JOURNAL OF
EMERGING INVESTIGATORS

Article

Retinal biomarkers for the detection of neurological disease using deep learning

Olivia Kim[1], Kwan-Woo Kim[1]
[1] Henry M Gunn High School, Palo Alto, California, United States

SUMMARY

Neurological disorders affect billions of people worldwide, yet curative treatments remain elusive, making timely identification critical to improving patient prognoses. This study explored whether features detectable through accessible retinal imaging can serve as biomarkers for neuro-ophthalmological diseases that share biological mechanisms with neurodegeneration and related neurological conditions. We hypothesized that three distinct retinal phenotypes would demonstrate strong diagnostic performance as candidate biomarkers for cataracts, glaucoma, macular degeneration, and pathological myopia. To test this prediction, we trained a hybrid deep learning model combining patient demographic data with retinal image analysis using the Ocular Disease Intelligence Recognition Five Thousand (ODIR5K) database. Model robustness was evaluated using the area under the receiver operating characteristic curve (AUROC) and externally validated with the Retinal Fundus Multi-Disease Image (RFMiD) dataset. Lens clarity demonstrated the strongest diagnostic performance, with cataract detection reaching an AUROC of 0.92. External validation supported the generalizability of retinal features across datasets. These findings suggest that retinal phenotypes are measurably associated with specific neuro-ophthalmological conditions and may serve as accessible complementary tools in neurological risk assessment. Future longitudinal studies are needed to determine whether these features can support earlier detection of disease progression.

INTRODUCTION

Approximately one in three people worldwide will experience a neurological disorder at some point in their life, resulting in billions of people experiencing disability and possibly eventual death (1). Though estimates vary, tens of billions of dollars are invested annually in palliative and hospice care in the United States alone, placing a substantial burden on healthcare systems, especially in regions without well-established infrastructures and limited budgets (2). These figures reflect the reality that few neurological conditions have curative treatments, making timely identification the most promising strategy to improve patient prognoses (3). However, many neurological diseases remain difficult to detect using definitive, accessible diagnostic methods (4).

One existing method involves a comprehensive neurological exam, which assesses a variety of cognitive functions and serves as a vital step in diagnosis. The process entails a trained physician evaluating a patient's cognitive status, motor function, sensory function, reflexes, and coordination, among other factors, to determine nervous system integrity (5). Additional methods include neuroimaging (e.g., magnetic resonance imaging, computed tomography, positron emission tomography), electrodiagnostic tests (e.g., electroencephalogram, electromyography, evoked potential testing), laboratory culture tests, cerebrospinal fluid analysis, genetic testing, and biopsies (5). While these procedures can provide valuable clinical insight, they often require sophisticated equipment, trained personnel, and financial resources, limiting their practicality for widespread screening.

Retinal imaging has emerged as a promising window into the brain for neurological assessment. The retina is a thin layer of tissue lining the back of the eye and converts light into electrical signals that are relayed by the optic nerve to be processed by the brain (6). Because it is anatomically an extension of the central nervous system via the optic nerve with a shared vascular supply and structural similarities, the retina is uniquely sensitive to neurological changes (6). As such, certain ophthalmic diseases are comorbid with neurodegenerative processes. This relationship has been supported in conditions such as Alzheimer's disease, Parkinson's disease, optic nerve atrophy, neuroinflammation, and broader neurological deterioration (7). Retinal changes related to these conditions, such as retinal nerve fiber thinning and retinal ganglion cell degeneration, are detectable using noninvasive imaging techniques such as optical coherence tomography (8). Retinal screening does not demand sophisticated infrastructure beyond handheld imaging hardware nor extensively trained medical staff, and scans can be analyzed remotely through cloud-based software (8). This makes it a relatively low-cost, portable, and accessible complement to traditional methods.

In parallel, certain retinal features measurable using standard imaging techniques may serve as indicators of neuro-ophthalmological pathology. Lens clarity refers to the transparency of the eye's crystalline lens, with reduced clarity being a hallmark sign of lens opacification and leading to blurred vision (9). Optic nerve cupping describes structural changes in the optic nerve head, where enlargement of the central depression is commonly associated with pressure-related damage (10). Hyperreflective foci are small, highly reflective spots observed on retinal imaging, often linked to inflammatory or degenerative processes in retinal diseases (11).

The four ophthalmic diseases examined in this study are cataracts, glaucoma, macular degeneration, and pathological myopia. These conditions each affect different structures within the eye and entail diverse risk profiles. Cataracts lead to blurred vision as proteins in the eye's lens aggregate, a process linked to neurodegenerative processes through shared pathways involving oxidative stress, vascular

https://doi.org/10.59720/25-246

dysfunction, and age-related cellular damage throughout the eye (9, 12). Glaucoma typically causes progressive visual defects and eye pain from the neurodegeneration of retinal ganglion cells in the optic nerve due to increased eye pressure (7). Macular degeneration entails increasing central vision loss as photoreceptors and retinal cells deteriorate, with aging, genetic factors, and chronic inflammatory processes contributing to neurodegeneration (7). Lastly, pathological myopia involves axial elongation of the eye, leading to progressive retinal degeneration and vision impairments (7). Each of these conditions is also associated with neurological factors, motivating investigation into their shared pathological mechanisms (6).

In this study, we hypothesized that lens clarity would demonstrate strong diagnostic performance as a candidate biomarker for cataracts, optic nerve cupping for glaucoma, and hyperreflective foci for macular degeneration. Specifically, we predicted that we would obtain higher diagnostic accuracy for these neuro-ophthalmological diseases, along with pathological myopia, when using these accessible retinal features as biomarkers. We reasoned that we would achieve greater diagnostic accuracy for each condition using these features as cataracts may cause reduced lens transparency, glaucoma-related nerve damage may cause structural changes to the optic nerve head, and hyperreflective foci may reflect inflammation related to retinal degeneration. By testing these predictions against two large multimodal clinical datasets, we aimed to determine whether these retinal features could serve as biomarkers for these neuro-ophthalmological diseases, potentially extending to neurodegeneration and broader neurological risk.

To address our research question, we used the Ocular Disease Intelligence Recognition Five Thousand (ODIR5K), a fully anonymized database publicly released by the Ocular Disease Intelligence Recognition group, which contains over 5000 samples encompassing diverse sexes and ages (13). Combined with a hybrid deep learning framework, we sought to identify which retinal feature was most strongly associated with our selected neuro-ophthalmological conditions. To further assess the robustness of these relationships, we applied the trained model to an external retinal imaging dataset, the Retinal Fundus Multi-Disease Image Dataset (RFMiD).

After evaluating retinal biomarkers across clinical data, we found that lens clarity was most strongly associated with one of our chosen neuro-ophthalmological diseases, based on the relatively high accuracy of cataract detection compared to the other conditions. Our findings supported the potential of incorporating retinal imaging as a complement to traditional diagnostic strategies for neurological conditions, bridging critical gaps in current accessibility standards. Because the datasets available were cross-sectional, this study focuses on neuro-ophthalmological disease association with retinal phenotypes. Overall, our work provides a foundation for future studies investigating whether retinal biomarkers may directly contribute to earlier identification of neuro-ophthalmological and neurological risk when longitudinal data becomes available.

RESULTS

To test our prediction that lens clarity would be associated with cataracts, optic nerve cupping with glaucoma, and hyperreflective foci with macular degeneration, we extracted retinal features from fundus scans showing the inner, back surface of the eye (**Figure 1**). For this process, we used a convolutional neural network (CNN). A CNN is a type of

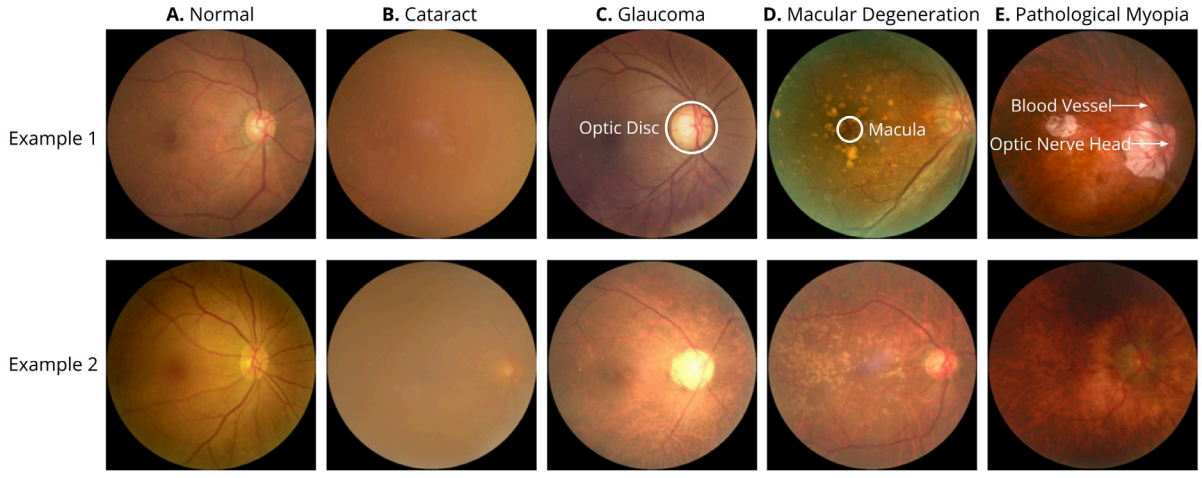

Figure 1: Representative sample fundus images illustrating anatomical features and phenotypic variability from the Ocular Disease Intelligence Recognition Five Thousand database following preprocessing. Two representative retinal images for each diagnostic class after preprocessing, reflecting variability in retinal appearance both within and across disease categories. Key anatomical structures and disease-associated features relevant to the retinal biomarkers evaluated by the model are labeled. **A)** Normal retinal images. **B)** The hazy appearance of cataract scans reflects reduced lens clarity caused by lens clouding. **C)** The enlarged optic disc in glaucoma scans suggests optic nerve cupping associated with retinal nerve fiber thinning and retinal ganglion cell loss. **D)** In macular degeneration scans, pigmentary changes near the macula reveal retinal abnormalities that are associated with the appearance of hyperreflective foci under certain imaging conditions. **E)** For pathological myopia, elongation of the eye can stretch the retina around the optic nerve head, leading to thinning that exposes lighter tissue underneath the retina; thinner blood vessels may also indicate reduced retinal blood flow and other vascular changes. The images demonstrate consistent resolution and centering achieved through preprocessing, ensuring uniform input for effective feature extraction by the convolutional neural network.

https://doi.org/10.59720/25-246

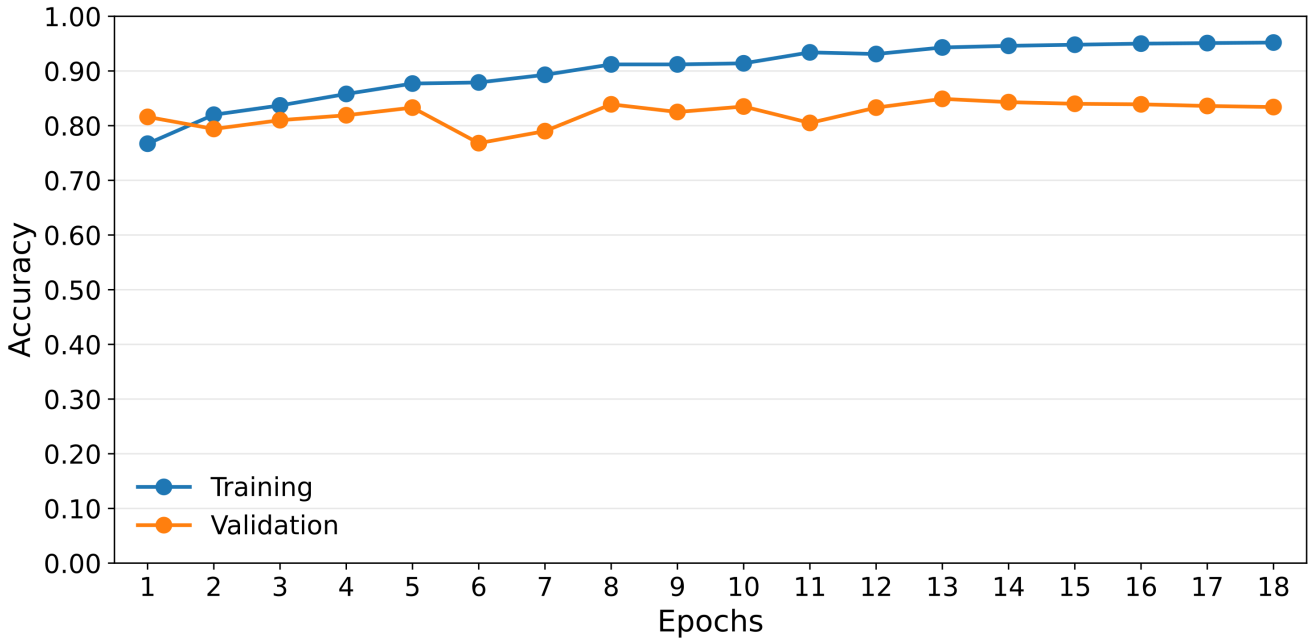

Figure 2: Model performance stabilizes after 13 training epochs without major indications of overfitting. Training and validation accuracy of the final EfficientNetB5 model across successive training epochs. Accuracy (unitless proportion) is plotted as a function of the number of epochs, where one epoch represents a complete pass through the training data. Training accuracy steadily increased with additional epochs, while validation accuracy began deteriorating after epoch 13 (13 complete passes through the training data). Beyond this point, the divergence between training and validation accuracy indicated overfitting, where the model's performance on the training data continued to improve but did not generalize well to unseen data. Epoch 13 yielded the highest validation accuracy observed during model training of 0.85, with a corresponding training accuracy of 0.94. Thus, the model at epoch 13 was selected as the optimal model state, and early stopping was applied to preserve it.

machine learning architecture modelled after the human visual cortex to extract features, including imaging data, making it well-adapted to image classification tasks (14). Afterward, it offered a probability breakdown of its classification results to indicate which neuro-ophthalmological conditions it predicted from the lens clarity, optic nerve cupping, and hyperreflective foci it identified in the input fundus scans (15).

Though the ODIR5K dataset included limited chart information, we complemented our CNN with patient age and sex data with a Keras Sequential model, a simple neural network with a linear structure (16). In doing so, we aimed to address age and sex differences in neuro-ophthalmological disease prevalence, as women and older adults are at significantly higher risk. This made it beneficial to err on the side of caution for certain demographics by interpreting disease probabilities more generously (17).

Optimizing model performance through iterative training

We conducted several iterations of deep learning to optimize our model. Specifically, we refined our model extensively through repeated training, validation, and testing cycles. We visualized this process with a learning curve, which displayed overall trends in model accuracy for both training and validation in relation to the number of passes through the training data, known as epochs. We monitored both training and validation accuracy to ensure the model was generalizing well to unseen data. Our model achieved optimal performance after 13 epochs before plateauing, where training and validation accuracy reached 0.94 and 0.85, respectively (**Figure 2**). These accuracy rates supported our model's ability to identify associations between retinal biomarkers and our four selected neuro-ophthalmological diseases.

Investigating the predictive value of retinal features across neuro-ophthalmological diseases

We selected five clinically relevant diagnostic classes that each represented at least 5% of the complete dataset, excluding three of the original eight categories (diabetes, hypertension, other abnormalities) that were not directly relevant to our study objectives. In our analysis, we ultimately examined the following five classes: normal, cataracts, glaucoma, macular degeneration, and pathological myopia. To evaluate the relative diagnostic accuracy of each class, we determined the area under the receiver operating characteristic curve (AUROC). For each disease class, the model returned a prediction score representing the probability that a given image belonged to that class. The model compared these scores with the known disease labels across different classification thresholds to calculate the true-positive and false-positive rates. Using these rates, it constructed the receiver operating characteristic (ROC) curves. The model then calculated AUROC values as the area under the ROC curves, encapsulating its ability to distinguish between images with and without the disease. These values range from 0.5 to 1, where a value of 0.5 is equivalent to random classification, and a value of 1 represents perfect discrimination (18).

https://doi.org/10.59720/25-246

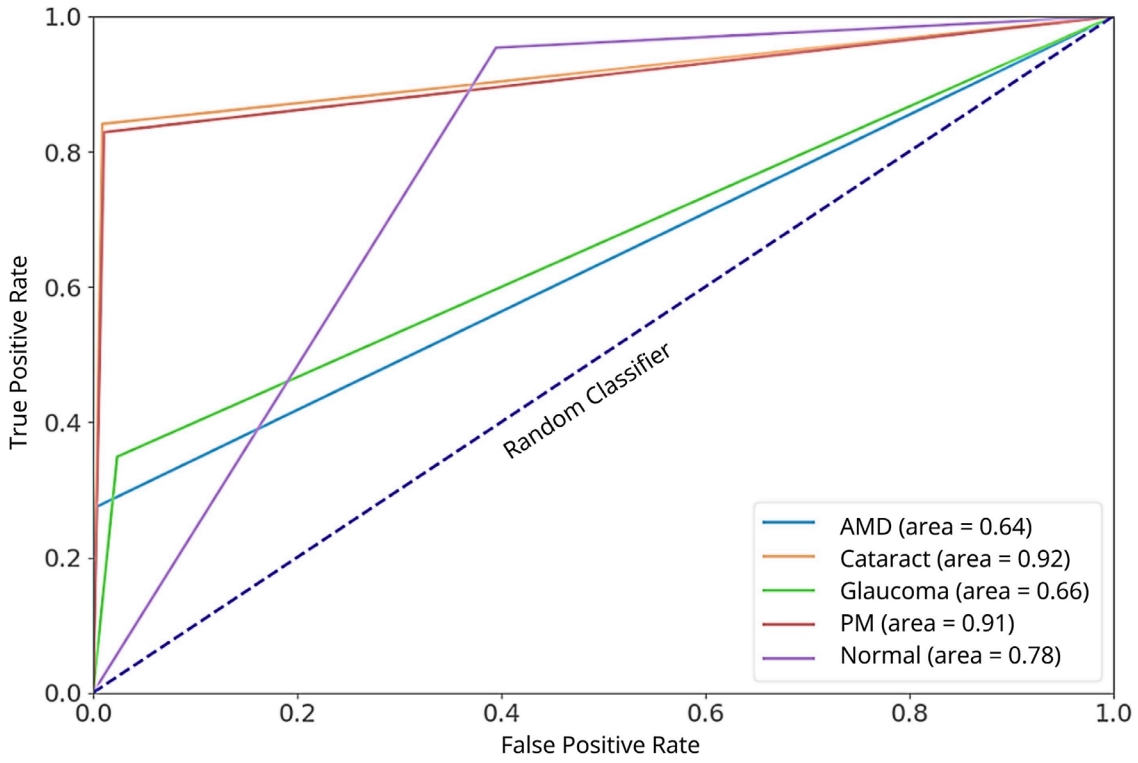

Figure 3: Retinal features demonstrate varying discriminative power across neuro-ophthalmological disease classes. Receiver operating characteristic (ROC) curves for each diagnostic class generated by the final EfficientNetB5 model trained on the Ocular Disease Intelligence Recognition Five Thousand dataset. ROC curves were produced using one-vs-rest classification on the held-out test set (n = 592 retina scans). Each curve represents the model's ability to distinguish one disease class from all others using the corresponding retinal biomarker: lens clarity for cataracts (orange), optic nerve cupping for glaucoma (green), and hyperreflective foci for age-related macular degeneration (AMD, blue). Pathological myopia (PM, red) and normal controls (purple) were classified using composite retinal structural features. Area under the receiver operating characteristic curve (AUROC) values are reported in the legend and indicate strong discrimination for cataracts (AUROC = 0.92) and pathological myopia (AUROC = 0.91), while glaucoma (AUROC = 0.66) and macular degeneration (AUROC = 0.64) demonstrated relatively weak discrimination. The dashed diagonal line (slope = 1) provides a reference for random classification, representing a model that distinguishes disease from non-disease cases no better than chance and has an AUROC of 0.5.

We evaluated each selected retinal feature as a potential biomarker of its corresponding disease based on our earlier predictions: lens clarity for cataracts, optic nerve cupping for glaucoma, and hyperreflective foci for macular degeneration. Lens clarity was the most informative biomarker, with cataract detection achieving an AUROC of 0.92, indicating a 0.92 probability of correctly distinguishing a cataract scan from a non-cataract scan (**Figure 3**). This relationship is biologically consistent with the clouding of the lens in cataracts, which can be detected in retina scans. In contrast, hyperreflective foci showed weaker predictive value, as indicated by the AUROC of 0.64 for macular degeneration detection (**Figure 3**). Optic nerve cupping also showed relatively weak value, where glaucoma detection reached an AUROC of 0.66 (**Figure 3**). Unlike cataracts, glaucoma, or macular degeneration, the model classified pathological myopia using composite retinal structural features rather than a single, isolated biomarker. While the other conditions had relatively identifiable phenotypes, pathological myopia produced a range of structural changes in the retina that may collectively provide stronger diagnostic information than any one predictor. Using a combination of features, including retinal stretching and abnormal blood vessel growth, pathological myopia detection reached an AUROC of 0.91 (**Figure 3**).

Testing results generalizability on an independent external dataset

We evaluated the generalizability of our findings beyond the ODIR5K database by applying our deep learning model to an independent dataset, the RFMiD. Without any additional fine-tuning of the model, we observed comparable overall performance, although the predictive value of individual biomarkers varied. Optic nerve cupping's predictive accuracy saw an increase to 0.80 (**Figure 4**). The predictive accuracy of hyperreflective foci also increased to 0.71 for macular degeneration detection (**Figure 4**). Meanwhile, the diagnostic accuracy of pathological myopia declined to 0.55, leading to an overall accuracy of 0.68 for the RFMiD dataset (**Figure 4**). Lens clarity was not an available biomarker in the RFMiD, and we were thus unable to investigate it. Overall, most categories did not deviate substantially from our previous observations, indicating that our selected retinal features retained their discriminatory value beyond the training data.

DISCUSSION

https://doi.org/10.59720/25-246

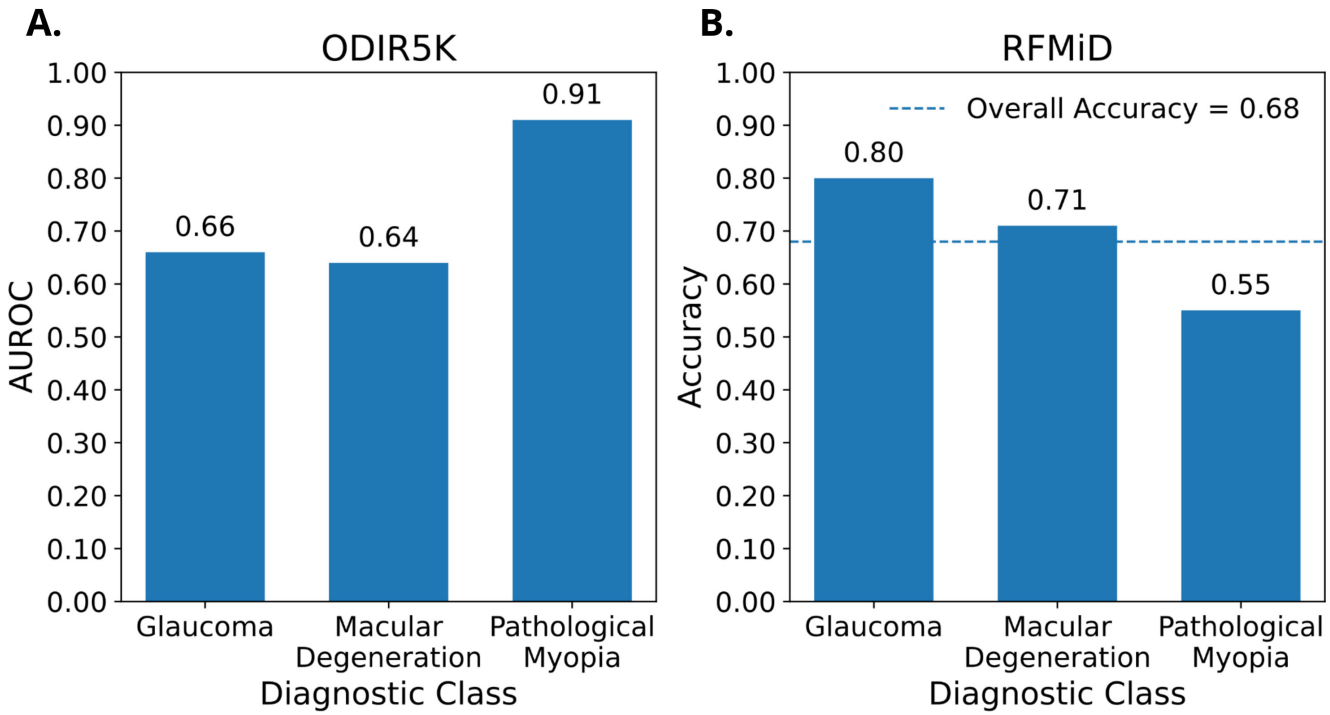

Figure 4: Model performance across the Ocular Disease Intelligence Recognition Five Thousand and independent Retinal Fundus Multi-Disease Image Dataset. Bar graphs showing model performance across three diagnostic classes shared between the Ocular Disease Intelligence Recognition Five Thousand (ODIR5K) and independent Retinal Fundus Multi-Disease Image Dataset (RFMiD). **A)** Area under the receiver operating characteristic curve (AUROC) values for the final model evaluated on the held-out ODIR5K test set. **B)** Accuracy values obtained when the same model was applied to the independent RFMiD dataset without additional fine-tuning. Model performance on the RFMiD was assessed using accuracy rather than AUROC due to differences in labelling structure. Cataract detection was excluded because lens clarity was not an available biomarker in the RFMiD dataset.

In this study, we investigated whether lens clarity, optic nerve cupping, and hyperreflective foci, three accessible retinal features, are associated with four neuro-ophthalmological conditions that share disease mechanisms with neurodegeneration. Using a deep learning framework iteratively trained on retinal imaging and demographic data, we first evaluated these associations within the ODIR5K dataset and later tested their robustness with an independent external dataset.

Our results partially supported our initial hypothesis that each retinal feature would show meaningful diagnostic performance for its corresponding condition: lens clarity for cataracts, optic nerve cupping for glaucoma, and hyperreflective foci for macular degeneration. Lens clarity emerged as the most informative biomarker, with cataract detection reaching an AUROC of 0.92, consistent with our hypothesis. This strong association is biologically plausible, as cataracts cause cloudy vision that may be explained by changes in lens clarity observable on retina scans (9).

In contrast, hyperreflective foci demonstrated a weaker predictive value than anticipated (macular degeneration AUROC = 0.64). This may reflect limitations in image quality, resolution, or lighting that could make hyperreflective foci less visually apparent, or alternatively, suggest that hyperreflective foci alone do not fully capture the pathology underlying macular degeneration. Optic nerve cupping showed a modest connection, as indicated by an AUROC of 0.66 for glaucoma detection. Notably, pathological myopia exhibited strong discriminative performance, with an AUROC of 0.91, indicating that the model effectively captured the retinal structural alterations associated with this condition despite its heterogeneous presentation. This strong performance may reflect the fact that pathological myopia produces pronounced structural changes to the retina that are readily detectable even when no single biomarker dominates.

These findings reinforced the potential of retinal biomarkers to reliably detect structural alterations in diseases with heterogeneous presentations. They also emphasized the need to consider disease-specific features and combinations of features when selecting predictive biomarkers to accommodate these diverse phenotypic changes. Overall, these outcomes highlight the importance of experimentally testing predictive hypotheses, as they allow us to confirm some expectations while revising our understanding of others.

Importantly, external validation on the RFMiD dataset demonstrated that our chosen retinal features retained meaningful predictive capacity when applied to retinal images outside the original database. Although our metrics shifted from the results of our original testing, this was an expected consequence of variability in imaging devices, population

https://doi.org/10.59720/25-246

characteristics, and disease prevalence across databases. In general, the preservation of meaningful discriminative ability across datasets supports the biological relevance of the selected retinal features. These findings emphasize the promising clinical implications of retinal biomarkers and the need for diverse, representative data when developing robust machine learning models.

Our findings largely align with existing studies, which have likewise noted the diagnostic potential of retinal imaging for neuro-ophthalmological and neurodegenerative disorders (19, 20). By integrating machine learning with accessible retinal imaging, our study extends this area of work by quantitatively comparing multiple biomarkers with a unified framework and assessing their generalizability across datasets. This approach helps mitigate human bias and enables the analysis of complex clinical data that would be difficult to interpret using traditional diagnostic methods alone, ensuring a holistic overview of factors that may contribute to disease.

Beyond their diagnostic implications, these insights raise broader related questions regarding the eye-brain axis and the extent to which visual structure anomalies and neurological conditions may reflect interconnected disease processes. While investigating these relationships presents challenges due to biological complexity and confounding variables, the potential clinical impact is immense. Integrating portable fundus imaging combined with machine learning-based triage systems into current practices could facilitate diagnosis, especially in settings with limited access to specialized neurological care.

However, our research is not without its limitations. The use of publicly available data restricted access to detailed clinical histories, including comorbid conditions, socioeconomic factors, and longitudinal outcomes, which may influence disease presentation and model performance. Additionally, since the data was cross-sectional, our model could only predict disease presence, rather than temporal aspects such as disease onset or pre-symptomatic diagnosis. Future work should include more granular, long-term clinical metadata, additional external datasets, and expanded evaluation metrics. Doing so would offer additional context and further refine model robustness while enabling longitudinal analyses to determine whether retinal biomarkers can directly inform the prediction of neurological diseases.

In sum, we investigated how certain biomarkers in the eye relate to multiple neuro-ophthalmological conditions. Specific retinal features, particularly lens clarity, were strongly associated with select diseases and may serve as sensitive indicators of neuro-ophthalmological conditions. Future studies can build on our work by directly evaluating whether such retinal features can also serve as biomarkers of broader neurological diseases. By combining multi-dimensional clinical data with deep learning, our findings generally supported our hypothesis and suggest that eye-based screening may be useful as a complementary tool in assessing neuro-ophthalmologic risk. Simultaneously, our results call for more representative data that capture greater diversity in patient demographics, geographic regions, and disease presentations, as the ODIR5K and RFMiD datasets may not fully represent the broader populations affected by these conditions. This limits the generalizability of findings needed to eventually draw the larger connection to neurological risk and enable translational impact.

MATERIALS AND METHODS

Data acquisition and preprocessing

The selection of retina scans, accompanied by age and sex information for 5000 unique patients, was first curated from the ODIR5K, a fully anonymized database publicly released by the ODIR group (13). The dataset was used as provided without additional sampling or demographic balancing. The images in the unprocessed multimodal database were first normalized by scaling pixel intensity values to a range of zero to one, followed by data cleaning. With the aid of the NumPy library (version 1.26.4) in Python (version 3.11) for image loading, all images were resized from 448 x 448 to 224 x 224 pixels for more efficient processing, in addition to being normalized for brightness and noise and centered for uniform resolution (**Figure 1**, 21).

After data preprocessing, a considerable class imbalance in the data was observed through visualizations made with the Python Pandas library (version 2.1.4) (22). The data was balanced using a variety of techniques, including oversampling, undersampling, data augmentation, and class weighting. Though these methods helped reduce inaccuracies, three of the original eight diagnostic categories (diabetes, hypertension, other abnormalities) with less than 5% representation and clinical irrelevance to the study's objectives were discarded, leaving five classes: normal, cataracts, glaucoma, macular degeneration, and pathological myopia.

Model architecture and training

EfficientNetB5 is a CNN architecture from the Python TensorFlow library (version 2.20.0) designed to balance classification accuracy with computational efficiency (23). It was selected after cross-experimentation with six other models for its optimal balance between classification performance, runtime, and memory usage. EfficientNetB5 operates by identifying visual patterns and combining these features to recognize structural patterns in unseen images.

When constructing the image-compatible CNN, multiple hyperparameter configurations were tested over several iterations. Hyperparameters are user-defined settings that control how a machine learning model learns from data, including factors such as learning rate, number of trainable layers, and optimization strategy. The Python Scikit-learn library (version 1.3.2) was used to split the data into a 70:15:15 ratio for training, validating, and testing, respectively (24). One-hot encoding was applied to the diagnostic labels to reduce bias during classification. Only one layer of the pretrained EfficientNetB5 was made trainable, and seven additional layers activated by rectified linear units (ReLU) were added for the model to learn complex patterns within the retinal images. Subsequently, a granular learning rate of 0.0001 with the Adam optimizer was implemented with a categorical cross-entropy loss function to promote stable and efficient training.

Finally, the said model was combined with a 13-layer TensorFlow sequential model incorporating age and sex data. Associated hyperparameters included the Adam optimizer, a

https://doi.org/10.59720/25-246

learning rate of 0.001, and a categorical cross-entropy loss function. This linear secondary model allowed non-image clinical information to contribute to predictions by learning relationships between demographic factors and disease prevalence.

Evaluation and External Validation

The complete model's performance was evaluated with accuracy and AUROC, a metric that measures how well the model distinguishes between diagnostic classes across a range of classification thresholds. AUROC was chosen for its robustness with imbalanced datasets, where a value of 0.5 indicates performance equivalent to random guessing, while a value of 1.0 represents flawless classification. For interpretation, AUROC values of 0.90–1.00 were considered excellent, 0.80–0.90 good, 0.70–0.80 moderate, 0.60–0.70 poor, and 0.50–0.60 indicative of no better than random discrimination. Accuracy and AUROC values were calculated with the Scikit-learn library and visualized with the Python Matplotlib library (version 3.8.4) (25). All analyses were conducted using Deepnote software (accessed July 2024), with graphics processing units (26).

To evaluate model robustness and generalizability, the finalized model was applied to the RFMiD dataset, an independent and publicly available dataset consisting of retinal fundus images labeled across multiple ocular disease categories (27). The images were preprocessed using the same pipeline applied to the ODIR5K retinal scans. No additional fine-tuning was performed prior to inference, ensuring a strict external validation protocol. Model performance on the RFMiD was assessed using accuracy metrics rather than AUROC due to differences in labelling structure.

Received: September 4, 2025
Accepted: May 7, 2026
Published: September 4, 2026

REFERENCES

1. World Health Organization. "Over 1 in 3 People Affected by Neurological Conditions, the Leading Cause of Illness and Disability Worldwide." *World Health Organization*, 14 Mar. 2024. https://www.who.int/news/item/14-03-2024-over-1-in-3-people-affected-by-neurological-conditions--the-leading-cause-of-illness-and-disability-worldwide. Accessed 16 Aug. 2025.
2. Raphael, Carol, et al. "Financing End-of-Life Care in the USA." *Journal of the Royal Society of Medicine*, vol. 94, no. 9, 2001, pp. 458–61. https://doi.org/10.1177/014107680109400912.
3. American Brain Foundation. "The Global Prevalence of Brain Disease." *American Brain Foundation*, 2 July 2024. https://www.americanbrainfoundation.org/the-global-prevalence-of-brain-disease/. Accessed 16 Aug. 2025.
4. *AMA Journal of Ethics*. "How Should Clinicians Minimize Harms and Maximize Benefits When Diagnosing and Treating Disorders Without Biomarkers?" *AMA Journal of Ethics*, 1 July 2021. https://edhub.ama-assn.org/ama-journal-of-ethics/module/2781715#section-249759696. Accessed 16 Aug. 2025.
5. National Institute of Neurological Disorders and Stroke. "Neurological Diagnostic Tests and Procedures." *National Institute of Neurological Disorders and Stroke*, 15 Jan. 2025. https://www.ninds.nih.gov/health-information/disorders/neurological-diagnostic-tests-and-procedures. Accessed 1 Sept. 2025.
6. London, Alon, et al. "The Retina as a Window to the Brain—From Eye Research to CNS Disorders." *Nature Reviews Neurology*, vol. 9, no. 1, 2013, pp. 44–53. https://doi.org/10.1038/nrneurol.2012.227.
7. Marchesi, Nicoletta, et al. "Ocular Neurodegenerative Diseases: Interconnection between Retina and Cortical Areas." *Cells*, vol. 10, no. 9, 2021, p. 2394. https://doi.org/10.3390/cells10092394.
8. Zimmermann, Hanna, et al. "Optical Coherence Tomography for Retinal Imaging in Multiple Sclerosis." *Degenerative Neurological and Neuromuscular Disease*, vol. 4, 9 Dec. 2014, pp. 153–62. https://doi.org/10.2147/DNND.S73506.
9. Cleveland Clinic. "Eye Lens (Crystalline Lens)." *Cleveland Clinic*, 13 Aug. 2024. https://my.clevelandclinic.org/health/body/eye-lens-crystalline-lens. Accessed 21 July 2026.
10. University of Illinois Hospital & Health Sciences System. "Optic Nerve Cupping." *University of Illinois Hospital & Health Sciences System*, 2026. https://hospital.uillinois.edu/primary-and-specialty-care/ophthalmology/ophthalmology-services/optic-nerve-cupping. Accessed 21 July 2026.
11. Amer, Mona, et al. "Hyperreflective Choroidal Foci: A Comprehensive Review." *Journal of Ophthalmic & Vision Research*, vol. 20, 13 Nov. 2025, https://doi.org/10.18502/jovr.v20.18397.
12. Jefferis, J. M., et al. "Cataract and Cognitive Impairment: A Review of the Literature." *The British Journal of Ophthalmology*, vol. 95, no. 1, 2011, pp. 17–23. https://doi.org/10.1136/bjo.2009.165902.
13. Peking University. "Peking University International Competition on Ocular Disease Intelligent Recognition (ODIR-2019)." *ODIR-2019 Grand Challenge*. https://odir2019.grand-challenge.org/. Accessed 24 July 2026.
14. Celeghin, Alessia, et al. "Convolutional Neural Networks for Vision Neuroscience: Significance, Developments, and Outstanding Issues." *Frontiers in Computational Neuroscience*, vol. 17, 6 July 2023, p. 1153572. https://doi.org/10.3389/fncom.2023.1153572.
15. Vakalopoulou, Maria, et al. "Deep Learning: Basics and Convolutional Neural Networks (CNNs)." *Machine Learning for Brain Disorders*, edited by Olivier Colliot, Humana, 2023, Chapter 3. https://doi.org/10.1007/978-1-0716-3195-9_3.
16. Chollet, François. "The Sequential Model." *Keras*, 12 Apr. 2020. https://keras.io/guides/sequential_model/. Accessed 1 Sept. 2025.
17. Sen, Hatice Nida, et al. "Gender Disparities in Ocular Inflammatory Disorders." *Current Eye Research*, vol. 40, no. 2, 2015, pp. 146–61. https://doi.org/10.3109/02713683.2014.932388.
18. Çorbacıoğlu, Şeref Kerem, and Gökhan Aksel. "Receiver Operating Characteristic Curve Analysis in Diagnostic Ac-

https://doi.org/10.59720/25-246

curacy Studies: A Guide to Interpreting the Area under the Curve Value." *Turkish Journal of Emergency Medicine*, vol. 23, no. 4, 3 Oct. 2023, pp. 195–98. https://doi.org/10.4103/tjem.tjem_182_23.
19. Zeppieri, Marco, et al. "Optical Coherence Tomography (OCT): A Brief Look at the Uses and Technological Evolution of Ophthalmology." *Medicina (Kaunas, Lithuania)*, vol. 59, no. 12, 3 Dec. 2023, p. 2114. https://doi.org/10.3390/medicina59122114.
20. Li, Xiaotao, et al. "Computer Vision for Brain Disorders Based Primarily on Ocular Responses." *Frontiers in Neurology*, vol. 12, 21 Apr. 2021, p. 584270. https://doi.org/10.3389/fneur.2021.584270.
21. Harris, Charles R., et al. "Array Programming with NumPy." *Nature*, vol. 585, no. 7825, 16 Sept. 2020, pp. 357–362. https://doi.org/10.1038/s41586-020-2649-2
22. McKinney, Wes. "Data Structures for Statistical Computing in Python." *Proceedings of the 9th Python in Science Conference*, 2010, pp. 51–56. https://doi.org/10.25080/Majora-92bf1922-00a
23. Abadi, Martín, et al. "TensorFlow: A System for Large-Scale Machine Learning." *12th USENIX Symposium on Operating Systems Design and Implementation (OSDI 16)*, 2016, pp. 265–283. https://www.usenix.org/conference/osdi16/technical-sessions/presentation/abadi. Accessed 27 Aug. 2026.
24. Pedregosa, Fabian, et al. "Scikit-learn: Machine Learning in Python." *Journal of Machine Learning Research*, vol. 12, 2011, pp. 2825–2830. https://jmlr.org/papers/v12/pedregosa11a.html. Accessed 27 Aug. 2026.
25. Hunter, John D. "Matplotlib: A 2D Graphics Environment." *Computing in Science & Engineering*, vol. 9, no. 3, 2007, pp. 90–95. https://doi.org/10.1109/MCSE.2007.55
26. Deepnote Team. "Deepnote: The Data Notebook for the AI Era." *Deepnote*. https://github.com/deepnote/deepnote. Accessed 27 Aug. 2026.
27. Pachade, Samiksha, et al. "Retinal Fundus Multi-Disease Image Dataset (RFMiD): A Dataset for Multi-Disease Detection Research." *Data*, vol. 6, no. 2, 2021, p. 14. https://doi.org/10.3390/data6020014