

Article

Not peer-reviewed version

---

# Topological Data Analysis Driven fNIRS Signal Processing for Alzheimer's Disease Stage Identification

---

[Siyuan Liu](#) , Hangcheng Wu , Cheng Sun , Yuanbin Qiu , Haoliang Wu , Yucong Wei , [Yang Lv](#) , [Zheng Yang](#) \*

Posted Date: 5 May 2026

doi: 10.20944/preprints202605.0121.v1

Keywords: Topological Data Analysis (TDA); persistent homology; persistence images; fNIRS; feature extraction; Alzheimer's disease; granger causality; brain networks



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC, OpenAlex.

Copyright: This open access article is published under a [Creative Commons CC BY 4.0 license](#), which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

Disclaimer/Publisher's Note: The statements, opinions, and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions, or products referred to in the content.

Article

# Topological Data Analysis Driven fNIRS Signal Processing for Alzheimer's Disease Stage Identification

Siyuan Liu <sup>1</sup>, Hangcheng Wu <sup>1</sup>, Cheng Sun <sup>1</sup>, Yuanbin Qiu <sup>1</sup>, Haoliang Wu <sup>2</sup>, Yucong Wei <sup>1</sup>, Yang Lv <sup>3</sup> and Zheng Yang <sup>1,\*</sup>

<sup>1</sup> Pittsburgh Institute, Sichuan University, Chengdu 610207, China

<sup>2</sup> West China School of Clinical Medicine, Sichuan University, Chengdu 610041, China

<sup>3</sup> Department of Geriatrics, The First Affiliated Hospital of Chongqing Medical University, Chongqing 400016, China

\* Correspondence: zhengyang2018@scu.edu.cn

## Abstract

This paper proposes a novel Topological Data Analysis (TDA) pipeline to extract robust structural features from functional near-infrared spectroscopy (fNIRS) signals for the classification of Alzheimer's Disease (AD) stages. Alzheimer's disease is increasingly understood as a disconnection syndrome, where the disruption of functional brain networks precedes gross anatomical atrophy. However, traditional graph-theoretic approaches rely on arbitrary connectivity thresholds, which can obscure critical multi-scale topological information and are sensitive to noise. To address this, our framework leverages Persistent Homology (PH) to analyse the topological evolution of brain networks across a continuous range of scales. By modeling 48-channel hemoglobin concentration time-series as high-dimensional point clouds via Granger causality metrics, we construct filtration sequences of Vietoris-Rips complexes. The resulting topological invariants—specifically 0-dimensional connected components, 1-dimensional loops, and 2-dimensional voids—are captured in Persistence Diagrams and subsequently vectorized into Persistence Images (PIs) using Gaussian kernel smoothing. This transformation enables the integration of complex topological features into standard machine learning workflows. Our experimental results on 284 recordings demonstrate that this topology-driven feature extraction method yields high discriminative power, achieving 77% accuracy in multi-class diagnosis (NC vs. MCI vs. AD). This study validates the efficacy of TDA as a sophisticated signal processing tool for revealing intrinsic neurodegenerative patterns in hemodynamic data, offering a potential non-invasive biomarker for early detection.

**Keywords:** Topological Data Analysis (TDA); persistent homology; persistence images; fNIRS; feature extraction; Alzheimer's disease; granger causality; brain networks

## 1. Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative disorder and the leading cause of dementia worldwide. As the global population ages, the prevalence of AD is projected to rise dramatically, posing a significant burden on healthcare systems [1]. The disease pathology is characterized by the accumulation of  $\beta$ -amyloid plaques and neurofibrillary tangles, which lead to synaptic loss and neuronal death [2]. Clinically, the disease progresses through a continuum: from Normal Cognition (NC) to Mild Cognitive Impairment (MCI)—a prodromal stage where intervention is most effective—and finally to AD dementia [3,4]. Consequently, developing accurate, non-invasive, and cost-effective tools for early-stage classification, particularly distinguishing MCI from NC, is a critical research priority.

While neuroimaging modalities such as Magnetic Resonance Imaging (MRI) and Positron Emission Tomography (PET) are the gold standards for diagnosis [5,6], they are often expensive, immobile, and contraindicated for patients with metallic implants or claustrophobia. Functional Near-Infrared Spectroscopy (fNIRS) has emerged as a promising alternative. By monitoring hemodynamic responses (oxy-hemoglobin HbO and deoxy-hemoglobin HbR) in the cortex, fNIRS provides a measure of neurovascular coupling similar to fMRI but with higher temporal resolution, lower cost, and greater portability [7–9].

However, analyzing fNIRS data for AD classification presents unique signal processing challenges. The "disconnection hypothesis" suggests that cognitive decline in AD arises from the disruption of functional integration between distributed brain regions rather than damage to a single area. Traditional network analysis typically involves constructing a functional connectivity graph by thresholding a correlation matrix. This approach suffers from two major limitations:

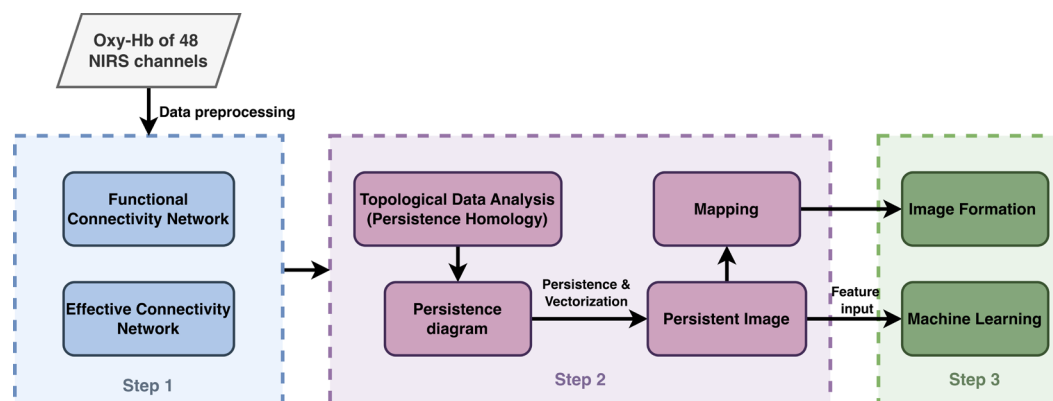
- 1) **Threshold Arbitrariness:** Choosing a threshold to binarize the network is subjective. A high threshold may fracture the network, losing weak but biologically significant connections, while a low threshold may introduce noise.
- 2) **Scale Sensitivity:** Brain networks exhibit multiscale organization. Fixed-scale graph metrics (e.g., path length, clustering coefficient) fail to capture the hierarchical topology of neural degeneration.

To overcome these limitations, we propose a Topological Data Analysis (TDA) framework based on Persistent Homology (PH). TDA treats data as a shape and analyzes its connectivity across all possible thresholds simultaneously (a process called filtration) [10,11]. This renders the analysis robust to noise and independent of arbitrary parameter choices. While TDA has been applied to MRI and PET data [12,13], its application to fNIRS-derived effective connectivity networks remains underexplored.

In this work, we develop a comprehensive pipeline that:

- Transforms raw fNIRS time-series into **Effective Connectivity** networks using Granger Causality, capturing the directionality of information flow.
- Extracts topological invariants (Betti numbers) via **Persistent Homology**, identifying stable structures (clusters, loops and higher-order voids) that persist across scales.
- Vectorizes these abstract topological features into **Persistence Images (PIs)**, creating a stable, fixed-dimensional input for machine learning classifiers.

We demonstrate that this topology-aware machine learning approach significantly outperforms baseline methods, providing a robust framework for automated AD staging. The workflow is illustrated in Figure 1.



**Figure 1.** Proposed TDA-based data analysis pipeline: From fNIRS signal acquisition to topological feature vectorization and classification. The pipeline emphasizes the transformation from temporal signal space to topological feature space.

## 2. Literature Review

### 2.1. Machine Learning in AD Diagnosis

The application of machine learning to neuroimaging has been extensive. Early approaches relied on hand-crafted features fed into Support Vector Machines (SVMs) [14]. More recently, deep learning, particularly Convolutional Neural Networks (CNNs), has achieved state-of-the-art results on structural MRI [15–17]. For instance, Suk et al. [18] utilized Deep Boltzmann Machines for multimodal fusion, while Liu et al. [19] explored manifold learning. Despite these successes, deep learning models often function as "black boxes," lacking interpretability regarding the underlying network failure mechanisms. Furthermore, their performance on early-stage MCI detection often plateaus between 60% and 80% [6,20], suggesting that standard morphological or intensity-based features may miss subtle global connectivity changes.

### 2.2. fNIRS in Cognitive Neuroscience

fNIRS has gained traction for studying neurodegenerative diseases due to its sensitivity to neurovascular coupling changes. Research indicates that AD patients exhibit reduced hemodynamic activation in the prefrontal and parietal cortices during cognitive tasks [8]. However, most fNIRS studies rely on localized activation analysis or simple functional connectivity metrics (Pearson correlation). These methods often overlook the directional and causal relationships between brain regions. Effective connectivity, assessed via Granger Causality, provides a more rigorous description of neural communication [21], yet its combination with advanced topology remains rare in fNIRS literature.

### 2.3. Topological Data Analysis (TDA) in Medicine

TDA has recently emerged as a powerful tool for quantifying the "shape" of complex biological data. In the context of AD, Mueller et al. [10] demonstrated that TDA could distinguish the branching structures of the hippocampus in AD versus MCI patients. On a network level, Hampel et al. [7] and Bai et al. [4] used TDA to map the disintegration of the cholinergic system. The core advantage of TDA is its ability to separate signal from noise based on the "persistence" of topological features. Unlike graph theory, which captures local properties, TDA captures global properties (holes, voids, connectivity) that are invariant under continuous deformation. This study aims to bridge the gap between fNIRS signal processing and algebraic topology.

## 3. Materials & Methodology

### 3.1. Participants and Experimental Protocol

Data were collected from a cohort of 284 participants, categorized into three groups: Normal Cognition (NC), Mild Cognitive Impairment (MCI), and Alzheimer's Disease (AD). All participants provided informed consent. The study utilized the Clock Drawing Test (CDT) as the cognitive activation paradigm, which is a validated screening tool for executive function and visuospatial abilities.

Experiments were conducted in a sound-attenuated, dimly lit room to minimize environmental distractions.

#### 3.1.1. Dataset Balancing via Augmentation

The original dataset ( $N = 284$ ) was imbalanced, with fewer MCI cases compared to AD and NC. Class imbalance can lead to biased classifiers that favor the majority class. To mitigate this, we applied a data augmentation technique specifically designed for time-series data.

We generated synthetic samples for the minority classes by injecting Gaussian noise into the raw fNIRS time series:

$$X_{\text{aug}}(t) = X_{\text{raw}}(t) + \mathcal{N}(0, \sigma_{\text{noise}}^2) \quad (1)$$

where  $\sigma_{\text{noise}}$  was set to 0.05 times the standard deviation of the original signal. This preserves the underlying temporal structure and topological properties while introducing sufficient variation. After augmentation, the total dataset size was expanded to 520 samples, ensuring a more balanced distribution for training.

### 3.1.2. Hyperparameter Optimization

To maximize model performance, we performed a Grid Search with 5-fold cross-validation on the training set. The search space included:

Random Forest: Number of trees  $\in [50, 100, 200]$ , Max depth  $\in [None, 10, 20]$ .

XGBoost: Learning rate  $\in [0.01, 0.1, 0.2]$ , Gamma  $\in [0, 0.1, 0.2]$ .

PI Resolution: Grid sizes of  $3 \times 3$ ,  $5 \times 5$ , and  $10 \times 10$  were tested. The  $3 \times 3$  grid provided the best trade-off between information density and overfitting.

### 3.1.3. Evaluation Metrics

We evaluated the classification performance using standard metrics derived from the confusion matrix elements (True Positives  $TP$ , False Positives  $FP$ , False Negatives  $FN$ ):

Accuracy: The ratio of correctly predicted observations to the total observations.

Precision: The ratio of correctly predicted positive observations to the total predicted positives:

$$\text{Precision} = \frac{TP}{TP+FP}$$

Recall (Sensitivity): The ratio of correctly predicted positive observations to all observations in actual class:  $\text{Recall} = \frac{TP}{TP+FN}$ .

F1-Score: The weighted average of Precision and Recall:

$$F1 = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \quad (2)$$

Given the multi-class nature of the problem, we report the macro-average for these metrics.

## 3.2. fNIRS Data Acquisition and Detailed Preprocessing

Hemodynamic signals were recorded using a multi-channel fNIRS system (NirScan, Danyang Huichuang Medical Equipment Co., Ltd, China). The system utilized two wavelengths of near-infrared light, 760 nm and 850 nm, with a sampling rate of 11 Hz. The probe geometry consisted of 15 sources and 16 detectors, arranged to create 48 measurement channels with a source-detector separation of 3 cm. The optodes were positioned according to the international 10-20 system, covering the bilateral prefrontal cortex (PFC), which is heavily implicated in executive dysfunction in AD.

The raw light intensity data contains significant physiological noise (e.g., Mayer waves, respiration, heart rate) and instrumental noise. Our preprocessing pipeline was rigorously designed as follows:

#### 1) Optical Density Conversion

The raw intensity  $I(t)$  was converted to optical density change relative to a baseline reference  $I_0$ :

$$\Delta OD(t) = -\log_{10} \left( \frac{I(t)}{I_0} \right) \quad (3)$$

#### 2) Motion Artifact Correction (Moving Standard Deviation):

Motion artifacts are often characterized by sudden spikes or baseline shifts. We employed a Moving Standard Deviation (MSD) method. For a window size  $w$ , the local standard deviation  $\sigma(t)$



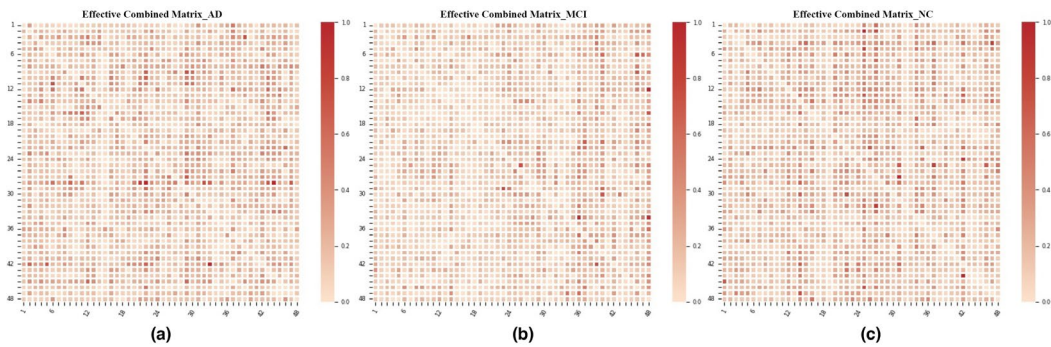
The Unrestricted Model (incorporating history of both variables):

$$X_1(t) = \sum_{j=1}^p A_{11,j} X_1(t-j) + \sum_{j=1}^p A_{12,j} X_2(t-j) + E_{1U}(t) \quad (5)$$

The Restricted Model (incorporating history of  $X_1$  only):

$$X_1(t) = \sum_{j=1}^p A'_{11,j} X_1(t-j) + E_{1R}(t) \quad (6)$$

Here,  $p$  is the model order, selected via the Akaike Information Criterion (AIC) to balance model complexity and goodness of fit. In this study, we determined the optimal order to be  $p=15$  (approximately 1.4 seconds), sufficient to capture hemodynamic lags.



**Figure 3.** Effective connectivity network constructed using Granger Causality Analysis (GCA), illustrating the directional information flow and causal influence, which is crucial for analysing the "disconnection syndrome" in Alzheimer's Disease. Subpicture (a), (b), and (c) show different results of AD, MCI and NC.

The magnitude of Granger causality  $F_{2 \rightarrow 1}$  is quantified as:

$$F_{2 \rightarrow 1} = \ln \left( \frac{\text{var}(E_{1R})}{\text{var}(E_{1U})} \right) \quad (7)$$

where  $\text{var}(\cdot)$  denotes the variance of the residual errors. If  $X_2$  contains unique information about the future of  $X_1$ , the unrestricted error  $E_{1U}$  will be significantly smaller than  $E_{1R}$ , yielding a positive F value.

This computation results in a  $48 \times 48$  asymmetric adjacency matrix for each subject. For TDA input, which typically operates on metric spaces, we symmetrized this matrix to define edge weights  $W_{ij}$

$$W_{ij} = \frac{F_{i \rightarrow j} + F_{j \rightarrow i}}{2} \quad (8)$$

To convert these weights (strength) into distances (dissimilarity), we applied the inversion:  $D_{ij} = 1/W_{ij}$  (normalized).

### 3.4. Topological Data Analysis: Persistent Homology

The core of our proposed method is the extraction of robust topological features. While traditional graph-theoretic metrics (e.g., degree, efficiency, small-worldness) provide valuable insights, they rely heavily on a fixed threshold to binarize the connectivity matrix. This thresholding process is often arbitrary and can lead to the loss of critical information: a high threshold may fracture weak but biologically significant connections, while a low threshold introduces noise. To overcome this limitation, we employ Persistent Homology (PH), a method rooted in algebraic topology. PH

tracks the evolution of topological features across a continuous range of scales (filtration), thereby providing a multi-scale summary of the brain network's intrinsic structure that is invariant to continuous deformations and robust to noise.

### 3.4.1. Mathematical Foundations: From Simplicies to Homology Groups

We model the functional brain network as a discrete combinatorial structure known as a Simplicial Complex.

Simplex means given a set of vertices  $V$  (in our case, the 48 fNIRS channels), a  $k$ -simplex  $\sigma_k$  is a convex hull of  $k + 1$  affinely independent vertices  $\{v_0, v_1, \dots, v_k\}$ . Specifically, 0-simplex is a vertex (brain region), 1-simplex is an edge (functional connection), 2-simplex is a triangle (functional triplet/cliue).

A simplicial complex  $K$  is a collection of simplices such that: If  $\sigma \in K$ , then every face of  $\sigma$  is also in  $K$ . The intersection of any two simplices in  $K$  is a face of both.

To rigorously quantify the "holes" or voids in the complex, we introduce the concept of Homology Groups over a field  $\mathbb{F}_2$  (typically  $\mathbb{Z}_2$  for computational efficiency). Let  $C_k$  be the vector space generated by the  $k$ -simplices of  $K$ . We define the Boundary Operator  $\partial_k: C_k \rightarrow C_{k-1}$  as a linear map that sends a simplex to the formal sum of its  $(k - 1)$ -dimensional faces. A fundamental property of the boundary operator is  $\partial_{k-1} \circ \partial_k = 0$ , which implies that the boundary of a boundary is empty.

Based on this, we define two subspaces:

Cycle Group ( $Z_k$ ): The kernel of the boundary operator,  $Z_k = \ker \partial_k = \{c \in C_k \mid \partial_k c = 0\}$ . Elements of  $Z_k$  represent cycles (loops, voids) that have no boundary.

Boundary Group ( $B_k$ ): The image of the boundary operator,  $B_k = \text{im } \partial_{k+1} = \{c \in C_k \mid \exists d \in C_{k+1}, \partial_{k+1} d = c\}$ . Elements of  $B_k$  are cycles that are "filled in" by higher-dimensional simplices.

The  $k$ -th Homology Group  $H_k(K)$  is defined as the quotient group:

$$H_k(K) = \frac{Z_k}{B_k} \quad (9)$$

This algebraic structure effectively filters out "trivial" cycles (those that are boundaries of solid shapes) and retains "non-trivial" holes. The rank of the homology group is the Betti number  $\beta_k = \dim(H_k)$ , which serves as a topological invariant.

By analyzing the Betti numbers across this filtration process, we can derive meaningful neurobiological insights:

$\beta_0$ : Number of connected components. In the context of AD, a higher  $\beta_0$  at late filtration stages indicates network fragmentation and a lack of global integration.

$\beta_1$ : Number of 1-dimensional holes (loops). Physically, these represent redundant pathways for information flow. A reduction in  $\beta_1$  suggests a loss of cognitive reserve and network resilience.

$\beta_2$ : Number of 2-dimensional voids or cavities. Physically, these represent higher-order functional cliques and complex 3D integration pathways among multiple brain regions. A reduction in  $\beta_2$  indicates a breakdown of higher-order cognitive coordination and a disrupted global functional architecture.

### 3.4.2. The Vietoris-Rips Filtration

Since the brain network is a weighted graph derived from Granger Causality, it naturally possesses a metric structure rather than a fixed binary structure. To analyze the topology across all possible thresholds, we construct a Filtration.

Let  $V$  be the set of channels and  $D$  be the distance matrix, where  $D_{ij}$  represents the dissimilarity (inverse of causal strength) between channel  $i$  and  $j$ . For a scale parameter  $\epsilon \geq 0$ , the Vietoris-Rips Complex  $Rips_\epsilon(V)$  is defined as:

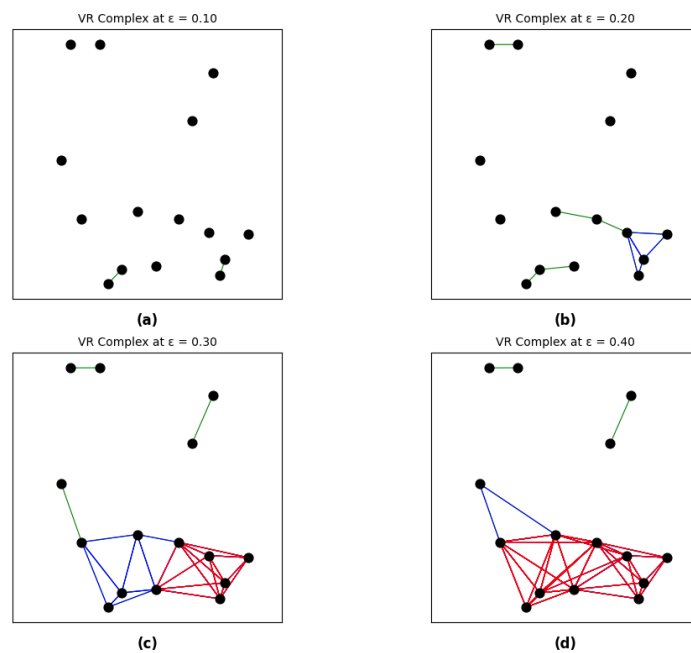
$$\sigma \in Rips_\epsilon(V) \Leftrightarrow \forall u, v \in \sigma, D(u, v) \leq \epsilon \quad (10)$$

This definition implies that a simplex is formed if and only if the pairwise distances between all its vertices are at most  $\epsilon$ . As  $\epsilon$  increases from 0 to  $\epsilon_{max}$ , we generate a nested sequence of simplicial complexes, known as the filtration:

$$\emptyset \subseteq K_{\epsilon_0} \subseteq K_{\epsilon_1} \subseteq \dots \subseteq K_{\epsilon_n} = K_{final} \quad (11)$$

As illustrated in Figure 4, this filtration simulates the dynamic evolution of the brain's functional architecture by systematically varying the connection threshold:

- 1) At  $\epsilon = 0$  (or low  $\epsilon$ , e.g.,  $\epsilon = 0.10$  in Figure 4a): The network starts as a cloud of disconnected points. In our study, this corresponds to  $\beta_0 = 48$  and  $\beta_1 = 0$ , where no functional integration has yet occurred.
- 2) As  $\epsilon$  increases (e.g.,  $\epsilon = 0.20$  to  $0.30$  in Figure 4b and c): Edges begin to form between channels with strong causal links. This process leads to the merging of connected components (death of H0 features) and the formation of cyclic pathways or loops (birth of H1 features), representing the emergence of local information processing clusters.
- 3) At sufficiently large  $\epsilon$  (e.g.,  $\epsilon = 0.40$  in Figure 4d): Most loops become "filled" by triangles or higher-order simplices (death of H1), and subsequent voids become "filled" by tetrahedrons or higher-order simplices (death of H2). In this stage, the network transitions into a single giant component, reflecting a state of global integration.



**Figure 4.** Schematic of the Rips Filtration process. As the scale parameter  $\epsilon$  increases, points are connected to form simplicial complexes. We track the 'birth' ( $\epsilon_{birth}$ ) and 'death' ( $\epsilon_{death}$ ) of topological features.

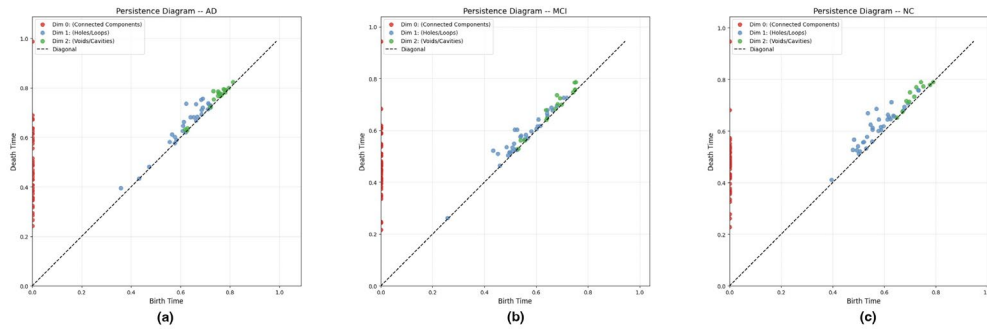
By tracking the 'birth' ( $\epsilon_{birth}$ ) and 'death' ( $\epsilon_{death}$ ) of these topological features across the filtration, we can quantify the persistence of brain network structures, providing a robust signature of cognitive health.

### 3.4.3. Persistence Diagrams and Stability Theory

The filtration process allows us to track the lifespan of each topological feature. If a homology class  $\gamma \in H_k$  is created at filtration value  $\epsilon_b$  (birth) and merges into a trivial class at  $\epsilon_d$  (death), its Persistence is defined as  $p_\gamma = \epsilon_d - \epsilon_b$ . This birth-death pair  $(\epsilon_b, \epsilon_d)$  is represented as a point in the 2D Persistence Diagram (PD).

To resolve ambiguities during component merging, we apply the "Elder Rule": when two components merge, the one generated earlier (older) persists, while the younger one dies. This ensures that the most prominent structural features have the longest lifespans.

In the resulting PDs (Figure 5), signal is the points with high persistence (far from the diagonal) represent robust global topological structures that are invariant to small deformations, and the noise is the points near the diagonal ( $p \approx 0$ ) represent transient, local connections often attributed to measurement noise.



**Figure 5.** Persistence Diagrams (PDs) for AD(a), MCI(b), and NC(c) subjects. Blue dots represent H0 features (components), red dots represent H1 features (loops), and green dots represent H2 features (voids). The distribution differences imply distinct topological organizations.

The theoretical stability is a critical advantage of using PH for fNIRS data is the Stability Theorem. It states that for two functions (or distance matrices)  $f$  and  $g$ , the Bottleneck distance  $d_B$  between their persistence diagrams is bounded by the  $L_\infty$  norm of their difference:

$$d_B(\text{Dgm}(f), \text{Dgm}(g)) \leq \|f - g\|_\infty \quad (12)$$

This theorem provides a rigorous mathematical guarantee that small perturbations in the input hemodynamic signals (e.g., physiological noise) result in only bounded, small changes in the topological signature. This robustness makes TDA superior to traditional graph metrics which can be unstable near threshold boundaries.

#### 3.4.4. Feature Vectorization: Persistence Images

While Persistence Diagrams capture the full topological information, they are multisets of points with varying cardinalities and do not possess a Hilbert space structure. Consequently, they cannot be directly used as input for standard machine learning algorithms (e.g., SVM, Random Forest, Neural Networks), which require fixed-length feature vectors. To bridge the gap between algebraic topology and statistical learning, we employ Persistence Images (PIs), a stable vector representation method.

The generation of PIs involves a three-step transformation mapping the PD  $\mathcal{D}$  to a vector  $V \in \mathbb{R}^d$ :

##### Step 1: Coordinate Transformation

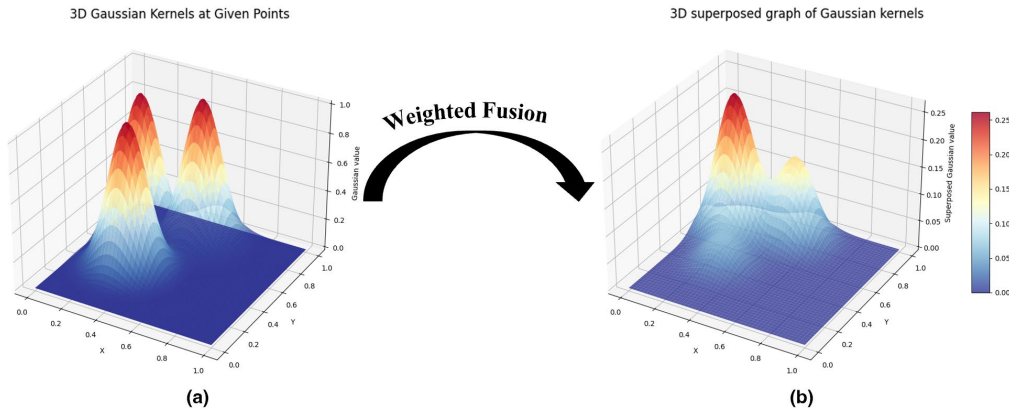
We first map each point  $(b, d) \in \mathcal{D}$  to birth-persistence coordinates  $(b, p)$  via the linear transformation  $T(b, d) = (b, d - b)$ . This transformation aligns the features such that the vertical axis explicitly represents the lifespan (persistence) of the topological feature.

##### Step 2: Weighted Surface Construction

To construct a continuous representation, we model the diagram as a sum of probability distributions. A differentiable kernel function  $k_u(x, y)$  (typically a 2D Gaussian) is centered at each transformed point  $u$ .

As demonstrated in Figure 6, individual topological features are first represented by discrete 3D Gaussian kernels. Through a "Weighted Fusion" process, these kernels are integrated into a continuous scalar field  $\rho(x, y)$ . Crucially, to emphasize significant topological features and suppress

noise near the diagonal, we apply a weighting function  $w(p)$  that is strictly increasing with persistence  $p$ .



**Figure 6.** Illustration of the Weighted Fusion process. The discrete topological features (left) are transformed into a continuous surface (right) using Gaussian kernels weighted by their persistence, bridging the gap between discrete homology and continuous vector spaces.

The scalar field  $\rho: \mathbb{R}^2 \rightarrow \mathbb{R}$  is defined as:

$$\rho(x, y) = \sum_{u \in T(\mathcal{D})} w(p_u) \cdot \frac{1}{2\pi\sigma^2} \exp\left(-\frac{\|(x, y) - u\|^2}{2\sigma^2}\right) \quad (13)$$

$$w(p) = \begin{cases} 0 & p \leq 0 \\ p/p_{max} & 0 < p < p_{max} \\ 1 & p \geq p_{max} \end{cases} \quad (14)$$

Here,  $\sigma$  is the bandwidth of the Gaussian kernel, which controls the smoothing scale. A properly tuned  $\sigma$  accounts for the uncertainty in the birth/death values derived from noisy biological data.

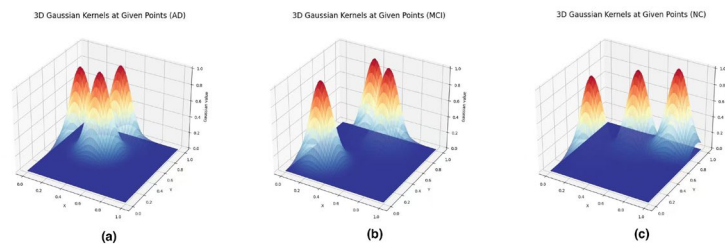
### Step 3: Discretization (Pixelation)

Finally, to obtain a finite-dimensional vector, the continuous surface  $\rho(x, y)$  is discretized by integrating it over a regular grid of size  $m \times n$ . The value of the pixel at index  $(i, j)$  is computed as:

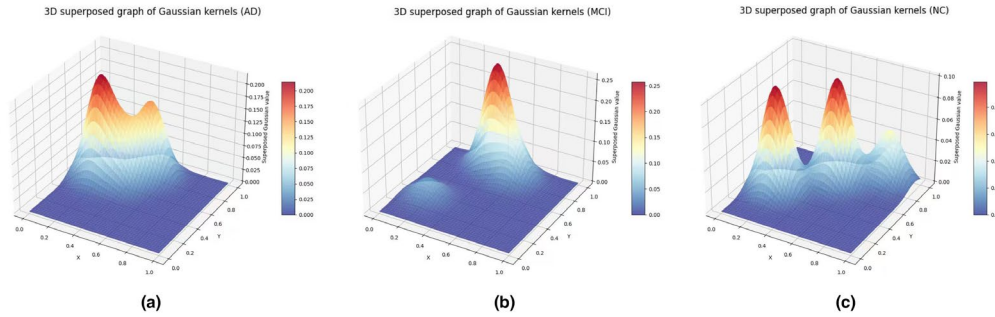
$$PI_{i,j} = \iint_{\text{pixel}_{i,j}} \rho(x, y) dx dy \quad (15)$$

Figures 7 and 8 illustrate the resulting continuous topological surfaces for representative subjects across the three clinical groups (AD, MCI, and NC).

Specifically, Figure 7 shows the distribution of raw Gaussian kernels before final integration, while Figure 8 presents the superposed scalar fields that form the basis of the Persistence Images. These visualizations reveal distinct topological "signatures": the AD group (Figure 8a) typically exhibits a different density of high-persistence peaks compared to the NC group (Figure 8c), reflecting alterations in brain network robustness.



**Figure 7.** 3D Gaussian kernels at given points for clinical groups. Representative plots for (a) Alzheimer's Disease (AD), (b) Mild Cognitive Impairment (MCI), and (c) Normal Controls (NC) showing the initial placement of kernels in the birth-persistence plane.



**Figure 8.** 3D superposed graph of Gaussian kernels (Persistent Surfaces). The integrated scalar field  $\rho(x,y)$  for (a) AD, (b) MCI, and (c) NC. These surfaces capture the collective topological strength of the brain networks, which are subsequently pixelated into Persistence Images for classification.

This process maps the topological space into a vector space. In this study, to prevent the "curse of dimensionality" given our sample size ( $N \approx 500$ ), we optimized the grid resolution to  $3 \times 3$ . Flattening this grid yields a compact 9-dimensional feature vector for each subject, effectively encapsulating the global topological signature of the brain network in a format suitable for the Voting Ensemble classifier.

#### 3.4.5. Machine Learning Classification Framework

The 9-dimensional topological feature vectors serve as inputs to our classification module. We employed a Voting Ensemble Classifier to predict the disease stage (NC, MCI, AD), which is proved have better performance than traditional machine learning and deep learning methods.

1) Base Classifiers: The ensemble aggregates predictions from three robust algorithms:

**Random Forest (RF):** An ensemble of decision trees. RF uses bagging (bootstrap aggregating) to reduce variance and is highly effective for high-dimensional data. We configured it with 100 trees and Gini impurity as the splitting criterion.

**XGBoost (Extreme Gradient Boosting):** A scalable implementation of gradient boosting. It sequentially builds trees to correct the residual errors of previous trees. We used a learning rate of 0.1 and a maximum depth of 4 to prevent overfitting.

**LightGBM:** A gradient boosting framework that uses tree-based learning algorithms. It is optimized for speed and efficiency, using leaf-wise tree growth.

2) Ensemble Strategy: We utilized a Soft Voting mechanism. For a given input sample  $x$ , each base classifier  $C_i$  outputs a probability distribution over the classes  $P_i(y|x)$ . The final prediction is the class with the highest average probability:

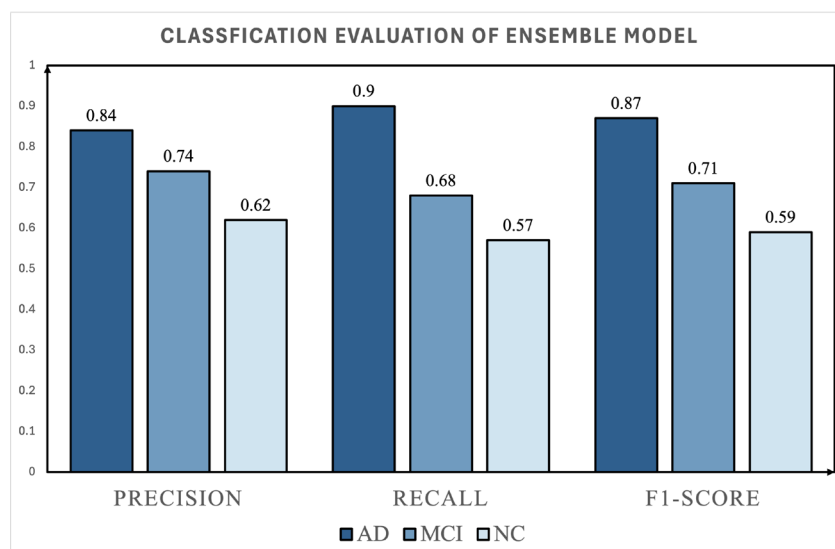
$$\hat{y} = \arg \max_{c \in \{NC, MCI, AD\}} \frac{1}{N} \sum_{i=1}^N P_i(y = c | x) \quad (16)$$

This approach typically yields better performance than hard voting (majority rule) as it takes into account the confidence of each classifier.

## 4. Results

### 4.1. Ensemble Model Performance

The proposed Voting Ensemble model achieved a significantly higher test accuracy of 77% on the held-out test set ( $N = 130$ ). The detailed performance breakdown is shown in Table 1 and graphically compared in Figure 9. Contrasting Figure 9 with the baseline in Figure 9, the ensemble approach demonstrates a substantial and balanced performance leap across all evaluation metrics for every diagnostic group.



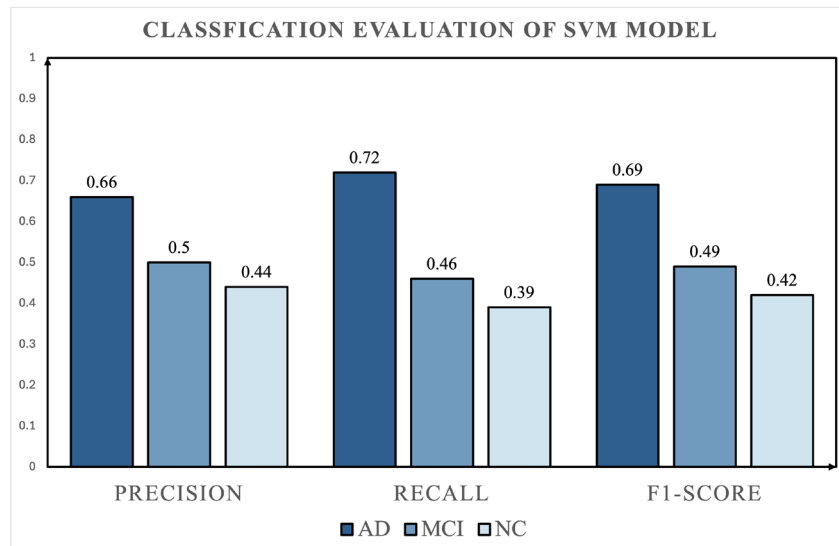
**Figure 9.** Classification evaluation of the proposed Voting Ensemble model. Corresponding to Table 1, this chart illustrates the robust performance of the ensemble approach. Compared to the baseline, the model shows significant improvements in Precision, Recall, and F1-Score across all classes, particularly excelling in the identification of Alzheimer's Disease .

**Table 1.** Performance of the proposed Ensemble Model on topological features. The model shows strong capability in distinguishing AD from other groups.

Classes	Total volumes	Testing set size	Precision	Recall	F1-Score
AD	325	65	0.84	0.90	0.87
MCI	185	37	0.74	0.68	0.71
NC	140	28	0.62	0.57	0.59
Total	650	130	0.77	—	—

#### 4.2. Baseline Comparison

To validate the effectiveness of our ensemble approach, we compared it against a baseline Support Vector Machine (SVM) with an RBF kernel. As shown in Table 2 and visually represented in Figure 10, the SVM achieved an accuracy of only about 58%. The bar chart clearly illustrates the model's struggle, displaying notably low precision and recall scores across all categories, particularly for the MCI and NC groups. This suggests that the decision boundary in the topological feature space is highly non-linear, and a single hyperplane is insufficient for separation.



**Figure 10.** Classification evaluation of the baseline SVM model. The bar chart highlighting the baseline model's difficulty in accurately classifying the topological features, especially within the intermediate MCI and NC stages.

**Table 2.** Baseline classification performance using SVM. The low accuracy highlights the difficulty of separating MCI from NC using a simple linear boundary in topological space.

Classes	Total volumes	Testing set size	Precision	Recall	F1-Score
AD	325	65	0.66	0.72	0.69
MCI	185	37	0.50	0.46	0.49
NC	140	28	0.44	0.39	0.42
Total	650	130	0.58	—	—

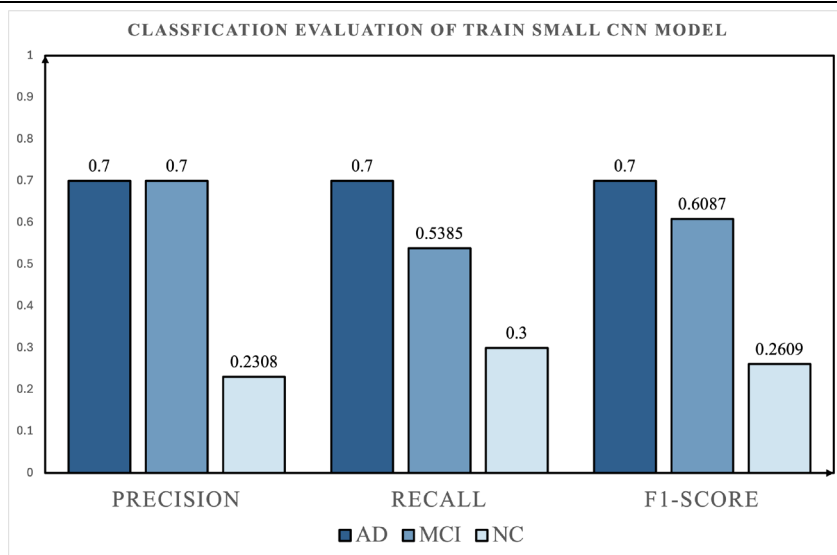
#### 4.3. Comparison with Deep Learning Baselines

To further validate the effectiveness of the proposed topology-based framework, we additionally compared our method with two deep learning baselines trained directly on the hemoglobin time-series data: a lightweight Small CNN and a CNN+Transformer model. Their class-wise evaluation results are shown in Table 3, Figure 11 and Figure 12. Both baselines were configured using the same DXY input and the same fixed temporal length setting. Among them, the Small CNN achieved the better performance, reaching a test accuracy of 55.81% with a Macro-F1 of 0.5232. Its class-wise F1-scores were 0.7000 for AD, 0.6087 for MCI, and 0.2609 for NC, indicating that although the model could capture part of the discriminative temporal information, its recognition ability for the NC group remained limited. In contrast, the CNN+Transformer model performed worse, with a test accuracy of only 48.84% and a Macro-F1 of 0.3841. Notably, the model failed to correctly identify any NC samples in the test set, resulting in an F1-score of 0.0000 for the NC category.

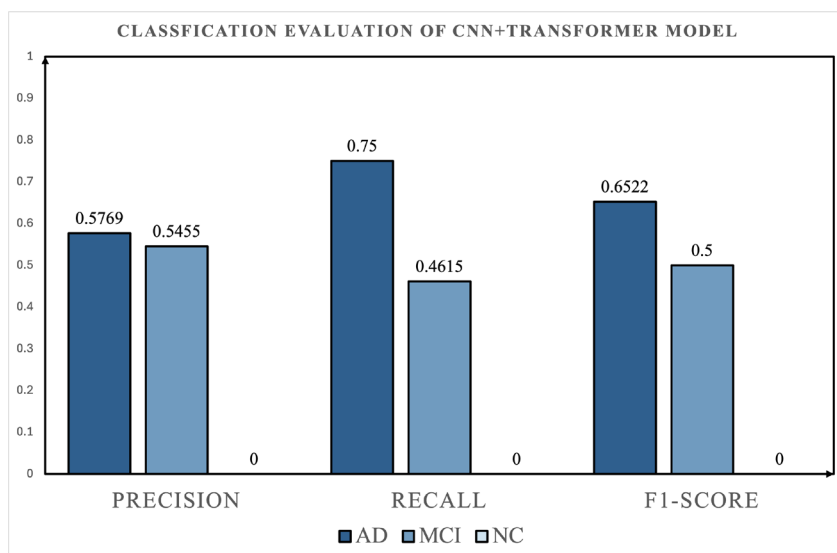
**Table 3.** Comparison between the proposed topology-based ensemble model and two end-to-end deep learning baselines.

Method	Classes	Precision	Recall	F1-Score
Ensemble Model	AD	<u>0.84</u>	<u>0.90</u>	<u>0.87</u>
	MCI	<u>0.74</u>	<u>0.68</u>	<u>0.71</u>

	NC	<u>0.62</u>	<u>0.57</u>	<u>0.59</u>
	Total	<u>0.7692</u>	—	—
Small CNN	AD	0.70	0.70	0.70
	MCI	0.70	0.54	0.61
	NC	0.23	0.30	0.26
	Total	0.5436	—	—
CNN+Transformer	AD	0.58	0.75	0.65
	MCI	0.55	0.46	0.50
	NC	0.00	0.00	0.00
	Total	0.3741	—	—



**Figure 11.** Classification evaluation of the Small CNN model. The bar chart shows the model's moderate performance in classifying the three diagnostic groups, with stronger results for AD and MCI than for the NC stage.



**Figure 12.** Classification evaluation of the CNN+Transformer model. The bar chart highlights the model's difficulty in achieving balanced classification, especially its failure to effectively distinguish the NC stage.

## 5. Discussion

### 5.1. Comparative Advantage of Topology-Based Representation

These results suggest that, under the current dataset size and class distribution, introducing a more complex Transformer-based temporal modeling module does not improve performance and may even aggravate class bias and optimization difficulty. By comparison, our proposed Voting Ensemble model based on topological features achieved a substantially higher test accuracy of 77%, with class-wise F1-scores of 0.87, 0.71, and 0.59 for AD, MCI, and NC, respectively. This comparison demonstrates that the proposed topology-based representation is more robust and discriminative than direct end-to-end temporal modeling in the present three-class AD classification setting.

Detailed Class-wise Analysis on Ensemble Model Performance based on topological features is showed below:

**AD Classification:** The model demonstrated high efficacy in identifying AD cases, with a Precision of 0.84 and Recall of 0.90. This indicates that the topological signature of late-stage Alzheimer's—characterized by severe network fragmentation—is distinct and well-captured by the Persistent Images.

**MCI Classification:** Distinguishing MCI from NC is historically challenging due to subtle prodromal changes. Our model achieved a Precision of 0.74 for MCI (a significant improvement from the SVM's 0.50). While lower than AD detection, this is a promising result, suggesting that TDA can detect early topological perturbations before they manifest as gross cognitive deficits.

**NC Classification:** The precision for NC was 0.62. The confusion matrix analysis revealed that misclassifications primarily occurred between NC and MCI groups. This overlap is clinically expected, as MCI is a transitional continuum rather than a discrete state.

### 5.2. Ablation Discussion

We conducted ablation studies to verify the contribution of different components.

#### 1) **Connectivity Metric:**

Replacing Granger Causality with Pearson Correlation resulted in a 5.2% drop in accuracy. This confirms that directional information flow is a critical biomarker for AD.

#### 2) **PI Resolution:**

Increasing the PI resolution to  $10 \times 10$  (100 features) reduced test accuracy by 3%, likely due to overfitting on the limited dataset. The  $3 \times 3$  resolution proved to be a robust low-dimensional embedding.

### 5.3. Limitations and Future Work

While promising, this study has limitations. The sample size, though augmented, is relatively small for deep learning generalization. Additionally, integrating this topological data with demographic factors (age, education) or multimodal data (MRI) could further boost classification accuracy, particularly for the challenging MCI group.

## 6. Conclusions

In this study, we presented a novel framework for Alzheimer's Disease classification that fuses functional near-infrared spectroscopy (fNIRS) with Topological Data Analysis (TDA). By moving beyond traditional graph theory and adopting a filtration-based approach, we successfully extracted robust, scale-invariant features that characterize the breakdown of brain networks in AD and MCI. To be specific, the key findings can be concluded as:

- 1) **Topological Biomarkers:** The Persistence Images revealed that AD brains exhibit a "topological simplification"—a loss of high-persistence loops and an increase in fragmented components. This aligns with the disconnection hypothesis.

- 2) **Methodological Robustness:** Our pipeline avoids the critical pitfall of arbitrary thresholding in network neuroscience. The use of Granger Causality further enhanced the model by incorporating directional coupling.
- 3) **Clinical Potential:** With an accuracy of nearly 77% using only 48 fNIRS channels, this approach offers a low-cost, portable screening tool that could be deployed in community clinics, unlike MRI or PET.
- 4) **Comparative Validation with Deep Learning Baselines:** Supplementary experiments using a Small CNN and a CNN+Transformer model showed that direct temporal modeling did not outperform the proposed topology-based ensemble framework under the current setting. In particular, the CNN+Transformer model exhibited poorer class balance and weak NC recognition, further highlighting the robustness of the proposed TDA-based representation.

**Author Contributions:** Conceptualization, S.L. and H.W. (Hangcheng Wu); methodology, S.L. and C.S.; software, S.L. and C.S.; validation, S.L., H.W. (Hangcheng Wu) and C.S.; formal analysis, S.L., H.W. (Hangcheng Wu), C.S., H.W. (Haoliang Wu) and Y.Q.; investigation, S.L., H.W. (Hangcheng Wu), C.S. and Y.Q.; resources, Y.L.; data curation, S.L. and Y.W.; writing—original draft preparation, S.L.; writing—review and editing, S.L., H.W. (Hangcheng Wu), Y.Q., H.W. (Haoliang Wu) and Z.Y.; visualization, S.L. and Y.Q.; supervision, Z.Y.; project administration, S.L.; funding acquisition, Z.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by Science and Technology Department of Sichuan Province, grant number 2026YFHZ0021.

**Institutional Review Board Statement:** Ethical review and approval were waived for this study as it is a retrospective analysis utilizing strictly de-identified and aggregated data. The current study poses no risk to subjects and involves no access to protected health information or direct human subject interaction.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The datasets generated and/or analyzed during the current study are not publicly available due to ethical considerations/institutional policies but are available from the corresponding author on reasonable request.

**Acknowledgments:** During the preparation of this manuscript, the authors used Gemini for the purposes of language refinement, formatting, and editorial organization. The authors have reviewed and edited the output and take full responsibility for the content of this publication.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## Reference

1. Nawaz, A.; Anwar, S.M.; Liaqat, R.; Iqbal, J.; Bagci, U.; Majid, M. Deep Convolutional Neural Network Based Classification of Alzheimer's Disease Using MRI Data. In 2020 IEEE 23rd International Multitopic Conference (INMIC), Bahawalpur, Pakistan, 2020; pp. 1–6. <https://doi.org/10.1109/INMIC50486.2020.9318172>
2. Cai, H.; Sheng, X.; Wu, G.; Hu, B.; Cheung, Y.-M.; Chen, J. Brain Network Classification for Accurate Detection of Alzheimer's Disease via Manifold Harmonic Discriminant Analysis. *IEEE Trans. Neural Netw. Learn. Syst.* **2024**, *35*, 17266–17280. <https://doi.org/10.1109/TNNLS.2023.3301456>.
3. Bhandarkar, A.; Naik, P.; Vakkund, K.; Junjappanavar, S.; Bakare, S.; Pattar, S. Deep Learning Based Computer Aided Diagnosis of Alzheimer's Disease: A Snapshot of Last 5 Years, Gaps, and Future Directions. *Artif. Intell. Rev.* **2024**, *57*, 30. <https://doi.org/10.1007/s10462-023-10644-8>.
4. Aisen, P.S.; Cummings, J.; Jack, C.R., Jr.; Morris, J.C.; Sperling, R.; Frölich, L.; Jones, R.W.; Dowsett, S.A.; Matthews, B.R.; Raskin, J.; Scheltens, P.; Dubois, B. On the Path to 2025: Understanding the Alzheimer's Disease Continuum. *Alzheimers Res. Ther.* **2017**, *9*, 60. <https://doi.org/10.1186/s13195-017-0283-5>.

5. Faux, N.G.; Rembach, A.; Wiley, J.; Ellis, K.A.; Ames, D.; Fowler, C.J.; Martins, R.N.; Pertile, K.K.; Rumble, R.L.; Trounson, B.; Masters, C.L.; The AIBL Research Group; Bush, A.I. An Anemia of Alzheimer's Disease. *Mol. Psychiatry* **2014**, *19*, 1227–1234. <https://doi.org/10.1038/mp.2013.178>.
6. Lu, D.; Popuri, K.; Ding, G.W.; Balachandar, R.; Beg, M.F.; Alzheimer's Disease Neuroimaging Initiative. Multimodal and Multiscale Deep Neural Networks for the Early Diagnosis of Alzheimer's Disease Using Structural MR and FDG-PET Images. *Sci. Rep.* **2018**, *8*, 5697. <https://doi.org/10.1038/s41598-018-22871-z>.
7. Pinti, P.; Tachtsidis, I.; Hamilton, A.; Hirsch, J.; Aichelburg, C.; Gilbert, S.; Burgess, P.W. The Present and Future Use of Functional Near-Infrared Spectroscopy (fNIRS) for Cognitive Neuroscience. *Ann. N. Y. Acad. Sci.* **2020**, *1464*, 5–29. <https://doi.org/10.1111/nyas.13948>.
8. Xie, L.; Liu, Y.; Gao, Y.; Zhou, J. Functional Near-Infrared Spectroscopy in Neurodegenerative Disease: A Review. *Front. Neurosci.* **2024**, *18*, 1469903. <https://doi.org/10.3389/fnins.2024.1469903>.
9. Pinti, P.; Tachtsidis, I.; Hamilton, A.; Hirsch, J.; Aichelburg, C.; Gilbert, S.; Burgess, P.W. The Present and Future Use of Functional Near-Infrared Spectroscopy (fNIRS) for Cognitive Neuroscience. *Ann. N. Y. Acad. Sci.* **2020**, *1464*, 5–29. <https://doi.org/10.1111/nyas.13948>.
10. Ferrer, I.; Gómez, A.; Carmona, M.; Huesa, G.; Porta, S.; Riera-Codina, M.; Biagioli, M.; Gustincich, S.; Aso, E. Neuronal Hemoglobin Is Reduced in Alzheimer's Disease, Argyrophilic Grain Disease, Parkinson's Disease, and Dementia with Lewy Bodies. *J. Alzheimers Dis.* **2011**, *23*, 537–550. <https://doi.org/10.3233/JAD-2010-101485>.
11. Ali, D.; Asaad, A.; Jimenez, M.-J.; Nanda, V.; Paluzo-Hidalgo, E.; Soriano-Trigueros, M. A Survey of Vectorization Methods in Topological Data Analysis. *IEEE Trans. Pattern Anal. Mach. Intell.* **2023**, *45*, 14069–14080. <https://doi.org/10.1109/TPAMI.2023.3308391>.
12. Bhandarkar, A.; Naik, P.; Vakkund, K.; Junjappanavar, S.; Bakare, S.; Pattar, S. Deep Learning Based Computer Aided Diagnosis of Alzheimer's Disease: A Snapshot of Last 5 Years, Gaps, and Future Directions. *Artif. Intell. Rev.* **2024**, *57*, 30. <https://doi.org/10.1007/s10462-023-10644-8>.
13. Suk, H.-I.; Shen, D. Deep Learning-Based Feature Representation for AD/MCI Classification. In *Medical Image Computing and Computer-Assisted Intervention—MICCAI 2013*; Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N., Eds.; Lecture Notes in Computer Science; Springer: Berlin, Heidelberg, **2013**; Vol. 8150; pp. 583–590. [https://doi.org/10.1007/978-3-642-40763-5\\_72](https://doi.org/10.1007/978-3-642-40763-5_72)
14. Klöppel, S.; Stonnington, C.M.; Chu, C.; Draganski, B.; Scahill, R.I.; Rohrer, J.D.; Fox, N.C.; Jack, C.R.; Ashburner, J.; Frackowiak, R.S.J. Automatic Classification of MR Scans in Alzheimer's Disease. *Brain* **2008**, *131*, 681–689. <https://doi.org/10.1093/brain/awm319>.
15. Liu, S.; Liu, S.; Cai, W.; Pujol, S.; Kikinis, R.; Feng, D. Early Diagnosis of Alzheimer's Disease with Deep Learning. In *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*, Beijing, China, **2014**; pp. 1015–1018. <https://doi.org/10.1109/ISBI.2014.6868045>.
16. Korolev, S.; Safiullin, A.; Belyaev, M.; Dodonova, Y. Residual and Plain Convolutional Neural Networks for 3D Brain MRI Classification. In *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, Melbourne, VIC, Australia, **2017**; pp. 835–838. <https://doi.org/10.1109/ISBI.2017.7950647>.
17. Hazarika, R.A.; Kandar, D.; Maji, A.K. A Novel Machine Learning Based Technique for Classification of Early-Stage Alzheimer's Disease Using Brain Images. *Multimed. Tools Appl.* **2024**, *83*, 24277–24299. <https://doi.org/10.1007/s11042-023-16379-6>.
18. Suk, H.-I.; Lee, S.-W.; Shen, D.; the Alzheimer's Disease Neuroimaging Initiative. Hierarchical Feature Representation and Multimodal Fusion with Deep Learning for AD/MCI Diagnosis. *NeuroImage* **2014**, *101*, 569–582. <https://doi.org/10.1016/j.neuroimage.2014.06.077>.
19. Liu, S.; Liu, S.; Cai, W.; Che, H.; Pujol, S.; Kikinis, R.; Feng, D.; Fulham, M.J.; ADNI. Multimodal Neuroimaging Feature Learning for Multiclass Diagnosis of Alzheimer's Disease. *IEEE Trans. Biomed. Eng.* **2015**, *62*, 1132–1140. <https://doi.org/10.1109/TBME.2014.2372011>.
20. Xu, H.; Zhong, S.; Zhang, Y. Multi-Level Fusion Network for Mild Cognitive Impairment Identification Using Multi-Modal Neuroimages. *Phys. Med. Biol.* **2023**, *68*, 095018. <https://doi.org/10.1088/1361-6560/accac8>.
21. Rao, K.N.; Gandhi, B.R.; Rao, M.V.; Javvadi, S.; Vellela, S.S.; Basha, S.K. Prediction and Classification of Alzheimer's Disease Using Machine Learning Techniques in 3D MR Images. In *2023 International*

Conference on Sustainable Computing and Smart Systems (ICSCSS), Coimbatore, India, 2023; pp. 85–90.  
<https://doi.org/10.1109/ICSCSS57650.2023.10169550>.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.