

Article

Personalized Explanations for Early Diagnosis of Alzheimer's Disease Using Explainable Graph Neural Networks with Population Graphs

So Yeon Kim ^{1,2} ¹ Department of Artificial Intelligence, Ajou University, Suwon 16499, Republic of Korea; jebi1771@ajou.ac.kr² Department of Software and Computer Engineering, Ajou University, Suwon 16499, Republic of Korea

Abstract: Leveraging recent advances in graph neural networks, our study introduces an application of graph convolutional networks (GCNs) within a correlation-based population graph, aiming to enhance Alzheimer's disease (AD) prognosis and illuminate the intricacies of AD progression. This methodological approach leverages the inherent structure and correlations in demographic and neuroimaging data to predict amyloid-beta ($A\beta$) positivity. To validate our approach, we conducted extensive performance comparisons with conventional machine learning models and a GCN model with randomly assigned edges. The results consistently highlighted the superior performance of the correlation-based GCN model across different sample groups in the Alzheimer's Disease Neuroimaging Initiative (ADNI) dataset, suggesting the importance of accurately reflecting the correlation structure in population graphs for effective pattern recognition and accurate prediction. Furthermore, our exploration of the model's decision-making process using GNNExplainer identified unique sets of biomarkers indicative of $A\beta$ positivity in different groups, shedding light on the heterogeneity of AD progression. This study underscores the potential of our proposed approach for more nuanced AD prognoses, potentially informing more personalized and precise therapeutic strategies. Future research can extend these findings by integrating diverse data sources, employing longitudinal data, and refining the interpretability of the model, which potentially has broad applicability to other complex diseases.



Citation: Kim, S.Y. Personalized Explanations for Early Diagnosis of Alzheimer's Disease Using Explainable Graph Neural Networks with Population Graphs.

Bioengineering **2023**, *10*, 701.

<https://doi.org/10.3390/bioengineering10060701>

Academic Editor: Larbi Boubchir

Received: 22 May 2023

Revised: 5 June 2023

Accepted: 6 June 2023

Published: 8 June 2023



Copyright: © 2023 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: graph neural networks; alzheimer's disease; amyloid-beta positivity; population graph; explainable graph neural networks

1. Introduction

Alzheimer's disease (AD) is a progressive neurodegenerative disorder characterized by a series of changes in the brain that occur years or even decades before the first symptoms of cognitive decline become evident [1–4]. Amyloid-beta ($A\beta$) is a protein that is implicated in AD, one of the most common forms of dementia. $A\beta$ deposition, which can be observed via amyloid positron emission tomography (PET) imaging, is one of the earliest detectable pathological changes and a pathological hallmark of AD. It precedes other biomarkers, such as tau pathology, neuronal injury or neurodegeneration, and cognitive symptoms [5]. Once amyloid plaques start to build up, there is a cascade of events, including the accumulation of tau tangles inside neurons and eventual cell death, leading to brain atrophy, which can be observed through magnetic resonance imaging (MRI). Detecting the presence or predicting the onset of $A\beta$ positivity can, therefore, be instrumental in the early diagnosis and prevention of this debilitating disease.

Numerous studies have employed machine learning and deep learning methodologies for predicting amyloid pathology and other AD phenotypes [1,6–20], typically classifying individuals into the categories of cognitively normal (CN), mild cognitive impairment (MCI), and Alzheimer's disease (AD) [11–14]. Furthermore, several studies have proposed

machine learning frameworks designed to predict the conversion from MCI to AD [15–19], highlighting the potential of these methodologies in forecasting AD progression. Despite these advances, a crucial area of focus remains the early detection of AD pathology, especially in cognitively unimpaired individuals. A β deposition, distinguishable before any cognitive impairment becomes apparent, serves as a pivotal biomarker for identifying individuals predisposed to AD. It is noteworthy that cognitively unimpaired individuals may not yet demonstrate significant A β deposition and are often categorized as A β negatives. However, AD typically harbors an extended preclinical phase where, despite an absence of overt cognitive symptoms, individuals may already possess considerable A β deposition [21,22]. Thus, predicting A β positivity could significantly expedite AD diagnosis, optimize participant selection for clinical trials of disease-modifying therapies, and facilitate proactive monitoring and potential early intervention. Many researchers are increasingly targeting the early stages and the preclinical phase of AD in an attempt to curtail disease progression [1,8,20,23].

Various biomarkers, including demographic traits, genetic factors, and MRI imaging features, are key tools in predicting A β positivity [24,25]. Age, a significant demographic factor, is intrinsically linked with the risk of developing AD. As age advances, the probability of both AD and other dementia forms, as well as A β plaque accumulation, increases. Sex is another demographic characteristic of note, with women statistically more likely to develop AD compared to men, a fact which is an active area of investigation. Educational attainment, gauged as years of formal education completed, is associated with AD risk. Those with higher education levels often exhibit a lower risk of AD, potentially attributed to a bolstered cognitive reserve, which allows for increased tolerance of brain damage before dementia symptoms manifest.

Genetic factors, particularly the apolipoprotein E (APOE) ϵ 4 allele, have a significant role in AD risk prediction [26–28]. Carriers of this allele are at elevated risk of developing AD and often experience earlier onset of symptoms. This allele is believed to influence AD by modifying A β processing or clearance in the brain. MRI features, including brain atrophy and other structural changes linked to AD, offer powerful predictive tools for A β positivity [29,30]. Changes such as specific regional brain shrinkage, ventricular expansion, and alterations in white matter integrity can be detected by MRI. The combined utilization of these biomarkers offers a comprehensive approach to predicting A β positivity. The multifaceted strategy facilitates earlier, more precise diagnosis; improved prognostic predictions; and the potential for personalized treatment plans. It can also guide the design of clinical trials and the development of new therapeutic interventions, underlining the enhanced predictive model offered by their combined use.

Effective capture of the collective power of these biomarkers for the efficient diagnosis and prognosis of AD can be achieved through graph-based machine learning, particularly graph neural networks (GNNs). GNNs have made significant strides in the healthcare sector, modeling interactions between biological entities, predicting potential disease-associated genes, constructing patient similarity networks, and even playing a crucial role in drug discovery [31–41]. GNNs have been utilized to model intricate correlations between multiple biomarkers, such as genetic, clinical, and neuroimaging features, offering valuable insights into the underlying mechanisms of AD progression [42–46]. The strength of GNNs lies in their ability to interpret the interconnections between brain regions and the impacts of changes in these regions on AD progression. These capabilities can enhance understanding of the complex disease trajectory, allowing for more precise prediction of AD prognosis.

One such tool that has proven to be powerful in graph-based deep learning is the subset of GNNs known as graph convolutional networks (GCNs) [47]. They extend the concept of convolutional operations from regular, grid-like structures—typical in images—to irregular graph structures. Notably, GCNs have been shown to be useful in disease prediction, particularly for autism spectrum disorder and AD [38]. They operate on a population graph, where nodes represent individuals and edges symbolize similarity in certain characteristics,

thereby facilitating the deciphering of population-level patterns and individual variations in brain images.

Despite these advances, their inherent black-box nature poses a challenge due to limited transparency. This opacity can hinder understanding of the models' internal decision-making processes—a significant concern in the medical field, where model interpretability is crucial. To combat this issue, researchers are turning to explainable artificial intelligence (XAI) to foster more comprehensible and transparent GNNs. Various techniques, such as GradCAM-based explanation [48,49], PGExplainer [50], PGMEExplainer [51], XGNN [52], and GNNExplainer [53], have been developed to enhance the interpretability of GNNs. However, the use of explainable GNNs remains largely restricted to medical image analysis [54–56] and drug discovery [57–60], suggesting a need for broader application and integration. Moreover, the comprehensive prioritization of personalized biomarkers, crucial for personalized medicine in AD diagnosis and treatment, is a largely unexplored area.

This study aims to bridge this gap by leveraging GCNs [47] to offer accurate predictions alongside interpretable results, thus contributing to a more holistic understanding of individual AD prognoses. This study is motivated by the hypothesis that cohorts with analogous clinical or neuroimaging characteristics may show a correlation that extends beyond the influence of prevalent biomarkers, such as the APOE genotype. Thus, we can build a population graph where nodes symbolize individuals at risk and edges depict similarities in demographic, genetic, and neuroimaging attributes. Our study highlights the utility of GCNs in predicting A β positivity, a crucial early indicator of AD, by demonstrating our proposed correlation-based population graph of cognitively unimpaired individuals. Furthermore, we utilize GNNExplainer [53], an explainable GNN model, which optimizes a subgraph within an individual's neighborhood and pinpoints a set of crucial features integral for the prediction. For each individual, we prioritize personalized AD risk factors, allocating risk scores derived from the average importance values garnered from their neighbors. This elucidation process further unveils a significant variation in the biomarkers identified for AD prognosis across different sample groups. The overview of the proposed model is illustrated in Figure 1. The main contributions of this study include:

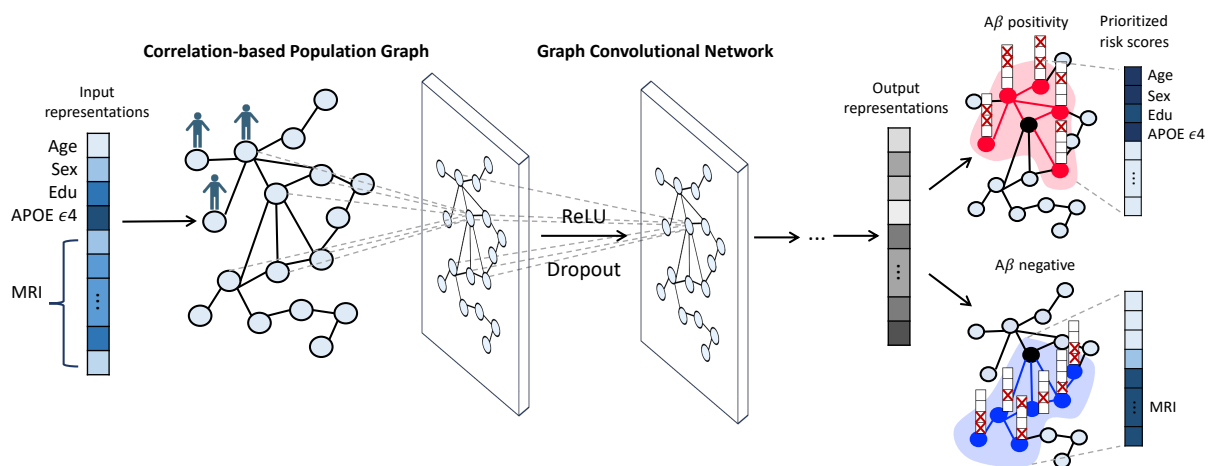


Figure 1. Overview of the proposed model. First, we construct a population graph in which the vertices represent individuals and are characterized by demographic features (age, sex, years of education), genetic information (APOE ϵ 4 status), and MRI imaging features (average cortical thickness values for 69 brain regions of interest (ROIs)). Edges are assigned when there is a high correlation between a pair of individuals. Next, we employ graph convolutional networks (GCNs) to analyze the population graph and predict the A β positivity for each individual. Finally, GNNExplainer provides an explanation for each prediction, optimizing a subgraph of the individual's neighborhood and identifying a set of crucial features for the prediction. For each individual, we prioritize the top 10 personalized biomarkers by assigning risk scores based on the average importance values obtained from their neighbors.

- We demonstrate the effectiveness of using graph neural networks on population graphs for early AD diagnosis in cognitively unimpaired individuals;
- We provide explanations of graph neural network predictions, offering sample-level interpretations using demographic and neuroimaging features;
- We prioritized personalized risk factors for AD by explaining graph neural network predictions, thereby characterizing groups of individuals based on their risk factors in AD prognosis.

2. Materials and Methods

2.1. Dataset

In this study, we leveraged the Alzheimer’s Disease Neuroimaging Initiative (ADNI) GO/2 dataset (adni.loni.usc.edu, accessed on 21 May 2023). The ADNI was launched in 2003 as a public–private partnership led by principal investigator Michael W. Weiner, MD. The primary goal of ADNI is to test whether serial MRI, PET, other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer’s disease (AD). For up-to-date information, see www.adni-info.org (accessed on 21 May 2023). Our selection targeted individuals identified as either cognitively normal (CN) or with mild cognitive impairment (MCI) status. We utilized a total of 506 samples from the ADNI cohort encompassing 214 CN and 292 MCI samples, with each sample characterized by 73 features.

These features included 3 demographic aspects (age, sex, and years of education), APOE $\epsilon 4$ status, and 69 neuroimaging features derived from MRI scans. The APOE $\epsilon 4$ status, denoted by the count of $\epsilon 4$ alleles (0, 1, or 2), functions as a critical genetic biomarker associated with elevated AD risk, with two $\epsilon 4$ alleles conferring the highest susceptibility. The neuroimaging features derived from MRI provide a thorough representation of neuroanatomical alterations, encompassing metrics such as cortical thickness. In this study, we utilized quantitative ADNI MRI data. The quantitative MRI data specifically represent the cortical thickness from T1-weighted MRI images obtained from the University of California, San Francisco, and archived at the LONI Image and Data Archive (IDA). We used average cortical thickness values for each of the 69 brain regions of interest (ROIs). A comprehensive description of the MRI image data acquisition process can be found here: <https://adni.loni.usc.edu/methods/mri-tool/mri-analysis/> (accessed on 21 May 2023). As a result, there were 69 distinct numerical MRI features available for each individual.

Our study primarily focuses on predicting $A\beta$ positivity in cognitively unimpaired individuals. The levels of $A\beta$ were quantified from ^{18}F -florbetapir PET scans, which specifically bind to $A\beta$ plaques present in the brain. The measurements of $A\beta$ levels were obtained from the LONI Image and Data Archive (IDA) at the University of California, Berkeley. The burden of $A\beta$ deposits was evaluated using the averaged value of the standardized uptake value ratio (SUVR). A detailed description of the PET image analysis method can be found here: <https://adni.loni.usc.edu/methods/pet-analysis-method/> (accessed on 21 May 2023). If the averaged value exceeded a cutoff of 1.11, the individual was classified as $A\beta$ -positive, indicative of AD pathology [61,62]. The ADNI cohort contained 291 $A\beta$ -negatives and 215 $A\beta$ -positives. Summary statistics of the clinical features for the CN and MCI samples are presented in Table 1.

Table 1. Demographic and neuroimaging characteristics of cognitively normal (CN) and mild cognitive impairment (MCI) groups from the Alzheimer’s Disease Neuroimaging Initiative (ADNI) dataset.

	CN (N = 214)	MCI (N = 292)
Age	74.18 $\pm$ 6.14	73.14 $\pm$ 7.48
Sex	96 F/118 M	173 F/119 M
Education (years)	16.7 $\pm$ 2.54	16.13 $\pm$ 2.64
APOE $\epsilon 4$	144 neg/70 pos	173 neg/119 pos
$A\beta$ positivity	139 neg/75 pos	152 neg/140 pos

N: Number of samples; F: Female; M: Male; pos: Positive; neg: Negative

2.2. Population Graph Construction

In our study, we constructed a graph denoted by $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, where the vertex set $\mathcal{V} = (v_1, \dots, v_n)$ signifies the collection of individuals. Each vertex v_i is associated with the i -th individual and is characterized by a feature vector x_{v_i} , which is composed of demographic and neuroimaging features. Specifically, for each individual, we concatenate features such as age, sex, years of education, APOE $\epsilon 4$ status (0, 1, or 2), and quantitative MRI imaging features (average cortical thickness values for 69 brain ROIs), resulting in a 73-feature vector. The set $\mathcal{E}$ denotes the collection of undirected edges linking the vertices in $\mathcal{V}$.

Pearson's correlation coefficient and associated p -values are computed for each vertex pair's feature vectors x_v and x_w . These metrics elucidate the magnitude, direction, and statistical significance of the linear relationship between the data linked to each vertex pair. Following this calculation, we rank all vertex pairs using two criteria: the absolute correlation coefficient value (in descending order) and the p -value (in ascending order). We identify the top M pairs, those exhibiting the highest absolute correlation and lowest p -values, and assign them edges, $e_{vw} = (v, w) \in \mathcal{E}$. These assigned edges signify the most statistically significant and the strongest correlations within our population graph.

We introduced the proposed model, which employs GCNs and was termed GCN-corr, to this correlation-based population graph. Its performance was contrasted with an equivalently sized graph but with edges assigned randomly termed GCN-random. This comparison elucidated the advantages of utilizing a correlation-based graph for predicting $A\beta$ positivity in cognitively unimpaired individuals as opposed to a graph with randomly assigned edges. In the GCN-corr model, edges mirror the correlations between individual nodes, thus enabling more precise prediction congruent with the population structure.

Through an ablation study, we ascertained the optimal number of edges (M) by manipulating the network density. This density, calculated as $\frac{2M}{N(N-1)}$, represents the proportion of actual edges compared to the maximum possible in a fully connected network of N nodes. We varied the density from a sparse 1% connectivity network to a maximally interconnected network (100% connectivity), with each increment representing a 10% increase. This methodology aids in striking an equilibrium between network complexity and predictive accuracy, thereby refining the graph's structure for enhanced prediction efficacy using the GCN model. Additional discussions on the ablation study are expounded upon in the Supplementary Materials.

2.3. Graph Convolutional Networks

In this study, we leveraged graph convolutional networks (GCNs), which were proposed by [47], with a population graph $\mathcal{G}$ to discern the connections between the demographic and neuroimaging features of individuals. This study addresses the problem of predicting $A\beta$ positivity as a node (individual)-level prediction task within the GCN model. An adjacency matrix $A \in [0, 1]^{N \times N}$ representing pairwise correlations between nodes in a population graph $\mathcal{G}$, a feature matrix $X \in \mathbb{R}^{N \times p}$, and labels $y \in [0, 1]^N$ ($A\beta$ positivity) are used as input to train the model.

GCNs are designed to learn robust node representations by aggregating information from the local neighborhoods within the graph [47]. The core operation in GCNs, graph convolution, operates as a message-passing mechanism that facilitates information exchange between adjacent nodes. Within each GCN layer, nodes gather, process messages from their neighbors, and subsequently transform this information using a learnable weight matrix. Diverging from the convolution concept used in convolutional neural networks (CNNs), which apply a filter to localized input data segments, GCNs redefine convolution as the process of aggregation of neighboring node information. This notion retains the integral principle of incorporating local information, a concept fundamental to CNNs, thus justifying its appellation as graph convolution. Within a single GCN layer, the operation can be denoted as:

$$H^{(l+1)} = \sigma\left(\tilde{D}^{-\frac{1}{2}} \tilde{A} \tilde{D}^{-\frac{1}{2}} H^{(l)} W^{(l)}\right) \quad (1)$$

In this equation, $H^{(l)}$ and $H^{(l+1)}$ represent the feature matrices at the l -th and $(l + 1)$ -th layers, respectively, each encapsulating node representations. The matrix $\tilde{A}$, derived by adding the identity matrix I (representing self-loops, which are edges from nodes to themselves) to the original adjacency matrix A , is referred to as a self-loop inclusive adjacency matrix. The degree matrix of $\tilde{A}$, $\tilde{D}$, is a diagonal matrix representing the degree (number of connections) of each node. The learnable weight matrix at the l -th layer, $W^{(l)}$, serves to transform the aggregated neighbor node information. σ signifies the activation function; in this study, a rectified linear unit (ReLU).

The model is then trained to minimize the loss $\mathcal{L}$ using the ADAM optimizer and learn the optimal parameters for the prediction task. The loss $\mathcal{L}$ represents a cross-entropy function, which is computed as follows:

$$\mathcal{L} = -(w_{y_0} y \log(h) + w_{y_1} (1 - y) \log(1 - h)) \quad (2)$$

In this equation, h denotes the model's prediction output following softmax activation, and y represents the ground-truth label. Given the higher number of $A\beta$ -negative samples in the cognitively unimpaired individuals dataset, the model could be impacted by class imbalance. To address this issue, we assigned rescaling weights during model training, which are inversely proportional to class frequencies as follows:

$$w_{y_c} = \frac{N}{2N_{y_c}} \quad (3)$$

Here, N represents the total number of samples, while N_{y_c} denotes the number of samples belonging to class c .

2.4. Interpretation with GNNExplainer

GNNExplainer [53] is a model-agnostic method specially designed to elucidate the predictions made by graph neural networks (GNNs), including the GCN model. The primary goal of GNNExplainer is to offer insights into the model's decision-making process, highlight significant features and relationships within the graph, and foster trust in the model's predictions.

The GNNExplainer method works by learning to extract a concise subgraph from the original input graph. This subgraph is optimized to best explain the GNN model's prediction for a specific target node or graph. This is accomplished by formulating an optimization problem where the objective is to minimize the difference between the original GNN model's prediction and the prediction made using the extracted subgraph. This optimization process involves the use of a binary mask for nodes, edges, and features. This mask determines whether to include or exclude graph elements, contingent on their contribution to the prediction.

In this research, we employed GNNExplainer to provide interpretable explanations for the predictions generated by our GCN model. This tool allowed us to explore individual biomarkers that significantly contribute to the prediction of $A\beta$ positivity. Additionally, we identified distinct groups of individuals who share common biomarkers yet contain unique prioritized features that differentiate them from other groups. We also detailed the characteristics of each group based on their significant biomarkers.

2.5. Performance Evaluation

In this study, we formulated a supervised node classification problem, where our objective was to predict the $A\beta$ positivity for each individual within the test sets. We assessed the classification performance by employing stratified five-fold cross-validation repeated 10 times. This method ensured that the ratio of $A\beta$ -positive to $A\beta$ -negative samples was maintained across all sets. During each iteration, we set aside one fold as the test set, while the remaining four folds were randomly divided into an 80% training set and a 20% validation set. We carried out this stratified partitioning to optimize and validate our

model. Consequently, the data were split into a 64% training set, a 16% validation set, and a 20% test set. We report the final classification performance as the mean area under the curve (AUC) values over a total of 50 iterations.

3. Results

3.1. Experimental Setting

In the experiments, we assessed the performance of our proposed correlation-based population graph model (GCN-corr) and compared it with an equivalently sized graph but with edges assigned randomly (GCN-random). Furthermore, to provide a broader perspective on the efficacy of GCN models applied to population graphs, our comparison was not limited to GCN-random and GCN-corr. We expanded our analysis to include a comparison with traditional machine learning algorithms. These encompassed the support vector machine (SVM) using a radial basis function (RBF) kernel, the random forest (RF) classifier, logistic regression (LR) with the L2 penalty (also known as ridge regularization), and the multi-layer perceptron (MLP).

To achieve optimal performance in our analyses, we set certain hyperparameters empirically for both the GCN and the MLP models. The GCN model was structured as a two-layer network, which included a hidden layer comprising 32 units. The model underwent training for a maximum of 550 epochs, with a learning rate set at 0.005. The L2 loss was set to 5×10^{-10} , and a dropout rate of 0.5 was employed to prevent overfitting and enhance generalization. Simultaneously, the MLP model was set up with the same hyperparameters as the GCN model for a fair comparison.

Due to the predominance of $A\beta$ -negative samples in our dataset, we utilized class weights as outlined in Equation (3) to balance the classes. For fairness in comparison, the same rescaling weights were applied to the conventional machine learning models previously mentioned. This adjustment in weights ensured a balanced evaluation of each model's performance, despite the unequal sample sizes between the $A\beta$ -positive and $A\beta$ -negative groups.

It is worth noting that, although we employed stratified cross-validation, a technique which guarantees a proportionate class distribution across all validation folds, it did not directly influence the model's learning process. Conversely, class weights were incorporated during model training, conferring more emphasis to the less represented class and thereby ameliorating the effects of imbalanced data. Therefore, these two methodologies, while serving distinct roles, complement each other: stratified cross-validation augments the accuracy of performance estimation, while class weighting refines the model's capacity to learn from imbalanced datasets.

3.2. Performance of Prediction of $A\beta$ Positivity

We evaluated the performance in predicting $A\beta$ positivity with comparative analyses conducted across three distinct groups: cognitively normal (CN) individuals, individuals with mild cognitive impairment (MCI), and a combined group of both CN and MCI individuals. The performances for each model were evaluated across these sample groups and are comprehensively displayed in Table 2 and Figure 2.

Upon analysis, we observed that the GCN models, when applied to the population graph constructed in this study, consistently outperformed the conventional machine learning models across all sample groups. Moreover, the correlation-based population graph model (GCN-corr) consistently demonstrated superior performance compared to the GCN model with randomly assigned edges (GCN-random). The GCN-corr model outperformed all other models across all groups, achieving the highest mean AUCs of 0.8851, 0.8741, and 0.8632 for the CN, MCI, and CN + MCI groups, respectively. This was particularly evident in the combined CN + MCI group, where the performance of the GCN-random model showed a significant drop (0.7160 ± 0.0135), not just in comparison to the GCN-corr model but also against most of the conventional machine learning models. This suggests the potential of GCN models in significantly enhancing the predictive accuracy for

$A\beta$ positivity when used with well-crafted population graphs, affirming their proficiency in handling complex biological data.

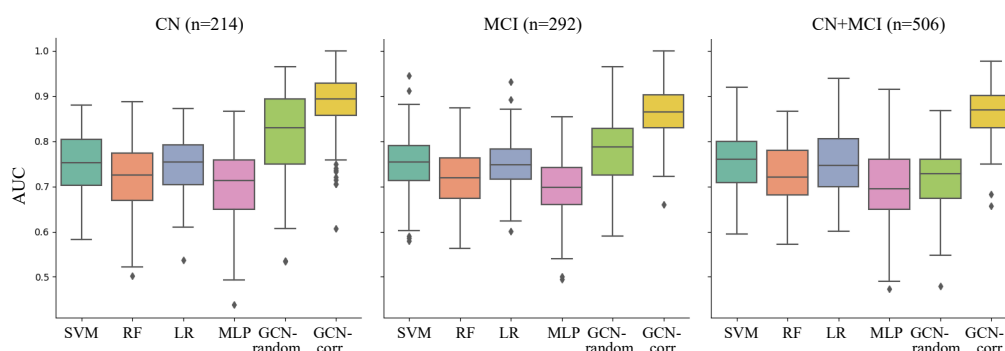


Figure 2. Comparison of the performance of each model (represented in x-axis and different colors) in predicting $A\beta$ positivity across three sample groups: cognitively normal (CN) individuals, those with mild cognitive impairment (MCI), and a combined group (CN + MCI). The performance was measured with the area under the ROC curve (AUC) derived from 10 repetitions of five-fold cross-validation. The compared models are support vector machine using a radial basis function kernel (SVM), random forest (RF) classifier, logistic regression with ridge regularization (LR), multi-layer perceptron (MLP), GCN model on a graph with randomly assigned edges (GCN-random), and GCN model on a correlation-based population graph (GCN-corr).

Table 2. Summary of the performance of each model in predicting $A\beta$ positivity across three sample groups: cognitively normal (CN) individuals, those with mild cognitive impairment (MCI), and a combined group (CN + MCI). The performance metrics were computed as the mean area under the ROC curve (AUC), along with a 95% confidence interval, derived from 10 repetitions of five-fold cross-validation. The highest performing result for each sample group is highlighted in bold text.

Model	AUC (Mean $\pm$ 95% CI)		
	CN	MCI	CN + MCI
SVM (RBF)	0.7515 $\pm$ 0.0131	0.7531 $\pm$ 0.0137	0.7537 $\pm$ 0.0129
RF	0.7205 $\pm$ 0.0143	0.7226 $\pm$ 0.0140	0.7238 $\pm$ 0.0134
LR (ridge)	0.7490 $\pm$ 0.0129	0.7480 $\pm$ 0.0112	0.7500 $\pm$ 0.0144
MLP	0.7009 $\pm$ 0.0158	0.7013 $\pm$ 0.0137	0.7009 $\pm$ 0.0159
GCN-random	0.8110 $\pm$ 0.0185	0.7768 $\pm$ 0.0153	0.7160 $\pm$ 0.0135
GCN-corr	0.8851 $\pm$ 0.0154	0.8741 $\pm$ 0.0114	0.8632 $\pm$ 0.0115

To further illustrate the proficiency of the GCN model in accurately classifying $A\beta$ -positive and -negative samples within a population graph, we offer a visualization of the final embedding of each GCN model in Figure 3. Notably, the GCN model with randomly assigned edges (GCN-random) demonstrated difficulty differentiating between the two classes across all three groups: cognitively normal (CN), mild cognitive impairment (MCI), and the combined CN + MCI group.

In contrast, the correlation-based population graph model (GCN-corr) effectively distinguished between the two classes across all sample groups, emphasizing the significant contribution of our proposed correlation-based population graph to the enhancement of $A\beta$ positivity prediction using the GCN model.

The strength of these findings lends considerable support to our research hypothesis. It suggests the existence of shared biomarkers within groups of individuals whose demographic and neuroimaging features strongly correlate, thus positively influencing $A\beta$ positivity prediction and enriching our understanding of AD prognosis. These results also suggest variability in AD risk factors across different individual groups, hinting at the potential benefits of tailoring prediction models and preventative strategies.

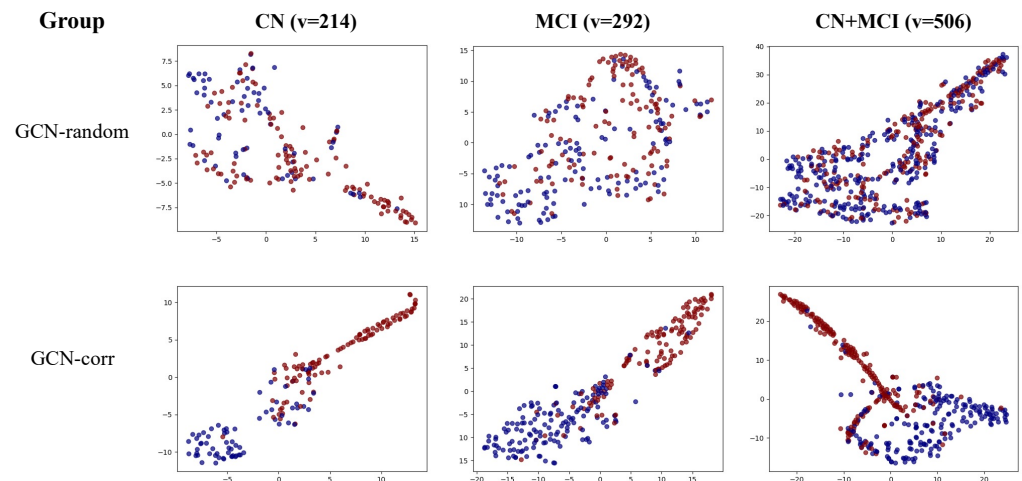


Figure 3. Visual representation of the final node embedding of the GCN-random and GCN-corr models across three groups: cognitively normal (CN), mild cognitive impairment (MCI), and a combined CN + MCI group. Nodes representing A β -positive samples are colored blue, while those representing A β -negative samples are in red. v represents the number of vertices (individuals).

3.3. Interpreting Predictions of Graph Neural Networks

In our endeavor to understand the intricate interplay between demographic and neuroimaging features, we utilized GNNExplainer. This tool helped us identify key biomarkers that significantly contribute to the prediction of A β positivity within GCN models. This investigation deepens our understanding of the model's decision-making process and bolsters our confidence in its predictive capabilities.

Figure 4 offers a visual representation of these prioritized biomarkers and their associated importance scores for each individual. For an improved visualization, we randomly selected 50 individuals from the test set that delivered the most accurate predictions across all cross-validation splits. Notably, these samples were correctly classified, which facilitated the identification of the most influential biomarkers for accurate predictions. To further enhance data interpretability, we employed both feature-wise and sample-wise clustering. This analysis led to the identification of four distinct groups (A, B, C, and D), each distinguished by the significance of their biomarkers. We delineated the top 10 biomarkers based on their averaged importance scores, uncovering unique patterns of significance across the four groups, as detailed in Table S1.

In group A, the left precentral gyrus, known for its involvement in motor function, emerged as the most significant biomarker with an average score of 0.9954. Additional notable biomarkers included the right precentral gyrus, the APOE ϵ 4 gene variant—associated with an increased risk of AD—and the left caudal middle frontal gyrus, all scoring above 0.89. In group B, the left precentral gyrus was also the top biomarker. However, the demographic feature of age was highlighted, receiving an average score of 0.7768. This underlines the well-established link between advancing age and increased risk of AD. In group C, the left precuneus, a brain region involved in episodic memory, stood out as the most significant biomarker with a score of 0.7137. This group also prioritized demographic features, such as the level of education (years) and the individual's sex, with scores of 0.6725 and 0.6160, respectively. These results may indicate a potential influence of educational attainment and biological sex on the disease's onset and progression. Group D highlighted the value of education (years) as the top biomarker, scoring a near-perfect 0.9970. This aligns with the cognitive reserve theory, which suggests that higher levels of education may offer a protective effect against cognitive decline. Neuroimaging biomarkers, such as the right precuneus and left pars orbitalis, also held high priority in this group.

underscores the efficacy of harnessing the inherent structure and correlations in the data to augment predictive accuracy.

Our findings comprehensively demonstrate that a population graph built on correlations rather than random assignment can significantly elevate the model's ability to discern patterns and predict accurately. This suggests that the interplay between demographic and neuroimaging features is not random; instead, it exhibits a specific correlation structure that is instrumental for the prediction task.

Moreover, the improved performance lends credence to the hypothesis of the existence of common biomarkers among groups of individuals who exhibit high correlations in their demographic and neuroimaging features. This implies that different groups may possess distinctive sets of biomarkers that are particularly indicative of A β positivity, thereby illuminating the heterogeneity of AD progression.

Thus, our research underscores the potential of utilizing correlation-based population graphs in tandem with GCN models for a more nuanced and effective AD prognosis. This approach could potentially inform the development of more personalized, precise therapeutic strategies.

Future research could build upon our findings and explore several promising directions. This could include enriching the GCN model by incorporating diverse data sources, such as genetic, proteomic, and lifestyle factors, alongside demographic and neuroimaging features, to enhance the predictive performance. Transitioning from a cross-sectional model to one that exploits longitudinal data could offer a more dynamic understanding of AD progression. A deeper exploration of model interpretability and biomarker validation could provide greater transparency for the decision-making process, ensuring the identified biomarkers' relevance. Understanding the unique contributions of biomarkers to AD progression in different groups could inform the development of personalized treatment strategies. Lastly, the proposed framework in this study could be expanded to other neurological disorders or complex diseases, potentially offering valuable insights into disease mechanisms and therapeutic targets.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/bioengineering10060701/s1>, Figure S1: Performances of the GCN-corr and GCN-random models as network density varies from a sparse 1% connectivity to a fully connected 100% connectivity in increments of 10%. Performance metrics were calculated as the mean area under the ROC curve (AUC), with a 95% confidence interval, derived from 10 repetitions of five-fold cross-validation. The confidence interval's upper and lower bounds are visually represented as a shaded area surrounding the line; Figure S2: Visual representation of the final node embedding of the GCN-random and GCN-corr models at selected network densities of 1%, 10%, and 50%. These densities were chosen to provide a clear comparison across varying levels of network densities. Nodes representing A β -positive samples are colored blue, while those representing A β -negative samples are in red; Table S1: The top 10 prioritized biomarkers from demographic and neuroimaging features, along with their corresponding averaged feature importance scores, listed in descending order. The biomarkers are divided into four groups (A–D) based on the results of heatmap clustering.

Funding: This work was supported by a National Research Foundation of Korea (NRF) grant funded by the Korean government (MSIT) (NRF-2022R1C1C1012060, NRF-2020R1I1A1A01071671).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data used in the preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu, accessed on 21 May 2023). As such, the investigators within the ADNI contributed to the design and implementation of the ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf (accessed on 21 May 2023).

Acknowledgments: Data collection and sharing for this project were funded by the Alzheimer’s Disease Neuroimaging Initiative (ADNI) (National Institutes of Health Grant U01 AG024904) and DOD ADNI (Department of Defense award number W81XWH-12-2-0012). ADNI is funded by the National Institute on Aging and the National Institute of Biomedical Imaging and Bioengineering and through generous contributions from the following: AbbVie, Alzheimer’s Association; Alzheimer’s Drug Discovery Foundation; Araclon Biotech; BioClinica, Inc.; Biogen; Bristol-Myers Squibb Company; CereSpir, Inc.; Cogstate; Eisai Inc.; Elan Pharmaceuticals, Inc.; Eli Lilly and Company; EuroImmun; F. Hoffmann-La Roche Ltd and its affiliated company Genentech, Inc.; Fujirebio; GE Healthcare; IXICO Ltd.; Janssen Alzheimer Immunotherapy Research & Development, LLC.; Johnson & Johnson Pharmaceutical Research & Development LLC.; Lumosity; Lundbeck; Merck & Co., Inc.; Meso Scale Diagnostics, LLC.; NeuroRx Research; Neurotrack Technologies; Novartis Pharmaceuticals Corporation; Pfizer Inc.; Piramal Imaging; Servier; Takeda Pharmaceutical Company; and Transition Therapeutics. The Canadian Institutes of Health Research provides funds to support ADNI clinical sites in Canada. Private sector contributions are facilitated by the Foundation for the National Institutes of Health (www.fnih.org, accessed on 21 May 2023). The grantee organization was the Northern California Institute for Research and Education, and the study was coordinated by the Alzheimer’s Therapeutic Research Institute at the University of Southern California. ADNI data are disseminated by the Laboratory for Neuro Imaging at the University of Southern California.

Conflicts of Interest: The authors declare no conflict of interest.

References

- Cullen, N.C.; Leuzy, A.; Janelidze, S.; Palmqvist, S.; Svenningsson, A.L.; Stomrud, E.; Dage, J.L.; Mattsson-Carlsson, N.; Hansson, O. Plasma biomarkers of Alzheimer’s disease improve prediction of cognitive decline in cognitively unimpaired elderly populations. *Nat. Commun.* **2021**, *12*, 3555. [\[CrossRef\]](#) [\[PubMed\]](#)
- Jeremic, D.; Jiménez-Díaz, L.; Navarro-López, J.D. Past, present and future of therapeutic strategies against amyloid- β peptides in Alzheimer’s disease: A systematic review. *Ageing Res. Rev.* **2021**, *72*, 101496. [\[CrossRef\]](#) [\[PubMed\]](#)
- Nakamura, A.; Kaneko, N.; Villemagne, V.L.; Kato, T.; Doecke, J.; Doré, V.; Fowler, C.; Li, Q.X.; Martins, R.; Rowe, C.; et al. High performance plasma amyloid- β biomarkers for Alzheimer’s disease. *Nature* **2018**, *554*, 249–254. [\[CrossRef\]](#) [\[PubMed\]](#)
- Selkoe, D.J.; Schenk, D. Alzheimer’s disease: Molecular understanding predicts amyloid-based therapeutics. *Annu. Rev. Pharmacol. Toxicol.* **2003**, *43*, 545–584. [\[CrossRef\]](#)
- Oddo, S.; Caccamo, A.; Kitazawa, M.; Tseng, B.P.; LaFerla, F.M. Amyloid deposition precedes tangle formation in a triple transgenic model of Alzheimer’s disease. *Neurobiol. Aging* **2003**, *24*, 1063–1070. [\[CrossRef\]](#)
- Shirbandi, K.; Khalafi, M.; Mirza-Aghazadeh-Attari, M.; Tahmasbi, M.; Shahvandi, H.K.; Javanmardi, P.; Rahim, F. Accuracy of deep learning model-assisted amyloid positron emission tomography scan in predicting Alzheimer’s disease: A systematic review and meta-analysis. *Inform. Med. Unlocked* **2021**, *25*, 100710. [\[CrossRef\]](#)
- Ezzati, A.; Harvey, D.J.; Habeck, C.; Golzar, A.; Qureshi, I.A.; Zammit, A.R.; Hyun, J.; Truelove-Hill, M.; Hall, C.B.; Davatzikos, C.; et al. Predicting amyloid- β levels in amnesic mild cognitive impairment using machine learning techniques. *J. Alzheimer’s Dis.* **2020**, *73*, 1211–1219. [\[CrossRef\]](#)
- Vergallo, A.; Mégret, L.; Lista, S.; Cavedo, E.; Zetterberg, H.; Blennow, K.; Vanmechelen, E.; De Vos, A.; Habert, M.O.; Potier, M.C.; et al. Plasma amyloid β 40/42 ratio predicts cerebral amyloidosis in cognitively normal individuals at risk for Alzheimer’s disease. *Alzheimer’s Dement.* **2019**, *15*, 764–775. [\[CrossRef\]](#)
- Kim, S.; Lee, P.; Oh, K.T.; Byun, M.S.; Yi, D.; Lee, J.H.; Kim, Y.K.; Ye, B.S.; Yun, M.J.; Lee, D.Y.; et al. Deep learning-based amyloid PET positivity classification model in the Alzheimer’s disease continuum by using 2-[18F] FDG PET. *EJNMMI Res.* **2021**, *11*, 56. [\[CrossRef\]](#)
- Chang, C.H.; Lin, C.H.; Lane, H.Y. Machine learning and novel biomarkers for the diagnosis of Alzheimer’s disease. *Int. J. Mol. Sci.* **2021**, *22*, 2761. [\[CrossRef\]](#)
- Basheera, S.; Ram, M.S.S. Deep learning based Alzheimer’s disease early diagnosis using T2w segmented gray matter MRI. *Int. J. Imaging Syst. Technol.* **2021**, *31*, 1692–1710. [\[CrossRef\]](#)
- Venugopalan, J.; Tong, L.; Hassanzadeh, H.R.; Wang, M.D. Multimodal deep learning models for early detection of Alzheimer’s disease stage. *Sci. Rep.* **2021**, *11*, 3254. [\[CrossRef\]](#) [\[PubMed\]](#)
- Houria, L.; Belkhamisa, N.; Cherfa, A.; Cherfa, Y. Multi-modality MRI for Alzheimer’s disease detection using deep learning. *Phys. Eng. Sci. Med.* **2022**, *45*, 1043–1053. [\[CrossRef\]](#)
- Ramzan, F.; Khan, M.U.G.; Rehmat, A.; Iqbal, S.; Saba, T.; Rehman, A.; Mehmood, Z. A deep learning approach for automated diagnosis and multi-class classification of Alzheimer’s disease stages using resting-state fMRI and residual neural networks. *J. Med. Syst.* **2020**, *44*, 37. [\[CrossRef\]](#)
- Zhang, T.; Liao, Q.; Zhang, D.; Zhang, C.; Yan, J.; Ngetich, R.; Zhang, J.; Jin, Z.; Li, L. Predicting MCI to AD conversion using integrated sMRI and rs-fMRI: machine learning and graph theory approach. *Front. Aging Neurosci.* **2021**, *13*, 688926. [\[CrossRef\]](#) [\[PubMed\]](#)

16. Grassi, M.; Rouleaux, N.; Caldirola, D.; Loewenstein, D.; Schruers, K.; Perna, G.; Dumontier, M.; Initiative, A.D.N. A novel ensemble-based machine learning algorithm to predict the conversion from mild cognitive impairment to Alzheimer's disease using socio-demographic characteristics, clinical information, and neuropsychological measures. *Front. Neurol.* **2019**, *10*, 756. [\[CrossRef\]](#) [\[PubMed\]](#)
17. Spasov, S.; Passamonti, L.; Duggento, A.; Lio, P.; Toschi, N.; Alzheimer's Disease Neuroimaging Initiative. A parameter-efficient deep learning approach to predict conversion from mild cognitive impairment to Alzheimer's disease. *Neuroimage* **2019**, *189*, 276–287. [\[CrossRef\]](#)
18. Lee, G.; Nho, K.; Kang, B.; Sohn, K.A.; Kim, D. Predicting Alzheimer's disease progression using multi-modal deep learning approach. *Sci. Rep.* **2019**, *9*, 1952. [\[CrossRef\]](#)
19. Ocasio, E.; Duong, T.Q. Deep learning prediction of mild cognitive impairment conversion to Alzheimer's disease at 3 years after diagnosis using longitudinal and whole-brain 3D MRI. *PeerJ Comput. Sci.* **2021**, *7*, e560. [\[CrossRef\]](#) [\[PubMed\]](#)
20. Li, J.; Bur, A.M.; Villwock, M.R.; Shankar, S.; Palmer, G.; Sykes, K.J.; Villwock, J.A. Olfactory phenotypes differentiate cognitively unimpaired seniors from Alzheimer's disease and mild cognitive impairment: A combined machine learning and traditional statistical approach. *J. Alzheimer's Dis.* **2021**, *81*, 641–650. [\[CrossRef\]](#)
21. Masters, C.L.; Selkoe, D.J. Biochemistry of amyloid β -protein and amyloid deposits in Alzheimer disease. *Cold Spring Harb. Perspect. Med.* **2012**, *2*, a006262. [\[CrossRef\]](#)
22. Engler, H.; Forsberg, A.; Almkvist, O.; Blomquist, G.; Larsson, E.; Savitcheva, I.; Wall, A.; Ringheim, A.; Långström, B.; Nordberg, A. Two-year follow-up of amyloid deposition in patients with Alzheimer's disease. *Brain* **2006**, *129*, 2856–2866. [\[CrossRef\]](#)
23. Leffa, D.T.; Ferrari-Souza, J.P.; Bellaver, B.; Tissot, C.; Ferreira, P.C.; Brum, W.S.; Caye, A.; Lord, J.; Proitsi, P.; Martins-Silva, T.; et al. Genetic risk for attention-deficit/hyperactivity disorder predicts cognitive decline and development of Alzheimer's disease pathophysiology in cognitively unimpaired older adults. *Mol. Psychiatry* **2023**, *28*, 1248–1255. [\[CrossRef\]](#)
24. Gottesman, R.F.; Schneider, A.L.; Zhou, Y.; Chen, X.; Green, E.; Gupta, N.; Knopman, D.S.; Mintz, A.; Rahmim, A.; Sharrett, A.R.; et al. The ARIC-PET amyloid imaging study: Brain amyloid differences by age, race, sex, and APOE. *Neurology* **2016**, *87*, 473–480. [\[CrossRef\]](#) [\[PubMed\]](#)
25. Wu, K.Y.; Hsiao, I.T.; Chen, C.S.; Chen, C.H.; Hsieh, C.J.; Wai, Y.Y.; Chang, C.J.; Tseng, H.J.; Yen, T.C.; Liu, C.Y.; et al. Increased brain amyloid deposition in patients with a lifetime history of major depression: Evidenced on 18 F-florbetapir (AV-45/Amyvid) positron emission tomography. *Eur. J. Nucl. Med. Mol. Imaging* **2014**, *41*, 714–722. [\[CrossRef\]](#)
26. Genin, E.; Hannequin, D.; Wallon, D.; Slegers, K.; Hiltunen, M.; Combarros, O.; Bullido, M.J.; Engelborghs, S.; De Deyn, P.; Berr, C.; et al. APOE and Alzheimer disease: A major gene with semi-dominant inheritance. *Mol. Psychiatry* **2011**, *16*, 903–907. [\[CrossRef\]](#) [\[PubMed\]](#)
27. Meyer, M.R.; Tschanz, J.T.; Norton, M.C.; Welsh-Bohmer, K.A.; Steffens, D.C.; Wyse, B.W.; Breitner, J. APOE genotype predicts when—not whether—one is predisposed to develop Alzheimer disease. *Nat. Genet.* **1998**, *19*, 321–322. [\[CrossRef\]](#) [\[PubMed\]](#)
28. Strittmatter, W.J.; Roses, A.D. Apolipoprotein E and Alzheimer disease. *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 4725–4727. [\[CrossRef\]](#) [\[PubMed\]](#)
29. Lerch, J.P.; Pruessner, J.C.; Zijdenbos, A.; Hampel, H.; Teipel, S.J.; Evans, A.C. Focal decline of cortical thickness in Alzheimer's disease identified by computational neuroanatomy. *Cereb. Cortex* **2005**, *15*, 995–1001. [\[CrossRef\]](#)
30. Lerch, J.P.; Pruessner, J.; Zijdenbos, A.P.; Collins, D.L.; Teipel, S.J.; Hampel, H.; Evans, A.C. Automated cortical thickness measurements from MRI can accurately separate Alzheimer's patients from normal elderly controls. *Neurobiol. Aging* **2008**, *29*, 23–30. [\[CrossRef\]](#)
31. Sun, Z.; Yin, H.; Chen, H.; Chen, T.; Cui, L.; Yang, F. Disease prediction via graph neural networks. *IEEE J. Biomed. Health Inform.* **2020**, *25*, 818–826. [\[CrossRef\]](#) [\[PubMed\]](#)
32. Cui, H.; Dai, W.; Zhu, Y.; Li, X.; He, L.; Yang, C. Brainnnexplainer: An interpretable graph neural network framework for brain network based disease analysis. *arXiv* **2021**, arXiv:2107.05097.
33. Zhang, X.M.; Liang, L.; Liu, L.; Tang, M.J. Graph neural networks and their current applications in bioinformatics. *Front. Genet.* **2021**, *12*, 690049. [\[CrossRef\]](#)
34. Bessadok, A.; Mahjoub, M.A.; Rekik, I. Graph neural networks in network neuroscience. *IEEE Trans. Pattern Anal. Mach. Intell.* **2022**, *45*, 5833–5848. [\[CrossRef\]](#)
35. Shen, Z.A.; Luo, T.; Zhou, Y.K.; Yu, H.; Du, P.F. NPI-GNN: Predicting ncRNA—Protein interactions with deep graph neural networks. *Briefings Bioinform.* **2021**, *22*, bbab051. [\[CrossRef\]](#)
36. Lv, G.; Hu, Z.; Bi, Y.; Zhang, S. Learning unknown from correlations: Graph neural network for inter-novel-protein interaction prediction. *arXiv* **2021**, arXiv:2105.06709.
37. Jha, K.; Saha, S.; Singh, H. Prediction of protein—Protein interaction using graph neural networks. *Sci. Rep.* **2022**, *12*, 8360. [\[CrossRef\]](#)
38. Parisot, S.; Ktena, S.I.; Ferrante, E.; Lee, M.; Guerrero, R.; Glocker, B.; Rueckert, D. Disease prediction using graph convolutional networks: Application to autism spectrum disorder and Alzheimer's disease. *Med. Image Anal.* **2018**, *48*, 117–130. [\[CrossRef\]](#)
39. Gu, Y.; Yang, X.; Tian, L.; Yang, H.; Lv, J.; Yang, C.; Wang, J.; Xi, J.; Kong, G.; Zhang, W. Structure-aware siamese graph neural networks for encounter-level patient similarity learning. *J. Biomed. Inform.* **2022**, *127*, 104027. [\[CrossRef\]](#) [\[PubMed\]](#)

40. Quan, Z.; Guo, Y.; Lin, X.; Wang, Z.J.; Zeng, X. Graphcpi: Graph neural representation learning for compound-protein interaction. In Proceedings of the 2019 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), San Diego, CA, USA, 18–21 November 2019; pp. 717–722.
41. Zhang, Z.; Chen, L.; Zhong, F.; Wang, D.; Jiang, J.; Zhang, S.; Jiang, H.; Zheng, M.; Li, X. Graph neural network approaches for drug-target interactions. *Curr. Opin. Struct. Biol.* **2022**, *73*, 102327. [[CrossRef](#)]
42. Qin, Z.; Liu, Z.; Zhu, P. Aiding Alzheimer’s Disease Diagnosis Using Graph Convolutional Networks Based on rs-fMRI Data. In Proceedings of the 2022 15th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI), Beijing, China, 5–7 November 2022; pp. 1–7.
43. Kolahkaj, S.; Zare, H. A connectome-based deep learning approach for Early MCI and MCI detection using structural brain networks. *Neurosci. Inform.* **2023**, *3*, 100118. [[CrossRef](#)]
44. Song, T.A.; Chowdhury, S.R.; Yang, F.; Jacobs, H.; El Fakhri, G.; Li, Q.; Johnson, K.; Dutta, J. Graph convolutional neural networks for Alzheimer’s disease classification. In Proceedings of the 2019 IEEE 16th international symposium on biomedical imaging (ISBI 2019), Venice, Italy, 8–11 April 2019; pp. 414–417.
45. Kim, M.; Kim, J.; Qu, J.; Huang, H.; Long, Q.; Sohn, K.A.; Kim, D.; Shen, L. Interpretable temporal graph neural network for prognostic prediction of Alzheimer’s disease using longitudinal neuroimaging data. In Proceedings of the 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Virtual, 9–12 December 2021; pp. 1381–1384.
46. Ma, J.; Zhu, X.; Yang, D.; Chen, J.; Wu, G. Attention-guided deep graph neural network for longitudinal Alzheimer’s disease analysis. In Proceedings of the Medical Image Computing and Computer Assisted Intervention—MICCAI 2020: 23rd International Conference, Lima, Peru, 4–8 October 2020; pp. 387–396.
47. Kipf, T.N.; Welling, M. Semi-supervised classification with graph convolutional networks. *arXiv* **2016**, arXiv:1609.02907.
48. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-cam: Visual explanations from deep networks via gradient-based localization. In Proceedings of the IEEE International Conference on Computer Vision, Venice, Italy, 22–29 October 2017; pp. 618–626.
49. Pope, P.E.; Kolouri, S.; Rostami, M.; Martin, C.E.; Hoffmann, H. Explainability methods for graph convolutional neural networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Long Beach, CA, USA, 15–20 June 2019; pp. 10772–10781.
50. Luo, D.; Cheng, W.; Xu, D.; Yu, W.; Zong, B.; Chen, H.; Zhang, X. Parameterized explainer for graph neural network. *Adv. Neural Inf. Process. Syst.* **2020**, *33*, 19620–19631.
51. Vu, M.; Thai, M.T. Pgm-explainer: Probabilistic graphical model explanations for graph neural networks. *Adv. Neural Inf. Process. Syst.* **2020**, *33*, 12225–12235.
52. Yuan, H.; Tang, J.; Hu, X.; Ji, S. Xgmn: Towards model-level explanations of graph neural networks. In Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Virtual, 23–27 August 2020; pp. 430–438.
53. Ying, Z.; Bourgeois, D.; You, J.; Zitnik, M.; Leskovec, J. Gnnexplainer: Generating explanations for graph neural networks. *Adv. Neural Inf. Process. Syst.* **2019**, *32*, 9240–9251.
54. Van der Velden, B.H.; Kuijff, H.J.; Gilhuijs, K.G.; Viergever, M.A. Explainable artificial intelligence (XAI) in deep learning-based medical image analysis. *Med. Image Anal.* **2022**, *79*, 102470. [[CrossRef](#)] [[PubMed](#)]
55. Singh, A.; Sengupta, S.; Lakshminarayanan, V. Explainable deep learning models in medical image analysis. *J. Imaging* **2020**, *6*, 52. [[CrossRef](#)]
56. Borys, K.; Schmitt, Y.A.; Nauta, M.; Seifert, C.; Krämer, N.; Friedrich, C.M.; Nensa, F. Explainable AI in Medical Imaging: An overview for clinical practitioners—Saliency-based XAI approaches. *Eur. J. Radiol.* **2023**, *162*, 110787. [[CrossRef](#)]
57. Jiménez-Luna, J.; Grisoni, F.; Schneider, G. Drug discovery with explainable artificial intelligence. *Nat. Mach. Intell.* **2020**, *2*, 573–584. [[CrossRef](#)]
58. Preuer, K.; Klambauer, G.; Rippmann, F.; Hochreiter, S.; Unterthiner, T. Interpretable deep learning in drug discovery. In *Explainable AI: Interpreting, Explaining and Visualizing Deep Learning*; Springer: Cham, Switzerland, 2019; pp. 331–345.
59. Dey, S.; Luo, H.; Fokoue, A.; Hu, J.; Zhang, P. Predicting adverse drug reactions through interpretable deep learning framework. *BMC Bioinform.* **2018**, *19*, 476. [[CrossRef](#)]
60. Karimi, M.; Wu, D.; Wang, Z.; Shen, Y. DeepAffinity: Interpretable deep learning of compound–protein affinity through unified recurrent and convolutional neural networks. *Bioinformatics* **2019**, *35*, 3329–3338. [[CrossRef](#)] [[PubMed](#)]
61. Clark, C.M.; Schneider, J.A.; Bedell, B.J.; Beach, T.G.; Bilker, W.B.; Mintun, M.A.; Pontecorvo, M.J.; Hefti, F.; Carpenter, A.P.; Flitter, M.L.; et al. Use of florbetapir-PET for imaging β -amyloid pathology. *JAMA* **2011**, *305*, 275–283. [[CrossRef](#)] [[PubMed](#)]
62. Joshi, A.D.; Pontecorvo, M.J.; Clark, C.M.; Carpenter, A.P.; Jennings, D.L.; Sadowsky, C.H.; Adler, L.P.; Kovnat, K.D.; Seibyl, J.P.; Arora, A.; et al. Performance characteristics of amyloid PET with florbetapir F 18 in patients with Alzheimer’s disease and cognitively normal subjects. *J. Nucl. Med.* **2012**, *53*, 378–384. [[CrossRef](#)] [[PubMed](#)]

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.