MAULANA AZAD

NATIONAL INSTITUTE OF TECHNOLOGY BHOPAL, 462003

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

MAJOR PROJECT ON

Emotion Recognition using Machine Learning on DEAP Dataset

UNDER THE GUIDANCE OF

Dr Mitul Kumar Ahirwal

SUBMITTED IN PARTIAL FULFILMENT FOR THE DEGREE OF BACHELOR OF TECHNOLOGY

SUBMITTED BY:

VIPUL SHARMA	171112040
SARTHAK RAGHUWANSHI	171112051
ANIRUDH GOYAL	171112102
BHASKAR KUMAR	171112096

SESSION: 2020-21

MAULANA AZAD

NATIONAL INSTITUTE OF TECHNOLOGY BHOPAL, 462003

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

This is to certify that **VIPUL SHARMA, SARTHAK RAGHUWANSHI, ANIRUDH GOYAL AND BHASKAR KUMAR**, students of B.Tech. Final Year (VIII SEMESTER), have successfully completed their project entitled **“Emotion Recognition using Machine Learning on DEAP Dataset”** in partial fulfilment of their major project in Computer Science & Engineering.

Dr Mitul Kumar Ahirwal
(Project Guide)

MAULANA AZAD NATIONAL INSTITUTE OF TECHNOLOGY BHOPAL

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING\

DECLARATION

We, hereby, declare that the following report which is being presented in the Major Project Documentation entitled **"Emotion Recognition using Machine Learning on DEAP Dataset"** is the partial fulfilment of the requirements of the final year (eighth semester) **Major Project** in the field of **Computer Science & Engineering**. It is authentic documentation of our original work carried out under the guidance of **Dr Mitul Kumar Ahirwal**. The work presented here is carried out entirely at **Maulana Azad National Institute of Technology, Bhopal**. The following project and its report, in part or whole, has not been presented or submitted by us for any purpose in any other organization or institution.

We, hereby, declare that the facts mentioned above are true to the best of our knowledge. In case of unlikely discrepancy that may occur, we will be the ones to take responsibility.

NAME	SCHOLAR NO.	SIGNATURE
VIPUL SHARMA	171112040	
SARTHAK RAGHUWANSHI	171112051	
ANIRUDH GOYAL	171112102	
BHASKAR KUMAR	171112096	

ACKNOWLEDGMENT

With due respect, we express our deep sense of gratitude and thanks to our respected guide **Dr Mitul Kumar Ahirwal**, for his valuable help and guidance. We are thankful for the encouragement that he has given us in completing this project successfully. His rigorous evaluation and constructive criticism were of great assistance.

It is imperative for us to mention the fact that this major project could not have been accomplished without the periodic advice and suggestions of our project coordinator **Dr S K Saritha** and **Dr Manish Pandey**.

We are also grateful to our director **Dr N. S. Raghuwanshi,** for permitting us to utilize all the necessary facilities in the college.

Needless to mention is the additional help and support extended by respected HOD, **Dr Nilay Khare**, in allowing us to use the departmental laboratories and other services.

We are also thankful to all the other faculty, staff members and laboratory attendants of our department for their kind cooperation and help. Last but certainly not least, we would like to express our deep appreciation towards our family members and batch mates for providing the much-needed support and encouragement.

ABSTRACT

In the proposed project an emotion recognition or classification system which is based on the valence/arousal model is proposed to be created. Electroencephalography (EEG) signals will primarily be used for creating this model. Different stimuli elicit different responses in the EEG signals. Different kinds of video stimulus will be used and its corresponding emotional effect which is determined using EEG signals. Our goal is to create an objective system which can determine the type of response the videos will create in any subject and thus can classify the video in terms of their emotional category, which is defined using the valence/arousal scale.

This project will create an emotion classification system as well as a classification system for videos i.e., will enable a user to objectively determine the kind of video using EEG signals from a person who is watching the video. Such a system may be beneficial in creating a recommendation system that is truly objective in nature.

The project uses 1D Convolutional Neural Network model for this classification purpose. The EEG signals from 32 channels are first reduced to 14 channels of symmetric differences in the EEG channels, then a model was applied on these 14 channels giving 75% weighted accuracy for emotional classification into 4 classes.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Emotion is a complex reaction pattern, involving behavioral, experiential and physiological elements. It can be associated with personality, mood or temperament. There has been extensive research done on the relation of EEG with emotions, much of these methods involve researchers manually looking at EEG signals to identify regions corresponding to a certain type of emotion.

Recently the application of Machine Learning to automate this process of classification of emotion from EEG has been widely worked upon. In this project, the DEAP dataset is used. Music clips are used as visual stimuli to obtain different types of emotions from subjects. The emotions are analyzed and classified into 4 regions using a Valence-Arousal Model of emotional classification.

The project utilizes the raw EEG signals with some preprocessing to classify the emotional response of a participant to a video stimulus. Many deep learning models have been proposed in this field, but most models first apply transformations like Discrete Wavelet Transform or Power Spectral Density calculation over EEG, in this project a 1D CNN is applied over the raw EEG signal achieving a weighted accuracy of 75% illustrating the power of Deep Learning Models.

1. THEORETICAL ASPECTS

1. Valence-Arousal Model

The proposed project plans to use valence-Arousal model for describing emotions more quantitatively. In this model, one can place each emotional state on a 2-D plane with arousal and valence as the x and y axis respectively (it can be seen in fig 1). One can also include dominance as a third dimension for deeper understanding of emotional state. Valence can range from unpleasant to pleasant, whereas arousal ranges from inactive to active. The valence-arousal space can be divided into four quadrants, low arousal/low valence, low arousal/high valence, high arousal /low valence and high arousal/high valence. Valence-Arousal can also be determined using Linear Regression.

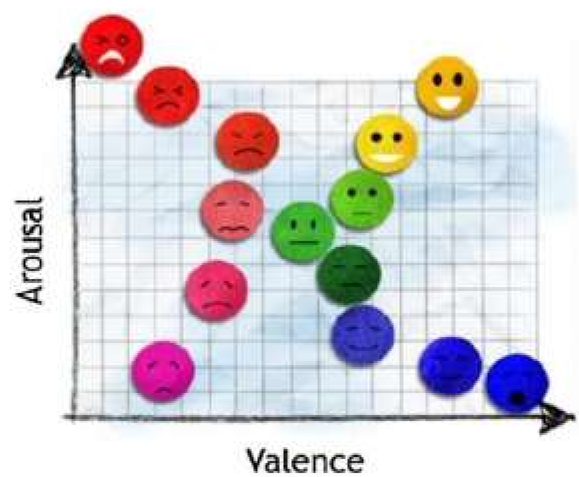

Figure 1: Valence-Arousal Model

2. DEAP Dataset

The DEAP dataset is obtained in the following way. At the start the researchers used a novel stimuli selection method and gathered a kind of large set of video clips. Then a 1 min segment of video clips used in the experiment were extracted. After which a subjective test was performed to select most appropriate video clips. Participants then took part in the dataset collection and their EEG signals were recorded at a sampling rate of 512 Hz using 32 active AgCl electrodes. After which each participant rated them based on valence and arousal. To maximize the strength of obtained emotions, the strongest volunteer ratings and at the same time a small variation type of videos were selected. The DEAP dataset has the highest number of peoples which participated in publicly available dataset for spontaneous emotions analysis from physiological signals. It is also the only database that uses music clips as emotion stimulation.

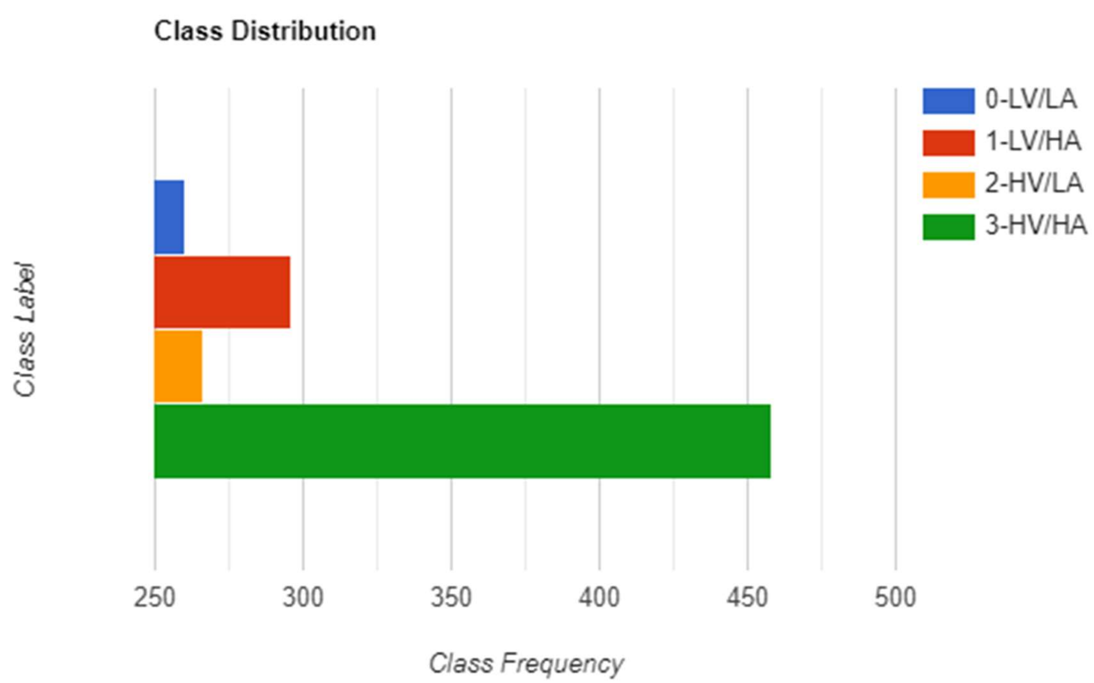

Figure 2: Frequency Distribution of Classes in the Dataset

3. Correlation between EEG and Emotions

Human brain insights can be provided using EEG technique. As EEG helps in detecting the smallest modulations which occur in the brain, to study emotion variance EEG is found to be a useful technique. EEG can detect various emotions like for example calm, happiness, stress, sadness, fear, surprise, etc. Valence has shown a very strong association with EEG signals and correlates 11 found across all frequency bands. It is observed that the association was seen to be less consistent with the observations made in the pilot study.

4. 1D-CNN models

The proposed project uses CNN on the raw EEG. As shown in Figure 3, in a 1D CNN a kernel with some width and height is passed over the time series data. While Passing over the time series it performs simple array operations of convolutions to learn discriminative features from the EEG signals.

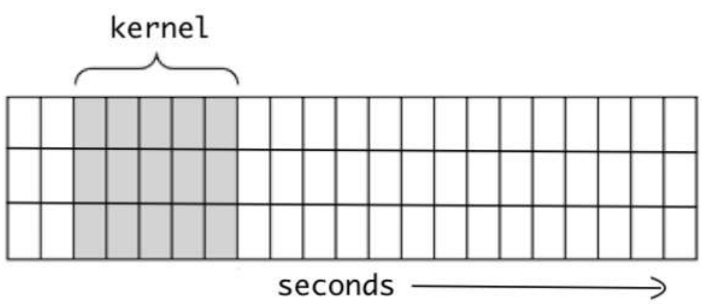

Figure 3: The 1D-CNN Kernel Passing over 3 channels of Time Series

Due to the low computational complexity of 1D CNN, they do not require specialized hardware to train and run in real-time. It is easier to train them, and they have shown good performance with shallow architectures as compared to very deep 2D-CNN architectures. There are many parameters to be decided in each CNN layer, the kernel size, number of filters, the stride length, the choice of activation functions etc.

3. LITERATURE REVIEW AND RESEARCH GAPS IDENTIFIED

Automatic EEG signal-based emotion recognition and classification is a recent problem which has mostly seen traditional methods applied over it. Various methods have been proposed like using Power Spectrum Density from EEG Signals and applying Naïve Bayes over it, or by applying K-Nearest Neighbor on any other handcrafted features.
Discrete Wavelet Transform is also a common technique used to extract features in the time-frequency domain from EEG signals, over such features SVMs etc. have been applied.
Some methods extract the statistical features from the signal like their mean, standard deviation, kurtosis etc. and apply Models over them.
Overall, the focus has been over the statistical features present in the EEG, researchers have reduced the time series to such features and then applied classical machine learning models over them.
Such models also focus on binary classification task of High Valence/Low Valence or High Arousal/Low Arousal. Individual models for each of them have been made to classify emotions.

In this project, these 2 issues have been dealt with. First, we used the raw EEG signals to obtain good classification accuracy as compared to using hand-engineered features. 1D-CNN is applied over raw EEG signal to obtain 75% accuracy. Secondly, a multi-classification problem has been proposed instead of having 2 separate binary classification problems. Multiclass classification is a more challenging problem, and it is tougher for a model to differentiate between 4 classes as compared to 2 as done by previous researchers.

4. PROPOSED WORK AND METHODOLOGY

1. Proposed Work

Proposed Architecture:

No.	Layer name	Kernel Size	Layer Parameters	Number of Parameters	Output Shape
1	Conv1D	128 x 64	Stride=1, Activation= ReLU	114816	7617 x 128
2	Max Pooling	2	-	0	3808 x 128
3	Conv1D	64 x 16	Stride=1, Activation= ReLU	131136	3793 x 64
4	Conv1D	16 x 4	Stride=1, Activation= ReLU	4112	3790 x 16
5	Conv1D	4 x 3	Stride=1, Activation= ReLU	196	3788 x 4
6	Max Pooling	2	-	0	1894 x 4
7	Flatten	-	-	0	7576
8	Dense	64	Activation= ReLU	484928	64
9	Dropout	-	Rate=0.4	0	64
10	Softmax	4	Activation= Softmax	260	4

Table 1: Description of the layers and parameters used in 1D-CNN model.

Figure 4: Layers of the 1D-CNN model used to classify EEG signals

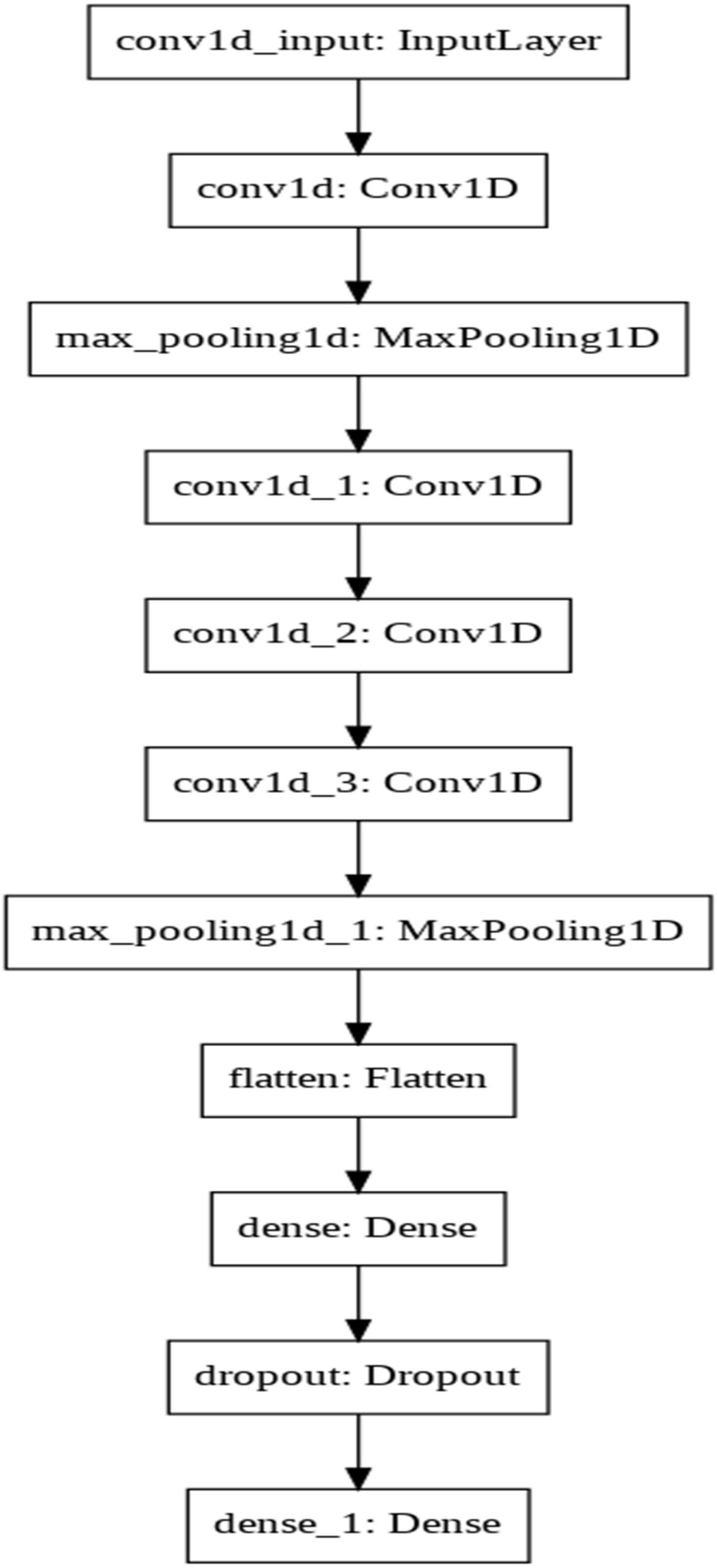

Table 1 and Figure 4 show the proposed model architecture in detail. The proposed model is a 1D-Convolutional Neural Network.

A 1D Convolution layer is the first layer of the model after input, it uses a kernel size of 64 with 128 filters and stride length as 1. ReLU activation is used at the output of this layer. 1 Dimensional Max Pooling is done on the output of the first layer with pool size of 2. Then 1D Convolution is repeated 3 times with number of filters as 64, 16 and 4 respectively, the kernel size as 16, 4 and respectively. All three of them have stride length as 1 and activation as ReLU. Another Max Pooling Layer with pool size of 2 is applied and its output is then flattened.

A Dense Layer with 64 neurons is then applied with ReLU activation as part of the classification part of the model. A dropout with rate 0.4 is applied for regularization purposes and to prevent coadaptation in the neurons. Finally, a Softmax layer with 4 neurons does the classification task into 4 classes.

Hyperparameter tuning was done to obtain the parameters listed in Table 2. Adam optimizer is used to reach optimal point for Loss function. A Learning rate of 0.01 is used with a decay of 1e-3. A custom decay scheduler was used which is described in the code. Categorical Cross-Entropy with Softmax layer was used for classification purposes. Accuracy of the model was monitored.

No.	Parameters	Values
1	Optimizer	Adam, beta1=0.9 and beta2=0.999
2	Learning Rate	0.01
3	Decay	1e-3
4	Loss Function	Categorical cross-entropy
5	Metrics	Accuracy
6	Batch Size	128
7	Epoch	100

Table 2: Hyperparameters used to get least Cross-Entropy Loss for model

2. METHODOLOGY

Algorithms and Flow Charts:

In the project, the aim is to create a classifier which will classify the emotion in a subject using the EEG signals generated in the brain of the subject during a trial of watching a video stimulus.

a. Input Instances (X) for Classifier:

The subject when shown a video during a trial generates EEG signals and these are a time series of length 8064 since sampling is done at 128 Hz for 63 seconds. 32 electrodes are used to read the EEG signals from the subject's brain. There are 40 videos, and each video is shown to each of the 32 subjects in the experiment. Hence for each subject data of 40*32*8064 is generated which is relevant to us.

b. Labels (Y):

After the trial, each subject rates the videos for the arousal and valence they felt after watching it. The rating is done on a continuous 1-9 scale. For the generation of labels from this data we have done thresholding of the values on the middle value of the scale i.e., 5. This provides us with two labels: High and Low Valence and similarly two labels for Arousal.

c. Metrics:

Accuracy:

It is simply a measure of the number of instances which were correctly predicted over the total number of instances.

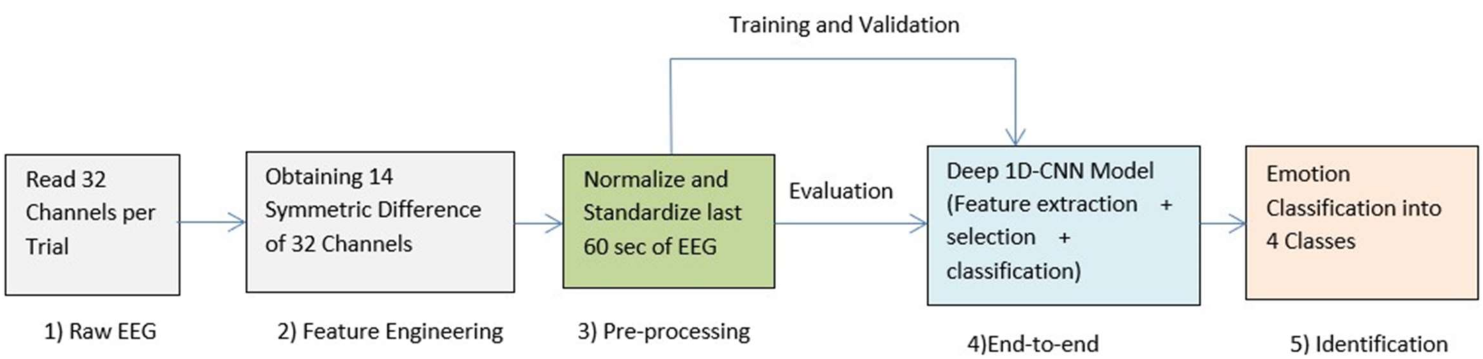

Figure 5: Flowchart for automatic emotion classification

Various Steps followed in the classification process have been illustrated in Figure 5. First a total of 1280 instances are read; each instance has data of 32 EEG channels each having length 8064 corresponding to 63 seconds of data recording. The first 3 seconds of data was removed, and the last 60 seconds were kept. The data was already preprocessed to remove artifacts and down sampled to 128 Hz.

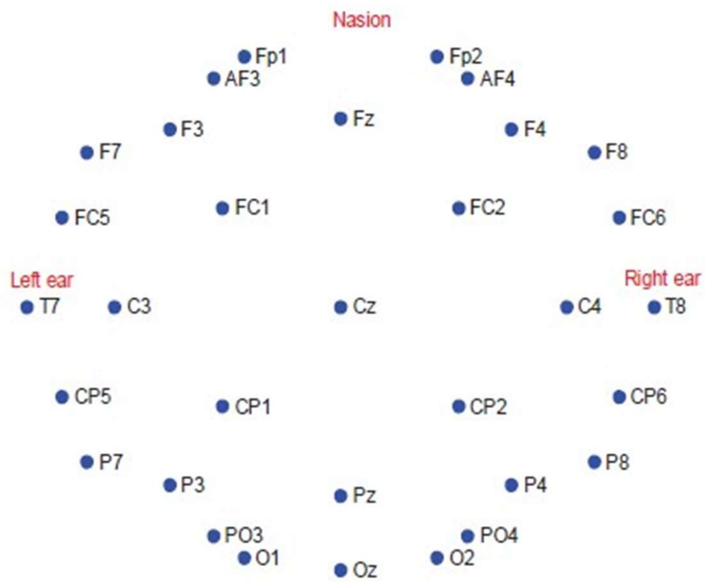

Figure 6: The location of each EEG electrode as described in [4]

Now using the EEG placement location described in [4] symmetrically opposite located pair of EEGs were found. These EEGs' data was subtracted to obtain 14 time series per sample from original 32.
These samples were then standardized to 0 to 1 range and then normalized to remove mean and make standard deviation 1.

After this minimal preprocessing, the EEG channels were split into training, validation and testing sets with 864, 288 and 128 samples each. A 1D CNN, whose architecture is already described, was then used to learn from this data and do classification. Figure 7 describes the algorithm of the deep learning model.

```python
def create_model(metrics = METRICS):
  model = Sequential()
  model.add(Conv1D(filters = 128, kernel_size = 64, activation = 'relu', input_shape=(7680,14)))
  model.add(MaxPooling1D(pool_size = 2))
  model.add(Conv1D(filters = 64, kernel_size = 16, activation = 'relu'))
  model.add(Conv1D(filters = 16, kernel_size = 4, activation = 'relu'))
  model.add(Conv1D(filters = 4, kernel_size = 3, activation = 'relu'))
  model.add(MaxPooling1D(pool_size=2))
  model.add(Flatten())
  model.add(Dense(64, activation='relu'))
  model.add(Dropout(rate = 0.4, seed = RANDOM_SEED))
  #mutliclass classification
  model.add(Dense(4, activation = 'softmax'))
  model.compile(optimizer= tf.keras.optimizers.Adam(learning_rate=0.01) , \
                loss = tf.keras.losses.CategoricalCrossentropy(), metrics= metrics)

  print("CNN Model defined.")
  return model

```

Figure 7: Algorithm for Building the Proposed Model

The model was implemented using Keras API with Tensorflow Backend in Python language.

5. RESULTS AND DISCUSSION

DEAP [6] dataset was used in this project to classify emotions into 4 classes namely HV/HA, HV/LA, LV/HA, LV/LA in the valence-arousal model. 14 pairs of EEG channels were identified [4] and obtained from the input of 32 channels per trial. The total of 1280 samples were divided into training, validation and testing sets of size 864, 288 and 128, respectively.
The model described in Table 1 was utilized for the classification purposes, after hyperparameter tuning the parameters present in Table 2 were obtained and used.

Figures 8 and 9 represent the plots of accuracy and loss versus epochs. The plots indicate that after an accuracy of 75%, the model ceases to learn and the accuracy and loss stagnate. The study concludes that for the given type of deep learning model and the given dataset 75% is the highest accuracy obtainable using raw EEG signals. This accuracy is for 4 class classification problem which is significantly harder than binary classification. The baseline given by [6] using Feature Extraction methods over EEG signal averages to about 58.3% classification accuracy for binary classification, the proposed method significantly exceeds this.

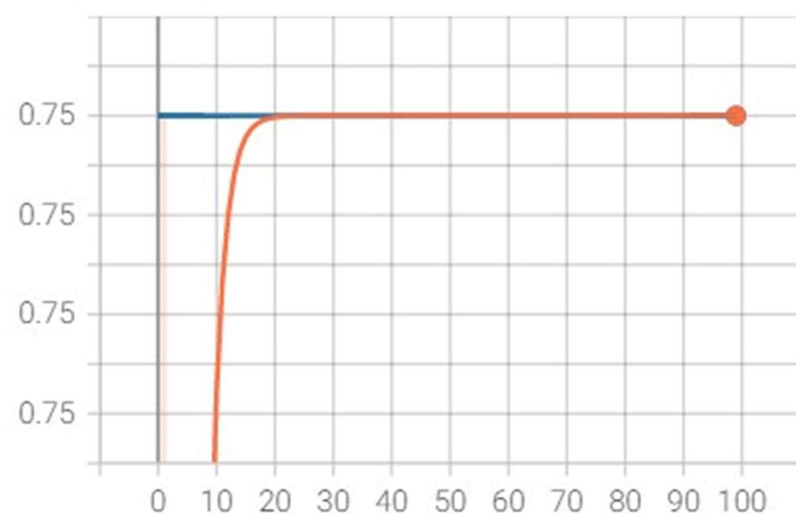

Figure 8: Smoothed Plot of Accuracy v/s Epochs, Validation measures in Blue, training in orange

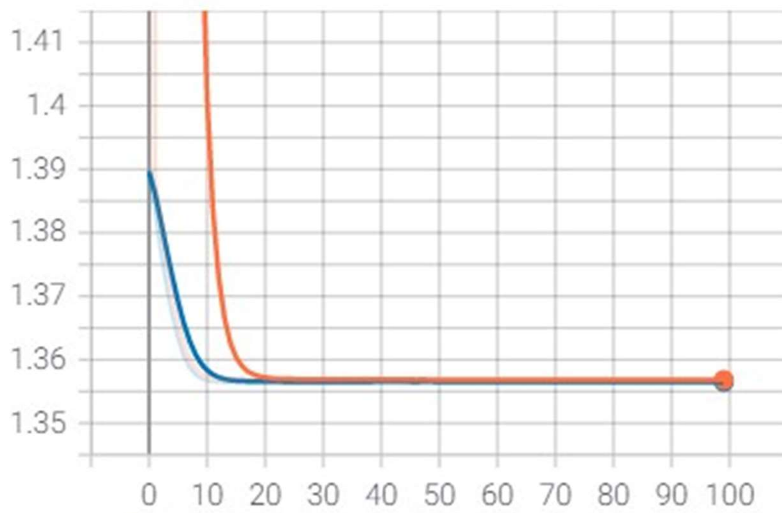

Figure 9: Smoothed Plot of Loss v/s Epochs, Validation Measures in Blue, training in orange

Model	Features	Accuracy
S. Koelstra et al. [6]	Power Spectrum Density	58.3% (2-Class)
Huang et al. [13]	Asymmetry Spatial Pattern	74.25% (2-Class)
Chung and Yoon [14]	Power Spectral Features	66.5% (2-Class)
Abeer et al. [15]	PSD & Pre-Frontal Asymmetry	82% (2-Class)
Proposed Method	1D-CNN	75.0% (4-Class)

Table 3: The accuracy comparison of the proposed method

Most of the research done on the DEAP Dataset for emotional classification focuses on using complex feature engineering methods like Discrete Wavelet Transform, PCA, Power Spectral Density of EEG signals etc. along with models like Naïve Bayes, Support Vector Machines and KNN classifiers. The proposed method does not use any complicated feature engineering techniques. 1D-CNN is directly applied over the raw EEG signals. Minimal Pre-Processing was done, on the first 60 seconds of the EEG signals, standardization to 0-1 range was done and then normalization was performed to remove mean and make standard deviation 1. This model on the raw EEG signal performs with good accuracy for multi-class classification while the comparable studies only get around 85% accuracies and that too on binary classification.

There seems to be a need to do feature augmentation to increase the number of samples available to train the model on. Further different architectures of model can be also more tuned to gain better accuracy.

6. SYSTEM REQUIREMENTS

Google Colab was used to do the project. The dataset was uploaded to Google Drive and Google Colab's GPU was used to train the models online. The experimental setup can be duplicated by using the following software and hardware requirements.

6.1) Software Requirement

Language Used-

1. Python

Modules or Libraries Used:

1. Numpy
2. Pandas
3. Tensorflow
4. Keras
5. Scikit-Learn
6. Pickle
7. Scipy
8. Matplotlibd
9. Seaborn
10. Imblearn
11. GC
12. Collections
13. Datetime

6.2) Hardware Requirement

- CPU: Hyper-Threaded Intel(R) Xeon(R) CPU @ 2.20GHz
- GPU: 1xTesla K80, compute 3.7, having 2496 CUDA cores, 12GB GDDR5 VRAM
- RAM: 13GB
- Hard Disk Space of 50 GB

7. CONCLUSION

The automated classification and recognition of EEG signal is a tough task in the field of machine learning. The project presents an end-to-end deep learning neural network model for the classification of the emotion experienced by a person by a musical video clip into 4 classes. A 1D-CNN was used for this task, it learns features from raw EEG signals. The proposed method achieves a weighted accuracy of 75% on the test data compared to an average binary classification accuracy of 58.3% in [6]. The model illustrates the power of Deep Learning models to learn features from raw data without any complicated feature extraction presented in traditional techniques.

8. REFERENCES

1. Yıldırım, Ö., Baloglu, U.B. & Acharya, U.R. A deep convolutional neural network model for automated identification of abnormal EEG signals. Neural Comput & Applic 32, 15857–15868 (2020). https://doi.org/10.1007/s00521-018-3889-z

2. Salma Alhagry, Aly Aly Fahmy and Reda A. El-Khoribi, “Emotion Recognition based on EEG using LSTM Recurrent Neural Network” International Journal of Advanced Computer Science and Applications(IJACSA), 8(10), 2017. http://dx.doi.org/10.14569/IJACSA.2017.081046

3. Emad-ul-Haq Qazi, Muhammad Hussain, Hatim AboAlsamh, & Ihsan Ullah. (2019). Automatic Emotion Recognition (AER) System based on Two-Level Ensemble of Lightweight Deep CNN Models. https://arxiv.org/abs/1904.13234

4. M. K. Ahirwal, A. Kumar, N. D. Londhe and H. Bikrol, "Scalp connectivity networks for analysis of EEG signal during emotional stimulation," 2016 International Conference on Communication and Signal Processing (ICCSP), Melmaruvathur, India, 2016, pp. 0592-0596, doi: 10.1109/ICCSP.2016.7754208.

5. Samarth Tripathi, Shrinivas Acharya, Ranti Dev Sharma, Sudhanshi Mittal, and Samit Bhattacharya. 2017. Using deep and convolutional neural networks for accurate emotion classification on DEAP dataset. In Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI'17). AAAI Press, 4746–4752.

6. "DEAP: A Database for Emotion Analysis using Physiological Signals (PDF)", S. Koelstra, C. Muehl, M. Soleymani, J.-S. Lee, A. Yazdani, T. Ebrahimi, T. Pun, A. Nijholt, I. Patras, IEEE Transactions on Affective Computing, vol. 3, no. 1, pp. 18-31, 2012

7. C. Qing, R. Qiao, X. Xu and Y. Cheng, "Interpretable Emotion Recognition Using EEG Signals," in IEEE Access, vol. 7, pp. 94160-94170, 2019, doi: 10.1109/ACCESS.2019.2 928691.
8. Doma, V., Pirouz, M. A comparative analysis of machine learning methods for emotion recognition using EEG and peripheral physiological signals. J Big Data 7, 18 (2020). https://doi.org/10.1186/s40537-020-00289-7

9. O. Bazgir, Z. Mohammadi and S. A. H. Habibi, "Emotion Recognition with Machine Learning Using EEG Signals," 2018 25th National and 3rd International Iranian Conference on Biomedical Engineering (ICBME), Qom, Iran, 2018, pp. 1-5, doi: 10.1109/ICBME.2018.8703559.

10. Pion-Tonachini L, Kreutz-Delgado K, Makeig S. Iclabel: an automated electroencephalographic independent component classifier, dataset, and website. NeuroImage. 2019; 198:181–97.

11. Asghar, Muhammad Adeel et al. “EEG-Based Multi-Modal Emotion Recognition using Bag of Deep Features: An Optimal Feature Selection Approach.” Sensors (Basel, Switzerland) vol. 19,23 5218. 28 Nov. 2019, doi:10.3390/s19235218

12. Pallavi Pandey, K.R. Seeja, Subject independent emotion recognition from EEG using VMD and deep learning, Journal of King Saud University - Computer and Information Sciences, 2019, ISSN 1319-1578, https://doi.org/10.1016/j.jksuci.2019.11.003

13. D. Huang, C. Guan, K. K. Ang, H. Zhang, and Y. Pan, “Asymmetric spatial pattern for EEG-based emotion detection,” in Proceeding of the International Joint Conference on Neural Networks (IJCNN '12), pp. 1–7, Brisbane, Australia, June 2012

14. Chung, Seong Youb, and Hyun Joong Yoon. ”Affective classification using Bayesian classifier and supervised learning.” Control, Automation and Systems (ICCAS), 2012 12th International Conference on. IEEE, (2012)

15. Abeer Al-Nafjan, Areej Al-Wabil, Manar Hosny, Yousef Al-Ohali, “Classification of Human Emotions from Electroencephalogram (EEG) Signal using Deep Neural Network”, IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 8, No. 9, 2017