Academia
Neuroscience and Brain Research

10.20935/AcadNeurosci8659

Research Article

Propagation-based seizure detection for wearable EEG using planar and spatial EEG modeling

Dhanushka Wijesinghe[1], Ivan T. Lima, Jr.[1,*]

[1] Department of Electrical and Computer Engineering, North Dakota State University, Fargo, ND, USA
* Correspondence: ivan.lima@ndsu.edu

Received: 18 August 2026 | **Accepted:** 16 November 2026 | **Published:** 25 November 2026

Abstract

Introduction: This study proposes a propagation-based seizure detection framework for wearable-oriented electroencephalography (EEG) monitoring and investigates its performance across different electrode configurations and independent EEG datasets.

Materials and methods: Seizure propagation was modeled using two-dimensional (2D) and three-dimensional (3D) representations for planar and spatial electrode configurations, respectively. The framework was evaluated using leave-one-subject-out cross-validation on the Temple University Hospital EEG Seizure Corpus (TUSZ) and further assessed retrospectively on the independent Children's Hospital Boston–Massachusetts Institute of Technology (CHB-MIT) scalp EEG database using a common four-channel bipolar montage with limited operating-threshold calibration. Cross-dataset performance was examined using a reduced feature set and temporal post-processing, with these configurations selected based on the CHB-MIT evaluation results.

Results: The proposed framework outperformed conventional feature-based seizure detection while using fewer electrodes. For spatial electrode configurations, the complete 3D propagation representation significantly improved focal seizure sensitivity compared with the 2D representation under matched-electrode conditions. When adapted to a wearable-oriented four-channel bipolar montage, the framework achieved sensitivities of 92.34%, 83.94%, and 92.01% with area under the receiver operating characteristic curve (AUC) of 0.975, 0.959, and 0.962 for generalized, focal, and combined seizure detection, respectively. Retrospective cross-dataset evaluation demonstrated the potential for transfer of the TUSZ-trained classifier to CHB-MIT with limited threshold calibration, while feature refinement and temporal post-processing improved the observed sensitivity–false-alarm trade-off.

Conclusions: The findings support the discriminative value of geometry-aware propagation representations across planar and spatial electrode configurations and demonstrate promising seizure detection performance using a wearable-oriented four-channel bipolar montage. The retrospective cross-dataset findings further suggest potential transferability, although independent validation of the selected feature and temporal-correction configurations remains necessary. Future prospective evaluation using wearable EEG hardware is needed to establish practical monitoring performance.

Keywords: electroencephalography; machine learning; propagation modeling; seizure detection; seizure propagation; wearable electroencephalography

1. Introduction

Epilepsy is a chronic neurological disorder characterized by recurrent seizures arising from abnormal, excessive, or synchronous neuronal activity in the brain [1]. Seizures vary considerably in their clinical manifestation, duration, spatial origin, and propagation pattern, ranging from brief alterations in awareness to prolonged convulsive episodes. Globally, epilepsy affects more than 50 million people and represents a substantial neurological and public-health burden [2]. Reliable seizure monitoring is therefore important not only for diagnosis but also for assessing seizure frequency, evaluating treatment effectiveness, and reducing risks associated with unrecognized or inadequately controlled seizure events [3].

Electroencephalography (EEG) remains the principal modality for objectively characterizing epileptic brain activity because it directly measures electrical activity generated by neuronal populations [4, 5]. Conventional clinical EEG monitoring typically relies on multichannel scalp recordings obtained using standardized electrode arrangements, providing broad spatial coverage that enables clinicians to identify abnormal waveforms and evaluate their distribution across the scalp. However, prolonged multichannel EEG monitoring is resource-intensive and generally requires specialized equipment, trained personnel, and multiple electrodes attached to the scalp. These requirements restrict patient mobility and make conventional systems inconvenient for continuous monitoring in everyday environments [3, 6]. Consequently, there is increasing interest in portable and wearable EEG systems capable of providing long-term seizure monitoring outside conventional clinical settings.

Automated seizure detection has consequently received substantial attention, with approaches ranging from conventional signal-processing and machine-learning methods to increasingly sophisticated deep-learning architectures [7–11]. Traditional approaches generally extract handcrafted descriptors from EEG signals, including amplitude statistics, spectral power, entropy, wavelet coefficients, and connectivity measures, followed by classification using algorithms such as support vector machines, neural networks, decision trees, or ensemble methods [12–14]. More recently, convolutional neural networks, recurrent architectures, attention mechanisms, and other deep-learning approaches have been employed to automatically learn discriminative representations from raw EEG or time–frequency transformations [15–19]. Recent developments in EEG foundation models, including single-channel approaches such as SingLEM, have further expanded the available methods for learning transferable EEG representations [20]. Although many of these methods report high detection performance under controlled experimental conditions, increasing model complexity does not necessarily translate directly to practical wearable implementation. Large models may require substantial memory, computational resources, and power, and their performance can also depend on extensive training data and long temporal contexts. Although lightweight deep-learning architectures have demonstrated promising seizure detection performance with reduced computational complexity [21, 22], achieving reliable detection under the combined constraints of computational efficiency and limited electrode availability remains an important consideration for wearable applications.

Wearable seizure monitoring introduces an additional and closely related constraint—the number and placement of EEG electrodes. Reducing electrode count can improve patient comfort, simplify sensor placement, reduce hardware complexity, and decrease the computational burden associated with signal acquisition and processing. Accordingly, several studies have investigated minimal-channel and even single-channel EEG approaches to seizure detection [23–26]. Such configurations can effectively detect seizure patterns that produce strong or widespread electrophysiological signatures. However, reducing the number of electrodes inevitably reduces spatial information, which may become particularly important for seizures with localized onset or heterogeneous propagation. Therefore, the challenge is not simply to reduce the number of electrodes

but also to preserve the physiologically relevant information required for reliable seizure characterization when spatial sampling is limited while considering the influence of electrode placement on the available spatial information.

This issue is especially important because epileptic seizures are inherently spatiotemporal phenomena. Generalized seizures typically involve widespread bilateral activity, whereas focal seizures originate within more localized cortical regions and may subsequently recruit neighboring or distant regions as the seizure evolves [1]. Experimental and clinical studies have demonstrated that epileptiform discharges can propagate across cortical regions with structured timing and directional characteristics [27, 28]. Recent investigations of seizure and interictal activity have further emphasized that propagation pathways, directionality, and effective connectivity contain physiologically meaningful information regarding the organization of epileptic networks [29–31]. These observations suggest that the relative timing and spatial progression of epileptiform activity across electrodes may provide discriminative information beyond that contained in the morphology or spectral content of individual EEG channels. However, scalp EEG reflects a mixture of distributed neural activity influenced by volume conduction, source orientation, tissue conductivity, and electrode referencing. Consequently, timing differences observed across scalp electrodes should be interpreted as scalp-level spatiotemporal patterns rather than direct measurements of cortical seizure propagation.

Nevertheless, most automated seizure detection approaches do not explicitly model this propagation behavior. Conventional handcrafted features primarily characterize local signal properties such as amplitude, frequency content, entropy, and waveform complexity, while connectivity measures generally quantify statistical or phase relationships between channels [12, 32]. Deep-learning models can implicitly learn spatial and temporal relationships from multichannel EEG, but the resulting representations may be difficult to interpret, and their computational requirements vary depending on the underlying architecture. Explicitly characterizing the relative timing and spatial organization of epileptiform activity at the scalp level therefore provides an alternative strategy: spatiotemporal information can be encoded directly into a compact and interpretable feature representation, complementing approaches that rely on implicitly learned representations or more computationally intensive models.

In our previous work, we developed a compact propagation- and morphology-based seizure detection framework that characterizes epileptic activity using features derived from spike morphology, scalp-level estimates of apparent propagation speed, and directional organization [33]. Using the Temple University Hospital (TUH) EEG Seizure Corpus (TUSZ), the framework demonstrated that seizure-related spatiotemporal characteristics can provide strong discrimination using short EEG windows and a reduced number of electrodes. Importantly, the resulting feature representation remained low-dimensional and was evaluated using a lightweight gradient-boosted decision tree classifier, supporting its potential for resource-constrained seizure monitoring. The study also showed that propagation-informed features remained effective across generalized, focal, and heterogeneous seizure categories, establishing the feasibility of explicitly incorporating scalp-derived spatiotemporal information into automated detection. However, the propagation model considered the selected electrodes within a two-dimensional (2D) planar representation. The underlying spike-detection procedure, morphology- and propagation-derived features, composite feature construction, and classification framework developed in that study form the methodological foundation of the present work.

A planar representation is convenient for reduced-electrode systems, but it does not fully account for the three-dimensional (3D) geometry of electrodes distributed over the scalp. This distinction may be particularly important for focal seizures, for which localized onset and spatial progression can make detection more dependent on the relative positions of the available electrodes. However, differences in detection performance

between planar and spatial electrode configurations may also arise from variations in electrode count and spatial coverage, making it necessary to distinguish the contribution of electrode geometry from that of additional recording channels. Moreover, wearable EEG devices frequently employ bipolar channels rather than referential multichannel configurations, and the behavior of propagation-based features under such configurations has not been sufficiently investigated. Finally, high performance within a single EEG corpus does not necessarily establish robustness across independent datasets because differences in patient populations, recording hardware, electrode montages, annotation procedures, and signal characteristics can substantially affect model performance. Therefore, cross-dataset validation represents an important step toward determining whether the learned seizure characteristics are transferable beyond the dataset used for model development.

The present study addresses these limitations by extending propagation-based seizure detection from planar representations toward spatially informed and wearable-oriented EEG configurations. First, 2D and 3D propagation representations are systematically evaluated across multiple electrode configurations to determine how electrode geometry and spatial coverage influence seizure detection, with particular attention to focal seizures. Additional matched-electrode comparisons are conducted using identical electrode sets under both coordinate representations to better isolate the contribution of geometric modeling from electrode count. Furthermore, a spike-presence-only ablation is performed to investigate whether propagation-derived information provides additional discriminative value beyond the spatial distribution of detected spikes. Second, the framework is evaluated using a compact four-channel bipolar configuration designed to preserve frontal and temporal information while reducing the number of channels required for monitoring. Third, the generalizability of the resulting model is investigated through cross-dataset evaluation with target-dataset threshold calibration via training on the TUSZ and testing on the Children's Hospital Boston–Massachusetts Institute of Technology (CHB-MIT) Scalp EEG Database. Calibration and temporal correction (TC) strategies are further examined to account for dataset-dependent differences and improve continuous seizure-event detection.

Building on our previous propagation-based seizure detection framework, the present study extends the approach toward spatially informed, wearable-oriented, and cross-dataset-evaluated seizure detection. Specifically, this study

- Systematically evaluates 2D and 3D scalp-level propagation representations across multiple reduced-electrode configurations, with particular attention to the role of spatial information in focal seizure detection, including matched-electrode comparisons to distinguish the contribution of geometric modeling from electrode count
- Investigates the additional discriminative contribution of propagation-derived features through an ablation study comparing the complete feature representation against a spike-presence-only baseline
- Investigates a wearable-oriented four-channel bipolar EEG configuration designed to preserve informative frontal and temporal coverage while substantially reducing the number of required channels
- Evaluates the generalizability of the proposed framework through cross-dataset evaluation with target-dataset threshold calibration, using the TUSZ for model development and the CHB-MIT Scalp EEG Database for held-out evaluation
- Examines calibration and temporal correction strategies for improving cross-dataset and event-level seizure detection while retaining a compact machine-learning framework with potential applicability to resource-constrained monitoring.

The remainder of this paper is organized as follows. Section 2 describes the EEG datasets, signal processing procedures, propagation-based feature extraction, electrode configurations, and evaluation framework. Section 3

presents the experimental results, followed by their interpretation and implications in Section 4. Finally, Section 5 summarizes the main findings and conclusions of the study.

2. Materials and methods

This study extends our previously developed propagation-based seizure detection framework toward spatially informed, reduced-channel, and threshold-calibrated cross-dataset EEG analysis [33]. The framework integrates EEG preprocessing, spike detection, estimation of apparent scalp-level spatiotemporal patterns from inter-channel spike arrival times, morphology- and propagation-based feature extraction, and gradient-boosted seizure classification. Building on the previously established feature extraction and classification methods, the present study investigates 2D and 3D electrode representations through matched-electrode comparisons and feature ablation studies, evaluates a wearable-oriented four-channel bipolar montage, and assesses cross-dataset transfer from TUSZ to CHB-MIT with limited operating-threshold calibration.

2.1. Datasets

Two publicly available scalp EEG databases were used: the TUSZ and the CHB-MIT Scalp EEG Database [34, 35]. The TUSZ contains approximately 504 h of expert-annotated clinical EEG recordings from more than 300 subjects, including approximately 36 h of seizure activity. Recordings predominantly follow the international 10–20 electrode system and include seizure onset, offset, and type annotations. Generalized non-specific seizures (GNSZs) and focal non-specific seizures (FNSZs) were analyzed separately, while a pooled "ANY" category included all eligible seizure types.

The CHB-MIT database was used for cross-dataset evaluation with target-dataset threshold calibration. It contains long-duration bipolar EEG recordings sampled at 256 Hz, with expert seizure annotations, from 24 recording cases corresponding to 23 distinct pediatric subjects with epilepsy. Cases chb01 and chb21 represent recordings from the same individual, acquired approximately 1.5 years apart. All 24 cases were retained for evaluation, with these two cases excluded from threshold calibration to maintain participant-level independence. Unlike the predominantly referential TUSZ recordings, CHB-MIT uses bipolar montages and differs in patient population and acquisition protocols, providing a distinct dataset for assessing cross-dataset transferability.

2.2. EEG preprocessing

EEG preprocessing was performed consistently across the evaluated analyses, with adaptations according to the feature representation and electrode montage. Recordings containing severe artifacts, corrupted intervals, or insufficient usable EEG data were excluded. Signals were linearly detrended and band-pass filtered between 0.5 and 70 Hz for the propagation-based features to preserve transient and higher-frequency components associated with epileptiform activity, whereas a 1–40 Hz passband was used for the traditional feature baseline. A fourth-order Butterworth band-pass filter was implemented using zero-phase forward–backward filtering. A 60 Hz notch filter with a quality factor of 30 was applied to suppress power-line interference.

Spatial re-referencing was performed according to the available electrode configuration. When sufficient bilateral coverage was available, hemispheric common average referencing (hemi-CAR) was applied using the median activity of the available channels within each hemisphere, selected from Fp1/F7/T3/T5/O1/C3 for the left

hemisphere and Fp2/F8/T4/T6/O2/C4 for the right hemisphere, according to the electrode configuration being evaluated. For configurations with insufficient channels for hemi-CAR, global common average referencing was used, while small Laplacian referencing was additionally applied where neighboring electrodes were available to enhance local spatial gradients [36]. The small Laplacian representation was obtained by subtracting the mean activity of the available predefined neighboring electrodes from each re-referenced channel. For the bipolar TUSZ experiments, bipolar channels were generated by subtracting the corresponding pairs of preprocessed referential EEG signals; CHB-MIT recordings were already provided in bipolar form and therefore required no additional bipolar derivation.

2.3. Epoch segmentation and data selection

Following preprocessing, EEG recordings were divided into non-overlapping 2 s windows for propagation-based features and 30 s windows for the traditional feature baseline, based on their previously established performance [33]. An epoch was labeled as seizure only when fully contained within an expert-annotated seizure interval, excluding boundary-overlapping windows from the seizure class. For TUSZ, non-seizure windows were required to be at least 60 s from seizure onset or offset, with additional baseline epochs obtained from background recordings of the same subject when necessary.

For seizure-type-specific analyses, only subjects with at least one annotated seizure of the corresponding type were included. GNSZs and FNSZs were analyzed separately, while the "ANY" dataset pooled all eligible seizure types. All 24 CHB-MIT recording cases were retained for cross-dataset evaluation with target-dataset threshold calibration, with chb01 and chb21 excluded from calibration to maintain participant-level independence. The resulting subject and epoch counts are summarized in **Table 1**.

Table 1. Summary of the datasets after epoch segmentation and subject selection.

Dataset	Window (s)	Subjects	No. of epochs		
			Seizure	Non-seizure	Total
GNSZ (TUSZ)	2	76	25,267	632,837	658,104
GNSZ (TUSZ)	30	73	2008	38,922	40,930
FNSZ (TUSZ)	2	86	45,279	706,626	751,905
ANY (TUSZ)	2	137	78,106	929,880	1,007,986
ANY (CHB-MIT)	2	24	5518	1,747,125	1,752,643

GNSZ: generalized non-specific seizures; FNSZ: focal non-specific seizures; TUSZ: Temple University Hospital EEG Seizure Corpus; ANY: pooled seizure category comprising all eligible seizure types included in TUSZ; CHB-MIT: Children's Hospital Boston–Massachusetts Institute of Technology scalp electroencephalography database.

2.4. EEG electrode configurations

Eight electrode configurations containing 4–12 electrodes from the international 10–20 system were designed to investigate the effects of electrode count, spatial coverage, and geometry on seizure detection (**Figure 1**). Configurations I, II, IV, and VI comprise planar arrangements of 4, 6, 8, and 10 electrodes, respectively,

while Configurations III, V, VII, and VIII additionally incorporate C3 and C4 to examine apparent scalp-level spatiotemporal patterns using 3D electrode geometry. To distinguish differences in spatial modeling from those associated with electrode count and selection, matched-electrode comparisons were performed using identical electrode sets under the complete 2D and 3D propagation representations.

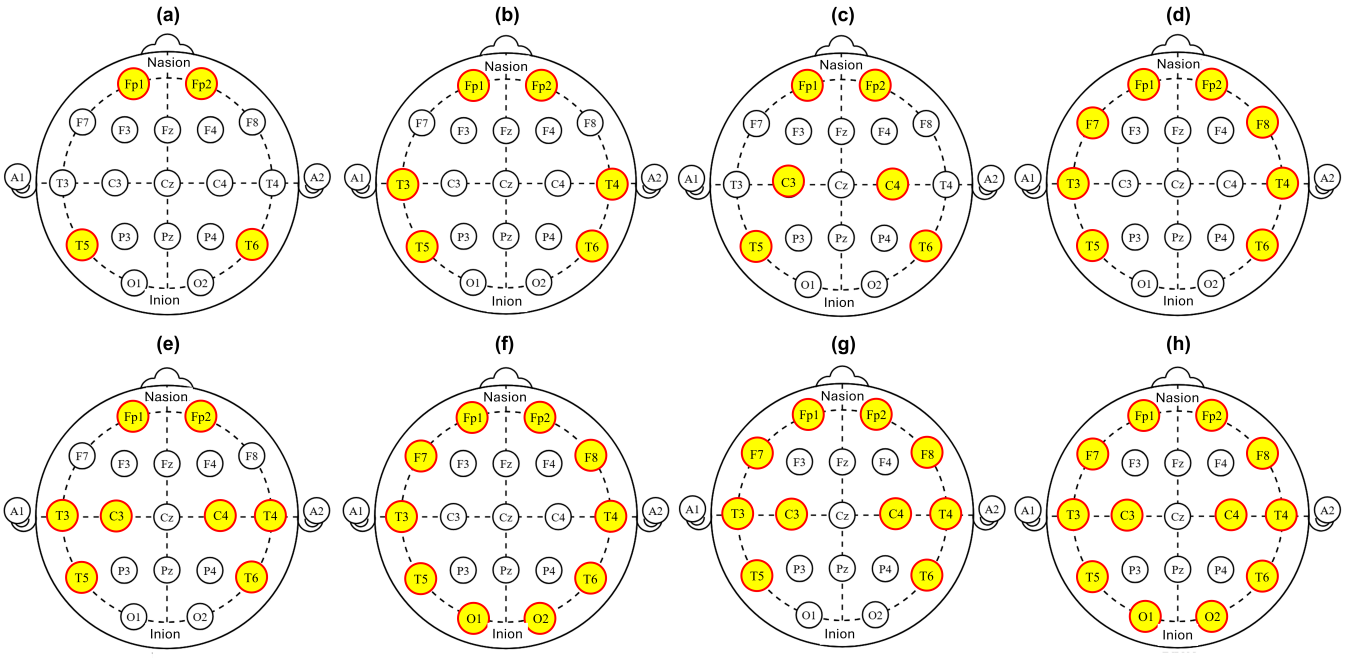

Figure 1. Reduced electrode configurations evaluated in this study. (**a**,**b**,**d**,**f**) show the planar Configurations I, II, IV, and VI comprising 4, 6, 8, and 10 electrodes, respectively. (**c**,**e**,**g**,**h**) show the corresponding spatial Configurations III, V, VII, and VIII, obtained by adding the central electrodes C3 and C4 to Configurations I, II, IV, and VI, respectively. Highlighted electrodes (yellow with red outlines) indicate those included in each configuration.

A wearable-oriented four-channel bipolar montage comprising Fp1–F7, T3–T5, Fp2–F8, and T4–T6 was evaluated on the TUSZ to investigate seizure detection using reduced bilateral frontal and temporal coverage. For cross-dataset evaluation on CHB-MIT, the corresponding bipolar montage comprised Fp1–F7, T7–P7, Fp2–F8, and T8–P8, following the updated electrode nomenclature of the international 10–20 system.

2.5. Propagation- and morphology-based feature extraction

The feature extraction framework builds on our previously published propagation-based seizure detection method [33]. The present study extends the original planar propagation framework to 3D electrode configurations and bipolar EEG channels while retaining the characterization of transient epileptiform activity through its morphology and apparent scalp-level spatiotemporal progression. The individual features and their calculations are described below.

2.5.1. Spike detection

Feature extraction begins by identifying transient, high-slope EEG activity that may correspond to epileptiform discharges. Following spatial referencing, the absolute temporal derivative of each EEG channel is computed and aggregated across channels to emphasize spatially coherent transients while reducing sensitivity to isolated

channel fluctuations. Robust normalization using the median and median absolute deviation is then applied within each analysis window [37]. Candidate spikes are identified as local maxima exceeding an adaptive threshold, with a refractory interval of 100 ms to prevent multiple detections of the same transient. The detected spike locations are subsequently used to extract morphology- and propagation-based features.

2.5.2. Spike arrival-time estimation and wavefront modeling

For each detected spike, its arrival time is estimated independently at each available EEG channel. A candidate timing estimate is first obtained from the dominant negative lobe within a short neighborhood surrounding the detected transient. Local interpolation is used to improve the temporal precision of the peak location. Inter-channel timing is further characterized using generalized cross-correlation with phase transform (GCC-PHAT), which provides a delay estimate between a selected reference channel and the remaining channels [38]. The peak-based and correlation-based estimates are combined using weights of 0.6 and 0.4, respectively, to obtain the final channel-wise arrival-time estimate. The resulting arrival-time pattern provides a spatial representation of the apparent spatiotemporal progression of each transient across the scalp electrode array. This representation characterizes relative timing differences among recorded EEG signals and should not be interpreted as a direct measurement of physical cortical seizure propagation.

For the planar model, let $r_c = [x_c, y_c]^\mathsf{T}$ denote the projected coordinate of electrode c, and let t_c denote the corresponding estimated spike arrival time. A planar wavefront is fitted using ordinary least squares, as expressed in **Equation (1)**:

$$t_c \approx ax_c + by_c + d, \tag{1}$$

where a and b describe the spatial gradient of the arrival-time surface and d represents its temporal offset. The gradient $\nabla t = [a, b]^\mathsf{T}$ therefore characterizes both the orientation and magnitude of propagation across the planar electrode arrangement.

To account for the spatial geometry of electrodes distributed over the scalp, the model was extended to 3D. For an electrode with spatial coordinate $r_c = [x_c, y_c, z_c]^\mathsf{T}$, the arrival-time surface is represented by **Equation (2)**:

$$t_c \approx ax_c + by_c + cz_c + d. \tag{2}$$

The gradient is defined in **Equation (3)**:

$$\nabla t = [a, b, c]^\mathsf{T}, \tag{3}$$

and therefore describes the propagation wavefront in 3D space. This formulation retains spatial differences that are lost when electrodes located at different scalp depths are projected onto a common 2D plane.

For bipolar EEG, an individual channel represents the potential difference between two physical electrodes and therefore does not correspond directly to a single scalp location. To incorporate bipolar signals into the same propagation framework, each bipolar channel was represented by the midpoint of its constituent electrode pair. For electrodes i and j, the midpoint is calculated using **Equation (4)**:

$$r_{ij} = \frac{r_i + r_j}{2}, \tag{4}$$

where r_i and r_j denote their respective spatial coordinates. The resulting midpoint coordinates were used when fitting the propagation wavefront to the bipolar channel arrival times.

2.5.3. Propagation speed and wavefront coherence

The fitted arrival-time surface provides a quantitative estimate of the spatial progression of each detected transient. Propagation speed is obtained from the reciprocal magnitude of the fitted spatial gradient, as expressed in **Equation (5)**:

$$v_i = \frac{R_h}{\|\nabla t_i\|}, \tag{5}$$

where R_h denotes the effective head-radius scaling factor and ∇t_i corresponds to either the 2D or 3D wavefront gradient for spike i. To avoid numerical instability associated with an effectively zero spatial gradient, fits with $\|\nabla t_i\| < 10^{-9}$ were excluded from the propagation-speed calculation. This threshold was used as a numerical tolerance to identify degenerate near-zero gradients rather than as an empirically tuned parameter. Therefore, faster propagation corresponds to a smaller change in arrival time across spatially separated electrodes.

Within each 2 s analysis window, the spike-wise propagation speed distribution is summarized using the median propagation speed and the 95th percentile propagation speed, capturing both typical and rapidly propagating activity. In addition, the median absolute difference between the observed channel-wise arrival times and those predicted by the fitted wavefront is computed as the median wavefront residual. This feature quantifies how well the observed activity conforms to an organized propagating wavefront; smaller residuals indicate greater spatial consistency of the estimated propagation pattern.

2.5.4. Propagation directionality

In addition to propagation speed, the directional organization of successive spikes is used to characterize the spatial coherence of EEG activity. The underlying hypothesis is that seizure activity produces more organized and repeatable propagation trajectories, whereas non-seizure transient activity exhibits comparatively dispersed directions [27, 28]. Representative planar and 3D direction distributions derived from seizure and non-seizure EEG are illustrated in **Figure 2**.

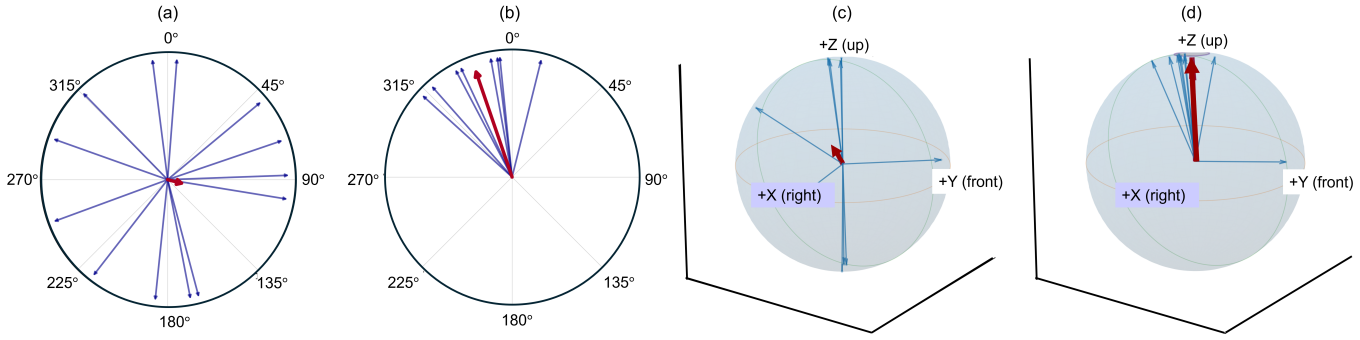

Figure 2. Representative propagation direction distributions obtained from real EEG recordings using 2 s analysis windows. (**a**,**b**) Planar propagation model for representative non-seizure (**a**) and seizure (**b**) epochs, showing a more dispersed directional distribution during the non-seizure epoch and a more concentrated propagation pattern during the seizure epoch. (**c**,**d**) Three-dimensional propagation model for representative non-seizure (**c**) and seizure (**d**) epochs, similarly demonstrating more dispersed spatial propagation during the non-seizure epoch and a more dominant propagation direction during the seizure epoch. Thin purple/blue arrows represent individual spike-wise propagation directions, while thick red arrows indicate the resultant mean propagation direction. The light-purple elliptical region in (**d**) represents the directional concentration region. The X-, Y-, and Z-axes represent the rightward, forward, and upward directions, respectively. EEG: electroencephalography.

For the planar model, the propagation direction of spike i is represented by the heading angle θ_i obtained from its fitted gradient. Given N detected spikes within an epoch, directional concentration is quantified using the mean resultant length, as expressed in **Equation (6)**:

$$R = \sqrt{\left(\frac{1}{N}\sum_{i=1}^{N}\cos\theta_i\right)^2 + \left(\frac{1}{N}\sum_{i=1}^{N}\sin\theta_i\right)^2}. \tag{6}$$

Values of R approaching one indicate strong directional alignment, whereas values approaching zero indicate dispersed propagation. The corresponding circular standard deviation is computed using **Equation (7)**:

$$\sigma_{\text{circ}} = \sqrt{-2\ln(R)}. \tag{7}$$

For the 3D propagation model, each fitted spatial gradient is converted to a unit propagation vector, as expressed in **Equation (8)**:

$$u_i = \frac{[a_i,\, b_i,\, c_i]^{\mathsf{T}}}{\sqrt{a_i^2 + b_i^2 + c_i^2}}. \tag{8}$$

Directional alignment on the unit sphere is then quantified using the 3D resultant length [39], as expressed in **Equation (9)**:

$$R_{3\text{D}} = \left\| \frac{1}{N}\sum_{i=1}^{N} u_i \right\|_2. \tag{9}$$

A concentration parameter is additionally estimated from the resultant length, as expressed in **Equation (10)**:

$$\kappa \approx \frac{R_{3\text{D}}(3 - R_{3\text{D}}^2)}{1 - R_{3\text{D}}^2}, \tag{10}$$

following standard approximations for directional data [39, 40]. Larger $R_{3\text{D}}$ and κ values indicate stronger clustering of the spike-wise propagation directions.

2.5.5. Spike morphology and composite features

Propagation characteristics were complemented by spike-activity and morphology-based descriptors of the detected transients. The number of detected spikes within each analysis window was included as a measure of transient activity. For each spike, a dominant channel was identified based on spike timing and waveform characteristics, with intensity quantified by the absolute amplitude of the dominant negative peak and width representing spike sharpness. Three morphology features were extracted within each window: median spike intensity, median spike width, and median top-k spike intensity, defined as the median of the k largest spike amplitudes ($k = 3$), or all available amplitudes when fewer than k spikes were detected.

Five composite features were also retained from our previous framework to capture interactions among spike intensity, width, propagation speed, and directional organization. The first two use normalized multiplicative combinations, while the remaining three use standardized additive combinations. Normalization and standardization parameters were estimated exclusively from the training data within each validation fold and applied to

the corresponding test data to prevent information leakage. The mathematical definitions of these unchanged composite features are provided in [33].

2.5.6. Final feature representation

The final representation integrates spike occurrence and morphology with wavefront propagation speed, goodness-of-fit, and directional organization. The feature dimensionality remains compact and does not increase with the number of EEG channels, allowing the same classifier architecture to be applied across the evaluated electrode configurations. The complete representation contains 16 feature positions, as summarized in **Table 2**. However, each model uses only 14 of these 16 positions, comprising 12 shared features and two model-specific directionality features. The planar and 3D models use different directionality descriptors: planar configurations use the mean resultant length and circular standard deviation, whereas spatial configurations use the 3D resultant length and concentration parameter. All remaining morphology, propagation speed, residual, and composite features are shared between the two representations.

Table 2. Summary of the features used in the planar and three-dimensional propagation-based representations.

Idx	Feature	Idx	Feature
1	Spike count per window	7	95th percentile propagation speed (m/s)
2	Median spike intensity (μV)	8	Propagation directionality R *
3	Median top-k spike intensity (μV)	9	Circular standard deviation *
4	Median spike width (ms)	10	3D resultant length R_{3D} †
5	Median wavefront residual (ms)	11	Concentration parameter κ †
6	Median propagation speed (m/s)	12–16	Composite features (see [33])

Idx: feature index; 3D: three-dimensional; R: mean resultant length (planar directional concentration); k: number of highest-intensity spikes considered ($k = 3$); R_{3D}: three-dimensional mean resultant length; κ: directional concentration parameter. Features marked with * are computed only for the planar propagation model, whereas features marked with † are computed only for the three-dimensional propagation model. The table lists 16 possible feature positions, of which 14 are used in each model: 12 shared features and two model-specific directionality features.

2.6. Traditional EEG feature extraction

A conventional handcrafted EEG feature set comprising amplitude, temporal, spectral, wavelet, nonlinear, and high-frequency descriptors was used as a baseline for comparison [12, 41]. A total of 43 features were extracted per channel, with their categories and corresponding counts summarized in **Table 3**. For configurations containing six or more electrodes, nine additional connectivity features were extracted using the Pearson correlation coefficient (PCC), phase lag index (PLI), and phase locking value (PLV) across three channel pairs (Fp1–Fp2, T3–T5, and T4–T6) [42, 43]. Thus, the traditional feature dimensionality is $D_{\text{traditional}} = 43N$ for $N < 6$ and $D_{\text{traditional}} = 43N + 9$ for $N \geq 6$, where N denotes the number of electrodes.

To reduce dimensionality, permutation-based feature importance was computed using only the training subjects within each validation fold. The 50 highest-ranked features were then retained and applied to the held-out subject without reselection to prevent information leakage.

Table 3. Summary of traditional EEG features used for baseline seizure detection.

Feature category	No.	Description
Per-channel features		
Amplitude-related	5	RMS amplitude change, mean-to-peak ratio, mean absolute amplitude change, peak-to-peak amplitude, and peak-to-RMS ratio
Time-domain morphology	9	Hjorth mobility and complexity, zero-crossing rate, line length, skewness, kurtosis, envelope variability, relative energy of the second half of the window, and fractal dimension
Spectral power ratios	14	Band power and ratios across δ, θ, α, and β bands, including δ/θ, θ/α, α/β, and $\delta/(\alpha+\beta)$
Wavelet-based features	10	Relative energy and entropy of discrete wavelet coefficients across selected decomposition levels
HFO proxy measures	5	HFO-to-broadband energy ratio, RMS amplitude, mean energy, burst count, and total burst duration
Inter-channel features		
Connectivity metrics	9	Pearson correlation coefficient (PCC), phase lag index (PLI), and phase locking value (PLV), computed for selected channel pairs

A total of 43 features were extracted per EEG channel. Nine additional inter-channel connectivity features were computed only for configurations containing six or more electrodes. EEG: electroencephalography; RMS: root mean square; HFO: high-frequency oscillation; No.: number of features; δ: delta frequency band; θ: theta frequency band; α: alpha frequency band; β: beta frequency band.

2.7. Model training and subject-independent evaluation

A gradient-boosted decision tree classifier (XGBoost, version 3.0.5) was employed for all experiments because it provides an efficient and scalable framework for learning from the extracted tabular EEG features [44]. The same classifier framework was used across electrode configurations and feature representations to enable consistent comparisons. The XGBoost hyperparameters were fixed across all experiments, as summarized in **Table 4**. The positive-class weight was calculated separately within each training fold using the negative-to-positive epoch ratio after downsampling.

Performance on the TUSZ was evaluated using subject-independent leave-one-subject-out (LOSO) cross-validation. In each fold, all epochs from one subject were reserved for testing, while the remaining subjects were used for model development. Data from the held-out subject were not used for training, feature selection, normalization, or estimation of composite-feature parameters. LOSO evaluation was selected instead of the predefined TUSZ train/development/evaluation partition to assess inter-subject generalization across the eligible subjects and obtain subject-level performance estimates for subsequent paired statistical comparisons. This approach ensures that each eligible subject contributes to an independent test fold while maintaining strict separation between training and test subjects.

To address class imbalance, features were initially extracted from all eligible seizure and non-seizure epochs, and downsampling was performed only after separating the training and test subjects. For each subject in the training partition, all seizure epochs were retained, while non-seizure epochs were randomly sampled up to a maximum of five times the number of seizure epochs. Thus, the 1:5 seizure-to-non-seizure ratio represents a

subject-wise upper limit rather than an exact global ratio. When fewer non-seizure epochs were available, all available epochs were retained. Sampling was performed using a fixed random seed of 42, whereas all eligible epochs from the held-out subject were retained for testing under their naturally imbalanced distribution.

The classification threshold was determined independently within each fold using the training data by maximizing the weighted F1 score over a candidate set comprising 0.3, 0.5, 0.7, and an additional threshold derived from the training-set precision–recall curve. Candidate thresholds were evaluated using in-sample predictions from the fitted model, and the threshold yielding the highest training-set weighted F1 score was selected. Although the held-out subject was not involved in threshold selection, using in-sample predictions rather than a separate validation subset may introduce optimistic bias in the selected operating threshold. The selected threshold was then applied without modification to the held-out subject. Performance was quantified using accuracy, sensitivity, specificity, weighted F1 score, and area under the receiver operating characteristic curve (AUC). Subject-level metrics were computed independently for each LOSO fold and subsequently summarized across subjects.

Table 4. XGBoost hyperparameters used across all experiments.

Hyperparameter	Value	Hyperparameter	Value
`n_estimators`	900	`reg_lambda`	1.0
`max_depth`	7	`reg_alpha`	0.0
`learning_rate`	0.05	`tree_method`	`hist`
`subsample`	0.8	`eval_metric`	`logloss`
`colsample_bytree`	0.8	`random_state`	42
`min_child_weight`	3	`scale_pos_weight`	$N_{\text{negative}}/N_{\text{positive}}$

The positive-class weight was recalculated within each LOSO training fold after subject-wise downsampling. All other listed hyperparameters remained fixed across experiments. LOSO: leave-one-subject-out; N_{negative}: number of negative-class training samples; N_{positive}: number of positive-class training samples.

2.8. Independent cross-dataset evaluation

To assess generalization beyond TUSZ, the model was developed using the pooled TUSZ "ANY" seizure dataset and evaluated on the independent CHB-MIT database without classifier retraining. This evaluation was designed as a threshold-calibrated cross-dataset assessment rather than a fully zero-shot transfer experiment, since a limited subset of CHB-MIT recordings was used exclusively to adjust the classification operating threshold. The wearable-oriented four-channel bipolar configuration was used because corresponding frontal and temporal channel pairs could be established across the two databases. To account for differences in the classifier operating point between datasets, four randomly selected CHB-MIT recording cases were used for threshold calibration, while the remaining 20 recording cases were reserved exclusively for evaluation. Cases chb01 and chb21, representing the same individual, were excluded from calibration and retained for evaluation to ensure participant-level independence. This procedure was repeated five times using different calibration cases. For each repetition, the classification threshold was selected using only the calibration cases and subsequently

applied without modification to the corresponding 20 held-out evaluation cases. No CHB-MIT data were used to retrain the classifier or modify its learned parameters.

Two feature representations were examined. The full representation retained all 14 features of the bipolar propagation framework, whereas the reduced representation contained nine features, replacing the spike count with a binary spike-presence indicator and excluding the five composite features. Separate classifiers were trained on the TUSZ for the two feature representations and evaluated on the CHB-MIT using their respective calibrated thresholds. Epoch-based performance was evaluated using the same metrics as the TUSZ experiments, while the continuous recordings and seizure annotations available in the CHB-MIT were additionally used for event-based evaluation. Accordingly, the reported cross-dataset results reflect generalization to held-out CHB-MIT recording cases following limited threshold calibration rather than performance under a completely adaptation-free deployment setting. Because the reduced feature representation and temporal correction settings were selected based on CHB-MIT results, their performance requires further independent validation.

2.9. Temporal correction and event-based evaluation

A simple temporal correction procedure was applied to the CHB-MIT epoch-wise binary predictions obtained after threshold calibration to reduce short isolated classification errors without modifying the trained classifier or its predicted probabilities. A single isolated non-seizure epoch occurring within an otherwise continuous sequence of predicted seizure epochs was relabeled as seizure. Conversely, short isolated sequences of predicted seizure epochs surrounded by non-seizure activity were removed. Four correction levels, TC #1–#4, were evaluated by removing isolated seizure sequences containing up to one, two, three, or four consecutive 2 s epochs, respectively, allowing the trade-off between seizure sensitivity and false-alarm suppression to be assessed.

In addition to epoch-based metrics, event-based sensitivity, the false alarm rate per hour (FAR/h), and detection latency were calculated. A seizure event was considered detected when at least one positive prediction occurred within its expert-annotated interval, and event sensitivity was defined as the proportion of annotated seizure events detected. FAR/h was calculated from the number of contiguous false-positive seizure segments occurring outside annotated seizure intervals and normalized by the corresponding non-seizure monitoring duration. Detection latency was measured from the annotated seizure onset to the first associated positive prediction for each successfully detected seizure event. These measures characterize the performance of the framework under continuous monitoring conditions beyond individual epoch classification.

3. Results

3.1. Evaluation of electrode configurations

This subsection presents the subject-independent LOSO cross-validation results obtained on the TUSZ. The proposed propagation-based framework was first evaluated on the GNSZ dataset and compared with the best-performing traditional feature-based approach. As established in our previous study, the traditional feature representation achieved its highest performance using 30 s epochs, whereas the propagation-based framework performed best using substantially shorter 2 s epochs. Accordingly, **Figure 3** compares the two approaches at their respective best-performing temporal resolutions.

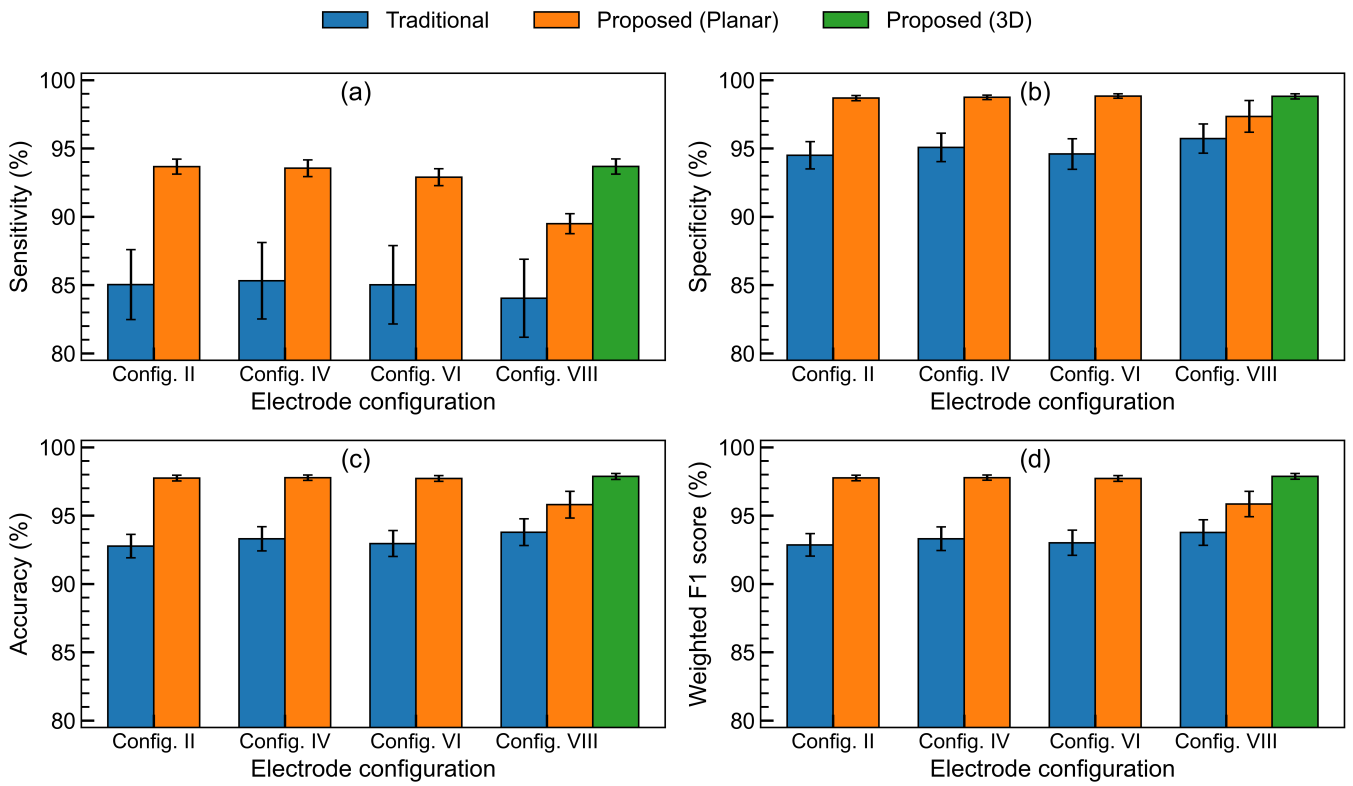

Figure 3. LOSO performance for the GNSZ seizure type on the TUSZ comparing the traditional feature-based method (30 s epochs) and the proposed propagation-based method (2 s epochs): (**a**) sensitivity, (**b**) specificity, (**c**) accuracy, and (**d**) weighted F1 score. Configurations II, IV, and VI are planar, whereas Configuration VIII is spatial and evaluated using both the 2D planar and proposed 3D propagation models. The propagation-based method consistently outperforms the traditional method across the evaluated metrics, while 3D spatial modeling further improves the performance of Configuration VIII. Accuracy is included for illustrative completeness; however, due to the highly imbalanced nature of the seizure detection problem, it may not adequately reflect classification performance. Error bars indicate the standard error across subjects. LOSO: leave-one-subject-out; GNSZ: generalized non-specific seizures; TUSZ: Temple University Hospital EEG Seizure Corpus; 2D: two-dimensional; 3D: three-dimensional.

Configurations II, IV, and VI correspond to planar electrode arrangements containing 6, 8, and 10 electrodes, respectively, whereas Configuration VIII represents the 12-electrode spatial arrangement. For Configuration VIII, performance was evaluated using both the original 2D propagation model and the proposed 3D propagation model to directly examine the effect of preserving the spatial electrode geometry. As shown in **Figure 3**, the propagation-based framework consistently outperformed the traditional feature-based approach across the evaluated performance metrics, including when fewer planar electrodes were used. When the 2D propagation model was applied to the 12-electrode spatial configuration, sensitivity decreased noticeably. In contrast, incorporating the 3D electrode coordinates restored the detection performance to a level comparable with that obtained from the planar configurations, suggesting a potential contribution from the spatial representation. This observation is investigated more systematically in the subsequent ablation study using a matched 12-electrode configuration.

To determine whether the observed improvements over the traditional approach were statistically supported, paired subject-level comparisons were performed for Configurations II, IV, VI, and VIII. The comparisons involved the traditional feature representation and the proposed 2D propagation-based representation, excluding the 3D model, which is examined separately in the ablation study. Two-sided Wilcoxon signed-rank tests were

performed, with Holm correction applied across all 16 comparisons (four configurations and four performance metrics). Hereafter, Holm-adjusted p-values are denoted by p_{adj}. The mean paired differences and their 95% confidence intervals, estimated using 10,000 paired subject-level percentile-bootstrap resamples, are summarized in **Table 5**.

Table 5. Paired statistical comparison between the traditional feature-based method and the proposed 2D propagation-based method for GNSZs across four electrode configurations.

Configuration	Accuracy		Weighted F1		Sensitivity		Specificity	
	95% CI	p_{adj}	95% CI	p_{adj}	95% CI	p_{adj}	95% CI	p_{adj}
II (6E)	[3.39, 6.98]	<0.0001	[3.38, 6.75]	<0.0001	[4.36, 14.59]	0.0216	[2.28, 6.34]	0.0030
IV (8E)	[2.99, 6.49]	<0.0001	[3.02, 6.46]	<0.0001	[3.87, 14.90]	0.0251	[1.83, 5.94]	0.0101
VI (10E)	[3.03, 6.89]	0.0002	[3.03, 6.75]	0.0002	[2.93, 14.10]	0.0474	[2.15, 6.77]	0.0369
VIII (12E, 2D)	[−1.82, 4.85]	0.0704	[−1.55, 4.83]	0.0704	[0.85, 11.93]	0.3757	[−2.94, 4.58]	0.1504

The 95% confidence intervals (CIs) represent the mean paired difference (Proposed − Traditional), expressed in percentage points. Adjusted p-values were obtained using two-sided Wilcoxon signed-rank tests with Holm correction across all 16 comparisons. GNSZ: generalized non-specific seizures; CI: confidence interval; E: electrodes; 2D: two-dimensional; p_{adj}: Holm-adjusted p-value.

The proposed method achieved statistically significant improvements in accuracy, weighted F1 score, sensitivity, and specificity for Configurations II, IV, and VI (all $p_{\text{adj}} < 0.05$). Although the mean performance differences favored the proposed method across all four metrics for Configuration VIII, none remained statistically significant after Holm correction. These findings provide statistical support for the performance advantages of the proposed planar propagation-based representation, particularly for the 6-, 8-, and 10-electrode configurations, while indicating that the strength of the evidence varies with electrode arrangement.

To further examine whether the performance differences between the proposed and traditional feature representations persist under matched experimental conditions, an additional comparison was conducted using the eight-electrode Configuration IV on the GNSZ dataset. Both approaches were evaluated using identical preprocessing, filtering, 2 s epochs, and the same 76 subjects. Paired subject-level comparisons showed that the proposed features improved mean accuracy from 85.16% to 97.78%, sensitivity from 62.04% to 93.55%, specificity from 90.27% to 98.74%, and weighted F1 score from 84.73% to 97.78%. The corresponding mean paired improvements were 12.62 percentage points (95% confidence intervals (CIs): [10.17, 15.36]), 31.50 percentage points (95% CI: [26.15, 37.15]), 8.47 percentage points (95% CI: [5.44, 12.14]), and 13.05 percentage points (95% CI: [10.67, 15.81]), respectively. All four improvements were statistically significant after Holm correction ($p_{\text{adj}} < 0.0001$). These findings provide additional evidence that the performance advantage of the proposed propagation-based features persists when epoch duration, preprocessing, electrode configuration, and subject population are held constant. However, threshold selection based on in-sample training predictions remains a limitation that may affect the estimated operating point despite maintaining subject-level separation between training and testing.

Figure 4 further summarizes the LOSO sensitivity, specificity, and weighted F1 score obtained using all eight electrode configurations. Planar configurations (I, II, IV, and VI) were evaluated using the 2D propagation model, whereas spatial configurations (III, V, VII, and VIII) were evaluated using the proposed 3D model. For GNSZs, performance remained consistently high across the different configurations, with only minor changes in

sensitivity, specificity, and weighted F1 score as electrode count and geometry were varied. This indicates that generalized seizure activity can be detected reliably even with relatively sparse electrode coverage.

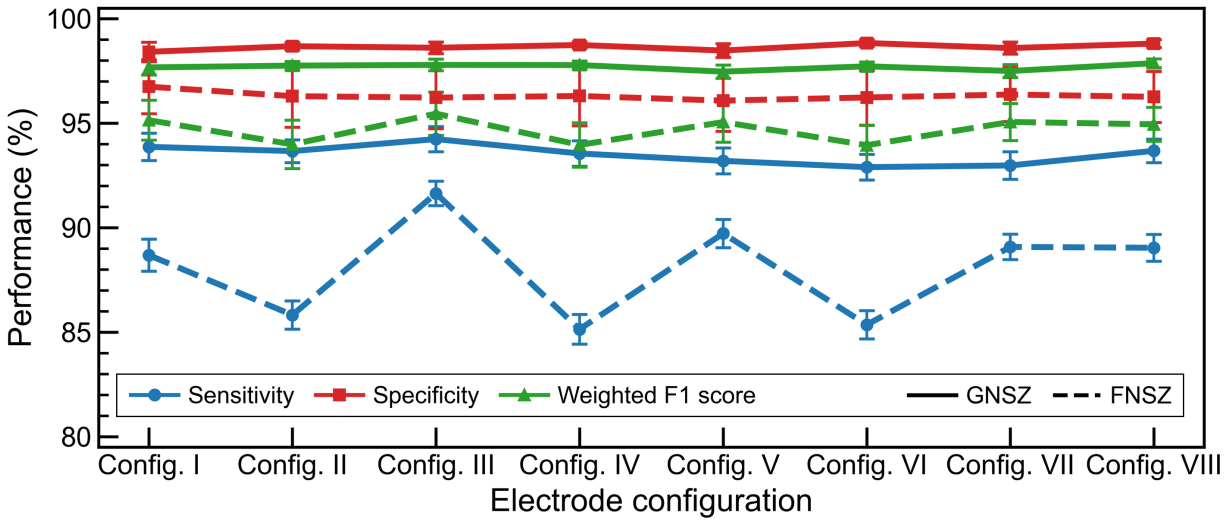

Figure 4. LOSO sensitivity, specificity, and weighted F1 score of the eight electrode configurations on the TUSZ. Solid and dashed lines represent GNSZs and FNSZs, respectively. Planar and spatial configurations use the 2D and proposed 3D propagation models, respectively. Performance remains consistently high across electrode configurations for GNSZs, whereas FNSZs exhibit greater sensitivity to electrode configuration, with improved sensitivity observed for spatial configurations incorporating central electrodes. Specificity and weighted F1 score remain comparatively stable for both seizure types. Error bars denote the standard error across subjects. LOSO: leave-one-subject-out; TUSZ: Temple University Hospital EEG Seizure Corpus; GNSZ: generalized non-specific seizures; FNSZ: focal non-specific seizures; 2D: two-dimensional; 3D: three-dimensional.

A more pronounced configuration-dependent effect was observed for FNSZs. Spatial configurations consistently achieved higher sensitivity than their corresponding planar configurations with comparable electrode counts, suggesting that electrode coverage and spatial representation may influence focal seizure detection. The highest FNSZ sensitivity, 91.48%, was obtained using the 6-electrode spatial Configuration III. However, the 12-electrode spatial Configuration VIII provided the strongest overall balance of performance, improving sensitivity and weighted F1 score by 3.22% and 0.96%, respectively, relative to the 6-electrode planar Configuration II, while maintaining nearly identical specificity (96.26% versus 96.29%).

Overall, these results show that the proposed framework maintains reliable seizure detection across a range of electrode configurations and can support substantial electrode reduction when the remaining channels preserve informative spatial coverage. To further investigate the contributions of individual methodological components, two complementary ablation studies were conducted. The first examined whether the propagation-derived features provide additional discriminative information beyond spike presence alone, while the second evaluated whether 3D spatial modeling offers performance advantages over the 2D representation when the electrode configuration is held constant.

3.2. Ablation studies

Two complementary ablation studies were conducted to investigate the contributions of the proposed propagation-based framework. The first examined whether propagation-derived features provide additional discriminative information beyond the presence of detectable spikes. The second evaluated the contribution of

3D electrode geometry by comparing 2D and 3D propagation models using identical electrode configurations. Both studies employed subject-independent LOSO cross-validation.

3.2.1. Contribution of propagation-derived features

The first ablation study was conducted on both the GNSZ and FNSZ datasets using the 12-electrode 3D configuration (Configuration VIII). This configuration was selected as a representative electrode arrangement for evaluating the contribution of the proposed features. The full propagation-based model was compared against two reduced models: a spike-presence-only model and a spike-presence-plus-morphology model. The same spike-detection procedure described in Section 2.5.1 was applied to both seizure and non-seizure epochs, producing 12 binary indicators representing spike presence at the corresponding electrodes. The second reduced model additionally incorporated three spike-morphology features: median spike intensity, median top-k spike intensity, and median spike width. Both reduced models excluded the explicitly engineered propagation descriptors, allowing the contribution of the complete feature representation to be evaluated beyond spike presence and morphology.

Table 6 summarizes the paired comparisons for GNSZs and FNSZs. Compared with the spike-presence-only model, the full 3D model achieved significantly higher accuracy, sensitivity, and weighted F1 score for both seizure types ($p_{\text{adj}} < 0.0001$). The sensitivity improvements were particularly substantial, with 95% confidence intervals of [92.11, 94.48] and [87.49, 90.05] percentage points for GNSZs and FNSZs, respectively. When spike morphology was additionally included in the reduced model, the full 3D model continued to demonstrate significantly higher performance across all four metrics for both seizure types ($p_{\text{adj}} < 0.05$). The corresponding sensitivity improvements were [30.62, 32.73] and [37.28, 38.94] percentage points, respectively.

Table 6. Ablation comparison of the full 3D propagation-based model against the spike-presence-only and spike-presence-plus-morphology models for GNSZs and FNSZs using Configuration VIII (12 electrodes).

Seizure	Metric	Ablation model vs. full 3D	
		Spike presence	Spike presence + morphology
GNSZ	Accuracy	$<0.0001, [15.46, 19.44]$	$<0.0001, [11.63, 12.32]$
	Sensitivity	$<0.0001, [92.11, 94.48]$	$<0.0001, [30.62, 32.73]$
	Specificity	$0.0711, [-0.84, 0.04]$	$0.0018, [0.23, 0.98]$
	Weighted F1 score	$<0.0001, [22.80, 27.61]$	$<0.0001, [6.18, 7.16]$
FNSZ	Accuracy	$<0.0001, [14.37, 20.48]$	$<0.0001, [17.37, 18.47]$
	Sensitivity	$<0.0001, [87.49, 90.05]$	$<0.0001, [37.28, 38.94]$
	Specificity	$0.0064, [-5.91, -1.00]$	$0.0005, [1.23, 3.38]$
	Weighted F1 score	$<0.0001, [22.65, 29.69]$	$<0.0001, [11.76, 13.11]$

Each cell reports the Holm-adjusted p-value followed by the 95% confidence interval (CI) of the mean paired difference (Full 3D − Ablation Model), expressed in percentage points. Confidence intervals were calculated using paired t-tests, with Holm correction applied across the four performance metrics within each seizure-type/model comparison. Statistical significance was defined as $p_{\text{adj}} < 0.05$. Values below 0.0001 are reported as $p_{\text{adj}} < 0.0001$. Positive differences indicate higher performance for the full 3D model. GNSZ: generalized non-specific seizures; FNSZ: focal non-specific seizures; CI: confidence interval; 3D: three-dimensional; p_{adj}: Holm-adjusted p-value.

The extremely low sensitivity of the spike-presence-only model can be partly explained by the high prevalence of detected spike-like events in both seizure and non-seizure epochs. Specifically, at least one spike-positive electrode was identified in 95.93% of seizure epochs and 94.07% of non-seizure epochs for GNSZs compared with 94.88% and 94.23%, respectively, for FNSZs. This widespread detection of local spike-like events under the z-score-based detection procedure limits the discriminative information provided by spike occurrence alone. Combined with the substantial class imbalance, this limited discrimination may have contributed to the classifier predominantly predicting the non-seizure class, resulting in high specificity but extremely low sensitivity.

3.2.2. Contribution of three-dimensional spatial modeling

The second ablation study investigated whether incorporating 3D electrode geometry provides additional benefits over the original 2D propagation representation. Four spatial electrode configurations (III, V, VII, and VIII) were evaluated using both the 2D and 3D models on the GNSZ and FNSZ datasets. For each configuration, the same electrodes were used in both models, allowing the complete 2D and 3D propagation-based representations to be compared without introducing differences in electrode count or placement. The directionality and directional concentration features were defined within their respective 2D and 3D coordinate systems, as these descriptors are integral components of each propagation representation. Thus, the comparison evaluates the combined effects of spatial dimensionality and its associated propagation descriptors rather than isolating coordinate dimensionality alone. Paired subject-level comparisons were performed separately for each seizure type and configuration.

Table 7 summarizes the paired statistical comparisons across all four performance metrics, revealing distinct patterns for the two seizure types. For GNSZs, the differences between the 2D and 3D models were relatively small for Configurations III, V, and VII, with none reaching statistical significance after Holm correction. However, Configuration VIII demonstrated a significant improvement in sensitivity from 91.67% to 93.67%, corresponding to a mean paired increase of 2.01 percentage points ($p_{\text{adj}} = 0.0005$). For this configuration, accuracy and weighted F1 score also improved significantly by 0.96 and 0.94 percentage points, respectively (both $p_{\text{adj}} = 0.0047$), whereas the specificity difference was not statistically significant ($p_{\text{adj}} = 0.0936$).

A more consistent benefit from 3D modeling was observed for FNSZs. Sensitivity improved significantly across all four configurations, with mean paired increases of 3.65, 4.36, 4.44, and 4.26 percentage points for Configurations III, V, VII, and VIII, respectively (all $p_{\text{adj}} < 0.0001$). Accuracy and weighted F1 score also improved significantly across all four configurations (all $p_{\text{adj}} < 0.0001$). Although the 3D model yielded slightly lower mean specificity than the 2D model for FNSZs, none of these differences was statistically significant after correction.

These findings indicate that the contribution of 3D spatial modeling depends on seizure type and electrode configuration. For GNSZs, the 2D representation generally provided comparable performance for the smaller configurations, while the 12-electrode Configuration VIII benefited significantly from incorporating spatial geometry. In contrast, for FNSZs, there were consistent improvements in sensitivity, accuracy, and weighted F1 score across all four matched configurations. This supports the value of preserving 3D electrode geometry for representing apparent scalp-level spike progression, particularly in the detection of focal seizures.

Table 7. Paired statistical comparison between the 2D and 3D propagation models using identical electrode configurations for GNSZs and FNSZs.

Seizure	Metric	Configuration			
		III	V	VII	VIII
GNSZ	Accuracy	$1.0000, [-0.34, 0.21]$	$1.0000, [-0.37, 0.42]$	$0.6329, [-0.14, 0.64]$	$0.0047, [0.36, 1.56]$
	Sensitivity	$0.8898, [-1.26, 0.30]$	$1.0000, [-1.12, 0.94]$	$0.6329, [-0.24, 1.41]$	$0.0005, [1.01, 3.00]$
	Specificity	$1.0000, [-0.24, 0.28]$	$1.0000, [-0.44, 0.44]$	$0.6329, [-0.26, 0.60]$	$0.0936, [-0.12, 1.48]$
	Weighted F1	$1.0000, [-0.34, 0.20]$	$1.0000, [-0.34, 0.41]$	$0.6329, [-0.11, 0.63]$	$0.0047, [0.37, 1.51]$
FNSZ	Accuracy	$0.0246, [0.19, 1.22]$	$0.0002, [0.54, 1.57]$	$0.0001, [0.56, 1.55]$	$0.0134, [0.25, 1.51]$
	Sensitivity	$<0.0001, [2.70, 4.59]$	$<0.0001, [3.59, 5.12]$	$<0.0001, [3.58, 5.31]$	$<0.0001, [3.56, 4.96]$
	Specificity	$0.1863, [-1.08, 0.21]$	$0.3194, [-0.56, 0.18]$	$0.8666, [-0.35, 0.29]$	$0.5075, [-0.66, 0.33]$
	Weighted F1	$0.0246, [0.15, 1.14]$	$0.0001, [0.56, 1.48]$	$<0.0001, [0.64, 1.50]$	$0.0065, [0.34, 1.48]$

Each cell reports the Holm-adjusted p-value followed by the 95% confidence interval (CI) of the mean paired difference (3D − 2D), expressed in percentage points. Confidence intervals were calculated using paired t-tests, with Holm correction applied across the four performance metrics within each seizure-type/configuration comparison. Statistical significance was defined as $p_{adj} < 0.05$. Values below 0.0001 are reported as $p_{adj} < 0.0001$. The paired sample sizes were $N = 76$ for the GNSZs and $N = 86$ for all FNSZ configurations. Weighted F1: weighted F1 score; GNSZ: generalized non-specific seizures; FNSZ: focal non-specific seizures; CI: confidence interval; 2D: two-dimensional; 3D: three-dimensional; p_{adj}: Holm-adjusted p-value; N: number of paired subjects.

3.3. Evaluation of the four-channel bipolar configuration

The preceding electrode-configuration analysis and ablation studies demonstrated that the proposed propagation-derived features provide additional discriminative information beyond spike presence alone and that incorporating 3D electrode geometry can improve detection performance, particularly for focal seizures. These findings motivate further investigation of whether the framework can maintain reliable performance using a substantially reduced, wearable-oriented electrode arrangement. However, the TUSZ primarily contains referential EEG recordings, whereas the CHB-MIT database used for independent validation provides bipolar recordings. Therefore, before performing cross-dataset evaluation, the proposed framework was evaluated on the TUSZ using a wearable-oriented four-channel bipolar montage comprising Fp1–F7, T3–T5, Fp2–F8, and T4–T6. This configuration preserves bilateral frontal and temporal coverage while substantially reducing the number of EEG channels required for seizure detection.

Figure 5 presents the pooled confusion matrices and subject-wise averaged receiver operating characteristic (ROC) curves obtained using LOSO cross-validation on the TUSZ for the GNSZ, FNSZ, and ANY seizure datasets. The corresponding quantitative results are summarized in **Table 8**. The proposed framework achieved sensitivities of 92.34%, 83.94%, and 92.01% for the GNSZ, FNSZ, and ANY databases, respectively, with corresponding AUC values of 0.975, 0.959, and 0.962. Specificity exceeded 96.5% for all three datasets, while the weighted F1 score remained above 95.3%. The pooled confusion matrices similarly show that more than 96% of non-seizure epochs and at least 84.7% of seizure epochs were correctly classified across the three evaluation settings.

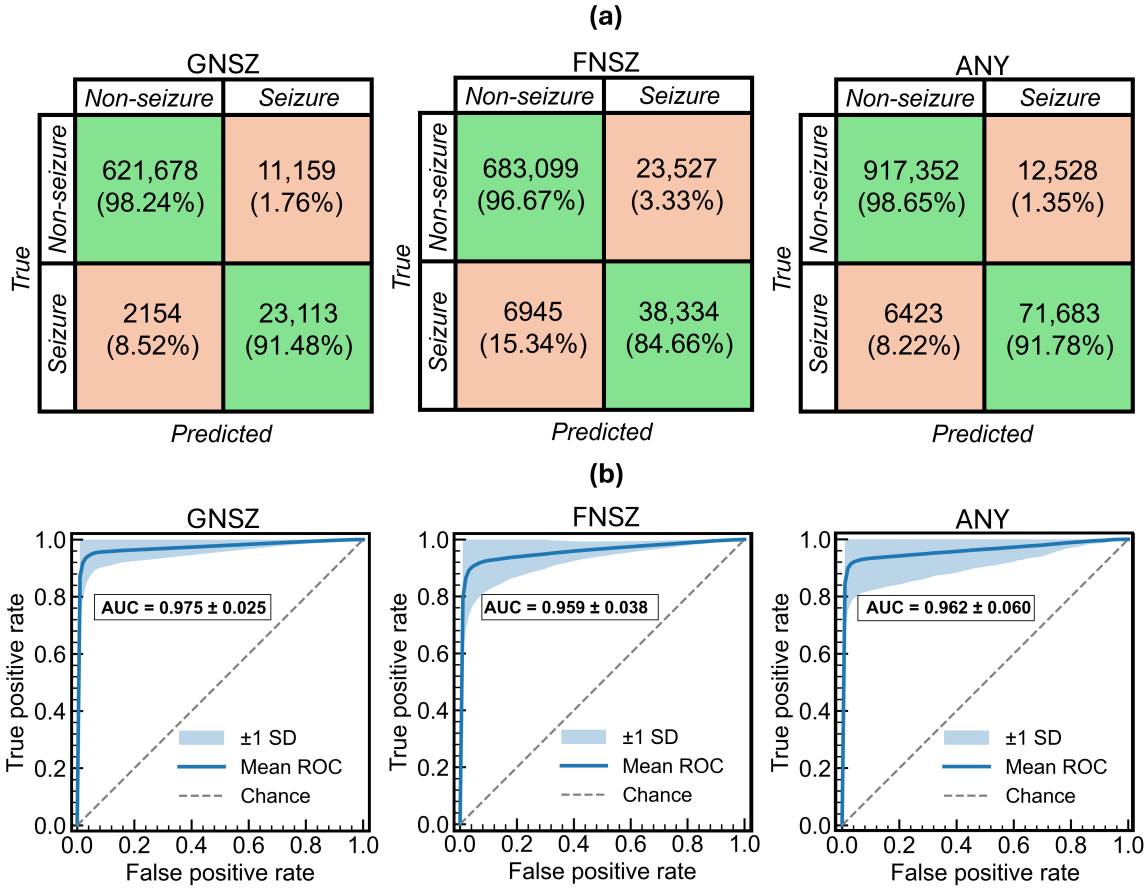

Figure 5. LOSO performance of the proposed four-channel bipolar configuration on the TUSZ. (**a**) Pooled confusion matrices across all subjects, with percentages normalized within each true class. Green cells indicate correct classifications, whereas salmon-colored cells indicate misclassifications. High correct classification rates are observed for both seizure and non-seizure epochs, although there is a comparatively higher seizure misclassification rate for FNSZs. (**b**) Subject-wise averaged ROC curves for the GNSZ, FNSZ, and ANY datasets, demonstrating strong discrimination across all three evaluations, with mean AUC values exceeding 0.95. Blue solid lines represent the mean ROC curves, light-blue shaded regions indicate ±1 standard deviation across subjects, and gray dashed lines represent chance-level classification. LOSO: leave-one-subject-out; TUSZ: Temple University Hospital EEG Seizure Corpus; EEG: electroencephalography; GNSZ: generalized non-specific seizures; FNSZ: focal non-specific seizures; ANY: pooled category including all eligible seizure types; ROC: receiver operating characteristic; AUC: area under the receiver operating characteristic curve; SD: standard deviation.

Performance remained particularly strong for GNSZs and for the heterogeneous ANY dataset, despite the substantial reduction to four bipolar channels. For FNSZs, sensitivity was lower at 83.94%, consistent with the greater sensitivity of focal seizure detection to electrode coverage and spatial representation observed in the preceding electrode-configuration analysis and ablation studies. Nevertheless, the corresponding AUC remained high at 0.959, indicating strong overall discrimination between seizure and non-seizure epochs.

Overall, these results demonstrate that the propagation-based framework maintains reliable seizure detection using only four bipolar channels while retaining bilateral frontal and temporal information. Having established the effectiveness of this compact bipolar configuration within the TUSZ, the following subsection evaluates whether the learned seizure characteristics generalize to the independent CHB-MIT database.

Table 8. Performance comparison for TUSZ LOSO and TUSZ-trained/CHB-MIT-tested evaluations using wearable-oriented four-channel bipolar montages with and without temporal correction (TC).

Evaluation	Type	Method	Epoch-based metrics					Event-based metrics		
			Accuracy (%)	Sensitivity (%)	Specificity (%)	Weighted F1 (%)	AUC	FAR/h	Sensitivity (%)	Latency (s)
TUSZ	GNSZ	LOSO	98.08 ± 0.28 [97.52, 98.64]	92.34 ± 0.75 [90.85, 93.83]	98.56 ± 0.23 [98.10, 99.02]	98.31 ± 0.23 [97.85, 98.77]	0.975 ± 0.003 [0.969, 0.981]	–	–	–
	FNSZ	LOSO	95.07 ± 1.10 [92.88, 97.26]	83.94 ± 0.80 [82.35, 85.53]	96.60 ± 1.26 [94.09, 99.11]	95.38 ± 0.96 [93.47, 97.29]	0.959 ± 0.004 [0.951, 0.967]	–	–	–
	ANY	LOSO	96.27 ± 0.87 [94.55, 97.99]	92.01 ± 0.45 [91.12, 92.90]	96.56 ± 1.12 [94.35, 98.77]	96.38 ± 0.83 [94.74, 98.02]	0.962 ± 0.005 [0.952, 0.972]	–	–	–
CHB-MIT	ANY	Full set	99.37 ± 0.08 [99.20, 99.54]	78.54 ± 1.51 [75.42, 81.66]	99.47 ± 0.08 [99.30, 99.64]	99.51 ± 0.06 [99.39, 99.63]	0.832 ± 0.010 [0.811, 0.853]	5.772 ± 0.470	100.00 ± 0.00	0.764 ± 0.059
		Reduced	99.29 ± 0.09 [99.10, 99.48]	84.05 ± 1.48 [80.99, 87.11]	99.37 ± 0.09 [99.18, 99.56]	99.47 ± 0.06 [99.35, 99.59]	0.882 ± 0.009 [0.863, 0.901]	6.835 ± 0.535	100.00 ± 0.00	0.506 ± 0.014
		TC #1	99.56 ± 0.09 [99.37, 99.75]	93.15 ± 1.25 [90.56, 95.74]	99.60 ± 0.09 [99.41, 99.79]	99.67 ± 0.06 [99.55, 99.79]	0.882 ± 0.009 [0.863, 0.901]	1.801 ± 0.269	100.00 ± 0.00	0.807 ± 0.073
		TC #2	99.63 ± 0.08 [99.46, 99.80]	92.64 ± 1.48 [89.58, 95.70]	99.67 ± 0.08 [99.50, 99.84]	99.71 ± 0.06 [99.59, 99.83]	0.882 ± 0.009 [0.863, 0.901]	1.223 ± 0.201	99.49 ± 0.29	1.299 ± 0.099
		TC #3	99.69 ± 0.08 [99.52, 99.86]	89.46 ± 2.87 [83.52, 95.40]	99.74 ± 0.08 [99.57, 99.91]	99.76 ± 0.05 [99.66, 99.86]	0.882 ± 0.009 [0.863, 0.901]	0.793 ± 0.136	93.27 ± 0.37	2.034 ± 0.227
		TC #4	99.73 ± 0.07 [99.59, 99.87]	87.82 ± 3.61 [80.35, 95.29]	99.78 ± 0.07 [99.64, 99.92]	99.78 ± 0.05 [99.68, 99.88]	0.882 ± 0.009 [0.863, 0.901]	0.572 ± 0.112	91.10 ± 0.69	2.143 ± 0.166

For TUSZ, the montage comprises Fp1–F7, T3–T5, Fp2–F8, and T4–T6. For CHB-MIT, the anatomically equivalent montage comprises Fp1–F7, T7–P7, Fp2–F8, and T8–P8. Epoch-based metrics are reported as mean ± standard error (SE), with 95% confidence intervals (CI) shown beneath each estimate. Event-based metrics are reported as mean ± standard deviation (SD) across five evaluation repetitions. Temporal correction (TC) #1–#4 fills a single isolated non-seizure epoch within seizure events and removes isolated seizure segments of one, two, three, and four consecutive epochs, respectively. "Full set" refers to the same 14-feature propagation-based representation used in the TUSZ LOSO evaluation. "Reduced" replaces the spike-count feature with a binary spike-presence indicator and removes the composite features derived from training-set statistics. Accuracy is included for comparison with the existing literature; however, because seizure detection is a highly imbalanced classification task, metrics such as sensitivity, specificity, weighted F1 score, AUC, and event-based measures provide a more informative evaluation of model performance. TUSZ: Temple University Hospital EEG Seizure Corpus; CHB-MIT: Children's Hospital Boston–Massachusetts Institute of Technology EEG Database; GNSZ: generalized non-specific seizures; FNSZ: focal non-specific seizures; ANY: pooled category including all eligible seizure types; LOSO: leave-one-subject-out cross-validation; TC: temporal correction; Weighted F1: weighted F1 score; FAR/h: false alarms per hour; AUC: area under the receiver operating characteristic curve; SE: standard error; CI: confidence interval; SD: standard deviation; –: not applicable.

3.4. Cross-dataset evaluation

To evaluate the generalizability of the proposed framework beyond the dataset used for model development, the classifier trained on the TUSZ using the four-channel bipolar montage Fp1–F7, T3–T5, Fp2–F8, and T4–T6 was evaluated on the independent CHB-MIT database without model retraining. For CHB-MIT evaluation, the corresponding bipolar channels were Fp1–F7, T7–P7, Fp2–F8, and T8–P8. Under the standard 10–20 electrode nomenclature, T3, T4, T5, and T6 correspond to T7, T8, P7, and P8, respectively. Therefore, the bipolar derivations are anatomically equivalent across the two datasets, despite differences in electrode labeling. Nevertheless, the two datasets differ in acquisition protocols and patient populations [34, 35]. To account for potential differences in the classifier operating point, four randomly selected CHB-MIT recording cases were used exclusively for threshold calibration, while the remaining 20 cases were reserved for evaluation. To ensure participant-level independence, chb01 and chb21, which correspond to the same individual, were excluded from threshold calibration and retained exclusively in the evaluation set. This procedure was repeated five times using different calibration cases without retraining the TUSZ-trained classifier. Notably, the same operating threshold was selected in all five repetitions, consistent with the original calibration results. Furthermore, excluding these two cases from calibration did not change the reported evaluation performance. Thus, this experiment represents cross-dataset transfer with limited participant-independent threshold calibration rather than a strictly zero-shot evaluation.

The first cross-dataset experiment employed the complete feature set developed using the TUSZ. As summarized in **Table 8**, the model achieved an epoch-level sensitivity of 78.54%, specificity of 99.47%, and an AUC of 0.832. These tabulated values represent the mean performance across the five evaluation repetitions, with the corresponding standard errors reported in the table. In contrast, **Figure 6** provides a pooled epoch-level visualization across the CHB-MIT recording cases. For this visualization, the confusion-matrix counts were first averaged separately for each recording case across the repetitions in which it appeared in the held-out evaluation set. The resulting case-level matrices were then summed to construct the final pooled confusion matrix, preventing repeated counting while preserving the total number of eligible epochs. Notably, the identical threshold selected in all five repetitions ensured a consistent operating point across these evaluations. Consequently, the pooled confusion-matrix percentages may differ slightly from the mean performance values in **Table 8** because the two summaries use different aggregation procedures. Although the model maintained high specificity when transferred to CHB-MIT, the confusion matrix in **Figure 6** shows a comparatively large proportion of missed seizure epochs, indicating reduced seizure sensitivity under the cross-dataset setting.

To improve robustness across datasets, a second model was trained after modifying the feature representation. Specifically, the spike-count feature was replaced with a binary spike-presence indicator, and the composite features derived from training-set statistics were removed. With this reduced feature set, epoch-level sensitivity increased from 78.54% to 84.05% (mean paired increase: 5.51 percentage points; 95% CI: 4.64–6.38; $p_{\text{adj}} < 0.0001$), while the AUC improved from 0.832 to 0.882 (mean paired increase: 0.051; 95% CI: 0.043–0.058; $p_{\text{adj}} < 0.0001$). Although specificity decreased slightly from 99.47% to 99.37% ($p_{\text{adj}} < 0.0001$), it remained above 99%. The corresponding confusion matrix in **Figure 6** shows a clear reduction in false-negative seizure epochs compared with the full feature set. These results indicate improved transfer of the reduced representation between the two independent EEG datasets.

Unlike the segmented TUSZ recordings used for model development, the CHB-MIT database contains continuous EEG recordings, allowing temporal consistency among consecutive predictions to be incorporated as a post-processing step. Four temporal correction strategies were therefore evaluated using the predictions obtained with the reduced feature set. Both the reduced feature representation and the temporal correction strategies were examined based on their observed performance on CHB-MIT, without a separate independent dataset for configuration selection. Consequently, the reported improvements may be influenced by this evaluation-guided selection and should not be interpreted as independently validated performance gains. In each strategy, an

isolated non-seizure epoch occurring within a predicted seizure event was relabeled as a seizure epoch. Conversely, short sequences of predicted seizure epochs surrounded by non-seizure activity were progressively removed by correcting sequences containing up to one, two, three, or four consecutive epochs, corresponding to TCs #1–#4, respectively.

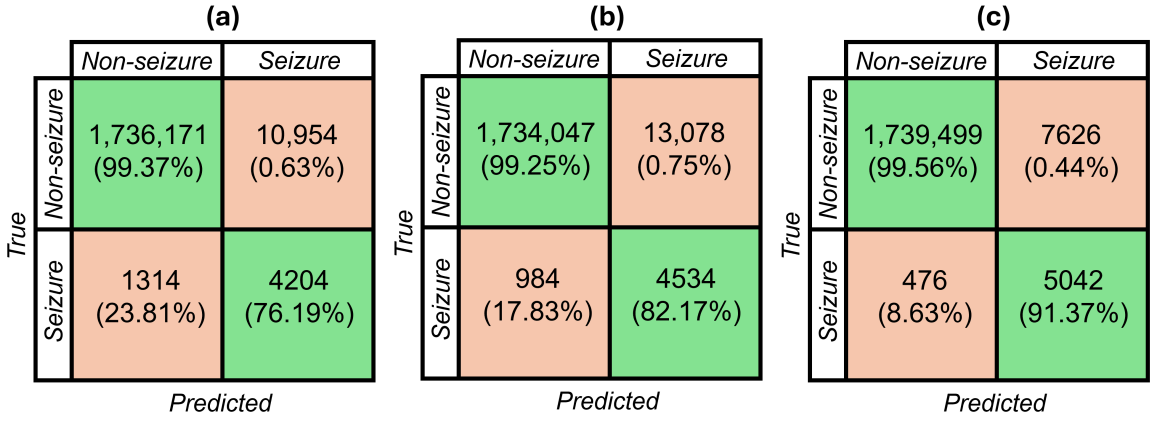

Figure 6. Pooled confusion matrices for the cross-dataset evaluation using TUSZ-trained models tested on the CHB-MIT. (**a**) Full feature set; (**b**) reduced feature set; (**c**) temporal correction (TC) #1. The full feature set comprises all 14 features, whereas the reduced set replaces the spike-count feature with a binary spike-presence indicator and removes the composite features, followed by model retraining on the TUSZ. TC #1 fills isolated non-seizure epochs within predicted seizure events and removes isolated single-epoch seizure predictions. Percentages are normalized within each true class. Green cells indicate correct classifications, whereas salmon-colored cells indicate misclassifications. For each CHB-MIT recording case, confusion-matrix counts were averaged across its held-out evaluation repetitions before summing across all 24 cases, avoiding repeated counting. Consequently, the pooled percentages may differ slightly from the mean performance values in **Table 8** due to different aggregation procedures. TUSZ: Temple University Hospital EEG Seizure Corpus; CHB-MIT: Children's Hospital Boston–Massachusetts Institute of Technology EEG Database.

Temporal post-processing substantially improved both epoch- and event-level performance. As shown in **Table 8**, TC #1 increased epoch-level sensitivity from 84.05% to 93.15% (mean paired increase: 9.10 percentage points; 95% CI: 8.09–10.12; $p_{\text{adj}} < 0.0001$) while maintaining 100% event sensitivity. Accuracy, specificity, and weighted F1 score also improved significantly (all $p_{\text{adj}} < 0.0001$), while AUC remained unchanged. The corresponding confusion matrix in **Figure 6** shows a marked reduction in missed seizure epochs compared with both the full and reduced feature representations. At the same time, the false alarm rate decreased from 6.835 to 1.801 false alarms/h, demonstrating that the temporal correction simultaneously improved seizure detection and suppressed short false-positive events.

Increasing the correction level further reduced the false alarm rate, reaching 0.572 false alarms/h at TC #4 with a corresponding specificity of 99.78%. This improvement was accompanied by a gradual reduction in epoch- and event-level sensitivity as progressively longer positive sequences were removed. Across all temporal correction levels, however, the average detection latency remained below 2.2 s. These latency values were calculated retrospectively from the timestamps of corrected predictions and do not represent real-time alarm availability. In particular, zero-phase forward–backward filtering requires future signal samples, while temporal correction may require subsequent predictions to determine whether short seizure-positive sequences should be retained or removed. Therefore, the reported latency may underestimate the delay of a causal implementation. These findings suggest that the compact propagation-based representation retains seizure-discriminative information when transferred from TUSZ to CHB-MIT, while feature refinement and temporal post-processing improve its performance under retrospective cross-dataset evaluation.

Among the evaluated configurations, the reduced feature representation with TC #2 was selected as the representative cross-dataset result because it provided a favorable balance between seizure sensitivity and false alarm reduction. Compared with TC #1, TC #2 reduced the false alarm rate from 1.801 to 1.223 false alarms/h (a 32.1% reduction), while epoch-level sensitivity decreased only slightly from 93.15% to 92.64% and event-level sensitivity from 100% to 99.49%. Nevertheless, a false alarm rate of 1.223 false alarms/h corresponds to approximately 29 false alarms per day if sustained during continuous monitoring, which remains substantial for practical clinical use. Although TC #3 and TC #4 achieved further reductions in false alarms, these improvements were accompanied by more substantial losses in seizure sensitivity, particularly at the event level. Therefore, TC #2 was considered a practical compromise between maintaining high seizure detection sensitivity and limiting false alarms. This configuration was selected based on the observed CHB-MIT evaluation results, without a separate independent dataset for selecting the temporal-correction level; accordingly, its reported performance should be interpreted as a representative operating point rather than an independently validated optimum. Furthermore, the standard deviations reported across five overlapping evaluation repetitions do not represent patient-level confidence intervals. Consequently, these results should be interpreted as retrospective performance estimates rather than evidence of clinical readiness or validated real-time wearable operation.

4. Discussion

4.1. Interpretation of the main findings

This study demonstrates that explicitly modeling the apparent spatiotemporal progression of seizure-related activity across scalp EEG electrodes enables accurate seizure detection while substantially reducing the number of required EEG electrodes. Unlike conventional feature-based approaches that rely primarily on the statistical, spectral, and morphological characteristics of individual channels, the proposed framework incorporates the spatiotemporal progression of seizure-related activity through propagation-based features. Across multiple electrode configurations, the framework consistently achieved high performance on the TUSZ and was subsequently transferred to the independent CHB-MIT database using a comparable four-channel bipolar montage with limited operating-threshold calibration. Collectively, these findings indicate that propagation-based features provide a compact and informative representation that remains effective across variations in electrode configuration and recording montage while also supporting transfer between independently acquired EEG datasets.

The electrode-configuration experiments highlight the importance of selecting a propagation model that is consistent with the underlying electrode geometry. As shown in **Figure 3**, applying the conventional 2D propagation model to the 12-electrode spatial configuration reduced detection performance, particularly sensitivity, whereas incorporating the 3D electrode coordinates restored performance to a level comparable with that obtained using the planar configurations. These observations suggest that a planar approximation can be effective for electrode arrangements with limited spatial variation but may become less suitable when electrodes distributed across different scalp regions are incorporated. Under such conditions, the 3D model preserves spatial relationships that would otherwise be lost through planar projection, allowing the estimated propagation vectors to more accurately represent the geometry of the recorded activity. However, because electrode coverage and geometric representation can vary simultaneously across configurations, the independent contribution of modeling dimensionality is examined further through the matched-electrode ablation study.

The importance of spatial modeling became particularly evident for focal seizures. As shown in **Figure 4**, consistently high performance across both planar and spatial configurations was maintained for generalized seizures, suggesting that their more widespread electrophysiological manifestation can be captured effectively

even using compact planar electrode arrangements. In contrast, for FNSZs, there was a stronger dependence on electrode configuration. Spatial configurations analyzed using the 3D propagation model consistently achieved higher sensitivities than corresponding planar configurations with comparable electrode counts. Since focal seizures originate within localized cortical regions and may subsequently propagate to neighboring or more distant regions [31, 45], their detection is more likely to depend on whether the available electrodes adequately capture the resulting spatiotemporal patterns at the scalp level. The observed improvement with spatial configurations therefore suggests that preserving 3D electrode relationships can provide additional information for characterizing the apparent progression of focal seizure-related activity. Importantly, these results also indicate that electrode count alone does not determine detection performance; the spatial distribution of the selected electrodes and the propagation model used to represent them are also important considerations.

The spike-presence ablation further demonstrates that detecting local spike-like activity alone is insufficient to reproduce the discriminative performance of the proposed framework. Replacing the complete feature representation with binary spike-presence indicators resulted in a substantial reduction in seizure sensitivity, approaching zero, despite maintaining high specificity. This reduction was observed for both GNSZs and FNSZs, with statistically significant differences in sensitivity, accuracy, and weighted F1 score compared with the full 3D model (all $p_{\text{adj}} < 0.0001$). One plausible explanation is that the spike-detection procedure identifies transient activity in both seizure and non-seizure epochs, making the mere presence of detected spikes insufficient for distinguishing the two classes. As illustrated in **Figure 2**, however, the detected spikes exhibit different spatiotemporal patterns between seizure and baseline activity. Without information describing these relationships, the classifier predominantly assigned epochs to the majority non-seizure class, resulting in very low seizure sensitivity. The additional ablation incorporating spike-presence and morphological features improved the representation of local spike characteristics but remained significantly inferior to the full 3D model in sensitivity, accuracy, and weighted F1 score for both seizure types (all $p_{\text{adj}} < 0.0001$). In contrast, the full model incorporates spike-arrival timing, apparent propagation direction and propagation speed, and spatial coherence, allowing the classifier to characterize differences in the organization of detected activity rather than relying solely on its presence. Together, these findings support the additional discriminative contribution of propagation-related features beyond spike presence and morphology, particularly for seizure sensitivity, without independently establishing that the estimated propagation characteristics represent physiological seizure propagation.

The matched-electrode ablation provides further evidence regarding the contribution of geometric representation independently of electrode selection. By evaluating the complete 2D and 3D propagation models using identical electrode sets, this experiment controls for differences in electrode count and spatial coverage while retaining the directionality and directional concentration features defined within each model's respective coordinate system. Therefore, the observed differences reflect the combined effects of spatial dimensionality and its associated propagation descriptors rather than coordinate dimensionality alone. For FNSZs, the 3D representation produced statistically significant improvements in sensitivity across all four spatial configurations, with mean gains of approximately 3.65–4.44 percentage points. These findings suggest that the complete 3D propagation representation may better characterize discriminative spatiotemporal patterns associated with the spatially localized and potentially heterogeneous scalp-level manifestations of focal seizures. In contrast, improvements for GNSZs were less consistent, with the clearest benefit observed for the 12-electrode configuration, while specificity did not exhibit a consistent significant improvement. This difference may reflect the more widespread manifestation of generalized seizure activity, for which a planar representation can retain sufficient discriminative information in many electrode arrangements. Collectively, these findings suggest that the relative benefit of the complete 3D propagation representation depends on seizure type and electrode configuration rather than representing a uniform improvement across all configurations.

Another important finding is that reliable seizure detection was maintained using only four bipolar channels. The selected montage retains bilateral frontal and temporal coverage while substantially reducing the

number of signals required for acquisition and processing. Such electrode reduction may simplify sensor placement, decrease preparation time, and support the development of lightweight ambulatory EEG systems, although these practical benefits require validation using actual wearable hardware. Despite this reduction, the proposed framework maintained high specificity and AUC across the GNSZ, FNSZ, and heterogeneous ANY seizure datasets. The comparatively lower sensitivity observed for FNSZs is also consistent with the electrode-configuration experiments, in which focal seizure detection showed greater dependence on spatial coverage. Nevertheless, the overall performance demonstrates that a carefully selected reduced montage can retain sufficient propagation information for reliable seizure detection. These findings support further investigation of the proposed framework for wearable EEG applications, particularly when frontal and temporal coverage is preserved, rather than establishing its performance under real-world wearable recording conditions.

The cross-dataset evaluation provides an additional assessment of the robustness of the proposed representation. Transferring a model developed using the TUSZ to the independent CHB-MIT database introduces substantial differences in patient populations, recording conditions, and available bipolar montages. Although four CHB-MIT recording cases per repetition were used exclusively to calibrate the operating threshold, the classifier itself was not retrained using CHB-MIT data, and the remaining 20 cases were reserved for held-out evaluation in each repetition. The complete TUSZ-derived feature representation retained high specificity but exhibited reduced seizure sensitivity, illustrating the difficulty of direct transfer between heterogeneous EEG datasets. Importantly, generalization improved when the spike-count feature was replaced by a binary spike-presence indicator and composite features dependent on training-set statistics were removed. The paired statistical analysis further confirmed significant improvements in seizure sensitivity and AUC with the reduced representation, although specificity decreased slightly while remaining above 99%. This suggests that features encoding the presence, morphology, and propagation of seizure-related activity may transfer more reliably than features whose numerical distributions are more strongly influenced by dataset-specific recording characteristics.

Temporal post-processing further improved the retrospective cross-dataset performance of the proposed framework. EEG seizure activity is temporally continuous, whereas epoch-wise classification treats consecutive windows independently and can therefore produce isolated false-negative or false-positive predictions. Incorporating this temporal consistency substantially increased epoch-level sensitivity and reduced the false alarm rate while maintaining 100% event sensitivity under TC #1. Stronger temporal correction further reduced false alarms, although at the cost of progressively lower seizure sensitivity. TC #2 provided a practical compromise, reducing the false alarm rate by approximately 32.1% relative to TC #1 (from 1.801 to 1.223 FAR/h), while maintaining 92.64% epoch-level sensitivity and 99.49% event-level sensitivity. Although TC #3 and TC #4 achieved further false-alarm reductions, their event-level sensitivities decreased to 93.27% and 91.10%, respectively. This trade-off is important for continuous monitoring applications, where excessive false alarms can substantially reduce the practical usefulness of an otherwise sensitive detector. Across all correction levels, the retrospectively estimated average detection latency remained below 2.2 s; however, these values do not account for the additional delays associated with non-causal filtering, and temporal correction and should not be interpreted as real-time alarm latency. The optimal degree of temporal correction may therefore depend on the intended application and the relative importance assigned to seizure sensitivity and false-alarm reduction. However, because both the reduced feature representation and TC #2 were selected after examining CHB-MIT results, the reported performance may be optimistic and requires validation on an additional independent dataset.

Our previous study established the feasibility of compact propagation- and morphology-based feature representation for seizure detection using referential EEG and within-dataset evaluation [33]. The present study extends that framework in several important directions. Specifically, it demonstrates how propagation modeling can be adapted to both 2D and 3D electrode geometries, examines the effect of electrode configuration on generalized and focal seizure detection, evaluates the framework using a wearable-oriented four-channel bipolar

montage, and assesses its transfer to an independent EEG database. Furthermore, the ablation studies establish the discriminative contribution of propagation-related information beyond binary spike presence and quantify the additional benefits of 3D geometric modeling under matched electrode configurations. The cross-dataset experiments further examine feature transferability and the sensitivity–false-alarm trade-off introduced by temporal correction. Together, these extensions move the framework beyond the original within-dataset evaluation and provide further evidence that explicitly representing the apparent spatiotemporal progression of seizure-related scalp EEG activity may be useful for reduced-channel and future wearable-oriented EEG monitoring.

4.2. Comparison with existing seizure detection approaches

Table 9 summarizes representative seizure detection approaches, including lightweight deep learning models [21, 46], reduced-channel systems [23], cross-dataset domain generalization [47], and emerging montage-flexible EEG representations [20]. However, direct performance comparisons are limited by differences in datasets, electrode configurations, preprocessing, epoch durations, and evaluation protocols. In particular, our TUSZ experiments employed leave-one-subject-out cross-validation, while the cross-dataset evaluation involved training on the TUSZ and testing on the CHB-MIT with limited threshold calibration. Thus, the comparison provides methodological context rather than a direct performance ranking.

Table 9. Comparison of representative seizure detection approaches.

Study/model	Dataset/evaluation	Acc. (%)	Spe. (%)	Epoch sen. (%)	Event sen. (%)	FAR/h	Latency (s)
Ayodele et al. (2020) [47]	CHB-MIT + TUSZ → external evaluation	-	-	-	74.50	0.840	2.320
LightSeizureNet (2023) [21]	CHB-MIT; patient-independent	97.09	-	-	-	-	-
Chung et al. (2024) [23]	CHB-MIT; patient-specific, single-channel	98.18	98.19	96.76	99.62	0.220	3.300
PSD-LW-DCN (2026) [46]	CHB-MIT; cross-subject	85.84	92.12	79.57	-	0.330	-
PSD-LW-DCN (2026) [46]	TUSZ; cross-subject	83.21	88.39	78.02	-	1.090	-
Proposed model (TC #2)	TUSZ → CHB-MIT; calibrated	99.63	99.67	92.64	99.49	1.223	1.299

Performance values are reported as provided in the original studies. The symbol → indicates training on the dataset to its left and evaluation on the dataset to its right; the symbol + indicates that both datasets were used in the study. A hyphen (-) indicates that the corresponding value was not reported in the original study. The sensitivity reported by Ayodele et al. [47] represents the combined evaluation result rather than an exclusively external-cohort estimate. The proposed model was trained on the TUSZ without classifier retraining on the CHB-MIT; limited operating-threshold calibration was performed using CHB-MIT recording cases. Results are not directly comparable due to differences in datasets, subject selection, preprocessing, electrode configurations, and evaluation protocols. Acc.: accuracy; Spe.: specificity; Sen.: sensitivity; FAR/h: false alarms per hour; TUSZ: Temple University Hospital EEG Seizure Corpus; CHB-MIT: Children's Hospital Boston–Massachusetts Institute of Technology EEG Database; TC: temporal correction; EEG: electroencephalography; PSD-LW-DCN: Power Spectral Density–based Lightweight Deep Convolutional Neural Network.

The proposed framework achieved 99.63% accuracy, 99.67% specificity, 92.64% epoch-level sensitivity, and 99.49% event-level sensitivity under TC #2, with 1.223 false alarms per hour and an average detection latency of 1.299 s. These results demonstrate promising cross-dataset transfer using four bipolar EEG channels without

retraining on the CHB-MIT. Unlike approaches focused primarily on classification performance, this study also investigates propagation-derived features, spatial modeling, and cross-dataset generalization. Nevertheless, the false-alarm burden remains substantial, and computational efficiency and wearable feasibility require direct hardware benchmarking and prospective validation.

4.3. Limitations and future work

Several limitations should be acknowledged. First, although the framework was evaluated using two independent publicly available scalp EEG datasets, retrospective database evaluation cannot fully reproduce the conditions encountered during prospective ambulatory or wearable monitoring. In particular, recordings containing severe artifacts, corrupted intervals, or insufficient usable EEG data were excluded from the analysis. Consequently, the present evaluation does not establish robustness against the physiological and non-physiological artifacts commonly encountered in wearable recordings, including motion-related disturbances, electrode displacement, and variations in electrode–skin contact. Furthermore, no actual wearable hardware or ambulatory recordings were evaluated. The reduced-channel configuration should therefore be regarded as a wearable-oriented methodological design rather than a demonstrated wearable implementation. Further validation using prospectively acquired long-term EEG recordings is therefore required before clinical deployment. Future work should specifically investigate artifact robustness under realistic ambulatory conditions, including the effects of motion, signal degradation, and intermittent electrode contact.

Second, the TUSZ and CHB-MIT databases differ substantially in patient populations and recording protocols, and four CHB-MIT recording cases per repetition were required for operating-threshold calibration. Future work should investigate calibration-independent or adaptive thresholding strategies to further improve generalization across previously unseen datasets. Although temporal correction substantially reduced false alarms, the observed rates of approximately 1–7 false alarms per hour across the evaluated configurations may remain excessive for prolonged patient-facing monitoring. Even the representative TC #2 configuration produced approximately 1.223 false alarms per hour, corresponding to roughly 29 false alarms over 24 h if the observed rate were sustained. Therefore, the reported improvements should not be interpreted as establishing clinically acceptable false-alarm performance. Further reductions through improved artifact rejection, adaptive decision thresholds, and longer-term temporal modeling are necessary, together with prospective evaluation of the sensitivity–false-alarm trade-off under realistic monitoring conditions.

Third, the fixed four-channel bipolar montage provides broad bilateral frontal and temporal coverage but may not optimally capture seizures originating outside these regions, which is particularly relevant for focal seizure detection. Patient-specific or adaptive electrode placement may therefore provide additional benefits. Additionally, the decision threshold in each TUSZ LOSO fold was selected using in-sample training predictions rather than an independent validation subset. Although the held-out subject remained excluded from this procedure, threshold selection may still introduce optimistic bias. Future work should investigate nested subject-wise validation or independent calibration subsets to provide less biased estimates of threshold-dependent performance. Finally, the present study evaluates the proposed framework retrospectively rather than on embedded wearable hardware. Inference latency, memory requirements, computational load, and energy consumption were not measured on a target device, and no prospective patient-facing evaluation was conducted. Accordingly, the present findings establish the potential of the proposed reduced-channel representation but do not demonstrate its readiness for real-time wearable deployment. Future work will investigate real-time implementation, computational and power requirements, and prospective evaluation of the framework using wearable EEG systems.

5. Conclusions

This study extended propagation-based seizure detection toward spatially informed and wearable-oriented EEG monitoring. Across multiple electrode configurations on the TUSZ, the proposed framework maintained high detection performance while demonstrating the importance of matching the propagation model to the underlying electrode geometry. In particular, the complete 3D propagation representation provided a benefit for focal seizure detection, highlighting the potential value of incorporating 3D electrode geometry and its associated propagation descriptors. The ablation studies further demonstrated that propagation-related features provide substantial discriminative information beyond binary spike presence and morphology, while the complete 3D representation significantly improved focal seizure sensitivity under matched-electrode configurations. Importantly, the framework also demonstrated promising seizure detection performance using a wearable-oriented four-channel bipolar montage with bilateral frontal and temporal coverage, supporting the feasibility of reduced-channel EEG representations. The TUSZ-trained classifier was further evaluated retrospectively on the independent CHB-MIT database with limited operating-threshold calibration, demonstrating the potential for cross-dataset transfer while highlighting the need for further independent validation. Overall, these findings support propagation-based EEG representations as a promising approach to developing compact, reduced-channel seizure monitoring systems and motivate future prospective evaluation using wearable EEG hardware.

Acknowledgments

The authors acknowledge the use of ChatGPT (GPT-5.6 Sol, OpenAI) for language editing. The tool was applied to enhance clarity and scholarly tone, and all edits were carefully reviewed and validated by the authors to ensure accuracy, compliance with academic standards, and research integrity. The authors fully support Academia.edu Journals' adherence to COPE guidelines on AI in publication ethics and confirm that this use has been managed responsibly and ethically.

Funding

This research received no external funding.

Author contributions

Conceptualization, I.T.L.Jr. and D.W.; methodology, D.W. and I.T.L.Jr.; software, D.W.; validation, D.W.; formal analysis, D.W.; investigation, D.W.; writing—original draft preparation, D.W.; writing—review and editing, I.T.L.Jr.; project administration, I.T.L.Jr. All authors have read and agreed to the published version of the manuscript.

Conflict of interest

The authors declare that they have no competing interests.

Data availability statement

The electroencephalography (EEG) data supporting the findings of this publication are available through two publicly accessible repositories. The Temple University Hospital EEG Seizure Corpus (TUSZ) is available at https:

//isip.piconepress.com/projects/nedc/html/tuh_eeg/. The Children's Hospital Boston–Massachusetts Institute of Technology (CHB-MIT) Scalp EEG Database (version 1.0.0) is available through PhysioNet at https://physionet.org/content/chbmit/1.0.0/ (doi: 10.13026/C2K01R). Both datasets were accessed and analyzed in accordance with their respective data usage conditions. The source code used in this study can be made available upon request.

Institutional review board statement

The study was conducted according to the guidelines of the Declaration of Helsinki and presented to the Institutional Review Board of North Dakota State University. The study was declared exempt from full board review (Category 4, 45 CFR 46.104(d)(4)) due to its being a secondary analysis of previously collected, publicly available, de-identified EEG recordings, with no recruitment of or interaction with participants and no access to identifiable private information (protocol code IRB0006055). The original datasets were collected under approval from the Temple University Institutional Review Board (IRB No. 20774) for the TUH EEG Corpus, and at Children's Hospital Boston for the CHB-MIT Scalp EEG Database.

Informed consent statement

Informed consent for participation in the study is not required as per local legislation 45 CFR 46.104(d)(4).

Publisher's note

Academia.edu Journals stays neutral regarding jurisdictional claims in published maps and institutional affiliations. All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright

© 2026 copyright by the authors. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license.

References

1. Stafstrom CE, Carmant L. Seizures and epilepsy: an overview for neuroscientists. Cold Spring Harb Perspect Med. 2015; 5 (6): a022426. doi: 10.1101/cshperspect.a022426
2. World Health Organization. Epilepsy. Geneva, Switzerland: World Health Organization; 2024. Available from: https://www.who.int/news-room/fact-sheets/detail/epilepsy
3. Elger CE, Hoppe C. Diagnostic challenges in epilepsy: seizure under-reporting and seizure detection. Lancet Neurol. 2018; 17 (3): 279–88. doi: 10.1016/S1474-4422(18)30038-3
4. Giannakakis G, Sakkalis V, Pediaditis M, Tsiknakis M. Methods for seizure detection and prediction: an overview. In: Sakkalis V, editor. Modern electroencephalographic assessment techniques. New York, NY, USA: Humana Press; 2014. p. 131–57. doi: 10.1007/7657_2014_68

5. Ein Shoka AA, Dessouky MM, El-Sayed A, Hemdan EE. EEG seizure detection: concepts, techniques, challenges, and future trends. Multimed Tools Appl. 2023; 82 (27): 42021–51. doi: 10.1007/s11042-023-15052-2
6. Jerger KK, Netoff TI, Francis JT, Sauer T, Pecora L, Weinstein SL, et al. Early seizure detection. J Clin Neurophysiol. 2001; 18 (3): 259–68. doi: 10.1097/00004691-200105000-00005
7. Alotaiby TN, Alshebeili SA, Alshawi T, Ahmad I, Abd El-Samie FE. EEG seizure detection and prediction algorithms: a survey. EURASIP J Adv Signal Process. 2014; 2014 (1): 183. doi: 10.1186/1687-6180-2014-183
8. Dash DP, Kolekar M, Chakraborty C, Khosravi MR. Review of machine and deep learning techniques in epileptic seizure detection using physiological signals and sentiment analysis. ACM Trans Asian Low-Resour Lang Inf Process. 2024; 23 (1): 16. doi: 10.1145/3552512
9. Alkan A, Koklukaya E, Subasi A. Automatic seizure detection in EEG using logistic regression and artificial neural network. J Neurosci Methods. 2005; 148 (2): 167–76. doi: 10.1016/j.jneumeth.2005.04.009
10. Jana GC, Praneeth MS, Agrawal A. A multi-view SVM approach for seizure detection from single channel EEG signals. IETE J Res. 2023; 69 (6): 3120–31. doi: 10.1080/03772063.2021.1913074
11. Chanu MM, Singh NH, Thongam K. An automated epileptic seizure detection using optimized neural network from EEG signals. Expert Syst. 2023; 40 (6): e13260. doi: 10.1111/exsy.13260
12. Boonyakitanont P, Lek-uthai A, Chomtho K, Songsiri J. A review of feature extraction and performance evaluation in epileptic seizure detection using EEG. Biomed Signal Process Control. 2020; 57: 101702. doi: 10.1016/j.bspc.2019.101702
13. Aayesha, Qureshi MB, Afzaal M, Qureshi MS, Fayaz M. Machine learning-based EEG signals classification model for epileptic seizure detection. Multimed Tools Appl. 2021; 80 (12): 17849–77. doi: 10.1007/s11042-021-10597-6
14. Fergus P, Hussain A, Hignett D, Al-Jumeily D, Abdel-Aziz K, Hamdan H. A machine learning system for automated whole-brain seizure detection. Appl Comput Inform. 2016; 12 (1): 70–89. doi: 10.1016/j.aci.2015.01.001
15. Zhang X, Yao L, Dong M, Liu Z, Zhang Y, Li Y. Adversarial representation learning for robust patient-independent epileptic seizure detection. IEEE J Biomed Health Inform. 2020; 24 (10): 2852–9. doi: 10.1109/JBHI.2020.2971610
16. Einizade A, Nasiri S, Mozafari M, Sardouie SH, Clifford GD. Explainable automated seizure detection using attentive deep multi-view networks. Biomed Signal Process Control. 2023; 79: 104076. doi: 10.1016/j.bspc.2022.104076
17. Ahmad I, Wang X, Javeed D, Kumar P, Samuel OW, Chen S. A hybrid deep learning approach for epileptic seizure detection in EEG signals. IEEE J Biomed Health Inform. 2026; 30 (2): 1019–29. doi: 10.1109/JBHI.2023.3265983
18. Li C, Li H, Dong X, Zhong X, Cui H, Ji D, et al. CNN-Informer: a hybrid deep learning model for seizure detection on long-term EEG. Neural Netw. 2025; 181: 106855. doi: 10.1016/j.neunet.2024.106855
19. Sun J, Xiang J, Dong Y, Wang B, Zhou M, Ma J, et al. Deep learning for epileptic seizure detection using a causal-spatio-temporal model based on transfer entropy. Entropy. 2024; 26 (10): 853. doi: 10.3390/e26100853
20. Sukhbaatar J, Imamura S, Inoue I, Murakami S, Hassan KM, Han S, et al. SingLEM: single-channel large EEG model. IEEE J Biomed Health Inform. 2026; 1–14. doi: 10.1109/JBHI.2026.3732223
21. Qiu S, Wang W, Jiao H. LightSeizureNet: a lightweight deep learning model for real-time epileptic seizure detection. IEEE J Biomed Health Inform. 2023; 27 (4): 1845–56. doi: 10.1109/JBHI.2022.3223970
22. Padmaja B, Maram B, Mammadova K, Makharov K, Qasim MT, Khishe M. A lightweight CNN–transformer framework for EEG seizure detection with ablation, interpretability, and noise-robustness analysis. Intell-Based Med. 2026; 15: 100446. doi: 10.1016/j.ibmed.2026.100446

23. Chung YG, Cho A, Kim H, Kim KJ. Single-channel seizure detection with clinical confirmation of seizure locations using CHB-MIT dataset. Front Neurol. 2024; 15: 1389731. doi: 10.3389/fneur.2024.1389731
24. Moctezuma LA, Molinas M. EEG channel-selection method for epileptic-seizure classification based on multi-objective optimization. Front Neurosci. 2020; 14: 593. doi: 10.3389/fnins.2020.00593
25. Refoyo M, Cano-Broncano F, Luengo D. Information-theoretical channel selection to enable efficient single-channel EEG-based epileptic seizure detection in ambulatory monitoring. Sci Rep. 2026. doi: 10.1038/s41598-026-69842-3
26. Romaine JB, Pereira Martín M, Salvador Ortiz JR, Manzano Crespo JM. EEG—single-channel envelope synchronisation and classification for seizure detection and prediction. Brain Sci. 2021; 11 (4): 516. doi: 10.3390/brainsci11040516
27. Karoly PJ, Freestone DR, Boston R, Grayden DB, Himes D, Leyde K, et al. Interictal spikes and epileptic seizures: their relationship and underlying rhythmicity. Brain. 2016; 139 (4): 1066–78. doi: 10.1093/brain/aww019
28. Matarrese MAG, Loppini A, Fabbri L, Tamilia E, Perry MS, Madsen JR, et al. Spike propagation mapping reveals effective connectivity and predicts surgical outcome in epilepsy. Brain. 2023; 146 (9): 3898–912. doi: 10.1093/brain/awad118
29. Smith EH, Liou JY, Merricks EM, Davis T, Thomson K, Greger B, et al. Human interictal epileptiform discharges are bidirectional traveling waves echoing ictal discharges. eLife. 2022; 11: e73541. doi: 10.7554/eLife.73541
30. Sun J, Niu Y, Wang C, Dong Y, Wang B, Wei J, et al. Exploring the propagation pathway in individual patients with epilepsy: a stepwise effective connection approach. Biomed Signal Process Control. 2024; 90: 105811. doi: 10.1016/j.bspc.2023.105811
31. Karimi-Rouzbahani H, McGonigal A. Directionality of neural activity in and out of the seizure onset zone in focal epilepsy. Netw Neurosci. 2025; 9 (2): 798–823. doi: 10.1162/netn_a_00454
32. Saadoon YA, Khalil M, Battikh D. Machine and deep learning-based seizure prediction: a scoping review on the use of temporal and spectral features. Appl Sci. 2025; 15 (11): 6279. doi: 10.3390/app15116279
33. Wijesinghe D, Lima IT Jr. Compact propagation and morphology-based EEG features for real-time seizure detection using machine learning. J Interdiscip Res Appl Med. 2026; 6 (2): 8. doi: 10.3390/jdream6020008
34. Shah V, von Weltin E, Lopez S, McHugh JR, Veloso L, Golmohammadi M, et al. The Temple University Hospital seizure detection corpus. Front Neuroinform. 2018; 12: 83. doi: 10.3389/fninf.2018.00083
35. Guttag J. CHB-MIT scalp EEG database. Version 1.0.0. PhysioNet; 2010. Available from: https://physionet.org/content/chbmit/1.0.0/
36. Kayser J, Tenke CE. On the benefits of using surface Laplacian (current source density) methodology in electrophysiology. Int J Psychophysiol. 2015; 97 (3): 171–3. doi: 10.1016/j.ijpsycho.2015.06.001
37. Leys C, Ley C, Klein O, Bernard P, Licata L. Detecting outliers: do not use standard deviation around the mean, use absolute deviation around the median. J Exp Soc Psychol. 2013; 49 (4): 764–6. doi: 10.1016/j.jesp.2013.03.013
38. Knapp C, Carter G. The generalized correlation method for estimation of time delay. IEEE Trans Acoust Speech Signal Process. 1976; 24 (4): 320–7. doi: 10.1109/TASSP.1976.1162830
39. Mardia KV, Jupp PE. Directional statistics. Chichester, UK: John Wiley & Sons; 1999. doi: 10.1002/9780470316979
40. Sra S. A short note on parameter approximation for von Mises-Fisher distributions: and a fast implementation of $I_s(x)$. Comput Stat. 2012; 27 (1): 177–90. doi: 10.1007/s00180-011-0232-x
41. Logesparan L, Casson AJ, Rodriguez-Villegas E. Optimal features for online seizure detection. Med Biol Eng Comput. 2012; 50 (7): 659–69. doi: 10.1007/s11517-012-0904-x

42. Stam CJ, Nolte G, Daffertshofer A. Phase lag index: assessment of functional connectivity from multi channel EEG and MEG with diminished bias from common sources. Hum Brain Mapp. 2007; 28 (11): 1178–93. doi: 10.1002/hbm.20346
43. Tian Z, Hu B, Si Y, Wang Q. Automatic seizure detection and prediction based on brain connectivity features and a CNNs meet Transformers classifier. Brain Sci. 2023; 13 (5): 820. doi: 10.3390/brainsci13050820
44. Chen T, Guestrin C. XGBoost: a scalable tree boosting system. In: Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; 13–17 August 2016; San Francisco, CA, USA. New York, NY, USA: Association for Computing Machinery; 2016. p. 785–94. doi: 10.1145/2939672.2939785
45. Ghulaxe Y, Joshi A, Chavada J, Huse S, Kalbande B, Sarda PP. Understanding focal seizures in adults: a comprehensive review. Cureus. 2023; 15 (11): e48173. doi: 10.7759/cureus.48173
46. Gu P, Zhang M, Xu M, Chen D, Shou J, Guo J, et al. PSD-LW-DCN: a generalizable power spectral density based lightweight deep convolutional neural network for seizure detection. Sci Rep. 2026; 16 (1): 14073. doi: 10.1038/s41598-026-44536-y
47. Ayodele KP, Ikezogwo WO, Komolafe MA, Ogunbona P. Supervised domain generalization for integration of disparate scalp EEG datasets for automatic epileptic seizure detection. Comput Biol Med. 2020; 120: 103757. doi: 10.1016/j.compbiomed.2020.103757