

RASR-Net: Lightweight Cross-Subject EEG Emotion Recognition for Hearing-Impaired Individuals

Wenying Wang

College of Computer Science and Technology, Qingdao University, Qingdao 266071, China

Email: wangwenying1@qdu.edu.cn

How to cite this paper: Wang, W. (2026). RASR-Net: Lightweight Cross-Subject EEG Emotion Recognition for Hearing-Impaired Individuals. Journal of Computer Science and Frontier Technologies, 4(2), 12–19. ISSN Print: 3104-4204, ISSN Online: 3104-4212.

<https://doi.org/10.63313/JCSFT.9099>

Published: 2026-09-28

Copyright © 2026 by author(s) and Erytis Publishing Limited.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Abstract

Cross-subject electroencephalography (EEG) emotion recognition for hearing-impaired individuals is limited by inter-subject distribution shifts, the acquisition burden of high-density systems, and overfitting on small cohorts. This study proposes RASR-Net, a Riemannian-anchored spectral-region lightweight residual network that combines unlabeled within-subject Euclidean alignment, robust five-band statistics, frontal-parietal attention and gated fusion, and a constrained neural correction to a Riemannian logistic anchor. The model was evaluated on 30 hearing-impaired participants using strict leave-one-subject-out validation, with model selection confined to training subjects and three random seeds (90 outer folds). RASR-Net achieved 58.563% accuracy and 54.859% macro-F1 with 7,227 trainable neural parameters. It outperformed same-channel SGD logistic regression by 6.07 percentage points in accuracy and EEGNet by 10.35 points. Using 23 frontal-parietal channels retained 98.94% of the accuracy of a parameter-matched 62-channel counterpart while reducing electrodes by 62.90%, the FLOPs lower bound by 58.09%, and single-thread CPU latency by 47.02%. These results indicate that geometry-anchored residual learning can improve robust cross-subject decoding while preserving a practical lightweight acquisition configuration.

Keywords

EEG emotion recognition; hearing impairment; cross-subject learning; Riemannian geometry; spectral attention; lightweight neural network

1. Introduction

Emotion recognition is a core capability in affective computing and human-computer interaction. Compared with facial, speech, and textual cues, EEG directly reflects central nervous activity, is difficult to disguise voluntarily, and offers high temporal resolution [1-3]. Most existing systems, however, have been developed for normal-hearing populations. Objective affective interfaces for people with hearing impairment may support more inclusive communication and adaptive services [4,5]. The HIED dataset provides EEG responses from 30 hearing-impaired

participants exposed to subtitled film clips, but its cross-subject decoding remains difficult because of individual neurophysiological variability, nonstationary noise, and limited cohort size.

A central methodological risk is subject leakage. Randomly splitting temporal windows can place samples from one person in both training and test sets, producing optimistic results that do not represent deployment to an unseen user. Strict leave-one-subject-out (LOSO) evaluation better reflects this setting but usually lowers performance [6-8]. Meanwhile, high-density EEG increases preparation, maintenance, and computation costs, whereas large convolutional, Transformer, or graph networks may amplify estimation variance in a 30-subject study [9,10].

We therefore propose RASR-Net, which combines stable geometric priors with low-capacity residual learning instead of increasing network depth. The method aligns unlabeled subject distributions, models five-band and frontal-parietal structure explicitly, and uses a Riemannian logistic classifier as a decision anchor. Its main contributions are: (i) a 7,227-parameter neural branch that captures spectral importance, regional complementarity, and short-term context; (ii) a strict LOSO30 protocol with validation-only model selection and three training seeds; and (iii) a parameter-matched 23/62-channel comparison plus paired ablation and class-wise error analyses.

2. Materials and Methods

2.1. Dataset and Evaluation Protocol

HIED contains recordings from 30 hearing-impaired university students (12 women and 18 men; mean age, 22 years). Thirty Chinese-subtitled film clips elicited six emotions. Following the dataset protocol, happiness and inspiration were merged as positive, calmness as neutral, and anger, fear, and sadness as negative. EEG was recorded with a 62-channel Neuroscan SynAmps2 system at 200 Hz. Public preprocessing comprised 1-75 Hz band-pass filtering, 49-51 Hz line-noise suppression, bad-channel interpolation, TP9/TP10 rereferencing, and ICA-based artifact removal [4,5]. The final 180 s of each trial were divided into nonoverlapping 3-s windows, yielding 1,800 samples per participant and 54,000 samples overall.

The lightweight montage contains 14 frontal electrodes (Fp1, Fpz, Fp2, AF3, AF4, F7, F5, F3, F1, Fz, F2, F4, F6, F8) and nine parietal electrodes (P7, P5, P3, P1, Pz, P2, P4, P6, P8). Each LOSO fold held one participant out for testing; 26 of the remaining participants formed the training set and three formed the validation set. Split seed 42 and training seeds 42, 3407, and 2026 produced 90 outer-fold evaluations. Test labels were never used for feature selection, scaling, early stopping, or hyperparameter selection. Euclidean alignment and within-subject standardization used only the target participant's unlabeled EEG; the reported setting is therefore offline unsupervised transductive calibration rather than a fully inductive no-target-data scenario [6].

Table 1. Dataset and experimental configuration.

Item	Configuration
Participants and task	30 hearing-impaired participants; three emotion classes
Samples	$15 \times 23 \times 600$ input context; 54,000 three-second windows
Channels	23 frontal-parietal electrodes; 62-channel matched control
Validation	Strict LOSO30; 26 train, 3 validation, 1 test
Repeated runs	Three training seeds; 90 outer folds
Optimization	AdamW; lr 3×10^{-4} ; weight decay 10^{-3} ; batch 256; max 30 epochs

2.2. RASR-Net

After unlabeled Euclidean alignment, each 3-s window is encoded in five canonical bands: delta (1-4 Hz), theta (4-8 Hz), alpha (8-14 Hz), beta (14-31 Hz), and gamma (31-50 Hz). Robust differential entropy, log power, and relative power form three statistics per band. A compact frequency-attention pathway (Linear 15-16, GELU, Linear 16-5, Softmax) weights the bands, while a Linear 3-32 projection with learned band embeddings provides a 32-dimensional representation. Separate channel-attention blocks summarize the 14 frontal and nine parietal electrodes. A learned gate combines the two regions, and their frontal, parietal, and fused representations are concatenated and projected to 32 dimensions.

$$g = \text{sigmoid}(W_g[h_F \oplus h_P] + b_g), \quad h_M = g \odot h_F + (1 - g) \odot h_P$$

The resulting tensor U has shape $B \times 15 \times 32$. Parameter-free averaging across the 15-window (45-s) context yields e , which enters a residual head composed of Linear 32-32, GELU, dropout 0.15, and Linear 32-3. The parallel anchor branch computes one shrinkage-regularized 23×23 covariance matrix per band, applies trace normalization and a Log-Euclidean map, and vectorizes the upper triangle with the isometric $\sqrt{2}$ convention [11-15]. Concatenating five 276-dimensional vectors gives 1,380 features; the mean of the most recent five windows is standardized using training data only and classified by L2-regularized SGD logistic regression [16,17].

$$C_b = \text{Cov}(X_b) + \varepsilon I, \quad L_b = \log(C_b), \quad v_b = \text{vech}(\sqrt{2}(L_b))$$

Validation chooses a bounded residual scale s from $\{0, 0.5, 1\}$; test logits are obtained by adding the scaled neural correction to the anchor logits. The neural branch is trained with AdamW, gradient clipping at 5, early stopping patience 6, a KL-consistency weight of 0.15, and an L2 residual penalty of 0.01 [18,19]. This bounded formulation keeps the stable geometric decision surface dominant when the neural correction is unreliable.

$$z = z_A + s(0.5r), \quad s \in \{0, 0.5, 1\}, \quad p(y|X) = \text{softmax}(z)$$

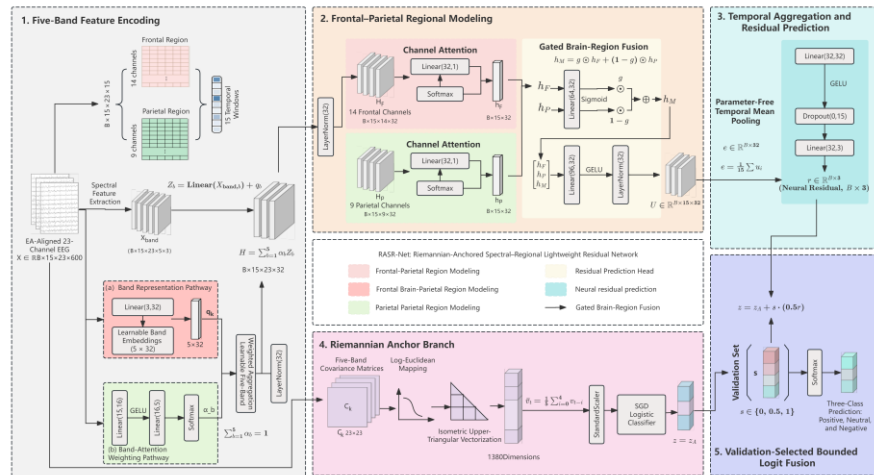


Figure 1. Overall architecture of RASR-Net. The dashed validation path selects the bounded residual scale without using test labels.

3. Results and Discussion

3.1. Recognition Performance and Efficiency

Across three seeds, RASR-Net achieved $58.563 \pm 0.077\%$ accuracy, $56.407 \pm 0.436\%$ balanced accuracy, $54.859 \pm 0.285\%$ macro-F1, $57.711 \pm 0.098\%$ weighted-F1, and $\kappa = 0.3291 \pm 0.0037$. On the identical 23-channel input, it exceeded SGD logistic regression by 6.07 percentage points in accuracy and 6.72 points in macro-F1, and exceeded EEGNet by 10.35 and 8.59 points, respectively. The gain over the lightweight neural residual baseline was 3.02 accuracy points (paired $p = 0.033$) and 2.24 macro-F1 points ($p = 0.168$). These results support the value of constraining neural learning around a stable geometric classifier rather than relying on an unconstrained deep model.

Table 2. Same-channel LOSO30 comparison (%).

Method	Accuracy	Macro-F1
TSception23	43.35	37.43
IAG23	47.48	45.27
EEGNet23	48.21	46.27
Linear SVM	49.81	46.19
SGD logistic	52.49	48.14
Lightweight neural residual	55.46	52.32
RASR-Net	58.56	54.86

The parameter-matched 62-channel model reached 59.100% accuracy and 55.044% macro-F1, compared with 58.476% and 54.563% for the 23-channel model under seed 42. The paired differences were not significant (accuracy $p = 0.452$; macro-F1 $p = 0.718$). Thus, 23 channels retained 98.94% of full-channel accuracy while reducing the electrode count by 62.90%, the FLOPs lower bound by 58.09%, checkpoint size by 24.82%, and single-thread CPU latency by 47.02%. The same 7,227 neural

parameters in both variants isolate the contribution of channel count from model capacity.

Table 3. Parameter-matched channel and efficiency comparison.

Configuration	Accuracy (%)	FLOPs (M)	CPU (ms)	Parameters
23 channels	58.476	0.853	0.566	7,227
62 channels	59.100	2.036	1.068	7,227
23/62 ratio	98.94%	41.90%	52.98%	100%

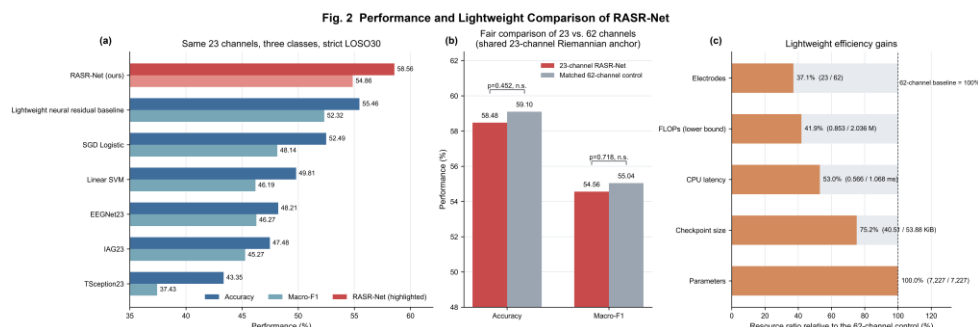


Figure 2. Main recognition performance and lightweight comparison of RASR-Net under strict LOSO30 evaluation.

3.2. Ablation and Error Analysis

Exact seed-42 ablation showed complementary contributions. Removing frequency attention reduced accuracy by 2.14 points and macro-F1 by 3.65 points; removing frontal-parietal fusion reduced them by 2.34 and 3.57 points. Removing the Riemannian anchor caused the largest accuracy loss (4.11 points), whereas the anchor-only model lost 4.24 macro-F1 points. Holm-corrected paired tests were significant for both frequency attention and regional fusion, for anchor removal in accuracy, and for anchor-only macro-F1 [20]. The pattern indicates that the anchor stabilizes cross-subject decisions while the spectral-region residual improves class balance.

Table 4. Exact module ablation under seed 42 (%).

Variant	Accuracy	Δ Accuracy	Macro-F1	Δ Macro-F1
Complete RASR-Net	58.48	—	54.56	—
Without frequency attention	56.34	-2.14	50.91	-3.65
Without regional fusion	56.14	-2.34	50.99	-3.57
Without Riemannian anchor	54.36	-4.11	52.77	-1.79
Riemannian anchor only	56.33	-2.14	50.32	-4.24

The three-seed mean confusion matrix yielded recalls of 60.3% for positive, 48.0% for neutral, and 60.9% for negative emotion. Neutral samples were most frequently confused with negative emotion (32.5%), making the neutral class the main error source. This likely reflects the lower neutral sample count and weaker physiological

separation. Although the proposed design compares favorably with conventional and compact neural baselines, the study remains limited to one dataset and an offline transductive calibration protocol. External cohorts, a strictly inductive setting, and montage optimization below 23 electrodes should therefore be tested before clinical or real-time use. Recent frequency-selection, domain-generalization, and contrastive methods offer relevant directions [22-29].

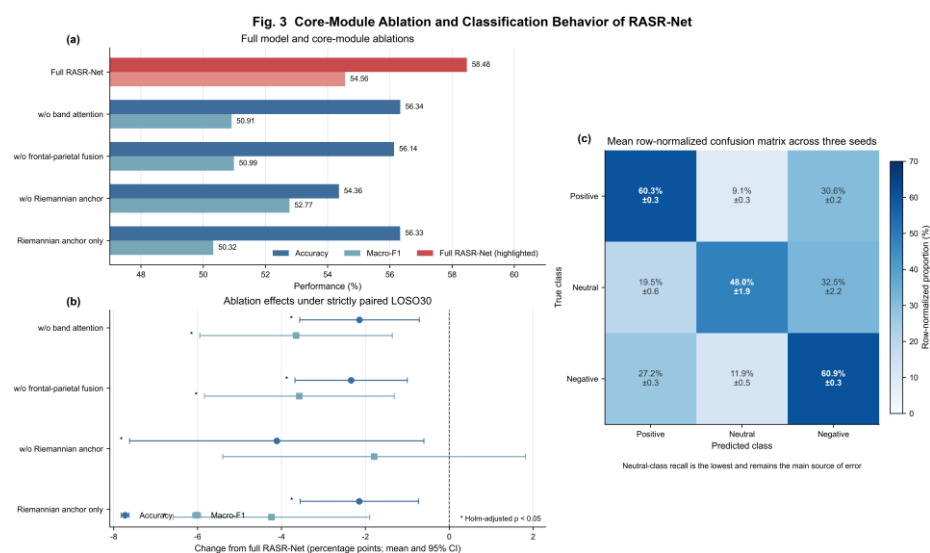


Figure 3. Core-module ablation and classification behavior. Error bars denote paired LOSO30 uncertainty; the confusion matrix is averaged over three seeds.

4. Conclusion

RASR-Net combines unlabeled Euclidean alignment, five-band frontal-parietal modeling, a Log-Euclidean Riemannian anchor, and bounded neural residual correction for cross-subject emotion recognition in hearing-impaired individuals. Under strict LOSO30 evaluation, it achieved 58.563% accuracy with 7,227 trainable neural parameters and substantially outperformed same-channel classical and compact neural baselines. The 23-channel configuration retained 98.94% of the accuracy of a parameter-matched 62-channel model while materially reducing acquisition and inference costs. The ablation and error analyses support the complementarity of the spectral, regional, geometric, and residual components, while identifying neutral-emotion recognition and external generalization as priorities for future work.

Acknowledgements

[To be completed by the authors, including funding information if applicable.]

Data Availability

The preprocessed HIED EEG data are available subject to the dataset provider's

license and application procedure. The authors do not redistribute the raw or preprocessed signals. Code and anonymous review materials: [repository link to be added].

Conflict of Interest

The authors declare no conflict of interest.

References

- [1] PICARD R W. *Affective Computing*. Cambridge: MIT Press, 1997.
- [2] KOELSTRA S, MÜHL C, SOLEYMANI M, et al. DEAP: A database for emotion analysis using physiological signals. *IEEE Transactions on Affective Computing*, 2012, 3(1): 18-31. DOI: 10.1109/T-AFFC.2011.15.
- [3] ZHENG W L, LU B L. Investigating critical frequency bands and channels for EEG-based emotion recognition with deep neural networks. *IEEE Transactions on Autonomous Mental Development*, 2015, 7(3): 162-175. DOI: 10.1109/TAMD.2015.2431497.
- [4] BAI Z L, LIU J J, HOU F Z, et al. Emotion recognition with residual network driven by spatial-frequency characteristics of EEG recorded from hearing-impaired adults in response to video clips. *Computers in Biology and Medicine*, 2023, 152: 106344. DOI: 10.1016/j.combiomed.2022.106344.
- [5] BAI Z L, HOU F Z, SUN K X, et al. SECT: A method of shifted EEG channel Transformer for emotion recognition. *IEEE Journal of Biomedical and Health Informatics*, 2023, 27(10): 4758-4767. DOI: 10.1109/JBHI.2023.3301993.
- [6] HE H, WU D R. Transfer learning for brain-computer interfaces: A Euclidean space data alignment approach. *IEEE Transactions on Biomedical Engineering*, 2020, 67(2): 399-410. DOI: 10.1109/TBME.2019.2913914.
- [7] DUAN R N, ZHU J Y, LU B L. Differential entropy feature for EEG-based emotion classification. In: 2013 6th International IEEE/EMBS Conference on Neural Engineering. San Diego: IEEE, 2013: 81-84. DOI: 10.1109/NER.2013.6695876.
- [8] LI X, SONG D, ZHANG P, et al. Exploring EEG features in cross-subject emotion recognition. *Frontiers in Neuroscience*, 2018, 12: 162. DOI: 10.3389/fnins.2018.00162.
- [9] LAWHERN V J, SOLON A J, WAYTOWICH N R, et al. EEGNet: A compact convolutional neural network for EEG-based brain-computer interfaces. *Journal of Neural Engineering*, 2018, 15(5): 056013. DOI: 10.1088/1741-2552/aace8c.
- [10] DING Y, ROBINSON N, ZHANG S, et al. TSception: Capturing temporal dynamics and spatial asymmetry from EEG for emotion recognition. *IEEE Transactions on Affective Computing*, 2023, 14(3): 2238-2250. DOI: 10.1109/TAFFC.2022.3169001.
- [11] BARACHANT A, BONNET S, CONGEDO M, et al. Multiclass brain-computer interface classification by Riemannian geometry. *IEEE Transactions on Biomedical Engineering*, 2012, 59(4): 920-928. DOI: 10.1109/TBME.2011.2172210.
- [12] ZANINI P, CONGEDO M, JUTTEN C, et al. Transfer learning: A Riemannian geometry framework with applications to brain-computer interfaces. *IEEE Transactions on Biomedical Engineering*, 2018, 65(5): 1107-1116. DOI: 10.1109/TBME.2017.2742541.
- [13] F. Yger, M. Berar and F. Lotte, "Riemannian Approaches in Brain-Computer Interfaces: A Review," in *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 25, no. 10, pp. 1753-1762, Oct. 2017, doi: 10.1109/TNSRE.2016.2627016.
- [14] HUANG Z, VAN GOOL L. A Riemannian network for SPD matrix learning. In: *Proceedings of the AAAI Conference on Artificial Intelligence*. San Francisco: AAAI Press, 2017, 31(1): 2036-2042.
- [15] LEDOIT O, WOLF M. A well-conditioned estimator for large-dimensional covariance matrices. *Journal of Multivariate Analysis*, 2004, 88(2): 365-411. DOI:

10.1016/S0047-259X(03)00096-4.

- [16] WELCH P D. The use of fast Fourier transform for the estimation of power spectra. *IEEE Transactions on Audio and Electroacoustics*, 1967, 15(2): 70-73. DOI: 10.1109/TAU.1967.1161901.
- [17] PEDREGOSA F, VAROQUAUX G, GRAMFORT A, et al. Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research*, 2011, 12: 2825-2830.
- [18] LOSHCHILOV I, HUTTER F. Decoupled weight decay regularization. In: *International Conference on Learning Representations*. New Orleans: ICLR, 2019.
- [19] HENDRYCKS D, GIMPEL K. Gaussian error linear units (GELUs). (2016-06-27)[2026-08-07]. <https://arxiv.org/abs/1606.08415>.
- [20] HOLM S. A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics*, 1979, 6(2): 65-70.
- [21] COHEN J. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 1960, 20(1): 37-46. DOI: 10.1177/001316446002000104.
- [22] ZHU M, BAI Z L, WU Q Z, et al. CFBC: A network for EEG emotion recognition by selecting the information of crucial frequency bands. *IEEE Sensors Journal*, 2024, 24: 30451-30461. DOI: 10.1109/JSEN.2024.3440340.
- [23] CAI M P, CHEN J X, HUA C C, et al. EEG emotion recognition using EEG-SWTNS neural network through EEG spectral image. *Information Sciences*, 2024, 680: 121198. DOI: 10.1016/j.ins.2024.121198.
- [24] YU P, HE X P, LI H Y, et al. FMLAN: A novel framework for cross-subject and cross-session EEG emotion recognition. *Biomedical Signal Processing and Control*, 2025, 100: 106912. DOI: 10.1016/j.bspc.2024.106912.
- [25] HU M, XU D, HE K, et al. Cross-subject emotion recognition with contrastive learning based on EEG signal correlations. *Biomedical Signal Processing and Control*, 2025, 104: 107511. DOI: 10.1016/j.bspc.2025.107511.
- [26] SHEN X, LIU X, HU X, et al. Contrastive learning of subject-invariant EEG representations for cross-subject emotion recognition. *IEEE Transactions on Affective Computing*, 2023, 14(3): 2496-2511.
- [27] ZHU M, JIN H N, BAI Z L, et al. Image-evoked emotion recognition for hearing-impaired subjects with EEG signals. *Sensors*, 2023, 23(12): 5461. DOI: 10.3390/s23125461.
- [28] WANG J C, SONG Y, MAO Z M, et al. EEG-based emotion identification using 1-D deep residual shrinkage network with microstate features. *IEEE Sensors Journal*, 2023, 23: 5165-5174. DOI: 10.1109/JSEN.2023.3239507.
- [29] GKINTONI E, AROUTZIDIS A, ANTONOPOULOU H, et al. From neural networks to emotional networks: A systematic review of EEG-based emotion recognition in cognitive neuroscience and real-world applications. *Brain Sciences*, 2025, 15(3): 220. DOI: 10.3390/brainsci15030220.