MULTIRATE ANALYSIS AND NEURAL NETWORK BASED CLASSIFICATION OF HUMAN EMOTIONS USING FACIAL ELECTROMYOGRAPHY SIGNALS

G. Charlyn Pushpa Latha and M. Mohana Priya
Faculty of Engineering, Karpagam University, Coimbatore, India
E-Mail: charlyn.latha@gmail.com

ABSTRACT

Facial electromyography is a modality for designing emotion recognition system which is gaining popularity as a human machine interface to control the devices. In this research, we analyse the Facial Electromyography (FEMG) signals for the six emotions namely anger, disgust, fear, happy, neutral and sad using the multirate features. Multirate signal processing is a technique which alters the rate of the discrete-time signals, either by adding or deleting a portion of the signal samples. The advantages of such multirate features are that it increases the processing efficiency and reduces DSP hardware requirements. Twenty subjects took part in this experimental study. Three multirate features are used to derive the significant features. Six emotions were identified by applying the multirate features as input to neural network models. Two network models namely Cascade Network and Fitting Network were used and compared to identify an efficient network for emotion identification. The performance of the networks identified the six emotions in the range of 78.56% to 98.72%.

Keywords: facial electromyography, multirate features, cascade neural network, fitting neural network.

1. INTRODUCTION

Emotions are an important aspect of human life and the basic research on emotions of the past in the few decades have produced several discoveries that have led to important real world applications. Human face is the richest source of information for revealing the affective state of the individual [1]. It provides an interface to exchange the information with the real world. In the process of information passing, the facial muscles play a dominant role to accomplish the information acquiring and message transmitting [2]. Many research around the world is devoted to the development of Facial Emotion Recognition Systems. FEMG is a precise and sensitive method to measure emotional expression.

FEMG does not depend upon language and does not require cognitive effort or memory. It is capable of registering the response even when subjects are instructed to inhibit their emotional expression. It is also able to measure the facial muscle activities to even weakly evocative emotional stimuli. FEMG has been studied to assess its utility as a tool for measuring emotional reaction. It is also used as a diagnostic tool in Advertising research. CERA (Continuous Emotional Response Analysis) which is based on facial EMG technique, permits second- by- second measurement of response to specific commercial elements. New application also uses FEMG to measure emotional response while playing video games. It has also been used in Human Computer Interaction studies.

FEMG can provide valuable reference in clinical diagnosis and biomedical applications. Before applying FEMG to any application, the natural features of facial muscles have to be taken into account. After understanding the facial muscle kinematics, it is inferred that the facial muscle is a small - three dimension combination of muscular slips carrying out a variety of complex orofacial functions. All facial functions such as speech, mastication and facial expression are all accomplished by individual muscles. A special neural feature of the facial muscles is that their contractions are not only voluntary but also emotional control [3]. Hence, a complete FEMG measurement requires simultaneous recording of multiple facial muscles. FEMG systems are designed by the classification or recognition of facial emotions by the expanding and contraction of facial muscles. With proper training and motivation, most of the subjects exhibit the emotions namely anger, disgust, fear, happy, neutral and sad within the specified frequency bands which can be used as a communication signal.

The content of this paper is organized as follows: Section 2 describes the research assistance necessary for the sampling frequency analysis of the FEMG signals. Section 3 illustrates the FEMG signal acquisition process. It also discusses on the feature extraction and the FEMG signal classification along with their results. Section4 deals with the conclusion and the future work of this paper.

2. RELATED WORK

Emotions are highly independent and explicit to an individual occurrence or state. All individuals do not experience the same emotional strength for similar situations. Experiments dealing with EMG techniques started in 1666. G. Rigas *et al* [4] used Random Forest classifiers and K - Nearest Neighbour (K-NN) classification algorithms for happiness, disgust and fear emotions and they designed an emotion recognition system with the Particle Swarm Optimization of synergetic neural classifier and obtained the recognition rate of 69% for these emotions. Wee Ming Wong *et al* [5] derived the highest recognition rate of 89.25%. Using the emotional features namely sad, anger, pleasure and joy. Shanxiao Yang *et al* [6] used the emotions happiness and disgust using SVM and Backpropagation network and derived

91.67% recognition rates. A Fuzzy C-Means Classifier using emotions namely rest, smile, frown, rage, and pulling up eyebrows was proposed by Hamedi. M *et al* [7]. Principal Component Analysis and higher order statistics were adopted by Jerritta *et al* to design a technique for categorizing the emotions from the facial EMG signals. Jonghwa. K and Ande .E [8] identified 4 emotional conditions namely anger, sad, pleasure and joy and achieved 95% using extended LDA classifier.

Gibert *et al* [9] analyzed the FEMG signal and obtained the mean recognition rate of 92% for six emotions with a sampling frequency of 25Hz. Eun-Hye Jang *et al* [10] et al analyzed the FEMG signals with a sampling frequency of 250Hz. Classification and Regression Tree, Linear Discriminant Analysis, Self Organizing Map and Naïve Bayes classifiers were used for negative emotions like sadness, fear, surprise and stress by Ira Cohen *et al* [11] using a sampling frequency of 30Hz. Shih-Chung Hsu *et al* [12] used a sampling frequency of 100Hz to analyze the FEMG signals. They used 3 databases namely Cohn- Kanade Database, MMI Database and Lab078 database and obtained an average accuracy of 89%. Literature reveals that very few studies on neural networks have been conducted on identifying emotions from FEMG signals. In this study, we propose to use both static and dynamic networks to provide a novel approach in emotion identification.

3. METHODS

In this FEMG signal analysis, the signals are collected for six emotions namely anger, disgust, fear, happy, neutral and sad. During the recording, the subjects are instructed not to move and keep his / her hands relaxed. FEMG signals are recorded from subjects comprising of 11 males and 9 females in the age group of 18-50 years. The audio - visual method for inducing emotions is used in our experiments. These signals are acquired using an ADI Instrument Bio Signal Amplifier using 5 gold plated cup shaped electrodes placed at the locations namely Corrugator Supercilli, Levator Labii Superioris, Orbicularis Oris and Depressor Anguli Oris as shown in Figure-1. During the time of data recording, the subjects are free from illness or medication. FEMG data are recorded in different sessions on different days. FEMG signals are recorded for all the six emotions namely happy, anger, disgust, fear, sad and neutral. Separate audio-visual clips are used for each emotion. Ten trials are recorded for each emotion per session. Data from two sessions are collected. Subjects are given a break of 15 minutes in-between the recording. A notch filter is applied during the recording process to remove power artifacts. Sampling frequency is fixed as 200Hz. Figure-2 depicts the types of emotions under study.

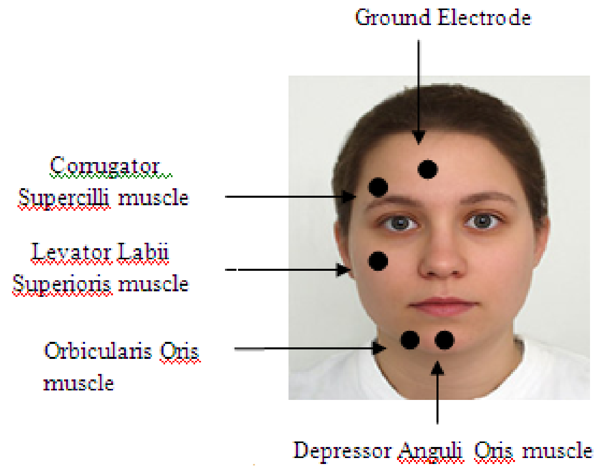

Figure-1. FEMG electrode placement.

Figure-2. Types of emotions.

3.1 FEMG feature extraction

Feature Extraction is very important in any signal processing methodology for efficient pattern recognition or classification. The feature extraction process results in a much smaller and richer set of attributes after combining attributes into a new reduced set of features. Multirate systems play a central role in many areas of signal processing, such as filter bank theory and multiresolution theory which are essential in various standard signal-

VOL. 11, NO. 21, NOVEMBER 2016 ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

processing techniques such as signal analysis, denoising, compression and so on. During the last decades, multirate features have increasingly found applications in new and emerging areas of signal processing, as well as in digital communications. Multirate feature extraction techniques namely multidownsample, multidecimate, and upfirdn are used for extracting the most prominent features.

Downsampling is a process of removing some samples, without the lowpass filtering which decreases the sampling rate by an integer factor. It decreases the sampling rate of x by keeping every n-th sample starting with the first sample. x can be a vector or a matrix. If x is a matrix, each column is considered a separate sequence. A total of 5 features is obtained for each emotion per trial. Hence a data set of 300 samples is collected for a single subject.

Decimation is a process of reducing the sampling rate of a signal which is a specific case of sample rate conversion in a multi-rate digital signal processing system. The decimation factor is an integer or a rational fraction greater than one. Decimation by an integer factor M, can be explained as a 2-step process, with an equivalent implementation that is more efficient.

a) Reduce high-frequency signal components with a digital low pass filter.
b) Downsample the filtered signal by M by keeping only every M^{th} sample.

The formula used for calculation of decimation is

$$y[n] = \sum_{k=0}^{k-1} x[nM - k].h[k] \quad (1)$$

where the h[•] sequence is the impulse response, and k is its length. x[•] represents the input sequence being downsampled. A total of 11 features are obtained for each emotion per trial. Hence a dataset of 660 samples are collected for an individual subject.

Upfirdn is a process of upsampling by applying FIR filter, and then downsample. upfirdn performs a cascade of three operations:

a) Upsampling the input data in the matrix xin by a factor of the integer p (inserting zeros)

b) FIR filtering the upsampled signals data with the impulse response sequence given in the vector or matrix h

c) Downsampling the result by a factor of the integer q (throwing away samples)

The total number of 7 features are obtained for each emotion per trial. Hence a dataset of 420 samples are collected for an individual subject.

3.2 FEMG signal classification

In the FEMG signal recognition phase, two dynamic neural networks are compared to model an efficient neural network algorithm for emotion identification

Cascade forward networks (CFNN) include a weight connection from the input layer to each layer and from each layer to the consecutive layers. It also uses the back propagation algorithm for updating weights, and each layer of neurons relate to all the previous layer of neurons [13].

Fitting networks train a neural network on a set of inputs in order to produce an associated set of target outputs. After constructing the network with the desired hidden layer and the training algorithm, the network is trained using a set of training data. Once the neural network fitted the data, it will be framed a generalization of the input-output relationship. After that the trained network is used to generate outputs for the corresponding inputs which was not trained.

Data from twenty subjects are analyzed and networks are developed for each subject data and 120 networks are modeled to investigate the best network model for emotion identification. A total of 5, 11 and 7 features which are obtained using the statistical features namely multidown sample, multidecimate and multiupfirdn respectively are given as input to each network. All the networks are designed using the hidden neurons namely 4, 7 and 5 with 3 output neurons. The hidden neurons are chosen by trial and error process. 75% of the data set is used for training and 100% of the data set is used for testing the networks. The testing and training error tolerances are fixed as 0.05 and 0.0001 respectively.

The classification performances of the network models chosen in this study are illustrated graphically in Figure 3-5. Figure-3 shows the comparison chart of classification accuracies of the network models using multidown sample features. From Figure-3 it is observed that the highest accuracy of 98.72% is achieved for Subject5 using Fitting network and the lowest accuracy of 78.56% is achieved for Subject18. From our study, it is analyzed that male subjects performed exceedingly well than the females in depicting emotions.

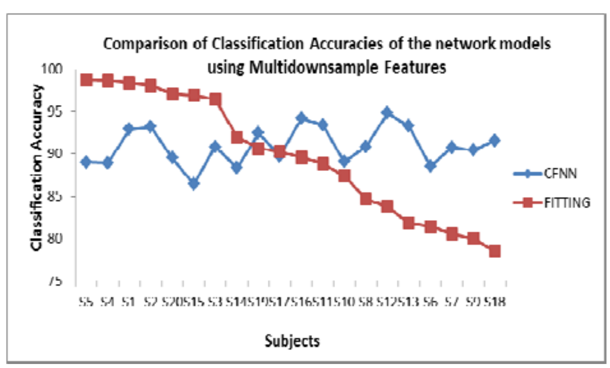

Figure-3. Comparison chart of classification accuracies of CFNN and Fitting Network using Multidownsample features.

ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

The performance of the network models based on multidecimate features are illustrated in Figure-4. On inferring Figure-4, it is observed that Subject11 achieved the maximum classification accuracy of 97.17% using Fitting network model and Subject3 achieved the lowest classification accuracy of 87.62%. Subject11 is also a male subject which reacted well to the emotive stimuli.

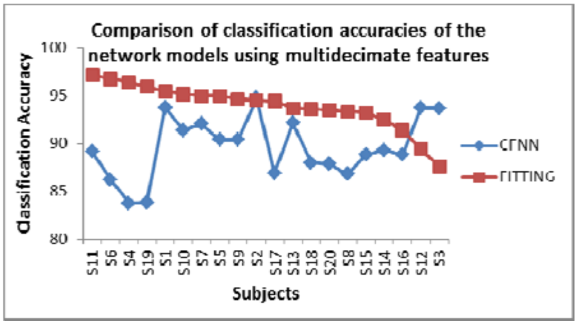

Figure-4. Comparison chart of classification accuracies of CFNN and Fitting Network using Multidecimate features.

Classification performance of the network models using upfirdn features is depicted in Figure-5. From Figure-5 it is inferred that the maximum classification accuracy of 96.92% is achieved for Subject12 using fitting network model and the lowest accuracy of 89.67% is achieved for Subject14. Unlike the other features namely multidownsample and multidecimate features, for the upfirdn features, Subject12 achieved the maximum classification accuracy which is a female subject.

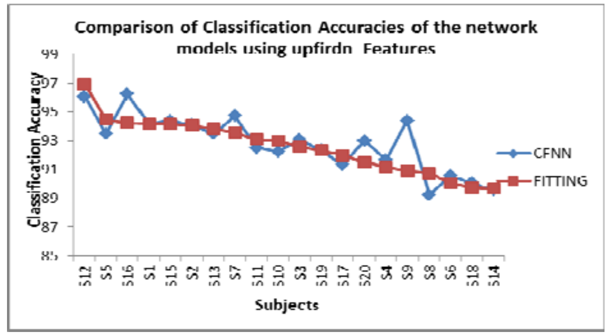

Figure-5. Comparison chart of classification accuracies of CFNN and Fitting Network using upfirdn Features.

Comparing both the network models, CFNN and Fitting Network, Fitting network model achieved the highest classification accuracy for all the features obtained.

3.3 Comparison of classification accuracies based on age factor

Age factor is a major criteria in differentiating strength of the emotions experienced. The 20 subjects volunteered in this study are grouped into 4 categories as 15-20, 21-30, 31-40 and 41-50.

Figure 6 shows the comparison of classification accuracies based on the age factor. On inferring Figure6, it is observed that for the three features namely multidownsample, multidecimate and upfirdn, the maximum classification accuracies fall in the age group of 21-30. The subjects in this age group have undergone regular mental exercises like meditation and these subjects have a positive outlook on the emotional stimuli.

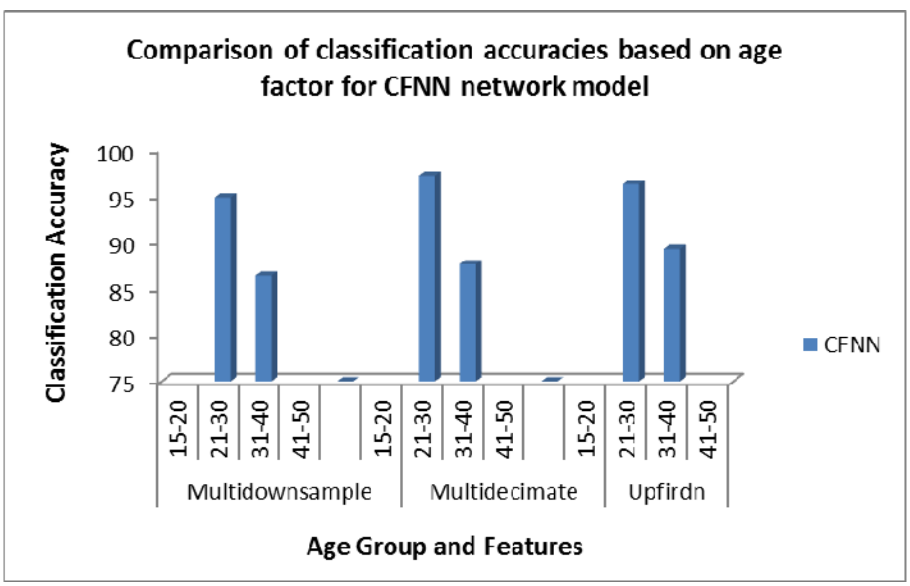

Figure-6. Comparison of classification accuracies based on age factor for CFNN model.

ISSN 1819-6608
ARPN Journal of Engineering and Applied Sciences

Figure-7 depicts the classification accuracies based on age factor for the Fitting Network model. Like the CFNN network model, subjects in the age group of 21-30 reacted well to the emotive stimuli compared to all other subjects in the prescribed age group.

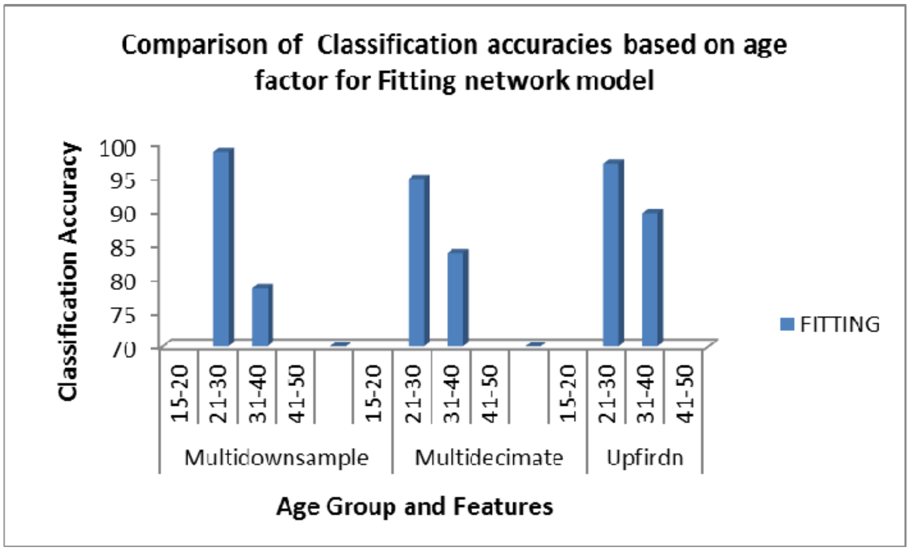

Figure-7. Comparison of classification accuracies based on age factor for Fitting network model.

Concluding, subjects in the age group of 21-30 outperformed other subjects in the prescribed age groups. This is also due to the fact that the people in this age group are altruistic in nature who could experience the emotions exceedingly well.

3.4 Single trial analysis of classification accuracy for the features and the neural networks under study

Single Trial Analysis illustrates the degree to which the subjects experience the emotions. It is a cross check experimentation process for the classification accuracy obtained. The following section illustrates the single trial analysis of the three features namely multidecimate, multidownsample and upfirdn with respect to the network models namely CFNN and Fitting network.

3.4.1 Single Trial analysis for Fitting network using multidecimate features

The single trial FEMG signals for the six emotions classified by Fitting Network model using multidecimate features are shown in Figure8. Comparing all the subjects, S11 performed well in the single trial analysis and S3 reacted worse. S5, S6, S7, S8, S10, S11, S12, S13, S14, S15, S17 and S19 attained the maximum classification accuracies for atleast one emotion. From the figure, it is also inferred that, S11 gives the maximum accuracy for the emotions namely disgust, happy and sad whereas S3 gives the least accuracy for the emotions namely anger, fear and sad. S1, S3, S4, S10, S13, S16 and S19 also gives the least accuracy for atleast one emotion. Most of the subjects performed well for the emotions namely sad, happy and disgust. Also, the emotions namely anger and fear could not be exhibited by most of the subjects. S11 is a male subject who performed well and S3 is a female subject who reacted worse to the emotive stimuli. Further training of the subjects to react to the emotive stimuli could provide improved performance.

ARPN Journal of Engineering and Applied Sciences

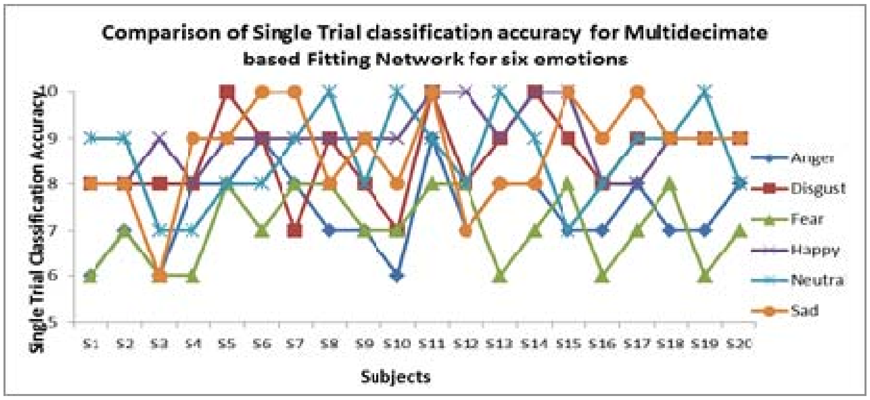

Figure-8. Comparison of Single Trial Classification Accuracy for multidecimate feature based fitting network for six emotions.

3.4.2 Single trial analysis for CFNN Network using multidecimate features

The single trial performance for the six emotions for CFNN network using multidecimate features are depicted in Figure-9. Subject2 responded well to the emotive stimuli which were reflected in single trial analysis. Subject4 reacted worse to the stimuli. It was able to achieve only 60% accuracy for anger and fear emotions, 70% accuracy for disgust, happy and neutral emotions, 80% accuracy for sad emotions. Like Fitting network, the subjects performed less for anger and fear emotions. 100% accuracy for sad emotion were depicted well for 4 subjects namely S1, S2, S14 and S20. 100% accuracy was achieved by S2 and S3 for disgust emotions and S3 and S12 for happy emotion. S2 achieved 100% accuracy for neutral emotion. All subjects depicted sad and disgust emotions exceedingly well in this analysis. S2 is a male subject who performed well and S4 is a female subject who reacted worse to the emotive stimuli. The performance of the subjects could be further improved by intensive training with the emotive stimuli.

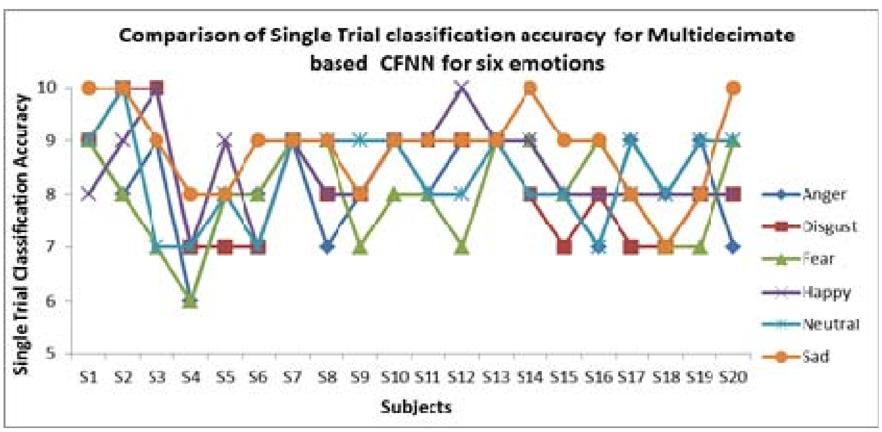

Figure-9. Comparison of Single Trial Classification Accuracy for multidecimate feature based CFNN network for six emotions.

3.4.3 Single trial analysis for Fitting Network using multidownsample features

Figure10 shows the classification performance for fitting network using multidownsample features by way of single trial analysis. S5 responded well to the audio - visual stimuli achieving 100% accuracy for 4 emotions namely disgust, happy, neutral and sad emotions. S18 responded worse comparatively than the other subjects. It achieved 70% accuracy for the emotions namely anger, disgust, fear and neutral. From the figure, it is also inferred that only few subjects namely S1, S5, S6 and S9 attained the highest accuracy for some emotions. Most of the subjects had an average performance for sad and neutral emotions. Fear and anger emotions were not very well performed by the subjects. Excessive training of the subjects for anger and fear emotions could pave way for the maximum accuracy in single trial analysis. Also the video clips which stimulate the anger and fear emotions could be chosen accordingly.

ARPN Journal of Engineering and Applied Sciences

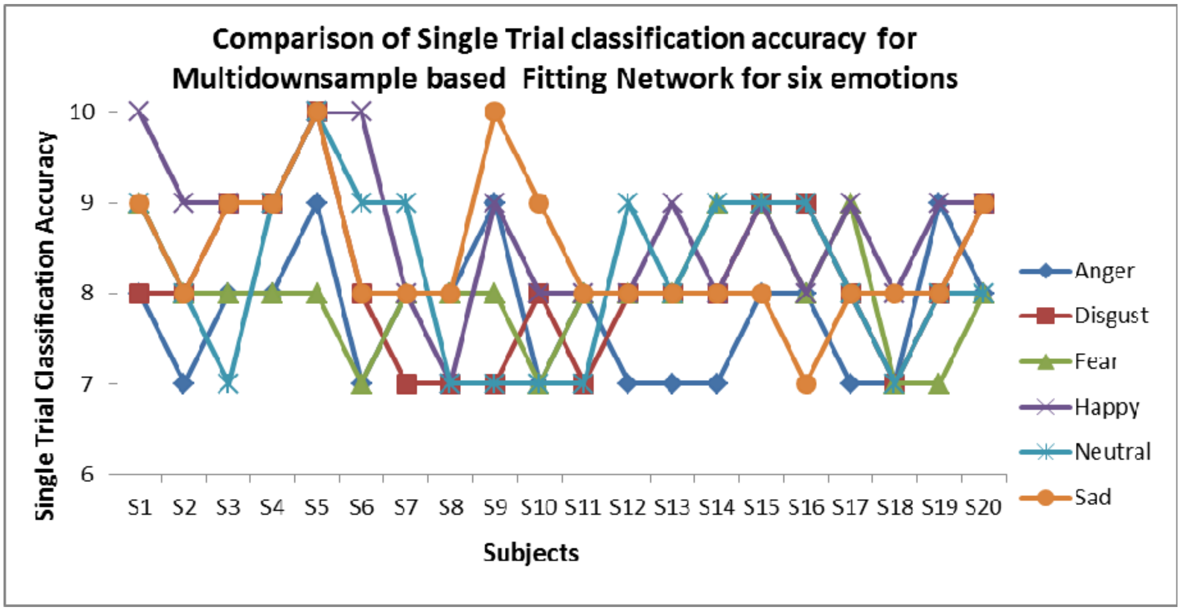

Figure-10. Comparison of Single Trial Classification Accuracy for multidownsample feature based Fitting network for six emotions.

3.4.4 Single trial analysis for CFNN Network using multidownsample features

Classification performance for fitting network is using multidownsample features by single trial analysis: It is illustrated in Figure11. S12 achieved 100% accuracy for 4 emotions namely disgust, happy, neutral and sad. Subject15 reacted worse to the stimuli achieving 60% for anger, 70% for disgust, fear and neutral emotions and 80% for happy and sad emotions respectively. For emotions namely disgust, happy, neutral and sad, 100% accuracy is achieved by S12, (S7, S12), S12 and S10 respectively. Most of the subjects performed well for sad emotion. Disgust emotion was performed in an average manner among the subjects. Comparing anger and fear emotions, the classification accuracy for the fear emotions was better. S12 is a male subject which responded well to the audio- visual stimuli and S15 is a female subject which reacted worse.

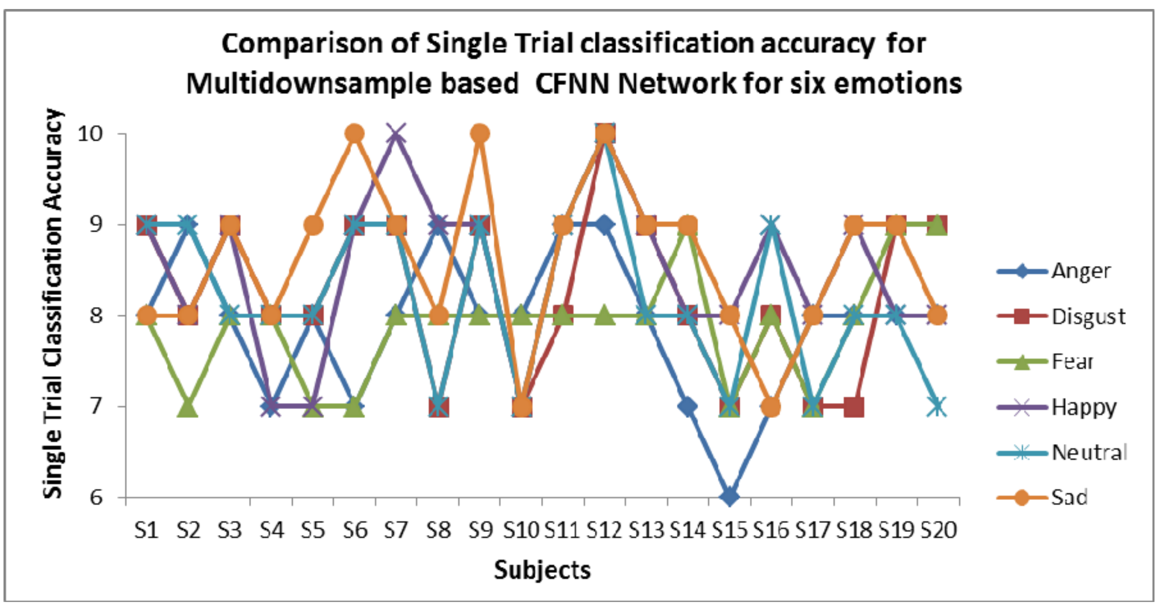

Figure-11. Comparison of Single Trial Classification Accuracy for multidownsample feature based CFNN network for six emotions.

ARPN Journal of Engineering and Applied Sciences

3.4.5 Single trial analysis for Fitting network using upfirdn features

The single trial classification performance for fitting network using upfirdn features are illustrated in Figure12. S12 reacted well to the emotive stimuli achieving 100% accuracy for 4 emotions namely disgust, happy, neutral and sad. S14 reacted worse to the stimuli achieving 70% accuracy for emotions namely anger, fear, neutral and sad. It achieved 80% accuracy for emotions namely disgust and happy. Apart from S12, S16 and S20 achieved 100% accuracy for sad emotion. Most of the subjects performed well for the neutral and sad emotions. Fear emotion is less depicted as the subjects did not find any stronger fearfulness. Hence, the classification accuracy using single trial could be improved using repeated training as well as by using more inducing video clips for fear emotion. S12 is also a male subject which responded well to the emotive stimuli.

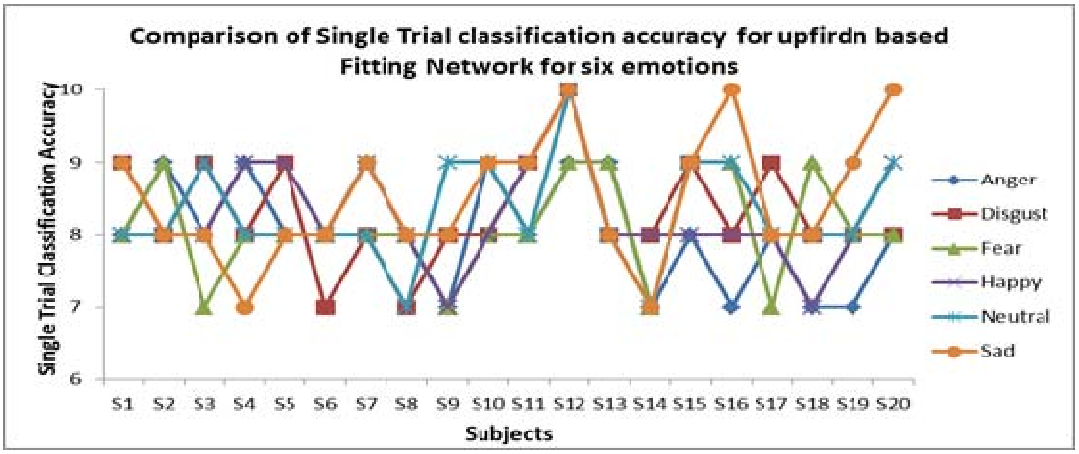

Figure-12. Comparison of Single Trial Classification Accuracy for upfirdn features based Fitting network for six emotions.

3.4.6 Single trial analysis for CFNN network using upfirdn features

The single trial classification performance of CFNN network using upfirdn features is illustrated in Figure-13. S16 achieved the maximum accuracy of 100% for 3 emotions namely disgust, happy and sad. S8 responded less to the emotive stimuli. It achieved only 60% accuracy for anger and fear emotions, 70% for neutral and sad emotions, 80% for disgust emotion and 90% for happy emotion. Only 2 emotions namely happy and sad achieved 100% accuracy for 5 subjects namely S3, S7, S12, S16 and S19. Sad emotion was depicted in average among all the subjects. This is because the subjects are too sentimental in nature and these subjects are very well affected by the video clips shown for sad emotion. From the single trial results, it is observed that none of the subjects performed exceedingly well for the emotions namely anger and fear. The classification accuracy is better for happy and sad emotions when compared with all the other emotions.

ARPN Journal of Engineering and Applied Sciences

www.arpnjournals.com

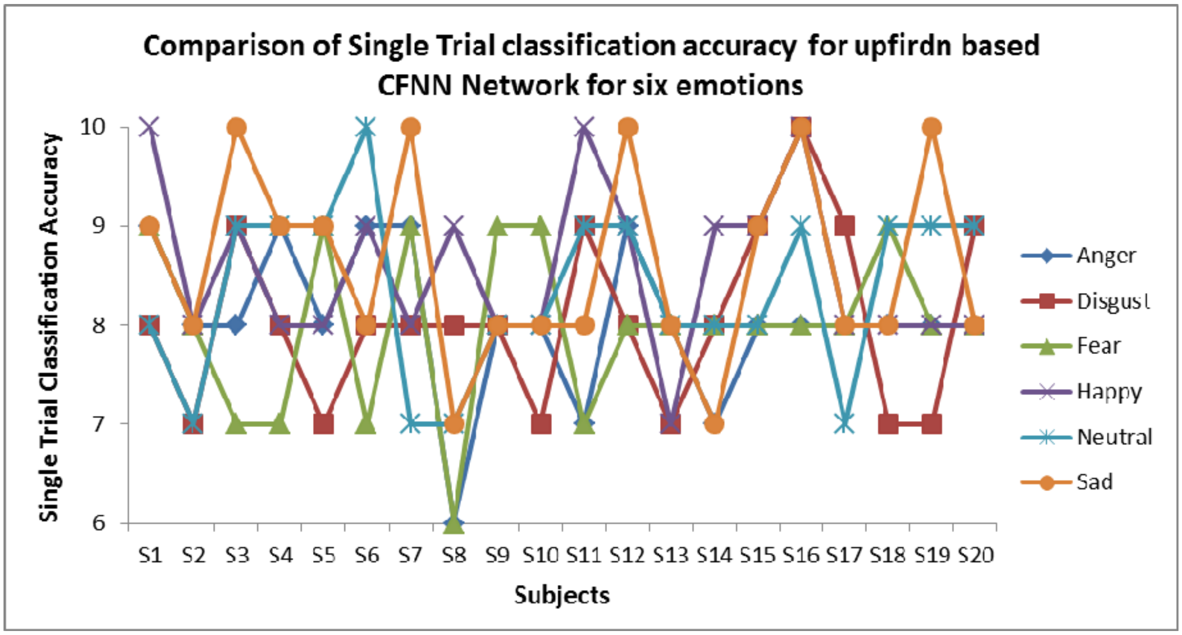

Figure-13. Comparison of Single Trial Classification Accuracy for upfirdn features based CFNN network for six emotions.

4. CONCLUSIONS

Identifying the person's emotional state through the FEMG signal has drawn increasing attention. FEMG recording may be considered a sensitive technique for inferring subjective mood states or affective responses. However, it has limitations for many applications under natural life circumstances due to its obtrusiveness and the fact that facial activity is influenced by many other, nonaffective, behavioural factors. In this research, mutirate features are proposed to recognize the six facial emotions namely anger, disgust, fear, happy, neutral and sad using two neural network models. Data from twenty subjects are used in this study. Maximum recognition rates of 98.72% and 97.17% are achieved for Fitting Network and CFNN model with multidown sample and multidecimate features respectively. The four emotions namely disgust, happy, neutral and sad were exhibited well by all the subjects. However, the two emotions namely anger and fear were not well when compared with the other four emotions.

After intensive training and also by choosing suitable video clips for anger and fear emotions specifically, both the emotions could also be classified well in single trial analysis. Hence, it is evident that a six state FEMG system could be developed using the emotions namely anger, disgust, fear, happy, neutral and sad. The focus of our future work will be based on increasing the recognition accuracy of the facial emotion recognition system using better features and classifiers. Also the performances of the proposed algorithms are to be verified for online FEMG signals in order to develop a real time FEMG based emotion recognition system.

REFERENCES

[1] Anton Van Boxtel. 2010. Facial EMG as a Tool for Inferring Affective States. Proceedings of Measuring Behaviour 2010 (Eindhoven, the Netherlands, August 24-27, Tilburg University.

[2] Cheng- Ning Huang, Chun-Han Chen, Hung-Yuan Chung. The Review of Applications and the Measurements in the Facial Electromyography. Journal of Medical and Biological Engineering 25(I):15-20, Department of Electrical Engineering, National Central University, Chung-Li, Taiwan, 320 ROC.

[3] Lapatki B.G, Stegeman D.F and Jonas I.E. 2003. A surface EMG electrode for the simultaneous observation of multiple facial muscles. Journal of Neuroscience Methods. 123(2): 117-128.

[4] G. Rigas, C. D. Katsis, G. Ganiatsas and D. I. Fotiadis. 2007. A User Independent, Biosignal Based Emotion Recognition Method. Proceedings of the11th International Conference, UM 2007, Corfu, Greece. pp. 314-318.

[5] Wee Ming Wong, Ayer Keroh, Malaysia Tan, A. W. C; Chu Kiong Loo; Wei Shiung Liew. 2010. PSO optimization of Synergetic Neural classifier for multichannel emotion recognition. Second World

ARPN Journal of Engineering and Applied Sciences

Congress on Nature and Biologically Inspired Computing (NaBIC). pp. 316-321, 15-17.

[6] Shanxiao Yang, Guangying Yang. 2011. Emotion Recognition of EMG Based on Improved L-M BP Neural Network and SVM. Journal of Software. 6(8): 1529-1536.

[7] Hamedi, M. Tehran, Iran Rezazadeh, I.M; Firoozabadi. 2011. Facial Gesture recognition using two- channel bio- Sensors configuration and fuzzy classifier: A pilot study. M. Electrical, Control and Computer Engineering (INECCE), International Conference. 21, 22, pp. 338-343.

[8] Jonghwa K and E. Ande. 2008. Emotion Recognition Based on Physiological Changes in Music Listening. IEEE Transactions on Pattern Analysis and Machine Intelligence. 30(12):2067-2083.

[9] Guillaume Gibert, Martin Pruzinec, Tanja Schultz, Catherine Stevens. 2009. Enhancement of Human Computer Interaction with facial Electromyographic sensors. OZCHI 2009, November 23-27, 2009, Melbourne, Australia. pp. 421-424.

[10] Eun-Hye Jang, Byoung-Jun Park, Sang-Hyeob Kim, Myoung-Ae Chung, Mi-Sook Park, Jin-Hun Sohn. 2013. Classification of Human Emotions from Physiological signals using Machine Learning Algorithms. ACHI 2013: The Sixth International Conference on Advances in Computer-Human Interactions, Copyright (c) IARIA, 2013. ISBN: 978-1-61208-250-9, pp. 395-400.

[11] Ira Cohen, Nicu Sebe, Ashutosh Garg, Larry Chen and Thomas S. Huang. 2002. Facial Expression Recognition from Video Sequences: Temporal and Static Modeling, Computer Vision and Image Understanding.

[12] Shih-Chung Hsu, Ch-Ting Hsu and Chung-Lin Huang. 2013. Video-Based Facial Expression Recognition Using Hough Forest, Signal and Information Processing, Association Annual Summit and Conference (APSIPA), Asia-Pacific,1-9.

[13] Hema C.R. 2010. Recognizing Motor Imagery using Dynamic Cascade Feed-forward Neural Networks. 6th International Colloquium Signal Processing and Its Applications. pp. 1-3.

[14] Wim H. J. Meeus*et al*. 2016. Motor affective and cognitive empathy in adolescence: interrelations between facial electromyography and self- reported trait and state measures. Article in Cognition and Emotion.pp. 745-761.