ONLINE RECOGNITION OF ORAL ACTIVITIES

JAYA RUPINI VIJAYARAGAVAN

4783395

Delft University of Technology

August 2020

SUPERVISORS

Prof. Dr. Frans. C.T. van der Helm   Delft University of Technology

Prof. Dr. sc. ETH. Luigi M. Gallo   University of Zürich.

PREFACE

I have always been interested in biology since high school. I wanted to pursue an engineering degree that focussed on the human body and hence enrolled in a four years long B.E. Biomedical engineering programme. The programme gave me a very broad exposure to what all biomedical engineering encompasses. I discovered my interest towards research and development and hence looked for an MSc programme in TU Delft. I chose to pursue Musculoskeletal Biomechanics to deepen my knowledge on it. One of the subjects that interested me was Neuromechanics and motor control. The desire to implement the theory I learnt into practice led me to Prof. Luigi Gallo from University of Zürich, Switzerland. I was fortunate to have met Prof. Frans van der Helm, from the department, seeking his guidance for my thesis project. Prof. Gallo had been working on temporomandibular joint, its mechanics and joint disorders. Having a sibling who is a dentist, I was able to have a better clarity of what help I can be to dentistry. Prof. Gallo explained to me how temporomandibular disorders are widely prevalent and how dentists are trying to treat the different disorders in patients. But the most common scenario in a dental clinic was patients lining up with intense pain not knowing the reason for the pain. Oral activities during sleep could be one main reason. Usually, bed partners are the ones enquired about it but when even they can't answer, a separate monitoring device is needed to track the oral activities during sleep. The idea was to design a monitoring device that can be used by subjects during wakefulness and during sleep. The device had to record the oral activities performed by the subjects and alert them about the different activities, especially the ones damaging their temporomandibular joints or teeth, they have been performing. My thesis work involves developing a machine learning algorithm that can classify oral activities performed by subjects as the first step towards a real-time monitoring set up.

Contents

I. **Introduction** **1**

II. **Methodology** **2**

A. Recorder 2

B. Acquisition of data 2

C. Signal pre-processing 3

D. Feature extraction 3

E. Classification using SVM 4

III. **Results** **4**

A. Offline recognition 4

B. Online recognition 6

IV. **Discussion** **7**

V. **Limitation** **7**

VI. **Conclusion** **7**

Acknowledgement **8**

References **8**

Experimental protocol **9**

A. Wavelet Transform 10

B. Support Vector Machine 13

ONLINE RECOGNITION OF ORAL ACTIVITIES

Jaya Rupini Vijayaragavan
4783395
MSc Biomedical Engineering
Delft University of Technology

***Abstract*—Temporomandibular disorders (TMD) affect about 5-12 percentage of individuals with consequences such as jaw noises, clicking, myofascial pain, discomfort, limited mandibular range of motion and stress. Treatments depend on the cause and extent of the damage and part of the joint or jaw affected. When exact aetiology of TMD is unclear, generic treatments (splint therapy) are offered. Different oral activities performed on a daily-basis result in different loading conditions on the joint, possibly triggering TMD. These need to be investigated to know the usage of the masticatory system and the potential damage, in order to perform specific treatments. Our work aims at developing an online algorithm that can classify oral tasks performed by individuals. It can be used during daytime or overnight's sleep to see how often different activities are performed by subjects. A 4-stage wavelet decomposition was employed to the signals and then subjected to feature extraction to train a support vector machine algorithm with. The prediction accuracy was found to be 90 percent for a group of selected oral activities (static, jaw opening, chewing and maximal voluntary clenching). The algorithm had about 80 percent prediction accuracy when classifying both functional (chewing, jaw opening and static) and parafunctional activities (grinding, incisal biting, maximal voluntary clenching, protrusion and laterotrusion) together. However, 80 percent accuracy is regarded as a set back due to the lack of more data. On reviewing the recognised activities, further research on any overuse of muscles or loading on the jaw joint during each activity can be conducted to give specific treatment and therapy preventing any deteriorating actions. Thus, the developed classification algorithm works as a prototype for future studies on online recognition of oral activities.**

Key words: TMJ, TMD, myofascial pain, time-frequency analysis of sEMG, online classification.

I. INTRODUCTION

Some of the most complex joints in the human body are the temporomandibular joints (TMJ), commonly called jaw joints, situated on each side of the cranium. These twin joints connect the mandible to the temporal bone. Since they articulate the same body segment, i.e. the mandible, to the skull, they move together as a single unit. The TMJ is a very mobile joint providing a variety of functional orofacial activities, such as mastication and speech, as well as parafunctional activities (i.e. not necessary and not vital but possibly damaging), such as lip and cheek biting and bruxism including tooth grinding and clenching during sleep (sleep bruxism) or wakefulness (awake bruxism). An articular disc is connected to the mandible and the temporal bone by loose fibres, together forming articular capsule. The mandible can move with six degrees-of-freedom as the articular capsule is slack and articular surfaces are not rigid [1]. However, there are ligaments that together with other anatomical structures around TMJ restrict border movements of the jaw. Border movements are those parafunctional movements of the mandible at its maximum in a direction or a plane [2]. These ligaments supporting TMJ are collateral, temporomandibular, stylomandibular and sphenomandibular ligaments.

Continuous and intensified use of the TMJ leads to joint degradation and malfunction, often with muscular involvement, termed as temporomandibular disorder (TMD). It is commonly noted that there is a relationship between parafunctions and signs and symptoms of TMD as most parafunctions become unintentional habit overtime. Gavish et al. (2000) studied a parafunction, namely jaw play, small mandibular movements with no contact of the upper and lower teeth, in teenage girls in an attempt to see how a parafunction led to TMD [3]. They found that jaw play led to increased tension in the jaw and increased joint noises when compared to controls who did not perform this parafunction. Another most common parafunction associated with TMD is bruxism [4]. It is characterised by involuntary mandibular movements of grinding and/or clenching of the upper and lower teeth. Bite force in humans, a parameter to measure the general status of the masticatory muscles, during sleep bruxism events can be much larger than that recorded when the subject is awake. This is because the neuromuscular reflexes that protect humans from extreme and damaging actions, usually active during wakefulness, are suppressed during sleep. While clenching is a static activity, grinding is a dynamic maxillomandibular activity. During grinding, the upper and the lower jaws move to extreme positions, damaging the masticatory system. Thus, loading in the teeth and the muscles of mastication increases, causing detrimental effects on TMJ.

TMD includes osteoarthritis, orofacial pain, a disturbance of the internal arrangement of the joint components or a combination of these conditions at a time [5] [6]. Along with pain, TMD causes limitation of oral functions in daily activities. TMD is associated with health issues such as insomnia, headache, neck pain and in some cases, fibromyalgia [7] [8]. The disorder can be caused as a secondary health issue by conditions such as rheumatoid arthritis [9]. A whiplash injury in motor vehicle accidents can also result in TMD [10] [11]. Some of the most common consequences of TMD are jaw noises, clicking, myofascial pain, discomfort, limited mandibular range of motion and stress. Treatments depend on the extent of the damage, the cause that led to the damage and the part of the joint or jaw affected. However, the exact aetiology of TMD is difficult to explain and hence generic treatments, namely splint therapy, maybe offered with the limited knowledge. Different oral functions and parafunctions performed on a daily-basis result in different loading conditions on the joint, that can trigger TMD. On the onset of TMD, in certain patients compensatory oral movements and loading can be identified. Oral activities need to be investigated to know the usage of the masticatory system and the potential damage that could result, in order to perform specific treatments.

It has been reported by the National Institute of Dental and Craniofacial Research (NIDCR), one of the National Institutes of Health (NIH) of the United States of America, in September 2017 that the rough estimate of people suffering with TMD is about 10 million in US with a larger number of women than men [12]. To study the onset and progress of TMD, it is necessary to know the usage of the masticatory system by tracking different oral activities for more information about temporal and mechanical aspects of loading conditions on TMJ. In our study, only temporal aspect is considered. Among the different techniques of monitoring, recording the electromyogram of the muscles, such as temporalis, masseter and suprahyoid (a group of four muscles namely, mylohyoid, geniohyoid, stylohyoid and digastric), is chosen for this study.

Surface electromyography (sEMG) of the masticatory system has been considered the best non-invasive measurement, considering the comfort of the subjects. Recorded electromyography indicates the muscle activation of the system while performing different oral tasks. sEMG recording has been widely done during polysomnographic studies (PSG) conducted in laboratories. However, such studies are expensive and cause discomfort in some subjects. Gallo et al. (1997) designed biosignal recorders, e.g. an 8-bit single chip processor fitted in a box of dimension 95x58x25mm as an attempt towards making ambulatory measurements of EMG activities during sleep [13]. The study not only showed that the ambulatory measurement was in concordance with PSG, but also was cost effective and gave more insight on the subjects natural behaviour during night. Gallo et al. (1998) classified the different masticatory muscle activities, including both their functions and parafunctions using multivariate discriminant analysis, with a high κ-value (0.80) [14]. For an effective computation, an optimization technique is required rather than an analytical approach of discriminant analysis as this study.

This study aims at developing an online algorithm that can classify the oral tasks performed by individuals in an experimental set up. In our study, oral activities were recorded using a portable recorder with a sampling frequency of 2 kHz. It could be done during daytime or during an overnight's sleep to see how often different activities are performed by subjects. The data were then processed and analysed on a desktop. The support vector machine (SVM) algorithm used for this classification was developed offline using MATLAB. Data were collected from 8 healthy individuals during various recording sessions of performing a set of oral activities. A 4-stage wavelet decomposition was employed to obtain time-frequency information from the signals and then subjected it to feature extraction to train SVM algorithm with.

II. Methodology

A. Recorder

The portable EMG signal recorder employed here is a data logger system. The recorder was designed to monitor biosignals in daily life for research studies. The recorder is integrated with a band-pass filter that filters out 10 Hz to 400 Hz and an amplifier with a gain of 4000. The signal-to-noise ratio of the microcontroller used in the recorder is 67 dB. The data acquisition frequency of the recorder is set to 2 kHz. The recorded data is stored in a microSD memory card in .raw format. The memory card can be read later on desktop for offline signal visualisation and processing. For online recognition, the Bluetooth unit in the recorder is activated to transfer the data from the recorder to a device that can run MATLAB to perform pre-processing and classification of oral activities.

B. Acquisition of data

A total of 8 healthy subjects (6 women and 2 men) were recorded. Each subject was given a time slot of 25 minutes for the EMG recording. Once, the subject arrived, he/she was asked to exfoliate facial skin to reduce impedance. The subject was then seated on the dental chair and bilaterally palpated to find the temporalis, masseter and the suprahyoid muscles. The bellies of the muscles were located to place disposable surface electrodes (Ag/AgCl) clipped to the cables of the portable recorder. Thus, a total of 6 channels were recorded with the left mastoid process as the reference. Figure 1 shows a model of the portable recorder that was used (left top corner) and a profile of a subject under study.

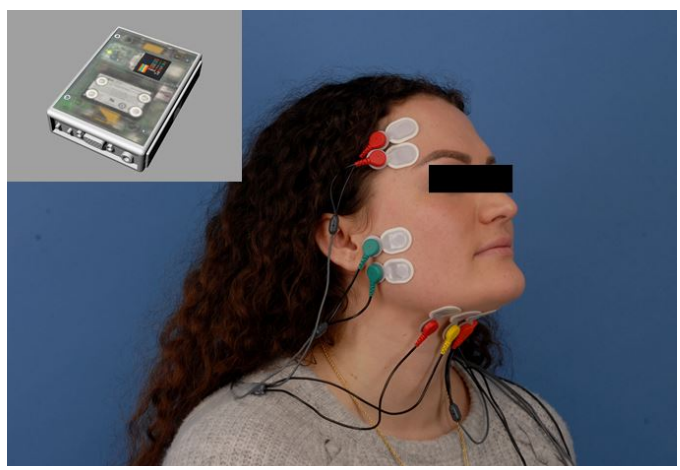

Figure 1: Electrode placement for recording subjects and the portable recorder in use (left top corner).

Oral activities namely, static, jaw opening, protrusion, laterotrusion right, laterotrusion left, incisal biting, grinding, maximal voluntary clenching, chewing gum on right and left side of the mouth were recorded 3 times for each activity and each subject. The following were the instructions provided to the subjects for each activity.

1) Static - Resting position of the jaw, where the teeth are not in contact.
2) Jaw opening - Opening the mouth as much as possible and returning to the resting posture.
3) Protrusion - Protruding the mandible to the maximum and returning to the resting posture.
4) Laterotrusion right - Moving the mandible as much as possible to the right and returning back to the resting posture.
5) Laterotrusion left - Moving the mandible as much as possible to the left and returning back to the resting posture.
6) Grinding - Beginning with bringing the upper and lower jaws to contact, grinding on the right once and the same on the left without losing teeth contact.
7) Incisal biting - Bringing the incisors of the upper and lower jaw in contact and pressing to the maximum for 3 seconds and then returning to the resting posture.
8) Maximal voluntary clenching - Clenching the jaws as much as possible for 3 seconds and returning to the resting posture.
9) Chewing gum right - A chewing gum was given to the subject to first break it down completely by chewing. Once the recording began, the gum was chewed on the right side for 10 cycles and the jaw was returned to the resting posture. It was noted that the generated saliva was swallowed only after the recording ended.
10) Chewing gum left - The chewed gum was taken to the left from the right. The 10 cycles of chewing on the left side were recorded.

Subjects were provided with the option of using cotton rolls to bite on during the incisal biting and maximal voluntary clenching activities in order to prevent damage to the teeth. Numbers were given to each task as a label for use in developing the classification algorithm. The recorded sEMG signal was stored as .raw file in the memory card of the portable recorder under the filename named after the date and time of each recording.

C. Signal pre-processing

The sEMG data from the memory card were loaded in a personal computer in which the rest of the research was carried out. MATLAB from MathWorks Ltd. was the software used to develop the algorithm. A study on the optimal bandwidth of surface EMG signals of facial, oral, neck and jaw muscles found the frequency range of muscles involved in oral activities to be between 25 Hz and 500 Hz [15]. The analog filters in the recorder filtered out the frequency range of 10 Hz to 400 Hz. A digital IIR Butterworth filter was used to remove the power line interference of 50 Hz. Since the sEMG signal can vary for the same muscle and same activity under different time instances and for the same and different subjects, the filtered signal was subjected to normalisation so that all values are within a single range for easy computation. The sEMG signal was normalised to mean amplitude for each activity and for each subject as suggested by Mark Halaki et al. (2012) [16]. Rectification of the signal was not employed here in order to avoid any alteration of frequency content of the recorded sEMG signals as suggested by Osmar Neto et al. (2010) [17].

D. Feature extraction

This research employed a 4-stage discrete wavelet transform (DWT) to extract both time and frequency information of the signal. DWT decomposes a signal into detail and approximation information by analysing different frequency bands of the signal with different resolutions. In this technique, the sEMG signal undergoes a series of low pass filtering and a high pass filtering and hence, a low pass branch and a high pass branch result at each stage of decomposition. The detail coefficients resulting from the high pass branch are noted while the low pass branch is further decomposed into low pass and high pass branches. The process is repeated at each stage of decomposition, giving information on how power of the signal changes over time. Thus, the high frequency components have good resolution in time and low frequency components have better frequency resolution.

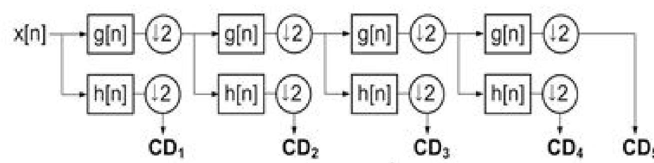

Figure 2: Schematic representation of 4-level wavelet decomposition (as presented by Orguc et al.) [18].

The mother wavelet used for this wavelet decomposition as the transformation function was Daubechies wavelet of order 7 (db7) as suggested by Orguc et al. (2018) [18]. The coefficients were grouped in subsets as CD_1-CD_5, as shown in Figure 2. Here, x[n] represents the sampled EMG signal (here, sample rate, 2 kHz) to be decomposed, upper branch is the low pass filter branch (g[n]) and lower is the high pass filter branch (h[n]). The low pass filter allows frequencies till half of the highest frequency of the signal and the high pass filter allows the other half. After each low pass filtering process, subsampling is done to eliminate half of the samples in accordance with Nyquist theorem, as the highest frequency is now half of the original highest frequency. Thus, signal information with half the time resolution and double the frequency resolution will result at each stage. CD_4 and CD_5 subsets were chosen for further computations as these two subsets have the least number of coefficients, reducing the computation cost but maintaining the accuracy that would have been obtained if all coefficient subsets were used, as studied by Orguc and colleagues. CD_4 and CD_5 subsets and the pre-processed sEMG signal (x) were then separately subjected to the calculation of the moving average value (MAV) and standard deviation (STD) using the formulae,

$$MAV_{CDi} = \frac{1}{M}\sum_{k=1}^{M}|CD_i|(k) \tag{1}$$

$$STD_{CDi} = \sqrt{\frac{1}{M-1}\sum_{k=1}^{M}|CD_i(k) - MAV_{CDi}|^2} \qquad (2)$$

where M is the length of the CD subsets.

Thus, the feature set was composed of 6 elements from each channel. The final feature set was computed for all the 6 channels in use to create a data set containing 36 elements for one recording of an oral activity.

E. Classification using SVM

A multiclass Support Vector Machine (SVM) is a supervised machine learning algorithm mostly used in classification and regression analysis. SVM is the most commonly used in classification problems because of its robustness [19] [20]. The goal of the algorithm is to define a hyperplane, a border that can distinguish the data set into respective classes. The input to the algorithm is the feature data pertaining to different classes and the output expected is a line or hyperplane that separates those classes. The data points closest to the hyperplane are called the support vectors. The shortest distance between the hyperplane and the support vectors is called the margin and the aim is to maximize this margin as much as possible. Optimal hyperplane is the one that has the maximum margin and is of interest for this research work. The support vectors are considered critical since they influence the position of the hyperplane the most. Hence, they are more important than the entire data set to define the hyperplane. This makes it clear that SVM provides an optimal solution with reduced computation cost in comparison with an analytical technique. The training data is of dimension pxn where p represents the total number of observations and n represents the number of dimensions of each feature vector. Thus hyperplane in n-dimensional Euclidean space can now be regarded as a flat plane cutting through the n-dimensional space. It is necessary to convert the data points to be linearly separable in higher dimensions and then project back the decision boundary to the original dimensions. Hence kernelling, a process of mapping data points into higher dimensions using mathematical functions called kernels, is employed. The most commonly used kernel function in sEMG analysis is radial basis function (RBF) with the formula

$$K(x, x') = exp(-\gamma * \|x - x'\|^2) \qquad (3)$$

The value K determines the influence of one feature vector (x) on the other (x'), representing the similarity between x and x'. Here $\|x - x'\|$ is the Euclidean distance between the feature vectors, x and x' and gamma (γ) is an important hyperparameter whose value determines the overfitting (increased γ) or underfitting (decreased γ) of the model. The RBF kernel gives the relationship of the two feature vectors in infinite dimensions. It doesn't transform the feature vectors to infinite dimensions but only calculate their relationship in infinite dimensions. Another such hyperparameter to be taken care of is the cost of misclassification in the training data set (C). This value determines if the model overfits (increased C) or underfits (decreased C). The best fitting γ and C values chosen for the final model would be the ones that gave the best cross validation accuracy.

This work uses LIBSVM, an open source machine learning library for SVM developed by National Taiwan University that can be incorporated on MATLAB [21]. It implements one-vs-one approach for multiclass problem. The data set was divided into training data and testing data. The training data was subjected to cross validation to estimate the performance of the SVM model. Among the different methods of cross validation, leave-one-out cross validation technique was performed because, here, the data set was small, the computations were inexpensive and almost the entire training data set could be used in each iteration. After performing cross validation, a grid search was performed to find the best parameters C and γ, to generate a model that resulted in the best accuracy. The generated model was tuned to incorporate these parameters and then subjected to testing using the testing data.

III. RESULTS

A. Offline recognition

Initially, the developed algorithm was subjected to binary classification taking different oral activities against the static position. On observing successful distinction, the algorithm was subjected to multi-class classification involving 4 oral activities, namely static, jaw opening, maximal voluntary clenching and chewing (right and left), keeping Orguc's work using single channel information as a reference. Our work however is bilateral with 6 channels and hence to test the robustness of the developed algorithm, the activities were grouped into 4, under the names Group 1, Group 2, Group 3 and Group 4.

- Group 1 - Static, jaw opening, maximal voluntary clenching and chewing right.
- Group 2 - Static, jaw opening, maximal voluntary clenching and chewing left.
- Group 3 - Static, jaw opening, maximal voluntary clenching, chewing right and chewing left.
- Group 4 - Static, jaw opening, maximal voluntary clenching, chewing right and chewing left together under one term, chewing.

The feature set of each recording of an activity had 36 attributes (6 attributes for each channel information). The cross validation accuracy of the training data was found for each of the above-mentioned groups as shown in Figure 3.

2 sets of data from each activity in each group were assigned to be test data and the rest in that group were used as training data for the model. The cross validation accuracy of the four groups were more than 80 percent because of unique muscle activation during each task recorded bilaterally. However, on comparing Group 3 with the other groups, there was a drop in the cross validation accuracy of the model because during chewing left and chewing right activities, masseter muscles on either side are the most activated in both the tasks. But

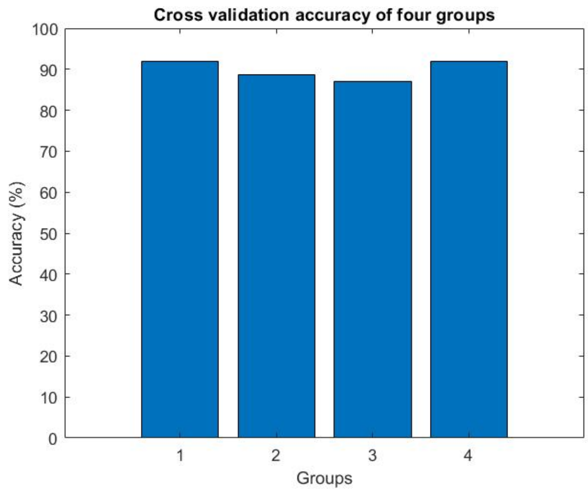

Figure 3: Cross validation accuracy in percentage for the different groups of oral activities considered.

when terming both the tasks under one umbrella term of chewing (Group 4), the cross validation accuracy increased. Thus, similar bilateral activation of muscles lead to confusion for the training SVM algorithm.

Table 1: Reference number for each oral activity recorded.

Activity number	Oral task performed
1	Static
2	Jaw opening
3	Protrusion
4	Laterotrusion right
5	Laterotrusion left
6	Grinding
7	Incisal biting
8	Maximal voluntary clenching
9	Chewing right
10	Chewing left

Table 1 is provided above as a recall for further references. Further, different combinations of these activities were grouped to test the robustness of the developed algorithm under different situations.

In order to find the performance of the classification algorithm, a confusion chart was created, as shown in Figure 4. 2 sets of data from each activity of Group 3 were assigned to be test data and the rest were used as training data for the model. The model had a prediction accuracy of 90 percent. It was seen that chewing right and chewing left tasks were confused and one reasoning could be that they have similar feature vectors.

In the confusion chart, the shades of blue represent correct classifications and shades of orange represent the wrong classifications. The number inside the box represents the number of data sets classified as a particular activity. The percentage of each classification shade is given adjacent to the confusion chart. Similarly, other activities were grouped keeping in mind similarities in muscle activation during different activities that could potentially reduce the robustness of the developed algorithm. These groups involve more parafunctions as compared to the previously discussed groups and are as follows:

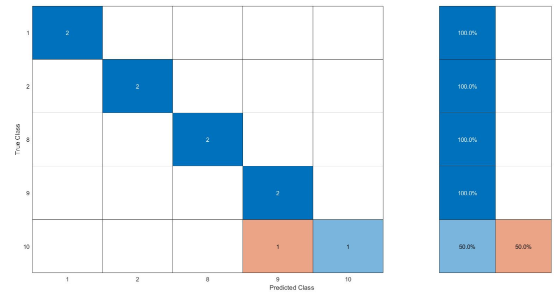

Figure 4: Confusion chart of Group 3 test data.

- Group 5 - Static, jaw opening, grinding, incisal biting, maximal voluntary clenching, chewing right and chewing left.

The above-mentioned activities were put under a single group as these activities are highly likely to be performed during daytime by a subject either as unintentional habits (grinding, incisal biting and maximal voluntary clenching) or as part of routine (static, jaw opening and chewing). The confusion chart for Group 5 is shown in Figure 5. One data set of incisal biting is wrongly classified as maximal voluntary clenching and one data set of chewing right as chewing left. The overall prediction accuracy for Group 5 is 86 percent.

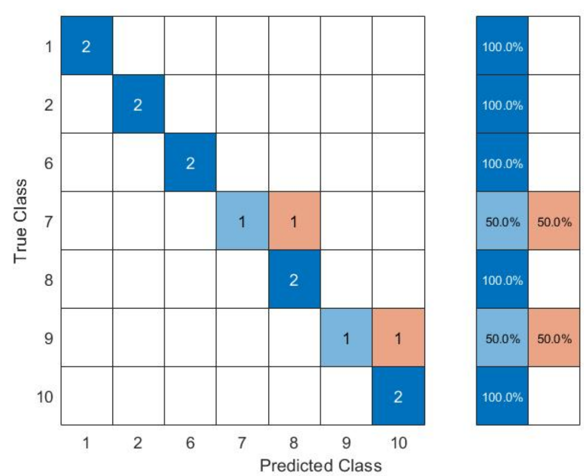

Figure 5: Confusion chart of Group 5 test data.

- Group 6- Static, jaw opening, grinding and incisal biting and maximal voluntary clenching.

Keeping in mind that the recorder could be used to track common parafunctions during sleep, the algorithm was trained with only the activities under Group 6. The prediction accuracy was found to be 80 percent. The confusion chart for Group 6 is shown in Figure 6. It could be seen that static was confused with jaw opening and chewing right with chewing left.

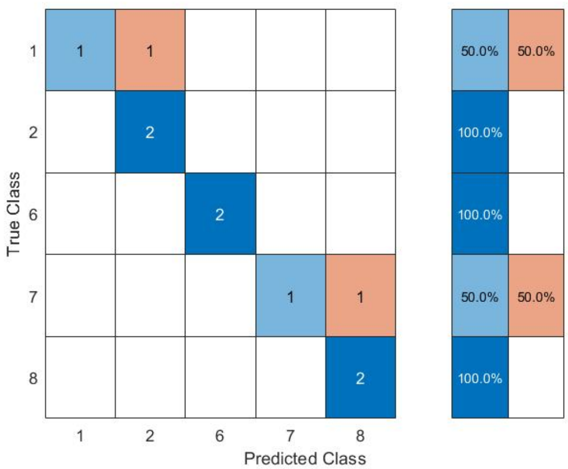

Figure 6: Confusion chart of Group 6 test data.

In order to see how the algorithm can distinguish among parafunctions that involve border movements, the following group was created.

- Group 7 - Protrusion, Laterotrusion left and laterotrusion right.

Group 7 was classified with a prediction accuracy of 67 percent. The confusion chart for Group 7 is shown in Figure 7.

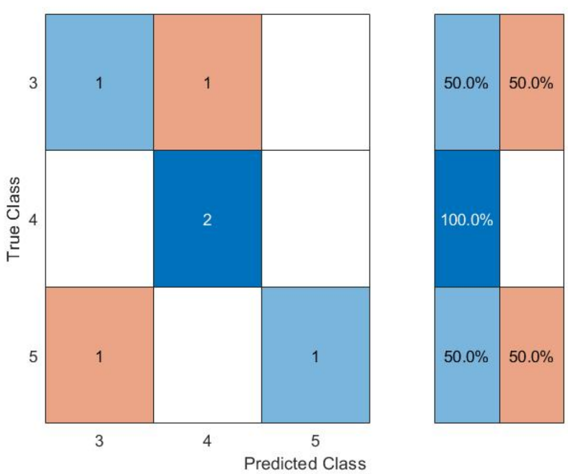

Figure 7: Confusion chart of Group 7 test data.

The temporalis, masseter, geniohyoid and digastric muscles on either side of the jaw are activated during protrusion. During laterotrusion, on the ipsilateral side, digastric, mylohyoid, geniohyoid and posterior fibres of the temporalis are active. Similar activation of muscles could be the reason for the confusion of protrusion with laterotrusion.

Finally, one group with all the oral activities recorded was subjected to classification to get an overall cross-validation accuracy of the developed model.

- Group 8 - All activities.

Finally, showed a low cross validation accuracy of about 58 percent mainly because of the increased number of activities the model was learning in comparison to the other groups. But still the algorithm was able to give a prediction accuracy of 80 percent on the test data, shown in Figure 8.

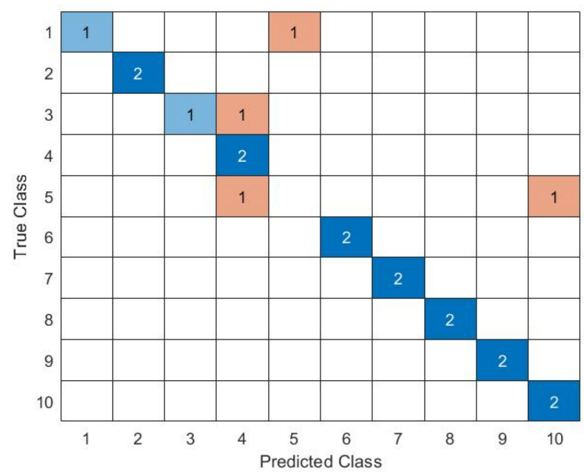

Figure 8: Confusion chart of Group 8 test data.

Looking at the wrong classifications, it can be seen that static was confused with laterotrusion left and protrusion with laterotrusion right. The laterotrusion left task was confused with laterotrusion right and chewing left.

In the end, a record of the classified oral activities can be derived to check if certain oral functions or parafunctions are being performed by the subjects too often or for prolonged duration that can potentially lead to overuse of TMJ and onset of TMD.

B. Online recognition

Theoretically, online recognition can be performed as the recorder transfers data to a computer connected to it via Bluetooth (115200 Baud) and can run MATLAB. The original signal was recorded at a sampling frequency of 2 kHz. However, in order to speed up the transfer of data via Bluetooth to a computer, it is necessary to down-sample the original signal, so that the new sampling frequency is 1 kHz, taking care to avoid aliasing. Another support vector machine algorithm was developed to train with the modified signals. To test the online recognition capability of the algorithm, a data set belonging to the chewing left task (duration of almost 10 seconds) was taken. The pre-processing and classification were made to run at different time intervals, e.g. every 1 second, every 2 seconds and every 5 seconds.

In the case of a classification every second, the algorithm was able to identify portions of static, maximal voluntary clenching, laterotrusion left in addition to guessing it could be chewing left. The sequence of predicted activity numbers per second was 5,10,10,10,10,8,1,10,8. The confusion chart is given in Figure 9.

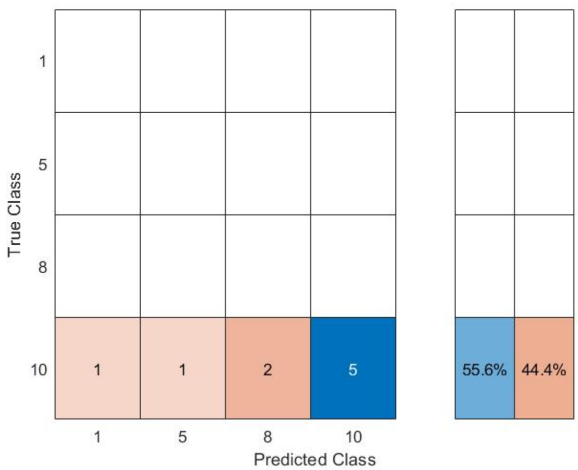

Figure 9: Confusion chart for chewing left task (1 second interval).

In the case of a classification every 2 seconds, the algorithm was able to identify that it could be chewing left or maximal voluntary clenching. The sequence of predicted activity numbers every 2 seconds was 10,10,8,10. The confusion chart is given in Figure 10.

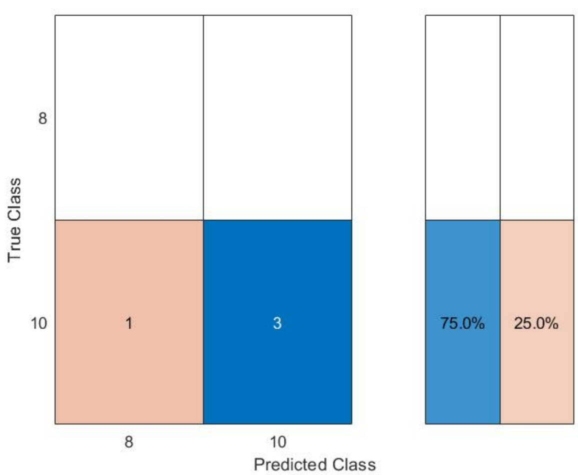

Figure 10: Confusion chart for chewing left task (2 seconds interval).

Finally, in the case of a classification every 5 seconds, the algorithm rightly classified the test signal to be chewing left. Thus, depending on the application, the time interval at which the algorithm has to classify the oral activities can be decided and implemented.

IV. Discussion

We present an online recognition of oral activities using 6 channel sEMG information from selected masticatory muscles. It has been developed to keep track of oral functions and parafunctions that can have deteriorating effects inducing discomfort and/or pain in the TMJ. Our work uses a portable recorder that can record sEMG signals from the oral muscles. Data pre-processing, feature extraction and classification of the oral activities are carried out in a separate computer. Wavelet analysis is employed to generate feature sets for each activity. Our work proposes to use only two coefficient subsets of DWT technique to reduce the computation cost. As 6 channels are used, there is more information available about the bilateral muscle activation, improving the accuracy of classification of different activities. An optimization algorithm (SVM) is chosen rather than an analytical technique to further reduce the computations. However, the algorithm has reduced accuracy in classifying certain activities. It confuses between chewing and clenching, static and jaw opening, incisal biting and clenching and chewing left and chewing right. The reason for this error could be similar muscle groups being activated during those activities or similarity in feature vectors. Nevertheless, the wrong classifications as shown in the confusion chart of Group 8 that cannot be explained citing similar muscle activation are of major concern. A conceivable explanation could be that the algorithm had outliers during training on those activities. One way to boost the performance of the algorithm would be to increase the training data. However with the available data, we were able to obtain an overall 80 percent accuracy and the classified activities can be saved in a report for further analysis on the loading conditions on the joint and usage of the oral muscles.

V. Limitations

One of the main limitations of this study is the number of subjects and hence of recordings. Usually, a machine learning algorithm requires huge data sets to train on. Similarly, we believe that the algorithm needs more training data to classify oral activities with better accuracy. Another limitation is that the work involved data recorded in a laboratory while the real challenge is to test the algorithm with field data, that is, data recorded from the subjects in their natural environment and not in a laboratory set up. By rectifying these limitations, it will be possible to have better classification results.

VI. Conclusion

The classified activities report generated at the end of the monitoring session can be reviewed to classify the oral activities into high-level, medium-level or low-level activities depending on the duration, type of activity performed and the loading conditions during such activities. The methodology in this work with the healthy subjects lays foundation to conduct a similar study in patient groups with similar TMD signs along with other monitoring techniques of interest. The algorithm developed is able to classify the data recorded in the laboratory set up. This work serves as a prototype which can further be modified according to the research goal. Future work would be to program the portable recorder as a stand alone wearable device that can classify oral activities in real-time and notify the subjects in their own environment about the recognised activities, so that necessary preventive measures can be taken.

ACKNOWLEDGEMENT

I would like to thank Prof. Dr. Frans CT van der Helm, TU Delft, The Netherlands and Prof. Dr. sc. ETH. Luigi M Gallo, University of Zurich, Switzerland for their timely support and mentorship in this work. I also thank the entire team under Prof. Gallo for helping me in the project with great team spirit and care.

REFERENCES

[1] Koolstra, J. H. (2002). Dynamics of the human masticatory system. Critical reviews in oral biology medicine, 13(4), 366-376.
[2] Yu, J. H., Lin, C. C., Chen, Y. H. (2017). Changes in mandibular border movements in adult patients after correction of functional anterior crossbite. Journal of Dental Sciences, 12(2), 139-144.
[3] Gavish, A., Halachmi, M., Winocur, E., Gazit, E. (2000). Oral habits and their association with signs and symptoms of temporomandibular disorders in adolescent girls. Journal of oral rehabilitation, 27(1), 22-32.
[4] Thie, N. M., Lavigne, G. J. (2001). Bruxism and orofacial movements during sleep. Dental, clinics of north america, 45(4), 657.
[5] Teachmeanatomy.info. (2019). The Temporomandibular Joint - Structure - Function - TeachMeAnatomy. [online] Available at: https://teachmeanatomy.info/head/joints/temporomandibular/
[6] Ingawale, S., Goswami, T. (2009). Temporomandibular joint: disorders, treatments, and biomechanics. Annals of biomedical engineering, 37(5), 976-996.
[7] Kim, T. Y., Shin, J. S., Lee, J., Lee, Y. J., Kim, M. R., Ahn, Y. J., ... Ha, I. H. (2015). Gender difference in associations between chronic temporomandibular disorders and general quality of life in Koreans: a cross-sectional study. PLoS One, 10(12), e0145002.
[8] Song, H. S., Shin, J. S., Lee, J., Lee, Y. J., Kim, M. R., Cho, J. H., ... Kim, S. (2018). Association between temporomandibular disorders, chronic diseases, and ophthalmologic and otolaryngologic disorders in Korean adults: A cross-sectional study. PloS one, 13(1), e0191336.
[9] Bagheri, Shahrokh C. Clinical Review of Oral and Maxillofacial Surgery: a Case-Based Approach. Elsevier, 2014.
[10] Peterson, C. Headache in Whiplash: A Comprehensive Overview, Elsevier (2011, March 29).
[11] Burgess, J. (1991). Symptom characteristics in TMD patients reporting blunt trauma and/or whiplash injury. Journal of Craniomandibular Disorders, 5(4).
[12] Bagis, B., Ayaz, E. A., Turgut, S., Durkan, R., zcan, M. (2012). Gender difference in prevalence of signs and symptoms of temporomandibular joint disorders: a retrospective study on 243 consecutive patients. International journal of medical sciences, 9(7), 539.
[13] Gallo, L. M., Lavigne, G., Rompre, P., Palla, S. (1997). Reliability of scoring EMG orofacial events: polysomnography compared with ambulatory recordings. Journal of sleep research, 6(4), 259-263.
[14] Gallo, L. M., Guerra, P. O., Palla, S. (1998). Automatic on-line one-channel recognition of masseter activity. Journal of dental research, 77(7), 1539-1546.
[15] Taha, Z., Deboucha, A., Ahamed, N. U., Ahmed, N., Ariffin, R., Ghazilla, R., ... Wong, Y. (2015). IIR Filter Order and Cut-off Frequency Influences on EMG Signal Smoothing.
[16] Halaki, M., Ginn, K. (2012). Normalization of EMG signals: to normalize or not to normalize and what to normalize to. Computational intelligence in electromyography analysis-a perspective on current applications and future challenges, 175-194.
[17] Neto, O. P., Christou, E. A. (2010). Rectification of the EMG signal impairs the identification of oscillatory input to the muscle. Journal of Neurophysiology, 103(2), 1093-1103.
[18] Orguc, S., Khurana, H. S., Stankovic, K. M., Leel, H. S., Chandrakasan, A. P. (2018, July). EMG-based real time facial gesture recognition for stress monitoring. In 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (pp. 2651-2654). IEEE.
[19] Crawford, B., Miller, K., Shenoy, P., Rao, R. (2005, July). Real-time classification of electromyographic signals for robotic control. In AAAI (Vol. 5, pp. 523-528).
[20] Kaur, G., Arora, A. S., Jain, V. K. (2009). Multi-class support vector machine classifier in EMG diagnosis. WSEAS Transactions on Signal Processing, 5(12), 379-389.
[21] Chih-Chung Chang and Chih-Jen Lin, LIBSVM : a library for support vector machines. ACM Transactions on Intelligent Systems and Technology, 2:27:1–27:27, 2011. Software available at http://www.csie.ntu.edu.tw/ cjlin/libsvm

Experiment Protocol

The subjects selected were given time slots 25 minutes each, totally spanned over 3 days for 8 subjects. Before beginning the recording, the subject was asked to wash the face off of any make-up and then exfoliate the skin. On making the subject sit on the dental chair, electrodes were placed on the muscles of interest, namely, masseter, temporalis and suprahyoid bilaterally according to the schematic given below in Figure 11.

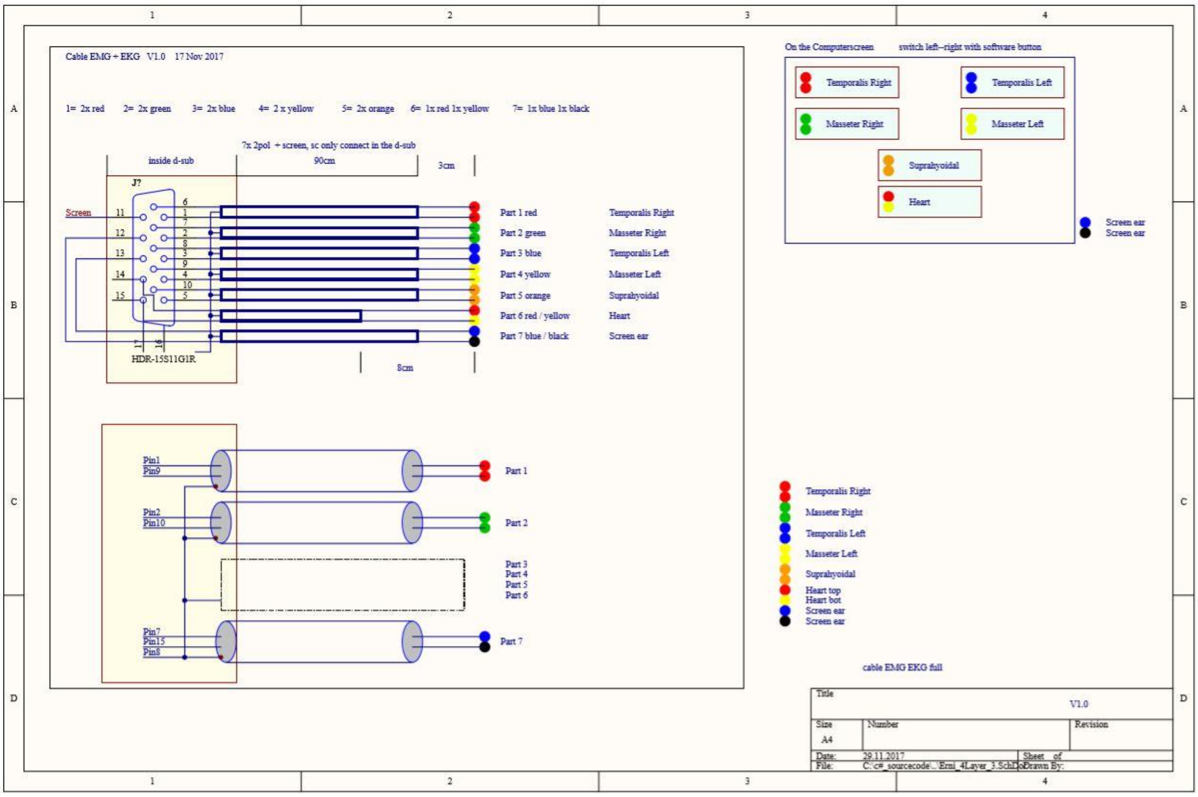

Figure 11: Guide for electrodes placement from the recorder cable.

The channel to be placed on the heart for EKG recording along with the suprahyoidal channel mentioned in the guide was made into 2 bilateral channels for recording the suprahyoid muscle. The part number in the guide represents the channel number.

The recorder has the facility to display each channel information as a bar with the respective colours shown in the guide. The bars move according to the amplitude of the EMG signals recorded at each channel. On placing the electrodes at desirable positions, the subject was made to perform a maximal voluntary clenching task once to see if the recorder can display the muscle activation of masseter and temporalis as these muscles are the most active during clenching. Adjustments are made if a particular channel is not able to pick the signal. Similar tests are done for jaw opening task to check suprahyoidal activation.

Once the channels were tested for picking signals up, instruction for each task was given followed by recording. The recorder was switched on, the activity was performed and recorded. Care was taken to see there was no swallowing and the activity was done as per instructions. Once finished with the activity, the recorder was switched off and the data was saved in a .raw file (named after the date and time of recording) on the memory card of the recorder. This process was repeated three times for each type of oral activity and hence, resulted in three recordings.

After recording the subjects, the data were transferred to a personal computer from the memory card of the recorder. All the recordings from all the subjects were grouped according to the activities performed. For example, a folder named ‘jaw opening’ had the recordings of jaw opening activity from all subjects. This was done for convenience in selecting files from MATLAB while training and testing the machine learning algorithm.

The flowchart summarizing the overall work done is given in Figure 12.

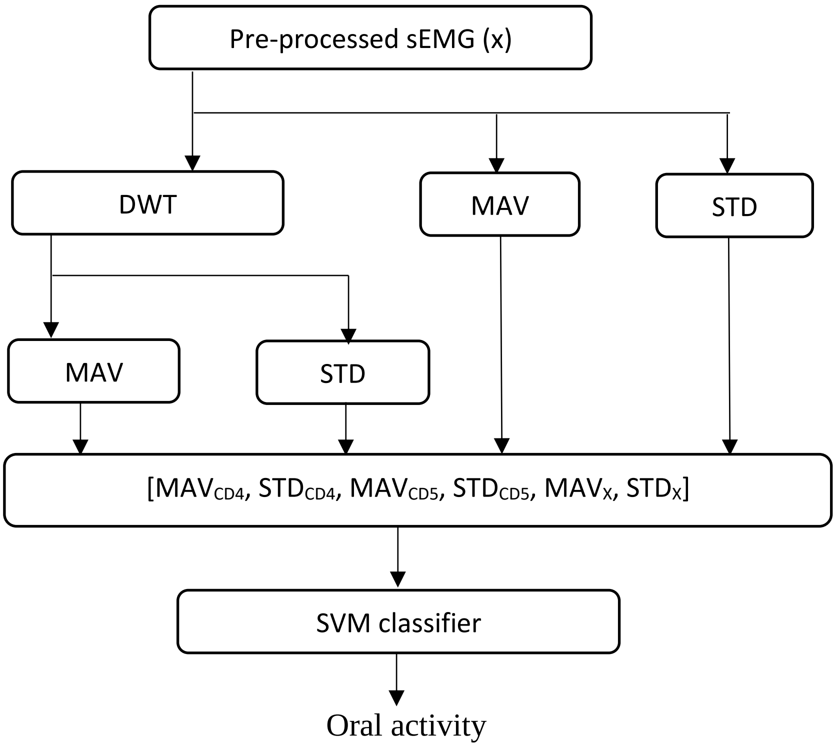

Figure 12: Flowchart summarizing the overall work done.

The pre-processed sEMG (x) signal here is the resultant signal after the removal of power line interference followed by mean normalization. DWT represents discrete wavelet transform of 4-stage performed. MAV and STD represent the Moving Average Value and Standard Deviation value of the two coefficients subsets, CD_4 and CD_5, given by MAV_{CD4}, STD_{CD4}, MAV_{CD5} and STD_{CD5} and the pre-processed signal, x, given by, MAV_X and STD_X. These form the feature sets that are the inputs to the trained Support Vector Machine (SVM) classifier. The output from the SVM classifier would be the oral activity recognised. During online recognition, the whole process from pre-processing till classification are done at regular intervals, for example, every 5 seconds.

A. Wavelet Transform

The idea behind a discrete wavelet transform, a time-frequency analysis technique, is to study the signal using different cut-off frequencies and different scales. The signal under study is subjected to a series of low pass and high pass filters to study the low frequencies and high frequencies respectively. The signal, x[n], is first passed through a low pass filter whose impulse response is given by g[n]. g[n] removes all the frequencies above half of the highest frequency of x[n]. Suppose our sampling frequency is 1 kHz, the highest frequency is assumed to be 500 Hz. Thus g[n] in the first stage of decomposition filters out frequency range from 0

Hz to 250 Hz. At the same time, a high pass filter, with impulse response h[n], is employed in the first stage which filters out the rest, from 250 Hz to 500 Hz. The filtered signals are subsampled by 2, reducing the data length. Suppose the original signal is 512-samples long. The g[n] and h[n] in the first stage of decomposition followed by subsampling will result in 256 samples each. Now since only half the frequencies are analysed in each branch of filtering, the frequency resolution is doubled and since only half the samples result from each branch, the time resolution is halved. This is repeated at each stage of decomposition. The g[n] in the next stage of decomposition filters out frequencies from 0 Hz to 125 Hz and h[n] from 125 Hz to 250 Hz and so on. The coefficients from the high pass filter branch are known as the detail coefficients and are noted down. The coefficients resulting from a low pass filter branch are called the approximation coefficients. These are then sent into the next level of decomposition into high and low pass branches. The maximum number of decomposition levels is determined by the length of the signal. Finally, the coefficients from the frequency range of interest can be separated from the rest for use, reducing the computation cost.

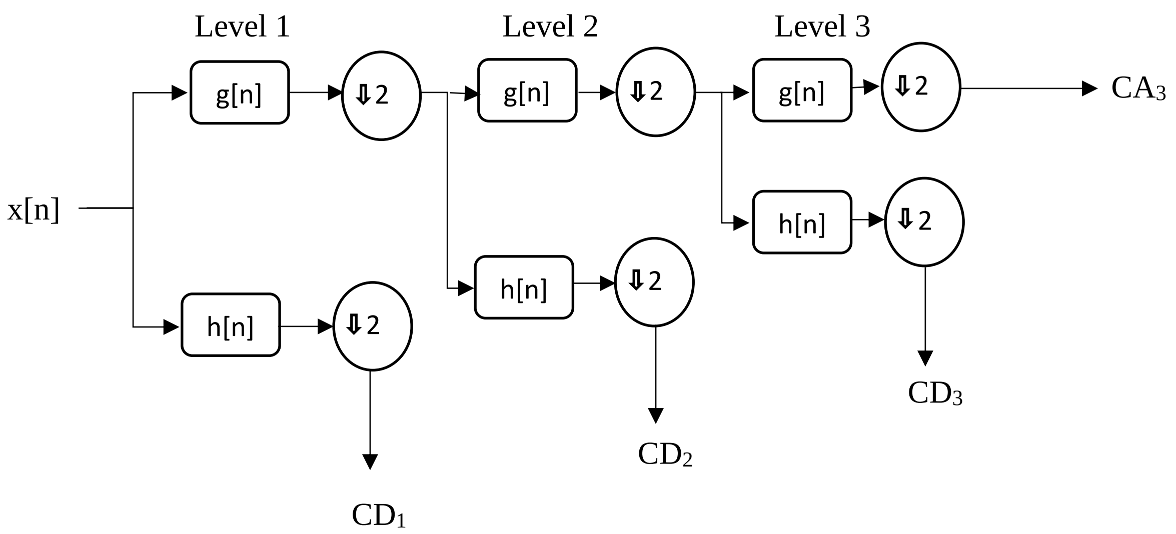

Figure 13: Representation of a 3-stage discrete wavelet transform.

For further understanding, 3-stage discrete wavelet transform is given in Figure 13. Here CD represent the detail coefficients and CA represent the approximation coefficients. When a wavelet decomposition is done on MATLAB, all these coefficient subsets are placed in ‘c’ and the lengths of each coefficient subsets are placed in ‘l’ in a particular order as shown in Figure 14.

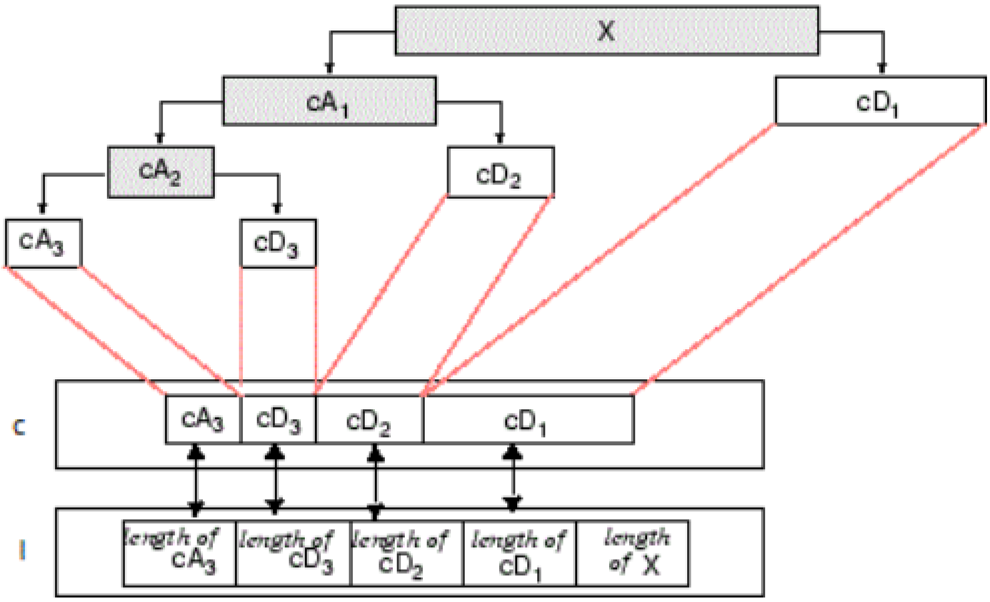

Figure 14: Storage of coefficient subsets in a 3-stage wavelet decomposition in MATLAB.

Here X is the original signal, cD the detail coefficients and cA the approximation coefficients. Care needs to be taken while selecting subsets needed for further computations. The idea behind a wavelet transform is the use of a basis function called mother wavelet which can be scaled and shifted to resemble the original signal. The mother wavelet used in our study is Daubechies 7 (db7) whose low pass (g[n]) and high pass decomposition filters (h[n]) are given in Figure 15.

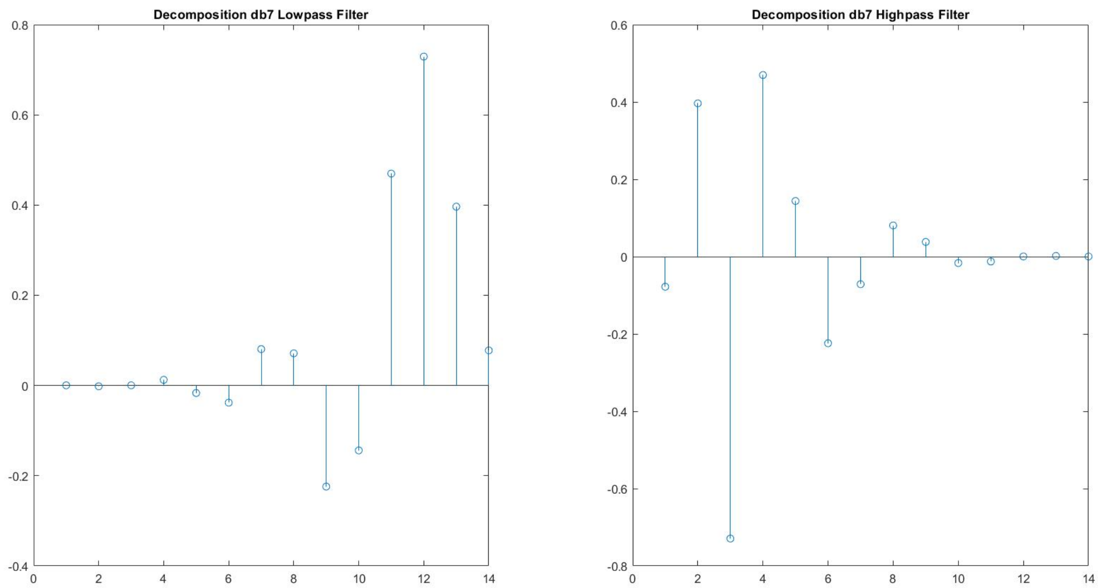

Figure 15: Daubechies (db7) low pass and high pass decomposition filters.

B. Support Vector Machine

Support Vector Machine (SVM) is gaining popularity among the other machine learning tools used for classification. SVM aims to define a hyperplane that can distinguish data from different classes. SVM was originally developed as a binary classifier, thus involving two classes only. In such a case, the hyperplane can be imagined as a line with a margin on either side. The data points closest to these margins are called the support vectors. What makes SVM interesting is that, instead of the entire training data set, SVM looks only for the support vectors in order to define the hyperplane. In an n-dimensional space, a hyperplane is n-1-dimensional. Thus, for example in a 2-dimensional space, a hyperplane is a line and in a 3-dimensional space, it is a plane. Beyond that, it is very difficult to imagine how a hyperplane will look. Our work involves 36-dimensional data. Thus, it is difficult to illustrate what exactly happens. The aim of the hyperplane is to separate the classes but in cases where such a separation is impossible, the data points are transformed to higher dimensions to see if in those new dimensions, the classes are separable, by a process called kernelling. The most commonly used kernel is the Radial Basis Function (RBF). The basic idea behind it is to find the influence on one feature vector on another considering the infinite dimensional inner product.

$$K(x, x') = \exp(-\gamma * ||x-x'||\hat{}2)$$

Here x and x' are the two feature vectors under study. Suppose there is a new feature vector to be classified. The feature vectors closest to this new feature vector have a lot of influence on it as compared to the ones farther away. The RBF kernel helps in finding this influence of other feature vectors on the new feature vector. The amount of influence the feature vectors, x and x', have on each other is a function of the squared distance between the two, ||x-x'||^2. Since gamma (γ) is multiplied, it acts like a scaling factor to the influence. Since RBF kernel works with infinite dimensions, K represents the relationship between the two feature vectors in infinite dimensions, that is, it doesn't transform the features into infinite dimensions but only represents the infinite dimension inner product to give a scalar determining the relationship between the two feature vectors in consideration.

LIBSVM is the SVM library we use in our work. It performs 'one-against-one' approach in multiclass classification. If there are p number of classes, there will be p(p-1)/2 number of binary classifiers where each classifier trains on two classes. Each binary classification is given a vote and the data point to be classified is assigned to a class with the maximum number of votes.

The equation of the hyperplane can be given by

$$w^T x_n + b$$

where w is the vector of weights, x_n is a data point and b is the bias. Suppose x_t is the data point under study and there are two classes, i and j, the optimization problem is given by the following equation as quoted in the LIBSVM guide.

$$\min_{w^{ij},\, b^{ij},\, \xi^{ij}} \ \tfrac{1}{2}(w^{ij})^T w^{ij} + C \sum_t (\xi^{ij})_t$$

$$\text{subject to } \ (w^{ij})^T \Phi(x_t) + b^{ij} \geq 1 - \xi^{ij}_t, \text{ if } x_t \text{ in the } i^{th} \text{ class,}$$

$$(w^{ij})^T \Phi(x_t) + b^{ij} \leq -1 + \xi^{ij}_t, \text{ if } x_t \text{ in the } j^{th} \text{ class,}$$

$$\xi^{ij}_t \geq 0.$$

Here $\Phi(x_t)$ maps x_t into a high-dimensional space, C is the regularisation parameter (cost of misclassification) and ξ_t is the slack variable introduced to skip a few outliers while trying to separate the data points into their respective classes. This is how a decision is made by the SVM classifier.