Empowering Speech-Impaired Individuals: EEG-Driven Cognitive Expression Translated into Speech

Jayavrinda Vrindavanam
Dayananda Sagar University
Bengaluru, India

Roshni M. Balakrishnan
Dayananda Sagar University
Bengaluru, India

Raghav Nanjappan
Dayananda Sagar University
Bengaluru, India

Gaurav Kamath
Dayananda Sagar University
Bengaluru, India

ABSTRACT

In the realm of communication, individualized treatment for persons with disabilities remains paramount. Roughly 5% of the population experiences communication impairments rooted in health conditions affecting speech, language comprehension, auditory processing, reading, writing, or social interaction skills. This spectrum encompasses lifelong instances seen in cerebral palsy, acquired aphasia, amyotrophic lateral sclerosis, and traumatic brain injuries. Although current technology adeptly translates neural activity into speech for those who have lost their innate vocal capabilities due to neurological illnesses or injuries, it does not address congenital speech disabilities.

Persons bearing communication disabilities often express being subjected to generalization. Thus, the imperative of supporting individuals with speech impairments emerges. At present, engineers have a distinctive opportunity to introduce innovative, cost-effective technological solutions to aid those with speech disabilities in effectively communicating with others. Electroencephalogram (EEG) signals, collected from the brain's scalp, play a pivotal role. These signals are commonly categorized based on their frequency, amplitude, and waveform characteristics.

This paper centers on a significant endeavor: enhancing the quality of life for individuals with speech impairments. The primary focus involves deciphering select cognitive expressions of speech-impaired individuals and translating them into speech. Accomplishing this objective necessitates the fusion of Electroencephalogram data with advanced machine learning algorithms, facilitating the accurate classification of intended thoughts within specified time frames.

General Terms

Thought Recognition

Keywords

Electroencephalogram, Signals, Support Vector Machine, K-Nearest Neighbors, Long Short-Term Memory

1. INTRODUCTION

Communication is the process of transferring information. It is the key to the existence and survival of humans. It helps to share and express ideas, information, views, facts, feelings and at the identical time, and help us to grasp the emotion and thoughts of others to succeed in a common understanding. People with speech disability have the necessity for support to complex communication needs, this is applicable whether the person encompasses a mobility impairment, a speech impairment, or a cognitive impairment. The existing technology serves the aim of converting brain activity to speech, for those who had lost their natural speech because of illness or injuries which affected the parts of the brain chargeable for speech. The prevailing technology isn't for those that have the matter of speech disability since birth. Famously, the late theoretical physicist Stephen Hawking was able to communicate after he was diagnosed with Amyotrophic Lateral Sclerosis (ALS) at the age of 21 by using a speech-generating device. Initially, he used a handheld clicker that enabled him to select words from a computer to speak, but later he used a speech-generating device that was controlled by his eye movements. The device was a computer that was controlled by his eye movements. Our work "Empowering Speech-Impaired Individuals: EEG-Driven Cognitive Expressions Translated into Speech" focuses on decoding the few thoughts of a person having a speech impairment to words by making use of Electroencephalography (EEG) with a machine learning algorithm to classify into their intended thought for the desired duration. This overcomes the gap and acts as a bridge of communication between normal people and folks having speech disabilities.

2. LITERATURE SURVEY

Keeping in view the objective of bringing together different techniques that can support surveillance under a machine-learning model, we observed that, the works in this direction are fairly spread out. Accordingly, several papers with a slightly different focus of the study were considered for this work. [1] demonstrated that the speech imagery EEG possesses significant discriminative information about the intended articulatory movements that are responsible for natural speech synthesis. [2] noticed a potential choice for restoration of discourse handicapped patients through cerebrum machine interfaces which utilize their dynamic comprehension capacities to convey the fanciful contemplations in expressed structure has been proposed. The results obtained indicate an average classification range of 93.1% to 94.74%. A study conducted by [3] has looked at the viability of syllable-level unit recognition of the temporal structure of speech as it is represented in the EEG data. [3] were able to prove that there is indeed a correlation between the EEG data and the syllable structure of speech. [4] propose using hybrid domain features in the EEG signals classification problem using Multiclass Support Vector Machines with New Kernel (MSVM). In comparison to the techniques that are currently in use, the new classification method has a higher classification accuracy and a lower computational complexity. [5] propose that the Brain Computer Interface (BCI) competition III data set be used for data collection. The subject was asked to perform an imaginary task, such as using their left small finger or tongue. The data in the set is mostly 3D data. [5] use artificial neural networks—three types of support vector machines (SVM) and K-Nearest Neighbors (KNN) as classifiers for EEG signals. The proposed system is accurate to 90.6%. [6] propose a" pattern recognition" method for distinguishing between EEG signals recorded under various cognitive conditions. The proposed approach used machine learning classifiers to produce much better classification results than the existing quantitative feature extraction techniques. [7] classified multiclass electroencephalogram (EEG) signals using an error-correcting output code and a multiclass support vector machine (SVM). The performance of the probabilistic neural network (PNN) and the multilayer perceptron neural network in classifying EEG signals was also evaluated and compared. The study demonstrated that the features that will represent the EEG signals are the wavelet coefficients and the Lyapunov exponents and that the multiclass SVM and PNN trained on these features achieved high classification accuracies. [8] have used two feature extraction algorithms for EEG signals. When these features are fed into support vector combination strategies, the classification performance can be effectively improved when the order of the autoregressive model is greater than 5, and the second strategy is superior to the first strategy in terms of classification accuracy for EEG signal classification. [9] The paper reviewed classification performance metrics for identifying epileptic episodes using EEG signals. Raw EEG signal-based computations and wavelet coefficient-based computations were effective in capturing seizures. A proposed system identifies specific thoughts and transmits relevant signals to individuals.

3. HUMAN BRAIN AND ELECTROENCEPHALOGRAM SIGNALS

3.1 Lobes of the Human Brain

The brain is the most vital functional organ of the human body. It is split into two hemispheres: namely the left brain and the right brain. The two hemispheres diverge into the Frontal, Temporal, Parietal, and Occipital lobes.

3.2 Electroencephalogram

EEG waveforms are categorized based on their frequency, their amplitude, and their shape. Analysis of frequency (Hertz, Hz) can help detect normal or abnormal rhythms. Frequency bands are used to categorize the brain's continuous rhythms, also known as brain waves. Distinct brain wave frequencies connect to different brain functions and mental states. There are 5 different frequency bands that exist during an EEG signal, and which incorporate the 5 waves: namely Delta (0.1 - 4 Hz), Theta (4 - 8 Hz), Alpha (8 - 13 Hz), Beta (13 - 30 Hz) and Gamma (30 - 100 Hz).

4. PROPOSED METHODOLOGY

We intend to employ a novel strategy to address all the issues outlined in the problem statement. Our method makes it possible for people with speech impairments to communicate with other people and lead an everyday life.

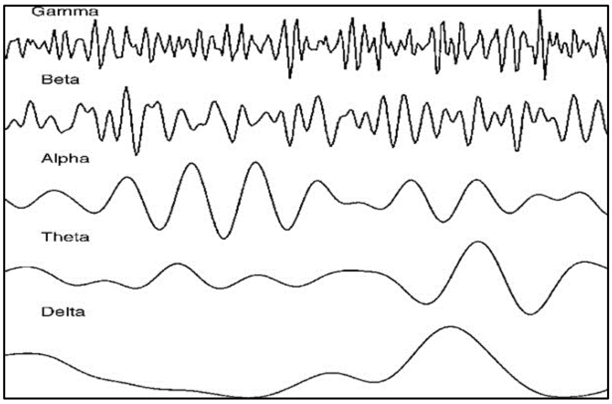

Fig 1: EEG Frequency bands

Our study utilizes the EEG signals for six different words, collected from various individuals. The signals were obtained with four electrodes. The feature extraction from the data is the next step. From the pre-processed data signal, nearly twenty features are extracted. Choosing the ML algorithm in accordance with the project's objective is critical. Using collected EEG signals, this model is trained from 10 people for 6 different words. This model is trained with the EEG data with the selected features. Later it includes testing the accuracy, specificity, sensitivity, and precision of the model. Then test the model for unseen data. The EEG signal of a particular person for a particular thought is loaded to train and test. The SVM classifier is used for the classification. Comparative analysis is done using the KNN algorithm. Evaluating the performance of the model for unseen data. The plot for the accuracy is plotted to know the comparative analysis. Then it is calculated to know which algorithm gives better results.

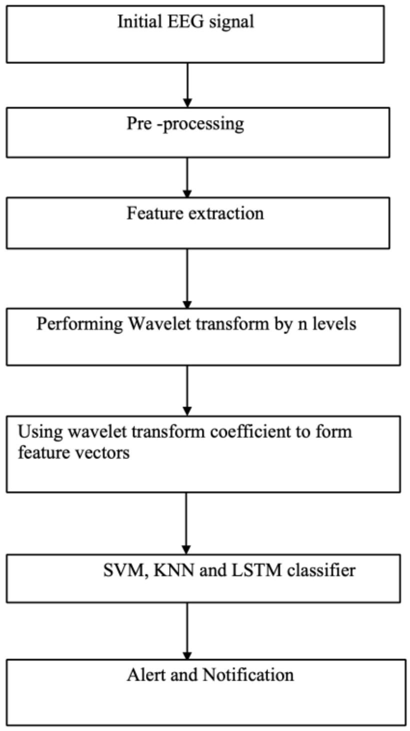

Fig 2: Block diagram of the proposed methodology

5. FEATURE EXTRACTION FROM ELECTROENCEPHALOGRAM

5.1 Introduction

Signals in the time domain usually only provide the information of signal amplitude and time information. A raw signal in the time domain needs to be processed to find out more information in that signal. Discrete wavelets transform (DWT) is a technique that provides enough information for both signal analysis and synthesis while reducing computation time significantly. The scaling function [$\Phi_{j,k}(n)$] characterizes the basic wavelet scale and allows expression of the needed details of the approximated function in the domain of interest, while the wavelet function [$\Psi_{j,k}(n)$] characterizes the basic wavelet shape and covers the entire domain of interest. These are denoted as follows:

$$\phi_{j,k}(n) = 2^{j/2}\,\phi(2^{jn} - k)$$

$$\Psi_{j,k}(n) = 2^{j/2}\,\Psi(2^{jn} - k)$$

Where n = 0, 1, 2, ..., M - 1; j = 0, 1, 2, ..., J - 1; k = 0, 1, 2, ..., 2j - 1; J = 5, and M is the length of the signal. Approximation coefficients and Detailed Coefficients are produced when this process is carried out repeatedly. These coefficients help obtain the value of the signal's frequency band power feature.

5.2 Details of the data

The session is comprised of 2 different phases the training and the test phase respectively. Each phase comprises 10 different individuals with 6 different thoughts. Each thought comprises 50 trials each with a duration of 3 seconds. Cumulatively, each comprises 9000 secs of EEG recorded. Therefore, the entire duration of the session is 18000 seconds.

6. ALGORITHMS

6.1 Support Vector Machine

For classification problems of two groups, a support vector machine (SVM) is implemented. SVM is a supervised machine learning model that uses classification algorithms. They can classify new text after providing an SVM model with sets of labeled training data for each category. They outperform more recent algorithms like neural networks in two main ways: greater efficiency and speed with a smaller number of samples. Thus, SVM is excellent for text classification problems in which a dataset consisting of no more than a few thousand tagged samples is typically available.

6.2 K-Nearest Neighbors

K-Nearest Neighbors (KNN) is known to be one of the simplest learning algorithms, which is based on the Supervised Learning technique. The KNN algorithm places the new case in the category that is most like the available categories and assumes that the new case is like the available cases. The KNN algorithm stores all the data that is available and uses similarity to classify a new data point. This indicates that the KNN algorithm can easily classify new data into a well-suited category when it appears.

6.3 Long Short Term Memory

LSTM is a type of RNN that excels at learning from sequences like time-series, text, and speech. It can capture long-term dependencies by selectively retaining or forgetting information using three gates. LSTM is useful in applications where input data has a temporal structure and requires context. Making it a good choice for EEG signal analysis.

7. RESULTS AND DISCUSSION

7.1 Result and Accuracy of Prediction

The model is trained to identify 6 different thoughts of speech-disabled people: namely –

- Urge to use the restroom.
- Assistance required.
- Hungry.
- Thirsty
- Hurting
- Sleeping

Model	Accuracy	Sensitivity	Specificity	Precision
SVM	97.4%	96.5%	97.2%	96.3%
KNN	72.8%	72.5%	73.1%	72.1%
LSTM	98.83%	99%	98%	99%

Fig 3: Accuracy of the models

7.2 Summary of Results

A comparative analysis of SVM and KNN algorithms is shown in Fig 3. The following 5 parameters are used as a basis for comparison between the 2 algorithms:

a) Accuracy: Accuracy is the proportion of forecasts that our model successfully anticipated. When the accuracy is 1, it means that the instances are predicted correctly. [11]

$$Accuracy = \frac{T_p + T_n}{T_p + T_n + F_p + F_n}$$

where T_p = True Positive, T_n = True Negative, F_p = False Positive and F_n = False Negative.

b) Sensitivity: Sensitivity is the proportion of actual positives that are correctly identified as such. It is also known as the true positive rate (TPR) or recall. A high sensitivity means the model is correctly identifying most of the positive findings, whereas a low sensitivity means the model is significantly underreporting the positive findings. [11]

$$Sensitivity = \frac{T_p}{T_p + F_n}$$

c) Specificity: Specificity is the proportion of actual negatives that are correctly identified as such. It is also known as the true negative rate (TNR). A high specificity means the model is correctly identifying most of the negative findings, whereas a low specificity means the model is significantly underreporting the negative findings. [11]

$$Specificity = \frac{T_n}{T_n + F_n}$$

d) Precision: Precision is the proportion of positive identifications that were correct. A precision score is used to assess the model's accuracy in accurately counting the number of genuine positives out of all positive predictions generated. [11]

$$Precision = \frac{T_p}{T_p + F_p}$$

8. CONCLUSION

We proposed a system to identify the thoughts of speech-disabled people using EEG signals. SVM and KNN models were applied for feature analysis, with SVM achieving 97.4% prediction and 96.3% accuracy, and KNN achieving 72.8% prediction and 72.1% accuracy. Additionally, an LSTM model was used, achieving 98.83% accuracy. The system can assist speech-impaired individuals in their daily life by identifying their thoughts.

8.1 Scope for Future Enhancement

The study brings to focus that it would be difficult to arrive at a generalized conclusion by suggesting a generalized model. The model must be specific for every speech impaired person. And in future we can also find the correlation between the EEG signals of different people for the same thought.

9. ACKNOWLEDGMENTS

We gratefully acknowledge Dayananda Sagar University for the constant encouragement and support for nurturing the much-needed research environment. We thank all the management, staff, and students at Dayananda Sagar University. We would like to thank the volunteers who helped us with data acquisition, without which this project wouldn't have been possible. Lastly, we would like to thank our family and friends for their continued support.

10. REFERENCES

[1] P. Saha, S. Fels and M. Abdul-Mageed, "Deep Learning the EEG Manifold for Phonological Categorization from Active Thoughts," ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019, pp. 2762-2766, doi: 10.1109/ICASSP.2019.8682330.

[2] M. P. Paulraj, S. B. Yaacob, C. R. Hema, A. H. Adom and S. K. Nataraj, "Voiceless voice: An intelligent communication system for physically retarded communities," 2010 IEEE International Conference on Computational Intelligence and Computing Research, 2010, pp. 1-4, doi: 10.1109/ICCIC.2010.5705789.

[3] R. A. Sharon, S. Narayanan, M. Sur and H. A. Murthy, "An Empirical Study of Speech Processing in the Brain by Analyzing the Temporal Syllable Structure in Speech-input Induced EEG," ICASSP 2019 - 2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2019, pp. 4090-4094, doi: 10.1109/ICASSP.2019.8683572.

[4] A. S. M. Murugavel, and S. Ramakrishnan, "Hierarchical multi-class SVM with ELM kernel for epileptic EEG signal classification" Forman, G. 2003. An extensive empirical study of feature selection metrics for text classification. J. Mach. Learn. Res. 3 (Mar. 2003), 1289-1305.

[5] M. Kanimozhi and R. Roselin, "Statistical Feature Extraction and Classification using Machine Learning Techniques in Brain-Computer Interface", International Journal of Innovative Technology and Exploring Engineering (IJITEE), ISSN: 2278-3075, Volume-9 Issue-3, January 2020.

[6] H. U. Amin, W. Mumtaz, A. R. Subhani, M. N. M. Saad, and A. Malik, "Classification of EEG Signals Based on Pattern Recognition Approach", Frontiers in Computational Neuroscience

[7] I. Guler and E. D. Ubeyli, "Multiclass Support Vector Machines for EEG-Signals Classification," in IEEE Transactions on Information Technology in Biomedicine, vol. 11, no. 2, pp. 117-126, March 2007, doi: 10.1109/TITB.2006.879600

[8] Y. Zhang, X. Ji and S. Zhang, "An approach to EEG-based emotion recognition using combined feature extraction method", Neuroscience Letters, vol. 633, 2019, pp. 152-157, ISSN 0304-3940, doi: 10.1016/j.neulet.2016.09.037.

[9] P. Boonyakitanont, A. Lek-uthai, K. Chomtho, and J. Songsiri, "A review of feature extraction and performance evaluation in epileptic seizure detection using EEG", Biomedical Signal Processing and Control, vol. 57, 2020, ISSN 1746-8094, doi: 10.1016/j.bspc.2019.101702

[10] S. Kabiraj et al., "Breast Cancer Risk Prediction using XGBoost and Random Forest Algorithm," 2020 11th International Conference on Computing, Communication and Networking Technologies (ICCCNT), 2020, pp. 1-4, doi: 10.1109/ICCCNT49239.2020.9225451.