Developing a Hybrid Deep Learning Model Combining CNNs and Graph Neural Networks for Enhanced Glaucoma Detection in OCT Images by Integrating Local and Global Context.

Mark Josh N. Alvear*[1], Raudmon Yvhan Fermin*[2], Nigel Jan L. Naniong*[3], Andrei Christian L. Tonido*[4]

*Department of Computer Science, University of Santo Tomas-Manila,
España Blvd, Sampaloc, Manila, 1008 Metro Manila, Philippines

[1]markjosh.alvear.cics@ust.edu.ph
[2]raudmonyvhan.fermin.cics@ust.edu.ph
[3]nigeljan.naniong.cics@ust.edu.ph
[4]andreichristian.tonido.cics@ust.edu.ph

***Abstract*—Glaucoma remains one of the most significant risk factors for uncorrectable, irreversible blindness. Thus, accurate and timely detection is required in order to achieve effective intervention in glaucoma. OCT-based imaging is of prime importance to detect glaucoma by its high-resolution retinal images, including images of the optic nerve head. Convolutional neural networks have demonstrated success in capturing local features and often lack capabilities to describe essential global spatial features in an accurate diagnosis of glaucoma. This study proposes a novel hybrid deep learning model that combines CNNs and Graph Neural Networks (GNNs) to leverage both local and global contextual information from OCT images. The CNN extracts local features, while the GNN models the spatial relationships between different regions of the ONH, enhancing the model's ability to detect subtle structural changes associated with glaucoma. This proposed approach is towards increasing the diagnostic accuracy, robustness, and efficiency of diagnosing glaucoma, leading to earlier detection and better management.**

***Keywords*— Glaucoma Detection, Optical Coherence Tomography (OCT), Convolutional Neural Networks (CNNs), Graph Neural Networks (GNNs), Hybrid Deep Learning, Medical Image Analysis**

I. INTRODUCTION

Glaucoma is a progressive optic nerve disorder leading to irreversible loss of vision, if not identified and treated at an early stage. It has been estimated that millions of people worldwide are prone to developing glaucoma. Therefore, an early and accurate diagnosis is considered to be an important intervention measure [1]. Optical Coherence Tomography has emerged as the cornerstone in glaucoma diagnosis, providing cross-sectional images with high resolution of the retina and optic nerve head (ONH) [9]. These images can now enable clinicians to appreciate fine morphological changes associated with the disease, such as thinning of the RNFL and optic disc cupping. The subtle nature and complexity, however, make it challenging for conventional diagnostic techniques.

Convolutional neural networks (CNNs) have shown promise for medical image analysis, such as OCT scans, as they are capable of local feature extraction including texture and edges [2]. CNNs fail to capture many of the important global spatial dependencies between the regions of the ONH for appropriate glaucoma diagnosis in many cases [4]. For example, variations in RNFL thickness between different sites may correlate with the degree of optic disc cupping, a correlation that CNNs alone may not easily represent.

In an effort to bypass this limitation, the current work presents a novel hybrid deep learning model that merges CNNs with Graph Neural Networks (GNNs). Here, the CNN would extract the local features of OCT images and GNN model spatial relationships across the different regions of the ONH, hence capturing the global context often overlooked by CNN alone [7]. The integrated model would involve local and global information to enhance the detection of accuracy and robustness into glaucoma diagnosis, leading to earlier disease detection and better management. To come up with the hybrid deep learning model by merging CNNs with GNNs for the automatic detection of glaucoma using OCT images to test the capabilities of the introduced model in gathering local and global contextual information as well as for the comparison between the hybrid and CNNs-GNNs with a standalone based on accuracy, sensitivity, and specificity, then analyze the strength of the developed hybrid model as an approach in improving the possibility of early glaucoma diagnosis and management.

II. LITERATURE SURVEY

Title of Study (Year)	Research Problem	Implemented Methodology	Study's Contribution to Computing	Gaps/Recommendations in their Study	Relevance to Proposed Study
Convolutional Neural Networks in Medical Image Understanding (2021)	CNNs' limitations in capturing global spatial relationships in medical images.	CNN-based models for segmentation and classification	Highlighted CNNs' strengths in local feature extraction and limitations in global context.	Did not address the need for global context in glaucoma detection.	Highlights the need for global context, which our hybrid model addresses.
A Feature Agnostic Approach for Glaucoma Detection in OCT Volumes (2019)	CNNs' inability to capture spatial relationships in OCT images.	3D CNNs trained on raw OCT volumes	Reduced reliance on manual annotation and improved scalability.	Missed crucial spatial relationships in OCT images.	Our study addresses this gap by incorporating GNNs for global context.
Assessing the Efficacy of 2D and 3D CNN Algorithms in OCT-Based Glaucoma Detection (2024)	CNNs' struggle to model complex relationships between anatomical regions.	Compared 2D and 3D CNNs for glaucoma detection	Showed 3D CNNs' effectiveness in leveraging volumetric data.	Still struggled with global spatial relationships.	Our hybrid model aims to overcome this limitation by integrating GNNs.
Deep Learning-Assisted Detection of Glaucoma Progression in SD-OCT (2023)	CNNs' reliance on paired scans and lack of global context.	Trained a CNN to predict glaucoma progression using paired OCT scans.	Demonstrated high performance in predicting glaucoma progression.	Did not address global context or spatial relationships.	Our study focuses on capturing global context, which is crucial for accurate diagnosis.
Detecting glaucoma with only OCT: Implications for the clinic, research, screening, and AI development (2022)	Challenges in glaucoma detection using OCT, including artifacts and the need for automated analysis.	Explored glaucoma detection using OCT, focusing on probability (p-) maps and incorporating both clinician-based and deep learning methods.	Highlighted the potential of OCT for glaucoma detection and discussed the challenges of artifacts, emphasizing the utility of deep learning for automated analysis.	Did not focus on specific architectures or address the limitations of CNNs in capturing global context, which is critical for accurate glaucoma detection.	This study underscores the importance of OCT in glaucoma detection and the need for advanced AI methods to address its limitations. Our proposed hybrid CNN-GNN model aims to fill the gap by capturing global context, which is not addressed in this study.
A Vision-GNN Framework for Retinopathy Classification Using Optical Coherence Tomography (2023)	GNNs' potential for OCT image analysis	Proposed a GNN-based pipeline for retinopathy classification.	Achieved high accuracy in multi-class retinopathy classification.	Focused on retinopathy, not glaucoma, and did not compare with CNNs. Highlights the need for global context, which our hybrid model addresses.	Our study applies GNNs specifically for glaucoma detection and compares with CNNs.
G2ViT: Graph Neural Network-Guided Vision Transformer Enhanced Network for retinal vessel and coronary angiograph segmentation (2024)	Need for global context in medical image segmentation, particularly for retinal vessels and coronary angiography.	Proposed G2ViT, a hybrid architecture combining CNNs, GNNs, and Vision Transformers for vessel segmentation. Used a GNN-guided transformer module to capture global context and a multi-scale edge feature attention module to preserve edge information.	Demonstrated the effectiveness of combining CNNs, GNNs, and Transformers for medical image segmentation. Achieved strong performance on retinal vessel and coronary angiography datasets.	Focused on vessel segmentation, not glaucoma detection. The specific architecture and modules (e.g., MEFA, MLF2) are tailored to vessel segmentation and may not be directly transferable to glaucoma detection.	The hybrid approach combining CNNs and GNNs is highly relevant to our proposed study. Our research adapts a similar hybrid approach but specifically targets glaucoma detection, addressing the need for global context in OCT image analysis.

The literature primarily focuses on CNNs for the detection of glaucoma while leaving aside the critical importance of global spatial relationships in OCT images. The work here fills the gap by suggesting a hybrid CNN-GNN model which captures local as well as global contextual information for the task of improving accuracy and robustness of the detection task for glaucoma. This contribution is important within the area of medical image analysis because it addresses a significant limitation of current methods in providing a more detailed approach for the diagnosis of glaucoma.

III. METHODOLOGY

The proposed methodology is a hybrid architecture for binary image classification. It uses a pre-trained convolutional neural network to extract feature maps from OCT images, capturing local texture and structural details. The feature maps will then be processed using a Graph Neural Network to model spatial relationships and global context within the image [8]. This allows the model to keep the fine details of the features, yet apply graph-based structural representations for classification improvement. GNN will represent OCT images as graphs where nodes represent the main segmented regions of interest (such as optic nerve head and retinal nerve fiber layer), and edges encode spatial and feature-based relationships [11].

IV. EXPERIMENTS

This study first offers benchmarking for standalone CNN models and GNNs as baseline models. Then, it compares the proposed hybrid CNN-GNN model with these baseline models in order to provide evidence on the superiority of the hybrid model for both local and global context capture. The hyperparameters of the CNN model and the GNN model are thus tuned such that their corresponding optimal performance is maximized. An ablation study is also performed to

investigate the contribution of each component- CNN and GNN-to achieve the final overall performance, so that the value of each one in the hybrid model is assured. The above justifications entail setting performance benchmarks with standalone models, showing capabilities of the proposed hybrid model and ensuring optimal performances through hyperparameter tuning, besides validating the value of each individual component by carrying out an ablation study.

V. EXPECTED RESULTS AND PLANNED ANALYSIS

A.) Expected Results

The hybrid CNN-GNN model will outperform individual CNNs and GNNs in terms of accuracy, sensitivity, and specificity. The hybrid model is anticipated to improve on the robustness of detection, especially of mild structural changes occurring with glaucoma.

B.) Planned Analysis

- Performance Metrics - AUC, sensitivity, specificity and F1-score analysis of the model
- Visualization – It draws a region in an OCT image that contributes most to the decision made by the model through Heatmaps.
- Discussion – It comments on how well the model has grasped local as well as wider context. Elaborate the consequences for diagnosis of glaucoma in light of the above through Model Performance Insights.

VI. SCOPE AND DELIMINAITON

The proposed work is centered on the binary classification of Optical Coherence Tomography (OCT) images with a view to distinguish between healthy eyes and glaucoma-affected eyes. This model is suited for 2D OCT images, although in future work, it could be extended to 3D volumes. In the study, no other diseases or conditions related to the eyes, which could be detected through OCT, have been covered. Because the model was trained and tested on a particular dataset, further validation might be necessary fo r generalization to other datasets. In addition, real-time deployment or how the model might be incorporated into clinical workflows are not addressed in this study; such could be further investigated in future work.

VII. REFERENCES

[1] Glaucoma Worldwide: A Growing Concern - Glaucoma Research Foundation. 10 Jan. 2024, https://glaucoma.org/articles/glaucoma-worldwide-a-growing-concern.

[2] Sarvamangala, D. R., & Kulkarni, R. V. (2022). Convolutional neural networks in medical image understanding: A survey. Evolutionary Intelligence, 15(1), 1–22. https://doi.org/10.1007/s12065-020-00540-3

[3] Maetschke, S., Antony, B., Ishikawa, H., Wollstein, G., Schuman, J., & Garnavi, R. (2019). A feature agnostic approach for glaucoma detection in OCT volumes. PLOS ONE, 14(7), e0219126. https://doi.org/10.1371/journal.pone.0219126

[4] Rasel, R. K., Wu, F., Chiariglione, M., Choi, S. S., Doble, N., & Gao, X. R. (2024). Assessing the efficacy of 2D and 3D CNN algorithms in OCT-based glaucoma detection. Scientific Reports, 14(1), 11758. https://doi.org/10.1038/s41598-024-62411-6

[5] Mariottoni, E. B., Datta, S., Shigueoka, L. S., Jammal, A. A., Tavares, I. M., Henao, R., Carin, L., & Medeiros, F. A. (2023). Deep learning–assisted detection of glaucoma progression in spectral-domain oct. *Ophthalmology Glaucoma*, *6*(3), 228–238. https://doi.org/10.1016/j.ogla.2022.11.004

[6] Hood, D. C., La Bruna, S., Tsamis, E., Thakoor, K. A., Rai, A., Leshno, A., de Moraes, C. G. V., Cioffi, G. A., & Liebmann, J. M. (2022). Detecting glaucoma with only OCT: Implications for the clinic, research, screening, and AI development. *Progress in Retinal and Eye Research*, *90*, 101052. https://doi.org/10.1016/j.preteyeres.2022.101052

[7] Hu, M., Wang, J., Wynne, J., Liu, T., & Yang, X. (2023). A vision-GNN framework for retinopathy classification using optical coherence tomography. *Medical Imaging 2023: Computer-Aided Diagnosis*, *12465*, 200–206. https://doi.org/10.1117/12.2653615

[8] Xu, H., & Wu, Y. (2024). G2vit: Graph neural network-guided vision transformer enhanced network for retinal vessel and coronary angiograph segmentation. *Neural Networks*, *176*, 106356. https://doi.org/10.1016/j.neunet.2024.106356

[9] CDC. (2024, October 17). *About glaucoma*. Vision and Eye Health. https://www.cdc.gov/vision-health/about-eye-disorders/glaucoma.html

[10] Xu, Y., Yan, X., Sun, B., & Liu, Z. (2022). Global contextual residual convolutional neural networks for motor fault diagnosis under variable-speed conditions. *Reliability Engineering & System Safety*, *225*, 108618. https://doi.org/10.1016/j.ress.2022.108618

[11] Chang, C.-W., Chang, C.-Y., Lin, Y.-Y., Su, W.-W., & Chen, H. S.-L. (2023). A glaucoma detection system based on generative adversarial network and incremental learning. *Applied Sciences*, *13*(4), 2195. https://doi.org/10.3390/app13042195