IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁒ Peer-reviewed & Refereed journal ⁒ Vol. 10, Issue 12, December 2023

DOI: 10.17148/IARJSET.2023.101204

A Survey of Artificial Intelligence (AI) and Brain Computer Interface Techniques (BCI) for Translating Brain Signals into Text

Dr Vijayalaxmi Mekali[1], Anusha Phaniraj[2], Kartik Bhatt[3], Sahithi Bhashyam[4], Vipul Kant Tripathi[5]

Professor and Head, Dept of Artificial Intelligence and Machine Learning, K S Institute of Technology, Bengaluru, Karnataka[1]

Dept of Artificial Intelligence and Machine Learning, K S Institute of Technology, Bengaluru, Karnataka[2-5]

Abstract: This survey paper explores the intersection of Artificial Intelligence (AI) and Brain-Computer Interface (BCI) technologies, with a specific focus on recent advancements in thought-to-text conversion systems. These systems, powered by Electroencephalography (EEG) signals, represent a groundbreaking leap in assistive technologies, particularly for individuals with speech impediments. By decoding and interpreting imagined speech, AI-BCI systems empower users facing profound disabilities, offering a voice through silent communication. The technical intricacies involve diverse datasets and a combination of signal processing, machine learning, and deep learning techniques. Beyond understanding the technical aspects, the paper emphasizes the importance of evaluating the functional efficacy of these systems. The ultimate success lies in the positive impact on users' lives, as AI-BCI systems redefine communication paradigms and break down barriers for individuals with unique challenges. In enabling silent communication, these technologies aim to create a world where everyone can freely express themselves.

Keywords: Brain-Computer Interface (BCI), Electroencephalography (EEG), Thought-to-text conversion, Speech impairment, Assistive technology, Signal processing, Machine learning, Deep learning, Natural Language Processing (NLP)

I. INTRODUCTION

For millions worldwide, spoken language forms the bridge to social connection, professional opportunities, and personal fulfillment. Yet, for those with speech impairments, this bridge can crumble, leaving them isolated, marginalized, and facing diminished quality of life. Enter AI-powered thought-to-text conversion systems – a transformative technology poised to revolutionize communication for this population.

These systems, a synergistic blend of AI and BCI technologies, represent a paradigm shift. By deciphering the intricate patterns of brain activity associated with imagined speech, they allow individuals to directly translate their thoughts into text, bypassing the limitations of traditional vocalization.

Pioneering work like Think2Type's [15] groundbreaking EEG-to-text translation laid the foundation. Subsequent advancements, including EmoWrite's sentiment-aware approach [3] and Silent EEG-Speech Recognition's impressive word classification accuracy [4], showcase the field's rapid progress. Deep learning techniques [19] and robust EEG pattern understanding research [22] further refine these systems, pushing the boundaries of accuracy and efficacy.

This survey delves deeper than ever into these systems' intricate workings, dissecting the diverse methodologies employed. We explore the vital role of varied datasets and cutting-edge signal processing techniques (investigated by Arteiro et al. [16] and Vorontsova et al. [8]) in extracting meaningful information from complex EEG patterns.

The deployment of sophisticated AI algorithms, pioneered by Ravi et al. [23] and Nieto et al. [10], plays a crucial role in deciphering these patterns and translating them into coherent text. Additionally, the analysis tackles the inherent technical challenges, including capturing subtle EEG signals, overcoming noise and computational limitations, and paving the way for continuous improvements in performance and usability (Saminu et al., 2021 [11]; Meng et al., 2021 [12]; Lee et al., 2021 [13]; Hamedi et al., 2020 [14]).

 This work is licensed under a Creative Commons Attribution 4.0 International License

DOI: 10.17148/IARJSET.2023.101204

Beyond technical specifications lies the crucial question: how effectively do these systems facilitate meaningful communication in the real world? Rigorous assessments of real-world performance metrics like accuracy, speed, and vocabulary size (Lee et al., 2020 [17]; Sarmiento et al., 2019 [18]) provide a clear understanding of their current capabilities and limitations. However, the impact of this technology transcends mere technical parameters.

These systems hold the immense potential to empower individuals with speech impairments to fully participate in the tapestry of society. They can unlock doors to previously inaccessible professional avenues, enrich educational opportunities, and foster meaningful social connections (Rezaei Tabar & Halici, 2017 [21]). By bridging the communication gap and empowering individuals to express themselves freely, these systems offer a beacon of hope for a more inclusive and equitable future.

As these systems continue to evolve and refine, drawing upon insights from referenced works like Wang and Jung's exploration of Independent Component Analysis in improving Brain-Computer Interfaces [24], they pave the way for a future where silence no longer signifies a barrier but rather an avenue for empowered expression and inclusivity. We stand on the precipice of a new era in communication, one where everyone has the voice to be heard and the potential to be realized.

II. RELATED WORK

[Think2Type: Thoughts to Text using EEG Waves]:

In the groundbreaking paper titled "Think2Type: Thoughts to Text using EEG Waves," researchers introduce an innovative application of Brain-Computer Interface (BCI) technology. This application is specifically designed to empower individuals with visual impairments to securely enter sensitive information into their devices. By leveraging the power of electroencephalography (EEG) to capture brain activity associated with thoughts, the study offers a solution that addresses the limitations of existing input methods and potential privacy concerns associated with external assistance.

The authors of the paper devote careful attention to reviewing prior studies in the field. They emphasize the challenges faced in differentiating brain activities for distinct alphabet characters and explore the use of Morse code as a means to digitize and reduce feature dimensions. This comprehensive literature review sets the stage for their novel approach, which focuses on utilizing a unified deep learning model to extract higher-level spatial variations in the EEG signal.

The study showcases the development of an innovative implementation strategy through meticulous experimentation. The dataset used for training and evaluation consists of over 1500 one- and two-minute EEG recordings collected from 109 volunteers. The recorded EEG waves undergo a meticulous process, including analog to digital conversion, denoising filters, FFT transformation, and subsequent classification using an Ensemble Deep Learning model. This classification, represented by Morse code, is then converted into alphanumeric text, providing individuals with a secure and independent means of data entry.

The paper reports a remarkable accuracy rate of 97.7% for the deep learning model utilizing a Convolutional Neural Network (CNN) to extract higher-level spatial variations in the EEG signal. This model successfully classifies the signals as a series of zeros and ones, representing Morse code. The authors of the study also emphasize the potential impact of this technology on individuals with visual impairments, highlighting its usability and relevance in improving their quality of life.The authors advocate for further research to refine the model and assess its broader practicality.

[Imagined Speech Classification with EEG Signals for Silent Communication: A Preliminary Investigation into Synthetic Telepathy]:

In this paper, the authors delve into the fascinating realm of using EEG signals to classify imagined speech, with the ultimate goal of enabling silent communication. The study delves into the inherent challenges associated with EEG signal classification for imagined speech, such as the low signal-to-noise ratio, the existence of artifacts, and the variability both between and within subjects.

To overcome these obstacles, the propose an extensive preprocessing approach that aims to reduce noise and eliminate artifacts from the EEG data. Expanding upon prior research in the classification of EEG signals related to imagined speech, the paper draws insights from experiments involving imagined vocalizations, mouthing of vowels, and the imagination of syllabic speech without associated muscle movements.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ※ Peer-reviewed & Refereed journal ※ Vol. 10, Issue 12, December 2023

DOI: 10.17148/IARJSET.2023.101204

For the study, they collected EEG datasets from volunteers who were given instructions to imagine speaking two different syllables, namely /ba/ and /ku/, at a specific rhythm. These datasets were recorded using a 128 Channel Sensor Net at a sampling rate of 1024 Hz, and comprised a total of 120 trials per syllable, per subject.

To extract informative features from the EEG signals, they employed various classification techniques. These techniques included asymmetry power ratios, autoregressive model coefficients, and the Common Spatial Patterns method. The reported results of the classification experiments encompass the average accuracy for each subject and the different combinations of subject data.

In an effort to identify undesirable signals, they employed the Hurst exponent and marked electrodes as "not useful" if their exponent value fell below a certain threshold. Trials with a high percentage of "useful" electrodes were retained for further analysis. The subsequent classification results, which are summarized in tables, indicate varying degrees of success across different subjects and combinations of subject data.

The paper makes significant contributions to the field of EEG-based imagined speech classification. They have addressed the challenges associated with this area, provided a detailed methodology, and showcased the potential for improvement through trial rejection based on quantitative measures of electrode quality. While they believe the study offers valuable insights, they also advocate for further investigations to refine features, explore noise reduction methods, and assess the possibility of generalization across a broader spectrum of subjects.

[EmoWrite: A Sentiment Analysis-Based Thought to Text Conversion]:

In the realm of Brain-Computer Interface (BCI) technology, a groundbreaking research paper titled "EmoWrite: A Sentiment Analysis-Based Thought to Text Conversion" has emerged. This paper focuses on addressing the communication barriers faced by individuals with speech impediments, particularly paralytic patients. By delving into the shortcomings of existing BCI systems, the authors introduce an innovative approach that combines sentiment analysis and thought-to-text conversion.

Individuals with speech impediments, such as paralytic patients, often face considerable difficulties expressing their thoughts and emotions. Traditional BCI systems, while offering some respite, still fall short in capturing the nuances of human sentiment. This is where "EmoWrite" comes into play, offering a ray of hope for those seeking effective communication solutions.

At the heart of "EmoWrite" lies sentiment analysis - a fascinating technique that deciphers emotions encoded in a person's thoughts. By leveraging this technology, the researchers behind "EmoWrite" have opened up new frontiers in understanding human sentiment. This breakthrough enables individuals with speech impairments to express not only the content of their thoughts but also the intricate tapestry of their feelings.

Imagine this as thoughts course through the mind of an individual with a speech impairment, "EmoWrite" captures and analyzes these thoughts in real time. Through a seamless integration of sentiment analysis, the system decodes the underlying emotions behind these thoughts. The resulting output is an authentic and emotionally rich stream of text using 14 channels.

The experimental facet of the paper involves assessing the integration of emotion-based prediction. Results underscore the practicality of this approach, showcasing its potential to significantly streamline thought-to-text conversion by minimizing typing time through emotion-related word predictions.

The research paper on "EmoWrite" signifies a promising step forward in the field of BCI technology. Its pioneering integration of sentiment analysis and thought-to-text conversion has the potential to transform the lives of individuals with speech impairments. As this technology evolves, we can envision a world where effective communication becomes a reality for all, providing greater inclusivity and empowerment.

The evaluation results of EmoWrite further substantiate its prowess. With an accuracy rate of 90.36%, a typing speed of 6.58 words per minute (WPM), and a commendable Information Transfer Rate (ITR) of 87.55 bits/min with commands, EmoWrite outshines conventional systems. The dynamic keyboard's contextualized appearance of characters emerges as a pivotal factor, significantly enhancing overall productivity.

© IARJSET This work is licensed under a Creative Commons Attribution 4.0 International License

DOI: 10.17148/IARJSET.2023.101204

[Silent EEG-Speech Recognition Using Convolutional and Recurrent Neural Network with 85% Accuracy of 9 Words Classification]:

The recent foray into silent speech recognition (SSR) holds immense promise for communication accessibility and human-computer interaction. However, a recent study using myography to decipher silent speech raises intriguing questions that deserve closer examination.

One key concern revolves around potential bias in the dataset. Focusing solely on a single individual could limit the generalizability of the findings. Imagine training a translator solely on Shakespeare – its proficiency in everyday conversations would be questionable. Similarly, this SSR system might struggle with diverse speech patterns beyond the single user's repertoire. Additionally, the possibility of the individual adapting to the classifier's quirks further muddies the water. Is the reported 72% accuracy a true reflection of the system's capabilities, or does it merely represent the person's clever manipulation of the technology? Addressing these issues through diverse subject recruitment and rigorous control conditions is essential for building robust and generalizable SSR technology.

The study's choice of myography – recording muscle activity – also sparks fascinating avenues for discussion. While EEG, measuring brain activity, is the more common approach, myography offers its own advantages. It may be less susceptible to external noise and potentially allows for real-time feedback in BCI applications. However, its dependence on subtle muscle movements poses challenges. Imagine whispering – the minimal muscle activity might fall below the system's detection threshold. Comparing the pros and cons of myography and EEG, and exploring ways to combine them, could unlock the full potential of SSR.

Another point of consideration is the dataset's limitations. The English-centric vocabulary confines the technology's reach, rendering it unusable for non-English speakers. Imagine a silent translator only conversant in Shakespearean prose – its usefulness in a globalized world would be severely hampered. Expanding datasets to encompass diverse languages and broadening the word range are crucial steps towards making SSR truly inclusive.

The study's focus on BCI development for communication aid offers a beacon of hope for those with impaired speech. The discovery of common brain patterns during silent speech suggests that a universal "language of thought" might exist, paving the way for BCI interfaces accessible to a wider range of individuals. However, challenges remain. Refining these interfaces for user-friendliness and ensuring their affordability will be crucial for real-world implementation.

[Converting Your Thoughts to Texts Enabling Brain Typing via Deep Feature Learning of EEG Signals]:

The research focuses on using EEG signals for brain typing, specifically in the context of motor imagery EEG (MI-EEG) for Brain-Computer Interface (BCI) applications. The proposed approach is a novel deep neural network-based learning framework that aims to decode raw MI-EEG signals and establish a relationship between the EEG data and brain activities. The model combines the benefits of Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) to effectively decode EEG signals. Additionally, an Autoencoder layer is employed to eliminate artifacts from the raw EEG signals.

The proposed approach is evaluated using a large-scale public MI-EEG dataset and a local dataset collected by the researchers. The results demonstrate that the proposed approach outperforms a series of baselines and competitive state-of-the-art methods, achieving a classification accuracy of 95.53%. The model's applicability is further demonstrated through the development of a practical BCI system for typing.

The research also discusses the challenges associated with EEG decoding, such as the variability of EEG signals between individuals and the presence of noise and artifacts. The proposed approach addresses these challenges and achieves high levels of accuracy and precision, making it suitable for practical applications. The study also highlights the potential for future work, including improving accuracy in person-independent scenarios and addressing limitations related to EEG headset attributes.

Two diverse datasets fuel this study, offering complementary strengths. The public eegmmidb, a treasure trove from PhysioNet, contains 280,000 meticulously annotated brain signals from 10 subjects performing various mental tasks. Each signal, meticulously captured by 64 EEG channels, provides rich insights. Alongside this robust resource, a smaller dataset of 241,920 samples emerges from the Emotiv Epoc+, a readily available commercial headset. This local dataset, collected from multiple subjects, showcases the model's adaptability to real-world conditions using accessible equipment. By harnessing both comprehensive and practical data sources, the study achieves both generalizability and relevance to everyday scenarios.

DOI: 10.17148/IARJSET.2023.101204

[Robust Understanding of EEG Patterns in Silent Speech]:

The document provides a comprehensive overview of the role of Electroencephalography (EEG) data and Principal Component Analysis (PCA) in the development of Brain-Computer Interface (BCI) systems. It emphasises the importance of EEG signals in non-invasive BCI systems and highlights the specific signals used, such as slow motor imagery potentials, P300 potentials, and steady-state visual evoked potentials (SSVEP). These signals are crucial for different functions of BCI systems, such as motor imagery and communication.

The dataset they use consists of EEG signals collected from 20 individuals (17 men and 3 women) at the National University of Colombia's Clinical Electrophysiology Laboratory. The signals were recorded using a neuro-headset with 21 electrodes, a ground, and a reference electrode, all placed on the left hemisphere Werincke's and Broca's area. The recordings were conducted with the individuals' eyes closed and under consistent noise and brightness levels.

It discusses the technical aspects of feature extraction and analysis methods used in BCI systems. It provides a detailed table of the most frequently used features in different analysis methods, including time domain, frequency domain, wavelet transform, and cepstral transform. This highlights the diverse range of features that can be extracted from EEG signals to capture relevant information for BCI applications.

Furthermore, it delves into the application of PCA for feature extraction in each channel of the EEG signal. It explains how PCA retains maximum variance and provides robust features that are resistant to noise, making it a valuable tool for processing EEG data in BCI systems. Additionally, it briefly mentions Robust Principal Component Analysis (RPCA) as a method to address issues with corrupted observations, indicating the consideration of robustness in feature extraction.

[Brain-Computer Interface and Silent Speech Recognition on Decentralised Messaging Applications]:

Online communication has become increasingly common in daily lives, thanks to technological advancements in instant messaging platforms. Efforts have been made to include disabled individuals in peer-to-peer communication, but those with hand or arm impairments face a lack of support in mainstream applications. Furthermore, current solutions relying on speech-to-text techniques fail to provide privacy when used in public. Concerns over the privacy and security of centralised systems have led to the traction of alternatives using blockchain technology.

In this inclusive project, an alternative approach to human-computer interaction for individuals with disabilities is presented. By utilising a combination of a brain-computer interface and a silent speech recognition system, efficient application navigation and text input are enabled without the use of hands or arms. The brain-computer interface allows users to interact with the platform using their thoughts, while the silent speech recognition system enables text input by capturing activity from articulatory muscles, eliminating the need to audibly speak. This innovative combination empowers individuals with hand or arm impairments to communicate seamlessly in their daily lives.

To ensure privacy and secure data exchange between peers, users of the application are part of a decentralised system. This system prioritises privacy and addresses concerns regarding centralised platforms.

The project utilises the EMG-UKA corpus dataset, developed at the Karlsruhe Institute of Technology. This dataset includes EMG and acoustic recordings, specifically designed for researching speech recognition based on EMG signals. The EMG-UKA Trial corpus, a subset of recordings from four speakers, provides 1 hour and 52 minutes of data. Recordings encompass audible, whispered, and silent speech modes, capturing variations in muscle activity intensity. EMG data was collected using a six-channel electrode setup, synchronised with acoustic data using marker signals.

[Continuous Silent Speech Recognition using EEG]:

This paper explores continuous silent speech recognition using electroencephalography (EEG) signals. The researchers implemented a connectionist temporal classification (CTC) automatic speech recognition (ASR) model to translate EEG signals recorded in parallel while subjects were reading English sentences in their minds without producing any voice to text. The results demonstrate the feasibility of using EEG signals for performing continuous silent speech recognition.

Continuous silent speech recognition technology can enable people with severe cognitive disabilities to use virtual assistants like Siri, Alexa, and Bixby by improving technology accessibility. It can also enable people with cognitive disabilities to communicate with other people, soldiers, and scientists to perform covert communication in sensitive working environments. Continuous silent speech recognition technology can introduce a new form of thought-based communication for able-bodied people.

© IARJSET This work is licensed under a Creative Commons Attribution 4.0 International License

International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.066 ⧖ Peer-reviewed & Refereed journal ⧖ Vol. 10, Issue 12, December 2023
DOI: 10.17148/IARJSET.2023.101204

Electronography (EEG) is a non-invasive way of measuring electrical activity of the human brain by placing EEG sensors on the scalp of the subject. EEG signals have high temporal resolution but poor spatial resolution. Electrocorticography (ECoG) is an invasive procedure requiring brain surgery to implant electrodes.

In this work, non-invasive EEG signals are used to decode thoughts or perform continuous silent speech recognition. Previous studies have demonstrated isolated and continuous speech recognition using EEG signals recorded in parallel while subjects were speaking out loud and listening to English utterances for a limited English vocabulary. However, these experiments did not involve subjects explicitly reading sentences in their minds, making it unclear whether the work described in this study studies the continuous silent speech recognition problem.

[Imagined Speech Classification using EEG]:

The research article discusses the use of Electroencephalogram (EEG) signals for classifying imagined speech, aiming to explore the potential for communication between individuals. EEG signals were recorded from 13 subjects as they imagined the English vowels 'a', 'e', 'i', 'o', and 'u' in response to visual stimuli. Preprocessing techniques were used to remove artifacts and noise from the recorded signals. Common features such as average power, mean, variance, and standard deviation were computed and classified using a bipolar neural network. The study achieved a maximum classification accuracy of 44%, indicating that EEG contains distinct information for classifying imagined speech across subjects.

The methodology involved subject preparation, stimulus presentation, and EEG recording. Preprocessing techniques were used to address artifacts and noise in the EEG signals. Feature extraction involved computing statistical parameters such as mean, variance, standard deviation, and average power. These features were computed for both relaxation and stimulation instances, and the difference values were used as the final features for classification. Classification was performed using a Back Propagation Neural Network with 4 input parameters and 5 output nodes representing the characters 'a', 'e', 'i', 'o', and 'u'. The results showed significant classification rates for some vowels, with an average classification rate of 44% across the subjects.

The features extracted from the preprocessed EEG signals include average power, mean, variance, and standard deviation. These features are computed for both relaxation and stimulation instances and are used for classification. The difference values between the relaxation and stimulation periods are also computed and used as the final features for classification. These features are crucial for differentiating the imagined speech vowels and are used in the classification process using a Back Propagation Neural Network.

The study concluded that EEG signals contain distinctive information for classifying imagined speech and demonstrated the potential for across-subject classification. The authors suggested that increasing relaxation time could improve classification accuracy. The research has implications for developing communication devices for paralyzed individuals and those unable to speak. Future work could involve expanding the set of characters and assessing the feasibility of decoding all major alphanumeric characters through EEG signals.

The article provides valuable insights into the potential use of EEG for imagined speech classification, with implications for assistive communication devices and the possibility of text input through imagined speech.

[Thinking out loud, an open-access EEG-based BCI dataset for inner speech recognition]:

The study "Thinking out loud, an open-access" introduces an EEG-based BCI dataset designed to facilitate inner speech recognition research. This dataset encompasses recordings from ten participants across various conditions, including inner speech, pronounced speech, and visualized conditions. The study's significance lies in its provision of an openly accessible resource, which has the potential to advance understanding of inner speech commands and related brain mechanisms within the BCI field.

The study's key findings encompass the acquisition of diverse EEG data from participants, the development of processing scripts, and the analysis of neural correlates associated with inner speech. By offering a comprehensive dataset for the development and assessment of inner speech recognition algorithms, the study addresses existing knowledge gaps and has the potential to advance understanding of automatic brain pattern detection related to inner speech.

The subject matter of the study revolves around the creation of an open-access EEG-based BCI dataset for inner speech recognition, aiming to address the scarcity of publicly available datasets in this area.

© IARJSET This work is licensed under a Creative Commons Attribution 4.0 International License

IARJSET

ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⁜ Peer-reviewed & Refereed journal ⁜ Vol. 10, Issue 12, December 2023

DOI: 10.17148/IARJSET.2023.101204

The study's objectives include EEG data acquisition under various speech conditions, script development, and neural correlates analysis. The evaluation involves the generation of a multiclass EEG dataset, which holds promise for advancing knowledge in the realm of inner speech recognition.

III. LITERATURE SURVEY

Sl. no	Paper	Dataset	Methodology	Remarks
1	Aditya Srivastava, Sameer Ahmed Ansari, Tanvi Shinde, Prashant Kanade and Prateek Mehta, "Think2Type: Thoughts to Text using EEG Waves", International Journal of Engineering Research & Technology (IJERT), 2020	EEG motor movement/imagery database available on PhysioNet	FFT transformation, Ensemble Deep Learning model.	However, further validation and usability testing are necessary for real-world application
2	K. Brigham and B. V. K. V. Kumar, "Imagined Speech Classification with EEG Signals for Silent Communication: A Preliminary Investigation into Synthetic Telepathy," 2010 4th International Conference on Bioinformatics and Biomedical Engineering 2010	EEG signals recorded from 7 volunteer subjects imagining the syllables /ba/ and /ku/	K - Nearest Neighbors (KNN) classification algorithm	Small sample size, the lack of diversity in the subject pool, and the limited number of imagined syllables
3	Shahid, Aisha, Imran Raza and Syed Asad Hussain. "EmoWrite: A Sentiment Analysis-Based Thought to Text Conversion." ArXiv, 2021	Four separate datasets were created, each containing words associated with a specific emotional class.	Recurrent Neural Networks (RNN)	It's vocabulary is constrained and it currently outputs single words. Moreover, its focus on sentiment analysis, while valuable, might not always capture the full nuance of thought.
4	Vorontsova D, Menshikov I, Zubov A, Orlov K, Rikunov P, Zvereva E, Flitman L, Lanikin A, Sokolova A, Markov S, Bernadotte A. Silent EEG-Speech Recognition Using Convolutional and Recurrent Neural Network with 85% Accuracy of 9 Words Classification. Sensors . 2021	270 subjects recorded 8 Russian commands and 1 pseudoword	Convolutional and Recurrent Neural Network	Expand the dataset to include a more diverse group of participants, russian words were used which limits the population, and the accuracy rate was 85% which can be further improved.
5	X. Zhang, L. Yao, Q. Z. Sheng, S. S. Kanhere, T. Gu and D. Zhang, "Converting Your Thoughts to Texts: Enabling Brain Typing via Deep Feature Learning of	Public MI-EEG Dataset and dataset collected in the lab	Joint Convolutional and Recurrent Neural Network, Autoencoder layer, Public MI-EEG Dataset	Further improvements can be made by focusing on person-independent approach where some individuals data is used for training

© IARJSET This work is licensed under a Creative Commons Attribution 4.0 International License

DOI: 10.17148/IARJSET.2023.101204

	EEG Signals," 2018			and others for testing.
6	P. Ghane, G. Hossain and A. Tovar, "Robust understanding of EEG patterns in silent speech," National Aerospace and Electronics Conference (NAECON), 2015	Consists of vowels for 20 subjects	Principal Component Analysis (PCA) for Feature Extraction, Dataset consists of vowels for 20 subjects,	The dataset is limited to only vowels, data of only 20 subjects were collected, and they used PCA which assumes that data is linear and can be considered as a major disadvantage.
7	Arteiro, L., Lourenço, F., Escudeiro, P., Ferreira, C. (2020). Brain-Computer Interaction and Silent Speech Recognition on Decentralized Messaging Applications. In: Stephanidis, C., Antona, M. (eds) HCI International, 2020	The EMG-UKA corpus, a collection of EMG data, was gathered using a setup with six channels.	EMG data as words using HMM or LSTM	The paper lacks a practical implementation of a messaging platform using BCI and SSR, and since it relies on a single dataset of EMG signals, it has a limited scope of diversity of use cases and data sources.
8	Vorontsova D, Menshikov I, Zubov A, Orlov K, Rikunov P, Zvereva E, Flitman L, Lanikin A, Sokolova A, Markov S, et al. Silent EEG-Speech Recognition Using Convolutional and Recurrent Neural Network with 85% Accuracy of 9 Words Classification. Sensors, 2021	EEG signals recorded from four male subjects in their early to mid-twenties. Three of the subjects were non-native English speakers, and one was a native English speaker. Each subject was asked to silently read 30 English sentences from the USC-TIMIT database while their EEG signals were recorded.	Connectionist Temporal Classification (CTC) Automatic Speech Recognition (ASR) model	The paper's decoding model suffers from small vocabulary size, subject-dependency, and lack of comparison with other methods for EEG-based silent speech recognition.
9	Ravi, Kamalakkannan & Rajkumar, R. & Raj, M.M. & Devi, S.S.. Imagined Speech Classification using EEG. Advances in Biomedical Science and Engineering. 2014	EEG signals were recorded from 13 volunteers, 10 male, and 3 female, with an average age of 21 years. The subjects were instructed to imagine the English vowels 'a', 'e', 'i', 'o', and 'u' in response to visual stimuli.	Back Propagation Neural Network.	Maximum classification accuracy of 44%, indicating room for improvement. The study's exclusive focus on classifying English vowels may limit the generalizability of the findings to a broader range of speech sounds..
10	Nieto, N., Peterson, V., Rufiner, H.L. et al. Thinking out loud, an open-access EEG-based BCI dataset for inner speech recognition, 2022	EEG recordings during inner speech tasks, including pronounced speech, inner speech, and visualized conditions	Surface electroencephalography system	Limited sample size, focus solely on Spanish speakers, and potential confounds due to mixing imagined and actual speech

IARJSET
ISSN (O) 2393-8021, ISSN (P) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.066 ∺ Peer-reviewed & Refereed journal ∺ Vol. 10, Issue 12, December 2023
DOI: 10.17148/IARJSET.2023.101204

IV. METHODOLOGY

Imagined speech recognition is the process of decoding the words or phrases that a person is thinking of using brain signals. This is a challenging and novel task that requires a combination of hardware and software techniques. In this project, we've a methodology for non-invasive imagined speech using electroencephalogram (EEG) signals. These signals are basically electrical recordings from the scalp. me walk you through the steps of our methodology:

Preprocessing: The process of real-time EEG signal acquisition from a multi-channel device, noise reduction using bandpass and notch filters, and mathematical reconstruction of signals for a clean aggregate signal.

Feature Extraction: The conversion of preprocessed EEG signals to text by extracting relevant features considering their spectral, temporal, spatial characteristics.

Signals Modeling: The use of machine learning or deep learning models to understand and interpret the extracted features from the EEG signals.

Feature Matching: The process of training and fine-tuning the model using smart optimization algorithms and hyperparameter tuning to match the features with the corresponding output.

Decisions: The integration of preprocessing and model inference steps onto a resource-constrained IoT device for practical application and real-time decision making.

EEG Signals Dataset: The creation of a dataset by designing tasks that elicit non-invasive imagined speech and capturing the corresponding EEG signals, including spoken speech EEG signals.

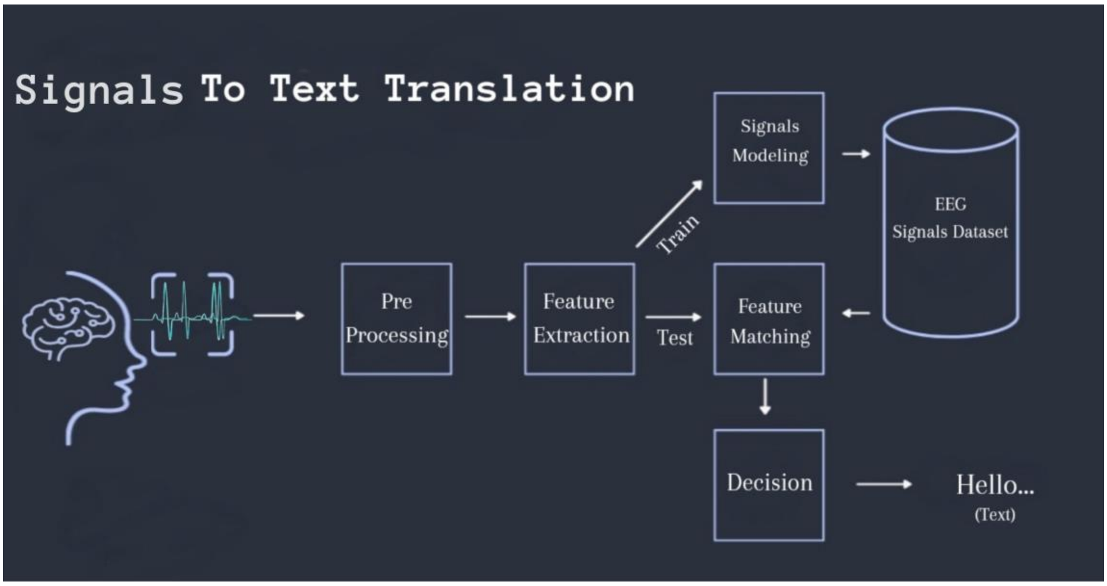

Fig 1. Methodology flow chart

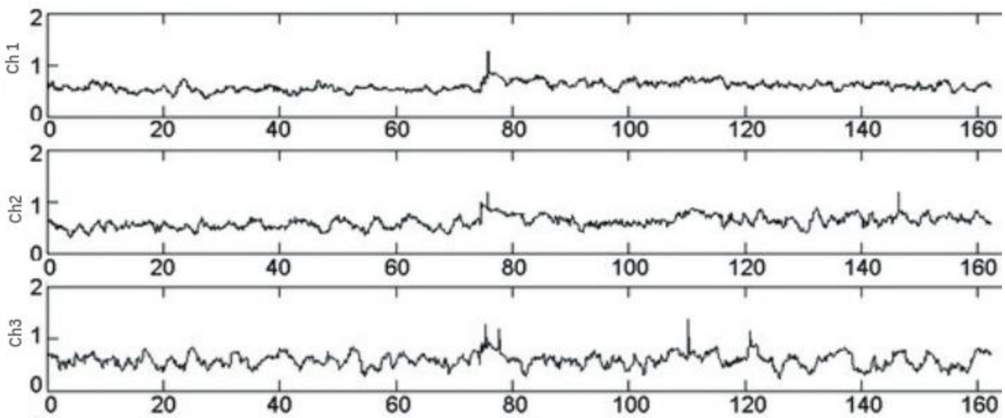

Fig 2. 3 channel EEG signals visualization

© IARJSET This work is licensed under a Creative Commons Attribution 4.0 International License

DOI: 10.17148/IARJSET.2023.101204

V. ACKNOWLEDGEMENT

We wish to extend our sincere appreciation to **Dr Vijayalaxmi Mekali** for the invaluable and constructive input provided throughout the planning and development of this project. We are truly grateful for her generous dedication of time.

Additionally, we'd like to express our thanks to the esteemed professors of KSIT for their unwavering support and encouragement.

REFERENCES

[1] N. Kobayashi, T. Nemoto and T. Morooka, "High Accuracy Silent Speech BCI Using Compact Deep Learning Model for Edge Computing," 2023 11th International Winter Conference on Brain-Computer Interface (BCI), Gangwon, Korea, Republic of, 2023, pp. 1-6, doi: 10.1109/BCI57258.2023.10078589.

[2] Nieto, N., Peterson, V., Rufiner, H.L. et al. Thinking out loud, an open-access EEG-based BCI dataset for inner speech recognition. Sci Data 9, 52 (2022). https://doi.org/10.1038/s41597-022-01147-2

[3] Shah U, Alzubaidi M, Mohsen F, Abd-Alrazaq A, Alam T, Househ M. The Role of Artificial Intelligence in Decoding Speech from EEG Signals: A Scoping Review. Sensors (Basel). 2022 Sep 15;22(18):6975. doi: 10.3390/s22186975. PMID: 36146323; PMCID: PMC9505262.

[4] C. Cooney, R. Folli and D. Coyle, "A Bimodal Deep Learning Architecture for EEG-fNIRS Decoding of Overt and Imagined Speech," in IEEE Transactions on Biomedical Engineering, vol. 69, no. 6, pp. 1983-1994, June 2022, doi: 10.1109/TBME.2021.3132861.

[5] Houssein, Essam & Hamad, Asmaa & Ali, Abdelmgeid. (2022). Human emotion recognition from EEG-based brain–computer interface using machine learning: a comprehensive review. Neural Computing and Applications. 34. 10.1007/s00521-022-07292-4.

[6] M. Rekrut, A. Fey, M. Nadig, J. Ihl, T. Jungbluth and A. Krüger, "Classifying Words in Natural Reading Tasks Based on EEG Activity to Improve Silent Speech BCI Training in a Transfer Approach," 2022 IEEE International Conference on Metrology for Extended Reality, Artificial Intelligence and Neural Engineering (MetroXRAINE), Rome, Italy, 2022, pp. 703-708, doi: 10.1109/MetroXRAINE54828.2022.9967665.

[7] Vorontsova D, Menshikov I, Zubov A, Orlov K, Rikunov P, Zvereva E, Flitman L, Lanikin A, Sokolova A, Markov S, et al. Silent EEG-Speech Recognition Using Convolutional and Recurrent Neural Network with 85% Accuracy of 9 Words Classification. Sensors. 2021; 21(20):6744. https://doi.org/10.3390/s21206744

[8] M. Rekrut, M. Sharma, M. Schmitt, J. Alexandersson and A. Krüger, "Decoding Semantic Categories from EEG Activity in Silent Speech Imagination Tasks," 2021 9th International Winter Conference on Brain-Computer Interface (BCI), Gangwon, Korea (South), 2021, pp. 1-7, doi: 10.1109/BCI51272.2021.9385357.

[9] Shahid, Aisha, Imran Raza and Syed Asad Hussain. "EmoWrite: A Sentiment Analysis-Based Thought to Text Conversion." ArXiv abs/2103.02238 (2021): n. Pag.

[10] Vorontsova D, Menshikov I, Zubov A, Orlov K, Rikunov P, Zvereva E, Flitman L, Lanikin A, Sokolova A, Markov S, Bernadotte A. Silent EEG-Speech Recognition Using Convolutional and Recurrent Neural Network with 85% Accuracy of 9 Words Classification. Sensors (Basel). 2021 Oct 11;21(20):6744. doi: 10.3390/s21206744. PMID: 34695956; PMCID: PMC8541074.

[11] Saminu, Sani & Xu, Guizhi & Shuai, Zhang & Abd El Kader, Isselmou & Halilu Jabire, Adamu & Karaye, Ibrahim & Ahmad, Isah & Abdulkarim, A.. (2021). Electroencephalogram (EEG) Based Imagined Speech Decoding and Recognition. Journal of Applied Materials and Technology. 2. 74-84. 10.31258/Jamt.2.2.74-84.

[12] K. Meng, D. B. Grayden, M. J. Cook, S. Vogrin and F. Goodarzy, "Identification of discriminative features for decoding overt and imagined speech using stereotactic electroencephalography," 2021 9th International Winter Conference on Brain-Computer Interface (BCI), Gangwon, Korea (South), 2021, pp. 1-6, doi: 10.1109/BCI51272.2021.9385355.

[13] D. -Y. Lee, M. Lee and S. -W. Lee, "Decoding Imagined Speech Based on Deep Metric Learning for Intuitive BCI Communication," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 29, pp. 1363-1374, 2021, doi: 10.1109/TNSRE.2021.3096874.

[14] N. Hamedi, S. Samiei, M. Delrobaei and A. Khadem, "Imagined Speech Decoding From EEG: The Winner of 3rd Iranian BCI Competition (iBCIC2020)," 2020 27th National and 5th International Iranian Conference on Biomedical Engineering (ICBME), Tehran, Iran, 2020, pp. 101-105, doi: 10.1109/ICBME51989.2020.9319439.

[15] Aditya Srivastava, Sameer Ahmed Ansari, Tanvi Shinde, Prashant Kanade and Prateek Mehta, "Think2Type: Thoughts to Text using EEG Waves", International Journal of Engineering Research & Technology (IJERT), ISSN: 2278-0181, vol. 9 Issue 06, June-2020, DOI: 10.17577/IJERTV9IS060431

[16] Arteiro, L., Lourenço, F., Escudeiro, P., Ferreira, C. (2020). Brain-Computer Interaction and Silent Speech Recognition on Decentralized Messaging Applications. In: Stephanidis, C., Antona, M. (eds) HCI International 2020

© IARJSET This work is licensed under a Creative Commons Attribution 4.0 International License

DOI: 10.17148/IARJSET.2023.101204

- Posters. HCII 2020. Communications in Computer and Information Science, vol 1226. Springer, Cham. https://doi.org/10.1007/978-3-030-50732-9_1

[17] Lee, SH., Lee, M., Lee, SW. (2020). EEG Representations of Spatial and Temporal Features in Imagined Speech and Overt Speech. In: Palaiahnakote, S., Sanniti di Baja, G., Wang, L., Yan, W. (eds) Pattern Recognition. ACPR 2019. Lecture Notes in Computer Science(), vol 12047. Springer, Cham. https://doi.org/10.1007/978-3-030-41299-9_30

[18] L. C. Sarmiento, J. B. Rodríguez, O. López, S. I. Villamizar, R. D. Guevara and C. J. Cortes-Rodriguez, "Recognition of silent speech syllables for Brain-Computer Interfaces," 2019 IEEE International Conference on E-health Networking, Application & Services (HealthCom), Bogota, Colombia, 2019, pp. 1-5, doi: 10.1109/HealthCom46333.2019.9009438.

[19] X. Zhang, L. Yao, Q. Z. Sheng, S. S. Kanhere, T. Gu and D. Zhang, "Converting Your Thoughts to Texts: Enabling Brain Typing via Deep Feature Learning of EEG Signals," 2018 IEEE International Conference on Pervasive Computing and Communications (PerCom), Athens, Greece, 2018, pp. 1-10, doi: 10.1109/PERCOM.2018.8444575.

[20] A. Kapur, S. Kapur, and P. Maes, "AlterEgo: A Personalized Wearable Silent Speech Interface." 23rd International Conference on Intelligent User Interfaces (IUI 2018), pp 43-53, March 5, 2018.

[21] Rezaei Tabar Y, Halici U. Brain Computer Interfaces for Silent Speech. European Review. 2017;25(2):208-230. doi:10.1017/S1062798716000569

[22] P. Ghane, G. Hossain and A. Tovar, "Robust understanding of EEG patterns in silent speech," 2015 National Aerospace and Electronics Conference (NAECON), Dayton, OH, USA, 2015, pp. 282-289, doi: 10.1109/NAECON.2015.7443084.

[23] Ravi, Kamalakkannan & Rajkumar, R. & Raj, M.M. & Devi, S.S.. (2014). Imagined Speech Classification using EEG. Advances in Biomedical Science and Engineering. 1. 20-32.

[24] Wang, Yijun & Jung, Tzyy-Ping. (2012). Improving Brain–Computer Interfaces Using Independent Component Analysis. 10.1007/978-3-642-29746-5_4.

[25] K. Brigham and B. V. K. V. Kumar, "Imagined Speech Classification with EEG Signals for Silent Communication: A Preliminary Investigation into Synthetic Telepathy," 2010 4th International Conference on Bioinformatics and Biomedical Engineering, Chengdu, China, 2010, pp. 1-4, doi: 10.1109/ICBBE.2010.5515807.

© IARJSET This work is licensed under a Creative Commons Attribution 4.0 International License