Study on Glaucoma Detection using Retinal Imaging

Piyush Bhokre, Rahul Khandate, Shivam Rakhunde and Tanmay Sorte Guided by Mrs.Kirti Digholkar
Pune Institute of Computer Technology, Pune 411043, India

Abstract. Glaucoma, a leading cause of irreversible blindness, progresses without obvious symptoms, making early detection essential. This project focuses on developing an automated system that analyzes retinal images to classify them as glaucomatous or non-glaucomatous using machine learning techniques. By extracting key retinal features and leveraging a labeled dataset, the system aims to provide accurate and timely glaucoma detection. The solution will offer a user-friendly interface for easy integration into clinical workflows, enhancing diagnostic accuracy and supporting early intervention to prevent vision loss. ...

Keywords: Fundus Image, OCDR, CNN, U-NET, Deep Learning, Optical Detection, Image Processing, Diagnostic Solutions.

1 INTRODUCTION

Glaucoma is a progressive eye disease that has devastating consequences if left undiagnosed and untreated. As one of the leading causes of permanent blindness globally, it affects an estimated 80 million people, with projections suggesting an increase due to aging populations. Often referred to as the "silent thief of sight," glaucoma progresses stealthily, without noticeable symptoms or pain, making it difficult for individuals to recognize the onset of the disease. By the time vision impairment is detected, significant damage to the optic nerve has already occurred, and unfortunately, this damage is irreversible. This delayed detection means that many patients only seek treatment when their vision is already permanently impaired, and the possibility of restoring lost sight is no longer an option.

The global impact of glaucoma is immense, particularly in older adults, where vision loss can lead to loss of independence, reduced mobility, and diminished quality of life. The effects are even more pronounced in regions with limited access to eye care services. In such areas, the lack of early detection and timely intervention contributes to a higher prevalence of undiagnosed and untreated glaucoma, further increasing the incidence of preventable blindness. For healthcare systems already under strain, this growing burden underscores the critical need for more effective, efficient, and accessible diagnostic methods.

Early detection is vital in the fight against glaucoma, as timely diagnosis allows for treatment strategies—such as medications, laser procedures, or surgery— that can slow the disease's progression and, in some cases, prevent further vision loss.However, current standard detection methods, such as measuring intraocular pressure (IOP) or assessing peripheral vision, often fail to detect glaucoma in its early stages. This is because the disease can develop even when IOP readings are normal, and vision loss in the initial stages may not be noticeable. As a result, traditional diagnostic tools are limited in their ability to identify glaucoma early enough for intervention to be most effective. This diagnostic gap highlights the urgent need for more reliable, sensitive, and accessible screening techniques that can detect the disease before significant damage occurs.

Advances in medical technology, particularly in the fields of imaging and artificial intelligence (AI), present new opportunities to address these challenges. In recent years, machine learning (ML)—a subset of AI where systems learn from large datasets—has emerged as a powerful tool[12] for enhancing diagnostic accuracy in healthcare. Specifically, ML algorithms can be trained to analyze retinal images, detecting minute changes in the optic nerve and retinal structure that signal the onset of glaucoma. These advanced systems can outperform traditional diagnostic methods by identifying early warning signs that are invisible to the naked eye, making it possible to diagnose glaucoma with greater precision and speed.

By automating the process of glaucoma detection, these technologies offer significant benefits to both patients and healthcare providers. For patients, early detection can lead to earlier treatment, potentially preventing the irreversible vision loss that characterizes glaucoma. Early diagnosis can also reduce the psychological and emotional burden associated with vision impairment, helping individuals maintain their quality of life for longer. For healthcare providers, automated systems can streamline the diagnostic process, allowing for faster and more accurate screenings. This can be particularly beneficial in settings where ophthalmologists are scarce, as these systems can extend the reach of eye care services to remote and underserved populations. In regions where resources are limited, the ability to conduct mass screenings with minimal infrastructure can transform the approach to managing glaucoma, ensuring that more people are screened and treated before the disease progresses.

The focus of this research is to develop an innovative and efficient system that leverages retinal imaging and machine learning to detect glaucoma at an early stage. By utilizing cutting-edge algorithms capable of analyzing retinal features with high accuracy, this system aims to provide a tool that not only improves diagnostic precision but also integrates seamlessly into clinical workflows[6]. The proposed system will be designed with ease of use in mind, making it suitable for widespread implementation in both well-resourced and resource-constrained environments. Through early and accurate detection, this system has the potential to significantly reduce the global burden of glaucoma, improving patient outcomes and preserving the vision of millions of people around the world. Moreover, by enhancing the accessibility of glaucoma screening, it can contribute to reducing health disparities and improving eye care for populations that currently face barriers to receiving timely diagnosis and treatment. In conclusion, this research represents a critical step forward in the application of machine learning to healthcare, particularly in addressing the challenges of glaucoma detection. With the potential to prevent blindness and improve the quality of life for those at risk, the development of an automated glaucoma detection system offers a promising solution to one of the most pressing public health concerns in ophthalmology today.

2 LITERATURE SURVEY

The reviewed papers here for glaucoma detection from fundus[11] using machine learning concepts and OCTA images [4] have shown significant promise in enhancing diagnostic accuracy. Several studies present a variety of techniques that leverage deep learning, image processing, and optimization strategies to improve the detection and classification of glaucoma. One robust framework combines Retinex theory with convolutional neural networks (CNNs) and Design of Experiments (DOE)[13] to enhance image quality and optimize CNN parameters. This approach achieved a sensitivity of 0.95, specificity of 0.98, and accuracy of 0.97 across 1,450 images, showcasing the potential for high-precision automated screening. Another study focuses on determining the vertical cup-to-disc ratio (VCDR)[7], a critical marker for glaucoma, using shape-selective filters and generalized matrix learning vector quantization. This method demonstrated a mean VCDR difference of 0.11 compared to manual annotations across 1,712 images.

In a further development, an ensemble of CNNs—including ResNet, DenseNet, and Inception—was employed for joint segmentation of the optic disc and cup, achieving high segmentation accuracy with a Dice score of 0.961 for the optic disc and 0.894 for the optic cup. This ensemble model also accurately classified glaucoma stages based on the cup-to-disc ratio, further advancing the precision of glaucoma detection. Additionally, Self-Organized Operational Neural Networks (Self-ONNs)[3] were introduced for real-time glaucoma detection, outperforming traditional CNNs in terms of both accuracy and computational efficiency on datasets like ACRIMA and RIM-ONE, offering lower complexity for real-world application.

Alongside these machine learning innovations, advanced image processing techniques continue to play a crucial role. Preprocessing methods such as adaptive histogram equalization are used to correct illumination variations and reduce noise, improving the quality of retinal fundus images. Techniques like morphological operations and edge detection remove blood vessels to facilitate a clearer analysis of key anatomical structures. Localization of the optic nerve head (ONH) remains central to identifying anatomical changes associated with glaucoma. Methods such as Principal Component Analysis (PCA), Hough Transform, and Pyramidal Decomposition are used to accurately pinpoint the ONH, while segmentation techniques help define the boundaries of the optic disc (OD), which is critical for monitoring disease progression. Models like Geometric Active Contour, Genetic Active Contour, and Gradient Vector Flow (GVF), along with Hu's Circular Model, improve segmentation precision, contributing to early glaucoma detection and reducing the risk of visual impairment.

Another study uses an Attention U-Net with ResNet50 for segmentation and a modified Inception V3 for classification. The segmentation model achieves high accuracy on the RIM-ONE dataset, with 99.58% for optic disc segmentation and 98.05% for optic cup segmentation. Heatmaps generated through GradCAM and Grad-CAM++ visualize regions influencing the diagnosis, enhancing model interpretability. The classification subsystem, based on segmented images, achieves 98.97% accuracy, 99.42% sensitivity, and 95.59% specificity, making it a strong candidate for automated glaucoma diagnosis. This two-subsystem approach, combining feature extraction and transfer learning, offers better performance than previous models.

Additionally, other deep learning architectures, such as EfficientNet, MobileNet, DenseNet, and GoogLeNet, have been tested for glaucoma detection. EfficientNet-b3 yielded the best results, with a test accuracy of 96.52% and an ROC AUC of 0.9574. Blood vessel segmentation[8] from fundus images was also explored as a potential alternative, with a MobileNet v3 achieving 83.48% accuracy. An alternative method uses a Computer-Aided Detection (CADe)[10] system with eye-tracking to detect visual field loss in glaucoma patients during real-life activities. The system tracks eye movements and displays performance maps like Gaze Fusion Map(GFM),Gaze Fusion Reaction Time Map (GFRT), and Gaze Convex Hull Map(GCHM), providing healthcare professionals with visual guides for screening. The system's explainability techniques, such as Gaze Health Exploration (GE-i)[9], distinguish glaucoma from normal conditions, further aiding diagnosis.

In parallel, CNN-based models specifically designed for OCTA imaging[4] are proving effective in diagnosing glaucoma. These models feature a two-stage architecture, with the first stage focused on feature extraction through convolutional layers and the second stage dedicated to classification. These models achieve high diagnostic accuracy by analyzing critical regions like the optic nerve head[5] and retinal nerve fiber layers, and they demonstrate superior performance over traditional models such as VGG16 and ResNet50, particularly in cases with low-quality images.

Publicly available datasets, such as RIGA and DRISHTI-GS [14], have been instrumental in training and validating machine learning models for glaucoma detection. These datasets help enhance segmentation and classification tasks, and model performance is typically evaluated using metrics such as accuracy, sensitivity, and specificity. Despite these advances, challenges remain in fully automating the detection process, especially in maintaining consistent accuracy across varying image qualities and addressing the complexities of real-time screening in resource-limited settings. Nonetheless, these machine learning and image processing techniques represent a significant step forward in the development of efficient, accurate, and automated systems for glaucoma detection and diagnosis.

3 METHODLOGIES EXPLORED

CNN :-

This reviewed model here is designed to classify OCTA images of the eye into glaucomatous or non-glaucomatous categories. The CNN architecture consists of two main sections: Feature-Extraction Section: This section has three convolutional layers (CONV), each followed by a Rectified Linear Unit (ReLU) and a Max Pooling (MP) layer. The convolutional layers capture important image features by filtering the input OCTA images, highlighting specific regions like the optic nerve head and retinal nerve fiber layers. The outputs from these layers undergo max pooling, which selects the most prominent features by reducing the dimensionality and focusing on essential visual elements. These layers help extract dominant patterns in the image that are indicative of glaucoma, such as reduced blood vessel density. Classification Section: After feature extraction, the network has five fully connected (FC) layers. These layers take the extracted features, flatten them into a vector, and then classify the image into one of two categories: glaucomatous or non-glaucomatous. The final layer uses a sigmoid activation function to predict the probability of the image belonging to each class.

1. Input Data and Pre-processing - The CNN processes retinal images (307x307 pixels) from OCTA scans. To enhance the robustness of the model, various image augmentation techniques, such as horizontal flipping, rotation, and brightness adjustments, are applied to the training data. This ensures that the model can handle different imaging conditions and settings.

2. Image Features and Focus Areas - The network focuses on the optic nervehead and retinal nerve fiber layers, two key areas impacted by glaucoma. Specifically, the feature extraction layers highlight regions with reduced blood vessel density, a sign of glaucoma progression. By emphasizing these areas, the CNN is able to detect subtle changes that may indicate the presence of the disease.

3. Robustness to Image Artifacts - The system is tested on both high-qualityand poor-quality images. Even with noisy or artifact-ridden images (due to eye movement, defocus, or shadowing), this model maintains accuracy above 80%, showing its ability to handle real-world variations.

4. Training and Validation - The model is trained using a large dataset ofOCTA images, split into training and testing sets. During training, 5-fold crossvalidation is used to fine-tune the architecture and select optimal hyperparameters, ensuring the model performs well across different data samples. The model was trained for 100 epochs using the Adam optimizer, with a learning rate of 0.0001, ensuring stable convergence during training.

5. Performance Evaluation - The CNN's performance is measured using metrics such as sensitivity, specificity, and accuracy. This model achieves high sensitivity (88.9%) and specificity (89.6%), which indicates its ability to correctly identify both glaucomatous and non—glaucomatous eyes.

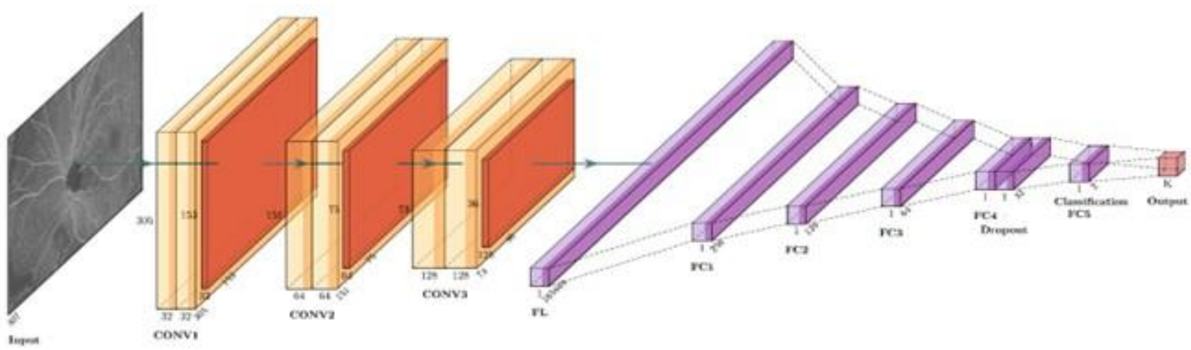


Fig.1: CNN Architecture [13]

U-Net for Segmentation :-

U-Net is a convolutional neural network (CNN) architecture specifically designed for image segmentation, meaning it divides an image into meaningful parts or regions. For glaucoma detection, it is used to segment the optic disc (OD) and optic cup (OC) from fundus images. Here is how U-Net achieves this:

1) Encoder (Down-sampling Path): This part is similar to a typical CNN where the input image is passed through several convolutional layers. Each layer applies convolution filters to the image, followed by activation functions (like ReLU) and max pooling to reduce the image size (down-sampling). The purpose of this down-sampling is to extract high-level features while reducing the spatial resolution. These features help identify key structures, like the boundaries of the optic disc and cup.

2) Bottleneck Layer: The deepest part of the U-Net, this layer captures the most abstract features from the image while maintaining important information about object boundaries.

After down-sampling, the image needs to be reconstructed back to its original resolution for pixel-wise classification (segmentation). This is done by upsampling the feature maps through transposed convolutions. During this process, the decoder concatenates the corresponding feature maps from the encoder (called skip connections). These skip connections ensure that the decoder has access to both high-level abstract features and lower-level, detailed features, which improves the accuracy of segmentation.

3) Decoder (Up-sampling Path): After down-sampling, the image needs to be reconstructed back to its original resolution for pixel-wise classification (segmentation). This is done by up-sampling the feature maps through transposed convolutions. During this process, the decoder concatenates the corresponding feature maps from the encoder (called skip connections). These skip connections ensure that the decoder has access to both high-level abstract features and lowerlevel, detailed features, which improves the accuracy of segmentation. The result is a pixel-wise classification where each pixel is assigned a class, i.e., whether it belongs to the optic cup, optic disc, or background.

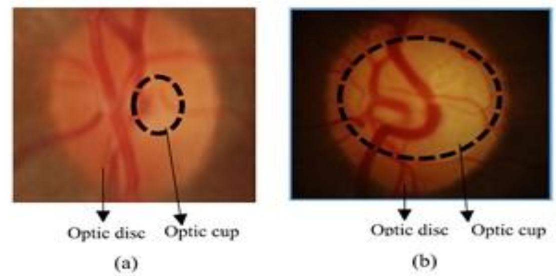


Fig.2: Retinal Fundus Image showing Optic cup and disc [1]

4) Final Output Layer: The last layer applies a 1x1 convolution to the output of the decoder, which assigns each pixel a class (optic cup, optic disc, or background). In glaucoma detection, this step isolates the optic cup and optic disc regions from the rest of the fundus image.

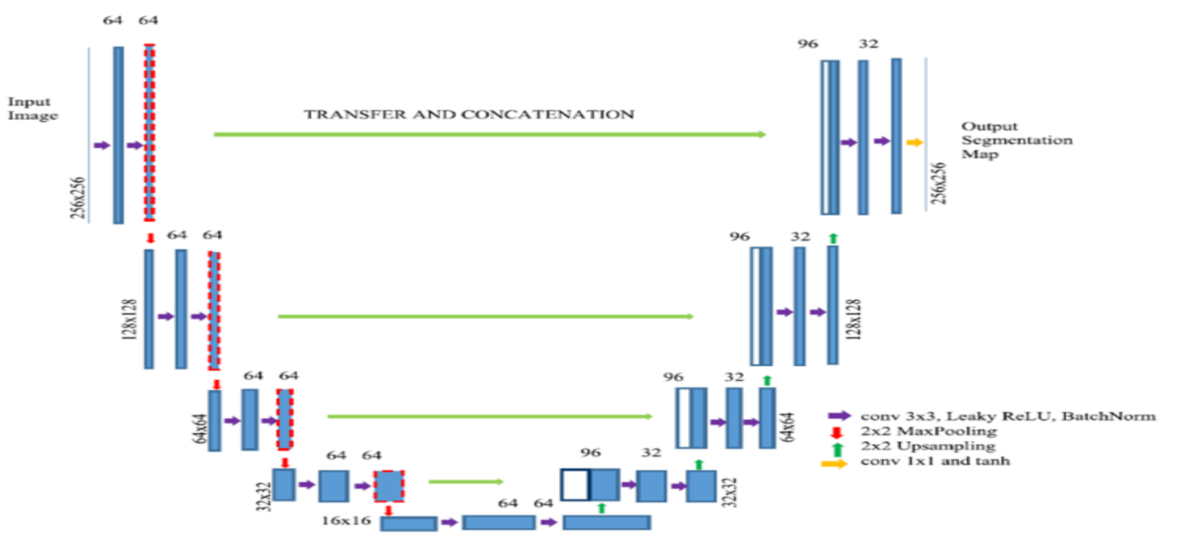


Fig.3: U-Net Architecture [1]

XGBoost (XGB) for Classification[1] :-

Once U-Net has segmented the optic disc and cup from the fundus images, Extreme Gradient Boosting (XGB) is used to classify the images into glaucomatous or non-glaucomatous. XGB is a machine learning algorithm that belongs to the family of gradient boosting methods, which are powerful for classification tasks. Here‘s how it works in this context:

1) Feature Extraction: After segmentation, specific features are extracted from the optic disc and cup. These features typically include: - Vertical, horizontal, and diagonal lengths of the optic cup and disc. - Cup-to-Disc Ratio (CDR): A key indicator of glaucoma where a larger optic cup relative to the disc may indicate the disease. - Vertical separations between the optic disc and cup are also used to capture small changes in the optic nerve structure, which are important for detecting glaucoma.

2) Input to XGB: These extracted features are provided as input to the XGB model. The model doesn’t process images directly; instead, it works with numerical features representing the geometry of the optic disc and cup.

3) Boosting Technique: XGB works by creating an ensemble of weak decision trees, each trained to

correct the errors of the previous ones. The idea is to build a strong predictive model by combining multiple weak learners. At each step, XGB assigns a weight to each sample and adjusts these weights based on the errors made by the model. The samples that were wrongly classified by the previous tree get higher weights, and the model gives them more attention in subsequent iterations. Through gradient boosting, the model minimizes the error function by updating itself using the gradient of the loss function (which measures how far the model is from the correct classification).

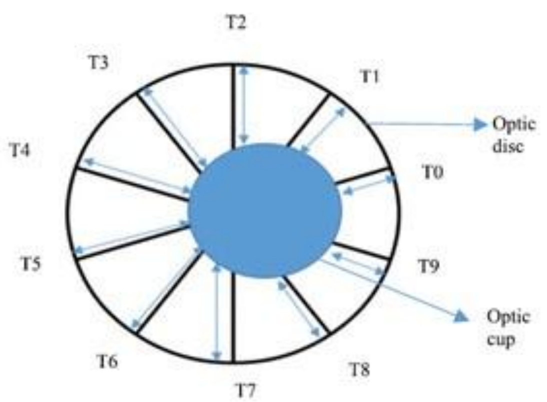
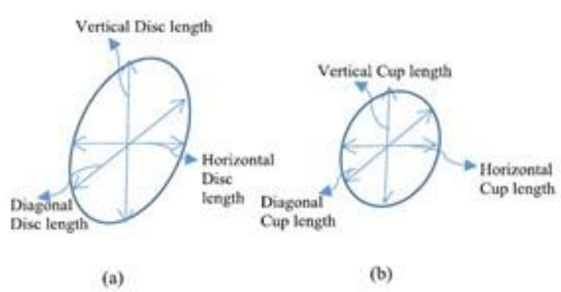


Fig.4: Retinal Disc and Cup Measures [1]



4) Classification: Finally, XGB outputs a probability or score indicating whether a given fundus image is likely to be glaucomatous or non-glaucomatous. It does this by aggregating the outputs of all the decision trees, where each tree contributes to the final decision.The model achieves high accuracy due to its ability to handle complex relationships between features, making it ideal for medical image classification tasks like glaucoma detection.

Grad-CAM and Grad-CAM++ :-

Method: Explainable AI (XAI) techniques like Grad-CAM and Grad-CAM++ are used to make the predictions of deep learning models interpretable by generating heatmaps that highlight important areas of an image. These methods work by calculating the gradients of the output (for a specific class, such as "glaucoma") with respect to the feature maps of the last convolutional layer in the network. This provides insight into which regions of the image contributed most to the model's decision. Grad-CAM does this by weighting and combining feature maps based on these gradients, creating a coarse localization map that shows the key areas responsible for the prediction. *Grad-CAM++* refines this process, providing better localization, especially when multiple regions in an image are contributing to the decision. These heatmaps give clinicians a visual explanation of the model's focus, allowing them to understand why the model made a particular diagnosis.

Implementation: Grad-CAM is utilized on the final convolutional layers of the Inception V3 model to generate a localization map that highlights areas of the image contributing to the diagnosis of glaucoma. By analyzing gradients from the target class, it identifies important regions in the fundus images. Grad-CAM++ enhances this process by providing finer localization, particularly useful in images where multiple features may indicate the presence of glaucoma. This refinement allows for more accurate highlighting of relevant areas, improving interpretability. Together, these techniques enhance the transparency and trustworthiness of the model's predictions in clinical settings.

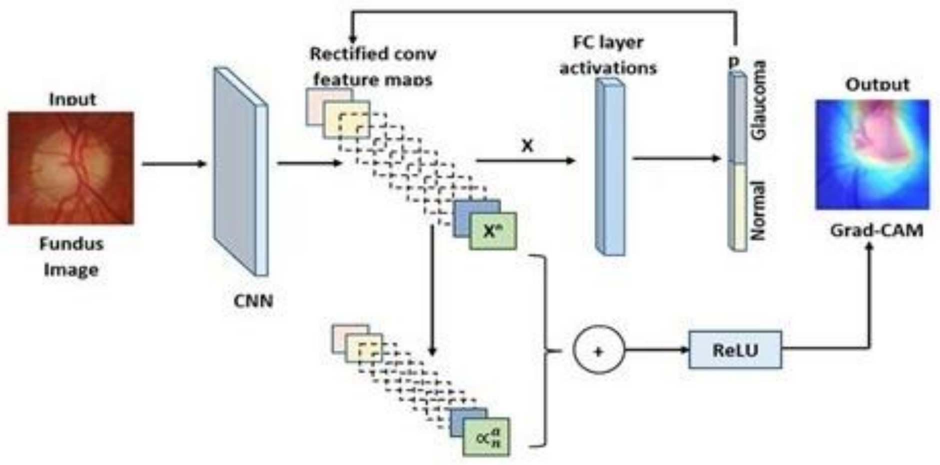


Fig.5: Grad Cam Classification Flow [2]

4 RESEARCH GAP

Limited Datasets: Current glaucoma detection models primarily rely on a few widely-used datasets, such as REFUGE[5], Drishti-GS[2], and RIM-ONE[4][3]. These datasets, while useful, are often limited in size and lack sufficient diversity in patient demographics like age, ethnicity, and different stages of disease progression. As a result, models trained on these datasets may struggle to generalize effectively to diverse real-world populations. The gap lies in the absence of large, representative datasets that can capture a broader spectrum of patient variability and disease progression. This project could address this issue by contributing to the development or incorporation of more comprehensive datasets that include diverse demographic data and multi-stage glaucoma annotations, thereby improving the generalizability of machine learning models.

Real-Time Processing and Clinical Integration [1]: Many existing glaucoma detection models achieve high accuracy, but they are often not optimized for realtime processing, making them impractical for integration into clinical workflows. These models are typically tested in controlled environments, without addressing the complexities of real-world clinical settings, such as the need for fast, realtime diagnosis and ease of use. The gap in current research is the lack of focus on practical deployment, particularly in terms of processing speed and the ability to integrate seamlessly into clinical practices. This project aims to bridge this gap by developing a lightweight, efficient model that can be integrated into clinical workflows, enabling real-time glaucoma detection with minimal computational resources.

Explainability and Model Transparency[5][12][15]: Although deep learning models, especially convolutional neural networks (CNNs), have shown success in glaucoma detection, they often function as вЂќblack boxesвЂќ that offer little insight into how decisions are made. This lack of explainability can reduce trust among medical professionals, who need to understand which retinal features the model is focusing on to make a diagnosis. The gap here is the absence of explainable AI (XAI) techniques in existing models that could make the decision-making process transparent. This project intends to address this gap by incorporating explainability features such as saliency maps or feature importance heatmaps, which can provide clear visual justifications for model predictions, fostering greater trust and acceptance among healthcare providers.

Multi-Modality Data Integration [1][2][9] Most glaucoma detection models rely solely on retinal fundus images, which limits their ability to provide a comprehensive diagnosis. Other diagnostic tools, such as Optical Coherence Tomography (OCT) scans, intraocular pressure readings, and patient medical history, are often excluded from these models, despite their potential to enhance diagnostic accuracy. The research gap lies in the failure to incorporate multi-modality data, which could improve the robustness of detection systems. This project seeks to fill this gap by exploring multi-modal machine learning techniques that combine retinal images with additional diagnostic data, offering a more holistic and accurate detection method.

Early-Stage Glaucoma Detection[8] Most current models excel at detecting advanced glaucoma but struggle to identify early-stage glaucoma due to the subtle nature of early disease indicators. The visual changes in retinal images during the early stages are often not pronounced enough for existing models to reliably distinguish between healthy and glaucomatous eyes. The gap in this area is the limited ability of current machine learning models to detect glaucoma at its onset. This project will focus on enhancing feature extraction techniquesto better capture subtle early-stage glaucoma indicators, improving early diagnosis and potentially preventing vision loss through timely intervention.

5 WORKFLOW

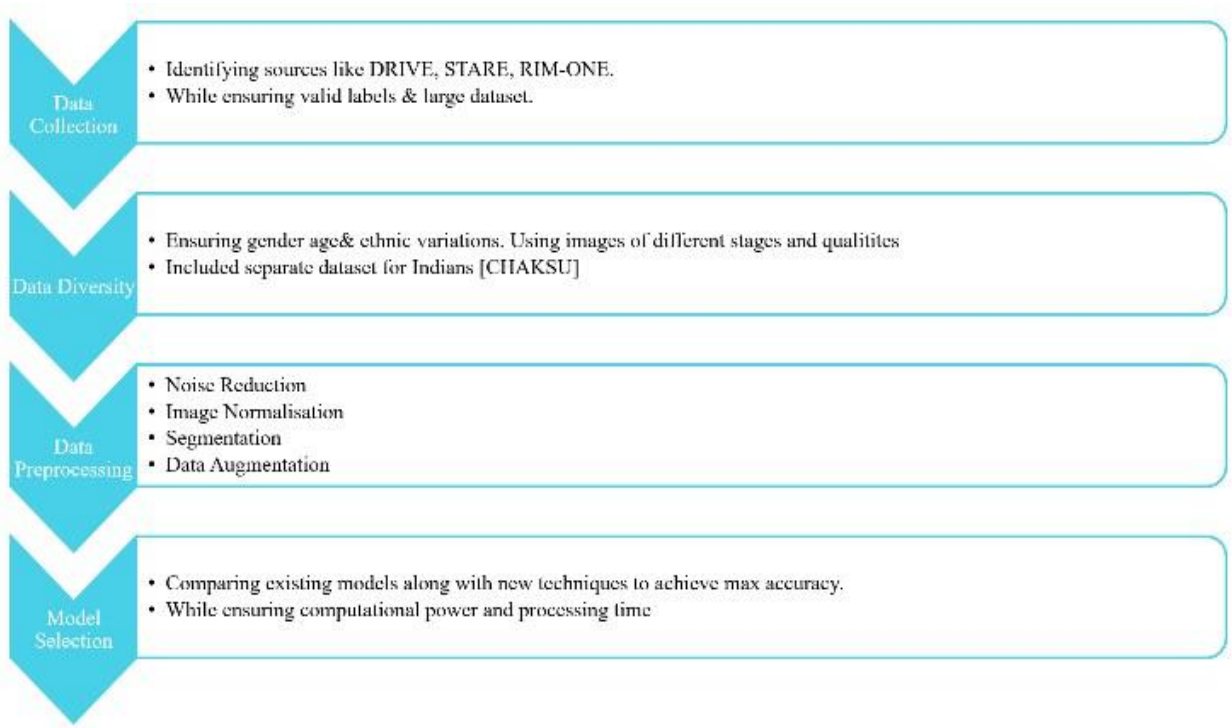


Fig.6: Worflow Diagram [1]

6 CONCLUSION

In conclusion, this paper provides a comprehensive analysis of existing machine learning techniques for glaucoma detection using retinal fundus images. By reviewing and comparing various feature extraction methods, machine learning models, and datasets, the study highlights the strengths and limitations of current approaches. The analysis identifies critical gaps, such as the need for larger, more diverse datasets, improved real-time processing capabilities, and the incorporation of explainable AI methods. This paper serves as a foundational step toward understanding the current landscape of automated glaucoma detection systems and offers insights for future research aimed at enhancing diagnostic accuracy, scalability, and clinical integration.

REFERENCE

[1] H O. J. Afolabi, G. P. Mabuza-Hocquet, F. V. Nelwamondo, and B. S. Paul. The use of u-net lite and extreme gradient boost (xgb) for glaucoma detection. *IEEE Access*, 9:47411–47424, 2021.

[2] T. K. Behera, S. Bakshi, M. Nappi, and P. K. Sa. Superpixel-based multiscale cnn approach toward multiclass object segmentation from uavcaptured

aerial images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 16:1771–1784, 2023.

[3] O. C. Devecioglu, J. Malik, T. Ince, S. Kiranyaz, E. Atalay, and M. Gabbouj. Real-time glaucoma detection from digital fundus images using self-onns. *IEEE Access*, 9:140031–140041, 2021.

[4] A. Manassakorn et al. Glaunet: Glaucoma diagnosis for octa imaging using a new cnn architecture. *IEEE Access*, 10:95613–95622, 2022.

[5] F. Abdullah et al. A review on glaucoma disease detection using computerized techniques. *IEEE Access*, 9:37311–37333, 2021.

[6] F. Guo et al. Yanbao: A mobile app using the measurement of clinical parameters for glaucoma screening. *IEEE Access*, 6:77414–77428, 2018.

[7] J. Guo, G. Azzopardi, C. Shi, N. M. Jansonius, and N. Petkov. Automatic determination of vertical cup-to-disc ratio in retinal fundus images for glaucoma screening. *IEEE Access*, 7:8527–8541, 2019.

[8] M. T. Islam, S. T. Mashfu, A. Faisal, S. C. Siam, I. T. Naheen, and R. Khan. Deep learning-based glaucoma detection with cropped optic cup and disc and blood vessel segmentation. *IEEE Access*, 10:2828–2841, 2022.

[9] S. Krishnan, J. Amudha, and S. Tejwani. Gaze exploration index (ge i)explainable detection model for glaucoma. *IEEE Access*, 10:74334–74350, 2022.

[10] R. Leonardo, J. GonBёcalves, A. Carreiro, B. SimЛњoes, T. Oliveira, and F. Soares. Impact of generative modeling for fundus image augmentation with improved and degraded quality in the classification of glaucoma. *IEEE Access*, 10:111636–111649, 2022.

[11] M. C. V. Stella Mary, E. B. Rajsingh, and G. R. Naik. Retinal fundus image analysis for diagnosis of glaucoma: A comprehensive survey. *IEEE Access*, 4:4327–4354, 2016.

[12] T. Shyamalee, D. Meedeniya, G. Lim, and M. Karunarathne. Automated tool support for glaucoma identification with explainability using fundus images. *IEEE Access*, 12:17290–17307, 2024.

[13] W. T. Song, I.-C. Lai, and Y.-Z. Su. A statistical robust glaucoma detection framework combining retinex, cnn, and doe using fundus images. *IEEE Access*, 9:103772–103783, 2021.

[14] S. VirbukaitЛ™e, J. BernataviЛ‡cienЛ™e, and D. ImbrasienЛ™e. Glaucoma identification using convolutional neural networks ensemble for optic disc and cup segmentation. *IEEE Access*, 12:82720–82729, 2024.