

Received April 28, 2022, accepted May 19, 2022, date of publication May 27, 2022, date of current version June 6, 2022.

Digital Object Identifier 10.1109/ACCESS.2022.3178372

A Comprehensive Review of Deep Learning Strategies in Retinal Disease Diagnosis Using Fundus Images

BALLA GOUTAM¹, MOHAMMAD FARUKH HASHMI¹, (Senior Member, IEEE), ZONG WOO GEEM², (Senior Member, IEEE), AND NEERAJ DHANRAJ BOKDE^{3,4}

¹Department of Electronics and Communication Engineering, National Institute of Technology, Warangal 506004, India

²College of IT Convergence, Gachon University, Seongnam 13120, South Korea

³Department of Civil and Architectural Engineering, Aarhus University, 8000 Aarhus, Denmark

⁴Center for Quantitative Genetics and Genomics, Aarhus University, 8000 Aarhus, Denmark

Corresponding author: Zong Woo Geem (geem@gachon.ac.kr)

This work was supported by the National Research Foundation of Korea (NRF) Grant funded by the Korean Government through the Ministry of Science and ICT (MSIT) under Grant 2020R1A2C1A01011131.

ABSTRACT In recent years, there has been an unprecedented growth in computer vision and deep learning implementation owing to the exponential rise of computation infrastructure. The same was also reflected in retinal image analysis and successful artificial intelligence models were developed for various retinal disease diagnoses using a wide variety of visual markers obtained from eye fundus images. This article presents a comprehensive study of different deep learning strategies employed in recent times for the diagnosis of five major eye diseases, i.e., Diabetic retinopathy, Glaucoma, age-related macular degeneration, Cataract, and Retinopathy of prematurity. This article is organized according to the deep learning implementation process pipeline, where commonly used datasets, evaluation metrics, image pre-processing techniques, and deep learning backbone models are first illustrated followed by an extensive review of different strategies for each of the five mentioned retinal diseases is presented. Finally, this article summarizes eight major research directions available in the field of retinal disease diagnosis and outlines key challenges and future scope for the present research community.

INDEX TERMS Computer vision, deep learning, fundus image, retinal disease diagnosis, artificial intelligence, diabetic retinopathy, glaucoma, AMD, cataract, ROP.

I. INTRODUCTION

To investigate the human eye, many imaging modalities have been developed over the years, out of which, 'Fundus Imaging' has gained in popularity due to its non-invasive and cost-effective nature. Fundus photography involves capturing the projection of the fundus (the rear portion of an eye) onto a two-dimensional plane using a monocular camera. Several ocular structures and biomarkers including various abnormalities can be identified from a captured 2D fundus image (Figure 1). Many of these visual markers play an important role in identifying retinal diseases.

The tiny red dot-like structures, known as Microaneurysms (MAs), are normally developed due to a lack of oxygen supply and bulging capillaries. Sometimes, when the supply

completely shuts down due to certain arteriolar blockages, white soft patches are formed which are indicated as Soft Exudates (SEs). Retinal vessels sometimes burst because of the built-up pressure in arterioles and manifest as dark red patches, known as Hemorrhages. Hard Exudates (HEs) are formed when proteins and fat leak from abnormal vessel walls and appear as hard yellow waxy structures. Examining the presence of these lesions along with other retinal biomarkers like an optic disc (OD), optic cup (OC), macular region, fovea, and blood vessels can provide valuable insights into some of the major retinal diseases and aid in their diagnosis.

Diabetic retinopathy (DR), Glaucoma, age-related macular degeneration (AMD), Diabetic macular edema (DME), retinopathy of prematurity (ROP), and Cataract are some of the major eye diseases that can cause blindness if not treated appropriately. The screening process for such retinal diseases generally requires expert attention and substantial skill [1].

The associate editor coordinating the review of this manuscript and approving it for publication was Gustavo Olague⁵.

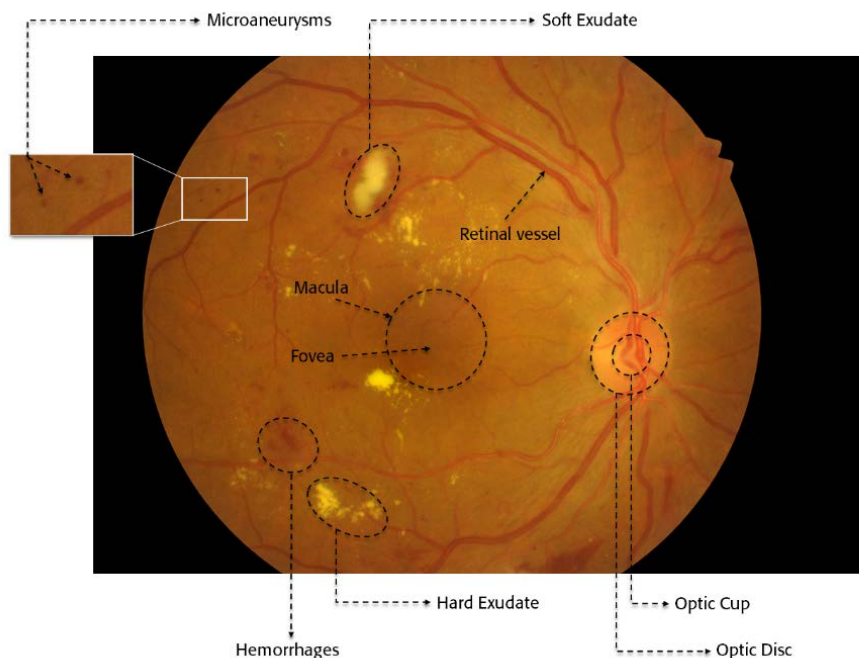


FIGURE 1. A fundus image with pathologies.

In densely populated countries like India, there is a severe lack of trained ophthalmologists, who can perform such time-consuming tasks [2]. Due to recent exponential growth in digital processors and data-driven technologies, artificial intelligence (AI) based medical screening systems are becoming more prevalent and offer feasible and cost-effective solutions for automatic diagnosis of retinal diseases [3]. In particular, computer vision and deep learning (DL) techniques have shown immense growth and promise in fundus image analysis.

DL tasks in retinal disease diagnosis mainly fall into two categories - classification and segmentation tasks. The classification task refers to a direct classification of input images into various disease categories. Similarly, identifying important biomarkers and crucial lesions through segmentation tasks from a given fundus image of the patient can reveal many details about the nature and type of retinal diseases. Many DL architectures have been developed and tested for such tasks are well illustrated in [4]. An overall DL framework for retinal disease diagnosis is shown in Figure 2.

The list of abbreviations used in this article are tabulated in Table 1.

A. FEATURES OF THE PROPOSED REVIEW

The proposed review focuses mainly on providing an in-depth review of various DL strategies recently implemented for retinal disease diagnosis using fundus images. This study also intends to outline possible future directions for new researchers interested in AI-based retinal disease diagnosis.

- Moreover, contrary to recently published review articles on this topic [5]–[9], this article takes a DL process pipeline approach to retinal disease diagnosis and surveys recent articles on a diagnosis of five major eye diseases, i.e., Diabetic retinopathy, Glaucoma, Age-related macular degeneration, cataract and Retinopathy of prematurity.
- It also comprehensively outlines all the datasets available for the above-mentioned diseases along with their ground truth descriptions.
- It provides knowledge about widely used image pre-processing techniques, evaluation metrics, and commonly used DL backbone models for retinal disease diagnosis tasks.
- It also contains an extensive literature study on DL implementation of five major retinal diseases along with tabulating their comparative performances.
- It also discusses various research directions currently available in this field.

The rest of this article is organized as follows. Section 2 covers datasets and evaluation metrics for retinal disease diagnosis. Commonly used fundus image pre-processing techniques are illustrated in Section 3. Most widely used DL strategies along with specific backbone models for fundus image-based classification tasks and segmentation tasks are outlined in Section 4. Section 5 presents a literature review along with comparisons of the performance of recent research work on DR, Glaucoma, AMD, cataract, and ROP diagnoses. Finally, Future research

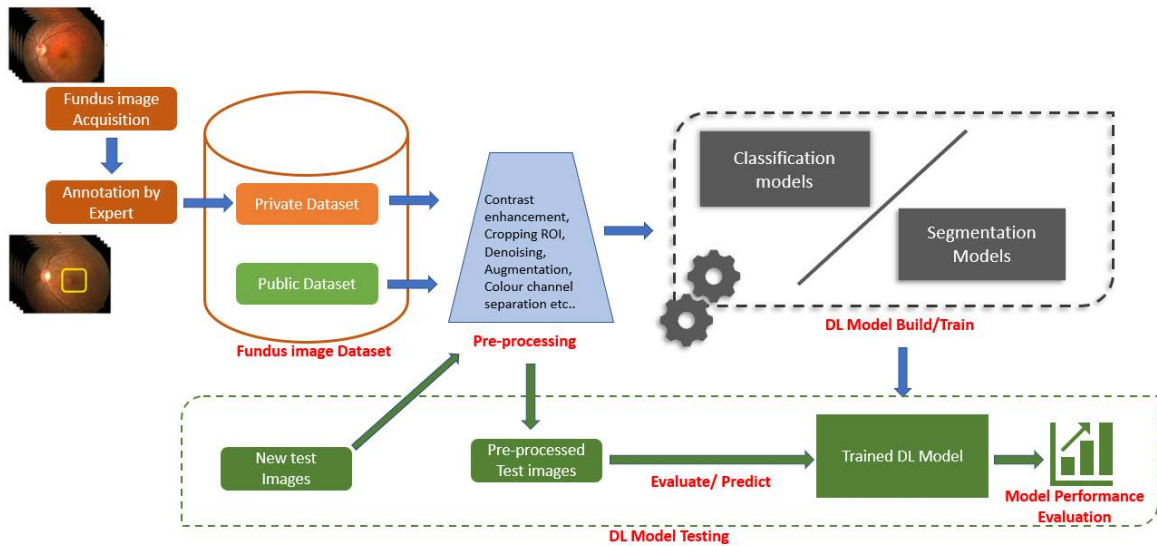


FIGURE 2. Overall DL framework for retinal disease diagnosis.

directions and conclusions are presented in Sections 6 and 7, respectively.

II. DATASETS AND EVALUATION METRICS FOR RETINAL DISEASE DIAGNOSIS

Fundus photography is the process of retrieving a two-dimensional image of the 3D ocular retinal fundus, using reflected light projected onto an image plane. Table 2 shows an overview of widely used fundus image datasets that are utilized in DL-based diagnosis for the above-mentioned retinal diseases. The table lists out datasets, the number of fundus images it contains, image size, format, and its ground truth description. Disease diagnosis task utilization of these datasets is presented through color-coding. For ease of comprehension and comparative study, all the datasets are presented in one table.

A. MODEL PERFORMANCE EVALUATION METRICS

Several performance evaluation metrics are used in evaluating the DL model in the retinal disease diagnosis task. Table 3 lists the most commonly used metrics along with their description. During the literature review in this paper, most of these metrics will be used for comparing various DL architectures. It must be noted that there are slight variances in the performance evaluation metrics for different retinal diseases. Metrics like ACC, PR, SE, SP, AUC and F1 are commonly used as performance indicators for both classifications as well as segmentation tasks, whereas additional metrics like IoU and DSC are used to indicate segmentation performance involved in retinal disease diagnosis.

III. PRE-PROCESSING TECHNIQUES

To improve the training process and build robust prediction models, fundus images are generally pre-processed before

the training phase. This is done to compensate for the noise-induced due to the variety of image capturing hardware used in varied illumination settings during the imaging. Considering the complexity of the retinal structure, many important biomarkers and lesions may not be identified due to the poor quality of the images, as shown in Figure 3. Apart from removing unwanted noise, pre-processing techniques are also used to enhance the fundus image features before DL model implementation. Some of the widely used pre-processing techniques on color fundus images for retinal disease diagnosis are presented in Table 4.

IV. DEEP LEARNING CONCEPTS

Deep learning (DL) is a sub-class of artificial intelligence methods that are based on artificial neural networks (learning methods inspired by the biological structure of the human brain). In the DL process, the latent and intrinsic relation of the input data is learned automatically through mathematical representations. Contrary to traditional machine learning (ML) methods, DL can execute with far less human guidance, as they directly extract useful features from the data without depending on hand-crafted features. This makes DL suitable for medical image analysis, where the features can be learned automatically from complex visual information. In the following section, we discuss the architectures of some of the frequently used backbone models, especially for classification and segmentation tasks in retinal disease diagnosis.

A. BACKBONE MODELS FOR CLASSIFICATION TASK

1) CONVOLUTION NEURAL NETWORKS (CNNs)

One of the most widely used DL architectures for efficient training through multiple layers is a convolution neural network (CNN) [45]. Figure 4 depicts the general architecture

TABLE 1. A list of abbreviations.











Abbreviations	
2D	Two Dimensional
ACC	Accuracy
AI	Artificial Intelligence
AMD	Age Related Macular Degeneration
AREDS	Age-Related Eye Disease Studies
AUC	Area Under the Curve
BPF	Band Pass Filter
CASA	Channel and Spatial Attention
CCS	Cross-Connection Subnetwork
CDR	Cup To Disc Ratio
CE	Contrast Enhancement
CECED	Contrast Enhanced Canny Edge Detection
CFE	Combined Feature Extraction
CGSA	Chinese Glaucoma Study Alliance
CLAHE	Contrast Limited Adaptive Histogram Equalization
CNN	Convolution Neural Network
DCNN	Deep Convolution Neural Network
DL	Deep Learning
DM	Dice Metric
DME	Diabetic macular edema
DSC	Depth-Wise Separable Convolution
DSC	Dice Similarity Coefficient
FCN	Fully convolution Network
FDS	Feature Detection Subnetwork
FN	False Negative
FP	False Positive
FPR	False Positive Rate
GANs	Generative Adversarial Networks
GLCM	Gray Level Co-occurrence Matrix
GMM	Gaussian Mixture Model
HEs	Hard Exudates
HSI	Hue, Saturation, Intensity
ICDRS	International Clinical Diabetic Retinopathy scale
IoU	Intersection Over Union
ISNT	Inferior, Superior, Nasal, Temporal
LoG	Laplacian Of Gaussian
LSTM	Long Short-Term Memory
MAs	Microaneurysms

TABLE 1. (Continued.) A list of abbreviations.

Abbreviations	
MCA	Multiple Correspondence Analysis
MCDR	Mean Cup-To-Disk Ratio
MDC	Multiple Dilated Convolution
ML	Machine Learning
MRCEV	Multi Logistic Regression Controlled Entropy Variance
MSA	Multi Scale Weight Shared Attention
nAMD	Neovascular AMD
OC	Optic Cup
OCT	Optical Coherent Tomography
OCT-A	OCT Angiography
OD	Optic Disc
ONH	Optic Nerve Head
PCV	Polypoidal Choroidal Vasculopathy
PR	Precision
RCGA	Real-Coded Genetic Algorithm
RCNN	Regions With CNN
REFUGE	Retinal Fundus Glaucoma Challenge
ResNet	Residual Network
RF	Random Forest
RGB	Red Green Blue
RNFL	Retinal Nerve Fiber Layer
RNN	Recurrent Neural Network
ROI	Region Of Interest
ROP	Retinopathy of Prematurity
SE	Sensitivity
SEs	Soft Exudates
SLIC	Simple Linear Iterative Clustering
SP	Specificity
SVM	Support Vector Machine
TN	True Negative
TP	True Positive
TPR	True Positive Rate
VGG	Visual Geometry Group
YOLO	You Only Look Once

of CNN. Convolution layers, pooling layers, and fully connected layers are three major components of a CNN. The training process consists of two stages. The first one is known

as a forward stage where the input image is represented with appropriate weights and biases in each layer, then the predicted output is used to measure the loss function by

TABLE 2. Overview of Fundus image Datasets for retinal disease diagnosis. (DL Tasks: -Vessel segmentation, -Exudates Detection, -MA Detection, -OD/OC segmentation, -HE Detection, -DR Diagnosis, -Glaucoma Diagnosis, -AMD Diagnosis, -Cataract Diagnosis, -ROP Diagnosis.)
























































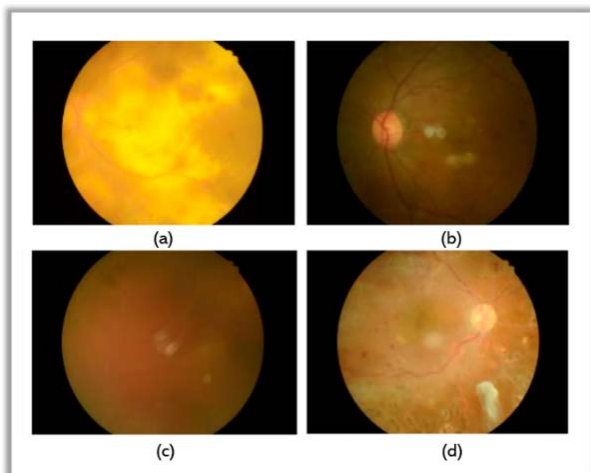
Dataset Name	Number of Images	Image Size (in pixels)	Image Format	Ground Truth Description	Disease diagnosis Task Utilization
DRIVE [10]	40	768 × 584	JPEG	33 - Normal, 7 - Mild early DR	 
STARE [11]	400	605 × 700	-	20 - Blood vessel segmentations annotated, 10 - Images, Annotated with pathologies, 10 - Images with Artery/Vein Labels	
HRF [12]	45	3504 × 2336	-	15 - Healthy Images, 15 - DR Images, 15 - Glaucomatous Images	 
CHASE_DB1 [13]	28	1280 × 960	TIFF	Blood Vessel Demarcation	
DIARETDB0 [14]	130	1500 × 1152	PNG	110 - DR signs (EXs, SEs, HEs, Mas and neovascularization), 20 - Normal Images	
DIARETDB1 [15]	89	1500 × 1152	PNG	84 - Mild NPDR signs (MAs), 5 - Normal Images	 
RC-RGB-MA [16]	250	2595 × 1944	-	MA Annotated-Bounding box	 
RC-SLO-MA [16]	58	1024 × 1024	-	MA Labels	 
ROC (Retinopathy online Challenge) [17]	100	-	-	MA centre location Labels	 
E-Optha EX [18]	82	2544 × 1696, 1440 × 960, 2048 × 130	JPEG	47 - Exudates Marked, 35 -Normal Images	 
E-Optha MA [18]	381	2544 × 169, 1440 × 960	JPEG	148 - MAs or small HEs marked, 233 - Normal images	 
MESSIDOR [19]	1200	1440 × 960, 2240 × 1488, 2304 × 1536	TIFF	-	 
MESSIDOR-2 [20]	1748	1440 × 960, 2240 × 1488, 2304 × 1536	TIFF	-	
CLEOPATRA [21]	298	-	-	EXs, HEs, MAs Marked	  
Kaggle/ EyePACS [22]	9963	-	JPEG	5-Stage DR Grading	
IDRiD [23]	516	4288 × 2848	JPEG	81 - MA, HE, SE, EX, OD area Marked 516 - DR and DME severity Grades	   
DDR [24]	13673	-	-	13673 - DR severity 5 classes, 757 - MA, EX, SE, HE Bounding box annotations	   
ONHSD [25]	100	640 × 480	-	96 - images with ONH Marked	
Drishiti-GS [26]	101	2896 × 1944	-	-	 
SINDI [27]	5783	-	-	5670 - Normal Images, 113 - Glaucomatous	
REFUGE [28]	1200	2124 × 2056, 1634 × 1634	-	-	
SEED [29]	235	-	-	43 - glaucomatic	
DrionsDB [30]	110	600 × 400	-	Images from glaucoma, eye hypertension Patients	
ORIGA [31]	650	3072 × 2048	-	168 - Glaucomatous, 482 - healthy images. OD and OC Boundaries, CDR Values	 
RIGA [32]	750	2240 × 1488, 2743 × 1936	-	-	
RIM-ONE [33]	169 ONH	-	-	118 - Normal, 40 - Glaucomatous, 11 - Ocular hypertension	 
ACHIKO-K [34]	258	-	-	144 - Glaucomatous, 114 - Normal	
LAG [35]	11760	3456 × 5184	-	2392 - Glaucoma	 
SCES [36]	1676	3072 × 2048	-	46 - Glaucomatous	 
AREDS [37]	206500	-	-	-	
iChallenge-AMD [38]	1200	-	-	924 - From non-AMD patient, 276 - From AMD Patients	 
KORA [39]	Images from 2840 individuals	-	-	7	

TABLE 3. Performance evaluation metrics.

Metric	Formula	Description
Sensitivity/Recall/TPR	$SE = \frac{TP}{(TP+FN)}$	Ratio of classified true positives to the actual number of true positives in the ground truth.
Specificity/FPR	$SP = \frac{TN}{(TN+FP)}$	Ratio of classified true negatives to the actual true negatives in the ground truth. False-positive rate (FPR) = (1-SP)
Precision	$PR = \frac{TP}{(TP+FP)}$	Precession indicates what proportion of positive findings was actually correct. Higher value of PR, indicates better system Performance.
Accuracy	$ACC = \frac{(TP+TN)}{(TP+FP+TN+FN)}$	Accuracy indicates the ratio of correct predictions to the total number of predictions.
F1 -Score	$F1 = \frac{2TP}{(2TP+FP+FN)}$	Represents the harmonic mean of recall value and precision. Higher value of F1-score indicates better system Performance. [40]
AUC	-	Indicates the Area under the Receiver Operating Characteristic Curve (AUC) Higher value of AUC indicates better system Performance.
IoU/ Jaccard similarity index	$IoU = \frac{TP}{(TP+FN+FP)}$	Intersection-over-union (IoU), widely used measure for understanding how accurate a proposed image segmentation is, compared to a known/ground-truth.
DSC (Dice Similarity Coefficient/ Dice Metric)	$DSC = \frac{2TP}{(2TP+FP+FN)}$	The value of a DSC ranges from 0 to 1, 0 - indicating no spatial overlap between two sets of binary segmentation results, and 1- indicating complete overlap

**FIGURE 3.** Examples of fundus images with the poor quality a) overexposure, b) underexposure, c) obscure, and d) postoperative.

comparing it with ground truth values. In the second stage, known as the backward stage, gradients of each parameter are computed based on the loss function. The parameters are then updated and initialized for the next forward stage. This is repeated as multiple iterations until the network presents accurate classification results.

2) VGGNet

Another commonly used backbone network in retinal disease classification is VGG Network (VGGNet). This was proposed by Karen Simonyan and Andrew Zisserman in 2014 [46]. Figure 5 shows the architecture of a VGGNet. VGG stands for Visual Geometry Group, which released various versions of Convolution network models for various image classification tasks starting from VGG-16 to VGG-19. The original intention behind the development of VGG is

to research how the depth of CNN impacts the accuracy of image classification. A small 3×3 kernel is used in all layers of the model to increase the depth of the network while avoiding too many parameters. In VGGNet, the input is set to a size of 224×224 RGB image. A 3×3 filter is used with a fixed convolution step. There are three fully connected layers which can vary from VGG-11 to VGG-19 depending on the total number of convolutions plus fully connected layers. VGG-11 has eight convolution layers followed by three fully connected layers. On the other hand, VGG-19 has sixteen convolution layers and three fully connected layers. In VGGNet, each convolution layer is not followed by a pooling layer; instead, a total of five pooling layers are distributed throughout the network as shown in Figure 5.

3) ResNet


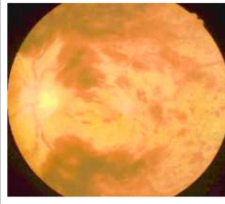

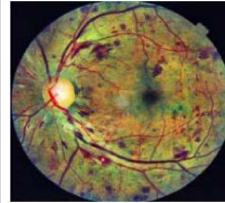

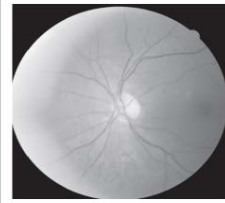
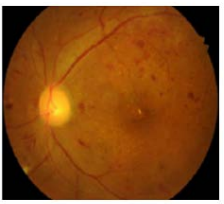
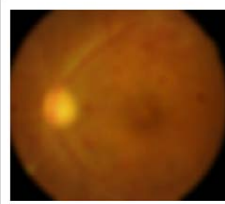
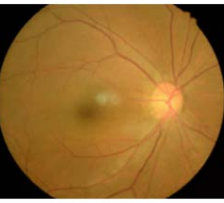

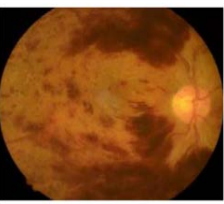
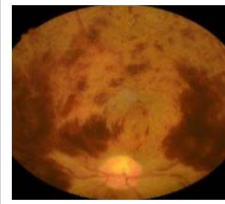
Residual network (ResNet) [47] consists of a total of 152 layers, which are built by stacking individual residual blocks shown in Figure 6 (a) and (b). Each of these residual blocks consists of two convolution layers (3×3). Periodically, the number of filters is doubled and down sampled spatially using a stride of 2. This network employs special skip connections along with batch normalization after each convolution layer. Skip connections are used to optimize such deep models as they take activation from one layer and directly feed it to another layer. This has the effect of training deep networks without encountering vanishing gradient problems. To reduce the number of parameters, ResNet has a fully connected layer that outputs 1000 classes.

B. BACKBONE MODELS FOR SEGMENTATION IN FUNDUS IMAGES

1) FULLY CONVOLUTION NETWORKS (FCNs)

Long *et al.* [48] proposed a modified CNN network, by replacing fully connected layers with upsampling layers (shown in Figure 7). The extracted feature maps from initial

TABLE 4. Commonly used pre-processing techniques for fundus image analysis.

Fundus image Pre-processing technique	Description	Original Image	After Pre-processing
Contrast enhancement (Histogram equalization) [41]	To highlight foreground pixels from the background. Histogram equalization increases the overall (global) contrast of the image.		
Contrast enhancement (CLAHE) [42], [43]	Contrast Limited adaptive histogram (CLAHE) is a widely used technique, especially for fundus images. Eliminates the problem of over-amplifying in constant pixel areas of the image. Enhances minute lesions and markers like microaneurysms in fundus images.		
Colour space transformation	In DL model implementation for fundus images, in certain cases, the model performance maybe improved by utilizing single color from RGB Channels. The extraction of green channels from the fundus images is famous as they offer high contrast images with rich visual information.		
Noise Removal [44]	Many denoising algorithms like Gaussian filters, median filters, non-local means denoising, etc. are utilized for removing unwanted noise. One tradeoff is, denoising can also blur the image and degrades by removing fine details of the image		
Cropping and Extracting Region of Interest (ROI)	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 reducing the unwanted learning burden.		
Augmentation	Augmentation techniques like image rotation (shown here), rescaling, flipping, translation, etc. are employed to balance the image datasets. This helps in improving model performance and robustness.		

layers are up-sampled to the equivalent size of the input image. A fully convolution network (FCN) can perform dense pixel-wise prediction, making it better suited for segmentation tasks compared to CNN.

2) U-NET

Ronneberger *et al.* [49] proposed a network shown in Figure 8, which has symmetrical encoder and decoder structures along with several skip connections from the encoding

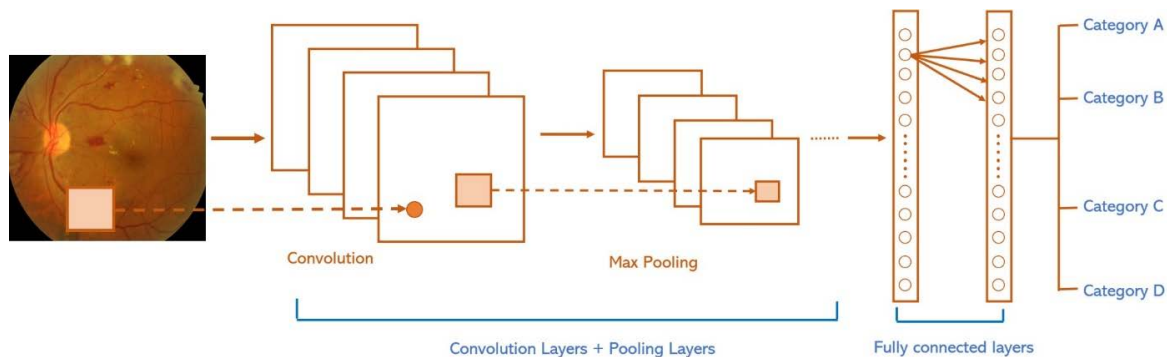


FIGURE 4. Basic CNN architecture pipeline.

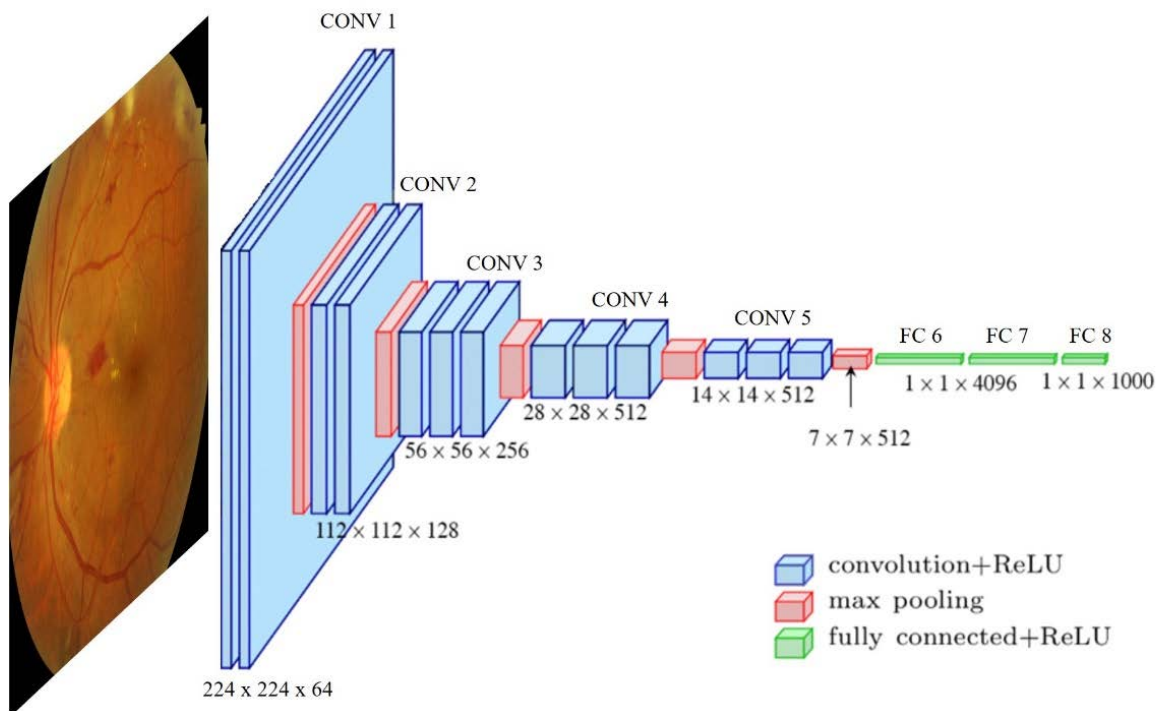


FIGURE 5. The architecture of VGGNet.

path to the decoding path. The encoder is responsible for extracting features from input images while the decoder reconstructs the images for the final output. The skip connections allow the network to make improved predictions by directly connecting low-level feature maps from encoder to decoder.

V. DEEP LEARNING IN RETINAL DISEASE DIAGNOSIS

A. DIABETIC RETINOPATHY (DR) DIAGNOSIS

Diabetic retinopathy is one of the most common retinal diseases that can cause blindness, if not treated in time. This is a complication seen in one-third of diabetes patients [50]. A survey estimates nearly 93 million people worldwide suffer from DR [51]. Any diabetes patient can develop DR causing

vascular disruption in the retina. These numbers are expected to grow higher considering the rapid growth in the number of diabetes patients worldwide [52]. International Clinical Diabetic Retinopathy scale (ICDRS) has classified the severity of DR into five categories, 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. Many DL models were implemented to design a robust model for DR diagnosis using fundus images. Recent papers that deal with this are discussed in the following section. The experimental results are listed in Table 5.

Wang et al. [53] proposed a network that jointly performs multiple tasks of increasing image resolution, various DR lesion segmentation, and DR grading. Their method leverages

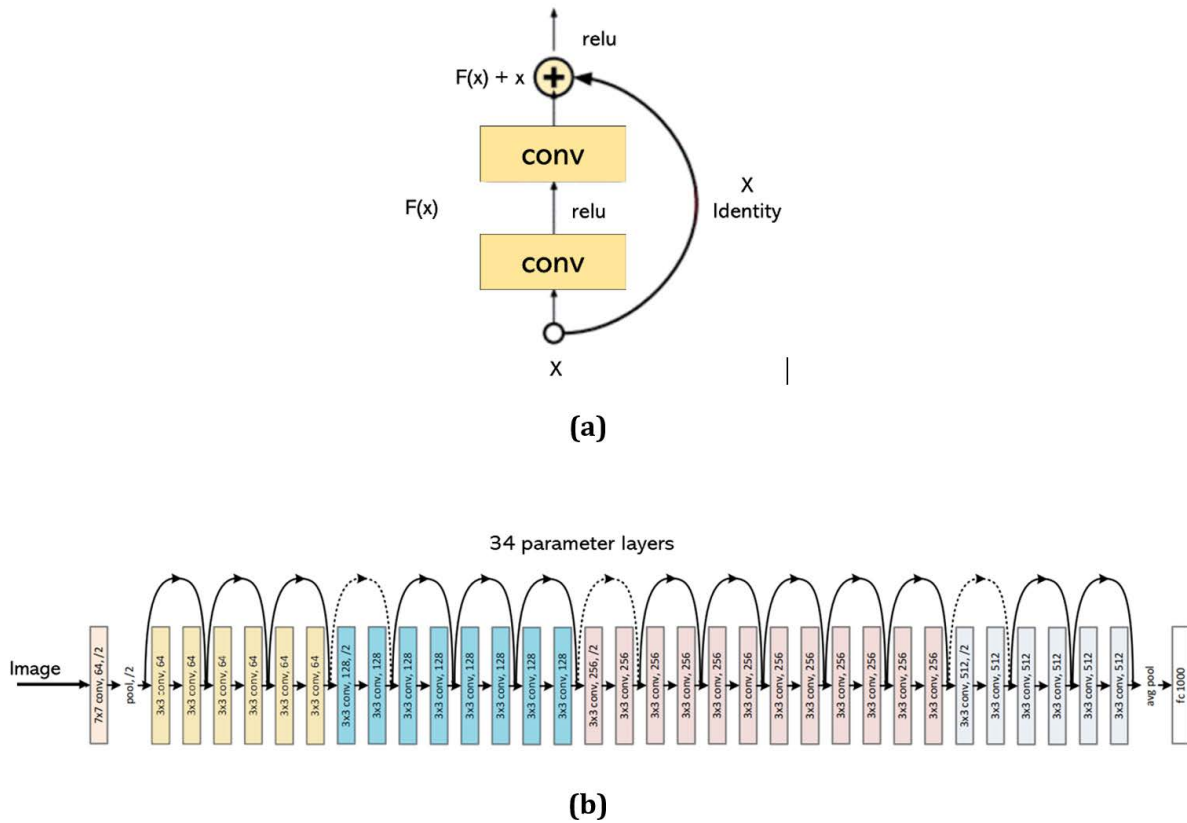


FIGURE 6. a) Residual block, and b) architecture of ResNet.

the fact of high-resolution images are suitable for accurate grading due to appropriate lesion segmentation. For each of the tasks, they employed CNN-based methods, where a robust feedback mechanism is established by utilizing task-aware loss functions. Li *et al.* [54] used an ensemble approach for developing a DR diagnosis method that utilizes enhanced Inception-V4 on a privately collected dataset which is then generalized using the Messidor-2 dataset. They investigated the effects of input image size and its number on model performance. Automatic DR diagnosis presents the difficult task of handling fundus images that are captured at different illuminations.

Kaushik *et al.* [55] proposed to handle these irregularities using image desaturation techniques in the pre-processing stage. They stacked three CNNs for their training process, where optimum weights from these networks are fused for classifying fundus images for DR diagnosis. Das *et al.* [56] proposed a DL method that examines the branching of retinal blood vessels and abnormal vessel growth to identify DR from fundus images. After the pre-processing stage, they used the maximal principal curvature technique for segmenting blood vessels followed by histogram equalization and a morphological operation for further refining the results. A CNN-based classifier was developed to work on the segmented blood vessels to classify for DR diagnosis.

Alyoubi *et al.* [8] presented a method consisting of two DL models working simultaneously. The first one was based on CNN which classifies the image into five DR categories and the second one for detecting DR lesions based on YOLOv3. Finally, both the models were combined to achieve improved accuracies. Shankar *et al.* [57] proposed a method where the input images are first treated for noise removal followed by histogram-based segmentation to retrieve salient regions for DR grading. Then a Synergic DL is employed for classification that consists of three sub-modules.

Shankar *et al.* [58] in their model for DR screening utilized Bayesian optimization for hyperparameter tuning of the DL architecture based on Inception-V4. First, they pre-processed the given images for contrast enhancement using the CLAHE algorithm, and then histogram-based segmentation was performed to generate suitable input images from which various features were extracted and utilized for DR classification. For early diagnosis of non-proliferative DR, Qiao *et al.* [59] utilized a CNN-based network along with the use of various pre-processing techniques like LoG, BPF, Match filters, and curve transform. They also developed a microaneurysm detection method based on principal component analysis. Wang *et al.* [60], in their work, developed a DL framework incorporating multiple tasks for DR diagnosis (Classification into different severity levels) and DR features

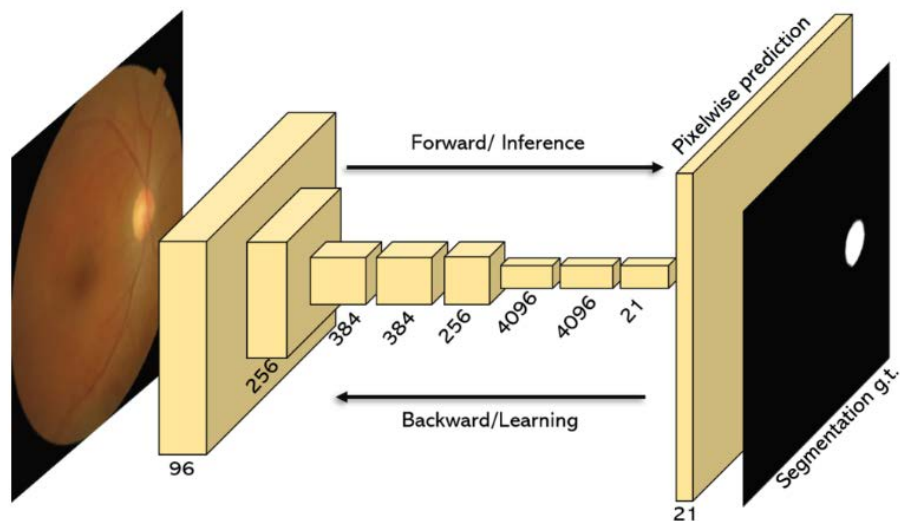


FIGURE 7. Architecture of an FCN.

simultaneously, which can act as supporting information. It comprises squeeze and excitation (SE) as the backbone for feature extraction at higher scales and two heads, one for DR severity classification and another for DR feature detection.

Araújo *et al.* [61] proposed a DR grading system that can support its decision by providing a grade uncertainty parameter. The network consists of convolutional blocks from which a lesion map is generated that are indicative of the presence of a lesion. Multiple instance learning along with gaussian sampling was utilized for computing grade-wise explanation maps. For detecting referable DR, Sahlsten *et al.* [62] developed a DL framework based on the Inception-V3 model, which had already been trained on the ImageNet dataset. High-resolution images were used for the training process, which tends to yield better results with comparatively smaller training samples. As training of such high-resolution images may take a lot of time intervals, they also took an ensemble approach where six DL networks were working with low-resolution images and performed comparative analysis.

Qummar *et al.* [63] used an ensemble approach for improving DR severity classification in the early stages. They utilized five DCNN models for extracting salient features and generating probabilities that indicate the image's adherence to a particular DR class. The ensemble was achieved by stacking. Nneji *et al.* [64] employed two separate DL models, Inception-V3 and VGG-16 to work on two individual channels of input fundus image. One channel is derived from applying CLAHE and another from the CECED pre-processing technique. Outputs of both the DL

models were weighted and merged for final DR classification. Bora *et al.* [65] developed two types of DL systems based on Inception-V3 architecture for predicting the growth of DR in patients with diabetes. The DL systems were categorized based on the one-field (primary only) or three-field (primary, temporal and nasal) fundus images that they take as input. A five-stage DR classification network was proposed by Majumder and Kehtarnavaz [66], where they implemented a multitask model with two separate models. One classification model with cross-entropy loss function and another was a regression model with a mean square error loss function. After training both of them separately, the extracted features were concatenated and utilized by a final perceptron network for five-stage DR classification.

1) DR DIAGNOSIS USING MICROANEURYSMS (MAs)

Microaneurysms (MAs) are one of the earliest visual indications of DR and have gained a lot of research interest in the field of fundus image analysis. The challenges of automatic detection and segmentation of MAs are their invariance to other lesions, their low contrast nature in fundus images, and extremely low pixel count compared to background pixels. Recent DL models, that handle the problem of MA segmentation/detection are discussed below. The Performance comparison is presented in Table 6.

Xia *et al.* [69] proposed a two-stage network, one for efficient feature extraction that employs residual learning from multiple scales of input images and the second stage for filtering out the false-positive patches. Liao *et al.* [70] used

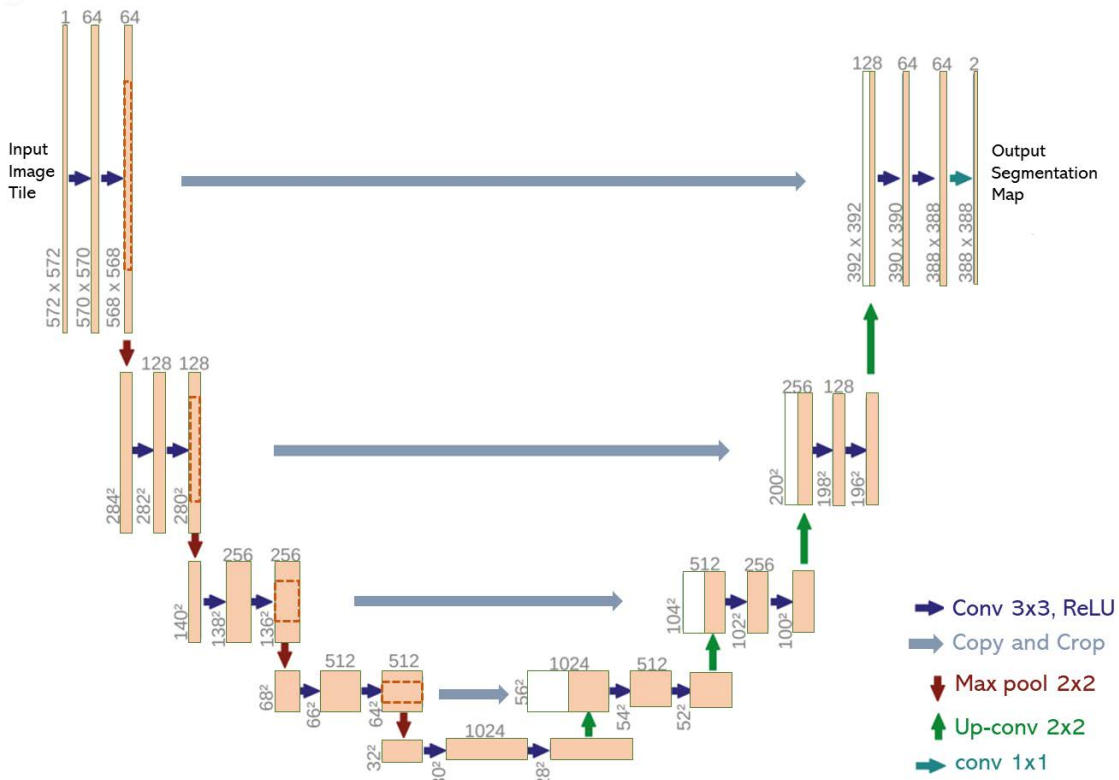


FIGURE 8. Architecture of U-Net.

an encoder-decoder network for MA detection by utilizing the difference between skip connection layers. A customized loss function was used (smooth dice loss) for allowing the network to concentrate more on hard samples during the training process. They also modified the standard activation function to achieve a very precise probability distribution for MA detection. Zhang *et al.* [71] developed training and test samples consisting of green and blue channels of the original fundus image and two additional samples, one with enhanced contrast and another with a suppressed background. These were then used in a feature transfer network for detecting MAs, where the optimized weights from the previous phase were carried forward for the next learning phase.

2) DR DIAGNOSIS USING EXUDATES

Another important biomarker for detecting DR is Exudates. Hence Hard and Soft Exudates (HEs and SEs) segmentation is another widely researched area in fundus image analysis. Some of the recent works in this direction are discussed below. Table 7 shows the experimental results of different articles on Exudates segmentation.

Huang *et al.* [74] in their work regarding hard exudates detection first used the Simple Linear Iterative Clustering (SLIC) algorithm to generate superpixels at each

input image. Then various pixel and superpixel level features were derived and training patches were produced from each feature. These patches were applied to a CNN model to classify them into HE pixels or background pixels. Many DL detection/segmentation methods focusing on pixel-by-pixel annotation often lead to a danger of catastrophic interference, where the model abruptly forgets the previously learned attributes while learning new information. He *et al.* [75] in their work used incremental learning to avoid this problem, where the knowledge of the previous model is utilized to perfect the present model.

Kurilová *et al.* [76] presented a method that utilizes a machine learning technique for filtering input images before employing them for the DL task. A Support Vector Machine (SVM) classifier along with a faster R-CNN network was used for preliminary scanning of input image patches. Image patches without exudates were discarded while others were used for the object detection network. This helped in improving the speed and detection accuracy. To deal with the segmentation issues due to class imbalance and vast size variations in HE lesions, Liu *et al.* [77] proposed a double branch network, where the easy task of large HE segmentation is performed first and then gradually shifted the attention to the hard task of small HE segmentation. This is

TABLE 5. DR diagnosis performance comparison.

References	Dataset	ACC	PR	SE	SP	AUC	F1	Kappa
[53]	DDR	0.836	0.831				0.83	0.802
[53]	EyePACs	0.869	0.871				0.857	0.852
[54]	Private			0.923	0.947	0.977		
[55]	EyePACs	0.979		0.977	1.00			
[56]	DIARETDB1	0.987	0.972	0.996	0.982			
[8]	DDR	0.890		0.890	0.973	0.970		
[57]	MESSIDOR	0.992		0.985	0.993			
[58]	MESSIDOR	0.994		0.988	0.996			
[60]	Private							0.808
[61]	Kaggle DR							0.740
[62]	MESSIDOR			0.896	0.974	0.987		
[63]	Kaggle DR	0.808	0.638		0.867		0.537	
[64]	MESSIDOR	0.985		0.989	0.98			
[64]	Kaggle DR	0.98		0.987	0.978			
[66]	EyePACS	0.82	0.69	0.64			0.66	
[67]	Kaggle DR			0.925	0.907	0.968		
[68]	APTOS	0.994	0.934	0.982	0.995			

TABLE 6. MA segmentation performance comparison.

Ref.	Dataset	Task	ACC	SE	SP	F1
[69]	E-Optha MA	Segmentation	0.999	0.705	0.999	0.619
[70]	E-Optha -MA	Detection		0.781		
	ROC	Detection		0.559		
[71]	ROC	Detection	0.723	0.983		
	DIARETDB1	Detection	0.964	1.00		
[72]	e-Optha -MA	Detection	0.911	0.993		
	E-Optha MA	Segmentation	0.997	0.67	0.998	
[73]	DDR		0.105			
	E-Optha MA	Segmentation	0.168			

achieved through carefully guiding the training process by a customized Dice loss.

Choice of color space of an image can also influence the accuracy of Exudates detection, as was demonstrated by Khojasteh *et al.* [78]. They first applied principal component analysis on three basic color spaces (RGB, LUV, and HSI) for contrast enhancement, and then a set of training samples were generated in all the color spaces to train a CNN for detecting exudates. Through the study, they also proposed a new color space named PHS space, for accurate detection. Kou *et al.* [79] implemented a modified U-Net structure, consisting of one encoding path and three decoding paths. They replaced the general convolutional block of a U-Net with residual blocks for detailed feature extraction during the learning process. Zong *et al.* [80] also proposed a few modifications to the U-Net network for handling the uneven

distribution of HE in input image patches. The inception module replaced the basic convolution blocks for deriving features from various scales. Also, residual connections were used for generating the output. For loss function, they used Focal loss, which suitably tackles the data imbalance problem. Mohan *et al.* [81] proposed an exudate detection process based on altered KAZE features that effectively extracts feature points. Machine autoencoders with extreme learning capability were used for exudate localization.

Liu *et al.* [82] tackled the extreme size variation and class imbalance problem in the hard exudate (HE) segmentation task, through a dual branch network. One for large exudates and another for small exudates. They also utilized a custom loss function (dual sampling modulated) for the training process to segment HEs in different sizes. Mohan *et al.* [83] demonstrated a unique feature extraction method based on Hessian Matrix approximation. The model was tested on multiple datasets including a privately collected dataset.

3) DR DIAGNOSIS USING HEMORRHAGES

Hemorrhages are one of the visual signs of DR, which develop due to the burst of retinal blood vessels under extreme pressure build-up inside the vessels. The hemorrhage segmentation task is another direction taken by many researchers for DR diagnosis. Some of the recent works are discussed below and the experimental results are presented in Table 8.

Maqsood *et al.* [84] introduced a method for hemorrhage detection, where initially, edge details of the input image are enhanced through contrast modification and then passed onto a second stage that employs a 3D-CNN for segmentation. A modified VGG-19-CNN is also used to implement

a transfer learning strategy for extracting features. Finally, before sending for feature fusion and classification, MRCEV-based feature selection is performed to mitigate redundancies. Lahmiri [85] combined CNN with a machine learning technique for detecting and classifying hemorrhage in fundus images. The task was performed in three stages, beginning with a CNN for feature extraction, followed by utilizing a Student *t*-test for further filtering and selecting key features. In the third stage, the selected features were passed through a support vector machine classifier for segregating images with hemorrhage from healthy ones.

4) DIAGNOSIS USING RETINAL VESSEL

Retinal blood vessels serve as a prominent biomarker for indicating the health of the eye. A variety of geometric characteristics like branch lengths, branch angles, vessel diameter among others can be derived from the retinal vessel map. These characteristics are used for the diagnosis of diseases like DR and Glaucoma. Researchers have concentrated on retinal vessel segmentation and achieved excellent results. Some of the Recent articles are discussed in the sections below. Experimental results of several works on vessel segmentation are presented in Table 9.

Yang *et al.* [86] proposed a method where initially a separate module based on U-Net is used for accurate segmentation of thin and thick vessels, which uses a common encoder for feature extraction followed by two decoders that use corresponding ground truth images of thick and thin vessels independently. Then a fusion module is employed to combine the two segmentation results from the previous module. Both U-Net architectures use additional skip connections to improve the context information during the training. To minimize computational time, Boudegga *et al.* [87] presented a new architecture, where the first image patches are extracted after pre-processing and augmentation before the training. Their method utilizes a U-shaped structure, implemented through lightweight convolution modules (LCMs). Segmented image patches were then merged in a post-processing stage to obtain the final results. Another U-Net-based DL model was presented by Fukutsu *et al.* [88] for vessel segmentation along with arteriovenous classification using probability maps. In addition to the traditional skip connections between down sampling and up sampling block, their network has short connections for minimizing gradient loss. They also implemented a multiple dilated convolution (MDC) block between encoder-decoder for extracting global features.

In their architecture for vessel segmentation, Atli and Gedik [89] modified the conventional encoder-decoder network by first performing up-sampling followed by a down-sampling operation. Their model attempts to continue the learning process during the progression of sampling operations.

For reducing computational complexity and improving model generalization Gegundez-Arias *et al.* [90] presented an altered U-Net model which works on image patches

derived from fundus images and employs a unique loss function during the training that considers each pixel distance to the vascular tree structure. It uses probability-based prediction for vessel segmentation. They also decreased the convolution count in each layer along with overall network depth for minimizing the model parameters. Building on the base VGG-16 network, Samuel and Veeramalai [91] proposed a method that can segment retinal blood vessels from both fundus images as well as from coronary angiograms. This architecture consists of two feature extraction layers on top of the base network. Each of these is responsible for localizing blood vessels, passing important vessel features through the intermediate layers, and summing up the feature maps from earlier stages.

To achieve a satisfactory segmentation from relatively small datasets Chen *et al.* [92] developed a method that uses a semi-supervised learning strategy along with U-Net architecture. Their encoder-decoder structure first uses a relatively small number of labeled data for training and later updates the old dataset using a custom-built strategy. Tian *et al.* [93] took a multi-path approach for vessel segmentation. The first path, consisting of convolution sampling blocks, utilizes the low-frequency image characteristics for deriving global features and the second path, composed of a coding and decoding region, concentrates on high-frequency components to get local feature details. Final segmentation results were obtained by fusing both information. Wang *et al.* [94] used an attention-driven encoder-multi decoder network for the segmentation task. A basic U-Net structure is first implemented for generating a rough vessel segmentation map, which is compared with available ground truth to identify hard and easy segmentation tasks in the image. This information serves as a basis for two additional decoders which focus on extracting features for hard and easy regions independently. All three outputs of the decoders are combined and finally fed into a light U-Net to yield the final results.

Biswas *et al.* [95] proposed a model that utilizes a dilated convolution to increase the receptive fields (amount of input image visible to the innermost network layer) without stacking the convolution layers linearly. This in turn helped in providing a better context for the segmentation task, without necessarily increasing the computations. Wang *et al.* [96] presented their work, where image patches were generated after the pre-processing stage and fed into an encoder-decoder structure, but through two separate paths, one for grabbing more receptive fields and another for storing spatial information. A unique Attention mechanism was employed to improve the original features, and a Fusion module to merge the features from the two paths. For accurate segmentation of capillaries of retinal fundus images, Wu *et al.* [97] demonstrated a cascaded deep network. The first network generates retinal vessel maps (probabilistic) based on the input image patches. The second network connected in series with the first one uses these maps to produce refined segmentation results. For avoiding the loss of information caused by the downsampling of the maps, skip connections were arranged

TABLE 7. Exudate segmentation performance comparison.

References	Dataset	Lesion	ACC	SE	SP	AUC	F1	IoU	
[74]	IDRiD	HE	0.981	0.984	0.906	0.967			
	e-Optha EX	HE	0.975	0.979	0.908	0.968			
[76]	e-Optha EX	HE				1.00	0.947		
	DiaretDB1	HE				0.972	0.881		
	Messidor	HE				0.885			
[77]	DDR	HE						40.47	
	IDRiD	HE						61.19	
[78]	DiaretDB1 + e-Optha	Exudates	0.982	0.99	0.96				
[79]	DDR	Exudates	0.93	0.871	0.926	0.961			
	e-Optha EX	Exudates	0.996	0.983	0.976				
[80]	IDRiD	HE	0.979	0.963	0.971				

TABLE 8. Hemorrhage segmentation performance comparison.

References	Dataset	Task	ACC	SE	SP	AUC
[84]	HRF, DRIVE, STARE, MESSIDOR, DIARETDB0, DIARETDB1		0.982	0.975	0.978	
[85]	STARE	Classification	0.991	0.991	0.990	0.973

between the two cascaded networks. In their attempt at reducing mis-segmentations and computational complexities, Xiuqin *et al.* [98] combined U-Net with a residual learning scheme. This enabled them to make the network deeper which is helpful for accurate segmentation, while the inbuilt residual module handles the network degradation caused due to the network depth.

It is difficult to segment retinal vessels in the presence of lesions or to identify microvessels due to low contrast in fundus images. Dharmawan *et al.* [99] presented a hybrid algorithm for segmentation tasks that involves a contracting path and an expansive path. Their architecture does not use a fully connected layer and hence reduces a substantial computational load. A provision for concatenation operation in their model helps it to train from both local and global features, yielding better results.

B. GLAUCOMA DIAGNOSIS

Glaucoma is another leading cause of irreversible blindness around the world [100]. Like so many other retinal diseases, researchers have concentrated on developing various DL models to diagnose glaucoma from fundus images. Recent developments in this direction are discussed in the following section. Various experimental results for DL-based glaucoma diagnosis are presented in Table 10.

Xu *et al.* [101] developed a DL framework for glaucoma diagnosis, with a relatively small number of training samples through the extraction of OD, OC, and retinal nerve fiber layer (RNFL) characteristics. Their framework consists of a pre-diagnosis classification phase based on a general

fundus image (global attributes). In the next phase, the above-mentioned biomarker segmentation is performed and ISNT and MCDR scores are calculated. The final diagnosis was performed by utilizing all the segmentation results. Shanmugam *et al.* [102] used the cup to disc ratio (CDR) for identifying glaucoma in a fundus image. Primarily, their method concentrates on accurate segmentation of OC and OD, which was performed by a modified U-Net. Due to the addition of adaptive convolution in their framework, the computational burden was reduced as it used fewer filters than the standard U-Net. The morphometric attributes derived from the segmentation results were then utilized by a random forest classifier for the classification of glaucoma images from the healthy ones. In another work, Wang *et al.* [103] employed a transfer learning approach using VGG-16 and AlexNet for model training and glaucoma classification. They collected ONH images from various publicly available datasets and constructed two sets. One set where various data augmentation techniques like random scaling, cropping, rotation, and flipping were performed to expand the dataset. In another set, three-dimensional topography maps of ONH were constructed from shading information of 2D images (SHS method). Both sets, when evaluated for Glaucoma classification, yielded improved performance than regular CNN classification models. Gheisari *et al.* [104] utilized fundus image sequence (video) for extracting temporal features along with spatial features from static images to improve glaucoma classification accuracy. They implemented a fusion method that combines CNN with a Recurrent neural network (RNN). For CNN, two DL models (ResNet-50 and VGG-16)

TABLE 9. Retinal vessel segmentation performance comparison.

Ref.	Dataset	ACC	SE	SP	AUC	F1
[86]	DRIVE	0.957		0.975		0.829
	CHASE_DB1	0.9632		0.977		0.799
	STARE	0.962		0.987		0.815
[87]	STARE	0.981	0.806	0.992		
	DRIVE	0.981	0.844	0.99		
[88]	DRIVE	0.97	0.78	0.99		
[89]	DRIVE	0.968	0.826	0.982		
	CHASE_DB1	0.967	0.785	0.984		
	STARE	0.971	0.677	0.994		
[90]	DRIVE	0.954			0.983	
	CHASE_DB1	0.966			0.988	
	STARE	0.975			0.992	
[91]	DRIVE	0.962		0.982	0.978	
	STARE	0.973		0.981	0.99	
[92]	DRIVE	0.963			0.976	
[93]	CHASE_DB1	0.96	0.877	0.968	0.957	
	DRIVE	0.958	0.863	0.969	0.956	
[94]	DRIVE	0.958	0.799	0.981	0.982	0.829
	CHASE_DB1	0.967	0.823	0.981	0.987	0.819
	STARE	0.967	0.818	0.984	0.988	0.837
[95]	DRIVE	0.956	0.782	0.981	0.979	
[96]	CHASE_DB1	0.970	0.842	0.983	0.982	0.810
	DRIVE	0.956	0.807	0.978	0.98	0.825
[97]	DRIVE	0.958	0.799	0.981	0.983	
	CHASE	0.968	0.800	0.988	0.989	
	STARE	0.967	0.796	0.986	0.987	
[98]	DRIVE	0.965	0.831	0.986	0.981	
[99]	STARE		0.792	0.982		0.816
	DRIVE		0.831	0.972		0.823

were implemented and tested. LSTM-based RNN was used as it can select and retain useful information from the image sequence. A fully connected layer is established at the end of the fusion module for enhanced glaucoma classification.

To avoid problems like overfitting and the necessity of large datasets, Nayak *et al.* [105] have proposed a network that utilizes a feature optimization technique based on biological phenomena, known as a real-coded genetic algorithm (RCGA). Once the improved features are derived through this technique, various classifiers are utilized for identifying glaucoma-based images. RCGA algorithm along with SVM classifier provided the best results. Li *et al.* [106] proposed a CNN-ResNet architecture, with 101 layers that utilize a total of 26,585 images, for testing and training the model. They avoided the vanishing gradient problem by applying skip connections between the layers during the training process. Hemelings *et al.* [107] developed a CNN-based method that combines transfer learning with active learning strategies for

accurate glaucoma diagnosis. They worked with 8433 fundus images for developing their classifier. After basic image pre-processing techniques like ROI extraction and data augmentation, a pre-trained ResNet-50 was utilized for transfer learning, which consists of 182 layers. ‘Uncertainty sampling’ technique was employed as an active learning method for the classification system. In addition to this, saliency maps have been generated that support the model decisions.

Juneja *et al.* [108] proposed a DL model, where, after certain pre-processing techniques like image cropping, augmentation and denoising, images were sent into a CNN-based model (76 Layers deep). To compensate for the loss of data, they used ‘Add layer’ in every block that combines the previous block output with the next block. Martins *et al.* [109] developed a glaucoma diagnosis pipeline that can be executed offline on mobile devices. They mainly relied on OD and OC segmentation (U-shaped model) for generating useful morphological features which are used by a separate classification network (based on MobileNet-V2 as the backbone). All the results along with morphological calculations are gathered for constructing a diagnosis pipeline that was integrated with a mobile application. Liu *et al.* [110] developed a DL framework for glaucoma detection by utilizing 241032 images collected from the Chinese glaucoma study alliance (CGSA). Down sampled images are fed into a CNN architecture (based on ResNet). For accurate generalization, a unique ‘online DL system’ was developed where experts confirm the model classification results and then the confirmed samples were utilized for retuning the model before the next prediction.

Bajwa *et al.* [111] performed glaucoma classification in two stages. The first stage utilizes ‘Regions with CNN’ (RCNN) for OD extraction and localization. It also includes a semi-automatic ground truth generation part, that helps in creating ground truths containing the location of OD for training the RCNN. The second stage is composed of four convolution layers and three fully connected layers and uses the ROI images (generated after OD extraction) for classification. Kim *et al.* [112] proposed a two-task network that utilizes various CNNs for glaucoma classification and ‘Gradient weighted class activation mapping’ for localizing most suspicious glaucomatous regions for a given fundus image. Out of various CNN variants, the ResNet-152-M model achieved the most promising results. As an extension, the authors also developed a web-based application incorporating the model in the background, which provides decision, diagnostic confidence score, and suspicious location for a given input fundus image. Singh *et al.* [113] conducted a detailed study on a variety of DL methods for classifying fundus images into glaucomatous and normal categories. ORIGA, HRF, and ACRIMA were used as training and validation datasets. This study indicates Xception models and Inception-ResNet-V2 which yield better performance compared to others.

Ovriu *et al.* [114] proposed a dense network consisting of 201 layers for improving the performance of glaucoma classification. Each layer of this network gathers

additional inputs from all previous layers. In another work, Saravanan *et al.* [115] demonstrated an autoencoder architecture for glaucoma detection along with AVP recognition; they specifically concentrated on reducing classification errors through multi-modal learning implementation. Shoukat *et al.* [116] compared the performance of three pre-trained CNN-based models for early glaucoma detection. The test was conducted on RIM-ONE, G1020, and REFUGE datasets. Pretrained- EfficientNet-B7 yielded the highest accuracy on the G1020 dataset. Islam *et al.* [117] compared the performances of different DL models like DenseNet, MobileNet, EfficientNet, and GoogleNet on a private dataset consisting of 643 fundus images. The best performance was achieved by EfficientNet-b3 when cropped OD and OC images were used for training. As an alternative to automatic glaucoma detection, they also developed a vessel segmentation model using U-Net architecture. For early detection of glaucoma, Shoukat *et al.* [118] utilized G1020 and DRISHTI-GS datasets for their EfficientNet based model training. Image enhancement through various pre-processing techniques was carried out in the initial stage to highlight crucial features in the fundus image.

1) OD/OC SEGMENTATION

Other important retinal biomarkers used for the diagnosis of glaucoma are Optic Disc (OD) and Optic Cup (OC). The cup-to-Disc ratio is calculated from vertical cup diameter and vertical disc diameter. Hence accurate segmentation of OD/OC has become crucial for glaucoma diagnosis and a lot of research has been carried out in this direction. Some of the recent articles on DL-based OD/OC segmentation are discussed in the following sections along with experimental results in Table 11.

Wang *et al.* [124] developed a DL network, based on the U-Net framework, consisting of two subnetworks, a feature detection subnetwork (FDS) and a cross-connection subnetwork (CCS). The first subnetwork is responsible for extracting desired objects and necessary image features, while the second subnetwork is used for object segmentation. The presence of two subnetworks introduced several multiscale features into both the encoding and decoding process by element-wise subtraction. This in turn made the model more sensitive to edge information and played a crucial role in accurately segmenting the Optic Disc. Veena *et al.* [125] proposed two individual CNN models for OD and OC segmentation. First, they located the optic nerve region using basic edge detection methods like ‘Sobel’ and ‘watershed algorithm’, and then the image was cropped to the desired region. The cropped images were fed for the segmentation task into two separate CNN models composed of 39 layers each. The increased number of layers in each CNN model contributed to the extraction of ample image features and also helped in retaining the image resolution in the output image which improved the segmentation results. Kumar and Bindu [126] presented a U-Net-based architecture consisting of three subsections, namely ‘contraction’, ‘bottleneck’, and ‘expansion’.

The kernel size in the first subsection starting at 16, got doubled with each block and reached 256, and this helped the architecture to learn more complicated features from the input image. The expansion subsection contains the same number of blocks as the contraction, and every input is combined with the feature maps of the subsequent contraction layer. This approach alleviated the problem of gradient vanishing during the training of the model.

Natarajan *et al.* [127] developed a lightweight network for OD segmentation, where they used a gaussian mixture model (GMM) superpixel segmentation algorithm followed by a ‘simple linear iterative clustering’ to extract the region of interest (RoI) from the fundus image. These ROIs are then fed into a U-Net architecture for the semantic segmentation of OD; to smoothen the isolated points and coarse edges of the output, a regularization term is introduced to the loss function. This helped in improving the model generalization for the segmentation task. Panda *et al.* [128] presented their work on OD and OC segmentation using a model involving residual learning built on CNN-based architecture. They carried out several image pre-processing techniques like region of interest (RoI) selection around the OD center, image normalization, and contrast enhancement. This was followed by random patch (32×32 pixels) extraction from both the pre-processed images and their corresponding ground truths as inputs for the training process. The unique combination of convolutional layers with residual blocks, processed on patch-level data, allowed the model to focus on localized structure similarities. This process flow helped in improved OD and OC segmentation, considering the limited availability of training samples.

Fu *et al.* [129] demonstrated a fusion method for OD segmentation to improve model performance and avoid the distraction in images caused by bright lesions and illumination variations. First, two separate U-nets were used for detecting the OD and retinal blood vessels independently. Hough transform method is employed to fit the blood vessels by line segments, and then the joint probability is derived by combining the OD detection of U-net and probability bubbles from the intersection points of hough line segments. This in turn is used for OD center and radius estimation. Zhao *et al.* [130] proposed a model to decrease the computational load by combining transfer learning with U-Net segmentation architecture. First, the segmentation accuracy was boosted by utilizing attention gate learning as an intermediate between the encoder and decoder phase of the classical U-Net architecture. Then the algorithm was implemented with transfer learning, where the weights were initially trained on a renowned dataset before they learned from the fundus dataset. The scarcity of sufficient labeled images was tackled using the above-mentioned transfer of learning. This approach has shown a significant reduction in network inference time compared to its contemporaries.

Xiang *et al.* [131] presented their work on OD and OC segmentation, concentrating on improving the model performance over multiple datasets. This was achieved by

TABLE 10. Glaucoma diagnosis performance comparison.

References	Dataset	ACC	SE	SP	AUC	F1
[101]	Private-Tongren		0.961	0.939	0.981	
	Private -Tibet and Ningxia		0.956	0.941	0.983	
[103]	DRIONS-DB, HRF,	0.943	0.907	0.979	0.991	
	RIM-ONE, Drishti-GS1					
[104]	Private		0.940	0.860		
[105]	Private- Kasturba Medical College, Manipal	0.980	0.974	0.988		0.983
[106]	Private	0.953	0.96	0.939	0.994	
[107]	Private		0.980	0.910	0.995	
[108]	DRISHTI-GS, RIM-ONE	0.975	0.987	0.962		
[109]	Origa, Drishti-GS, iChallenge, RiM-ONE, RIGA	0.870	0.85		0.93	
[110]	CGSA	0.996	0.962	0.977		
[111]	ORIGA, HRF, OCT & CFI		0.717		0.874	
[112]	Samsung Medical Centre, Seoul.	0.96	0.95		0.99	0.97
[116]	G1020	0.992	0.98	0.97		
[117]	Private Dataset – Bangladesh eye hospital	0.965			0.957	0.951
[118]	G1020	0.98	0.951	0.94		
[119]	RIM-ONE	0.852	0.848	0.855	0.916	
[120]	ORIGA				0.92	
[121]	ORIGA				0.88	
[35]	LAG	0.953	0.954	0.952		
[122]	Private				0.945	
[110]	Private		0.962	0.977	0.996	
[123]	ORIGA	0.957	0.949	0.947	0.978	0.963

introducing a multi-scale weight shared attention (MSA) module after the encoder phase which enhances the OD/OC feature extraction process and a depth-wise separable convolution (DCS) module after the decoder phase that accurately concentrates on the target features. Model performance was tested for generalization across five fundus datasets, which achieved improved results compared to other state-of-the-art architectures. Jin *et al.* [132] proposed a method for OD segmentation that involves a DenseNet-based encoder for feature extraction, for dealing with small datasets. In the decoder stage, they used various layers of features to drive the attention process, and combine feature information from multiple scales for the upsampling process. At the end of their network, they combined the cross-entropy and the dice coefficient to generate an improved loss function for optimizing the model during the training. Bengani *et al.* [133] handled the problem of the lack of a large labeled dataset by using a two-tire approach for the OD segmentation task. First, a convolution autoencoder was trained on a large number of unlabelled fundus images, implementing semi-supervised learning. Later transfer learning was applied to the above-mentioned pre-trained model with OD ground truth images. Bhatkalkar *et al.* [134] proposed an encoder-decoder-based generic regression model for simultaneous segmentation of the fovea center and OD. For training of the model, central

coordinates of fovea and OD from ground truth images were transformed into heatmaps.

Nazir *et al.* [135] utilized EfficientNetB0 as the base model to develop a network for glaucoma detection using OD and OC lesions. First, the relevant features were extracted using the base network, then a unique bidirectional feature module was employed to fuse the key points multiple times. And finally, glaucoma localization was achieved along with class prediction. Xiong *et al.* [136] proposed OD segmentation by leveraging Hough transform annotations. They used a Bayesian U-Net based on weak labels for the segmentation. A probability-based graphical model was built and implemented on U-Net. The expectation-maximization method was used for OD estimation and subsequent weight assignment. Hervella *et al.* [137] demonstrated a multi-task architecture that simultaneously performs OD and OC segmentation along with glaucoma classification. Both image-level and pixel-level labels were utilized in the training process. The simultaneous classification and segmentation task increased the number of shared parameters.

C. AMD DIAGNOSIS

Age-related macular degeneration (AMD) is one of the leading causes of blindness among the elderly population [138]. AMD generally affects the macular region of the retina.

A study shows that by 2020 the number of patients with AMD has reached approximately 196 million at the global level and it is expected to reach 288 million by the year 2040 [139]. In the following section, we discuss various DL-based methods employed for the automatic diagnosis of AMD in recent times. Various publicly available datasets utilized for AMD diagnosis are presented in Table 2. Similarly, experimental results of recent research work on AMD diagnosis are shown in Table 12.

Chou *et al.* [140] utilized a stacking technique for combining a fundus image-based DL model with biomarkers derived from optical coherence tomography (OCT) for distinguishing neovascular AMD (nAMD) from polypoidal choroidal vasculopathy (PCV), as they both manifest similar image properties. A novel method called Multiple Correspondence Analysis (MCA) was introduced for converting OCT biomarkers into continuous principal components. EfficientNet-B3 was used for the training and validation of Fundus images. The ensemble stacking strategy yields the best mixture from the above two paths, for accurate predictions on new input images. Yan *et al.* [141] 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. A total of 31,262 fundus images from the AREDS dataset were used for the project. Inception-V3 CNN was used as a base model for the training process. The extracted features were fed to a fully connected layer for AMD severity classification. This severity mixed with 52 genetic variants was again fed into another FC layer for predicting the probability of late AMD if it exceeded the inquired years.

Xu *et al.* [142] proposed a dual deep CNN model which utilizes fundus and OCT image pairs for categorizing AMD and PCV. Transfer learning was employed by first utilizing the weights from ResNet-50 onto two individual models that separately take OCT images and fundus images as inputs. After refining the weights on new data, they were transferred to corresponding convolutional blocks. In the end, an FC layer was established that classified input pairs into Wet AMD, Dry AMD, PCV, and nAMD. Another work was proposed, based on drusen segmentation for AMD detection, by Pham *et al.* [143] where they tried to tackle the data imbalance problem, as the number of non-drusen pixels was very high compared to drusen pixels. Their model consists of two networks, one an Image level network that uses a Deeplabv3+ base architecture to generate drusen probability maps and a patch-level network that works on corresponding patch images and their probability maps for final prediction. The Patch level network employs U-Net-based segmentation. A total of 775 fundus images from Kangbuk Samsung hospital, were used for training. Model performance was also tested on the STARE dataset. Vaghefi *et al.* [144] introduced a multimodal approach for dry AMD diagnosis, where the DL model utilizes three modalities - fundus images, optical coherence tomography (OCT), and OCT angiography (OCT-A) for improving accuracy. Input samples from

75 individuals were collected and grouped into three categories. Inception-ResNet-V2 based CNN was used as a base model and further modifications were made to facilitate training on multiple modalities. It was demonstrated that higher accuracies can be achieved with DL models through the suitable utilization of images with different modalities.

Peng *et al.* [145] presented a two-step DL model framework for accurately estimating the risk of late AMD at the individual level. Initially, a classification network was implemented and trained over 80,000 manually annotated fundus images, collected from AREDS and AREDS2 datasets. The second part of the architecture, known as the survival model, was responsible for predicting late AMD progression probability depending on the grading results or on the extracted features of the previous section. This stage also avails an option for including genotype information. Their model achieved an accuracy of 0.864 when validated with an independent test dataset. In another work, Heo *et al.* [146] developed a CNN-based classification model, that uses VGG-16 as a base architecture for classifying dry AMD and Neovascular AMD. In the pre-processing phase, image cropping was performed (around the macula center) followed by border adjustments. Feature maps from the final CNN layer were derived and weights were computed by a class activation map, for generating a heatmap. González-Gonzalo *et al.* [147] collected 600 fundus images with AMD and DR from three different medical centers (Europe) and conducted performance validation of the RetCAD-DL model (commercially available DL system), for combined detection of AMD and DR from fundus images. The model was additionally assessed with AREDS and Messidor datasets to establish further validation. The RetCAD model executes joint detection by first taking RGB and Contrast-Enhanced (CE) images of the original fundus input image, and then utilizes two ensembles (each with three CNN) one for DR and one for AMD. These ensembles generate DR and AMD scores indicating the probability of referable DR and AMD.

Bridge *et al.* [148] developed a prognostic model that predicts the future progression of AMD based on multiple longitudinal images (whose timepoints are spaced unevenly). In stage one of this particular method, a CNN (Inception-V3) was used for generating feature vectors from the input image, in the next stage, the vectors were merged through interval scaling which compensates for uneven image time points. Finally, in the third stage, a recurrent unit provides the probability of AMD progression by employing the sigmoid activation function. Another DL method was proposed by Yoo *et al.* [149], which utilizes fundus images and OCT to diagnose AMD. This multimodal approach improved the diagnostic results compared to results obtained by using any one imaging modality. VGG-19, which was pre-trained on the ImageNet dataset was used for feature extraction from both OCT and fundus images; later a Random Forest (RF) model was operated on those features for final classification. Delong test was performed, which implies that the multimodal approach significantly boosted model performance.

TABLE 11. OD/OC segmentation performance comparison.

References	Dataset	OD/OC	ACC	SE	SP	AUC	IOU	DSC
[102]	DRISHTI-GS	OD	0.99	0.865	0.935			
[124]	MESSIDOR, ORIGA, REFUGE	OD		0.9744				0.9377
[125]	DRISHTI-GS	OD		-			0.932	0.987
[125]	DRISHTI-GS	OC		-			0.921	0.971
[126]	DRIONS-DB	OD					0.983	
[126]	RIM-ONE	OD					0.979	
[126]	IDRiD	OD					0.976	
[127]	RIM-ONEv2	OD		0.995	0.994	0.998	0.997	
[129]	Kaggle	OD				0.977		
[129]	MESSIDOR	OD				0.991		
[130]	DRISHTI-GS	OD	0.997	0.948	0.993			
[130]	DRISHTI-GS	OC	0.995	0.876	0.997			
[131]	DRISHTI-GS	OD					0.950	0.974
[131]	MESSIDOR	OD					0.943	0.970
[131]	DRISHTI-GS	OC					0.834	0.900
[133]	DRISHTI-GS	OD	0.995	0.953	0.999			0.967
[133]	RIM-ONE	OD	0.994	0.873	0.998			0.902
[137]	REFUGE	OD						0.958
[137]	DRISHTI-GS	OD						0.971
[137]	REFUGE	OC						0.882
[137]	DRISHTI-GS	OC						0.910

Building on their previous work, Chen *et al.* [150], utilized DeepSeeNet [151] for developing a classification network that classifies fundus images into AREDS prescribed Nine-step AMD severity scale. The initial stage contains 10 Inception-V3 blocks for feature extraction, second stage has an average pooling layer followed by a dense and a dropout layer. The third stage consists of four layers that run in parallel (multi-task) for detecting four major AMD characteristics that are later combined to map the image into a nine-step severity score. The training was performed using the AREDS dataset and the model was later evaluated on both AREDS and AREDS2.

Pham *et al.* [152] developed a framework for monitoring AMD disease progression from early-stage images by synthesizing future AMD images. They utilized a GAN network with two discriminators for producing realistic future fundus images. In another work, Yellapragada *et al.* [153] presented a method for training the model without labeled data. They first used a self-supervised NPID training technique on fundus images and then tested its performance using a classifier (supervised) for grading severity levels of AMD. Wu *et al.* [154] presented a comparative study on the performance of a model in predicting AMD progression by utilizing automatic OCT imaging biomarkers versus manually graded color fundus images. Govindaiah *et al.* [155] showed the overall performance of late AMD prediction models may

improve by adding genetic, clinical, and socio-demographic data while training the model.

D. CATARACT DIAGNOSIS

Cataracts are among the serious retinal diseases, if not identified and treated in time, and may lead to irreversible vision loss [156]. A recent study shows that nearly 33.6 million cases of blindness (45 percent of global blindness) are due to cataracts [157]. Recently many attempts have been made to automatically diagnose cataracts from fundus images. Recent developments in this direction are discussed below along with a performance comparison (Table 13) of various DL models.

For reducing training parameters and computational burden while training the model for cataract detection, Junayed *et al.* [158], adjusted the activation function and loss function of their CNN-based architecture. They also experimented with three different models which use various numbers of CNN blocks (3, 4, and 5 number of blocks) and tested them for detection accuracy. The model with 4 blocks presented optimal results without any overfitting. Imran *et al.* [159] proposed a cataract classification model (severe, moderate, mild, normal) by combining CNN with recurrent neural network (RNN). After the pre-processing stage, each fundus image from the dataset was subdivided into 12 patches and each of these patches was processed through pre-trained CNN models (GoogleNet, AlexNet,

TABLE 12. AMD diagnosis performance comparison.

References	Dataset	Task	Base Architecture	ACC	SE	SP	AUC	F1	Dice Score
[140]	Private	AMD and PCV differentiation	Efficient-Net-B3	0.836	0.807	0.847	0.885		
[141]	AREDS	Late AMD prediction	Inception-V3				0.85		
[142]	Private	AMD classification	ResNet-50	0.874	0.888	0.956		0.878	
[143]	Kangbuk Samsung Hospital	Drusen Segmentation for AMD	DeeplabV3+, U-Net	0.99	0.725	0.992			0.517
	STARE	Drusen Segmentation for AMD	DeeplabV3+, U-Net	0.975	0.317	0.992			0.322
[144]	Private	Dry AMD diagnosis	Inception-ResNet-V2	0.998	1.00	0.992			
[145]	AREDS, AREDS2	Predicting Risk of late AMD	-	0.864					
[146]	Private-Ulsan University Hospital	Classifying nAMD and dAMD	VGG-16	0.913	0.879	0.944			
[147]	Private DR-AMD	Detection of referable AMD	-		0.918	0.875	0.949		
[148]	AREDS	Detection of AMD Progression	Inception-V3		0.878	0.887	0.950		
[149]	Private	AMD Diagnosis	VGG-19+ RF	0.905	0.910	0.896	0.969		
[150]	AREDS	AMD Classification (AREDS 9-step score)	-	0.614				0.597	
[153]	AREDS	AMD Classification (Advanced AMD 2- Step)	ReSNet-50	0.958					

VGGNet, ResNet) for feature extraction. LSTM (bidirectional) based RNN was utilized to process feature vectors for cataract classification. To deal with noise-affected fundus images which are common due to image acquisition complexities, Pratap and Kokil [160] employed two independent DCNNs for a combined feature extraction (CFE) strategy. Pre-trained AlexNet was modified and used for implementing CFE. Support vector machine (SVM) classifiers (independently trained, at different noise levels) were used and features extracted in an earlier stage were fed into a specific classifier by considering the noise level in that particular image. In another work, Imran *et al.* [161] also combined an SVM classifier with several DL models for cataract classification. After pre-processing phase which involves image resizing, green channel extraction, and image normalization, the Transfer Learning mechanism was implemented with pre-trained ResNet, VGGNet, and AlexNet for the feature extraction stage. Next, the first fully connected layer was replaced by a pooling layer (global averaging), and individual SVMs were employed for the final four-level cataract classification.

Hossain *et al.* [162] proposed architecture, used ResNet-50 as their base module for the cataract detection task. They collected fundus images from various sources and utilized a total of 3048 cataract infected images and 2670 non-cataract images for their research. Zhang and He [163] used the stacking technique for grading cataracts into six different levels. This was achieved by employing ResNet-18 for extracting high-level features combined with manual extraction of texture features using GLCM (Gray level co-occurrence matrix).

These two feature sets were fed into two separate SVM Classifiers to learn the base level probabilities of each input image, followed by a fully connected neural network for generating the final cataract classification label.

To overcome some of the shortcomings of standard CNN architectures like overfitting, high computational burden, and fading gradient problem, Raza *et al.* [164] utilized transfer learning based on the Inception-V4 variant to classify cataracts disease fundus images. Similarly, Khan *et al.* [165] utilized a pre-trained VGG-19 for binary classification of cataract and non-cataract images. To emphasize local image features for the cataract classification task, Xu *et al.* [166] proposed an attention network that focuses on global features as well as local features before final grading.

E. ROP DIAGNOSIS

Retinopathy of prematurity (ROP) is a retinal disease that mainly affects the fundus vasculature of infants. With neo-vascularization, the effect of this disease may present severe consequences like retinal detachment and complete blindness among children. For timely treatment, it is important to identify an early symptom called plus disease-causing morphological changes to retinal blood vessels of preterm infants. Recent developments in DL implementation for ROP diagnosis and their experimental results are presented in Table 14.

Ramachandran *et al.* [167] introduced a framework for identifying ROP by detecting plus disease in infant fundus images. In their semi-supervised approach, the network

TABLE 13. Cataract detection/classification performance comparison.

References	Dataset	Task	Base Architecture	ACC	SE	SP	AUC	F1	PR
[158]	HRF, FIRE, ACHIKO-I, IDRiD, DRIVE	Detection	CNN	0.991		0.991		0.990	0.990
[159]	Private- Tongren Hospital, China	4 Class Classification	CNN+RNN	0.973	0.976	0.977	0.975		
[161]	Private- Tongren Hospital, China	4 Class Classification	CNN+SVM	0.956	0.956	0.960	0.974	0.956	0.965
[162]	Private, HRF, DRIVE, IDRiD, STARE	Detection	CNN (ResNet-50)	0.957	0.944	0.980	0.982		
[163]		6 Class Classification	CNN (ResNet-18)	0.929					
[166]	Private data	2 Class Classification	CNN	0.906					

generates bounding boxes around the twisted vessels and the count of these boxes indicates the presence of plus disease in the retinal image. This is achieved by employing a fully convolution neural network (inspired by YOLO architecture) for detecting the twisted vessels. A twofold training approach has been adopted, where the model is trained with manually labeled images to generate images with bounding boxes (pseudo labeled images), then both manually labeled and pseudo labeled images are used for retraining the model which is finally used for predicting ROP. For establishing ROP diagnosis along with an assisted medical follow-up mechanism, Agrawal *et al.* [168] developed a network that uses an ensemble of U-Net and hough transform techniques to detect various zones (I, II, III) in fundus images of infants (premature). These zones are used to indicate ROP severity and assist in scheduling the next screening. Two U-Nets are used for OD and retinal vessel segmentation tasks, from which the zones can be estimated.

Lei *et al.* [169] developed an ROP detection network that also generates supporting evidence about its decision. Initially, ResNet-50 (backbone network) was modified and improved by the addition of a Channel and Spatial Attention (CASA) Module, which extracts ROP lesion features and subsequently through a fully connected layer detects ROP in the fundus image. Parallely, Gradient weighted class activation mapping is implemented on the extracted features for visualizing the extracted features and locating the retinal structures that can explain the model decision. To deal with the relatively high obscurity of ROP features, compared to other retinal features in fundus images, Chen *et al.* [170] proposed a network that learns at multiple instances and classifies ROP into different stages. A fully convolutional network (FCN) is used for obtaining local features and producing a spatial score map of ROP lesions. These are again converted into patches to augment the dataset. A separate CNN network (Multiple instance learning networks) utilizes these patches for retraining to achieve even better performance.

Finally, through the attention pooling mechanism, ROP classification was performed. Yildiz *et al.* [171] utilized spatial attention maps, which bring out regions that are crucial for ROP classification. This helped CNN to concentrate more on disease-related regions and extract valuable features. As opposed to the standard attention learning framework, where the regions are learned using class labels, this method incorporated specific domain (Structural) information to improve the maps. The attention maps also helped in highlighting the specific areas of the image, which are utilized for model prediction. Huang *et al.* [172] tested five DL models for ROP classification by applying transfer learning. After applying suitable pre-processing, normalization, and augmentation techniques to the dataset, models were trained with hyperparameter tuning. Out of VGG-16, VGG-19, MobileNet, Inception-V3, and DenseNet backbone models, VGG-19 produced the most accurate results for ROP classification.

Coyner *et al.* [173] investigated the viability of utilizing synthetically generated fundus images for the diagnosis of plus disease in ROP. Generative adversarial networks (GANs) were utilized for fundus image 'synthesis'. Pix2pix pipeline method was implemented for realistic image generation from retinal vessel maps. This work indicates that GANs can be effectively utilized for dataset augmentation for improved model training in ROP classification networks. Tan *et al.* [174] utilized a private dataset consisting of 6974 fundus images to train an AI model for classifying normal and plus diseased images. Their algorithm also showed promise in detecting a comparatively less severe pre-disease stage known as a pre-plus disease from a fundus image. In another research, Redd *et al.* [175] assessed a DL severity screening system developed for ROP. The DL system was developed to generate a 1-9 scale score which indicates the severity based on retinal vascular abnormalities, which in turn was used to predict an overall ROP disease category. Wang and Chen [176] developed an automated system for

TABLE 14. ROP diagnosis performance comparison.

References	Dataset	Task	Base Architecture	ACC	PR	SE	SP	AUC	F1
[167]	Private-KIDROP	Plus, disease detection		0.99		0.99	0.98		0.98
[169]	Private	ROP Detection	ResNet-50	0.990		0.948	0.994	0.993	0.945
[170]	Private	ROP Classification	FCN	0.944		0.927	0.942	0.972	0.922
[171]	Private	ROP Classification	CNN	0.950				0.987	0.974
[172]	Private	ROP Classification	VGG-19	0.96		0.966	0.952	0.97	
[174]	Private	ROP Classification	-	0.973		0.966	0.98	0.993	
[41]	Private	Normal and ROP classification	ResNet-50	0.927	0.899				0.899
[41]		Mild-ROP and Severe-ROP classification	ResNet-50	0.736	0.791				0.811

identifying the existence of ROP in fundus images along with understanding its severity level. They utilized 6030 data samples for the training process where a median frequency balancing technique was employed to handle the data imbalance problem.

VI. RESEARCH DIRECTIONS

As discussed in the previous sections, retinal disease diagnosis using DL methods has progressed amazingly in terms of testing and evaluating various network architectures for retinal disease diagnosis. However, there is a lot of scope and unexplored areas open for future research. From the review conducted in this paper, we see the following directions for future research:

- Weakly supervised learning models:** Even though many fundus image datasets exist in the public domain, when compared to natural image datasets like ImageNet which has nearly 14 million images, the availability of labeled fundus images is quite limited. The available fundus datasets are also diverse in terms of their ground truth labeling. Although there are other techniques like image synthesis, that are parallelly investigated, which can generate artificial fundus images, researchers can explore weakly supervised learning models for training on original fundus images that have different ground truth labeling. Through weakly supervised training techniques robust model performances can be achieved for retinal disease diagnosis even with datasets that are partially labeled or inaccurately labeled.
- Fundus image synthesis:** Recent popularity of Generative Adversarial Networks (GANs) has shown potential in generating synthetic fundus images which can be used to augment the training dataset. This can effectively eliminate the lack of good quality labeled data and improve prediction performance. Although some of the recent research showed the synthesis of images for DR, glaucoma, and AMD, the field is still relatively new and presents ample scope for future research.
- Lightweight Network design:** Most of the DL models developed for retinal disease diagnosis perform well at the expense of high consumption of computational resources. This is a major hurdle in implementing such models on portable edge devices. Another open research direction in this field is developing novel lightweight models to reduce the computational parameters while maintaining performance.
- Improving generalization:** It is observed that due to the difference in image acquisition settings for different datasets, DL models showed varied performance across them, i.e., some models performed well on specific datasets, whereas they failed on others. Researchers can focus on improving model generalization performance by exploring various domain adaption techniques. These techniques essentially aim to minimize the distribution difference between target and source data domains. Several adaption methods are already explored, one method known as moment matching where distribution differences are minimized at the feature space level, another method utilizes adversarial learning to align both source and target domains. Considering the complexities in acquiring retinal fundus images, for improving model generalization the area of domain adaption presents ample future scope for the researchers.
- Implementing federated learning:** Due to various data privacy laws, most hospitals and other research institutes hesitate in sharing fundus images with others. This adds up to the data scarcity problem and restricts the model training to only available public datasets, depriving them of training on rich and diverse private fundus data available at the hospitals. Schemes like federated learning can be explored where the models can be trained on private data locally and then the learned weights are transferred to a global model.
- Multiple disease diagnosis:** Another promising research direction is simultaneously detecting multiple retinal diseases using DL. This can be helpful for

clinicians to identify patients with more than one retinal disease. Although studies have been carried out in this direction such as simultaneous ‘DME and DR diagnosis’, simultaneous ‘AMD, DR and glaucoma diagnosis’, etc. it is still a relatively less explored area.

- **Smartphone-based retinal disease diagnosis:** The majority of current work in this field utilizes fundus images captured through high-resolution fundoscopy. There is ample scope for researchers to develop models that can learn from fundus images captured through smartphones. This will help in developing a remote eye screening facility.
- **Generating evidence maps:** One of the major concerns of DL implementation for retinal disease diagnosis is acquiring the approval of professional doctors for the AI-based model. Very limited research has been carried out in the direction of making the predictions more interpretable. One possible solution for this could be, generating evidence maps for the predictions made by the DL model and showing or highlighting the crucial regions of the fundus image the deep network used to arrive at the final decision. Recently some approaches have shown progress in this direction but there is still vast scope for research in terms of improving the accuracy of these evidence maps. For example, DR diagnosis depends on finding various lesions and markers on fundus image, so one can perform accurate lesion segmentation and simultaneously grade DR to generate quality evidence maps.

VII. CONCLUSION

There is a pressing need for automated systems for identifying eye-related diseases, considering the lack of medical experts when compared with the number of patients. A color fundus image, which contains a wide variety of eye-related pathologies in image format, opened up a lot of scope for research in terms of medical image analysis. A wide range of DL models are being implemented and tested for automatic disease diagnosis. Sophisticated image processing techniques can now be utilized for bringing out salient features from a given fundus image. Lesions like microaneurysms, exudates, hemorrhages, etc. which constitute significantly a smaller number of pixels in a fundus image are now utilized to diagnose diseases from an early stage. This review presented a process-based approach for understanding the latest DL approaches in the ophthalmic disease diagnosis process.

As the success of a DL model highly depends on the training dataset, a consolidation of all publicly available fundus image datasets is presented along with their ground truth description. It is observed that many datasets like IDRiD, Messidor, DRIVE, etc. contain high-quality fundus images, which are captured in a controlled environment. The models trained on these datasets may not perform well on other datasets. On the other hand, datasets like Kaggle, and Eye-PACS among others, contain images captured in diverse environmental conditions. These may not be suitable for efficient

model training, but as they mimic the real-world scenario, they can steer the model behavior toward the practical side. A balanced combination of datasets may help in developing a robust model which can be implemented clinically.

Most of the studies have shown that the application of image pre-processing techniques like contrast enhancement, color space transformation, image augmentation, filtering, etc. can help the DL model to better extract disease-relevant features during the training process.

The work published in recent years has used various backbone models to build solutions for disease diagnosis. Basic CNN, VGG, ResNet, Inception, etc. are utilized for classification tasks, while networks like U-Net, FCNs, Mask RCNN, Seg-Net, etc. are implemented for segmentation tasks. In most of the studies, these backbone models have only served as a base structure. Many learning paradigms like ensemble learning, transfer learning, multitask learning active learning, etc. have also been explored to improve the model performance and provide an accurate diagnosis. Among all retinal diseases, Diabetic retinopathy has been widely studied and explored and its actual clinical implementation has been examined. The current primary research for DR is now directed toward providing interpretable heatmaps along with disease classification. Similarly, glaucoma diagnosis is also studied considerably but most of them focus on direct classification or diagnosis based on CDR estimation. Compared to these two diseases, much less attention has been paid to AMD and one of the reasons is the lack of large datasets for AMD diagnosis tasks. Diagnosis of Cataract and ROP offers plenty of scope for future researchers as relatively very few studies have been carried out in that direction.

The retinal diseases focused in this review are of crucial importance as a delay in their treatment may lead to complete vision loss. The interest in implementing DL for retinal disease diagnosis has grown significantly in past few years. In some cases, the performance of DL models has surpassed that of human experts. However, the future scope is still wide open concerning efficient and effective patient care since DL systems must evolve and be integrated into clinical practice.

REFERENCES

- [1] M. M. Fraz, P. Remagnino, A. Hoppe, B. Uyyayanovara, A. R. Rudnicka, C. G. Owen, and S. A. Barman, “Blood vessel segmentation methodologies in retinal images—A survey,” *Comput. Methods Programs Biomed.*, vol. 108, no. 1, pp. 407–433, Oct. 2012.
- [2] D. Chawla and A. Deorari, “Retinopathy of prematurity prevention, screening and treatment programmes: Progress in India,” in *Seminars in Perinatology*, vol. 43. Amsterdam, The Netherlands: Elsevier, 2019, pp. 344–347.
- [3] M. D. Abramoff, M. K. Garvin, and M. Sonka, “Retinal imaging and image analysis,” *IEEE Rev. Biomed. Eng.*, vol. 3, pp. 169–208, 2010.
- [4] Y. Guo, Y. Liu, A. Oerlemans, S. Lao, S. Wu, and M. S. Lew, “Deep learning for visual understanding: A review,” *Neurocomputing*, vol. 187, pp. 27–48, Apr. 2016.
- [5] S. Sengupta, A. Singh, H. A. Leopold, T. Gulati, and V. Lakshminarayanan, “Ophthalmic diagnosis using deep learning with fundus images—A critical review,” *Artif. Intell. Med.*, vol. 102, Jan. 2020, Art. no. 101758.
- [6] M. Badar, M. Haris, and A. Fatima, “Application of deep learning for retinal image analysis: A review,” *Comput. Sci. Rev.*, vol. 35, Feb. 2020, Art. no. 100203.

- [7] T. Li, W. Bo, C. Hu, H. Kang, H. Liu, K. Wang, and H. Fu, "Applications of deep learning in fundus images: A review," *Med. Image Anal.*, vol. 69, Apr. 2021, Art. no. 101971.
- [8] W. L. Alyoubi, M. F. Abulkhair, and W. M. Shalash, "Diabetic retinopathy fundus image classification and lesions localization system using deep learning," *Sensors*, vol. 21, no. 11, p. 3704, May 2021.
- [9] D. Mirzania, A. C. Thompson, and K. W. Muir, "Applications of deep learning in detection of glaucoma: A systematic review," *Eur. J. Ophthalmol.*, vol. 31, no. 4, pp. 1618–1642, Jul. 2021.
- [10] J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE Trans. Med. Imag.*, vol. 23, no. 4, pp. 501–509, Apr. 2004.
- [11] A. D. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Trans. Med. Imag.*, vol. 19, no. 3, pp. 203–210, Mar. 2000.
- [12] A. Budai, R. Bock, A. Maier, J. Hornegger, and G. Michelson, "Robust vessel segmentation in fundus images," *Int. J. Biomed. Imag.*, vol. 2013, pp. 1–11, Dec. 2013.
- [13] M. M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A. R. Rudnicka, C. G. Owen, and S. A. Barman, "An ensemble classification-based approach applied to retinal blood vessel segmentation," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 9, pp. 2538–2548, Sep. 2012.
- [14] T. Kauppi, V. Kalesnykiene, J.-K. Kamarainen, L. Lensu, I. Sorri, H. Uusitalo, H. Kälviäinen, and J. Pietilä, "DIARETDB0: Evaluation database and methodology for diabetic retinopathy algorithms," in *Machine Vision and Pattern Recognition Research Group*, vol. 73. Lappeenranta, Finland: Lappeenranta Univ. Technol., 2006, pp. 1–17.
- [15] R. Kälviäinen and H. Uusitalo, "Diaretdb1 diabetic retinopathy database and evaluation protocol," in *Medical Image Understanding and Analysis*. Aberystwyth, U.K.: Univ. Wales, Citeseer, 2007, p. 61.
- [16] B. Dastbozorg, J. Zhang, F. Huang, and B. M. ter Haar Romeny, "Retinal microaneurysms detection using local convergence index features," *IEEE Trans. Image Process.*, vol. 27, no. 7, pp. 3300–3315, Jul. 2018.
- [17] M. Niemeijer, B. Van Ginneken, M. J. Cree, A. Mizutani, G. Quellec, C. I. Sánchez, B. Zhang, R. Hornero, M. Lamard, and C. Muramatsu, "Retinopathy online challenge: Automatic detection of microaneurysms in digital color fundus photographs," *IEEE Trans. Med. Imag.*, vol. 29, no. 1, pp. 185–195, Jan. 2010.
- [18] E. Decencière, G. Cazuguel, X. Zhang, G. Thibault, J. C. Klein, F. Meyer, B. Marcotegui, G. Quellec, M. Lamard, and R. Danno, "TeleOphta: Machine learning and image processing methods for teleophthalmology," *Irbm*, vol. 34, no. 2, pp. 196–203, Apr. 2013.
- [19] E. Decencière, X. Zhang, G. Cazuguel, B. Lay, B. Cochener, C. Trone, P. Gain, R. Ordóñez, P. Massin, and A. Erginay, "Feedback on a publicly distributed image database: The messidor database," *Image Anal. Stereol.*, vol. 33, no. 3, pp. 231–234, 2014.
- [20] M. D. Abramoff, J. C. Folk, D. P. Han, J. D. Walker, D. F. Williams, S. R. Russell, P. Massin, B. Cochener, P. Gain, L. Tang, M. Lamard, D. C. Moga, G. Quellec, and M. Niemeijer, "Automated analysis of retinal images for detection of referable diabetic retinopathy," *JAMA Ophthalmol.*, vol. 131, no. 3, pp. 351–357, Mar. 2013.
- [21] S. Sivaprasad, G. Arden, A. T. Prevost, R. Crosby-Nwaobi, H. Holmes, J. Kelly, C. Murphy, G. Rubin, J. Vasconcelos, and P. Hykin, "A multi-centre phase III randomised controlled single-masked clinical trial evaluating the efficacy and safety of light-masks at preventing dark-adaptation in the treatment of early diabetic macular oedema (CLEOPATRA): Study protocol for a randomised controlled trial," *Trials*, vol. 15, no. 1, pp. 1–10, Dec. 2014.
- [22] *Diabetic Retinopathy Detection* | Kaggle. Accessed: Apr. 22, 2022. [Online]. Available: <https://www.kaggle.com/c/diabetic-retinopathy-detection>
- [23] P. Porwal, S. Pachade, R. Kamble, M. Kokare, G. Deshmukh, V. Sahasrabudde, and F. Meriaudeau, "Indian diabetic retinopathy image dataset (IDRiD): A database for diabetic retinopathy screening research," *Data*, vol. 3, no. 3, p. 25, 2018.
- [24] T. Li, Y. Gao, K. Wang, S. Guo, H. Liu, and H. Kang, "Diagnostic assessment of deep learning algorithms for diabetic retinopathy screening," *Inf. Sci.*, vol. 501, pp. 511–522, Oct. 2019.
- [25] J. Lowell, A. Hunter, D. Steel, A. Basu, R. Ryder, E. Fletcher, and L. Kennedy, "Optic nerve head segmentation," *IEEE Trans. Med. Imag.*, vol. 23, no. 2, pp. 256–264, Feb. 2004.
- [26] J. Sivaswamy, S. R. Krishnadass, G. D. Joshi, M. Jain, and A. U. S. Tabish, "Drishti-GS: Retinal image dataset for optic nerve head (ONH) segmentation," in *Proc. IEEE 11th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2014, pp. 53–56.
- [27] H. Fu, J. Cheng, Y. Xu, C. Zhang, D. W. K. Wong, J. Liu, and X. Cao, "Disc-aware ensemble network for glaucoma screening from fundus image," *IEEE Trans. Med. Imag.*, vol. 37, no. 11, pp. 2493–2501, Nov. 2018.
- [28] J. I. Orlando, H. Fu, J. B. Breda, K. van Keer, D. R. Bathula, A. Diaz-Pinto, R. Fang, P.-A. Heng, J. Kim, and J. Lee, "REFUGE challenge: A unified framework for evaluating automated methods for glaucoma assessment from fundus photographs," *Med. Image Anal.*, vol. 59, Jan. 2020, Art. no. 101570.
- [29] Y. Zheng, C.-Y. Cheng, E. L. Lamoureaux, P. P. Chiang, A. R. Anuar, J. J. Wang, P. Mitchell, S.-M. Saw, and T. Y. Wong, "How much eye care services do Asian populations need? Projection from the Singapore epidemiology of eye disease (SEED) study," *Investigative Ophthalmol. Vis. Sci.*, vol. 54, no. 3, pp. 2171–2177, 2013.
- [30] E. J. Carmona, M. Rincón, J. García-Feijóo, and J. M. Martínez-de-la-Casa, "Identification of the optic nerve head with genetic algorithms," *Artif. Intell. Med.*, vol. 43, no. 3, pp. 243–259, 2008.
- [31] Z. Zhang, F. S. Yin, J. Liu, W. K. Wong, N. M. Tan, B. H. Lee, J. Cheng, and T. Y. Wong, "ORIGA-light: An online retinal fundus image database for glaucoma analysis and research," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol.*, Aug. 2010, pp. 3065–3068.
- [32] A. Almazroa, S. Alothayb, E. Osman, E. Ramadan, M. Hummadi, M. Dlain, M. Alkateer, K. Raahemifar, and V. Lakshminarayanan, "Retinal fundus images for glaucoma analysis: The RIGA dataset," in *Proc. Med. Imag., Imag. Informat. Healthcare, Res., Appl.*, vol. 10579, 2018, Art. no. 105790B.
- [33] F. Fumero, S. Alayón, J. L. Sanchez, J. Sigut, and M. Gonzalez-Hernandez, "RIM-ONE: An open retinal image database for optic nerve evaluation," in *Proc. 24th Int. Symp. Comput. Med. Syst. (CBMS)*, Jun. 2011, pp. 1–6.
- [34] Z. Zhang, J. Liu, F. Yin, B.-H. Lee, D. W. K. Wong, and K. R. Sung, "ACHIKO-K: Database of fundus images from glaucoma patients," in *Proc. IEEE 8th Conf. Ind. Electron. Appl. (ICIEA)*, Jun. 2013, pp. 228–231.
- [35] L. Li, M. Xu, H. Liu, Y. Li, X. Wang, L. Jiang, Z. Wang, X. Fan, and N. Wang, "A large-scale database and a CNN model for attention-based glaucoma detection," *IEEE Trans. Med. Imag.*, vol. 39, no. 2, pp. 413–424, Feb. 2020.
- [36] M. Baskaran, R. C. Foo, C.-Y. Cheng, A. K. Narayanaswamy, Y.-F. Zheng, R. Wu, S.-M. Saw, P. J. Foster, T.-Y. Wong, and T. Aung, "The prevalence and types of glaucoma in an urban Chinese population: The Singapore Chinese eye study," *JAMA Ophthalmol.*, vol. 133, no. 8, pp. 874–880, 2015.
- [37] Age-Related Eye Disease Study Research Group, "The age-related eye disease study (AREDS): Design implications AREDS report no. 1," *Controlled Clin. Trials*, vol. 20, no. 6, p. 573, 1999.
- [38] H. Fu, F. Li, J. I. Orlando, H. Bogunovic, X. Sun, J. Liao, Y. Xu, S. Zhang, and X. Zhang, "Adam: Automatic detection challenge on age-related macular degeneration [data set]," in *Proc. IEEE DataPort*, 2020. [Online]. Available: <https://doi.org/10.48550/arXiv.2202.07983>, doi: 10.48550/arXiv.2202.07983.
- [39] C. Brandl, V. Breinlich, K. J. Stark, S. Enzinger, M. Aßenmacher, M. Olden, F. Grassmann, J. Graw, M. Heier, A. Peters, H. Helbig, H. Küchenhoff, B. H. F. Weber, and I. M. Heid, "Features of age-related macular degeneration in the general adults and their dependency on age, sex, and smoking: Results from the German KORA study," *PLoS ONE*, vol. 11, no. 11, Nov. 2016, Art. no. e0167181.
- [40] V. Mayya, S. Kamath, and U. Kulkarni, "Automated microaneurysms detection for early diagnosis of diabetic retinopathy: A comprehensive review," *Comput. Methods Programs Biomed. Update*, vol. 1, Jan. 2021, Art. no. 100013.
- [41] Y. Wang and Y. Chen, "Automated recognition of retinopathy of prematurity with deep neural networks," *J. Phys., Conf.*, vol. 1187, no. 4, Apr. 2019, Art. no. 042057.
- [42] C. Lam, D. Yi, M. Guo, and T. Lindsey, "Automated detection of diabetic retinopathy using deep learning," *AMIA Summits Transl. Sci.*, vol. 2018, no. 1, p. 147, 2018.
- [43] B. Wu, W. Zhu, F. Shi, S. Zhu, and X. Chen, "Automatic detection of microaneurysms in retinal fundus images," *Comput. Med. Imag. Graph.*, vol. 55, pp. 106–112, Jan. 2017.

- [44] C. Swathi, B. K. Anoop, D. A. S. Dhas, and S. P. Sanker, "Comparison of different image preprocessing methods used for retinal fundus images," in *Proc. Conf. Emerg. Devices Smart Syst. (ICEDSS)*, Mar. 2017, pp. 175–179.
- [45] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998.
- [46] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*. [Online]. Available: <https://doi.org/10.48550/arXiv.1409.1556>, doi: 10.48550/arXiv.1409.1556.
- [47] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [48] J. Long, E. Shelhamer, and T. Darrell, "Fully convolutional networks for semantic segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 3431–3440.
- [49] O. Ronneberger, P. Fischer, and T. Brox, "U-Net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput. Assist. Intervent.* Cham, Switzerland: Springer, 2015, pp. 234–241.
- [50] A. J. Jenkins, M. V. Joglekar, A. A. Hardikar, A. C. Keech, D. N. O'Neal, and A. S. Januszewski, "Biomarkers in diabetic retinopathy," *Rev. Diabetic Stud.*, vol. 12, nos. 1–2, p. 159, 2015.
- [51] 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, "Global prevalence and major risk factors of diabetic retinopathy," *Diabetes Care*, vol. 35, no. 3, pp. 556–564, 2012.
- [52] 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 England J. Med.*, vol. 362, no. 12, pp. 1090–1101, 2010.
- [53] X. Wang, M. Xu, J. Zhang, L. Jiang, L. Li, M. He, N. Wang, H. Liu, and Z. Wang, "Joint learning of multi-level tasks for diabetic retinopathy grading on low-resolution fundus images," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 5, pp. 2216–2227, May 2022.
- [54] F. Li, Y. Wang, T. Xu, L. Dong, L. Yan, M. Jiang, X. Zhang, H. Jiang, Z. Wu, and H. Zou, "Deep learning-based automated detection for diabetic retinopathy and diabetic macular oedema in retinal fundus photographs," *Eye*, pp. 1–9, Jul. 2021. [Online]. Available: <https://doi.org/10.1038/s41433-021-01552-8>, doi: 10.1038/s41433-021-01552-8.
- [55] H. Kaushik, D. Singh, M. Kaur, H. Alshazly, A. Zaguia, and H. Hamam, "Diabetic retinopathy diagnosis from fundus images using stacked generalization of deep models," *IEEE Access*, vol. 9, pp. 108276–108292, 2021.
- [56] S. Das, K. Kharbanda, M. Suchetha, R. Raman, and E. Dhas, "Deep learning architecture based on segmented fundus image features for classification of diabetic retinopathy," *Biomed. Signal Process. Control*, vol. 68, Jul. 2021, Art. no. 102600.
- [57] K. Shankar, A. R. W. Sait, D. Gupta, S. K. Lakshmanaprabu, A. Khanna, and H. M. Pandey, "Automated detection and classification of fundus diabetic retinopathy images using synergic deep learning model," *Pattern Recognit. Lett.*, vol. 133, pp. 210–216, May 2020.
- [58] K. Shankar, Y. Zhang, Y. Liu, L. Wu, and C.-H. Chen, "Hyperparameter tuning deep learning for diabetic retinopathy fundus image classification," *IEEE Access*, vol. 8, pp. 118164–118173, 2020.
- [59] L. Qiao, Y. Zhu, and H. Zhou, "Diabetic retinopathy detection using prognosis of microaneurysm and early diagnosis system for non-proliferative diabetic retinopathy based on deep learning algorithms," *IEEE Access*, vol. 8, pp. 104292–104302, 2020.
- [60] J. Wang, Y. Bai, and B. Xia, "Simultaneous diagnosis of severity and features of diabetic retinopathy in fundus photography using deep learning," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 12, pp. 3397–3407, Dec. 2020.
- [61] T. Araújo, G. Aresta, L. Mendonça, S. Penas, C. Maia, Â. Carneiro, A. M. Mendonça, and A. Campilho, "DRIGRADUATE: Uncertainty-aware deep learning-based diabetic retinopathy grading in eye fundus images," *Med. Image Anal.*, vol. 63, Jul. 2020, Art. no. 101715.
- [62] J. Sahlsten, J. Jaskari, J. Kivinen, L. Turunen, E. Jaanio, K. Hietala, and K. Kaski, "Deep learning fundus image analysis for diabetic retinopathy and macular edema grading," *Sci. Rep.*, vol. 9, no. 1, pp. 1–11, Dec. 2019.
- [63] S. Qummar, F. G. Khan, S. Shah, A. Khan, S. Shamshirband, Z. U. Rehman, I. A. Khan, and W. Jadoon, "A deep learning ensemble approach for diabetic retinopathy detection," *IEEE Access*, vol. 7, pp. 150530–150539, 2019.
- [64] G. U. Nneji, J. Cai, J. Deng, H. N. Monday, M. A. Hossin, and S. Nahar, "Identification of diabetic retinopathy using weighted fusion deep learning based on dual-channel fundus scans," *Diagnostics*, vol. 12, no. 2, p. 540, Feb. 2022.
- [65] A. Bora, S. Balasubramanian, B. Babenko, S. Virmani, S. Venugopalan, A. Mitani, G. de Oliveira Marinho, J. Cuadros, P. Ruamviboonsuk, G. S. Corrado, L. Peng, D. R. Webster, A. V. Varadarajan, N. Hammel, Y. Liu, and P. Bavishi, "Predicting the risk of developing diabetic retinopathy using deep learning," *Lancet Digit. Health*, vol. 3, no. 1, pp. e10–e19, Jan. 2021.
- [66] S. Majumder and N. Kehtarnavaz, "Multitasking deep learning model for detection of five stages of diabetic retinopathy," *IEEE Access*, vol. 9, pp. 123220–123230, 2021.
- [67] X. Zhang, D. Li, Q. Wei, X. Han, B. Zhang, H. Chen, Y. Zhang, B. Mo, B. Hu, and D. Ding, "Automated detection of severe diabetic retinopathy using deep learning method," *Graefes Arch. Clin. Experim. Ophthalmol.*, vol. 260, pp. 1–8, Mar. 2021.
- [68] M. Toğaçar, "Detection of retinopathy disease using morphological gradient and segmentation approaches in fundus images," *Comput. Methods Programs Biomed.*, vol. 214, Feb. 2022, Art. no. 106579.
- [69] H. Xia, Y. Lan, S. Song, and H. Li, "A multi-scale segmentation-to-classification network for tiny microaneurysm detection in fundus images," *Knowl.-Based Syst.*, vol. 226, Aug. 2021, Art. no. 107140.
- [70] Y. Liao, H. Xia, S. Song, and H. Li, "Microaneurysm detection in fundus images based on a novel end-to-end convolutional neural network," *Biocybernetics Biomed. Eng.*, vol. 41, no. 2, pp. 589–604, Apr. 2021.
- [71] X. Zhang, J. Wu, M. Meng, Y. Sun, and W. Sun, "Feature-transfer network and local background suppression for microaneurysm detection," *Mach. Vis. Appl.*, vol. 32, no. 1, pp. 1–13, Jan. 2021.
- [72] J. Xue, S. Yan, J. Qu, F. Qi, C. Qiu, H. Zhang, M. Chen, T. Liu, D. Li, and X. Liu, "Deep membrane systems for multitask segmentation in diabetic retinopathy," *Knowl.-Based Syst.*, vol. 183, Nov. 2019, Art. no. 104887.
- [73] S. Guo, T. Li, H. Kang, N. Li, Y. Zhang, and K. Wang, "L-Seg: An end-to-end unified framework for multi-lesion segmentation of fundus images," *Neurocomputing*, vol. 349, pp. 52–63, Jul. 2019.
- [74] C. Huang, Y. Zong, Y. Ding, X. Luo, K. Clawson, and Y. Peng, "A new deep learning approach for the retinal hard exudates detection based on superpixel multi-feature extraction and patch-based CNN," *Neurocomputing*, vol. 452, pp. 521–533, Sep. 2021.
- [75] W. He, X. Wang, L. Wang, Y. Huang, Z. Yang, X. Yao, X. Zhao, L. Ju, L. Wu, L. Wu, H. Lu, and Z. Ge, "Incremental learning for exudate and hemorrhage segmentation on fundus images," *Inf. Fusion*, vol. 73, pp. 157–164, Sep. 2021.
- [76] V. Kurilová, J. Goga, M. Oravec, J. Pavlovičová, and S. Kajan, "Support vector machine and deep-learning object detection for localisation of hard exudates," *Sci. Rep.*, vol. 11, no. 1, pp. 1–9, Dec. 2021.
- [77] Q. Liu, H. Liu, Y. Zhao, and Y. Liang, "Dual-branch network with dual-sampling modulated dice loss for hard exudate segmentation in color fundus images," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 3, pp. 1091–1102, Mar. 2022.
- [78] P. Khojasteh, B. Aliahmad, and D. K. Kumar, "A novel color space of fundus images for automatic exudates detection," *Biomed. Signal Process. Control*, vol. 49, pp. 240–249, Mar. 2019.
- [79] C. Kou, W. Li, Z. Yu, and L. Yuan, "An enhanced residual U-Net for microaneurysms and exudates segmentation in fundus images," *IEEE Access*, vol. 8, pp. 185514–185525, 2020.
- [80] Y. Zong, J. Chen, L. Yang, S. Tao, C. Aoma, J. Zhao, and S. Wang, "U-net based method for automatic hard exudates segmentation in fundus images using inception module and residual connection," *IEEE Access*, vol. 8, pp. 167225–167235, 2020.
- [81] N. J. Mohan, R. Murugan, T. Goel, and P. Roy, "Fast and robust exudate detection in retinal fundus images using extreme learning machine autoencoders and modified KAZE features," *J. Digit. Imag.*, pp. 1–18, Feb. 2022. [Online]. Available: <https://doi.org/10.1007/s10278-022-00587-x>, doi: 10.1007/s10278-022-00587-x.
- [82] Q. Liu, H. Liu, Y. Zhao, and Y. Liang, "Dual-branch network with dual-sampling modulated dice loss for hard exudate segmentation in color fundus images," *IEEE J. Biomed. Health Informat.*, vol. 26, no. 3, pp. 1091–1102, Mar. 2022.
- [83] N. J. Mohan, R. Murugan, T. Goel, and P. Roy, "Exudate localization in retinal fundus images using modified speeded up robust features algorithm," in *Proc. IEEE-EMBS Conf. Biomed. Eng. Sci. (IECBES)*, Mar. 2021, pp. 367–371.
- [84] S. Maqsood, R. Damaševičius, and R. Maskeliūnas, "Hemorrhage detection based on 3D CNN deep learning framework and feature fusion for evaluating retinal abnormality in diabetic patients," *Sensors*, vol. 21, no. 11, p. 3865, Jun. 2021.

- [85] S. Lahmiri, "Hybrid deep learning convolutional neural networks and optimal nonlinear support vector machine to detect presence of hemorrhage in retina," *Biomed. Signal Process. Control*, vol. 60, Jul. 2020, Art. no. 101978.
- [86] L. Yang, H. Wang, Q. Zeng, Y. Liu, and G. Bian, "A hybrid deep segmentation network for fundus vessels via deep-learning framework," *Neurocomputing*, vol. 448, pp. 168–178, Aug. 2021.
- [87] H. Boudegga, Y. Elloumi, M. Akil, M. H. Bedoui, R. Kachouri, and A. B. Abdallah, "Fast and efficient retinal blood vessel segmentation method based on deep learning network," *Computerized Med. Imag. Graph.*, vol. 90, Jun. 2021, Art. no. 101902.
- [88] K. Fukutsu, M. Saito, K. Noda, M. Murata, S. Kase, R. Shiba, N. Isogai, Y. Asano, N. Hanawa, M. Dohke, M. Kase, and S. Ishida, "A deep learning architecture for vascular area measurement in fundus images," *Ophthalmol. Sci.*, vol. 1, no. 1, Mar. 2021, Art. no. 100004.
- [89] I. Atli and O. S. Gedik, "Sine-Net: A fully convolutional deep learning architecture for retinal blood vessel segmentation," *Eng. Sci. Technol., Int. J.*, vol. 24, no. 2, pp. 271–283, Apr. 2021.
- [90] M. E. Gegundez-Arias, D. Marin-Santos, I. Perez-Borrero, and M. J. Vasallo-Vazquez, "A new deep learning method for blood vessel segmentation in retinal images based on convolutional kernels and modified U-Net model," *Comput. Methods Programs Biomed.*, vol. 205, Jun. 2021, Art. no. 106081.
- [91] P. M. Samuel and T. Veeramalai, "VSSC Net: Vessel specific skip chain convolutional network for blood vessel segmentation," *Comput. Methods Programs Biomed.*, vol. 198, Jan. 2021, Art. no. 105769.
- [92] D. Chen, Y. Ao, and S. Liu, "Semi-supervised learning method of U-Net deep learning network for blood vessel segmentation in retinal images," *Symmetry*, vol. 12, no. 7, p. 1067, Jun. 2020.
- [93] C. Tian, T. Fang, Y. Fan, and W. Wu, "Multi-path convolutional neural network in fundus segmentation of blood vessels," *Biocybernetics Biomed. Eng.*, vol. 40, no. 2, pp. 583–595, Apr. 2020.
- [94] D. Wang, A. Haytham, J. Pottenburgh, O. Saecedi, and Y. Tao, "Hard attention net for automatic retinal vessel segmentation," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 12, pp. 3384–3396, Dec. 2020.
- [95] R. Biswas, A. Vasan, and S. S. Roy, "Dilated deep neural network for segmentation of retinal blood vessels in fundus images," *Iranian J. Sci. Technol., Trans. Electr. Eng.*, vol. 44, no. 1, pp. 505–518, Mar. 2020.
- [96] B. Wang, S. Wang, S. Qiu, W. Wei, H. Wang, and H. He, "CSU-Net: A context spatial U-Net for accurate blood vessel segmentation in fundus images," *IEEE J. Biomed. Health Informat.*, vol. 25, no. 4, pp. 1128–1138, Apr. 2021.
- [97] Y. Wu, Y. Xia, Y. Song, Y. Zhang, and W. Cai, "NFN+: A novel network followed network for retinal vessel segmentation," *Neural Netw.*, vol. 126, pp. 153–162, Jun. 2020.
- [98] P. Xiuqin, Q. Zhang, H. Zhang, and S. Li, "A fundus retinal vessels segmentation scheme based on the improved deep learning U-Net model," *IEEE Access*, vol. 7, pp. 122634–122643, 2019.
- [99] D. A. Dharmawan, D. Li, B. P. Ng, and S. Rahardja, "A new hybrid algorithm for retinal vessels segmentation on fundus images," *IEEE Access*, vol. 7, pp. 41885–41896, 2019.
- [100] R. R. A. Bourne, G. A. Stevens, R. A. White, J. L. Smith, S. R. Flaxman, H. Price, J. B. Jonas, J. Keeffe, K. Leasher, K. Naidoo, K. Pesudovs, S. Resnikoff, and H. R. Taylor, "Causes of vision loss worldwide, 1990–2010: A systematic analysis," *Lancet Global Health*, vol. 1, no. 6, pp. e339–e349, Dec. 2013.
- [101] 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, and N. Wang, "A hierarchical deep learning approach with transparency and interpretability based on small samples for glaucoma diagnosis," *NPJ Digit. Med.*, vol. 4, no. 1, pp. 1–11, Dec. 2021.
- [102] S. P., R. J., and P. R., "An automatic recognition of glaucoma in fundus images using deep learning and random forest classifier," *Appl. Soft Comput.*, vol. 109, Sep. 2021, Art. no. 107512.
- [103] P. Wang, M. Yuan, Y. He, and J. Sun, "3D augmented fundus images for identifying glaucoma via transferred convolutional neural networks," *Int. Ophthalmol.*, vol. 41, no. 6, pp. 2065–2072, Jun. 2021.
- [104] S. Gheisari, S. Shariflou, J. Phu, P. J. Kennedy, A. Agar, M. Kalloniatis, and S. M. Golzan, "A combined convolutional and recurrent neural network for enhanced glaucoma detection," *Sci. Rep.*, vol. 11, no. 1, pp. 1–11, Dec. 2021.
- [105] D. R. Nayak, D. Das, B. Majhi, S. V. Bhandary, and U. R. Acharya, "ECNet: An evolutionary convolutional network for automated glaucoma detection using fundus images," *Biomed. Signal Process. Control*, vol. 67, May 2021, Art. no. 102559.
- [106] F. Li, L. Yan, Y. Wang, J. Shi, H. Chen, X. Zhang, M. Jiang, Z. Wu, and K. Zhou, "Deep learning-based automated detection of glaucomatous optic neuropathy on color fundus photographs," *Graefes Arch. Clin. Experim. Ophthalmol.*, vol. 258, no. 4, pp. 851–867, Apr. 2020.
- [107] R. Hemelings, B. Elen, J. Barbosa-Breda, S. Lemmens, M. Meire, S. Pourjavan, E. Vandewalle, S. Van De Veire, M. B. Blaschko, P. De Boever, and I. Stalmans, "Accurate prediction of glaucoma from colour fundus images with a convolutional neural network that relies on active and transfer learning," *Acta Ophthalmol.*, vol. 98, no. 1, pp. e94–e100, Feb. 2020.
- [108] M. Juneja, N. Thakur, S. Thakur, A. Uniyal, A. Wani, and P. Jindal, "GC-NET for classification of glaucoma in the retinal fundus image," *Mach. Vis. Appl.*, vol. 31, no. 5, pp. 1–18, Jul. 2020.
- [109] J. Martins, J. S. Cardoso, and F. Soares, "Offline computer-aided diagnosis for glaucoma detection using fundus images targeted at mobile devices," *Comput. Methods Programs Biomed.*, vol. 192, Aug. 2020, Art. no. 105341.
- [110] 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 deep learning system to detect glaucomatous optic neuropathy using fundus photographs," *JAMA Ophthalmol.*, vol. 137, no. 12, pp. 1353–1360, 2019.
- [111] M. N. Bajwa, M. I. Malik, S. A. Siddiqui, A. Dengel, F. Shafait, W. Neumeier, and S. Ahmed, "Two-stage framework for optic disc localization and glaucoma classification in retinal fundus images using deep learning," *BMC Med. Informat. Decis. Making*, vol. 19, no. 1, pp. 1–16, Dec. 2019.
- [112] M. Kim, J. C. Han, S. H. Hyun, O. Janssens, S. Van Hoecke, C. Kee, and W. De Neve, "Medinoid: Computer-aided diagnosis and localization of glaucoma using deep learning," *Appl. Sci.*, vol. 9, no. 15, p. 3064, Jul. 2019.
- [113] L. K. Singh, H. Garg, and M. Khanna, "Deep learning system applicability for rapid glaucoma prediction from fundus images across various data sets," *Evolving Syst.*, pp. 1–30, Feb. 2022. [Online]. Available: <https://doi.org/10.1007/s12530-022-09426-4>, doi: 10.1007/s12530-022-09426-4.
- [114] S. Ovreiu, E.-A. Paraschiv, and E. Ovreiu, "Deep learning & digital fundus images: Glaucoma detection using DenseNet," in *Proc. 13th Int. Conf. Electron., Comput. Artif. Intell. (ECAI)*, Jul. 2021, pp. 1–4.
- [115] V. Saravanan, R. D. J. Samuel, S. Krishnamoorthy, and A. Manickam, "Deep learning assisted convolutional auto-encoders framework for glaucoma detection and anterior visual pathway recognition from retinal 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.
- [116] A. Shoukat, S. Akbar, S. A. E. Hassan, A. Rehman, and N. Ayesha, "An automated deep learning approach to diagnose glaucoma using retinal fundus images," in *Proc. Int. Conf. Frontiers Inf. Technol. (FIT)*, Dec. 2021, pp. 120–125.
- [117] M. Tanvir Islam, S. T. Mashfu, A. Faisal, S. Chowdhury Siam, I. Tahmid Naheen, and R. Khan, "Deep learning-based glaucoma detection with cropped optic cup and disc and blood vessel segmentation," *IEEE Access*, vol. 10, pp. 2828–2841, 2022.
- [118] A. Shoukat, S. Akbar, and K. Safdar, "A deep learning-based automatic method for early detection of the glaucoma using fundus images," in *Proc. Int. Conf. Innov. Comput. (ICIC)*, Nov. 2021, pp. 1–6.
- [119] L. Li, M. Xu, X. Wang, L. Jiang, and H. Liu, "Attention based glaucoma detection: A large-scale database and CNN model," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2019, pp. 10571–10580.
- [120] R. Zhao, W. Liao, B. Zou, Z. Chen, and S. Li, "Weakly-supervised simultaneous evidence identification and segmentation for automated glaucoma diagnosis," in *Proc. AAAI Conf. Artif. Intell.*, vol. 33, Jul. 2019, pp. 809–816.
- [121] W. Liao, B. Zou, R. Zhao, Y. Chen, Z. He, and M. Zhou, "Clinical interpretable deep learning model for glaucoma diagnosis," *IEEE J. Biomed. Health Informat.*, vol. 24, no. 5, pp. 1405–1412, May 2019.
- [122] S. Phene, R. C. Dunn, N. Hammel, Y. Liu, J. Krause, N. Kitade, M. Schaeckermann, R. Sayres, D. J. Wu, A. Bora, C. Semturs, A. Misra, A. E. Huang, A. Spitz, F. A. Medeiros, A. Y. Maa, M. Gandhi, G. S. Corrado, L. Peng, and D. R. Webster, "Deep learning and glaucoma specialists: The relative importance of optic disc features to predict glaucoma referral in fundus photographs," *Ophthalmology*, vol. 126, no. 12, pp. 1627–1639, 2019.

- [123] J. Latif, S. Tu, C. Xiao, S. U. Rehman, A. Imran, and Y. Latif, "ODGNet: A deep learning model for automated optic disc localization and glaucoma classification using fundus images," *Social Netw. Appl. Sci.*, vol. 4, no. 4, pp. 1–11, Apr. 2022.
- [124] L. Wang, J. Gu, Y. Chen, Y. Liang, W. Zhang, J. Pu, and H. Chen, "Automated segmentation of the optic disc from fundus images using an asymmetric deep learning network," *Pattern Recognit.*, vol. 112, Apr. 2021, Art. no. 107810.
- [125] H. N. Veena, A. Muruganandham, and T. S. Kumaran, "A novel optic disc and optic cup segmentation technique to diagnose glaucoma using deep learning convolutional neural network over retinal fundus images," *J. King Saud Univ. Comput. Inf. Sci.*, vol. 32, no. 2, pp. 1–15, Feb. 2021. [Online]. Available: <https://doi.org/10.1016/j.jksuci.2021.02.003>.
- [126] E. S. Kumar and C. S. Bindu, "Two-stage framework for optic disc segmentation and estimation of cup-to-disc ratio using deep learning technique," *J. Ambient Intell. Humanized Comput.*, pp. 1–13, Mar. 2021. [Online]. Available: <https://doi.org/10.1007/s12652-021-02977-5>, doi: 10.1007/s12652-021-02977-5.
- [127] D. Natarajan, E. Sankaralingam, K. Balraj, and S. Karuppusamy, "A deep learning framework for glaucoma detection based on robust optic disc segmentation and transfer learning," *Int. J. Imag. Syst. Technol.*, vol. 32, no. 1, pp. 230–250, Jan. 2022.
- [128] R. Panda, N. B. Puhan, B. Mandal, and G. Panda, "GlaucoNet: Patch-based residual deep learning network for optic disc and cup segmentation towards glaucoma assessment," *Social Netw. Comput. Sci.*, vol. 2, no. 2, pp. 1–17, Apr. 2021.
- [129] Y. Fu, J. Chen, J. Li, D. Pan, X. Yue, and Y. Zhu, "Optic disc segmentation by U-net and probability bubble in abnormal fundus images," *Pattern Recognit.*, vol. 117, Sep. 2021, Art. no. 107971.
- [130] X. Zhao, S. Wang, J. Zhao, H. Wei, M. Xiao, and N. Ta, "Application of an attention U-Net incorporating transfer learning for optic disc and cup segmentation," *Signal, Image Video Process.*, vol. 15, no. 5, pp. 913–921, Jul. 2021.
- [131] Q. Zhu, X. Chen, Q. Meng, J. Song, G. Luo, M. Wang, F. Shi, Z. Chen, D. Xiang, and L. Pan, "GDSCSeg-Net: General optic disc and cup segmentation network for multi-device fundus images," *Biomed. Opt. Exp.*, vol. 12, no. 10, pp. 6529–6544, 2021.
- [132] B. Jin, P. Liu, P. Wang, L. Shi, and J. Zhao, "Optic disc segmentation using attention-based U-Net and the improved cross-entropy convolutional neural network," *Entropy*, vol. 22, no. 8, p. 844, Jul. 2020.
- [133] S. Bengani, A. A. J. J., and V. S., "Automatic segmentation of optic disc in retinal fundus images using semi-supervised deep learning," *Multimedia Tools Appl.*, vol. 80, no. 3, pp. 3443–3468, Jan. 2021.
- [134] B. J. Bhatkalkar, S. V. Nayak, S. V. Shenoy, and R. V. Arjunan, "Fundus-PosNet: A deep learning driven heatmap regression model for the joint localization of optic disc and fovea centers in color fundus images," *IEEE Access*, vol. 9, pp. 159071–159080, 2021.
- [135] M. Nawaz, T. Nazir, A. Javed, U. Tariq, H.-S. Yong, M. A. Khan, and J. Cha, "An efficient deep learning approach to automatic glaucoma detection using optic disc and optic cup localization," *Sensors*, vol. 22, no. 2, p. 434, Jan. 2022.
- [136] H. Xiong, S. Liu, R. V. Sharan, E. Coiera, and S. Berkovsky, "Weak label based Bayesian U-Net for optic disc segmentation in fundus images," *Artif. Intell. Med.*, vol. 126, Apr. 2022, Art. no. 102261.
- [137] Á. S. Hervella, J. Rouco, J. Novo, and M. Ortega, "End-to-end multi-task learning for simultaneous optic disc and cup segmentation and glaucoma classification in eye fundus images," *Appl. Soft Comput.*, vol. 116, Feb. 2022, Art. no. 108347.
- [138] L. S. Lim, P. Mitchell, J. M. Seddon, F. G. Holz, and T. Y. Wong, "Age-related macular degeneration," *Lancet*, vol. 392, no. 10153, pp. 1147–1159, 2018.
- [139] W. L. Wong, X. Su, X. Li, C. M. G. Cheung, R. Klein, C.-Y. Cheng, and T. Y. Wong, "Global prevalence of age-related macular degeneration and disease burden projection for 2020 and 2040: A systematic review and meta-analysis," *Lancet Global Health*, vol. 2, no. 2, pp. e106–e116, Feb. 2014.
- [140] Y.-B. Chou, C.-H. Hsu, W.-S. Chen, S.-J. Chen, D.-K. Hwang, Y.-M. Huang, A.-F. Li, and H. H.-S. Lu, "Deep learning and ensemble stacking technique for differentiating polypoidal choroidal vasculopathy from neovascular age-related macular degeneration," *Sci. Rep.*, vol. 11, no. 1, pp. 1–9, Dec. 2021.
- [141] 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.
- [142] Z. Xu, W. Wang, J. Yang, J. Zhao, D. Ding, F. He, D. Chen, Z. Yang, X. Li, W. Yu, and Y. Chen, "Automated diagnoses of age-related macular degeneration and polypoidal choroidal vasculopathy using bi-modal deep convolutional neural networks," *Brit. J. Ophthalmol.*, vol. 105, no. 4, pp. 561–566, Apr. 2021.
- [143] Q. T. M. Pham, S. Ahn, S. J. Song, and J. Shin, "Automatic drusen segmentation for age-related macular degeneration in fundus images using deep learning," *Electronics*, vol. 9, no. 10, p. 1617, Oct. 2020.
- [144] E. Vaghefi, S. Hill, H. M. Kersten, and D. Squirrel, "Multimodal retinal image analysis via deep learning for the diagnosis of intermediate dry age-related macular degeneration: A feasibility study," *J. Ophthalmol.*, vol. 2020, pp. 1–7, Jan. 2020.
- [145] Y. Peng, T. D. Keenan, Q. Chen, E. Agrón, A. Allot, W. T. Wong, E. Y. Chew, and Z. Lu, "Predicting risk of late age-related macular degeneration using deep learning," *NPJ Digit. Med.*, vol. 3, no. 1, pp. 1–10, Dec. 2020.
- [146] T.-Y. Heo, K. M. Kim, H. K. Min, S. M. Gu, J. H. Kim, J. Yun, and J. K. Min, "Development of a deep-learning-based artificial intelligence tool for differential diagnosis between dry and neovascular age-related macular degeneration," *Diagnostics*, vol. 10, no. 5, p. 261, Apr. 2020.
- [147] C. González-Gonzalo, V. Sánchez-Gutiérrez, P. Hernández-Martínez, I. Contreras, Y. T. Lechanteur, A. Domanian, B. Ginneken, and C. I. Sánchez, "Evaluation of a deep learning system for the joint automated detection of diabetic retinopathy and age-related macular degeneration," *Acta Ophthalmol.*, vol. 98, no. 4, pp. 368–377, Jun. 2020.
- [148] J. Bridge, S. Harding, and Y. Zheng, "Development and validation of a novel prognostic model for predicting AMD progression using longitudinal fundus images," *BMJ Open Ophthalmol.*, vol. 5, no. 1, Oct. 2020, Art. no. e000569.
- [149] T. K. Yoo, J. Y. Choi, J. G. Seo, B. Ramasubramanian, S. Selvaperumal, and D. W. Kim, "The possibility of the combination of OCT and fundus images for improving the diagnostic accuracy of deep learning for age-related macular degeneration: A preliminary experiment," *Med. Biol. Eng. Comput.*, vol. 57, no. 3, pp. 677–687, Mar. 2019.
- [150] Q. Chen, Y. Peng, T. Keenan, S. Dharssi, and E. Agro, "A multi-task deep learning model for the classification of age-related macular degeneration," *AMIA Summits Translational Sci. Proc.*, vol. 2019, p. 505, Jun. 2019.
- [151] Y. Peng, S. Dharssi, Q. Chen, T. D. Keenan, E. Agrón, W. T. Wong, E. Y. Chew, and Z. Lu, "DeepSeeNet: A deep learning model for automated classification of patient-based age-related macular degeneration severity from color fundus photographs," *Ophthalmology*, vol. 126, no. 4, pp. 565–575, Apr. 2019.
- [152] Q. T. M. Pham, S. Ahn, J. Shin, and S. J. Song, "Generating future fundus images for early age-related macular degeneration based on generative adversarial networks," *Comput. Methods Programs Biomed.*, vol. 216, Apr. 2022, Art. no. 106648.
- [153] B. Yellapragada, S. Hornauer, K. Snyder, S. Yu, and G. Yiu, "Self-supervised feature learning and phenotyping for assessing age-related macular degeneration using retinal fundus images," *Ophthalmol. Retina*, vol. 6, no. 2, pp. 116–129, Feb. 2022.
- [154] Z. Wu, H. Bogunović, R. Asgari, U. Schmidt-Erfurth, and R. H. Guymer, "Predicting progression of age-related macular degeneration using OCT and fundus photography," *Ophthalmol. Retina*, vol. 5, no. 2, pp. 118–125, Feb. 2021.
- [155] A. Govindaiah, A. Baten, R. T. Smith, S. Balasubramanian, and A. Bhuiyan, "Optimized prediction models from fundus imaging and genetics for late age-related macular degeneration," *J. Personalized Med.*, vol. 11, no. 11, p. 1127, Nov. 2021.
- [156] R. Isaacs, J. Ram, and D. Apple, "Cataract blindness in the developing world: Is there a solution?" *J. Agromedicine*, vol. 9, no. 2, pp. 207–220, 2004.
- [157] S. Pawar, "Causes of blindness and vision impairment in 2020 and trends over 30 years, and prevalence of avoidable blindness in relation to VISION 2020: The right to sight: An analysis for the global burden of disease study," *SSRN Electron. J.*, vol. 9, no. 2, pp. e144–e160, 2021.
- [158] M. S. Junayed, M. B. Islam, A. Sadeghzadeh, and S. Rahman, "Cataract-Net: An automated cataract detection system using deep learning for fundus images," *IEEE Access*, vol. 9, pp. 128799–128808, 2021.
- [159] A. Imran, J. Li, Y. Pei, F. Akhtar, T. Mahmood, and L. Zhang, "Fundus image-based cataract classification using a hybrid convolutional and recurrent neural network," *Vis. Comput.*, vol. 37, no. 8, pp. 2407–2417, Aug. 2021.
- [160] T. Pratap and P. Kokil, "Deep neural network based robust computer-aided cataract diagnosis system using fundus retinal images," *Biomed. Signal Process. Control*, vol. 70, Sep. 2021, Art. no. 102985.

- [161] A. Imran, J. Li, Y. Pei, F. Akhtar, J.-J. Yang, and Y. Dang, "Automated identification of cataract severity using retinal fundus images," *Comput. Methods Biomechanics Biomed. Eng., Imag. Vis.*, vol. 8, no. 6, pp. 691–698, Nov. 2020.
- [162] M. R. Hossain, S. Afroze, N. Siddique, and M. M. Hoque, "Automatic detection of eye cataract using deep convolution neural networks (DCNNs)," in *Proc. IEEE Region Symp. (TENSYP)*, Jun. 2020, pp. 1333–1338.
- [163] H. Zhang, K. Niu, Y. Xiong, W. Yang, Z. He, and H. Song, "Automatic cataract grading methods based on deep learning," *Comput. Methods Programs Biomed.*, vol. 182, Dec. 2019, Art. no. 104978.
- [164] A. Raza, M. U. Khan, Z. Saeed, S. Samer, A. Mobeen, and A. Samer, "Classification of eye diseases and detection of cataract using digital fundus imaging (DFI) and inception-V4 deep learning model," in *Proc. Int. Conf. Frontiers Inf. Technol. (FIT)*, Dec. 2021, pp. 137–142.
- [165] M. S. Mahmud Khan, M. Ahmed, R. Z. Rasel, and M. Monirujjaman Khan, "Cataract detection using convolutional neural network with VGG-19 model," in *Proc. IEEE World AI IoT Congr. (AllIoT)*, May 2021, pp. 0209–0212.
- [166] X. Xu, J. Li, Y. Guan, L. Zhao, Q. Zhao, L. Zhang, and L. Li, "GLA-Net: A global-local attention network for automatic cataract classification," *J. Biomed. Informat.*, vol. 124, Dec. 2021, Art. no. 103939.
- [167] S. Ramachandran, P. Niyas, A. Vinekar, and R. John, "A deep learning framework for the detection of plus disease in retinal fundus images of preterm infants," *Biocybernetics Biomed. Eng.*, vol. 41, no. 2, pp. 362–375, Apr. 2021.
- [168] R. Agrawal, S. Kulkarni, R. Walambe, and K. Kotecha, "Assistive framework for automatic detection of all the zones in retinopathy of prematurity using deep learning," *J. Digit. Imag.*, vol. 34, no. 4, pp. 932–947, Aug. 2021.
- [169] B. Lei, X. Zeng, S. Huang, R. Zhang, G. Chen, J. Zhao, T. Wang, J. Wang, and G. Zhang, "Automated detection of retinopathy of prematurity by deep attention network," *Multimedia Tools Appl.*, vol. 80, no. 30, pp. 36341–36360, Dec. 2021.
- [170] S. Chen, R. Zhang, G. Chen, J. Zhao, T. Wang, G. Zhang, and B. Lei, "Attention-guided deep multi-instance learning for staging retinopathy of prematurity," in *Proc. IEEE 18th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2021, pp. 1025–1028.
- [171] V. Yildiz, S. Ioannidis, I. Yildiz, P. Tian, J. P. Campbell, S. Ostmo, J. Kalpathy-Cramer, M. F. Chiang, D. Erdogmus, and J. Dy, "Structural visual guidance attention networks in retinopathy of prematurity," in *Proc. IEEE 18th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2021, pp. 353–357.
- [172] Y.-P. Huang, S. Vadloori, H.-C. Chu, E. Y.-C. Kang, W.-C. Wu, S. Kusaka, and Y. Fukushima, "Deep learning models for automated diagnosis of retinopathy of prematurity in preterm infants," *Electronics*, vol. 9, no. 9, p. 1444, Sep. 2020.
- [173] A. S. Coyner, J. Chen, J. P. Campbell, S. Ostmo, P. Singh, J. Kalpathy-Cramer, and M. F. Chiang, "Diagnosability of synthetic retinal fundus images for plus disease detection in retinopathy of prematurity," in *Proc. AMIA Annu. Symp.*, 2020, p. 329.
- [174] Z. Tan, S. Simkin, C. Lai, and S. Dai, "Deep learning algorithm for automated diagnosis of retinopathy of prematurity plus disease," *Transl. Vis. Sci. Technol.*, vol. 8, no. 6, p. 23, Dec. 2019.
- [175] T. K. Redd, J. P. Campbell, J. M. Brown, S. J. Kim, S. Ostmo, R. V. P. Chan, J. Dy, D. Erdogmus, S. Ioannidis, J. Kalpathy-Cramer, and M. F. Chiang, "Evaluation of a deep learning image assessment system for detecting severe retinopathy of prematurity," *Brit. J. Ophthalmol.*, vol. 103, no. 5, pp. 580–584, May 2019.
- [176] Y. Wang and Y. Chen, "Automated recognition of retinopathy of prematurity with deep neural networks," *J. Phys., Conf.*, vol. 1187, no. 4, Apr. 2019, Art. no. 042057.



BALLA GOUTAM received the B.Tech. degree in applied electronics and instrumentation engineering from the Biju Patnaik University of Technology (BPUT), Odisha, and the M.Tech. degree in electronics and control engineering from SRM University, Chennai. He is currently pursuing the Ph.D. degree with the National Institute of Technology (NIT), Warangal, under the supervision of Dr. Md. Farukh Hashmi. His primary research interests include computer vision, deep learning, medical image analysis, and disease diagnosis.



Dr. Avinash G. Keskar. He is currently working as an Assistant Professor with the Department of Electronics and Communication Engineering, National Institute of Technology (NIT), Warangal. He has published up to 75 articles including 25 SCI indexed research papers in international/national journals/conferences of publishers like IEEE, Elsevier, and Springer. He has published one patent to his credit. He was a principal investigator of one research project of worth five lakhs funded by the Institute Seed Grant through TEQIP III. He has a teaching and research experience of 13 years. He has supervised two Ph.D. scholars. He is presently guiding four Ph.D. scholars. His current research interests include computer vision, machine vision, machine learning, deep learning, embedded systems, the Internet of Things, digital signal processing, image processing, and digital IC design. He is a Life Member of IETE, ISTE, and IAENG societies. He is also serving as an Active and Potential Technical Reviewer for IEEE Access, *IET Image Processing*, *IET Computer Vision*, *Wireless Personal Communications*, *IEEE SYSTEMS JOURNAL*, *Sensors (MDPI)*, *Electronics (MDPI)*, *Diagnostic (MDPI)*, *The Visual Computer*, *Applied Soft Computing (Elsevier)*, *Color Research and Application*, *Journal of Supercomputing*, and various other journals like Elsevier/Springer/IEEE TRANSACTIONS publishers of repute.



ZONG WOO GEEM (Senior Member, IEEE) received the B.Eng. degree from Chung-Ang University, the M.Sc. degree from Johns Hopkins University, and the Ph.D. degree from Korea University. He researched at Virginia Tech, the University of Maryland-College Park, and Johns Hopkins University. He is currently an Associate Professor with the College of IT Convergence, Gachon University, South Korea. He has invented a music-inspired optimization algorithm, harmony search, which has been applied to various scientific and engineering problems. His research interests include phenomenon-mimicking algorithms and their sustainable applications to energy, environment, water, and infrastructure. He has served for various journals as an Editor (an Associate Editor for *Engineering Optimization* and a Guest Editor for *Swarm and Evolutionary Computation*, *International Journal of Bio-Inspired Computation*, *Journal of Applied Mathematics*, *Applied Sciences*, *Complexity*, and *Sustainability*).



NEERAJ DHANRAJ BOKDE received the M.E. degree in embedded systems from the EEE Department, BITS Pilani, Pilani Campus, India, and the Ph.D. degree in data science from the Visvesvaraya National Institute of Technology, Nagpur, India. Then, he has worked as a Postdoctoral Researcher at different departments with Aarhus University, Aarhus, Denmark, where he is currently an Assistant Professor with the Center for Quantitative Genetics and Genomics. His major research interests include the domain of data science topics, focused majorly on data science, machine learning, time series analysis, software package development, and prediction applications in different domains (his detailed biography and research contributions are available at <https://neerajbokde.in/>).