2024 46th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) | 979-8-3503-7149-9/24/$31.00 ©2024 IEEE | DOI: 10.1109/EMBC53108.2024.10781747

Novel Alzheimer's Disease Stating Based on Comorbidities-Informed Graph Neural Networks

1st Ferial Abuhantash
Biomedical Eng. & Biotechnology
Khalifa University
AbuDhabi, UAE

2nd Mohd Khalil Abu Hantash
Biomedical Eng. & Biotechnology
Khalifa University
AbuDhabi, UAE

3rd Roy Welsch
Sloan School of Management and Statistics
Massachusetts Institute of Technology
Cambridge, Massachusetts, USA

4th Mohamed Lamine Seghier
Biomedical Eng. & Biotechnology
Khalifa University
AbuDhabi, UAE

5th Leontios Hadjileontiadis
Biomedical Eng. & Biotechnology
Khalifa University
AbuDhabi, UAE

6th Aamna Al Shehhi
Biomedical Eng. & Biotechnology
Khalifa University
AbuDhabi, UAE
aamna.alshehhi@ku.ac.ae

***Abstract*—Alzheimer's Disease (AD), the most prevalent form of dementia, requires early prediction for timely intervention. Leveraging data from the Alzheimer's Disease Neuroimaging Initiative (ADNI), our study employs Graph Neural Networks (GNNs) for multi-class AD classification. Initial steps involve creating a patient-clinical graph network considering latent relationships among cognitive normal (CN), mild cognitive impairment (MCI), and AD patients, followed by training several GNN-based techniques for building prediction models. Incorporating comorbidity data from electronic health records into the feature set yielded the most effective classification results. Notably, the GNN model with attention mechanisms outperforms state-of-the-art techniques in multi-class AD classification, achieving an accuracy = 0.92 [0.91,0.94], AUC = 0.96 [0.95,0.96], and F1-score = 0.92 [0.91,0.94]. This work highlights comorbidity data's impact on AD classification and suggests its potential to deepen disease understanding.**

***Index Terms*—Alzheimer's Disease, Multi-class Classification, Clinical Data, Comorbidity**

I. Introduction

Alzheimer's disease (AD) is a progressive neurological disorder characterized by memory loss, cognitive decline, and behavioral changes, ultimately leading to severe impairment in daily functioning [1]. According to the world health organization (WHO), each year, about 10 million new cases of dementia emerge among the 55 million people affected worldwide, with AD making up 60–70% of these cases [2]. Despite the absence of a cure, accurately identifying an individual's cognitive state can significantly aid in slowing the progression of Alzheimer's disease by facilitating lifestyle adjustments [3]. Detecting mild cognitive impairment (MCI), the prodromal stage of AD presents a significant challenge due to its subtle manifestations. The use of machine learning (ML) techniques for Alzheimer's prediction has been the subject of extensive research [4]. The majority of the research relies on two modalities: magnetic resonance imaging (MRI) and positron emission tomography (PET), since they are also used by physicians to diagnose Alzheimer's disease [5]. The existing literature often overlooked the potential of non-imaging clinical data, particularly data derived from routinely collected electronic health records (EHR), as a valuable, cost-effective, and non-invasive source for improving prediction of AD dementia [6]. The confidentiality constraints in medical records should be addressed by robust anonymization techniques, ethical compliance, and encouraging shared, anonymous EHR data for unbiased research access [6]. AD prediction is challenging due to shared risk factors with other comorbidities; hence, transforming comorbidity/EHR data into low-dimensional vectors is a common approach, but latent relationships between AD, comorbidities, and patients may impact prediction accuracy. To address this issue, Social Network Analysis (SNA) has been applied to healthcare data where they utilized network features extracted from the patient network, combined with individual patient characteristics, in machine learning algorithms for predicting the risk of chronic diseases [7]. Many of these studies use basic ML algorithms like logistic regression and random forests, and require redevelopment and feature recalculation when new data is introduced. As Graph Neural Network (GNN)-based approaches can seamlessly incorporate new data without the need for such recalculations, this makes GNN a very attractive alternative for complex multi-classification problems on clinical data. Employing Graph Neural Networks for Alzheimer's Disease (AD) classification offers a unique advantage by integrating patient-specific information and clinical test data within a graph structure.

In this study, we present a novel classification framework based on GNNs that utilizes complete clinical data, encompassing both imaging and non-imaging features, for the identification of individuals across multiple classes (cognitive normal, MCI, and AD). Unlike state-of-the-art methods that often depend on resource-intensive binary classification, our proposed approach excels in multi-class classification.This work contributes by exploring the utilization of valuable comorbidity data extracted from electronic health records as input to a patient GNN. The investigation showcases excep-

This work is supported by Khalifa University under Award no. FSU-2021-005.

TABLE I: Subject Demographic

	CN	MCI	AD	p
N	122	258	111	-
Age (mean (SD))	75.9 (4.7)	75.1 (7.2)	76.1 (7.4)	0.027
Male (%)	49.2	65.2	47.3	< 0.001*
Race-White (%)	90.2	91.7	83.0	<0.001*
Married (%)	67.4	76.2	82.5	<0.001*
Education years(mean (SD))	15.7 (3.2)	15.7 (2.8)	14.9 (2.9)	<0.001*
Ethnicity-not Hispanic(%)	98.9	96.3	98.0	<0.001*

* p<=0.05, ANOVA, Kruskal–Wallis test, or x^2 test, as appropriate.

tional diagnostic accuracy achieved by our proposed GNN framework, particularly when incorporating comorbidities.

II. Materials & Methods

A. Data Seletction and Prepatation

Data in this work come from the widely used Alzheimer's Disease Neuroimaging Initiative (ADNI), a unique resource for longitudinal AD research with unrestricted access to diverse data modalities [8]. In this study, all 491 ADNI participants with prognosis of CN, MCI, or AD and available medical history are included. The limited number is due to the availability of medical history data. The participants demographics are summarized in Table I. The clinical data from ADNI has a heterogeneous collection of information, including demographics (6 features), cognitive test scores (15 features), comorbidities (28 features), and brain imaging features (14 features). The ADNI clinical dataset has missing values; thus, any feature with over 50% missing values was removed (7 imaging features-likely due to incomplete imaging procedures). Given the dataset's simplicity and interpretive needs, one-hot encoding is employed for all categorical features, avoiding unnecessary complexity. Stratified k-fold is used to split the data into training (80%) and testing (20%) sets. Multiple imputation by chained equations *(mice)* [9] is employed to handle missing data. Subsequently, Min-Max scaling is applied to normalize the data. Both imputation and normalization are performed on the training set and then replicated on the test set using the prediction matrix derived from the training set. To address imbalances in class distribution, Synthetic Minority Over-sampling Technique (SMOTE) is applied to training data [10].

B. GNN Implementation

GNNs excel in processing graph-structured data, offering a promising approach for disease classification. This work was implemented using python packages: PyTorch-Geomteric (PyG) [11] and Scikit-learn [12]. The graph structure has nodes, each representing participants with node features reflecting their demographic details and comorbidity data derived from EHRs. The edges in the graph denote patient similarity derived from clinical tests (15 cognitive features and 7 brain imaging features). Node features encapsulate individual traits, while edge features convey clinician-administered test outcomes. An edge links two nodes if there are at least 10 edge features within a 20% distance of each other. This selected cut-off, a hyperparameter, was determined through optimization in the range of 5-20 edge features and 60-90% similarity, as there is no systematic method providing an optimal threshold value for maximizing classification output.

We employed graph convolution network (GCN), the ChebConv network (CCN), and Graph Attention Network (GAT). GCN uses trainable weights to learn features between nodes and their neighbors and transform the current node state (x) into a final node state (x_n) called "node embedding". The baseline model architecture consisted of a GCN that is followed with a ReLU activation function, drop out layer and finally another GCN layer to produce the final node embedding. The GCN layer is defined as [13]:

$$x_v^{(l+1)} = W^{(l+1)} \sum_{w \in N(v) \cup \{v\}} \frac{1}{c_{w,v}} . x_w^{(l)} \tag{1}$$

where W is the trainable weight matrix, x is the node features and finally C is a fixed normalization coefficient. A variant of GCNs, ChebConv [14], employs the Chebyshev polynomials to accomplish a type of spectral convolution that is computationally more effective. Graph Attention Networks (GAT) are GNNs that introduce attention mechanisms to weigh the importance of neighboring nodes when aggregating information. The mathematical formulation of GAT is as follows [15]:

$$h_i' = \sigma \left(\sum_{j \in \mathcal{N}(i)} \alpha_{ij} W h_j \right) \tag{2}$$

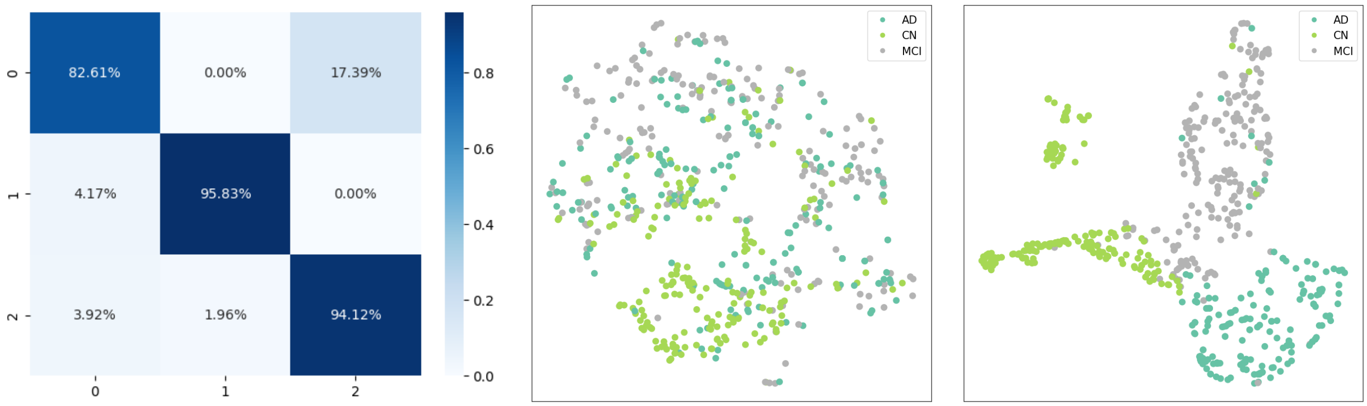
$$\alpha_{ij} = \frac{\exp\left(\text{LeakyReLU}(\vec{a}^T [W h_i || W h_j])\right)}{\sum_{k \in \mathcal{N}(i)} \exp\left(\text{LeakyReLU}(\vec{a}^T [W h_i || W h_k])\right)} \tag{3}$$

Here, h_i and h_j represent the node feature vectors of nodes i and j, respectively. The attention coefficients α_{ij} are computed based on the concatenation of the transformed feature vectors Wh_i and Wh_j. The attention mechanism is parameterized by learnable weights W and $\vec{a}$.

To enhance GNN performance, the hyperparameter optimization utilized the Tree-structured Parzen Estimator (TPE) sampler algorithm [16]. The optimized parameters included the learning rate, dropout rate, and the number of nodes in a layer. Adam optimization [17] and cross-entropy loss were consistently applied across all models.

C. Ablation Studies & Evaluation

Ablation studies in GNNs involve systematically altering node and edge features to understand their impact on model performance. We performed four ablation studies to analyze the impact of changing network features. The original setup has node features of demographics and comorbidities, and edge features of cognitive scores and imaging features. In Ablations 1, 2, and 3, we eliminate comorbidity features, cognitive scores, and imaging features, respectively. Lastly,



(a) (b) (c)

Fig. 1: Graph attention model performance visualized through (a) confusion matrix of multiclass classification for labels 0 (CN), 1 (MCI) and 2(AD). and through t-SNE visualization of node representation of (a) graph inputs, (b) graph outputs

in ablation 4, we swap the node and edge features to test our network design. To assess the classification performance we utilized three metrics: Area Under the Receiver Operating Characteristic curve (AUC-ROC), accuracy (ACC), and F1-score.

III. Results and Discussion

Multi-class classification poses a significant challenge in Alzheimer's subject classification. Despite satisfactory accuracy achieved in binary classification [18], extending it to multi-class remains a challenge. The 10-fold cross-validation results for multi-class classification of our GNN-based models against the well-known Multilayer Perceptron (MLP) artificial neural network are presented in Table II. The GNN-based models consistently outperform the MLP model, demonstrating their superior ability to handle clinical data in multi-class classification tasks. The GAT model achieves the highest accuracy (ACC = 0.92, AUC = 0.96, F1-score = 0.92) when using all clinical features. Ablation studies reveal that removing comorbidities leads to the poorest performances across various models, highlighting the high impact and relevance of comorbidities in node classification tasks. Utilizing comorbidities holds the potential to identify risk factors for future development of Alzheimer's disease in individuals currently classified as normal. Additionally, although the treatment of these comorbidities will not undo the cognitive dysfunction, they should be recognized as co-factors and treated immediately. This information could prove valuable in clinical settings, allowing for a focus on preventive care based on patients' medical history before resorting to costly modalities like MRI and PET data. When cognitive scores or imaging features are removed from the edge features, there is a noticeable decrease in performance, albeit not as significant as when manipulating the node features. As expected, the impact of removing cognitive scores is more pronounced compared to removing imaging data. This observation underscores a significant point – non-imaging clinical data alone can yield high accuracies (ACC = 0.91) without necessitating costly procedures like MRIs; however, it is noteworthy that the availability of only 14 imaging features may have limited the inclusion of relevant features initially. Despite utilizing all features in the node-edge swap ablation, the GNNs exhibit reduced performance (best model CCN with acc=0.79, auc=0.72, F1=0.79) compared to the original study. This decline is attributed to the network design change, where representing patients as edges between nodes of clinical tests is deemed less meaningful to both us and the models. This outcome may be attributed to the raw nature of cognitive scores and imaging data, which are not presented as a similarity matrix. The attention mechanism in GAT and its superior capability in capturing intricate graph patterns, likely contributed to it being the best-performing GNN-based model. The confusion matrix for the model reveals strong performance as seen in Fig. 1a. The diagonal elements indicate high percentages for each class: 82.6% for the first class, 95.8% for the second, and 94.1% for the third. Minimal off-diagonal values suggest effective classification across all classes. The t-SNE visualization provides valuable insights into the distribution of nodes before and after the training process, Fig. 1b and Fig. 1c respectively. Before training, nodes are scattered in a high-dimensional space based on their initial features. As the GAT undergoes training, the t-SNE visualization shows the model's ability to cluster nodes in a lower-dimensional space, highlighting the network's capacity to capture meaningful relationships between nodes. The visualization post-training demonstrates a more structured arrangement of nodes, reflecting the enhanced representation learning achieved by the GAT.

GNNs have been applied to ADNI dataset for binary classification tasks [19], [20]. However, for multi-class classifications with ADNI data, to the best of our knowledge no study applied GNNs with the same data modalities. Other studies have achieved high accuracy's using the ADNI dataset for the same task of 3-way classification: One study achieved an accuracy of 89.4% using MRI images with 3D convolutional neural networks [21] and another study achieved an accuracy of 79.8% using MRI images and clinical features with a KNN classifier [18]. Notably, our work surpasses these study,

TABLE II: Model Performance Comparison Across Four Ablation Studies with Bold Values Indicating the Highest Performance

Features	Model*	CN vs. MCI vs. AD		
		ACC	AUC	F1-Score
All features	MLP	0.53 [0.51,0.56]	0.72 [0.71,0.74]	0.53 [0.50,0.55]
	GCN	0.90 [0.88,0.91]	0.91 [0.89,0.93]	0.90 [0.88,0.91]
	CCN	0.68 [0.66,0.69]	0.85 [0.84,0.85]	0.64, [0.62, 0.67]
	GAT	**0.92 [0.91,0.94]**	**0.96 [0.95,0.96]**	**0.92 [0.91,0.94]**
Ablation 1 Without Comorbidity	MLP	0.52 [0.51,0.54]	0.68 [0.66,0.70]	0.49 [0.47,0.51]
	GCN	0.63 [0.60,0.66]	0.82 [0.81,0.83]	0.58 [0.52,0.63]
	CCN	**0.64 [0.63,0.64]**	**0.80 [0.78,0.81]**	**0.61 [0.60,0.63]**
	GAT	0.55 [0.51,0.59]	0.79 [0.78,0.81]	0.58[0.54,0.62]
Ablation 2 Without Cognitive Scores	MLP	0.55 [0.53,0.57]	0.72 [0.71,0.73]	0.55 [0.53,0.57]
	GCN	0.81 [0.78,0.83]	0.87 [0.85,0.88]	0.81 [0.79,0.83]
	CCN	0.60 [0.58,61]	0.77 [0.76,0.78]	0.60 [0.57,61]
	GAT	**0.82 [0.80,0.84]**	**0.92 [0.90,0.93]**	**0.82 [0.80,0.84]**
Ablation 3 Without Imaging	MLP	0.55 [0.53,0.57]	0.74 [0.73,0.75]	0.54 [0.52,0.57]
	GCN	0.90 [0.88,0.92]	0.92 [0.91,0.93]	0.90 [0.88,0.92]
	CCN	0.82 [0.80,0.84]	0.90 [0.90,0.91]	0.82 [0.81,0.83]
	GAT	**0.91 [0.90,0.92]**	**0.96 [0.95,0.97]**	**0.91 [0.90,0.92]**
Ablation 4 Node-Edge swap	MLP	0.64 [0.60,0.67]	0.60 [0.58,0.61]	0.63 [0.58,0.67]
	GCN	0.60 [0.60,0.60]	0.72 [0.72,0.72]	0.44 [0.44,0.44]
	CCN	**0.79 [0.77,0.80]**	**0.72 [0.71,0.75]**	**0.79 [0.77,0.80]**
	GAT	0.79 [0.79,0.79]	0.67 [0.67,0.67]	0.63 [0.63,0.63]

*MLP: Multilayer Perceptron, GCN: Graph Convolution Network, CCN: ChebConv Network, GAT: Graph Attention Network.

achieving an accuracy of 92% with the use of the clinical data through the GNN architecture.

IV. Conclusion

In conclusion, this paper proposes a GNN-based approach for multi-class classification of AD, CN, and MCI using clinical data. Results show that including comorbidity data in the feature set achieves comparable performance to top models, surpassing the inclusion of imaging data. A comprehensive feature set yields the best overall performance, aligning with prior research on the potential of longitudinal clinical EHR data in monitoring AD progression. Future work aims to incorporate longitudinal clinical data as input to GNN models and explore alternative architectures.

References

[1] R. J. Castellani, R. K. Rolston, and M. A. Smith, "Alzheimer disease," *Disease-a-month: DM*, vol. 56, no. 9, p. 484, 2010.

[2] "Dementia," https://www.who.int/news-room/fact-sheets/detail/dementia, accessed: 2023-12-11.

[3] D. M. Holtzman, J. C. Morris, and A. M. Goate, "Alzheimer's disease: the challenge of the second century," *Science translational medicine*, vol. 3, no. 77, pp. 77sr1–77sr1, 2011.

[4] M. Tanveer, B. Richhariya, R. U. Khan, A. H. Rashid, P. Khanna, M. Prasad, and C. Lin, "Machine learning techniques for the diagnosis of alzheimer's disease: A review," *ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM)*, vol. 16, no. 1s, pp. 1–35, 2020.

[5] A. D. Arya, S. S. Verma, P. Chakarabarti, T. Chakrabarti, A. A. Elngar, A.-M. Kamali, and M. Nami, "A systematic review on machine learning and deep learning techniques in the effective diagnosis of alzheimer's disease," *Brain Informatics*, vol. 10, no. 1, p. 17, 2023.

[6] S. Kumar, I. Oh, S. Schindler, A. M. Lai, P. R. Payne, and A. Gupta, "Machine learning for modeling the progression of alzheimer disease dementia using clinical data: a systematic literature review," *JAMIA open*, vol. 4, no. 3, p. ooab052, 2021.

[7] S. Karim, S. Uddin, T. Imam, and M. A. Moni, "A systematic review of network studies based on administrative health data," *International Journal of Environmental Research and Public Health*, vol. 17, no. 7, p. 2568, 2020.

[8] S. G. Mueller, M. W. Weiner, L. J. Thal, R. C. Petersen, C. R. Jack, W. Jagust, J. Q. Trojanowski, A. W. Toga, and L. Beckett, "Ways toward an early diagnosis in alzheimer's disease: the alzheimer's disease neuroimaging initiative (adni)," *Alzheimer's & Dementia*, vol. 1, no. 1, pp. 55–66, 2005.

[9] S. Van Buuren and C. G. Oudshoorn, "Multivariate imputation by chained equations," 2000.

[10] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "Smote: synthetic minority over-sampling technique," *Journal of artificial intelligence research*, vol. 16, pp. 321–357, 2002.

[11] M. Fey and J. E. Lenssen, "Fast graph representation learning with pytorch geometric," *arXiv preprint arXiv:1903.02428*, 2019.

[12] O. Kramer and O. Kramer, "Scikit-learn," *Machine learning for evolution strategies*, pp. 45–53, 2016.

[13] T. N. Kipf and M. Welling, "Semi-Supervised Classification with Graph Convolutional Networks," in *Proceedings of the 5th International Conference on Learning Representations*, ser. ICLR '17, 2017. [Online]. Available: https://openreview.net/forum?id=SJU4ayYgl

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

[15] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Lio, and Y. Bengio, "Graph attention networks," *arXiv preprint arXiv:1710.10903*, 2017.

[16] J. Bergstra, R. Bardenet, Y. Bengio, and B. Kégl, "Algorithms for hyper-parameter optimization," vol. 24, 2011.

[17] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[18] T. Altaf, S. M. Anwar, N. Gul, M. N. Majeed, and M. Majid, "Multi-class alzheimer's disease classification using image and clinical features," *Biomedical Signal Processing and Control*, vol. 43, pp. 64–74, 2018.

[19] S. Parisot, S. I. Ktena, E. Ferrante, M. Lee, R. Guerrero, B. Glocker, and D. Rueckert, "Disease prediction using graph convolutional networks: application to autism spectrum disorder and alzheimer's disease," *Medical image analysis*, vol. 48, pp. 117–130, 2018.

[20] Q. Li, X. Wu, L. Xu, K. Chen, L. Yao, and A. D. N. Initiative, "Classification of alzheimer's disease, mild cognitive impairment, and cognitively unimpaired individuals using multi-feature kernel discriminant dictionary learning," *Frontiers in computational neuroscience*, vol. 11, p. 117, 2018.

[21] A. Payan and G. Montana, "Predicting alzheimer's disease: a neuroimaging study with 3d convolutional neural networks," *arXiv preprint arXiv:1502.02506*, 2015.