Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 5, No.2, May : 2025

FROM DETECTION TO GRADING: A DEEP LEARNING APPROACH FOR GLAUCOMA STAGING

Pratibha Patil Department of Information Technology, I2IT, Pune pratibhap@isquareit.edu.in
Dr. Sonali Patil Department of Information Technology, I2IT, Pune sonalip@isquareit.edu.in
Manthan Titarmare Department of Information Technology, I2IT, Pune manthan.titarmare13@outlook.com
Somya Sikarwar Department of Information Technology, I2IT, Pune somyasikarwar21@gmail.com
Aanchal Takhtar Department of Information Technology, I2IT, Pune aanchaltakhtar@gmail.com
Abhishek Gaikwad Department of Information Technology, I2IT, Pune abhirajj99135@gmail.com

ABSTRACT
Glaucoma, a leading cause of irreversible blindness, necessitates early and accurate detection for effective treatment and prevention of vision loss. This study introduces a novel deep learning-based approach for automated glaucoma detection and staging using retinal fundus images. Leveraging the ResNet18 architecture, the system identifies glaucoma stages by extracting critical features from the images, while the AlexNet model determines whether the images represent healthy or glaucomatous eyes. The system is integrated into a web-based platform, enabling ophthalmologists to upload retinal images and instantly receive a comprehensive analysis of glaucoma presence and its stage. Key performance metrics such as accuracy, precision, recall, and F1-score are used to evaluate the model's efficacy. This research aims to support ophthalmologists by providing a robust, automated tool for enhanced diagnostic accuracy, contributing to improved patient outcomes and advancing the field of medical image analysis through artificial intelligence.

Keywords:
Glaucoma Detection, Deep Learning, Retinal Fundus Images, CNN, Optimization.

I. Introduction

Glaucoma is a group of progressive optic neuropathies that cause damage to the optic nerve, responsible for transmitting critical visual information from the eye to the brain [16]. Often associated with elevated intraocular pressure, glaucoma can also occur in individuals with normal pressure readings, making it a complex and multifactorial disease to manage [16]. It remains one of the leading causes of irreversible blindness worldwide, primarily affecting individuals over the age of 40 [4]. With the global population aging, the prevalence of glaucoma is expected to rise, placing a significant burden on public health systems [4]. The silent and gradual progression of glaucoma means that symptoms often go unnoticed until advanced stages, where vision loss becomes irreversible [16]. Therefore, early diagnosis and intervention are imperative for mitigating the disease's impact. Glaucoma manifests in several subtypes, including open-angle glaucoma—the most common form—characterized by a progressive obstruction of the eye's drainage system, leading to increased intraocular pressure and optic nerve damage [4][16]. Other forms, such as angle-closure glaucoma, normal-tension glaucoma, and congenital glaucoma, require tailored diagnostic and management approaches [16].

Early detection is crucial, especially for high-risk populations, such as individuals with a family history of glaucoma, those with diabetes, or certain ethnic groups predisposed to the disease [16]. Comprehensive ocular assessments, including intraocular pressure measurement, visual field testing, and optic nerve imaging, are essential for identifying glaucoma in its nascent stages [4][16]. Among these, retinal fundus imaging plays a pivotal role, providing detailed, high-resolution images of the retina and optic nerve head. These images help detect key indicators such as optic nerve cupping and retinal nerve fiber layer thinning, both of which are hallmarks of glaucomatous damage [15][16]. In recent years, advancements in artificial intelligence have transformed the field of medical imaging,

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 5, No.2, May : 2025

with deep learning frameworks such as convolutional neural networks (CNNs) proving particularly effective in automating and enhancing fundus image analysis [7][8][11]. These models excel at identifying subtle patterns in medical images that may be challenging for human experts to discern, thereby improving diagnostic precision and efficiency [8][11].
This study builds upon these advancements by proposing a novel hybrid framework for glaucoma detection and staging, leveraging the strengths of two deep learning models. ResNet18 is employed for staging glaucoma by extracting comprehensive features from retinal fundus images, while AlexNet classifies cases as glaucomatous or healthy. Additionally, a web-based platform has been developed, allowing ophthalmologists to upload fundus images and instantly receive diagnostic results, including both glaucoma staging and presence classification. By integrating these technologies, the proposed system aims to provide a user-friendly, automated solution that supports timely intervention and enhances diagnostic accuracy. This work contributes to the field of medical image analysis by advancing the early detection and management of glaucoma, a silent but formidable threat to vision [7][8][11][16].

II. Literature

2.1 GlauNet: Glaucoma Diagnosis for OCTA Imaging Using a New CNN Architecture (2022)

GlauNet is the first of its kind to mark an advancement in deep learning in the application for medical imaging diagnostics, particularly for diagnosing glaucoma with OCTA [17]. The study comprised 939 images obtained from 546 patients [17]. To ensure the diversity of cases, which is a must for robust models, there were 285 glaucomatous and 487 non-glaucomatous images [17]. The researchers made sure that 167 poor-quality images were removed so that only high-quality inputs feed into the model, thus making the model more reliable [17]. The study divided the dataset into a training set of 258 glaucomatous and 439 non-glaucomatous images and a test set of 27 glaucomatous and 48 non-glaucomatous images so as to have proper validation of the model's efficacy [17]. GlauNet's architecture was specifically designed to include innovative feature extraction and classification modules that leverage advanced techniques in CNNs [17]. The incorporation of data augmentation strategies not only expanded the effective size of the training dataset but also improved the model's ability to generalize, particularly in real-world scenarios where variability in image quality is common [17].
During validation, GlauNet showed 88.9% sensitivity and 89.6% specificity, with an AUC of 0.89, which are indicators of high diagnostic accuracy [17]. The performance of the model on the training set varied between 85.30% and 90.11%, whereas the test set results were uniform, ranging from 79.86% to 87.05% [17]. Surprisingly, GlauNet also performed well even when tested against poor-quality images, with over 80% sensitivity and specificity [17]. The robustness that GlauNet demonstrates would make it suitable for practical use in clinical settings, where images are not of optimal quality all the time [17]. Moreover, the research highlights the real-world application of machine learning in ophthalmology while pointing towards future research directions: one would be the fusion of GlauNet with multiple imaging modalities and another would be its augmentation with additional functionalities by continued learning from new data [17].

2.2 MTRA-CNN: A Multi-Scale Transfer Learning Framework for Glaucoma Classification in Retinal Fundus Images

The MTRA-CNN framework proposes a new methodology for glaucoma classification based on multi-scale transfer learning advantages [10]. At a time when low data presents significant hindrance to medical imaging progress, the proposed method shines bright in being able to learn effectively from a smaller dataset [10]. Adding the Residual Attention (RA) Block boosts the ability of the model to focus more on the pertinent features while suppressing the irrelevant noise to obtain a better feature representation [10]. Training MTRA-CNN on the ODIR dataset, which is larger and more diverse, would enable transferring the learned features to the smaller glaucoma datasets [10]. This technique highlights the strength of transfer learning to overcome data limitations that commonly limit the development of strong predictive models in healthcare [10].

The research achieved an accuracy level of 86.8%, which is not only a testament to the strength of the model but also to its establishment as a benchmark for further studies in the field [10]. This result is particularly impressive given the intrinsic challenges associated with small-sized samples, where traditional training methods often face challenges [10]. The findings suggest that the MTRA-CNN framework might be a valuable tool for practitioners in the field of ophthalmology, possibly improving diagnostic accuracy and preventing late intervention in glaucoma cases [10]. Besides this, this research also leaves scope for further studies regarding multi-stage transfer learning techniques and their usability in the various ocular conditions [10]. It thus increases the research horizon in this vital healthcare area [10].

2.3 Glaucoma Detection and Staging from Visual Field Images Using Machine Learning Techniques (2024)

This study investigates the performance of deep learning (DL) models in differentiating normal and glaucomatous visual fields (VFs) and staging glaucoma into early, moderate, and advanced stages using pattern deviation (PD) plots [17]. It compares DL models against machine learning (ML) classifiers trained on conventional VF parameters, such as mean deviation (MD), pattern standard deviation (PSD), and visual field index (VFI). A dataset comprising 265 PD plots and numerical datasets from Humphrey 24–2 VF images was collected from 119 normal eyes and 146 glaucomatous eyes [17]. Popular pre-trained DL models, ResNet18 and VGG16, were employed to classify PD images using five-fold cross-validation (CV). ResNet18 trained on balanced, pre-augmented data achieved high diagnostic accuracy with an F1-score of 96.8%, precision of 97.0%, recall of 96.9%, and specificity of 99.0%. Comparatively, VGG16 achieved an F1-score of 88.7% using feature extraction techniques [17]. Grad-CAM visualization was also employed to validate model predictions and locate affected VF loss in PD plots.

The study found that DL models trained on PD plots performed on par with conventional global indices-based methods. Additionally, the research demonstrated the effectiveness of DL models in staging glaucoma according to Mills criteria, offering a promising automated tool for precision glaucoma detection and progression management. Future research directions suggested integrating DL models with real-time screening tools to further enhance diagnostic workflows [17].

2.4 Glaucoma Detection Based on Deep Convolutional Neural Network (2015)

This study introduces a deep learning (DL) architecture utilizing convolutional neural networks (CNNs) for automated glaucoma diagnosis. The proposed system analyzes fundus images to discriminate between glaucoma and non-glaucoma patterns, leveraging the hierarchical representation of images learned by CNNs. The architecture comprises six layers: four convolutional layers and two fully-connected layers, with the inclusion of dropout and data augmentation techniques to enhance diagnostic performance [12]. Extensive experiments were conducted using the **ORIGA** and **SCES** datasets, achieving area under curve (AUC) values of **0.831** and **0.887**, respectively, for glaucoma detection. These results were found to outperform state-of-the-art algorithms at the time, demonstrating the effectiveness of DL methods for glaucoma diagnosis. The study highlights the utility of CNN-based approaches in identifying early signs of glaucoma, emphasizing their potential role in clinical applications for automated and efficient detection [12].

III. Dataset

The detection and staging of glaucoma rely heavily on the availability of diverse and high-quality retinal fundus image datasets. Each dataset contributes unique features and imaging conditions, providing a comprehensive foundation for training and evaluating deep learning models [1]. However, relying on a single dataset can limit generalization due to the inherent variability in imaging equipment, protocols, and patient demographics across clinical settings [8]. To overcome this limitation, we created a unified dataset, named Uni-Glauc, by combining four prominent publicly available datasets:

RIM-ONE, ORIGA, REFUGE, and DRISHTI-GS [19] [16]. By merging these datasets, Uni-Glauc provides a robust and diverse collection of retinal images, encompassing variations in resolution, illumination, clinical settings, and patient demographics. This combined dataset is especially valuable for addressing critical challenges in glaucoma detection, such as class imbalance, small dataset size, and variability in image quality [9]. It enables comprehensive model training for both detecting glaucoma and staging its severity into four categories: Healthy, Mild, Moderate, and Severe [17] [11].

Rationale for Combining Datasets

The combination of datasets was driven by the need to:

- **Enhance Diversity:** Individual datasets may lack diversity in imaging protocols, patient populations, or disease representation. Combining datasets ensures coverage of a broader range of clinical scenarios [2].
- **Increase Dataset Size:** Small datasets often lead to overfitting and poor generalization. Merging four datasets significantly increases the sample size, improving the robustness of the models [14].
- **Address Class Imbalance:** Certain glaucoma severity levels (e.g., Mild and Severe) are often underrepresented in individual datasets. The unified structure of Uni-Glauc allows for targeted preprocessing techniques, such as oversampling and augmentation, to mitigate these imbalances [18] [6].
- **Facilitate Model Generalization:** By leveraging data from multiple sources, the resulting models are better equipped to handle variations encountered in real-world clinical environments [7].

Structure of the Uni-Glauc Dataset

To standardize the data and facilitate training, the merged dataset was categorized into four folders representing glaucoma severity levels:

- **Healthy:** Image with no signs of glaucoma.
- **Mild:** Images showing early-stage glaucomatous changes.
- **Moderate:** Images reflecting intermediate stages of glaucoma progression.
- **Severe:** Images depicting advanced glaucomatous damage.

This structured organization not only simplifies data preprocessing and augmentation but also aids in evaluating the model's performance across different severity levels [12]. Moreover, the segmentation annotations available in some of the original datasets allow for the extraction of critical features, such as the cup-to-disc ratio (CDR), further enriching the dataset’s utility [10].

Importance of Uni-Glauc in Research

The creation of Uni-Glauc bridges the gap between the capabilities of existing datasets and the requirements of advanced deep learning models for glaucoma research. It provides an opportunity to develop highly accurate and generalizable models capable of detecting and staging glaucoma across diverse populations and clinical setups [13]. Additionally, by integrating segmentation data, the combined dataset supports multi-task learning, where models can simultaneously address classification and segmentation tasks [5].

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 5, No.2, May : 2025

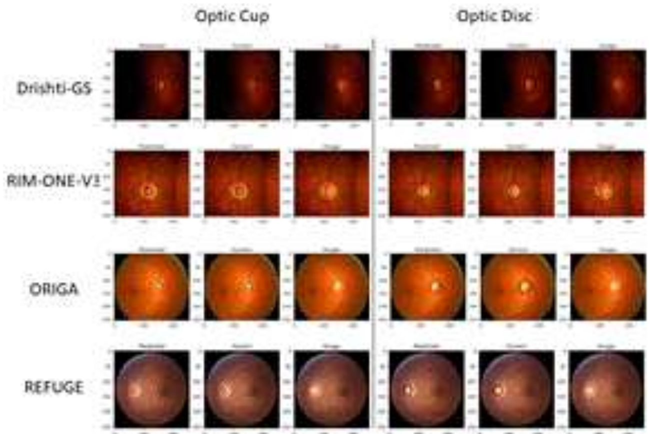


Figure 1: Sample Glaucoma Detection Algorithm. Adapted from [18].

3.1 RIM-ONE Dataset

The **RIM-ONE** dataset is a crucial resource for researchers and practitioners working on glaucoma detection through retinal fundus imaging. This dataset, containing **312 high-quality fundus images**, serves as a benchmark for developing and evaluating machine learning and deep learning models for glaucoma research. Of the total images, **172 are from clinically diagnosed glaucomatous eyes**, while the remaining are normal eyes. The balanced representation in the dataset supports the development of models capable of accurately distinguishing between glaucomatous and non-glaucomatous conditions [5] [9]. The dataset was carefully curated from three prominent hospitals in Spain: **Hospital Universitario de Canarias (HUC)** in Tenerife, **Hospital Universitario Miguel Servet (HUMS)** in Zaragoza, and **Hospital Clínico Universitario San Carlos (HCSC)** in Madrid. The diversity of data collected from these institutions enhances its representativeness and applicability to various populations and clinical settings, ensuring effective model training and evaluation [5].

To facilitate the analysis of glaucoma detection models, the RIM-ONE dataset is structured using two partitioning strategies:

- **Random Partitioning**: Images from all three hospitals are randomly assigned to the training and testing sets. This strategy ensures that the models train on a broad and diverse sample, minimizing biases and enhancing their ability to generalize.
- **Hospital-Based Partitioning**: Training is conducted solely on images from HUC, while images from HUMS and HCSC are reserved for testing. This setup simulates real-world scenarios where a model trained in one clinical setting is applied to another, highlighting the importance of robustness and transferability across healthcare environments [9] [8].

The **RIM-ONE** dataset contributes significantly to advancements in both ophthalmology and artificial intelligence, enabling the development of automated systems that assist clinicians in the early detection and effective management of glaucoma. This progress holds the potential to improve patient outcomes and advance the use of technology in healthcare [9].

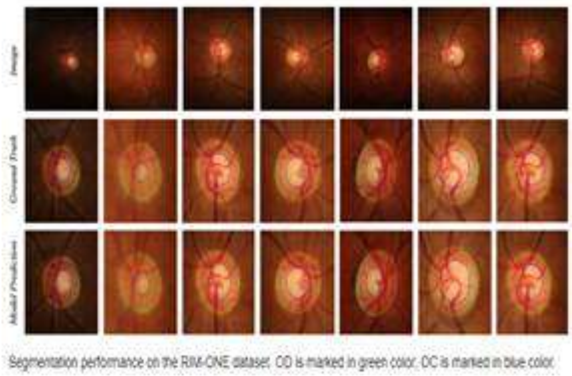


Figure 2: Segmentation performance on the RIM-ONE dataset. OD is marked in green, OC is marked in red [19].

3.2 ORIGA Dataset

The **ORIGA (Online Retinal Fundus Image Database for Glaucoma Analysis)** dataset is a key resource for advancing glaucoma research. It contains **650 high-resolution fundus images**, each carefully annotated with detailed boundaries of the optic disc and optic cup, critical for calculating the **cup-to-disc ratio (CDR)**—a key indicator in glaucoma diagnosis [1]. Of the 650 images, **168 are glaucomatous**, while the remainder are classified as normal, creating a balanced dataset for training and assessing diagnostic models [2].

In addition to annotated images, ORIGA includes clinical metadata such as intraocular pressure measurements and demographic details. These features allow researchers to explore holistic diagnostic methodologies and build state-of-the-art algorithms that integrate structural and clinical information [3] [4]. The dataset reflects diverse imaging scenarios, including variations in illumination, resolution, and patient demographics, ensuring generalizability of the models developed from it [5]. Expert annotations enhance its value for applications like optic nerve head segmentation, disease classification, and feature analysis [6].

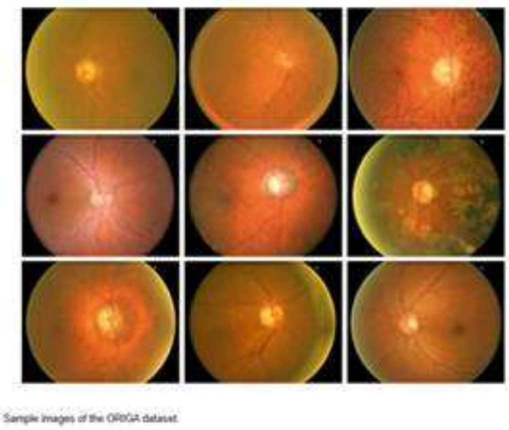


Figure 3: Sample images of the ORIGA dataset [20].

3.3 REFUGE Database

The REFUGE dataset has been specifically designed as a benchmarking resource for supporting glaucoma detection research and the development of optic nerve head analysis systems [7]. It contains **1,200 fundus images**, distributed into three folds: **400 for training**, **400 for validation**, and **400 for testing** [8]. Each image is annotated with precise boundaries of the optic disc and optic cup, enabling the computation of the **cup-to-disc ratio (CDR)**, a crucial indicator for diagnosing glaucoma [9].

The dataset includes images sourced from a variety of clinical environments, ensuring diversity in image quality, lighting conditions, and patient demographics [10]. Such variability reflects real-world clinical settings, enhancing the robustness and generalizability of machine learning models across different populations [11]. Additionally, the REFUGE dataset provides detailed clinical labels and segmentation maps, including positive and negative glaucoma indicators, further supporting various research applications such as segmentation, classification, and fully automated glaucoma detection [12][13]. The comprehensive annotations in the REFUGE dataset highlight structural differences crucial for accurate CDR calculation, aiding in the early detection and analysis of glaucomatous changes. Below is a sample image from the REFUGE dataset with the annotated optic disc and optic cup regions:

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 5, No.2, May : 2025

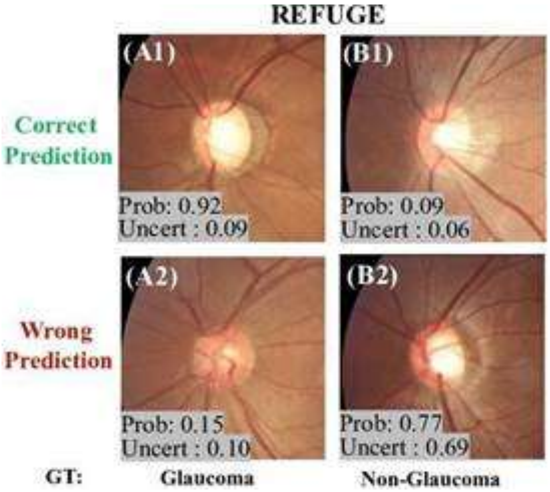


Figure 4: Sample images from the ORIGA dataset [21].

3.4 DRISHTI-GS Dataset

The **DRISHTI-GS** dataset is a specialized resource designed for glaucoma detection, particularly focusing on **optic disc (OD) and optic cup (OC) segmentation**. It consists of **101 high-resolution retinal fundus images**, categorized into **healthy** and **glaucomatous cases**. Each image is accompanied by expert annotations, including precise boundary markings for the optic disc and optic cup, enabling the accurate computation of the **cup-to-disc ratio (CDR)**—a key parameter in glaucoma diagnosis [7].

The dataset was collected under controlled clinical conditions, ensuring uniform image clarity and consistency. Despite this controlled acquisition, variations in patient demographics, imaging equipment, and clinical environments add diversity, allowing models trained on DRISHTI-GS to generalize effectively across different populations [8]. The segmentation annotations enhance the dataset's utility for **optic nerve head segmentation, glaucoma classification, and feature extraction** tasks [9].

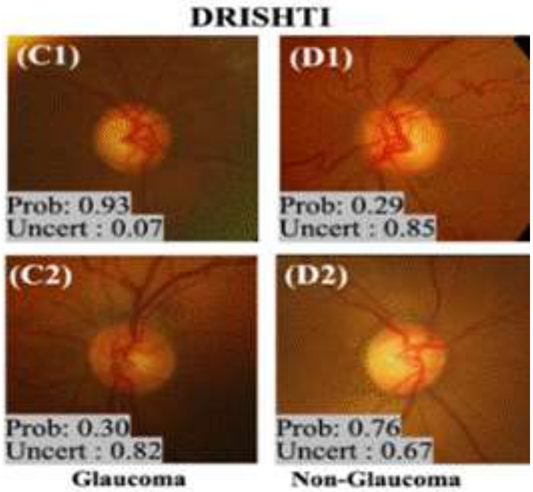


Figure 5: Sample images from the DRISHTI-GS dataset [21]

Importance of the Dataset The DRISHTI-GS dataset is widely used for developing deep learning models aimed at glaucoma assessment and segmentation. It is particularly valuable for:

- **Optic Disc and Cup Segmentation**: Providing clearly annotated boundaries for training segmentation models.
- **Classification**: Differentiating between healthy and glaucomatous eyes based on structural features.
- **Feature Analysis**: Assisting in identifying key retinal indicators critical for glaucoma progression tracking [10].

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 5, No.2, May : 2025

Its compact size and well-labeled ground truth data make it a crucial dataset for validating machine learning models and benchmarking segmentation accuracy across different methodologies.

IV. Methods

4.1 AlexNet

AlexNet, proposed by Krizhevsky et al. [6], is a pioneering deep convolutional neural network architecture that played a key role in the success of deep learning in visual recognition. It consists of five convolutional layers, some of which are followed by max-pooling layers, and three fully connected layers at the end. The architecture introduces Rectified Linear Unit (ReLU) activations for non-linearity and implements dropout in the fully connected layers to reduce overfitting [7]. In the context of this research, AlexNet is utilized as the initial detection model in a hybrid diagnostic pipeline. Its role is to perform binary classification on retinal fundus images and identify whether an input image is from a glaucomatous or healthy eye. Unlike traditional approaches that rely on manual feature extraction (e.g., cup-to-disc ratio or optic disc segmentation), AlexNet is trained end-to-end on the Uni-Glauc dataset [16][19], learning relevant discriminative features directly from raw pixel data.

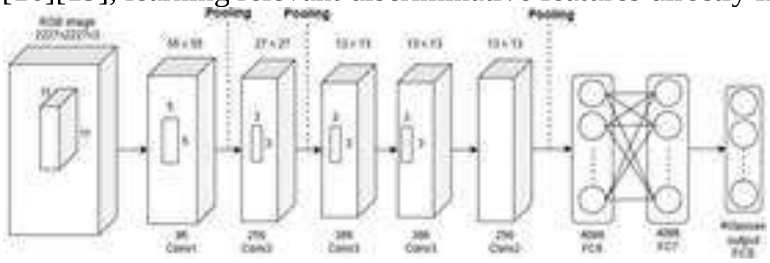


Fig. 6. Alex Net Architecture

The strength of AlexNet lies in its ability to capture spatial hierarchies of features, starting from low-level edges and textures to high-level structural patterns such as the deformation of the optic nerve head and changes in retinal texture. These patterns, though subtle, are critical indicators of early glaucomatous changes [9]. Furthermore, the use of data augmentation techniques (rotation, scaling, and flipping) enhances the generalization ability of the model across diverse clinical imaging conditions.

Training AlexNet on the unified Uni-Glauc dataset allows it to learn from a diverse set of imaging conditions and patient demographics, which improves its ability to detect glaucoma across real-world settings. The model achieved a detection accuracy of 94%, demonstrating strong capability in identifying glaucomatous patterns with high sensitivity and specificity. Given its relatively shallow depth and fast inference time, AlexNet is highly suitable for real-time integration into web-based diagnostic platforms, making it a practical choice for automated glaucoma screening in both urban and remote healthcare settings.

4.2 ResNet18

ResNet18 is a member of the Residual Network (ResNet) family developed by He et al. [11], designed to address the limitations of deep neural networks, particularly the vanishing gradient problem. Unlike traditional CNNs, ResNet18 introduces shortcut (identity) connections that bypass one or more layers, allowing the network to learn residual mappings instead of direct transformations. This design makes it possible to train deeper architectures effectively by stabilizing gradients during backpropagation [13]. In this study, ResNet18 is employed as the second phase of the hybrid framework, taking the output from AlexNet (i.e., glaucoma-positive images) and classifying them into one of three severity categories: Mild, Moderate, or Severe. This multiclass classification task is vital for clinical decision-making, as the stage of glaucoma directly influences treatment planning and follow-up protocols. The network comprises 17 convolutional layers and a final fully connected output layer, built with batch normalization, ReLU activations, and identity shortcuts. Its relatively shallow depth, compared to deeper variants like ResNet50 or ResNet101, strikes a balance between model complexity and performance, making it ideal for medium-sized datasets like Uni-Glauc [9][10]. The model is trained

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 54, Issue 5, No.2, May : 2025

to identify subtle, progressive structural changes in the optic disc and surrounding retina that are indicative of disease severity. These include features like the loss of retinal nerve fiber layer, neuroretinal rim thinning, and vascular displacement—visual cues that often go unnoticed during early clinical assessment. Unlike rule-based methods, ResNet18 learns these features automatically from labeled fundus images, enhancing the objectivity and consistency of glaucoma staging.
Additionally, ResNet18's architecture promotes better gradient flow and faster convergence, which is particularly advantageous in medical image analysis tasks where labeled data is limited and high precision is critical [13]. Trained on the stratified and augmented Uni-Glauc dataset, ResNet18 achieved a staging accuracy of 90%, proving effective at distinguishing between the mild, moderate, and severe stages of glaucoma — a classification that is often difficult even for trained ophthalmologists. By integrating ResNet18 into the diagnostic pipeline, the system supports fine-grained clinical assessment and contributes to personalized patient care. Its relatively lightweight nature also allows for deployment in resource-constrained environments, enabling widespread adoption of AI-assisted glaucoma staging.

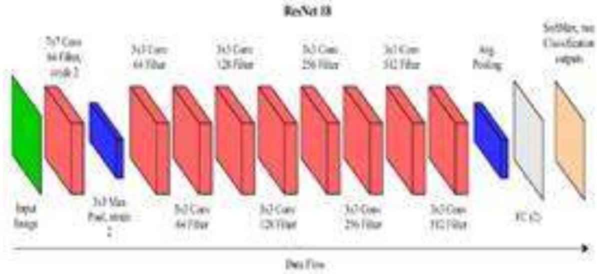


Fig.7. ResNet18 Architecture

4.4 Hybrid Model and Innovation

This research introduces a novel hybrid deep learning model for automated glaucoma detection and staging, offering a significant advancement over traditional single-model systems. The innovation lies in combining the strengths of multiple convolutional neural networks (CNNs), where each model is dedicated to a specific diagnostic task — detection and staging — to enhance both accuracy and interpretability. Rather than relying on a single network to handle the entire classification pipeline, this hybrid approach modularizes the process, allowing for optimized performance in each diagnostic phase.

Hybrid Model Architecture

The proposed framework integrates two well-established deep learning architectures — AlexNet and ResNet18 — into a unified diagnostic pipeline. Each model plays a distinct role: AlexNet performs initial detection of glaucoma, and ResNet18 is responsible for staging the condition based on image features.

Phase 1: Glaucoma Detection using AlexNet

In the first step of the pipeline, AlexNet is employed to classify retinal fundus images as either *glaucomatous* or *healthy*. AlexNet's architecture, comprising five convolutional layers and three fully connected layers, is known for its strong feature extraction capabilities. It is particularly effective in identifying important visual indicators of glaucoma such as optic disc cupping and cup-to-disc ratio (CDR) anomalies.

- Input: Retinal fundus images
- Output: Classification — *Glaucomatous* or *Healthy*

AlexNet was trained on the Uni-Glauc dataset — a unified and diverse compilation of four major retinal fundus image datasets (RIM-ONE, ORIGA, REFUGE, and DRISHTI-GS). This large and varied dataset enabled the model to learn from a wide range of clinical scenarios, imaging conditions, and severity levels. As a result, the detection model achieved high generalizability and robustness across different patient demographics.This initial detection step ensures that only glaucoma-positive cases are passed on for further staging, thereby streamlining the diagnostic process.

Phase 2: Glaucoma Staging using ResNet18

For the images classified as glaucomatous by AlexNet, ResNet18 is used to grade the severity of glaucoma into three distinct stages: *Mild*, *Moderate*, or *Severe*. ResNet18 utilizes residual learning blocks, which help train deeper networks without the vanishing gradient problem. This enables the model to capture more nuanced and detailed features from fundus images, essential for accurate staging.

- Input: Glaucoma-positive images (from AlexNet output)
- Output: Staging — *Mild*, *Moderate*, or *Severe*

The ResNet18 model was also trained on the Uni-Glauc dataset, leveraging its detailed image annotations and balanced class distribution. This helped the staging model learn subtle variations between severity levels — such as nerve fiber layer thinning or progressive optic nerve damage — that are often challenging to detect visually or with shallow networks.

This phase is critical in assisting ophthalmologists in determining the appropriate treatment strategy.

Real-Time Web-Based Integration

To ensure practical usability, the hybrid model is deployed within a web-based diagnostic platform. This platform enables ophthalmologists and healthcare professionals to upload retinal fundus images and receive instant diagnostic results, including both glaucoma classification and severity staging. The intuitive interface enhances clinical workflows and supports early intervention by providing automated insights that can assist in decision-making.

This hybrid deep learning approach not only enhances diagnostic efficiency and accuracy but also aligns with the growing need for AI-assisted tools in medical imaging. By decoupling detection and staging tasks and assigning them to models best suited for each, the system delivers more reliable and interpretable results, thereby contributing to improved patient outcomes and the advancement of AI in ophthalmology.

4.5 Invention and Novelty

The innovation of this work lies in the hybrid deep learning approach used for automated glaucoma detection and staging. Unlike conventional systems that either focus solely on glaucoma detection or apply a single model for both detection and classification, our hybrid methodology introduces a modular, multi-phase system that assigns specialized tasks to distinct models — AlexNet for detection and ResNet18 for severity staging. This structured pipeline enhances the system's diagnostic precision, scalability, and real-world usability.

Hybrid Two-Phase Architecture

The proposed framework breaks down the diagnostic process into two clearly defined stages:

- Phase 1: Detection using AlexNet, which classifies retinal fundus images as *glaucomatous* or *healthy*.
- Phase 2: Staging using ResNet18, which further categorizes glaucoma-positive images into *Mild*, *Moderate*, or *Severe* stages.

This division of responsibility allows each network to concentrate on its specialized task, improving accuracy and reducing computational complexity. The decoupling ensures that detection errors do not cascade into the staging phase, while staging benefits from more focused feature refinement.

Beyond Binary: Multiclass Glaucoma Staging

A significant novelty of this work is its emphasis on multiclass glaucoma staging, moving beyond traditional binary classification (Glaucoma vs. Non-glaucoma). By classifying glaucoma into three clinical severity levels — *Mild*, *Moderate*, and *Severe* — the system offers granular diagnostic insight, which is crucial for determining personalized treatment paths and monitoring disease progression over time. This level of classification is often overlooked in automated systems but is highly valued in clinical ophthalmology.

Real-Time and Scalable Deployment

The hybrid model has been integrated into a web-based diagnostic application, enabling real-time glaucoma analysis through image upload. This makes the solution both scalable and accessible, as healthcare providers can perform automated diagnostics without requiring high-end infrastructure or manual image interpretation. The system's architecture supports efficient inference, ensuring rapid feedback that is vital for timely medical intervention.

Deep Feature Extraction with ResNet18

A key innovation is the use of ResNet18 for the staging phase. Its deep residual architecture enables the extraction of subtle and complex patterns from fundus images — such as nerve fiber thinning and optic disc damage — that are difficult to detect through shallow networks or manual observation. This contributes significantly to the staging model's reliability and clinical relevance. In conclusion, the hybrid model proposed in this study introduces a novel hybrid solution that bridges the gap between accurate glaucoma detection and detailed severity staging. By combining the feature extraction strength of AlexNet and the deep refinement capability of ResNet18, the system represents a robust, interpretable, and scalable advancement in the field of ophthalmic AI diagnostics.

V. Results and Discussions

The performance of the proposed hybrid deep learning model was evaluated using the Uni-Glauc dataset — a merged collection of RIM-ONE, ORIGA, REFUGE, and DRISHTI-GS datasets — to ensure diversity in imaging conditions, clinical settings, and patient demographics. The model operates in two phases: glaucoma detection using AlexNet and glaucoma severity staging using ResNet18.

5.1 Glaucoma Detection Performance (AlexNet)

For the binary classification task (healthy vs. glaucomatous), AlexNet achieved an accuracy of 94%. This performance demonstrates the model's robustness in detecting glaucoma across a wide range of image types and clinical conditions present in the Uni-Glauc dataset. The relatively shallow yet powerful architecture of AlexNet allows for efficient feature extraction, especially when detecting prominent glaucoma indicators like optic disc cupping and cup-to-disc ratio anomalies.

The high detection accuracy confirms AlexNet's effectiveness in differentiating between healthy and glaucomatous eyes, making it a suitable first stage in the diagnostic pipeline. By correctly identifying glaucoma-positive cases early, the system ensures that only relevant images proceed to the staging phase, optimizing computational efficiency and clinical relevance.

5.2 Glaucoma Staging Performance (ResNet18)

In the second phase, ResNet18 was employed to classify glaucomatous images into three stages: Mild, Moderate, and Severe. The model achieved an overall accuracy of 90% in this multiclass classification task.

ResNet18's deep residual connections enabled the model to learn subtle differences in retinal features associated with disease progression — such as nerve fiber layer thinning and optic nerve head damage. The model showed particularly strong performance in distinguishing moderate and severe cases, where visual cues are more pronounced. However, as with many multiclass medical image classification tasks, early-stage (mild) cases posed more challenges due to their less obvious pathological features.

5.3 Interpretation and Clinical Implications

The hybrid model's two-phase architecture provides several practical advantages:

- **Improved Accuracy and Modularity:** By separating detection and staging tasks into two specialized models, the system maintains high accuracy in both phases without overburdening a single network.
- **Efficient Clinical Workflow:** Only glaucoma-positive images undergo staging, reducing computational time and focusing clinical attention where it's needed most.
- **Enhanced Diagnostic Insight**: Multiclass staging offers ophthalmologists more detailed insights into disease severity, aiding in personalized treatment planning and ongoing patient monitoring.

5.4 Comparative Advantage

Compared to traditional single-model approaches or prior studies relying on individual datasets, this hybrid system — trained on a unified and diverse dataset — demonstrates improved generalizability and clinical relevance. The results indicate that even relatively lightweight networks like AlexNet and

ResNet18 can outperform more complex models when trained on well-prepared, diverse data and structured in a task-optimized architecture.

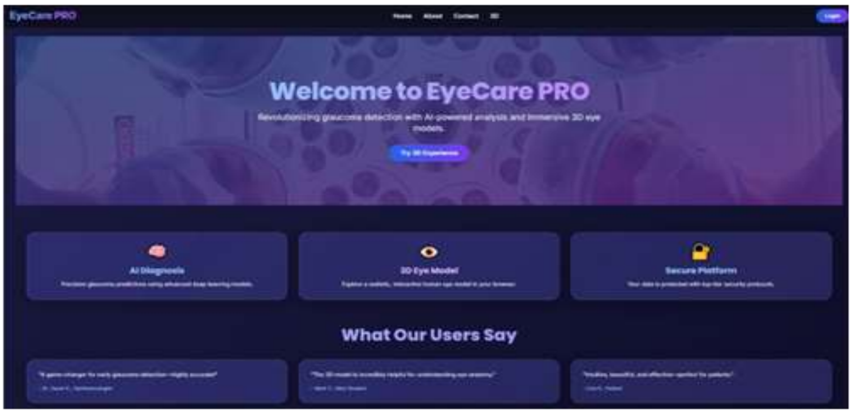
5.5 Web-Based Diagnostic Platform

To enhance clinical usability and support real-time diagnosis, the trained hybrid deep learning models were deployed through a custom-built web-based application. This platform allows ophthalmologists, clinicians, and researchers to upload retinal fundus images and instantly receive diagnostic feedback including:

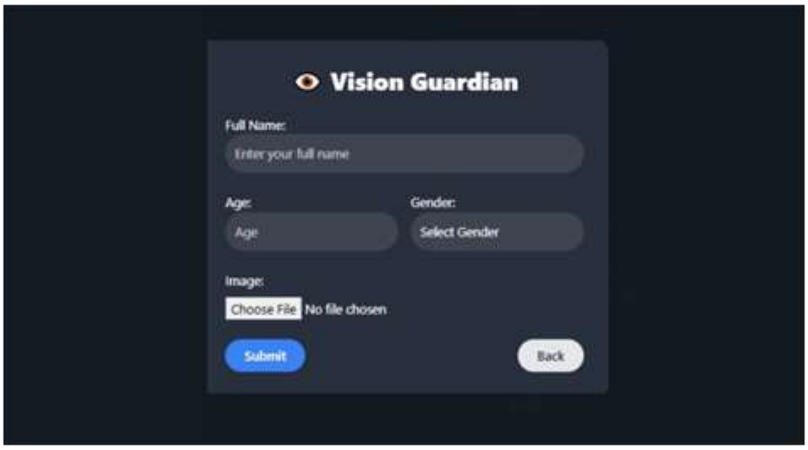
- Glaucoma Detection (Healthy or Glaucomatous) — powered by AlexNet
- Glaucoma Severity Staging (Mild, Moderate, Severe) — powered by ResNet18

Visual Overview of the EyeCare PRO Web Application

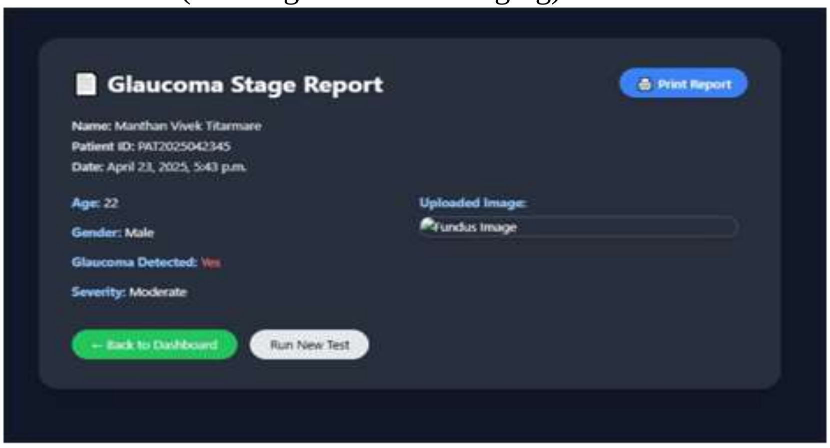
1. Homepage/Dashboard view



2. Image Upload interface



3. Prediction results screen (showing detection + staging)



4. Report Generation & Printing

5. Record Searcher

VI. Conclusion

This research presents a structured, hybrid deep learning-based system for automated glaucoma detection and staging, integrating AlexNet and ResNet18 into a unified diagnostic pipeline. The proposed two-phase model ensures efficient and accurate classification by delegating glaucoma detection to AlexNet and severity staging to ResNet18. This modular design not only enhances diagnostic precision but also supports interpretability by assigning specific roles to each model.

Experimental results demonstrate that both AlexNet and ResNet18 perform effectively within their respective phases, validating the importance of using specialized architectures tailored to distinct medical imaging tasks. The hybrid approach successfully addresses the limitations of single-model classifiers by enabling detailed glaucoma staging — Mild, Moderate, Severe — which is critical for timely intervention and treatment planning.

Future improvements may include expanding the dataset diversity, incorporating explainable AI (XAI) techniques to improve model transparency, and further optimizing the web-based deployment for real-time clinical applications. This study contributes to the growing field of AI in ophthalmology by offering a scalable, accurate, and interpretable solution for automated glaucoma diagnosis.

References

[1] Lindberg D V & Lee H K H (2015), ‘Optimization under Constraints by Applying an Asymmetric Entropy Measure’, *Journal of Computational and Graphical Statistics*, pp. 379–393.

[2] Rieder B (2020), *Engines of Order: A Mechanology of Algorithmic Techniques*, Amsterdam University Press, Amsterdam, Netherlands.

[3] Boglaev I (2016), ‘A Numerical Method for Solving Nonlinear Integro-Differential Equations of Fredholm Type’, *Journal of Computational Mathematics*, pp. 262–284.

[4] GBD 2019 Blindness and Vision Impairment Collaborators (2021), 'Causes of Blindness and Vision Impairment in 2020 and Trends Over 30 Years, and Prevalence of Avoidable Blindness in Relation to VISION 2020: The Right to Sight: An Analysis for the Global Burden of Disease Study', *Lancet Global Health*, pp. e144–e160.

[5] Elangovan P & Nath M K (2021), 'Glaucoma Assessment from Color Fundus Images Using Convolutional Neural Network', *International Journal of Imaging Systems and Technology*, pp. 955–971.

[6] Deng J, Dong W, Socher R, Li L-J, Li K & Fei-Fei L (2009), 'ImageNet: A Large-Scale Hierarchical Image Database', *IEEE Conference on Computer Vision and Pattern Recognition*, pp. 248–255.

[7] Hao J, Li F, Hao H, Fu H, Xu Y, Higashita R, Zhang X, Liu J & Zhao Y (2021), 'Hybrid Variation-Aware Network for Angle-Closure Assessment in AS-OCT', *IEEE Transactions on Medical Imaging*, pp. 123–134.

[8] Yi S, Zhou L, Ma L & Shao D (2023), 'MTRA-CNN: A Multi-Scale Transfer Learning Framework for Glaucoma Classification in Retinal Fundus Images', *IEEE Access*, pp. 142689–142701.

[9] Aamir M, Irfan M, Ali T, Ali G, Shaf S A, Saeed A A, Al-Beshri A, Alasbali T & Mahnashi M H (2020), 'An Adoptive Threshold-Based Multi-Level Deep Convolutional Neural Network for Glaucoma Eye Disease Detection and Classification', *Diagnostics*, pp. 602–617.

[10] Plested J & Gedeon T (2022), 'Deep Transfer Learning for Image Classification: A Survey', *arXiv*, pp. 1–15.

[11] Liu J, Guo F, Gao H, Huang Z, Zhang Y & Zhou H (2021), 'Image Classification Method on Class Imbalance Datasets Using Multi-Scale CNN and Two-Stage Transfer Learning', *Neural Computing and Applications*, pp. 14179–14197.

[12] Chen X, Xu Y, Wong D W K, Wong T Y & Liu J (2015), 'Glaucoma Detection Based on Deep Convolutional Neural Network', *Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 715–718.

[13] Fan R, Bowd C, Brye N, Christopher M, Weinreb R N, Kriegman D J & Zangwill L M (2020), 'One-Vote Veto: Semi-Supervised Learning for Low-Shot Glaucoma Diagnosis', *IEEE Transactions on Medical Imaging*, pp. 3764–3778.

[14] [14] Liu H, Li L, Wormstone I M, Qiao C, Zhang C & Xu J (2019), 'Development and Validation of a Deep Learning System to Detect Glaucomatous Optic Neuropathy Using Fundus Photographs', *JAMA Ophthalmology*, pp. 1353–1360.

[15] Joshi S, Partibane B, Hatamleh W A, Tarazi H, Yadav C S & Krah D (2022), 'Glaucoma Detection Using Image Processing and Supervised Learning for Classification', *Journal of Healthcare Engineering*, pp. 1–8.

[16] Devecioglu O C, Malik J, Ince T, Kiranyaz S, Atalay E & Gabbouj M (2021), 'Real-Time Glaucoma Detection from Digital Fundus Images Using Self-ONNs', *IEEE Access*, pp. 140031–140041.

[17] Akter N, Gordon J, Li S, Poon M, Perry S, Fletcher J, Chan T, White A & Roy M (2025), 'Glaucoma Detection and Staging from Visual Field Images Using Machine Learning Techniques', *PLOS ONE*, pp. e0316919.

[18] Liu Y, Zhang D & Sun C (2023), 'Glaucoma Detection Using Machine Learning', *Sensors*, pp. 4668–4684.

[19] Liu J (2019), 'Segmentation Performance on the RIM-ONE Dataset, OD Is Marked in Green Color, OC Is Marked in Red Color', *ResearchGate*, pp. 1–10.

[20] ResearchGate (2020), 'Sample Images of the ORIGA Dataset', *ResearchGate*, pp. 1–5.

[21] Springer (2020), 'Sample Images of the ORIGA and DRISHTI-GS Datasets', *Springer*, pp. 1–10.