Decoding Bilingual EEG Signals With Complex Semantics Using Adaptive Graph Attention Convolutional Network

Chengfang Li, *Student Member, IEEE*, Yang Liu, *Graduate Student Member, IEEE*, Jielin Li, Yuhao Miao, *Graduate Student Member, IEEE*, Jing Liu, *Member, IEEE*, and Liang Song, *Senior Member, IEEE*

Abstract— Decoding neural signals of silent reading with Brain-Computer Interface (BCI) techniques presents a fast and intuitive communication method for severely aphasia patients. Electroencephalogram (EEG) acquisition is convenient and easily wearable with high temporal resolution. However, existing EEG-based decoding units primarily concentrate on individual words due to their low signal-to-noise ratio, rendering them insufficient for facilitating daily communication. Decoding at the word level is less efficient than decoding at the phrase or sentence level. Furthermore, with the popularity of multilingualism, decoding EEG signals with complex semantics under multiple languages is highly urgent and necessary. To the best of our knowledge, there is currently no research on decoding EEG signals during silent reading of complex semantics, let alone decoding silent reading EEG signals with complex semantics for bilingualism. Moreover, the feasibility of decoding such signals remains to be investigated. In this work, we collect silent reading EEG signals of 9 English Phrases (EP), 7 English Sentences (ES), 10 Chinese Phrases (CP), and 7 Chinese Sentences (CS) from the subject within 26 days. We propose a novel Adaptive Graph Attention Convolution Network (AGACN) for classification. Experimental results demonstrate that our proposed method outperforms state-of-the-art methods, achieving the highest classification accuracy of 54.70%, 62.26%, 44.55%, and 57.14% for silent reading EEG signals of EP, ES, CP, and CS, respectively. Moreover, our results prove the feasibility of complex semantics EEG signal decoding. This work will aid aphasic patients in achieving regular communication while providing novel ideas for neural signal decoding research.

Index Terms— Brain-computer interface, EEG signals, complex semantics, silent reading, multiple languages.

I. INTRODUCTION

THE aphasia caused by nerve damage is irreversible. Fortunately, Brain-Computer Interface (BCI) technology for decoding neural signals offers an efficient communication method for aphasic patients [1], [2], [3]. Numerous studies [4], [5], [6], [7], [8], [9] have successfully decoded Electrocorticogram (ECoG) signals of silent reading words, phrases, and sentences. Compared with ECoG, Electroencephalogram (EEG) [10] is non-invasive, more user-friendly, has a high temporal resolution, and is widely used for decoding neural signals. Due to its low signal-to-noise ratio, the units of early studies on decoding silent reading EEG signals primarily focused on syllables and phonemes [11], [12], [13], [14]. Research [12] also used non-English to replace the imaginary pronunciation of English words and decode their EEG signals to help aphasic patients achieve simple communication. However, decoding limited sets of syllables, phonemes, and non-English words is inadequate for patients with language disorders to communicate daily. With the advancement of deep learning techniques and the enhancement of acquisition equipment [15], [16], researchers [14], [17] have attempted to decode silent reading EEG signals for multiple categories of words and improved classification accuracy. For example, Vorontsova et al. [17] decoded silent reading EEG signals with 85% accuracy using 8 meaningful words. Although the results were not as promising for out-of-sample subjects, good performance was achieved on a single subject. These findings provide an encouraging incentive for investigating the decoding of silent reading EEG signals. However, decoding EEG signals word-by-word is time-consuming and requires more cognitive effort. A more efficient and natural approach is directly decoding entire phrases and sentences from neural signals. This method will significantly enhance

Chengfang Li, Yuhao Miao, Jing Liu, and Liang Song are with the Academy for Engineering and Technology, Fudan University, Shanghai 200433, China (e-mail: cfli20@fudan.edu.cn; yhmiao22@m.fudan.edu.cn; jingliu19@fudan.edu.cn; songl@fudan.edu.cn).

Yang Liu is with the Academy for Engineering and Technology, Fudan University, Shanghai 200433, China, also with the Department of Computer Science, University of Toronto, Toronto, ON M5S 1A1, Canada, and also with the Division of Natural and Applied Sciences, Duke Kunshan University, Suzhou 215316, China (e-mail: yang_liu20@fudan.edu.cn).

Jielin Li is with the Department of Electrical and Computer Engineering, University of Toronto, Toronto, ON M5S 3G4, Canada (e-mail: jielin.li@mail.utoronto.ca).

Digital Object Identifier 10.1109/TNSRE.2023.3348981

the communication efficiency of aphasic patients. Additionally, it is necessary to decode complex semantic EEG signals of multiple languages in a multilingual social environment. Decoding silent reading multiple EEG signals provides a more realistic and smooth communication experience. It avoids delays, misinterpretation, and information leakage caused by translation software, which relies on external input and lacks communication depth. Furthermore, decoding silent reading bilingual EEG signals can help deepen the understanding of the neural mechanisms in the brain related to multilingual processing. It thus can promote neuroscience research in bilingual interaction, memory, and language processing [18].

Compared with words, phrases and sentences contain more complex information. Therefore, the corresponding silent reading EEG signals are more intricate. Decoding neural signals with complex semantics under bilingualism requires consideration not only of the differences in brain use under different languages but also of the complexity of semantics and changes in neural signals over time [19]. To effectively extract meaningful information from EEG signals with complex semantics and decode them. The proposed algorithm must be capable of selecting channels that contain the most useful information corresponding to different languages. Additionally, it should capture fluctuations in both temporal and spatial relationships between the involved channels. Therefore, decoding bilingual silent reading EEG signals with complex semantics is methodologically more demanding. Furthermore, whether it is possible to decode them directly has yet to be studied for feasibility.

EEG signals are typically non-Euclidean structured data [20], [21]. Representing EEG signals with graphs provides a more comprehensive representation of the spatial connections between channels compared to using temporal sampling points for each channel [22]. On this basis, we use a two-dimensional matrix and its adaptive graph to represent each EEG signal jointly. The two-dimensional matrix represents the sampling points of each channel in time, and the graph provides a richer representation of the spatial connections between channels. Using a fixed graph for all EEG signals does not provide a more comprehensive representation of the variability of EEG signals [23] due to the differences in the corresponding EEG signals of the same subject doing the same task at different times. Therefore, we construct an adaptive graph structure for each EEG signal based on its characteristics to capture real-time changes in both temporal and spatial domains.

This study collects silent reading EEG signals of 9 English Phrases (EP), 7 English Sentences (ES), 10 Chinese Phrases (CP), and 7 Chinese Sentences (CS) from the subject over 26 days. We propose a novel Adaptive Graph Attention Convolution Network (AGACN) for classifying our datasets. Our proposed method achieves the accuracy of $51.05\pm3.62\%$, $57.85\pm4.41\%$, $40.00\pm4.55\%$, and $52.20\pm4.94\%$ for silent reading EEG signals of EP, ES, CP, and CS, respectively. Our work demonstrates the feasibility of decoding silent reading EEG signals of phrases and sentences, opening up a novel representation method for developing neural signal language decoding techniques. This study is expected to provide more communication options and improve the quality of life for aphasic individuals. Our main contributions are summarized as follows:

- We collect comprehensive datasets that contain silent reading EEG signals of 9 English phrases, 7 English sentences, 10 Chinese phrases, and 7 Chinese sentences from the subject over 26 days. These datasets fill the gap for the research of decoding silent reading EEG signals with complex semantics in multiple languages.

- We propose a novel Adaptive Graph Attention Convolution Network (AGACN) for classifying our datasets. Our method uses each feature matrix and its adaptive graph structure as inputs. AGACN uses the attention mechanism on the feature matrix according to the weights on the edges of the graph to effectively capture the temporal and spatial features of EEG signals.

- Our method performs superior to state-of-the-art methods on four collected datasets. Extensive experiments demonstrate the feasibility of decoding bilingual silent reading EEG signals with complex semantics.

II. RELATED WORK

Studies [4], [5], [6], [7], [8], [24], [25] have shown that it is possible to decode certain aspects of continuously spoken speech from ECoG signals, including phonemes, words, and sentences, or reconstruct acoustic properties of perception and speech. In 2021, Willett et al. [19] successfully decoded handwriting movements from neural activity in the motor cortex, achieving typing speeds of 90 characters per minute with an accuracy of 94.10% in one individual. Moreover, multiple studies have proven direct decoding of EEG signals associated with the silent reading of phonemes and words. Early research [11] attempted to classify covert speech EEG, focused on small language units such as vowels and syllables. In 2009, DaSalla et al. [11] recognized EEG signals of English phonemes /a/ and /u/. In 2010, Brigham and Kumar [26] classified the EEG signals of two imaginary syllables /ba/ and /ku/, and initial results suggest that it is possible to identify imagined speech. Following these, several researchers proposed classifying complete and meaningful words, but most are non-English. For instance, Jahangir and Sepulveda [12] took 'le', 'ry','ba' and 'fo' as the abbreviation of 'left', 'right', 'back' and 'forward'. Zhao and Rudzicz [13] used 7 phonemic syllabic prompts (i.e., /iy/, /uw/, /piy/, /tiy/, /diy/, /m/, and /n/) and 4 words phonetically-similar pairs (i.e., *pat, pot, knew*, and *gnaw*) with 3 modalities (EEG, facial, and audio) for classification. Nguyen et al. [14] tried to classify the unspoken speech EEG signals of vowels (/a/, /i/, and /u/), long words ('cooperate' and 'independent'), and short words ('in', 'out', and 'up'). Sereshke et al. [27] classified 'yes' and 'no' to investigate the reliability of EEG signals in discriminating between different covert speech tasks. Lee et al. [28] classified meaningful Spanish words: 'arriba', 'abajo', 'derecha', 'izquierda', 'adelante', and 'atrás'. In 2021, Darya et al. [17] expanded the category of EEG signals to 8 words (i.e., 'forward', 'backward', 'up', 'down', 'help', 'take', 'stop', and 'release'). Balaji et al. [29] classified English and Hindi silent speech EEG signals for 'yes'/'no' and 'Haan'/'Na' respectively. Li et al. [30] decoded silent reading EEG signals of 7 Chinese and 9 English words.
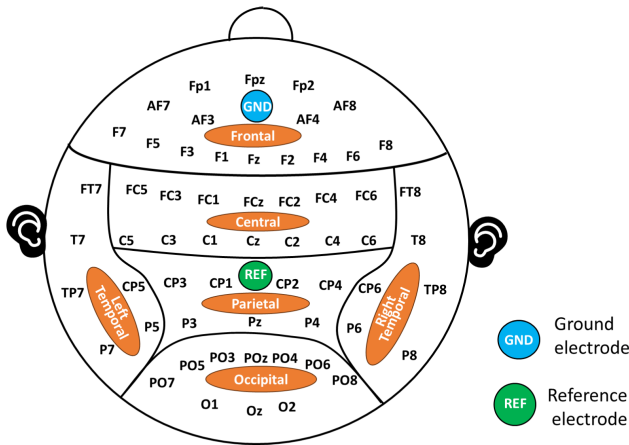
Fig. 1. Brain topography and regions covered by 59 brain channel electrode.

III. Data Collection

A. Preliminary Works

We use a 64-channel wet electrode wireless EEG equipment device with a sampling frequency of $1000 Hz$. The 64 channels include 59 brain and 5 body functional signal channels. The device is equipped with a wireless amplification, which can amplify the signal at the same point in time by an equal factor 2×10^4. In our study, we select channels corresponding to distinct brain regions for the comparative analysis in Section V-E. Therefore, we refer to [31] and categorize the 59 channels into six clusters. These clusters represent the frontal, central, temporal, parietal, and occipital lobe regions, as detailed in [32]. Notably, the temporal lobe is further divided into the left and right temporal lobe regions. The positions of the corresponding 59 brain channels and their regions, following the 10–20 system electrode placement method [33] are shown in Fig. 1.

The subject is 21 years old, male, right-handed, and has no neurological or other diseases. His native and second language are Chinese and English. He began learning English at 10, has a high English level, and can communicate skillfully. We collect EEG signals on 16 days during the 26 days, specifically on days 1, 4–8, 13–16, 18–22, and 26, with 3 to 5 daily trials conducted in each category. Data collection involving multiple days can enrich the data diversity because there are some changes in EEG signals when the same subject performs the same task on different days.

To ensure the selection phrases and sentences are appropriate, we refer to the corpus used in [24] and consult several individuals with language disorders. After consulting with numerous aphasia patients, we summarize the most frequently used phrases and sentences they desire to express daily. Subsequently, psychologists and linguists evaluate the scientific validity of these phrases and sentences employed in the experiment. Table I shows the corpus used in each category.

Before the experiment, the experimental monitor helps the subject wear the EEG cap and fill conductive paste into channels to reduce the impedance level within an acceptable error tolerance. During the experiment, the subject wears headphones and faces a white wall to minimize the effects of noise and visual stimulation. Two computers are used in the experiment. According to the set program, computer A passes the marks wirelessly to computer B. Computer B is used to record EEG signals and corresponding marks.

TABLE I
CORPUS OF EACH TASK CATEGORY

Task category	Corpus
EP	1. *close the window*, 2. *good morning*, 3. *go out*, 4. *happy birthday*, 5. *open the door*, 6. *pick up*, 7. *sit down*, 8. *slow down*, 9. *stand up*.
ES	1. *Have a nice weekday!* 2. *How is the weather today?* 3. *I have a lot of problems.* 4. *It is time to get up.* 5. *Long time no see.* 6. *Please call me before you come.* 7. *You win some, you lose some.*
CP	1. 出去 (*go out*), 2. 坐下 (*sit down*), 3. 开门 (*open the door*), 4. 早上好 (*good morning*), 5. 杯子 (*cup*), 6. 生日快乐 (*happy birthday*), 7. 看书 (*read a book*) 8. 站起来 (*stand up*), 9. 红色 (*red*), 10. 苹果 (*apple*).
CS	1. 今天天气怎么样？ (*Same as the 2nd in ES.*) 2. 你吃饭了吗？ (*Have you eaten yet?*) 3. 你来的时候给我打电话。(*Same as the 6th in ES.*) 4. 周末愉快！ (*Same as the 1st in ES.*) 5. 我有很多问题。(*Same as the 3rd in ES.*) 6. 胜败乃兵家常事。 (*Same as the 7th in ES.*) 7. 该起床了。(*Same as the 4th in ES.*)

The experimental design aims to collect EEG signals during silent reading. Depending on the characteristics of the used device, we set up *'listening'* before silent reading and *'speaking'* after silent reading. The *'listening'* reminds the subject to read what is pronounced silently, and *'speaking'* ensures that the subject does the silent reading of the audio he has heard. Ensure the subject has sufficient reception and reaction time to read phrases and sentences silently. Each phrase and sentence is listened to three times, silently read once, and spoken once, and EEG signals are collected throughout the entire process. The operator monitors the experiment process according to the instructions on the screen and marks the unqualified data. For example, this data will be marked if the subject does not speak or the *'speaking'* is inconsistent with the *'listening'*.

The experimental process is illustrated by collecting EEG signals of 9 English phrases, as shown in Fig. 2. There are a 'beep' and automatic marks at the beginning and end of *'listening'*, *'silent reading'*, and *'speaking'*. One trial consists of listening (3 times), silent reading (once), and speaking (once) 9 phrases in sequence. The duration for *'listening'*, *'silent reading'*, and *'speaking'* is set at 4 seconds each. 1–9 represent the beginning of *'listening'* 9 phrases. 21 and 22 represent the beginning and end of silent reading, and 30 and 31 represent the beginning and end of speaking. For example, marks 21, 22, 30, and 31 after 4 represent the beginning and end of silent reading and speaking the fourth phrase.

The process of collecting EEG data for sentences is identical to that for phrases. The duration of *'listening'*, *'silent reading'*, and *'speaking'* for sentences is each 5 seconds. Before the experiment, we introduce the process to the subject and sign an informed consent form. The experimental procedure is conducted under the ethical standards encoded in the latest Declaration of Helsinki [34]. This research received approval
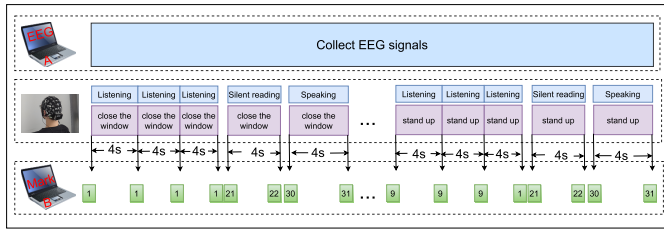
Fig. 2. Experiment procedure. The operator monitors the experiment process on two screens according to the instructions. The subject wears headphones and faces a white wall. 9 phrases, in turn, listening (3 times, each takes 4 seconds, all use 12 seconds), silent reading (once, takes 4 seconds), and speaking (once, takes 4 seconds). 1–9 represents listening 9 English phrases in turn. 21 and 22 represent the beginning and end of the silent reading corresponding phrase. 30 and 31 represent the beginning and end of speaking the corresponding phrase.

TABLE II
SIZE OF EACH DATASET

Dataset	Training size	Test size
SREP	460	116
SRES	391	106
SRCP	460	110
SRCS	392	91

from the Ethics Committee of Fudan University (IRB number: FE23167I).

B. Datasets

We cut 'silent reading' EEG signals of each phrase or sentence according to the marks in each trial. To reduce computation, we apply a sliding window with a size of 2 and a step of 2, downsampling the data within each window on average. The average of the numbers in the sliding window is used as a new sampling point. Thus, 2000 and 2500 sampling points are in the EEG signals of the downsampled silent phrases and sentences, respectively. Eventually, there are four datasets: Silent Reading English Phrases (SREP), Silent Reading English Sentences (SRES), Silent Reading Chinese Phrases (SRCP), and Silent Reading Chinese Sentences (SRCS). We remove the unqualified data marked in the experiment and randomly divide the qualified data into the training and test sets in the ratio of 8 : 2. The size of each dataset is shown in Table II. The dataset will be publicly available at: https://github.com/cfli20/EEG.

IV. METHOD

A. Graph Construction

Cumulative changes in neural activity can lead to differences in corresponding EEG signals of doing the same task [35], [36], even if the same subject performs the same task at different times. In this work, we propose a novel method for representing EEG signals to fully capture the temporal and spatial connections between channels in real-time. Specifically, we utilize a feature matrix and its corresponding graph structure to represent each EEG signal at different times. The feature matrix captures the temporal dynamics of EEG signals, while its graph structure represents the spatial correlation between each channel. The feature matrix ($F \in \mathbb{R}^{N \times T}$) is a two-dimensional matrix of the sampling points of each channel in time. The adjacency matrix ($G \in \mathbb{R}^{N \times N}$) is used to represent the graph structure, where

N is the number of channels, and T is the data points. In the graph structure $G(D, E)$, D is the set of nodes, and $D = \{v_{ti} | t = 1, \cdots, T, i = 1, \cdots, N\}$. Where E is the set of edges, the node set $E_s = \{E_{ij} | (i, j) \in N\}$, E_s is the weight value of edges i and j. We hypothesize that if channels exhibit strong correlations with each other, it indicates that their corresponding neural activities are relatively consistent. Thus, the more strongly correlated channels there are among the key channels, the more consistent the neural activity of their counterparts will be. Therefore, we calculate the Pearson Correlation Coefficients (PCC) [37] of any two channels and use it as the weight between the corresponding nodes. PCC is calculated as follows:

$$\rho = \frac{E[(X - \mu_X)(Y - \mu_Y)]}{\sigma_X \sigma_Y}, \tag{1}$$

where ρ is the PCC value corresponding to two nodes, μ_X and μ_Y are the mean value of channels X and Y, and σ_X, σ_Y are the standard deviation of the two channels. We treat each channel as a graph node and use ρ as the weights on the edges between two nodes to enhance the channel relationship. The selection of weights $E(i, j)$ on the corresponding edges of nodes i, j follows the following equation.

$$E(i, j) = \begin{cases} \rho & , & \rho > 0, \\ 0 & , & \rho \leq 0. \end{cases} \tag{2}$$

Therefore, we construct an adaptive graph structure for each EEG signal. Extensive experiments show that choosing the top 12 Key channels (AK-channels) using Analysis of Variance (ANOVA) [38] can get the highest accuracy. The specific selection process for the 12 AK-channels is detailed in Section V-D. Each silent reading task takes 5 seconds, and each data has 2500 points after performed downsampling, so its corresponding feature matrix is $F \in \mathbb{R}^{12 \times 2500}$. The adjacency matrix is $G \in \mathbb{R}^{12 \times 12}$, and the graph structure differs in the weights of each edge. In Fig. 3, we present the temporal and spatial variations of EEG signals between channels during the same silent reading task on the SREP. The visualization of the adjacency matrix further illustrates that the graph structure varies based on different EEG signals, mainly manifested in the weight differences among individual channels. Due to the fluctuations in EEG signals, the constructed graph structure also changes accordingly. As a result, the correlations between any two channels also vary.

Overall, the feature matrix and adjacency matrix keep the spatial and temporal connection between channels and help explore the changes in temporal and spatial dynamics of the EEG signals during different tasks and over different periods. By incorporating the graph structure and the feature matrix, the models can better capture the spatial correlation between the EEG channels, improving classification accuracy.

B. Adaptive Graph Attention Convolution Network

The adjacency matrix (G) is symmetric in an undirected weighted graph, and the principal diagonal is zero. Therefore, we introduce the unit matrix to represent the characteristics of each node adequately. The adjacency matrix ($\tilde{G}$) is as follows:

$$\tilde{G} = G + I, \tag{3}$$
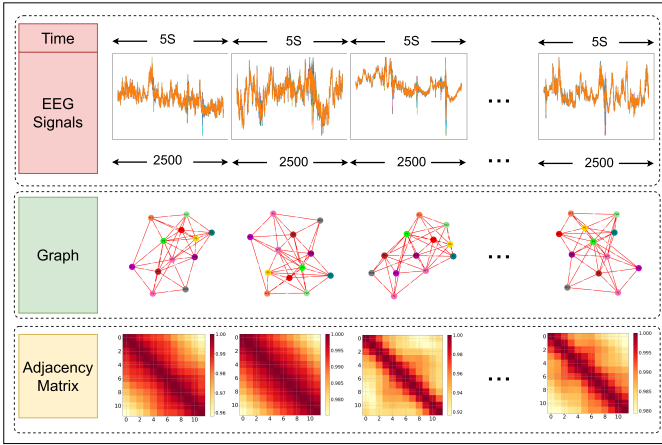
where, I is the unit matrix.

Fig. 3. Visualization of the data from the same silent reading task on the SREP. Each silent reading task lasts 5 seconds, resulting in 2500 down-sampled data points. An adaptive graph structure is constructed for each 5-second EEG signal, and its corresponding adjacency matrix is visualized.
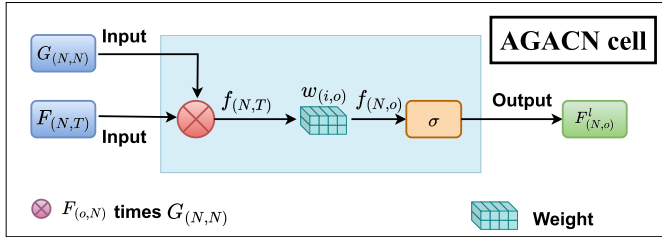

Fig. 4. The AGACN cell. $F_{(N,T)}$ is feature matrix, $G_{(N,N)}$ is the corresponding adjacency matrix of $F_{(N,T)}$. $w(i, o)$ is the training weight. σ is an activation function.

The range of the weights on edges is [0,1]. Therefore, no normalization of the adjacency matrix can ensure that different input data have similar values and can find the optimal solution faster by gradient descent [39]. In this work, each feature matrix (F) and its adaptive graph structure ($\tilde{G}$) are inputs of AGACN. AGACN uses $\tilde{G}$ as attention, F to convolve with $\tilde{G}$, and $\tilde{G}$ is the convolution kernel with the perceptual field of the whole F. According to different weights on the edges of $\tilde{G}$, different attention is given to its corresponding channels. This can preserve the important features of each channel in time more richly and can effectively capture and strengthen the spatial connection of channels. Then, a hidden layer network maintains useful information and uses the activation function for nonlinear expressions. Therefore, the AGACN cell layer is:

$$H = \sigma\left(\tilde{G}FW(i, o)\right), \qquad (4)$$

where, H is the output of the current layer, σ is the activation function, W is the training weight matrix, i is the input dimension, and o is the output dimension. The AGACN cell layer is shown in Fig. 4.

For the continuation of the features captured by the previous network layer to be passed on, we take the output of the previous layer as the feature matrix of the current layer. The transfer relationship between the AGACN cell layer is as follows:

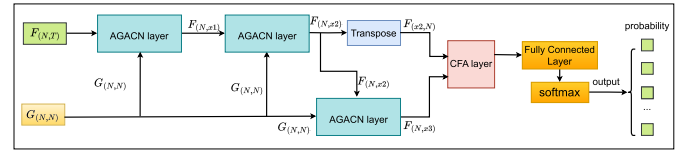$$H^{(l+1)} = \sigma\left(\tilde{G}^{(d)}H^{(l)}W(i, o)\right), \qquad (5)$$

Fig. 5. Structure of the AGACN. The inputs of AGACN are each feature matrix and its adaptive adjacency matrix. The outputs of the second and third AGACN cell layers are cross-fusion convolution and following FC layer and softmax layer.

where $H^{(l)}$ is the output of the previous layer as the char-acteristic matrix of the current layer. The feature of the first AGACN cell layer is F. $\tilde{G}^{(d)}$ is the d-th adjacency matrix that corresponding to the d-th characteristic matrix. In this work, $\tilde{G}^{(d)}$ is equal $\tilde{G}$.

We use three AGACN cell layers to classify complex seman-tics. The output of the second layer is transposed $((H^{(2)})^T)$, and as Attention to Cross Fusion (CFA layer) convolves with the output of the third layer to fully capture the connections of dynamic features between channels. Using $(H^{(2)})^T$ as the convolutional kernel with a receptive field corresponding to all the data in the third AGACN cell layer ($H^{(3)}$) ensures com-prehensive integration of the valuable information captured in both $(H^{(2)})^T$ and $H^{(3)}$. Therefore, the CFA layer is as follows:

$$H^{(f)} = (H^{(2)})^T \times H^{(3)}. \qquad (6)$$

After the CFA layer, a Fully Connected (FC) layer and a softmax layer are added. The structure diagram of the AGACN network is shown in Fig. 5.

V. RESULTS AND DISCUSSION

A. Experiment Setup

This study implements the AGACN using PyTorch frame-work [40], [41]. Parameters are kept the same for classifying all datasets, demonstrating the effectiveness and robustness of our proposed method. Specifically, the output dimension sizes for three AGACN cell layers are 126, 64, and 132, respectively. The activation functions for the three AGACN cell layers are Relu, Tanh and Relu [42]. We use a learning rate of 0.0001 and a dropout rate of 0.05. We use 5-fold cross-validation [43] to evaluate the performance of the model comprehensively. Namely, the training set is equally divided into 5 equal parts, and 4 of them are used in turn as the training set and the remaining one as the validation set. The Adam optimizer is used with a decay of 0.0005, a batch size of 80, and epochs of 500. The loss function is cross-entropy [44] as follows in Eq. (7), which measures the discrepancy between predicted and ground truth values.

$$H(p, q) = -\sum_{i=1}^{n} p(x_i) \cdot \log(q(x_i)), \qquad (7)$$

where, $p(x_i)$ represents the true distribution of ground truth sample, and $q(x_i)$ represents the distribution predicted by the model.

To illustrate the data flow changes during the attention convolution process, we show a feature capture schematic in Fig. 6. The input feature matrix is $F \in \mathbb{R}^{(12,2500)}$. The size of the graph structure in each AGACN cell layer is $G \in \mathbb{R}^{(12,12)}$. The feature matrices from the first to the third layers are denoted as $F^1 \in \mathbb{R}^{(12,126)}$, $F^2 \in \mathbb{R}^{(12,64)}$,
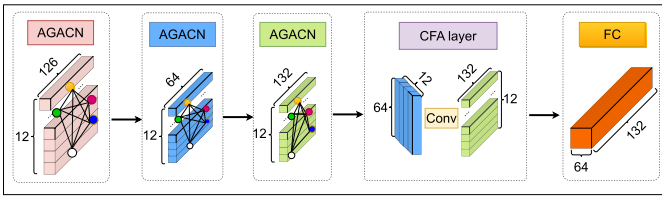
Fig. 6.   Data flow changes on SREP using AGACN. Feature matrices: $F \in R^{12 \times 2500}$, adjacency matrix: $G \in R^{12 \times 12}$.

and $F^3 \in \mathbb{R}^{(12,132)}$, respectively. Each layer of the AGACN cell uses an adaptive graph as attention. Layer by layer, the feature matrix containing richer features is gradually selected. Then, the feature matrices selected by the second layer of the AGACN cell are transposed as $F^2 \in \mathbb{R}^{(64,12)}$) and further convolved with the third AGACN cell layer. The CFA layer fully fuses features from the second and third layers of the AGACN cell, capturing both temporal and spatial aspects. It aligns the signals captured by the last two layers temporally and integrates features spatially, effectively highlighting useful characteristics.

B. Key Channels Selection

Zheng and Lu [45] observed higher classification accuracy when using key channels selected based on calculated features compared to using all channels. Therefore, identifying the most potent characteristic channels and using relatively few channels to achieve the highest classification accuracy is essential. We select key channels from global features on the training set in the following steps. Firstly, we apply a Butterworth [46] 7th-order filter to eliminate noise above $50Hz$, ensuring all the selected key channels contain more neural signals. Secondly, we normalize the filtered data to reduce the impact of different impedances on channel signals. For each one-dimensional feature x_i, the normalization formula for the eigenvalue $\tilde{x_{ik}}$ of the k-th sample x_{ik} in Eq. (8). Finally, we use the SelectKBest [47] function in Keras [48] to achieve it. The selectKBest function scores the features of each channel on the training set, with ANOVA employed as the criterion for feature scoring in this study.

$$\tilde{x_{ik}} = \frac{x_{ik} - min(x_i)}{max(x_i) - min(xi)}, \qquad (8)$$

where $min(x_i)$ and $max(x_i)$ represent the minimum and maximum values across all data samples x_{ik}.

We sort the 59 channels in descending order of their scores using the SelectKBest function. The selection process for the 12 channels is as follows. Firstly, we extract EEG data from channels 2 to 59 in training and test sets, forming feature matrices. Subsequently, we construct an adaptive graph for each feature matrix. Then, we divide the training set, which comprises channels 2 to 59, into new training and validation sets in a ratio of 7 : 3, respectively. Finally, we utilize the AGACN to classify the new training set and record the highest decoding results of the validation set. We also experiment with the classification accuracy of the validation set using N-fold cross-validation, with N values of 5, 8, and 10. The experimental results show that the validation set accuracy is relatively higher for the four datasets when using 5-fold cross-validation with the top 12 AK-channels. Therefore, we chose

TABLE III
THE TOP 12 AK-CHANNELS ON EACH DATASET

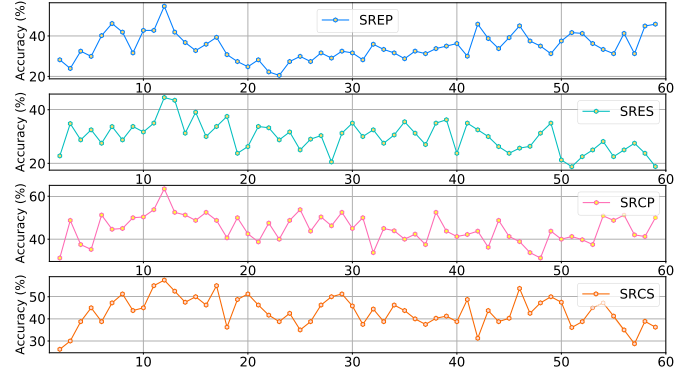| Dataset	Key Channels
SREP	F8, FCz, Fp1, AF7, AF3, C1, FC4, F1, Pz, F2, P5, P6
SRES	AF8, FC6, F8, T8, C6, AF7, Fp2, CP6, Fpz, F4, TP8
SRCP	O2, F1, P5, AF3, FC1, P3, FC3, FCz, PO5, Fz, F4, P4
SRCS	F8, FCz, Fp1, AF7, AF3, C1, FC4, F1, Pz, F2, P5, P6

Fig. 7.   The validation accuracy for the new training set using 2 to 59 channels on different datasets, respectively.

the top 12 AK-channels and 5-fold cross-validation for classification. The top 12 AK-channels are shown in Table III. The classification accuracy of each validation set is shown in Fig. 7.

As can be seen from Fig. 7, the validation accuracy is relatively high when using 12 AK-channels on our four datasets. When the number of channels is less than 12, the included features are not comprehensive, resulting in relatively lower accuracy. EEG signals have a low signal-to-noise ratio. More channels will have more noise, and thus, accuracy will not always increase with more channels. If the newly added channel contains more noise than the neural activity signal, the noise interferes with useful features, causing a decrease in accuracy. Conversely, if the newly added channel contains more useful neural signal features than noise, additional valuable features are introduced, increasing accuracy.

C. Classification Results Using AK-Channels

We simultaneously use EEGNet [49], DeepCovNet (DCN) [50], ShallowConvNet (SCN) [50], Support Vector Machine (SVM) [51] for classifying on our datasets. We also utilize Power Spectral Density [52] and Discrete Wavelet Transform [53] to extract features from EEG signals. Subsequently, we use SVM (PSD-SVM and DWT-SVM) to classify each case. We further utilize Multivariate Fast and Adaptive Empirical Mode Decomposition (MFAEMD) to extract features and employ the Light Gradient Boosting Machine (LGBM) algorithm (MFAEMD-LGBM) for classification [54]. They are widely used for EEG classification and have proven highly effective across various EEG classification tasks. In this study, SVM is trained using a Gaussian kernel function [55]. In the PSD-SVM experiments, the sampling frequency is selected as 100, the length of each data segment is 300, and the overlap between two analogous data segments is 10. Discrete wavelet transform fifth level decomposition using Db2 [56]. Classification results of our datasets using 12 AK-channels are shown in Table IV.

TABLE IV
CLASSIFICATION RESULTS USING 12 AK-CHANNELS (%)

Networks	SREP	SRES	SRCP	SRCS
SVM [51]	32.20±3.14	42.37±7.33	22.73±3.16	29.67±2.36
DCN [50]	35.47±3.85	40.31±3.57	24.09±6.81	35.92±1.85
SCN [49]	10.85±0.36	14.36±0.07	10.55±0.56	14.15±0.14
EEGNet [50]	35.47±1.28	42.33±2.55	28.18±1.82	31.32±1.65
PSD-SVM [52]	21.09±4.55	25.00±2.00	18.00±2.00	24.50±1.50
DWT–SVM [53]	21.50±7.50	34.50±7.50	18.75±2.25	28.00±1.00
MFAEMD-LGBM [54]	11.11±0.43	14.00±0.15	11.46±0.36	13.74±0.55
AGACN	**51.05±3.62**	**57.85±4.41**	**40.00±4.55**	**52.20±4.94**

(a) SREP   (b) SRES

(c) SRCP   (d) SRCS
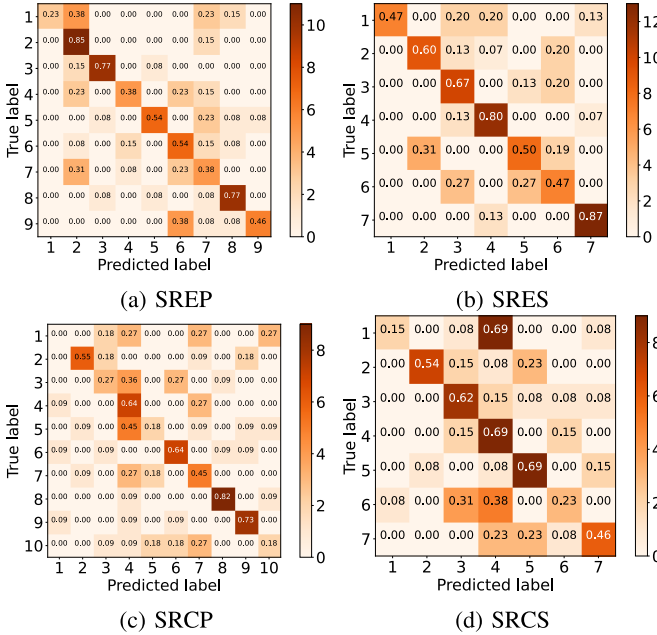
Fig. 8. Confusion matrix of four datasets classification using AGACN. The numbers corresponding to the horizontal and vertical coordinates represent the EEG signal category of the corresponding silent reading task in each data set in turn.

The confusion matrix is widely used to evaluate the performance of classification networks, which consists of True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN). The number on the diagonal of the confusion matrix indicates the probability that the category corresponding to the true label is correctly classified. In Fig. 8, we show the classification confusion matrix using the AGACN for four dataset classifications. For example, in the confusion matrix of SREP, the classification accuracy of silent EEG signals for 2 ('*good morning*') and 3 ('*go out*') reach 85% and 77%. However, the classification accuracy of 1 ('close the window') is lower, with 38% being incorrectly classified as 2 ('good morning'). After engaging in discussions with the subject, we conclude that an imbalance in accuracy across categories may be attributed to inconsistent attention to silent reading during various collection times within the same low-accuracy category. The subject exhibited difficulty concentrating and associating different elements with the same low-accuracy category during each collection time, which increased noise and inconsistencies, obscuring category features. The limited data also contributes to reduced feature distinctiveness within the same low-accuracy class.

In Table IV, the classification accuracy of four datasets using our proposed method is higher than that of the other networks. The main reason is that in the representation, we use a feature matrix and its adaptive graph structure jointly representing each EEG signal, fully preserving the temporal
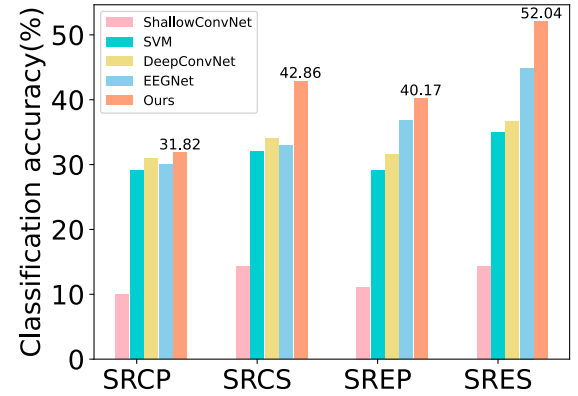

Fig. 9. Classification results using CTK-channels.

TABLE V
THE TOP 12 CTK-CHANNELS ON EACH DATASET

Dataset	CTK-channels
SREP	F1, P5, F4, AF7, PO5, FC1, FCz, Fp2, Fz, PO3, TP8, P4
SRES	F1, P4, F4, AF3, CP6, P5, Fz, FC2, FC1, P6, PC6, F2
SRCP	P5, F1, AF3, FC3, P3, F4, FC1, PO5, F3, P4, PO7, PO4
SRCS	AF3, F1, P5, F4, FC1, POz, Fz, FC3, Pz, FCz, PO8, F3

and spatial connections between the original data. In the convolution process, the graph is used as the convolution kernel to add attention to each channel according to the magnitude of the edge weights on the graph, capturing the temporal and spatial connections between channels. The network effectively captures the spatial continuity between channels by utilizing the output of the previous layer as the feature matrix of the current layer. The second and third layers of the network perform cross-fusion convolution of their outputs, a further fusion of features between each channel in both temporal and spatial.

D. Classification Results Using CTK-Channels and 59 Brain Channels

ANOVA and chi-square test [57] are widely used statistical methods across various data types. We use ANOVA as the criterion to select AK-channels since it is suitable for continuous data and categorical independent variables and can compare means and detect differences between groups. The chi-square test is appropriate for analyzing categorical data and determining independence between variables, which assesses the association between categorical variables. We also utilize the chi-square test as a feature selection criterion, employing the SelectKBest function to choose key channels for comparison. These selected channels are referred to as Chi-square Test Key channels (CTK-channels). The 12 CTK-channels selected in each dataset are shown in Table V. We construct an adaptive graph structure for each feature matrix on each dataset and then use AGACN, EEGNet [49], DeepCovNet [50], ShallowConvNet [50], and Support Vector Machine [51] for classification. The classification results are presented in Fig. 9. Results demonstrate that our proposed method gets higher decoding accuracy than other comparison networks when using CTK-channels but lower than AK-channels. This further demonstrates the effectiveness and robustness of our proposed method.

We also utilize 59 brain channels to decode the four datasets. The decoding accuracy of AGACN for EP, ES, CP, and CS
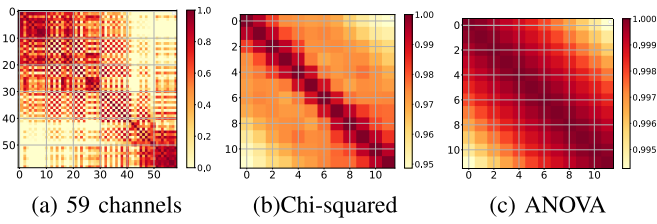
(a) 59 channels   (b) Chi-squared   (c) ANOVA

Fig. 10. PCC value between any two channels. (a) is the PCC values of any two channels from 59 channels. (b) is the PCC values of any channels from AK-channels (c) is the PCC values of any channels from CTK-channels.

TABLE VI
CLASSIFICATION RESULTS (%) USING DIFFERENT FUNCTIONAL AREA CHANNELS

Channels type	SREP	SRES	SRCP	SRCS
Frontal	39.32	31.63	27.37	30.33
Central	36.75	42.85	34.55	50.55
Parietal	28.21	23.47	14.55	20.88
Occipital	27.35	33.67	17.27	34.07
Left Temporal	28.21	34.70	21.82	33.33
Right Temporal	28.21	39.80	24.55	36.26
BW-channels	40.17	46.94	34.55	46.15

is 44.92%, 49.15%, 18.80%, and 22.45%, respectively. The decoding accuracy for all of them is lower than that of using AK-channels. We count the correlation between any two of the 12 AK-channels, 12 ATK-channels, and 59 channels of reading the same task silently EEG signals. The statistical results show that while the correlation between any two channels in the CTK-channels is strong in any of the datasets, the strength of the correlation of AK-channels is higher. This suggests that all AK-channels contain rich information on neural activity. Consequently, using AK-channels holds the potential for achieving higher decoding accuracy. We randomly visualize the size of the PCC between any two channels on the 59 brain channels, AK-channels, and CTK-channel dataset of the same silent reading tasks, as shown in Fig. 10. In Fig. 10, the strength of correlations between individual channels can be visually assessed, indicating differences in the consistency of neural activity corresponding to each channel. Furthermore, it demonstrates that the key channels selected using ANOVA are more suitable for AGACN than those selected using the chi-square test.

E. Classification Results Using Different Functional Area Channels

Using different functional area channels to decode EEG signals is mainstream [58], [59], [60]. To prove the scientificity of AK-channels, as a comparison, we select channels corresponding to the auditory brain center and auditory language center (Broca and Wernicke areas) [61], namely AF3, F3, F5, FC3, FC5, T7, C5, TP7, CP5, and P5 for classifying. We refer to the channels selected from Broca and Wernicke as BW-channels. Meanwhile, we also use the channels from the frontal, central, temporal, parietal, and occipital lobe regions for classifying, respectively. The channels corresponding to each lobe are illustrated in Fig. 1. All the classification results are shown in Table VI.

The classification results show that using AK-channels is higher than using different functional area channels on four datasets. These results are because language processing

TABLE VII
THE RESULTS (%) OF ABLATION STUDY ON EACH DATASET

Dataset	AGACN cell	AGACN cell	AGACN cell	AGACN cell	CFA layer	Accuracy
SREP	✓					22.22
	✓	✓				24.58
	✓	✓			✓	25.64
	✓	✓	✓			33.90
	✓	✓	✓	✓		38.46
	✓	✓	✓		✓	40.17
SRES	✓					33.19
	✓	✓				34.91
	✓	✓			✓	48.11
	✓	✓	✓			54.72
	✓	✓	✓	✓		50.94
	✓	✓	✓		✓	53.78
SRCP	✓					20.91
	✓	✓				22.73
	✓	✓			✓	23.64
	✓	✓	✓			27.28
	✓	✓	✓	✓		28.18
	✓	✓	✓		✓	24.55
SRCS	✓					31.87
	✓	✓				32.97
	✓	✓			✓	32.97
	✓	✓	✓			43.96
	✓	✓	✓	✓		40.66
	✓	✓	✓		✓	41.76

involves multiple brain regions, and the distribution of EEG signals on the scalp is complex, with possible cross-talk. Furthermore, the signal collected by the functional brain area may contain noise signals caused by the quality of the collection equipment, as well as hair and skin interference. The 12 AK-channels with the strongest features contain more useful information than different functional area channels. This also illustrates the scientific validity and effectiveness of AK-channels. As a result of these factors, the classification accuracy using channels from different functional language areas is lower than that of AK-channels.

F. Ablation Study

The AGACN achieves higher classification accuracy than state-of-the-art methods on four datasets due to the role of each layer. We conduct ablation experiments to explore the role of each layer in decoding. Table VII shows the ablation experiment results, with the '✓' indicating that the corresponding network layer is used.

Ablation experiment results demonstrate that classification results using three AGACN cells are higher than one or two AGACN cells on four datasets. Three AGACN layers can more effectively capture temporal and spatial characteristics than one or two AGACN cells. Using less than three AGACN cells does not fully capture the characteristics within the signals. The classification results using three AGACN cells with a CFA layer are superior to using three AGACN cells, indicating the significant role of the CFA layer in further capturing signal features. The accuracy of classification using four AGACN cells is not necessarily higher than that using three AGACN cells on our four datasets. The primary reason is that excessive network layers can lead to over-capturing features between signals, potentially misidentifying noise features as actual signal features. This also leads to the reason that the accuracy of using four AGACN cells with the CFA layer for the classifications is lower than that of three AGACN cell layers followed by the CFA layer. The results show that each layer of AGACN is important for decoding complex semantics silent reading EEG signals.

G. K-Fold Cross-Validation Experiments

K-fold cross-validation divides the training dataset into K equally sized subsets, rotating with each subset used once
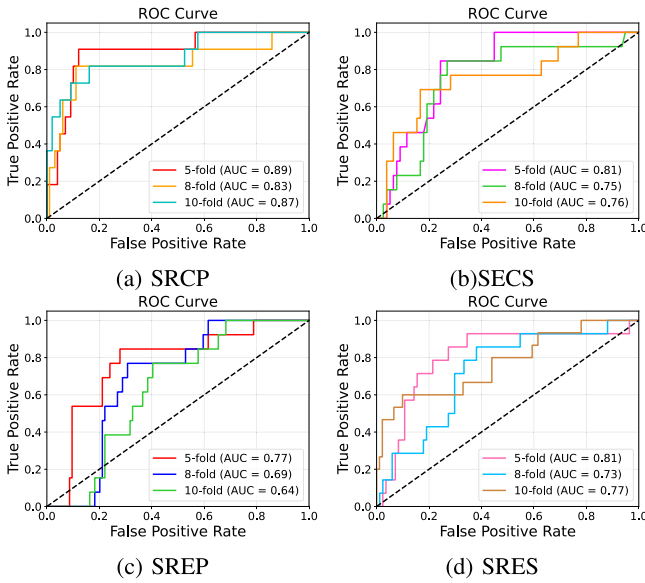
(a) SRCP (b)SECS

(c) SREP (d) SRES

Fig. 11.   ROC curves of 5-fold, 8-fold, and 10-fold cross-validation.

as the validation set and the remaining K-1 subset used for training. It can comprehensively evaluate the model and prevent over-fitting due to small data. The choice of K plays a significant role in the performance of the model.

Experiments prove that the AGACN performs best using 5-fold cross-validation. The Receiver Operating Characteristic (ROC) curve is a visual representation used to assess the quality of the classification model. TP and TN are the correctly predicted positive and negative classes. FP and FN are the incorrectly predicted positive and negative classes. The ROC curve illustrates the trade-off between the True Positive Rate (TPR) and the False Positive Rate (FPR) at various classification thresholds. The formulas of TPR and FPR are as follows:

$$TPR = \frac{TP}{TP + FN}, \tag{9}$$

$$FPR = \frac{FP}{FP + TN}. \tag{10}$$

Therefore, the higher the value of the Area Under the Curve (AUC), the better the decoding ability of the model [62]. Fig. 11 shows the ROC curve of 5-fold, 8-fold, and 10-fold cross-validation classifying using AGACN on four datasets. The results show that AGACN performs better using 5-fold cross-validation than 8-fold and 10-fold cross-validation on four datasets.

VI. CONCLUSION AND FUTURE WORKS

In this study, we collect SREP, SRES, SRCP, and SRCS datasets and fill the blank of decoding bilingual silent reading EEG signals with complex semantics. A novel AGACN is proposed for EEG signals with complex semantics classification. We use each feature matrix and its adaptive graph structure to represent each EEG signal. Further, they are used as inputs to the AGACN to capture the relationships among channels in both temporal and spatial domains. Extensive experiments demonstrate that our proposed method achieves superior classification accuracy compared to state-of-the-art methods.

It is crucial to note that while the current study exhibits the feasibility of decoding silent reading EEG signals with complex semantics, it is not a fully developed, clinically applicable system. Additionally, it should be acknowledged that the results of this work achieve the best performance among the three subjects examined. In addition, the AGACN experiments are limited to one-subject prediction. Further research is necessary to propose a more robust method that can overcome individual differences to achieve the decoding of silent neural signals, which is also the focus of our future work. We will also expand the count of phrases and sentences and further propose methods for decoding silent EEG signals involving more languages.

REFERENCES

[1] I. Lazarou, S. Nikolopoulos, P. C. Petrantonakis, I. Kompatsiaris, and M. Tsolaki, "EEG-based brain–computer interfaces for communication and rehabilitation of people with motor impairment: A novel approach of the 21st century," *Frontiers Hum. Neurosci.*, vol. 12, p. 14, Jan. 2018.

[2] L. Song, X. Hu, G. Zhang, P. Spachos, K. N. Plataniotis, and H. Wu, "Networking systems of AI: On the convergence of computing and communications," *IEEE Internet Things J.*, vol. 9, no. 20, pp. 20352–20381, Oct. 2022.

[3] F. Nijboer et al., "An auditory brain–computer interface (BCI)," *J. Neurosci. Methods*, vol. 167, no. 1, pp. 43–50, 2008.

[4] E. M. Mugler et al., "Direct classification of all American English phonemes using signals from functional speech motor cortex," *J. Neural Eng.*, vol. 11, no. 3, Jun. 2014, Art. no. 035015.

[5] S. Kellis, K. Miller, K. Thomson, R. Brown, P. House, and B. Greger, "Decoding spoken words using local field potentials recorded from the cortical surface," *J. Neural Eng.*, vol. 7, no. 5, Oct. 2010, Art. no. 056007.

[6] H. Akbari, B. Khalighinejad, J. L. Herrero, A. D. Mehta, and N. Mesgarani, "Towards reconstructing intelligible speech from the human auditory cortex," *Sci. Rep.*, vol. 9, no. 1, pp. 1–12, Jan. 2019.

[7] B. N. Pasley et al., "Reconstructing speech from human auditory cortex," *PLoS Biol.*, vol. 10, no. 1, Jan. 2012, Art. no. e1001251.

[8] C. Herff, G. Johnson, L. Diener, J. Shih, D. Krusienski, and T. Schultz, "Towards direct speech synthesis from ECoG: A pilot study," in *Proc. 38th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2016, pp. 1540–1543.

[9] P. Shenoy, K. J. Miller, J. G. Ojemann, and R. P. N. Rao, "Generalized features for electrocorticographic BCIs," *IEEE Trans. Biomed. Eng.*, vol. 55, no. 1, pp. 273–280, Jan. 2008.

[10] N. Kannathal, U. R. Acharya, C. M. Lim, and P. Sadasivan, "Characterization of EEG—A comparative study," *Comput. Methods Programs Biomed.*, vol. 80, no. 1, pp. 17–23, 2005.

[11] C. S. DaSalla, H. Kambara, M. Sato, and Y. Koike, "Single-trial classification of vowel speech imagery using common spatial patterns," *Neural Netw.*, vol. 22, no. 9, pp. 1334–1339, Nov. 2009.

[12] A. Jahangiri and F. Sepulveda, "The contribution of different frequency bands in class separability of covert speech tasks for BCIs," in *Proc. 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2017, pp. 2093–2096.

[13] S. Zhao and F. Rudzicz, "Classifying phonological categories in imagined and articulated speech," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process. (ICASSP)*, Apr. 2015, pp. 992–996.

[14] C. H. Nguyen, G. K. Karavas, and P. Artemiadis, "Inferring imagined speech using EEG signals: A new approach using Riemannian manifold features," *J. Neural Eng.*, vol. 15, no. 1, Feb. 2018, Art. no. 016002.

[15] F. Riaz, A. Hassan, S. Rehman, I. K. Niazi, and K. Dremstrup, "EMD-based temporal and spectral features for the classification of EEG signals using supervised learning," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 24, no. 1, pp. 28–35, Jan. 2016.

[16] M. Jochumsen, I. Khan Niazi, D. Taylor, D. Farina, and K. Dremstrup, "Detecting and classifying movement-related cortical potentials associated with hand movements in healthy subjects and stroke patients from single-electrode, single-trial EEG," *J. Neural Eng.*, vol. 12, no. 5, Oct. 2015, Art. no. 056013.

[17] D. Vorontsova et al., "Silent EEG-speech recognition using convolutional and recurrent neural network with 85% accuracy of 9 words classification," *Sensors*, vol. 21, no. 20, p. 6744, Oct. 2021.

[18] C. Li and L. Song, "GCN-LSTM for EEG classification based on unspoken speech of bilinguals," in *Proc. 24th Int. Conf. Digit. Signal Process. (DSP)*, Jun. 2023, pp. 1–4.

[19] F. R. Willett, D. T. Avansino, L. R. Hochberg, J. M. Henderson, and K. V. Shenoy, "High-performance brain-to-text communication via handwriting," *Nature*, vol. 593, no. 7858, pp. 249–254, May 2021.

[20] P. K. Singh, "Data with non-Euclidean geometry and its characterization," *J. Artif. Intell. Technol.*, vol. 2, no. 1, pp. 3–8, Dec. 2021.

[21] C. H. Nguyen and P. Artemiadis, "EEG feature descriptors and discriminant analysis under Riemannian manifold perspective," *Neurocomputing*, vol. 275, pp. 1871–1883, Jan. 2018.

[22] B. A. Miller, N. T. Bliss, and P. J. Wolfe, "Toward signal processing theory for graphs and non-Euclidean data," in *Proc. IEEE Int. Conf. Acoust., Speech Signal Process.*, Mar. 2010, pp. 5414–5417.

[23] Q. Chang et al., "Classification of first-episode schizophrenia, chronic schizophrenia and healthy control based on brain network of mismatch negativity by graph neural network," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 29, pp. 1784–1794, 2021.

[24] D. A. Moses et al., "Neuroprosthesis for decoding speech in a paralyzed person with anarthria," *New England J. Med.*, vol. 385, no. 3, pp. 217–227, Jul. 2021.

[25] X. Pei, D. L. Barbour, E. C. Leuthardt, and G. Schalk, "Decoding vowels and consonants in spoken and imagined words using electrocorticographic signals in humans," *J. Neural Eng.*, vol. 8, no. 4, Aug. 2011, Art. no. 046028.

[26] K. Brigham and B. V. K. V. Kumar, "Imagined speech classification with EEG signals for silent communication: A preliminary investigation into synthetic telepathy," in *Proc. 4th Int. Conf. Bioinf. Biomed. Eng.*, Jun. 2010, pp. 1–4.

[27] A. Rezazadeh Sereshkeh, R. Trott, A. Bricout, and T. Chau, "EEG classification of covert speech using regularized neural networks," *IEEE/ACM Trans. Audio, Speech, Language Process.*, vol. 25, no. 12, pp. 2292–2300, Dec. 2017.

[28] D.-Y. Lee, M. Lee, and S.-W. Lee, "Classification of imagined speech using Siamese neural network," in *Proc. IEEE Int. Conf. Syst., Man, Cybern. (SMC)*, Oct. 2020, pp. 2979–2984.

[29] A. Balaji et al., "EEG-based classification of bilingual unspoken speech using ANN," in *Proc. 39th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2017, pp. 1022–1025.

[30] C. Li, H. Wang, Y. Liu, X. Zhu, and L. Song, "Silent EEG classification using cross-fusion adaptive graph convolution network for multilingual neurolinguistic signal decoding," *Biomed. Signal Process. Control*, vol. 87, Jan. 2024, Art. no. 105524.

[31] S. Shahdadian, X. Wang, H. Wanniarachchi, A. Chaudhari, N. C. D. Truong, and H. Liu, "Neuromodulation of brain power topography and network topology by prefrontal transcranial photobiomodulation," *J. Neural Eng.*, vol. 19, no. 6, Dec. 2022, Art. no. 066013.

[32] E.-U.-H. Qazi, M. Hussain, and H. A. AboAlsamh, "Electroencephalogram (EEG) brain signals to detect alcoholism based on deep learning," *Comput., Mater. Continua*, vol. 67, no. 3, pp. 3329–3348, 2021.

[33] R. W. Homan, "The 10–20 electrode system and cerebral location," *Amer. J. EEG Technol.*, vol. 28, no. 4, pp. 269–279, Dec. 1988.

[34] M. D. Goodyear, K. Krleza-Jeric, and T. Lemmens, "The declaration of Helsinki," *Bmj*, vol. 335, no. 7621, pp. 624–625, 2007.

[35] T. W. James and I. Gauthier, "Repetition-induced changes in BOLD response reflect accumulation of neural activity," *Hum. Brain Mapping*, vol. 27, no. 1, pp. 37–46, Jan. 2006.

[36] H. P. Goodkin, "Status epilepticus increases the intracellular accumulation of GABAA receptors," *J. Neurosci.*, vol. 25, no. 23, pp. 5511–5520, Jun. 2005.

[37] W. DeGroat, V. Venkat, W. Pierre-Louis, H. Abdelhalim, and Z. Ahmed, "Hygieia: AI/ML pipeline integrating healthcare and genomics data to investigate genes associated with targeted disorders and predict disease," *Softw. Impacts*, vol. 16, May 2023, Art. no. 100493.

[38] L. Sthle and S. Wold, "Analysis of variance (ANOVA)," *Chemometrics Intell. Lab. Syst.*, vol. 6, no. 4, pp. 259–272, Nov. 1989.

[39] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *Proc. 5th Int. Conf. Learn. Represent.*, 2017, pp. 1–14.

[40] A. Paszke et al., "PyTorch: An imperative style, high-performance deep learning library," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 32, Dec. 2019, pp. 8024–8035.

[41] N. Ketkar, J. Moolayil, N. Ketkar, and J. Moolayil, "Introduction to PyTorch," in *Deep Learning With Python*. Berkeley, CA, USA: Apress, 2021, pp. 27–91.

[42] S. Sharma, S. Sharma, and A. Athaiya, "Activation functions in neural networks," *Towards Data Sci.*, vol. 6, no. 12, pp. 310–316, 2017.

[43] L. Dora, S. Agrawal, R. Panda, and A. Abraham, "Nested cross-validation based adaptive sparse representation algorithm and its application to pathological brain classification," *Expert Syst. Appl.*, vol. 114, pp. 313–321, Dec. 2018.

[44] D. Oliva, S. Hinojosa, E. Cuevas, G. Pajares, O. Avalos, and J. Gálvez, "Cross entropy based thresholding for magnetic resonance brain images using crow search algorithm," *Expert Syst. Appl.*, vol. 79, pp. 164–180, Aug. 2017.

[45] W.-L. Zheng and B.-L. Lu, "Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks," *IEEE Trans. Auto. Mental Develop.*, vol. 7, no. 3, pp. 162–175, Sep. 2015.

[46] S. K. Jagtap and M. D. Uplane, "The impact of digital filtering to ECG analysis: Butterworth filter application," in *Proc. Int. Conf. Commun., Inf. Comput. Technol. (ICCICT)*. London, U.K.: Butterworth, Oct. 2012, pp. 1–6.

[47] E. Bisong, "More supervised machine learning techniques with scikit-learn," in *Building Machine Learning and Deep Learning Models on Google Cloud Platform*. Cham, Switzerland: Springer, 2019, pp. 287–308.

[48] N. Ketkar and N. Ketkar, "Introduction to keras," in *Deep Learning With Python*. Berkeley, CA, USA: Apress, 2017, pp. 97–111.

[49] V. J. Lawhern, A. J. Solon, N. R. Waytowich, S. M. Gordon, C. P. Hung, and B. J. Lance, "EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces," *J. Neural Eng.*, vol. 15, no. 5, Oct. 2018, Art. no. 056013.

[50] R. T. Schirrmeister et al., "Deep learning with convolutional neural networks for EEG decoding and visualization," *Hum. Brain Mapping*, vol. 38, no. 11, pp. 5391–5420, Nov. 2017.

[51] C.-C. Chang and C.-J. Lin, "Training V-support vector classifiers: Theory and algorithms," *Neural Comput.*, vol. 13, no. 9, pp. 2119–2147, Sep. 2001.

[52] M. Alsolamy and A. Fattouh, "Emotion estimation from EEG signals during listening to Quran using PSD features," in *Proc. 7th Int. Conf. Comput. Sci. Inf. Technol. (CSIT)*, Jul. 2016, pp. 1–5.

[53] M. Zuhair and S. Thomas, "Classification of patient by analyzing EEG signal using DWT and least square support vector machine," *Adv. Sci., Technol. Eng. Syst. J.*, vol. 2, no. 3, pp. 1280–1289, Aug. 2017.

[54] S. Dash, R. K. Tripathy, D. K. Dash, G. Panda, and R. B. Pachori, "Multiscale domain gradient boosting models for the automated recognition of imagined vowels using multichannel EEG signals," *IEEE Sensors Lett.*, vol. 6, no. 11, pp. 1–4, Nov. 2022.

[55] S. S. Keerthi and C.-J. Lin, "Asymptotic behaviors of support vector machines with Gaussian kernel," *Neural Comput.*, vol. 15, no. 7, pp. 1667–1689, Jul. 2003.

[56] D. Chen, S. Wan, and F. S. Bao, "Epileptic focus localization using discrete wavelet transform based on interictal intracranial EEG," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 5, pp. 413–425, May 2017.

[57] Y.-T. Chen and M. C. Chen, "Using chi-square statistics to measure similarities for text categorization," *Expert Syst. Appl.*, vol. 38, no. 4, pp. 3085–3090, Apr. 2011.

[58] T. Alotaiby, F. E. A. El-Samie, S. A. Alshebeili, and I. Ahmad, "A review of channel selection algorithms for EEG signal processing," *EURASIP J. Adv. Signal Process.*, vol. 2015, no. 1, pp. 1–21, Dec. 2015.

[59] S.-H. Lee, M. Lee, and S.-W. Lee, "Neural decoding of imagined speech and visual imagery as intuitive paradigms for BCI communication," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 28, no. 12, pp. 2647–2659, Dec. 2020.

[60] P. Wang, A. Jiang, X. Liu, J. Shang, and L. Zhang, "LSTM-based EEG classification in motor imagery tasks," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 26, no. 11, pp. 2086–2095, Nov. 2018.

[61] H.-J. Ahn, D.-H. Lee, J.-H. Jeong, and S.-W. Lee, "Multiscale convolutional transformer for EEG classification of mental imagery in different modalities," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 31, pp. 646–656, 2023.

[62] A. P. Bradley, "The use of the area under the ROC curve in the evaluation of machine learning algorithms," *Pattern Recognit.*, vol. 30, no. 7, pp. 1145–1159, Jul. 1997.