EMOTION AND LIE RECOGNITION FROM EEG SIGNALS USING DEEP LEARNING

Enrollment No. : 16103012, 16103198, 16103237

Name of Students : Vanshika Mittal, Tushar Jain, Abhishek Verma

Name of Supervisor : Dr. Suma Dawn

DECEMBER -2020

Submitted in partial fulfillment of Degree of
Bachelor of Technology
In
Computer Science Engineering

DEPARTMENT OF COMPUTER SCIENCE ENGINEERING & INFORMATION TECHNOLOGY
JAYPEE INSTITUTE OF INFORMATION TECHNOLOGY, NOIDA

TABLE OF CONTENTS

Topics	Page No.
Chapter-1 Introduction	**9-13**
1.1 General Introduction	9
1.2 Problem Statement	10
1.3 Novelty of problem	10-11
1.4 Empirical Study	11-12
1.5 Brief Description of the Solution Approach	12
1.6 Comparison of existing approach to the problem framed	12-13
Chapter-2 Literature Survey	14-25
2.1 Summary of papers studied	14-24
2.2 Integrated summary of the literature studied	24-25
Chapter 3: Requirement Analysis and Solution Approach	**25-26**
3.1 Overall description of the project	25
3.2 Requirement Analysis	25-26
3.3 Solution Approach	26
Chapter-4 Modeling and Implementation Details	**27-35**
4.1 Design Diagrams	27-29
4.2 Implementation details and issues	29-35
4.3 Risk Analysis and Mitigation	35

Chapter-5 Testing **35-38**

5.1 Testing Plan 35-37

5.2 Component Decomposition and type of testing required 37-38

5.3 List all test cases in prescribed format 38

5.4 Limitations of the solution 38

5.5 Error and Exception Handling 38

Chapter-6 Conclusion, and Future Work **39-46**

6.1 Findings 39-42

6.2 Conclusion 42-44

6.3 Future Work 45-46

References 46-47

DECLARATION

We hereby declare that this submission is our own work and that, to the best of my knowledge and belief, it contains no material previously published or written by another person nor material which has been accepted for the award of any other degree or diploma of the university or other institute of higher learning, except where due acknowledgment has been made in the text.

Place : Noida   Signature:

Date : 17/12/2019   Name: Vanshika Mittal, Tushar Jain, Abhishek Verma

Enrollment No.: 16103012, 16103198, 16103237

CERTIFICATE

This is to certify that the work titled "**EMOTION AND LIE RECOGNITION FROM EEG SIGNALS USING DEEP LEARNING**" submitted by "**Vanshika Mittal, Tushar Jain and Abhishek Verma** " of Bachelors of Technology of Jaypee Institute of Information Technology University, Noida has been carried out under my supervision. This work has not been submitted partially or wholly to any other University or Institute for the award of any other degree or diploma.

Signature of Supervisor :

Name of Supervisor: Dr. Suma Dawn

Date: 17/12/2019

ACKNOWLEDGEMENT

We take this opportunity to express our profound sense of gratitude and appreciation to all those who have helped us throughout the duration of this project. We would like to express our special thanks of gratitude to our mentor **Dr. Suma Dawn,** who gave us the golden opportunity to be a part of this wonderful project in the domain of Deep Learning and for providing guidance and expert supervision for this project. She has set a spectacular example on us young impressionable minds in the duration of creating this project. We are really thankful to her. We are making this project not only for fulfilling a college requirement but to also increase our knowledge. Thanks again to all who helped us in making of this project.

Signature:

Name: Vanshika Mittal, Tushar Jain, Abhishek Verma

Enrollment No.: 16103012, 16103198, 16103237

Date: 17/12/2019

SUMMARY

As a challenging pattern recognition task, automatic real time emotion recognition based on multichannel EEG signals is becoming an important computer aided method for emotion disorder diagnosis in neurology and psychiatry. Traditional machine learning approaches require to design and extract various features from single or multiple channels based on comprehensive domain knowledge.

Consequently , these approaches may be an obstacle for non domain experts. In contrast , deep learning approaches have been used successfully in many recent literature to learn features and classify different types of data.

A hybrid neural network which combines 'Convolutional Neural Network (CNN)' and 'Recurrent Neural Network (RNN)' has been applied to classify human emotion states by effectively learning compositional spatial temporal representation of raw EEG streams . The CNN module is used to mine the inter channel correlation among physically adjacent EEG signals by converting the chain like EEG sequence into 2D frame sequence. The LSTM module is adopted to mine contextual information.

Experiments are carried out in a segment level emotion identification task, on the DEAP benchmarking dataset for all five emotions.

Signature:

Name: Vanshika Mittal, Tushar Jain, Abhishek Verma

Enrollment No.: 16103012, 16103198, 16103237

Date: 17/12/2019

Signature:

Name of Supervisor: Dr. Suma Dawn

LIST OF FIGURES:

1. EEG Electrodes
2. Dataset contents
3. Performance comparison
4. Flowchart of pre-processing
5. Converting 1D EEG signals to 2D EEG frames
6. Parallel convolutional recurrent neural network
7. Workflow
8. Model Architecture
9. Summary of Model
10. Accuracies of emotions
11. Emotional Accuracy
12. Maximum accuracy for arousal for subject 1
13. Minimum loss for arousal for subject 1

LIST OF TABLES:

1. Data format
2. Accuracies for valence emotion
3. Accuracies for arousal emotion
4. Accuracies for familiarity emotion
5. Accuracies for liking emotion
6. Accuracies for dominance emotion

1. Introduction

1.1 General Introduction

Emotion plays an important role in human daily life, it reflects human feelings of things. The mental health status even influences human interpersonal interaction and decision making. In medical fields of psychiatry and neurology, the detected emotional states of patients can be adopted as an indicator of the certain functional emotional disorders, such as post traumatic stress disorder and major depression. Most recently, researchers analyzed the emotional characteristics of a smartphone overuse group and a healthy group through EEG signals.

Human emotions can be detected by facial expressions, speech, eye blinking and physiological signals. However, the first three approaches are susceptible to subjective influences of the participants, that is, participants can deliberately disguise their emotions.

There already exists extensive studies using machine learning to identify emotion states with EEG signals. Traditional machine learning based methods have shown effective in classifying emotional states, while the shortage of these kinds of approaches is that the researchers must devote numerous efforts to find and design various emotion-related features from origin noisy signals. And the computation of these features is time consuming. In recent years, deep learning has drawn wide attention due to its great success in the visual field. Some deep learning based approaches have also achieved competitive accuracy in EEG-based recognition task.

Here, in this project we used DEAP Dataset were an effective pre-processing method has been proposed to improve the recognition accuracy. Meanwhile, we used a hybrid neural network which combines 'Convolutional Neural Network (CNN)' and 'Recurrent Neural Network (RNN)' has been applied to classify human emotion states by effectively learning compositional spatial-temporal representation of raw EEG streams.

1.2 Problem Statement

As a challenging pattern recognition task, automatic real-time emotion recognition based on multi-channel EEG signals is becoming an important computer-aided method for emotion disorder diagnosis in neurology and psychiatry. Traditional machine learning approaches require to design and extract various features from single or multiple channels based on comprehensive domain knowledge. Consequently, these approaches may be an obstacle for non-domain experts. In contrast, deep learning approaches have been used successfully in many recent literature to learn features and classify different types of data.

We need effective emotion recognition methods for our military purposes too. Since people cannot fake their emotions and this can be used as a reliable lie detector.

1.3 Novelty of problem

In our approach, we used DEAP dataset. We proposed a hybrid neural network which combines 'Convolutional Neural Network (CNN)' and 'Recurrent Neural Network (RNN)'. After preprocessing our dataset several neural network models were trained for different emotions. We used 5 different emotions i.e. arousal, valence, liking , dominance, and familiarity for training of our model.

From most of the literature survey done the accuracies shown by our model were better as shown in Fig.3.

Mostly papers studied focused on 2 emotions that are valence and arousal. But we trained our model on all of the 5 emotions namely valence, arousal, familiarity, likely and dominance.

The model was implemented with Keras framework and trained on an Overclockable NVIDIA GeForce GTX 1060Ti GPU. The Adam optimizer is adopted to minimize the cross-entropy loss function. The keep probability of dropout operation is 0.5. The hidden states of the LSTM cell d is 32. All fully connected layers have the same size of 1024.

Keras is better suited to deploy CNN and RNN model in deep learning. We have refrained from forced reshaping of tensors, forcefully tweaking the learning rate and applying L2 regularization to the cost like most authors and used keras' adam optimiser. We have also used separate testing and validation data to get better performance from the model. For batch size we have chosen the highest common factor of training-validation split samples so that no data sample is ignored by the model during training process.

-Most literatures show that preprocessing can improve the accuracies of the model significantly. My focus would be to preprocess the baseline signals (produced for the initial amount of time as the subject puts on the EEG sensor cap).

-Further novelty would be fine - tuning the model to introduce early stopping of iteration, and experimenting with optimisers and activation functions.
- Introducing scope for further analysis on the predicted data.
-Working towards making the technology more deployable.

1.4 Empirical Study

An electroencephalogram (EEG) is a test that detects electrical activity in your brain using small metal discs (electrodes) attached to your scalp. Your brain cells communicate via electrical impulses and are active all the time, even when you're asleep. This activity shows up as wavy lines on an EEG recording. EEG has 64 channels in total and DEAP dataset contains values of 32 channels.

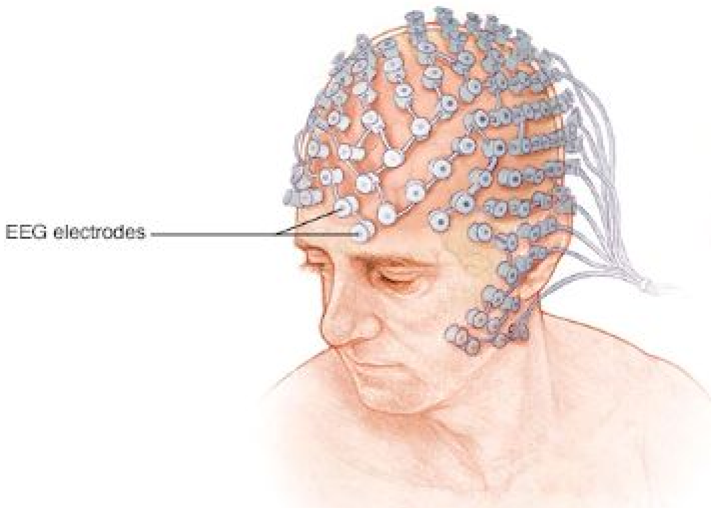


Fig 1. EEG Electrodes

We used multimodal data set for the analysis of human affective states. The electroencephalogram (EEG) and peripheral physiological signals of 32 participants were recorded as each watched 40 videos each 63 seconds long. Participants rated each video in terms of the levels of arousal, valence, like/dislike, dominance, and familiarity. During the experiment, the EEG signals were sampled at 512 Hz and then down-sampled to 128 Hz.

Column name	Column contents
Participant_id	The unique id of the participant (1-32).
Trial	The trial number (i.e. the presentation order).
Experiment_id	The video id corresponding to the same column in the video_list file.
Start_time	The starting time of the trial video playback in microseconds (relative to start of experiment).
Valence	The valence rating (float between 1 and 9).
Arousal	The arousal rating (float between 1 and 9).
Dominance	The dominance rating (float between 1 and 9).
Liking	The liking rating (float between 1 and 9).
Familiarity	The familiarity rating (integer between 1 and 5). Blank if missing.

Fig 2. Dataset contents

1.5 Brief Description of the Solution Approach

In the solution approach an effective pre-processing method has been proposed to improve the recognition accuracy. Meanwhile, a hybrid neural network which combines 'Convolutional Neural Network (CNN)' and 'Recurrent Neural Network (RNN)' has been applied to classify human emotion states by effectively learning compositional spatial-temporal representation of raw EEG streams. The CNN module is used to mine the inter-channel correlation among physically adjacent EEG signals by converting the chain-like EEG sequence into 2D-like frame sequence. The LSTM module is adopted to mine contextual information. Experiments are carried out in a segment-level emotion identification task, on the DEAP benchmarking dataset.

1.6 Comparison of existing approach of the problem framed

There already exists extensive studies using machine learning to identify emotion states with EEG signals. Traditional machine learning based methods have shown effective in classifying emotional states, while the shortage of these kinds of approaches is that the researchers must devote numerous efforts to find and design various emotion-related features from origin noisy signals. And the computation of these features is time consuming. A variety of feature extraction methods are proposed in recent years [1]. While the most commonly used methods are Fourier Transform (FT), Power Spectral Density (PSD) and Wavelet Transform (WT) [2]. Chen and Zhang compared two different feature extraction approaches and four different machine learning classifiers and found that nonlinear dynamic features lead to higher accuracy [3]. Yin et al. proposed a transfer recursive feature elimination (T-RFE) approach that selects features of EEG to determine the optimal feature subset regarding a cross subject emotion classification issue [4].

In recent years, deep learning has drawn wide attention due to its great success in the visual field [5]. Some deep learning based approaches have also achieved competitive accuracy in EEG-based recognition task. In [6], CNN was used to extract time, frequency and location information features of EEG and Stacked Auto Encoders (SAE) was employed to improve the classification accuracy. Zhang et al. proposed both cascade and parallel convolutional recurrent neural network for EEG based movement intention recognition and achieved a perfect performance [7]. Li et al. proposed a pre-processing method that transforms the multi-channel EEG data into 2D frame representation and integrates CNN and RNN to recognition emotion states in the trial-level [8]. Li et al. extracted Power Spectral Density from different EEG channels and mapped it to two-dimensional plane to construct the EEG Multidimensional Feature Image (EEG MFI), then CNN was adopted to learn temporary image patterns from EEG MFI sequences, while the LSTM RNN was used to classify human emotions [9]. Tang et al. used Bimodal Deep Denoising AutoEncoder and BimodalLSTM to classify emotion states and achieved the state-of-the art performance [10].

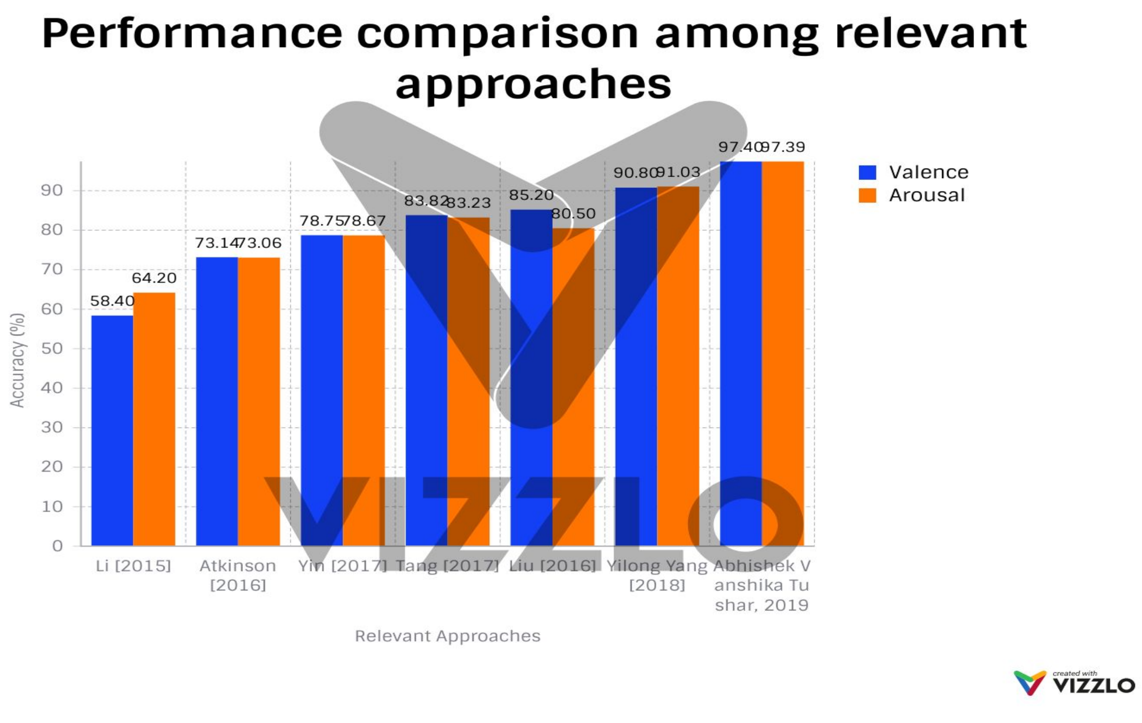


Fig 3. Performance comparison

2. Literature Survey

2.1 Summary of papers studied

Paper 1

Paper Title: Real-Time EEG-Based Happiness Detection System

Authors: Noppadon Jatupaiboon, Setha Pan-ngum, and Pasin Israsena

In this paper, the authors classify happy and unhappy emotions using real-time EEG signals provoked by different pictures which were taken from Geneva Affective Picture Database (GAPED) and sounds from classical emotion elicitation.They have used 50 highest valence scored pictures to be happy stimulus and 50 lowest valence scored pictures to be unhappy stimulus. They also develop games based on the happiness detection system to recognize and control happiness.Power Spectral Density (PSD) have been used as a feature, temporal window duration of 1s to achieve optimal emotion recognition. Classification algorithm known as Gaussian Support Vector Machine (SVM) classifier with leave-one-trial-out cross-validation (LOTO-CV) and leave-one-subject-out cross-validation (LOSO-CV) for evaluating subject-dependent and subject-independent models, respectively. SVM were used due to its good generalization properties and insensitive to overtraining and curse of dimensionality. The average accuracies of 75.62% and 65.12% were estimated for two different models: subject-dependent model and subject-independent model respectively. They used Dimensional model (i-e, valence and arousal) of emotions were used in this study because it is easier to express emotions in terms of arousal and valence rather than basic emotion model (anger, disgust, fear, joy, sadness, and surprise). Emotions are related with the centre of brain called limbic system which includes hippocampus, hypothalamus, thalamus and amygdala. The main factors that we include for EEG-based emotion recognition systems are: Participants (subject-dependent model and subject-independent model), Model of Emotion (Basic model or Dimensional model), Stimulus for emotion elicitation, Features are the EEG signal characteristics, Temporal Window Length which depends on a type of emotion and physiological signal, and a type of Classifier.Emotive headset of 14 channels (AF3, AF4, F3, F4, F7, F8, FC5, FC6, P7, P8, T7, T8, O1, and O2) with sampling rate of 128Hz , 16-bit resolution and 10 participants (i.e., 1male and 9 females; average age is 34.60)were used for EEG recordings. There were 5 trials and each trial consists of one happy (60s) and one unhappy (60s) stimulus. 6000 total samples were labelled either happy or unhappy depending on the stimulus.

A fifth order sinc filter to notch out power line noise at 50 Hz and 60Hz were used for preprocessing EEG signal. For each channel, a baseline has been removed from EEG signal so that the values of this signal distributed around 0. Wavelet transform has been used to decompose EEG signal into 5 frequency bands (Delta (0–4Hz), Theta (4–8Hz), Alpha (8–16Hz), Beta (16–32Hz), and Gamma (32–64 Hz)) with 1s window. Then the Power Spectral Density (PSD) has been calculated from each band. The total number of features were 70 (14 channels and 5 frequency bands) which were normalized for each participant to reduce inter-participant variability by scaling between 0 and 1. Then the normalized features are inserted into classification model, to classify emotion. The selected appropriate parameters are derived from LOTO-CV method. The system detects the happy emotion every 5 seconds. Since emotion is classified every second, there are 5 classifications. Majority vote among classifications is used for system detection output. If the number of classifications during consecutive 5 seconds is happy more than unhappy, the detected emotion is happy. Otherwise, the detected emotion is unhappy. They also divide the level of emotion from happy to unhappy depending on the number of happy classifications.

The advantage of real time EEG data analysis is the minimum computation time.

Paper 2

Paper Title: Using Deep and Convolutional Neural Networks for Accurate Emotion Classification on DEAP Dataset.

Authors: Samarth Tripathi, Shrinivas Acharya, Ranti Dev Sharma, Sudhanshi Mittal, and Samit Bhattacharya.

In this paper two types of neural networks(Deep Neural Network and Convolutional Neural networks) were used for the classification of emotions (Valence and Arousal dimensions) using EEG signals from multimodal EEG benchmark dataset, DEAP (Database for Emotion Analysis Using Physiological Signals).The technique called Dropout is an advanced machine learning technique which has been used in this study. Emotions plays an important role in cognitive processes, so in order to extract these emotions from the human brain instead of using facial, speech or text, EEG (Electroencephalography) signals are used. Extracting emotions using EEG are affordable, non-invasive and promote ubiquitous adoption in industries. A standardised DEAP dataset of human affective states contains the physiological signals of 32 subjects using 32 EEG Channels and with sampling rate of 512Hz.These signals were extracted by using the emotional stimulus of 40 music videos, which were watched by each subject for one-minute in terms of levels of like/dislike, valence, arousal, dominance, and familiarity. Valence (Inactive to Active) and Arousal (Unpleasant to Pleasant) scale was developed by Russel for describing human emotions. There were 40 experiments for each 32 participants and a labels array which contains the valence, arousal, dominance and liking ratings.

This paper uses two neural networks: First is Deep Neural Network which is comprising of 4 neural layers with 5000 neurons, 500 neurons, 1000 neurons and 2 or 3 neurons in output layer respectively. These layers are connected using Softmax which acts as an activator and Dropout method for rectifying the outputs. The second model is Convolution Neural Network (CNN) which were used for classifying images. In this model two convolution layers with Tan Hyperbolic acting as activator followed by Max Pooling the output. Then the Dropout technique is applied for flattening the output results, before being fed into a fully connected neural layer which feeds the output to the final layer of classification classes' size, using Softplus as activator.

2 Classes (High/Low) and 3Classes (High/Normal/Low) classification models were used for Valence and Arousal emotion classification. For 2 classes the accuracies achieved using Deep Neural Networks are 75.58% and 73.28% and for 3 classes the accuracies achieved are 58% and 54% for Valence and Arousal respectively. For 2 classes the accuracies achieved using Convolution Neural Networks (CNN) are 81.41% and 73.36% and for 3 classes the accuracies achieved are 66.79% and 57.58% for Valence and Arousal respectively.

Paper 3

Paper Title: Affective classification using Bayesian classifier and supervised learning.

Authors: Chung, Seong Youb, and Hyun Joong Yoon.

This paper proposes an affective classification method(2 or 3 classes in Valence and Arousal dimensions) using EEG signals from multimodal EEG benchmark dataset, DEAP (Database for Emotion Analysis Using Physiological Signals). The two main machine learning approaches: Bayes Classifier and supervised Learning technique were used. Instead of using the posterior function or likelihood function with Bayes classifier they have introduced a new function known as weighted-log-posterior function. This function is represented as the sum of likelihood function of each feature and the bias factor under the assumption of feature independence. The weights and the bias factors keep computed using supervisory technique such as perceptron convergence algorithm. The maximum value of this function is used by the Bayesian Classifier to classify the affective state. For 2 classes the achieved accuracies are 66.6% and 66.4% and for 3 classes the achieved accuracies are 53.4% and 51.0% for Valence and Arousal respectively.

The human brain with external stimuli is responsible for the efficient affective activities. These affective activities are recognised through different methods (e.g. Electroencephalography) using various machine learning techniques such as Neural Networks, Naïve Bayes Classification, Hidden Markov Model, Support Vector Machines and so on. This study uses the Naïve Bayes Classification method for recognition of affective activities. The unnormalized posterior probability is given as:

$P\ (|\ x) = p\ (x|)\ p\ ()$

Where C is the finite set of affective states {,, ……., }, x is the feature vector with D elements, P (x|) is the likelihood function and p () is the prior probability. The prior probability p () is hard to find in many real applications, therefore under the assumption of feature independence, this study focuses on the unnormalized posterior probability as given:

$P\ (|\ x) = p\ (x|)\ p\ () \Rightarrow P\ ()\ p\ (x_i|)^{w}{}_{c,i}$

Where $w_{c,\ i}$ is Wight factor for c^{th} affective state and i^{th} feature. The Bayes classifier finds an affective state ε with the maximum value of the weighted-log posterior function, that is,

$\in$ = arg max ψ (| x). Then the perceptron Convergence algorithm has been used to obtain the weight factors and bias factor of the weighted-log posterior function which is given as:

$\Psi\ (\ |\ x) = \psi_{c\,=}\ w_c\ z + b_c$and the output of the model is defined by the hard limit transfer function:

$y_c = \text{hardlim}\ (\psi_c) = \{0 \text{ if } \psi_c < 0 \text{ and } 1 \text{ if } \psi_c >= 0\}$

The Bayes Classifier is simple and intuitive but there has not been successful applications yet as compared to other machine learning techniques such as Support Vector Machine.

Paper 4

Paper Title: Emotion Recognition from EEG Signals Using Multidimensional Information in EMD Domain

Authors: Ning Zhuang, Ying Zeng, Li Tong, Chi Zhang, Hanming Zhang, and Bin Yan

This paper introduces a new method for feature extraction and emotion recognition system which is based on Empirical Mode Decomposition (EMD) using EEG signals from multimodal EEG benchmark dataset, DEAP (Database for Emotion Analysis Using Physiological Signals). EMD is a technique which is suitable for the analysis of nonlinear and nonstationary signals such as brain signals and used for decomposing the signal into a set of elemental signals known as Intrinsic Mode Functions (IMFs). These intrinsic functions represents the multiple frequency components of the original signal with band-limited characteristics. To get the instantaneous phase information of IMF,the Hilbert Transform can be applied.The high frequency component from IMF has the significant effect on the different emotion states. The multidimensional information of these functions are utilized for features. This study shows that the three main features: first difference of timeseries, the first difference of phase, and the normalized energy are effective for recognizing emotions from EEG Signals. In this study, a feature vector with these three features were fed to the SVM Classifier for the classification of emotions. Also the informative channels based on EMD strategy are investigated and 8 channels, namely, FP1, FP2, F7, F8, T7, T8, P7, and P8, are selected for feature extraction.

There are different feature extraction methods for emotion recognition from EEG Signals such as: time domain techniques, frequency domain techniques, time-frequency analysis techniques, and other strategies. The time-frequency method is based on the spectrum of EEG signals; then the energy, power, power spectral density (PSD), and differential entropy of certain sub-band are usually utilized as features. Short-time Fourier transform (STFT), Hilbert-Huang transform (HHT) and discrete wavelet transform (DWT) are the most commonly used techniques for spectrum calculating. It has been commonly tested and verified that higher frequency sub-band such as Beta (16–32Hz) and Gamma (32–64 Hz) bands outperforms lower sub-band for emotion recognition The statistics of EEG data include: First and second difference, Mean Value, and Power. The nonlinear features for emotion recognition includes the Fractal Dimension (FD), Sample Entropy, Non-stationary Index, Hjorth features and Higher Order Crossings (HOC).

Paper 5

Paper Title: Investigating Critical Frequency Bands and Channels for EEG-Based Emotion Recognition with Deep Neural Networks.

Authors: Wei-Long Zheng, Student Member, IEEE, and Bao-Liang Lu, Senior Member, IEEE

In this paper the critical frequency bands and channels of EEG were identified using Deep Belief Networks (DBNs) for developing emotion recognition system for three types of emotions (Positive, Negative, and Neutral). The DBNs were trained using Differential Entropy (DE) features from the multichannel EEG dataset which were extracted from 15 Subjects. The reduced electrode set positions are identified using some statistical approaches such as correlation coefficient, F-Score and accuracy rate. Instead of these statistical approaches, the critical frequency bands and channels were investigated by examining the weights of the trained DBNs. An optimal reduction of electrode set position has been investigated which is computationally faster, more feasible and outperforms the original set of channels. The four profiles of EEG channels (4, 6, 9 and 12) were selected among the original 62 channels of EEG, which were relatively stable with the best accuracy of 86.65% than the original 62 channels. The neural signatures which are associated with different emotions were investigated according to the weight distributions obtained from the trained deep neural networks. This study shows the comparison between the deep learning models with shallow models such as DBNs and SVM, K-NN, LR-L2 respectively.

Six different features were used in this study: Differential Entropy (DE), Asymmetry (ASM), Rational Asymmetry (RASM), Differential Asymmetry (DASM), Power Spectral Density (PSD), and Differential Caudality (DCAU). Differential Entropy (DE) feature has the ability to discriminate between low and high frequency energy of EEG patterns. It is defined as:

Differential Entropy (DE) = h (X) = -

If the continuous random variable follows the Gaussian distribution (N, ^{2}), then a fixed length EEG segment, differential entropy is equivalent to the logarithm energy spectrum in a certain frequency band:

$h(X) = \frac{1}{2} \log 2e^2$

Differential entropy feature were calculated on the five frequency bands (Alpha, Beta, Gamma, Theta, and Delta) of EEG signals which were extracted using a 256-point STFT. In the emotion experiments, the EEG data has been extracted from fifteen subjects and each subject has done the experiments twice at intervals of about one week. There are totally 30 experiments evaluated in this study. The training data and the test data are from different sessions of the same experiment. The training data contains nine sessions of data while the test data contains other six sessions of data from the same experiment. The mean accuracies and standard deviations (%) of SVM and DNN for different kinds of features and five frequency bands were presented. Using the DE features from five frequency bands as inputs, the means and standard deviations of accuracies of, K-NN, LR, SVM, and DBN are 72.60%/13.16%, 82.70%/10.38%, 83.9%/9.72%, 86.08%/8.34%, respectively. The results show that the DBN models outperform over other models with higher mean accuracy and lower standard deviations.

Paper 6

Paper Title: Identifying Stable Patterns over Time for Emotion Recognition from EEG.

Authors: Wei-Long Zheng, Student Member, IEEE, Jia-Yi Zhu, and Bao-Liang Lu, Senior Member, IEEE

This paper investigates the stable-neural patterns of Electroencephalogram (EEG) with emotion elicitation over time, which are related to the activities of critical areas of brain and frequency bands. The machine learning technique known as Discriminative Graph Regularized Extreme Learning Machine (GELM) has been adopted for identifying these patterns.The main issues for recognizing the emotions using EEG such as How to discriminate between different emotions based on EEG signals, which patterns of neural oscillations of EEG signals are identifying the emotions in the brain, and to check the performance of models based on machine learning techniques were addressed in this paper. For evaluating these patterns, the new dataset based on SEED (SJTU Emotion EEG Dataset),has been developed in this study. Also a systematic comparison and qualitative evaluation for feature extraction, selection and smoothing and pattern classification were carried out on the DEAP and SEED dataset.

This study reveals that neural signatures exist for three emotions (Positive, Negative, and Neutral) and at critical frequency bands and brain areas, EEG patterns are relatively stable. The time-frequency analysis were used for evaluating these signatures and stable patterns of different emotions. There were 15 participants (7 Males and 8 Females) which perform emotion experiments three times with an interval of a week or longer and 15 trials for each experiment. For emotion elicitation, three types of (Positive, Negative, and Neutral) 15 clip films with audio and video were used. Six different features from EEG were used such as: Differential Entropy (DE), Asymmetry (ASM), Rational Asymmetry (RASM), Differential Asymmetry (DASM), Power Spectral Density (PSD), and Differential Caudality (DCAU). Power Spectral Density (PSD) features were computed using Short Time Fourier Transform with 1s long window based on the five different frequency bands: Alpha (8-13Hz), Beta (14-30Hz), Gamma (31-50Hz), Theta (4-7Hz), and delta (1-3Hz). As this study observes the EEG generation from hidden emotional states, the dynamic characteristics of emotional changes were introduced to emotion recognition. The components of the signal which are not associated with emotional states were discarded by using the technique called Linear Dynamic System (LDS) and for comparison the performance of conventional moving average method has been evaluated.

For dimensionality reduction, the minimal redundancy maximal relevance (MRMR) algorithm has been used. This algorithm uses mutual information as the relevance measure with the max-dependency criterion and minimal redundancy criterion. Max-Relevance searches for features satisfying with the mean value of all the mutual information values between the individual feature x and class c. The three conventional classifiers (K-NN, SVM, and Logistic Regression) and a newly developed classifier Discriminative Graph Regularized Extreme Learning Machine (GELM) were used to build the Emotion Recognition System. The discriminative Graph regularized Extreme Learning Machine (GELM) based on the idea that similar samples should share similar properties. GELM yields a much better performance in comparison with other models for face recognition and emotion classification. From the experimental results, the GELM with the differential entropy features outperforms other methods. The best average achieved classification accuracies of 69.67% and 91.07% on the DEAP and SEED datasets, respectively.

Paper 7

Paper Title: Differential Entropy Feature for EEG-based Emotion Classification.

Authors: Ruo-Nan Duan, Jia-Yi Zhu and Bao-Liang Lu_ Senior Member, IEEE

In this paper, a new EEG based feature called Differential Entropy (DE) were used to represent the characteristics that are associated with emotional states. The combination of this feature on symmetrical electrodes (DASM and RASM) were compared with the Energy Spectrum (ES) feature of the frequency domain. Using these features (DE, DASM, RASM, and ES) on EEG data, the classification accuracies are 84.22%, 80.96%, 83.28%, and 76.56%, respectively. Therefore DE feature is well suited for the recognition of emotions as compared to the traditional frequency domain feature, ES. It is also confirmed that EEG signals on frequency band Gamma relates to emotional states more closely than other frequency bands. Feature smoothing method - linear dynamical system (LDS), and feature selection algorithm - the principal component analysis (PCA) and the minimal redundancy maximal relevance (MRMR) algorithm to reduce the feature dimension, to save the storage space and speed up the classification procedure.

A 512-point short-time Fourier transform with a non-overlapped Hanning window of 1s was used to calculate the frequency domain features. Four kinds of features were compared, which were ES, DE, DASM, and RASM. Two kinds of classifiers were employed, which were linear-kernel support vector machine (SVM) and k-nearest neighbors (K-NN) algorithm. All the features used in this study had been smoothed by LDS, and SVM models were applied as classifiers. The result confirms that the emotional states related to EEG in Gamma frequency band more closely than other frequency bands. It can be implied from this result that DE is more suited for EEG-based emotion classification than traditional ES features.

Paper 8

Paper Title: Differential Entropy Feature for EEG-based Vigilance Estimation.

Authors: Li-Chen Shi, Ying-Ying Jiao and Bao-Liang Lu_ Senior Member, IEEE

In this paper the Differential Entropy (DE) feature has been used for the vigilance estimation based on EEG signals. A physical interpretation of logarithmic form of spectrum feature in EEG signal analysis and a new feature: differential entropy has been presented. Differential entropy is proposed to characterize the level of vigilance, as a complexity measure for continuous EEG signals. For the analysis of EEG signals, the entropy measure has been widely used. From the perspective of information theory, it is defined as the amount of information stores in a more general probability distribution. This paper represents that the Differential Entropy (DE) feature outperforms other features such as Energy Spectrum (ES), auto-regressive (AR) parameters, fractal dimension and sample entropy.

For each session, the visual stimulus sequence and subject's response sequence were recorded by the Neuro-Scan Scan software. At the same time, a total of 62 EEG channels were recorded by the Neuro-Scan system at sampling rate 500Hz and then re-sampled down to 100Hz for the simplicity of data processing. In this study, only 6 EEG channels (PO3, POz, PO4, O1, O2, and Oz) around the occipital region were used for vigilance estimation to improve the feasibility of EEG-based vigilance estimation system. In the sustained visual task, when the subject becoming sleepy, they were more likely to press the wrong button or miss some responses. To evaluate the subject's vigilance level, the local error rate e (t) of a subject's performance was used as the reference of vigilance levels, which was defined as the rate that the subject made a false response (including lapse) within a time window with a constant width. The performance of the differential entropy were evaluated by the values of RMSE (Root Mean Square Error), and compared to other four features, which were energy spectrum, auto-regressive parameters, sample entropy and fractal dimension. The RMSE value indicated the mean of each subject performance. The smaller the value of RMSE, the more precise the vigilance estimation.

Conclusion:

As human brain is highly non-stationary, the EEG signals which are extracted from the human brain are asymmetric and non-stationary in nature. Analysing these signals are challenging and a crucial task in the area of affective computing and require specific domain knowledge such as channels and frequency bands that are very important for understanding brain response. In the field of affective neuroscience, we should know how to discriminate the neural activities of EEG signals between different types of emotions, how to analyse stability of EEG patterns which reflects the emotions elicited in the human brain, and how deep learning and shallow learning techniques affect the performance of developing models. Also we should evaluate the performance of different popular methods for feature extraction, feature selection, feature smoothing and pattern classification for emotion recognition on different datasets. The human brain is a very complicated, and mysterious information processing system, so we need to know how to analyze neurocognitive data efficiently.

2.2 Integrated summary of the literature studied

There already exists extensive studies using machine learning to identify emotion states with EEG signals. Traditional machine learning based methods have shown effective in classifying emotional states, while the shortage of these kinds of approaches is that the researchers must devote numerous efforts to find and design various emotion-related features from origin noisy signals. And the computation of these features is time consuming. A variety of feature extraction methods are proposed in recent years . While the most commonly used methods are Fourier Transform (FT), Power Spectral Density(PSD) and Wavelet Transform (WT) . Chen and Zhang compared two different feature extraction approaches and four different machine learning classifiers and found that nonlinear dynamic features lead to higher accuracy . Yin et al. proposed a transfer recursive feature elimination (T-RFE) approach that selects features of EEG to determine the optimal feature subset regarding a cross subject emotion classification issue.

In recent years, deep learning has drawn wide attention due to its great success in the visual field. Some deep learning based approaches have also achieved competitive accuracy in EEG-based recognition task. In CNN was used to extract time, frequency and location information features of EEG and Stacked Auto Encoders (SAE) was employed to improve the classification accuracy. Zhang et al. proposed both cascade and parallel convolutional recurrent neural network for EEG based movement intention recognition and achieved a perfect performance. Li et al. proposed a pre-processing method that transforms the multi-channel EEG data into 2D frame representation and integrates CNN and RNN to recognition emotion states in the trial-level. Li et al. extracted Power Spectral Density from different EEG channels and mapped it to two-dimensional plane to construct the EEG Multidimensional Feature Image (EEG MFI), then CNN was adopted to learn temporary image patterns from EEG MFI sequences, while the LSTM RNN was used to classify human emotion. Tang et al. used Bimodal Deep Denoising Auto Encoder and Bimodal-LSTM to classify emotion states and achieved the state-of-the art performance.

3. Requirement Analysis and Solution Approach

3.1 Overall description of the project

As a challenging pattern recognition task, automatic real-time emotion recognition based on multi-channel EEG signals is becoming an important computer-aided method for emotion disorder diagnosis in neurology and psychiatry. Traditional machine learning approaches require to design and extract various features from single or multiple channels based on comprehensive domain knowledge. Consequently, these approaches may be an obstacle for non-domain experts. On the contrast, deep learning approaches have been used successfully in many recent literatures to learn features and classify different types of data. In this paper, baseline signals are considered and a simple but effective pre-processing method has been proposed to improve the recognition accuracy. Meanwhile, a hybrid neural network which combines 'Convolutional Neural Network (CNN)' and 'Recurrent Neural Network (RNN)' has been applied to classify human emotion states by effectively learning compositional spatial-temporal representation of raw EEG streams. The CNN module is used to mine the inter-channel correlation among physically adjacent EEG signals by converting the chain-like EEG sequence into 2D-like frame sequence. The LSTM module is adopted to mine contextual information. Experiments are carried out in a segment-level emotion identification task, on the DEAP benchmarking dataset.

3.2 Requirement Analysis

Python 3

numpy

matplotlib

pandas

scipy

pickle

gc

sklearn

tensorflow

Keras

3.3 Solution Approach

In the solution approach baseline signals are considered and a simple but effective pre-processing method has been proposed to improve the recognition accuracy. Meanwhile, a hybrid neural network which combines ‘Convolutional Neural Network (CNN)’ and ‘Recurrent Neural Network (RNN)’ has been applied to classify human emotion states by effectively learning compositional spatial-temporal representation of raw EEG streams. The CNN module is used to mine the inter-channel correlation among physically adjacent EEG signals by converting the chain-like EEG sequence into 2D-like frame sequence. The LSTM module is adopted to mine contextual information. Experiments are carried out in a segment-level emotion identification task, on the DEAP benchmarking dataset.

4. Modeling and Implementation Details

4.1 Design Diagrams

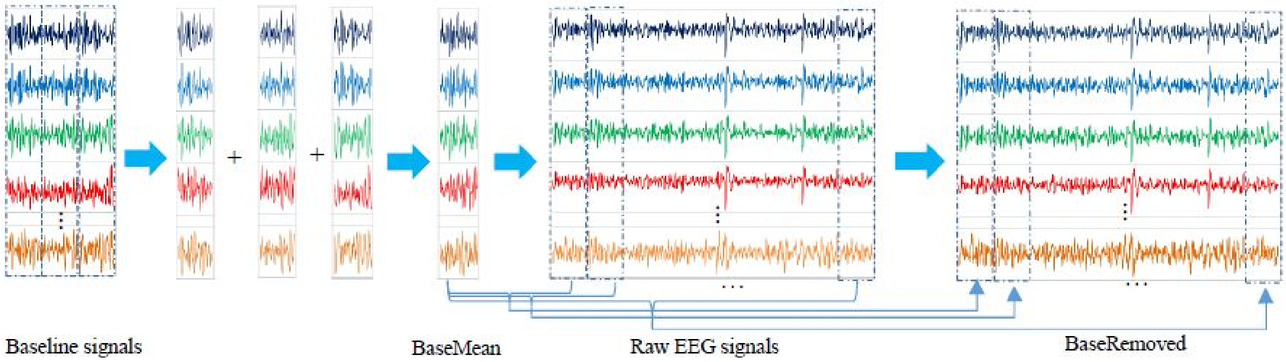


Fig. 4. Flowchart of pre-processing.

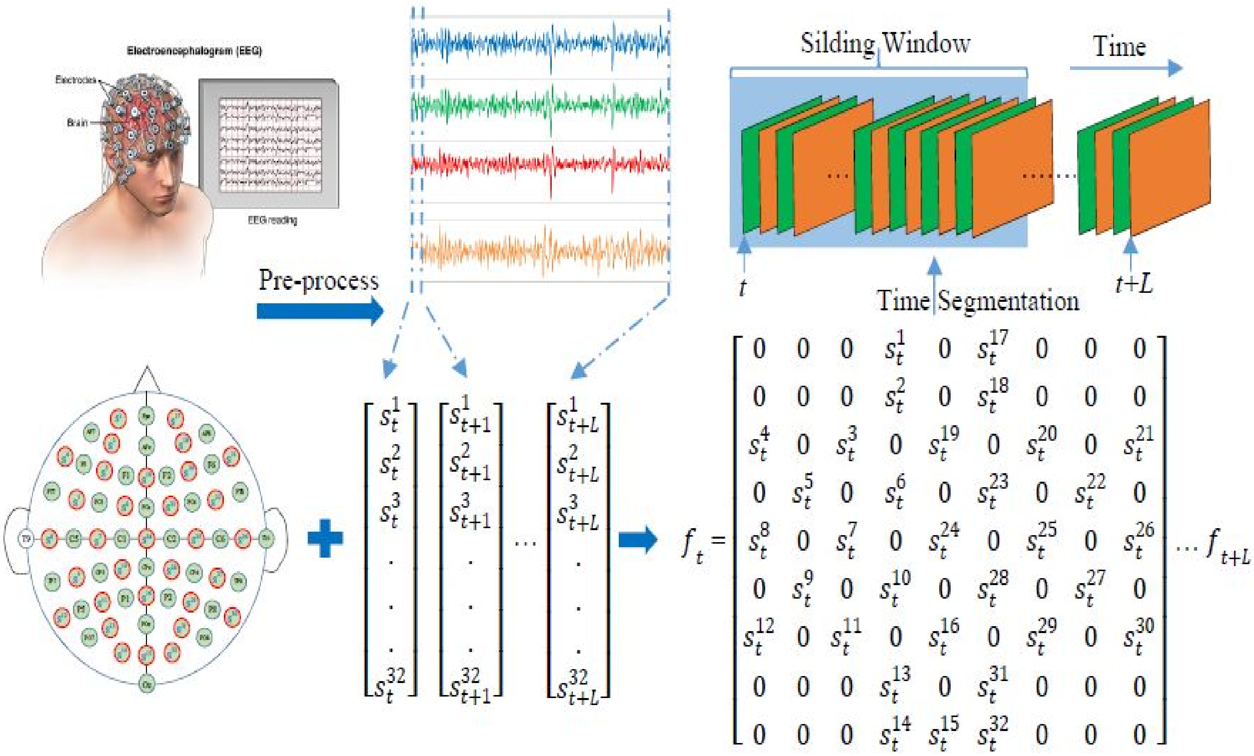


Fig. 5. Converting 1D EEG signals to 2D EEG frames

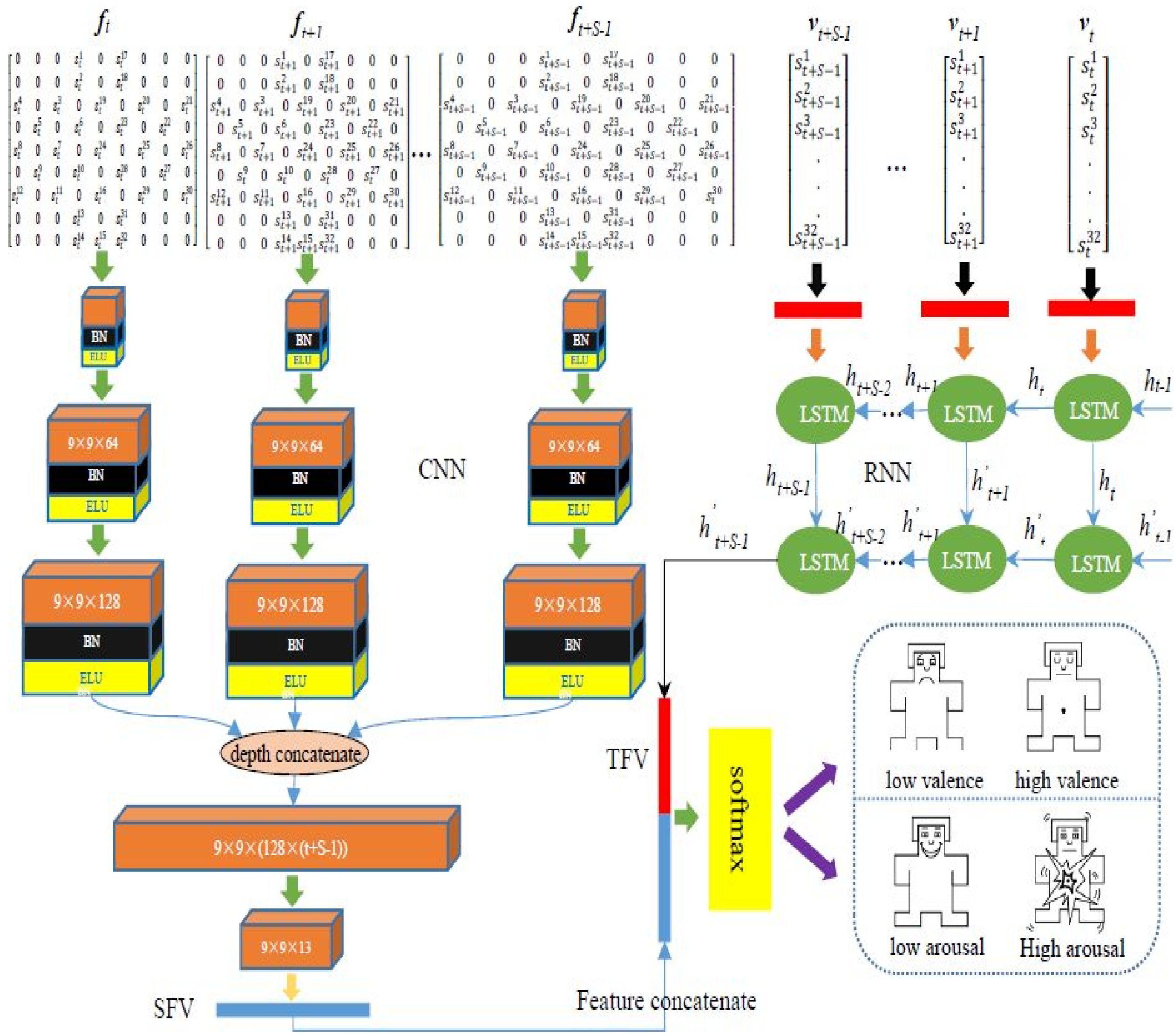


Fig. 6. Parallel Convolutional Recurrent Neural Network.

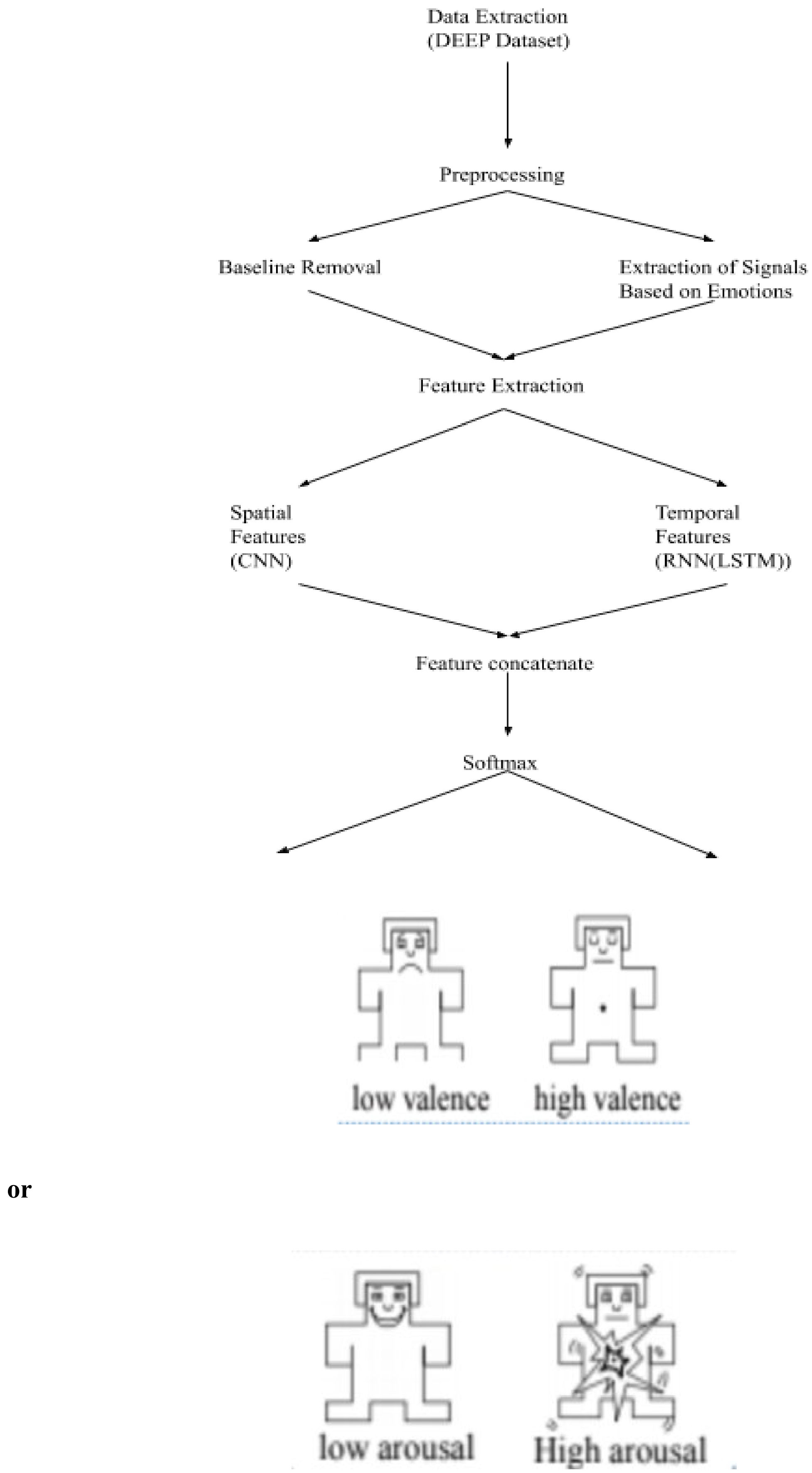


Fig. 7. Workflow

4.2 Implementation details and issues

The Datasets

In this paper, we use DEAP dataset to validate our proposed approach. The DEAP dataset was first introduced in [11]. EEG signals and peripheral physiological signals of 32 participants were recorded when they were watching 40 pieces of music videos. The dataset contains 32 channel EEG signals and 8 channel peripheral physiological signals. Here the EEG signals are used for emotional recognition and the peripheral physiological signals are abnegated. During the experiment, the EEG signals were sampled at 512 Hz and then down-sampled to 128 Hz. EOG artefacts were removed. A bandpass frequency filter from 4.0-45.0 Hz was applied. The preprocessed EEG data contains 60s trial data and 3s baseline data. The emotional music videos include 40 one-minute long clips and participants were asked to rate the levels of arousal, valence, liking and dominance for each video. Each subject file contains two arrays, the data format of file is illustrated in Table I. In order to compare the performance of our proposed method with previous result in [4],[9], [12], [13], [14], we choose 5 as threshold to divide the trials into two classes according to the rated levels of arousal and valence. Then the task can be treated as two binary classification problems, namely high or low arousal and valence.

Pre-processing

In order to improve the recognition accuracy, we use pre-trial data to measure the differences between baseline signals and signals which are recorded while participants are under stimulation. First, we take out pre-trial signals from all C channels and cut it into N segments with a same length L. After the first step, we can get N (C × L) matrixes. Second, we do the element-wise addition for all of these matrices and calculate the mean value. This step can be formulated as:

$$\text{BaseMean} = \frac{\sum_1^N mat_i}{N}$$

Here mat i denotes the ith matrix. After these two steps, we obtain a C × L matrix named BaseMean, which is used to represent subjects' basic emotional state without any stimulation. Third, we segment the raw EEG signals into M (C × L) matrixes named RawEEG and then minus the BaseMean for each matrix. The data we use to represent the difference between experiment signals and baseline signals is named as BaseRemoved, it is created as follows:

$BaseRemoved_j = RawEEG_j - BaseMean$

The last step is concatenating all of these BaseRemoved matrices into a big matrix of which size is the same as raw EEG signals. The flowchart of pre-processing is shown in Fig. 5.

Converting 1D EEG signals into 2D EEG frames

The overall EEG data acquisition and transformation flowchart is shown in Fig. 6. The EEG based BCI system uses a wearable headset with multiple electrodes to capture EEG signals. The International 10-20 System is an internationally recognized method of describing and applying the location of scalp electrode and the underlying area of the cerebral cortex. The “10” and “20” refer to the fact that the actual distance between the adjacent electrodes are either 10% or 20% of the total front front-back or right-left distance of the skull. The data from the EEG signal acquisition system at time index t is a 1D data vector where each element is a pre-processed data of the ith electrode channel and the acquisition system totally contains n channels. With the DEAP dataset, n equals 32. For the observation period [t, t + L], there are (L + 1) 1D data vectors, each of which contains n elements corresponding to n electrodes of the acquisition headset. We can see that the lower left corner of Fig. 6 is a plan view of the International 10-20 System, where the EEG electrodes circled in red are the test points used in the DEAP dataset. In this study, we generalized the International 10-20 System with test electrodes used in the DEAP dataset to form a matrix (h×w), where h is the maximum point number of the vertical test points and w is the maximum point number of the horizontal test points. With the DEAP dataset, h equals w equals 9. From the EEG electrode map, each electrode is physically neighboring multiple electrodes which records the EEG signals in a certain area of the brain, while the elements of the chain-like 1D EEG data vector are restricted to two neighbors. For the purpose of maintaining spatial information among multiple adjacent channels, we convert the 1D EEG data vectors to 2D EEG frames according to the electrode distribution map. The corresponding 2D data frame f_t of the 1D data vector v_t at time index t is denoted as follows:

$$
f_t = \begin{bmatrix}
0 & 0 & 0 & s_t^1 & 0 & s_t^{17} & 0 & 0 & 0 \\
0 & 0 & 0 & s_t^2 & 0 & s_t^{18} & 0 & 0 & 0 \\
s_t^4 & 0 & s_t^3 & 0 & s_t^{19} & 0 & s_t^{20} & 0 & s_t^{21} \\
0 & s_t^5 & 0 & s_t^6 & 0 & s_t^{23} & 0 & s_t^{22} & 0 \\
s_t^8 & 0 & s_t^7 & 0 & s_t^{24} & 0 & s_t^{25} & 0 & s_t^{26} \\
0 & s_t^9 & 0 & s_t^{10} & 0 & s_t^{28} & 0 & s_t^{27} & 0 \\
s_t^{12} & 0 & s_t^{11} & 0 & s_t^{16} & 0 & s_t^{29} & 0 & s_t^{30} \\
0 & 0 & 0 & s_t^{13} & 0 & s_t^{31} & 0 & 0 & 0 \\
0 & 0 & 0 & s_t^{14} & s_t^{15} & s_t^{32} & 0 & 0 & 0
\end{bmatrix}
$$

Zero is used to represent the signals from the channels that are unused in DEAP dataset, which has no effects on neural network. By this transformation, the pre-processed 1D data vector sequences [v t , v t+1 , ..., v t+L] is converted to 2D data frame sequences [f t , f t+1 , ..., f t+L]. For the observation duration [t, t + L], the quantity of 2D data frames is still (L + 1). After transformation, each data frame is normalized across the non-zero elements using Z-score normalization by the following equation:

$$z = \frac{x-\mu}{\sigma}$$

Here x denotes a non-zero element from a certain position of the frame, μ denotes the mean of all non-zero elements and σ denotes the stand deviation of these elements. Finally, we apply the sliding window approach to segment the streaming 2D frames to individual frame-group as shown in the last step of Fig. 6. Each frame-group is a sequence of 2D frames with a fixed length without any overlap between consecutive neighbors. The data frames segment S j is created as follows:

S j = [f t , f t+1 , ..., f t+S-1]

Where S denotes the window size and subscript j is used to identify the different segments during the observation period. The goal of this paper is to develop an effective model to recognize a set of human emotions E = [e 1 , e 2 , ..., e n] T from each windowed data frames segment S j .

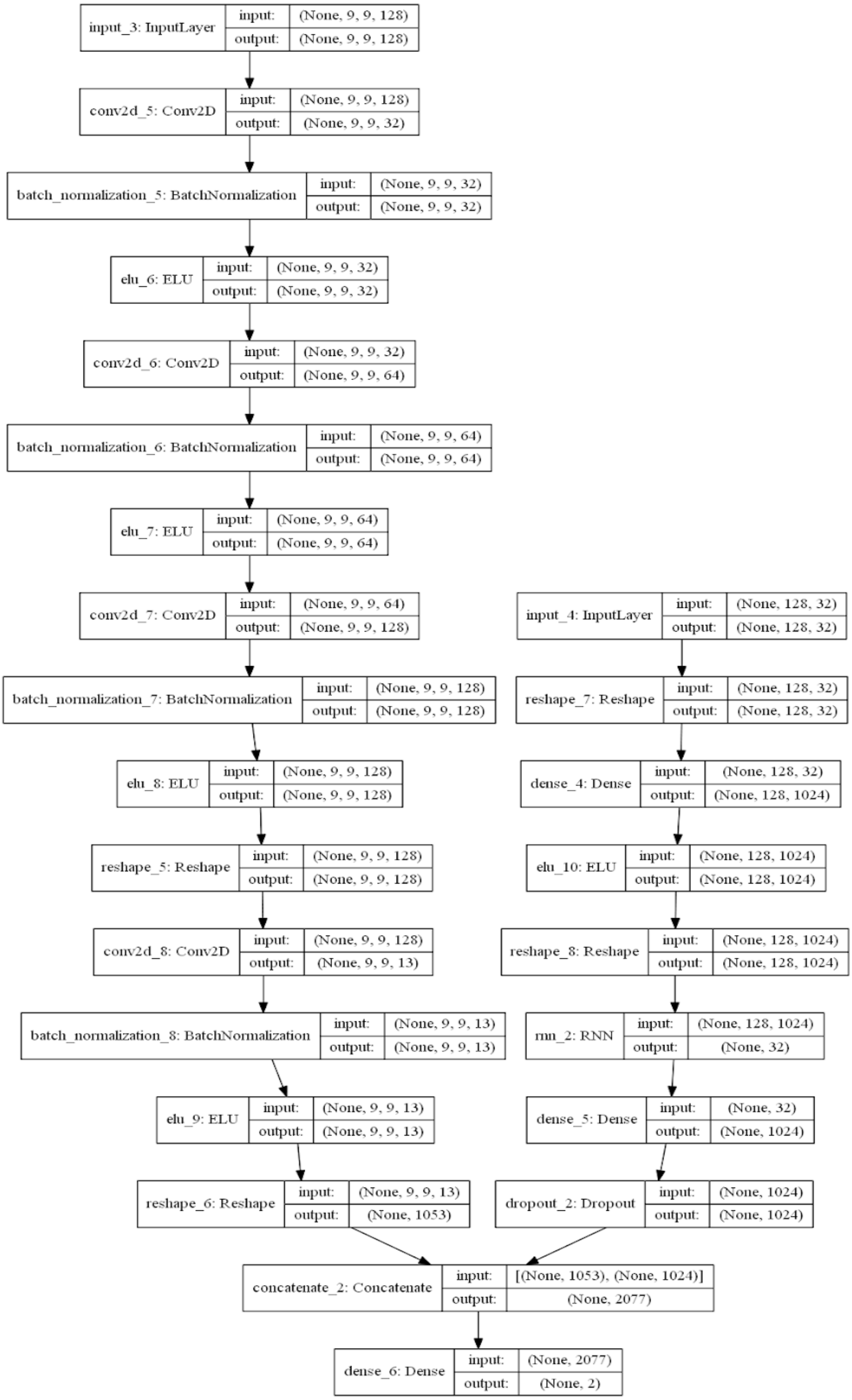


Fig.8. Model Architecture

Parallel Convolutional Recurrent Neural Network

Besides the pre-processing method mentioned above, we also adopt a hybrid deep learning model to classify emotion states, named "Parallel Convolutional Recurrent Neural Network". The model is a composition of two kinds of deep learning structures. It combines the powerful ability of CNN and RNN in extracting spatial and temporal features respectively. The CNN unit works for mining cross-channel correlation and extracting features from 2D frames. The refined RNN structure named "Long Short-Term Memory (LSTM)" models the context information for streaming 1D data vectors. Following these two units, a feature fusion method is used to fuse the extracted features at last for final emotion recognition. The structure of the parallel convolutional recurrent neural network is depicted in Fig. 7. For CNN part, there are three continuous 2D convolutional layers with a same kernel size of 4×4 for spatial feature extraction. While the 3×3 kernel is widely used in computer vision field, we choose 4×4 filter as the signals in 2D EEG frame is sparse. So 4 by 4 filter can mine the correlation among more channels than 3 by 3 kernel. In each convolutional layers, we use zero-padding to prevent missing information at the edge of input data frame. Then we start the first convolutional layer with 32 feature maps and double the feature maps in each of the following convolutional layers. Hence, there are 64 and 128 feature maps in the second and the third convolutional layers respectively. Although a convolutional layer is often followed by a pooling layer in classical CNN architectures, it is not necessary in our model. The pooling layer is usually added for reducing data dimensional at the cost of missing some information. While in this EEG recognition task, the size of data frame is much smaller than that used in computer vision field. Thus, in order to keep all information, we do not use pooling operation in our model. Following each convolution operation, a batch normalization (BN) operation is applied to accelerate the model training. After these three convolutional layers, for the purpose of fusing spatial feature vectors and temporal feature vector which is extracted by RNN, a depth concatenate operation is applied to combine the spatial feature maps into a large cube. And then we use 13 1×1 convolution kernels to shrink the cube into $13\times9\times9$, and further flatten it into a spatial feature vector SFV (dimension 1053) . The purpose of 1 by 1 convolution operation is to fuse feature maps at different times and play a role in dimensionality reduction. The input of CNN part is a pre-processed segment of 2D data frame, while due to the RNN part being responsible for the temporal feature extraction, 1D data vectors are not converted to 2D frame sequences.

TABLE I. DATA FORMAT

Array name	*array shape*	*Array contents*
data	40 × 40 × 8064	video/trial × channel × data
labels	40×4	video/trial × label (valence, arousal, dominance, liking)

Model Implementation

An appropriate time window length is critical for classification performance on streaming data. Wang et al. have found that 1 second long time window is the most suitable window length for emotion recognition [20]. Hence, 1 s is chosen as the time window length in this paper. Since the signals were down-sampled to 128 Hz, the pre-trial baseline signals are segmented to 3 32×128 matrixes to calculate the BaseMean matrix. Then the pre-process operation is applied to trial signals, as shown in Fig. 1. After the preprocessing stage, in each trial, we get 32 channels' pre-processed EEG signals and use a sliding window to divide it into 60 segments with 1 s length, which means the window size S is 128. After segmenting, we get a total of 2400 samples (40 trials×60 segments) for each subject. Then the 2D data frames are transformed with the size of 9×9 as shown in Fig. 6.

Extracting Spatial Feature Vector from CNN and Temporal Feature Vector from RNN (LSTM) we concatenate them and apply softmax activation to classify the emotions as low/high valence/arousal/liking/dominance/familiarity.

4.3 Risk analysis and Mitigation

EEG has several limitations. Most important is its poor spatial resolution. EEG is most sensitive to a particular set of postsynaptic potentials those generated in superficial layers of the cortex, on the crests of gyri directly abutting the skull and radial to the skull.

5. Testing

5.1 Testing Plan

Cnn_datasets.shape: (2400, 128, 9, 9) -->(2400, 9, 9, 128)

Rnn_datasets.shape: (2400, 128, 32)

Labels.shape: (2400, 2)

Total Train Samples: 2160

- Training Samples: 1728
- Validation Samples: 432
- Batch Size: 128

Test Samples: 240

Epochs: 50

Layer (type)	Output Shape	Param #	Connected to
input_3 (InputLayer)	(None, 9, 9, 128)	0	
conv2d_5 (Conv2D)	(None, 9, 9, 32)	65568	input_3[0][0]
batch_normalization_5 (BatchNor	(None, 9, 9, 32)	128	conv2d_5[0][0]
elu_6 (ELU)	(None, 9, 9, 32)	0	batch_normalization_5[0][0]
conv2d_6 (Conv2D)	(None, 9, 9, 64)	32832	elu_6[0][0]
batch_normalization_6 (BatchNor	(None, 9, 9, 64)	256	conv2d_6[0][0]
elu_7 (ELU)	(None, 9, 9, 64)	0	batch_normalization_6[0][0]
conv2d_7 (Conv2D)	(None, 9, 9, 128)	131200	elu_7[0][0]
input_4 (InputLayer)	(None, 128, 32)	0	
batch_normalization_7 (BatchNor	(None, 9, 9, 128)	512	conv2d_7[0][0]
reshape_7 (Reshape)	(None, 128, 32)	0	input_4[0][0]
elu_8 (ELU)	(None, 9, 9, 128)	0	batch_normalization_7[0][0]
dense_4 (Dense)	(None, 128, 1024)	33792	reshape_7[0][0]
reshape_5 (Reshape)	(None, 9, 9, 128)	0	elu_8[0][0]
elu_10 (ELU)	(None, 128, 1024)	0	dense_4[0][0]
conv2d_8 (Conv2D)	(None, 9, 9, 13)	1677	reshape_5[0][0]
reshape_8 (Reshape)	(None, 128, 1024)	0	elu_10[0][0]

```
________________________________________________________________________________________________
batch_normalization_8 (BatchNor (None, 9, 9, 13)     52          conv2d_8[0][0]
________________________________________________________________________________________________
rnn_2 (RNN)                     (None, 32)           143616      reshape_8[0][0]
________________________________________________________________________________________________
elu_9 (ELU)                     (None, 9, 9, 13)     0           batch_normalization_8[0][0]
________________________________________________________________________________________________
dense_5 (Dense)                 (None, 1024)         33792       rnn_2[0][0]
________________________________________________________________________________________________
reshape_6 (Reshape)             (None, 1053)         0           elu_9[0][0]
________________________________________________________________________________________________
dropout_2 (Dropout)             (None, 1024)         0           dense_5[0][0]
________________________________________________________________________________________________
concatenate_2 (Concatenate)     (None, 2077)         0           reshape_6[0][0]
                                                                 dropout_2[0][0]
________________________________________________________________________________________________
dense_6 (Dense)                 (None, 2)            4156        concatenate_2[0][0]
================================================================================================
Total params: 447,581
Trainable params: 447,107
Non-trainable params: 474
________________________________________________________________________________________________
None
```

Fig.9 Summary of Model

5.2 Component Decomposition and type of testing required

Dimensions of emotion:

Valence (positive, negative - the pleasantness of the stimulus)

9: happy, pleased, satisfied, contented, hopeful

1: unhappy, annoyed, unsatisfied, melancholic, despaired, or bored

Arousal (strong, weak - the intensity of emotion provoked by the stimulus)

9: stimulated, excited, frenzied, jittery, wide-awake, or aroused

1: relaxed, calm, sluggish, dull, sleepy, or unaroused;

Control (dominance - the degree of control exerted by the stimulus)

9: in control, influential, important, dominant, autonomous, or controlling

1: controlled, influenced, cared-for, awed, submissive, or guided

Arousal and Valence from the basic axis for most of emotions.

Dominance detects the control on the mind.

Liking and Familiarity detect how much a person likes or is familiar with something. This is used for lie detection.

5.3 List all test cases

Specialised military and police applications.

For training (tracking cadet's mental state).

Detecting stress in drivers, pilots, cadets.

For criminal interrogation (will keep a check on crime rate).

Lie detection (polygraphs test can be faked but brain signals can't be).

5.4 Limitations of the solution

Subject 27 felt dominant at all times so we could not train the model to identify its dominant emotion.

Here, we see that if a subject is completely biased towards feeling a particular emotion then emotion classification for such particular individual if performed, would be illogical as the model would always predict presence of that emotion.

5.5 Error and Exception Handling

Subject 27 felt dominant at all times so we could not train the model to identify its dominant emotion.

Readings for familiarity emotion for subject 2, 15 and 23 were missing in the official DEAP Dataset.

6. Findings, Conclusion, and Future Work

6.1 Findings

Our analysis revealed the following accuracies for Arousal, Valence, Dominance, Liking and Familiarity.

sub: Subject
acc: Accuracy (in %)

<u>Valence Emotion</u>

Average - 97.4

sub	acc	sub	acc	sub	acc	sub	acc	sub	acc	sub	acc	sub	acc	sub	acc
1	99.58	5	97.08	9	95.83	13	98.33	17	89.16	21	97.5	25	96.24	29	100
2	95.83	6	97.91	10	97.5	14	94.99	18	99.58	22	96.24	26	95.83	30	98.33
3	98.33	7	99.58	11	97.91	15	99.58	19	98.75	23	99.58	27	97.58	31	99.16
4	92.5	8	97.08	12	95.66	16	98.33	20	99.16	24	96.24	28	98.75	32	95.83

Table 2. Accuracies for valence emotion

<u>Arousal Emotion</u>

Average - 97.39

sub	acc	sub	acc	sub	acc	sub	acc	sub	acc	sub	acc	sub	acc	sub	acc
1	100	5	97.5	9	95.41	13	99.16	17	95.83	21	98.75	25	97.08	29	98.24
2	93.75	6	97.91	10	96.66	14	97.5	18	97.91	22	96.24	26	93.75	30	99.16
3	99.58	7	99.16	11	92.41	15	98.75	19	98.33	23	99.58	27	97.16	31	97.5
4	95.83	8	97.91	12	97.91	16	91.16	20	99.58	24	99.16	28	96.24	32	98.75

Table 3. Accuracies for arousal emotion

<u>Familiarity Emotion</u>

Average - 98.26

sub	acc	sub	acc	sub	acc	sub	acc	sub	acc	sub	acc	sub	acc	sub	acc
1	99.58	5	100	9	98.75	13	99.58	17	96.66	21	97.91	25	100	29	97.91
2	NA	6	98.33	10	98.75	14	95.41	18	97.5	22	93.33	26	95.41	30	99.16
3	98.33	7	99.58	11	95.41	15	NA	19	98.57	23	NA	27	100	31	99.16
4	98.33	8	98.33	12	97.08	16	100	20	99.58	24	97.08	28	97.5	32	97.5

Table 4. Accuracies for familiarity emotion

<u>Liking Emotion</u>

Average - 97.90

sub	acc	sub	acc	sub	acc	sub	acc	sub	acc	sub	acc	sub	acc	sub	acc
1	99.58	5	96.24	9	99.16	13	99.16	17	94.16	21	97.91	25	97.5	29	99.58
2	96.24	6	97.91	10	97.91	14	96.66	18	98.75	22	97.91	26	97.5	30	99.16
3	99.58	7	100	11	94.16	15	98.33	19	99.16	23	99.16	27	100	31	98.75
4	93.75	8	100	12	97.50	16	98.75	20	99.16	24	97.5	28	94.16	32	97.5

Table 5. Accuracies for liking emotion

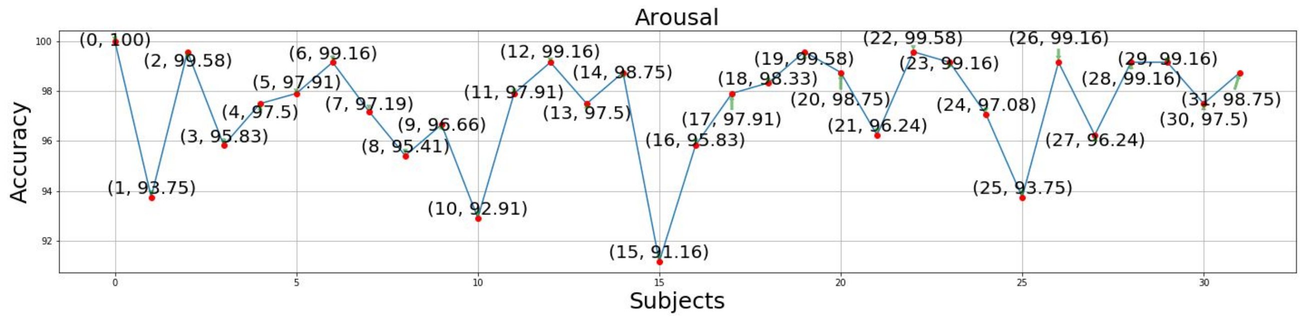
Dominance Emotion

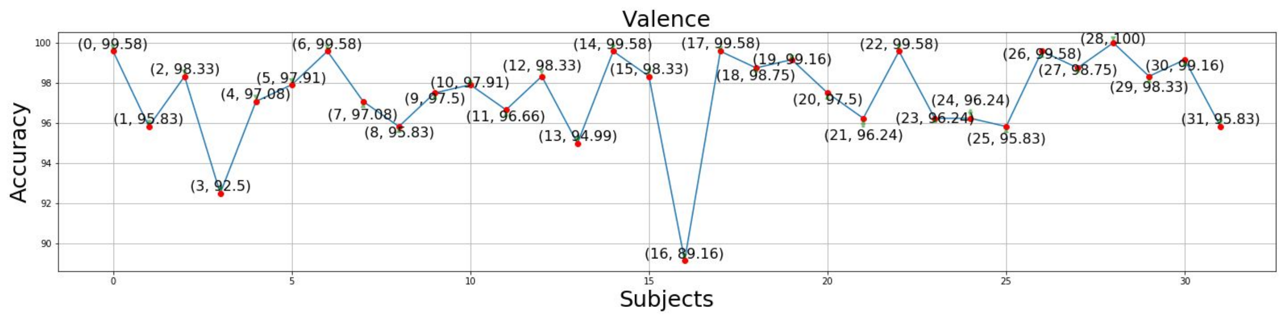
Average - 97.68

sub	acc	sub	acc	sub	acc	sub	acc	sub	acc	sub	acc	sub	acc	sub	acc
1	100	5	94.99	9	98.33	13	97.91	17	98.75	21	98.33	25	99.58	29	100
2	94.58	6	96.66	10	97.5	14	94.99	18	100	22	94.99	26	92.91	30	98.33
3	98.33	7	99.16	11	96.66	15	99.16	19	98.33	23	99.58	27	NA	31	98.75
4	95.41	8	96.24	12	97.08	16	98.33	20	98.75	24	97.5	28	97.91	32	99.16

Table 2. Accuracies for dominance emotion

All of this data was plotted and the results are as follows:





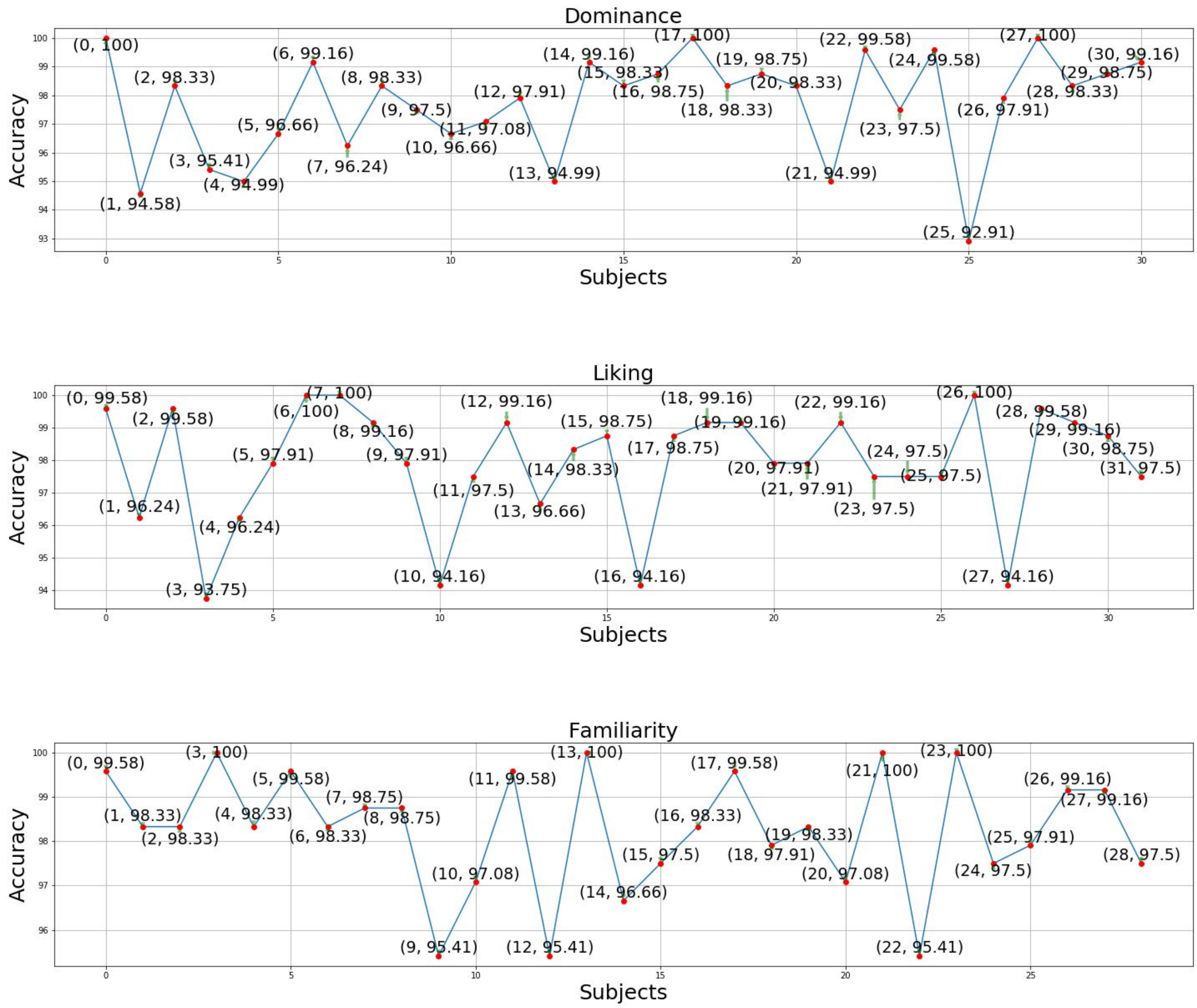


Fig. 10 Accuracies for Emotions.

6.2 Conclusion

We took baseline signals into account and a simple, computation cheap pre-processing method has been proposed to improve recognition accuracy. In addition, we are trying to apply a hybrid neural network to classify human emotion states by effectively learning compositional spatial-temporal representation of raw EEG streams. Experimental results have shown that the pre-processing approach can improve the accuracy by almost 30% and the model could achieve a high accuracy for valence and arousal classification task, respectively.

Our model has an accuracy of 97.4, 97.39, 97.68, 97.90, and 98.26 for valence, arousal, dominance, liking and familiarity respectively which is current state of the art as shown in Fig 3.

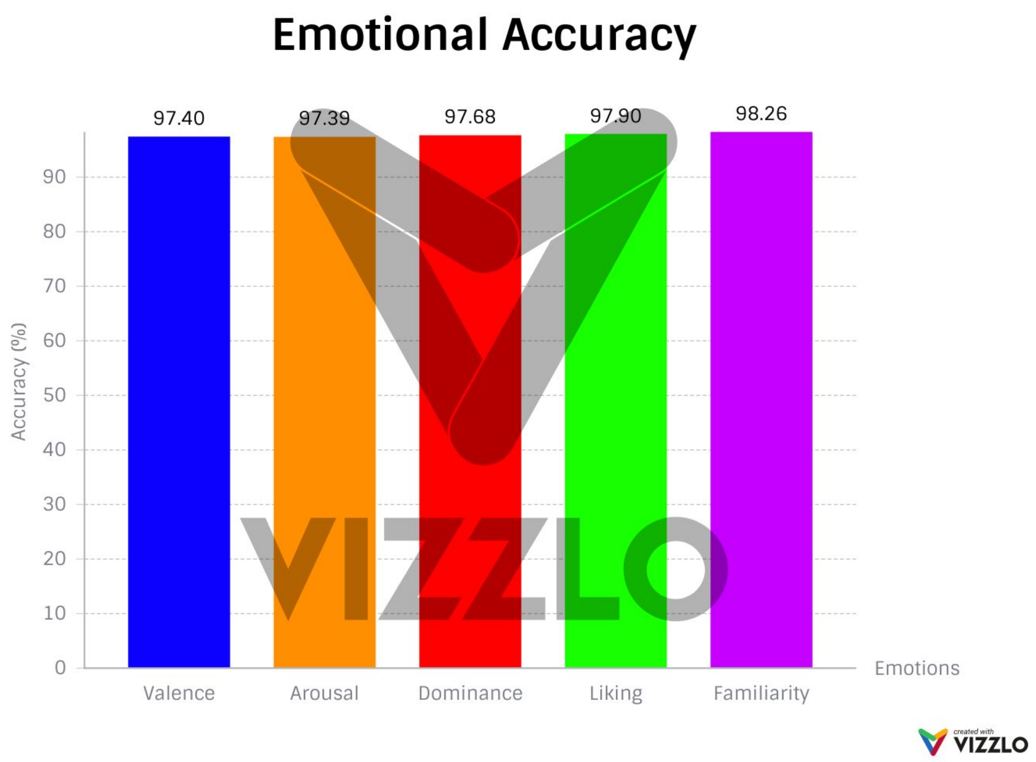


Fig.11 Emotional Accuracy

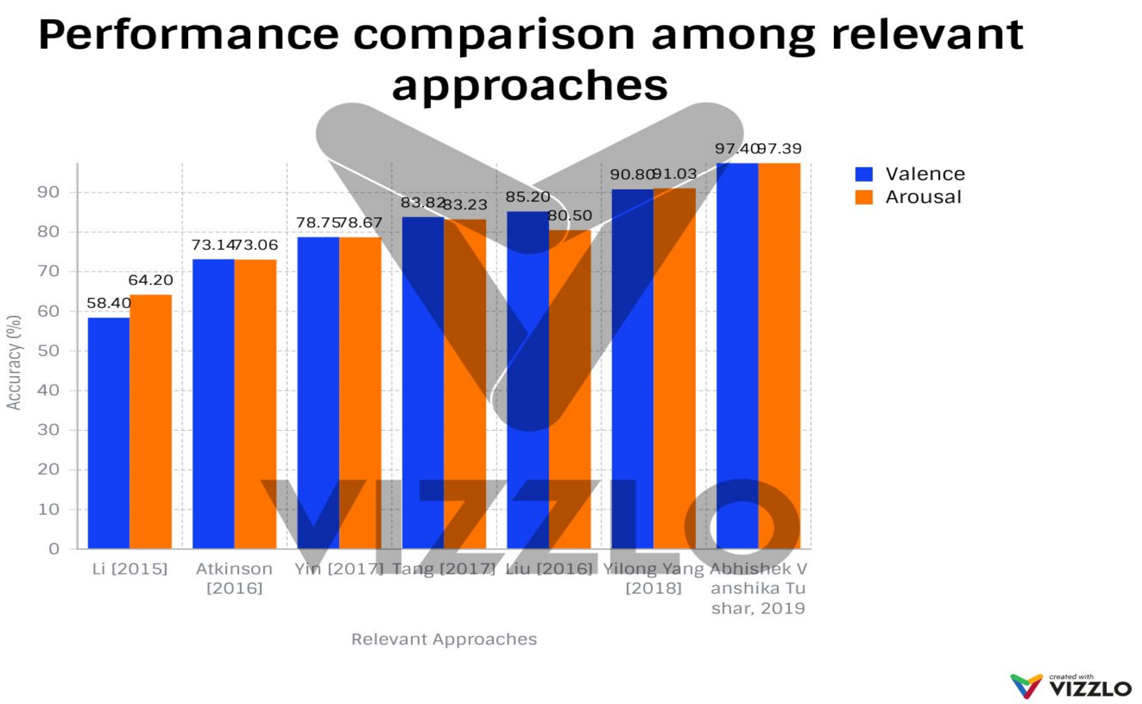


Fig 3. Performance comparison

Accuracies of valence, arousal, dominance, liking and familiarity are also shown in Fig.9.

For each emotion we generated two graphs of each subject displaying maximum accuracy and minimum loss. As of now maximum accuracy and minimum loss for subject 01 for arousal is shown in Fig 10 and Fig 11.

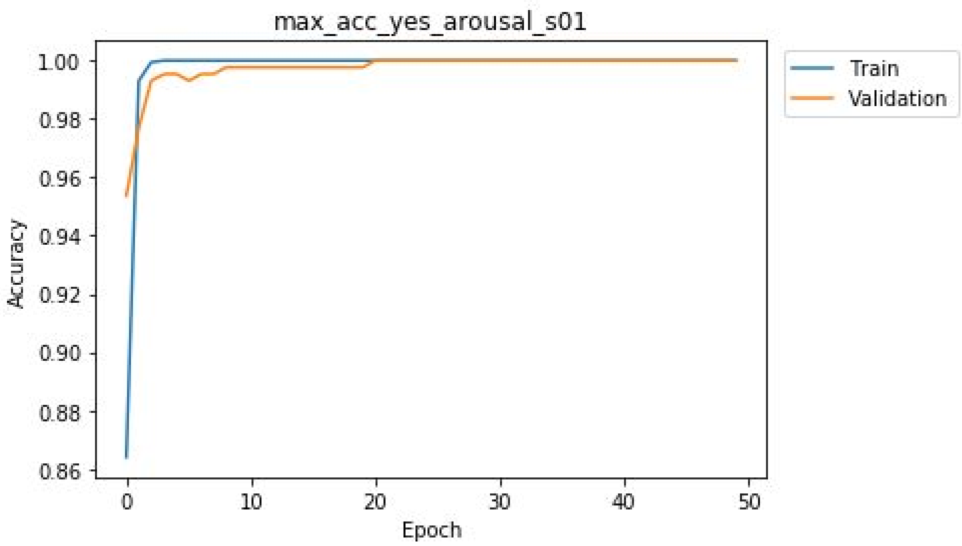


Fig 12. Maximum accuracy for arousal for subject 1

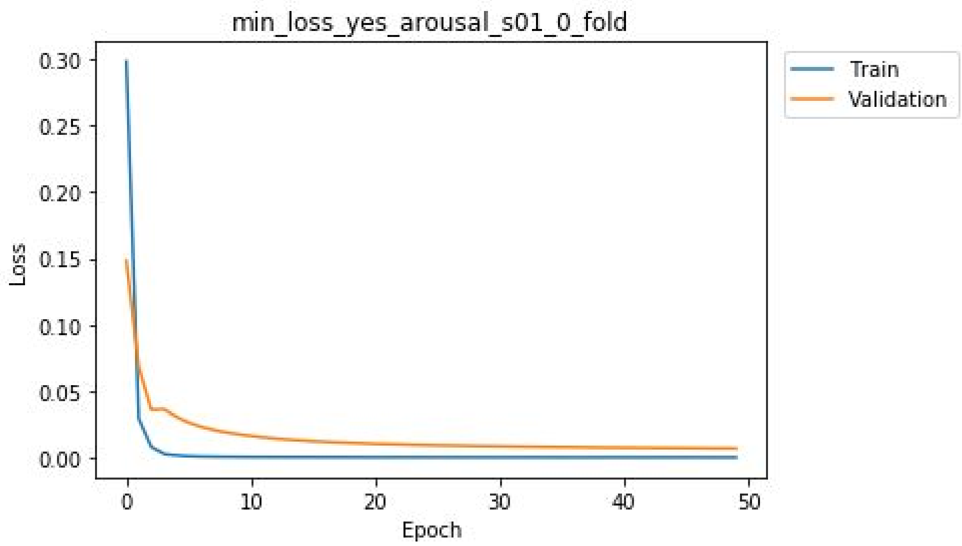


Fig 13. Minimum loss for arousal for subject 1

6.3 Future Work

Military

- Specialised military and police applications.
- For training (tracking cadet's mental state).
- Detecting stress in drivers, pilots, cadets.
- For criminal interrogation (will keep a check on crime rate).
- Lie detection (polygraphs test can be faked but brain signals can't be).

Health Sector

- Detecting mental state of babies, autistic / mentally paralysed patients / specially abled people / patients in coma / trauma patients / feedback for how someone is being treated.
- Psychology and Neurology in diagnosis of patients.
- Helping psychiatrists to deal with patients better as they'll have a better understanding of their emotions.
- Indicating post traumatic stress disorders, major depression and anxiety.
- Research shows with the help of valence, arousal and dominance, we can measure levels of burnout and productivity.
- Fits / Seizure diagnosis.

Education

- Detecting confusion in students (would enhance e-learning).
- Webinar feedback.

Honest Feedbacks

- Sentiments towards politicians, products, countries, ideas, personalities, ideologies,movies, music, games, etc.

Luxury Uses

- Automating lighting conditions and room environment (fragrance, temperature etc.) according to the emotions user is experiencing.

Evolutionary in Entertainment & Gaming Industry

- Could generate great revenues by having a gripping effect on the players / viewers.
- Making games more interactive based on user's emotional state and indulgence. Example: first person shooters, strategy games, racing, arcade, fighting, RPG, open-world games, virtual reality, altering the storyline to player's interests.

REFERENCES

[1] Jenke, Robert, Angelika Peer and Martin Buss. "Feature extraction and selection for emotion recognition from EEG." IEEE Transactions on Affective Computing 5.3 (2014)

[2] Alarcao, Soraia M., and Manuel J. Fonseca. "Emotions recognition using EEG signals: a survey." IEEE Transactions on Affective Computing (2017).

[3] Chen, Peng, and Jianhua Zhang. "Performance Comparison of Machine Learning Algorithms for EEG-Signal-Based Emotion Recognition." International Conference on Artificial Neural Networks. Springer, Cham, 2017.

[4] Yin, Zhong, et al. "Cross-subject EEG feature selection for emotion recognition using transfer recursive feature elimination." Frontiers in neurorobotics 11 (2017).

[5] LeCun, Yann, Yoshua Bengio, and Geoffrey Hinton. "Deep learning." Nature 521.7553 (2015)

[6] Tabar, Yousef Rezaei, and Ugur Halici. "A novel deep learning approach for classification of EEG motor imagery signals." Journal of neural engineering 14.1 (2016)

[7] Zhang, Dalin, et al. "EEG-based Intention Recognition from SpatioTemporal Representations via Cascade and Parallel Convolutional Recurrent Neural Networks." arXiv preprint arXiv:1708.06578 (2017)

[8] Li, Xiang, et al. "Emotion recognition from multi-channel EEG data through Convolutional Recurrent Neural Network." Bioinformatics and Biomedicine (BIBM), 2016 IEEE International Conference on. IEEE, 2016.

[9] Li, Youjun, et al. "Human Emotion Recognition with Electroencephalographic Multidimensional Features by Hybrid Deep Neural Networks." Applied Sciences 7.10 (2017): 1060.

[10] Tang, Hao, et al. "Multimodal Emotion Recognition Using Deep Neural Networks." International Conference on Neural Information Processing. Springer, Cham, 2017.

[11] Koelstra, Sander, et al. "Deap: A database for emotion analysis; using physiological signals." IEEE Transactions on Affective Computing 3.1 (2012): 18-31.

[12] Li, Xiang, et al. "EEG based emotion identification using unsupervised deep feature learning." (2015).

[13] Atkinson, John, and Daniel Campos. "Improving BCI-based emotion recognition by combining EEG feature selection and kernel classifiers." Expert Systems with Applications 47 (2016).

[14] Liu, Wei, Wei-Long Zheng, and Bao-Liang Lu. "Emotion recognition using multimodal deep learning." International Conference on Neural Information Processing. Springer International Publishing, 2016.

[15] Wang, Xiao-Wei, Dan Nie, and Bao-Liang Lu. "Emotional state classification from EEG data using machine learning approach." Neurocomputing 129 (2014): 94-106.