**A Comprehensive Review on Deep Learning Techniques used in Diagnosing Retinal Diseases on Fundus Images**

**Author’s VIJAYA MADHAVI V1  Author’s Dr. P. LALITHA SURYA KUMARI2\***

# ***1*Research Scholar, Department of CSE, Koneru Lakshmaiah Education Foundation, Hyderabad, Telangana, India.**

Email: madhavikrishnat@klh.edu.in

# **2Professor, Department of CSE, Koneru Lakshmaiah Education Foundation, Hyderabad, Telangana, India**

#  Email:vlalithanagesh@gmail.com

 **Abstract: Due to the exponential growth of the computing infrastructure, there has been an unheard increase in the deployment of computer vision and deep learning in recent years. The same was true for retinal image processing, and effective artificial intelligence models were created employing a range of visual markers extracted from eye image data to identify various retinal diseases. The detection of five important eye diseases—diabetic retinopathy, glaucoma, age-related macular degeneration, cataract, and retinopathy of prematurity—has been made possible in recent years by the use of a variety of deep learning techniques. This article is structured along the pipeline for implementing deep learning, with an extensive review of various approaches for each of the five retinal diseases mentioned after an illustration of commonly used datasets, evaluation metrics, image pre-processing techniques, and deep learning backbone models. The article concludes by listing the eight main study areas that are now being studied in the area of diagnosing retinal diseases, outlining their primary difficulties and potential future prospects for the research community at large.**

***Keyterms:- Computer vision, deep learning, fundus images, retinal disease diagnosis, artificial intelligence, diabetic retinopathy, glaucoma, cataract***

I.INTRODUCTION

Numerous imaging techniques have been developed over the years to study the human eye, but "Fundus Imaging" has become more well-known because of its non-invasive and affordable attributes. Fundus photography includes employing a monocular camera to record the projection of the fundus, or the back of the eye, onto a two-dimensional plane. A 2D fundus scan can be used to identify a variety of ocular structures, biomarkers, and anomalies (Figure 1). The identification of retinal disorders is aided greatly by many of these visual indicators.Microaneurysms (MAs), which resemble tiny red dots, typically develop as a result of constricted capillaries and a lack of oxygen. White soft patches that are referred to as Soft Exudates occur when the supply entirely stops working due to certain arteriolar obstructions (SEs). Hemorrhages, which appear as dark red spots when retinal arteries rupture as a result of arteriole pressure buildup, are occasionally seen. Hard Exudates (HEs), which are hard yellow waxy substances, are produced when proteins and fat seep from defective vessel walls. Examining the existence of these lesions in conjunction with other retinal biomarkers, such as the macular area, the fovea, the optic disc, and blood vessels, can offer important insights into some of the main retinal illnesses and help with their diagnosis.

There is a major shortage of qualified ophthalmologists in highly populated nations like India who can carry out such laborious tasks [1].

Some of the major eye conditions that, if not properly managed, might result in blindness include diabetic retinopathy (DR), glaucoma, age-related macular degeneration (AMD), diabetic macular edoema(DME), retinopathy of prematurity (ROP), and cataract. Such retinal illnesses typically require specialist care and considerable competence to screen [2].



**Fig1: Fundus Image**

Classification and segmentation tasks are the two main types of DL tasks used in the diagnosis of retinal diseases. The task of direct classification of input photos into multiple illness categories is referred to as the classification task. Similar to this, segmenting isimportant for biomarkers and significant lesions from a patient's fundus image can offer a wealth of information on the nature and classification of retinal illnesses. For such purposes, numerous DL architectures have been created and evaluated, as is extensively illustrated in [3].

Digital processors have grown exponentially recently, and data-driven technologies have made artificial intelligence (AI)-based medical screening systems increasingly common. These systems provide workable and affordable options for the automatic identification of retinal illnesses. In particular, fundus image analysis has shown tremendous growth and promise in computer vision and deep learning (DL) techniques. [4].

Additionally, this article adopts a DL process pipeline approach to retinal disease diagnosis and surveys recent articles on a diagnosis of five major eye diseases, namely diabetic retinopathy, glaucoma, age-related macular degeneration, cataract, and retinopathy of prematurity. This is in contrast to recently published review articles on this topic [5]-[9].

 In addition, it fully describes all datasets that are available for the aforementioned illnesses together with their ground truth descriptions.

It contains an extensive literature study on the DL implementation of five major retinal diseases along with tabulating their comparative performances It also discusses various research directions that are currently available in this field. It provides knowledge about broadly used image pre-processing methods, evaluation methods, and commonly used DL backbone approaches for retinal disease diagnosis.

II.DATASETS AND EVALUATION METRICS

Using reflected light that is projected onto an image plane, fundus photography is the method of obtaining a two-dimensional image of the three-dimensional ocular retinal fundus. The fundus image datasets that are frequently used for DL-based diagnosis of the retinal disorders. The datasets are used for disease diagnosis, and the results are displayed using color coding. All the datasets are given in a single table for simplicity of understanding and comparative analysis.

III.PREPROCESSING TECHNIQUES

Fundus images are typically pre-processed before training in order to enhance learning and create strong prediction models.The learning stag is done to make up for the noise that was brought on by the different types of image-capture devices employed in the different illumination settings during the imaging. Due to the intricacy of the retinal structure and the low picture quality, many significant biomarkers and lesions may go undetected, as seen in Figure 2. Prior to the execution of the DL model, pre-processing techniques are also utilized to improve the fundus picture features in addition to eliminating undesirable noise. In Table 2, some of the pre-processing methods that are frequently employed to diagnose retinal diseases from color fundus images are listed.



**Fig:- Retinal Fundus Images**

|  |  |
| --- | --- |
| **Fundus Image Preprocessing Technique** | **Description** |
| CLAHE(Contrast Enhancement) | Contrast Limited Adaptive Histogram Equalization is a widely used technique, especially for Fundus images. |
| Colour Space Transformation | The extraction of green channels from fundus images is famous as they offer high contrast images with rich visual information |
| Noise Removal | Many Denoising algorithms like Gaussian filters, median filters, non-local means denoising etc. are utilized for removing unwanted noise. |
| Cropping and Extracting Region of Interest | Cropping is done to extract the exact region of interest from the entire fundus image. For example to investigate optic disc size, only that portion of the image is cropped and utilized as ROI in model training in reducing the unwanted learning burden |
| Augmentation  | Augmentation techniques like image rotation, rescaling, flipping, translation etc. Are employed to balance the image datasets. |

**Table2:Commonly used Preprocessing techniques for Image Enhancement**

IV.DEEP LEARNING TECHNIQUES

Artificial neural network-based deep learning (DL) is a subclass of artificial intelligence techniques (learning methods inspired by the biological structure of the human brain). The latent and intrinsic relationship of the input data is automatically learned through mathematical representations in the DL process. As opposed to typical machine learning (ML) techniques, deep learning (DL) methods directly extract meaningful features from the data without relying on manually created characteristics. This qualifies DL for medical image analysis, as the features can be automatically learned from intricate visual data. In this section we go over the designs of some of the backbone models that are often employed, particularly for classification and segmentation tasks in the diagnosis of retinal diseases.

V.MODELS OF CLASSIFICATION

A.CONVOLUTIONAL NEURAL NETWORKS(CNN)

Convolution neural network (CNN) architectures are among the most often utilised DL designs for effective training over several layers [10]. The overall layout of CNN is shown in Figure 3. A CNN is made up of three main parts: convolution layers, pooling layers, and fully connected layers. Two steps make up the training process. The first stage is referred to as the "forward stage," in which the input image is represented with the proper weights and biases in each layer. The anticipated output is then used to calculate the loss function by comparing it against values from the ground truth. The second stage, referred to as the "backward stage," computes the gradients of each parameter using the loss function. The parameters are setup and modified for the succeeding forward.



Figure 3:CNN Architecture

B. VGGNET

VGG Network is yet another backbone network that is frequently utilised to classify retinal disorders (VGGNet). In 2014 [11], Karen Simonyan and Andrew Zisserman made this suggestion. The architecture of a VGGNet is seen in Figure 4. The VGG acronym stands for Visual Geometry Group, which starting with VGG-16 through VGG-19 produced numerous versions of Convolution network models for various image categorization applications. Researching how the depth of CNN affects the precision of picture classification was the initial motivation for the creation of VGG. In order to improve the depth of the network without using too many parameters, a tiny kernel is employed in all levels of the model.



**Figure4: VGGNET**

C.RESNET

The 152 layers that make up the residual network (ResNet) [12] are created by stacking the individual residual blocks depicted in Figure5 (a) and (b). These residual blocks are made up of two convolution layers (3 3). The number of filters are periodically doubled, and a stride of 2 is used for spatial downsampling. After each convolution layer, this network uses batch normalisation and specific skip connections. Since these deep models take activation from one layer and directly feed it to another layer, skip connections are employed to optimise them. As a result, deep networks can be trained without running into vanishing gradient issues. ResNet has a fully connected layer that outputs 1000 classes to reduce the amount of parameters.

-

**Figure5:a) Residual Block b) Architecture of ResNet**

D.BACKBONEMODELSFORSEGMENTATIONINFUNDUSIMAGES

1. FULLYCONVOLUTIONNETWORKS(FCNs)

A modified CNN network was proposed by Long et al. [13] by substituting up sampling layers for fully linked layers (shown in Figure 6). The initial layers' extracted feature maps are up-sampled to the input image's size. In comparison to CNN, a fully convolution network (FCN) is more suitable for segmentation tasks because to its ability to execute dense pixel-wise prediction.



**Figure 6: FCN Architecture**

2)U-NET

A network with symmetrical encoder and decoder structures and several skip connections from the encoding path to the decoding path was presented by Ronneberger et al. [14] and is depicted in Figure 7. While the decoder reconstructs the images for the final output, the encoder is in charge of extracting features from the input images. By directly connecting low-level feature maps from encoder to decoder, the skip configurations enable the network to generate better predictions.



**Figure7: U-Net Architecture**

VI.DEEP LEARNING RETINAL DISEASE DIAGNOSIS PROCESS

1. DIABETIC RETINOPATHY DIAGNOSIS

One of the most prevalent retinal illnesses that can result in blindness is diabetic retinopathy. One-third of diabetic people experience this problem [15]. According to a survey, DR affects around 93 million people worldwide [16]. Any diabetic patient can get DR, which disrupts the retina's blood vessels. Given the sharp increase in diabetes patients worldwide, these figures are anticipated to rise [17]. The severity of DR has been divided into five categories by the International Clinical Diabetic Retinopathy Scale (ICDRS), namely Class 0 for No DR, Class 1 for Mild DR, Class 2 for Moderate DR, Class 3 for Severe DR, and Class 4 for Proliferated DR. To create a reliable model for DR diagnosis utilizing fundus pictures, many DL models were implemented.

1. GLAUCOMA DIAGNOSIS

Another major factor contributing to permanent blindness worldwide is glaucoma [18]. Researchers have focused on creating multiple DL models to diagnose glaucoma using fundus images, much like they have with so many other retinal illnesses. The next section discusses recent advancements made in this direction. Through the extraction of OD, OC, and retinal nerve fibre layer (RNFL) characteristics, Xu et al. [19] created a DL framework for glaucoma diagnosis with a relatively small number of training samples. Pre-diagnosis classification is based on a general fundus image in their framework (global attributes). The above-mentioned biomarker segmentation is carried out in the following stage, and the ISNT and MCDR scores are computed. All the segmentation data were used to perform the final diagnosis. The cup to disc ratio (CDR) was utilized by Shanmugam et al. [20] to detect glaucoma in a fundus picture. Their approach primarily focuses on the precise segmentation of OC and OD, which was carried out through a modified U-Net. The computational load was lightened by the inclusion of adaptive convolution in their framework because it utilized fewer filters than the traditional U-Net. A random forest classifier was used to separate the glaucoma images from the healthy ones using the morphometric attributes gained from the segmentation findings. In a different study, Wang et al. [21] used a transfer learning strategy for model training and glaucoma classification using VGG-16 and Alex Net. They created two sets of ONH photos by compiling them from several publically accessible sources. One dataset was expanded using a variety of data augmentation techniques, such as random scaling, cropping, rotation, and flipping.

Nayak et al. [22] have developed a network that uses a feature optimization method based on biological phenomena, known as a real-coded genetic algorithm, to overcome issues like overfitting and the need for big datasets (RCGA). After enhanced features have been obtained using this method, glaucoma-based images are identified using a variety of classifiers. Results were best when the RCGA algorithm and SVM classifier were used. For testing and training the model, Li et al. [23] developed a CNN-ResNet architecture with 101 layers and 26,585 pictures in total. By implementing skip connections across the layers throughout the training phase, they were able to circumvent the vanishing gradient problem. A CNN-based technique was created by Hemelings et al. [24] that combines active learning techniques with transfer learning for correct diagnosis of glaucoma.

According to Juneja et al [25].'s proposed DL model, images were delivered into a CNN-based model after being subjected to specific pre-processing methods such image cropping, augmentation, and denoising (76 Layers deep). They employed a "Add layer" in each block, which combines the output of the previous block with the output of the following block, to make up for the lost data. A glaucoma diagnosis pipeline that can be used offline on mobile devices was created by Martins et al. [26]. They primarily used the U-shaped model (OD and OC segmentation) to produce beneficial morphological traits that are utilised by a different classification network (based on MobileNet-V2 as the back- bone).

Glaucoma classification was carried out by Bajwa et al. [27] in two steps. "Regions using CNN" (RCNN) is used in the initial stage for OD extraction and localization. In order to create ground truths containing the position of OD for training the RCNN, it also has a semi-automatic ground truth creation component. The second stage employs the ROI pictures (produced after OD extraction) for classification and is made up of four convolution layers and three completely linked layers. A two-task network was suggested by Kim et al. [28] that uses different CNNs for glaucoma classification and "Gradient weighted class activation mapping" to identify the most suspect glaucomatous areas in a given fundus image. The ResNet-152-M model produced the most encouraging outcomes among the other CNN variations.

Bajwa et al[29] .'s categorization of glaucoma was accomplished in two parts. OD extraction and localization are done in the early step utilizing "regions employing CNN" (RCNN). It also has a semi-automatic ground truth creation component for building ground truths with the location of OD for training the RCNN. The second stage, which consists of four convolutional layers and three fully connected layers, uses the ROI images (generated after OD extraction) for classification. Kim et al. [30] proposed a two-task network that employs various CNNs for glaucoma classification in addition to "Gradient weighted class activation mapping" to pinpoint the most likely glaucomatous regions in a given fundus image. Among the different CNN iterations, the ResNet-152-M model gave the most encouraging results.

A dense network with 201 layers was proposed by Ovreiu et al. [31] to enhance the performance of the categorization of glaucoma. This network's layers are built up from the inputs of the layers before them. In a different study, Saravanan et al. [32] showed an autoencoder architecture for glaucoma diagnosis and AVP recognition; they paid particular attention to lowering classification errors by the implementation of multi-modal learning. The effectiveness of three pre-trained CNN-based models for early glaucoma detection was compared by Shoukat et al. [33]. The RIM-ONE, G1020, and REFUGE datasets were used to run the test. On the G1020 dataset, pretrained EfficientNet-B7 produced the best results. On a private dataset made up of 643 fundus photos, Islam et alstudy.'s [34] examined the performances of various DL models like DenseNet, MobileNet, EfficientNet, and Google Net.

Ovreiu et al. [31] suggested a thick network with 201 layers to improve the effectiveness of the categorization of glaucoma. The inputs from the layers preceding them are used to construct the layers of this network. In a different work, Saravanan et al. [32] demonstrated an autoencoder architecture for diagnosing glaucoma and identifying AVP. They gave particular focus to reducing classification errors by incorporating multi-modal learning. Shoukat et al. [33] examined the efficacy of three pre-trained CNN-based models for early glaucoma detection. The test was carried out using the datasets from RIM-ONE, G1020, and REFUGE. EfficientNet-B7 that had been pretrained performed the best on the G1020 dataset.The performance of techiques by different authors is defined the table[3].

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **References** | **Dataset** | **ACC** | **SE** | **SP** | **AUC** | **F1** |
| [19] | Private-Tongren |  | 0.961 | 0.939 | 0.981 |  |
| [20] | DRIONS-DB,HRF | 0.943 | 0.907 | 0.979 | 0.991 |  |
| [21] | Private |  | 0.940 | 0.860 |  |  |
| [22] | Private-Kasthurba Medical College,Manipal | 0.980 | 0.974 | 0.988 |  | 0.983 |
| [23] | Private | 0.953 | 0.96 | 0.939 | 0.944 |  |
| [24] | Private |  | 0.980 | 0.910 | 0.995 |  |
| [25] | DRISHTI-GS,RIM-ONE | 0.870 | 0.85 |  | 0.93 |  |
| [26] | Origa,Drishti,iChalenge,RIM-ONE,RIGA | 0.870 | 0.85 |  | 0.93 |  |
| [27] | ORIGA,HRF,OCF&CFI |  | 0.717 |  | 0.874 |  |
| [28] | SamsungMedical Centre,Soeul | 0.96 | 0.95 |  | 0.99 | 0.97 |
| [33] | G1020 | 0.992 | 0.98 | 0.97 |  |  |
| [34] | G1020 | 0.98 | 0.951 | 0.94 |  |  |

**Table3:Glaucoma Diagnosis Performance Comparision**

1)OC/OD SEGMENTATION

Optic Disc (OD) and Optic Cup are two additional significant retinal indicators utilised for the diagnosis of glaucoma (OC). The vertical cup diameter and the vertical disc diameter are used to compute the cup-to-disc ratio. As a result, correct OD/OC segmentation has become essential for glaucoma diagnosis, and much study has been done in this area. Along with experimental findings a few recent studies on DL-based OD/OC

 segmentation are discussed in the sections that follow

C) AMD DIAGNOSIS

One of the main causes of blindness in the senior population is age-related macular degeneration (AMD) [35]. AMD typically damages the retina's macular area. According to a study, there would be 288 million AMD sufferers worldwide by the year 2040, up from an estimated 196 million people in 2020 [36]. In the section that follows, we go over various DL-based approaches used recently for the automatic diagnosis of AMD. In order to distinguish between neovascular AMD (NAMD) and polypoidal choroidal vasculopathy (PCV), Chou et al. [37] combined a fundus image-based DL model with biomarkers obtained from optical coherence tomography (OCT) using a stacking technique. Multiple Correspondence Analysis (MCA), a brand-new technique, was developed to transform OCT biomarkers into continuous main components. Fundus pictures were trained and validated using EfficientNet-B3. For precise predictions on new input photos, the ensemble stacking technique produces the optimal combination from the aforementioned two paths. Yan et al. [38] presented a framework for predicting late AMD progression using a modified Deep CNN. Apart from fundus images, their model also considers genotypes for improving accuracy. Chou et al. [37] used a stacking strategy to integrate a fundus image-based DL model with biomarkers collected from optical coherence tomography (OCT) in order to differentiate between neovascular AMD (NAMD) and polypoidal choroidal vasculopathy (PCV). OCT biomarkers were converted into continuous major components using a novel technique called Multiple Correspondence Analysis (MCA). EfficientNet-B3 was used to train and validate fundus images. The ensemble stacking method creates the ideal mix from the aforementioned two approaches to exact predictions on new input images. Using a modified Deep CNN, Yan et al. [38] proposed a methodology for forecasting the evolution of late AMD. Their model takes genotypes into account in addition to fundus photos to increase accuracy.

In order to classify AMD and PCV, Xu et al. [39] suggested a dual deep CNN model that makes use of fundus and OCT picture pairs. In order to use transfer learning, weights from ResNet-50 were first applied to two independent models that separately took input from fundus images and OCT images. The weights were transferred to the corresponding convolutional blocks after being improved on new data. In the end, an FC layer was created and input pairs were divided into Wet AMD, Dry AMD, PCV, and nAMD categories. Another study based on drusen segmentation for AMD detection was put out by Pham et al. [40], who attempted to address the issue of data imbalance because the number of non-drusen pixels was significantly higher than the number of drusen pixels. the use of images in many modalities.

Xu et al. [39] proposed a dual deep CNN model using fundus and OCT image pairs to classify AMD and PCV. Weights from ResNet-50 were first applied to two distinct models that separately incorporated input from fundus images and OCT images in order to use transfer learning. The weights were enhanced using fresh input and then assigned to the corresponding convolutional blocks. In the end, input pairs were separated into Wet AMD, Dry AMD, PCV, and nAMD categories and an FC layer was built. Pham et al. [40] published another study using drusen segmentation for AMD detection.

D) CATARACT DIAGNOSIS

If a retinal disease is identified and treated promptly, cataract is one of the major retinal diseases, can cause irreversible vision loss [41]. According to a recent study, cataracts are responsible for roughly 33.6 million cases of blindness, or 45% of all occurrences of blindness worldwide [42]. There have recently been numerous attempts to automatically detect cataracts from fundus images. Below is a discussion of recent advancements in this area and a comparison of the performance of several DL models.

The activation function and loss function of their CNN-based architecture were modified by Junayed et al. [43] to reduce training parameters and computational load when training the model for cataract detection. Additionally, they assessed the detection accuracy of three alternative models that employ 3, 4, and 5 CNN blocks, depending on the model. The four-block model delivered the best outcomes without any overfitting. Combining CNN and recurrent neural network, Imran et al. [44] suggested a cataract categorization model (severe, moderate, mild, and normal) (RNN). Each fundus picture from the dataset was separated into 12 patches after pre-processing, and each patch was then processed through pre-trained CNN models (GoogleNet, AlexNet, VGGNet, and ResNet) for feature extraction.

E) ROP DIAGNOSIS

Infants' fundus vasculature is most impacted by the retinal illness known as retinal retinopathy of prematurity (ROP). Children who have this illness may experience serious outcomes like retinal detachment and total blindness due to neo-vascularization. Identification of an early symptom known as well as disease-causing morphological abnormalities to preterm infants' retinal blood vessels is crucial for prompt treatment

A framework for recognising ROP by spotting additional disease in baby fundus pictures was introduced by Ramachandran et al. [45]. The network creates bounding boxes around the twisted vessels in their semi-supervised approach, and the number of these boxes identifies the presence of illness in the retinal image. This is accomplished by using a fully convolution neural network for recognising the twisted vessels, which was influenced by YOLO architecture. The model is first trained using manually labelled images to create images with bounding boxes (pseudo labelled images), and then it is retrained using both manually labelled and pseudo labelled images. The model is then used to forecast ROP. to establish a diagnosis of ROP and a system for supported medical follow-up.

VI RESEARCH DIRECTIONS

As it was covered in the previous sections, testing and assessing different network topologies for the diagnosis of retinal diseases has advanced impressively using DL approaches. Future research has a lot of potential and uncharted territory, nevertheless.

***Weekly supervised Learning Models:***

Eventhoughmanyfundusimagedatasetsexistinthepublicdomain,whencomparedtonaturalimagedatasetslikeImageNetwhich has nearly 14 million images, the availability oflabelled fundus images is quite limited. The availablefundusdatasetsarealsodiverseintermsoftheirgroundtruth labelling. Although there are other techniques likeimage synthesis, that are parallelly investigated, whichcan generate artificial fundus images, researchers canexplore weakly supervised learning models for trainingon original fundus images that have different groundtruth labeling. Through weakly supervised training techniques robust model performances can be achieved forretinaldiseasediagnosisevenwithdatasetsthatarepartiallylabelledorinaccuratelylabeled.

***Fundus Image Synthesis***

RecentpopularityofGenerativeAdversarialNetworks(GANs)hasshownpotentialin generating synthetic fundus images which can beused to augment the training dataset. This can effectively eliminate the lack of good quality labeled data andimprove prediction performance. Although some of therecentresearchshowedthesynthesisofimagesforDR,glaucoma,andAMD,thefieldisstillrelativelynewandpresentsamplescopeforfutureresearch.

***Light Weight Network Design***

The majority of DL models created for diagnosing retinal diseases work well but use up a lot of computational resources. Implementing such models on portable edge devices faces this significant obstacle. The creation of new, lightweight models to cut down on computing requirements while retaining performance is another open research area in this area.

***Improving Generalization***

It was shown that because different datasets had variable picture acquisition settings, the performance of DL models differed amongst them, with certain models succeeding on some datasets while failing on others. By investigating alternative domain adaptation approaches, researchers can concentrate on enhancing the performance of models for generalisation. The basic objective of these methods is to reduce the distribution gap between the source and target data domains. Existing adaptation techniques include moment matching, which minimises distribution differences at the feature space level, and adversarial learning, which aligns the source and target domains. Given the difficulties in obtaining retinal fundus pictures, the field of domain adaptation offers plenty of potential opportunities for researchers to improve model generalisation.

***Implementing Federated Learning***

Most hospitals and other research institutions are reluctant to share fundus photos with others because of various data protection restrictions. This compounds the issue of data scarcity and limits model training to only publicly available datasets, denying them access to training on rich and varied private fundus data available at the hospitals. Models can be trained locally on private data using techniques like federated learning, and then the learnt weights are applied to a global model.

***Multiple Disease Diagnosis***

The simultaneous detection of many retinal disorders with DL is another intriguing study area. Clinicians may find it useful to identify patients who have multiple retinal diseases. Even though studies in this field have been done, such as simultaneous "DME and DR diagnosis," simultaneous "AMD, DR, and glaucoma diagnosis," etc., this area is still not very well researched.

***Smart phone based Retinal Disease Diagnosis***

Most recent research in this area makes use of fundus pictures obtained using high-resolution fundoscopy. Researchers have plenty of room to create models that can learn from fundus photos taken with smartphones. This will assist in creating a facility for remote eye examination.

***Generating Evidence Maps***

Getting the professional doctors' clearance for the AI-based model is one of the main issues with DL implementation for retinal disease detection. Only a small amount of study has been done to increase the predictability of the results. Making evidence maps for the predictions made by the DL model and displaying or emphasising the key fundus image regions that the deep network relied on to make the final judgement could be one solution to this problem. Although certain methods have recently made headway in this area, there is still much that can be learned about how to increase the precision of these evidence maps. For instance, accurate lesion segmentation and simultaneous grading of DR are possible to provide high-quality evidence maps since DR diagnosis rely on the discovery of numerous lesions and markers on fundus images.

VII.CONCLUSION

Given the imbalance between the number of patients and medical professionals, there is an urgent need for automated solutions for identifying eye-related disorders. In terms of medical image analysis, a colour fundus picture, which presents a wide range of eye-related disorders in image format, has greatly expanded the field of study. For automatic illness diagnosis, a wide variety of DL models are being applied and tested. Salient features from a given fundus image can now be extracted using sophisticated image processing techniques. These days, it is possible to diagnose diseases at an early stage using lesions like microaneurysms, exudates, haemorrhages, etc. that make up a substantially smaller number of pixels in a fundus image. For understanding the most recent DL techniques in the process of diagnosing ocular diseases, this review provided a process-based approach.

A compilation of all publicly accessible fundus picture datasets is supplied together with a ground truth description because the performance of a DL model heavily depends on the training dataset. It has been noted that numerous databases, like IDRiD, Messidor, DRIVE, etc., provide excellent fundus photos that were taken in a controlled setting. On other datasets, the models that were trained on these datasets might not perform well. On the other side, photos taken in various environmental situations can be found in databases like Kaggle and Eye-PACS, among others.

These might not be appropriate for effective however, by simulating a real-world situation, they can direct the model behaviour toward the useful side. A robust model that may be used in clinical settings may be developed with the aid of a well-balanced collection of datasets.

The majority of research has demonstrated that using image pre-processing techniques, such as contrast enhancement, colour space transformation, picture augmentation, filtering, etc., can improve the DL model's ability to extract disease-relevant features during the training phase.

The work that has been published recently has built solutions for disease diagnostics using a variety of skeletal models. For classification tasks, networks like Basic CNN, VGG, ResNet, Inception, etc. are used, whereas networks like U-Net, FCNs, Mask RCNN, Seg-Net, etc. are used for segmentation tasks. These backbone models have typically only been used as the foundation for studies. To enhance the model's performance and deliver an accurate diagnosis, other learning paradigms, including group learning, transfer learning, multitask learning, active learning, etc., have also been investigated. Diabetic retinopathy is one of the retinal diseases that has received the most attention in terms of research and clinical application. In addition to disease classification, the current core research for DR is focused on producing interpretable heatmaps.

The retinal illnesses that are the subject of this review are extremely important since delaying treatment could result in total vision loss. In the past few years, there has been a noticeable increase in interest in using DL to diagnose retinal diseases. In several instances, DL models' performance has surpassed that of professional humans. DL systems must develop and be incorporated into clinical practise, therefore the future of efficient and successful patient care is still very uncertain.

REFERENCES

1. D.ChawlaandA.Deorari,‘‘Retinopathyofprematurityprevention,screeningandtreatmentprogrammes:ProgressinIndia,’’in*Seminarsin Perinatology*, vol. 43. Amsterdam, The Netherlands: Elsevier, 2019,pp. 344–347.
2. M.M.Fraz,P.Remagnino,A.Hoppe,B.Uyyanonvara,A.R.Rudnicka,

C. G. Owen, and S. A. Barman, ‘‘Blood vessel segmentation methodolo-giesinretinalimages—Asurvey,’’*Comput.MethodsProgramsBiomed.*,vol. 108, no. 1,pp. 407–433, Oct.2012.

1. Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, and M. S. Lew, ‘‘Deeplearningforvisualunderstanding:Areview,’’*Neurocomputing*,vol.187,pp. 27–48, Apr. 2016.
2. M. D. Abràmoff, M. K. Garvin, and M. Sonka, ‘‘Retinal imaging andimageanalysis,’’ *IEEE Rev. Biomed. Eng.*,vol. 3, pp. 169–208, 2010.
3. S. Sengupta, A. Singh, H.A. Leopold, T. Gulati,and V. Lakshminarayanan,‘‘Ophthalmicdiagnosisusingdeeplearningwithfundusimages—Acriticalreview,’’*ArtifIntell.Med.*,vol.102,Jan. 2020, Art. no. 101758.
4. M. Badar, M. Haris, and A. Fatima, ‘‘Application of deep learning forretinal image analysis: A review,’’ *Comput. Sci. Rev.*, vol. 35, Feb. 2020,Art. no. 100203
5. T.Li,W.Bo,C.Hu,H.Kang,H.Liu,K.Wang,andH.Fu,‘‘Applicationsofdeeplearninginfundusimages:Areview,’’*Med.ImageAnal.*,vol.69,Apr. 2021, Art. no. 101971.
6. W.L.Alyoubi,M.F.Abulkhair,andW.M.Shalash,‘‘Diabeticretinopa-thyfundusimageclassificationandlesionslocalizationsystemusingdeeplearning,’’ *Sensors*, vol. 21, no. 11, p. 3704, May 2021.
7. D. Mirzania, A. C. Thompson, and K. W. Muir, ‘‘Applications of deeplearning in detection of glaucoma: A systematic review,’’ *Eur. J. Oph-thalmol.*, vol.31, no.4, pp.1618–1642,Jul. 2021.
8. Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, ‘‘Gradient-based learn-ingappliedtodocumentrecognition,’’*Proc.IEEE*,vol.86,no.11,pp. 2278–2324, Nov. 1998.
9. K.SimonyanandA.Zisserman,‘‘Verydeepconvolutionalnetworksforlarge-scaleimagerecognition,’’2014,*arXiv:1409.1556*.[Online].Available:https://doi.org/10.48550/arXiv.1409.1556,doi:[10.48550/arXiv.1409.1556.](http://dx.doi.org/10.48550/arXiv.1409.1556)
10. K.He,X.Zhang,S.Ren,andJ.Sun,‘‘Deepresiduallearningforimage recognition,’’ in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.(CVPR)*, Jun. 2016, pp.770–778.
11. J.Long,E.Shelhamer,andT.Darrell,‘‘Fullyconvolutionalnetworksfor semantic segmentation,’’ in *Proc. IEEE Conf. Comput. Vis. PatternRecognit. (CVPR)*,Jun. 2015,pp. 3431–3440.
12. O.Ronneberger,P.Fischer,andT.Brox,‘‘U-Net:Convolutionalnetworksfor biomedical image segmentation,’’ in *Proc. Int. Conf. Med. ImageComput. Comput.-Assist. Intervent.* Cham, Switzerland: Springer, 2015,pp. 234–241.
13. J.Jenkins,M.V.Joglekar,A.A.Hardikar,A.C.Keech,D.N. O’Neal,and A. S. Januszewski, ‘‘Biomarkers in diabetic retinopathy,’’ *Rev. Dia-betic Stud.*, vol.12, nos. 1–2,p. 159, 2015.
14. J.W.Yau,S.L.Rogers,R.Kawasaki,E.L.Lamoureux,J.W.Kowalski,T. Bek, S.-J. Chen, J. M. Dekker, A. Fletcher, and J. Grauslund, ‘‘Globalprevalenceandmajorriskfactorsofdiabeticretinopathy,’’*DiabetesCare*,vol. 35, no. 3, pp.556–564, 2012.
15. W. Yang, J. Lu, J. Weng, W. Jia, L. Ji, J. Xiao, Z. Shan, J. Liu, H. Tian,and Q. Ji, ‘‘Prevalence of diabetes among men and women in China,’’*New EnglandJ.Med.*,vol.362,no.12,pp.1090–1101,2010.
16. R.R.A.Bourne,G.A.Stevens,R.A.White, J.L.Smith,S.R.Flaxman,H.Price,J.B.Jonas,J.Keeffe,J.Leasher,K.Naidoo,K.Pesudovs,S. Resnikoff,andH.R.Taylor,‘‘Causesofvisionlossworldwide,1990–2010:Asystematicanalysis,’’*LancetGlobalHealth*,vol.1,no.6,pp. e339–e349, Dec. 2013.
17. Y.Xu, M. Hu, H. Liu, H. Yang, H. Wang, S.Lu, T. Liang, X. Li, M. Xu,L.Li,H.Li,X.Ji,Z.Wang,L.Li,R.N.Weinreb,andN.Wang,‘‘Ahier-archical deep learning approach with transparency and interpretabilitybasedonsmallsamplesforglaucomadiagnosis,’’*NPJDigit.Med.*,vol.4,no. 1, pp. 1–11, Dec.2021.
18. S. P., R. J., and P. R., ‘‘An automatic recognition of glaucoma in fundusimages using deep learning and random forest classifier,’’ *Appl. SoftComput.*, vol. 109,Sep. 2021,Art. no.107512.
19. P.Wang,M.Yuan,Y.He,andJ.Sun,‘‘3Daugmentedfundusimagesforidentifyingglaucomaviatransferredconvolutionalneuralnetworks,’’*Int.Ophthalmol.*, vol.41,no.6, pp.2065–2072,Jun.2021.
20. D. R. Nayak, D. Das, B. Majhi, S. V. Bhandary, and U. R. Acharya,‘‘ECNet:Anevolutionaryconvolutionalnetworkforautomatedglaucomadetectionusingfundusimages,’’*Biomed.SignalProcess.Control*,vol.67,May 2021, Art. no. 102559.
21. F. Li, L.Yan, Y. Wang,J. Shi, H.Chen, X. Zhang,M. Jiang, Z.Wu, andK. Zhou, ‘‘Deep learning-based automated detection of glaucomatousoptic neuropathy on color fundus photographs,’’ *Graefe’s Arch. Clin.Experim.Ophthalmol.*,vol.258,no.4,pp.851–867,Apr.2020.
22. R.Hemelings,B.Elen,J.Barbosa-Breda,S.Lemmens,M.Meire,S.Pourjavan,E.Vandewalle,S.VanDeVeire,M.B.Blaschko,P. De Boever, and I. Stalmans, ‘‘Accurate prediction of glaucoma fromcolourfundusimageswithaconvolutionalneuralnetworkthatreliesonactiveandtransferlearning,’’*ActaOphthalmol.*,vol.98,no.1,pp. e94–e100, Feb. 2020.
23. M.Juneja,N.Thakur,S.Thakur,A.Uniyal,A.Wani,andP.Jindal,‘‘GC-NET for classification of glaucoma in the retinal fundus image,’’ *Mach.Vis. Appl.*, vol. 31, no. 5,pp. 1–18, Jul. 2020.
24. J. Martins, J. S. Cardoso, and F. Soares, ‘‘Offline computer-aided diag-nosis for glaucoma detection using fundus images targeted at mobiledevices,’’ *Comput. Methods Programs Biomed.*, vol. 192, Aug. 2020,Art. no. 105341.
25. H.Liu,L.Li,I.M.Wormstone,C.Qiao,C.Zhang,P.Liu,S.Li,H. Wang, D. Mou, and R. Pang, ‘‘Development and validation of a deeplearning system to detect glaucomatous optic neuropathy using fundusphotographs,’’*JAMAOphthalmol.*,vol.137,no.12,pp. 1353–1360,2019.
26. M.N.Bajwa,M.I.Malik,S.A.Siddiqui,A.Dengel,F.Shafait,W.Neumeier,andS.Ahmed,‘‘Two-stageframeworkforopticdisclocal-ization and glaucoma classification in retinal fundus images using deeplearning,’’ *BMC Med. Informat. Decis. Making*, vol. 19, no. 1, pp. 1–16,Dec. 2019.
27. M.Kim,J.C.Han,S.H.Hyun,O.Janssens,S.VanHoecke,C.Kee,andW.De Neve,‘‘Medinoid:Computer-aideddiagnosisandlocalizationofglaucomausingdeeplearning,’’*Appl.Sci.*,vol.9,no.15,p.3064,Jul. 2019.
28. L. K. Singh, H. Garg, and M. Khanna, ‘‘Deep learning system applica-bility for rapid glaucoma prediction from fundus images across variousdatasets,’’*EvolvingSyst.*,pp. 1–30,Feb.2022.[Online].Available:https://doi.org/10.1007/s12530-022-09426-4, doi: [10.1007/s12530-022-09426-4.](http://dx.doi.org/10.1007/s12530-022-09426-4)
29. S. Ovreiu, E.-A. Paraschiv, and E. Ovreiu, ‘‘Deep learning & digitalfundus images: Glaucoma detection using DenseNet,’’ in *Proc. 13th Int.Conf.Electron.,Comput.Artif.Intell.(ECAI)*,Jul.2021,pp.1–4.
30. V. Saravanan, R. D. J. Samuel, S. Krishnamoorthy, and A. Manickam,‘‘Deeplearningassistedconvolutionalauto-encodersframeworkforglaucomadetectionandanteriorvisualpathwayrecognitionfromreti-nal fundus images,’’ *J. Ambient Intell. Humanized Comput.*, pp. 1–11,Jan.2022.[Online].Available:https://doi.org/10.1007/s12652-021-02928-0, doi:[10.1007/s12652-021-02928-0.](http://dx.doi.org/10.1007/s12652-021-02928-0)
31. A.Shoukat,S.Akbar,S.A.E.Hassan,A.Rehman,andN.Ayesha,‘‘Anautomateddeeplearningapproachtodiagnoseglaucomausingretinal fundus images,’’ in *Proc. Int. Conf. Frontiers Inf. Technol. (FIT)*,Dec. 2021, pp. 120–125.
32. M.TanvirIslam,S.T.Mashfu,A.Faisal,S.ChowdhurySiam,I.TahmidNaheen, and R. Khan, ‘‘Deep learning-based glaucoma detection withcroppedopticcupanddiscandbloodvesselsegmentation,’’*IEEEAccess*,vol. 10, pp. 2828–2841, 2022.
33. A. Shoukat, S. Akbar, and K. Safdar, ‘‘A deep learning-based automaticmethod for early detection of the glaucoma using fundus images,’’ in*Proc.Int. Conf. Innov. Comput. (ICIC)*, Nov. 2021, pp. 1–6.
34. Q.Yan,D.E.Weeks,H.Xin,A.Swaroop,E.Y.Chew,H.Huang,Y. Ding, and W. Chen, ‘‘Deep-learning-based prediction of late age-related macular degeneration progression,’’ *Nature Mach. Intell.*, vol. 2,no. 2, pp. 141–150, Feb. 2020.
35. Z.Xu,W.Wang,J.Yang,J.Zhao,D.Ding,F.He,D.Chen,Z.Yang,X.Li,W.Yu,andY.Chen,‘‘Automateddiagnosesofage-relatedmaculardegenerationandpolypoidalchoroidalvasculopathyusingbi-modaldeepconvolutionalneuralnetworks,’’*Brit.J.Ophthalmol.*,vol.105,no.4,pp. 561–566, Apr. 2021.
36. Q.T.M.Pham,S.Ahn,S.J.Song,andJ.Shin,‘‘Automaticdrusensegmentationforage-relatedmaculardegenerationinfundusimagesusing deep learning,’’*Electronics*, vol. 9,no. 10,p. 1617, Oct.2020.
37. R.Isaacs,J.Ram,andD.Apple,‘‘Cataractblindnessinthedeveloping world: Is there a solution?’’ *J. Agromedicine*, vol. 9, no. 2, pp. 207–220,2004.
38. S.Pawar,‘‘Causesofblindnessandvisionimpairmentin2020and trendsover30years,andprevalenceofavoidableblindnessinrelationto VISION 2020: The right to sight: An analysis for the global burden ofdiseasestudy,’’ *SSRNElectron. J.*, vol.9, no.2, pp. e144–e160,2021.
39. M.S.Junayed,M.B.Islam,A.Sadeghzadeh,andS.Rahman,‘‘Cataract- Net: An automated cataract detection system using deep learning forfundus images,’’*IEEE Access*,vol.9, pp.128799–128808,2021.
40. Imran,J.Li,Y.Pei,F.Akhtar,T.Mahmood,andL.Zhang,‘‘Fun-dus image-based cataract classification using a hybrid convolutional andrecurrent neural network,’’ *Vis. Comput.*, vol. 37, no. 8, pp. 2407–2417,Aug. 2021.
41. S.Ramachandran,P.Niyas,A.Vinekar,andR.John,‘‘Adeeplearn-ingframeworkforthedetectionofplusdiseaseinretinalfundusimages of preterm infants,’’ *Biocybernetics Biomed. Eng.*, vol. 41, no. 2,pp. 362–375, Apr. 2021.