ECoGNet: an **E**EG-based Effective **Co**nnectivity **G**raph Neural **Net**work for Brain Disorder Detection

Jacopo Burger*, Vittorio Cuculo†, Alessandro D'Amelio*, Giuliano Grossi*, Raffaella Lanzarotti*
*Department of Computer Science, University of Milan, Italy
Email: {name.surname}@unimi.it
†Department of Computer Engineering, University of Modena and Reggio Emilia, Italy
Email: {name.surname}@unimore.it

***Abstract*—Alzheimer's Disease (AD) and Frontotemporal Dementia (FTD), among the most prevalent neurodegenerative disorders, disrupt brain activity and connectivity, highlighting the need for tools that can effectively capture these alterations. Effective Connectivity Networks (ECNs), which model causal interactions between brain regions, offer a promising approach to characterizing AD and FTD related neural changes. In this study, we estimate ECNs from EEG traces using a state-of-the-art causal discovery method specifically designed for time-series data, to recover the causal structure of the interactions between brain areas. The recovered ECNs are integrated into a novel Graph Neural Network architecture (ECoGNet), where nodes represent brain regions and edge features encode causal relationships. Our method combines ECNs with features summarizing local brain dynamics to improve AD and FTD detection. Evaluated on a publicly available EEG dataset, the proposed approach demonstrates superior performance compared to models that either use non-causal connectivity networks or omit connectivity information entirely.**

I. Introduction

Alzheimer's Disease (AD) and Frontotemporal Dementia (FTD) are among the most prevalent forms of neurodegenerative disorders. Projections estimate that the number of individuals affected by AD will rise to 75 million by 2030 and 131 million by 2050 [1]. Similarly, estimates for FTD suggest that dementia rates will double approximately every 20 years, surpassing 115 million cases by 2050 [2]. There is a substantial overlap of symptoms between Alzheimer's disease and frontotemporal dementia, making differential diagnosis challenging in many cases. AD is characterized by progressive neuronal loss due to the accumulation of abnormal proteins which disrupt brain function. These changes primarily affect the temporal lobe and hippocampus, often altering the brain before clinical symptoms of dementia appear [3]. FTD is characterized by progressive and relatively selective atrophy of the frontal and/or temporal lobes [4]. Its variants are associated with four distinct types of histopathology, all sharing common features such as neuronal loss and astrocytic gliosis. However, they differ in the abnormal processing and deposition of specific proteins.

According to neuroscience literature [5], AD and FTD lead to synaptic loss and cellular death, disrupting both the activity within individual brain regions and the communication among them. Consequently, the dynamics of neural activity within a given region, as well as the structural mechanisms governing information exchange across regions, are expected to show distinct patterns when comparing healthy subjects (CN) with AD and FTD patients. The connectivity patterns revealing the relationships among brain regions are typically referred to as Brain Connectivity Networks (BCNs). These can generally be categorized into either functional connectivity (FCNs) or effective connectivity networks (ECNs) [6]. FCNs aim at summarizing the statistical dependence between brain areas [7], hence by design FCNs lack an explicit modeling of the direction of information flow between brain regions. In contrast, ECNs model how activity in one brain region causally influences another, aiming to determine the direction of information flow at multiple time lags. The hypothesis of this study is that accurately estimated ECNs provide a more nuanced representation of relevant brain activity patterns. When combined with an appropriate computational model for data classification, such as Graph Neural Networks (GNNs), this can enhance the discrimination of Alzheimer's disease (AD) and frontotemporal dementia (FTD) patients from healthy controls. Notably, GNNs have been successfully applied across a wide range of health and medical applications [8], [9].

ECNs can be estimated from any functional brain imaging modality (e.g. MEG, fMRI, PET or EEG) and then employed to characterize the spatio-temporal brain organization and to identify the relevant biomarkers for a given neurodegenerative disease. Among these modalities, electroencephalography (EEG) stands out as a non-invasive, cost-effective, and portable method for analyzing brain activity.

Recent literature has proposed several solutions for estimating ECNs from EEG data (*cfr.* Section II-B), ranging from well-established techniques (*e.g.* Granger Causality) to more complex and data-hungry end-to-end neural network-based solutions. However, selecting the most suitable approach for estimating an ECN, and more importantly, identifying the appropriate computational model to extract meaningful information for disease diagnosis, remains an open challenge.

This work assumes that constructing an ECN involves solving a causal discovery problem on the EEG multivariate time-series data. To this end, we employ the PCMCI (Peter-Clark Momentary Conditional Independence) [10] causal discovery method. PCMCI provides state-of-the-art causal graph recovery from time-series data by leveraging conditional in-

dependence tests rather than data-intensive parametric models [11]. This feature is particularly desirable in contexts with limited data availability, such as the one considered here.

The recovered causal structure is embedded into a novel Graph Neural Network architecture (ECoGNet), where each node represents a brain area (EEG electrode). Connections between nodes are initialized with edge features derived from the recovered ECN, while node features capture local brain dynamics.

Our main contributions can be summarized as follows:

1) We evaluate the importance of adopting ECN in conjunction with features summarizing the local brain dynamics for the recognition of AD and FTD. To estimate ECNs from EEG data we adopt state-of-the-art causal discovery methods for time series data.
2) We propose a novel GNN architecture (ECoGNet) explicitly devised to properly model both local brain dynamics and the causal relationships (direction of information flow) between different brain areas (ECN), for the objective of discriminating AD and FTD patients from healthy controls.

The proposed approach is evaluated on the publicly available dataset introduced in [12]. Results demonstrate that the estimated ECNs, when processed by the proposed GNN architecture, achieve the overall best performance for the detection of the considered brain disorders. Notably, using a FCN as input to the GNN (thereby ignoring causal information) significantly reduces the approach's performance. Similarly, excluding brain connectivity information entirely, leads to suboptimal results.

II. Related Works

Numerous studies have explored the use of EEG features to improve the diagnostic accuracy of AD and other forms of dementia [13]. These efforts encompass a wide array of methodologies, including both traditional ML models [14] and Deep Learning (DL) approaches [15]. ML models typically rely on engineered features that capture temporal, spectral, or statistical characteristics of EEG signals, while DL models process raw data in an end-to-end manner. Additionally, a growing body of research focuses on modeling Brain Connectivity Networks (BCNs) as either Functional Connectivity Networks (FCNs) or Effective Connectivity Networks (ECNs) to better capture interactions between brain regions [16].

This section provides a brief review of the relevant literature in this context, and concludes with a summary of how causal discovery can be achieved using the PCMCI method, as this approach is employed in our methodology.

A. Machine and Deep learning models

A wide variety of ML algorithms have been utilized in the literature for EEG classification in dementia detection. Traditional methods, such as Support Vector Machines (SVM) [17], k-Nearest Neighbors (kNN) [18], logistic regression [19], and random forests [18], remain relevant and widely used for classifying Alzheimer's Disease and other types of dementia. However, deep learning methodologies have gained significant traction in EEG-based dementia research due to their ability to extract and learn features directly from raw data without the need for hand-crafted features or prior knowledge. Notable examples of deep learning models applied to EEG classification include Convolutional Neural Networks (CNN) [20], Recurrent Neural Networks (RNN) [21], Transformers [22], and graph-based methods [23]. These models have demonstrated promising results in classifying EEG signals and identifying biomarkers for AD.

B. Brain Connectivity Networks

Recent years have seen a considerable amount of research focused on modeling BCNs in the form of either FCNs [24], [25] or ECNs [26]–[28], with possible applications to the diagnosis of brain disorder mechanisms.

Historically, the most widely adopted technique to derive the causal connections between brain areas has been Granger Causality [29], [30]. However, more recently, a wider variety of approaches has been employed. For instance, in [31] authors adopted Sugihara causality analysis to develop novel EEG biomarkers for discriminating normal aging from Mild Cognitive Impairment (MCI) and early Alzheimer's disease.

Successively, Li et al. [32] derived the ECN via an ultra-group Lasso method for diagnosing MCI. Their experimental findings revealed that the ECN-based approach significantly outperformed the traditional functional connectivity FCN-based method. Similarly, Chen et al. [33] proposed a message-passing algorithm to estimate the directional flow of information between brain regions, demonstrating that ECN-based methodologies offer a distinct advantage over FCN-based approaches in identifying neuroimaging biomarkers associated with diseases.

In [34] authors proposed an Effective Temporal-Lag Neural Network (ETLN) designed to simultaneously infer causal relationships and temporal-lag values between brain regions, with the advantage of end-to-end trainability. ETLN is applied to functional Magnetic Resonance Imaging (fMRI) signals from distinct brain regions to assess direct or indirect interactions between them, with the ultimate goal of detecting AD.

More recently, in [35] a directed structure learning GNN (DSL-GNN) was developed to estimate the ECN while learning to discriminate between Alzheimer's disease, Parkinson's disease and Healthy controls.

III. Background: Causal Discovery with PCMCI

PCMCI (Peter-Clark Momentary Conditional Independence) analyzes multivariate time series data (e.g., EEG signals) to infer causal relationships among each of the variables composing it. Discovered causes can be eventually represented via a Directed Acyclic Graph (DAG). As any causal discovery method, PCMCI comes with a set of standard assumptions on the data that ingests; beyond the causal Markov condition and faithfulness assumption [36], PCMCI assumes that causal effects between EEG traces do not change over time (causal stationarity), time series do not have contemporaneous causal

links and all the relevant variables are measured (causal sufficiency, or unconfoundedness) [10].

Here we apply PCMCI to non overlapping temporal segments (epochs) of the EEG traces. Given a time series $x_i(t)$, representing a segment of the EEG trace of the i-th electrode, PCMCI identifies causal dependencies using a Conditional Independence (CI) test. Causality between time series from to two different electrode segments $x_i(t)$ and $x_j(t)$ is assessed based on their conditional dependency, typically denoted by:

$$x_i(t) \perp x_j(t) \,|\, \mathbf{z}(t),$$

where $\mathbf{z}(t)$ is the set of conditional variables influencing their relationship.

PCMCI requires the definition of a limited numbers of fundamental hyperparameters:

- *The CI Test:* Determines the eventual causal effect between variables; options include Partial Correlation (ParCorr), robust-ParCorr (r-ParCorr), Chi-squared or G-Test. The appropriate CI test is selected based on data type, distribution, and sample size.
- *Time Lag (τ_{min}, τ_{max}):* Defines the range of temporal lags at which a causal effect should be determined. Wider time lag ranges allow for more exhaustive inferences, while requiring significantly higher computational resources.
- *Significance Level (α):* Sets the significance level of the CI independence test for pruning spurious dependencies.

PCMCI identifies causal relationships between time series by detecting directed, lagged links among the measured variables (EEG electrodes). It outputs a p-value matrix indicating statistical significance, test statistics quantifying dependency strength, and confidence intervals to assess result reliability. The resulting causal graph captures lagged interactions, providing insights into the system's causal structure.

IV. The Model

The proposed method employs a multi-step pipeline for analyzing EEG recordings, as illustrated in Figure 1, which outlines the key stages of the process. Briefly, the pipeline begins with an epoching phase, where the continuous EEG signal is segmented into discrete epochs for further analysis. Subsequently, feature extraction captures essential temporal and spectral characteristics from the raw EEG data. Concurrently, causal graphs are constructed to represent the dynamic relationships between different brain regions over time. The extracted information is structured as graphs with both node and edge features, capturing the relationships between brain regions. Finally, a graph convolutional network-based classifier (ECoGNet) processes these graphs, leveraging both node features and the underlying causal structure for classification.

A. Graph Construction

For each epoch, two key operations are performed: node feature extraction and causal graph estimation which defines the edge features.

i. Node Feature Extraction

In the feature extraction phase [37], a set of EEG biomarkers, well-established in the literature [38], [39] and associated with AD and FTD, is computed for each epoch and each electrode signal. These features include the power within the alpha, beta, theta, and delta frequency bands, along with statistical measures such as the mean, standard deviation, and interquartile range (IQR) of the signal. A visual comparison of the differences in average feature values is provided in Figure 2 for the AD vs. CN case and in Figure 3 for FTD vs. CN case.

From a qualitative standpoint, Figure 2 shows how Delta and Alpha power reveal minimal differences in central and posterior regions, with more variation in the frontal and occipital areas. Theta power exhibits broader differences, especially in the frontal and parietal regions, suggesting potential alterations in AD patients; conversely, Beta power differences are relatively low. Statistical measures such as mean, standard deviation, and interquartile range reveal more widespread differences, reflecting increased EEG variability in AD vs controls.

Concerning FTD vs. CN case (Figure 3), the image suggests substantial EEG differences between FTD patients and healthy controls, particularly in frontal regions, aligning with the disease's known impact on the frontal and temporal lobes. FTD may be associated with increased power in the lower frequency Delta and Theta bands, potentially indicating slowed brain activity or abnormal slow waves. A possible reduction in Alpha power, especially frontally, could reflect reduced alertness and cognitive slowing. Beta power changes are less clear but present. Variations in the spread of these differences, as seen in standard deviation and interquartile range maps, may reflect disease heterogeneity.

ii. ECN Estimation

The PCMCI algorithm is applied to each of the segmented EEG signals (epochs) using both the Partial Correlation (ParCorr) and the Robust Partial Correlation (r-ParCorr) as the conditional independence (CI) tests [40]. Both these CI tests ensure a robust evaluation of conditional dependencies between time-series variables, leveraging their strengths in managing linear relationships. The algorithm iteratively removes non-significant connections and identifies significant lagged and contemporaneous interactions, ultimately generating a causal graph that captures the directed causal relationships within the data. Figure 4 shows the highest causal links found by PCMCI across all subjects and time lags.

As can be observed, in healthy controls (Figure 4, left) numerous connections are evident, suggesting active communication between different brain regions. Frontal and temporal areas, seem to be highly involved, with many connections terminating there. This suggests a central role for the frontal cortex and temporal lobe in the overall network organization. We also observe connections spanning longer distances across the circle, implying communication between more spatially separated brain regions, suggesting integration of information across different functional areas.

FTD subjects (Figure 4, center) seem to present a noticeable shift in the network causal organization. The overall density of

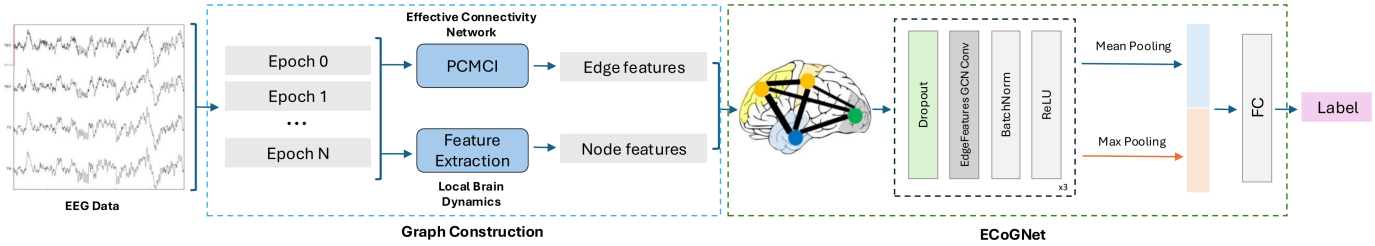

Fig. 1. Overview of the pipeline for causal graph construction and classification. First, node feature extraction and identification of causal relationships from EEG time-series data using the PCMCI algorithm are performed (Graph Construction). Subsequently, the obtained graph is processed by ECoGNet, which incorporates both node and edge features.

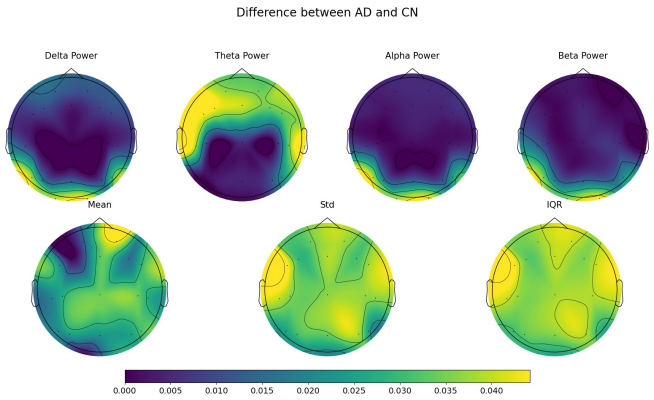

Fig. 2. Plots showing the differences between the average values of each feature across electrodes for the AD and CN groups.

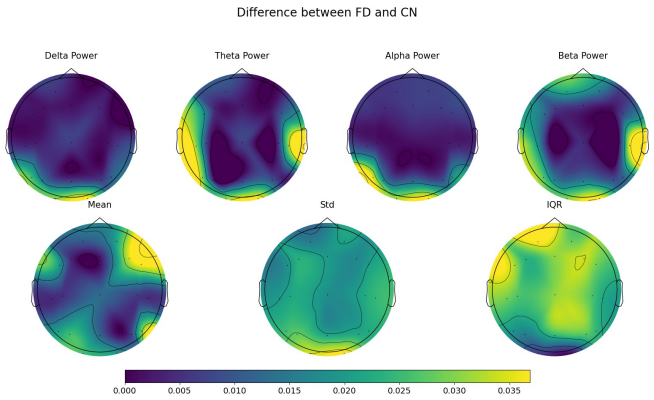

Fig. 3. Plots showing the differences between the average values of each feature across electrodes for the FTD and CN groups.

connections appears reduced compared to the CN group, suggesting a decrease in communication between brain regions. The connections to frontal areas, which were prominent in the CN network, seem to exhibit a slight decrease.

Eventually, AD subjects (Figure 4, right), exhibit the most sparse connectivity of the three. The number of connections is substantially reduced compared to both CN and FTD, indicating a further decline in inter-regional communication. Long-range connections, which were present in the CN network, are now greatly diminished, implying a breakdown in the integration of information across distant brain regions.

B. ECoGNet

A graph for each epoch is then constructed: each node v_i of the graph corresponds to an electrode of the EEG recording, with its features computed as previously reported (*cfr.* Section IV-A.i), resulting in a feature vector $\mathbf{f}_i \in \mathbb{R}^{F_{\text{node}}}$.

The graph is fully connected, and for each edge e_{ij} between two nodes (v_i, v_j), a feature vector is constructed based on the causal graph generated in the ECN Estimation (*cfr.* Section IV-A.ii). Specifically, for each edge and each time lag, a one-hot encoding is used to represent the three possible edge types: causal, undetermined, or no causal relationship. Furthermore, the causal strength is included as an additional numerical feature. The feature vectors corresponding to different time lags are then concatenated to form a comprehensive feature vector $\mathbf{e}_{ij} \in \mathbb{R}^{F_{\text{edge}}}$ for the edge e_{ij}. This concatenated vector encapsulates the temporal dynamics and causal relationships between the nodes across the various lags.

The graph convolutional network (GCN) that we propose is designed to process graph-structured data with edge features, so as to learn both node and edge representations through message passing, where each node aggregates information from its neighbors, considering both node and edge features.

The proposed layer, termed *EdgeFeatureGCNConv*, extends the traditional graph convolutional operation to incorporate edge attributes, enabling a richer representation of the graph structure. Let $\mathbf{F} \in \mathbb{R}^{N \times F_{\text{node}}}$ represent the node feature matrix with N nodes and F_{node} features, and let $\mathbf{E} \in \mathbb{R}^{E \times F_{\text{edge}}}$ denote the edge feature matrix with E representing the number of edges. The node features are transformed by a learnable linear layer:

$$\mathbf{F}' = \mathbf{F}\mathbf{W}_n, \tag{1}$$

where $\mathbf{W}_n \in \mathbb{R}^{F_{\text{node}} \times F_{\text{out}}}$ is a weight matrix. Similarly, the edge features, if provided, are linearly transformed:

$$\mathbf{E}' = \mathbf{E}\mathbf{W}_e, \tag{2}$$

where $\mathbf{W}_e \in \mathbb{R}^{F_{\text{edge}} \times F_{\text{out}}}$ is another weight matrix.

During message passing, the layer propagates information along the edges using the transformed features. For an edge between nodes v_i and v_j, the message is defined as:

$$\mathbf{m}_{ij} = \mathbf{f}'_j \cdot \mathbf{e}'_{ij},$$

where $\cdot$ denotes element-wise multiplication, $\mathbf{f}'_j$ represents the embedding of node j, obtained according to Equation (1), and

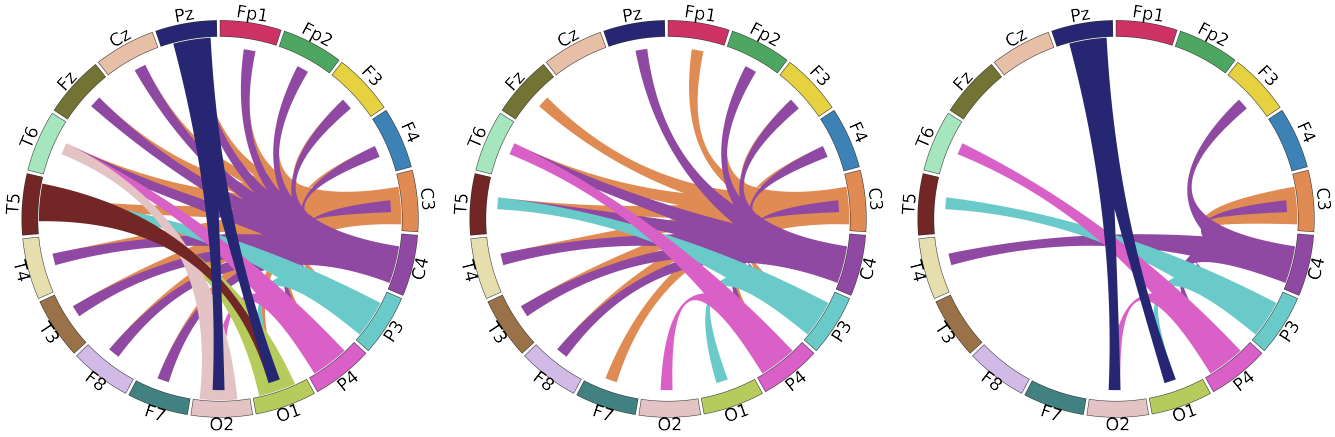

Fig. 4. Chord diagrams illustrating the highest causal links recovered by PCMCI from the adopted EEG traces. For visualisation purposes only arcs with a causal strength above a threshold are shown. Casual connections between various brain areas are shown across the three different groups: Healthy Controls (CN, **left**), Frontotemporal Dementia (FTD, **center**), and Alzheimer's Disease (AD, **right**). Each arc of the circle represents a causal link between two brain areas, while the color indicates the direction of information flow (source of the connection).

$\mathbf{e}'_{ij}$ represents the embedding of edge (i, j), obtained according to Equation (2). The aggregated messages for each node are computed based on the specified aggregation method (e.g., summation), and a learnable bias term $\mathbf{b} \in \mathbb{R}^{F_{\text{out}}}$ is added:

$$\mathbf{F}'_{\mathbf{i}} = \text{Aggregate}(\{\mathbf{m}_{ij} : j \in \mathcal{N}(i)\}) + \mathbf{b},$$

where $\mathcal{N}(i)$ represents the neighbors of node i. This layer, repeatedly applied throughout the architecture, effectively integrates node and edge information, enabling the model to capture both local and relational properties within the graph.

At the output stage, the features are aggregated using both global maximum pooling and global mean pooling across all nodes in the graph. These pooled features are concatenated and passed through a linear classifier (LC) to generate the final predictions. The classifier is designed to distinguish between classes, such as Alzheimer's patients and control subjects, based on the learned node and edge features. The final output is computed as:

$$\hat{y} = \text{LC}(\text{concat}(\text{GlobalMaxPool}(\mathbf{F}), \text{GlobalMeanPool}(\mathbf{F}))),$$

where $\hat{y}$ represents the predicted class label and $\mathbf{F}$ is the node feature matrix at the last layer.

V. Experimental analysis

In this section, we present a detailed analysis of the experiments conducted to evaluate the proposed methodology, the parameters used in the model, the evaluation metrics, and the experimental setup. Finally, we report and discuss the results, highlighting key findings and their implications.

A. Dataset

Experimental analysis has been conducted using the publicly available dataset presented in [12]. It includes recordings from 36 patients with Alzheimer's disease, 23 patients with frontotemporal dementia, and 29 healthy age-matched controls. The EEG data were acquired using a clinical system with 19 scalp electrodes while participants were in a resting state with their eyes closed. Signals were recorded using a monopolar montage, and the dataset includes both raw and preprocessed EEG in the standard BIDS format. For preprocessing, artifact subspace reconstruction and independent component analysis were employed to ensure high-quality, denoised signals. Additionally, the dataset provides each subject's Mini-Mental State Examination (MMSE) score. Each recording lasted approximately 13.5 minutes for the AD group (mean = 5.1, max = 21.3), 12 minutes for the FTD group (mean = 7.9, max = 16.9), and 13.8 minutes for the cognitively normal (CN) group (mean = 12.5, max = 16.5). In total, the dataset includes 485.5 minutes of AD recordings, 276.5 minutes of FTD recordings, and 402 minutes of CN recordings. In this study, we used all three groups—control subjects, patients with Alzheimer's disease, and those with frontotemporal dementia.

To ensure independence from the preprocessing phase, the analysis used the preprocessed data provided within the dataset.

B. Model configuration

The model architecture (*cfr.* Figure 1) is composed of three sequential blocks, each consisting of four distinct components arranged in the following order: a dropout layer, an EdgeFeatureGCNConv layer, a batch normalization layer, and a ReLU activation function. At the input stage, the node feature vector has an initial dimensionality of $F_{\text{node}} = 7$, while the edge feature vector starts with a dimensionality of $F_{\text{edge}} = 20$. As the data propagates through the network, the node feature vector undergoes transformations across hidden layers, with dimensions set to 32, 64, and 16, respectively. To promote generalization and reduce the risk of overfitting, a dropout layer with a probability of 0.5 is applied at the start of each block. Before producing the final output, the network applies both global max pooling and global mean pooling to aggregate information across all nodes. These pooled representations

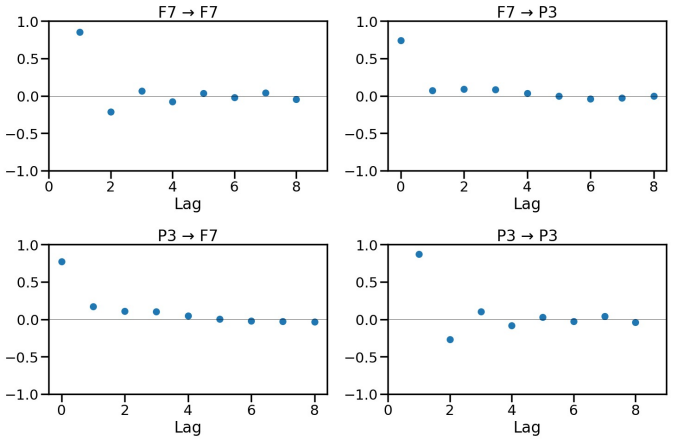

Fig. 5. Lagged unconditional dependencies for a representative subset of two electrodes (F7 and P3). The full analysis, which includes all 19 electrodes, was used to determine the optimal maximal time lag $\tau_{\max} = 4$.

are concatenated, resulting in a combined feature vector with a dimensionality of $2 \times F_{\text{hidden}}$, where $F_{\text{hidden}} = 16$ is the dimensionality of the last hidden layer. This concatenated representation is then passed through a final linear classifier, which outputs a vector of size $F_{\text{out}} = 2$.

The model was trained over a total of 500 epochs to ensure sufficient convergence. During training, a batch size of 512 was used to balance between memory efficiency and the stability of the optimization process. The Adam optimizer was selected with a learning rate of $\text{lr} = 0.0001$. Additionally, a weight decay of 1×10^{-5} was applied to regularize the model and prevent overfitting by penalizing large weights. These hyperparameter choices were made based on a series of preliminary experiments, the details of which are not included here as they did not significantly alter the outcome and were deemed unnecessary for the scope of this work.

C. Hyperparameters Tuning and Ablation Studies

The performance of the PCMCI algorithm in recovering directed causal relationships depends on several key hyperparameters. To optimize the method and ensure robust and interpretable results, we focused on the maximum time lag $\tau_{\max}$ and the significance level α of the PC algorithm.

In order to determine the appropriate value for the $\tau_{\max}$ parameter in the causal discovery algorithm, we visualized the lagged conditional dependencies (*i.e.*lagged correlations) from the ParCorr conditional independence test. This approach allows to gain insights into the temporal relationships within the data. Figure 5 reports the lagged dependencies for a subset of 2 electrodes, as a representative example. The analysis of the full set of 19 electrodes revealed an optimal value of $\tau_{\max} = 4$ as the conditional dependencies at lags greater than 4 were mostly zero or negligibly small, thus allowing us to ignore eventual causal relationships above such lag.

For what concerns the α parameter, setting the significance level of the PCMCI's conditional independence tests, we conducted an ablation analysis to assess its impact on the resulting causal graph and the downstream classification performance. By default, the PCMCI algorithm automatically evaluates a predefined set of candidate values for α and selects the one that best fits the input time series data. Alternatively, it is possible to specify a specific value for α to enforce consistency across analyses. To evaluate the implications of this choice, we compared the performance of the PCMCI algorithm when using its adaptive selection mechanism versus fixed values of α in the AD vs CN classification task. For a fair comparison, the fixed values considered in the ablation study matched the candidate values used by the algorithm during its adaptive selection process. The results of such analysis showed that the overall best performances are achieved when the significance level α is set to either 0.05, 0.1, or 0.5, with such vaues all delivering similar performances. Among these, $\alpha = 0.05$ is selected in order to maintain a stricter threshold for conditional independence tests, thereby reducing the likelihood of including spurious connections in the causal graph.

D. Results

To assess the performance of our framework, we employed Leave-One-Subject-Out Cross Validation (LOSO) to thoroughly evaluate the model's ability to generalize to unseen subjects. In LOSO, each subject in the dataset is treated as an independent entity, and the model is trained on data from all other subjects, leaving the data from the selected subject out for testing. Since each subject is associated with multiple epochs (data segments), this approach ensures that the model is tested on all available epochs for that specific subject, providing a more comprehensive evaluation of its ability to generalize to unseen individuals. Binary classification tasks were performed comparing AD vs. CN, FTD vs. CN, and AD vs. FTD. To evaluate our framework, we employed standard metrics such as Accuracy, ROC-AUC and F1 score. The metrics were computed at both an epoch-level - *i.e.* by treating each epoch as a separate instance - and subject-level - *i.e.* the predicted label is set as the most frequent prediction across all the epochs composing a subject.

To assess the specific contribution of the proposed approach, we trained two different instances of ECoGNet: one that relies on the recovered ECN (*cfr.* Section IV-B) and one that encodes mere correlational information (functional connectivity, FCN). Specifically, in the vein of [24], we devised an instance of ECoGNet which retained the same connections and node features as defined by the ECN causal graph, but replaced the edge weights with either Pearson correlation or Mutual Information values computed from the electrode pairs. In addition to this, we also benchmarked our model against two standard classifiers not implementing any network connectivity strategy: Support Vector Machine (SVM) and Decision Tree. For these models, the features from all nodes are concatenated into a single vector of size 133, representing the combined information from each node.

The detailed results are summarized in Table I, which provides a comprehensive overview of the model's performance

TABLE I

COMPARISON OF METHODS USING ECNS, FCNS, OR NO STRUCTURAL CONNECTIVITY. RESULTS ARE REPORTED AT THE SUBJECT-LEVEL AND EPOCH-LEVEL ACROSS DIFFERENT METRICS: ACCURACY (ACC), AREA UNDER THE CURVE (AUC) AND F1-SCORE (F1). BEST RESULTS ARE IN **BOLD**, SECOND-BEST ARE UNDERLINED.

Methods	**CN vs. AD**						**CN vs. FTD**						**AD vs. FTD**					
	Subject			**Epoch**			**Subject**			**Epoch**			**Subject**			**Epoch**		
	ACC	AUC	F1	ACC	AUC	F1	ACC	AUC	F1	ACC	AUC	F1	ACC	AUC	F1	ACC	AUC	F1
ECN-based																		
ParCorr	**83.08**	**86.57**	**85.24**	70.16	**77.55**	**70.62**	71.15	72.21	67.12	**66.03**	**69.14**	**65.75**	**67.80**	**65.21**	**58.20**	**60.86**	52.90	39.16
r-ParCorr	75.38	75.81	71.20	67.41	71.69	67.07	**76.92**	**74.44**	**71.11**	64.23	68.07	64.10	64.41	57.49	44.44	59.08	**54.80**	44.13
FCN-based																		
Pearson Corr.	73.85	73.04	69.09	68.47	73.20	62.29	63.46	64.47	59.37	62.26	61.40	58.89	59.32	49.94	35.56	56.76	51.19	37.40
Mutual Info.	73.85	73.04	69.09	68.72	72.69	62.40	73.85	62.29	55.81	57.71	61.91	53.53	57.63	52.90	40.91	53.64	45.67	21.92
No Structure																		
SVM	76.92	84.24	82.35	**70.55**	68.41	68.18	61.53	59.30	50	57.32	55.73	55.50	59.32	42.39	54.47	42.03	51.25	36.71
Decision Tree	76.92	81.37	81.08	64.35	63.97	63.96	63.43	76.61	65.12	61.96	60.99	61.03	61.01	57.07	52.09	46.17	48.75	**46.05**

across all evaluation criteria.

VI. DISCUSSION AND CONCLUSIONS

In this study, we present a novel approach for Alzheimer's disease (AD) and Frontotemporal Dementia (FTD) detection that leverages state-of-the-art causal discovery techniques to estimate effective connectivity networks (ECNs) from EEG data and integrates them with well established EEG features, capturing local brain dynamics. Such information is processed by a novel graph neural network architecture, namely ECoGNet. The results demonstrate that ECNs, which capture the direction of information flow between brain regions, provide a richer and more nuanced representation of brain dynamics compared to functional connectivity networks (FCNs). By jointly modeling both local brain activity and causal interactions, our method achieves superior performance in distinguishing AD and FTD patients from healthy controls.

Qualitative results show how the adopted features representation (*cfr.* Figure 2 and Figure 3) and, notably, the recovered causal structure of the brain (see Figure 4) offer great discrimination potential between healthy controls and patients affected by either AD or FTD. Actual classification results, summarized in Table I, showcase the effectiveness of different approaches based on ECNs, FCNs, and methods that do not incorporate structural information across different metrics. A clear trend emerges: our approach (ECoGNet) when ingesting ECNs estimated with PCMCI (using either ParCorr or r-ParCorr as conditional independence tests), consistently achieves the highest performance across all three classification tasks (CN vs. AD, CN vs. FTD, and AD vs. FTD). This highlights the crucial role of capturing causal information flow between brain regions in distinguishing these groups.

In contrast, when relying on functional connectivity (FCN-based) estimated via either Pearson Correlation or Mutual Information, ECoGNet demonstrates moderate performance, generally falling short of ECN-based approaches. While FCN captures statistical dependencies between brain regions, it proves to be less effective in this context.

Interestingly enough, the methods assuming no structural dependence between brain areas, but relying solely on features summarizing local brain dynamics, (SVM and Decision Tree classifiers), exhibit the lowest performance overall. This underscores the importance of connectivity-based features for successful classification. Simply relying on statistical features without considering the underlying network structure of the brain significantly diminishes the ability to differentiate between healthy controls and either AD or FTD.

In summary, our findings highlight the limitations of approaches that ignore causality or rely solely on statistical dependencies between brain regions. While FCN-based and feature-only baselines offer some predictive capacity, their inability to represent directional information significantly reduces their diagnostic performance for AD and FTD.

Future work could extend this approach to other imaging modalities, assess its robustness across diverse datasets, and explore its potential for tracking disease progression or predicting treatment responses. Moreover, some PCMCI assumptions for brain causal structure discovery may eventually be relaxed by some of its extensions (*e.g.*, [41]) at the expense of significantly higher computational requirements.

ACKNOWLEDGMENTS

This work was carried out at the Department of Computer Science "Giovanni degli Antoni", University of Milan – Via Celoria 18, Milan. It was partially funded by the National Plan for NRRP Complementary Investments (PNC, established with the decree-law 6 May 2021, n. 59, converted by law n. 101 of 2021) in the call for the funding of research initiatives for technologies and innovative trajectories in the health and care sectors (Directorial Decree n. 931 of 06-06-2022) - project n. PNC0000003 - AdvaNced Technologies for Human-centrEd Medicine (project acronym: ANTHEM). This work reflects only the authors' views and opinions, neither the Ministry for University and Research nor the European Commission can be considered responsible for them.

REFERENCES

[1] B. Winblad, P. Amouyel, S. Andrieu, C. Ballard, C. Brayne, H. Brodaty, A. Cedazo-Minguez, B. Dubois, D. Edvardsson, H. Feldman, *et al.*, "Defeating alzheimer's disease and other dementias: a priority for european science and society," *The Lancet Neurology*, vol. 15, no. 5, pp. 455–532, 2016.

[2] J. Bang, S. Spina, and B. L. Miller, "Frontotemporal dementia," *The Lancet*, vol. 386, no. 10004, pp. 1672–1682, 2015.

[3] H. Braak and E. Braak, "Neuropathological stageing of alzheimer-related changes," *Acta neuropathologica*, vol. 82, no. 4, pp. 239–259, 1991.

[4] J. D. Warren, J. D. Rohrer, and M. N. Rossor, "Frontotemporal dementia," *Bmj*, vol. 347, 2013.

[5] J. Zhou and W. W. Seeley, "Network dysfunction in alzheimer's disease and frontotemporal dementia: implications for psychiatry," *Biological psychiatry*, vol. 75, no. 7, pp. 565–573, 2014.

[6] K. J. Friston, "Functional and effective connectivity in neuroimaging: a synthesis," *Human brain mapping*, vol. 2, no. 1-2, pp. 56–78, 1994.

[7] L. Ioulietta, G. Kostas, N. Spiros, O. P. Vangelis, T. Anthoula, K. Ioannis, T. Magda, and K. Dimitris, "A novel connectome-based electrophysiological study of subjective cognitive decline related to alzheimer's disease by using resting-state high-density eeg egi ges 300," *Brain sciences*, vol. 10, no. 6, p. 392, 2020.

[8] S. Patania, G. Boccignone, S. Buršić, A. D'Amelio, and R. Lanzarotti, "Deep graph neural network for video-based facial pain expression assessment," in *Proceedings of the 37th ACM/SIGAPP Symposium on Applied Computing*, pp. 585–591, 2022.

[9] P. Veličković, "Everything is connected: Graph neural networks," *Current Opinion in Structural Biology*, vol. 79, p. 102538, 2023.

[10] J. Runge, P. Nowack, M. Kretschmer, S. Flaxman, and D. Sejdinovic, "Detecting and quantifying causal associations in large nonlinear time series datasets," *Science advances*, vol. 5, no. 11, p. eaau4996, 2019.

[11] J. Runge, A. Gerhardus, G. Varando, V. Eyring, and G. Camps-Valls, "Causal inference for time series," *Nature Reviews Earth & Environment*, vol. 4, no. 7, pp. 487–505, 2023.

[12] A. Miltiadous, K. D. Tzimourta, T. Afrantou, P. Ioannidis, N. Grigoriadis, D. G. Tsalikakis, P. Angelidis, M. G. Tsipouras, E. Glavas, N. Giannakeas, *et al.*, "A dataset of scalp eeg recordings of alzheimer's disease, frontotemporal dementia and healthy subjects from routine eeg," *Data*, vol. 8, no. 6, p. 95, 2023.

[13] L. Tomasello, L. Carlucci, A. Laganà, S. Galletta, C. V. Marinelli, M. Raffaele, and P. Zoccolotti, "Neuropsychological evaluation and quantitative eeg in patients with frontotemporal dementia, alzheimer's disease, and mild cognitive impairment," *Brain Sciences*, vol. 13, no. 6, p. 930, 2023.

[14] N. Sharma and M. H. Kolekar, "Dementia diagnosis with eeg using machine learning," in *Artificial Intelligence for Neurological Disorders*, pp. 107–129, Elsevier, 2023.

[15] M.-j. Kim, Y. C. Youn, and J. Paik, "Deep learning-based eeg analysis to classify normal, mild cognitive impairment, and dementia: Algorithms and dataset," *NeuroImage*, vol. 272, p. 120054, 2023.

[16] V. Sakkalis, "Review of advanced techniques for the estimation of brain connectivity measured with eeg/meg," *Computers in biology and medicine*, vol. 41, no. 12, pp. 1110–1117, 2011.

[17] N. Sharma, M. H. Kolekar, and K. Jha, "Eeg based dementia diagnosis using multi-class support vector machine with motor speed cognitive test," *Biomedical Signal Processing and Control*, vol. 63, p. 102102, 2021.

[18] R. Bakare, V. Shete, M. Tsolaki, and I. Kompatsiaris, "Machine learning approach for dementia classification using eeg signal analysis," in *2023 1st International Conference on Cognitive Computing and Engineering Education (ICCCEE)*, pp. 1–6, IEEE, 2023.

[19] M. Şeker, Y. Özbek, G. Yener, and M. S. Özerdem, "Complexity of eeg dynamics for early diagnosis of alzheimer's disease using permutation entropy neuromarker," *Computer Methods and Programs in Biomedicine*, vol. 206, p. 106116, 2021.

[20] C. Ieracitano, N. Mammone, A. Bramanti, A. Hussain, and F. C. Morabito, "A convolutional neural network approach for classification of dementia stages based on 2d-spectral representation of eeg recordings," *Neurocomputing*, vol. 323, pp. 96–107, 2019.

[21] M. Alessandrini, G. Biagetti, P. Crippa, L. Falaschetti, S. Luzzi, and C. Turchetti, "Eeg-based alzheimer's disease recognition using robust-pca and lstm recurrent neural network," *Sensors*, vol. 22, no. 10, p. 3696, 2022.

[22] A. Miltiadous, E. Gionanidis, K. D. Tzimourta, N. Giannakeas, and A. T. Tzallas, "Dice-net: a novel convolution-transformer architecture for alzheimer detection in eeg signals," *IEEE Access*, 2023.

[23] F. V. Farahani, W. Karwowski, and N. R. Lighthall, "Application of graph theory for identifying connectivity patterns in human brain networks: a systematic review," *frontiers in Neuroscience*, vol. 13, p. 585, 2019.

[24] D. Klepl, F. He, M. Wu, D. J. Blackburn, and P. Sarrigiannis, "Eeg-based graph neural network classification of alzheimer's disease: An empirical evaluation of functional connectivity methods," *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, vol. 30, pp. 2651–2660, 2022.

[25] T. Barbera, S. Zini, S. Bianco, and P. Napoletano, "Lightweight graph neural network for dementia assessment from eeg recordings," in *2024 IEEE 8th Forum on Research and Technologies for Society and Industry Innovation (RTSI)*, pp. 190–195, IEEE, 2024.

[26] S. Shimizu, P. O. Hoyer, A. Hyvärinen, A. Kerminen, and M. Jordan, "A linear non-gaussian acyclic model for causal discovery.," *Journal of Machine Learning Research*, vol. 7, no. 10, 2006.

[27] B. J. Frey and N. Jojic, "A comparison of algorithms for inference and learning in probabilistic graphical models," *IEEE TPAMI*, vol. 27, no. 9, pp. 1392–1416, 2005.

[28] S. M. Smith, K. L. Miller, G. Salimi-Khorshidi, M. Webster, C. F. Beckmann, T. E. Nichols, J. D. Ramsey, and M. W. Woolrich, "Network modelling methods for fmri," *Neuroimage*, vol. 54, no. 2, pp. 875–891, 2011.

[29] S. L. Bressler and A. K. Seth, "Wiener–granger causality: a well established methodology," *Neuroimage*, vol. 58, no. 2, pp. 323–329, 2011.

[30] A. B. Barrett, M. Murphy, M.-A. Bruno, Q. Noirhomme, M. Boly, S. Laureys, and A. K. Seth, "Granger causality analysis of steady-state electroencephalographic signals during propofol-induced anaesthesia," *PloS one*, vol. 7, no. 1, p. e29072, 2012.

[31] J. C. McBride, X. Zhao, N. B. Munro, G. A. Jicha, F. A. Schmitt, R. J. Kryscio, C. D. Smith, and Y. Jiang, "Sugihara causality analysis of scalp eeg for detection of early alzheimer's disease," *NeuroImage: Clinical*, vol. 7, pp. 258–265, 2015.

[32] Y. Li, H. Yang, K. Li, P.-T. Yap, M. Kim, C.-Y. Wee, and D. Shen, "Novel effective connectivity network inference for mci identification," in *Machine Learning in Medical Imaging: 8th International Workshop, MLMI 2017, Held in Conjunction with MICCAI 2017, Quebec City, QC, Canada, September 10, 2017, Proceedings 8*, pp. 316–324, Springer, 2017.

[33] N. Chen, M. Guo, Y. Li, X. Hu, Z. Yao, and B. Hu, "Estimation of discriminative multimodal brain network connectivity using message-passing-based nonlinear network fusion," *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, vol. 20, no. 4, pp. 2398–2406, 2021.

[34] Z. Xia, T. Zhou, S. Mamoon, A. Alfakih, and J. Lu, "A structure-guided effective and temporal-lag connectivity network for revealing brain disorder mechanisms," *IEEE Journal of Biomedical and Health Informatics*, vol. 27, no. 6, pp. 2990–3001, 2023.

[35] J. Cao, L. Yang, P. G. Sarrigiannis, D. Blackburn, and Y. Zhao, "Dementia classification using a graph neural network on imaging of effective brain connectivity," *Computers in Biology and Medicine*, vol. 168, p. 107701, 2024.

[36] P. Spirtes, C. Glymour, and R. Scheines, *Causation, prediction, and search*. MIT press, 2001.

[37] L.-M. Sánchez-Reyes, J. Rodríguez-Reséndiz, G. N. Avecilla-Ramírez, M.-L. García-Gomar, and J.-B. Robles-Ocampo, "Impact of eeg parameters detecting dementia diseases: A systematic review," *IEEE Access*, vol. 9, pp. 78060–78074, 2021.

[38] A. H. Meghdadi, M. Stevanović Karić, M. McConnell, G. Rupp, C. Richard, J. Hamilton, D. Salat, and C. Berka, "Resting state eeg biomarkers of cognitive decline associated with alzheimer's disease and mild cognitive impairment," *PloS one*, vol. 16, no. 2, p. e0244180, 2021.

[39] A. H. Meghdadi, A. Verma, M. F. Mendez, and C. Berka, "Comparative eeg biosignatures in alzheimer's disease and frontotemporal dementia: A pilot study," *Alzheimer's & Dementia*, vol. 20, p. e090543, 2024.

[40] L. Silfwerbrand, Y. Koike, P. Nyström, and M. Gingnell, "Directed causal effect with pcmci in hyperscanning eeg time series," *Frontiers in Neuroscience*, vol. 18, p. 1305918, 2024.

[41] A. Gerhardus and J. Runge, "High-recall causal discovery for autocorrelated time series with latent confounders," *Advances in Neural Information Processing Systems*, vol. 33, pp. 12615–12625, 2020.