Rochester Institute of Technology

RIT Scholar Works

Theses

7-2020

# Using Photoplethysmography for Simple Hand Gesture Recognition

Karthik Subramanian
kxs8997@rit.edu

Follow this and additional works at: https://scholarworks.rit.edu/theses

Recommended Citation
Subramanian, Karthik, "Using Photoplethysmography for Simple Hand Gesture Recognition" (2020). Thesis. Rochester Institute of Technology. Accessed from

This Thesis is brought to you for free and open access by RIT Scholar Works. It has been accepted for inclusion in Theses by an authorized administrator of RIT Scholar Works. For more information, please contact ritscholarworks@rit.edu.

# Using Photoplethysmography for Simple Hand Gesture Recognition

by

**Karthik Subramanian**

A Thesis Submitted in Partial Fulfillment of the Requirements for the Degree of Master of Science
in Electrical Engineering

Supervised by

Dr. Ferat Sahin
Department of Electrical and Microelectronic Engineering
Kate Gleason College of Engineering
Rochester Institute of Technology
Rochester, New York
July 2020

Approved by:

---

Dr. Ferat Sahin, Professor
*Thesis Advisor, Department of Electrical and Microelectronic Engineering*

---

Dr. Gill Tsouri, Professor
*Committee Member, Department of Electrical and Microelectronic Engineering*

---

Dr. Jamison Heard, Assistant Professor
*Committee Member, Department of Electrical and Microelectronic Engineering*

---

Dr. Ferat Sahin, Professor
*Department Head, Department of Electrical and Microelectronic Engineering*

# Thesis Release Permission Form

Rochester Institute of Technology
Kate Gleason College of Engineering

Title:

Using Photoplethysmography for Simple Hand Gesture Recognition

I, Karthik Subramanian, hereby grant permission to the Wallace Memorial Library to reproduce my thesis in whole or part.

---

Karthik Subramanian

---

Date

# Dedication

I dedicate my thesis work to my family and many friends. A special feeling of gratitude to my loving parents, Subramanian and Meenakshi, whose words of encouragement and push for tenacity ring in my ears. I will always appreciate all they have done. Both of you have been my best cheerleaders.

# Acknowledgments

I would first like to thank my thesis advisor Dr.Ferat Sahin for his constant support. He consistently allowed this paper to be my own work, but steered me in the right the direction whenever he thought I needed it.

I would also like to thank the members of MABL. Without their passionate participation and input, this research could not have been successfully conducted.

I would also like to extend special thanks to Celal Savur as the second reader of this thesis, and I am gratefully indebted to his very valuable comments on this thesis.

Finally, I must express my very profound gratitude to my parents for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them.

For the rest not mentioned **I** thank you **all**.

Karthik Subramanian

# Abstract

**Using Photoplethysmography for Simple Hand Gesture Recognition**

**Karthik Subramanian**

**Supervising Professor: Dr. Ferat Sahin**

A new wearable band is developed which uses three Photoplethysmography (PPG) sensors for the purpose of hand gesture recognition (HGR). These sensors are typically used for heart rate estimation and detection of cardiovascular diseases. Heart rate estimates obtained from these sensors are disregarded when the arm is in motion on account of artifacts. This research suggests and demonstrates that these artifacts are repeatable in nature based on the gestures performed. A comparative study is made between the developed band and the Myo Armband which uses surface-Electromyography (s-EMG) for gesture recognition. Based on the results of this paper which employs supervised machine learning techniques, it can be seen that PPG can be utilized as a viable alternative modality for gesture recognition applications.

# List of Contributions

- Developed a novel wrist wearable device that can track heart rate and activity
- Explored a new Modality for Hand Gesture Recognition in the form of PPG
- Compared PPG to already existing modality like s-EMG
- Proved PPG is a viable alternative to s-EMG for Simple Hand Gesture Recognition
- **Publication**

  K. Subramanian, C. Savur, F. Sahin: â Using Photoplethysmography for Simple Hand Gesture Recognitionâ in the proceedings of, IEEE 15th International Conference on Systems of Systems Engineering. June 2-4, 2020

# Contents

# List of Tables

# List of Figures

# Chapter 1

# Introduction

Research in the field of Human-Computer Interaction (HCI) has seen a meteoric rise in the last few years by virtue of advancements in technology leading to better sensors and improved computing capabilities. Hand Gesture Recognition (HGR) is one aspect of HCI which is explored in this paper. Lately, wearable devices have amassed a lot of attention from both the research as well as the industrial community on account of their feasibility. The smartwatch is one such device which fits the billing. Most smart watches come equipped with green light PPG sensors. PPG is an uncomplicated and inexpensive measurement method is usually used for heart rate monitoring purposes. PPG reflects the blood movements which goes from the heart to the finger tips and toes through blood vessels in a wave like motion [8]. In comparison to the various types of PPG based devices, wrist band based devices are considered the most popular and preferred ones[9].

As PPG sensors are being well utilized already in the consumer wearable sector with the advent of smartwatches and other fitness wearables, they are more easy to integrate into existing products for the purpose of exploring HGR compared to other modalities. HGR modalities can be broadly classified into two sections, vision and non-vision based methodologies, Section II explores available technologies for HGR in more detail. In this work, we focus solely on non-vision based technologies for recognizing hand gestures, specifically using s-EMG and PPG. There is sufficient proof to suggest EMG based devices can be utilized for HGR However not many research articles exist to showcase PPG as a viable modality. The purpose of this paper is to compare the results obtained from both s-EMG and PPG on the set of same gestures performed, and provide more weight to existing research thus inferring PPG can be utilized as a cheaper and more readily available alternative to EMG.

Four visibly distinguishable gestures are selected. These gestures represent commonly used non verbal expressions, An additional reject class is included, this is done so as to be able test real time applications of the developed system. To facilitate robust data collection, Lab Stream Layer is employed along with time synchronised event markers which automate the process of separating the performed gesture with general PPG measurements.

Experiments are conducted on eight participants, performing the same gestures using both EMG and PPG devices. The experiments involve the participants responding to on screen stimuli by performing the necessary gestures. Chapter '2' delves into relevant literature to help explain the need for HGR and using PPG as an alternative. Chapter '4' follows up with the description of the methodology of this research to recognize hand gestures. This section further explains the experimental setup, the hardware used, the process of data acquisition and segmentation, followed by feature extraction and the classification processes in sufficient detail. Chapter '5' reports the results of PPG data analysis and classification and compares it to s-EMG for HGR. Chapter '6' is a brief conclusion and summary of this work and Chapter '7' explores all the future possibilities.

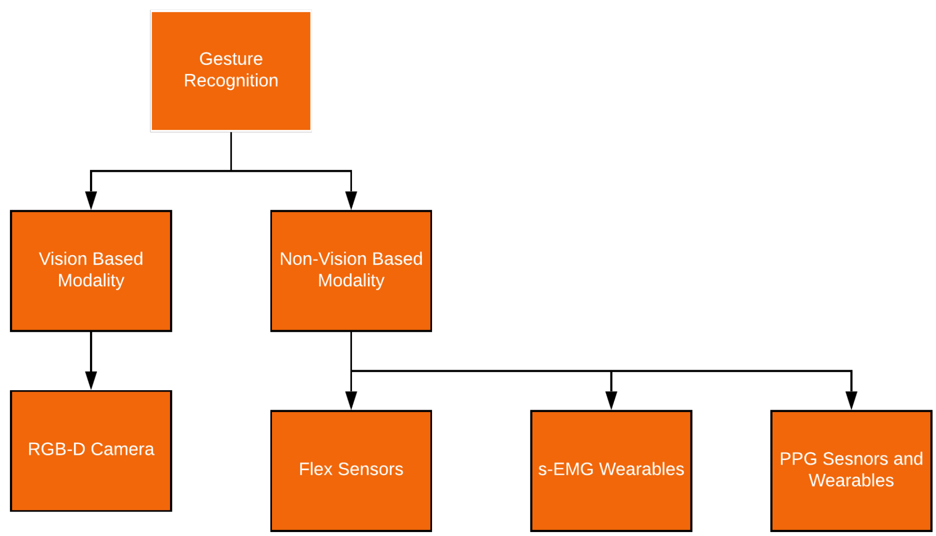
# Chapter 2

# Background Literature

This chapter presents a literature review of the alternative methods used for Hand gesture recognition.

## Hand Gesture Recognition

**Figure 2.1** Hand Gesture Recognition Modalities




HGR has gained a lot of traction in the recent years existing approaches may be divided into two main categories: Vision and Non-Vision based modalities. Vision based modalities mainly include RGB-D cameras. Non-Vision based modalities include bio-sensors and flex sensors.

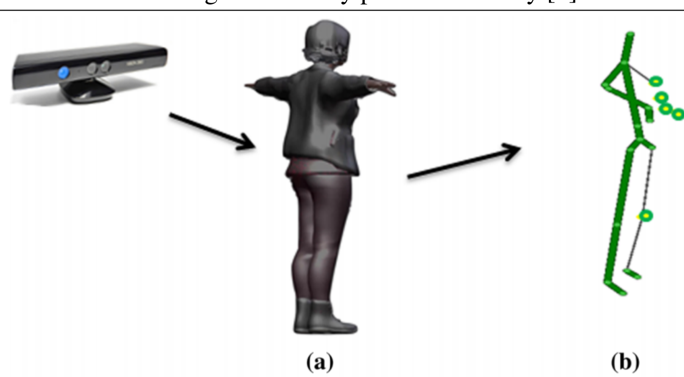
## 2.1 Vision Based Modalities

### RGB-D Cameras

**Figure 2.2** RGB-D Camera sensors [1]


An RGB-D camera like the kinect or 'Time of flight' sensor is utilized in vision based HGR approaches. The main challenges to vision based HGR are the type of system and the environment it is used in. Systemic challenges involve the potential for real time applications and the total cost the entire system may incur. Environmental challenges like backdrop in which the gestures are being performed, body occlusions,ambient lighting have all been found to affect the performance [1].

**Figure 2.3** RGB-D Camera sensors being occluded by parts of the body [1]

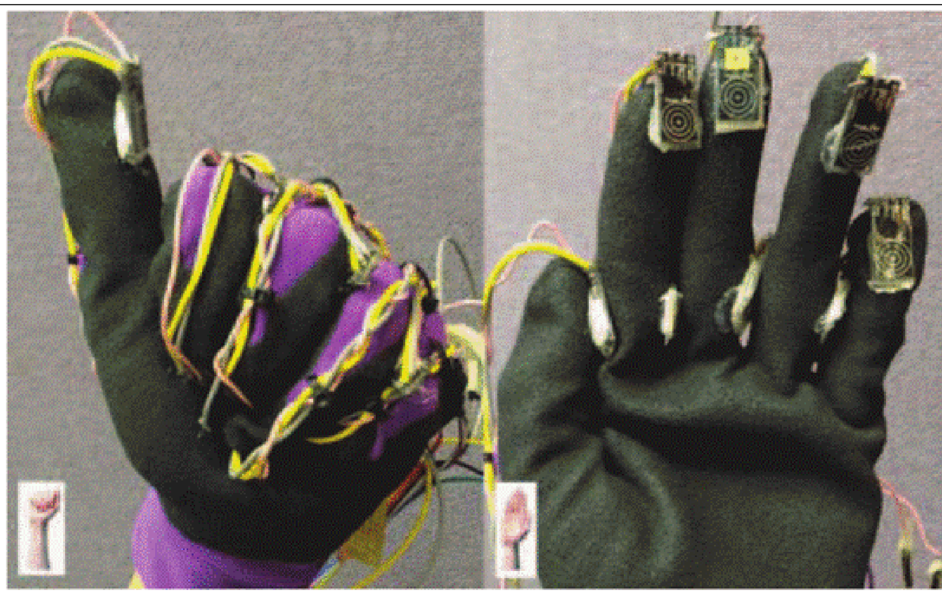



In the above figure (b) showcases how part of the body is being occluded due to the postion of the camera and the subject.

## 2.2 Non Vision Based Modalities

### Flex Sensors and Cyber Gloves

**Figure 2.4** Cumbersome Nature of Flex Gloves



Non Vision based modalities include action gloves which use flex sensors, However at their current state, a lot of work is required to establish that these systems can provide reliable support for a real-life engineering simulation [10]. They also require to be worn at all times covering the digits of the hand thereby making them cumbersome.

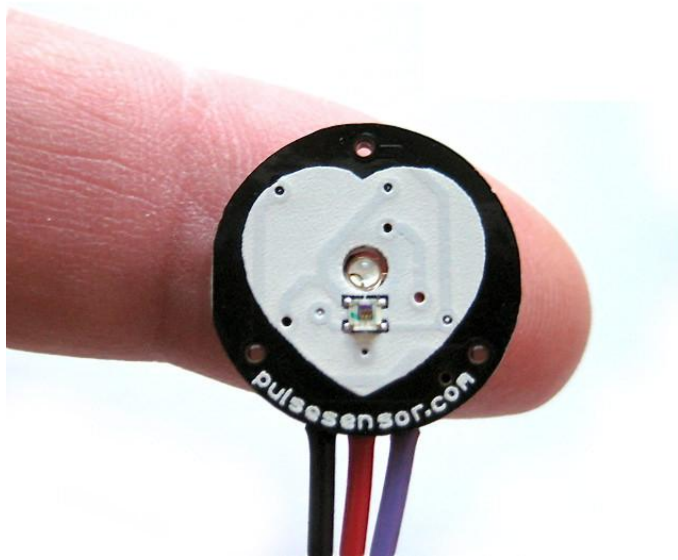
## s-EMG Signals

**Figure 2.5** Myo-Arm Band, A s-EMG Device [2]

A widely used technology for HGR is using surface-EMG. Extensive research has already been conducted in this particular domain [7][11][12]. The general challenges to using s-EMG for HGR is computational intensity along with the requirement of hardware that is not commonly available. Geng has been able to develop a generalized model with the help of deep learning to learn spacial features from s-Emg signals. To be able to do so they required a dense array of a 128 sensors [12]. The inherent nature of s-EMG signals is inconsistent. s-EMG depends on intrinsic factors like Physiology anatomy, number of motor units, fiber type, composition, depth and location of active fibers and the amount of tissue between surface and the electrode. This makes the s-EMG obtained from one person significantly different to others [13].

## PPG Signals

**Figure 2.6** PPG Pulse sensor [3]




PPG is an uncomplicated and inexpensive optical measurement that is often used for heart rate monitoring purposes. This method utilizes infrared light or green light to measure volumetric changes in blood flow. PPG devices are susceptible to motion artifacts [9] These PPG devices are generally used in wrist bands with the sensors position on or above the radial and ulnar arteries [9]. Research suggests that these artifacts are repeatable in nature based on the type of movement of the wrist, Zhao and company have performed a preliminary feasibility study on this potential technology by employing machine learning on data acquired from 10 different subjects they could classify 10 different hand gestures with 88% accuracy [14]. According to this research, they concluded that the repeatable nature of the hand gesture PPG signals were from motion artifacts.

# Chapter 3

# Supporting Information

This chapter contains all the necessary supporting information pertaining to pre-processing data and validating collected data, types of features to be extracted from the processed data followed by a brief explanation of methods and algorithms used for classification of the extracted features.

## 3.1 Preprocessing

### 3.1.1 Zero-Mean

Raw data usually contains unwanted noise. This may affect the performance of classifiers as the features may contain irrelevant information. Especially frequency based features. '0' Hz Frequency is generally observed in battery operated data collection devices. The removal of the DC component in a signal can be obtained by a process called Zero mean. The mathematical representation can be seen in 3.1. Here $X$ is a single recorded hand gesture

$$X_{zero} = X - mean(X) \tag{3.1}$$

## 3.2 Feature Extraction

### 3.2.1 Features and Formulae

In machine learning, having distinguishable features are the most important keys as they improve accuracy of the system significantly. In order to have a good HGR system, following feature extraction techniques were used:

**Table 3.1** Table of extracted features

| Category | Features |
|---|---|
| Time Domain | Standard deviation<br>Variance<br>Maximum value<br>Minimum value<br>Root mean square |
| Frequency Domain | Skewness<br>Kurtosis<br>Mean frequency |

#### Standard Deviation

The standard deviation is a measure of how far the signal fluctuates from the mean. ... By definition, the standard deviation only measures the AC portion of a signal, while the rms value measures both the AC and DC components. If a signal has no DC component, its rms value is identical to its standard deviation.

$$\sigma = \sqrt{\frac{\sum_{i=1}^{N}(X_i - \mu)^2}{N}} \tag{3.2}$$

In (3.2) $\sigma$ is the standard deviation of one recorded gesture. $\mu$ is the mean. $N$ is the number of data points in the recorded gesture signal.

### Variance

Variance of a signal is the difference between the normalized squared sum of instantaneous values with the mean value. In other words it provides you with the deviation of the signal from its mean value. It gives you the spread of your signal's data set.

$$\sigma^2 = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2 \tag{3.3}$$

### Minimum and Maximum Values

The maximum amplitude and the minimum amplitude values are obtained from recorded Hand Gestures and stored as features.

$$\max(X) \tag{3.4}$$

$$\min(X) \tag{3.5}$$

### Root Mean Square (RMS

The RMS value of a signal calculates the average of a signal over time.the signal value (amplitude) is squared, averaged over a period of time, then the square root of the result is calculated. The result is a value, that when squared, is related (proportional) to the effective power of the signal. In (3.6) $X_r$ is the Root mean square value of the recorded hand gesture.

$$X_r = \sqrt{\sum_{i=1}^{N}\frac{(x_i^2)}{N}} \tag{3.6}$$

**Skewness**

Skewness indicates the symmetry of the probability density function (PDF) of the amplitude of a time series. A time series with an equal number of large and small amplitude values has a skewness of zero. A time series with many small values and few large values is positively skewed (right tail), and the skewness value is positive. A time series with many large values and few small values is negatively skewed (left tail), and the skewness value is negative

$$skewness = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^3 \tag{3.7}$$

**Kurtosis**

Kurtosis measures the peakedness of the PDF of a time series. A kurtosis value close to three indicates a Gaussian-like peakedness. PDFs with relatively sharp peaks have kurtosis greater than three. PDFs with relatively flat peaks have kurtosis less than three

$$Kurtosis = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^4 \tag{3.8}$$

## 3.3 Dataset Validation Methods

Correlation describes the mutual relationship which exists between two or more things. The same definition holds good even in the case of signals. That is, correlation between signals indicates the measure up to which the given signal resembles another signal. The Pearson correlation measures how two continuous signals co-vary over time and indicate the linear relationship as a number between -1 (negatively correlated) to 0 (not correlated) to 1 (perfectly correlated) Variables $x$ and $y$ are two independently recorded hand gestures of the same class.

$$\mu_x = \frac{1}{N}\sum_{i=1}^{n} x_i \tag{3.9}$$

Equation (3.9) solves for the mean of the recorded Hand Gesture $x$

$$S_{xx} = \frac{1}{(n-1)}\sum_{i=1}^{n}(x_i - \mu_x)^2 \tag{3.10}$$

Equation (3.10) solves for the Sample Variance of $x$

$$\mu_y = \frac{1}{N}\sum_{i=1}^{n} y_i \tag{3.11}$$

Equation (3.11) solves for the mean of the recorded Hand Gesture to be compared with $y$

$$S_{yy} = \frac{1}{(n-1)}\sum_{i=1}^{n}(y_i - \mu_y)^2 \tag{3.12}$$

Equation (3.12) solves for the Sample Variance of $y$

$$S_{xy} = \frac{1}{(n-1)}\sum_{i=1}^{n}(x_i - \mu_y)(y_i - \mu_y) \tag{3.13}$$

Equation(3.13) solves for the Sample Covariance between $x$ and $y$

$$Pcc = \frac{S_{xy}}{\sqrt{(S_{xX})(S_{yY})}} \tag{3.14}$$

Equation (3.14) solves for the sample Pearson correlation coefficient for measuring the association between variables $x$ and $y$

## 3.4 Principal Component Analysis (PCA)

Principal Component Analysis is a form of feature extraction, it allows you to drop the least important variables and at the same time keep the relevant variables. After PCA all the variables are independent of each other. Principal component analysis can be used to analyze the structure of a data set or allow the representation of the data in a lower dimensional dataset (as well as many other applications).

Let $\{\vec{x}_i\}$ be a set of $N$ column vectors of dimension $D$. Define the scatter matrix $S_x$ of the data set as

$$S_x = \sum_{i=1}^{N} (\vec{x}_i - \vec{\mu}_x)(\vec{x}_i - \vec{\mu}_x)^T$$

where $\vec{\mu}_x$ is the mean of the dataset

$$\vec{\mu}_x = \frac{1}{N} \sum_{i=1}^{N} \vec{x}_i$$

The $d$ largest principle components are the eigenvectors $\vec{w}_i$ corresponding to the $d$ largest eigenvalues. $d$ can be chosen arbitrarily with $d < D$. The eigenvectors of $\mathbf{S}$ can usually be found by using singular value decomposition.

The dominant eigenvectors describe the main directions of variation of the data. For example, if a dataset had 2 large eigenvalues, then the data variation is described largely by linear combinations of the 2 corresponding eigenvectors (ie. the data is largely coplanar).

The $d$ eigenvectors can also be used to project the data into a $d$ dimensional space. Define

$$W = [\vec{\mu}_1, \vec{\mu}_2, \ldots, \vec{\mu}_d]$$

The projection of vector $\vec{x}$ is $\vec{y} = W^T \vec{x}$. The corresponding scatter matrix $S_y$ of the vectors

$\{\vec{y}_i\}$ is:

$$S_y = W^T S_x W$$

The matrix W maximizes the determinant of $S_y$ for a given $d$. [15]

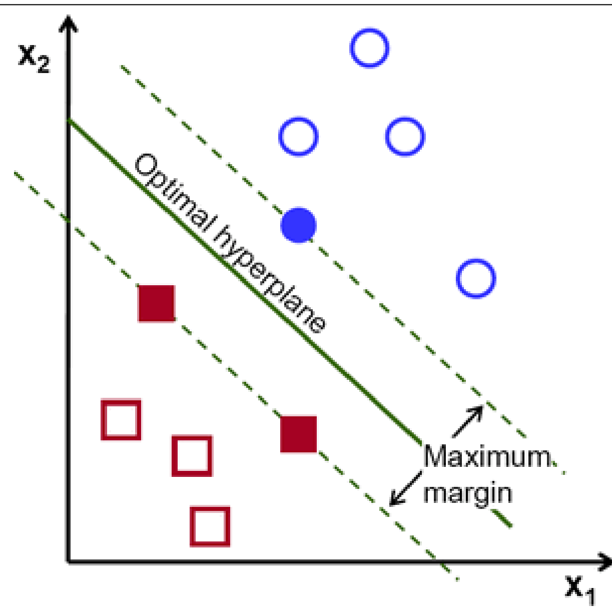
## 3.5 Classification Methods

In this research following Machine learning algorithm were used:

- Support Vector Machine
- Xg-Boosted Trees

### 3.5.1 Support Vector Machine (SVM)

**Figure 3.1** Support Vector Machine [4]




The SVM is the model that represents samples as points in the space, then divides them into separate groups with as large gap as possible as shown in Figure 3.1. There are kernel functions can be used to separate classes in higher dimensions. In this study, Radial Basis

Function (RBF) kernels were used with SVM. The RBF is non-linear kernel which is a good choice for HGR dataset with multiple classes. The SVM is a binary classification. To achieve multi class classification, one verses all approach was taken.

### 3.5.2 Extreme Gradient Boosted Trees

**Figure 3.2** Why XG-Boost for Boosted Trees [5]


Gradient boosting is a machine learning technique for regression and classification problems, which produces a prediction model in the form of an ensemble of weak prediction models, typically decision trees. [5]

A general model of tree boosting can be described by the following equation:

$$y_i = \sum_{k=1}^{K} f_k(x_i), f_k \epsilon (F) \tag{3.15}$$

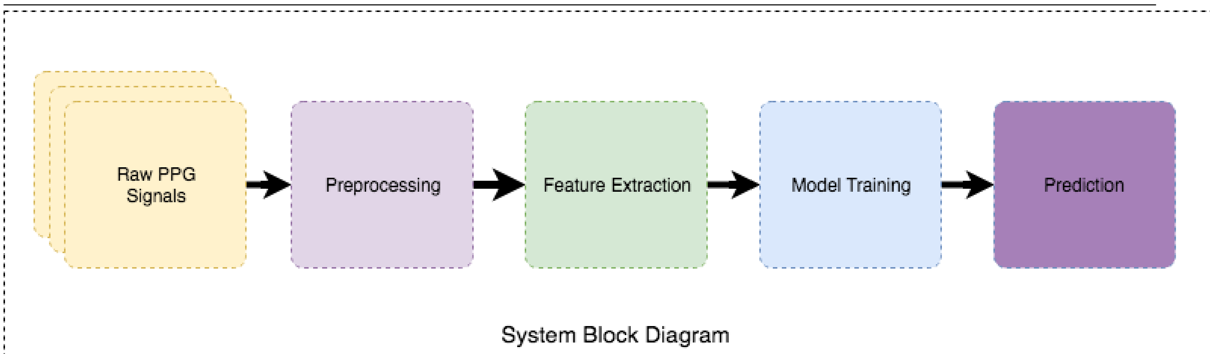
In the above Equation, the desired final model is $y_i$ and the assumption is that there are $k$ trees and $F$ is the space containing all trees. [5]

# Chapter 4

# Proposed Method and Experiements

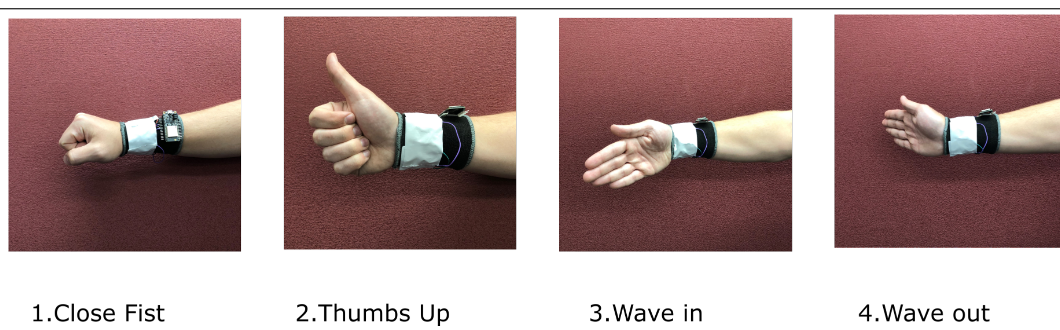
## 4.1 Methodology

**Figure 4.1** System block diagram




A set of three channel signals are recorded while performing four different static hand gestures and a reject class as seen in figure 4.2.

**Figure 4.2** Selected Hand Gestures



To confirm if the hand gestures repeatable nature were based on motion artifacts or blood flow, a small experiment was performed. The subject wore the wearable with the wrist covered with a cloth/sleeve. The subject would rest their arm to ensure there was no movement. In such a scenario, without the cloth/sleeve, The PPG wave forms should resemble regular heartbeats. With the sleeve, there would be no pulsating signal as the sensor is not in direct contact with the wrist. It was noticed that on repeating a the same hand gesture, there was a repeatable signal, as long as the sensors was placed in the same relative position compared to the previous attempt. The bearing of how blood flow affects these wave forms is inconclusive.

The general approach can be seen as a system block diagram in figure 4.1. Raw data is obtained from the device. It is followed by pre-processing, Relevant features are obtained by applying feature extraction. A supervised learning classifier is trained to generate a model with these obtained features. The then trained model is used to make predictions on unseen test data not utilized in the training process.

### 4.1.1 Hardware

The device essentially comprises of three off the shelf PPG sensors and a micro controller. To ensure similar sensor placements for a wide variety of subjects with different wrist sizes, the sensors are stitched on an elastic wrist band. This ensures that the sensors maintain similar relative distances between themselves every time it is worn and are positioned over the important arteries in the wrist viz. the radial and the ulnar arteries. This is illustrated in figure 4.4

**Figure 4.3** Developed PPG wrist band

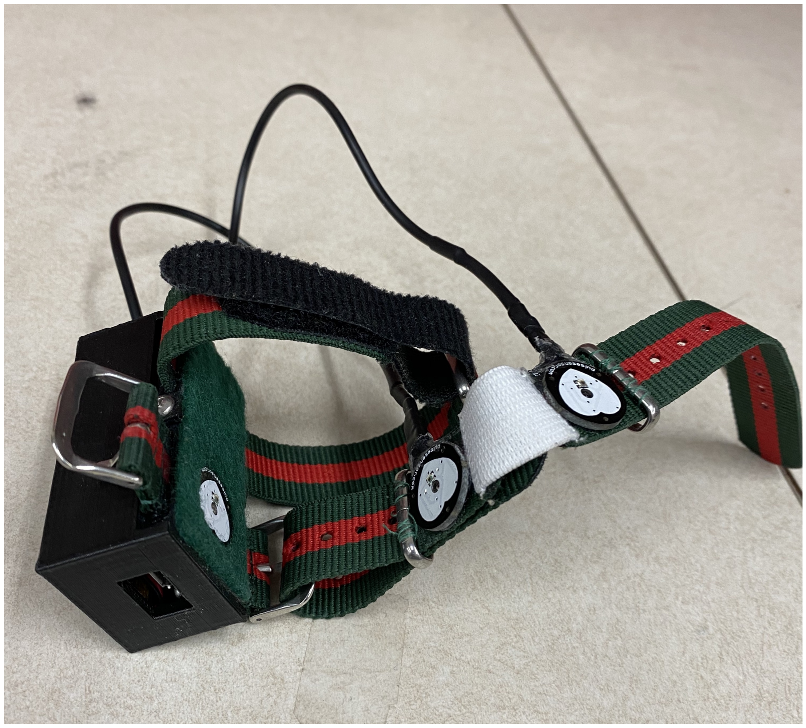
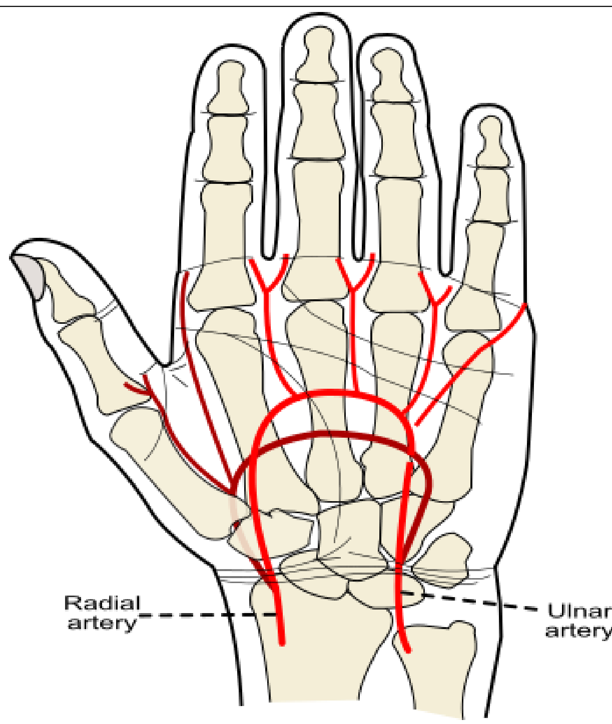


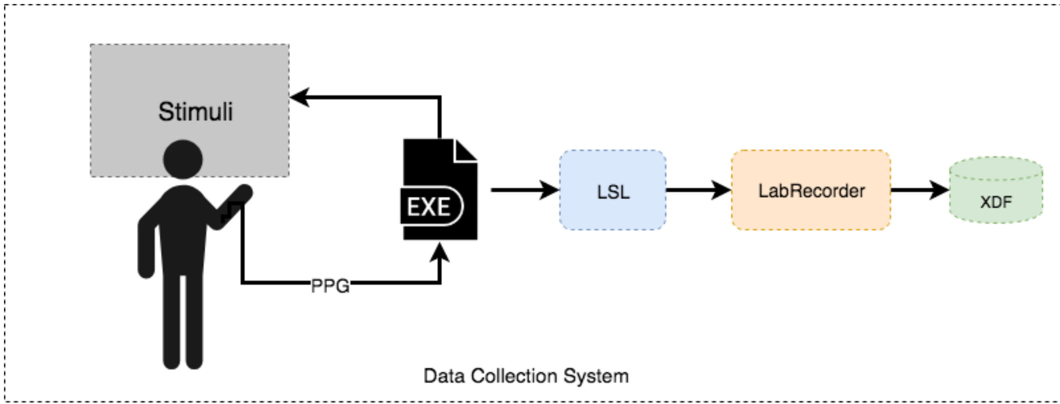
**Figure 4.4** Targeted Arteries [6]



The micro controller used, is a Adafruit feather ESP 32. The ESP 32 has both Wi-Fi and Bluetooth support, which enables wireless operation. It supports upto 12 ADC inputs and has I2C support which allows for future improvements with potential of adding more sensors like an accelerometer. The sampling frequency is set at 60 Hz. For better protection the micro controller is housed inside a 3-D printed enclosure.

### 4.1.2 Experimental Setup

**Figure 4.5** Data Acquisition System




The developed band continuously streams information to a computer which runs LSL. Lab Stream Layer(LSL) allows event marking with (near-)millisecond level accuracy, This helps diminish the chances of error which may be caused due to manual event marking. LSL is a system for the unified collection of measurement time series in research experiments that handles both the networking, time-synchronization, (near-) real-time access as well as optionally the centralized collection, viewing and disk recording of the data. [16] Use of LSL ensures there is no loss during the process of data acquisition. Necessary data is segmented with event markers as to ascertain the relevant parts of the recording. Subjects responds to visual stimuli from a computer program developed for the process of data collection. Event marking is automated using LSL. Figure 4.5 helps visualize this process.

Every time a number is visible as stimuli, LSL marks the start of a gesture recording. Once the number dissappears LSL marks the recording as end of gesture. While the entire stream is always recorded, these event markers serve to easily identify the relevant part of the stream in which the gestures are recorded. The program window with the visual cues are available in Fig 4.6

**Figure 4.6** Stimuli Software for Data Collection


### 4.1.3 Data Segmentation and Pre-processing

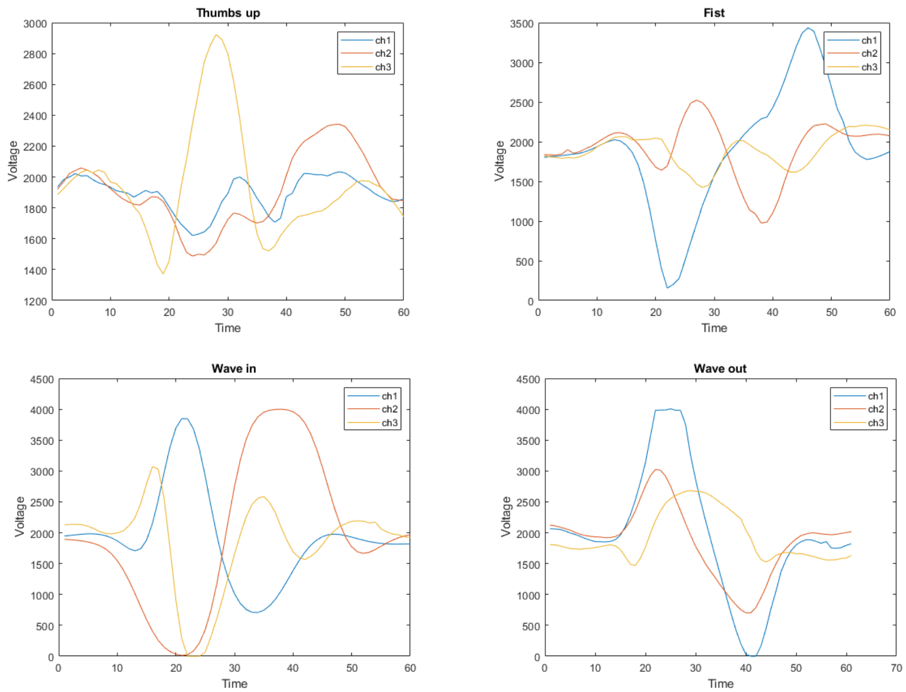
It was observed that the average time taken to perform any of the selected gestures is less than one second in duration, can be observed in Figure 4.7. To keep uniformity in data sample length, the data sample recording duration was set at one second, This resulted in the sample dimension size to be 60x3 where, 60 is the number of data points and 3 being the number channels. The amplitude is voltage in millivolts (mV)

Frequency analysis of the acquired signals can be seen in Figure 4.8. This suggests that all relevant information is available between 1-20 Hz. The zero Hz component is removed to eliminate any DC offset and a low pass filter is applied to the signal to ensure all non essential frequencies are ignored.

This was further verified by programming the wearable sensor to acquire data at higher frequencies. At 150Hz and 300 Hz the results are the same, No relevant information is available in frequencies greater than 20 Hz and can be seen in Fig 4.9

Before extracting any features it must be confirmed that the collected data is good. The Pearson Correlation Coefficient is calculated between samples of the same class to show

**Figure 4.7** PPG Wave forms of all selected hand gestures

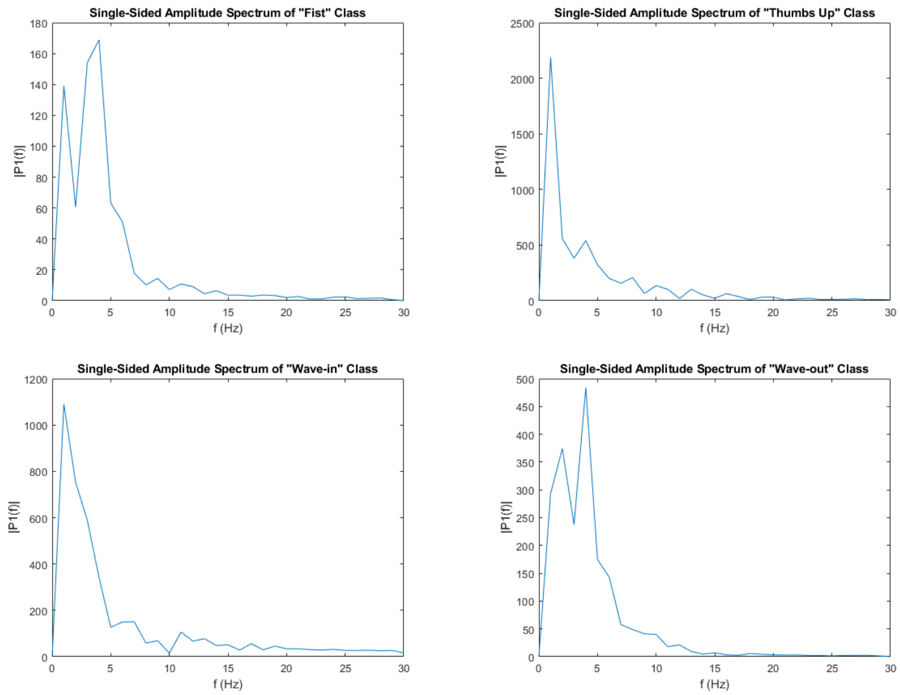



high association between themselves. Table 4.1 represents Pearson Correlation Coefficient from data obtained from a male able bodied subject. Pearson Correlation Coefficient is not utilized as a feature, it is used here to showcase the validity of the data set. A Pearson Correlation Coefficient value closer to one is desirable.

**Table 4.1** Pearson Correlation Coefficient

| **Type of Gesture** | **PCC** |
|---|---|
| Thumbs up | 0.8714 |
| Fist | 0.6932 |
| Wave in | 0.7912 |
| Wave out | 0.7989 |

**Figure 4.8** Fast Fourier Transforms of all selected PPG hand gestures

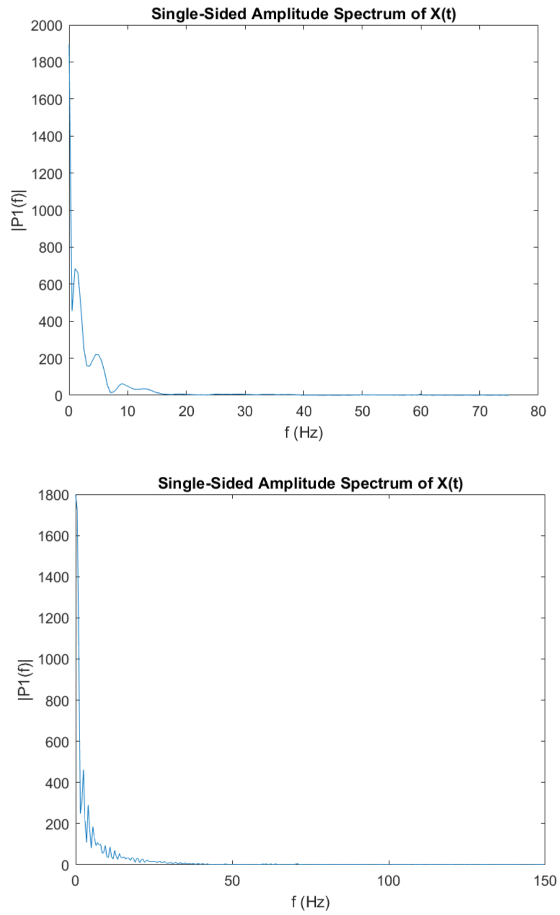



### 4.1.4 Feature Extraction

A feature is an individual measurable property of the process being observed. Using a set of features, any machine learning algorithm can perform classification [17]. Standard time and frequency series features are extracted from every recorded gesture. A feature set is generated and labelled with its class. This is used to train supervised learning classifiers. A total of 8 features are extracted per channel. The list of features are available in Table II for reference.

Figure 4.10 shows a plot of the extracted features reduced to 3-dimensions using Principle Component Analysis. The three best components of PCA can explain 70% variance of the extracted feature set. It can be observed from this figure that the classes begin to form small clusters. It can be better visualized with single subject data.

**Figure 4.9** Fast Fourier Transforms of 'Thumbs up Gesture' at Higher Frequencies



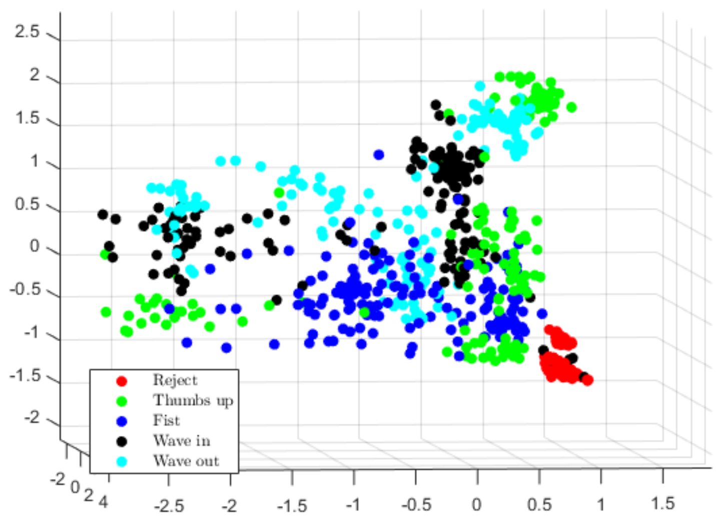

### 4.1.5 Feature selection and Classification

Feature selection techniques show that more information is not always good in machine learning applications [17]. After the process of feature extraction, it could be observed that a certain set of feature(s) may degrade or add no value to the performance of the classifier. A good measure for feature selection can be identifying the number of times a feature splits a tree. One of the classifiers used to identify hand gestures in this research is a gradient boosted tree. The xg-boost package is used to train the classifier. It is a scalable tree boosting system that is widely used by data scientists and provides state-of-the-art results

**Table 4.2** Table of extracted features

| Category | Features |
|---|---|
| Time Domain | Standard deviation (f0-f2)<br>Variance (f3-f5)<br>Maximum value (f12-f14)<br>Minimum value (f15-17)<br>Root mean square (f18-f20) |
| Frequency Domain | Skewness (f6-f8)<br>Kurtosis (f9-f11)<br>Mean frequency (f21-f23) |

**Figure 4.10** Extracted PCA features from all subjects using PPG




on many problems [5].

Figure 4.12 exhibits the features which split the tree most number of times after training, of the 24(8x3) extracted features, table 4.2 serves as a legend. Of the total features extracted, 18 are observed to be non-zero. Another observation that adds weight to this finding is that 99.9% of the total variance of the feature set can be explained by 18 principle components obtained from PCA. These PCA features are used to classify hand gestures

**Figure 4.11** Best Class Separation for Single-Subject using PPG

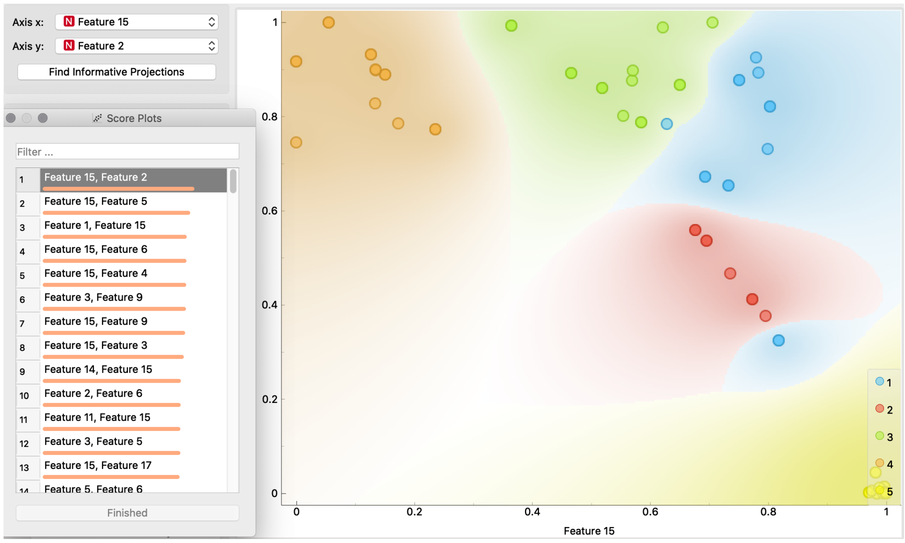
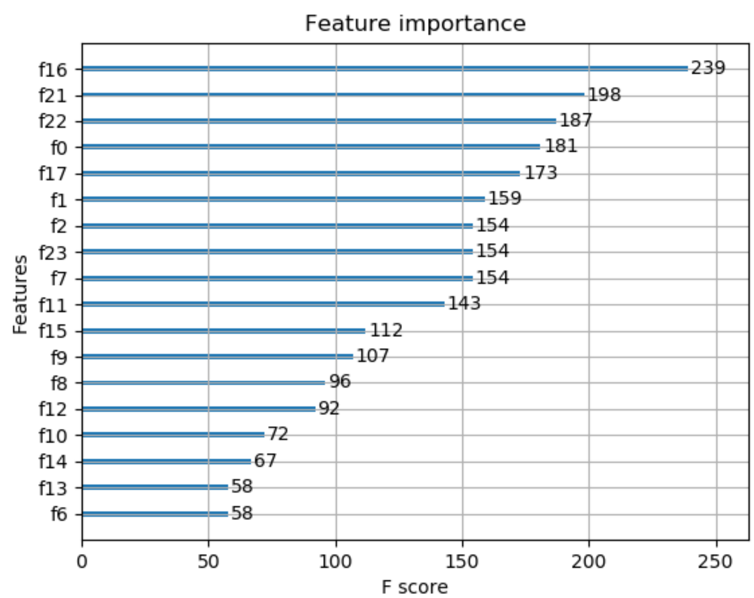
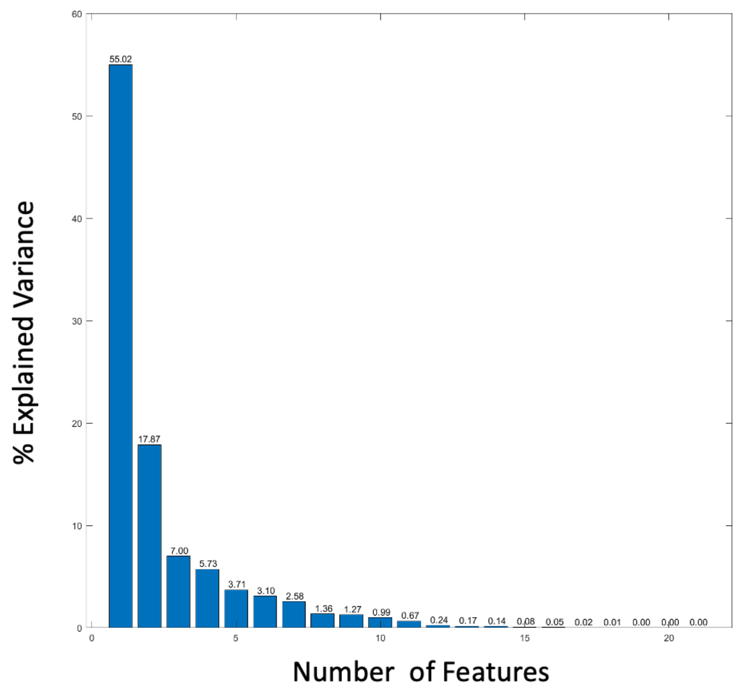



**Figure 4.12** Ranking Extracted features based on their importance




using a Support Vector Machine (SVM) type classifer. A cubic type kernel was best suited for this situation by observation.

**Figure 4.13** Ranking Extracted features based on their importance




PCA allows to reduce the dimensionality of the dataset. First 3 components of the PCA contain 77% of the explained variance of the total feature set, As seen in figure 4.13 Once the PCA components are identified, it eliminates all dependant variables, thus making the PCA features more separable for classifiers. The first three components are plotted in figure 4.8, It contains the features from every subject. Class clusters can be seen to form in three dimensions.

# Chapter 5

# Results

### 5.0.1 PPG Gesture Recognition

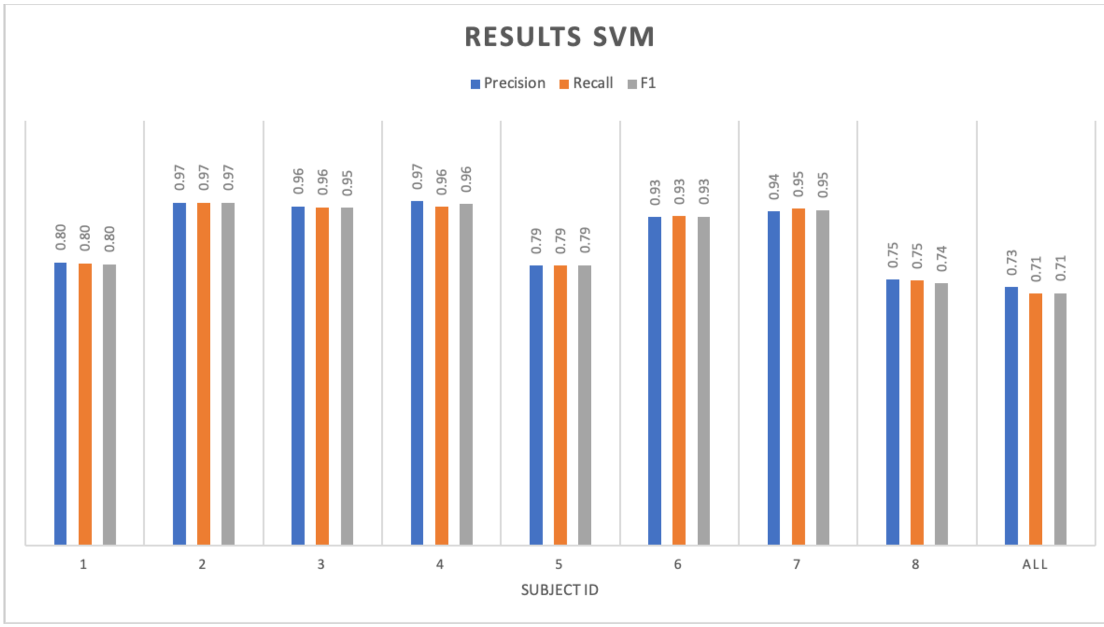
Two distinct classifiers are used to classify the hand gestures from the extracted features. A total of 1200 samples were recorded over 8 different subjects performing 4 gestures with an additional reject class. Table 5.1 displays the description of the collected data. Of the 8 subjects, 6 subjects were male and the remaining 2 subjects were female.

**Table 5.1** Data description

| Hand Gestures | | | | | | |
|---|---|---|---|---|---|---|
| Id | Thumbs up | Fist | Wave in | Wave out | Reject | Total |
| 1 | 50 | 50 | 50 | 50 | 50 | 250 |
| 2 | 50 | 50 | 50 | 50 | 50 | 250 |
| 3 | 25 | 25 | 25 | 25 | 50 | 150 |
| 4 | 25 | 25 | 25 | 25 | 25 | 125 |
| 5 | 15 | 15 | 15 | 15 | 15 | 75 |
| 6 | 25 | 25 | 25 | 25 | 25 | 125 |
| 7 | 30 | 30 | 30 | 30 | 30 | 150 |
| 8 | 15 | 15 | 15 | 15 | 15 | 75 |
| Total | 235 | 235 | 235 | 235 | 260 | 1200 |

The collected data is shuffled and split into a training and test sets. Of the total data 66% is used for training and the remaining data is used for testing purposes. One performs k-fold cross-validation using k = 5 or k = 10, as these values have been shown empirically to yield test error rate estimates that suffer neither from excessively high bias nor from very high variance [18]. Hence k was chosen to be equal to five. Using k-fold cross validation are good strategies to avoid an over fitting bias.

**Figure 5.1** PPG results for SVM classifier




Precision, recall and f-1 scores are common metrics used to tell how well a classifier has trained [19]. Figures 5.1 and 5.2 presents the average validation precision, recall and f-1 scores for every individual subject and the collective model which used Support Vector Machine and Boosted trees respectively. The F1-score is obtained from both recall and precision. The obtained results are all in the same range. The average accuracy is about 88% across all subjects. The average leave one out accuracy tested on subjects not included in the training process is about 74% between the two classifiers.

Figure 5.1 and 5.2 exhibit the results obtained from SVM and Boosted Trees respectively.

**Figure 5.2** Results for xg-boost classifier for classifying hand gestures using PPG


### 5.0.2 Comparative Study

The same set of gestures are trained to develop a classification model using s-EMG as the modality. The Myo arm band is a commercially available wrist band which facilitates the collection of s-EMG. It has eight channels and samples at a frequency of 200 Hz. The entire process exhibited in Figure 4.1 is repeated with s-EMG. The same set of features are extracted for the purpose of training a classifier.

Completely identical Data collection process is maintained for s-EMG. The collected data is split into a train and test set. All standard over-fit avoidance strategies observed for the experiment with PPG are maintained for s-EMG. The extracted feature set is utilized to train a SVM classifier. The first three PCA components can be seen in Figure 5.3 The confusion matrix comparisons for single subjects can be seen in Figures 5.4 and 5.5 On a test data obtained from 1 subject, s-EMG can classify the selected hand gestures with an

**Figure 5.3** Single Subject s-EMG PCA class separation

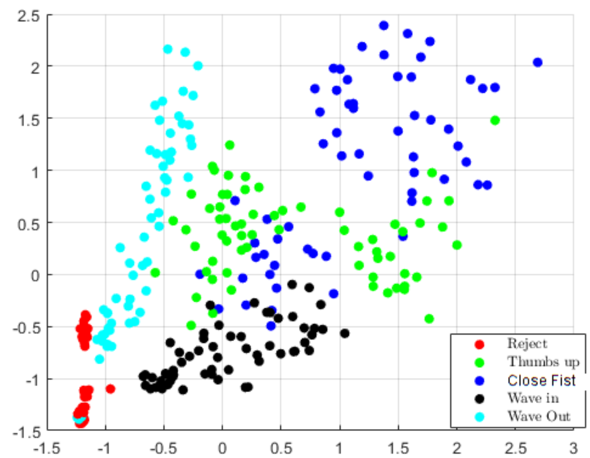
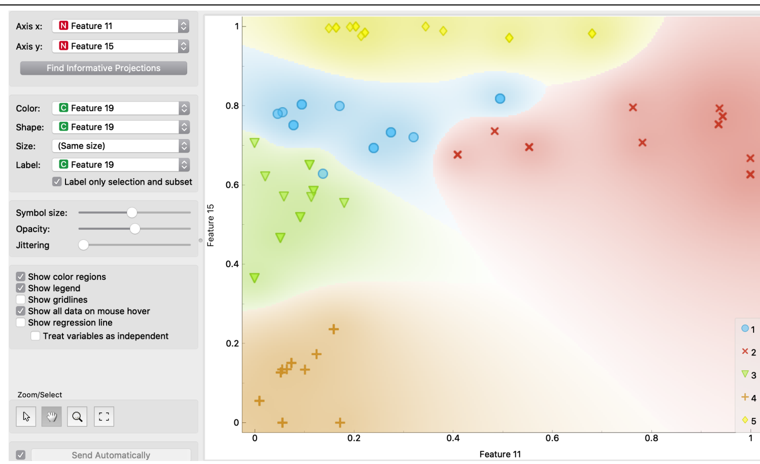



**Figure 5.4** Single-Subject PPG PCA class separation

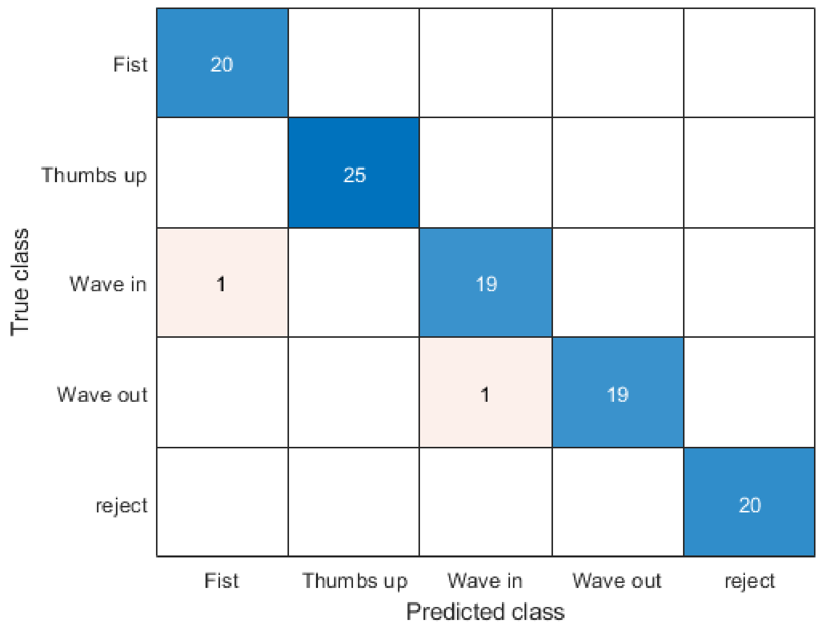



accuracy of 92.8% It is however observed that the the best three PCA components could only explain 58% of the total explained variance on account of higher channel count leading to more number of features, 64 instead of 24. To obtain similar accuracy comparable to PPG, 28 PCA features are necessary to explain more than 95% of the total variance of the extracted feature set. Thus making s-EMG more computationally intensive over machine learning methodologies to classify simple hand gestures.

### 5.0.3 s-EMG vs. PPG

As expected both s-EMG and PPG are highly accurate over single-subject testing. This can be observed in the confusion matrix of the results obtained from the same subject using two different modalities and same classification algorithm.

**Figure 5.5** Confusion Matrix for Single Subject SVM classifier using s-EMG



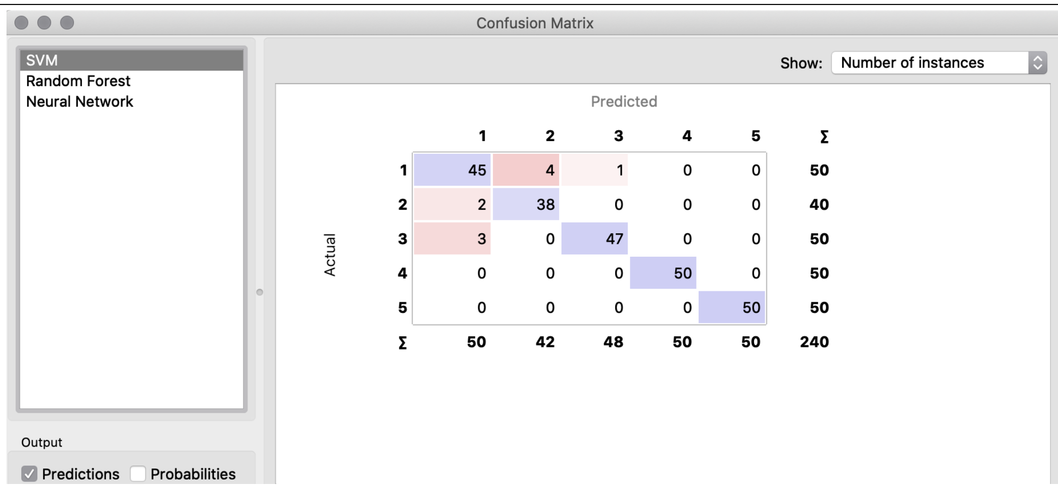

PPG can be seen to have a short training time like s-EMG and can be deployed for applications after a short training calibration phase. This can be seen in the results section. Acceptable levels of testing accuracy are obtained on single subjects with as little as 10 instances per class for training data. A similar approach is maintained in the use of the Myo-Arm Band which is a commercially available consumer product.

**Figure 5.6** Multi Subject Results for s-EMG, Generalized Model [7]

| **Feature** | **Validation Accuracy (%)** | **Trained Subject Test Accuracy (%)** | **Leave One Out Test Accuracy (%)** |
|---|---|---|---|
| MAV | 55.72 | 58.94 | 27.06 |
| STD | 53.66 | 57.49 | 26.93 |
| RMS | 53.84 | 57.11 | 26.37 |
| LOG | 55.98 | 59.18 | 29.32 |
| Power | 42.73 | 46.80 | 20.67 |
| WL | 55.15 | 58.93 | 27.85 |
| MFL | 51.53 | 54.44 | **30.12** |
| SampEn | 23.95 | 25.75 | 14.05 |
| MFL + LOG | 56.04 | 59.63 | 31.76 |
| LOG + SampEn | 56.08 | 59.35 | 31.71 |
| MFL + SampEn | 53.01 | 56.4 | **33.05** |
| LOG + MFL + SampEn | 56.69 | 59.96 | **33.66** |

Multi-subject Leave one out accuracy for s-EMG is only as high as 33% [7]. PPG comfortably beats out s-EMG for multiple subjects in our testing. A leave one out accuracy of 74% is achieved in this work of 8 subjects. The gestures are referenced in figure 4.2

**Figure 5.7** Confusion Matrix for Single subject PPG using SVM classifier



# Chapter 6

# Conclusion

Based on the findings of this research, it can be seen that PPG can achieve similar results as s-EMG for HGR. S-EMG is a well proven, tried and tested modality which has many commercially available products. The PPG wrist band developed for this research was directly compared with the Myo-Arm band on 4 hand gestures and a reject class. The findings of this research show that despite having lower sensor count and sampling rate, the classifiers are able to learn the gestures on PPG signals and were comparable to a classifier learned on s-EMG. s-EMG is more computationally challenging to train on multiple subjects for a simple hand gestures The experiments conducted in this research demonstrate that the learned models on PPG can accurately classify four simple hand gestures over eight different individuals with an accuracy of 88% . These findings suggest that PPG may be used for HGR in future applications.

# Chapter 7

# Future Work

There can be further potential advantages of adding an accelerometer to the developed band. An accelerometer can be used with PPG signals in a fusion scheme for identifying the hand orientation. Accurately recognizing the same gestures over multiple subjects requires a good band design. The band used in this research work is a prototype which will be improved to house a flexible PCB to replace the breakout board and an accelerometer as an additional sensor. A robust data set can be created by increasing the number of participants. Deep learning algorithms such as 1-D CNNs or LSTMs can be employed to obtain better results over multiple subjects. The developed wearable device can be used for translation of American Sign Language as an application. However we are aware that the current modality would restrict the functionality to only static gestures. On testing with subjects who were fluent in ASL, simple static gesture had repeatable wave forms. Dynamic gestures would require some form of fusion with an accelerometer. It would also require a sliding window approach for data collection, the inter-class transitions will be required to be included in the data set and any ML model intended to run a real time application will require to learn these transitions. While the intentions of the development of this band is recognizing hand gestures, it can also be utilized for other applications in health monitoring by analysing heart rate patterns. In order to develop this wearable as a sensor suite platform, the PPG sensor can be replaced with a better PPG sensor which also incorporates a pulse-oxymeter which can gather oxygen levels in the blood stream. This can be used in sport performance monitoring along with the accelerometer.

# Bibliography

[1] A. S. Al-Shamayleh, R. Ahmad, M. A. M. Abushariah, K. A. Alam, and N. Jomhari, "A systematic literature review on vision based gesture recognition techniques," *Multimedia Tools and Applications*, vol. 77, no. 21, pp. 28 121–28 184, Nov 2018. [Online]. Available: https://doi.org/10.1007/s11042-018-5971-z

[2] T. Labs, "Myo armband @ONLINE," https://myo.com/.

[3] P. Sensor, "Pulse sensor @ONLINE," https://pulsesensor.com/.

[4] "Support Vector Machine Plot," 2015. [Online]. Available: http://docs.opencv.org/2.4/doc/tutorials/ml/introduction_to_svm/introduction_to_svm.html

[5] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining - KDD '16*. ACM Press, pp. 785–794. [Online]. Available: http://dl.acm.org/citation.cfm?doid=2939672.2939785

[6] R.H.Castilhos, "Illustration of arteries @ONLINE," https://commons.wikimedia.org/wiki/File:Gray1237.svg.

[7] C. Derr and F. Sahin, "Signer-independent classification of american sign language word signs using surface EMG," in *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*. IEEE, pp. 665–670. [Online]. Available: http://ieeexplore.ieee.org/document/8122683/

[8] M. Elgendi, "On the analysis of fingertip photoplethysmogram signals," vol. 8, no. 1, pp. 14–25. [Online]. Available: http://www.eurekaselect.com/openurl/content.php?genre=articleissn=1573-403Xvolume=8issue=1spage=14

[9] M. Ghamari, "A review on wearable photoplethysmography sensors and their potential future applications in health care," vol. 4, no. 4. [Online]. Available: https://medcraveonline.com/IJBSBE/IJBSBE-04-00125.php

[10] O. Luzanin and M. Plancak, “Hand gesture recognition using low-budget data glove and cluster-trained probabilistic neural network,” vol. 34, no. 1, pp. 94–105. [Online]. Available: http://www.emeraldinsight.com/doi/10.1108/AA-03-2013-020

[11] C. Savur and F. Sahin, “Real-time american sign language recognition system using surface EMG signal,” in *2015 IEEE 14th International Conference on Machine Learning and Applications (ICMLA)*. IEEE, pp. 497–502. [Online]. Available: http://ieeexplore.ieee.org/document/7424365/

[12] W. Geng, Y. Du, W. Jin, W. Wei, Y. Hu, and J. Li, “Gesture recognition by instantaneous surface EMG images,” vol. 6, no. 1. [Online]. Available: http://www.nature.com/articles/srep36571

[13] M. B. I. Reaz, M. S. Hussain, and F. Mohd-Yasin, “Techniques of EMG signal analysis: detection, processing, classification and applications,” vol. 8, no. 1, pp. 11–35. [Online]. Available: http://www.springerlink.com/index/10.1251/bpo115

[14] T. Zhao, J. Liu, Y. Wang, H. Liu, and Y. Chen, “PPG-based finger-level gesture recognition leveraging wearables,” in *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications*. IEEE, pp. 1457–1465. [Online]. Available: https://ieeexplore.ieee.org/document/8486006/

[15] M. OpenCourseware, “Principal component analysis @ONLINE,” http://people.csail.mit.edu/tieu/notebook/pca.tex.

[16] S. C. for Computational Neuroscience, “Lab stream layer @ONLINE,” https://github.com/sccn/labstreaminglayer.

[17] G. Chandrashekar and F. Sahin, “A survey on feature selection methods,” vol. 40, no. 1, pp. 16–28. [Online]. Available: https://linkinghub.elsevier.com/retrieve/pii/S0045790613003066

[18] G. James, D. Witten, T. Hastie, and R. Tibshirani, *An Introduction to Statistical Learning: With Applications in R*. Springer Publishing Company, Incorporated, 2014.

[19] J. Lever, M. Krzywinski, and N. Altman, “Classification evaluation,” vol. 13, p. 603. [Online]. Available: https://doi.org/10.1038/nmeth.3945