FEATURE-ENHANCED CONSENSUS GRAPH MODEL FOR EEG-BASED IMAGINED WORD RECOGNITION

Syed Hur Abbas[1], *Peter Birkholz*[1], *Muhammad Arif*[1]

[1]*Institute of Acoustics and Speech Communication, Technische Universität Dresden, Dresden, Germany*

muhammad.arif1@tu-dresden.de

Abstract: For people with speech impairments, imagined speech provides a way to communicate by mentally formulating words without engaging vocal organs. The good temporal resolution, portability, safety, and cost-effectiveness of Electroencephalography (EEG) have made it a prominent non-invasive method for studying imagined speech. Recently, a multi-view graph fusion model was proposed, named as KGMV, which learned a consensus graph from multiple EEG features in a Reproducing Kernel Hilbert Space (RKHS). An accuracy of 81.73% was reported on 5-class imagined word classification using a limited set of temporal, statistical, and frequency-domain views. Much of the potentially informative feature space thus remained unexplored. In this work, we extended KGMV to a Feature Enhanced KGMV (FE-KGMV) framework that (1) augmented the feature pool with additional temporal, statistical, and frequency-domain features and (2) applied a subject-wise view-ranking step based on a fused discriminability score (MV-Fisher, ANOVA F, CV-LDA) to retain only the most informative features before graph learning. On the same dataset, FE-KGMV reached a mean accuracy of 86.27%, corresponding to a 4.54 percentage-point improvement over the baseline KGMV.

1 Introduction

Imagined speech refers to the silent production of words in a person's mind, where the intended utterance is formed neurally without any overt articulation or movement of the speech organs (e.g., tongue, lips, or larynx). For individuals with severe speech or motor impairments, such mentally produced speech offers a potential control signal for brain–computer interfaces (BCIs) that aim to restore communication without relying on residual muscular activity. EEG is particularly attractive for this purpose because it is non-invasive, portable, relatively inexpensive, and offers good temporal resolution for tracking fast speech-related neural dynamics. However, decoding imagined speech from EEG remains substantially more challenging than conventional automatic speech recognition, largely due to low signal-to-noise ratios, strong inter-subject variability, and the open question of which EEG features and modeling strategies best capture speech imagery processes [1, 2].

Existing EEG-based imagined and covert speech decoding approaches broadly fall into two families [1, 2, 3, 4, 5, 6]. The first family relies on engineered EEG feature representations (e.g., temporal, statistical, etc.) combined with classical machine-learning classifiers (e.g., support vector machines (SVMs), linear discriminant analysis (LDA), etc.) [1, 2, 5, 7]. Existing literature reports that subject-dependent accuracies for vocabularies of about 5–10 imagined words often fall in the 60–80% range, while performance typically decreases when the task is scaled to larger vocabularies or more demanding cross-session or cross-subject validation protocols [2, 4, 5]. This pattern suggests that, despite careful feature engineering, imagined-speech decoders remain sensitive to task design and recording conditions, and that there is still room to

improve how EEG features are organized and combined for a given subject. The second family employs deep neural networks to learn task-relevant representations directly from minimally processed EEG or from structured inputs such as topographic maps and scalograms. Representative approaches include convolutional and hybrid convolutional–recurrent architectures that capture spatial and temporal patterns, reporting average accuracies in the mid-70% range on imagined-speech datasets [8, 9, 7]. Although these models are highly expressive, their performance is strongly dependent on architectural design, input representation, and training strategy, particularly given the limited size of available imagined-speech EEG datasets [3, 4, 5, 6]. This sensitivity complicates reproducibility and limits interpretability, as the learned representations are not easily linked to established EEG features [4, 5].

A more recent study organized the diverse EEG feature domains explicitly through multi-view learning and graph-based modeling. Zhao et al. [10] introduced the kernel-based multi-view consensus graph model (KGMV), which represented temporal, statistical, and frequency-domain features as separate views and learned a shared similarity graph in a reproducing kernel Hilbert space (RKHS) jointly with non-negative view weights. On the BCI2020 imagined-word dataset [11], KGMV reported a mean validation accuracy of 81.73% for five distinct classes, outperforming single-view baselines and providing a strong reference point for multi-view EEG modeling in imagined speech decoding. However, the proposed KGMV relied on a relatively small, fixed set of feature views that was applied uniformly across all subjects. As a result, it did not systematically probe the broader feature space that could be derived from EEG time series and did not account for potential subject-specific differences in which feature families were most informative. These limitations motivated the present work, in which we extended the KGMV framework to a feature-enhanced and subject-adaptive variant.

In this paper, we introduce a Feature-Enhanced KGMV (FE-KGMV) framework for EEG-based imagined word recognition. Building on the RKHS-based multi-view consensus graph model of Zhao et al. [10], our contribution is threefold: (i) we substantially expand the feature pool beyond the baseline KGMV configuration by extracting 18 temporal, 27 frequency-domain, and 13 statistical features, (ii) we construct subject-wise view configurations via a fused discriminability ranking that combines a multivariate Fisher criterion, univariate ANOVA F-statistics, and cross-validated LDA accuracy into a single score per view, retaining only the top-ranked views for each subject prior to consensus-graph learning; and (iii) we perform a systematic grid search over the key kernel-graph hyperparameters to select an optimal configuration under the same dataset and validation protocol as the baseline. The proposed FE-KGMV is evaluated on the BCI2020 imagined-word corpus (15 subjects, 5 word classes) [11] and compared against the baseline KGMV. FE-KGMV increases the mean validation accuracy from 81.73% to 86.27%, yielding improvements for 12 of the 15 subjects while maintaining matched performance for the remaining three. These results indicate that a carefully enriched, subject-adaptive feature–view design can substantially strengthen consensus-graph-based decoding of imagined words from EEG.

2 Methodology

This section described the pipeline of the proposed FE-KGMV framework. The central idea was to transform each imagined-speech EEG trial into a compact set of subject-adaptive multi-view representations and to perform semi-supervised decoding through RKHS-based consensus graph learning, following the optimization framework of Zhao *et al.* [10]. The pipeline comprised three main stages. First, the EEG was preprocessed to suppress irrelevant spectral components and improve signal-to-noise characteristics, after which complementary feature views were constructed from the temporal, frequency-domain, and statistical perspectives. These fea-

tures were computed channel-wise and then concatenated to form view-specific trial vectors, yielding a multi-view representation that preserved both spatial (across channels) and domain-specific information. Second, to retain only the most informative features for a given subject, we applied a ranking-based retention step: an initially broad candidate pool was scored and ordered using a fused discriminability criterion that combined class-separability and predictive utility signals, and only the top-ranked elements were carried forward. Third, the selected views were integrated by RKHS consensus-graph learning to infer labels for unlabeled samples in a semi-supervised setting. The overall processing flow was illustrated in Figure 1.

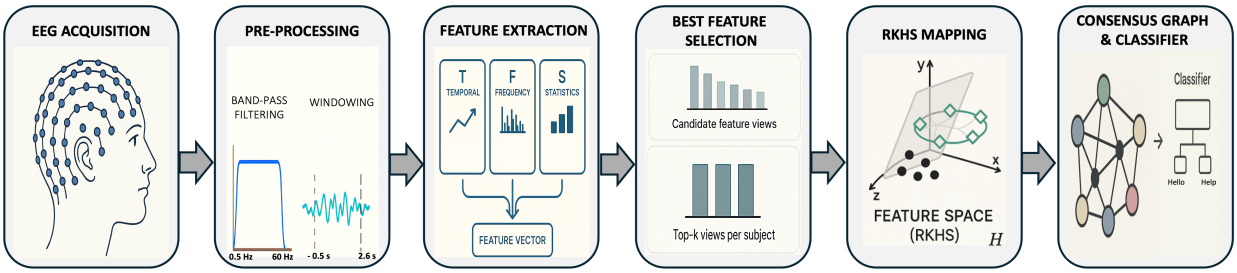

Figure 1 – General flow of the FE-EEG imagined-speech decoding pipeline

2.1 Preprocessing and Feature Extraction

To reduce irrelevant spectral components and enhance the signal-to-noise ratio, all EEG channels for each trial were filtered using a zero-phase Butterworth band-pass filter with fixed cutoff frequencies applied uniformly across the dataset. No subject- or class-specific tuning was introduced at this stage. After preprocessing, each trial was transformed into a set of feature-domain views. For each channel, we extracted three complementary types of features: temporal-domain features that captured waveform morphology and short-range dynamics, frequency-domain features that characterized spectral energy distribution and shape, and statistical features that described distributional and complexity properties. The complete set of features explored was summarized in Tables 1–3 and was computed using standard signal processing definitions in EEG and biosignal analysis. Feature extraction was implemented in Python using `NumPy`, `SciPy`, and the built-in `math` module. Following feature extraction, each feature dimension was normalized independently for each subject using min-max scaling to the range $[0, 1]$. This normalization prevented features with larger numerical ranges from dominating kernel distances and ensured balanced contributions across different feature domains.

Table 1 – Temporal-domain EEG features.

No.	Feature	No.	Feature	No.	Feature
1	Zero-Crossing Rate	7	Mean Absolute Value	13	Autoregressive Coefficient 1
2	Waveform Length	8	Root Mean Square	14	Autoregressive Coefficient 2
3	Slope Sign Changes	9	Peak-to-Peak Amplitude	15	Autoregressive Coefficient 3
4	Willison Amplitude	10	Crest Factor	16	Autoregressive Coefficient 4
5	Hjorth Activity	11	Hjorth Complexity	17	Autoregressive Coefficient 5
6	Hjorth Mobility	12	Difference Absolute Standard Deviation	18	Autoregressive Coefficient 6

Multi-view construction. All features were computed independently per channel and concatenated to preserve channel-wise structure within each view. Let $\mathbf{f}_m^{(i),v} \in \mathbb{R}^{d_v}$ denote the feature vector extracted from channel m of trial i for feature domain v. The corresponding trial-level

Table 2 – Frequency-domain EEG features.

No.	Feature	No.	Feature	No.	Feature
1	Delta Band Power	10	Relative Delta Power	19	Theta-to-Alpha Ratio
2	Theta Band Power	11	Relative Theta Power	20	Beta-to-Alpha Ratio
3	Alpha Band Power	12	Relative Alpha Power	21	Theta-to-Beta Ratio
4	Low-Beta Band Power	13	Relative Low-Beta Power	22	Spectral Centroid
5	High-Beta Band Power	14	Relative High-Beta Power	23	Spectral Entropy
6	Low-Gamma Band Power	15	Relative Low-Gamma Power	24	Spectral Spread
7	Mid-Gamma Band Power	16	Relative Mid-Gamma Power	25	Spectral Roll-off
8	High-Gamma Band Power	17	Relative High-Gamma Power	26	Peak Frequency
9	Total Power	18	Spectral Flatness	27	Total Spectral Power

Table 3 – Statistical and complexity EEG features.

No.	Feature	No.	Feature	No.	Feature
1	Skewness	6	Sample Entropy	11	Interquartile Range
2	Kurtosis	7	Approximate Entropy	12	Mean Absolute Deviation
3	Shannon Entropy	8	Lempel–Ziv Complexity	13	Detrended Fluctuation Analysis
4	Permutation Entropy	9	Higuchi Fractal Dimension		
5	Petrosian Fractal Dimension	10	Hurst Exponent		

representation was

$$\mathbf{x}^{(i),v} = \left[(\mathbf{f}_1^{(i),v})^\top, \ldots, (\mathbf{f}_M^{(i),v})^\top\right]^\top \in \mathbb{R}^{Md_v}. \tag{1}$$

Stacking all trials yielded the view matrix

$$X^v = \left[\mathbf{x}^{(1),v}, \ldots, \mathbf{x}^{(n),v}\right]^\top \in \mathbb{R}^{n\times(Md_v)}. \tag{2}$$

The resulting multi-view representation was $\mathscr{X} = \{X^1, X^2, \ldots, X^p\}$, which was subsequently subjected to subject-wise scoring and view retention as described in Section 2.2.

2.2 Feature Scoring and Selection

Following the construction of multi-view representations in Section 2.1, a subject-wise feature scoring and selection step was applied to reduce redundancy and retain the most discriminative features prior to graph learning. This step was necessary due to the broad initial feature pool obtained from multiple channels and domains. Importantly, feature selection was performed independently for each subject, allowing the retained feature subset to adapt to subject-specific neural characteristics. For each feature dimension k, three complementary criteria were computed: the Fisher score J_k, the one-way ANOVA F-statistic F_k, and the cross-validated linear discriminant analysis (LDA) accuracy CV_k. These criteria respectively quantified class separability via between-class mean dispersion, variance ratio, and predictive utility under a simple linear classifier. Specifically, the Fisher score was defined as

$$J_k = \frac{\sum_{c=1}^{C} N_c \, (\mu_{c,k} - \mu_k)^2}{\sum_{c=1}^{C} N_c \, \sigma_{c,k}^2}, \tag{3}$$

and the ANOVA F-statistic was given by

$$F_k = \frac{\frac{1}{C-1}\sum_{c=1}^{C} N_c \, (\mu_{c,k} - \mu_k)^2}{\frac{1}{N-C}\sum_{c=1}^{C}\sum_{i:y_i=c} \left(x_{i,k} - \mu_{c,k}\right)^2}, \tag{4}$$

where C denotes the number of classes, N_c the number of samples in class c, $\mu_{c,k}$ and $\sigma_{c,k}^2$ the class-wise mean and variance of feature k, μ_k the global mean, and N the total number of

samples. The cross-validated LDA score was computed as

$$CV_k = \frac{1}{K}\sum_{j=1}^{K} \mathrm{Acc}_k^{(j)}, \tag{5}$$

where $\mathrm{Acc}_k^{(j)}$ denotes the classification accuracy obtained using only feature k on fold j of a K-fold cross-validation.

To ensure comparability across heterogeneous scoring criteria, each score was normalized across feature dimensions in a subject-wise manner. Concretely, J_k, F_k, and CV_k were independently rescaled to the unit interval via min–max normalization over all features, yielding the normalized scores $\tilde{J}_k$, $\tilde{F}_k$, and $\widetilde{CV}_k$. The three criteria were then fused into a single ranking score,

$$\mathrm{Score}_k = \frac{1}{3}\tilde{J}_k + \frac{1}{3}\tilde{F}_k + \frac{1}{3}\widetilde{CV}_k, \tag{6}$$

and features were ranked accordingly. For each subject, only the top-ranked subset was retained. The resulting feature set was referred to as the *best feature set* (X_b) and the retained features were regrouped into p domain-specific views, yielding the subject-specific multi-view representation

$$\mathscr{X}_b = \{X_b^1, X_b^2, \ldots, X_b^p\},$$

which served as the input to the subsequent consensus graph learning stage. Note that despite correlations between scoring criteria, preliminary findings revealed subject-dependent variability; therefore, an equal-weight fusion was employed as a conservative ranking heuristic to stabilize feature selection prior to graph learning.

2.3 Consensus Graph Learning with Multi-View Best Features

Based on the selected best multi-view feature representation $\mathscr{X}_b$, we followed the consensus graph learning framework proposed by Zhao *et al.* [10] to construct a unified affinity graph for semi-supervised classification. Each view $X_b^v \in \mathbb{R}^{n\times d_b^v}$ corresponded to the retained features of domain v, where d_b^v denotes the dimensionality after subject-wise feature selection.

Let $\mathbf{Y} \in \mathrm{Ind}^{n\times c}$ be the class indicator matrix with c classes. In the semi-supervised setting, the dataset consisted of l labeled and u unlabeled samples ($n = l+u$), and the objective was to infer the unknown labels $\mathbf{F}_u \in \mathrm{Ind}^{u\times c}$ for the unlabeled samples.

Single-view graph learning. For a given view, similarity between samples $x_{b,i}$ and $x_{b,j}$ was modeled via an affinity matrix $S \in \mathbb{R}^{n\times n}$ by solving

$$\min_S \sum_{i,j}^{n} \left(\|x_{b,i} - x_{b,j}\|_2^2 s_{ij} + \lambda s_{ij}^2 \right), \quad \text{s.t. } C_1(S) = \{s_{ij} \geq 0,\ s^i 1 = 1\},\ \mathrm{rank}(S) = c, \tag{7}$$

where λ is a regularization parameter and s^i denotes the i-th row of S.

Multi-view consensus graph learning. To integrate information from all best-feature views, a consensus graph S and adaptive view weights α were jointly learned by minimizing

$$\min_{S,\alpha} \sum_{v=1}^{p} \alpha^v \sum_{i,j}^{n} \|x_{b,i}^v - x_{b,j}^v\|_2^2 s_{ij} + \lambda \|S\|_F^2 + \beta \|\alpha\|_2^2, \quad \text{s.t. } C_1(S),\ C_2(\alpha),\ \mathrm{rank}(S) = c, \tag{8}$$

where $C_2(\alpha) = \{\alpha^v \geq 0,\ \sum_v \alpha^v = 1\}$ enforces convex combination of views.

RKHS-based formulation. To capture nonlinear relations, each view was implicitly mapped into a reproducing kernel Hilbert space (RKHS) via $\phi(\cdot)$. The objective became

$$\min_{S,\alpha} \sum_{v=1}^{p} \alpha^v \sum_{i,j} h_{ij}^v s_{ij} + \lambda \|S\|_F^2 + \beta \|\alpha\|_2^2, \quad \text{s.t. } C_1(S),\, C_2(\alpha),\, \text{rank}(S) = c, \tag{9}$$

where

$$h_{ij}^v = \|\phi(x_{b,i}^v) - \phi(x_{b,j}^v)\|_2^2 = \kappa(x_{b,i}^v, x_{b,i}^v) - 2\kappa(x_{b,i}^v, x_{b,j}^v) + \kappa(x_{b,j}^v, x_{b,j}^v), \tag{10}$$

and $\kappa(\cdot,\cdot)$ denotes the kernel function.

Ky Fan relaxation and label inference. The rank constraint was relaxed using Ky Fan's theorem, yielding a label-consistency regularization term $\text{Tr}(F^\top L_S F)$, where L_S is the graph Laplacian. The final optimization problem was

$$\min_{S,F_u,\alpha} \sum_{v=1}^{p} \alpha^v \sum_{i,j} h_{ij}^v s_{ij} + \lambda \|S\|_F^2 + \beta \|\alpha\|_2^2 + \sigma\, \text{Tr}(F^\top L_S F), \tag{11}$$

subject to $C_1(S)$, $C_2(\alpha)$, $F = [F_l; F_u]$, $F_u \geq 0$, and $F_u 1 = 1$. The variables S, α, and F_u were alternately updated to obtain a discriminative consensus graph over the selected best-feature views. For detailed intermediate steps, please refer to [10].

3 Results

We evaluated FE-KGMV on the BCI-2020 imagined-speech corpus under the same subject-dependent validation protocol used in KGMV [10]. The task was 5-class imagined-word recognition (*hello*, *help me*, *stop*, *thank you*, *yes*) with 15 subjects (S01–S15), recorded with 64 EEG channels at 256 Hz. For each subject, each imagined word was repeated 70 times (70 trials per class), of which 60 trials were used for training and 10 trials were used for validation, resulting in 350 trials per subject in total. Each subject was evaluated using this fixed split in which training data were used to construct the kernelized multi-view representations and the consensus graph, while the held-out trials were used only for validation-time model selection and reporting. In all runs, extracted features were normalized to $[0, 1]$.

We began with a broad set of subject-wise extracted multi-domain features comprising 18 temporal, 27 frequency-domain, and 13 statistical features, as detailed in Tables 1–3. From this initial feature pool, we conducted a preliminary evaluation to assess the performance of FE-KGMV across multiple feature subsets. Specifically, we examined temporal, frequency, and statistical features individually, as well as their combination. The results revealed that at the individual feature level, temporal features provided the best results, while the combined feature set achieved the highest overall performance. To further reduce redundancy and computational complexity, the feature-ranking criterion was applied jointly to the temporal, frequency, and statistical features described in Section 2.2 for each subject, without imposing domain-specific quotas. Features were retained in descending order of score while monitoring validation accuracy, and the selection process was terminated once performance saturated and additional features consistently contributed negligibly to the learned multi-view weighting. This subject-adaptive selection was subsequently used to construct the final set of views for the graph learner.

In addition to selecting the subject-specific best feature configuration, we conducted an extensive grid search over the kernel–graph hyperparameters $(\gamma, \lambda, \beta, \delta)$ to identify the optimal setting based on validation accuracy. Specifically, the Gaussian kernel scale γ was searched over the values 0.5, 1.0, 2.0, 5.0, and 10.0. The view-weight regularization parameter β was

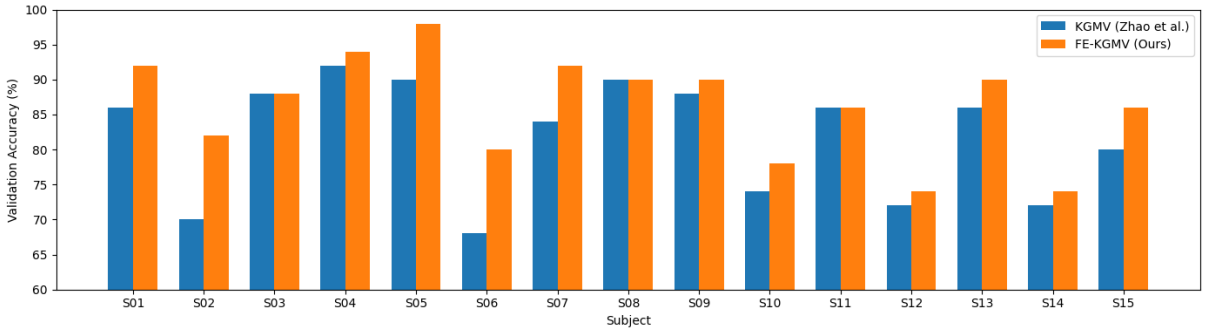

Figure 2 – Per-subject validation accuracy comparison between KGMV and FE-KGMV

selected from 0.0625, 0.125, 0.25, 0.5, 1.0, and 2.0, the graph regularization parameter λ from 0.03125, 0.0625, 0.125, 0.25, 0.5, 1.0, 2.0, and 4.0, and the graph sparsity parameter δ from 16, 32, 64, 128, 256, 512, 1024, and 2048. For each subject, the hyperparameter combination yielding the highest validation accuracy was retained for subsequent evaluation.

The performance of the proposed FE-KGMV and the baseline KGMV, measured in terms of per-subject validation accuracy, was summarized in Table 4. A corresponding comparative bar chart was presented in Figure 2. It could be observed that FE-KGMV improved performance for 12 out of 15 subjects and achieved matched performance for the remaining 3 subjects. This indicated that the enriched feature space and the feature selection strategy introduced additional discriminative information without degrading performance in cases that were already well separated. The largest improvements were observed for subjects S02 and S06, with accuracies increasing from 70% to 82% and from 68% to 80%, respectively. Further notable gains were achieved for subjects S05 and S07, where performance improved from 90% to 98% and from 84% to 92%. For subjects in which the baseline KGMV already performed strongly (e.g., S03, S08, and S11), FE-KGMV maintained comparable accuracy, suggesting that the subject-wise ranking strategy was conservative and did not replace strong evidence with weaker alternatives. Overall, FE-KGMV increased the mean validation accuracy from 81.73% (KGMV) to **86.27%**, corresponding to an absolute improvement of **4.54** percentage points.

Table 4 – KGMV vs. FE-KGMV — Per-subject validation accuracy (%).

Subject	S01	S02	S03	S04	S05	S06	S07	S08	S09	S10	S11	S12	S13	S14	S15	**Mean**
KGMV	86	70	88	92	90	68	84	90	88	74	86	72	86	72	80	81.73
FE-KGMV	**92**	**82**	88	**94**	**98**	**80**	**92**	90	**90**	**78**	86	**74**	**90**	**74**	**86**	**86.27**
Gain	**+6**	**+12**	0	**+2**	**+8**	**+12**	**+8**	0	**+2**	**+4**	0	**+2**	**+4**	**+2**	**+6**	**+4.54**

4 Conclusion

We introduced Feature-Enhanced KGMV (FE-KGMV) for EEG-based imagined word recognition by extending the RKHS multi-view consensus graph framework of Zhao et al. [10] with an expanded feature pool and subject-adaptive view selection. Specifically, we extracted complementary temporal, frequency-domain, and statistical features, and constructed compact, per-subject view configurations using a discriminability score that integrated MV-Fisher criteria, ANOVA F-statistics, and cross-validated LDA accuracy. The proposed FE-KGMV increased the mean validation accuracy from 81.73% to 86.27%, a gain of 4.54 percentage points compared to baseline KGMV. These results indicated that strengthening the input representation through principled feature enrichment and tailored view retention could substantially enhance

consensus-graph-based decoding without altering the underlying kernel-graph learning mechanism. Since the performance of FE-KGMV is subject-dependent, future work will focus on extending the framework by incorporating deep neural networks to enable a generalized learning setting in which subjects are modeled jointly rather than independently. This direction aims to bridge subject-specific graph learning with cross-subject generalization, thereby improving scalability and practical deployment in real-world BCI systems.

References

[1] LOPEZ-BERNAL, D., D. BALDERAS, P. PONCE, and A. MOLINA: *A state-of-the-art review of EEG-based imagined speech decoding*. *Frontiers in Human Neuroscience*, 16, p. 867281, 2022. doi:10.3389/fnhum.2022.867281.

[2] PANACHAKEL, J. T. and A. G. RAMAKRISHNAN: *Decoding covert speech from EEG: A comprehensive review*. *Frontiers in Neuroscience*, 15, p. 642251, 2021. doi:10.3389/fnins.2021.642251.

[3] ZHANG, L., Y. ZHOU, P. GONG, and D. ZHANG: *Speech imagery decoding using EEG signals and deep learning: A survey*. *IEEE Transactions on Cognitive and Developmental Systems*, 17(1), pp. 22–39, 2024. doi:10.1109/TCDS.2024.3431224.

[4] JIN, Z., D. LI, and S. HUANG: *A systematic review of EEG-based imagined speech decoding*. *Applied Soft Computing*, 183, p. 113563, 2025. doi:10.1016/j.asoc.2025.113563.

[5] ALZAHRANI, S., H. BANJAR, and R. MIRZA: *Systematic review of EEG-based imagined speech classification methods*. *Sensors*, 24(24), p. 8168, 2024. doi:10.3390/s24248168.

[6] SU, K. and L. TIAN: *Systematic review: Progress in EEG-based speech imagery brain–computer interface decoding and encoding research*. *PeerJ Computer Science*, 11, p. e2938, 2025. doi:10.7717/peerj-cs.2938.

[7] ALHARBI, Y. F. and Y. A. ALOTAIBI: *Decoding imagined speech from EEG data: A hybrid deep learning approach to capturing spatial and temporal features*. *Life*, 14(11), p. 1501, 2024. doi:10.3390/life14111501.

[8] LI, F., W. CHAO, Y. LI, B. FU, Y. JI, H. WU, and G. SHI: *Decoding imagined speech from EEG signals using hybrid-scale spatial–temporal dilated convolution network*. *Journal of Neural Engineering*, 18(4), p. 0460c4, 2021. doi:10.1088/1741-2552/ac13c0.

[9] PARK, H.-J. and B. LEE: *Multiclass classification of imagined speech EEG using densely connected convolutional networks with multi-receptive fields*. *Frontiers in Human Neuroscience*, 17, p. 1186594, 2023. doi:10.3389/fnhum.2023.1186594.

[10] ZHAO, Z., Y. PENG, K. CAMILLERI, W. KONG, and A. CICHOCKI: *Imagined speech decoding by learning consensus graph from RKHS-based multi-view EEG features*. *IEEE Signal Processing Letters*, 32, pp. 3944–3948, 2025. doi:10.1109/LSP.2025.3619947.

[11] JEONG, J.-H., J.-H. CHO, Y.-E. LEE, S.-H. LEE, G.-H. SHIN, Y.-S. KWEON, J. D. R. MILLÁN, K.-R. MÜLLER, and S.-W. LEE: *2020 international brain–computer interface competition: A review*. *Frontiers in human neuroscience*, 16, p. 898300, 2022.