

Think2Type: Thoughts to Text using EEG Waves

Aditya Srivastava

Department of Computer Engineering
Vivekanand Education Society Institute of Technology
Mumbai

Tanvi Shinde

Department of Computer Engineering
Vivekanand Education Society Institute of Technology
Mumbai

Sameer Ahmed Ansari

Department of Computer Engineering
Vivekanand Education Society Institute of Technology
Mumbai

Mr. Prashant Kanade

Department of Computer Engineering
Vivekanand Education Society Institute of Technology
Mumbai

Prateek Mehta

Department of Computer Engineering
Vivekanand Education Society Institute of Technology
Mumbai

Abstract—Visually impaired and differently abled people who can't use their hands have to depend on a helper when they wish to use a computer or a smartphone and they wish to enter any data. Depending on someone also causes privacy and security threats as all information including confidential information like passwords, etc also have to be shared so that the helper can write the data into the device. Think2Type is a BCI solution which will translate a user's active intent to text represented by morse code. BCI or 'Brain Computer Interface' is a technology that enables its users to interact with computers by means of brain-activity only, this activity being generally measured by Electroencephalography (EEG). Our solution helps them to be more independent as they merely have to think about the characters in the password and use active intent to enter them. The intents correspond to 'dot' and 'dash' of Morse code representation of alpha numerics. It will be converted to the corresponding text and provide them with complete confidentiality.

Keywords—Brain Computer Interface; Electroencephalography; Deep Learning; Python Development; Solution for visually impaired;

I. INTRODUCTION

Thoughts to Text using BCI (Brain-Computer Interface) is a relatively new and emerging field. The electroencephalogram (EEG) is the depiction of the electrical activity occurring at the surface of the brain. This activity appears as EEG waveforms of varying frequency and amplitude measured in voltage (specifically microvoltages).

When a person thinks about something, voltage fluctuations are generated, resulting from ionic current within the neurons of the brain. These EEG signals typically occur in microvolts and can be measured from various spots on the head, in a non-invasive manner.

The difficult question is to differentiate between the "brain activity" of different alphabets as our brain perceives words as phonetics and not alphabetically. To counter that problem we are using morse code to digitize the data. So now instead of thinking about an "a", the user needs to think of the morse code of 'a'.

We take input in the form of EEG signals, denoise it both in the hardware and the software.

The processed signal is analysed using a deep learning model which is explained further. This model understands the latent features of these signals and classifies them as a series of zeros and ones. This series is Morse code representation of the character which the person wishes to enter. Thus alphanumeric text is obtained from EEG signals after decoding this morse code.

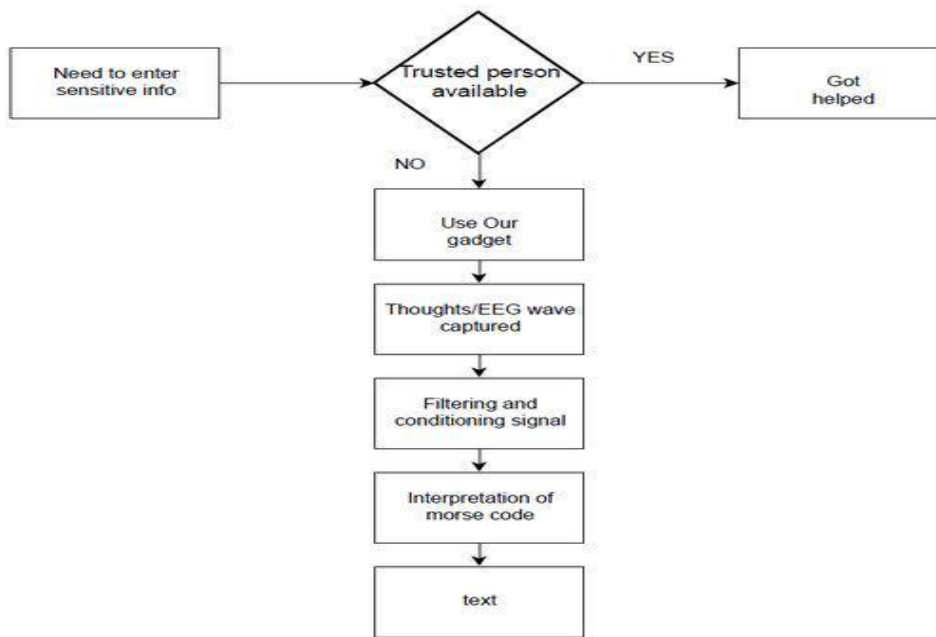


Fig. 1. Block Diagram of the system

Uniqueness:

Brain wave activities of alphabet characters are very similar and recognition and classification is a very difficult task. Our Idea is quite unique as there were no papers / patents / articles which had a direct implementation of our idea for this problem.

B. Background:

According to W.H.O, there are around 39 million blind and 285 million visually impaired people [4]. A majority of them are from India, That is India currently has around 12 million blind people which makes India home to one-third of the world's blind population. When such people want to use a computer or a smartphone, they have to rely on their caretakers to enter all data which is required by the applications. It is more problematic when it is sensitive information like passwords and bank details as issues of trust and security come into play.

Our solution is a Brain-Computer Interface(BCI) which will read the subject's active intent in the form of EEG signals and then convert it into alphanumeric data which will be entered in the required text fields . This will be used by differently-abled people to enter sensitive information in devices.

C. Connection:

In order to input the active intent of the user, he will think about moving his left arm for zero and right arm for one and both the arms for '/' or space. Thus he will

input the morse code[6] of the intended text. This will be captured, using electrodes and thus obtained by the hardware. Denoising filters will be applied to it along with analog to digital converters. This will be passed to the Deep Learning model. Latent features will be extracted by the model which will be used to classify the signals into a series of ones, zeros and slashes or spaces. This will be the inputted morse code.The last step will be decoding the morse code into alphanumeric text and sending it to the intended place.

D. Goal of the work:

The main aim of our product is to help them become more independent, enabling them to live a fuller and more secure life. By using our device they can keep their confidential details to themselves without the risk of anyone misusing them. Though this solution is designed for people with impairments, it can be used by anyone who wishes for confidentiality of their data.

II. LITERATURE SURVEY

As stated in papers [1] , [2], [3] , [5] and [6] , Deep Feature engineering is a proven method for processing brain waves as many researchers have tried and tested neural network approach to classification of EEG waves.

And further studies in [1] , [3] , and [4] show that Motor Imagery produces brain waves which can be generalized and hence a deep learning model can be trained upon the signals corresponding to imagined motor movements.

Also due to the miniscule differences and diversity in the thoughts of different subjects, EEG waves of alphabets, Brain wave activities corresponding to alphabet characters cannot be recognized or classified from signals obtained by non-invasive means with current progress in the field.

III. IMPLEMENTATION STRATEGY:

The modular diagram describes the journey of the thoughts of the user in the form of user intent to characters which will appear in the input field. The EEG waves captured from the user via electrodes will be sent to the hardware kit where the signal will undergo analog to digital conversion and

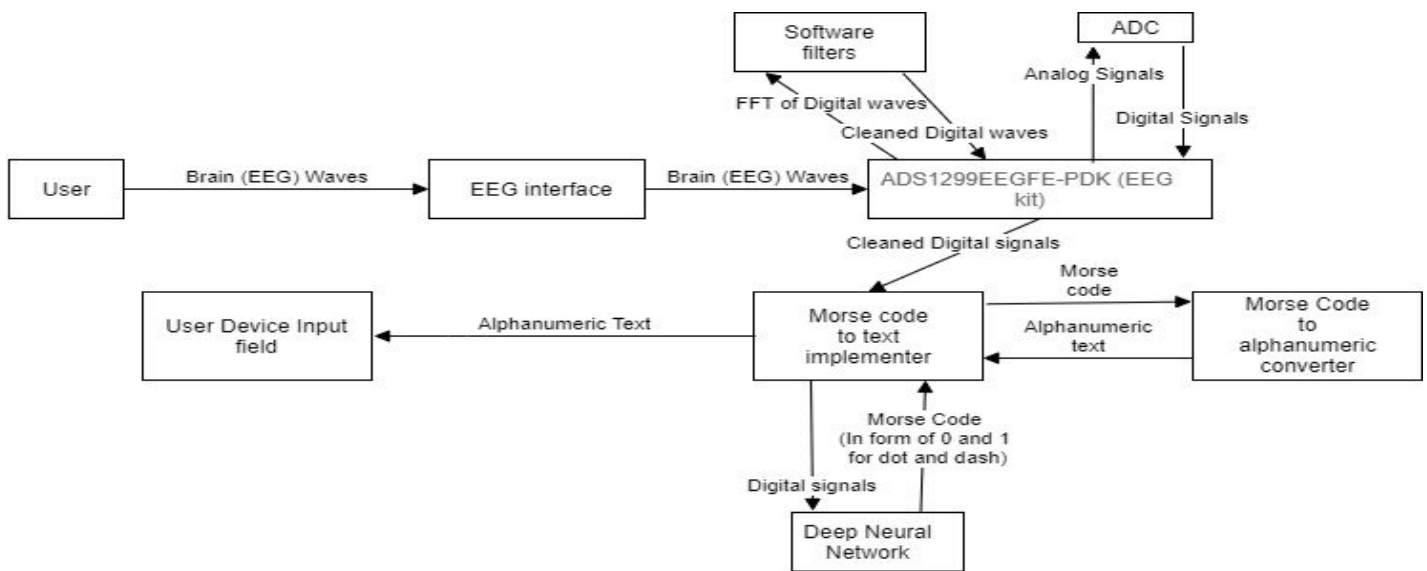


Fig 2. Modular Diagram

software filters will be applied to it as well in the form of FFT transformation. A fast Fourier transform (FFT) is an algorithm that computes the discrete Fourier transform (DFT) of a sequence, or its inverse (IDFT). Fourier analysis converts a signal from its original domain (often time or space) to a representation in the frequency domain and vice versa. This cleaned digital signal is further sent as an input to the Ensemble Deep Learning model which extracts important features from it and classifies the signals as a series of three characters. 0, 1 or '/'. This is further passed on to a morse code to text converter, and the output of which is passed on to the input field.

Input Data:

EEG waves will be collected from the subjects using the hardware. This includes the process of capturing the intent of the individual in terms of morse code. A band will be worn by the end user. We are using a velcro band which will be worn around the

forehead. The user will think about the input text in the form of its morse code. For Eg. if 'password' is the word the user wants to enter in the given field securely, he will have to think about the morse code of the word, and thus think '0110/01/000/000/011/111/010/100'. In order to enter a zero, the user will think about moving the left arm, for a one, right hand and both the hands together for a '/' or a space. Three electrodes will be attached to the band which will capture the EEG waves. Thus we will have three channel-data which will be forwarded to the hardware kit.

E. Deep Learning:

We designed a unified deep learning model. It uses - Convolutional Neural Network (CNN): is used to extract higher-level spatial variations of signal. Convolutional Neural Networks (CNN) is one of the variants of neural networks used heavily in the field of Computer Vision. It derives its name from the type of hidden layers it consists of. The hidden layers of a CNN typically consist of

convolutional layers, pooling layers, fully connected layers, and normalization layers.

Long Short-Term Memory (LSTM): is designed to model temporal sequences and long-range dependencies. They work tremendously well on a large variety of problems, and are now widely used. LSTMs are explicitly designed to avoid the long-term dependency problem. All recurrent neural networks have the form of a chain of repeating modules of neural networks

Artificial Neural Network (ANN): is used to extract features from the data provided by LSTM and CNN. An ANN is based on a collection of connected units or nodes called artificial neurons, which loosely model the neurons in a biological brain. Each connection, like the synapses in a biological brain, can transmit a signal to other neurons.

Extreme Gradient Boosting (XGBoost): is used for classification. XGBoost is an optimized distributed gradient boosting library designed to be highly efficient, flexible and portable. It implements machine learning algorithms under the Gradient Boosting framework.

The Dataset consists of EEG waves reading and the corresponding intent of the subject. A CNN is trained on the data and features of the second last layer are extracted from it. A LSTM RNN is trained on the data and features of the second last layer are extracted from it. The extracted features are combined and this is trained on an ANN and features of the second last layer are extracted. This is done to reduce the dimensions of the extracted features. The features extracted from ANN are trained using XGB which predicts the intent of the waves.

F. Integration:

The intent of the user will be captured via EEG waves which will be sent to the hardware. The software part of the project is a python desktop application with a deep learning model built using Keras and Tensorflow.

G. Dataset:

EEG data from PhysioNet eegmmidb (EEG motor movement/imagery database) database makes up the

major part of our dataset. This data set consists of over 1500 one- and two-minute EEG recordings, obtained from 109 volunteers. Subjects performed different motor/imagery tasks while 64-channel EEG were recorded using the BCI2000 system². Each subject performed 14 experimental runs: two one-minute baseline runs (one with eyes open, one with eyes closed), and three two-minute runs of each of the tasks described on their website¹.

The data was available in the EDF format. We created our own data set for training of the model which was a mixture of both, the one described above and other collected by us using the NeuroSky MindWave Mobile 2 kit³. We extracted and transformed data such that we finally have 3 labels to train our model on. Left wrist (open / close) indicating a 0. Right wrist (open / close) indicating a 1 and both wrists together corresponding to a '/' used as a delimiter.

IV. EXAMPLE SET

Case 1:

We input the eeg signals of a particular character to the function 'run' which runs through the entire pipeline and verifies if the model correctly interpreted its waveform and gave back the same result. In the given scenario, the EEG wave form of 'h' is inputted to the model and it returns the same letter after decoding it.

Case 2:

We input a text phrase to the function 'phrase' which converts that text to its morse code, and inputs the wave forms, corresponding to the characters, to the model. It then receives the morse code output and decodes it to alphanumeric text which is outputted. The phrase entered, must be the output phrase.

Result:

```
5]: 1 run()
Enter file name => h.csv
5]: 'h'
7]: 1 phrase review two
7]: 'review two'
1. gmmidb/
2. http://www.bci2000.org
https://store.neurosky.com/pages/m
3. indwave
```

Fig 3 Test Result
<https://archive.physionet.org/pn4/ee>

A. CALCULATING ACCURACY:

Accuracy for various stages in the model-

1. CNN accuracy ~ 97.7%
2. RNN accuracy ~ 89.8%
3. ANN accuracy ~ 97.6%
4. Final model(XGB) accuracy ~ 97.7%

We used the categorical cross entropy loss given by the following equation for training the CNN, RNN and ANN models.

$$L(y, \hat{y}) = - \sum_{j=0}^M \sum_{i=0}^N (y_{ij} * \log(\hat{y}_{ij}))$$

Fig 4. Categorical Loss

For the XGB classifier, since it's a multi-class classifier the softmax loss function shown below is used.

$$\text{softmax}(y)_i = \frac{\exp(y_i)}{\sum_j \exp(y_j)}$$

Fig 5. Softmax Equation

For the CNN, RNN and ANN keras's accuracy metric implementation was used and the final model(XGB) accuracy was calculated using sklearn's accuracy_score function.

B. LEARNING CURVE:

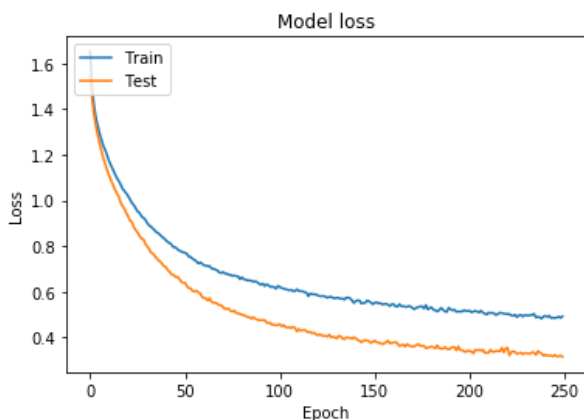


Fig 5 Model loss

C. ACCURACY GRAPH:

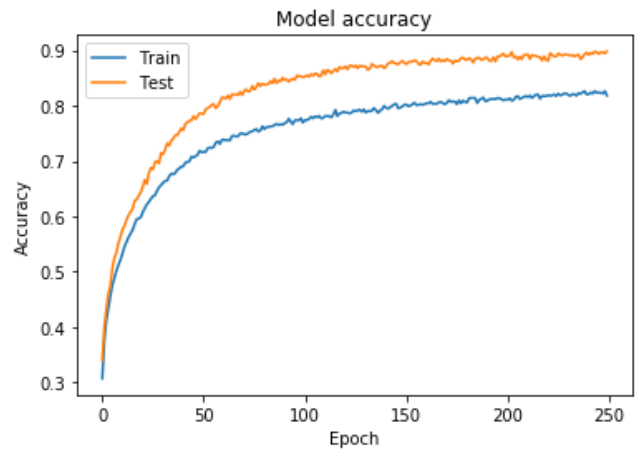


Fig 6 Model accuracy

V. INFERENCE DRAWN

The model captures the spatial features (CNN) with a very high accuracy of 97.7% and the temporal features are learnt by the RNN. Looking at the Model Loss (Fig.) we can infer that the learning rate is appropriate and there is no problem of either variance or bias. Hence we can say that the model neither overfits the training data nor underfits it. Classification of every class is done unbiased.

VI. CONCLUSION AND FUTURE SCOPE

For the general population to be able to use Think2Type, the cost effectiveness of the product can be improved. The device can be made more portable and hassle free to use, by making the design more compact. Wired connections can be replaced with a wireless connection like Bluetooth. The GUI of the software can be improved and it can be made easy to use for everyone.

With the little step of users, learning the morse code and adapting to the methodology used, the problem of entering sensitive data especially for visually impaired and other differently abled people and the fantasy of typing what you think can be achieved.

REFERENCES

- [1] Xiang Zhang, et. al. Converting Your Thoughts to Texts: Enabling Brain Typing via Deep Feature Learning of EEG Signals
- [2] Mohammad H. Alomari, et. al. Automated Classification of L/R Hand Movement EEG Signals using Advanced Feature Extraction and Machine Learning (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 4, No. 6, 2013
- [3] G. Pfurtscheller and C. Neuper, "Motor imagery and direct brain-computer communication," in Proceedings of the IEEE, vol. 89, no. 7, pp. 1123-1134, July 2001.
- [4] GLOBAL DATA ON VISUAL IMPAIRMENTS 2010 by World Health Organization (WHO).(<https://www.who.int/blindness/GLOBALDATAFINALforweb.pdf>)
- [5] International Morse Code Reference. (https://en.wikipedia.org/wiki/Morse_code).
- [6] Y. Ren and Y. Wu, "Convolutional deep belief networks for feature extraction of eeg signal," in Neural Networks (IJCNN), 2014 International Joint Conference on. IEEE, 2014, pp. 2850–2853.
- [7] J. Sleight, P. Pillai, and S. Mohan, "Classification of Executed and Imagined Motor Movement EEG Signals," Ann Arbor: University of Michigan, 2009
- [8] Y. R. Tabar and U. Halici, "A novel deep learning approach for classification of eeg motor imagery signals," Journal of neural engineering, vol. 14, no. 1, p. 016003, 2016.
- [9] B. Graimann, G. Pfurtscheller, and B. Allison, "Brain-Computer Interfaces: A Gentle Introduction," in Brain-Computer Interfaces: Springer Berlin Heidelberg, 2010, pp. 1-27.
- [10] Physionet Data set (<https://archive.physionet.org/pn4/eegmidb/>)