



A Comparative Analysis of Time-frequency Feature Extraction Techniques for Large Scale Electroencephalogram Data

Farzana Kabir Ahmad¹, Aysar Thamer Naser Tuaimah²

¹School of Computing, College of Arts and Science, Universiti Utara Malaysia, Sintok Kedah, Malaysia.

farzana58@uum.edu.my

²Department of Computer Science, Al Mustansiriya University, Baghdad, Iraq.

acer_era@yahoo.com

ABSTRACT

Recognition of human emotions is a fascinating research field that motivates many researchers to use various approaches, such as facial expression, speech or gesture of the body. Electroencephalogram (EEG) is another approach of recognizing human emotion through brain signals and has offered promising findings. Although EEG signals provide detail information on human emotional states, the analysis of non-linear and chaotic characteristics of EEG signals is a substantial problem. The main challenge remains in analyzing EEG signals to extract relevant features in order to achieve optimum classification performance. Various feature extraction methods have been developed by researchers, which mainly can be categorized under time, frequency or time-frequency based feature extraction methods. Yet, there are numerous setting that could affect the performance of any model. In this paper, we investigated the performance of Discrete Wavelet Transform (DWT) and Discrete Wavelet Packet Transform (DWPT), which are time-frequency domain methods using Support Vector Machine (SVM) and k-Nearest Neighbor (KNN) classification techniques. Different SVM kernel functions and distance metrics of KNN are tested in this study by using subject-dependent and subject-independent approaches. The experiment is implemented using publicly available DEAP dataset. The experimental results show that DWT is mostly suitable with weighted KNN classifier while DWPT reported better results when tested using Linear SVM classifier to accurately classify the EEG signals on subject-dependent approach. Consistent results are observed for DWT-KNN on subject-independent approach, however SVM works better in the setting of quadratic kernel functions. These results indicate that further investigation is significant to examine the impact of different setting of methods in analyzing large scale of EEG data.

Key words: Emotion recognition, Electroencephalogram Feature extraction, DWT, DWPT, Subject-independent, Subject-dependent

1. INTRODUCTION

Human emotion is mainly associated with thought, feeling, and behavioural responses and play a vital role in decision making, communication and learning process. Emotion recognition is a fascinating research field that aims to get insight into human emotional states and gives the ability to recognize them properly. The need of computers to be able to detect and interact with user's emotion states is a growing research interest [1]. Last decades, many researchers have attempted to recognize human emotion through various approaches, such as facial expression, speech recognition or gesture of the body [2]. However, these approaches were considered unreliable since users can control their facial expressions, prosody or body movements which could affect the ultimate emotion recognition results.

Human emotion recognition by using brain signals has become an emerging research area. EEG is one of the approaches of recognizing human emotion through brain signals. Such signals are recorded and monitored based on the brain electricity that is generated from electrodes places on human scalp. Signals obtained from EEG have been known to provide effective information on both mental and emotional activities as it measures brain waves produce when neurons are communicating. In the past, EEG device is widely used in medical domain typically to detect epilepsy and seizures [3]. Since the availability of wireless EEG that is portable, easy to operate and inexpensive, it has attracted many recent researchers to use EEG in recognizing human emotion [4]. In addition, several public repositories have been established to permit researchers to work on analysing and enhancing techniques that could handle large scale of EEG signals data.

Although EEG signals provide detail information about human emotional states, it is not quite easy to identify the user's mental state from the same signals.

The nature of EEG signals that are complicated, unsteady, nonlinear and random is mainly caused by the complex way of neurons interconnection. Hence, the analysis of non-linear and chaotic characteristics of EEG signals is a substantial problem. As a result, a thorough analysis is required to process and analyse such complex EEG signals. Moreover, the EEG signals are subject to high dimensionality features, which require further attention. The process of EEG classification starts with the stimulus that provokes brain signals which is later recorded. The recorded EEG signals undergoes a pre-processing phase in which the noisy data is eliminated and ready to be used as an input for feature extraction phase. The features that are extracted are then used to classify the signals.

Emotion recognition is built either based on discrete emotion model or valence-arousal emotion model. Discrete emotion model is a model-built base on a discrete emotion theory that suggests that there is a minimal number of fundamental emotions such as, surprise interest, joy, rage, fear, disgust, shame, and anguish. Valence-arousal emotion model on the other hand, is a two main dimensions for emotions: Valence (Pleasant-ness) and Arousal (low/high intensity of perception) [5]. Using this model, emotional states are located on two dimensions which are valence as horizontal axe and arousal as vertical axe.

The dimension of valence ranges from unpleasant (e.g., fearful- negative valence and high arousal) to pleasant (e.g., happy – positive valence and high arousal), while in the context of arousal, it ranges from inactive (e.g., calm – low arousal and positive valence) to active (e.g., anxious – high arousal and negative valence). Since the valence-arousal model provides more comprehensive information on emotion recognition [6], this model is used in the context of this research.

In addition, studies on EEG-based emotion recognition can be divided into subject-dependent and subject-independent approaches. In the subject-dependent approach, the data from each subject used to train the classifier, whereas, in the subject-independent approach, the data from all the subjects are used to train the classifier. Since there are so many variations in emotion recognition model, the performance of feature extraction techniques in handling large-scale EEG data makes it incomparable and complicated. This study aims to conduct comparative analysis of time-frequency feature extraction techniques namely DWT and DWPT on KNN and SVM using different setting to comprehend it outcomes on valence-arousal model. This study also examined the results

of emotion recognition in the context of subject-dependent and subject-independent approaches. The comparative analysis of DWT and DWPT is tested on publicly available DEAP dataset. The following sections are organized as follows; Section 2 explains on the related work, while Section 3 enlightens on the methodology used in this research. Experimental results and its corresponding discussion are presented in Section 4. Finally, concluding annotations is given in Section 5.

2. RELATED WORK

Crucial features of EEG signals assist in identifying different emotions of human. The process of extracting features is called feature extraction, in which some algorithms are applied to allow the extraction of the concealed information's in the EEG signals. This process is considered as the most important process in the EEG signal processing. The objective of feature extraction is to transform the pre-processed EEG signals into different function vectors that characterize different consciousness tasks.

Classification accuracy of human emotions is highly impacted by the feature extraction methods. There are three methods for feature extraction which are (1) time domain analysis, (2) frequency domain analysis and (3) time-frequency domain analysis. Time domain analysis using statistical parameters such as mean, variance, power and standard deviation. The other time domain analysis techniques are Hjorth parameters, Fractal Dimension (FD) and Event Related Potentials (ERP). In the context of frequency domain analysis, features are extracted based on band power using Fourier Transform (FT) and Discrete Fourier Transform (DFT) or by calculating the Power Spectral Density (PSD). The third approach is a combination of the above approaches known as Time-Frequency domain analysis. In this method features such as energy and entropy can extract by applying techniques like Discrete Wavelet Transform (DWT) or Discrete Wavelet Transform (DWPT).

Table 2.1 shows the comparative analysis of various feature extraction methods which have been tested on different classifiers. Some studies have run their own experiment set-up to obtain the EEG signal, like [4] [7][14], while others have used existing public dataset like DEAP. Although few studies have tested different feature extraction methods on DEAP dataset, their results are not comparable since they are implemented in various setting form, e.g., in term of subject-dependent or subject-independent approach, and different classifiers.

Previous research has tested Wavelet Transform feature extraction technique on different classifiers such as SVM, KNN, Naïve Bayes and Random Forest [8]. The study is based

on subject dependent approach, whereby the data for each subject were analyzed and classified. On the other hand, [9] have tested 14 feature extraction techniques on SVM and Random Forest using subject-independent approach. However, this study does not include DWPT in their

experiment testing. Furthermore, both studies applied Principal Component Analysis (PCA) as a method of feature reduction, which eventually lead to a smaller number of features and has led to information loss.

Table 2.1. Comparative Analysis Studies

<i>Author</i>	<i>Domain</i>	<i>Dataset</i>	<i>No. of Channel</i>	<i>No. of Trails</i>	<i>No. Subject s</i>	<i>Approache s (SID or SD)</i>	<i>Feature extraction (FE) methods</i>	<i>Classifier</i>
[4]	Steady state visually evoked potentials (SSVEP)	Experiment based	2	NP	NP	NP	✓ Principle Component Analysis (PCA) ✓ Independent Component Analysis (ICA) Linear Discriminant Analysis (LDA).	SVM KNN
[7]	Steady state visually evoked potentials (SSVEP)	Experiment based	16	8	7	SD	✓ Filter bank ✓ Short Fourier Transform ✓ Welch's method	LDA. SVM. Extreme Learning Machine (ELM)
[8]	Secondary data	DEAP	32	40	32	SD	✓ Wavelet Transform	SVM KNN Native Bayes Random Forest
[9]	Secondary data	DEAP	32	40	32	SID	✓ 14 FE methods	SVM Random Forest
[14]	Epileptic Disorder	Experiment based	NP	NP	4	NP	✓ Wavelet Transform ✓ Wavelet Packet Transform	NP

*NP: Not Provided; SID: Subject-Independent; SD: Subject-Dependent

3. METHODOLOGY

This study used an experimental research design to compare the performance of DWT and DWPT techniques on SVM and KNN classifiers. In addition, the performance of DWT and DWPT are also measured in the context of subject-dependent and subject-independent approach. This study is implemented in MATLAB 2018b environment using wavelet processing toolbox.

3.1 Research Framework

Figure 3.1 shows the proposed research framework of this study. This experimental research consists of three main phases, namely (1) Data Acquisition (2) Feature Extraction and (3) Classification. The following sections elaborate in detail each of these phases.

3.1.1 Phase 1- Data Acquisition

This research is implemented on the DEAP EEG dataset, which was built by Queens Mary University of London. This data set can be retrieved from

<http://www.eecs.qmul.ac.uk/mmv/datasets/deap/index.html>. The DEAP database has been pre-processed whereby the data is sampled from 512Hz decreasing to 128Hz. The electrooculographic (EOG) objects are removed by applying a band-pass filter, resulting signals with 4-45Hz frequency range. The EEG signals are recorded from 32 subjects using music-video as a stimulus to induce subjects' emotions. The 32 EEG electrode channels are based on Bio semi-Active.

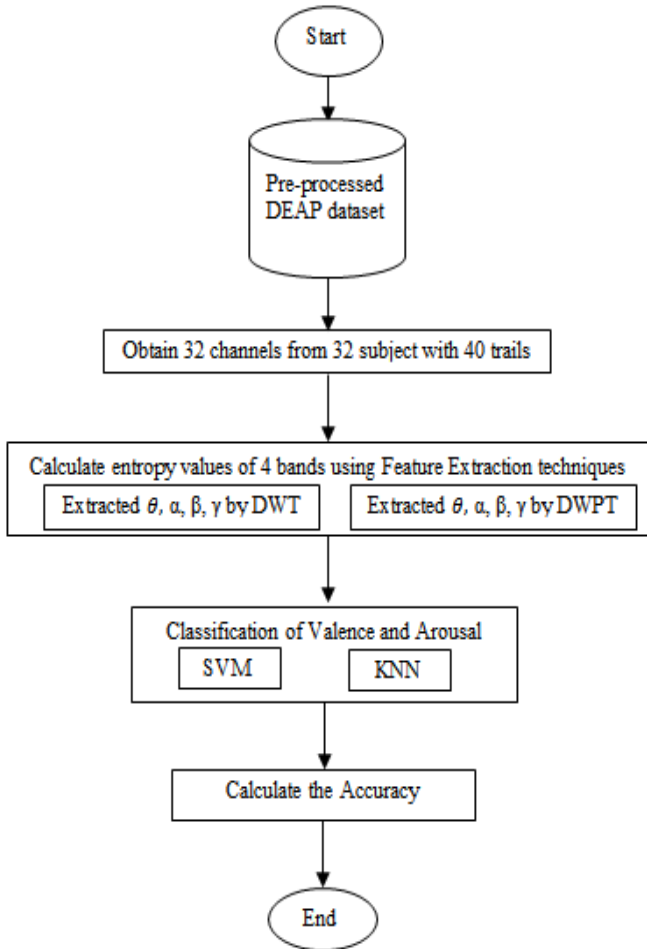


Figure 3.1: Research Framework

3.1.2 Phase 2- Feature Extraction Technique

Phase 2 on the other hand aims to extract the signals bands, delta (δ), theta (θ), alpha (α), beta (β), and gamma (γ) from the EEG DEAP dataset. According to the literature review, there waveforms that occur mainly within five frequency bands in the EEG spectrum which are δ (<4 Hz), θ (4–8 Hz), α (8–13 Hz), and β (13–30 Hz), γ (>31 Hz). Since DEAP dataset only consists of frequency from 4-45Hz, the delta δ band cannot be extracted in this study.

DWT and DWPT are two feature extraction methods that belong to time-frequency domain. This study proposed to use DWT and DWPT feature extraction techniques in order to extract signals and conduct the comparative analysis.

According to prior studies, there is no work have been done in comparing these two techniques on DEAP dataset, which has motivated this study. The obtained feature space from each adopted data subject includes 163,840 features. The feature space is illustrated in Table 3.1. The extracted features space contains high number of features which are originated from the facts that $(32 \times 4 \times 40 \times 32)$; number of electrodes \times number of frequency band \times number of the trails per subject \times number of subjects).

Table 3.1: The Extracted Features Space.

Channel Trial	Channel 1				Channel 32			
	Band (entropy)				Band (entropy)	Band (entropy)			
Tria 11	θ	α	β	γ		θ	α	β	γ
.
.
Tria 140
.
Tria 11	θ	α	β	γ		θ	α	β	γ
.
.
Tria 140

This research used entropy as a feature. Entropy is a numerical measure of the randomness of a signal, and it is the statistical descriptor of the variability within the EEG signals. It is also a strong feature representation for the emotion classification [10].

3.1.2.1 Discrete Wavelet Transform

The DWT is a linear conversion of data in the form of vector, within the length power of two, resulting separated frequency components. Later each component is studied with resolution matched to its scale. The process of computing DWT is accomplished with a filtering cascade followed by a factor 2 sub-sampling as shown in Figure 3.2, where A is an approximation coefficient and D is a detail coefficient value.

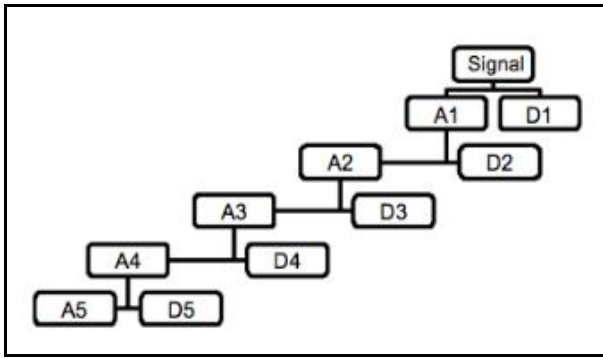


Figure 3.2: Discrete Wavelet Tree Decomposition

Figure 3.3 on the other hand shows the high and low filter pass filter mechanism for DWT decomposition tree. *H* signifies high-pass filter and *L* signifies low-pass filter, meanwhile $\downarrow 2$ represents sub-sampling. Results of these filters are acquired by the two following equations.

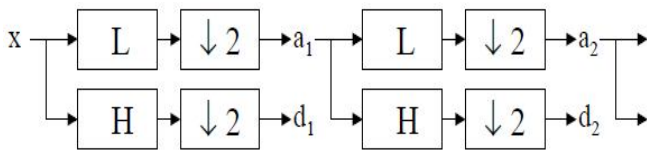


Figure 3.3: High and Low Pass Filter Mechanism of DWT Decomposition

Factors a_j are used in the later scale in the transformation, while factors d_j known as wavelet coefficients, predicts output of the transform. Low-pass $l(n)$ and high-pass $h(n)$ are both – coefficients used as filters.

$$a_{j+1}[p] = \sum_{n=-\infty}^{+\infty} l[n - 2p]a_j[n] \quad (3.1)$$

$$d_{j+1}[p] = \sum_{n=-\infty}^{+\infty} h[n - 2p]a_j[n] \quad (3.2)$$

3.1.2.2 Discrete Wavelet Packet Transform

DWPT is another feature extraction technique that belongs to time-frequency domain. Distinct from DWT, DPWT is able to decompose the signal details further, while the DWT only decompose the resulted approximation signals in the form of new detailed and approximated signals, as illustrated in Figure 3.2. One of the advantages of DWPT is that DWPT provides an efficient localisation of frequency band and has ability to produce approximation and detail coefficients (refer to Figure 3.4), different from the DWT algorithm as presented in Figure 3.2. The extraction of signal using DWPT results in the following decomposition tree (Figure 3.4), where this technique takes into account both A and D coefficient signals.

The lowpass filter process the enter signal $L(m)$ as well as the highpass filter $H(m)$ to get both estimate and detail coefficients. In reconsideration, the coefficients of estimate are applied as inputs for repeated filtering. The algorithm used here is designed to stop when one coefficient represents the final estimate which eventually represents the wavelet decomposition. The EEG signal is in the range of 4-45 Hz and five level of decomposition has been achieved. This decomposition tree is mainly used to extract the four bands of signals, θ (4–8 Hz), α (8–13 Hz), and β (13–30 Hz), γ (>31Hz). For instance, the yellow box in the Figure 3.7 indicates for theta wave (4–8 Hz). The resulting entropy values form the feature vectors that will be used as input for emotion classification.

Detail explanation of SVM and KNN are given in the following section along with performance matrix.

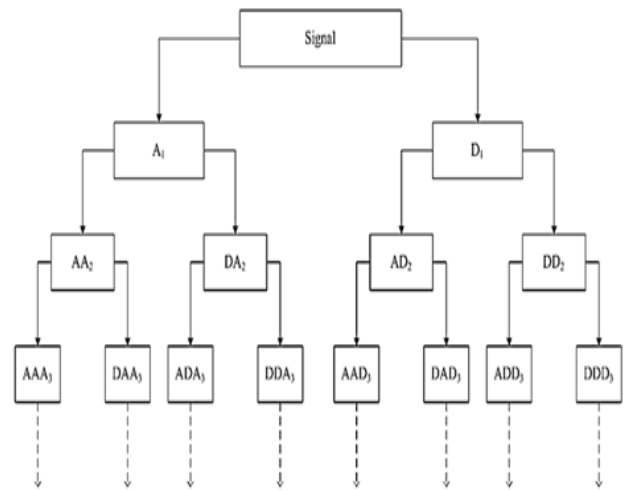


Figure 3.4: Three Level DWPT Decomposition Tree

3.1.3 Phase 3- Classification Techniques

In the EEG-based emotion recognition, after obtaining feature space from the feature extraction techniques, it needs to be classified through the classification phase. Numerous classification algorithms have been used to classify human emotion, which include Bayesian, LDA, Neural Network and other classifiers. This study used Support Vector Machine (SVM) and k-Nearest Neighbor (KNN) to classify emotion into dimensional model of arousal and valence. Different kernel functions of SVM and distance metrics of KNN are tested in this study. The detail explanation of SVM and KNN is provided in the following sections. The experiment outputs are measured using accuracy. The accuracy results are computed using Equation 3.3 where True Negative (TN), True Positive (TP), False Negative (FN) and False Positive (FP) are declared.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3.3)$$

3.1.3.1 Support Vector Machine Classifier

SVM is known to be a supervised learning technique which uses a discriminant hyperplane to identify classes. The chosen hyperplane increases the margins, indicating the area length from the closest training points of two diverse classes. SVM is popular in data mining tasks mainly when dealing with high dimensional feature space. The optimal classifier function can be derived as in the following equation by utilizing the Lagrange transformation:

$$f(x) = \text{sgn} \left(\sum_{i=1}^1 a_i^* y_i K(x, x_i) + b \right) \quad (3.4)$$

where i is the training sample eigenvector, and x is the recognising sample eigenvector, and present the Lagrange operator, and for the kernel function $K(x, x_i) = \varphi(x_i) \cdot \varphi(x)$. Six kernel functions have been used in this study such as linear, quadratic, cubic, fine Gaussian, medium Gaussian and coarse Gaussian.

3.1.3.2 K-Nearest Neighbor Classifier

KNN is a common machine learning algorithm that is used to in classification problem. In this algorithm, the number of integer k is determined through the distance measurement on the input feature vector training samples. Data is classified computationally by identifying the most mutual class based on number of the k nearest neighbors.

MATLAB provided six diverse distance functions of KNN that are applied to classify data; such as Fine, Medium and Coarse KNN algorithms which use Euclidian distance to find the nearest neighbors whereas the rest depend of different distance metrics.

Their details as given in MATLAB functions are illustrated in Table 3.2 below.

Table 3.2: KNN Distance Metric Details

Number	Distance Metric	Description
i.	Fine KNN	A nearest neighbor classifier that makes finely detailed distinctions between classes with the number of neighbors set to 1.
ii.	Medium	A nearest neighbor classifier that makes fewer distinctions than a Fine

	KNN	KNN with the number of neighbors set to 10.
iii.	Coarse KNN	A nearest neighbor classifier that makes coarse distinction between classes, with the number of neighbors set to 100.
iv.	Cosine KNN	A nearest neighbor classifier that uses the cosine distance metric.
v.	Cubic KNN	A nearest neighbor classifier that uses the cubic distance metric.
vi.	Weighted KNN	A nearest neighbor classifier that uses distance weighting

4. EXPERIMENTAL RESULTS AND DISCUSSION

In this section, the results of experimental are presented. This study has compared the performance of DWT and DWPT using two different classifiers namely SVM and KNN. In addition, we also ran these two feature extraction methods on subject-dependent and subject-independent approach. The main difference of both approaches is in the classification step. The subject-dependent trained a classifier for each subject whereas subject-independent trained a classifier using all subject at once.

In order to select the appropriate functions of SVM and KNN classifiers, an additional experiment has been performed. This comparative analysis has been conducted on the DEAP dataset with five-fold cross validation. The results are measured in terms of accuracy for valence-arousal model. The entropy features of the theta (θ), alpha (α), beta (β), and gamma (γ) bands through the 32 EEG electrode channels (Fp1, AF3, F3, F7, FC5, FC1, C3,T7, CP5, CP1, P3, P7, PO3, O1, Oz, Pz, Fp2, AF4, Fz, F4, F8, FC6, FC2, Cz, C4, T8, CP6, CP2, P4, P8, PO4, and O2) were extracted using DWT and DWPT decomposition.

Figure 4.1 and Figure 4.2 show an example of the entropy values of alpha band from Fp1 for the first subject. These values of theta (θ), alpha (α), beta (β), and gamma (γ) bands are later used to classify the human emotion into valence and arousal dimension. In this study, the value valence, in which 0 to ≤ 5 has been discretized to -1, while > 5 has been converted to 1. The identical transformation has been performed on arousal values as well.

DWT					
Linear SVM		Weighted KNN		Cosine KNN	
V	A	V	A	V	A
65.9%	64.9%	65.5%	65.3%	63.6%	64.5%
DWPT					
Quadratic SVM		Weighted KNN		Cosine KNN	
V	A	V	A	V	A
60.1%	62.8%	60.3%	62.8%	58.9%	60.3%
DWPT					
Quadratic SVM		Weighted KNN		Cosine KNN	
V	A	V	A	V	A
61.3%	62.9%	59.3%	61.8%	55.8%	60.5%

Figure 4.2: Entropy values for alpha band from Fp1 for Subject 1 using DWPT

Table 4.1 and Table 4.2 illustrate the summarized comparative analysis of DWT and DWPT on subject-dependent and subject-independent approaches using SVM and KNN classifier. Detail description is given below.

Table 4.1: DWT and DWPT based on Subject -dependent Approach

Table 4.2: DWT and DWPT based on Subject-independent Approach

4.1 Subject-dependent Approach on SVM Classifier

Figure 4.3 and Figure 4.4 present the results of subject dependent approach using DWT and DWPT techniques on different kernel functions of SVM classifier. Six SVM kernel functions have been tested in this experiment which are, (1) linear; (2) quadratic; (3) cubic; (4) fine Gaussian; (5) medium Gaussian; (6) coarse Gaussian. Based on Table 4.1 DWPT has obtained better accuracy in relative to DWT techniques. The results have showed that DWPT outperformed DWT in which it has achieved 66.41% of mean accuracy in valence and 66.05% of mean accuracy in arousal. Meanwhile DWT decomposition has achieved higher mean accuracy of 65.9% and 64.9% for valence and arousal respectively. These results are mainly due to the fact that DPWT can further decompose the signals into new approximation and detailed signal whereas the DWT only resulting approximation signals. In addition, DWPT is a more appropriate technique to extract EEG features as it provides efficient localization of frequency bands [11]. Hence, DWPT is preference over DWT technique.

Interestingly, it has been also noted that both feature extraction techniques work well on linear SVM kernel. This is mainly due to the linear SVM that is less prone to overfitting than non-linear functions and its ability to work with large number of features [12]. In this experiment, about 163,480 features have been generated to be trained by the classifier, which make the linear SVM performs better. Moreover, this result is aligned with other existing works [13] which have reported that linear SVM kernel works well on EEG data set.

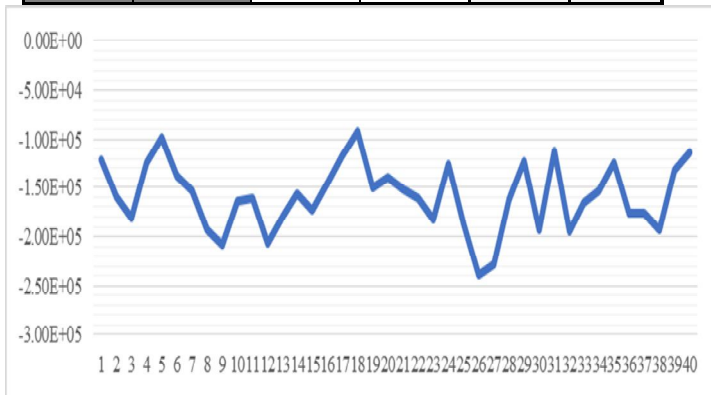


Figure 4.1: Entropy values for alpha band from Fp1 for Subject 1 using DWT



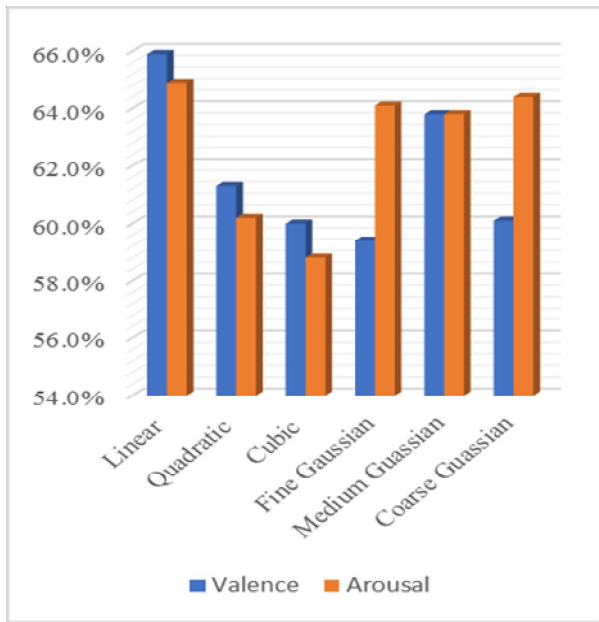


Figure 4.3: Mean accuracy of all SVM kernel functions using DWT

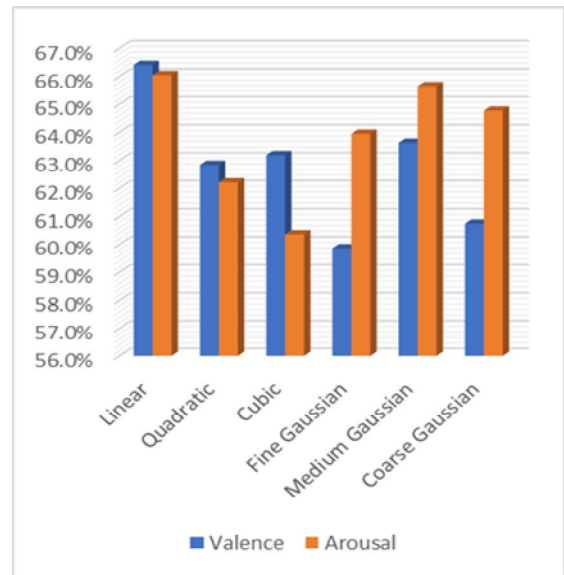


Figure 4.4: Mean accuracy of all SVM kernel functions using DWPT

Table 4.3: Subjects who achieve accuracy equals to or higher 80% using Linear SVM classifier

Subject	DWT		DWPT	
	Valence	Arousal	Valence	Arousal
3	65.0%	77.5%	70.00%	80.00%
12	60.0%	80.0%	62.5%	85.0%
13	75.0%	82.5%	80.0%	85.0%
18	80.0%	57.5%	80.0%	60.0%
20	72.5%	80.0%	72.5%	77.5%
21	62.5%	77.5%	65.0%	80.0%
24	55.0%	82.5%	67.5%	82.5%

Table 4.3 on the other hand shows the subjects who have achieved accuracy equals to or higher 80% using Linear SVM. In the context of subject-dependent approach, subject 13 has achieved highest accuracy of 75% in valence and 82.5% in arousal when tested using DWT, while in DWPT subject 13 has achieved 80.0% and 85% of accuracy for valence and arousal respectively. It is also noted that in DEAP data set, accuracy of arousal is higher compared to valence for every subject except for subject 18, which appear as an outlier.

4.2 Subject-dependent Approach on KNN Classifier

Figure 4.5 and Figure 4.6 shows the performance of DWT and DWPT when classified using KNN. In this experiment, another six distance functions are tested. It has been noticed that DWT and DWPT perform approximately similar when tested on Cosine KNN distance function with differences about $\pm 0.6\%$ for valence and $\pm 0.1\%$ for arousal. However, DWT outperform DWPT when the features space is trained using KNN weighted distance function as presented in Table 4.1. In weighted distance function, it makes use of all training examples, not just k as predetermined by the researchers. Hence, KNN with weighted distance function classifier has becomes a global one since all training instances are used. The only disadvantage is that this classifier runs more slowly [19].

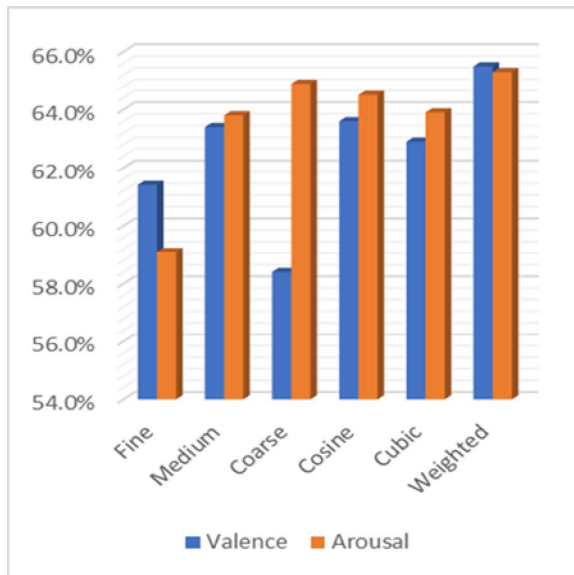


Figure 4.5: Mean accuracy of all KNN functions using DWT

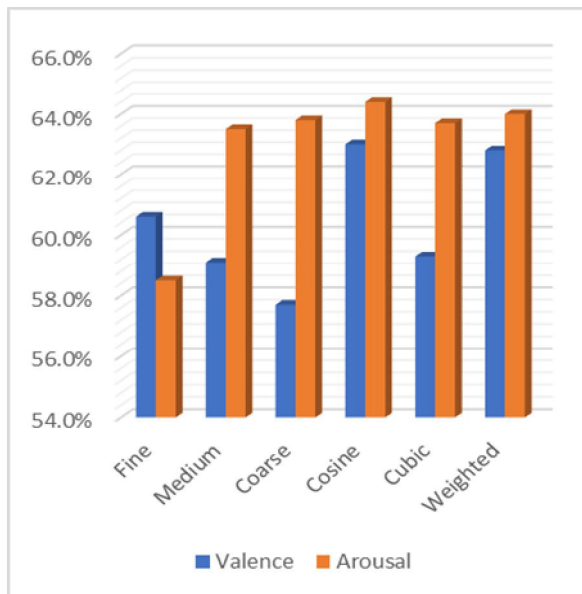


Figure 4.6: Mean accuracy of all KNN functions using DWPT

4.3 Subject-independent Approach on SVM Classifier

We evaluate the performance of DWT and DWPT with the different kernel functions of SVM classifiers by considering 32 participants from the DEAP dataset. In this subject independent experiment, all the data is trained at once using selected classifier (SVM and KNN on different setting).

The obtained results are illustrated in Figure 4.7 and Figure 4.8 Based on the given results on the Table 4.2, DWPT outperformed DWT when is tested on the Quadratic SVM with 61.3% for valence and 62.9% for arousal value. On the other hand, DWT with Quadratic SVM showed 60.1% and 62.8% for valence and arousal respectively. Similarly, the Linear SVM classifier gave a close result with 59.1% for valence and 62.0% for arousal when performed with DWT,

while, DWPT with Linear SVM gave 57.3% and 61.4% for valence and arousal respectively. Although Linear SVM have been discovered to perform well in subject-dependent approach, one must remember Linear SVM works well only for linear problems, however as the complexity and the size of the training set grows, in the case of subject independent (with all the data trained at once) nonlinear kernels such as quadratic kernel performs better for non-linear problems.

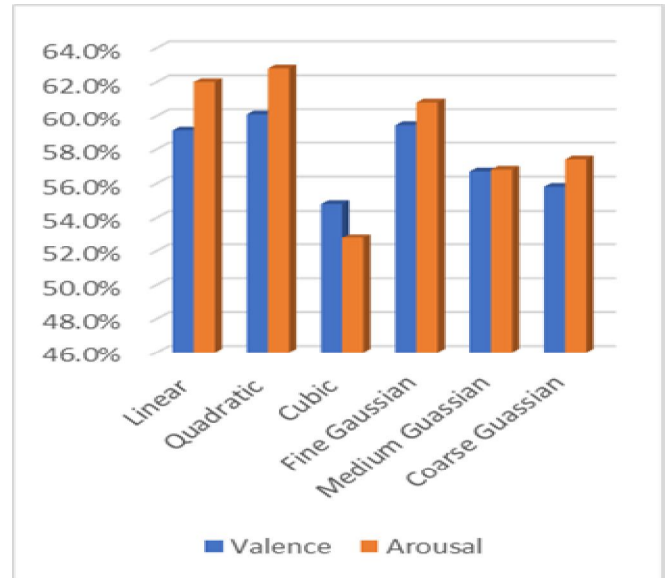


Figure 4.7: Accuracy of all SVM functions using DWT for independent approach.

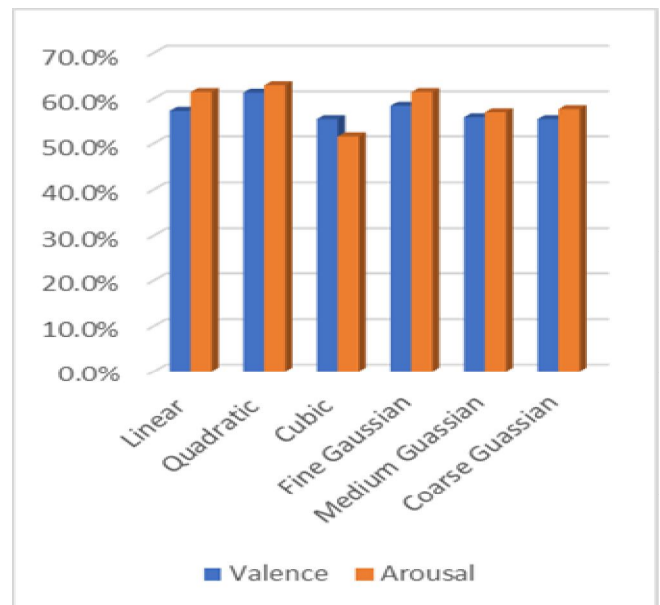


Figure 4.8: Accuracy of all SVM functions using DWPT for subject-independent approach.

4.4 Subject-independent Approach on KNN Classifier

The six distance functions of KNN are used to identify the performance of DWT and DWPT by using subject independent approach. Figure 4.9, Figure 4.10 show the

classification accuracy obtained across all participants. In Table 4.2, we noticed that the DWT with Weighted KNN classifier perform better than other procedures with maximum accuracy of 60.3% and 62.8% for valence and arousal respectively. This is due to advantage of Weighted KNN that minimize leave one out classification error on the given training set by assigning the weight of training samples.

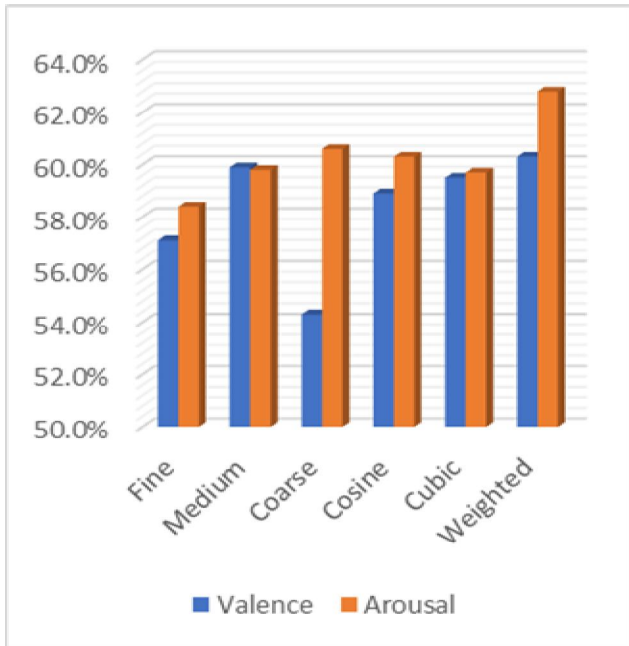


Figure 4.9: Accuracy of all KNN functions using DWT for independent approach.

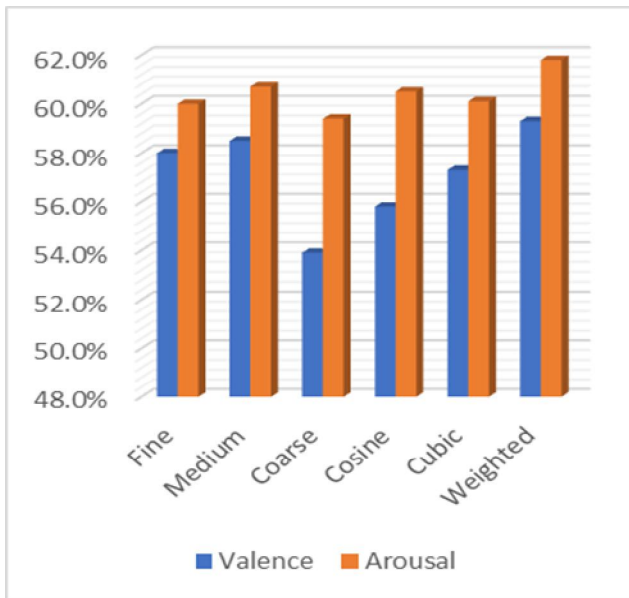


Figure 4.10: Accuracy of all KNN functions using DWPT for independent approach.

5. CONCLUSION

Emotion recognition is a fascinating research field that aims to get insight into human emotional states. EEG is one of the approaches of recognizing emotion through brain signals

obtained from the electrodes placed at human scalp. Although this device can offer better understanding of human emotion compare to facial recognition in which users can hide or control their feelings, the used of this magnificent technology has open a new challenge. EEG signals are generally complex, nonlinear, unstable, and random due to the complex interconnection among billions of neurons. Hence, the analysis of non-linear and chaotic characteristics of EEG signals is a substantial problem. Feature extraction technique is a crucial phase in emotion recognition study. To examine the effectiveness and the performance of two common time-frequency domain feature extraction, this study has conducted an extensive analysis on DWT and DWPT. Both techniques are tested on publicly available DEAP dataset, with 163,840 features have been extracted. The extracted features are then classified using SVM and KNN on different kernel and distance functions. This study has also examined the performance of DWT and DWPT when is executed on subject dependent and subject independent approach. The experimental results have shown that DWPT work well with Linear SVM on subject dependent approach compared to DWT, while in the subject independent approach; DWPT with Quadratic SVM has been seen to achieve better results. This is mainly due to the fact that DPWT can further decompose the signals into new approximation and detailed signal whereas the DWT only resulting approximation signals. Furthermore, as the complexity and the size of the training set grows, in the case of subject independent (with all the data trained at once) nonlinear kernels such as Quadratic kernel has performed better for non-linear problems. On the other hand, when tested with KNN classifier with weighted distance function, DWT has been reported to show better results for both subjects dependent and subject independent approach. In the future, this research aims to investigate feature selection phase prior to the classification to ensure the significant features are only considered for the training purpose while it could reduce the computational cost.

ACKNOWLEDGEMENT

We thank the anonymous referees for their useful suggestions.

REFERENCES

- Kim, S.H., Nguyen, N.A.T. (2017). **Passive EEG-based Emotional state Cognition Using Subject-independent Feature Extraction.** International Journal of Machine Learning and Computing, 7(4), 85–88. <https://doi.org/10.18178/ijmlc.2017.7.4.625>
- Jenke, R., Peer, A., Buss, M. (2014). **Feature extraction and selection for emotion recognition from EEG.** IEEE Transactions on Affective Computing, 5(3), 327–339. <https://doi.org/10.1109/TAFFC.2014.2339834>

3. Hernández, D. E., Trujillo, L., Z-Flores, E., Villanueva, O. M., Romo-Fewell, O. (2018). **Detecting Epilepsy in EEG Signals Using Time, Frequency and Time-Frequency Domain Features**. *Studies in Systems, Decision and Control*, 143, 167–182. https://doi.org/10.1007/978-3-319-74060-7_9
4. Saini, P. K., Dutta, M. (2017). **Literature Review of Feature Extraction Methods for Classification of EEG Signals**. *International Journal of Advanced Research in Science and Engineering (IJARSE)*, 551–560.
5. Russell, J. A. (1980). **A circumplex model of affect**. *Journal of Personality and Social Psychology*, 39(6), 1161–1178. <https://doi.org/10.1037/h0077714>
6. E. Hudlicka, H. Gunes. (2012). **Benefits and Limitations of Continuous Representations of Emotions in Affective Computing**. *International Journal of Synthetic Emotions*, 3(1).
7. Carvalho, S.N., Costa, T.B.S., Uribe, L.F.S., Soriano, D.C., Yared, G.F.G., Coradine, L.C., Attux, R. (2015). **Comparative analysis of strategies for feature extraction and classification in SSVEP BCIs**. *Biomedical Signal Processing and Control*, 21, 34–42. <https://doi.org/10.1016/j.bspc.2015.05.008>
8. Raghav, G.K., Nongmeikapam, K., Dixit, A., Bose, S., Singh, D. (2018). **Evaluating Classifiers for Emotion Signal on DEAP Dataset**, in International Conference on Machine Learning and Data Science (MLDS).
9. Liu, J., Meng, H., Li, M., Zhang, F., Qin, R., Nandi, A. (2018). **Emotion detection from EEG recordings based on supervised and unsupervised dimension reduction**. *Concurrency and Computation: Practice and Experience*, 30(23), 1–13. <https://doi.org/10.1002/cpe.4446>
10. Al-qammaz, A., Yuhanis, Y., Ahmad, F. K. (2016). **Optimization Of Least Squares Support Vector Machine Technique using Genetic Multi-Dimensional Signals**. *Jurnal Teknologi*, 78(5-10), 107–115.
11. Karim, S.A.A., Ismail, M.T., Hasan, M.K., Sulaiman, J., Muthuvalu, M.S., Josefina, J.B. (2015). **Electroencephalography data analysis by using discrete wavelet packet transform**, in AIP Conference Proceedings <https://doi.org/10.1063/1.4915737>
12. Lennon. (2013). **The Use Of Support Vector Machinre With Linear Kernal**. Retrieved from <https://stackoverflow.com/questions/20566869/where-is-it-best-to-use-svm-with-linear-kernel>
13. Rajaguru, H. (2017). **Machine and Gaussian Mixture Model for**, in International Conference on Electronics, Communication and Aerospace Technology ICECA 585–593.
14. Milon, M. H., & Islam, P. R. (2013). **Comparison on Wavelet and Wavelet Packet Transformation for an EEG Signal to Detecting an Epileptic Disorder**.