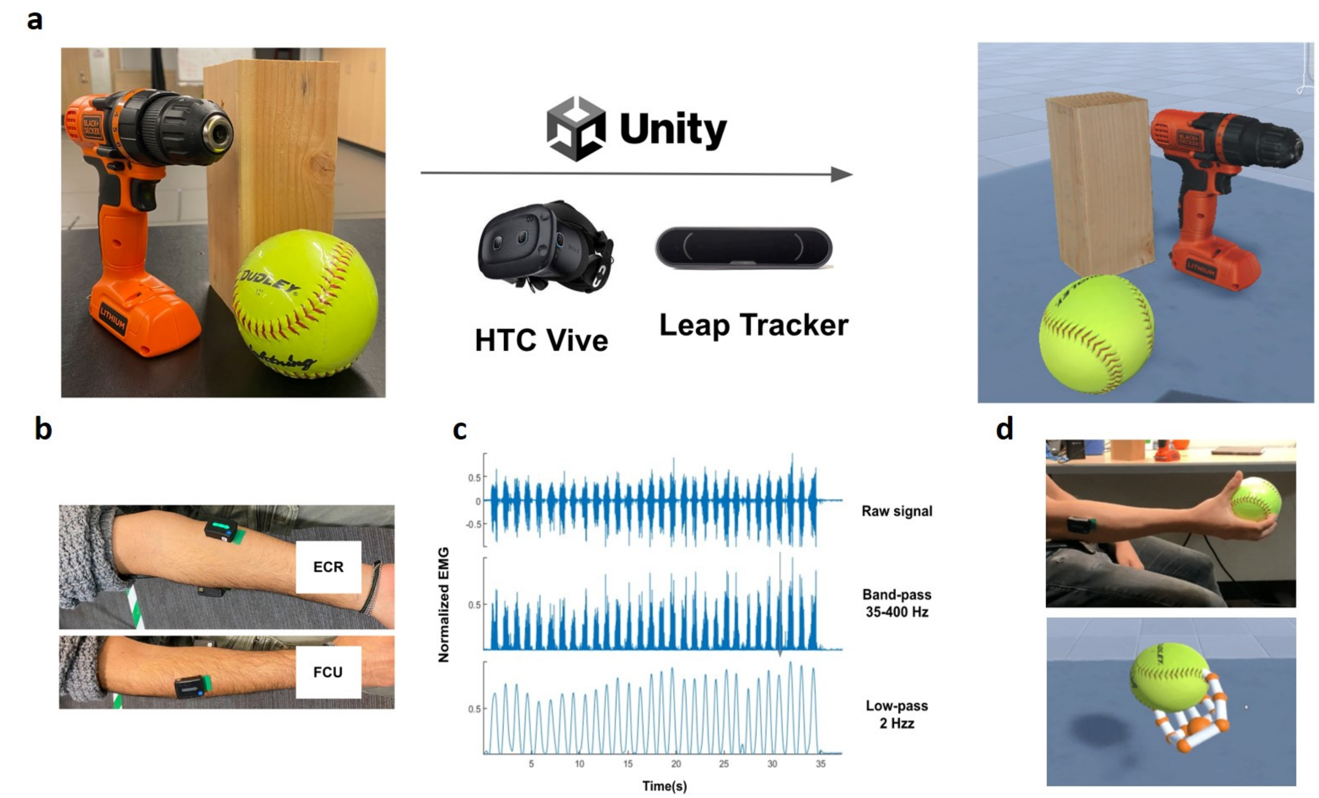

Figure 1. a) Virtual reality system with physical and virtual objects. b) EMG electrodes placement on Flexor Carpi Ulnaris (FCU) and Extensor Carpi Radialis (ECR). c) EMG processing pipeline. d) Grasping tasks in real and virtual environments.

ing for potential users of myoelectric prostheses. However, limited studies have been performed to demonstrate that VR - either immersive or not - can elicit muscle activity that could be evaluated when virtually and physically interacting with an object.

3. Data

3.1. Data collection

To validate the study, an experimental protocol was designed for data collection purposes. The HTC Vive Cosmos elite VR system and the UltraLeap Hand Motion tracking system, along with the Delsys EMG acquisition system were used to allow for EMG data collection in both environments (VE and RW) (Figure 1a). Three physical objects and their corresponding 3D models were chosen to perform the interaction task. The objects were gathered from the Yale-CMU-Berkeley (YCB) manipulation benchmark [4] (Figure 1a). The EMG signal was acquired from two muscles, Flexor Carpi Ulnaris (FCU) and Extensor Carpi Radialis (ECR) (Figure 1b), while performing a task in both the virtual environment (VE) and the real world (RW). (Figure 1d). Participants were prompted to remain seated for the entirety of the experiment. Sensor positioning was selected by palpitation while individually performing flexion and extension of the hand, and was confirmed by visual inspection of the EMG signals through the Delsys EMG acquisition software. Skin preparation involved cleaning the skin with alcohol. Subjects were instructed to interact with the virtual and physical objects as naturally as possible. The participant's hand position in the VE was projected as a virtual hand model through the use of the UltraLeap system. Contact with the virtual objects was controlled by a collision detection algorithm. Before the experimental task started, subjects were required to generate a maximal voluntary contraction (MVC) of both their flexors and extensor muscles. The procedure consisted of 3sec of resting phases and 2sec contraction phases with the users making a fist and pressing upwards against a table. As performed by Sapsanis et al. [16], the participant was then prompted to perform grasping and releasing movement 30 times for 40 seconds in both environments (Figure 1d).

3.2. Data processing

The final dataset included a total of 180 40-second-long, 2- channel EMG signals collected from the two muscles previously described. EMG signals were then processed to be ready for the feature extraction process. Firstly, manual data segmentation was performed to identify the EMG signals of each of the 30 repetitions. To easily identify the 30 segments we first applied a 7-th order bandpass Butterworth filter between 35Hz and 400Hz, each signal was then rectified and normalized to the maximum of each signal. Finally, we extracted the envelope of the signal by low pass filtering using a 5-th order lowpass Butterworth filter with a cutoff frequency of 2Hz. The entire pipeline for data segmentation is shown in Figure 1c. Each starting and end point of the 30 EMG segments characterizing the 30 repetitions for each object and subject was then visually identified. When the data segmentation process concluded, the raw signal of each segment was filtered using a 7-th order bandpass Butterworth filter between 35Hz and 400Hz, rectified, and normalized to the maximum voluntary contraction (MVC) of the subject. A total of 9 features for each segment were extracted. To select meaningful features we referred to the work of Sapsanis et al. [16]. We, therefore, extract the features as follow:

1. Zero crossing (ZC)
   $\sum_0^N (x_k > 0\&\&x_{k+1} < 0)||(x_k < 0\&\&x_{k+1} > 0)$
2. Slope Sign Changes (SSC)
   $\sum_0^N (x_k < x_{k+1}\&\&x_k < x_{k-1}) + \sum_0^N (x_k > x_{k+1}\&\&x_k > x_{k-1})$
3. Waveform Length (WL)
   $\sum_0^N |x_{k+1} - x_k|$
4. Willison Amplitude (WA)
   $\sum_0^N (|x_{k+1} - x_k| > \text{threshold})$
5. Variance (VAR)
   $\frac{1}{N}\sum_{samples}(x_k - \mu)^2$
6. Skewness (SKEW)
   $\frac{E(x-\mu)^3}{\sigma^3}$
7. Kurtosis (KURT)
   $\frac{E(x-\mu)^4}{\sigma^4}$
8. Mean (MEAN)
   $\frac{1}{N}\sum_{samples} x_k$
9. Standard deviation (SD)
   $\sqrt{\frac{1}{N}\sum_{samples}(x_k - \mu)^2}$

From inspecting our data, we chose a threshold of 0.06 for the Willison Amplitude.

All the features were organized in a python DataFrame of dimension 540 x 19 where the last column corresponds to the label. The features were labeled as follows:

- 0, if `fi` belongs to the RW
- 1, if `fi` belongs to the VE

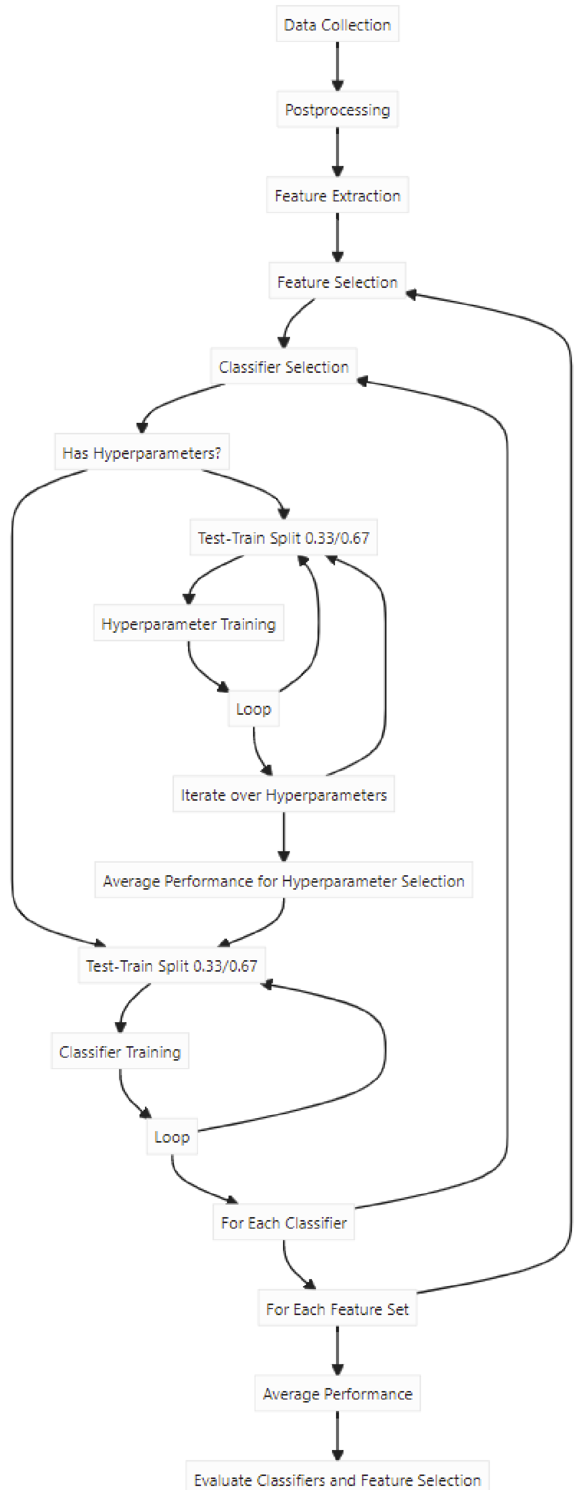

Figure 2. Classification Process

4. Methods

This study aimed to discriminate between the EMG signals produced in the RW and the VE. Using the data collected, we had 540 trials. We follow a fairly standard process for examination of classification, as outlined in the flow diagram in 2

The overall process started with data collection and processing. We then aimed to look at the performance of different feature sets and classifiers. For each step of training for hyperparameter tuning or classifier training, the dataset was split into train and test sets using the *sklearn* package *train_test_split*. We used a 67%/33% train-test split. The final dimension of the train set was therefore 361x19, while the test set was 179x19.

4.1. Classifiers

For this classification problem, we chose to evaluate the performance of six different classifiers using EMG signals from two different muscles after training. These algorithms were: K-nearest neighbors (KNN), support vector machines (SVM), logistic regression (LR), random forest (RF), gradient boosting (GB), and Naive Bayes (NB). For each of these, implementations from *sklearn* were used.

4.2. Feature Subsets

We also sought to determine what features would be the most useful for the classification of these EMG signals.

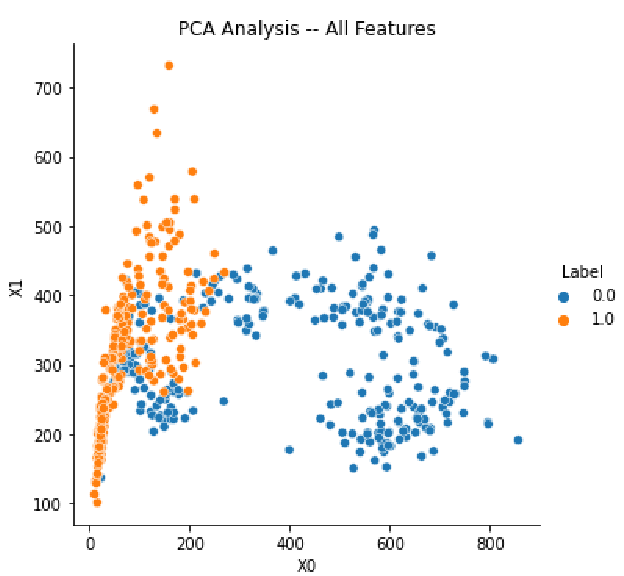

Figure 3. PCA of All Features shows a reasonable separation

After extracting the features from our experimental data, PCA was used to visualize how well our features captured the distinction between the RW and the VE. While there is a reasonable boundary, as shown in Figure 3, we wanted to determine which features were the most important for the classification problem.

	Avg Acc (ECR) [%]	Avg Acc (FCU) [%]
KUR	58.8	50.8
SKEW	60.9	54.7
SSC	81.6	68.7
WA	75.8	88.0
WL	78.0	80.7
SD	83.5	88.7
ZC	63.5	56.1
MEAN	77.3	94.5
VAR	74.4	89.2

Table 1. Average Accuracies of Training on a Single Feature

In order to try to get a first-order approximation of which features were the best at distinguishing between the two environments, all of our classifiers were trained on a single feature and evaluated over the average accuracy. The features that scored an average accuracy of approximately 77% or greater as our smaller feature set were then selected, ending up with nine features: SSC_ECR, WA_FCU, WL_ECR, WL_FCU, SD_ECR, SD_FCU, Mean_ECR, Mean_FCU, and Var_FCU. These accuracies and selected features are shown in 1. Running PCA on this subset, we see in 4 that the classification boundary is even more distinct.

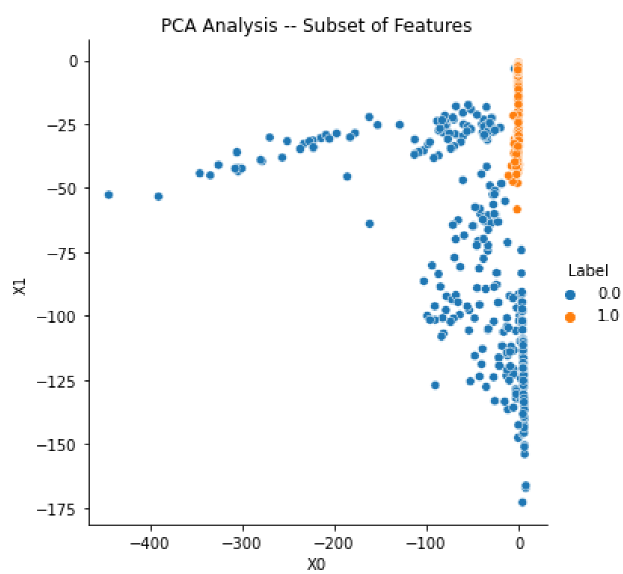

Figure 4. PCA of the Subset shows an even more distinct classification boundary

In total, we tested four different feature sets. These were all of the features, PCA on all of the features, our selected subset of features, and PCA on the subset of features. We chose 3 dimensions for PCA All and 2 dimensions for PCA Subset, as these accounted for 88.8% and 88.9% of the explaining power, respectively.

5. Experiments

5.1. Classifier Tuning

Of the classifiers, KNN, random forest, and gradient boosting all had hyperparameters. This was K for KNN and the number of estimators for random forest and gradient boosting. These hyperparameters for each of the classifiers and each of the feature sets were tuned. We iterated over a range of hyperparameters from 1 to 49 and trained and tested the classifier on 50 different train/test splits.

From the plots in 5, 6, and 7, we saw that performance was generally better for a small K, but a large number of estimators. This makes sense, as a larger K draws from further and further neighbors, which may not be indicative of the class. For random forests, a larger number of estimators helps average out bias and reduce underfitting, while overfitting is more closely related to tree depth. Gradient boosting sees the least consistent trend, although performance seems to improve with number of estimators. The final hyperparameters that we selected were shown in the 2.

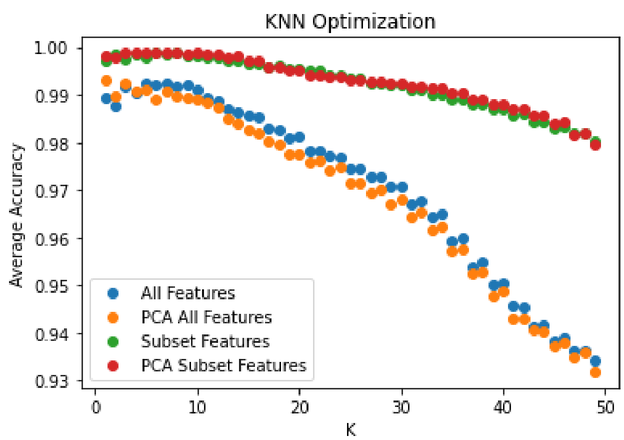

Figure 5. KNN K Optimization

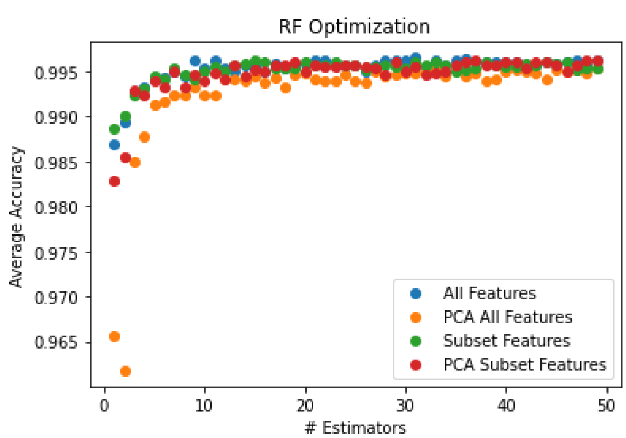

Figure 6. RF Number of Estimators Optimization

	KNN	RF	GB
All	7	31	23
PCA All	1	49	20
Subset	7	15	36
PCA Subset	5	49	44

Table 2. Final Hyperparameter Values

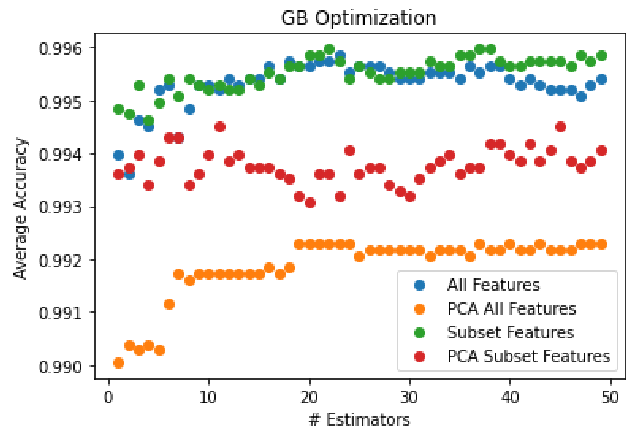

Figure 7. GB Number of Estimators Optimization

5.2. Accuracy Scoring

The primary metric tested in this study was accuracy. To do this, we trained and tested each classifier using the optimal hyperparameters, with each feature set for 50 different train/test splits. The distribution of accuracies was then plotted, both looking at the performance of the classifier and the impact of the feature set. A comparison of classifiers can be seen in 8.

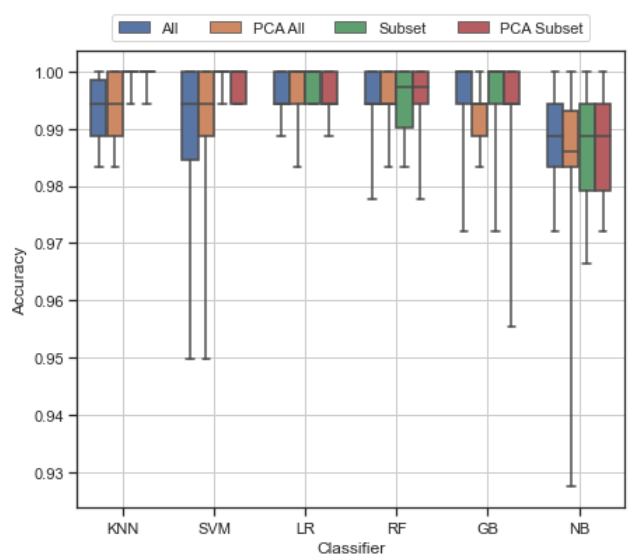

Figure 8. Classifier Accuracy

First, considering each type of classifier, we see that NB performs the worst out of the algorithms across all feature sets. We attribute this to the assumption from Naive Bayes that the features are independent, which is unlikely to hold given that our feature set came from only two muscles and some features, like the SD and VAR, are by definition not independent. We also see that LR performs fairly well across all feature sets, scoring highly and having few outliers. KNN and SVM both show significant improvement from only using a subset of the features. Both the ensemble methods, RF and GB, did fairly well, although they had some poor-performing outliers.

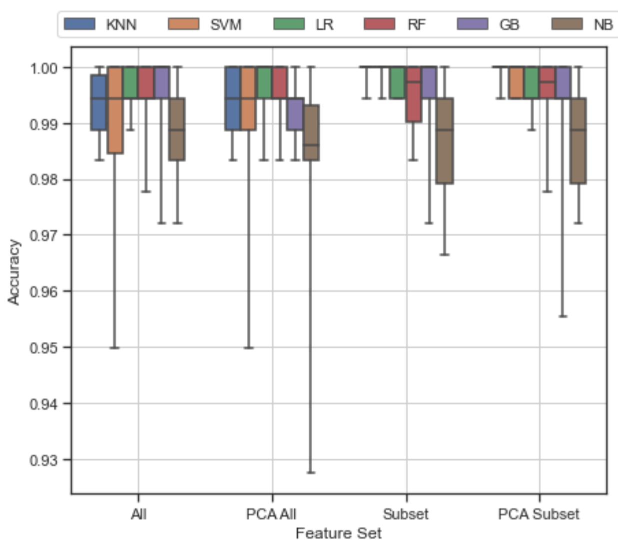

Figure 9. Feature Set Accuracy

A comparison of the feature sets is plotted in 9. We see that using a subset of the features improved the performance across most of the algorithms. While the performance of Naive Bayes was not greatly affected, its variance was reduced greatly. For both KNN and SVM, we see a performance improvement in both average accuracies as well as variance. The PCA feature sets have a slightly worse variance when compared to their full-dimension counterparts, which we would expect due to their loss of information. However, the performance drop is marginal, showing the possible benefits of dimensionality reduction.

All of the average accuracies can also be seen seen in 3.

	All	PCA All	Subset	PCA Subset
KNN	99.24	99.32	99.87	99.88
SVM	98.98	99.09	99.87	99.80
LR	99.60	99.80	99.73	99.58
RF	99.58	99.52	99.55	99.59
GB	99.58	99.23	99.58	99.42
NB	98.98	98.45	98.63	98.80

Table 3. Average Accuries

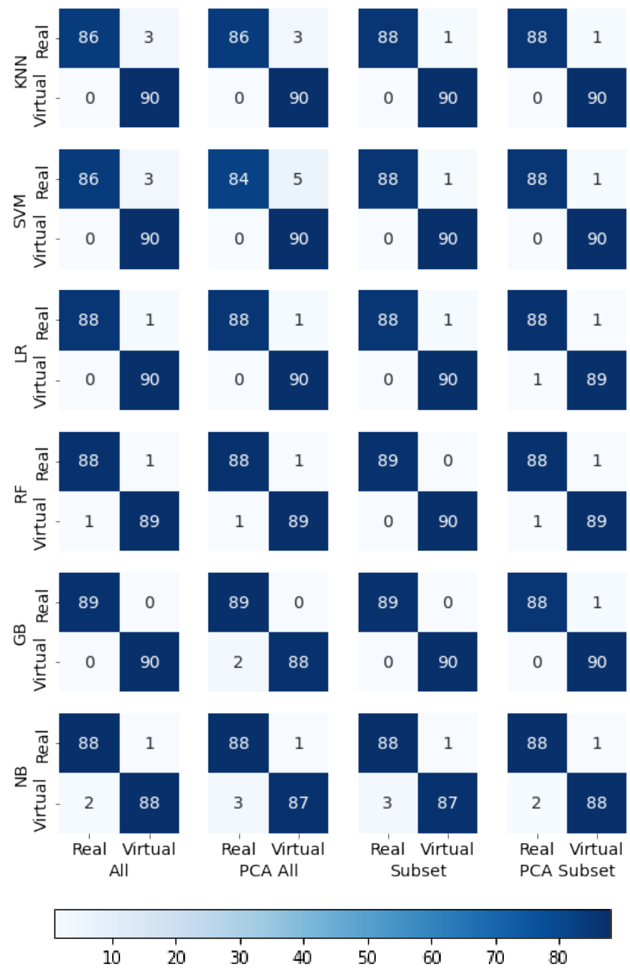

Figure 10. Confusion Matrices

5.3. Confusion Matrices

We also wanted to characterize where our classifiers were underperforming by looking at the confusion matrices. The confusion matrices were plotted in a matrix against the feature set and classifier shown in 10.

Again, we see the results of the classifiers are good overall. We also see some trends emerge with the types of misclassifications. For example, KNN and SVM only had false positives (for Real World identification). RF and NB were likely to have false negatives.

6. Conclusion

Personalized rehabilitation has been shown to improve subjects' motor performance [5] [10] [19]. The use of virtual reality, as a tool for personalized training, has also been extensively used in the past few years. This study aimed to show differences in muscle activation while performing a task in the real world or virtual environment resulting in different EMG patterns. To do that, we trained and tested six different classifiers in identifying EMG patterns from real and virtual environments. All of them showed an accuracy higher than 97% on the test set, proving a consequently substantial difference in the two sources of the EMG signals. Overall the Naive Bayes showed the worst performances compared to the other classifiers. Instead, logistic regression, random forest, and gradient boosting showed consistently high accuracy and low standard deviation. Given the results, its computational efficiency and ease of scaling to larger datasets, logistic regression would be the best classifier for the discrimination of EMG signals between real and virtual environments.

References

[1] A. Al-Jumaily and R. A. Olivares. Electromyogram (emg) driven system based virtual reality for prosthetic and rehabilitation devices. In *Proceedings of the 11th International Conference on Information Integration and Web-based Applications & Services*, pages 582–586, 2009.

[2] S. Arlati, N. Keijsers, G. Paolini, G. Ferrigno, and M. Sacco. Kinematics of aimed movements in ecological immersive virtual reality: a comparative study with real world. *Virtual Reality*, 26(3):885–901, 2022.

[3] B. S. Armour, E. A. Courtney-Long, M. H. Fox, H. Fredine, and A. Cahill. Prevalence and causes of paralysis—united states, 2013. *American journal of public health*, 106(10):1855–1857, 2016.

[4] B. Calli, A. Singh, J. Bruce, A. Walsman, K. Konolige, S. Srinivasa, P. Abbeel, and A. M. Dollar. Yale-cmu-berkeley dataset for robotic manipulation research. *The International Journal of Robotics Research*, 36(3):261–268, 2017.

[5] C. G. Canning, N. E. Allen, E. Nackaerts, S. S. Paul, A. Nieuwboer, and M. Gilat. Virtual reality in research and rehabilitation of gait and balance in parkinson disease. *Nature Reviews Neurology*, 16(8):409–425, 2020.

[6] D. Cano Porras, H. Sharon, R. Inzelberg, Y. Ziv-Ner, G. Zeilig, and M. Plotnik. Advanced virtual reality-based rehabilitation of balance and gait in clinical practice. *Therapeutic advances in chronic disease*, 10:2040622319868379, 2019.

[7] M. Capogrosso. Spinal cord stimulation for restoration of arm and hand function in people with subcortical stroke.

[8] B. A. De la Cruz-Sánchez, M. Arias-Montiel, and E. Lugo-González. Emg-controlled hand exoskeleton for assisted bilateral rehabilitation. *Biocybernetics and Biomedical Engineering*, 2022.

[9] C. T. Freeman, A.-M. Hughes, J. H. Burridge, P. H. Chappell, P. L. Lewin, and E. Rogers. A model of the upper extremity using fes for stroke rehabilitation. 2009.

[10] E. Klinger, A. Kadri, E. Sorita, J.-L. Le Guiet, P. Coignard, P. Fuchs, L. Leroy, N. Du Lac, F. Servant, and P.-A. Joseph. Agathe: A tool for personalized rehabilitation of cognitive functions based on simulated activities of daily living. *Irbm*, 34(2):113–118, 2013.

[11] M. S. H. Majid, W. Khairunizam, A. Shahriman, I. Zunaidi, B. Sahyudi, and M. Zuradzman. Emg feature extractions for upper-limb functional movement during rehabilitation. In *2018 international conference on intelligent informatics and biomedical sciences (ICIIBMS)*, volume 3, pages 314–320. IEEE, 2018.

[12] M. F. Montoya, J. E. Muñoz, and O. A. Henao. Enhancing virtual rehabilitation in upper limbs with biocybernetic adaptation: the effects of virtual reality on perceived muscle fatigue, game performance and user experience. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 28(3):740–747, 2020.

[13] F. Muri, C. Carbajal, A. M. Echenique, H. Fernández, and N. M. López. Virtual reality upper limb model controlled by emg signals. In *Journal of Physics: Conference Series*, volume 477, page 012041. IOP Publishing, 2013.

[14] J. R. Octavia and K. Coninx. Adaptive personalized training games for individual and collaborative rehabilitation of people with multiple sclerosis. *BioMed research international*, 2014, 2014.

[15] T. Rose, C. S. Nam, and K. B. Chen. Immersion of virtual reality for rehabilitation-review. *Applied ergonomics*, 69:153–161, 2018.

[16] C. Sapsanis, G. Georgoulas, A. Tzes, and D. Lymberopoulos. Improving emg based classification of basic hand movements using emd. In *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 5754–5757. IEEE, 2013.

[17] H. Singh, M. Shah, H. M. Flett, B. C. Craven, M. C. Verrier, and K. E. Musselman. Perspectives of individuals with subacute spinal cord injury after personalized adapted locomotor training. *Disability and rehabilitation*, 40(7):820–828, 2018.

[18] H. Sveistrup. Motor rehabilitation using virtual reality. *Journal of neuroengineering and rehabilitation*, 1(1):1–8, 2004.

[19] S. Tresser, T. Kuflik, I. Levin, and P. L. Weiss. Personalized rehabilitation for children with cerebral palsy. *User modeling and user-adapted interaction*, 31(4):829–865, 2021.

[20] D. Wen, Y. Fan, S.-H. Hsu, J. Xu, Y. Zhou, J. Tao, X. Lan, and F. Li. Combining brain–computer interface and virtual reality for rehabilitation in neurological diseases: A narrative review. *Annals of physical and rehabilitation medicine*, 64(1):101404, 2021.

[21] J. W. Yoo, D. R. Lee, Y. J. Sim, J. H. You, and C. J. Kim. Effects of innovative virtual reality game and emg biofeedback on neuromotor control in cerebral palsy. *Bio-medical materials and engineering*, 24(6):3613–3618, 2014.

[22] S. Zhao, J. Liu, Z. Gong, Y. Lei, X. OuYang, C. C. Chan, and S. Ruan. Wearable physiological monitoring system based on electrocardiography and electromyography for upper limb rehabilitation training. *Sensors*, 20(17):4861, 2020.