



A State-of-the-Art Review of EEG-Based Imagined Speech Decoding

Diego Lopez-Bernal*, David Balderas, Pedro Ponce and Arturo Molina

Tecnologico de Monterrey, National Department of Research, Mexico City, Mexico

Currently, the most used method to measure brain activity under a non-invasive procedure is the electroencephalogram (EEG). This is because of its high temporal resolution, ease of use, and safety. These signals can be used under a Brain Computer Interface (BCI) framework, which can be implemented to provide a new communication channel to people that are unable to speak due to motor disabilities or other neurological diseases. Nevertheless, EEG-based BCI systems have presented challenges to be implemented in real life situations for imagined speech recognition due to the difficulty to interpret EEG signals because of their low signal-to-noise ratio (SNR). As consequence, in order to help the researcher make a wise decision when approaching this problem, we offer a review article that sums the main findings of the most relevant studies on this subject since 2009. This review focuses mainly on the pre-processing, feature extraction, and classification techniques used by several authors, as well as the target vocabulary. Furthermore, we propose ideas that may be useful for future work in order to achieve a practical application of EEG-based BCI systems toward imagined speech decoding.

Keywords: EEG, BCI, review, imagined speech, artificial intelligence

OPEN ACCESS

Edited by:

Hiram Ponce,
Universidad Panamericana, Mexico

Reviewed by:

Juan Humberto Sossa,
Instituto Politécnico Nacional (IPN),
Mexico

Yaqi Chu,
Shenyang Institute of Automation
(CAS), China

*Correspondence:

Diego Lopez-Bernal
lopezbernal.d@tec.mx

Specialty section:

This article was submitted to
Brain-Computer Interfaces,
a section of the journal
Frontiers in Human Neuroscience

Received: 31 January 2022

Accepted: 24 March 2022

Published: 26 April 2022

Citation:

Lopez-Bernal D, Balderas D, Ponce P
and Molina A (2022) A
State-of-the-Art Review of EEG-Based
Imagined Speech Decoding.
Front. Hum. Neurosci. 16:867281.
doi: 10.3389/fnhum.2022.867281

1. INTRODUCTION

One of the main technological objectives in our current era is to generate a connected environment in which humans can be able to create a link between their daily and real life physical activities and the virtual world (Chopra et al., 2019). This type of applications are currently developed under a framework denominated as Future Internet (FI). There is a wide range of technological implementations that can benefit from FI, such as human-computer interaction and usability (Haji et al., 2020). For example, speech driven applications such as Siri and Google Voice Search are widely used in our daily life to interact with electronic devices (Herff and Schultz, 2016). These applications are based on a speech recognition algorithm, which allows the device to convert human voice to text. Nevertheless, there are certain health issues that may impede some people from using these applications.

Verbal communication loss can be caused by injuries and neurodegenerative diseases that affect the motor production, speech articulation, and language understanding. Few examples of these health issues include stroke, trauma, and amyotrophic lateral sclerosis (ALS) (Branco et al., 2021). In some cases, these neurodegenerative conditions may lead patients to fall into a locked-in syndrome (LIS), in which they are not capable to communicate due to the complete loss of motor control.

To address this problem, Brain Computer Interfaces (BCI) have been proposed as an assistive technology to provide a new communication channel for those individuals with LIS. BCI technologies offer a bridge between the brain and outer world, in such a way that it creates a bi-directional communication interface which reads the signals generated by the human brain and converts them into the desired cognitive task (Gu et al., 2021; Rasheed, 2021; Torres-García et al., 2022). In such manner, a thought-to-speech interface can be implemented so that people who are not able to speak due to motor disabilities can use their brain signals to communicate without the need of moving any body part.

Generally speaking, BCI for imagined speech recognition can be decomposed into four steps:

1. Signal acquisition: this step involves a deep understanding of the properties of the signals that are being recorded, as well as how the signals are going to be captured.
2. Pre-processing: the main objective of this step is to unmask and enhance the information and patterns within the signal.
3. Feature extraction: this step involves the extraction of the main characteristics of the signal.
4. Classification: this is the final step, in the different mental states are classified depending on their features.

Several methods, both invasive and non-invasive, have been proposed and studied in order to acquire the signals that the brain produce during the speech imagining process. Some of these methods are magnetoencephalography (MEG), functional magnetic resonance imaging (fMRI), functional near-infrared spectroscopy (fNIRS), electrocardiography (ECOG), and electroencephalography (EEG) (Sereshkeh et al., 2018; Angrick et al., 2019; Dash et al., 2020b; Fonken et al., 2020; Si et al., 2021). Invasive methods, such as ECOG, have proven to provide, in average, greater classifying accuracies than non-invasive methods (MEG, fMRI, fNIRS, and EEG) during imagined speech decoding. In fact, invasive techniques have more easily exceeded the threshold for practical BCI imagined speech application (70%), in contrast to non-invasive techniques (Sereshkeh et al., 2018). Among the mentioned techniques for imagined speech recognition, EEG is the most commonly accepted method due to its high temporal resolution, low cost, safety, and portability (Saminu et al., 2021). Nevertheless, speech-based BCI systems using EEG are still in their infancy due to several challenges they have presented in order to be applied to solve real life problems.

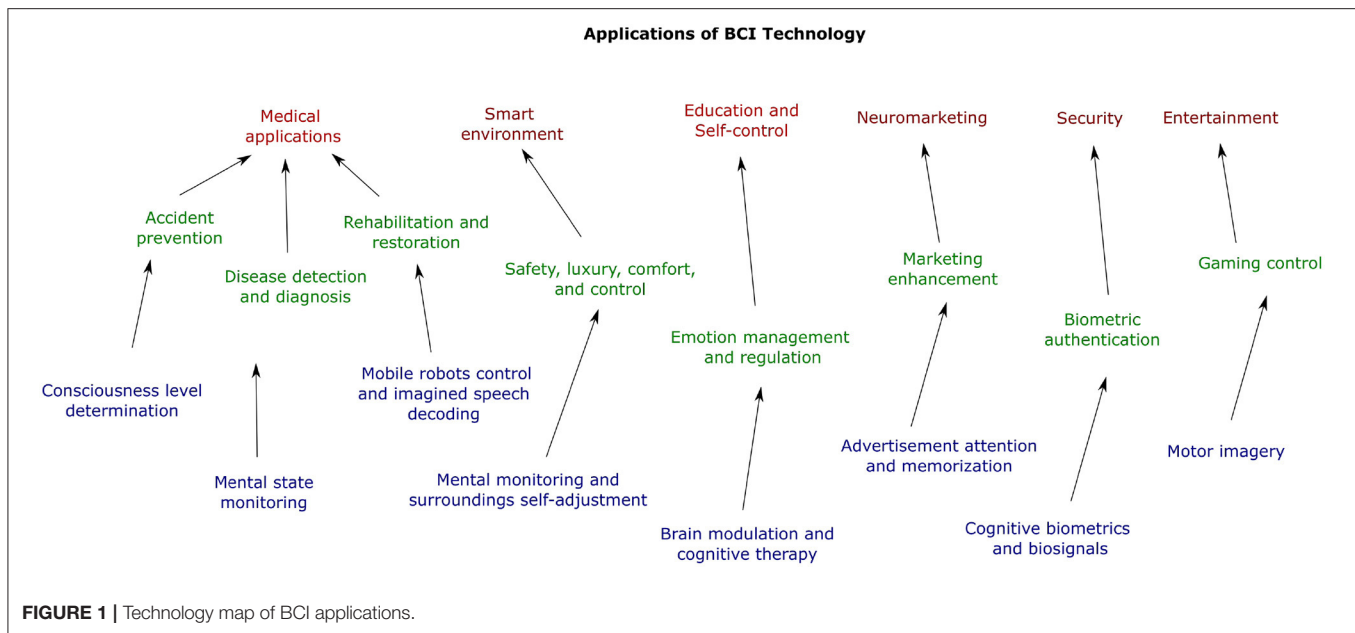
One of the main challenges that imagined speech EEG signals present is their low signal-to-noise ratio (SNR). This low SNR cause the component of interest of the signal to be difficult to recognize from the background brain activity given by muscle or organs activity, eye movements, or blinks. Furthermore, even EEG equipment is sensitive enough to capture electrical line noise from the surroundings (Bozhkov and Georgieva, 2018). Moreover, despite EEG having high temporal resolution, it lacks from spatial resolution which can lead to low accuracy on the source of information on the brain cortex, distortion of topographical maps by removing high spatial frequency, and difficulty to reject artifacts from the main signal (Kwon et al.,

2019). Because of these issues, classical machine learning (ML) methods that have proven to be successful in the recognition of motor imagery tasks have not obtained good performance when applied to imagined speech recognition. Thus, deep learning (DL) algorithms, along with various filtering and feature extraction techniques, have been proposed to enhance the performance of EEG-based BCI systems (Antoniades et al., 2016).

That being said, imagined speech recognition has proven to be a difficult task to achieve within an acceptable range of classification accuracy. Therefore, in order to help researchers to take the best decisions when approaching this problem, the main objective of the present review is to provide an insight about the basics behind EEG-based BCI systems and the most recent research about their application toward imagined speech decoding, as well as the most relevant findings on this area. The rest of the paper is organized as follows: Section 2 investigates the current applications of BCI systems and their classification. Section 3 discusses the characteristics of electroencephalography (EEG) and the different frequency bands that can be found in it. Section 4 presents the different prompts that have been studied in literature; while Sections 5, 6, and 7 discuss about the pre-processing, feature extraction and classification techniques, respectively. Section 8 offers a summary of the reviewed works and techniques. Finally, Section 9 presents the findings of this work and proposes future directions for the improvement of imagined speech recognition.

2. BRAIN COMPUTER INTERFACE

The advent of Future Internet has caused a widespread connectivity between everyday electronic devices and the human body (Zhang et al., 2018). One example is Brain Computer Interface, which is a technology that uses brain activity and signals to create a communication channel between external electronic devices and the human brain (Abiri et al., 2019). BCI has been used for several applications in various areas, as shown in **Figure 1**. For example, BCI systems have been applied toward neuromarketing, security, entertainment, smart-environment control, emotional education, among others (Abdulkader et al., 2015; Abo-Zahhad et al., 2015; Aricò et al., 2018; Padfield et al., 2019; Mudgal et al., 2020; Suhaimi et al., 2020; Moctezuma and Molinas, 2022). One of the most explored applications of BCI is toward the medical area to treat and diagnose neurological disorders such as epilepsy, depression, dementia, Alzheimer's, brain stroke, among others (Subasi, 2007; Morooka et al., 2018; Saad Zaghoul and Bayoumi, 2019; Hashimoto et al., 2020; Rajagopal et al., 2020; Sani et al., 2021). Moreover, it has also been used to recognize and classify emotions (Kaur et al., 2018; Suhaimi et al., 2020) and sleep stages (Chen et al., 2018), as well as to bring the opportunity of performing normal movements to people with motor disabilities (Antelis et al., 2018; Attallah et al., 2020; Al-Saegh et al., 2021; Mattioli et al., 2022). Furthermore, one of the most interesting, yet difficult, tasks that are being tried to be accomplished using BCI is imagined speech recognition, in which the objective is to convert the input brain signal to text, sound, or control commands. Different types of BCI systems have



been proposed by researchers to be able to use them in real-life scenarios. Some of the most important BCI classifications are: synchronous vs. asynchronous, online vs. offline, exogenous vs. endogenous, invasive vs. non-invasive (Portillo-Lara et al., 2021).

Synchronous BCI are systems that cannot be used freely by the users because they are fixed in determined periods of time. This means that, for imagined speech decoding, the user needs a cue that indicates when to begin the imagination process. Then, the selected time window is analyzed, discarding any other EEG signals that do not belong to that time constraint. On the other hand, asynchronous BCI can be used without any time constraint and they do not need any cue, meaning that it is a more natural process that can be more practical toward real-life applications. However, these systems have shown less accuracy than synchronous ones because of the difficulty on distinguishing intentional mental activity from unintentional one (Han et al., 2020).

Among BCI classification, there are also online and offline systems. Online BCI, just as asynchronous BCI, are promising toward real-life applications because they allow real-time data processing. In other words, during an online setting, the feature extraction and classification processes are done several times during each trial. However, because of this same advantage, the computational complexity that an online system can employ is limited. On the other hand, offline systems do not have this problem as they can use as much computational resources as needed because the feature extraction and classification processes are done until all trails are available and the sessions are over. Nevertheless, because of this same reason, an offline BCI system will be hardly applied under real-life circumstances (Chevallier et al., 2018).

Depending of the type of stimulus that the BCI uses, there can be exogenous and endogenous systems. Exogenous ones use external stimulus to generate the desired neural activation; while

endogenous ones can operate independently of any stimulus. For a real-life application of imagined speech decoding, the most appropriate between these two systems would be the endogenous BCI (Lee et al., 2021a).

Brain computer interfaces can also be classified as invasive and non-invasive. The invasive techniques, despite offering the best representation of the brain signals, have the risk of scaring brain tissue, at the same time that are more costly and difficult to use. On the other hand, non-invasive techniques, such as EEG, are used through scanning sensors or electrodes fixed on the scalp to record the brain signals. Due to its easiness to use, its portability and its safety, EEG based BCI have been broadly explored to be applied toward imagined speech recognition.

3. ELECTROENCEPHALOGRAPHY (EEG)

Electroencephalography, also known as EEG, is the most common non-invasive method to measure the electrical activity of the human brain. The signals are acquired by electrodes placed over the scalp that record the voltage difference generated during neural communication (Singh and Gumaste, 2021). The electrodes are then connected to an amplifier and are typically distributed in a standard 10–20 placement (Sazgar and Young, 2019). Commonly, EEG systems consist of 14–64 electrodes (also called channels), thus creating a multi-dimensional signal.

Along with its easiness to use and safety, EEG also has a high temporal resolution, characteristics that make it the most suitable option for imagined speech recognition. The reason behind this is that the analysis of imagined speech signals requires to track how the signal changes over time. However, one of the main disadvantages of EEG is that it can be easily contaminated by surrounding noise caused by external electronic devices. Hence, before being able to analyze EEG waves for imagined speech

tasks, they must be pre-processed to enhance the most important information within the signal.

3.1. EEG Waves

EEG waves consist of a mixture of diverse base frequencies. These frequencies have been arranged on five different frequency bands: gamma (>35 Hz), beta (12–35 Hz), alpha (8–12 Hz), theta (4–8 Hz), and delta (0.5–4 Hz) (Abhang et al., 2016). Each frequency band represents a determined cognitive state of the brain. Each of these frequency bands plays a determined role at the different stages of speech processing. Thus, recognizing them may aid to better analyze the EEG signal.

- **Gamma waves.** Changes in high gamma frequency (70–150 Hz) are associated with overt and covert speech. According to Pei et al. (2011), during overt speech the temporal lobe, Broca's area, Wernicke's area, premotor cortex and primary motor cortex present high gamma changes. On the other hand, this study also presents evidence of high gamma changes during covert speech in the supramarginal gyrus and superior temporal lobe.
- **Beta waves.** These waves are often related with muscle movement and feedback. Therefore, it can be considered that they are involved during auditory tasks and speech production (Bowers et al., 2013).
- **Alpha waves.** During language processing, these waves are involved in auditory feedback and speech perception. Moreover, alpha frequency during covert speech has been identified as weak in comparison to its behavior during overt speech (Jenson et al., 2014).
- **Theta waves.** According to Kösem and Van Wassenhove (2017), these waves become active during the phonemic restoration, and processing of co-articulation cues to compose words. Also, another study (Ten Oever and Sack, 2015), identified that theta waves can help to identify consonants in syllables.
- **Delta waves.** Intonation and rhythm during speech perception have been found to fall into frequency ranges that belong to the lower delta oscillation band (Schroeder et al., 2008). Also, diverse studies have found other speech processes in which delta waves are involved, such as prosodic phrasing, syllable structure, long syllables, among others (Peelle et al., 2013; Ghitza, 2017; Molinaro and Lizarazu, 2018; Boucher et al., 2019).

4. IMAGINED SPEECH PROMPTS IN LITERATURE

As said in Section 2, the main objective of applying BCI toward imagined speech decoding is to offer a new communication channel to people who are not able to speak due to any given motor disability. However, as language can be decomposed in several parts, as syllables, phonemes, vocals, and words, several studies have been carried on in order to classify these different parts of language.

In D'Zmura et al. (2009), Brigham and Kumar (2010), and Deng et al. (2010), volunteers imagined two syllables, /ba/ and

/ku/. For these studies, the volunteers were given an auditory cue indicating the syllable to be imagined. Another study done by Callan et al. (2000) focused on the imagined speech process of /a/, /i/, and /u/ vowels during a mental rehearsing process after speaking them out loud. DaSalla et al. (2009) also studied /a/, and /u/ vowels using a visual cue for both of them. Those vowels were chosen because of them causing similar muscle activation during real speech production. Also, in a study done by Zhao and Rudzicz (2015) seven phonetic/syllabic prompts were classified during a covert speech production process. In more recent works (Jahangiri et al., 2018, 2019) four phonemic structures (/ba/, /fo/, /le/, and /ry/) were analyzed. The difference between these studies was that in Jahangiri et al. (2018) they used a visual cue, while in Jahangiri et al. (2019) it was an auditory one. Some other studies such as Cooney et al. (2019), Tamm et al. (2020), and Ghane and Hossain (2020) have analyzed EEG signals produced during the imagined speech process of five vowels: /a/, /e/, /i/, /o/, and /u/. Besides phonemes, vowels, and syllables, there have been other studies that have worked with imagined words. For example, Wang et al. (2013) studied the classification of two imagined Chinese characters, whose meanings were “left” and “one.” In González-Castañeda et al. (2017), a study was done to classify five different imagined words: “up,” “down,” “left,” “right,” and “select.” Very similarly, the work done in Pawar and Dhage (2020) worked over the same prompts, with exception of the word “select.” Also, in the study done by Mohanchandra and Saha (2016), they used as prompts five words, being them, namely “water,” “help,” “thanks,” “food,” and “stop.” In Zhao and Rudzicz (2015), apart from the phonetic classification, they also worked toward the classification of the imagined words “pat,” “pot,” “knew,” and “gnaw”; where “pat”/“pot” and “knew”/“gnaw” are phonetically similar. Furthermore, in Nguyen et al. (2017) two different groups of imagined words (short and long) were analyzed. The former consisted on the words “in,” “out,” and “up,” while the latter consisted on “cooperate” and “independent.”

5. PRE-PROCESSING TECHNIQUES IN LITERATURE

As mentioned previously, EEG signals can be easily contaminated by external noise coming from electrical devices and artifacts such as eye blinks, breathing, etc. In order to diminish the noise and increase the SNR of the EEG waves, several pre-process techniques have been proposed in literature. Moreover, pre-processing is important because it can help to reduce the computational complexity of the problem and, therefore, to improve the efficiency of the classifier (Saminu et al., 2021). Generally speaking, pre-processing of EEG signals is usually formed by downsampling, band-pass, filtering, and windowing (Roy et al., 2019). However, the steps may vary depending on the situation and the data quality. For example, in Hefron et al. (2018) the pre-processing consisted on trimming the trials, downsampling them to 512 Hz and 64 channels to reduce the complexity of the problem. Also, a high-pass filter was applied to the data, at the same time that the PREP (an standardized early-stage EEG processing) pipeline was used to

calculate an average reference and remove line noise. On the other hand, the work carried in Stober et al. (2015) only applied a single pre-processing step of channel rejection. In the works done by Saha et al. (2019a,b) they used channel cross-covariance (CCV) for pre-processing; while in Cooney et al. (2019) they employed independent component analysis (ICA). Common average reference (CAR) method has also been employed to improve SNR from EEG signals by removing information that is present in all electrodes simultaneously (Moctezuma et al., 2019). Moreover, several studies have used temporal filtering as pre-process technique to focus on specific frequencies among the EEG signals (Jahangiri et al., 2018; Koizumi et al., 2018; Jahangiri and Sepulveda, 2019; Pawar and Dhage, 2020). Another preprocessing technique that has been applied is Laplacian filter (Zhao and Rudzicz, 2015), which is a spatial filter. However, this type of filters is not commonly used because it can lead to loss of important EEG information. In fact, most of pre-processing techniques can lead to loss of information, besides requiring an extra computational cost. Therefore, end-to-end learning methods that require minimum pre-processing are of currently of interest in EEG classification. However, classifying almost raw EEG signals is not an easy task and requires further study (Lee et al., 2020).

6. FEATURE EXTRACTION TECHNIQUES IN LITERATURE

During feature extraction, the main objective is to obtain the most relevant and significant information that will aid to correctly classify the neural signals. This process can be carried on the time domain, frequency domain, and spatial domain. In the time domain, the feature extraction process is often done through statistical analysis, obtaining statistical features such as standard deviation (SD), root mean square (RMS), mean, variance, sum, maximum, minimum, Hjorth parameters, sample entropy, autoregressive (AR) coefficients, among others (Riaz et al., 2014; Iqbal et al., 2016; AlSaleh et al., 2018; Cooney et al., 2018; Paul et al., 2018; Lee et al., 2019). On the other hand, the most common methods used to extract features from the frequency domain include Mel Frequency Cepstral Coefficients (MFCC), Short-Time Fourier transform (STFT), Fast Fourier Transform (FFT), Wavelet Transform (WT), Discrete Wavelet Transform (DWT), and Continuous Wavelet Transform (CWT) (Riaz et al., 2014; Salinas, 2017; Cooney et al., 2018; García-Salinas et al., 2018; Panachakel et al., 2019; Pan et al., 2021). Additionally, there is a method called Bag-of-Features (BoF) proposed by Lin et al. (2012), in which a time-frequency analysis is done to convert the signal into words using Sumbolic Arregate appRoXimation (SAX). In the case of spatial domain analysis, the most common method used in several works is Common Spatial Patterns (CSP) (Brigham and Kumar, 2010; Riaz et al., 2014; Arjestaan et al., 2016; AlSaleh et al., 2018; Lee et al., 2019; Panachakel et al., 2020). Moreover, it is important to mention that these feature extraction methods can be done in two different ways: from individual channels and simultaneously

from multiple channels. Despite individual channel analysis being easier, extracting features from diverse channels at the same time is more useful because it helps to analyze how information is transferred between the different areas of the brain. In order to do a simultaneous feature extraction, the most common method is the channel cross-covariance (CCV) matrix; in which the features of each channel are fused together to enhance the statistical relationship between the different electrodes (Nguyen et al., 2017; Saha and Fels, 2019; Singh and Gumaste, 2021). In fact, Riemannian geometry is an advanced feature extraction technique that has been used to manipulate covariance matrices. It has been successfully applied toward several applications, such as motor imagery, sleep/respiratory states classification, EEG decoding, etc. (Barachant et al., 2010, 2011; Navarro-Sune et al., 2016; Yger et al., 2016; Chu et al., 2020).

7. CLASSIFICATION TECHNIQUES IN LITERATURE

In order to classify the features extracted from the EEG signal, researchers have used both classical machine learning and deep learning algorithms. Both of them are methods that provide computers the capacity of learning and recognizing patterns. In the case of BCI, the patterns to be recognized are the features extracted from the EEG waves, and then, based on what the computer learnt, some predictions are made in order to classify the signals. Several classical machine learning techniques have been used to approach imagined speech decoding for EEG-based BCI systems. Some of the most common algorithms include Linear Discriminant Analysis (LDA) (Chi et al., 2011; Song and Sepulveda, 2014; Lee et al., 2021b), Support Vector Machines (SVM) (DaSalla et al., 2009; García et al., 2012; Kim et al., 2013; Riaz et al., 2014; Sarmiento et al., 2014; Zhao and Rudzicz, 2015; Arjestaan et al., 2016; González-Castañeda et al., 2017; Hashim et al., 2017; Cooney et al., 2018; Moctezuma and Molinas, 2018; Agarwal and Kumar, 2021), Random Forests (RF) (González-Castañeda et al., 2017; Moctezuma and Molinas, 2018; Moctezuma et al., 2019), k-Nearest-Neighbors (kNN) (Riaz et al., 2014; Bakhshali et al., 2020; Agarwal and Kumar, 2021; Rao, 2021; Dash et al., 2022), Naive Bayes (Dash et al., 2020a; Agarwal and Kumar, 2021; Iliopoulos and Papatotiriou, 2021; Lee et al., 2021b), and Relevance Vector Machines (RVM) (Liang et al., 2006; Matsumoto and Hori, 2014). Furthermore, deep learning approaches have recently taken a huge role for imagined speech recognition. Some of these techniques are Deep Neural Networks (DBN) (Lee and Sim, 2015; Chengaiyan et al., 2020), Correlation Networks (CorrNet) (Sharon and Murthy, 2020), Standardization-Refinement Domain Adaptation (SRDA) (Jiménez-Guarneros and Gómez-Gil, 2021), Extreme Learning Machine (ELM) (Pawar and Dhage, 2020), Convolutional Neural Networks (CNN) (Cooney et al., 2019, 2020; Tamm et al., 2020), Recurrent Neural Networks (RNN) (Chengaiyan et al., 2020), and parallel CNN+RNN with and without autoencoders (Saha and Fels, 2019; Saha et al., 2019a,b; Kumar and Scheme, 2021).

TABLE 1 | Imagined speech classification methods summary.

References	Task	Methods	Brain area and waves	Performance
D'Zmura et al. (2009)	Binary classification /ba/ and /ku/ syllables	- Hilbert transform - Matched filters	- All with exception of 18 electrodes most sensitive to electromyographic artifact -Alpha, beta, and theta waves	61% Average accuracy
DaSalla et al. (2009)	Binary classification between /a/, /u/ and rest state	- CSP	- All areas	71% Average accuracy
Brigham and Kumar (2010)	Binary classification /ba/ and /ku/ syllables	- SVM - ICA - AR model - KNN	- 1–45 Hz bandpass - All with exception of 18 electrodes most sensitive to electromyographic artifact - 4–25 Hz bandpass	68% Average accuracy
Deng et al. (2010)	Binary classification /ba/ and /ku/ syllables	- SOBI algorithm - Hilbert spectrum - FFT / STFT - LDA	- All areas - 3–20 Hz bandpass	67% Average accuracy
Chi et al. (2011)	Binary classification of five phoneme classes	- Naive Bayes - LDA - Spectrogram	- All areas omitting occipital and far frontal positions - 4–28 Hz bandpass	72% Average accuracy

TABLE 2 | Imagined speech classification methods summary (continuation).

References	Task	Methods	Brain area and waves	Performance
García et al. (2012)	Multi-class classification of five words	- Naive Bayes - SVM - Random Forests	- Wernicke's area - -25 Hz bandpass	26% Average accuracy
Matsumoto and Hori (2014)	Binary classification of /a/, /e/, /i/, /o/, and /u/ Japanese vowels	- CSP - RVM	- All areas - 0.1–300 Hz bandpass	- 79% RVM Average accuracy - 77% SVM Average accuracy
Sarmiento et al. (2014)	Binary classification of /a/, /e/, /i/, /o/, and /u/	- SVM - ANOVA - Power spectral density - SVM	- Broca's area and Wernicke's area - 2–50 Hz bandpass	- 79% RVM Average accuracy
Riaz et al. (2014)	Binary classification of /a/, /e/, /i/, /o/, and /u/	- AR coefficients - Hidden Markov Model - KNN / SVM - CSP / MFCC / LDA	- All areas -Alpha and beta waves	75% Average accuracy
Zhao and Rudzicz (2015)	Binary classification for presence of C/V, ± Nasal, ± Bilab, ± /uw/, ± /iy/	- DBN - SVM	- T7, FT8, FC6, C5, C3, CP3, C4, CP5, CP1, P3 - 1–50 Hz bandpass	C/V: 18% Accuracy ± Nasal: 63.5% Accuracy ± Bilab: 56.6% Accuracy ± /uw/: 59.6% Accuracy ± /iy/: 79.1% Accuracy

8. DISCUSSION, APPLICATIONS, AND LIMITATIONS OF PREVIOUS RESEARCH

Based on the previous sections and the diverse works mentioned in them, imagined speech classification can be summed up as in **Tables 1–6**.

As observed in the previous tables, there have been different attempts to achieve a good performance of imagined speech

recognition using EEG-based BCI. These attempts involve diverse feature extraction and classification methods. Therefore, in **Tables 7, 8** we offer a summary of the advantages and disadvantages of some of these methods.

The main objective of most imagined speech decoding BCI is to provide a new communication channel for those who have partial or total movement impairment (Rezazadeh Sereshkeh et al., 2019). Nevertheless, besides speech restoration, there are

TABLE 3 | Imagined speech classification methods summary (continuation).

References	Task	Methods	Brain area and waves	Performance
Arjestan et al. (2016)	Binary classification vowels, syllables, and resting state	- CSP - SVM	- All areas - 8–45 Hz bandpass	Vowels: 76.6% best accuracy Syllables: 76.4% best accuracy
Sereshkeh et al. (2017)	Binary classification of “yes,” “no,” and rest state	- Multilayer perceptron - DWT - RMS / SD	- All areas - 0–50 Hz bandpass	70% Average accuracy
Nguyen et al. (2017)	Classification of vowels, short, and long words	- Riemannian Manifold - RVM - CSP / WT	- Broca’s area, the motor cortex and Wernicke’s area - 8–70 Hz bandpass	- Vowels: 49% Average accuracy - Short words: 50.1% Average accuracy - Long words: 66.2% Average accuracy - S-L: 80.1% Average accuracy
Paul et al. (2018)	Classification of three Hindi vowels	- SVM - AR coefficients - Hjorth parameters - Sample entropy	- Broca’s and Wernicke’s area - 0.1–36 Hz bandpass	63% Average accuracy
Cooney et al. (2018)	Classification of seven phonemic prompts and four words	- SVM - Decision Tree - MFCC - ICA	- All areas - 1–50 Hz bandpass	- Phonemes: 20% Average accuracy - Pair of words: 44% Average accuracy

some other novel applications of imagined speech decoding that have been explored. In Kim et al. (2020), researchers proposed a BCI paradigm that combined event-related potentials and imagined speech to target individual objects in a smart home environment. This was done through EEG analysis and classification using regularized linear discriminant analysis (RLDA). Moreover, the work presented in Asghari Bejestani et al. (2022) focused on the classification of six Persian words through imagined speech decoding. These words, as said by the authors, can be used to control electronic devices such as a wheelchair or to fill a simple questionnaire form. Tøttrup et al. (2019) explored the possibility of combining motor imagery and imagined speech recognition for controlling an external device through EEG-based BCI and random forest algorithm. Furthermore, the work presented by Moctezuma and Molinas (2018) explored the application of imagined speech decoding toward subject identification using SVM.

Regardless of the rising interest on EEG-based BCI for imagined speech recognition, the development of systems that are useful for real-life applications is still in its infancy. In the case of syllables, vowels, and phonemes, the limited amount of vocabulary that has been analyzed impedes the possibility of applying BCI to allow people to speak through their thoughts. Among all the reviewed proposals, the one that seems closer to be applied in real life is the classification of words such as “up,” “down,” “left,” “right,” “forward,” “backward,” and “select.” The reason behind this is that those words can be used to control external devices such as a computer/cellphone screen and robotic

prosthesis. However, the fact of those words being classified by EEG-based BCI systems that are offline and synchronous makes the projects less scalable to real-life applications.

Also, it is important to mention that EEG-based BCI lacks from accuracy when compared with other methods such as ECoG and MEG. ECoG has been applied in several studies for either covert and overt speech decoding, achieving higher average accuracies than EEG-based BCI. For example, in Martin et al. (2016) imagined speech pairwise classification reached an accuracy of 88.3% through ECoG recording. Kanas et al. (2014) presented a spatio-spectral feature clustering of ECoG recordings for syllable classification, obtaining an accuracy of 98.8%. Also, a work performed by Zhang et al. (2012) obtained a 77.5% accuracy on the classification of eight-character Chinese spoken sentences through the analysis of ECoG recordings. Moreover, in the work presented by Dash et al. (2019) MEG was used for phrase classification, achieving a top accuracy of 95%. Finally, the study in Dash et al. (2020a) aimed to classify articulated and imagined speech on healthy and amyotrophic lateral sclerosis (ALS) patients. In this work the best articulation decoding accuracy for ALS patients was 87.78%, while for imagined decoding was 74.57%.

In summary, the past research allowed to observe the following current limitations of EEG-based BCI systems for imagined speech recognition:

- Limited vocabulary: Most of the reviewed studies focused on imagined vowels (/a/, /e/, /i/, /o/, /u/, /ba/, /ku/) and words

TABLE 4 | Imagined speech classification methods summary (continuation).

References	Task	Methods	Brain area and waves	Performance
Moctezuma and Molinas (2018)	Classification of “up,” “down,” “right,” “left,” “select”	- RF	- All areas	- RF: 64% Average accuracy
		- SVM	- 1–50 Hz bandpass	- SVM: 84% Average accuracy
		- Naive Bayes		- Naive Bayes: 68% Average accuracy
		- KNN		- KNN: 78% Average accuracy
		- EMD/IMF		
Saha and Fels (2019)	Classification of vowels, short and long words	- CCV	- All areas	- Vowels: 72% Average accuracy
		- CNN + LSTM + DAE		- Short words: 77% Average accuracy - Long words: 79% Average accuracy
Saha et al. (2019a)	Binary classification for presence of C/V, ± Nasal, ± Bilab, ± /uw/, ± /iy/	- CCV - CNN + TCNN + DAE	- All areas	- C/V: 85% Accuracy ± Nasal: 73% Accuracy ± Bilab: 75% Accuracy ± /uw/: 82% Accuracy ± /iy/: 73% Accuracy ± Multiclass: 28% Accuracy
Panachakel et al. (2019)	Classification of seven phonemic prompts and four words	- DNN - DWT	- C4, FC3, FC1, F5, C3, F7, FT7, CZ, P3, T7, C5 - 1–50 Hz bandpass	- 57% Average accuracy
Moctezuma et al. (2019)	Classification of “up,” “down,” “right,” “left,” “select”	- RF - CAR - DWT / Statistics	- All areas - 0–64 Hz bandpass	- 93% Average accuracy

such as “right,” “left,” “up,” and “down.” This shows how far away we are from truly decode enough vocabulary for a real-life application of covert speech decoding.

- Limited accuracy: Despite some works reaching +80% accuracy, this was achieved mostly for binary classification. Multi-class classification, which would be more viable for real-life application, demonstrated to have much lower classification rates than binary tasks. It is important to notice that even binary accuracy decreases or increases depending on the nature of the task to be done (for example: long vs. short words compared to words of the same length).
- Mental repetition of the prompt: The experimental design of most studies included the repeated imagination of the vowel, phoneme or word. This helps increasing the accuracy of the algorithm; however, mental repetition is not included on daily conversation tasks. Therefore, the design of some proposed experiments have low reliability when considering their practical application.
- Acquisition system: Most of the reviewed works used a high-density EEG system, which may be difficult to apply in real-life situations. Also, almost no work reviewed in here deals with an online and asynchronous BCI system, which, as mentioned earlier, is the feasible BCI option for practical applications.

9. CONCLUSIONS AND FUTURE WORK

The rapid development of the Future Internet framework has led to several new applications such as smart environments, autonomous monitoring of medical health, cloud computing, etc. (Zhang et al., 2019). Moreover, there are important future plans, such as Internet Plus and Industry 4.0, that require further integration of internet with other areas, such as medicine and economics. Therefore, technologies such as Brain Computer Interfaces seem to be promising areas to be explored and implemented to solve real-life problems.

Through this review, we analyzed works that involved EEG-based BCI systems directed toward imagined speech recognition. These works followed the decoding of imagined syllables, phonemes, vowels, and words. However, the study of each of those groups was individual, meaning that there was no work aiming to study vowels vs. words, phonemes vs. words, phonemes vs. vowels, etc. at the same time. Also, it is important to notice that each BCI was used for a single person, which would make difficult the implementation of a general and globalized system. It seems that each individual would need to train their own BCI system in order to use it successfully.

TABLE 5 | Imagined speech classification methods summary (continuation).

References	Task	Methods	Brain area and waves	Performance
Cooney et al. (2019)	Classification of /a/, /e/, /i/, /o/, and /u/	- CNN - Transfer learning	- All areas - 2–40 Hz bandpass	- 34% Average accuracy
Sharon and Murthy (2020)	Binary classification for presence of C/V, ± Nasal, ± Bilab, ± /uw/, ± /iy/	- CorrNet	- All areas - 1-50 Hz bandpass	C/V: 89% Accuracy ± Nasal: 76% Accuracy ± Bilab: 75% Accuracy ± /uw/: 82% Accuracy ± /iy/: 80% Accuracy
Bakhshali et al. (2020)	Binary classification for presence of C/V, ± Nasal, ± Bilab, ± /uw/, ± /iy/ Binary Classification of /pat/, /pot/, /gnaw/, and /knew/	- Riemannian distance - Correntropy Spectral Density	- Broca's area and Wernicke's area - 1–50 Hz bandpass	C/V: 86% Accuracy ± Nasal: 72% Accuracy ± Bilab: 69% Accuracy ± /uw/: 84% Accuracy ± /iy/: 75% Accuracy ± Word binary classification: 69% Average accuracy
Pawar and Dhage (2020)	Muti-class and binary classification of "left," "right," "up," and "down"	- Kernel ELM - Statistical features - DWT - ICA	- All areas - Prefrontal cortex, Wernicke's area, right inferior frontal gyrus, Broca's area - Prefrontal cortex, Wernicke's area, right inferior frontal gyrus, Broca's area, primary motor cortex - 0.5–128 Hz bandpass	Multi-class: 49% best accuracy Binary: 85% best accuracy

TABLE 6 | Imagined speech classification methods summary (continuation).

References	Task	Methods	Brain area and waves	Performance
Cooney et al. (2020)	- Classification of /a/, /e/, /i/, /o/, and /u/ - Classification of "left," "right," "up," "down," "forward," "backward" (in Spanish)	- CNN - ICA/LDA	- All areas - 2–40 Hz bandpass	- Vowels: 35% best accuracy - Words: 30% best accuracy
Tamm et al. (2020)	Classification of five vowels and six words	- CNN - Transfer learning	- F3, F4, C3, C4, P3, P4	- 24% Average accuracy
Chengaiyan et al. (2020)	Vowel classification	- RNN - DBN	- All areas - 2–40 Hz bandpass	- RNN: 72% Average accuracy - DBN: 80% Average accuracy
Jiménez-Guarneros and Gómez-Gil (2021)	Short and long words classification	- SRDA	- All areas	- Short: 61% Average accuracy - Long: 63% Average accuracy

Another thing to take into account is that several languages have been analyzed, such as English, Spanish, Chinese, and Hindi. However, there is not a comprehensive study that evaluates the impact of how a method performs toward an specific language.

Regarding feature extraction methods, there have been a large amount of proposed techniques such as DWT, MFCC, STFT, CSP, Riemannian space, etc. On the other hand, the most studied classification algorithm has been SVM, which is a classical

machine learning technique. Deep learning techniques such as CNN and RNN have also been explored by some authors. Despite deep learning showing promising accuracy improvements in comparison to classical ML, it is difficult to fully exploit it because of the limited amount of data available to train DL algorithms.

Additionally, currently there is not definitive information regarding the most important EEG recording locations of imagined speech recognition. Broca's and Wernicke's areas are

TABLE 7 | Comparison of feature extraction methods.

Method	Advantages	Disadvantages
AR	<ul style="list-style-type: none"> - Good frequency resolution - Limited spectral loss 	<ul style="list-style-type: none"> - Low performance when applied to non-stationary signals - The order of the model is difficult to select
FFT	<ul style="list-style-type: none"> - Good performance when applied to stationary signals - Appropriate for narrowband signals - Good speed for real-time applications 	<ul style="list-style-type: none"> - High noise sensitivity - Poor performance on non-stationary signals - Weak spectral estimation
WT	<ul style="list-style-type: none"> - Varying window size to analyze several frequencies - Good to analyze transient signal changes 	<ul style="list-style-type: none"> - Proper mother wavelet selection is not trivial

TABLE 8 | Comparison of classification methods.

Method	Advantages	Disadvantages
KNN	<ul style="list-style-type: none"> - Easy to understand and implement 	<ul style="list-style-type: none"> - Large storage capacity is needed - Error susceptibility and sensitivity to irrelevant features
SVM	<ul style="list-style-type: none"> - Effective in high dimensional spaces - Relatively low storage capacity is needed 	<ul style="list-style-type: none"> - Poor performance on noisy and large datasets
LDA	<ul style="list-style-type: none"> - Simple to understand and use 	<ul style="list-style-type: none"> - Requires a linear model
ANN	<ul style="list-style-type: none"> - Relatively high accuracy - Flexible and adaptable structure - Handles multidimensional data 	<ul style="list-style-type: none"> - Requires large datasets for it to be trained - High computational cost - Performance depends on several parameters, such as number of neurons and hidden layers
DL algorithms	<ul style="list-style-type: none"> - Robustness for adaptation - It can be adapted to different problems through transfer learning - Features can be automatically deduced and tuned 	<ul style="list-style-type: none"> - Requires large datasets for it to perform well - Need high computational resources - Difficult to implement for novices

well-known to be involved in speech production; however, some studies reviewed here showed that they are not the only zones that contain valuable information for covert speech decoding. Therefore, it seems a good idea to propose a method that helps selecting the EEG channels that better characterize a given task.

All things considered, we identified the following tasks as promising for the future development of EEG-based BCI systems for imagined speech decoding:

- Broaden the existing datasets in such a way that deep learning techniques could be applied to their full extent. Moreover, explore and propose prompts that could be more easily applied to solve real-life problems.
- Find and propose more varied prompts in order to enhance the difference between their EEG signatures and detect the most discriminative characteristic to enhance classification. This can be done by employing different rhythms, tones, overall structure, and language.
- Explore how a same proposed method performs over different languages.
- Recognize the best feature extraction and machine learning techniques to improve classification accuracy.

At the same time, there is still room for improvement in the identification of EEG frequency range that offers the most valuable information.

- Most of the current studies are offline-synchronous BCI systems applied in healthy subjects. Also, most experiments are highly controlled in order to avoid artifacts. Therefore, there is room for further work in these areas.
- Explore different imagery processes, such as Visual Imagery (Ullah and Halim, 2021).

AUTHOR CONTRIBUTIONS

DL-B: formal analysis, investigation, methodology, and writing—original draft. PP and AM: resources. DB, PP, and AM: supervision, validation, and writing—review and editing. All authors have read and agreed to the published version of the manuscript.

FUNDING

This work was funded by Fondo para el financiamiento para la publicación de Artículos Científicos of Monterrey Institute of Technology and Higher Education.

REFERENCES

- Abdulkader, S., Atia, A., and Mostafa, M. (2015). Brain computer interfacing: applications and challenges. *Egypt. Inform. J.* 16, 213–230. doi: 10.1016/j.eij.2015.06.002
- Abhang, P. A., Gawali, B., and Mehrotra, S. C. (2016). *Introduction to EEG and Speech-Based Emotion Recognition*. India: Academic Press. doi: 10.1016/B978-0-12-804490-2.00007-5
- Abiri, R., Borhani, S., Sellers, E. W., Jiang, Y., and Zhao, X. (2019). A comprehensive review of eeg-based brain-computer interface paradigms. *J. Neural Eng.* 16:011001. doi: 10.1088/1741-2552/aaf12e
- Abo-Zahhad, M., Ahmed, S. M., and Abbas, S. N. (2015). State-of-the-art methods and future perspectives for personal recognition based on electroencephalogram signals. *IET Biometr.* 4, 179–190. doi: 10.1049/iet-bmt.2014.0040
- Agarwal, P., and Kumar, S. (2021). *Transforming Imagined Thoughts Into Speech Using a Covariance-Based Subset Selection Method*. NISCAIR-CSIR.
- Al-Saegh, A., Dawwd, S. A., and Abdul-Jabbar, J. M. (2021). Deep learning for motor imagery EEG-based classification: a review. *Biomed. Signal Process. Control* 63:102172. doi: 10.1016/j.bspc.2020.102172
- AlSaleh, M., Moore, R., Christensen, H., and Arvaneh, M. (2018). “Discriminating between imagined speech and non-speech tasks using EEG,” in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, (USA: IEEE) 1952–1955. doi: 10.1109/EMBC.2018.8512681
- Angrick, M., Herff, C., Mugler, E., Tate, M. C., Slutzky, M. W., Krusienski, D. J., et al. (2019). Speech synthesis from ECOG using densely connected 3D convolutional neural networks. *J. Neural Eng.* 16:036019. doi: 10.1088/1741-2552/ab0c59
- Antelis, J. M., Gudiño-Mendoza, B., Falcón, L. E., Sanchez-Ante, G., and Sossa, H. (2018). Dendrite morphological neural networks for motor task recognition from electroencephalographic signals. *Biomed. Signal Process. Control* 44, 12–24. doi: 10.1016/j.bspc.2018.03.010
- Antoniades, A., Spyrou, L., Took, C. C., and Sanei, S. (2016). “Deep learning for epileptic intracranial EEG data,” in *2016 IEEE 26th International Workshop on Machine Learning for Signal Processing (MLSP)*, (Italy: IEEE) 1–6. doi: 10.1109/MLSP.2016.7738824
- Aricó, P., Borghini, G., Di Flumeri, G., Sciaraffa, N., and Babiloni, F. (2018). Passive BCI beyond the lab: current trends and future directions. *Physiol. Measure.* 39:08TR02. doi: 10.1088/1361-6579/aad57e
- Arjastan, M. A., Vali, M., and Farajdi, F. (2016). “Brain computer interface design and implementation to identify overt and covert speech,” in *2016 23rd Iranian Conference on Biomedical Engineering and 2016 1st International Iranian Conference on Biomedical Engineering (ICBME)*, (Iran) 59–63. doi: 10.1109/ICBME.2016.7890929
- Asghari Bejestani, M., Khani, M., Nafisi, V., Darakeh, F., et al. (2022). Eeg-based multiword imagined speech classification for Persian words. *BioMed Res. Int.* 2022:8333084. doi: 10.1155/2022/8333084
- Attallah, O., Abougharbia, J., Tamazin, M., and Nasser, A. A. (2020). A BCI system based on motor imagery for assisting people with motor deficiencies in the limbs. *Brain Sci.* 10:864. doi: 10.3390/brainsci10110864
- Bakhshali, M. A., Khademi, M., Ebrahimi-Moghadam, A., and Moghimi, S. (2020). EEG signal classification of imagined speech based on Riemannian distance of correntropy spectral density. *Biomed. Signal Process. Control* 59:101899. doi: 10.1016/j.bspc.2020.101899
- Barachant, A., Bonnet, S., Congedo, M., and Jutten, C. (2010). “Riemannian geometry applied to BCI classification,” in *International Conference on Latent Variable Analysis and Signal Separation* (Springer), (France: Springer) 629–636. doi: 10.1007/978-3-642-15995-4_78
- Barachant, A., Bonnet, S., Congedo, M., and Jutten, C. (2011). Multiclass brain-computer interface classification by Riemannian geometry. *IEEE Trans. Biomed. Eng.* 59, 920–928. doi: 10.1109/TBME.2011.2172210
- Boucher, V. J., Gilbert, A. C., and Jemel, B. (2019). The role of low-frequency neural oscillations in speech processing: revisiting delta entrainment. *J. Cogn. Neurosci.* 31, 1205–1215. doi: 10.1162/jocn_a_01410
- Bowers, A., Saltuklaroglu, T., Harkrider, A., and Cuellar, M. (2013). Suppression of the μ rhythm during speech and non-speech discrimination revealed by independent component analysis: implications for sensorimotor integration in speech processing. *PLoS ONE* 8:e72024. doi: 10.1371/journal.pone.0072024
- Bozhkov, L., and Georgieva, P. (2018). “Overview of deep learning architectures for EEG-based brain imaging,” in *2018 International Joint Conference on Neural Networks (IJCNN)*, (Brazil) 1–7. doi: 10.1109/IJCNN.2018.8489561
- Branco, M. P., Pels, E. G., Sars, R. H., Aarnoutse, E. J., Ramsey, N. F., Vansteensel, M. J., et al. (2021). Brain-computer interfaces for communication: preferences of individuals with locked-in syndrome. *Neurorehabil. Neural Repair* 35, 267–279. doi: 10.1177/1545968321989331
- Brigham, K., and Kumar, B. V. (2010). “Imagined speech classification with EEG signals for silent communication: a preliminary investigation into synthetic telepathy,” in *2010 4th International Conference on Bioinformatics and Biomedical Engineering*, (China) 1–4. doi: 10.1109/ICBBE.2010.5515807
- Callan, D. E., Callan, A. M., Honda, K., and Masaki, S. (2000). Single-sweep EEG analysis of neural processes underlying perception and production of vowels. *Cogn. Brain Res.* 10, 173–176. doi: 10.1016/S0926-6410(00)00025-2
- Chen, T., Huang, H., Pan, J., and Li, Y. (2018). “An EEG-based brain-computer interface for automatic sleep stage classification,” in *2018 13th IEEE Conference on Industrial Electronics and Applications (ICIEA)*, (China: IEEE) 1988–1991. doi: 10.1109/ICIEA.2018.8398035
- Chengaiyan, S., Retnapandian, A. S., and Anandan, K. (2020). Identification of vowels in consonant-vowel-consonant words from speech imagery based EEG signals. *Cogn. Neurodyn.* 14, 1–19. doi: 10.1007/s11571-019-09558-5
- Chevallier, S., Kalunga, E., Barthélemy, Q., and Yger, F. (2018). *Riemannian Classification for SSVEP Based BCI: Offline versus Online Implementations*. Versailles: HAL. doi: 10.1201/9781351231954-19
- Chi, X., Hagedorn, J. B., Schoonover, D., and D’Zmura, M. (2011). Eeg-based discrimination of imagined speech phonemes. *Int. J. Bioelectromagn.* 13, 201–206.
- Chopra, K., Gupta, K., and Lambora, A. (2019). “Future internet: the internet of things—a literature review,” in *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*, (India) 135–139. doi: 10.1109/COMITCon.2019.8862269
- Chu, Y., Zhao, X., Zou, Y., Xu, W., Song, G., Han, J., et al. (2020). Decoding multiclass motor imagery EEG from the same upper limb by combining Riemannian geometry features and partial least squares regression. *J. Neural Eng.* 17:046029. doi: 10.1088/1741-2552/aba7cd
- Cooney, C., Folli, R., and Coyle, D. (2018). “Mel frequency cepstral coefficients enhance imagined speech decoding accuracy from EEG,” in *2018 29th Irish Signals and Systems Conference (ISSC)*, (United Kingdom) 1–7. doi: 10.1109/ISSC.2018.8585291
- Cooney, C., Folli, R., and Coyle, D. (2019). “Optimizing layers improves CNN generalization and transfer learning for imagined speech decoding from EEG,” in *2019 IEEE International Conference on Systems, Man and Cybernetics (SMC)*, (Italy) 1311–1316. doi: 10.1109/SMC.2019.8914246
- Cooney, C., Korik, A., Folli, R., and Coyle, D. (2020). Evaluation of hyperparameter optimization in machine and deep learning methods for decoding imagined speech EEG. *Sensors* 20:4629. doi: 10.3390/s20164629
- DaSalla, C. S., Kambara, H., Sato, M., and Koike, Y. (2009). Single-trial classification of vowel speech imagery using common spatial patterns. *Neural Netw.* 22, 1334–1339. doi: 10.1016/j.neunet.2009.05.008
- Dash, D., Ferrari, P., Heitzman, D., and Wang, J. (2019). “Decoding speech from single trial MEG signals using convolutional neural networks and transfer learning,” in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, (Germany: IEEE) 5531–5535. doi: 10.1109/EMBC.2019.8857874
- Dash, D., Ferrari, P., Hernandez-Mulero, A. W., Heitzman, D., Austin, S. G., and Wang, J. (2020a). “Neural speech decoding for amyotrophic lateral sclerosis,” in *INTERSPEECH*, (China) 2782–2786. doi: 10.21437/Interspeech.2020-3071
- Dash, D., Ferrari, P., and Wang, J. (2020b). Decoding imagined and spoken phrases from non-invasive neural (MEG) signals. *Front. Neurosci.* 14:290. doi: 10.3389/fnins.2020.00290
- Dash, S., Tripathy, R. K., Panda, G., and Pachori, R. B. (2022). Automated recognition of imagined commands from EEG signals using multivariate fast and adaptive empirical mode decomposition based method. *IEEE Sensors Lett.* 6, 1–4. doi: 10.1109/LENS.2022.3142349

- Deng, S., Srinivasan, R., Lappas, T., and D'Zmura, M. (2010). EEG classification of imagined syllable rhythm using Hilbert spectrum methods. *J. Neural Eng.* 7:046006. doi: 10.1088/1741-2560/7/4/046006
- D'Zmura, M., Deng, S., Lappas, T., Thorpe, S., and Srinivasan, R. (2009). "Toward EEG sensing of imagined speech," in *International Conference on Human-Computer Interaction* (United States: Springer), 40–48. doi: 10.1007/978-3-642-02574-7_5
- Fonken, Y. M., Kam, J. W., and Knight, R. T. (2020). A differential role for human hippocampus in novelty and contextual processing: implications for p300. *Psychophysiology* 57:e13400. doi: 10.1111/psyp.13400
- García, A. A. T., García, C. A. R., and Pineda, L. V. (2012). "Toward a silent speech interface based on unspoken speech," in *Biosignals*, (Mexico) 370–373.
- García-Salinas, J. S., Villase nor-Pineda, L., Reyes-García, C. A., and Torres-García, A. (2018). "Tensor decomposition for imagined speech discrimination in EEG," in *Mexican International Conference on Artificial Intelligence* (Mexico: Springer), 239–249. doi: 10.1007/978-3-030-04497-8_20
- Ghane, P., and Hossain, G. (2020). Learning patterns in imaginary vowels for an intelligent brain computer interface (BCI) design. *arXiv preprint arXiv:2010.12066*. doi: 10.48550/arXiv.2010.12066
- Ghitza, O. (2017). Acoustic-driven delta rhythms as prosodic markers. *Lang. Cogn. Neurosci.* 32, 545–561. doi: 10.1080/23273798.2016.1232419
- González-Casta neda, E. F., Torres-García, A. A., Reyes-García, C. A., and Villase nor-Pineda, L. (2017). Sonification and textification: proposing methods for classifying unspoken words from EEG signals. *Biomed. Signal Process. Control* 37, 82–91. doi: 10.1016/j.bspc.2016.10.012
- Gu, X., Cao, Z., Jolfaei, A., Xu, P., Wu, D., Jung, T.-P., et al. (2021). EEG-based brain-computer interfaces (BCIs): a survey of recent studies on signal sensing technologies and computational intelligence approaches and their applications. *IEEE/ACM Trans. Comput. Biol. Bioinformatics* 18, 1645–1666. doi: 10.1109/TCBB.2021.3052811
- Haji, L. M., Ahmad, O. M., Zeebaree, S., Dino, H. I., Zebari, R. R., and Shukur, H. M. (2020). Impact of cloud computing and internet of things on the future internet. *Technol. Rep. Kansai Univ.* 62, 2179–2190.
- Han, C.-H., Müller, K.-R., and Hwang, H.-J. (2020). Brain-switches for asynchronous brain-computer interfaces: a systematic review. *Electronics* 9:422. doi: 10.3390/electronics9030422
- Hashim, N., Ali, A., and Mohd-Isa, W.-N. (2017). "Word-based classification of imagined speech using EEG," in *International Conference on Computational Science and Technology* (Malaysia: Springer), 195–204. doi: 10.1007/978-981-10-8276-4_19
- Hashimoto, Y., Kakui, T., Ushiba, J., Liu, M., Kamada, K., and Ota, T. (2020). Portable rehabilitation system with brain-computer interface for inpatients with acute and subacute stroke: a feasibility study. *Assist. Technol.* 1–9. doi: 10.1080/10400435.2020.1836067
- Hefron, R., Borghetti, B., Schubert Kabban, C., Christensen, J., and Estep, J. (2018). Cross-participant eeg-based assessment of cognitive workload using multi-path convolutional recurrent neural networks. *Sensors* 18:1339. doi: 10.3390/s18051339
- Herff, C., and Schultz, T. (2016). Automatic speech recognition from neural signals: a focused review. *Front. Neurosci.* 10:429. doi: 10.3389/fnins.2016.00429
- Iliopoulos, A., and Papatotiriou, I. (2021). Functional complex networks based on operational architectonics: application on EEG-BCI for imagined speech. *Neuroscience* 484, 98–118. doi: 10.1016/j.neuroscience.2021.11.045
- Iqbal, S., Shanir, P. M., Khan, Y. U., and Farooq, O. (2016). "Time domain analysis of EEG to classify imagined speech," in *Proceedings of the Second International Conference on Computer and Communication Technologies* (India: Springer), 793–800. doi: 10.1007/978-81-322-2523-2_77
- Jahangiri, A., Achancaray, D., and Sepulveda, F. (2019). "A novel EEG-based four-class linguistic BCI," in *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, (Germany: IEEE) 3050–3053. doi: 10.1109/EMBC.2019.8856644
- Jahangiri, A., Chau, J. M., Achancaray, D. R., and Sepulveda, F. (2018). "Covert speech vs. motor imagery: a comparative study of class separability in identical environments," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, (United States: IEEE) 2020–2023. doi: 10.1109/EMBC.2018.8512724
- Jahangiri, A., and Sepulveda, F. (2019). The relative contribution of high-gamma linguistic processing stages of word production, and motor imagery of articulation in class separability of covert speech tasks in EEG data. *J. Med. Syst.* 43, 1–9. doi: 10.1007/s10916-019-1379-1
- Jenson, D., Bowers, A. L., Harkrider, A. W., Thornton, D., Cuellar, M., and Saltuklaroglu, T. (2014). Temporal dynamics of sensorimotor integration in speech perception and production: independent component analysis of EEG data. *Front. Psychol.* 5:656. doi: 10.3389/fpsyg.2014.00656
- Jiménez-Guarneros, M., and Gómez-Gil, P. (2021). Standardization-refinement domain adaptation method for cross-subject EEG-based classification in imagined speech recognition. *Pattern Recogn. Lett.* 141, 54–60. doi: 10.1016/j.patrec.2020.11.013
- Kanas, V. G., Mporas, I., Benz, H. L., Sgarbas, K. N., Bezerianos, A., and Crone, N. E. (2014). Joint spatial-spectral feature space clustering for speech activity detection from ECG signals. *IEEE Trans. Biomed. Eng.* 61, 1241–1250. doi: 10.1109/TBME.2014.2298897
- Kaur, B., Singh, D., and Roy, P. P. (2018). EEG based emotion classification mechanism in BCI. *Proc. Comput. Sci.* 132, 752–758. doi: 10.1016/j.procs.2018.05.087
- Kim, H.-J., Lee, M.-H., and Lee, M. (2020). "A BCI based smart home system combined with event-related potentials and speech imagery task," in *2020 8th International Winter Conference on Brain-Computer Interface (BCI)*, (Korea) 1–6. doi: 10.1109/BCI48061.2020.9061634
- Kim, T., Lee, J., Choi, H., Lee, H., Kim, I.-Y., and Jang, D. P. (2013). "Meaning based covert speech classification for brain-computer interface based on electroencephalography," in *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)*, (United States: IEEE) 53–56. doi: 10.1109/NER.2013.6695869
- Koizumi, K., Ueda, K., and Nakao, M. (2018). "Development of a cognitive brain-machine interface based on a visual imagery method," in *2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, 1062–1065. (United States) doi: 10.1109/EMBC.2018.8512520
- Kösem, A., and Van Wassenhove, V. (2017). Distinct contributions of low- and high-frequency neural oscillations to speech comprehension. *Lang. Cogn. Neurosci.* 32, 536–544. doi: 10.1080/23273798.2016.1238495
- Kumar, P., and Scheme, E. (2021). "A deep spatio-temporal model for EEG-based imagined speech recognition," in *ICASSP 2021–2021 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, (Canada: IEEE) 995–999. doi: 10.1109/ICASSP39728.2021.9413989
- Kwon, M., Han, S., Kim, K., and Jun, S. C. (2019). Super-resolution for improving EEG spatial resolution using deep convolutional neural network-feasibility study. *Sensors* 19:5317. doi: 10.3390/s19235317
- Lee, D.-H., Jeong, J.-H., Ahn, H.-J., and Lee, S.-W. (2021a). "Design of an EEG-based drone swarm control system using endogenous BCI paradigms," in *2021 9th International Winter Conference on Brain-Computer Interface (BCI)*, (Korea) 1–5. doi: 10.1109/BCI51272.2021.9385356
- Lee, D.-H., Kim, S.-J., and Lee, K.-W. (2021b). Decoding high-level imagined speech using attention-based deep neural networks. *arXiv preprint arXiv:2112.06922*. doi: 10.1109/BCI53720.2022.9734310
- Lee, D.-Y., Lee, M., and Lee, S.-W. (2020). "Classification of imagined speech using siamese neural network," in *2020 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, (Canada: IEEE) 2979–2984. doi: 10.1109/SMC42975.2020.9282982
- Lee, S.-H., Lee, M., and Lee, S.-W. (2019). "EEG representations of spatial and temporal features in imagined speech and overt speech," in *Asian Conference on Pattern Recognition* (Korea: Springer), 387–400. doi: 10.1007/978-3-030-41299-9_30
- Lee, T.-J., and Sim, K.-B. (2015). Vowel classification of imagined speech in an electroencephalogram using the deep belief network. *J. Instit. Control Robot. Syst.* 21, 59–64. doi: 10.5302/J.ICROS.2015.14.0073
- Liang, N.-Y., Huang, G.-B., Saratchandran, P., and Sundararajan, N. (2006). A fast and accurate online sequential learning algorithm for feedforward networks. *IEEE Trans. Neural Netw.* 17, 1411–1423. doi: 10.1109/TNN.2006.880583
- Lin, J., Khade, R., and Li, Y. (2012). Rotation-invariant similarity in time series using bag-of-patterns representation. *J. Intell. Inform. Syst.* 39, 287–315. doi: 10.1007/s10844-012-0196-5
- Martin, S., Brunner, P., Iturrate, I., Millán, J. d. R., Schalk, G., Knight, R. T., et al. (2016). Word pair classification during imagined speech using direct brain recordings. *Sci. Rep.* 6, 1–12. doi: 10.1038/srep25803

- Matsumoto, M., and Hori, J. (2014). Classification of silent speech using support vector machine and relevance vector machine. *Appl. Soft Comput.* 20, 95–102. doi: 10.1016/j.asoc.2013.10.023
- Mattioli, F., Porcaro, C., and Baldassarre, G. (2022). A 1d cnn for high accuracy classification and transfer learning in motor imagery EEG-based brain-computer interface. *J. Neural Eng.* 18:066053. doi: 10.1088/1741-2552/ac4430
- Moctezuma, L. A., and Molinas, M. (2018). “EEG-based subjects identification based on biometrics of imagined speech using EMD,” in *International Conference on Brain Informatics* (Norway: Springer), 458–467. doi: 10.1007/978-3-030-05587-5_43
- Moctezuma, L. A., and Molinas, M. (2022). “EEG-based subject identification with multi-class classification,” in *Biosignal Processing and Classification Using Computational Learning and Intelligence* (Mexico: Elsevier), 293–306. doi: 10.1016/B978-0-12-820125-1.00027-0
- Moctezuma, L. A., Torres-García, A. A., Villase nor-Pineda, L., and Carrillo, M. (2019). Subjects identification using eeg-recorded imagined speech. *Expert Syst. Appl.* 118, 201–208. doi: 10.1016/j.eswa.2018.10.004
- Mohanchandra, K., and Saha, S. (2016). A communication paradigm using subvocalized speech: translating brain signals into speech. *Augment. Hum. Res.* 1, 1–14. doi: 10.1007/s41133-016-0001-z
- Molinari, N., and Lizarazu, M. (2018). Delta (but not theta)-band cortical entrainment involves speech-specific processing. *Eur. J. Neurosci.* 48, 2642–2650. doi: 10.1111/ejn.13811
- Morooka, R., Tanaka, H., Umahara, T., Tsugawa, A., and Hanyu, H. (2018). “Cognitive function evaluation of dementia patients using p300 speller,” in *International Conference on Applied Human Factors and Ergonomics* (Japan: Springer), 61–72. doi: 10.1007/978-3-319-94866-9_6
- Mudgal, S. K., Sharma, S. K., Chaturvedi, J., and Sharma, A. (2020). Brain computer interface advancement in neurosciences: applications and issues. *Interdiscip. Neurosurg.* 20:100694. doi: 10.1016/j.inat.2020.100694
- Navarro-Sune, X., Hudson, A., Fallani, F. D. V., Martinerie, J., Witon, A., Pouget, P., et al. (2016). Riemannian geometry applied to detection of respiratory states from EEG signals: the basis for a brain-ventilator interface. *IEEE Trans. Biomed. Eng.* 64, 1138–1148. doi: 10.1109/TBME.2016.2592820
- Nguyen, C. H., Karavas, G. K., and Artemiadis, P. (2017). Inferring imagined speech using EEG signals: a new approach using Riemannian manifold features. *J. Neural Eng.* 15:016002. doi: 10.1088/1741-2552/aa8235
- Padfield, N., Zabalza, J., Zhao, H., Masero, V., and Ren, J. (2019). eeg-based brain-computer interfaces using motor-imagery: techniques and challenges. *Sensors* 19:1423. doi: 10.3390/s19061423
- Pan, C., Lai, Y.-H., and Chen, F. (2021). “The effects of classification method and electrode configuration on eeg-based silent speech classification,” in *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, (Mexico: IEEE) 131–134. doi: 10.1109/EMBC46164.2021.9629709
- Panachakel, J. T., Ramakrishnan, A., and Ananthapadmanabha, T. (2019). “Decoding imagined speech using wavelet features and deep neural networks,” in *2019 IEEE 16th India Council International Conference (INDICON)*, (India: IEEE) 1–4. doi: 10.1109/INDICON47234.2019.9028925
- Panachakel, J. T., Ramakrishnan, A., and Ananthapadmanabha, T. (2020). A novel deep learning architecture for decoding imagined speech from EEG. *arXiv preprint arXiv:2003.09374*. doi: 10.48550/arXiv.2003.09374
- Paul, Y., Jaswal, R. A., and Kajal, S. (2018). “Classification of EEG based imagine speech using time domain features,” in *2018 International Conference on Recent Innovations in Electrical, Electronics & Communication Engineering (ICRIEECE)*, (India) 2921–2924. doi: 10.1109/ICRIEECE44171.2018.9008572
- Pawar, D., and Dhage, S. (2020). Multiclass covert speech classification using extreme learning machine. *Biomed. Eng. Lett.* 10, 217–226. doi: 10.1007/s13534-020-00152-x
- Peelle, J. E., Gross, J., and Davis, M. H. (2013). Phase-locked responses to speech in human auditory cortex are enhanced during comprehension. *Cereb. Cortex* 23, 1378–1387. doi: 10.1093/cercor/bhs118
- Pei, X., Leuthardt, E. C., Gaona, C. M., Brunner, P., Wolpaw, J. R., and Schalk, G. (2011). Spatiotemporal dynamics of electrocorticographic high gamma activity during overt and covert word repetition. *Neuroimage* 54, 2960–2972. doi: 10.1016/j.neuroimage.2010.10.029
- Portillo-Lara, R., Tahirbegi, B., Chapman, C. A., Goding, J. A., and Green, R. A. (2021). Mind the gap: state-of-the-art technologies and applications for EEG-based brain-computer interfaces. *APL Bioeng.* 5:031507. doi: 10.1063/5.0047237
- Rajagopal, D., Hemanth, S., Yashaswini, N., Sachin, M., and Suryakanth, M. (2020). Detection of Alzheimer’s disease using BCI. *Int. J. Prog. Res. Sci. Eng.* 1, 184–190.
- Rao, M. (2021). “Decoding imagined speech using wearable EEG headset for a single subject,” in *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, (Unites States: IEEE) 2622–2627.
- Rasheed, S. (2021). A review of the role of machine learning techniques towards brain-computer interface applications. *Mach. Learn. Knowl. Extract.* 3, 835–862. doi: 10.3390/make3040042
- Rezazadeh Sheresheh, A., Yousefi, R., Wong, A. T., Rudzicz, F., and Chau, T. (2019). Development of a ternary hybrid fNIRS-EEG brain-computer interface based on imagined speech. *Brain Comput. Interfaces* 6, 128–140. doi: 10.1080/2326263X.2019.1698928
- Riaz, A., Akhtar, S., Iftikhar, S., Khan, A. A., and Salman, A. (2014). “Inter comparison of classification techniques for vowel speech imagery using EEG sensors,” in *The 2014 2nd International Conference on Systems and Informatics (ICSAI 2014)*, (China) 712–717. doi: 10.1109/ICSAI.2014.7009378
- Roy, Y., Banville, H., Albuquerque, I., Gramfort, A., Falk, T. H., and Faubert, J. (2019). Deep learning-based electroencephalography analysis: a systematic review. *J. Neural Eng.* 16:051001. doi: 10.1088/1741-2552/ab260c
- Saad Zaghloul, Z., and Bayoumi, M. (2019). Early prediction of epilepsy seizures VLSI BCI system. *arXiv e-prints: arXiv-1906*. doi: 10.48550/arXiv.1906.02894
- Saha, P., Abdul-Mageed, M., and Fels, S. (2019a). Speak your mind! towards imagined speech recognition with hierarchical deep learning. *arXiv preprint arXiv:1904.05746*. doi: 10.21437/Interspeech.2019-3041
- Saha, P., and Fels, S. (2019). Hierarchical deep feature learning for decoding imagined speech from EEG. *Proc. AAAI Conf. Artif. Intell.* 33, 10019–10020. doi: 10.1609/aaai.v33i01.330110019
- Saha, P., Fels, S., and Abdul-Mageed, M. (2019b). “Deep learning the EEG manifold for phonological categorization from active thoughts,” in *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, (United Kingdom: IEEE) 2762–2766. doi: 10.1109/ICASSP.2019.8682330
- Salinas, J. S. G. (2017). *Bag of features for imagined speech classification in electroencephalograms*. Tesis de Maestría, Instituto Nacional de Astrofísica, Óptica y Electrónica. Available online at: <https://inaoe.repositorioinstitucional.mx/jspui/bitstream/1009/1253/1/GarciaSJS.pdf>
- Saminu, S., Xu, G., Shuai, Z., Isselmou, A. E. K., Jabire, A. H., Karaye, I. A., et al. (2021). Electroencephalogram (EEG) based imagined speech decoding and recognition. *J. Appl. Mat. Tech.* 2, 74–84. doi: 10.31258/jamt.2.2.74-84
- Sani, O. G., Yang, Y., and Shanchi, M. M. (2021). Closed-loop bci for the treatment of neuropsychiatric disorders. *Brain Comput. Interface Res.* 9:121. doi: 10.1007/978-3-030-60460-8_12
- Sarmiento, L., Lorenzana, P., Cortes, C., Arcos, W., Bacca, J., and Tovar, A. (2014). “Brain computer interface (BCI) with EEG signals for automatic vowel recognition based on articulation mode,” in *5th ISSNIP-IEEE Biosignals and Biorobotics Conference, Biosignals and Robotics for Better and Safer Living (BRC)*, (Salvador: IEEE) 1–4. IEEE. doi: 10.1109/BRC.2014.6880997
- Sazgar, M., and Young, M. G. (2019). “Overview of EEG, electrode placement, and montages,” in *Absolute Epilepsy and EEG Rotation Review* (Cham: Springer), 117–125. doi: 10.1007/978-3-030-03511-2_5
- Schroeder, C. E., Lakatos, P., Kajikawa, Y., Partan, S., and Puce, A. (2008). Neuronal oscillations and visual amplification of speech. *Trends Cogn. Sci.* 12, 106–113. doi: 10.1016/j.tics.2008.01.002
- Seresheh, A. R., Trott, R., Bricout, A., and Chau, T. (2017). EEG classification of covert speech using regularized neural networks. *IEEE/ACM Trans. Audio Speech Lang. Process.* 25, 2292–2300. doi: 10.1109/TASLP.2017.2758164
- Seresheh, A. R., Yousefi, R., Wong, A. T., and Chau, T. (2018). Online classification of imagined speech using functional near-infrared spectroscopy signals. *J. Neural Eng.* 16:016005. doi: 10.1088/1741-2552/aae4b9
- Sharon, R. A., and Murthy, H. A. (2020). Correlation based multi-phasal models for improved imagined speech EEG recognition. *arXiv preprint arXiv:2011.02195*. doi: 10.21437/SMM.2020-5

- Si, X., Li, S., Xiang, S., Yu, J., and Ming, D. (2021). Imagined speech increases the hemodynamic response and functional connectivity of the dorsal motor cortex. *J. Neural Eng.* 18:056048. doi: 10.1088/1741-2552/ac25d9
- Singh, A., and Gumaste, A. (2021). Decoding imagined speech and computer control using brain waves. *J. Neurosci. Methods* 358:109196. doi: 10.1016/j.jneumeth.2021.109196
- Song, Y., and Sepulveda, F. (2014). "Classifying speech related vs. idle state towards onset detection in brain-computer interfaces overt, inhibited overt, and covert speech sound production vs. idle state," in *2014 IEEE Biomedical Circuits and Systems Conference (BioCAS) Proceedings*, (Switzerland: IEEE) 568–571. doi: 10.1109/BioCAS.2014.6981789
- Stober, S., Sternin, A., Owen, A. M., and Grahn, J. A. (2015). Deep feature learning for EEG recordings. *arXiv preprint arXiv:1511.04306*. doi: 10.48550/arXiv.1511.04306
- Subasi, A. (2007). EEG signal classification using wavelet feature extraction and a mixture of expert model. *Expert Syst. Appl.* 32, 1084–1093. doi: 10.1016/j.eswa.2006.02.005
- Suhaimi, N. S., Mountstephens, J., and Teo, J. (2020). EEG-based emotion recognition: a state-of-the-art review of current trends and opportunities. *Comput. Intell. Neurosci.* 2020:8875426. doi: 10.1155/2020/8875426
- Tamm, M.-O., Muhammad, Y., and Muhammad, N. (2020). Classification of vowels from imagined speech with convolutional neural networks. *Computers* 9:46. doi: 10.3390/computers9020046
- Ten Oever, S., and Sack, A. T. (2015). Oscillatory phase shapes syllable perception. *Proc. Natl. Acad. Sci. U.S.A.* 112, 15833–15837. doi: 10.1073/pnas.1517519112
- Torres-García, A. A., Reyes-García, C. A., and Villase nor-Pineda, L. (2022). "A survey on EEG-based imagined speech classification," in *Biosignal Processing and Classification Using Computational Learning and Intelligence* (Mexico: Elsevier), 251–270. doi: 10.1016/B978-0-12-820125-1.00025-7
- Tøttrup, L., Leerskov, K., Hadsund, J. T., Kamavuako, E. N., Kæseler, R. L., and Jochumsen, M. (2019). "Decoding covert speech for intuitive control of brain-computer interfaces based on single-trial EEG: a feasibility study," in *2019 IEEE 16th International Conference on Rehabilitation Robotics (ICORR)*, (Canada: IEEE) 689–693. doi: 10.1109/ICORR.2019.8779499
- Ullah, S., and Halim, Z. (2021). Imagined character recognition through EEG signals using deep convolutional neural network. *Med. Biol. Eng. Comput.* 59, 1167–1183. doi: 10.1007/s11517-021-02368-0
- Wang, L., Zhang, X., Zhong, X., and Zhang, Y. (2013). Analysis and classification of speech imagery EEG for BCI. *Biomed. Signal Process. Control* 8, 901–908. doi: 10.1016/j.bspc.2013.07.011
- Yger, F., Berar, M., and Lotte, F. (2016). Riemannian approaches in brain-computer interfaces: a review. *IEEE Trans. Neural Syst. Rehabil. Eng.* 25, 1753–1762. doi: 10.1109/TNSRE.2016.2627016
- Zhang, D., Gong, E., Wu, W., Lin, J., Zhou, W., and Hong, B. (2012). "Spoken sentences decoding based on intracranial high gamma response using dynamic time warping," in *2012 Annual International Conference of the IEEE Engineering in Medicine and Biology Society (United States: IEEE)*, 3292–3295. doi: 10.1109/EMBC.2012.6346668
- Zhang, J., Huang, T., Wang, S., and Liu, Y.-j. (2019). Future internet: trends and challenges. *Front. Inform. Technol. Electron. Eng.* 20, 1185–1194. doi: 10.1631/FITEE.1800445
- Zhang, X., Yao, L., Zhang, S., Kanhere, S., Sheng, M., and Liu, Y. (2018). Internet of things meets brain-computer interface: a unified deep learning framework for enabling human-thing cognitive interactivity. *IEEE Intern. Things J.* 6, 2084–2092. doi: 10.1109/JIOT.2018.2877786
- Zhao, S., and Rudzicz, F. (2015). "Classifying phonological categories in imagined and articulated speech," in *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (Australia: IEEE)*, 992–996. doi: 10.1109/ICASSP.2015.7178118

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2022 Lopez-Bernal, Balderas, Ponce and Molina. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.