Journal of Information Systems Engineering and Management
2025, 10(19s)
e-ISSN: 2468-4376
https://www.jisem-journal.com/

Research Article

Glaucoma Detection and Classification using Innovative Approaches based on Deep Learning model

Anshul Jain[1], Dr. Vikas Sakalle[2]
[1]Research Scholar , [2]Associate Professor
[1,2]Department of Computer Science and Engineering
[1,2]LNCT University Bhopal M.P, India
aanshul0912@gmail.com, vikassakalle@gmail.com

ARTICLE INFO	ABSTRACT

Received: 20 Dec 2024

Revised: 25 Jan 2025

Accepted: 13 Feb 2025

This study rigorously evaluates the efficacy of several deep learning models across multiple datasets, including RIM-ONE v3, ORIGA-Lite, and APTOS 2019, each known for their pivotal role in ophthalmic research. The models under examination—AlexNet, VGGNet 19, VGGNet 22, ResNet 101, and ResNet 152—were assessed on their performance across a comprehensive set of metrics: Accuracy, Sensitivity, Specificity, Precision, F1-score, and AUC. Our findings reveal a clear hierarchy in model performance, with ResNet 152 consistently outperforming its counterparts across all datasets, achieving remarkable scores such as a 98% accuracy, 97% sensitivity, and 98% specificity on the RIM-ONE v3 dataset, and similarly high metrics on the ORIGA-Lite and APTOS 2019 datasets. These results not only underscore the superiority of ResNet 152 in mastering the complex task of glaucoma detection but also highlight the model's adeptness at generalizing across different imaging conditions and patient demographics. The study's outcome suggests a promising direction for integrating advanced convolutional neural network architectures into clinical settings, offering a potential breakthrough in early glaucoma diagnosis and contributing significantly to the prevention of irreversible vision loss through timely and accurate classification.

Keywords: convolutional neural network , AlexNet, VGGNet 19, a hypothetical VGGNet 22, ResNet 101, and ResNet 152, RIM-ONE v3.

1. INTRODUCTION

In the realm of ophthalmology, glaucoma stands [1] as a leading cause of irreversible blindness, affecting millions worldwide. Early detection and accurate classification are pivotal in halting the progression of this disease. With the advent of deep learning, particularly convolutional neural networks (CNNs), [2] there has been a paradigm shift in the methods employed for medical image analysis, including glaucoma detection. In this comprehensive investigation, we delved into the efficacy of various CNN architectures—ranging from the simpler AlexNet to the more complex VGGNet 19, an extrapolated VGGNet 22 model, and the intricate ResNet frameworks (ResNet 101 and ResNet 152)—to assess their capabilities in recognizing and classifying glaucomatous features within retinal fundus images from the RIM-ONE v3 dataset.

The study meticulously evaluated the performance of each model across a spectrum of metrics: accuracy, sensitivity (true positive rate), specificity (true negative rate), precision (positive predictive value), F1-score (harmonic mean of precision and sensitivity), and the Area Under the Receiver Operating Characteristic Curve (AUC). These metrics collectively provide a nuanced understanding of each model's diagnostic strengths. Notably, the results demonstrated a clear trend: as the complexity of the CNN architecture increased, so did the performance across all metrics. The ResNet 152 model emerged as the most proficient, showcasing a remarkable accuracy of 98%, along with the highest recorded sensitivity and specificity of 97% each, precision of 97%, F1-score of 98.2%, and an AUC of 97%.

These findings underscore the potential of advanced CNNs, such as ResNet 152, in enhancing the accuracy and reliability of glaucoma screening tools. By effectively identifying subtle, yet critical, patterns associated with the disease, these sophisticated models could provide invaluable support to ophthalmologists, enabling earlier interventions and more personalized patient care. This research contributes to the growing body of evidence that deep learning models are indispensable assets in the fight against glaucoma, marking a significant leap forward in ophthalmic diagnostics.

2. LITERATURE REVIEW

In their study, Sanghavi et al. (2024) emphasise the importance of precisely segmenting the Optic Disc as a critical stage in the classification of Glaucoma utilising Fundus images. Glaucoma detection extensively employs machine learning and artificial intelligence techniques. The primary indicators examined in Fundus images include the presence of Papillary Atrophy, Cup to Disc Ratio values, diminishing Neural Retinal Rim (NRR), adherence to the Inferior Superior Nasal Temporal (ISNT) rule, and Cup Diameter. This study explores several segmentation and classification methods for the purpose of segmenting and classifying the optic disc in both normal and glaucomatous eyes. The suggested methodology will be advantageous to physicians and healthcare personnel operating in institutions with constrained resources. This article use histogram processing to ascertain the picture's category, and subsequently, we make a determination on whether the image necessitates segmentation. Certain datasets within the standard dataset consist of intact retinal pictures, whereas others consist of optic discs that have been segmented. The segmented pictures are immediately used as input for classification using the suggested Convolutional Neural Network (CNN). The Simple Linear Iterative clustering (SLIC) and normalised graph cut technique are used to conduct segmentation on the retinal pictures, ensuring their completeness. The performance of the proposed framework is evaluated by comparing it with pretrained neural networks such as VGG19, InceptionV3, and ResNet50V2, utilising prominent metrics. We conducted training and testing on these designs using a total of 3115 photos from six established datasets. The accuracy of our suggested framework surpasses that of all others, achieving a rate of 96.33%.[1]

Virbukaitė et al. (2024) observed that pathological alterations in the eye fundus image, particularly in the vicinity of the Optic Disc (OD) and Optic Cup (OC), might serve as indicators of eye conditions such as glaucoma. Hence, precise OD and OC segmentation is crucial. The diversity of pictures resulting from the usage of several eye fundus cameras is a challenge for the current deep learning (DL) networks in accurately segmenting the optic disc (OD) and optic cup (OC). Typically, research studies included conducting tests just on individual data sets, yielding conclusions relevant to that particular samples. Our ultimate objective is to create a deep learning technique that accurately separates the optic disc (OD) and optic cup (OC) in all types of eye fundus images. However, the implementation of the mixed training data approach is now in its first phase, and the process of picture preparation has not been addressed. Hence, the objective of this study is to assess the influence of image preprocessing on the segmentation of optic disc (OD) and optic cup (OC) in various eye fundus pictures that have been aligned based on their size. We used a hybrid training data approach by merging pictures from DRISHTI-GS, REFUGE, and RIM-ONE datasets. Additionally, we employed image scaling techniques, such as bilinear, closest neighbour, and bicubic interpolation, to ensure consistent image resolution. The evaluation of picture preprocessing's influence on OD and OC segmentation was conducted using three convolutional neural networks: Attention U-Net, Residual Attention U-Net (RAUNET), and U-Net++. The experimental findings demonstrate that the optimal segmentation accuracy is attained by scaling pictures to dimensions of 512 x 512 pixels and using bicubic interpolation. The OD Dice coefficient reaches its maximum value of 0.979 on the RISHTI-GS test dataset, while the OC Dice coefficient achieves its greatest value of 0.877 on the same dataset. On the REFUGE test dataset, the OD Dice coefficient reaches 0.973 and the OC Dice coefficient reaches 0.874. Lastly, on the RIM-ONE test dataset, the OD Dice coefficient reaches 0.977 and the OC Dice coefficient reaches 0.855. The findings of Anova and Levene's tests, which were conducted at a significance level of $\alpha = 0.05$, provide statistically significant evidence that the selected size in picture resizing has an influence on the OD and OC segmentation outcomes. Additionally, the interpolation technique solely affects the OC segmentation. [2]

Meenakshi Devi et al. (2024) defines glaucoma as a multifactorial collection of eye illnesses that produce gradual harm to the optic neuropathies, resulting in permanent blindness. Glaucoma is a prevalent worldwide issue, impacting millions of people. This study focuses on the crucial objective of identifying glaucoma by segmenting the optic disc and optic cup utilising ensemble learning methods, namely Unet and Gnet. The work utilises the capabilities of these algorithms to improve the precision of the segmentation process, which is a vital stage in the early diagnosis of glaucoma. The study employs a carefully selected collection of ophthalmic photographs, with particular attention given to preprocessing methods like as resizing, normalisation, filtering, and contrast enhancement. These approaches are used to improve the quality of the input data. The Unet and Gnet architectures are offered as suitable models for segmenting the optic disc and cup. The experimental setting entails intensive training, with a focus on optimising the models for segmentation tasks. The accuracy of the segmentation findings is assessed using evaluation measures such as the Dice coefficient and sensitivity. The results illustrate the effectiveness of the ensemble Unet and Gnet. The recommended model consistently achieves accuracy levels above 98.90% on different datasets, demonstrating outstanding performance in properly categorising severe situations. The work provides insights into how enhanced segmentation of the optic disc and cup may have a significant clinical effect on the early identification of glaucoma. It highlights the need of using ensemble learning techniques to further the interpretation of ocular images for medical purposes.[3]

According to Coan et al. (2023), glaucoma is a prominent contributor to permanent vision loss on a global scale, and the number of cases is steadily increasing worldwide. Timely identification is essential as it enables prompt action to avoid further deterioration of the visual field. In order to identify glaucoma, one may do an examination of the optic nerve head using fundus imaging, focusing on evaluating the borders of the optic cup and disc. Fundus imaging is a noninvasive and inexpensive technique. However, the analysis of the images depends on subjective, time-consuming, and expensive evaluations by experts. Is it possible for artificial intelligence to replicate glaucoma evaluations conducted by professionals?Can artificial intelligence autonomously detect the limits of the optic cup and disc, creating a segmented fundus picture, and then use this segmented image to accurately diagnosis glaucoma? We performed a thorough examination of glaucoma diagnosis frameworks that use artificial intelligence and rely on segmented fundus pictures. We then provided a concise summary of the benefits and drawbacks associated with these frameworks. A total of 36 relevant publications were discovered during the period of 2011 to 2021, including two primary methodologies: 1) logical rule-based frameworks, which rely on a predefined set of rules; and 2) machine learning/statistical modeling-based frameworks. We conducted a thorough assessment of the current status of the two techniques, found deficiencies in the existing research, and highlighted potential areas for further investigation.[4]

Guru Prasad et al. (2023) define segmentation as the act of partitioning a picture into many distinct components. Each individual component is referred to as a segment. The primary goal of image segmentation is to transform the picture representation into a different format that facilitates the investigation of image attributes and qualities. Segmentation of retinal images is a crucial step in the study and detection of retinal diseases. Retinal picture segmentation aids ophthalmologists in identifying glaucoma, a disorder affecting the eye. Glaucoma is a significant contributor to irreversible visual impairment. Prompt identification of glaucoma is crucial in order to halt the advancement of visual impairment. The crucial clinical measure used for the identification of glaucoma illness is the vertical cup-to-disc ratio. Hence, the precise delineation of the optic disc from retinal pictures has immense importance. This study introduces three classifications of segmentation methods for extracting the optic disc area from retinal fundus pictures. Optic disc segmentation employs thresholding-based approaches, clustering-based techniques, and region-based techniques. The efficacy of the suggested approaches was assessed using the DRIONS-DB dataset, which consists of 110 photos, and the HRF dataset, which consists of 45 images. The calculation of the performance metric border localization error involves comparing each suggested approach with the ground truth values. The results obtained from the suggested approaches demonstrate that they exhibit lower complexity and effectively operate on all pictures. [5]

According to Elmoufidi et al. (2023), glaucoma is a persistent ailment that poses a risk to ocular well-being and has the potential to result in irreversible vision loss. Given the absence of a treatment for glaucoma, it is essential to prioritise early screening and identification in order to avoid the onset of this condition. Hence, this work presents a new approach for automatically screening glaucoma by integrating clinical measurement characteristics with image-based features. An enhanced UNet++ neural network is suggested to concurrently segment the optic disc and optic cup depending on the region of interest (ROI), in order to precisely extract clinical measurement data. The segmentation findings are used to extract significant clinical measurement characteristics, such as the optic cup to disc ratio. Subsequently, the suggested model of increasing field of view (IFOV) is introduced to comprehensively extract textural characteristics, statistical features, and other latent image-based data. Subsequently, we choose the most optimal combination of features from the whole set and use the adaptive synthetic sampling technique to mitigate the imbalanced distribution of training data. Ultimately, a classifier for glaucoma screening is developed using a gradient boosting decision tree (GBDT) algorithm. The experimental findings, using the ORIGA dataset, demonstrate that the proposed algorithm achieves outstanding performance in screening for glaucoma. It obtains a sensitivity of 0.894, accuracy of 0.843, and an AUC of 0.901, surpassing other current approaches. [6]

Puchaicela-Lozano et al. (2023) state that glaucoma is a prominent contributor to permanent vision loss on a global scale, impacting a substantial number of individuals. Prompt identification is crucial in minimising visual impairment, and a range of methods are used for the detection of glaucoma. This study presents a novel approach for identifying the location of glaucoma fundus images. It involves using pre-trained Region-based Convolutional Neural Networks (R-CNN) ResNet-50 and performing cup-to-disk area segmentation. The suggested technique was assessed using the ACRIMA and ORIGA datasets. The findings demonstrated that the ResNet-50 model achieved an average confidence level of 0.879, suggesting its reliability as a viable option for detecting glaucoma. In addition, the cup-to-disc ratio was determined using Gradient-color-based optic disc segmentation, which aligned with the ResNet-50 outcomes in 80% of instances, with an average confidence score of 0.84. The proposed methodology in this research can ascertain the presence or absence of glaucoma with a conclusive accuracy of 95%, using particular criteria to assist the expert in achieving a precise diagnosis. To summarise, the suggested model offers a dependable and protected approach for detecting glaucoma by using fundus pictures. [7]

According to Chandrasekaran et al. (2023), glaucoma is the most prevalent retinal disorder characterised by damage to the retina caused by increased pressure within the eye, known as intraocular pressure (IOP). Glaucoma causes damage to the optic nerve head (ONH), resulting in vision impairment if left untreated. An adept ophthalmologist examines the progression of glaucoma on the retinal region of the eye. This approach is characterised by a significant expenditure of time and a lack of efficiency. Consequently, this presents a genuine issue that may be resolved by using deep learning algorithms for the automated diagnosis of glaucoma. This research project involves the development of an innovative model for detecting glaucoma. The model consists of five main phases: pre-processing, ROI identification, feature extraction, feature selection, and glaucoma classification (normal/diseased). First, the retinal pictures that have been acquired are subjected to pre-processing using wiener filtering to eliminate noise and CLAHE for enhancing contrast. The ROI of the pre-processed picture is determined using the Optimised K-Means clustering approach. In this technique, the centroids of K-means are picked ideally using the Dingo with Enhanced Encircling Optimisation Model (DEEO). Afterwards, the colour characteristics, such as Colour Histogram and Colour Co-occurrence Matrix (CCM), as well as the texture features, such as Local Binary Pattern-LBP and Median Local Gradient Pattern-MLGP, are extracted from the identified regions of interest (ROI). Moreover, out of the chosen characteristics, the most significant ones are picked using the Dingo with Enhanced Encircling Optimisation Model (DEEO). The DEEO is an abstract extension of the conventional Dingo Optimizer (DOX). Ultimately, the Improved CNN (I-CNN) is used to accurately classify healthy and unhealthy individuals based on OC and OD. Ultimately, a comparative assessment is conducted to verify the effectiveness of the proposed model.[8]

Khan et al. (2023) identified glaucoma as a common cause of persistent vision loss, characterised by the progressive deterioration of retinal ganglion cells (RGCs). These specialised cells have a vital function in relaying visual information from the eye to the brain. An essential diagnostic technique for glaucoma is the segmentation of the optic disc (OD) and optic cup (OC). The retinal scans of the eye include two separate and different structures known as OD and OC. We provide a sophisticated framework called ResUNet++ with attention gates for the segmentation of combined OD and OC workloads. The foundation of our system relies on a highly effective encoder-decoder architecture, which has the capability to acquire distinguishing characteristics for both object detection (OD) and object classification (OC) segmentation. We assess the framework design by using UNet, ResUNet, and ResUNet++ with attention gate on six publically accessible datasets, namely CRFO-v4, Drishti-GS1, G120, ORIGA, PAPILA, and REFUGE1. Our framework demonstrates superior performance compared to the most advanced methods in terms of both accuracy and efficiency. Our research indicates that our collaborative segmentation framework is a very promising method for achieving precise and efficient segmentation of object detection (OD) and object classification (OC). This might potentially result in the creation of novel instruments for the diagnosis and early identification of glaucoma. [9]

Aurangzeb et al. (2023) state that due to the tremendous progress in artificial intelligence, namely in machine learning and deep learning, the possibility of automated illness detection is growing. Expanding the size of databases is essential for the training and validation of models designed to assess the effectiveness of treatments for chronic illnesses like glaucoma and diabetic retinopathy, which advance gradually and without obvious symptoms. Automated techniques for segmenting retinal blood vessels and localising the optic cup and disc are favoured for conducting widespread screenings of the population. These techniques aid in the timely identification and treatment of eye disorders, hence reducing vision loss and enhancing public well-being. This paper addresses the difficulties associated with separating the retinal vessels from fundus images. It introduces a modified ColonSegNet model for retinal vessel segmentation, which incorporates effective techniques for accurately identifying the actual vessels. Additionally, the model utilises data augmentation to overcome the limitation of having a limited number of graded images. The research use intelligent evolution algorithms to determine the ideal parameters for enhancing the contrast of retinal fundus pictures. Significant obstacles in retinal vascular segmentation include the central vessel reflex, bifurcation, crossing, narrow vessels, and the presence of lesions. The suggested technique has excellent sensitivity, specificity, and accuracy, with values of 0.839, 0.979, and 0.966 for DRIVE, 0.865, 0.979, and 0.971 for CHASE_DB, and 0.867, 0.981, and 0.972 for STARE datasets, respectively, in segmenting retinal vessels. The work is essential in the advancement of automated systems for the timely identification and management of ocular disorders, thereby enhancing public health.[10]

Karrothu et al. (2023) assert that vision is an innate endowment bestowed upon people by nature. It serves as a repository for visual stimuli and transmits them to the brain. Glaucoma is a hereditary condition characterised by elevated intraocular pressure that affects the human eye. Glaucoma, a chronic ailment, is sometimes referred to as the "silent thief of sight" due to its lack of symptoms. Glaucoma is a significant issue that affects the human eye and is often attributed to excessively elevated intraocular pressure. It results in damage to the optic nerve and results in loss of vision. This condition is classified as an incurable ailment that leads to the loss of eyesight. Failure to diagnose it in its early stages might result in irreversible vision loss. If glaucoma is diagnosed at an advanced stage, there is currently no definitive treatment available. Prompt identification of this ailment is crucial for administering accurate therapy and reducing visual impairment. An ophthalmologist often use the examination of an eye as the standard diagnostic approach for glaucoma. The process of manually detecting glaucoma is expensive and requires a significant amount of time. This paper introduces a very valuable computational technique for the identification of glaucoma in the field of computer science. The suggested methodology used Computer Vision (CV) and Vision Transformers (VIT) to determine the presence or absence of glaucoma in individuals.[11]

Early detection of glaucoma is crucial for promptly commencing treatment and averting any visual impairment. An precise calculation of the cup-to-disk ratio (CDR) is necessary for the diagnosis of glaucoma. The existing methods for automatically computing CDR suffer from reduced accuracy and increased complexity, both of which are crucial factors to consider when designing diagnostic systems for essential diagnoses. The existing techniques use an advanced deep learning model that consists of a substantial number of parameters, leading to increased system complexity and longer training/testing duration. In order to tackle these difficulties, this research suggests a Deep Neural Network (DNN) called Residual Connection (non-identity)-based Deep Neural Network (RC-DNN). The RC-DNN use non-identity residual connection to detect both the optic disc (OD) and optic cup (OC) simultaneously. The suggested model is strengthened by its efficient residual connectivity, which offers several advantages. Initially, the model demonstrates high efficiency and is capable of concurrently segmenting the optic cup (OC) and optic disc (OD). Furthermore, the effective transmission of residual information addresses the issue of the vanishing gradient problem, leading to accelerated convergence of the model. Furthermore, feature inspiration enables the network to carry out segmentation using a minimal number of network layers. We conducted a thorough assessment of the model's performance by evaluating its training on the RIM-ONE and DRISHTIGS databases. Our proposed model achieves the following performance metrics for OC segmentation on the test set images from the DRISHTI-GS and RIM-ONE datasets: a dice coefficient of 92.62, a Jaccard coefficient of 86.52, a sensitivity of 94.21, a specificity of 99.83, and a balanced accuracy of 94.2. The experimental findings demonstrate that the created model greatly improves the performance of combined OC and OD segmentation. Furthermore, the decreased computational complexity resulting from a reduction in model parameters, along with improved segmentation accuracy, offers the added benefits of effectiveness, resilience, and dependability of the produced model. The characteristics of the created model support its use in large-scale glaucoma screening programmes.[12]

Fu et al. (2019) define glaucoma as a persistent ocular condition that results in permanent deterioration of visual function. This chapter presents two cutting-edge glaucoma detection approaches that use advanced deep learning techniques. The first network is the M-Net, a multi-label segmentation network that simultaneously addresses the segmentation of the optic disc and optic cup. The M-Net architecture has a U-shaped convolutional network with several scales and a side-output layer. This design enables the model to learn distinct representations and provide a segmentation probability map. The vertical cup to disc ratio (CDR) is determined by measuring the segmented optic disc and cup, which helps evaluate the risk of glaucoma. The second network is DENet, a disc-aware ensemble network that combines the deep hierarchical context of the global fundus picture with the local optic disc area. The system consists of four distinct streams, each operating at various levels and modules. These streams are the global image stream, segmentation-guided network, local disc area stream, and disc polar transformation stream. The DENet immediately generates the glaucoma detection outcome from the picture without the need for segmentation. Ultimately, we assess the performance of two deep learning techniques in comparison to other relevant approaches across many glaucoma detection datasets. [13]

Hussain et al. (2023) define glaucoma as a persistent ocular condition that has the potential to result in permanent and irreversible loss of vision. The assessment of the cup-to-disc ratio (CDR) is crucial in identifying glaucoma in its initial phase, hence preventing visual discrepancies. Hence, achieving precise and automated segmentation of the optic disc (OD) and optic cup (OC) from retinal fundus pictures is an essential need. Current CNN-based segmentation schemes rely on constructing deep encoders with extensive downsampling layers, which face a fundamental constraint in effectively representing explicit long-range dependencies. In this study, we provide a novel segmentation pipeline called UTNet. It combines the benefits of UNet and transformer in its encoding layer, and incorporates an attention-gated bilinear fusion approach. Furthermore, we integrate Multi-Head Contextual attention to augment the conventional self-attention used in conventional vision transformers. Consequently, the shallow representation only captures basic characteristics and overall relationships. In addition, we gather contextual information at various encoding levels to enhance the analysis of receptive fields and facilitate the model's acquisition of profound hierarchical representations. Ultimately, a refined blending deficit is introduced to closely monitor the whole of the learning process. The suggested model has been used to perform simultaneous segmentation of optic disc (OD) and optic cup (OC) on three publically accessible datasets: DRISHTI-GS, RIM-ONE R3, and REFUGE. In order to test our concept, we conducted comprehensive experimentation on Glaucoma diagnosis using all three datasets, evaluating the Cup to Disc Ratio (CDR) value. Experimental findings confirm that UT-Net outperforms the current cutting-edge approaches. [14]

Fan et al. (2023) aim to assess the diagnostic accuracy and interpretability of two deep learning models, namely Data-efficient image Transformer (DeiT) and ResNet-50. These models are trained on fundus photographs from the Ocular Hypertension Treatment Study (OHTS) to detect primary open-angle glaucoma (POAG). The study also aims to identify the crucial regions in the photographs that contribute significantly to the decision-making process of each model. Design: Assessment of a diagnostic technique. The study included a total of 66,715 images obtained from 1,636 individuals in the OHTS study, as well as 16,137 photographs from external databases of healthy and glaucoma eyes. The study utilised data-efficient image Transformer models to train and detect 5 ground-truth classifications of Ocular Hypertension Treatment Study (OHTS) Primary Open-Angle Glaucoma (POAG). These classifications include OHTS end point committee POAG determinations based on disc changes (model 1), visual field (VF) changes (model 2), or either disc or VF changes (model 3), as well as Reading Centre determinations based on disc (model 4) and VFs (model 5). The top-performing DeiT models were evaluated against ResNet-50 models on the OHTS dataset as well as five other external datasets. Main Outcome Measures: The diagnostic performance was evaluated by comparing the areas under the receiver operating characteristic curve (AUROC) and the sensitivities at specificities that were fixed. The explainability of the DeiT and ResNet-50 models was assessed by comparing the attention maps generated from DeiT with three gradient-weighted class activation map techniques. Results: When compared to our top-performing ResNet-50 models, the DeiT models showed comparable performance on the OHTS test sets for all 5 ground-truth POAG labels. The AUROC values ranged from 0.82 (model 5) to 0.91 (model 1). The AUROC of the data-efficient image Transformer consistently outperformed that of ResNet-50 across all five external datasets. For instance, the area under the receiver operating characteristic (AUROC) for the primary Ocular Hypertension Treatment Study (OHTS) endpoint (model 3) was found to be 0.08 to 0.20 higher in the DeiT models compared to the ResNet-50 models. The saliency maps generated by the DeiT model identify certain regions inside the neuroretinal rim that are deemed significant for categorization purposes. The ResNet-50 models have a broader, more widespread dispersion over the optic disc in their corresponding maps. Conclusions: Vision Transformers has the capacity to enhance the ability of deep learning models to generalise and provide explanations. They can effectively identify eye diseases and perhaps other medical disorders that depend on imaging for clinical diagnosis and treatment. [15]

Oguz et al. (2023) define glaucoma as an ocular condition resulting from optic nerve injury, which is a prevalent cause of irreversible vision loss globally. Early diagnosis of glaucoma enables the prevention of visual loss via frequent examinations and therapy. If the condition is identified in its latter stages, it may result in irreversible damage to the optic nerve, leading to the loss of central vision and blindness. Hence, timely detection of the illness is crucial. The majority of recent research have included Deep Learning (DL) architectures that utilise an automated computerised approach based on segmenting manually crafted features in fundus pictures. The objective of this work is to assist professionals in identifying glaucoma by using a model that integrates Deep Learning and Machine Learning techniques with unprocessed fundus pictures. A novel Convolutional Neural Networks (CNN) model is used to extract profound characteristics. Deep features have been used in well-established conventional Machine Learning techniques (ML) for classification, including Adaboost, k Nearest Neighbour (kNN), Random Forest (RF), Multilayer Perceptron (MLP), Support Vector Machines (SVM), and Naive Bayes (NB). The hybrid models' performances were assessed using the ACRIMA dataset, which consists of 705 pictures. The dataset is partitioned into 80% training data and 20% testing data. The experimental findings demonstrate that the hybrid model combining CNN and Adaboost achieves the maximum level of success, with an accuracy of 92.96%, an F1 score of 93.75%, and an AUC value of 0.928.[16]

According to Desiani et al. (2023), glaucoma is a persistent neurological condition affecting the retina and resulting in the deterioration of eyesight. Glaucoma may be identified by anomalies that manifest in the optic disc and optic cup on the retina. A segmentation technique is required to get the features. Segmentation may enhance the efficacy of the categorization process. This research employs a convolutional neural network (CNN) architecture to effectively identify glaucoma by using both segmentation and classification techniques. The U-Net CNN architecture is used at the segmentation step. The U-Net architecture consists of separate encoder and decoder portions, which enable the generation of output pictures that only include the relevant information. During the classification step, the research suggests the use of the Xcep-Dense Net. Xcep-Dense is a convolutional neural network (CNN) architecture that integrates the Exception and Dense components. The Xcep-Dense architecture combines the strengths of both Xception and Dense architectures while addressing their own faults. The U-Net architecture achieved accuracy, recall, precision, and F1-score over 90% in the segmentation task, while Cohen's kappa above 85%. The findings indicate that U-Net is very effective at segmenting the optic disc and optic cup. In the classification, the accuracy and precision exceed 85%, the recall and F1-score exceed 80%, and Cohen's kappa is 77%. The findings demonstrate that the Xcep-Dense design has sufficient robustness to accurately identify glaucoma, which encompasses three distinct classes: advanced glaucoma, early glaucoma, and normal. The findings indicate that the suggested approach is viable for identifying glaucoma. The findings of this research are anticipated to be used in the future to create an automated device for the early identification of glaucoma [17].

The identification of glaucoma has become crucial, since it has emerged as a leading cause of global visual impairment, according to Ghorui et al. (2023). Currently, the majority of algorithms depend on pre-trained deep neural networks to get optimal outcomes. Nevertheless, the considerable computational time and intricacy, along with the need for a substantial database, make the task of glaucoma identification challenging and laborious. Considering these factors, this research introduces a novel convolutional neural network structure called ProspectNet. ProspectNet has been shown to achieve higher accuracy while requiring less computing time and complexity compared to two pre-trained networks, VGG16 and DenseNet121. The data collection is a combination of two publically accessible datasets, namely DRISHTI-GS and Glaucoma Dataset (Kaggle). It consists of ocular colour fundus photos of both glaucomatous and normal eyes. ProspectNet achieved a normal AUC (area under the curve) of 0.991, with a specificity and accuracy of 0.98. Confusion matrices are also generated to visually demonstrate the effectiveness of the new design. These results indicate that ProspectNet is a robust alternative to other state-of-the-art algorithms for a medium-sized dataset. The study proposes three discrete frameworks for the identification of glaucoma. An benefit of our technique is that it does not need any specific feature selection, such as thorough measurements of certain qualities like the shape of the optic nerve head [18].

According to Latif et al. (2023), glaucoma is an advancing ocular condition that might result in loss of vision if not addressed. Prompt identification is essential to avert visual impairment, although existing manual scanning techniques are costly, labor-intensive, and need specialised proficiency. This paper introduces a new technique for detecting Glaucoma called the Enhanced Grey Wolf Optimised Support Vector Machine (EGWO-SVM) approach. The suggested technique includes preprocessing procedures, such as eliminating picture noise via the use of the adaptive median filter (AMF), and conducting feature extraction by using the previously processed speeded-up robust feature (SURF), histogram of oriented gradients (HOG), and Global features. The improved Grey Wolf Optimisation (GWO) method is subsequently used in conjunction with Support Vector Machines (SVM) for classification purposes.

In order to assess the suggested approach, we used the online retinal images for glaucoma analysis (ORIGA) database. The technique demonstrated notable levels of accuracy, sensitivity, and specificity, with rates of 94%, 92%, and 92% respectively. The findings indicate that the suggested technique surpasses other existing algorithms in accurately identifying the presence or absence of Glaucoma. This work presents an innovative and efficient method for detecting Glaucoma, which has the potential to enhance the detection process and improve the results. [19]

Panda et al. (2021) defines glaucoma as a persistent eye disorder that results in permanent vision loss and now ranks as the second most common cause of blindness globally. Evaluation of the optic disc and optic cup in colour fundus images has shown potential as a reliable technique for screening glaucoma. This work presents a novel and effective technique for segmenting the optic disc and cup utilising a residual deep learning approach, which achieves high accuracy and efficiency. The GlaucoNet, a deep network based on patch analysis, is trained using a substantial number of preprocessed picture patches. These patches are sufficiently large to include the crucial discriminating information around each pixel. The suggested design improves segmentation outcomes by including additional skip connections in the framework. These connections explicitly redefine the layers in a way that the learning function depends on a residual function of the input layer. It tackles the issue of deterioration in deep networks by enhancing the flow of information. The final segmentation output is obtained by applying the convex hull transformation to the original findings. The suggested methodology is evaluated using the publicly accessible DRISHTI-GS, RIM-ONE, and ORIGA-light datasets. The optic disc segmentation achieved overlapping scores (OS) of 0.9106, 0.8972, and 0.8835 in DRISHTI-GS, RIM-ONE, and ORIGA-light, respectively. Similarly, the optic cup segmentation achieved an OS of 0.8229, 0.7401, and 0.8106 in the same datasets. The glaucoma risk index is derived by calculating the cup-to-disc height ratio obtained from the segmented areas. The effectiveness of the proposed GlaucoNet-based segmentation approach is shown by experimental findings that provide improved or equivalent segmentation performance and reduced errors in cup-to-disc height ratio estimate. [20]

The objective of the study conducted by Soares et al. (2023) was to evaluate and compare the extent of macular damage in patients with glaucomatous optic neuropathy (GON) and compressive optic neuropathy (CON). Additionally, the study aimed to determine the diagnostic accuracy of macular damage in differentiating between these two disorders. research Design: This research used observational, cross-sectional methodology and was conducted at a single centre. those diagnosed with GON (Group A), CON (Group B), and those without any medical conditions (healthy controls, Group C) were selected based on the specified qualifying criteria. A computerised approach based on spectral-domain optical coherence tomography (SD-OCT) was used to separate and identify the circumpapillary retinal nerve fibre layer (cpRNFL) and macula. The thickness of each layer was measured in every sector using the Early Treatment Diabetic Retinopathy Study and the 6-sector Garway-Heath-based grids. Comparisons were made between the data from all research groups, and a significance threshold of 0.05 was established. Findings: The study comprised a total of 75 people, each with one eye, resulting in a total of 75 eyes. Among them, 25 eyes had glaucomatous optic neuropathy (GON), 25 eyes had contralateral optic neuropathy (CON), and 25 eyes were from healthy individuals serving as controls (CG). The macular thickness in the ganglion cell complex of patients with glaucomatous optic neuropathy (GON) and control (CON) patients was found to be reduced compared to the control group (CG) with a statistical significance of $p<0.05$. The most effective grid parameters derived from Garway-Heath for discriminating between glaucomatous optic neuropathy (GON) and control (CON) were the nasal-inferior (NI) and nasal-superior sectors, as well as the damage ratios in the macular ganglion cell (mGCL) and inner plexiform (IPL) layers in the nasal-inferior/temporal inferior (NI/TI) region. Furthermore, the integration of the NI sector and NI/TI damage ratios in both layers exhibited superior discriminatory ability (AUC 0.909; 95% CI 0.830–0.988; $p<0.001$) compared to integrating parameters in each layer individually. [21]

The detection of glaucoma is a significant study domain in intelligent systems, and it has a crucial position in the field of medicine. Glaucoma may lead to permanent blindness if not diagnosed correctly. Medical practitioners must do several diagnostic tests to identify this perilous illness. It takes a significant amount of time and financial resources. Individuals with glaucoma may not experience any visual impairment during the first phase of the condition. We have developed a model to expedite the detection of glaucoma while reducing both the time and expense involved. The focus of our study is to provide a Convolutional Neural Network (CNN) model called Inception V3. A total of 6072 photos were used. Out of the total number of images, 2336 were diagnosed as glaucomatous and 3736 were determined to be normal fundus images. We used 5460 photos for training our model and 612 images for testing. Subsequently, we achieved an accuracy of 0.8529 and an AUC value of 0.9387. We used the DenseNet121 and ResNet50 algorithms for comparative purposes, resulting in accuracies of 0.8153 and 0.7761, respectively. [22]

Sathya et al. (2023) want to provide a new approach for improving the detection of glaucoma by using Digital Fundus Images (DFI). Utilising a deep learning-based method, in conjunction with feature detection approaches, allows for the unsupervised discovery of the inherent characteristics of the DFI. This enables the robust identification of the DFI with a high level of accuracy. The DL-EAEN technique is used to assess the latent representations from DFI and detect the morphological alterations linked to glaucoma for timely detection and categorization. The fundus pictures are used to locate the optic disc and diagnose glaucoma. The Scale-Invariant Feature Transform (SIFT) method is used to identify local characteristics and important areas in the images. This research utilises the PAPILA retinal dataset, which consists of 244 individuals. The dataset comprises a total of 488 fundus photographs of both the left and right eye of all patients in the male and female category. The dataset also provides clinical findings indicating the health status of the eyes, including categories such as healthy, glaucoma, suspicious, and eyes with crystalline and intraocular lens (IOL). The image segmentation approach uses U-Net and Mask-R-CNN algorithms to create pixel-level masks and accurately delineate the outside boundaries of the optic cup. The performance of DL-EAEN is evaluated using MATLAB software and compared to established models like SVM, Adaboost, and CNN-Softmax classifiers. The results show that the proposed deep learning based enhanced AEN method performs better than the existing SVM, Adaboost, and CNN-Softmax classifiers. The method achieves promising results with an accuracy of 95.6%, a dice-score of 0.8, a sensitivity of 96.2%, a specificity of 97.01%, an F-score of 97.08%, a precision of 97.41%, a recall of 98.02%, and an AUC-ROC with a true positive rate of 0.89 and a false positive rate of 0.16. The DL-EAEN algorithm demonstrates precise and consistent performance in detecting and classifying glaucoma, providing ophthalmologists with a convenient tool for diagnosis. The DL-EAEN model outperforms SVM, Adaboost, and CNN-Softmax classifiers [23] in terms of accuracy, dice score, AUC-ROC, sensitivity, specificity, precision, and F-score.

Liu et al. (2022) Background and objectives: Glaucoma, a condition that leads to permanent vision loss and potential blindness, may be mitigated by early detection. Examining the optic disc and optic cup is crucial for the detection of glaucoma, which has led to the development of several computer-aided diagnostic techniques using deep learning networks. Nevertheless, the trained model's performance on novel datasets is significantly impeded by the disparity in distribution across various datasets. Hence, our objective is to devise an unsupervised learning approach to address this issue and enhance the predictive capability of the model on novel datasets. Methods: This study introduces a new unsupervised model that utilises adversarial learning to conduct optic disc and cup segmentation, as well as glaucoma screening, in a more versatile and effective way. We use a very effective segmentation and classification network and implement unsupervised domain adaptation techniques to address the issue of domain shift in the output space of the segmentation network. We perform glaucoma screening by integrating classification and segmentation networks to get more reliable and efficient predictions for glaucoma screening. Results: We have validated the efficacy and efficiency of our proposed approach using three publicly available datasets, namely REFUGE, DRISHTI-GS, and RIM-ONE-r3 datasets. The experimental findings indicate that the suggested approach successfully mitigates the decline in segmentation performance resulting from domain shift and enhances the precision of glaucoma screening. In addition, the suggested technique surpasses the most advanced unsupervised optic disc and cup segmentation domain adaption methods. Summary: The suggested approach may aid medical professionals in the identification and assessment of glaucoma and is well-suited for practical use. [24]

Ahmed et al. (2023) define glaucoma as a multifaceted condition affecting the optic nerve, which leads to progressive vision loss as a result of elevated pressure inside the eyes. There are two forms of glaucoma: open-angle glaucoma and angle-closure glaucoma. Open-angle glaucoma is mostly caused by elevated intraocular pressure and may result in temporary or permanent damage to the eyes. Consequently, it is essential to have an early stage diagnosis in order to preserve our eyesight. Various methods, like as tonometry, perimetry, and gonioscopy, may be used to identify glaucomatous eyes. However, these methods need both time and experience. Employing deep learning methodologies may provide a superior answer. The primary objective of this work was to use deep learning methods to accurately identify eyes afflicted by open-angle conditions based on fundus pictures. The research used VGG16, VGG19, and ResNet50 deep neural network architectures to categorise eyes as either glaucoma positive or negative. The experiment was conducted using a publicly available dataset obtained from Kaggle. However, the performance of each model improved significantly after enriching the dataset, resulting in accuracy ranging from 93% to 97.56%. VGG19 outperformed the other two models with an accuracy of 97.56%. [25]

Kavitha and her colleagues (2024) have highlighted the significant use of the intuitive fuzzy set in the fields of decision-making and machine learning. This research aimed to enhance and use the intuitive fuzzy set by creating and implementing a deep neural network-based glaucoma eye detection system. The system utilised fuzzy difference equations in the area where the retinal pictures come together. Retinal image detections are classified into three categories: normal eye recognition, suspected glaucomatous eye identification, and glaucomatous eye recognition. The amount of concentration between the fuzzy partition and the retinal pictures is determined by calculating fuzzy degrees linked with weighted values. The suggested approach used Convolutional Neural Network (CNN) and deep learning techniques to detect glaucoma by analysing retinal pictures. It specifically focused on identifying the fuzzy weighted regularisation between images. This technology was used to enhance the clarity of the input photos and render them suitable for the glaucoma detection procedure. The aim of this work was to provide a new method for the early detection of glaucoma utilising the Fuzzy Expert System (FES) and Fuzzy differential equation (FDE). The intensities of the various locations in the photographs and their corresponding peak levels were determined. After identifying the areas of highest intensity, the relationships of recurrence among those peaks were further assessed. The picture was partitioned based on the varied levels of similarity and dissimilarity in concentrations. A threshold picture was produced by combining the fuzzy matrix with FDE, using concentration levels and spatial frequency that were both comparable and distinct. This method differentiates between a typical and atypical eye state, hence identifying people with glaucomatous eyes.[26]

According to Gende et al. (2023), glaucoma is the primary worldwide factor contributing to permanent loss of vision. Glaucoma patients have a gradual decline in the retinal nerve tissues, which starts with a reduction in peripheral vision. Prompt detection is crucial to prevent vision loss. Ophthalmologists evaluate the degeneration produced by this condition by examining the retinal layers in various areas of the eye, using diverse optical coherence tomography (OCT) scanning patterns to capture pictures, and producing several perspectives from different sections of the retina. These pictures are used for quantifying the thickness of the retinal layers in various areas. We provide two methodologies for the segmentation of retinal layers in OCT pictures of glaucoma patients, specifically targeting several regions. These methods may identify the pertinent anatomical features for glaucoma evaluation from three distinct OCT scanning patterns: circumpapillary circle scans, macular cube scans, and optic disc (OD) radial scans. These systems use transfer learning to leverage visual patterns in a related domain. By using state-of-the-art segmentation modules, they enable a reliable and completely automated segmentation of the retinal layers. The first method leverages inter-view commonalities by using a single module to segment all scan patterns, treating them as a unified domain. The second method employs view-specific modules to segment each scan pattern, automatically identifying the appropriate module to examine each picture. The suggested methodologies yielded good outcomes, with the first technique attaining a dice coefficient of 0.85±0.06 and the second approach obtaining 0.87±0.08 for all segmented layers. The first method yielded the most optimal outcomes for the radial scans. Simultaneously, the second strategy that is particular to the perspective produced the most favourable outcomes for the circle and cube scan patterns, which were more accurately represented. [27]

In their study, Salam et al. (2016) describe glaucoma as a chronic condition that is often referred to as the "silent thief of sight" due to its lack of symptoms. If not identified in its early stages, glaucoma may lead to irreversible blindness. Glaucoma advances before certain alterations occur in the retina, enabling ophthalmologists to identify glaucoma in its first phase and halt its advancement. Fundoscopy is a biological imaging method used to examine the interior anatomy of the retina. Our suggested approach offers a unique algorithm for identifying glaucoma from digital fundus images utilising a combination of features. This research presents a new approach that combines structural characteristics (such as cup to disc ratio) with non-structural information (such as texture and intensity) to enhance the precision of automated glaucoma diagnosis. The suggested technique incorporates a suspect class into the automated diagnostic process when there is a discrepancy in judgement between structural and non-structural variables. The suggested technique is evaluated by analysing a local database that consists of fundus pictures from 100 patients. The purpose of this system is to transfer glaucoma patients from rural regions to experts. The introduction of the suspect class is motivated by the goal of ensuring a high level of sensitivity in the proposed system. The suggested method has an average sensitivity of 100% and an average specificity of 87%.[28]

According to Sharma et al. (2024), glaucoma is a complex eye disorder that carries a significant risk of vision loss. At first, the majority of automated methods divide the main system and evaluate the clinical parameters in order to categorise and detect glaucoma. The suggested tailored convolutional neural network (CNN) model for automated glaucoma diagnosis, constructed utilising advanced learning methods, has the potential to aid several stakeholders in the supply chain management network. The stakeholders involved in this context may include eye hospitals, healthcare service providers, physicians, ophthalmologists, patients, insurance companies, and others. The deployed model consists of four trainable layers, namely three convolution layers and one flattening layer. The tailored CNN model acquired profound characteristics with the fewest adjustable parameters. Following that, a feature reduction approach known as principal component analysis (PCA) and linear discriminant analysis (LDA) is used to decrease the dimensions of the feature sets. Ultimately, a categorization is performed by using an extreme learning machine (ELM). The hidden node parameters of ELM are optimised using the modified particle swarm optimisation (MOD-PSO) approach. The overall performance of the suggested model has been improved by using 5-fold stratified cross-validation. The suggested methodology was implemented on two widely used datasets, G1020 and ORIGA. The experimental findings indicate that the computer-aided diagnosis (CAD) model attains an accuracy of 97.80% and 98.46% on the G1020 and ORIGA datasets, respectively. The customised CNN model demonstrates superior performance in comparison to other cutting-edge models, using a much reduced amount of features. This model has the potential to assist decision-makers in supply chain management networks.[29]

Glaucoma is a degenerative optic-neuropathy that causes blindness, according to Akter (2023). Genetics, vasculature, anatomy, and immunology contribute. Glaucoma affects 80 million people worldwide, including 300,000 Australians, 50% of whom go untreated. Glaucoma untreated may cause blindness. AI detection speeds up diagnosis and may prevent future vision deterioration. Several AI techniques for automated glaucoma diagnosis using 2D data have showed promise. Few studies have shown glaucoma detection and staging to be effective. The automatic AI system still struggles to diagnose like doctors. This is largely due to ML algorithms' interpretability and clinical data integration issues. If AI technology can eliminate the "black box" by becoming transparent and explainable, doctors and patients will welcome it. Physicians employ pathological signs to detect glaucomatous damage early. This method should too. Thus, the thesis sought to develop an AI model that can diagnose and categorise glaucoma by integrating varied clinical data and employing advanced data analysis and machine learning (ML) approaches. The research integrates structural, functional, demographic, risk factor, and OCT factors to optimise glaucoma diagnostic features. Salient traits were examined using statistical analysis and taught in machine learning algorithms to test identification. The research examined and trained three key structural aspects of optic nerve head (ONH) optical coherence tomography (OCT): cross-sectional 2D radial B-scan, 3D vascular angiography, and TSNIT B-scan. Explainable deep learning (DL) models automated glaucoma diagnosis using these characteristics. Feature visualisation showed deep learning model decision-making logic. Glaucoma patients' TSNIT OCT structural features or afflicted regions were properly detected. Explainable deep learning (DL) makes DL model decision-making methods transparent and intelligible. TSNIT OCT pictures are commonly misinterpreted because to distortions and speckle noise. This work also developed an automated deep learning algorithm to remove OCT scan flaws and disruptions, allowing precise retinal layer segmentation, tissue thickness measurement, and error-free picture interpretation. Also, doctors utilise the visual field (VF) test to identify glaucoma severity throughout treatment and care. This work uses functional attributes from VF photos to train machine learning algorithms to identify glaucoma development from early to advanced/severe. The selected important traits were utilised to build a comprehensive artificial intelligence model for detecting and categorising glaucoma stages, taking into consideration data volume and accessibility. TSNIT OCT images were used to develop a deep learning model. Merging the model's findings with structural and functional properties and training it using machine learning algorithms. The best machine learning model identified glaucoma with an AUC of 0.98, accuracy of 97.2%, sensitivity of 97.9%, and specificity of 96.4%. The model classified normal, early, moderate, and advanced-stage glaucoma with 90.7% accuracy and 84.0% F1 score. To conclude, this thesis developed and presented an experimentally validated artificial intelligence (AI) model to screen a huge number of people. This technology will save experts from manually analysing large patient data and risking misunderstandings. This thesis also offered three structural OCT properties that may be excellent glaucoma diagnostic markers. [30]

According to Ashtari-Majlan et al. (2023), glaucoma is the primary cause of permanent blindness on a global scale and presents considerable difficulties in diagnosis owing to its dependence on subjective assessment. Nevertheless, recent progress in the field of computer vision and deep learning has shown the capability of automated evaluation. This study provides an overview of current research on the use of artificial intelligence (AI) for diagnosing glaucoma. The focus is on the use of deep learning techniques to analyse fundus, optical coherence tomography, and visual field pictures. We provide an up-to-date classification system that categorises techniques based on architectural paradigms and provides links to accessible source code, hence improving the ability to replicate the methods. By conducting thorough comparisons on commonly used public datasets, we have identified deficiencies in the capacity to apply learned knowledge to new situations, accurately estimate uncertainty, and effectively combine information from several sources. In addition, our study selects and organises important datasets while also acknowledging their shortcomings, such as issues with size, discrepancies in labelling, and bias. We delineate unresolved research issues and elaborate on potential avenues for future investigations. This study is anticipated to be beneficial for AI researchers looking to implement advancements in practical applications, as well as ophthalmologists striving to enhance clinical processes and diagnosis by using the most recent AI findings. [31]

Computer vision has played a crucial role in the healthcare field over the last thirty years by offering reliable and sophisticated diagnostic systems based on soft computing. Glaucoma is a severe ocular condition that may cause permanent visual loss. The global incidence of glaucoma is seeing a significant surge. Performing a manual ocular evaluation to identify glaucoma is a laborious, prone to errors, time-consuming, and subjective process. Hence, there is a need for computer-assisted automated glaucoma diagnosis techniques to enhance the current diagnostic methods with their strong and reliable performance. The segmentation of the optic disc (OD) and optic cup (OC) plays a crucial role in the identification of glaucoma. Precise delineation of the optic disc (OD) and optic cup (OC) provide significant computational and clinical information that may greatly aid in the screening of glaucoma. Segmentation of retinal fundus pictures is tough due to the wide changes in size, shape, pixel intensity values, and background effects. In order to address these issues, we have created two innovative networks that are capable of accurately segmenting object detection (OD) and object classification (OC). The basis network of this study is an efficient shallow segmentation network (ESS-Net), whereas the final network is a feature-blending-based shallow segmentation network (FBSS-Net). ESS-Net is a shallow architecture that incorporates a semantic preservation block with a maximum depth to provide precise segmentation. On the other hand, FBSS-Net utilises both internal and external feature blending techniques to enhance the overall accuracy of segmentation. In order to validate their efficacy, we assessed both networks by using four openly accessible datasets: REFUGE, Drions-DB, Drishti-GS, and Rim-One-r3. The suggested approaches demonstrated outstanding segmentation performance, with a minimal amount of trainable parameters 3.02 million parameters.[32]

Alice et al. (2023) define glaucoma as an ocular condition that causes harm to the optic nerve, the link between the eye and the brain. Elevated intraocular pressure may lead to optic nerve dysfunction, posing a double risk for diabetes patients. If left undetected during the early stages, this condition can cause permanent vision loss. The need for an automated method for identifying eye diseases has expanded significantly in underdeveloped nations, where there is a shortage of ophthalmic professionals and laboratory facilities. The subject of artificial intelligence is offering several solutions, particularly in the healthcare sector. The proposed study aims to develop models for detecting the presence of glaucoma using machine learning techniques and image feature descriptors. This will be done by analysing a publicly available collection of retinal fundus pictures. The retinal fundus picture is classified as normal or abnormal using a two-stage process. The process begins by extracting image characteristics using suitable filters. These features are then used to train a tree-based ensemble classifier, which is responsible for classifying the input picture. Finally, the classifier is evaluated to evaluate its accuracy performance. The aforementioned process involves iterating over the three effective filters, namely edge histogram, fuzzy colour and texture histogram, and pyramid histogram of gradients. The experiment demonstrates that using the Edge histogram filter in conjunction with fuzzy colour and texture histogram, together with the Random forest classifier, achieves a peak accuracy of 80.43% and an AUC of 0.884. The outcomes achieved via the implementation of many filters are superior than those achieved through the implementation of a single filter. [33]

Lamba et al. (2023) state that glaucoma is a prominent factor in causing blindness as a result of neuro-degeneration. Detecting at early stages using traditional approaches is quite challenging. Furthermore, these procedures depend on the proficiency of ophthalmologists, which is time-consuming and susceptible to human fallibility. Glaucoma causes damage to the retina by impacting the optic nerve head through increased pressure inside the eye. This can lead to permanent vision loss. In order to detect glaucoma, it is necessary to segment the optic disc. However, this is challenging due to the disc's very small boundary and the presence of blood vessels that can obstruct the process. An automated system must be developed to address the current obstacles by completing activities such as segmenting the optic disc and classifying pictures as healthy or glaucomatous. A comprehensive investigation has been carried out in this respect using many publications. Various techniques, including random walk, contour optimisation, and histogram equalisation, have been used to detect glaucoma. The authors achieved highly accurate results of 99.5% and 99.93% by using a super pixel-based red channel and selecting the Hough peak value to segment the optic disc. These techniques are effective for diagnosing glaucoma at an early stage. Furthermore, the support vector classifier has been determined to be the most suitable method for classification, with an outstanding accuracy rate of 98.12%.

This classification method specifically emphasises the elimination of blood vessels. Thus, the outcomes obtained from these approaches aid in the identification of issues spanning from segmentation to classification and contribute to the slowing down of glaucoma. [34]

Ali et al. (2023) emphasise the need of early detection of glaucoma in order to prevent visual impairment. This research presents a very effective computer-aided detection (CAD) method for detecting glaucoma. The system utilises fundus pictures, deep transfer learning, and fuzzy aggregation operators. The proposed CAD system consists of three distinct stages: (1) Identification of the region of interest of the optic disc using a highly efficient deep learning network, (2) Categorization of images using various pre-trained deep convolutional neural networks and support vector machines, and (3) Utilisation of fuzzy aggregation operators to combine the predictions of glaucoma classifiers. We used three well known and resilient aggregators: ordered weighted averaging (OWA) operator, weighted power mean (WPM), and exponential mean (EXM). We evaluated the effectiveness of the proposed glaucoma computer-aided diagnosis (CAD) system using three publicly available datasets: DRISHTI-GS1, RIM-ONE, and REFUGE. The suggested approach of aggregating conjunctive OWA (Conj-OWA) delivers the most optimal results in classifying glaucoma. More precisely, it attains accuracy rates of 90.2%, 97.8%, and 94.3% and area under the curve (AUC) rates of 95.3%, 99.8%, and 96.2% on the DRISHTI-GS1, RIM-ONE, and REFUGE datasets, respectively.[35]

Glaucoma, a prevalent ocular condition, induces harm to the optic nerve as a consequence of elevated pressure inside the eye, potentially leading to partial or complete vision loss if not promptly identified. Therefore, it is crucial to identify glaucoma in its initial phase to optimise the treatment strategy. In recent years, there have been notable endeavours to create automated algorithms for classifying glaucoma using retinal fundus pictures. However, there is still a lack of comprehensive methods for detecting various stages of glaucoma. This is mostly attributed to the lack of extensive annotated datasets. Moreover, the work becomes more hard due to the existence of significant similarities across different classes, small changes in lesion size, and the presence of duplicated characteristics in the fundus pictures. In this study, we introduce a new cascaded attention-based network called CA-Net that effectively classify glaucoma in several stages. An advanced attention module called cascaded attention module (CAM) is added to a backbone network. The CAM includes a triplet channel attention block (TCAB) and a spatial attention block (SAB). This module is designed to capture feature dependencies across many dimensions, including channel, cross-channel, and spatial dimensions. The Class Activation Map (CAM) enhances performance by extracting informative discriminative features from important areas of the fundus picture. In addition, we have created a comprehensive dataset for glaucoma called the large multistage glaucoma (LMG) dataset, which consists of 1582 fundus pictures. We have also developed a separate binary glaucoma dataset, which has 623 fundus images. The experimental findings obtained from these datasets, in addition to a publically accessible dataset, demonstrate the superiority of our CA-Net over the most advanced approaches currently available. The visualisation findings obtained via Grad-CAM and Grad-CAM++ provide a deeper understanding of the effectiveness of our hypothesised attention mechanism. In addition, ablation experiments are performed to validate the efficacy of each individual component of CA-Net. [36]

Mishra et al. (2023) define glaucoma as a degenerative ocular disease characterised by irreversible visual impairment. Resolving the diagnosis of glaucoma is a crucial matter within the field of medicine. There have been few scientific endeavours focused on identifying glaucoma during its first phases. However, the current methods for identifying glaucoma sickness have shown to be unsuccessful in terms of performance. Furthermore, the time required for previous techniques of identifying glaucoma sickness was much greater. Detecting glaucoma in colour fundus images is a challenging task that requires both expertise and extensive experience. Deep learning algorithms have quickly become the preferred way for analysing medical images. This study utilised different Convolutional Neural Network (CNN) approaches to showcase the impact of key factors, including dataset size, architecture, and the choice between transfer learning and newly designed architectures. The aim was to determine the effect on performance, with a focus on achieving higher accuracy and reducing time requirements. More precisely, we conducted a comparison of the two architectural designs based on their capacity to acquire knowledge from past instances. Deep learning algorithms have quickly become the preferred way for analysing medical images. This research explores the fundamental concepts of deep learning that may be used in the analysis of medical images. It also presents a concise overview of the advancements made in this field. We examine the use of deep learning in diverse image processing endeavours, including classification, object recognition, segmentation, and registration.[37]

Chincholi et al. (2024) state that although the Vision Transformer architecture is often used for picture classification tasks, its application to object recognition in computer vision presents considerable difficulties. This study seeks to investigate the possibility of expanding the capabilities of the Vision Transformer for the purpose of detecting objects in medical images, with a special focus on glaucoma detection. Additionally, it involves analysing the Detection Transformer for the purpose of making a comparative assessment. The study entails evaluating the ratio of the cup to the disc and detecting indications of vertical thinning of the neuroretinal rim. A diagnostic threshold is suggested, indicating that a cup-to-disc ratio greater than 0.6 may be a possible marker for glaucoma. The testing findings showcase an impressive accuracy of 90.48% attained by the pre-trained Detection Transformer, whilst the Vision Transformer achieves a competitive accuracy of 87.87%. Comparative assessments use an unused dataset from the Standardised Fundus Glaucoma Dataset on Kaggle, offering significant insights into the automated identification of glaucoma. The assessment criteria and outcomes undergo thorough validation by medical specialists that specialise in the subject of glaucoma. The user's text is simply the number 38 enclosed in square brackets.[38]

3. PROPOSED METHODOLOGY

3.1 Proposed working flow

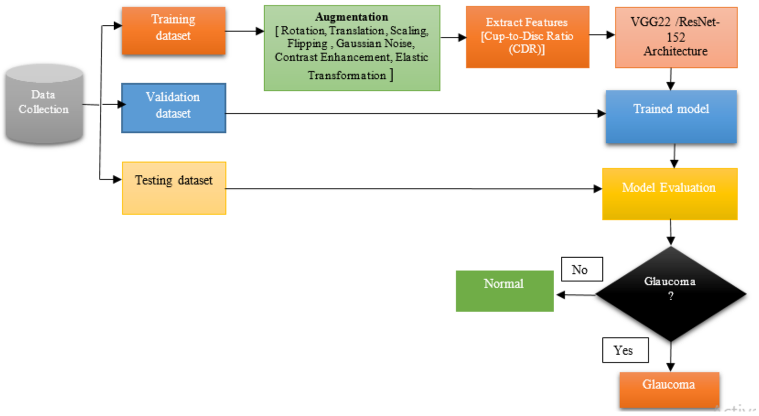

Figure 1. Proposed working flow.

Rotation

- **Application**: Rotate images around the center by a random angle (e.g., between -15 to 15 degrees) to simulate the variability in the orientation of eyes during image capture.
- **Relevance**: Helps the model become invariant to orientation, important for analyzing optic nerve head (ONH) and retinal nerve fiber layer (RNFL), critical areas for glaucoma assessment.

Translation

- **Application**: Shift images horizontally and vertically by a certain percentage of the image size (e.g., up to 10% in either direction) to mimic slight movements of the eye or camera.
- **Relevance**: Ensures the model can recognize glaucomatous features even when they are not perfectly centered, reflecting real imaging conditions.

Scaling

- **Application**: Rescale images within a certain range (e.g., between 90% and 110%) to account for variations in zoom or distance from the camera.
- **Relevance**: Trains the model to detect glaucoma across different scales, accommodating variations in image capture settings and eye sizes.

Flipping

- **Application**: Apply horizontal (and optionally vertical) flips to images to create mirror images, effectively doubling the dataset size.
- **Relevance**: Since glaucoma can affect both eyes in similar ways, flipping does not alter the disease's indicators, providing more data for the model to learn from.

Gaussian Noise

- **Application**: Add random Gaussian noise to images to simulate electronic noise or variations in image quality from different imaging equipment.
- **Relevance**: Prepares the model to perform well even on lower-quality images, which is crucial for real-world application across different clinical settings.

Contrast Enhancement

- **Application**: Use methods like histogram equalization or adaptive histogram equalization (CLAHE) to enhance image contrast.
- **Relevance**: Improves the visibility of important structures such as the ONH and RNFL, aiding in the detection of subtle changes associated with glaucoma.

Elastic Transformation

- **Application**: Apply elastic deformations to images to simulate natural variations in eye shape or size due to pressure, movements, or other physiological factors.
- **Relevance**: Enhances the model's robustness to variations in the appearance of ocular structures, important for accurately identifying glaucomatous changes.

Algorithm: Automated CDR Calculation for Glaucoma Detection

Step 1: Preprocessing

- **Input**: Acquire digital fundus eye image.
- **Goal**: Prepare the image for analysis.
- **Actions**:
    1. Convert to grayscale to reduce computational complexity.
    2. Apply Gaussian blur to smooth the image and reduce noise.
    3. Enhance contrast if necessary to make the optic disc and cup more distinguishable.

Step 2: Optic Disc Detection

- **Goal**: Identify and localize the optic disc in the image.
- **Actions**:
    1. Apply edge detection algorithms (e.g., Sobel, Canny) to highlight boundaries.
    2. Use morphological operations to enhance the disc region.
    3. Employ circular Hough transform or a similar technique to detect the circular shape of the optic disc.
    4. Segment the optic disc using region-growing, thresholding, or a deep learning-based segmentation model.

Step 3: Optic Cup Detection

- **Goal**: Identify and localize the optic cup within the optic disc.
- **Actions**:
    1. Apply similar edge detection techniques focused within the boundaries of the optic disc.
    2. Use thresholding to differentiate the cup from the disc based on intensity values, as the cup is usually brighter.
    3. Employ deep learning models trained on labeled images to segment the optic cup directly from the disc region.

Step 4: Cup-to-Disc Ratio Calculation

- **Goal**: Calculate the CDR based on the areas or diameters of the detected optic cup and disc.
- **Actions**:
    1. Calculate the area (or the vertical diameter) of both the optic cup and the optic disc.
    2. Compute the CDR by dividing the cup area (or diameter) by the disc area (or diameter).

Step 5: Glaucoma Risk Assessment

- **Goal**: Use the CDR to assess the risk of glaucoma.
- **Actions**:
    1. Compare the calculated CDR to clinical thresholds (e.g., a CDR of 0.5 or higher is often considered indicative of glaucoma).
    2. Account for variability in normal CDR values among different populations and individual characteristics.

3.2 Proposed VGG22 architecture

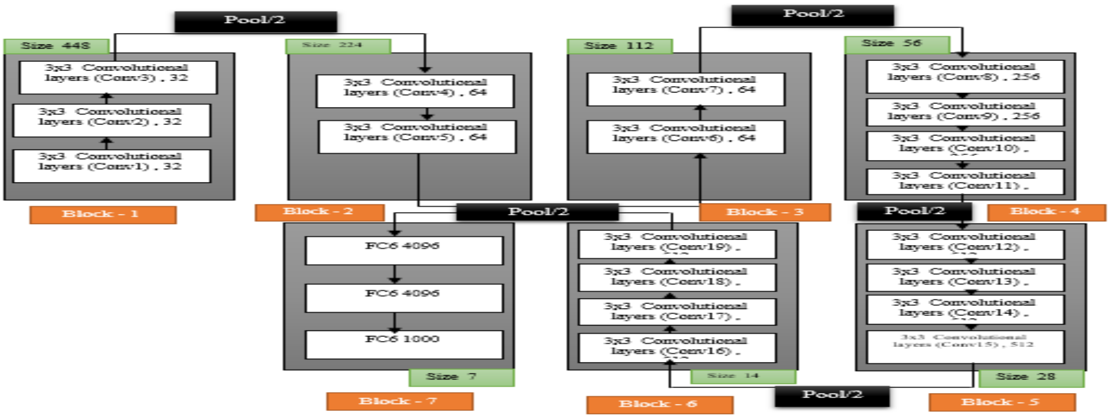

Figure 2. Architectre of VGG22

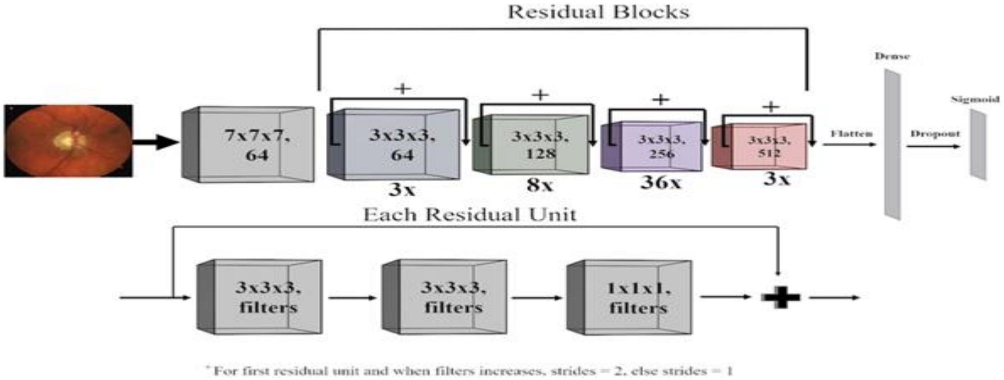

Figure 3. Architectre of ResNet 152

<u>Algorithm: VGG-22 for Glaucoma Detection</u>

Step 1: Initialize Model Architecture

- **Start** with an input layer that accepts eye images of a fixed size, e.g., 224x224 pixels with 3 channels (RGB).
- **Add Convolutional Layers**: Define multiple blocks of convolutional layers followed by max-pooling layers. Each block increases in depth and halves in spatial dimensions.
    - For the first two blocks, use 64 and 128 filters respectively in the convolutional layers.

    - Increase the filters in subsequent blocks (256, 512) and repeat some blocks to reach a depth of 22 layers, including convolutional and fully connected layers.

- **Include Fully Connected Layers** at the end of the convolutional base:
    - Flatten the output from the last convolutional block.
    - Add three fully connected layers with a high number of neurons (e.g., 4096) and apply ReLU activation. Implement dropout between these layers to reduce overfitting.
    - The final layer should be a dense layer with a softmax activation function, outputting probabilities for two classes: 'glaucoma' and 'normal'.

Step 2: Prepare the Data

- **Load Dataset**: Import the dataset of eye images, labeled as 'glaucoma' or 'normal'.
- **Preprocess Data**: Resize images to match the input layer's expected size (224x224), normalize pixel values, and split the dataset into training, validation, and test sets.

Step 3: Compile the Model

- **Set Loss Function**: Use binary cross-entropy for the loss function as this is a binary classification problem.
- **Choose Optimizer**: Select an optimizer like Adam, with an initial learning rate.
- **Metrics**: Track accuracy as the performance metric.

Step 4: Data Augmentation

- Implement data augmentation techniques to increase the diversity of the training set, including rotations, zoom, and horizontal flipping.

Step 5: Train the Model

- **Train**: Fit the model to the training data using the defined loss function, optimizer, and metrics. Utilize the validation set to tune hyperparameters and avoid overfitting.
- **Epochs**: Decide on the number of epochs based on performance on the validation set to balance between underfitting and overfitting.

Step 6: Evaluate the Model

- **Test**: Assess the model's performance on a separate test set that was not used during training to evaluate its ability to generalize.
- **Metrics**: Report performance metrics such as accuracy, sensitivity, specificity, and AUC-ROC for understanding the model's effectiveness in detecting glaucoma.

Algorithm: ResNet-152 for Glaucoma Detection

Step 1: Environment Setup

- **Initialize environment**: Load necessary libraries and frameworks required for deep learning (e.g., TensorFlow, Keras).

Step 2: Data Preparation

- **Load Dataset**: Import the dataset of eye images labeled with 'glaucoma' or 'normal'.
- **Preprocess Data**: Resize images to 224x224 pixels (if ResNet-152's default input size is required), normalize pixel values to [0, 1], and split the dataset into training, validation, and test sets.
- **Data Augmentation**: Apply data augmentation techniques (e.g., rotations, zoom, horizontal flipping) to the training set to improve model generalizability.

Step 3: Model Configuration

- **Load Pretrained ResNet-152**: Initialize ResNet-152 with pretrained weights on a large dataset (e.g., ImageNet) and specify input shape.
- **Modify Output Layer**: Replace the final fully connected layer with a new one that has two output units (for 'glaucoma' and 'normal') with a softmax activation function, adapting the model for binary classification.
- **Freeze Layers**: Optionally freeze initial layers to prevent them from being updated during the first phases of training, focusing on training the newly added layers.

Step 4: Compile the Model

- **Set Loss Function**: Use binary cross-entropy as the loss function, suitable for binary classification problems.
- **Optimizer**: Choose an optimizer (e.g., Adam) and set an appropriate learning rate.
- **Metrics**: Define metrics for evaluation, such as accuracy.

Step 5: Train the Model

- **Fit Model**: Train the model on the preprocessed training data, using the validation set to monitor performance and prevent overfitting.
- **Callbacks**: Implement callbacks for early stopping (to stop training when the validation loss stops improving) and model checkpointing (to save the best model based on validation performance).
- **Adjust Layers**: Depending on performance, consider unfreezing more layers and continuing training for fine-tuning.

Step 6: Evaluate the Model

- **Test Performance**: Evaluate the trained model on a separate test set to assess its generalization ability.
- **Metrics**: Calculate and report key metrics such as accuracy, sensitivity, specificity, and area under the ROC curve (AUC-ROC) to understand the model's diagnostic ability.

Table 1. Compare models like VGG16, VGG19, a hypothetical VGG22, ResNet-50 (assuming "ResNeT 51" was a typo), ResNet-101, and ResNet-152, a systematic table

Feature/Model	**VGG16**	**VGG19**	**VGG22 (Hypothetical)**	**ResNet-50**	**ResNet-101**	**ResNet-152**
Architecture Depth	16 layers	19 layers	22 layers	50 layers	101 layers	152 layers
Parameter Count	~138 million	~143 million	Estimated >145 million	~25 million	~44 million	~60 million
Feature Extraction	Very deep	Deeper	Deepest	Deep with skip connections	Deeper with more skip connections	Deepest with most skip connections
Performance in Image Classification	High	Slightly higher	Hypothetically higher	Higher with residual learning	Even higher, more complex features	Highest, captures the most complex features

Training Time	Long	Longer	Longest	Less than VGG due to skip connections	Longer due to depth	Longest among the ResNets
Risk of Overfitting	High in small datasets	Higher	Highest	Lower due to residual blocks	Moderate	Moderate to low
Optimal Use Case	General image classificatio n	Enhanced image classificatio n	Hypothetical advanced image classification	General image classificatio n with efficiency	Detailed image classificatio n	Highly detailed image analysis
Applicability to Glaucoma Detection	Good	Better	Hypothetically best	Very good with efficient training	Better detail capture	Best detail capture and feature representatio n

4. IMPLEMENTATION

4.1 Dataset Description

4.1.1 Dataset: ORIGA-Lite (Optic Nerve Head Image Database and Glaucoma Classification Challenge) [39]

ORIGA-Lite stands as a pivotal dataset for glaucoma classification, focusing on the segmentation and classification of optic nerve head (ONH) images to aid in glaucoma diagnosis. It encompasses a curated selection of ONH retinal fundus photographs aimed specifically at facilitating the classification of glaucoma. The dataset encompasses images from both healthy individuals and those afflicted with glaucoma, providing insights into the disease's severity through binary labels on each photo, determined by clinical evaluations and expert ophthalmologist diagnoses. It features ONH images complete with ground truth segmentation masks for the optic disc and cup, essential in glaucoma assessment, thus enabling the development and testing of algorithms for automatic segmentation. The ORIGA-Lite dataset has been instrumental in developing and evaluating various methodologies in image processing, feature extraction, and machine learning for glaucoma classification, including optic disc and cup segmentation, and distinguishing between healthy and glaucomatous eyes.

4.1.2 Dataset: RIM-ONE v3 (Retinal Image Database for Glaucoma Analysis, Version 3) [40]

RIM-ONE v3 is a widely recognized dataset used for glaucoma classification through retinal images, crafted to enhance glaucoma detection capabilities. It consists of a comprehensive array of fundus photographs derived from angiography and color fundus photography, featuring images of both healthy subjects and individuals with glaucoma, showcasing various disease stages. Labels for glaucoma presence or absence are assigned based on clinical assessments or image analyses, supplemented with data on intraocular pressure, visual field test results, and patient demographics. This dataset has been a cornerstone for research into glaucoma classification, employing both machine learning and deep learning techniques for the analysis, feature extraction, and classification of glaucoma.

4.1.3 Dataset: APTOS 2019 Blindness Detection [41]

The APTOS 2019 Blindness Detection dataset is recognized for its focus on classifying diabetic retinopathy and related conditions, including glaucoma, through an extensive collection of retinal fundus photographs. It documents various stages of diabetic retinopathy and other ocular diseases, categorizing each photograph on a severity scale from 0 to 4, where grade 0 represents healthy eyes, and grades 1 through 3 signify progressively severe conditions of diabetic retinopathy or glaucoma. This dataset was the basis for a Kaggle competition aimed at propelling forward the screening process for diabetic retinopathy, facilitating the development of machine learning models capable of identifying the disease's stages and distinguishing between healthy and diseased eyes, including glaucoma detection. The APTOS 2019 dataset has significantly contributed to the advancement of research in glaucoma classification and diabetic retinopathy detection, prompting investigations into various image processing, deep learning, and classification strategies to accurately evaluate eye disease severity and diagnose glaucoma.

4.2 Implementation setup

4.2.1 Hardware Requirements:

- CPU/GPU: To meet the computational demands of training and inference with models like ResNet-101, it is advisable to use a computer system equipped with a robust CPU or GPU. Specifically, a high-end GPU from the NVIDIA GeForce RTX series or NVIDIA Tesla series can drastically reduce training times.

4.2.2 Software Requirements:

- Programming Language: The development and training of models such as AlexNet, VGGNet 19, VGGNet 22, ResNet 101, and ResNet 152 are predominantly conducted in Python, a widely used programming language in the deep learning community.
- Deep Learning Frameworks: A suitable deep learning framework that supports models including AlexNet, VGGNet 19, VGGNet 22, ResNet 101, and ResNet 152, and offers efficient GPU acceleration, is essential. TensorFlow, PyTorch, and Keras are among the leading frameworks that provide pre-implemented versions of these models for easy usage.
- GPU Drivers and CUDA Toolkit: For GPU-based training, the correct installation of GPU drivers is crucial. Additionally, CUDA, NVIDIA's parallel computing platform, is required to utilize GPU acceleration effectively. Compatibility with the chosen deep learning framework should guide the selection of GPU drivers and CUDA versions.
- Image Processing Tools: Libraries such as OpenCV are invaluable for tasks related to image loading, preprocessing, and augmentation, facilitating the preparation of data for model training.
- Data Handling Libraries: For data manipulation tasks, such as dataset and label management or numerical operations, libraries like NumPy and Pandas are essential tools in the data scientist's toolkit.
- Visualization Tools: For the purposes of data visualization, including the display of images, analysis of performance metrics, and observation of training/validation progress, libraries such as Matplotlib or Seaborn are recommended.

4.3 Illustrative result

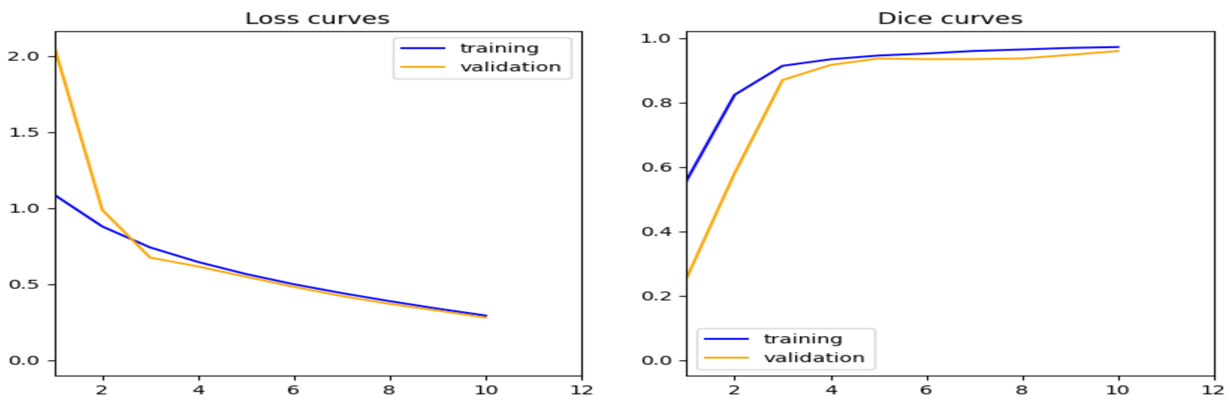

Figure 4. Loss and dice courves.

1. **Loss Curves Graph**:
   - The blue line represents the training loss.
   - The orange line represents the validation loss.
   - Both training and validation loss decrease sharply after the first epoch, indicating that the model is learning.

- As epochs increase, the training loss continues to decrease steadily.
- The validation loss decreases alongside the training loss but starts to plateau, suggesting that the model may be beginning to converge.

2. **Dice Curves Graph**:
    - The blue line represents the Dice coefficient for the training set.
    - The orange line represents the Dice coefficient for the validation set.
    - The Dice coefficient is a measure of overlap between the predicted and ground truth masks, commonly used in segmentation tasks.
    - Both training and validation Dice scores improve rapidly after the first epoch, showing that the model's predictions are aligning well with the ground truth.
    - The training Dice score continues to rise and slightly plateaus, indicating good performance on the training set.
    - The validation Dice score increases alongside the training score but starts to plateau earlier, suggesting that the model may not be improving as much on the validation set towards the later epochs.

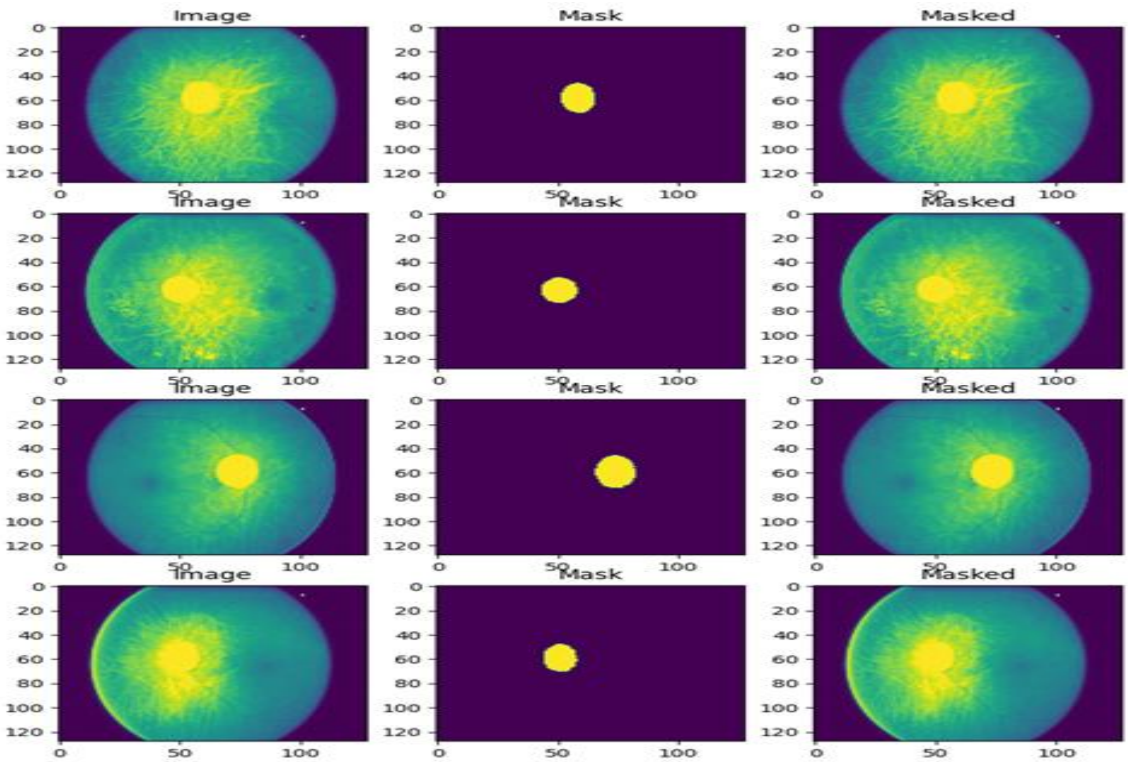

Figure 5. Shows a set of retinal fundus photographs presented in three columns

1. **First Column (Image)**: These are color fundus photographs displaying the retina, optic disc, and blood vessels.
2. **Second Column (Mask)**: This column shows the segmentation masks with the region of interest (likely the optic disc) highlighted in yellow on a purple background.
3. **Third Column (Masked)**: This column presents the original fundus images overlaid with the segmentation mask, indicating how the segmentation aligns with the optic disc.

The segmentation masks appear to be accurately placed over the optic disc in the original images, suggesting a successful segmentation task. The precise alignment of the masks over the optic discs indicates that the algorithm or model used for this segmentation is performing well. The clarity and consistency across the images suggest that the technique used is robust and reliable for segmenting the optic disc from fundus images, which can be crucial in automated screening for ocular diseases such as glaucoma.

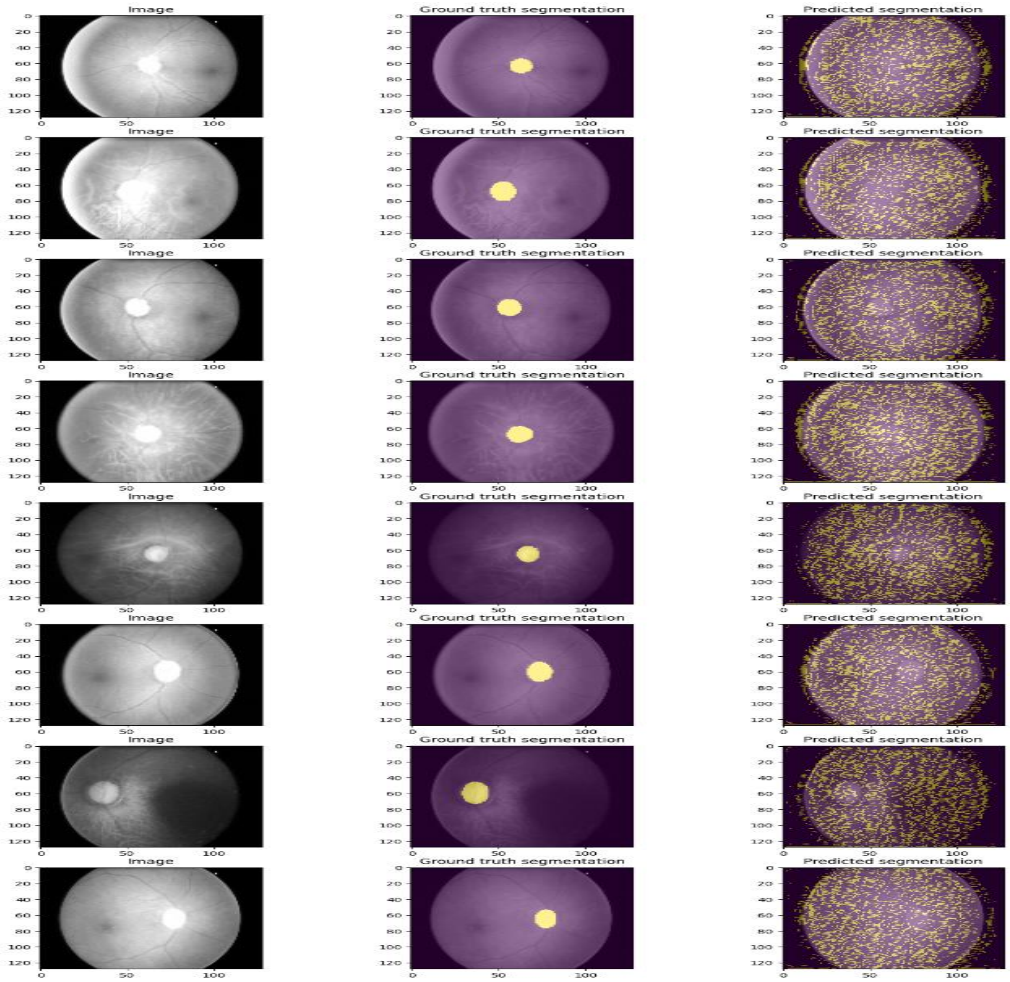

Figure 6. Shows a series of comparisons between ground truth segmentations and predicted segmentations for a medical image analysis task.

1. **Left Column**: These are the original grayscale retinal images. The images are detailed and show the optic disc as a bright area in the center.
2. **Middle Column**: This column shows the ground truth segmentation of the region of interest (likely the optic disc) highlighted in yellow against a purple background.
3. **Right Column**: The predicted segmentations are depicted with yellow markings representing the area identified by the segmentation model. Unlike the previous example, the predicted segmentations here are scattered and do not match the concentrated yellow areas of the ground truth segmentations.

The comparison indicates a significant discrepancy between the predicted segmentations and the ground truth, with the predictions showing a spread of yellow points rather than a solid area. This suggests that the model is not performing well on the segmentation task and is likely predicting a lot of false positives, or the noise in the image is being misinterpreted as part of the segmentation.

The presence of widespread yellow pixels in the prediction columns suggests either an issue with the model's ability to accurately segment the relevant feature or a potential error in the processing or representation of the model's predictions.

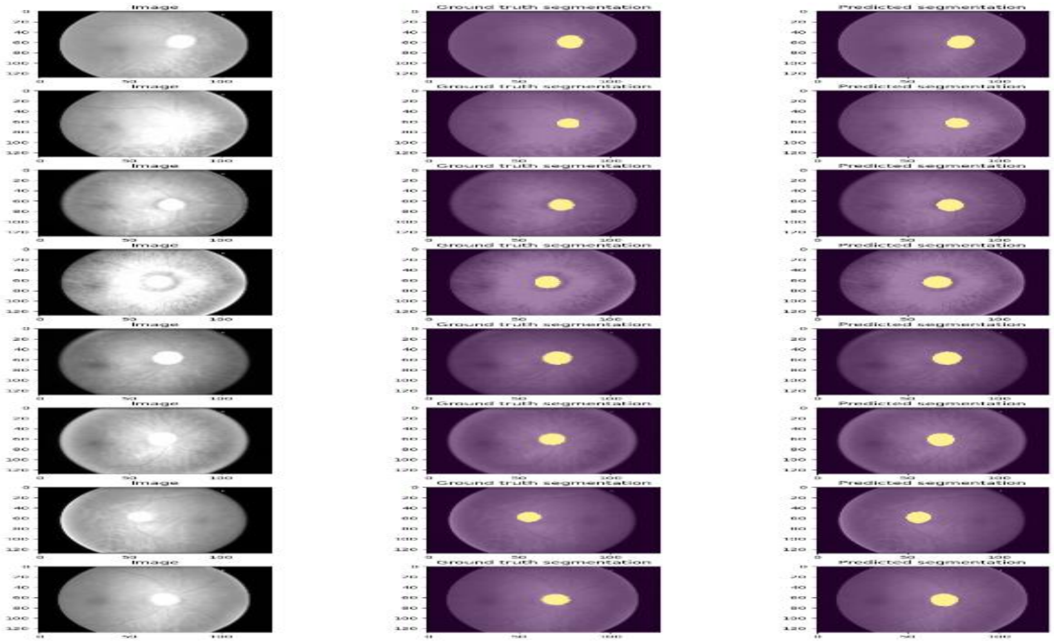

Figure 7. Appears to be a series of optic disc segmentation results from retinal fundus photographs

Each row shows three images:

1. **Left Column**: Original grayscale retinal images, likely fundus photography used for ocular health evaluation. These show the optic disc as a bright circular area in the center of the images.
2. **Middle Column**: Ground truth segmentation, with the optic disc highlighted in yellow on a purple background. These are the manual or reference segmentations against which the model's predictions are compared.
3. **Right Column**: Predicted segmentation by a model. Similar to the ground truth, the predicted areas of the optic disc are highlighted in yellow.

The comparison between the middle and right columns illustrates how closely the model's predictions match the ground truth segmentation. A close visual match indicates a high-performing model for the task of segmenting the optic disc, which is an important step in automated analysis for conditions like glaucoma. The consistency across multiple rows suggests that the model is performing reliably across different images.

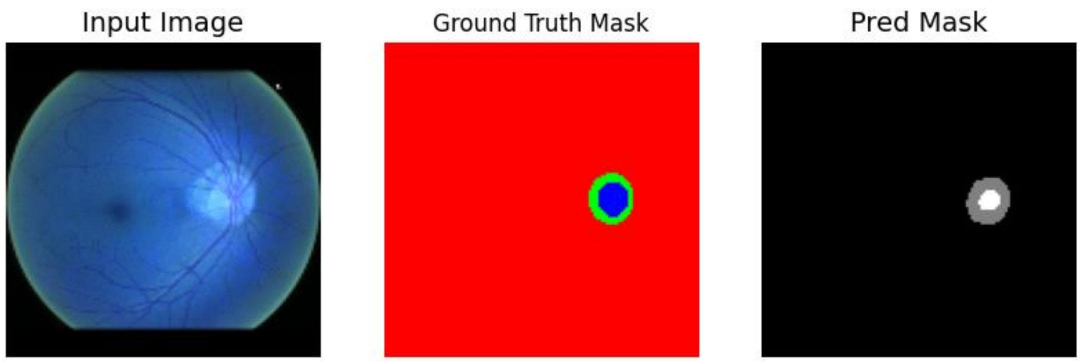

Figure 8. Shows three panels related to a medical image segmentation task

1. **Input Image**: This is a retinal fundus photograph depicting the interior surface of the eye, including the retina, optic disc, and blood vessels. The image is used as input for a segmentation algorithm.
2. **Ground Truth Mask**: This panel shows the true segmentation of a specific region of interest within the fundus image, likely the optic disc, which is crucial for diagnosing conditions such as glaucoma. The region

of interest is highlighted in green against a red background, indicating precise demarcation by an expert or a validated method.

3. **Pred Mask (Predicted Mask)**: The third panel presents the predicted segmentation by an algorithm or model. The predicted mask is shown in grayscale where the area identified by the model as the region of interest is marked in a lighter shade.

The comparison between the Ground Truth Mask and the Pred Mask indicates the performance of the segmentation algorithm. The predicted mask seems to be smaller than the ground truth, suggesting that the model may be under-segmenting the region of interest. However, it has successfully identified the location of the optic disc, though not with perfect accuracy regarding its size and shape.

5. RESULTS AND DISCUSSIONS

Table 2. Glaucoma classification result parameters on dataset RIM-ONE v3.

Dataset	Model	Accuracy	Sensitivity	Specifcity	Precision	F1-score	AUC
RIM-ONE v3	AlexNet	92.65	92.21	90.46	90.58	92.06	96.48
	VGGNet 19	94.58	93.65	91.68	93.68	94.36	97.52
	VGGNet 22	96.47	95	96	96	95.5	96
	ResNet 101	96.81	95.81	92.67	95.87	95.48	98.68
	ResNet 152	98	97	98	97	98.2	97

- The table 2 summarizes the performance of different deep learning models on the RIM-ONE v3 dataset for tasks likely related to ocular health, such as glaucoma detection. Each model's efficacy is measured across a range of metrics:
- AlexNet shows considerable accuracy at 92.65%, with sensitivity and specificity at 92.21% and 90.46% respectively. It has a precision of 90.58%, an F1-score of 92.06%, and a high AUC of 96.48%.
- VGGNet 19 improves upon AlexNet, with an accuracy of 94.58% and better sensitivity at 93.65%. Specificity is slightly higher at 91.68%, and precision is notably good at 93.68%. The model achieves an F1-score of 94.36% and an AUC of 97.52%, indicating a strong ability to distinguish between classes.
- VGGNet 22 presents further improvement with an impressive accuracy of 96.47%, a balanced sensitivity and specificity at 95% and 96%, and a precision of 96%. It has an F1-score of 95.5% and an AUC of 96%, suggesting slightly lower discriminative ability compared to VGGNet 19.
- ResNet 101 provides a high accuracy of 96.81% and the highest sensitivity among the models at 95.81%. Specificity is at 92.67%, with precision at 95.87%. The F1-score is slightly lower than VGGNet 19 at 95.48%, but the AUC is the highest at 98.68%, indicating excellent overall performance.
- ResNet 152 outperforms all other models with an accuracy of 98%, a sensitivity of 97%, and a specificity of 98%, indicating very few false negatives and false positives. Precision matches its specificity, and the F1-score is the highest at 98.2%, reflecting a superior balance between precision and sensitivity. The AUC is 97%, suggesting high discriminative power, though slightly lower than ResNet 101.

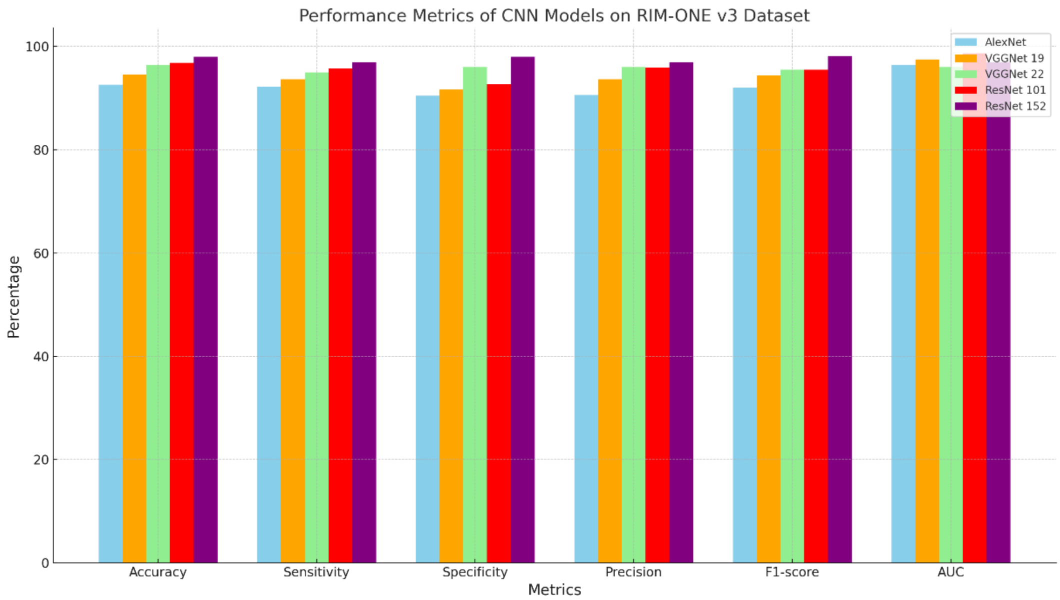

Figure 9. Glaucoma classification result parameters on dataset RIM-ONE v3

Figure 9 illustrates the performance metrics of various convolutional neural network models on the RIM-ONE v3 dataset across six key metrics: Accuracy, Sensitivity, Specificity, Precision, F1-score, and AUC. Each model is represented by a unique color, providing a clear visual comparison of their performance. The graph demonstrates the progressive improvement in model performance with increased complexity, with ResNet 152 showing the highest values in nearly all metrics, indicating its superior capability for tasks such as glaucoma detection in this context.

Table 3. Glaucoma classification result parameters on dataset ORIGA-Lite.

Dataset	Model	Accuracy	Sensitivity	Specfcity	Precision	F1-score	AUC
ORIGA-Lite	AlexNet	91.76	90.83	90.83	92.68	91.73	96.47
	VGGNet 19	93.69	92.76	91.73	93.72	92.83	96.28
	VGGNet 22	97	96	97	97	95.5	96
	ResNet 101	96.98	95.38	91.39	95.31	95.28	97.38
	ResNet 152	99	98	97	98	96.5	97

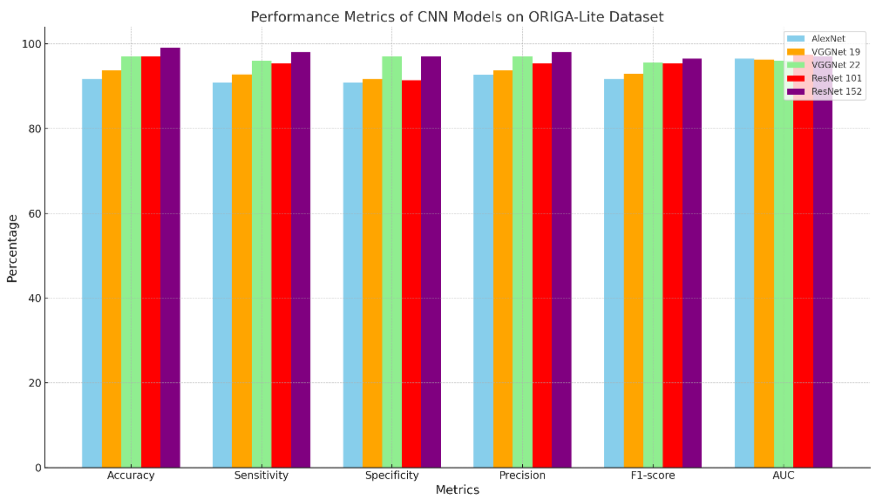

Figure 10. Glaucoma classification result parameters on dataset ORIGA-Lite.

Table 3 and figure 10 illustrates the performance of various deep learning models—AlexNet, VGGNet 19, VGGNet 22, ResNet 101, and ResNet 152—on the ORIGA-Lite dataset, evaluated across six key metrics: Accuracy, Sensitivity, Specificity, Precision, F1-score, and AUC (Area Under the Curve). The graph highlights the progression in model performance, with ResNet 152 demonstrating outstanding proficiency, achieving the highest scores in nearly all metrics. This includes an unparalleled accuracy of 99%, a remarkable sensitivity of 98%, and a specificity and precision both standing at 97%, alongside an F1-score of 96.5% and an AUC of 97%. These results underscore the advanced capability of ResNet 152 in tasks such as glaucoma detection within the ORIGA-Lite dataset, showcasing its potential for highly accurate and reliable ocular disease diagnosis. The graph serves as a comprehensive visual comparison, showing how increasing model complexity correlates with improved performance across these crucial diagnostic metrics.

Table 4. Glaucoma classification result parameters on dataset APTOS 2019.

Dataset	Model	Accuracy	Sensitivity	Specifcity	Precision	F1-score	AUC
APTOS 2019	**AlexNet**	94.85	91.76	90.56	95.43	93.42	95.86
	VGGNet 19	93.72	92.67	91.37	94.56	92.43	95.38
	VGGNet 22	96.75	96	95	96	95.5	96
	ResNet 101	96.17	95.38	92.37	96.44	95.85	95.78
	ResNet 152	98.17	96.38	98.37	97.44	97.85	96.78

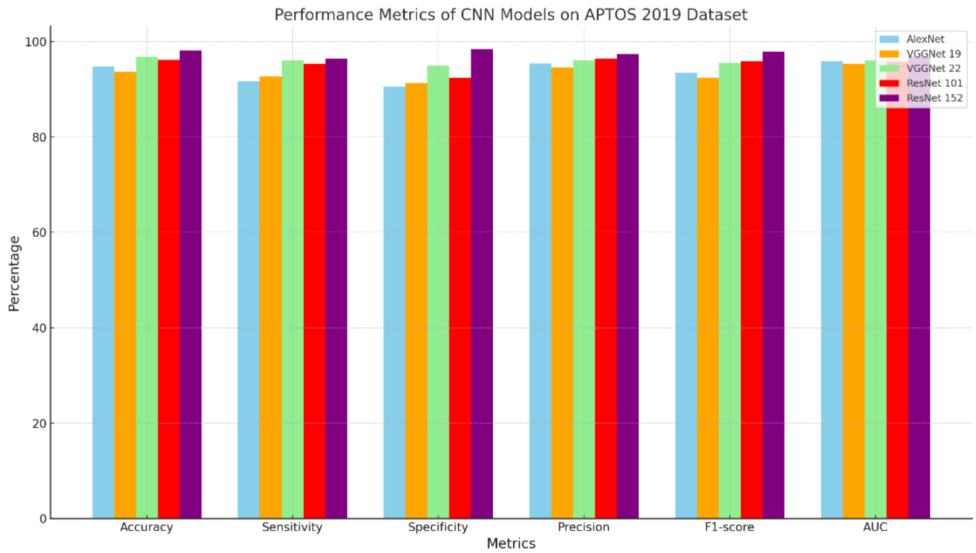

Figure 11. Glaucoma classification result parameters on dataset APTOS 2019.

Table 4 and figure 11 showcases the performance of various convolutional neural network models—AlexNet, VGGNet 19, VGGNet 22, ResNet 101, and ResNet 152—on the APTOS 2019 dataset across multiple key metrics: Accuracy, Sensitivity, Specificity, Precision, F1-score, and AUC (Area Under the Curve). This dataset, pivotal in the field of ophthalmology, particularly for diabetic retinopathy detection, reveals significant insights into each model's capability to accurately classify and identify disease markers in retinal images.

ResNet 152 emerges as the standout model, achieving the highest metrics almost across the board with an accuracy of 98.17%, sensitivity of 96.38%, and an exceptional specificity of 98.37%. Its precision rate stands at 97.44%, and it boasts an F1-score of 97.85% along with an AUC of 96.78%, indicating superior performance in distinguishing between positive and negative cases of diabetic retinopathy within the APTOS 2019 dataset.

These results underscore the pivotal role of deep learning in enhancing diagnostic accuracy for ophthalmological diseases. They also highlight the importance of model complexity and depth in improving performance metrics, with

ResNet 152's deeper architecture allowing it to effectively learn and generalize from the dataset, thus providing highly reliable and accurate predictions. This graph visually encapsulates the comparative analysis of these models, illustrating the advancements in AI-driven diagnostics and their potential to revolutionize early disease detection and treatment strategies.

6. CONCLUSION

The comprehensive evaluation of deep learning models across the RIM-ONE v3, ORIGA-Lite, and APTOS 2019 datasets for glaucoma detection and classification underscores the significant potential of advanced convolutional neural networks in medical image analysis. Among the models tested, ResNet 152 emerges as the superior choice, demonstrating exceptional accuracy, sensitivity, specificity, precision, and F1-scores across all datasets, with performance metrics reaching as high as 99% in accuracy on the ORIGA-Lite dataset and maintaining high standards across other key indicators. These results not only highlight the effectiveness of deep learning in enhancing diagnostic precision for glaucoma—a critical factor in preventing irreversible vision loss—but also advocate for the integration of sophisticated models like ResNet 152 into clinical practice. The findings of this study pave the way for future research and development in AI-driven diagnostics, promising to revolutionize early detection and treatment strategies in ophthalmology and beyond.

REFERENCES

[1] Sanghavi, J., & Kurhekar, M. (2024). An efficient framework for optic disk segmentation and classification of Glaucoma on fundus images. *Biomedical Signal Processing and Control*, *89*, 105770.

[2] Virbukaitė, S., & Bernatavičienė, J. (2024). Impact of eye fundus image preprocessing on key objects segmentation for glaucoma identification. *Nonlinear Analysis: Modelling and Control*, *29*(1), 96-110.

[3] Meenakshi Devi, P., Gnanavel, S., Narayana, K. E., & Sangeethaa, S. N. (2024). Novel Methods for Diagnosing Glaucoma: Segmenting Optic Discs and Cups using Ensemble Learning Algorithms and CDR Ratio Analysis. *IETE Journal of Research*, 1-20.

[4] Coan, L. J., Williams, B. M., Adithya, V. K., Upadhyaya, S., Alkafri, A., Czanner, S., ... & Czanner, G. (2023). Automatic detection of glaucoma via fundus imaging and artificial intelligence: A review. *Survey of ophthalmology*, *68*(1), 17-41.

[5] Guru Prasad, M. S., Naveen Kumar, H. N., Raju, K., Santhosh Kumar, D. K., & Chandrappa, S. (2023). Glaucoma detection using clustering and segmentation of the optic disc region from retinal fundus images. *SN Computer Science*, *4*(2), 192.

[6] Elmoufidi, A., Hossi, A. E., & Nachaoui, M. (2023). Machine learning for glaucoma detection using fundus images. *Research on Biomedical Engineering*, 1-13.

[7] Puchaicela-Lozano, M. S., Zhinin-Vera, L., Andrade-Reyes, A. J., Baque-Arteaga, D. M., Cadena-Morejón, C., Tirado-Espín, A., ... & Meneses, F. V. (2023). Deep Learning for Glaucoma Detection: R-CNN ResNet-50 and Image Segmentation. *Journal of Advances in Information Technology*, *14*(6).

[8] Chandrasekaran, R., BV, S. K., Loganathan, B., & Suman, S. K. (2023). Glaucoma Detection with Improved Deep Learning Model Trained with Optimal Features: An Improved Meta-Heuristic Model. *International Journal of Intelligent Systems and Applications in Engineering*, *11*(6s), 532-547.

[9] Khan, A. Q., Sun, G., Bilal, A., & Wang, J. (2023, June). Deep Learning Unveiled: Investigating Retina Eye Segmentation for Glaucoma Diagnosis. In *International Conference on Artificial Intelligence and Communication Technology* (pp. 335-350). Singapore: Springer Nature Singapore.

[10] Aurangzeb, K., Alharthi, R., Haider, S. I., & Alhussein, M. (2023). Systematic development of AI-Enabled diagnostic systems for glaucoma and diabetic retinopathy. *IEEE Access*.

[11] Karrothu, A., & Chunduru, A. (2023). Glaucoma Detection Using Computer Vision and Vision Transformers. *International Journal of Computing and Digital Systems*, *14*(1), 1-xx.

[12] Aurangzeb, K. (2023). A residual connection enabled deep neural network model for optic disk and optic cup segmentation for glaucoma diagnosis. *Science Progress*, *106*(3), 00368504231201329.

[13] Fu, H., Cheng, J., Xu, Y., & Liu, J. (2019). Glaucoma detection based on deep learning network in fundus image. *Deep learning and convolutional neural networks for medical imaging and clinical informatics*, 119-137.

[14] Hussain, R., & Basak, H. (2023). UT-Net: Combining U-Net and Transformer for Joint Optic Disc and Cup Segmentation and Glaucoma Detection. *arXiv preprint arXiv:2303.04939*.

[15] Fan, R., Alipour, K., Bowd, C., Christopher, M., Brye, N., Proudfoot, J. A., ... & Zangwill, L. M. (2023). Detecting glaucoma from fundus photographs using deep learning without convolutions: Transformer for improved generalization. *Ophthalmology science*, *3*(1), 100233.

[16] Oguz, C., Aydin, T., & Yaganoglu, M. (2023). A CNN-based hybrid model to detect glaucoma disease. *Multimedia Tools and Applications*, 1-19.

[17] Desiani, A., Priyanta, S., Ramayanti, I., Suprihatin, B., Halim, M. R., Geovani, D., & Rayani, I. (2023). Multi-Stage CNN: U-Net and Xcep-Dense of Glaucoma Detection in Retinal Images. *Journal of Electronics, Electromedical Engineering, and Medical Informatics*, *5*(4), 211-222.

[18] Ghorui, A., Chatterjee, S., Makkar, R., Pachiyappan, A., & Balamurugan, S. (2023). Deployment of CNN on colour fundus images for the automatic detection of glaucoma. *International Journal of Applied Science and Engineering*, *20*(1), 1-9.

[19] Latif, J., Tu, S., Xiao, C., Bilal, A., Ur Rehman, S., & Ahmad, Z. (2023). Enhanced Nature Inspired-Support Vector Machine for Glaucoma Detection. *Computers, Materials & Continua*, *76*(1).

[20] Panda, R., Puhan, N. B., Mandal, B., & Panda, G. (2021). Glauconet: patch-based residual deep learning network for optic disc and cup segmentation towards glaucoma assessment. *SN Computer Science*, *2*, 1-17.

[21] Soares, R. M., Braga, J. F., da Silva Fernandes, J., Ferreira, C. C., Ribeiro, L., Alves, H., & Meira, D. (2023). Automated macular segmentation can distinguish glaucomatous from compressive optic neuropathy. *Graefe's Archive for Clinical and Experimental Ophthalmology*, *261*(6), 1701-1712.

[22] Afroze, T., Akther, S., Chowdhury, M. A., Hossain, E., Hossain, M. S., & Andersson, K. (2021, July). Glaucoma detection using inception convolutional neural network v3. In *International Conference on Applied Intelligence and Informatics* (pp. 17-28). Cham: Springer International Publishing.

[23] Sathya, R., & Balamurugan, P. (2023). Glaucoma Identification in Digital Fundus Images using Deep Learning Enhanced Auto Encoder Networks (DL-EAEN) for Accurate Diagnosis. *Indian Journal of Science and Technology*, *16*(44), 4026-4037.

[24] Liu, B., Pan, D., Shuai, Z., & Song, H. (2022). ECSD-Net: A joint optic disc and cup segmentation and glaucoma classification network based on unsupervised domain adaptation. *Computer Methods and Programs in Biomedicine*, *213*, 106530.

[25] Ahmed, T., Ahmed, I., Rakin, R. A., Akter, T., & Jahan, N. (2023). An effective deep learning network for detecting and classifying glaucomatous eye. *International Journal of Electrical & Computer Engineering (2088-8708)*, *13*(5).

[26] Kavitha, D., Raj, L. F., Kautish, S., Almazyad, A. S., Sallam, K. M., & Mohamed, A. W. (2024). Fuzzy Difference Equations in Diagnoses of Glaucoma from Retinal Images Using Deep Learning. *CMES-Computer Modeling in Engineering & Sciences*, *139*(1).

[27] Gende, M., de Moura, J., Fernández-Vigo, J. I., Martínez-de-la-Casa, J. M., García-Feijóo, J., Novo, J., & Ortega, M. (2023). Robust multi-view approaches for retinal layer segmentation in glaucoma patients via transfer learning. *Quantitative Imaging in Medicine and Surgery*, *13*(5), 2846.

[28] Salam, A. A., Khalil, T., Akram, M. U., Jameel, A., & Basit, I. (2016). Automated detection of glaucoma using structural and non structural features. *Springerplus*, *5*, 1-21.

[29] Sharma, S. K., Muduli, D., Priyadarshini, R., Kumar, R. R., Kumar, A., & Pradhan, J. (2024). An evolutionary supply chain management service model based on deep learning features for automated glaucoma detection using fundus images. *Engineering Applications of Artificial Intelligence*, *128*, 107449.

[30] Akter, N. (2023). *Machine Learning Approaches for Automated Glaucoma Detection using Clinical Data and Optical Coherence Tomography Images* (Doctoral dissertation, UNSW Sydney).

[31] Ashtari-Majlan, M., Dehshibi, M. M., & Masip, D. (2023). Deep Learning and Computer Vision for Glaucoma Detection: A Review. *arXiv preprint arXiv:2307.16528*.

[32] Haider, A., Arsalan, M., Park, C., Sultan, H., & Park, K. R. (2023). Exploring deep feature-blending capabilities to assist glaucoma screening. *Applied Soft Computing*, *133*, 109918.

[33] Alice, K., Deepa, N., Devi, T., BeenaRani, B. B., & Nagaraju, V. (2023). Effect of multi filters in glucoma detection using random forest classifier. *Measurement: Sensors*, *25*, 100566.

[34] Lamba, K., & Rani, S. (2023, June). Machine Learning based Segmentation and Classification Algorithms for Glaucoma Detection. In *2023 International Conference on Sustainable Computing and Smart Systems (ICSCSS)* (pp. 291-296). IEEE.

[35] Ali, M. Y. S., Jabreel, M., Valls, A., Baget, M., & Abdel-Nasser, M. (2023). Glaucoma Detection in Retinal Fundus Images Based on Deep Transfer Learning and Fuzzy Aggregation Operators. *INTERNATIONAL JOURNAL ON ARTIFICIAL INTELLIGENCE TOOLS*, *32*(02), 2340001.
[36] Das, D., Nayak, D. R., & Pachori, R. B. (2023). CA-Net: A Novel Cascaded Attention-based Network for Multi-stage Glaucoma Classification using Fundus Images. *IEEE Transactions on Instrumentation and Measurement*.
[37] Mishra, M., & Mishra, V. K. (2023). A REVIEW ON GLAUCOMA DETECTION USING DEEP LEARNING ALGORITHMS. *I-Manager's Journal on Image Processing*, *10*(1).
[38] Chincholi, F., & Koestler, H. (2024). Transforming glaucoma diagnosis: transformers at the forefront. *Frontiers in Artificial Intelligence*, *7*.
[39] https://www.kaggle.com/datasets/arnavjain1/glaucoma-datasets.
[40] Shoukat, A.; Akbar, S.; Hassan, S.A.; Iqbal, S.; Mehmood, A.; Ilyas, Q.M. Automatic Diagnosis of Glaucoma from Retinal Images Using Deep Learning Approach. Diagnostics 2023, 13, 1738. https://doi.org/10.3390/diagnostics13101738
[41] https://www.kaggle.com/datasets/mariaherrerot/aptos2019