RB Journal of Lib & Information Science ISSN: 0972-2750
(UGC Care Group I Listed Journal) Vol-12 Issue-06 No.01: 2022

HUMAN EMOTION IDENTIFICATION WITH STACKED LSTM & RNN

K.N.V.SATYANARAYANA[1]

Research scholar Department of ECE Annamalai University Chidambaram Tamil Nadu 608002 India

Dr.T.SHANKAR[2]

Assistant Professor (Deputed from Annamalai University) Department of ECE, Government College of Engineering, Srirangam, Trichy, Tamilnadu, India.

Dr.P.V.RAMA RAJU[3]

Professor, ECE S.R.K.R. Engineering college, Bhimavarm, A.P India.

ABSTRACT

In modern day era the need of communication between the humans and machines is very essential for the technological advancement. In this paper we will thoroughly discuss about human emotion Identification by using EEG (electroencephalogram) signals with the help of stacked LSTM (Long Short-Term Memory and Recurrent Neural Networks, Basically LSTM model and is used for complex domains like speech recognition, machine translation, emotion identification etc. The model we use in this study gives the best performance in time series prediction problems. It gives us an accuracy of 97.19% and F1 score of 0.93

KEY WORDS: EMOTION IDENTIFICATION, EEG, DEAP DATASET, RNN, LSTM, STACKED LSTM

INTRODUCTION

Emotion recognition plays a vital role in our day-to-day life. Practical assessment of human emotions will improve people's life and is useful to do better. Emotion recognition plays a major role in 2 types of activities in one's life, primarily it is used in human machine interaction, it is a prime factor in BCI (Brain Computer Interface) [1],[2],[3] systems for better communication between humans and computers as well as in real world applications [8][9]. Secondly it is very important in medical sector [7] for those who are unable to express their feelings through words and those who are paralyzed. If emotion recognition is done successfully then doctors have a basic idea about the actual problems of their patients. Basically, human brain has too many emotions to be detected. After a long research some scientists classified the human emotions into 3 categories positive, negative and neutral emotions [11][12] Emotions like happiness ,pride ,love ,hope ,excitement ,enthusiasm all these comes under positive emotions where as emotions like anger ,sad ,fear ,guilt goes under negative category. Emotion recognition can be done based on the reaction that a person gives or feels when an incident happens. Basically various methods are introduced to find emotions using[6][20]

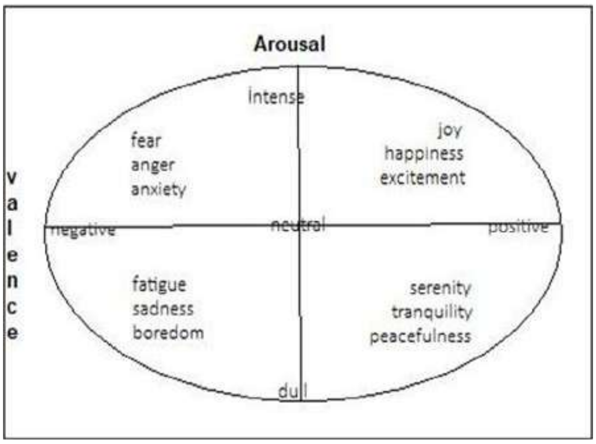


Figure 1: Arousal Model

EEG. Some of them are Time Domain techniques, Frequency Domain techniques, Joint Time Frequency Domain Techniques and some other techniques

INTRODUCTION TO EEG

Electroencephalogram (EEG) is a test that detects electrical activity in our brain using small, metal discs(electrodes) attached to your scalp. Emotions are classified based on the electric potentials of the brain signals. These signals which are caused by the US neurons are collected by using electrodes placed on the scalp. These signals are nothing, but the electrical pulses created by the difference in voltages between the minimum of two electrodes. This difference must be noted simultaneously to produce a better result of a particular event that is being recognized [17]. Feature extraction and emotion recognition based on empirical mode decompositions (EMD) [18][19]. Generally Time Frequency analysis is based on the spectrum of EEG signals ,power ,PSD ,and differential entropy of some specific sub bands are usually made use of as features. To calculate this spectrum some of the most common techniques are Short-Time Fourier Transform (STFT) and Hilbert-Huang Transform(HTT) and also Discrete Wavelet Transform (DWT)[1][4].Higher frequency Sub bands like Beta(16-32 Hz),Gamma(32-64 Hz) bands gives an impeccable results when compared to lower sub bands when used for emotion detection Statistics of EEG series, that includes First and Second difference, Mean value and power are generally used in time domain. Nonlinear features,[1] which includes fractal dimensions, sample entropy and non-stationary index are employed for emotion recognition [8][15]

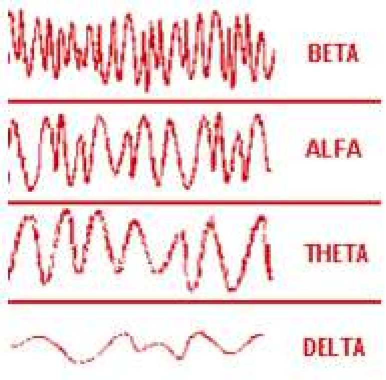


Figure 2: Samples of brain waves at different frequencies

EEG can be measured using different methods like LSTM, RNN and CNN. LSTM is an effective implementation of RNN. In this paper we used stacked LSTM which is the further implementation of LSTM.

LSTM

In this paper, an LSTM network model is used for predicting the emotions of individual persons by giving their EEG data as input. LSTM means Long Short-Term Memory Networks. LSTM is an upgraded version of recurrent neural networks (RNNs). Hochreiter and Schmidhuber introduced LSTM in the year 1997. Traditional RNNs are not good at capturing long-range dependencies i.e. when a huge dataset is given to the multiple RNN layer it leads to the vanishing gradient problem which means the gradient which is used to update the weights of the neural network decreases exponentially, sometimes it even stops the training of the neural network. LSTM can remember the RNNs weights and their inputs over a very long period. In natural language processing the sequential data is handled by LSTM network models. It is also used in speech recognition and in the prediction of time series. In LSTM in addition to the hidden state, the cell state is also passed down to the next block [12] The basic architecture of the LSTM cell consists of three gates. Below fig represents the architecture of the LSTM cell [5].

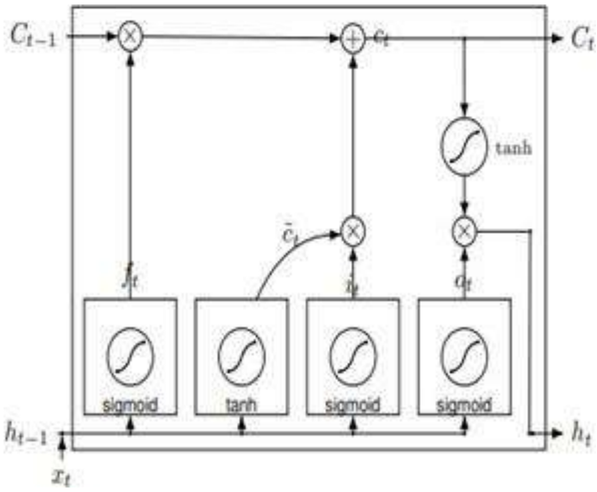


Figure 3: LSTM Cell Architecture

First one is a forget gate. It removes the data that is no longer necessary in the cell state.

This action is done by using sigmoid layer, which is illustrated by the following equations

$$ft = \sigma(Wf \cdot [ht-1, xt] + bf) \qquad (1)$$

The second is input gate, it adds further data to the cell state. It additionally consists of sigmoid layer, it decides that values ar to be updated. The new updated values ar regenerate into vector kind exploitation tanh layer. These equations ar explained exploitation below equations

$$it = \sigma(Wi \cdot [ht-1, xt] + bi) \qquad (2)$$

$$C\tilde{}\, t = \tanh(Wc \cdot [ht-1, xt] + bc) \quad (3)$$

And finally we have output gate. It adds useful information to the cell state.

c(t) obtained by the equations is

$$Ct = ft * Ct-1 + it * C\tilde{}\, t \qquad (4)$$

Finally using the below equations,

$$ot = \sigma(Wo \cdot [ht-1, xt] + bo) \qquad (5)$$

$$ht = ot * \tanh(Ct) \qquad (6)$$

the output of the current state is calculated using updated cell state and a sigmoid layer .it decides which parts of the cell state will gives the final output

Here, σ is activation function,

ht−1 is past hidden state,

Wf,Wi,Wc,Wo are weight matrices, bf , bi , bc, bo are biasing vectors, xt is input vector,

tanh is hyperbolic tangent.

STACKED LSTM

Stacked LSTM is a deep Learning technique. Generally, LSTM [4] model consists of a single hidden LSTM layer along with a feed-forward output layer. Whereas the stacked LSTM is a development of the LSTM model. To account for the intrinsic dependencies, the stacked LSTM network is used as a primary learning model. It has multiple hidden LSTM layers, and each LSTM layer consists of various memory cells. Piling up the hidden layers of the LSTM makes the model deeper, giving more accuracy. It is the depth of the neural networks that determine the success of the model for many problems that are faced in sequence predicting. Stacked LSTM workflow takes in the forward direction. Time-varying input is given to it. When a time-varying input is given, the same input is processed by other deep layers also but with different time steps. Information from deep-stacked LSTM is combined which produces the output for varying time steps[21]. Stacked LSTM provides the best performance with time-series data when compared with other forms of LSTM. The below figure represents the general architecture of the stacked LSTM model [6].

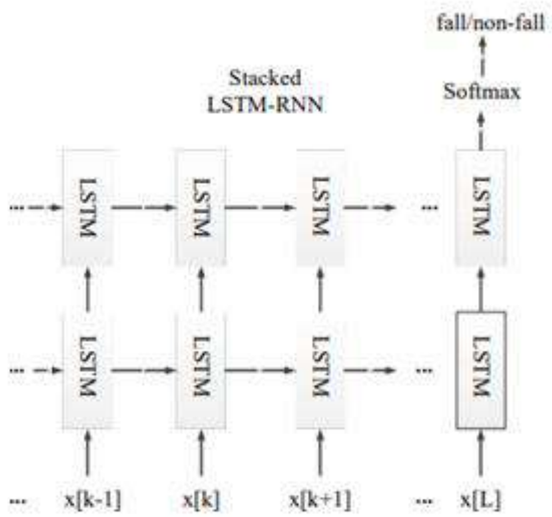


Figure 4: Stacked LSTM architecture

The stacked LSTM structure is achieved by combining the multiple layers of LSTM one above the other. A Stacked LSTM structure can be described as an LSTM mannequin comprised of more than one LSTM layer. An LSTM layer above presents a sequence output as an alternative to a single value output to the below LSTM layer [10]. To be precise, one output per entering time step is an alternative to one output time step for all enter time steps. Given that LSTM functions on sequential data, it means that the summation of layers provides stages of abstraction of given input observations over time. In effect, chunking observations over time or representing the trouble at distinct time scales. When compared with normal single layer LSTM structure and Stacked LSTM, the latter provides a better outlook for extracting rich features and increasing. By observing the fig we can say that the LSTM cell takes the time samples x[k] and the output from the previous time step.[14] Then the weighted inputs are added and passed through the activation function which is tan h. The final LSTM cell is connected to an output layer with a softmax activation function for classification. By using Keras Python's deep learning library, LSTM models can be created. For each LSTMs perceptron model, a 3D input is required. If one input sequence of time steps is processed by an LSTM, then it's each memory cell will output a single value for the whole sequence as a 2D array.

RNN

A recurrent neural network or RNN is a type of artificial neural network which is designed to recognize sequential characteristics in patterns of a given dataset, which can then be used to predict the next likely scenario. The sequential Data is Important in Recurrent neural networks. The sequential Data includes audio, Strings of data and even EEG data Feed-forward neural networks are composed of An input, hidden, and output layer, which passes information from each layer to the next. An RNN has a looping mechanism that allows for information to re-enter the hidden layer, this is information is called the hidden state, which is a representation of all the previous inputs. However, RNNs suffer from short-term memory, caused By the in famous vanishing gradient problem, which is predominant in neural networks that use gradient- based learning methods and back propagation [16][13].The first step to training a feed-forward neural network is a forward pass and making a prediction, this prediction is then compared to the ground truth using a loss function, which then returns an error value .This Error value is an indication of how bad the neural network has performed, which is then used to perform back propagation to calculate the gradients of each node of the neural network, starting from the Nodes of the layers closest to the output layer, and working its way back to the first hidden layer.

A gradient is a value that measures the change in weight with regard to the change in error ;this is used to adjust the network's internal weights, allowing the network to learn, the greater the gradient, faster the model can learn, and here is where the vanishing gradient problem comes in

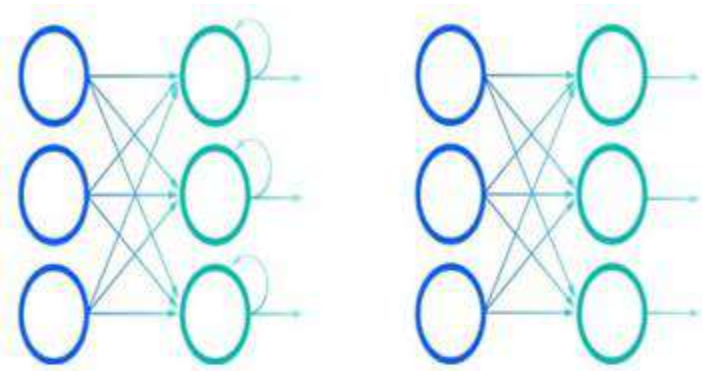


Figure 5 RNN Network

RNN vs FEED FORWARDNETWORK

RNNs use form of backpropagation called 'backpropagation through time', which is essentially tailored for RNNs by treating each loop as a hidden layer, and here, the gradients' value will exponentially shrink as it propagates through each loop. Due to the vanishing gradients, the RNN is unable to learn the dependencies from layers / loops further back, meaning there is a chance that the earlier segments of the EEG data are not considered when Making a prediction if the dataset is large enough, giving our model short-term memory. To address this problem, two types of specialized RNNs were created: LSTM (Long Short-Term Memory) and GRU (Gated Recurrent Unit). These work exactly like RNNs; however, they employ these mechanisms called 'gates' which allow them to learn long-term dependencies, unlike a regular RNN. These 'gates' are essentially different tensor operations that can learn which information to add or remove to the hidden state—and because of this ability, short-term memory [15].

APPROACH

IMPORTING LIBRARIES, the first and the most needed step of any machine Learning model is to import the necessary Packages. In this paper we have imported the Packages NumPy, Pandas, TensorFlow, Keras ,Sklearn, Matplotlib, Seaborn

NUMPY AND PANDAS

NumPy and Pandas plays a huge role in data Manipulation operations. NumPy plays a role in performing large numerical operations. Pandas Is used for import files in various formats.

TENSORFLOW AND KERAS

TensorFlow plays a role in deep learning techniques Like classification, prediction. Keras is used for Performing deep learning operations on Small Size datasets.

SKLEARN

Sklearn is the most widely used package As it contains many most important statistical Models like regression, clustering, dimensional Reductionality of machine learning.[22]

MATPLOTLIB AND SEABORN

These two packages are used for performing Data visualization techniques on data for better Understanding of results.

DATA COLLECTION AND UNDERSTANDING OF DATA

Data is collected and loaded using pd.read function in pandas and data is understood by seeing the head and tail of a dataset which gives the first five and last five rows of a dataset. Later we analyze the data using info and describe functions which give the information and five number summary of a dataset respectively.

PREPROCESSING OF DATA

In the preprocessing step we check for null values if there Are any null values, we impute them using imputation techniques. Then in the next step of preprocessing we encode the data Having more than two categories using label encoding as Positive, neutral, negative. Then we perform scaling to bring the values to a related Range to other columns.

BUILDING MODEL

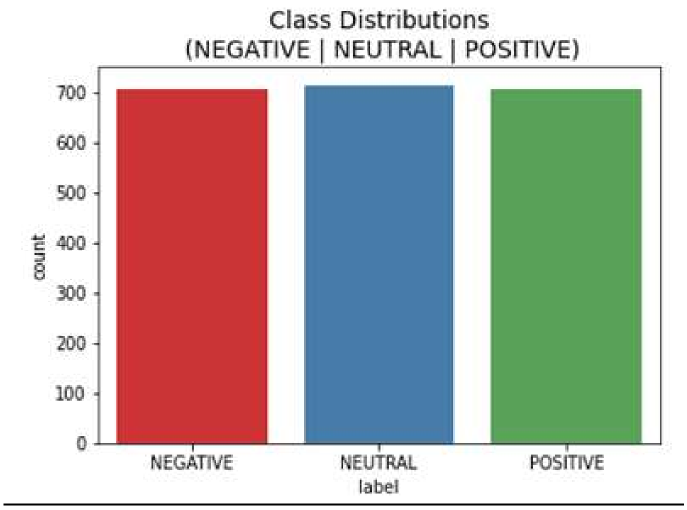
In this paper we have built the model using Stacked LSTM which is a sub classifier of RNN From the imported deep learning packages TensorFlow and Keras we have used various Sub packages like sequential, Dense, Dropout, Embedding, LSTM. Sequential is used for forming a stack of layers for which it will contain a single input and single Output. Dense function is used to link all outputs with every input. The main function of Dropout in keras is to prevent the model from overfitting it randomly allocates zeros to some values and balances other values to maintain the sum as a constant value. We embed the text to a machine understandable vector of fixed length. The main purpose of using LSTM over other methods Is that it is very efficient in sequential predictions Where it is able to store long term items and short-Term items. All these sub packages help us in building A better predictive model.

TRAINING AND TESTING THE DATA

In this step we train the data using train_test_split from sklearn. It is used in splitting the data in to 80-20 or 75-25percentage for training and testing respectively. The training data is loaded in to one Variable and test data in to other variable. The difference between them is the presence of target Variable in the training data.

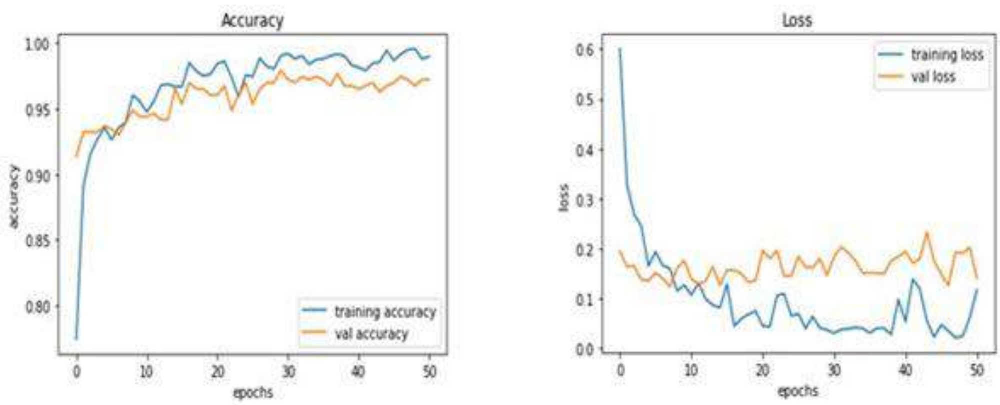
RESULTS

 ISSN: 0972-2750
Vol-12 Issue-06 No.01: 2022

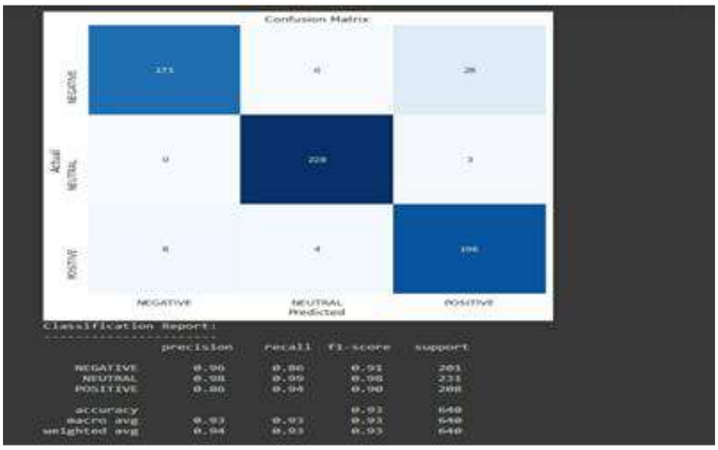


Using LSTM:-

Model evaluation metrics scores Gives a accuracy of 97.19% on a test data.



Using RNN :-




```
[ ] model_acc = model.evaluate(X_test, y_test, verbose=0)[1]
    print("Test Accuracy: {:.3f}%".format(model_acc * 100))

    Test Accuracy: 93.281%
```

ISSN: 0972-2750
Vol-12 Issue-06 No.01: 2022

Model evaluation metrics scores gives accuracy of 93.281% on a test data

CONCLUSION

Finally, the conclusion of this paper is that the model is best for its prediction of emotions and it categorized them into three categories as positive, negative and neutral with the help of method known as stacked LSTM using RNN which obtains a best accuracy of 97.13% The stacked LSTM using RNN is the learning technique used for storing data for a long run with the help of multiple LSTM layers which predicts the emotions based on the past data. The future extraction can be emotions can be further classified to improve the accuracy.

FUTURE SCOPE

Further these classifications can be done by using RGNN (Regularized graphical neural networks) and Other Hybrid Machine Learning Algorithms in future and we can get more F-1 score and Accuracy compared to LSTM using RNN Algorithm.

REFERENCES

[1] K.n.v.satyanarayana " A Survey on Assessment of Sentiments and Emotions Using EEG Signals" intentional journal of advanced science and technology(IJAST),vol.29,no.3,pp.717-728.2020.
[2] B. W. Schuller, ''Speech emotion recognition: Two decades in a nutshell, benchmarks, and ongoing trends,'' Commun. ACM, vol. 61, no. 5, pp. 90–99, 2018.
[3] M. S. Hossain and G. Muhammad, ''Emotion recognition using deep learning approach from audio–visual emotional big data,'' Inf. Fusion, vol. 49, pp. 69–78, Sep. 2019.
[4] M. Chen, P. Zhou, and G. Fortino, ''Emotion communication system,'' IEEE Access, vol. 5, pp. 326–337, 2016.
[5] Mouhannad Ali, Fadi Al Machot, Ahmad Haj Mosa, and Kyandoghere Kyamakya. 2016. A Novel EEG-based Emotion Recognition Approach for e-Healthcare Applications. In Proceedings of the 31st Annual ACM Symposium on Applied Computing (SAC '16). ACM, New York, NY, USA, 162--164.
[6] Salma Alhagry, Aly Aly, and Reda El-Khoribi. 2017. Emotion Recognition based on EEG using LSTM Recurrent Neural Network. International Journal of Advanced Computer Science and Applications 8, 10 (2017). https://doi.org/10.14569/IJACSA.2017.081046Google Scholarlstm
[7] Satyanarayana, K. N. V. "An Approach For Finding Emotions Using Seed Dataset With Knn Classifier." Turkish Journal of Computer and Mathematics Education (TURCOMAT) 12.10 (2021): 2838-2846.
[8] Satyanarayana, KN V., et al. "An Approach to EEG based Emotion Identification by SVM classifier." 2022 6th International Conference on Computing Methodologies and Communication (ICCMC). IEEE, 2022.
[9] H. Sadreazami, M. Bolic and S. Rajan, "On the Use of Ultra-Wideband Radar and Stacked LSTM-RNN for at Home Fall Detection," 2018 IEEE Life Sciences Conference (LSC), 2018, pp. 255-258.
[10] A. B. Nassif, I. Shahin, I. Attili, M. Azzeh, and K. Shaalan, ''Speech recognition using deep neural networks: A systematic review,'' IEEE Access, vol. 7, pp. 19143–19165, 2019.
[11] K. R. Scherer, ''What are emotions? And how can they be measured?'' Social Sci. Inf., vol. 44, no.4, pp. 695–729, 2005.
[12] Naga V. Kanuboyina, Satyanarayana, and Rama Raju Venkata Penmetsa. "Electroencephalograph based Human Emotion Recognition Using Artificial Neural Network and Principal Component Analysis." IETE Journal of Research (2021): 1-10.
[13] T. Balomenos, A. Raouzaiou, S. Ioannou, A. Drosopoulos, K. Karpouzis, and S. Kollias, ''Emotion analysis in man-machine interaction systems,'' in Proc. Int. Workshop Mach. Learn. Multimodal Interact. Springer, 2004,pp. 318–328.
[14] Satyanarayana, KN V., et al. "Human Emotion Classification using KNN Classifier and Recurrent Neural Networks with Seed Dataset." 2022 6th International Conference on Computing Methodologies and Communication (ICCMC). IEEE, 2022.

[15] Practical Approach to Electroencephalography-handbook.

[16] Hyun Joong Yoon and Seong Youb Chung. Eeg-based emotion estimation using bayesian weighted-log-posterior function and perceptron convergence algorithm. Computers in biology and medicine, 43(12):2230– 2237, 2013.

[17] Daimi Syed Naser and Goutam Saha. Recognition of emotions induced by music videos using dt-cwpt. In Medical Informatics and Telemedicine (ICMIT), 2013 Indian Conference on, pages 53–57. IEEE, 2013.

[18] Jingxin Liu, Hongying Meng, Asoke Nandi, and Maozhen Li. Emotion detection from eeg recordings. In Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD), 2016 12th International Conference on, pages 1722–1727. IEEE, 2016.

[19] Anuja R Bhagwat and AN Paithane. Human disposition detection using eeg signals. In Computing, Analytics and Security Trends (CAST), International Conference on, pages 366–370. IEEE, 2016.

[20] Sepideh Hatamikia and Ali Motie Nasrabadi. Recognition of emotional states induced by music videos based on nonlinear feature extraction and som classification. In Biomedical Engineering (ICBME), 2014 21st Iranian Conference on, pages 333–337. IEEE, 2014.

[21] Sepp Hochreiter and Jurgen Schmidhuber. Long short-term memory. ¨ Neural computation, 9(8):1735–1780, 1997.

[22] Francois Chollet et al. Keras. https://github.com/fchollet/keras, 2015.