GUIDE: A Responsible Multimodal Approach for Enhanced Glaucoma Risk Modeling and Patient Trajectory Analysis

Heman Shakeri
School of Data Science
University of Virginia
hs9hd@virginia.edu

Ahson Syed
School of Data Science
University of Virginia
as7ac@virginia.edu

Arjun Dirghangi
School of Medicine
University of Virginia
ad3bb@uvahealth.org

Behnaz. Moradi-Jamei
Department of Department of Mathematics and Statistics
James Madison University
moradibx@jmu.edu

Abstract

Glaucoma is a irreversible vision loss that disproportionately affects marginalized communities. Current diagnostic and management strategies often fail to account for individualized risks, leading to suboptimal patient outcomes and exacerbating health disparities. Here, we present **GUIDE** (Glaucoma Understanding and Integrated Data Evaluation), a conceptual framework for explainable multimodal AI framework that integrates diverse data sources—including clinical measurements, imaging data, unstructured electronic health records (EHR), and social determinants of health (SDOH)—to create a comprehensive and personalized view of glaucoma risk and progression. Our approach focuses on developing clinically interpretable, expert-tunable hierarchical fusion models that address key issues such as fairness, transparency, and robustness, aligning with responsible AI principles. By disentangling the embedding space using clinical supervision at each stage of modality fusion, we prevent model hallucinations and ensure that the embeddings can be decoded into physician-understandable clinical concepts. We also implement contextual transparency by engaging stakeholders and tailoring transparency measures according to the NIST's Contextual Transparency Playbook. Our framework handles data quality issues through pre-training strategies and hierarchical data fusion, and considers modalities with varying costs to optimize resource utilization. We demonstrate how GUIDE provides a comprehensive understanding of glaucoma progression, facilitates more accurate risk stratification, and enables personalized treatment plans.

1 Introduction

Glaucoma is the leading cause of irreversible vision loss worldwide, affecting over 3 million adults in the United States alone [5, 2]. The disease disproportionately impacts marginalized communities, exacerbating existing health disparities. Current diagnostic and management strategies often fail to account for individualized risks, leading to suboptimal patient outcomes. This misalignment results in inefficient allocation of healthcare resources, where low-risk individuals may receive more care than needed, while high-risk patients remain inadequately monitored and managed [12]. Recent research has highlighted how machine learning tools can offer new avenues for improved glaucoma detection

38th Conference on Neural Information Processing Systems (NeurIPS 2024).

and management [21, 8]. While these approaches have made significant strides in analyzing clinical measurements and imaging data, many current methods still face challenges in integrating diverse data types. They often lack the incorporation of unstructured electronic health records (EHR) and social determinants of health (SDOH) characteristics, which are crucial for identifying key predictors of glaucoma risk and progression and for accounting for health disparities in outcomes [8]. Moreover, the integration of these varied data sources with traditional clinical and imaging data remains a significant challenge in developing comprehensive, personalized risk assessment models.

To address these challenges, we present **GUIDE** (Glaucoma Understanding and Integrated Data Evaluation), an explainable multimodal AI framework that integrates diverse data sources to create a comprehensive and personalized view of glaucoma risk and progression. GUIDE combines available clinical measurements (such as intraocular pressure and visual field tests), high-resolution imaging data (including optical coherence tomography and fundus photography), unstructured electronic health records (EHR), and social determinants of health (SDOH) in a unified framework. This approach allows for a more holistic understanding of patient risk factors and disease trajectories. GUIDE focuses on developing clinically interpretable, expert-tunable hierarchical fusion models that address key issues such as fairness, transparency, and robustness, aligning with responsible AI principles while leveraging the full spectrum of available patient data.

Our approach tackles several key challenges in the field:
1) **Interpretability:** We decompose the modeling process into distinct, interpretable stages, each incorporating clinical context through semi-supervised learning. This ensures that the latent spaces generated at each stage are directly translatable into meaningful clinical concepts, enhancing transparency and alignment with clinical reasoning [13].
2) **Multimodal integration:** By transforming raw data into clinically meaningful metrics that reflect underlying disease mechanisms, we construct a comprehensive view of the disease state in a latent space regularized by clinical expertise. This facilitates more informed clinical decisions and advances personalized medicine.
3) **Temporal dynamics:** We address the complexities of longitudinal patient data by treating each visit's data as independent set elements, allowing us to uncover temporal connections and dynamic changes without enforcing a strict sequential order [1, 14].
4) **Disentanglement of the representation:** By disentangling the embedding space using clinical semi-supervision at each stage of modality fusion, we restrict the embeddings to be decoded into physician-understandable clinical concepts, preventing model hallucinations.
5) **Contextual transparency:** We implement contextual transparency by engaging stakeholders and tailoring transparency measures according to the NIST's Contextual Transparency Playbook [**?**].
6) **Data quality and resource optimization:** Our framework handles data quality issues through pre-training strategies and hierarchical data fusion, and considers modalities with varying costs to optimize resource utilization.

By addressing these challenges, GUIDE aims to provide a comprehensive understanding of glaucoma progression, facilitate more accurate risk stratification through the identified patient trajectories, and enable personalized treatment plans. Ultimately, this research seeks to improve patient outcomes, address disparities in eye health, and advance eye and vision health equity. By focusing on marginalized communities, this work contributes to reducing health disparities in eye care and informs community guidelines for the responsible design of next-generation foundational models in healthcare. Furthermore, GUIDE serves as a case study for broader applications of trustworthy multimodal data fusion in healthcare. The principles and methodologies developed in this work can potentially be adapted to other medical domains, demonstrating how responsible AI can be implemented to integrate diverse data sources while maintaining interpretability, fairness, and robustness.

1.1 Related Work

Recent advancements in AI for glaucoma management have shifted from single-modality models [21, 8] to multimodal approaches [20, 19], which better capture the disease's complexity. However, these often neglect non-clinical data like social determinants of health (SDOH). Our work extends these approaches by integrating SDOH with clinical and imaging data for a more comprehensive risk assessment. Interpretability in medical AI remains crucial for clinical adoption. Recent efforts focus on developing interpretable models [16], with specific applications in glaucoma [22]. We build on these techniques, extending them to a multimodal context and incorporating clinical expertise throughout the modeling process.

Responsible AI in healthcare addresses fairness, accountability, and transparency. Studies have highlighted gender and racial biases in medical AI [9, 3]. Our work applies these principles to glaucoma care, explicitly considering SDOH and implementing fairness constraints. Temporal modeling is essential for accurate risk prediction and personalized treatment. Recent works have explored various approaches for healthcare data [10], including glaucoma progression [4]. We adapt these techniques to our multimodal context, addressing challenges of irregular sampling and feature heterogeneity. Contextual transparency in AI systems, particularly in healthcare, has gained attention. Sloane et al. [15] argue for transparency measures tailored to specific stakeholders, while Mitchell et al. [11] proposed model cards for comprehensive AI reporting. We incorporate these principles to develop meaningful transparency mechanisms for both clinicians and patients. In summary, while AI applications in glaucoma care have progressed significantly, there remains a need for responsible, interpretable, and contextually transparent multimodal models that can integrate diverse data sources.

2 Conceptual Framework

Our approach, employs a hierarchical representation learning framework that integrates multimodal data to produce clinically interpretable representations for improved understanding and management of established glaucoma patients. The methodology consists of two main components: (1) Hierarchical Representation Learning and (2) Patient Trajectory Analysis.
Hierarchical Representation Learning: The hierarchical representation learning approach synthesizes diverse data sources into coherent representations, offering new insights into treatment responsiveness in a semi-supervised, clinically-driven manner. This approach is well-suited to address the challenges posed by the relatively low prevalence of glaucoma and the difficulty in distinguishing early glaucoma from normal variation. As illustrated in Figure 1, the model employs joint modalities in a hurdle-like fashion, prioritizing low-cost, common screening tests at lower levels and progressing to more comprehensive data integration at higher levels to identify higher risks.

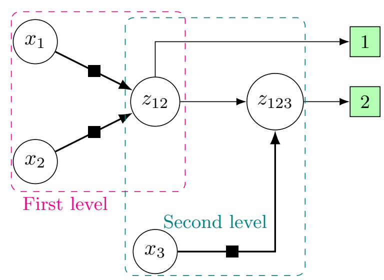

Figure 1: *The proposed multi-level hierarchical modality fusion. Clinical contexts 1 and 2 are regularizing the latent space in semi-supervised fashion. Solid squares illustrate the representation learning blocks.*

At each point in time, our model learns feature embeddings at multiple levels of granularity. Let $\mathbf{e}_{jk}$ represent the embedding of the j-th feature at time step k at level l. The updated embedding at level l is given by:

$$\mathbf{z}_k^{(l)} = \mathbf{z}_k^{(l-1)} + \Delta\mathbf{z}_k^{(l-1)} \tag{1}$$

where $\Delta\mathbf{z}_k^{(l-1)}$ is the change in the embedding from level $l-1$ to l, computed using an attention mechanism on the new features introduced at level l:

$$\Delta\mathbf{z}_k^{(l-1)} = \sum_{j=1}^{n_k} \alpha_{jk}^{(l-1)} \mathbf{e}_{jk}^{(l)} \tag{2}$$

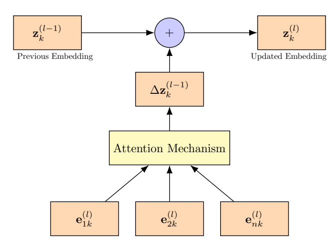

Figure 2: *Feature embedding update process at visit k. Each new feature at level l contributes to the update in the embedding using an attention mechanism.*

The attention weights $\alpha_{jk}^{(l-1)}$ are computed as:

$$\alpha_{jk}^{(l-1)} = \frac{\exp(\langle \mathbf{u}^{(l-1)}, \mathbf{e}_{jk}^{(l)} \rangle)}{\sum_{i=1}^{n_k} \exp(\langle \mathbf{u}^{(l-1)}, \mathbf{e}_{ik}^{(l)} \rangle)} \tag{3}$$

2.1 Semi-Supervised Regularization and Disentanglement

To enhance interpretability, we employ a semi-supervised approach to disentangle the latent space using available clinical supervision signals. At each stage l, we inject clinical knowledge to refine the latent process $\mathbf{z}^{(l)}$ with respect to defined clinical concepts $\mathbf{y} = \mathbf{y}_{1:T} \in \mathbb{R}^{P \times T}$. The regularization loss is formalized as [17]:

$$\text{regularization loss} = -\prod_{t=1}^{T}\prod_{g=1}^{G}\prod_{j\in\nu(g)} P(y_t^j | h_\gamma^j(z_t^{(l),\epsilon(g)}, \mathbf{c}_t)) \tag{4}$$

where $h_\gamma^j(z_t^{(l),\epsilon(g)}, \mathbf{c}_t)$ is the resultant hierarchical deep parameterized probability vector, and $\nu(g)$ and $\epsilon(g)$ correspond to the indices of the g-th guided medical concept, and the indices in the latent space defined for guided concept g at stage l, respectively.

2.2 Patient Trajectory Analysis

Building upon our hierarchical representation learning approach, we now focus on analyzing patient trajectories to enable dynamic risk scoring and patient subtyping. This analysis is crucial for enhancing personalized glaucoma management through the comprehensive examination of multimodal data trajectories. We represent the sequence of visit data up to time T for patient j as an unordered set $\mathcal{S}_j = s_i := (\tau_1, z_1), \cdots, (\tau_T, z_T)$, where τ is a positional time embedding normalized to a chosen dimension c, i.e., $\tau \in \mathbb{R}^c$ [18], and z_i is the feature embedding at embedded time τ_i obtained from our hierarchical representation learning process. Inspired by Horn et al. [6], we employ a set function approach to aggregate all elements (visit data) within each patient. This method is particularly well-suited for the irregularly-sampled and non-synchronized time series data common in glaucoma care. By treating each visit as an element in a set, we can naturally handle varying numbers of observations and unaligned measurements across different modalities. The patient's trajectory is embedded using a summary statistic of the set $\mathcal{S}_j$, defined as:

$$\Psi_i = \frac{1}{|\mathcal{S}i|}\sum j \leq i a_j(s_j, \mathcal{S})h(s_j) \tag{5}$$

where a_j are attention weights, and $h(s_j)$ is a feature function applied to each set element s_j. These summary statistics are then aggregated using a parameterized function g_θ, implemented as a neural network that operates over sets: $g_\theta(\Psi_i)$. Following Horn et al. [6], we express our set function f as:

$$f(\mathcal{S}) = g\left(\frac{1}{|\mathcal{S}|}\sum_{s_j\in\mathcal{S}} h(s_j)\right) \tag{6}$$

where $h : \Omega \to \mathbb{R}^d$ and $g : \mathbb{R}^d \to \mathbb{R}^C$ are neural networks, $d \in \mathbb{N}^+$ determines the dimensionality of the latent representation, and C is the number of classes. This formulation allows for learning dataset-specific summary statistics optimized for classification performance. To handle irregularly sampled time series data and quantify risks at intermediate time points, we incorporate a self-supervised learning approach through masking, as proposed by Trottet et al. [17]. This technique captures both local and global structures of glaucoma progression, particularly where trend changes are observed at a few visits. Our set function approach offers several advantages for glaucoma progression modeling:

Flexibility: It handles varying numbers of observations per patient without requiring imputation or fixed-length inputs. Permutation invariance: The order of observations doesn't affect the output, suitable for non-synchronized measurements. Scalability: The computational cost scales linearly with the number of observations, making it efficient for large datasets [6]. Interpretability: Using attention mechanisms allows us to quantify the importance of individual observations to the final prediction. By combining this set function approach with our hierarchical representation learning and semi-supervised disentanglement, we create a powerful and flexible framework for modeling glaucoma progression that can handle the complexities and irregularities of real-world clinical data.

2.3 Contextual Transparency

To ensure fairness and transparency in glaucoma risk assessment and subtyping, we align our approach with the Blueprint for an AI Bill of Rights [7] and the concept of contextual transparency [15]. Recognizing perfect transparency as unattainable, we adopt a tailored approach specific to glaucoma care, involving: i) Stakeholder Specificity: Identifying key stakeholders and their transparency needs. ii) ADS Specificity: Documenting our model's technical details, including its hierarchical structure and data processing methods. iii) Transparency and Outcome Specificity: Designing measures that directly address glaucoma risk assessment and patient subtyping goals. We prioritize transparency by documenting the algorithm's design and decision-making processes, ensuring accessibility to both healthcare providers and patients. This includes explaining data integration, handling of irregularly sampled data, and the derivation of risk assessments and patient subtypes. Our dynamic, customized approach surpasses a one-size-fits-all framework by: i) Developing interpretable visualizations of patient trajectories and risk factors; ii) Providing context-specific explanations tailored to user expertise; iii) Offering quantifiable measures of model certainty and factor importance. This approach fosters trust and understanding of the AI's role, crucial for addressing the diverse presentations and ensuring equitable, unbiased diagnostics in glaucoma care.

3 Experiments

To evaluate the effectiveness of the proposed GUIDE framework, we are conducting extensive experiments using a multicenter comprehensive electronic health record (EHR) dataset encompassing over 10,000 patients diagnosed with glaucoma. Our dataset comprises a rich array of multimodal data sources. Clinical measurements include intraocular pressure (IOP) readings, visual acuity scores, and visual field (VF) test results. The imaging data consists of optical coherence tomography (OCT) scans and fundus photographs, providing detailed structural information about the retina and optic nerve. We also incorporated unstructured clinical notes containing physician observations and patient-reported symptoms, which offer valuable contextual information and sometimes supervision signal for the semi-supervised disentanglement (Section 2.1). To address health disparities, we extracted social determinants of health (SDOH) data from patient records and linked community-level data sources. Additionally, we included patient health system interactions, such as records of appointments, missed visits, and prescription adherence, to capture patient engagement patterns.

Evaluation Metrics We employ a comprehensive set of evaluation metrics to assess various aspects of our model's performance. Predictive accuracy was primarily measured using the area under the Receiver Operating Characteristic curve (AUC) for risk prediction tasks. To evaluate interpretability, we conducted attention weight analysis and gathered feedback from clinicians on the model's outputs. Fairness is assessed using demographic parity and equalized odds metrics across different patient subgroups, ensuring that the model performed equitably across diverse populations. We also test the model's robustness against missing data scenarios and adversarial perturbations to simulate real-world challenges.

3.1 Illustrative Example: Representing Unstructured Visual Acuity Text Data as LogMAR

To clarify our concept of meaningful embeddings in the context of glaucoma management, we present an illustrative example using visual acuity data. This example demonstrates how our approach can transform unstructured clinical text into standardized, clinically interpretable values. Visual acuity is a critical measure in glaucoma assessment, but it often appears in clinical notes as unstandardized, textual descriptions. Our goal was to convert these unstructured descriptions into standardized LogMAR (Logarithm of the Minimum Angle of Resolution) scores, which are widely used in ophthalmology for quantifying visual acuity. We developed a pipeline to achieve this transformation, consisting of three main steps (illustrated in Figure 3): First, unstructured clinical text is converted into a machine-readable format through tokenization and transforms into an embedded representations that can be processed by downstream machine learning algorithms. Second, a Random Forest classifier is trained to differentiate between various descriptions of visual acuity and assign corresponding standardized labels. This step is analogous to the first level of our hierarchical model (as shown in Figure 1). Finally, the output from the classifier is normalized to produce either discrete categorical labels or continuous LogMAR values.

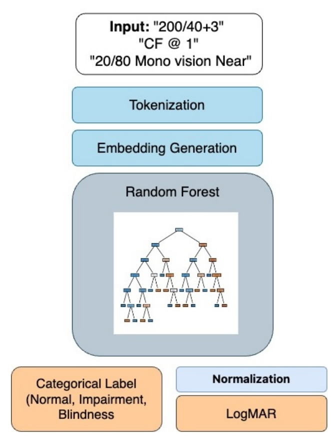

Figure 3: *Example of Level 1 integration in Figure 1: parsing clinical notes into machine-readable embeddings, classifying, and normalizing the output to endpoint scores.*

This approach demonstrates several key aspects of our broader GUIDE framework: First, by converting textual data into standardized numerical scores, we enable the integration of visual acuity information with other clinical measures. Second, the resulting LogMAR scores are directly interpretable by clinicians, maintaining the clinical relevance of the original data while standardizing it for computational analysis. In tests with a subset of visual acuity descriptions, our method achieved high classification accuracy, with 99% of the variability captured by the first 338 most unique frequent terms.

4 Conclusions

The GUIDE framework represents a conceptual advancement in applying multimodal AI to healthcare, particularly for managing chronic conditions like glaucoma. It addresses key challenges in responsibly building and deploying AI systems, aiming for high predictive accuracy while maintaining interpretability, fairness, and robustness – crucial for AI adoption in healthcare.

A key innovation of GUIDE is its approach to creating meaningful embeddings from complex data at each stage of modality fusion. This capability is crucial for handling heterogeneous and irregularly sampled data in glaucoma care and potentially other medical domains. Our framework recognizes that irregularity in both measurement modalities and visit timing carries valuable information about the patient's condition and risk state.

In chronic condition management, the patient's risk state evolves, sometimes becoming more uncertain at different disease stages. GUIDE is designed to capture this dynamic nature, recognizing when additional fine-tuned modalities, such as expensive imaging or invasive checks, may be necessary. By learning from existing patterns of irregularity, GUIDE can suggest optimal modalities for each visit and inform better timing for future follow ups.

This adaptive approach to data collection and visit scheduling represents a significant advancement in personalized care, allowing for more efficient resource use while ensuring critical changes in a patient's condition are not missed. The principles underlying GUIDE are applicable to a wide range of chronic conditions where long-term patient trajectories and diverse data sources are crucial.

Our emphasis on contextual transparency addresses critical ethical considerations in AI-assisted healthcare, aiming to foster trust among clinicians and patients by tailoring transparency measures to specific stakeholders and contexts.

This is an ongoing research with the following directions:

- Implementation and testing in glaucoma care settings, optimizing modality selection and visit timing.
- Extending the model to other chronic conditions with similar data requirements.
- Developing dynamic interfaces for guiding clinicians in test selection and follow-up scheduling.
- Investigating the impact of the model's suggestions on long-term outcomes and resource utilization.
- Exploring effective communication of the model's rationale to enhance shared decision-making.

While still a conceptual framework, GUIDE represents a significant step towards responsible AI in healthcare.

References

[1] Ahmed M Alaa, Scott Hu, and Mihaela Schaar. Learning from clinical judgments: Semi-markov-modulated marked hawkes processes for risk prognosis. In *International Conference on Machine Learning*, pages 60–69. PMLR, 2017.

[2] Karen Allison, Deepkumar Patel, and Omobolanle Alabi. Epidemiology of glaucoma: the past, present, and predictions for the future. *Cureus*, 12(11), 2020.

[3] Philippe Burlina, Neil Joshi, William Paul, Katia D Pacheco, and Neil M Bressler. Addressing artificial intelligence bias in retinal diagnostics. *Translational Vision Science & Technology*, 10(2):13–13, 2021.

[4] Rui Dai, Srijan Das, Kumara Kahatapitiya, Michael S Ryoo, and François Brémond. Ms-tct: Multi-scale temporal convtransformer for action detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 20041–20051, 2022.

[5] Shaili S Davuluru, Alison T Jess, Joshua Soo Bin Kim, Kristy Yoo, Van Nguyen, and Benjamin Y Xu. Identifying, understanding, and addressing disparities in glaucoma care in the united states. *Translational Vision Science & Technology*, 12(10):18–18, 2023.

[6] Max Horn, Michael Moor, Christian Bock, Bastian Rieck, and Karsten Borgwardt. Set functions for time series. In *International Conference on Machine Learning*, pages 4353–4363. PMLR, 2020.

[7] White House. *Blueprint for an ai bill of rights: Making automated systems work for the american people*. Nimble Books, 2022.

[8] Xiaoqin Huang, Md Rafiqul Islam, Shanjita Akter, Fuad Ahmed, Ehsan Kazami, Hashem Abu Serhan, Alaa Abd-Alrazaq, and Siamak Yousefi. Artificial intelligence in glaucoma: opportunities, challenges, and future directions. *BioMedical Engineering OnLine*, 22(1):126, 2023.

[9] Agostina J Larrazabal, Nicolás Nieto, Victoria Peterson, Diego H Milone, and Enzo Ferrante. Gender imbalance in medical imaging datasets produces biased classifiers for computer-aided diagnosis. *Proceedings of the National Academy of Sciences*, 117(23):12592–12594, 2020.

[10] Bryan Lim, Sercan Ö Arık, Nicolas Loeff, and Tomas Pfister. Temporal fusion transformers for interpretable multi-horizon time series forecasting. *International Journal of Forecasting*, 37(4):1748–1764, 2021.

[11] Margaret Mitchell, Simone Wu, Andrew Zaldivar, Parker Barnes, Lucy Vasserman, Ben Hutchinson, Elena Spitzer, Inioluwa Deborah Raji, and Timnit Gebru. Model cards for model reporting. In *Proceedings of the conference on fairness, accountability, and transparency*, pages 220–229, 2019.

[12] Jack Phu, Katherine Masselos, Michael Sullivan-Mee, and Michael Kalloniatis. Glaucoma suspects: the impact of risk factor-driven review periods on clinical load, diagnoses, and healthcare costs. *Translational Vision Science & Technology*, 11(1):37–37, 2022.

[13] Humberto Salazar, Vedant Misra, and Swarup S Swaminathan. Artificial intelligence and complex statistical modeling in glaucoma diagnosis and management. *Current opinion in ophthalmology*, 32(2):105–117, 2021.

[14] Satya Narayan Shukla and Benjamin M Marlin. Multi-time attention networks for irregularly sampled time series. *arXiv preprint arXiv:2101.10318*, 2021.

[15] Mona Sloane, Ian René Solano-Kamaiko, Jun Yuan, Aritra Dasgupta, and Julia Stoyanovich. Introducing contextual transparency for automated decision systems. *Nature Machine Intelligence*, 5(3):187–195, 2023.

[16] Erico Tjoa and Cuntai Guan. A survey on explainable artificial intelligence (xai): Toward medical xai. *IEEE transactions on neural networks and learning systems*, 32(11):4793–4813, 2020.

[17] Cécile Trottet, Manuel Schürch, Ahmed Allam, Imon Barua, Liubov Petelytska, Oliver Distler, Anna-Maria Hoffmann-Vold, Michael Krauthammer, et al. Modeling complex disease trajectories using deep generative models with semi-supervised latent processes. *arXiv preprint arXiv:2311.08149*, 2023.

[18] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.

[19] Zhizhou Wang, Jun Wang, Hongru Zhang, Chen Yan, Xingkui Wang, and Xin Wen. Mstnet: method for glaucoma grading based on multimodal feature fusion of spatial relations. *Physics in Medicine & Biology*, 68(24):245002, 2023.

[20] Yongli Xu, Run Sun, Man Hu, and Hui Zeng. A dual-modal fusion network using optical coherence tomography and fundus images in detection of glaucomatous optic neuropathy. *Current Eye Research*, pages 1–7, 2024.

[21] Siamak Yousefi. Clinical applications of artificial intelligence in glaucoma. *Journal of Ophthalmic & Vision Research*, 18(1):97, 2023.

[22] Zizhao Zhang, Han Zhang, Long Zhao, Ting Chen, Sercan Ö Arik, and Tomas Pfister. Nested hierarchical transformer: Towards accurate, data-efficient and interpretable visual understanding. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 36, pages 3417–3425, 2022.