SAHYADRI
COLLEGE OF ENGINEERING & MANAGEMENT
An Autonomous Institution
MANGALURU

Department of Computer Science & Engineering

SYNOPSIS

1.	**Title of the Project**	Developing a Brain-Computer Interface for Real-Time Neural Signal Decoding and Speech Conversion		
2.	**Department & College**	Department of Computer Science & Engineering, Sahyadri College of Engineering & Management		
3.	**Name of the Students**	GAGAN V	4SF21CS049	
		MISBAH ZOHAR	4SF21CS084	
		NEHA P ACHAR	4SF21CS096	
		PRATEEK MALAGUND	4SF21CS109	
4.	**Guide**	Dr. MUSTAFA BASTHIKODI	HoD, Dept of CSE, SCEM	

Abstract

A brain-computer interface (BCI) is a form of technology that develops a communication channel between a user and certain environmental objects using the user's brain impulses to develop a non-invasive brain-computer interface that can "translate ideas into voice" (that converts brain signals into speech). The goal of this project is to develop a framework for a brain-computer interface (BCI) that can translate neural signals into spoken words in real-time, so assisting those with serious brain injuries or illnesses like ALS who struggle with communication. The device uses electroencephalography (EEG) to record brain activity and then uses certain algorithms to translate neural impulses into phonetic representations and eventually into speech that sounds natural. The project uses high-performance computation and frameworks such as PyTorch and TensorFlow, and it needs big annotated datasets for training. The technology captures brain activity during speech or speech-like stimuli, filters out background noise, extracts meaningful characteristics, and trains models to map these features to phonetic outputs and generate speech synthesized from them. The system's performance is validated through rigorous testing and iterative refinement, aiming to provide a reliable, real-time communication tool that significantly enhances the quality of life for individuals with severe communication disorders.

Keywords:

Brain-computer interface (BCI), Neural Signal Decoding, Speech Synthesis, Electroencephalography (EEG), Machine Learning

1. Introduction

Brain-computer interfaces (BCIs) are a rapidly evolving technology with the potential to revolutionize human-computer interaction. BCIs measure brain activity and translate it into commands for computers or other devices, allowing users to control machines and devices using only their thoughts. These Neuro Gadgets range from moving robotic spiders and balls for entertainment to more practical applications for assisting people with disabilities, such as limb paralysis. The versatility and potential of BCIs are immense, promising significant advancements in both entertainment and assistive technologies.

BCIs are typically categorized into unidirectional and bidirectional systems based on the direction of information flow. Unidirectional BCIs either receive signals from the brain or send them to it, enabling tasks like moving a cursor on a screen or sending sensory feedback to the user. In contrast, bidirectional BCIs facilitate information exchange in both directions, allowing the brain to control external devices while receiving sensory feedback from them. This bidirectional communication is crucial for developing sophisticated assistive technologies that can enhance the quality of life for individuals with severe disabilities.

Research into feedback methods is ongoing, aiming to develop technologies that can transform external commands into electrical signals transmitted via the nervous system. One promising application involves enabling electrical stimulation of leg muscles in people with spinal cord injuries. This technology could allow individuals to regain mobility by controlling their movements through a tablet device, significantly enhancing their independence and quality of life. The potential for such applications underscores the importance of continued research and development in this field.

The utilization of neural networks and other machine learning algorithms in BCI signal processing is commonplace, given the variability in brain activity between individuals.

These systems require lengthy training sessions to accurately interpret commands from a specific user. The duration of the training depends on the number of commands that the BCI needs to recognize and the complexity of the tasks. Despite these challenges, the adaptability and precision of machine learning algorithms make them indispensable for developing effective BCIs.

While this technology is still in its early stages of development, recent advances have shown great promise across a range of applications, from medical rehabilitation to gaming and entertainment. For instance, BCIs are being explored for use in neurorehabilitation, helping stroke patients regain motor functions by retraining their brains through targeted exercises. In the gaming industry, BCIs offer an immersive experience by allowing players to control game elements with their thoughts, creating a new dimension of interaction and engagement.

In conclusion, BCIs represent a transformative technology with the potential to significantly impact various aspects of human life. From aiding individuals with disabilities to creating new forms of entertainment, the possibilities are vast and varied. Continued research and collaboration across disciplines will be essential to unlocking the full potential of BCIs and ensuring their safe and effective integration into society. This paper aims to contribute to this ongoing dialogue by providing a comprehensive overview of the latest advancements and future directions in BCI research.

2. Literature Survey

REFERENCE	METHODS	MERITS	DEMERITS	FINDINGS
[1]	• EEG and MEG for neural signal acquisition • Machine learning for neural signal decoding • Speech synthesis for generating audible speech	Enhancing communication for speech-impaired individuals through real-time neural signal translation to speech shows high potential for accuracy with advanced algorithms.	High computational demands. Variability in neural patterns across users. Challenges in noise reduction and signal clarity.	• Feasibility of BCI for speech synthesis • Importance of personalized models for accuracy • Critical factors: signal quality and latency
[2]	• EEG-based signal acquisition • Machine learning algorithms • Signal preprocessing and feature extraction	Non-invasive. High accuracy in pattern recognition. And, Reduces noise, improves signal clarity	Limited signal resolution. Requires large datasets for training. And, Computationally intensive	• BCIs show promise in enhancing communication • Integration of AI improves decoding efficiency • Personalized models yield better performance
[3]	• Advanced ML algorithms for decoding and synthesis • Continuous subject adaptation for accuracy improvement • Implanted BCI from EEG signals	High accuracy in neural signal capture. Real-time processing and minimal latency. And, Personalized adaptation to user's neural patterns	Invasive implantation procedure. Potential for neural signal degradation over time. And, Requires extensive training and calibration	• Achieved natural-soundin g speech synthesis • High intelligibility and user satisfaction • Demonstrated significant improvement in communication

[4]	• EEG to capture brain activity • Machine learning models to decode neural signals • Speech synthesis to convert phonetic outputs to speech	Non-invasive. Potential for real-time communication. And, Adaptability to individual neural patterns	Lower spatial resolution. Requires extensive training data. And, High computational demands	• BCI can facilitate speech communication • Achieved significant accuracy improvement
[5]	• EEG-based BCI systems • Signal processing techniques • Machine learning for neural signal decoding	Offers non-invasive neural signal acquisition. Facilitates real-time data processing. And, enables personalized system adaptation	Limited spatial resolution. Susceptible to noise and artifacts. And, Training requires large annotated datasets	• Significant progress in motor imagery tasks • Successful control of external devices • Challenges in long-term reliability
[6]	• EEG and MEG for neural signal acquisition • Machine learning for neural signal decoding • Speech synthesis for converting phonetics to speech	Enables non-invasive brain activity monitoring. Accurate mapping of neural patterns to phonetic representations. And, Natural-sounding speech output	Limited spatial resolution with EEG. Requires large, annotated datasets for training. And, Latency issues in real-time systems	• Neural signals can encode speech intent • Achieved real-time speech synthesis from thoughts • System accuracy and latency are critical for usability
[7]	• Cascade and Parallel Convolutional Recurrent Neural Networks • EEG-based intention recognition • Spatio-temporal representations	Effective for spatio-temporal EEG data analysis. Captures temporal dependencies in EEG signals. And, Utilizes both spatial and temporal EEG features	Requires substantial computational resources. Sensitivity to noise in EEG signals. And, Complexity in hyperparameter tuning	• Achieved high accuracy in intention recognition tasks • Improved latency in real-time intention prediction

[8]	• Invasive brain stimulation • Review of current research • Future BCI advancements	Precise neural activation control. Comprehensive understanding of challenges.And, Insights into potential advancements	Surgical risks. Limited non-invasive application. And, Ethical concerns	• Stimulates artificial speech perception • Identifies gaps in BCI technology • Calls for Interdisciplinary Collaboration
[9]	• Multiple Generator Conditional Wasserstein GAN • Data augmentation using GANs • Conditional training for emotion recognition	Improves EEG data diversity for emotion recognition. Increases dataset size without manual labeling. And, Tailors EEG data generation to specific emotional states	Enhances model robustness and generalization. Requires substantial computational resources. And, May introduce synthetic data bias	• Augmented EEG data enhances emotion recognition accuracy • GANs effectively generate diverse EEG samples • Conditional GANs improve emotion recognition
[10]	• EEG-based BCI, P300 potential for spelling. • Training on mental spelling tasks with feedback.	Enables communication for locked-in patients. Non-invasive, real-time communication.And a customizable and adaptable interface.	Requires intact P300 potential. Needs extensive calibration. And, Limited vocabulary and typing speed.	• Achieved high accuracy (70-90%) in letter selection • Users maintained performance over time.
[11]	• Implanted EEG on the brain surface • Used ML for neural decoding • Rigorous testing and validation	Direct neural signal acquisition. Improved signal translation accuracy. And, Ensured the reliability of the BCI system	Invasive procedure. Limited adaptability. And, High cost and complexity	• Enabled reliable communication with ALS • Demonstrated real-time control feasibility • Achieved high accuracy in decoding signals

[12]	• Microelectrode assembly • Neurotrophic electrode coating • Surgical implantation in speech cortex	Precise neural signal acquisition. Prolonged electrode lifespan. And, Targets specific brain areas	Invasive procedure. Tissue reaction risk. And, Surgical risks	• Neural signal acquisition in the speech cortex • Improved electrode longevity • Direct interface for speech production
[13]	• EEG and MEG for neural signal acquisition • Machine learning for neural signal decoding • Speech synthesis from phonetic representations	Natural-sounding speech output. High potential for personalized applications. And, Enables real-time neural decoding	Limited spatial resolution (EEG). Signal noise and artifacts. And, Latency issues	• Advancements in signal processing • Challenges in achieving high accuracy levels • Importance of dataset quality for model performance
[14]	• Review of EEG-based BCI applications • Analysis of historical and current trends • Future prospects and innovations	Comprehensive overview. Insight into EEG technology. And, Identifies emerging applications	They have limited recent advancements. Lack of standardized methodologies. And, Signal processing challenges	• Improved accuracy with ML algorithms • AI integration enhances neural decoding capabilities • Hybrid BCIs combining EEG with fNIRS are gaining interest
[15]	• Case studies of locked-in patients • Neurological assessments • Longitudinal observation	Provides detailed clinical insights. Comprehensive neurological data. And, Allows for longitudinal analysis	Limited sample size. Lack of quantitative metrics. And, Subjectivity in assessments	• Identified distinct locked-in syndrome types • Described varying degrees of communication ability

3. Problem Statement & Description

3.1 Problem Statement: Individuals with speech impairments, such as those suffering from conditions like ALS or severe brain injuries, often face significant challenges in communication. The project aims to bridge this gap by developing a system that translates neural signals directly into spoken language.

3.2 Description: The proposed project seeks to address this critical issue by developing a system that translates neural signals directly into spoken language. By leveraging the advancements in neurotechnology and machine learning, this system aims to interpret the brain's speech-related neural activity and convert it into audible words and sentences. This approach promises to provide a more direct and natural mode of communication for individuals who cannot speak, enabling them to convey their thoughts and needs in real-time

For individuals with ALS, severe brain injuries, or other debilitating conditions that impair speech, this system could restore a fundamental aspect of human interaction—the ability to communicate verbally. It would allow these individuals to engage more fully in social, educational, and professional settings, thereby enhancing their overall quality of life. Moreover, the successful implementation of such a system could pave the way for further advancements in brain-computer interface (BCI) technologies, opening new avenues for assisting individuals with various disabilities.

In summary, the development of a system that translates neural signals into spoken language represents a significant leap forward in assistive communication technologies. It holds the promise of transforming the lives of those with severe speech impairments by providing them with a new, more effective means of communication.

4. Objectives

The key objectives of the project are:

- Collect or create datasets using sensors to measure the brain's analog signals, then amplify and filter these signals to remove noise, and convert them into a digital format for further processing
- Design and develop a Signal Processing Unit incorporating algorithms for feature extraction, translation, and speech synthesis.
- Develop a system to adapt to individual users' neural patterns, enhancing accuracy and usability by tailoring the signal processing to each user's unique brain activity.
- Evaluate the developed algorithm and framework against the benchmark datasets, and publish the results in reputed journals.

5. Proposed Methodology

- Data Collection: Record neural activity using EEG devices while subjects speak or imagine speaking specific phrases. Collect paired datasets of neural signals and spoken language.
- Preprocessing: Apply signal processing techniques to clean and preprocess the raw neural data, removing noise and artifacts.
- Feature Extraction: Use algorithms to extract relevant features from the neural signals that correlate with speech components such as phonemes, intonations, and prosody.
- Model Training: Train machine learning models to map neural features to phonetic representations. Utilize supervised learning with paired datasets to improve the accuracy of the translation.
- Speech Synthesis: Implement a speech synthesis engine to convert phonetic outputs into audible speech. Ensure the synthesized speech is natural and intelligible.
- Validation and Testing: Validate the system using separate test datasets to evaluate performance metrics such as accuracy, latency, and intelligibility. Iteratively refine the models based on feedback and performance results

5.1 System Design

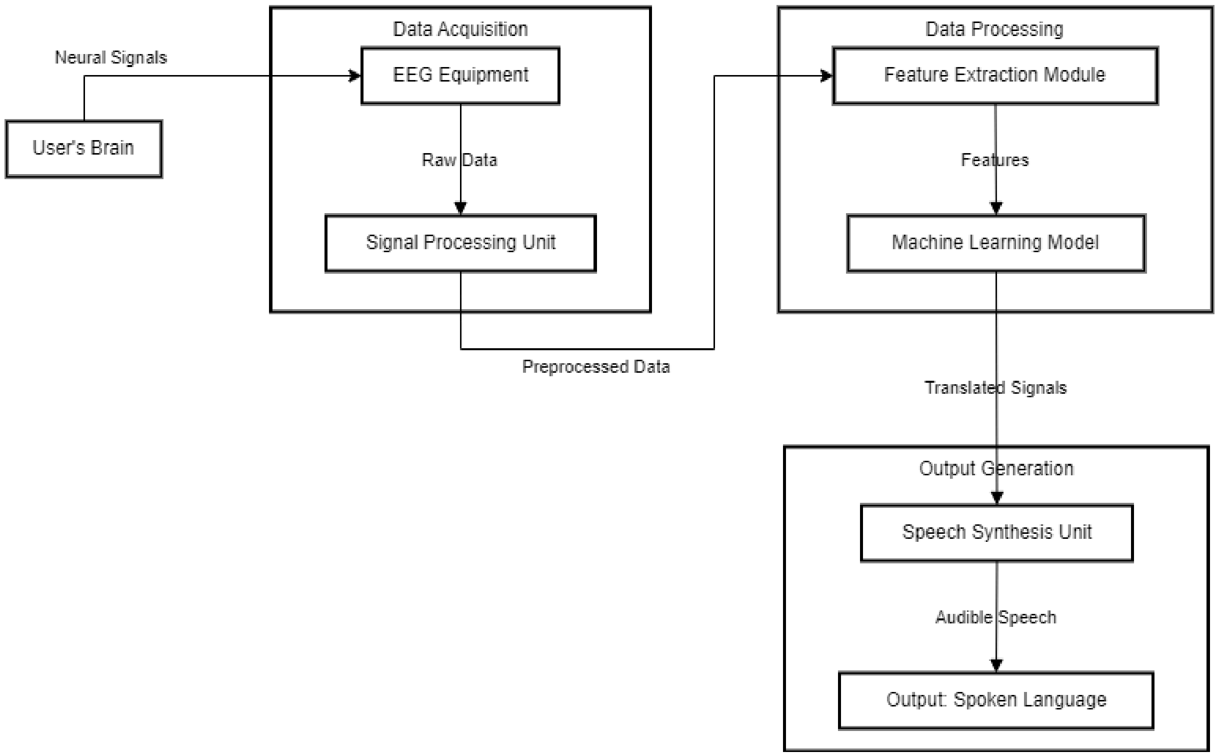

Fig 5.1.1: System Design of a Brain Computer Interface

5.2 Use Cases:

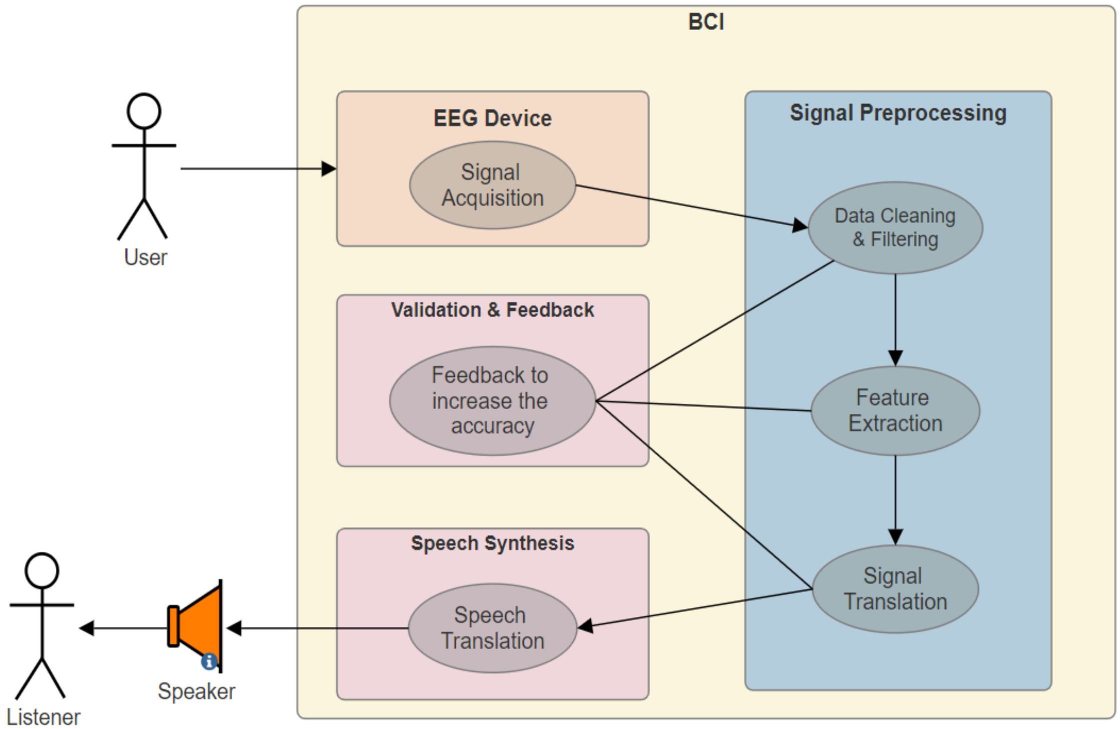

Fig 5.2.1: Use Case Diagram for a Brain Computer Interface

5.3 Functional & Non-Functional Requirements:

5.3.1 Functional Requirements:

- The system must acquire brain activity signals using EEG devices when a user attempts to speak or imagines speaking.
- The system must process the acquired neural signals using algorithms for feature extraction, translation, and speech synthesis to convert them into phonetic outputs.
- The system must convert the phonetic outputs into natural and intelligible speech in real-time.
- The system must adapt to individual users' neural patterns to improve translation accuracy over time.

5.3.2 Non-Functional Requirements:

- The system must process and translate neural signals with minimal latency to provide immediate feedback to the user.
- The system must achieve high accuracy in translating neural signals to speech, ensuring minimal errors in communication.
- The system must be user-friendly, with an intuitive interface that can be easily used by individuals with varying levels of technical proficiency.
- The system must ensure the security and privacy of the users' neural data.

5.4 Data Flow Diagram:

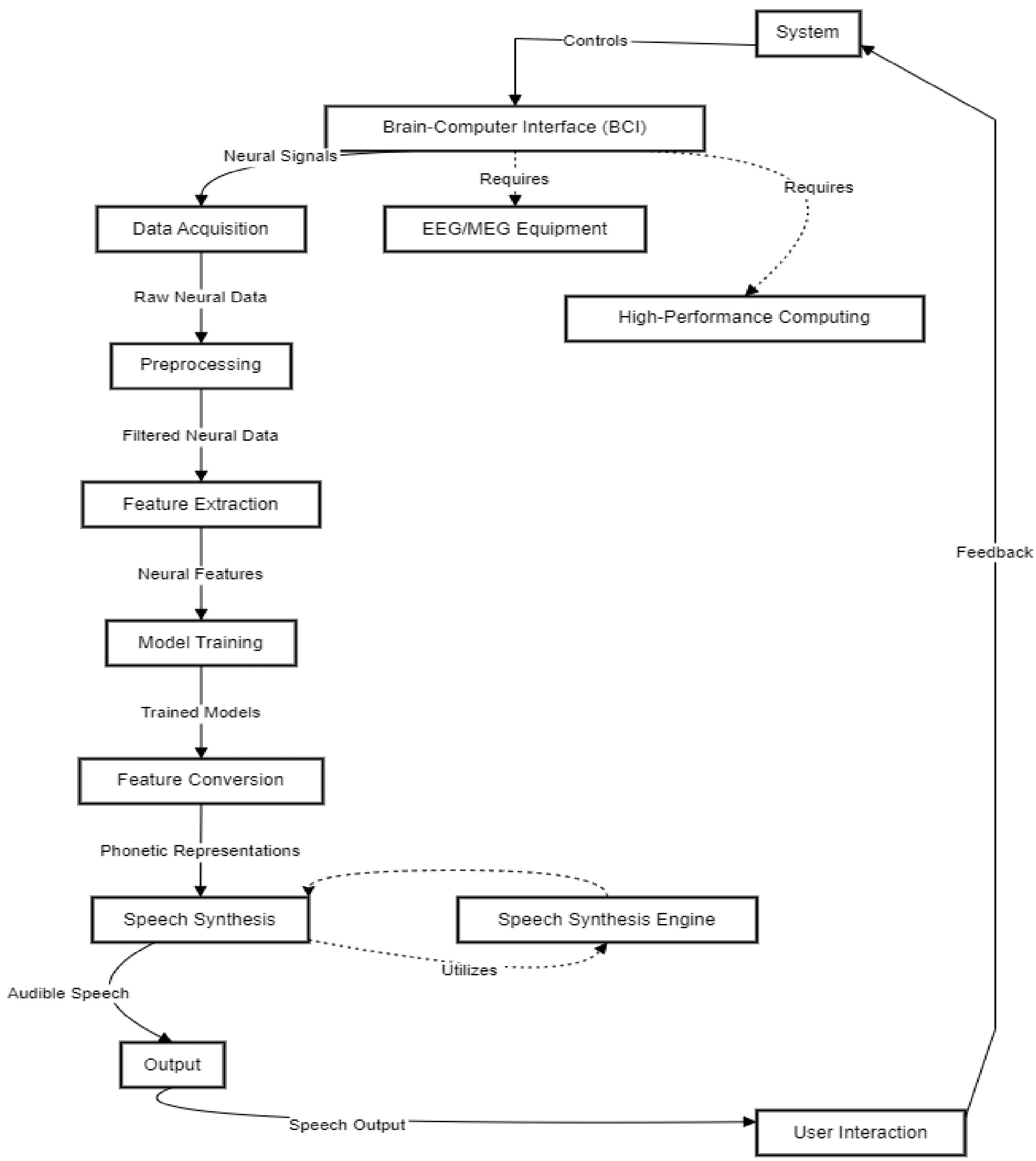

Fig 5.4: L0 Data Flow Diagram of a BCI

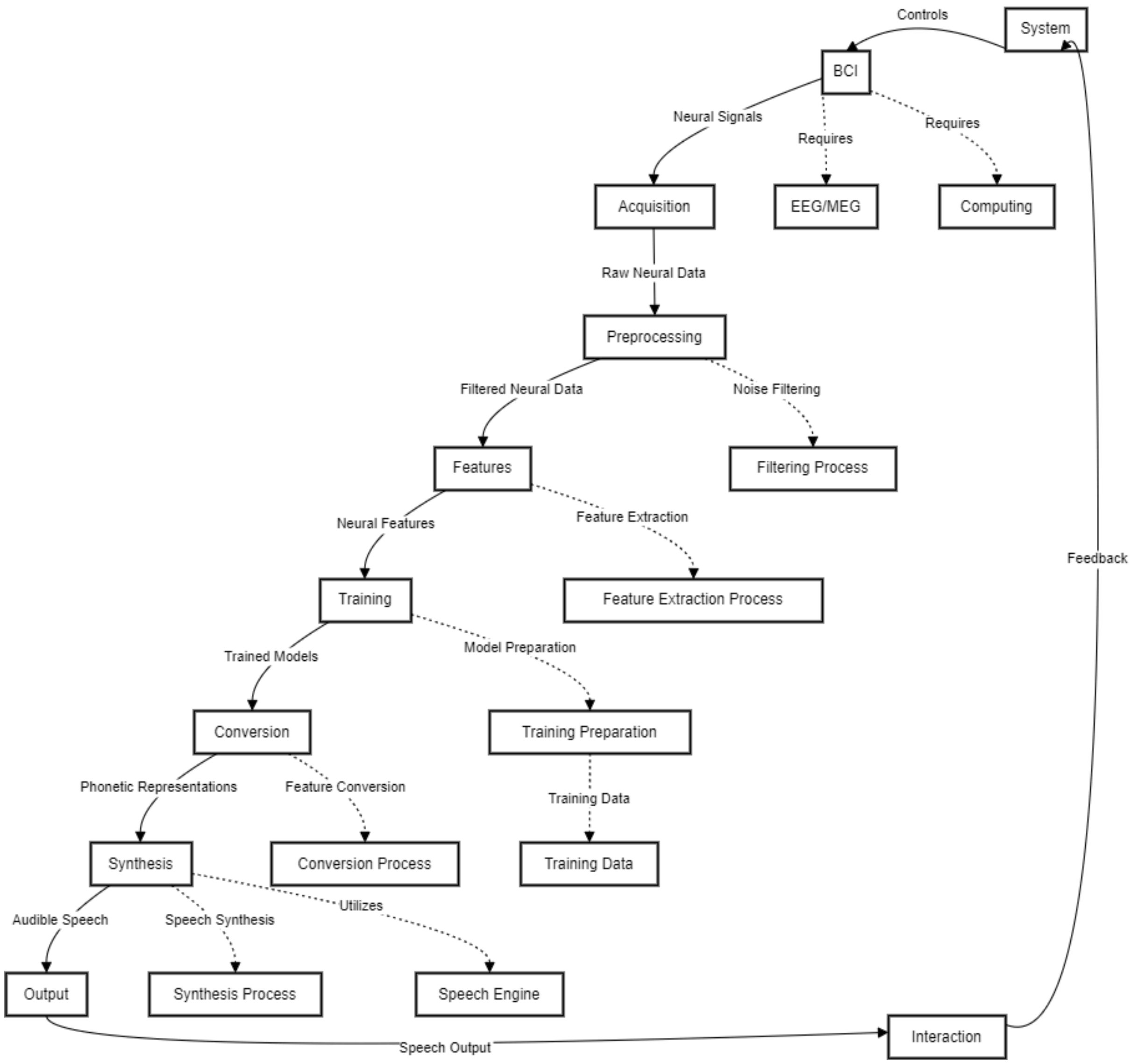

Fig 5.5: L1 Data Flow Diagram of a BCI

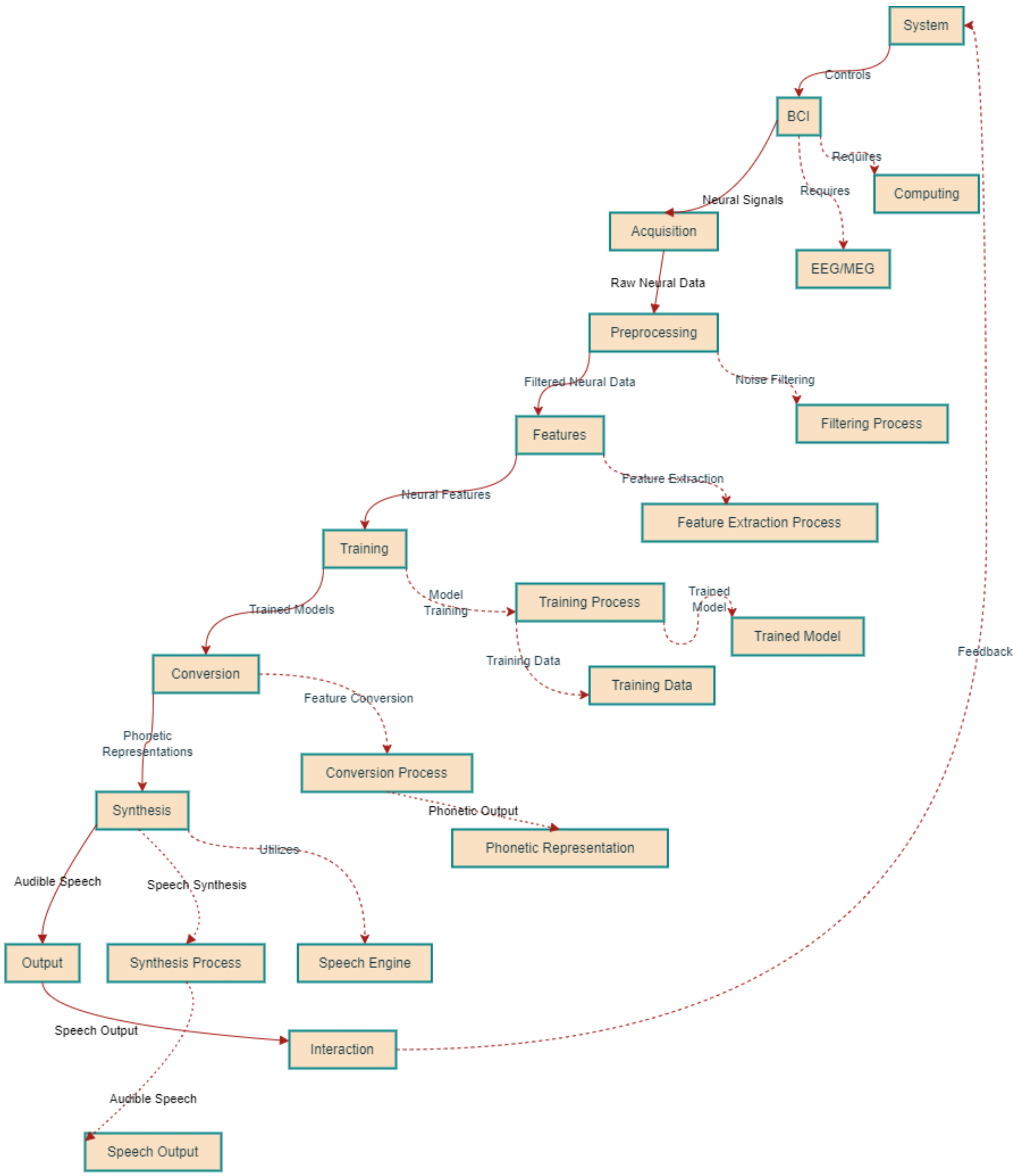

Fig 5.6: L2 Data Flow Diagram of a BCI

5.5 System Architecture Diagram:

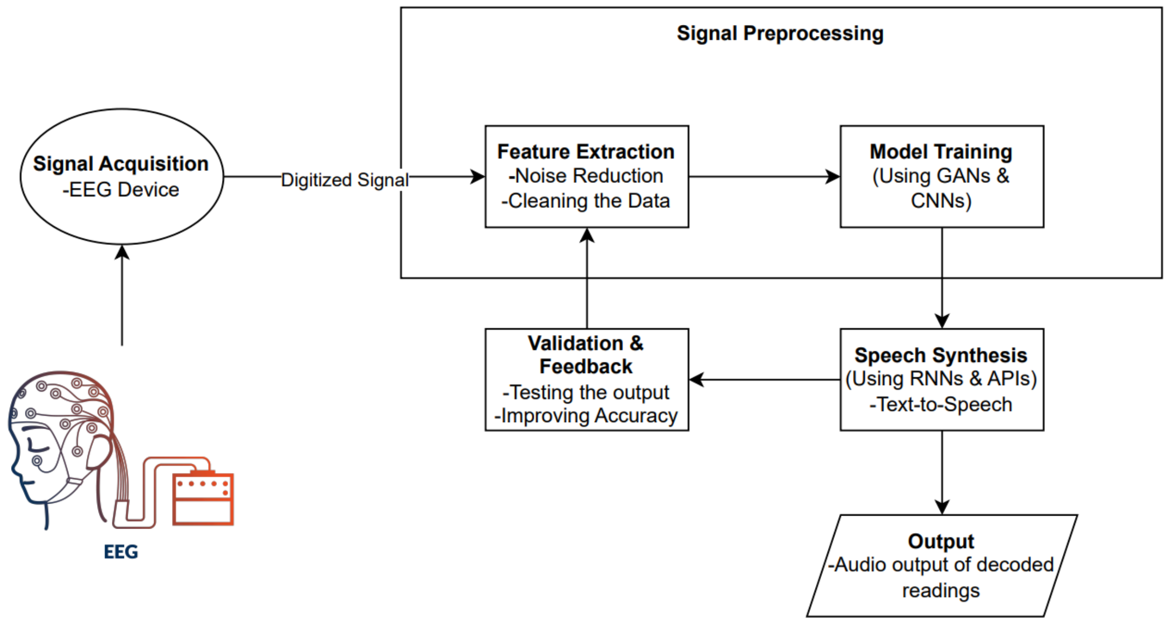

Fig 5.7: System Architecture Diagram of a BCI

6. Outcome of the Work

The project outcome for developing a Brain-Computer Interface (BCI) for real-time neural signal decoding and speech conversion can be summarized as:

- Reliable Neural Signal Acquisition and Processing: Implementation of EEG equipment to capture neural signals and advanced signal processing techniques to clean and preprocess the data, ensuring accurate feature extraction and translation.
- Machine Learning Models for Neural Signal Translation: Development and training of machine learning models that map neural signal features to phonetic representations, improving the accuracy and efficiency of the translation from neural signals to speech.
- Enhanced Communication for Speech-Impaired Individuals: A functional system that translates neural signals directly into spoken language, providing a new communication method for individuals with speech impairments, such as those suffering from ALS or severe brain injuries.

7. Work Plan/Gantt Chart

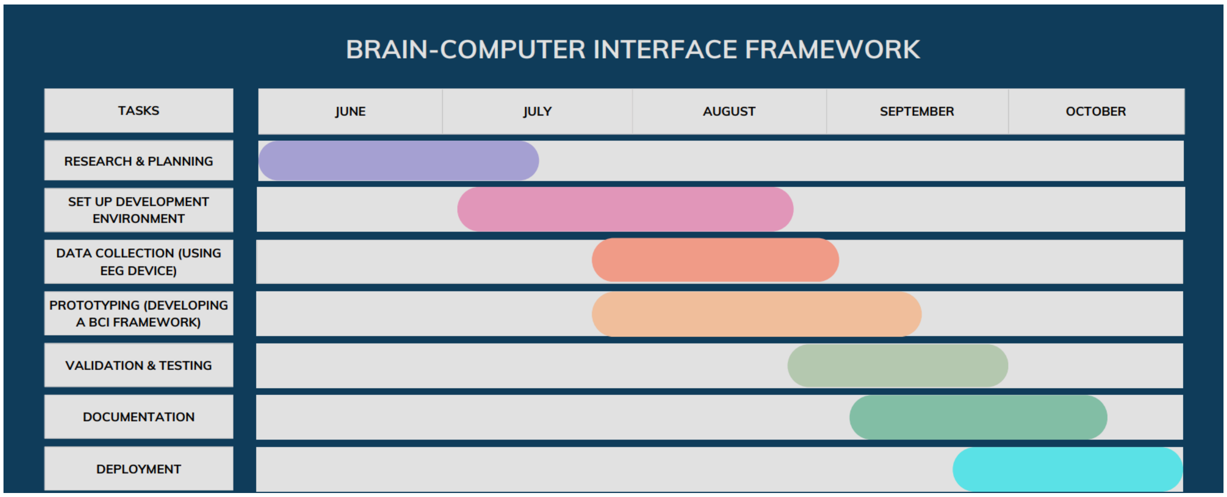

Fig 7.1: Project Development timeline for a BCI

8. Conclusion

In conclusion, this project aims to develop a fully functioning Brain Computer Interface (BCI) system, that translates neural signals into spoken language, providing a communication method for individuals with speech impairments, such as those suffering from ALS or severe brain injuries.

Key Milestones of the project include

- Reliable Neural Signal Acquisition and Processing: The implementation of high-performance EEG equipment and advanced signal processing techniques ensuring the accurate capture and preprocessing of neural data. This facilitates effective feature extraction and translation of neural signals into phonetic representations.
- Machine Learning Models for Neural Signal Translation: The development and training of sophisticated machine learning models will significantly improve the accuracy and efficiency of translating neural signals into speech. The models will be iteratively refined to adapt to individual users' neural patterns, enhancing usability and performance.
- Enhanced Communication for Speech-Impaired Individuals: The BCI system will provide a functional and natural communication method for individuals with severe speech impairments, enabling them to convey their thoughts and needs audibly in real-time. This innovation will greatly enhance their social, educational, and professional interactions, thereby improving their overall quality of life.

- **References**

[1] Luo, S., Rabbani, Q. and Crone, N.E., 2023. Brain-computer interface: applications to speech decoding and synthesis to augment communication. *Neurotherapeutics*, *19*(1), pp.263-273.

[2] Peksa, Janis, and Dmytro Mamchur. 2023. "State-of-the-Art on Brain-Computer Interface Technology" *Sensors* 23, no. 13: 6001. https://doi.org/10.3390/s23136001

[3] Angrick, M., Luo, S., Rabbani, Q. *et al.* Online speech synthesis using a chronically implanted brain-computer interface in an individual with ALS. *Sci Rep* 14, 9617 (2024). https://doi.org/10.1038/s41598-024-60277-2

[4] Brumberg, Jonathan & Nieto-Castanon, Alfonso & Kennedy, Philip & Guenther, Frank. (2010). Brain-Computer Interfaces for Speech Communication. Speech communication. 52. 367-379. 10.1016/j.specom.2010.01.001.

[5] Allison, Brendan Z., Elizabeth Winter Wolpaw, and Jonathan R. Wolpaw. "Brain-computer interface systems: progress and prospects." *Expert review of medical devices* 4, no. 4 (2007): 463-474.

[6] Warshi, A. (2023). BRAIN-COMPUTER INTERFACE FOR CONVERTING THOUGHTS TO SPEECH.

[7] Zhang, Dalin & Yao, Lina & Zhang, Xiang & Wang, Sen & Chen, Weitong & Boots, Robert. (2017). EEG-based Intention Recognition from Spatio-Temporal Representations via Cascade and Parallel Convolutional Recurrent Neural Networks.

[8] Hong, Y., Ryun, S., & Chung, C. K. (2024). Evoking artificial speech perception through invasive brain stimulation for brain-computer interfaces: Current challenges and future perspectives. *Frontiers in Neuroscience*, *18*, 1428256. https://doi.org/10.3389/fnins.2024.1428256

[9] Zhang, A., Su, L., Zhang, Y., Fu, Y., Wu, L. and Liang, S., 2021. EEG data augmentation for emotion recognition with a multiple generator conditional Wasserstein GAN. *Complex & Intelligent Systems*, pp.1-13.

[10] Kübler, A., Furdea, A., Halder, S., Hammer, E.M., Nijboer, F. and Kotchoubey, B., 2009. A brain–computer interface controlled auditory event‐related potential (P300) spelling system for locked‐in patients. Annals of the New York Academy of Sciences, 1157(1), pp.90-100.

[11] Vansteensel, M.J., Pels, E.G., Bleichner, M.G., Branco, M.P., Denison, T., Freudenburg, Z.V., Gosselaar, P., Leinders, S., Ottens, T.H., Van Den Boom, M.A. and Van Rijen, P.C., 2016. Fully implanted brain–computer interface in a locked-in patient with ALS. New England Journal of Medicine, 375(21), pp.2060-2066.

[12] Jess Bartels, Dinal Andreasen, Princewill Ehirim, Hui Mao, Steven Seibert, E. Joe Wright, Philip Kennedy, Neurotrophic electrode: Method of assembly and implantation into human motor speech cortex, Journal of Neuroscience Methods, Volume 174, Issue 2, 2008, Pages 168-176, ISSN 0165-0270, https://doi.org/10.1016/j.jneumeth.2008.06.030.

[13] Mridha, M.F., Das, S.C., Kabir, M.M., Lima, A.A., Islam, M.R. and Watanobe, Y., 2021. Brain-computer interface: Advancement and challenges. Sensors, 21(17), p.5746.

[14] Värbu, K., Muhammad, N. and Muhammad, Y., 2022. Past, present, and future of EEG-based BCI applications. Sensors, 22(9), p.3331.

[15] Bauer, G., Gerstenbrand, F. and Rumpl, E., 1979. Varieties of the locked-in syndrome. Journal of neurology, 221, pp.77-91.