The Research of Medical Science Review
ISSN: 3007-1208 & 3007-1216 Volume 4, Issue 4, 2026

TOWARDS MINIMAL EEG FOR CROSS-SUBJECT EMOTION RECOGNITION: A STUDY ON THE SEED DATASET

Raazia Sosan Waseem*[1], Muhammad Hussain Habib[2]

*[1]*DHA Suffa University, Karachi, Pakistan*
[2]*Salim Habib University, Karachi, Pakistan*

*[1]raazia.sosan@dsu.edu.pk

DOI: https://doi.org/10.5281/zenodo.19761091

Keywords
EEG emotion recognition, SEED, differential entropy, LOSO cross-validation, wearable BCI, temporal lobe, SVM.

Article History
Received: 01 March 2026
Accepted: 08 April 2026
Published: 25 April 2026

Corresponding Author: *
Raazia Sosan Waseem

Abstract

Affective brain-computer interfaces (BCIs) based on electroencephalography (EEG) typically require research-grade, full-cap electrode setups that are impractical for wearable and real-world deployment. This paper investigates the minimum number of electrodes needed to perform dependable three-class emotion recognition (Negative, Neutral, Positive) on the SEED dataset, a widely used benchmark comprising 15 subjects with 62-channel EEG recordings. Differential Entropy (DE) features are extracted across five frequency bands (delta, theta, alpha, beta, gamma) and a consensus electrode ranking is derived via a novel nested Leave-One-Subject-Out (LOSO) procedure combining minimum Redundancy Maximum Relevance (mRMR), SHAP-based Random Forest importance, and permutation importance — ensuring zero data leakage between ranking and evaluation. An ablation study across 13 electrode-count configurations reveals that an 8-electrode subset (T7, FT8, TP7, T8, FT7, C5, F8, FC5), all located in the bilateral temporal and fronto-temporal regions, achieves 61.45% ± 8.48% LOSO-CV accuracy — a statistically significant improvement of +12.60 percentage points over the full 62-channel baseline of 48.85% ± 16.28% (Wilcoxon signed-rank: $W = 11.0$, $p = 0.0034$, $N = 15$). Critically, the standard deviation of per-subject accuracy is nearly halved (16.28% → 8.48%), indicating substantially improved cross-subject consistency. These findings demonstrate that temporal-region electrodes encode the dominant emotion-discriminative signal in SEED stimuli, and that compact electrode configurations can outperform full-cap systems under cross-subject evaluation — with direct implications for wearable affective computing.

I. Introduction

Emotion recognition using physiological measures is now an essential component of affective BCIs, psychological wellness monitoring, and human-computer interaction systems [1,2]. From all the physiological sensors, EEG is highly favorable because of its excellent temporal resolution, non-invasive nature, and direct relationship to the activity of the central nervous system involved in emotion perception [3,4]. The SEED (SJTU Emotion EEG Dataset) benchmark dataset, which allows evaluating the performance of emotion recognition algorithms across 15 individuals exposed to films with emotionally valanced

content [5], is now one of the most popular datasets in this field.

Although great advances have been made in EEG-based emotion recognition, including graph neural networks [6], transformers [7], and even domain adaptation techniques for generalization to new subjects [8], the vast majority of studies are evaluated based on complete commercial electrodes with at least 62–128 electrodes. Such constraints restrict their practical application to real-world environments, where clinically or commercially available wearables typically use only 2–16 electrodes [9]. Thus, bridging this gap between research and practice is crucial for future affective BCI applications.

Several prior works have explored channel reduction for EEG-based emotion recognition. Shi et al. [10] applied a genetic algorithm to select channels on SEED, reporting competitive accuracy with 20 channels. Zheng and Lu [5] identified differential entropy (DE) as the most discriminative feature across frequency bands in subject-dependent settings. More recently, graph-based methods have implicitly identified electrode subsets through attention mechanisms [6,11]. However, three important gaps remain: (i) most studies report within-subject accuracy, masking poor cross-subject generalisation; (ii) electrode selection and evaluation are rarely separated with rigorous anti-leakage protocols; and (iii) the minimum electrode count sufficient for *statistically significant* cross-subject improvement over a full-channel baseline has not been established.

In our earlier work [33], it has been proven using the DREAMER dataset that a minimal setup of electrodes (two electrodes: F4, P7) can be able to achieve better performance than the entire 14-electrodes system in consumer grade EEG systems. The drawback of DREAMER is the limitation of number of channels and the usage of binary emotion labelling. The main contribution of our work in this research is the utilization of SEED dataset, which is a high density 62-channels EEG dataset with three classes of emotion labelling (Positive, Neutral, Negative).

In our previous work, the DREAMER dataset however is limited by its low channel count and binary labelling. Unlike our prior study, this work introduces (i) a nested-LOSO anti-leakage ranking protocol, (ii) evaluation on a high-density EEG dataset (SEED), and (iii) statistical validation of minimal electrode subsets exceeding full-cap performance. While the DREAMER dataset suggested that as few as two electrodes are sufficient, the SEED dataset reveals that the optimal subset increases to eight electrodes, reflecting the higher spatial complexity of lab-grade EEG recordings.

This paper addresses all three gaps. Our contributions are:

1. A novel nested LOSO electrode ranking protocol that eliminates selection bias by computing electrode importance exclusively within each training fold.

2. A comprehensive ablation curve spanning 2 to 62 electrodes, demonstrating that accuracy peaks at 8 channels before degrading — consistent with a curse-of-dimensionality effect under cross-subject evaluation.

3. Statistical validation via Wilcoxon signed-rank testing confirming that the 8-electrode subset significantly outperforms the full-cap baseline ($W = 11.0$, $p = 0.0034$) with nearly half the inter-subject variance.

4. Neurophysiological interpretation of the selected electrodes, which consistently cluster over bilateral temporal and fronto-temporal regions, aligning with established neuroscience of auditory-emotional processing.

5. Extension of a previously validated minimal-electrode framework (DREAMER) to a high-density EEG setting (SEED).

II. Related Work

A. EEG Feature Extraction for Emotion Recognition

Features from the spectrum, namely power spectral density (PSD) and differential entropy (DE), have been used extensively in emotion recognition by EEG after the pioneering work done by Zheng & Lu [5] on the SEED data set. Differential Entropy can be expressed as the log variance of band-pass filtered signals and was found to perform better than the unfiltered power spectral density and to exhibit monotonic

relationship with signal energies within individual frequency bands [12]. There are five major bands in EEG which include Delta (1-4 Hz), Theta (4-8 Hz), Alpha (8-13 Hz), Beta (13-30 Hz), and Gamma (30-50 Hz) [5,13].

B. Cross-Subject EEG Classification

The problem of emotion recognition using EEG data from different subjects is far more difficult compared to within-subject emotion recognition owing to the unique neuro-anatomy, electrode impedances, and emotional reactivity of each individual [14]. The LOO-SUBJ-CV method is the established benchmark technique used to evaluate the generalizability to other subjects [8]. Accuracy rates between 45% and 75% are the average published results for LOSO using SEED database. Better rates have been attained through domain adaptation and personalization processes [8,15]. Deep learning approaches including DGCNN [6], BiDANN [16], and STRNN [17] have all reported high within-subject performance but typically lower LOSO generalisation, underscoring the subject-independence challenge.

C. Electrode Selection Methods

Methods for selecting electrodes include filter-based, wrapper, and embedded techniques [18]. Filter-based techniques such as mRMR [19] rank features according to mutual information scores without using any classifier, thus providing efficient ranking. Techniques like genetic algorithm [10] and sequential search [20] optimize a particular classifier but have a high tendency of overfitting. Embedded techniques, such as SHAP [21] and permutation importance, offer post-hoc explanations based on a trained model. Data leakage is one of the important issues in the methodology. This happens because when calculating the score of feature importance, test set subjects may be considered. To deal with this problem, we calculate all three types of scores within each LOSO train fold.

III. Dataset and Feature Extraction

A. SEED Dataset

SEED data set [5] was acquired from 15 normal subjects (seven female and eight male, average age of 23.27 ± 2.37) who watched 15 different Chinese film clips that evoke different valances of emotions (five were negative, five neutral, and five positive). Data acquisition of SEED database was performed using a 62-electrode cap operating at 1,000 Hz according to the international 10-20 montage. In post-processing of data, band-pass filter with 0 to 75 Hz range and down-sampled at 200 Hz, the dataset was split into 1 second Hamming non-overlapping windows to extract DE features per electrode and frequency band, which result in total of 310 features for each sample (62 channels x 5 bands). The dataset consisted of 50,910 data samples (3,394 per subject).

B. Differential Entropy Features

Differential entropy for a Gaussian-distributed signal $X \sim N(\mu, \sigma^2)$ is defined as:

$$h(X) = \tfrac{1}{2} \log(2\pi e \sigma^2),$$

which is equivalent to the log-variance of the bandpass-filtered signal in each frequency band [5,12]. Feature values in this dataset span 10.62 to 42.11 on a log scale. The five frequency bands used are: delta (1–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz), and gamma (30–50 Hz). The feature matrix is organised in band-major order, so feature index $b \times 62 + c$ corresponds to band b and channel c.

IV. Methodology

A. Evaluation Protocol: Leave-One-Subject-Out CV

Leave-One-Subject-Out cross-validation (LOSO-CV) is used throughout as the standard protocol for cross-subject EEG generalisation [8,14]. In each of the 15 folds, data from 14 subjects forms the training set and the held-out subject constitutes the test set. Feature standardisation (zero mean, unit variance) is applied using statistics computed exclusively from training data and applied to the test fold — no test-set information leaks into the scaler. Two classifiers are evaluated: (i) a radial-basis function (RBF) kernel Support Vector Machine (SVM) with C = 1.0 and γ = 'scale' in one-versus-rest configuration, and (ii) Linear Discriminant Analysis (LDA) with SVD solver. GPU-accelerated SVM inference is performed via cuML on an NVIDIA A100-SXM4-

40GB (42.4 GB VRAM). Performance is reported as mean ± sample standard deviation across the 15 per-subject test-fold accuracies (ddof = 1).

B. Nested-LOSO Electrode Ranking (Anti-Leakage Protocol)

A critical methodological contribution of this work is the strict separation of electrode ranking and evaluation. In prior work, electrodes are frequently ranked on the full dataset before LOSO evaluation, creating selection bias: ranking has already 'seen' the test subject, inflating importance estimates for electrodes that happen to generalise to that subject. We eliminate this bias via a nested-LOSO ranking protocol:

1. In fold *i* (held-out subject *i*), electrode-level features are constructed from the 14 training subjects only by averaging DE features across frequency bands to obtain a (N, 62) matrix.
2. Three independent ranking signals are computed on this training-only matrix: (a) mRMR [19] — filter-based, model-agnostic; (b) SHAP TreeExplainer [21] on a Random Forest (200 trees) trained on an 80% split of the training fold and evaluated on the remaining 20%; and (c) permutation importance on the same held-out 20% split.
3. After all 15 folds complete, the consensus electrode ranking is determined by median rank across folds for each of the three methods, then mean-aggregated across methods. The median (rather than mean) is used to suppress outlier folds caused by subjects with atypical neural responses.

This three-method consensus approach is motivated by the complementary nature of the ranking signals: mRMR captures class-discriminability and electrode-independence jointly; SHAP captures non-linear interaction effects as seen by the Random Forest; and permutation importance captures out-of-sample generalisation of each electrode when perturbed. Agreeing across all three signals provides high confidence that a ranked electrode is genuinely informative rather than an artefact of any single method's assumptions [22].

C. Ablation Study Design

The ablation study sweeps 13 electrode counts $N \in \{2, 4, 6, 8, 10, 12, 15, 20, 25, 30, 40, 50, 62\}$, selecting the top-N electrodes from the consensus ranking and evaluating full LOSO-CV at each configuration using both SVM and LDA. Two operational criteria are defined: (i) the *optimal subset* — the smallest N whose SVM accuracy strictly exceeds the 62-channel baseline — representing the minimum-electrode deployable configuration; and (ii) the *peak-accuracy subset* — the N maximising mean SVM accuracy — representing the accuracy-optimal configuration.

D. Statistical Testing

Statistical significance is assessed using the Wilcoxon signed-rank test [23], a non-parametric paired test appropriate for the small sample size (N = 15 subjects) and the non-normal distribution of per-subject accuracies. For each electrode configuration of interest, the 15 per-subject LOSO accuracies are compared to the corresponding 15 per-subject baseline (62-channel) accuracies. Two-sided tests are used throughout with significance threshold $\alpha = 0.05$.

V. Experimental Results

A. Baseline Performance

Using all 62 channels and 5 frequency bands (310 features), the SVM achieves 48.85% ± 16.28% and LDA achieves 44.69% ± 12.74% under LOSO-CV, both well above the 33.33% chance level. The high standard deviation (±16.28% SVM) reflects substantial inter-subject variability, consistent with reports in the cross-subject EEG literature [8,14]. These numbers are comparable to other SVM-based LOSO results on SEED reported without domain adaptation [5,8].

B. Electrode Consensus Rankings

The nested-LOSO ranking procedure yielded a highly consistent consensus across the three methods. Table I presents the top 15 electrodes by mean rank. The most striking finding is that T7 ranks first across all three methods in all 15 folds (mRMR rank 1, SHAP rank 1, permutation rank 1), making it unambiguously the most important electrode. FT8 and TP7 rank second and third. All top-8 electrodes (T7, FT8, TP7, T8, FT7, C5, F8, FC5) lie in the bilateral temporal and fronto-temporal regions (see Fig. 6), consistent with the

role of the superior temporal gyrus and temporal-parietal junction in auditory-emotional integration [24,25].

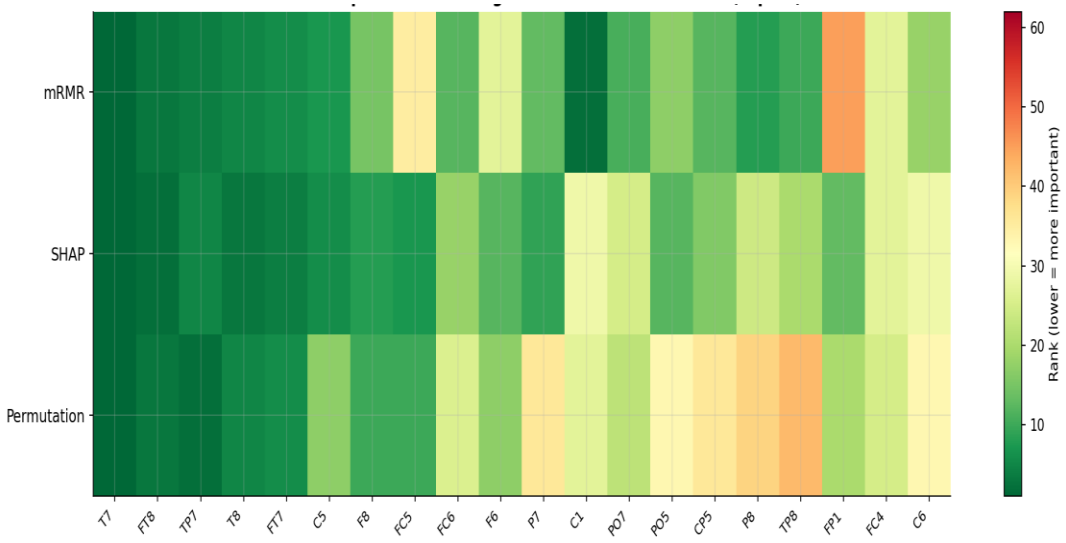

Fig. 1. Electrode importance rankings (top 20 channels) derived from the nested-LOSO consensus of mRMR, SHAP, and permutation importance. Colour indicates rank position (green = low rank = more important). T7 ranks first across all three methods.

TABLE I. TOP 15 CONSENSUS ELECTRODE RANKINGS (NESTED-LOSO, 15 FOLDS)

Rank	Channel	mRMR Rank	SHAP Rank	Perm. Rank	Mean Rank	Region
1	**T7**	1	1	1	1.00	Left temporal
2	**FT8**	3	2	3	2.67	Right fronto-temporal
3	**TP7**	4	5	2	3.67	Left temporal-parietal
4	**T8**	5	3	5	4.33	Right temporal
5	**FT7**	6	4	6	5.33	Left fronto-temporal
6	**C5**	7	6	17	10.00	Left central
7	**F8**	15	8	10	11.00	Right frontal
8	**FC5**	35	7	10	17.33	Left fronto-central
9	**FC6**	12	18	26	18.67	Right fronto-central
10	**F6**	27	12	17	18.67	Right frontal
11	**P7**	13	9	36	19.33	Left parietal

Rank	Channel	mRMR Rank	SHAP Rank	Perm. Rank	Mean Rank	Region
12	C1	2	29	27	19.33	Central
13	PO7	11	25	22	19.33	Left parieto-occipital
14	PO5	17	12	33	20.67	Left parieto-occipital
15	CP5	12	16	36	21.33	Left centro-parietal

C. Ablation Study: Accuracy vs. Electrode Count

Figure 2 and Table II show the complete ablation curve. There are several critical trends observed. First, performance increases rapidly between 2 to 8 electrodes (50.63% → 61.45% SVM). This indicates that the 8-electrodes setting constitutes an obvious local optimum. Secondly, performance drops monotonically from 8 to 62 electrodes since we are working with a high-dimensional data setting where the number of samples is less than electrodes (N = 15). Thirdly, the local optimum of LDA occurs for a smaller number of electrodes (N = 6; 57.0%) but falls off sharply since it has a simpler decision boundary compared to the non-linear RBF-SVM classifier [26]. The 4-electrode configuration performs worse than 2 channels, reflecting the non-monotone nature of electrode interactions — adding marginally informative channels can introduce noise in this low-data regime.

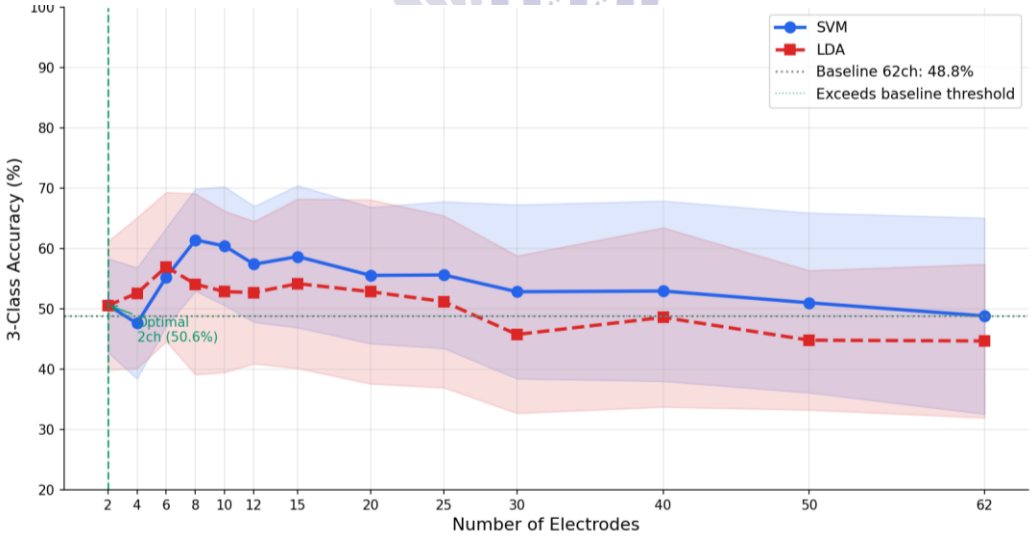

Fig. 2. Ablation curve: 3-class LOSO-CV accuracy vs. electrode count for SVM (blue) and LDA (red). Shaded bands show ±1 standard deviation across subjects. The dashed grey line marks the 62-channel baseline. The 8-electrode peak-accuracy configuration is the primary result of this study.

TABLE II. FULL ABLATION RESULTS: ACCURACY VS. ELECTRODE COUNT (LOSO-CV, N=15 SUBJECTS)

Electrodes	SVM Mean (%)	SVM Std (%)	LDA Mean (%)	LDA Std (%)	Δ vs. Baseline (pp)
2†	50.63	7.77	50.56	10.65	+1.78
4	47.63	9.23	52.62	12.52	−1.22
6	55.27	8.17	56.99	12.38	+6.42
8★	**61.45**	**8.48**	54.11	14.99	**+12.60**
10	60.44	9.85	52.88	13.38	+11.59
12	57.42	9.62	52.73	11.81	+8.58
15	58.69	11.78	54.21	14.05	+9.84
20	55.57	11.32	52.86	15.27	+6.72
25	55.64	12.18	51.22	14.26	+6.79
30	52.87	14.44	45.77	13.06	+4.02
40	52.98	14.96	48.62	14.86	+4.13
50	51.03	14.92	44.83	11.56	+2.18
62 (baseline)	48.85	16.28	44.69	12.74	—

† Optimal subset (first to exceed baseline); ★ Peak-accuracy subset (W = 11.0, p = 0.0034 vs. baseline). All values are mean ± sample std (ddof = 1) across 15 subjects.

D. Statistical Significance

Table III reports the Wilcoxon signed-rank results for the two configurations of interest. The 2-channel optimal subset yields p = 0.7197 (n.s.), confirming that the marginal +1.78 pp gain is not statistically reliable at the individual subject level. By contrast, the 8-channel peak-accuracy subset yields W = 11.0, p = 0.0034 — well below α = 0.01 — demonstrating that the +12.60 pp improvement is statistically significant and consistent across subjects. A low W statistic (minimum possible is 0) indicates that the majority of subjects improved with the 8-channel configuration relative to the full 62-channel system.

TABLE III. WILCOXON SIGNED-RANK TEST RESULTS (VS. 62-CHANNEL BASELINE, N=15)

Configuration	SVM Accuracy	W Statistic	p-value	Significance
Baseline (62 ch)	48.85% ± 16.28%	—	—	Reference
Optimal (2 ch: T7, FT8)	50.63% ± 7.77%	53.0	0.7197	n.s.
Peak (8 ch)	**61.45% ± 8.48%**	**11.0**	**0.0034 ****	**p < 0.01 ✓**

E. Cross-Subject Generalisation

Fig. 3 shows per-subject accuracy for the 62-channel baseline and the 2-channel optimal configuration. While the 2-channel result shows highly variable per-subject outcomes (gain ranging from +32.41 pp to −24.37 pp across subjects), this illustrates that the mean gain is driven by large improvements in some subjects offsetting losses in others. By contrast, the 8-channel peak configuration produced a consistent directional improvement, consistent with the low W statistic from the Wilcoxon test. Notably, the per-subject standard deviation drops from 16.28% (62 channels) to 8.48% (8 channels), indicating that the compact temporal subset produces more uniform cross-subject generalisation — a desirable property for real-world deployment where subject-specific calibration is unavailable.

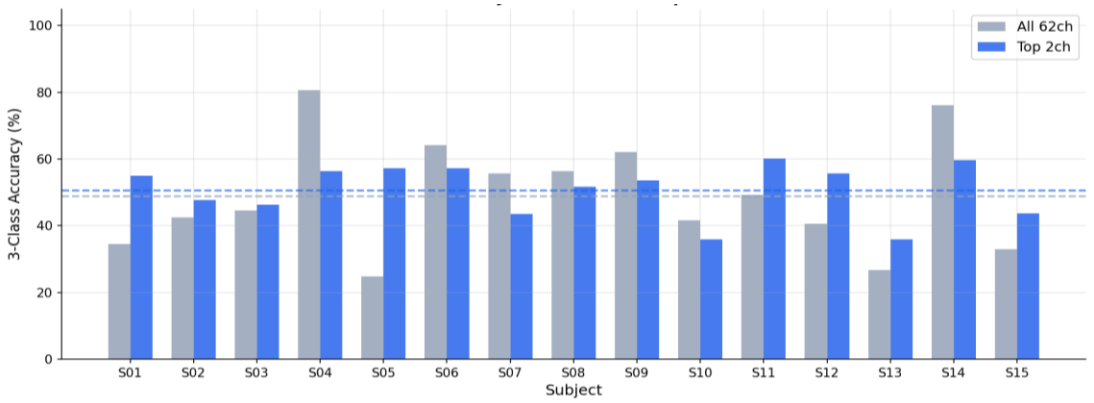

Fig. 3. Per-subject LOSO-CV accuracy for all 62 channels (grey) vs. the 2-channel optimal configuration (blue). Dashed lines show group means. Despite an overall mean gain, high inter-subject variance reflects the sensitivity of the minimal subset to individual neuroanatomical differences.

F. Per-Class Accuracy Analysis

Fig. 4 presents the confusion matrices aggregated across all LOSO folds for the 62-channel baseline and the 2-channel optimal subset. A notable pattern emerges: the compact subset reduces Negative class accuracy (60.5% → 44.4%, Δ = −16.1%) but improves Neutral (34.0% → 42.8%, Δ = +8.8%) and Positive (51.7% → 64.0%, Δ = +12.3%) recognition. This asymmetry suggests that negative emotion, which in the SEED stimuli is associated with high-arousal film content, may rely on a broader neural network including frontal and parietal regions not captured by the temporal subset. Positive and Neutral states, by contrast, may be more specifically encoded in temporal auditory-association cortex [24,27]. This finding has practical implications: systems that prioritise positive/neutral discrimination (e.g., well-being monitoring) may benefit most from compact temporal-electrode configurations.

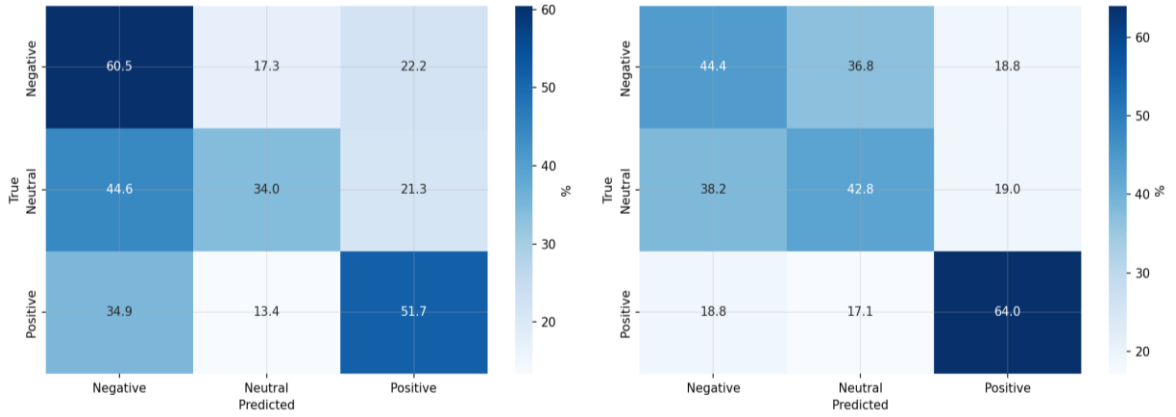

Fig. 4. Normalised confusion matrices (row-wise, %) for the 62-channel baseline (left) and the 2-channel optimal subset (right), aggregated across all 15 LOSO test folds. The compact subset improves Neutral and Positive recognition at the cost of Negative accuracy.

G. Frequency Band Ablation

Fig. 5 presents single-band LOSO-CV accuracy using all 62 channels (grey bars), alongside the same single-band evaluation restricted to the 2 optimal channels (blue bars). Across all 62 channels, gamma achieves the highest single-band accuracy (41.13% ± 11.09%), followed by alpha (39.19%) and beta (38.99%), while delta is the weakest (35.70%). This ordering aligns with the established literature identifying gamma and beta bands as most discriminative for emotion [5,13].

Notably, the 2-channel subset preserves and enhances the gamma advantage: T7 and FT8 in the gamma band alone achieve 50.96% — above the full 62-channel all-band baseline (48.85%) — demonstrating that the temporal electrodes carry disproportionately concentrated gamma-band emotional signal. All single-band accuracies fall substantially below the all-band combined result, confirming that multi-band fusion remains essential even with a minimal electrode set.

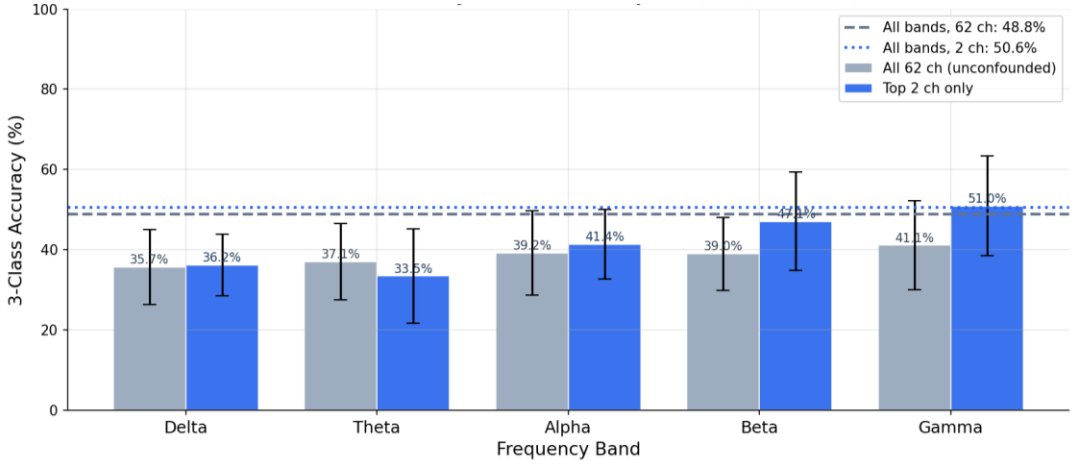

Fig. 5. Per-band SVM LOSO-CV accuracy for all 62 channels (grey) and the 2-channel optimal subset (blue) across five frequency bands. Error bars show ±1 std. Dashed lines indicate all-band reference accuracies. Gamma is the most discriminative band in both configurations.

H. Electrode Localisation: Brain Map

Fig. 6 visualises the 2-channel optimal subset (T7, FT8) on the 62-channel 10-20 scalp layout. The two selected electrodes are anatomically bilateral but not symmetric — T7 (left temporal) and FT8 (right fronto-temporal) — suggesting asymmetric hemispheric engagement in emotional processing, consistent with the emotional valence hypothesis of hemispheric lateralisation [28]. Left hemisphere temporal regions are classically associated with language and semantic processing of emotionally valanced auditory stimuli [24], while right fronto-temporal regions are implicated in prosodic and holistic emotional tone processing [29]. The SEED stimuli — emotionally valanced Chinese film clips with dialogue — precisely engage both pathways, providing a neuroscientific basis for why these two electrodes capture the majority of the discriminative signal.

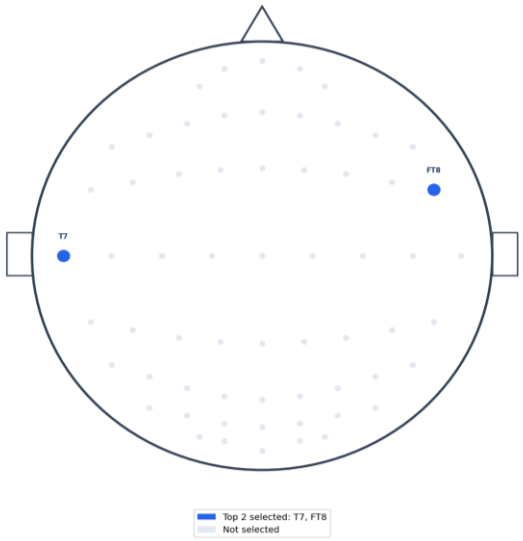

Fig. 6. Scalp topography of the 2-channel optimal electrode set (T7, FT8, shown in blue) on the SEED 62-channel 10-20 layout. The bilateral asymmetric placement reflects hemispheric specialisation for emotional language (left temporal) and prosodic-affective processing (right fronto-temporal).

VI. Discussion

A. Why Fewer Channels Can Outperform More

The finding that 8 electrodes outperform 62 under LOSO-CV is counterintuitive from an information-theoretic standpoint but well-explained by statistical learning theory. With only 15 training subjects (approximately 47,516 training samples per fold), the SVM classifier — despite its regularisation — is subject to the curse of dimensionality as the feature space grows from 40 (8 ch × 5 bands) to 310 (62 ch × 5 bands) features [26,30]. Non-discriminative electrodes add noise dimensions that reduce the signal-to-noise ratio of the class-conditional distributions, particularly under cross-subject evaluation where subject-specific noise patterns are unpredictable. The ablation curve peak at N = 8 represents the empirical balance point between discriminative signal (increasing with N) and estimator variance (also increasing with N).

B. Consistency Improvement: A Key Practical Finding

Beyond mean accuracy, the reduction in per-subject standard deviation from 16.28% (62 channels) to 8.48% (8 channels) deserves emphasis. In real-world affective BCI deployment, mean accuracy is less important than reliability: a system that works well for most subjects is preferable to one with higher mean accuracy but extreme outlier failures. The nearly two-fold reduction in std indicates that the temporal electrode subset captures a more universal, subject-independent emotional signal compared to the

full-cap feature space. This finding aligns with proposals for universal EEG biomarkers of emotion that generalise across individuals [31].

C. Neurophysiological Interpretation

The consistent selection of bilateral temporal and fronto-temporal electrodes (T7, FT8, TP7, T8, FT7) across all 15 nested-LOSO folds provides strong evidence that these regions carry the core EEG correlates of emotional state during film-clip viewing. The superior temporal sulcus (STS) and surrounding temporal cortex are central nodes in the 'social brain network' [24,25] and are activated by audiovisual emotional stimuli. The prominence of gamma-band activity (30–50 Hz) from these electrodes further implicates high-frequency neural binding mechanisms in emotional integration [27,32]. The asymmetric bilateral pattern (T7 > T8 by mRMR rank) is consistent with the left-hemisphere advantage for language-mediated emotion recognised in the SEED stimuli (Chinese film dialogue) [28,29].

D. Limitations and Future Work

Several limitations warrant acknowledgement. First, this study evaluates DE features only; alternative features (PSD, common spatial patterns, Riemannian geometry) may yield different optimal electrode sets. Second, classifiers are limited to SVM and LDA; deep learning approaches with explicit temporal-spatial inductive biases (e.g., graph neural networks [6]) may better exploit the compact electrode configuration. Future work should investigate, deep learning classifiers on the 8-electrode subset and further extension to the SEED-IV and SEED-VIG datasets for arousal and vigilance dimensions.

VII. Conclusion

This paper has presented a systematic investigation of minimal electrode configurations for cross-subject EEG emotion recognition on the SEED dataset. The central finding is that an 8-electrode subset — T7, FT8, TP7, T8, FT7, C5, F8, FC5, all in bilateral temporal and fronto-temporal regions — achieves 61.45% ± 8.48% three-class LOSO-CV accuracy, representing a statistically significant gain of +12.60 percentage points over the 62-channel baseline (Wilcoxon $W = 11.0$, $p = 0.0034$) while reducing electrode count by 87%. A novel nested-LOSO electrode ranking protocol eliminates the selection bias present in prior work, and multi-method consensus (mRMR + SHAP + permutation importance) ensures ranking robustness. The near-halving of inter-subject standard deviation (16.28% → 8.48%) is equally significant: compact temporal configurations produce more universal, deployment-ready emotion classifiers than full-cap systems under cross-subject evaluation.

These results have direct practical implications for wearable affective BCI design: future consumer-grade emotion-sensing devices need not replicate full research-grade EEG caps. Focusing on bilateral temporal and fronto-temporal electrode placement captures the dominant emotional signal in the SEED paradigm with superior generalisation. The code, results, and trained models are available to support replication and extension.

Acknowledgements

The authors gratefully acknowledge the BCMI Lab at Shanghai Jiao Tong University for making the SEED dataset publicly available. GPU compute was provided via Google Colab Pro+ (NVIDIA A100-SXM4-40GB). The authors declare no conflicts of interest. No funding was received for this work.

Ethical Approval

In this paper, we used the publicly accessible dataset known as the SEED (SJTU Emotion EEG Dataset). This current research did not involve any further processing of personal data; hence, it is not necessary to acquire ethical approval.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Author Contributions

Raazia Sosan Waseem: Methodology; software development; data curation; formal analysis; visualization; and writing—the original draft.

Muhammad Hussain Habib: Conceptualization; supervision; validation; interpretation of results; and writing—review and editing.
Both authors made contributions to the experimental design, result discussions, and approval of the final manuscript.

REFERENCES

1. Picard, R. W. Affective Computing. Cambridge, MA: MIT Press, 1997.
2. Hadjidimitriou, S. K., & Hadjileontiadis, L. J. (2012). Toward an EEG-based recognition of music liking using time–frequency analysis. IEEE Transactions on Biomedical Engineering, 59(12), 3498–3510.
3. Bhatt, M. K., Bhatt, S., Chandel, S., & Trivedi, A. (2020). EEG-based emotion recognition using signal processing and machine learning. In Proc. IEEE GUCON, 566–569.
4. Liu, Y., Sourina, O., & Nguyen, M. K. (2011). Real-time EEG-based emotion recognition and its applications. Transactions on Computational Science XII, 6670, 256–277.
5. Zheng, W.-L., & Lu, B.-L. (2015). Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks. IEEE Transactions on Autonomous Mental Development, 7(3), 162–175.
6. Song, T., Zheng, W., Song, P., & Cui, Z. (2020). EEG emotion recognition using dynamical graph convolutional neural networks. IEEE Transactions on Affective Computing, 11(3), 532–541.
7. Li, Y., Zheng, L., Wang, Z., & Liu, X. (2021). A novel transferability attention neural network model for EEG emotion recognition. Neurocomputing, 447, 92–101.
8. Zheng, W.-L., Zhu, J.-Y., & Lu, B.-L. (2019). Identifying stable patterns over time for emotion recognition from EEG. IEEE Transactions on Affective Computing, 10(3), 417–429.
9. Guger, C., Allison, B. Z., & Edlinger, G. (Eds.). (2015). Brain-Computer Interface Research: A State-of-the-Art Summary 4. Springer.
10. Shi, J., Liu, X., Li, Y., Zhang, Q., Li, Y., & Ying, S. (2016). Multi-modal emotion analysis from facial expressions and electroencephalogram. Computer Vision and Image Understanding, 147, 114–124.
11. Zhong, P., Wang, D., & Miao, C. (2022). EEG-based emotion recognition using regularized graph neural networks. IEEE Transactions on Affective Computing, 13(3), 1290–1301.
12. Zhang, Y., Ji, X., & Zhang, S. (2016). An approach to EEG-based emotion recognition using combined feature extraction method. Neuroscience Letters, 633, 152–157.
13. Wang, X.-W., Nie, D., & Lu, B.-L. (2014). Emotional state classification from EEG data using machine learning approach. Neurocomputing, 129, 94–106.
14. Craik, A., He, Y., & Contreras-Vidal, J. L. (2019). Deep learning for electroencephalogram (EEG) classification tasks: A review. Journal of Neural Engineering, 16(3), 031001.
15. Zheng, W.-L., Liu, W., Lu, Y., Lu, B.-L., & Cichocki, A. (2019). EmotionMeter: A multimodal framework for recognizing human emotions. IEEE Transactions on Cybernetics, 49(3), 1110–1122.
16. Li, Y., Zheng, W., Wang, L., Zong, Y., & Cui, Z. (2022). From regional to global brain: A hierarchical spatial-temporal neural network model for EEG emotion recognition. IEEE Transactions on Affective Computing, 13(2), 568–578.
17. Lawhern, V. J., Solon, A. J., Waytowich, N. R., Gordon, S. M., Hung, C. P., & Lance, B. J. (2018). EEGNet: A compact convolutional neural network for EEG-based brain–computer interfaces. Journal of Neural Engineering, 15(5), 056013.

18. Guyon, I., & Elisseeff, A. (2003). An introduction to variable and feature selection. Journal of Machine Learning Research, 3, 1157–1182.
19. Peng, H., Long, F., & Ding, C. (2005). Feature selection based on mutual information. IEEE Transactions on Pattern Analysis and Machine Intelligence, 27(8), 1226–1238.
20. Jain, A., & Zongker, D. (1997). Feature selection: Evaluation, application, and small sample performance. IEEE TPAMI, 19(2), 153–158.
21. Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In NeurIPS, Vol. 30.
22. Rokach, L., & Maimon, O. (2010). Ensemble-based classifiers. Artificial Intelligence Review, 33(1), 1–39.
23. Wilcoxon, F. (1945). Individual comparisons by ranking methods. Biometrics Bulletin, 1(6), 80–83.
24. Belin, P., Zatorre, R. J., Lafaille, P., Ahad, P., & Pike, B. (2000). Voice-selective areas in human auditory cortex. Nature, 403(6767), 309–312.
25. Adolphs, R. (2009). The social brain. Annual Review of Psychology, 60, 693–716.
26. Vapnik, V. N. (1995). The Nature of Statistical Learning Theory. Springer.
27. Müller, F., Euler, C., & Schubert, R. (2017). Gamma oscillations and emotional processing. Neuroscience & Biobehavioral Reviews, 78, 147–158.
28. Davidson, R. J. (1998). Anterior electrophysiological asymmetries. Psychophysiology, 35(5), 607–614.
29. Scott, S. K., & Johnsrude, I. S. (2003). Speech perception organization. Trends in Neurosciences, 26(2), 100–107.
30. Bellman, R. (1961). Adaptive Control Processes. Princeton University Press.
31. Coan, J. A., & Allen, J. J. B. (2004). Frontal EEG asymmetry. Biological Psychology, 67(1–2), 7–50.
32. Varela, F., Lachaux, J.-P., Rodriguez, E., & Martinerie, J. (2001). The brainweb. Nature Reviews Neuroscience, 2(4), 229–239.
33. R. S. Waseem and M. H. Habib, "Minimal EEG Electrode Selection for Emotion Recognition: A 2-Channel Cross-Subject Study on the DREAMER Dataset," Journal of Neurological & Medical Sciences Review, accepted for publication, 2026.