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.

Alvear, Mark Josh, Fermin, Raudmon Yvhan, Naniong, Nigel Jan, Tonido, Andrei Christian

Background of the Study

Glaucoma is an optic nerve disorder characterized by a progressive damage for the most part to the optic nerve. This may occur in almost anyone, with millions around the world at risk for irreversible vision loss. As a result, early and accurate diagnosis may be more important for intervention and slowing or halting of progression [1]. The recent advances in diagnostic imaging have resulted in the widespread acceptance of OCT as a valuable tool in glaucoma diagnosis. It allows high-resolution cross-sectional images of the retinal layers and ONH [9]. Such high resolution enables the subtle structural changes seen in the form of glaucomatous damage, like RNFL thinning and optic disc cupping, to be visualized. While CNNs have shown significant success in analyzing medical images, including OCT scans for glaucoma detection, they often fail to capture the complex spatial relationships and global context necessary for accurate diagnosis [2].

CNNs perform excellently at local feature extraction, like texture and edge information, but have a hard time modeling the subtle structural organization of the ONH and its adjacent tissues. For example, RNFL thickness variations between different locations may be associated with the optic disc cupping level, and this global dependency cannot be directly learned by conventional CNN architectures [4]. This study fills the important gap by presenting a novel hybrid deep learning model that unifies the benefits of both Convolutional Neural Networks (CNNs) and Graph Neural Networks (GNNs) [7]. The idea is to merge the capabilities of CNNs for local feature extraction with GNNs to capture global relations and hypothesize superior performance for glaucoma detection and classification from OCT images.

In particular, the CNN will learn to extract relevant local features from the OCT image, and the GNN will model relationships between different regions or segments of the ONH, capturing crucial global context often missed by CNNs alone [10]. The integrated approach should enhance the accuracy and robustness of glaucoma diagnosis and ultimately lead to earlier detection and more effective management of this sight-threatening disease [5].

Literature review:

Research Title/ Articles	Date of Publication	Summary of the Method	Advantages	Disadvantages
Convolutional Neural Networks in Medical Image Understanding: A Survey[2]	Jan 3, 2021	This paper surveys the use of CNNs in medical image analysis. It identifies strengths in local feature extraction and highlights CNNs' limitations in capturing global spatial relationships, critical for analyzing complex anatomical structures.	Demonstrates CNNs' effectiveness in tasks like segmentation and classification. Provides a strong foundation for their role in medical imaging.	Struggles with modeling spatial relationships and global dependencies, leading to inefficiencies in detecting structural abnormalities like the optic nerve's features.
A Feature Agnostic Approach for Glaucoma Detection in OCT Volumes[3]	July 1, 2019	Used 3D CNNs to classify healthy and glaucomatous eyes from raw OCT volumes without requiring segmentation. Focused on automation for scalability.	Reduces reliance on manual annotation. Efficient for large datasets and clinical applications.	CNNs alone may miss crucial spatial relationships in OCT images, leading to potential misclassification.

Assessing the Efficacy of 2D and 3D CNN Algorithms in OCT-Based Glaucoma Detection[4]	May 4, 2024	Compared 2D and 3D CNN architectures for glaucoma detection in OCT images, emphasizing the need for improved feature representation.	Highlights that 3D CNNs can leverage volumetric data more effectively than 2D CNNs.	CNNs still struggle to model complex relationships between different anatomical regions.
Deep Learning–Assisted Detection of Glaucoma Progression in Spectral-Domain OCT[5]	July, 2023	Trained a CNN to analyze paired (baseline and follow-up) SD-OCT RNFL thickness measurements and predict glaucoma progression based on the time between visits. Ground truth was based on expert consensus.	Demonstrated high performance (AUC 0.938, sensitivity 87.3%, specificity 86.4%), outperforming conventional trend-based analysis.	Relied on paired scans, which may not always be available. Did *not* address global context or spatial relationships between retinal regions, a key focus of our proposed research.
Detecting glaucoma with only OCT: Implications for the clinic, research, screening, and AI development [6]	September, 2022	Explored glaucoma detection using only OCT, focusing on probability (p-) maps and incorporating both clinician-based and deep learning methods to address artifact challenges in OCT scans.	Emphasized the potential of OCT for glaucoma detection and discussed the challenges of artifacts, highlighting the utility of deep learning for automated analysis.	Did not focus on specific architectures or address the limitations of CNNs in capturing global context, a critical aspect of our research.
A vision-GNN framework for retinopathy classification using optical coherence tomography[7]	April 7, 2023	This paper proposes an end-to-end Graph Neural Network (GNN) pipeline for multi-class retinopathy classification using OCT images. It constructs a graph from the OCT image and uses a GNN to extract graph-based features for classification.	Demonstrates the effectiveness of GNNs for OCT image analysis and achieves high accuracy (99.07%) for multi-class retinopathy classification. Applies Vision-GNN to OCT image analysis for the first time (according to the authors).	Focuses on *retinopathy* classification, not specifically *glaucoma* detection, although the methodology is highly relevant. The specific graph construction method and GNN architecture used may not be directly transferable to glaucoma detection, as the relevant anatomical features and relationships may differ. Doesn't explicitly compare against CNN-only approaches or address the limitations of CNNs in capturing global context, which is a central point of our research.

G2ViT: Graph Neural Network-Guided Vision Transformer Enhanced Network for retinal vessel and coronary angiograph segmentation[8]	August, 2024	This paper proposes G2ViT, a hybrid architecture combining CNNs, GNNs, and Vision Transformers for vessel segmentation. It uses a GNN-guided transformer module to capture global context and a multi-scale edge feature attention module to preserve edge information.	Demonstrates the effectiveness of combining CNNs, GNNs, and Transformers for medical image segmentation. Achieves strong performance on retinal vessel and coronary angiography datasets. Explicitly addresses the need for global context in vessel segmentation.	Focuses on *vessel segmentation*, not *glaucoma detection*, although the architectural principles are highly relevant. The specific architecture and modules (MEFA, MLF2) are tailored to vessel segmentation and may not be directly transferable to glaucoma. While it combines CNNs, GNNs, *and* Transformers, our proposed research focuses on the CNN-GNN combination, so we will explain why we are prioritizing that specific hybrid approach.

Tentative Research Problem:

Despite considerable progress in medical imaging and AI, glaucoma remains a significant cause of vision loss globally. Accurately diagnosing glaucoma using OCT images in a timely manner is crucial for effective treatment and prevention of blindness. However, current classification methods may not be robust or efficient enough to handle the complexities of glaucoma detection, specifically in distinguishing between healthy and glaucomatous eyes [6]. The structural changes associated with glaucoma, such as variations in the optic nerve head and retinal nerve fiber layer, present challenges for traditional classification methods. These complexities underscore the importance of developing innovative AI architectures that can effectively utilize both local and global information within OCT images to improve diagnostic accuracy and efficiency [3].

Tentative Proposed 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].

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