GEOMETRY-ALIGNED EEG SCALPOGRAMS: GRAPH HARMONICS ENABLE ACCURATE, CALIBRATED DEMENTIA SUBTYPING

Siddhant Ujjain[*]*, Ahmad Siraj Hashmi*[*]*, Pooja Singh, Ekta Srivastava,*
Sandeep Kumar, Tapan Kumar Gandhi

Department of Electrical Engineering, Indian Institute of Technology Delhi, India

ABSTRACT

Dementia subtyping from routine scalp EEG could enable low-cost decision support, but deep models often ignore 10–20 electrode geometry and can overfit small clinical cohorts. We introduce *graph-harmonic scalpograms*, a geometry-aligned representation that projects 19-channel EEG onto low-frequency scalp graph harmonics and converts harmonic activity into compact time-frequency images. Using OpenNeuro ds004504 ($N = 88$, eyes-closed; AD/FTD/CN), recordings are segmented into 8 s windows (75% overlap), projected onto the first $K = 12$ non-DC graph modes, transformed over 1–45 Hz using $F = 60$ log-spaced bins, and encoded as Band-RGB scalpograms ($\delta{+}\theta$, α, $\beta{+}\gamma$) under strict subject-wise 80/10/10 splits to prevent leakage. Models are trained with class-weighted cross-entropy, geometry-preserving spectro-temporal augmentations, and post-hoc temperature scaling to improve probability calibration; evaluation is performed at the subject level by averaging window posteriors with 95% BCa bootstrap confidence intervals. A lightweight CNN trained from scratch achieves the best subject-level performance (balanced accuracy 0.83 ± 0.07, macro-F1 0.847 ± 0.06, AUROC 0.95 ± 0.05), outperforming transfer baselines such as ResNet-34 (0.65 ± 0.05 / 0.667 ± 0.04 / 0.90 ± 0.03). Overall, aligning the model inductive bias with 10–20 sensor geometry supports accurate, well-calibrated subject-level dementia subtyping using lightweight CNNs that are promising for small clinical EEG cohorts.

1. INTRODUCTION

Dementia subtyping from routine scalp EEG could provide a low-cost, noninvasive adjunct to neuropsychological assessment and imaging, especially in settings where access to advanced biomarkers is limited [1]. Eyes-closed resting EEG is widely available and quick to acquire, yet its use in clinical decision support remains constrained by small datasets, heterogeneous preprocessing, and representations that either discard spatial structure or require large models that are poorly suited to small clinical cohorts. Classical EEG analyses for Alzheimer's disease (AD) and frontotemporal dementia (FTD) emphasize bandpower changes (alpha slowing; increased delta–theta) [2], band ratios, and regional asymmetries; other lines examine spectral aperiodic/periodic separation [3], functional connectivity (e.g., wPLI/imaginary coherence) [4], [5], and microstates. Deep learning work commonly projects channel-wise signals to spectrograms and applies CNNs or transformers; these approaches can be effective but often (i) ignore the 10–20 sensor geometry [6], (ii) rely on large backbones that overfit small clinical cohorts, and (iii) risk information leakage when splitting by windows instead of subjects [7]. Graph-signal processing has been explored [8] to encode spatial priors over electrode layouts, yet most methods either require heavy graph construction (subject-specific connectivity) or do not couple the graph basis with a compact time–frequency (TF) representation suitable for small-N training.

Problem Statement. We study routine 19-channel, eyes-closed EEG from OpenNeuro ds004504 [9] (88 subjects; AD/FTD/CN). Our goal is a geometry-aware representation that (i) preserves smooth scalp structure, (ii) produces small image-like inputs for lightweight models, and (iii) supports strict subject-wise evaluation with calibrated outputs.

We project signals onto the lowest non-DC *scalp graph harmonics*, i.e., eigenvectors [10] of a kNN graph with heat-kernel weights built from 10–20 electrode coordinates, which explicitly encodes spatial smoothness. The time-frequency energy of harmonic coefficients yields a compact $K \times F$ map. We summarize it as *Band-RGB scalpograms* $(\delta + \theta, \alpha, \beta + \gamma)$, then standardize and resize the resulting geometry-aligned image for common backbones. Training uses subject-wise splits (80/10/10), class-weighted cross-entropy, and safe augmentations (cutout; frequency masking along the F axis). We apply temperature scaling for calibration [11]. Evaluation is at the *subject* level by averaging window probabilities and reporting bootstrap confidence intervals [12].

Contributions. (i) An interpretable *graph-harmonic scalpogram* representation that enables small-capacity models. (ii) A calibrated subject-level evaluation protocol on ds004504 designed to avoid leakage. (iii) An ablation identifying practical defaults (K=12, combinatorial Laplacian, STFT) that balance spatial smoothness and discriminative detail.

[*]These authors contributed equally to this work

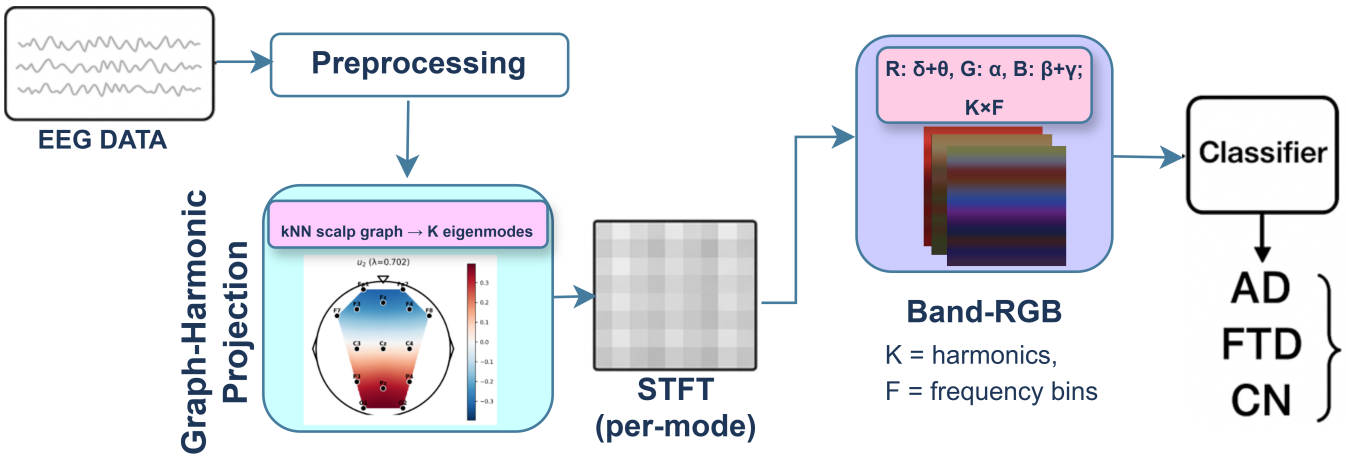

Fig. 1: **Geometry-aware EEG classification pipeline.** We preprocess 19-channel EEG, project signals onto the lowest non-DC scalp graph harmonics from a kNN 10–20 electrode graph, and compute per-mode STFT energy to obtain a compact $K \times F$ scalpogram. We then form a three-channel Band-RGB encoding (R: $\delta + \theta$, G: α, B: $\beta + \gamma$) and feed the resulting image to a lightweight classifier for dementia subtyping.

2. RELATED WORK

EEG Markers for Dementia & Dementia Subtyping Resting-state EEG has long been explored as a non-invasive and low-cost biomarker for neurodegenerative disorders. A large body of work reports characteristic spectral slowing in Alzheimer's disease (AD), typically reflected by reduced posterior α activity and increased δ–θ power, together with reduced dynamical complexity and altered coupling patterns. Comprehensive reviews have summarized these findings and the associated quantitative EEG (qEEG) pipelines based on bandpower, complexity, and synchronization/connectivity features [13], [14], [15]. More recently, expert recommendations have emphasized the importance of standardized resting-state EEG measures and evaluation practices for translation to clinical trials and decision support [16].

Beyond AD versus cognitively normal (CN) classification, distinguishing AD from frontotemporal dementia (FTD) remains clinically important and methodologically challenging. Prior work suggests that differences in topography and spectral power ratios may support AD–FTD separation, although reported performance depends on cohort characteristics and protocol design [17], [18].

Deep Learning on EEG Time-Frequency & Topographic Representations. Deep learning approaches for EEG commonly start from time–frequency representations (e.g., STFT or wavelet spectrograms) and apply convolutional networks that can be trained end-to-end. Compact architectures such as EEGNet are widely used as strong baselines in small-data EEG settings, while deeper ConvNet variants can learn more expressive representations when sufficient data are available [19], [20]. A key limitation of naïve spectrogram pipelines is that channels are often treated as independent "rows," which can underuse known electrode geometry. To better exploit spatial structure, topology-aware image-like encodings and recurrent-convolutional schemes have been proposed to learn from spectrograms while preserving (approximately) scalp neighborhood relations [21]. In dementia-focused applications, CNN-based frameworks operating on spectrogram-derived inputs remain common, including recent studies combining spectrogram representations with neural or hybrid learning components [22], [23], [24].

Graph-Based EEG Learning & Spectral Graph Methods. Graph signal processing and graph neural networks (GNNs) provide a principled way to model EEG channels on irregular domains, where nodes represent electrodes and edges encode spatial proximity or functional relations. Localized spectral graph filtering and graph convolutions have become standard tools for learning on graphs [25], [26], motivating EEG-specific GNN pipelines that explicitly leverage electrode graphs for classification [27]. Recent reviews further highlight the growing interest in GNNs for EEG analysis and discuss design choices such as graph construction, pooling, and dynamic connectivity estimation [28]. In parallel, spectral graph theory offers interpretable multiscale representations through graph wavelets and Laplacian eigenmodes, enabling geometry-aligned analysis of spatially smooth EEG patterns [29]. These perspectives motivate approaches that combine compact time–frequency features with geometry-aware spatial bases to improve robustness on small clinical datasets.

3. METHODS

Dataset & Preprocessing. We use the preprocessed derivatives of OpenNeuro ds004504, consisting of 19-channel eyes-closed resting-state EEG from three cohorts (AD/FTD/CN). Recordings are loaded with MNE. To ensure consistent sensor localization, legacy channel names (T3/T4/T5/T6) are renamed to T7/T8/P7/P8 before attaching the standard 10–20 montage, which provides 3-D electrode coordinates for all channels. Signals may be resampled to 250 Hz to reduce computation. To prevent information leakage due to cor-

related windows from the same subject, we perform strict *subject-wise* partitioning into train/validation/test splits using an 80/10/10 ratio, stratified by diagnosis; all windows and derived representations inherit the subject assignment.

Windowing & Frequency Discretization. Each continuous recording is segmented into 8-s windows with 75% overlap (hop = 2 s). Windows with more than 20% missing samples are discarded. We restrict analysis to 1–45 Hz and discretize the spectrum into $F = 60$ logarithmically spaced frequency bins $\{f_m\}_{m=1}^{F}$.

Scalp Sensor Graph & Graph-Harmonic Basis. Let $\{r_i\}_{i=1}^{19}$ denote the 3-D electrode coordinates. We construct a weighted undirected sensor graph using a k-nearest-neighbor rule (default $k = 3$) based on Euclidean distances $d_{ij} = \|r_i - r_j\|_2$. Edge weights follow a heat kernel:

$$w_{ij} = \begin{cases} \exp\left(-d_{ij}^2/\sigma^2\right), & j \in k\mathrm{NN}(i), \\ 0, & \text{otherwise}, \end{cases} \qquad (1)$$
$$\sigma = \mathrm{median}\{d_{ij} \,:\, j \in k\mathrm{NN}(i)\}.$$

Let $W = [w_{ij}]$ and D be the degree matrix. We use the (combinatorial) graph Laplacian $L = D - W$ with eigendecomposition $L = U\Lambda U^\top$. The eigenvectors $\{u_k\}$ form a graph Fourier (harmonic) basis. We retain the lowest non-DC modes

$$U_K = [u_1, \ldots, u_K] \in \mathbb{R}^{19\times K}, \qquad (2)$$

with default $K = 12$.

Graph-Spectral Projection of EEG Windows. For each time sample t within a window, let $x(t) \in \mathbb{R}^{19}$ be the vector of simultaneous sensor values. Because we exclude the DC eigenvector u_0, we remove the instantaneous sensor mean to suppress the DC component:

$$x'(t) = x(t) - \frac{1}{19}\mathbf{1}\mathbf{1}^\top x(t). \qquad (3)$$

This operation is equivalent to applying a per-time-sample common-average reference (CAR) on the 19 channels; we use it solely to suppress the DC graph mode and not as a separate clinical rereferencing step. We then project the demeaned signal onto the retained graph harmonics:

$$c(t) = U_K^\top x'(t) \in \mathbb{R}^K, \qquad (4)$$

yielding K harmonic coefficient time series $\{c_k(t)\}_{k=1}^{K}$ per window.

Time-Frequency Power On Harmonic Coefficients. For each harmonic coefficient $c_k(t)$, we compute a short-time Fourier transform (STFT) using a Hann window of length 2 s and hop 2 s:

$$S_k(f,\tau) = \sum_t c_k(t)\, h(t-\tau)\, e^{-j2\pi f t}. \qquad (5)$$

We summarize each window by the time-averaged power spectrum

$$P_k(f) = \mathbb{E}_\tau \left|S_k(f,\tau)\right|^2, \qquad (6)$$

sampled on the common frequency grid $\{f_m\}_{m=1}^{F}$. Stacking over k yields a fixed-size harmonic power map

$$P \in \mathbb{R}^{K\times F}, \qquad (7)$$

for every EEG window.

Scalpogram Encoding, Resizing, & Normalization. We use either (i) the full $P \in \mathbb{R}^{K\times F}$ as a single-channel "scalpogram", or (ii) a three-channel *Band-RGB* encoding derived from canonical EEG bands: δ(1–4), θ(4–8), α(8–13), β(13–30), $\gamma(30-45)$ Hz. To keep the representation image-like ($K \times F$) while injecting band semantics, we form three channels by masking the frequency axis into band groups (Red: $\delta \cup \theta$, Green: α, Blue: $\beta \cup \gamma$). Concretely, each RGB channel is formed by averaging $P_k(f)$ within a predefined band group (red: δ+θ, green: α, blue: β+γ) and replicating that summary along the frequency axis for backbone compatibility; this yields an RGB tensor $I \in \mathbb{R}^{K\times F\times 3}$ that preserves harmonic ordering while injecting band semantics. Optionally, for interpretation, we also compute band-averaged power per harmonic, $\bar{P}_k^{(b)} = \frac{1}{|B_b|}\sum_{f\in B_b} P_k(f)$, and summarize bands as $\bar{R} = \bar{P}^{(\delta)} + \bar{P}^{(\theta)}$, $\bar{G} = \bar{P}^{(\alpha)}$, $\bar{B} = \bar{P}^{(\beta)} + \bar{P}^{(\gamma)}$. We resize the resulting map with bicubic interpolation to 224×224 for ResNet/VGG/ViT (or 299×299 for InceptionV3). Each image is standardized independently using per-image z-scoring and clipping:

$$\hat{I} = \mathrm{clip}\left(\frac{I - \mu(I)}{\sigma(I) + \varepsilon}, -4, 4\right), \qquad \varepsilon = 10^{-6}. \qquad (8)$$

Data Augmentation. To preserve the semantics of the $K \times F$ representation, we apply augmentations limited to: (i) random erasing (small rectangles), (ii) frequency masking along the F axis (SpecAugment-style), and (iii) low-variance Gaussian noise in log-power. We do not use flips, which would alter harmonic order or invert frequency structure.

Models, Training, & Calibration. We evaluate standard image backbones (ResNet-18/34, VGG-16(BN), InceptionV3, ViT-S/16) and a lightweight CNN trained from scratch (four convolutional blocks + global average pooling + 3-way linear head). For single-channel inputs, we either replicate channels or use a learned 1×1 convolution to adapt to RGB backbones.

Class imbalance is handled with weighted cross-entropy using weights inversely proportional to class frequencies; we optionally apply label smoothing ($\varepsilon = 0.1$). Models are optimized with AdamW, cosine learning-rate decay with warmup, mixed precision, and early stopping based on validation macro-F1.

After training, we calibrate probabilities via temperature scaling. Given logits z, calibrated probabilities are

$$p_c = \mathrm{softmax}\left(\frac{z_c}{T}\right), \qquad T > 0, \qquad (9)$$

where T is fitted on the validation set by minimizing negative log-likelihood.

Table 1: Subject-level performance comparison with classical ML baselines and the proposed lightweight CNN (mean ± 95% CI).

Model	Input features / representation	Bal. Acc.	Macro-F1	AUROC
Logistic Regression	PCA on flattened Band-RGB scalpograms	0.36 ± 0.05	0.42 ± 0.04	0.66 ± 0.04
LDA	Region-averaged band powers and band-power ratios	0.43 ± 0.06	0.47 ± 0.05	0.78 ± 0.05
kNN ($k = 3$)	PCA on flattened Band-RGB scalpograms	0.43 ± 0.05	0.42 ± 0.05	0.75 ± 0.05
SVM (RBF)	Flattened Band-RGB scalpograms	0.57 ± 0.06	0.54 ± 0.05	0.86 ± 0.04
Decision Tree	Flattened Band-RGB scalpograms	0.478 ± 0.05	0.50 ± 0.05	0.83 ± 0.04
XGBoost	Flattened Band-RGB scalpograms	0.54 ± 0.05	0.55 ± 0.05	0.67 ± 0.03
Lightweight CNN (ours)	Band-RGB scalpograms	0.83 ± 0.07	0.847 ± 0.06	0.95 ± 0.05

Subject-Level Aggregation & Evaluation. Because clinical decisions are made per subject, we aggregate window-level predictions into a subject prediction. For subject s with N_s windows, we average the calibrated window posteriors (optionally using a 10% trimmed mean):

$$\bar{p}^{(s)} = \frac{1}{N_s} \sum_{i=1}^{N_s} p_i^{(s)}, \qquad \hat{y}^{(s)} = \arg\max_c \bar{p}_c^{(s)}. \tag{10}$$

We report subject-level balanced accuracy and macro-F1, as well as one-vs-rest AUROC computed from subject-level probabilities. We apply temperature scaling to improve probability calibration and compute calibration diagnostics (e.g., ECE) on the validation split; uncertainty is quantified using 95% BCa bootstrap confidence intervals over subjects.

4. RESULTS

4.1. Main Comparison.

Table 2 summarizes subject-level performance averaged over 1,000 BCa bootstrap samples (95% CIs shown). The *lightweight CNN* operating on *Band-RGB scalpograms* is the top performer, reaching **Bal. Acc. 0.83 ± 0.07**, **Macro-F1 0.847 ± 0.06**, and **AUROC 0.95 ± 0.05**. Among transfer-learned baselines, *ResNet-34* is the strongest (0.65 ± 0.05 bal-acc; 0.667 ± 0.04 macro-F1; 0.90 ± 0.03 AUROC), while *ResNet-18* and *ViT-S/16* form a middle tier. *InceptionV3* and *VGG-16(BN)* underperform on accuracy/F1 despite moderate AUROC values.

Two observations stand out. **(i) Capacity vs. inductive bias:** deeper or more complex backbones did not yield higher subject-level accuracy; the best model is the smallest one, suggesting that our $K \times F$ representation with Band-RGB already encodes a strong spatial–spectral prior that benefits from local convolutional biases rather than high-capacity feature mixers. **(ii) Ranking vs. decision quality:** *VGG-16(BN)* attains relatively high AUROC (0.91 ± 0.04) yet low Macro-F1 (0.32 ± 0.05), indicating good score ranking but poor operating-point behavior under class imbalance. In contrast, the light CNN pairs high AUROC with high Macro-F1, reflecting both separability and usable calibration for 3-way decisions. Decisions are made *per subject* by averaging window-level *probabilities* (not logits). Using a 10% trimmed mean changed results by < 0.5 Macro-F1 points on validation and remained within the reported CIs, so we fixed plain averaging for test reporting.

Table 2: Subject-level performance (mean ± 95% CI over bootstrap).

Model	Input	Bal. Acc.	Macro-F1	AUROC
ResNet-18	Band-RGB	0.54±0.04	0.50±0.03	0.78±0.02
ResNet-34	Band-RGB	0.65±0.05	0.667±0.04	0.90±0.03
ViT-S/16	Band-RGB	0.60±0.05	0.52±0.04	0.75±0.03
InceptionV3	Band-RGB	0.34±0.05	0.20±0.04	0.82±0.03
VGG-16(BN)	Band-RGB	0.40±0.06	0.32±0.05	0.91±0.04
CNN (light)	Band-RGB	**0.83**±0.07	**0.847**±0.06	**0.95**±0.05

Comparison with Classical Machine-Learning Baselines. Table 1 contrasts standard classical machine-learning (ML) baselines against our lightweight CNN when trained/evaluated using the proposed Band-RGB scalpogram representation. For the classical ML models, each scalpogram is converted into a fixed-length feature vector (flattened), and PCA is applied where indicated to reduce dimensionality; in addition, we include a handcrafted-feature baseline using region-averaged band powers and band-power ratios. In contrast, the lightweight CNN operates directly on the Band-RGB scalpogram images, preserving spatial–spectral structure and enabling end-to-end feature learning. All results are reported at the *subject level* as mean ± 95% confidence intervals (CI).

Ablation Studies. Table 3 explores three design axes. **(i) Graph harmonics K.** Peak at $\mathbf{K = 12}$ (Macro-F1 **0.84**); $K = 6-9$ underfit spatial structure, $K = 15$ adds high-frequency noise/overfit. **(ii) Laplacian.** *Combinatorial* > symmetric-normalized (**0.84** vs. 0.77 Macro-F1), likely due to better preservation of informative degree patterns on 19 nodes. **(iii) Time–frequency. STFT** > CWT > SST (Macro-F1: **0.8475**/0.6564/0.4722); fixed tiling aligns with resizing and per-image z-scoring.

Table 3: Ablation study (Macro-F1).

Setting	Variant	Macro-F1
# Harmonics K	6 / 9 / **12** / 15	0.76 / 0.774 / **0.84** / 0.699
Laplacian	**Comb.** / Norm.	**0.84** / 0.77
T–F Engine	**STFT** / CWT / SST	**0.8475** / 0.6564 / 0.4722

Interpretability, Calibration, and Robustness. Low-order harmonics ($u_1 - u_6$) form smooth global waves; beyond u_{12} they become localized/inconsistent, matching the K ablation (Fig. 2). Class scalpograms align with neurophysiology—CN shows posterior α (G), while AD/FTD exhibit elevated $\delta+\theta$ (R), broadened α, and weaker $\beta+\gamma$ (B) (Fig. 3). The pipeline is compact: derivatives → graph projection → STFT → Band-RGB → classifier (group-aware batching, temperature scaling) (Fig. 1). Uncertainty is reported as **95%**

BCa bootstrap CIs over subjects (1,000 resamples) in Tables 2–3. Post-hoc temperature scaling is used as a standard post-hoc step to improve the reliability of predicted probabilities; we fit it on the validation split and report metrics on held-out test subjects. Sanity checks confirmed strict subject-wise splits (no leakage), approximate energy conservation (sensor vs. graph; excluding DC), and well-conditioned low-frequency spatial modes.

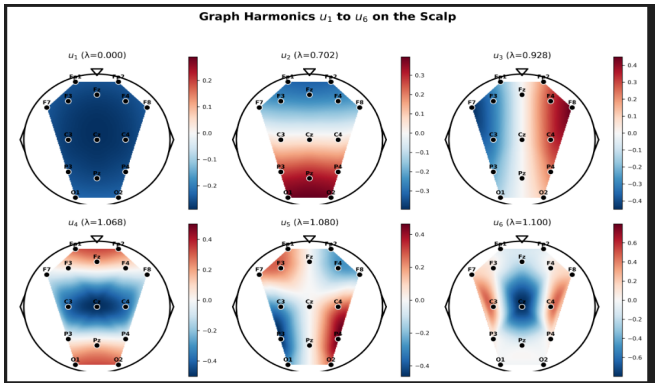

Fig. 2: Graph harmonics $u_1 \dots u_6$ on the scalp: smooth spatial waves validate the geometric prior.

(a) Healthy(CN) (b) AD (c) FTD

Fig. 3: Class-conditioned scalpograms (CN, AD, FTD). Band-RGB mapping $(\delta{+}\theta, \alpha, \beta{+}\gamma)$ emphasizes cross-band contrasts.

5. DISCUSSION

Our results indicate that *geometry-aware scalpograms* derived from graph-harmonic projections provide an effective and compact representation for subject-level dementia classification from routine EEG. On OpenNeuro ds004504, a lightweight CNN operating on Band-RGB images achieved the strongest performance (Bal. Acc. **0.83±0.07**, Macro-F1 **0.847±0.06**, AUROC **0.95±0.05**), outperforming deeper ImageNet-initialized backbones.

Inductive Bias vs. Capacity. With limited clinical data, locality and translation-equivariant inductive bias appeared more beneficial than additional depth or long-range token mixing: the light CNN exceeded ResNet-34, ViT-S/16, VGG-16, and InceptionV3, indicating that $K \times F$ energy maps benefit from convolutional biases more than heavy architectures.

Scalp Graph as Scaffold. Ablations peak at $K{=}12$; smaller K underfit spatial structure, larger K add high-frequency noise. The *combinatorial* Laplacian outperformed the symmetric-normalized variant on the 19-node montage. Under our standardization and Band-RGB mapping, *STFT* > CWT > SST.

Interpretability, Calibration, & Robustness. Low-order harmonics are smooth global waves (Fig. 2); class scalpograms align with expected posterior α in CN and broader low-frequency power in dementia (Fig. 3). Probabilities are well calibrated via temperature scaling. We use subject-wise splits to avoid leakage, average window posteriors for subject decisions, and report BCa bootstrap CIs over subjects ($N{=}88$). The pipeline requires only the standard 10–20 montage and can be applied under different referencing schemes when used consistently; we recommend reporting the reference choice and preprocessing settings for reproducibility.

Clinical and Scientific Impact. Graph-harmonic scalpograms couple a *biophysically plausible spatial prior* with a *computationally light* pipeline, yielding calibrated, subject-level decisions from routine EEG. Because they rely on standard clinical montages and produce interpretable intermediate maps, they are amenable to retrospective screening and workflow integration (e.g., flagging studies with dementia-like spectral profiles). Scientifically, the observed peak at $K{=}12$ supports the view that dementia effects manifest as *low-to-moderate spatial frequency* patterns rather than highly localized features.

Limitations. This single-site, eyes-closed dataset has a modest sample size ($N{=}88$), limiting external validity. We did not explicitly separate aperiodic/oscillatory components or model connectivity/microstate dynamics, and potential confounds (e.g., age, disease duration) were not adjusted in the classifier. Results may depend on the A1–A2 reference, TF hyperparameters, and Laplacian construction.

6. CONCLUSION

We propose a geometry-aware EEG pipeline that projects 10–20 sensor signals onto scalp graph harmonics and converts harmonic coefficients into compact time–frequency scalpograms for dementia subtyping. On OpenNeuro ds004504 (88 subjects; 19-channel, eyes-closed EEG), a lightweight CNN trained on Band-RGB scalpograms achieves the best subject-level results (Bal. Acc. 0.83±0.07, Macro-F1 0.847±0.06, AUROC 0.95±0.05), surpassing deeper ImageNet-pretrained backbones. Ablations support practical defaults ($K{=}12$ non-DC harmonics, combinatorial Laplacian, STFT), showing that an explicit geometry-aligned inductive bias can outperform higher-capacity models in the small-N clinical regime. The evaluation protocol is designed to resist leakage through strict subject-wise splitting and yields calibrated probabilities that are more clinically interpretable. Future work will validate robustness across sites and hardware, incorporate covariate-aware and causal analyses, model periodic/aperiodic structure, integrate connectivity and microstates, and scale via self-supervised pretraining and domain adaptation.

References

[1] P. M. Rossini, R. Di Iorio, F. Vecchio, et al., "Early diagnosis of alzheimer's disease: The role of biomarkers including advanced eeg signal analysis. report from the ifcn-sponsored panel of experts," *Clinical Neurophysiology*, vol. 131, no. 6, pp. 1287–1310, 2020.

[2] J. Jeong, "Eeg dynamics in patients with alzheimer's disease," *Clinical neurophysiology*, vol. 115, no. 7, pp. 1490–1505, 2004.

[3] T. Donoghue, M. Haller, E. J. Peterson, et al., "Parameterizing neural power spectra into periodic and aperiodic components," *Nature neuroscience*, vol. 23, no. 12, pp. 1655–1665, 2020.

[4] M. Vinck, R. Oostenveld, M. Van Wingerden, F. Battaglia, and C. M. Pennartz, "An improved index of phase-synchronization for electrophysiological data in the presence of volume-conduction, noise and sample-size bias," *Neuroimage*, vol. 55, no. 4, pp. 1548–1565, 2011.

[5] G. Nolte, O. Bai, L. Wheaton, Z. Mari, S. Vorbach, and M. Hallett, "Identifying true brain interaction from eeg data using the imaginary part of coherency," *Clinical neurophysiology*, vol. 115, no. 10, pp. 2292–2307, 2004.

[6] G. H. Klem, "The ten-twenty electrode system of the international federation. the international federation of clinical neurophysiology," *Electroencephalogr. Clin. Neurophysiol. Suppl.*, vol. 52, pp. 3–6, 1999.

[7] G. Varoquaux, "Cross-validation failure: Small sample sizes lead to large error bars," *Neuroimage*, vol. 180, pp. 68–77, 2018.

[8] D. I. Shuman, S. K. Narang, P. Frossard, A. Ortega, and P. Vandergheynst, "The emerging field of signal processing on graphs: Extending high-dimensional data analysis to networks and other irregular domains," *IEEE signal processing magazine*, vol. 30, no. 3, pp. 83–98, 2013.

[9] A. Miltiadous, K. D. Tzimourta, T. Afrantou, et al., "A dataset of eeg recordings from: Alzheimer's disease, frontotemporal dementia and healthy subjects," *OpenNeuro*, vol. 1, p. 88, 2023.

[10] M. Belkin and P. Niyogi, "Laplacian eigenmaps for dimensionality reduction and data representation," *Neural computation*, vol. 15, no. 6, pp. 1373–1396, 2003.

[11] C. Guo, G. Pleiss, Y. Sun, and K. Q. Weinberger, "On calibration of modern neural networks," in *International conference on machine learning*, PMLR, 2017, pp. 1321–1330.

[12] B. Efron, "Better bootstrap confidence intervals," *Journal of the American statistical Association*, vol. 82, no. 397, pp. 171–185, 1987.

[13] J. Dauwels, F. Vialatte, and A. Cichocki, "Diagnosis of alzheimer's disease from eeg signals: Where are we standing?" *Current Alzheimer Research*, vol. 7, no. 6, pp. 487–505, 2010.

[14] J. Dauwels, K. Srinivasan, M. Ramasubba Reddy, et al., "Slowing and loss of complexity in alzheimer s eeg: Two sides of the same coin?" *International journal of Alzheimer's disease*, vol. 2011, no. 1, p. 539 621, 2011.

[15] F. Vecchio, C. Babiloni, R. Lizio, et al., "Resting state cortical eeg rhythms in alzheimer's disease: Toward eeg markers for clinical applications: A review," *Supplements to Clinical neurophysiology*, vol. 62, pp. 223–236, 2013.

[16] C. Babiloni, X. Arakaki, H. Azami, et al., "Measures of resting state eeg rhythms for clinical trials in alzheimer's disease: Recommendations of an expert panel," *Alzheimer's & Dementia*, vol. 17, no. 9, pp. 1528–1553, 2021.

[17] R. Nardone, L. Sebastianelli, V. Versace, et al., "Usefulness of eeg techniques in distinguishing frontotemporal dementia from alzheimer's disease and other dementias," *Disease markers*, vol. 2018, no. 1, p. 6 581 490, 2018.

[18] J. Chang and C. Chang, "Quantitative electroencephalography markers for an accurate diagnosis of frontotemporal dementia: A spectral power ratio approach," *Medicina*, vol. 59, no. 12, p. 2155, 2023.

[19] 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," *Journal of neural engineering*, vol. 15, no. 5, p. 056 013, 2018.

[20] R. T. Schirrmeister, J. T. Springenberg, L. D. J. Fiederer, et al., "Deep learning with convolutional neural networks for eeg decoding and visualization," *Human brain mapping*, vol. 38, no. 11, pp. 5391–5420, 2017.

[21] P. Bashivan, I. Rish, M. Yeasin, and N. Codella, "Learning representations from eeg with deep recurrent-convolutional neural networks," *arXiv preprint arXiv:1511.06448*, 2015.

[22] K. Stefanou, K. D. Tzimourta, C. Bellos, et al., "A novel cnn-based framework for alzheimer's disease detection using eeg spectrogram representations," *Journal of Personalized Medicine*, vol. 15, no. 1, p. 27, 2025.

[23] F. B. Arikan, D. Cetintas, A. Aksoy, and M. Yildirim, "A deep learning approach to alzheimer's diagnosis using eeg data: Dual-attention and optuna-optimized svm," *Biomedicines*, vol. 13, no. 8, p. 2017, 2025.

[24] S. Ujjain, V. N. Soren, S. Kumar, and T. K. Gandhi, "Evidence-based epileptic seizure detection," in *2025 47th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, IEEE, 2025, pp. 1–6.

[25] M. Defferrard, X. Bresson, and P. Vandergheynst, "Convolutional neural networks on graphs with fast localized spectral filtering," *Advances in neural information processing systems*, vol. 29, 2016.

[26] T. Kipf, "Semi-supervised classification with graph convolutional networks," *arXiv preprint arXiv:1609.02907*, 2016.

[27] A. Demir, T. Koike-Akino, Y. Wang, M. Haruna, and D. Erdogmus, "Eeg-gnn: Graph neural networks for classification of electroencephalogram (eeg) signals," in *2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, IEEE, 2021, pp. 1061–1067.

[28] M. Graña and I. Morais-Quilez, "A review of graph neural networks for electroencephalography data analysis," *Neurocomputing*, vol. 562, p. 126 901, 2023.

[29] D. K. Hammond, P. Vandergheynst, and R. Gribonval, "Wavelets on graphs via spectral graph theory," *Applied and Computational Harmonic Analysis*, vol. 30, no. 2, pp. 129–150, 2011.