

Received 25 June 2024, accepted 9 July 2024, date of publication 12 July 2024, date of current version 22 July 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3427394

TOPICAL REVIEW

A Systematic Review on Fundus Image-Based Diabetic Retinopathy Detection and Grading: Current Status and Future Directions

AMNA IKRAM^{1,2}, AZHAR IMRAN^{1,2}, (Senior Member, IEEE),
JIANQIANG LI¹, (Senior Member, IEEE), ABDULAZIZ ALZUBAIDI³,
SAFA FAHIM², AMANULLAH YASIN², AND HANAA FATHI^{4,5}

¹School of Software Engineering, Beijing University of Technology, Beijing 100124, China

²Department of Creative Technologies, Air University, Islamabad 44000, Pakistan

³Computer Science Department, College of Computing in Al-Qunfudhah, Umm Al-Qura University, Makkah 28821, Saudi Arabia

⁴Applied Science Research Center, Applied Science Private University, Amman 11937, Jordan

⁵MEU Research Unit, Middle East University, Amman 11831, Jordan

Corresponding author: Azhar Imran (azhar.imran@mail.au.edu.pk)

This work was supported by the National Key Research and Development Program of China under Project 2020YFB2104402.

ABSTRACT Diabetic Retinopathy (DR) is a prevalent outcome of diabetic mellitus. This causes lesions to form on the retina, impairing eyesight. Most likely, blindness can be avoided if the DR condition is discovered at an initial stage. Since DR is a non-reversible condition, early detection and treatment can significantly reduce the chance of visual loss. Fundus images manually detect DR, which is a laborious and error-prone procedure. In assessing and categorizing medical images, machine learning and deep learning have emerged as the most efficient methods, surpassing human performance, common image processing methods, and other computer-aided detection systems. For this article, the most recent approaches for utilizing fundus images to classify and detect DR using machine learning and deep learning methods have been researched and evaluated. The freely accessible DR Datasets consisting of fundus images have also been discussed. We reviewed several DR pipeline components, including the datasets that researchers frequently used and the preprocessing and data augmentation steps, feature extraction methods, commonly used detection and classification algorithms, and the generally used performance metrics. This paper ends with a discussion of current challenges that have to be tackled by researchers working in this field to translate the research methodology into actual clinical practice. Finally, we conclude with a discussion of the future perspectives of DR.

INDEX TERMS Diabetic retinopathy, machine learning, fundus images, computer-aided diagnosis, retinal diseases.

I. INTRODUCTION

The eye is the human body's most essential and complex organ [1]. It assists us in visualizing objects in a light colour and depth perception. It consists of the following parts (Fig 1).

- Sclera: The eye's sclera, or white component, serves as the exterior covering and is a robust protective layer.
- Cornea: The cornea is the transparent front portion of the sclera. Through the cornea, light enters the eye.

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

- Iris: The iris is a ring-shaped black muscular tissue structure below the cornea. The eye's colour may be determined by looking at the iris. Changing the iris, the iris also contributes to regulating or adjusting exposure.
- Pupil: Pupil refers to the tiny opening of the iris. The iris aids in controlling its size. It controls the amount of light that reaches the eye.
- Lens: The lens is a transparent object behind the pupil. It transforms into a different shape to concentrate light on the retina through the action of ciliary muscles. When focusing on nearer items, it gets thicker and gets thinner as it gets farther away.

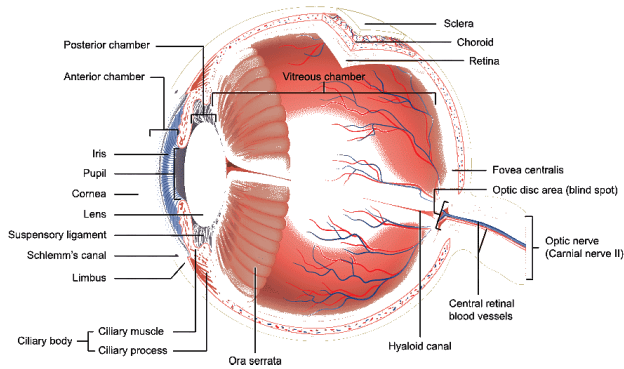


FIGURE 1. Eye anatomy.

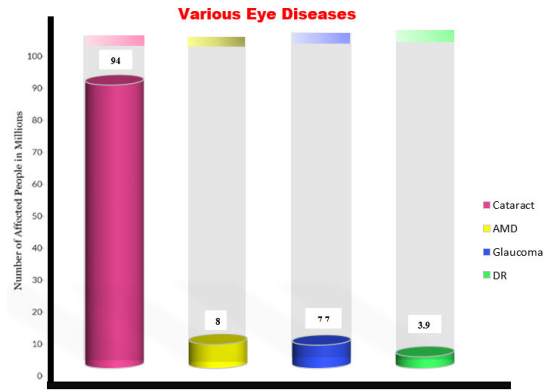


FIGURE 2. Graphical representation of rise in retinal diseases in 2022.

- Retina: This layer has numerous nerve cells, which are also light-sensitive. It alters the lens’s generated photos into electrical impulses. These electrical impulses are subsequently sent to the brain via the optic nerves.
- Optic Nerve: Optic nerves consist of cones and rods.
 - 1) **Cones:** Cones are the nerve cells more sensitive to bright light. They help in detailed central and colour vision.
 - 2) **Rods:** Rods are the optic nerve cells that are more sensitive to dim lights. They help in peripheral vision.

Any disorder in the eye can cause serious effects, from visual impairment to blindness. Although there are numerous eye disorders, the four most prevalent ones that result in vision loss or blindness are

- Age-related Macular Degeneration (AMD)
- Glaucoma
- Cataract
- Diabetic Retinopathy

Macular degeneration (age-related macular degeneration) is a visual disease that can harm central vision. The macula, the middle portion of the retina that enables us to see tiny details, is hurt by it. It is projected to be the primary cause of blindness in people over 60. Glaucoma is a visual disorder brought on by excessive eye fluid pressure. The pressure harms the optic nerve, which alters the visual data transmitted to the brain. Lens clouding in the eye is known as a cataract. One or both eyes may develop this cloudy lens.

A. DIABETIC RETINOPATHY AND ITS EPIDEMIOLOGY

Diabetic retinopathy, a severe complication of diabetes mellitus, results from prolonged periods of uncontrolled high blood sugar levels that progressively damage the retinal blood vessels. This condition, if left untreated, can lead to significant visual impairment and even blindness. As highlighted in Table 1, the symptoms and causes of diabetic retinopathy, alongside other retinal diseases, underscore the critical need for early detection and intervention. Fig 3 visually represents the impact of these diseases.

According to the World Health Organization (WHO) report from 2022, approximately 2.2 billion people globally suffer

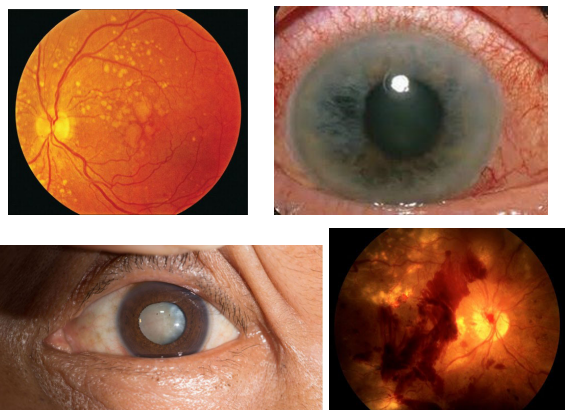


FIGURE 3. The appearance of various retinal diseases (a) Age-related macular degeneration (b) Glaucoma (c) Cataract (d) Diabetic Retinopathy.

from some form of visual impairment, whether at a near or distant range (Fig.2). Alarmingly, at least one billion of these cases involve vision loss that could have been prevented or remains untreated. Among these individuals, 94 million experience moderate to severe distance vision impairment or blindness due to untreated cataracts, 8 million are affected by age-related macular degeneration, 7.7 million suffer from glaucoma, and 3.9 million have diabetic retinopathy.

The prevalence of diabetic retinopathy, affecting nearly 4 million people worldwide, highlights a significant public health challenge. Given the increasing global prevalence of diabetes, the burden of diabetic retinopathy is expected to rise, exacerbating healthcare systems and economic resources. This condition not only diminishes the quality of life for millions but also imposes substantial costs on healthcare systems and economies worldwide. Therefore, understanding the epidemiology of diabetic retinopathy is crucial for developing effective prevention and treatment strategies.

This review aims to synthesize current knowledge on diabetic retinopathy, to inform healthcare professionals, policymakers, and researchers about the latest advancements and gaps in the field. By enhancing our understanding and response to diabetic retinopathy, we can mitigate its impact, improve patient outcomes, and ultimately contribute to reducing the global burden of visual impairment.

TABLE 1. Symptoms and effects of various eye diseases.

Eye Diseases	Symptoms	Effects
Age-related Macular Degeneration (AMD)	<ol style="list-style-type: none"> 1) Visual distortions, such as straight lines seeming bent. 2) Reduced central vision in one or both eyes. 3) The need for brighter light when reading or doing close-up work. 4) Increased difficulty adapting to low light levels, such as when entering a dimly lit restaurant or theatre. 5) Increased blurriness of printed words. 	<ul style="list-style-type: none"> • AMD affects the central vision and the ability to see fine details. • In AMD, a part of the retina called the macula is damaged. In advanced stages, people lose their ability to drive, see faces, and read smaller print.
Glaucoma	<ol style="list-style-type: none"> 1) Intense eye pain 2) Nausea and vomiting. 3) A red eye. 4) A headache. 5) Tenderness around the eyes. 6) Seeing rings around lights. 7) Blurred vision. 	<ol style="list-style-type: none"> 1) Glaucoma is a group of retinal diseases that can cause vision loss and blindness by damaging a nerve in the back of the eye called the optic nerve. 2) The symptoms can start so slowly that they go unnoticed. The only way to find out is the comprehensive dilated eye exam.
Cataract	<ol style="list-style-type: none"> 1) Vision is cloudy or blurry. 2) Colors look faded. 3) We can't see well at night. 4) Lamps, sunlight, or headlights seem too bright. 5) We see a halo around the lights. 6) We see double (this sometimes goes away as the cataract gets bigger) 	<ol style="list-style-type: none"> 1) Vision may be clouded, blurred, or dim vision. It increases difficulty with vision at night. 2) It causes sensitivity to light and glare. Need brighter light for reading and other activities.
Diabetic Retinopathy (DR)	<ol style="list-style-type: none"> 1) Gradually worsening vision. 2) Sudden vision loss. 3) Shapes floating in a field of vision (floaters) 4) Blurred or patchy vision. 5) Eye pain or redness. 6) Difficulty seeing in the dark. 	<ol style="list-style-type: none"> 1) Diabetic retinopathy can cause abnormal blood vessels to grow out of the retina and block fluid from draining out of the eye. 2) This causes glaucoma (a group of retinal diseases that can cause vision loss and blindness).

Among other retinal diseases, the most significant cause of blindness worldwide is diabetic retinopathy. It is that condition of the eye that diabetics experience when their excessive blood sugar levels harm the blood vessels in their retina. These blood vessels can expand and leak, or they can block the flow of blood. Abnormally new blood vessels on the retina can occasionally form and cause visual loss. The WHO estimates that the number of DR patients is increasing exponentially and will reach approximately 439 million by the year 2030 [2]. Blindness can result from diabetic retinopathy (DR), an optical symptom of diabetes. Diabetes patients have more chances of DR in between 40 and 45% of cases. [3]. Non-proliferative diabetic retinopathy (NPDR) and proliferative diabetic retinopathy (PDR) are the primary stages of diabetic retinopathy. A proliferative retinal disease causes the retina to develop unusual blood vessels. Tiny red dots are present in the retina at the starting stage of the condition, known as a non-proliferative retinal disease (NPDR). (Fig 4). These tiny spots could be microaneurysms (MA) or unusual pouching of blood vessels, which would be hemorrhage (HM). These blood vessel linings are exposed to damage, allowing fluid and fatty substances called exudates (EX) to flow out. Yellow lesions called hard exudates (HE) result from plasma leakage. They span the retina's outer

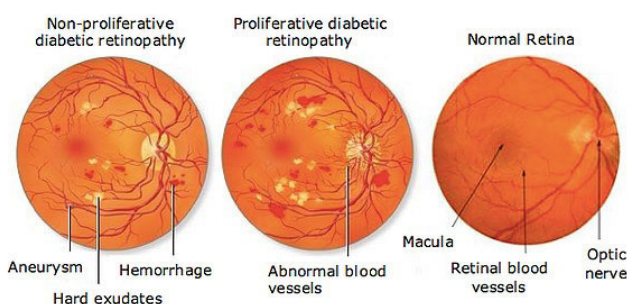


FIGURE 4. Retinal images (a) NDPR fundus image with Microaneurysms, Haemorrhage, and Exudate (b) PDR fundus image with abnormal blood vessels (c) Normal Retina.

layers and feature sharp edges. Due to swelling of the nerve fibres, Soft Exudates (SE) appear on the retina as white ovals. NPDR is further subdivided into mild, moderate, and severe NPDR (Fig 5); refer to Table 2 for their details.

To prevent blindness, the foveal avascular zone (FAZ), the optic disc (OD), and neovascularization (NV) are additional markers that can be utilized to identify and characterize DR at the starting stage. The lack of blood vessels and the inner retinal tissue that covers it is believed to improve the

TABLE 2. Brief description of stages of DR.

Stages of DR	Description
Mild NDPR Moderate NDPR Severe NDPR	Microaneurysms, i.e., small swellings in the retina’s tiny blood vessels, will be formed in this stage. As the disease progresses, some blood vessels that nourish the retina are blocked. Many more blood vessels are blocked, depriving several areas of the retina of their blood supply. The affected areas of the retina begin to show signs of ischemia (lack of oxygen), such as blot hemorrhages, bleeding of the veins, and intra-retinal microvascular abnormalities.
Proliferative retinopathy	At this advanced stage, the vaso-proliferative factors produced by the retina begin to trigger the growth of new blood vessels. These new blood vessels are abnormal and fragile

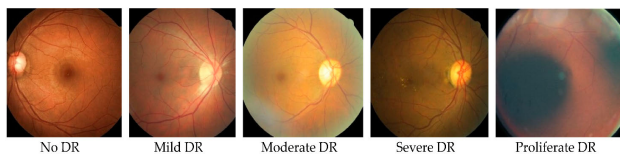


FIGURE 5. Different retinal images representing the severity levels of DR.

fovea pit’s optical quality by minimizing light scattering. The foveal avascular zone refers to this core avascular area. A vast, unstable, dispersed microaneurysm foveal avascular zone indicates diabetic retinopathy. The Optic Disc is the brightest, homogenous, circular structure in a usual eye fundus image and appears yellowish [4]. Identifying any anomaly in the OD’s area, size, shape, or structure will indicate the early changes leading to vision loss. Because of the severe oxygen deficiency in red blood cells (RBC), there is less blood flow to the eye’s tissues, which leads to the growth of new, unstable blood vessels and a dense OD. Neovascularization, the collective term for these new blood vessels, poses a major threat to vision [5]. The intimidations of NPDR can be minimized if it is detected and treated at early stages. Manually identifying DR patients is a very time-consuming and challenging process that is always subject to ophthalmologists’ judgment [5]. Further, with the demand for screening services, manual diagnosis for many DR patients is insufficient, and manual screening often leads to significant inconsistencies among retinal specialists [6]. Around seventy-five percent of DR patients come from low-income nations, where they don’t have the infrastructure and retinal specialists needed for DR screening [7] Therefore, automated methods of detection of DR are important to solve these problems with manual diagnosis and help ophthalmologists.

B. DR DETECTION IMAGING MODALITIES

There are various imaging modalities used for DR detection (see, (Fig 6)), which includes

- Slit lamp Images
- Ultrasonic Images
- Digital Images
- Fundus Images

A microscopic slit lamp with strong light is used during an eye exam. It allows ophthalmologists to look closely at the several structures in the eye and on the front of the eye.

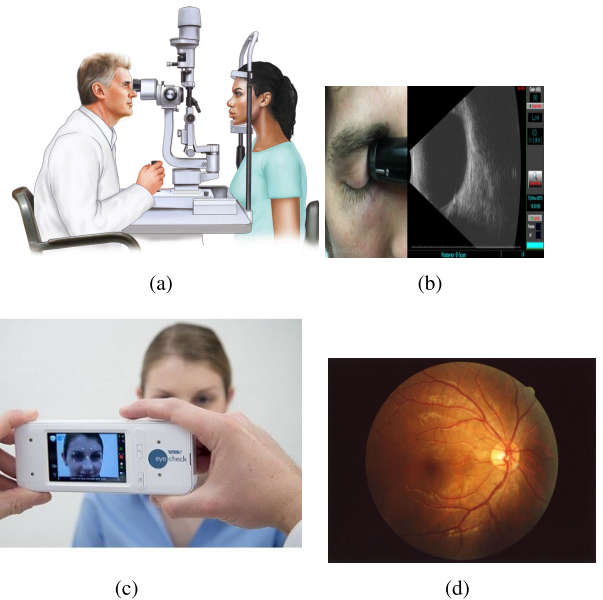


FIGURE 6. Imaging modalities used for examining eye (a) Slit lamp (b) Ultrasonic image (c) Digital image (d) Fundus image.

It is a crucial tool for assessing eye health and spotting eye problems. It is made up of expensive equipment. It can raise eye pressure, making people feel sick or hurt their eyes. Sound waves of high frequency are used in imaging ultrasound to observe the body’s interior. These real-time ultrasound images can display blood moving through BV and the body’s interior organs’ movement. Ultrasonic evaluation takes time, and a specialist is needed to ensure the evaluation’s quality. Creating a digital representation of an object’s visual qualities, such as the object’s inner structure, is called digital imaging. It doesn’t contain high-resolution images. Retinal images are also known as fundus images. A non-invasive diagnostic method called retinal imaging can produce detailed retina pictures. Specialized cameras and scanners are utilized to enlarge the eye’s retina, optic nerve, and blood vessels. Fundus photography has a sensitivity and specificity that is superior to ophthalmoscope methods.

Machine learning (ML) and deep learning (DL) techniques are utilized to detect and classify diabetic retinopathy. Many algorithms of machine learning are used in DR for classification, including neural networks (NNET), random forest (RF), K-Nearest Neighbor (KNN), Decision trees (DT), Naive Bayes (NB), and support vector machine (SVM).

DR detection and grading have extensively utilized deep learning methods. Even when numerous diverse sources are combined, it may still learn the characteristics of the incoming data. Numerous DL-based techniques exist, including convolutional neural networks (CNNs), autoencoders, restricted Boltzmann machines (RBMs), and sparse coding. Contrary to machine learning approaches, these techniques work better as the amount of training data increases with the increase of learned features. The difference between them is that DL methods don't require hand-crafted feature extraction, but the machine learning method requires it. DL methods require large data, but ML doesn't require large data [8].

The contributions of this study are as follows:

- 1) A study and analysis of the most recent machine learning and deep learning strategies for detecting and grading DR using fundus photographs.
- 2) We discussed various aspects of the DR pipeline, ranging from commonly used datasets by the researchers, the employed preprocessing methods and data augmentation steps, feature extraction methods, commonly used detection and classification algorithms, and the generally used performance metrics.
- 3) We thoroughly analyze the current trends and explore future trends of DR.
- 4) In addition, various open problems and limitations of DR that need further study are discussed.

A systematic literature review is presented in Section II, Section III presents a framework for DR detection and grading, and Section IV presents open problems and discussion.

II. LITERATURE REVIEW

The earlier detection of diabetic retinopathy has been proposed using various ML and DL techniques. This study provides a detailed review based on two categories, i.e. ML methods and DL methods for the detection and grading of DR. The difference between these two methods is given in Table 3.

A. LITERATURE SEARCH DETAILS

The systematic literature review was based on both ML and DL methods. The methods used for detecting and grading DR are discussed in detail (Table 4 and 5). Initially, we searched research papers related to diabetic retinopathy and then filtered them according to our required criteria. We want papers relevant to the detection and grading of DR. Secondly, we explore those research papers where authors used fundus images to evaluate their proposed methods. We searched on different repositories, i.e. Google Scholar, IEEE Xplore, Springer, and Science Direct. The keywords used for searching research papers were DR, DR Detection, DR grading, DR Classification, Lesions, Retinal Diseases, and PDR/NDPR. The chosen research papers range from 1995 to 2023. The search strategy used

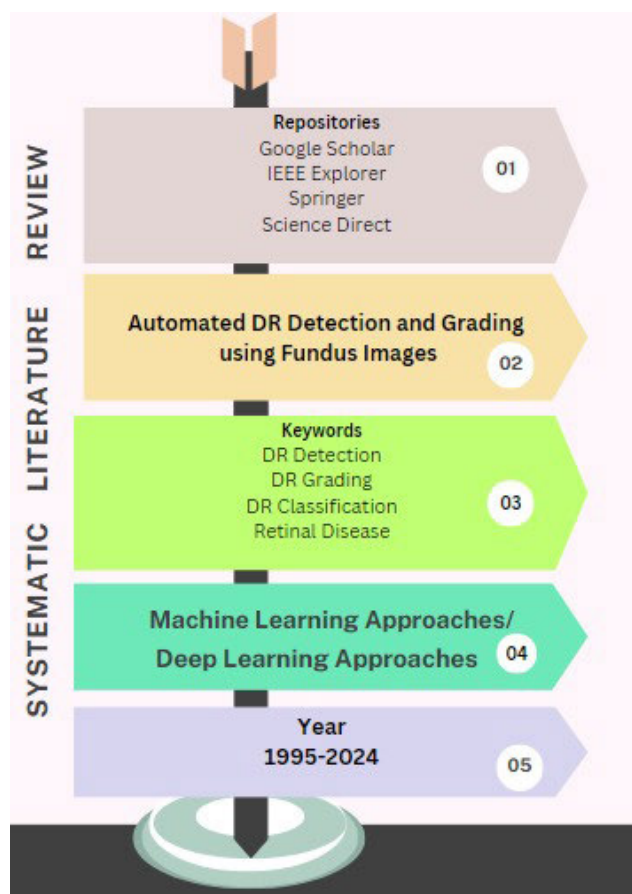


FIGURE 7. Flowchart of summarizing the Literature search.

to find relevant research publications required for the study involved these main steps (Fig 7).

B. MACHINE LEARNING APPROACHES

Machine learning-based DR classification approaches are divided into lesion-based and binary/multilevel classification. The details of these classifications are given in the subsequent sections.

1) LESION BASED CLASSIFICATION

An overview of the machine learning (ML) techniques used to diagnose and grade early-stage DR utilizing fundus images is provided in this section. Walter et al. suggested a method that automatically detects MA in fundus images. In a preprocessing step on the set of 5 fundus images, noise reduction, shade correction, image normalization, and contrast enhancement were performed. After this, image enhancement and Gaussian filtering were performed. Morphological top-hat transform is used to extract the dark details from the image. Automatic thresholding is then generated based on image quality. False positives are eliminated during the last phase. After an experiment, 86.4% of accuracy was achieved [9]. The study [10] develops a system for identifying exudates, hemorrhages, and microaneurysms in colour retinal pictures.

TABLE 3. Difference between machine learning and deep learning methods.

Aspect	Machine Learning	Deep Learning
Relationship	A subset of Artificial Intelligence	A subset of Machine Learning
Data Requirements	Can train on smaller datasets	Requires a large amount of data
Human Intervention	Requires more human intervention to correct and learn	Learns on its own from the environment and experience
Training Time and Accuracy	Shorter training time, generally lower accuracy	Longer training time, generally higher accuracy
Correlation Types	Makes simple linear correlations	Makes nonlinear complex correlations
Hardware Requirements	Can train on CPU (Central Processing Unit)	Needs a specialized GPU (Graphic Processing Unit) to train

Retina's photographs were taken from a clinical diabetic retinopathy screening service for 1273 patients in a row. The method employed segmentation to identify potential lesions, preparing images to uniformly color and improve contrast, and artificial neural network categorization of lesions. The system was tested against a set of test photos from 773 patients after being trained on a subset of photographs from 500 patients. On each patient, the maximum sensitivity for retinopathy identification was 95.1%, with a specificity of 46.3%. Although sensitivity decreased to 70.8%, specificity could be raised as high as 78.9%. No cases of retinopathy posing a threat to one's ability to see were overlooked at a location with 94.8% sensitivity and 52.8% specificity. Niemeijer et al. describe and assess an automated technique that utilizes ML to find exudates and soft exudates in digital color fundus images. It also differentiates soft exudates (cotton-wool spot) from hard exudates (drusen). A diabetic retinopathy teliagnosis database had 300 retinal pictures from one eye of 300 diabetic patients. A statistical classifier is used in the system to classify. The system obtained the ROC of 0.95. Sensitivity/specificity pairs of 0.95/0.88 for identifying any bright lesions and 0.95/0.86, 0.70/0.93, and 0.77/0.88 for the detection of exudates, soft exudates, and hard exudates, respectively, were also attained [11]. An automated system developed by Silberman et al. to detect diabetic retinopathy from retinal photographs. The images were preprocessed through global colour-balancing operation, increased pixel luminosity, customized derivative filter, and grayscale conversion. They extracted exudates from the dataset of 1000 fundus images through SIFT. SVM is further used for classification. After an experiment, they achieved 90% accuracy [12].

To execute a DR grading methodology based on FAZ enlargement utilizing fundus pictures of the retina, a less intrusive computerized DR system was designed [13]. An external fundus camera that can take high-definition retinal images is attached to a computer for image processing that can digitize and analyze retinal fundus images. For analysis, 315 fundus photos were collected. Through CLAHE (Contrast Limited Adaptive Histogram Equalization), the contrast of the RBV versus the background image is improved in the first procedure. Independent component analysis (ICA) is also implemented to get a more accurate estimation of vessel endpoints. Segmenting RBV in the fundus picture is the second step. The third step involves identifying and choosing retinal blood vessel endpoints in the perifoveal

capillary network to define and compute the FAZ area by linking the endpoints of the RBV. The performance of the classifier is assessed using V-Fold cross-validation (VFCV). For all DR phases, the system attained 84% sensitivity, 97% specificity, and 95% accuracy. The goal of the study [14] is to identify RBV, which is essential for minimizing the progression of diabetic retinopathy and preventing vision loss. The algorithm's primary modules include converting color images (RGB) to grayscale images, contrast improvement, background elimination, thresholding, and postfiltration. An automatic segmentation method for the RBV is provided by contrast enhancement and thresholding. The technique has been tested on collections of fundus photos taken from the Drive database, a freely accessible database. It consists of 40 retinal photographs. The proposed system achieves 96% accuracy.

Mahendran et al. identify exudates in five-colour fundus retinal pictures to detect diabetic retinopathy automatically. Additionally, they categorize how severe the lesions are. The preprocessing of the retinal images initially includes colour space conversion, image enhancement, and restoration. They suggested a technique in which the blood vessels and optic disc are first removed because their intensity level is the same as that of exudates. The presence and distribution of exudates are then determined using morphological processes, including dilatation and erosion. Additionally, the SVM classifier determines whether the patient is moderately impacted or seriously affected by the condition [15]. Different machine learning (ML) techniques were used to diagnose DR earlier [16]. ML techniques were applied to fundus photographs and DiaretDB0 datasets, which consist of 350 and 130 images, respectively. First, for the segmentation of blood vessels, image processing techniques such as grey channel extraction, adaptive histogram equalization (AHE), discrete wavelet transform, gaussian matched filter response, and fuzzy C-means clustering are utilized. Extraction of the green channels, thresholding, and operations of morphological dilation are just a few techniques of image processing used to retrieve features like blood vessels, hemorrhages, and exudates from raw pictures. The various classifiers (PNN, BC, and SVM) are then fed the retrieved features for classification. According to experimental findings, PNN has an accuracy rate of 89.6%, BC has a rate of 94.4%, and SVM has a rate of 97.6%. This implies that the SVM model performs better than every other model. A framework was designed by Antal et al. to classify colour fundus images

based on DR. They experimented by using the Messidor dataset consisting of 1200 images. Firstly, images were compressed with an FOV of 450 and various resolutions. Lesion detection and anatomical part recognition algorithms were used to extract features. A group of classifiers is then used to classify these features. This framework achieved 91% Specificity and 90% Sensitivity [17].

Bhatia et al. used parameters such as the diameter of the optic disc, lesion specific (microaneurysms, exudates), and image level (prescreening) from the output of several retinal image processing algorithms to focus on automatic identification of diabetic retinopathy (DR). They experimented by using the Messidor dataset consisting of 1200 images. Firstly, images were compressed with a FOV of 450 and several resolutions. After training several classifiers, the best individual classifier determines DR or non-DR categories. 94% of the accuracy was achieved by the experiment [18]. Mankar et al. detect haemorrhages and Exudates using fundus images for early DR detection. They used two image-based techniques of preprocessing i.e. Median Filtering and histogram equalization. From the gray level matrix for an image, texture features were derived. 100 fundus photographs are used for SVM training. To determine if an image is normal or has diabetic retinopathy, features are entered into the classifier. They achieved an accuracy of 89.50% [19].

The identification and grading of DR depend on the early and precise detection of microaneurysms (MAs) [20]. Several preprocessing techniques are used to create images appropriate for candidate and feature extraction. Two open databases were used to evaluate the suggested strategy: E-optha and ROC, which comprise 148 and 100 fundus images. Illumination equalization, CLAHE, and smoothing were performed on images as a preprocessing step. After Candidate extraction, Hessian matrix-based features, Shape and intensity features, and profile features were retrieved. Three supervised classifiers—K-Nearest Neighbour (KNN), Naive Bayes (NB), and AdaBoost—are the underlying classifiers to select an appropriate classifier for the feature collection. The ROC database is used to evaluate these three classifiers. According to the evaluation results, the KNN and AdaBoost classifiers perform similarly to one another and outperform the NB classifiers. This method achieved an FROC score of 0.202 and 0.273 on ROC and E-optha datasets, respectively. For the early diagnosis of diabetic retinopathy, a novel technique was put forth that uses several textural features and a machine learning classification algorithm [21]. They experimented by using the Messidor dataset consisting of 1200 images. Grayscale conversion of images was performed as a first step. Using the local ternary pattern (LTP) and a local energy-based shape histogram (LESH), two features, i.e. haemorrhage and exudates, are retrieved. LTP and LESH feature vectors are utilized to classify using SVM. 0.841 accuracy and 0.916 AUC of LTP, while 0.904 accuracy and 0.931 AUC of LESH were achieved.

Revathy et al. concentrated on a hybrid machine-learning model for auto-mated computer-aided identification of diabetic retinopathy. Edge zero padding, median filtering, and adaptive histogram equalization were all carried out during the pre-processed phase of images. They extracted the features of haemorrhage, micro-aneurysms, and exudates by using the Kaggle dataset. This dataset consists of 244 images. Exudates were segmented using smoothing, masking, and bitwise AND after image preprocessing. Median blurring, thresholding, image erosion, and image dilation were used to segment haemorrhages and micro-aneurysms. This proposed model employs a hybrid SVM, RF, and KNN classifier. The maximum accuracy value obtained from the experiment data was 82%. Precision, recall, and f-measure scores from the hybrid technique were each 0.8119, 0.8116, and 0.8028 [22]. Akif used ML methods to diagnose DR early. 300 fundus photographs are taken from credential sources. First, the original fundus image is converted to grayscale, enhancing image contrast. The canny detection algorithm is utilized to find OD and blood vessels. The image processing techniques like thresholding and morphological operation were utilized to retrieve features from photographs of the fundus. For classification, these photographic features are taken to the classifiers (PNN, BC, SVM, K-Means clustering). Following an experiment, it was determined that PNN, BC, SVM, and K-Means Clustering all had accuracy levels above 89%. The preliminary findings demonstrate that SVM was the best approach for DR detection [16].

2) BINARY/MULTILEVEL CLASSIFICATION

Identification of the several stages of DR utilizing 124 retinal optical Photos was performed [23]. The stages of DR include mild, moderate, severe, and proliferative DR. Firstly, pre-processing of images is performed, which essentially comprises histogram-based image contrast improvement. Morphological operations and thresholding are employed to extract features from retinal pictures. The extracted features are then inputted to the feedforward neural network for classification. The classifier has obtained a sensitivity of over 90% and a 100% specificity. Acharya et al. automatically identify the different stages of DR. The stages include normal, mild, moderate, severe, and prolific DR. Firstly; they perform the pre-processing step on the dataset of 300 retinal photographs. This increases image contrast via histogram equalization. For feature extraction, they have applied higher-order spectra (HOS). The SVM is then fed the extracted features (HM, MA, HE, and SE) for classification. 82% of sensitivity and 88% of specificity were obtained after an experiment [24]. AM-FM (amplitude-modulation, frequency-modulation) multiscale approaches have been proposed for differentiating between healthy and diseased retinal pictures. The online ETDRS (Early Treatment Diabetic Retinopathy Study) database was used to choose the images. They employ 120 sections of 40*40 pixels, each including two different types of normal retinal regions and four lesions frequently found in

diabetic retinopathy (DR). A professional analyst manually chose these regions. The region's types were MA, EX, NV on the retina, HM, normal retinal background, and typical vessel patterns. The textural characteristics vectors are the instantaneous amplitude's cumulative distribution functions, the instantaneous frequency magnitude (FM), and the relative instantaneous frequency angle from various scales. They evaluate inter-structure similarity using metrics of the distance between the retrieved feature vectors. The output shows that diseased lesions and normal structures of the retina can be statistically distinguished. This method achieves accuracy up to 92% [25].

Naveen et al. identified DR by using a technique for extracting blood vessels from photographs of the human eye's retina. In the preprocessing step, they converted the image RGB to a grey image. The HRF (High-Resolution Fundus) Dataset's fundus images are given a contrast enhancement using an Adaptive Histogram Equalization (AHE) image. This dataset is made up of 10 images. To prevent the over-noise amplification that AHE can bring about, CLAHE was created. The outcome demonstrated that affected DR has been identified in the fundus image and that DR is not detected in the healthy fundus image [26]. A novel model was proposed in which the sub-datasets were generated from Messidor fundus images [27]. This sub-dataset consists of 1151 images. Then the feature selection process is performed through Information Gain Attribute Evaluation (IGAE) and Wrapper Subset Evaluation (WSE). IGAE, often called entropy evaluation, recognizes the most informative characteristics for further processing in the classification model while ignoring those that provide less insightful information. The WSE technique compares the performance of various subsets of the original dataset's features using a fast-learning yet effective pre-defined algorithm. The ensemble framework was then constructed using a selection of specialized features. Three primary classification algorithms—Random Forest, Neural Network, and Support Vector Machine—were integrated into the ensemble to predict the final result of DR. The accuracy of the given model was 75.1%.

Machine learning in diabetic retinopathy detection utilizes algorithms like SVM, Decision Trees, and Random Forests to analyze retinal images and identify signs of the disease. Studies like those by Walter et al. demonstrated the effectiveness of SVM in detecting microaneurysms with high accuracy. Niemeijer et al. developed a machine learning-based system for detecting exudates in digital colour fundus photographs, achieving significant sensitivity and specificity. Machine learning models can efficiently handle structured data and provide clear insights into decision-making, facilitating easier clinical interpretation. They are generally less computationally intensive and faster to train than deep learning models. These approaches may struggle with the high dimensionality of image data and require extensive feature engineering to achieve optimal performance. They can suffer from overfitting, especially in cases with limited training data, affecting their ability to generalize

to new, unseen data. The potential of machine learning in diabetic retinopathy detection is substantial, especially in environments with limited computational resources. Future work should focus on combining machine learning with deep learning to leverage the strengths of both approaches, such as using machine learning for preliminary feature extraction followed by deep learning for detailed image analysis.

C. DEEP LEARNING APPROACHES

Same as ML-based approaches, deep learning approaches are split into lesion-based classification and binary/multilevel classification. The details of these DL-based approaches will be given later.

1) LESION BASED CLASSIFICATION

The different methods used to detect and grade DR through fundus images are presented in this section of DL. Gardner et al. took fundus images to detect DR. 179 fundus images, which were taken under consideration. Histogram equalization, edge detection, and median filtering were performed on fundus images as a preprocessing step. They used the squares produced by non-overlapped slicing a retinal image with a resolution of 700×700 pixels into several smaller squares of 20×20 pixels to train an ANN to distinguish between the presence and absence of a haemorrhage, exudate, or blood vessel. For identifying DR on fundus images, the network attained 88.4% sensitivity and 83.5% specificity [28]. A method presented in [29] used the "Moat Operator" segments haemorrhages and exudates and improves repetitive segmentation of regions of 30 digitized retinal pictures. An image's RGB components were converted into an intensity, hue, and saturation (IHS) model. The intensity band underwent adaptive, local contrast enhancement, and the resulting image was again changed into RGB to display. Ophthalmologists identified haemorrhages and exudates in multiple small, 10×10 square, non-overlapping cut pictures. The authors used segments rather than pixel segmentation to assess their segmentation instead of pixel segmentation. The NPDR method identified hemorrhages and micro-aneurysms with 77.5% sensitivity and 88.7% specificity and exudates with 88.5% sensitivity and 99.7% specificity.

Kande et al. use pixel categorization and mathematical morphology to identify microaneurysms and haemorrhages in red lesions. They randomly took 89 images from Stare, DiaretDB0, and DiaretDB1 datasets for an experiment. In the pre-processing step, the Red and green channels of images were extracted. Contrast stretching and median filtering were performed on images. To determine whether the image had red lesions, they examined the red and green extractions of an image. Then, candidate locations for red lesion containment are classified using the SVM technique. The suggested method had 100% sensitivity and 91% specificity [30]. Using the k-Nearest Neighbours algorithm as a splat-based feature classifier chosen using an envelope and filter approach, Tang et al. identified haemorrhages. They used a Messidor

TABLE 4. Systematic analysis of literature review of ML approaches.

R.No	Author /Year	Dataset	Pre-Processing	Feature Engineering	Results
1	Walter et al/2002	Photographic set /5	Noise Reduction, Shade Correction, Image Normalization, Contrast Enhancement, Image Enhancement and Gaussian Filtering.	Microaneurysms are detected through morphological top-hat transform, the bounding box closing, and automatic thresholding.	Accuracy=86.4%
2	Usher et al/ 2003	Retinal Images/ 1273	Standardize colour and Enhance contrast	MA, HE, and EX are extracted and ANN performed classification of these lesions.	Sensitivity= 94.8%and Specificity= 52.8%
3	Niemeijer et al/2007	Retinal images/300	NA	A machine learning-based, the automated system detects and soft exudates in digital colour fundus photographs. Also, soft exudates can be differentiated from hard exudates. A statistical classifier is used for classification.	Sensitivity /specificity =0.95/0.86, 0.70/0.93, and 0.77/0.88 for exudates, soft exudates, and hard exudates
4	Yun et al/2008	Retinal photographs/124	Image contrast improvement and Histogram Equalization	The features extracted from image processing techniques are fed to the feed-forward neural network for classification.	Sensitivity= 90% , Specificity= 100%
5	Acharya et al/2008	Retinal photographs/300	Image contrast improvement and Histogram Equalization	Higher-order spectra (HOS) is used for feature extraction. The extracted features (hemorrhages, microaneurysms, hard exudates, and cotton wool spots) are then fed to the SVM for classification.	Sensitivity=82%, Specificity=88%
6	Silberman et al/2010	Retinal images/1000	Global Color-balancing Operation, Increased pixel's Luminosity, Customized Derivative Filter and Gray Scale Conversion.	Extracted exudates from the images through SIFT. SVM is further used for classification.	Accuracy= 90%
7	Agurto et al/2010	ETDRS	NA	AM-FM methods used discriminating between normal and pathological retinal images.	Accuracy= 92%
8	Fadzil/ 2011	Fundus Photographic set/ 315	Contrast Enhancement	Segmented retinal blood vessels through Otsu's thresholding, then detect and determine FAZ.	Sensitivity> 84%, Specificity >97% and Accuracy >95%
9	Saleh/2011	DRIVE/40	Color image (RGB) to grey/green conversion and Contrast enhancement.	After preprocessing, Background exclusion, thresholding and postfiltration were performed. The contrast enhancement and thresholding offer an automated segmentation procedure for retinal blood vessels.	Accuracy= 96%
10	Mahendran et al/2013	Retinal fundus cameras/5	Colour Space Conversion, Median Filtering and Adaptive Histogram Equalization.	The input features were extracted using the Gray Level Co-occurrence Matrix (GLCM). SVM classifier is used to assess the severity of the disease.	NA
11	Priya et al/2013	DiaretDB0/130, Fundus Photographs/350	Grayscale Conversion, Adaptive Histogram Equalization, Discrete Wavelet Transform and Gaussian Matched Filter Response.	The features like blood vessels, haemorrhages and exudates are extracted from the raw images using the image processing techniques and then fed to the different classifiers (PNN, BC, SVM) for classification.	For DiaretDB0, Accuracy of PNN,BC,SVM= 89.6 %, 94.4%, 97.6% For fundus images Accuracy of PNN, BC, SVM= 87.69%, 90.76%, 95.38%
12	Antal et al./2014	Messidor/1200	Compressed images with 450 FOV and different resolutions.	Features were extracted using image-level, lesion-specific, and anatomical components. These features are then classified using an ensemble of classifiers.	Specificity= 91% Sensitivity: 90
13	Bhatia et al/2016	Messidor /1200	Compressed images with 450 FOV and different resolutions.	Features were extracted using image-level, lesion-specific, and anatomical components. Classifiers are trained individually.	Accuracy= 94%
14	Mankar /2016	Fundus Images/ 100	Median Filtering and Histogram Equalization	Derived texture from GLCM and classified through SVM.	Accuracy= 89.50%
15	Wu et al/2016	E-optha/381, ROC/100	Illumination equalization, Contrast enhancement and smoothing	After Candidate extraction, Hessian matrix-based features, Shape and intensity features and profile features were extracted.	ROC and E-optha FROC score 0.273 and 0.202.

TABLE 4. (Continued.) Systematic analysis of literature review of ML approaches.

16	Chetoui et al/2018	Messidor /1200	Convert the colour images to grayscale.	Haemorrhages and exudates are extracted using (LTP) and (LESH).	LTP Accuracy= 0.841 AUC= 0.916 , LESH Accuracy= 0.904 AUC= 0.931
17	Naveen et al/2019	HRF/10	RGB to Gray conversion Image.	Adaptive Histogram Equalization using CLAHE algorithm with open CV (Computer Vision) framework implemented to Extract blood vessels to detect exudates.	Accuracy= 98.25, Specificity= 99.18, Sensitivity= 86.22
18	Revathy et al./2020	Kaggle/244	Colour Space Conversion, Edge Zero Padding, Median Filtering, Adaptive Histogram Equalization.	Counting the number of white pixels from the segmented images and dividing it by the total number of pixels in the image to detect exudates. SVM, KNN and Random Forest are used for classification	SVM Accuracy= 68%, KNN Accuracy= 76%, Random forest Accuracy= 90% ,
19	Odeh et al/2021	Messidor /1151	NA	Feature selection process is performed through Information Gain Attribute Evaluation (IGAE) and Wrapper Subset Evaluation. Random Forest, Neural Network, and Support Vector Machine were used as an ensemble classifier.	Accuracy= 75.1%
20	Akif/2021	Fundus Photographs /300	Gray Scale Conversion and Image Contrast Enhancement.	Canny detection algorithm is used to find OD and blood vessels. The features are extracted from fundus images using the image processing techniques and then fed to the different classifiers i.e. SVM, PNN, BC and K-means Clustering.	PNN accuracy = 89%, BC 94%, SVM 97%, and K-Means Clustering 87%.

dataset consisting of 1200 images for an experiment. The FOV was automatically determined during a preprocessing stage, and the image was resized to 1026×681 pixels, with an FOV that was roughly 630 pixels in diameter throughout the whole dataset. With the Messidor data set, the experiment produced a 0.96 receiver operating feature (ROC) curve score [31].

To detect DR, the study [32] has presented a Retinal Blood Vessel (RBV) segmentation method utilizing a DNN (Deep Neural Network) that was trained on 400,000 samples. This method uses improved, contrast-normalized, and amplified images. Images of Drive, Stare, and Chase datasets were used. The network training was performed using backpropagation and dropout in this model. PLAIN BALANCED and NO-POOL are the two fundamental configurations proposed in the model. They rely on structure prediction for the concurrent classification of numerous pixels. The maximum AUC was attained by the PLAIN BALANCED model, which was 0.9738 on Drive and 0.9820 ± 0.0045 on Stare. The area under the ROC curve for this model was 99%, and its classification accuracy was 97%. The method has an 87% sensitivity rating for identifying fine vessels and is resistant to central vessel reflex phenomena. A CNN architecture was created by Grinsven et al. by using nine layers made up of carefully chosen samples and 41×41 -inch patches labelled with or without haemorrhage evidence; it is possible to detect haemorrhage. Messidor and Kaggle data sets consisting of 6679 and 1200 images were used for an experiment. In preprocessing, the FOV of the fundus images was segmented, and circular template matching was

used. Images were cropped and resized to 512×512 -pixel dimensions to reduce computational costs. Image contrast was improved. Haemorrhages were detected on images from the Messidor and Kaggle data sets with 84.8% sensitivity and 90.4% sensitivity, respectively [33]. A couple of CNN models are composed of a single CNN and heterogeneous CNN modules.

Both models were trained using gradient descent and backpropagation, respectively, proposed by Paul et al. [34]. Using the DiaretDB0 dataset, they are contrasted to determine how well they can identify DR. The multilayer perceptron network classifier utilized in the proposed framework and its output show normal pictures, MAs, HEs, hard EXs, soft EXs, and NV, respectively. The experiment was performed with heterogeneous CNN and single CNN with varied filter sizes and receptive field sizes using 130 colour images from the DiaretDB0 dataset. For MAs, HEs, hard EXs, soft EXs, and NV, the model's accuracy on a single CNN was 95%, 75%, 62.5%, 67.5%, and 95%, respectively. The feature was detected and extracted with 100% accuracy on the heterogeneous CNN compared to a single CNN. Using fundus images, a DL technique has been developed for the automated detection of DR and Diabetic Macular Edema (DME), and for the detection of DR, Inception-V3- architecture Neural Networks have been used to identify HEs and MAs [35]. The EyePACS-1 dataset of 9963 photos and the Messidor-2 dataset of 1748 photos have both been used in the proposed methodology. The model has undergone preprocessing, distributed Stochastic Gradient Descent (SGD) network weight optimization, and

hyperparameter optimization. An ensemble of 10 networks was utilized to process the whole developed set consisting of 128,175 photos, and the ensemble predictions were used to generate the final linear average prediction. With a 0.991 and 0.990 AUC for EyePACS-1 and Messidor-2, the algorithm has successfully recognized DR. To increase the accuracy of EXs identification for DR detection, **Prentašić and Sven Lončarić** [36] recommended a DCNN (Deep Convolutional Neural Network) that takes RBVs and OD into account to detect DR. The Frangi vesselness filter, Total Variation (TV) regularisation denoising, split Bregman algorithm for denoising, morphological operations, dynamic thresholding, pixel-wise feature extraction and classification, and clustering techniques are all included in the model's per-formed image preprocessing. The proposed approach combines various landmark detection algorithms to detect and locate EXs. The proposed model uses a variety of preprocessing, thresholding, localization, and object detection algorithms, including the entropy-based method, the method of brightness, the Simulated annealing method, and the Hough transformation of vessels. Using 50 images of the DRiDB dataset, the provided CNN model successfully distinguished between an EX and a non-EX, with 78% sensitivity, 78% Positive Predictive Value, and 0.78 F-score. A 10-layer multiclass neural network was developed to segment HM, MA, and EX in the retinal fundus images. CNN automatically performs segmentation and simultaneously discriminates these lesions. The images used belong to the Cleopatra dataset, which consists of 298 images. No preprocessing was performed on the photos. Hemorrhage, micro-aneurysms, and exudates segmentation had a sensitivity of 62.57%, 46.06%, and 87.58%, respectively [37].

According to **Quelleg et al.**, a CNN model was developed using the ConvNets network structure that produces heat maps to simultaneously detect lesions: MA, HM, HE, and SE. Three CNNs were trained to classify each image in the dataset as referable DR (refer to moderate stage or more) or non-referable DR (no DR or mild stage). Three datasets, Kaggle, DiaretDB1, and private E-ophtha, each containing 88702, 89, and 463 photos, were used. The photos were reduced in size, cropped to 448 by 448 pixels, normalized, and had the FOV degraded by five percent in the pre-processing stage. The data were augmented and a big filter of Gaussian was utilized. CNN architecture pertained to the AlexNet and the two networks of the o O solution. The CNNs found MA, HM, SE, and HE. For hemorrhages, hard exudates, soft exudates, and microaneurysms, the model showed AUC values of 0.614, 0.735, 0.809, and 0.500, respectively [38]. **Lam et al.** investigated five CNN models, including AlexNet, VGG16, GoogLeNet, ResNet, and Inception-v3 to find various types of lesions in retinal pictures. The ophthalmologist looks at the photos of the fundus that belong to the Kaggle retinopathy dataset and the eOphtha dataset, which contain 31,126 and 463 photos, respectively. They generate patches from photographs of haemorrhages, microaneurysms, exudates,

retinal neovascularization, or structures that appear to be normal. CNN is trained using these image patches to forecast the appearance of these 5 categories. The window of sliding approach is utilized to produce a probability map for the full image. The suggested model successfully detected exudates and microaneurysms with an AUC of 0.94, 0.95, and ROC of 0.86, 0.64 [39].

By combining DL approaches with domain knowledge for feature learning, red lesions were found in DR images [40]. Following that, the photos were categorized using the Random Forest technique. The green band was extracted from the datasets of Messidor (1200 images), E-ophtha (463 images), and DiaretDB1 (89 images) datasets. The FOV was also expanded. A Gaussian filter, an r-polynomial transformation, a thresholding operation, and numerous morphological closure functions were also applied. Then, patches of red lesions were enhanced for CNN training and scaled to 32*32 pixels. The custom CNN comprises 4 CONV layers, 3 pooling layers, and 1 FC layer. They obtained Competition Metrics (CPM) of 0.4874 on DiaretDB1 and 0.3683 on E-ophtha, respectively. Utilizing a modified DCNN model based on VGG-Net has been considered to identify DR features such as drusen, EXs, MAs, SEs, and HEs [41]. They employed the Kaggle, e Messidor-2 dataset for DR detection, which consists of 88,696 photographs. The 81,670 photographs were utilized for training, and the remaining pictures, combined with the Messidor- 2 dataset, which consists of 1748 enhanced images, were used for testing. Preprocessing, assessment module of image quality, augmentation of random images, localization and segmented features, dropout, and retinal lesions classification make up the diagnostic pipeline of the proposed model. The model has scored 0.923 AUROC, 92% sensitivity, and 72% specificity with the Kaggle dataset while operating at the high sensitivity. The 80% sensitivity and 92% specificity of the suggested model with the Kaggle dataset at the operational point of high specificity. The 99%, 87% of sensitivity and 71%, 92% of specificity were recorded at a high sensitivity and high specificity, respectively with Messidor-2. Deep Residual Learning (DRL) was used to develop a CNN to detect DR automatically in the proposed data-driven DL technique for deep feature retrieval and to classify images. Testing of the model was performed with the Messidor-2 dataset and E-Ophtha dataset, having 1748 and 463 photos, respectively. It was also trained with EyePACS dataset (75,137 images). Image Preprocessing, dataset augmentation, batch normalization, ReLU activation, and categorical cross-entropy loss function for class discrimination utilizing gradient boosting classifiers have all been carried out in the model. The convolutional approach was used to retrieve 1024 deep features from the model. The model has identified retinal HMs, HE, and NV by displaying heat maps. The proposed model achieved 0.97 AUC with the EyePACS dataset with an average of 94% sensitivity and 98% specificity, compared to AUCs of 0.94 on the Messidor-2 dataset with an average

of 93% sensitivity and 87% specificity and 0.95 on the E-Ophtha dataset with an average 90% sensitivity and 94% specificity [42]. To distinguish MA from DR images, a distinctive CNN architecture was applied [43].

This study included three datasets with 100, 381, and 89 pictures each: ROC, E-ophtha, and DiaretDB1. The green plane was extracted from these datasets before being processed with cropping, scaling, and Otsu thresholding to create a mask, weighted sum, and morphological functions. After that, random alterations were applied to the collected MA patches. The employed CNN has 18 CONV layers, four skip connections between the two routes after each CONV layer, a batch normalization layer, three max-pooling layers, and three fundamental up-sampling levels. The experiment received FROC (Free-response ROC) scores of 0.355 for the ROC dataset, 0.392 for the DiaretDB1 dataset, and 0.562 for the E-ophtha dataset, respectively. To locate EX in DR images, **Adem** employed a customized CNN with Circular Hough Transformation (CHT). The three public datasets they utilised were the DiaretDB0, DiaretDB1, and DrimDB, each containing 130, 89, and 125 images. The datasets were all made into grayscale images. Next, adaptive histogram equalization functions and canny edge detection were utilized. CHT found the OD and eliminated it from the datasets. 3 CONV layers, 3 max-pooling layers, and an FC layer that employs softMax as a classifier make up the custom CNN. This CNN was given the photos' 1152*1152 pixel data. The accuracy of detecting EX was 99.17 with DiaretDB0, 98.53 with DiaretDB1, and 99.18 with DrimDB [44].

Mo et al. used deep residual networks to segregate and classify the exudates to identify exudate lesions in the publicly accessible E-ophtha dataset and HEI-MED dataset which contains 381 and 169 pictures, respectively. The EX were grouped using a fully convolutional residual network. This network includes two modules, up-sampling and down-sampling. A deep residual network was then used to classify the exudates. This network consists of 5 residual blocks, 1 CONV layer, and 1 max-pooling layer. The up-sampling module uses CONV and deconvolutional layers to enlarge the image as input. In contrast, the down-sampling module uses a CONV layer, a max pooling layer, and 12 residual blocks. The residual block contains three CONV layers for batch normalization. This work achieved a sensitivity of 0.9227 and 0.9255 and an AUC of 0.9647 and 0.9709 for the E-ophtha and HEI-MED datasets, respectively [45]. A ten-layered CNN was presented to detect DR. On the fundus photographs, it employs patch image-based analysis. The model consisted of 284 retinal photographs from the DiaretDB1 and e-ophtha datasets; 75 of these images were used for patch-based analysis training on the DiaretDB1 dataset, and the remaining photos were used for image-based analysis testing from both the DiaretDB1 and e-ophtha datasets. The model performed contrast enhancement to retrieve EX, HE, and MA. Next, it divided 50×50 patch sizes to create rule-based probability maps. 0.96 Sensitivity, 0.98 specificity and

0.98 accuracy were achieved for identifying EXs by model, correspondingly during the patch-based analysis, as well as HEs with sensitivity, specificity, and accuracy of 0.84, 0.92, and 0.90 and MAs with 0.85, 0.96, and 0.94 respectively. On the DiaretDB1 test set, the proposed method segmented EXs, HEs, and MAs with accuracy of 0.96, 0.98, and 0.97, respectively, and error rates of 3.9%, 2.1%, and 2.04%. On the e-Ophtha dataset, EXs and MAs were segmented with accuracy of 0.88 and 3.0, and error rates of 4.2% and 3.1%, respectively. It has been found that feature detection performed simultaneously rather than individually, without redundancy, can more accurately reduce potential blunders. It has been found that feature detection performed simultaneously rather than individually, without redundancy, can more accurately reduce potential mistakes. Additionally, it has been noted that by considering the surroundings and history of potential lesions, the Post-processed phase has improved image quality and reduced mistake rates [46].

A 5-stage expert-guided statistical model to resolve an imbalanced MAs detection problem has been proposed by the study [47] to close the semantic gap between fundus images and clinical data through image-to-text mapping. The proposed methodology is a "partition frequency-inverse lesion frequency" model that depicts and forecasts specific lesions. An 'rxr' patch with an r-value of 64 has been taken by the MS-CNN, upon which selection of candidates, filtering, and MA segmentation through Gaussian filter and top-hat transform are carried out. A cascaded CNN classifier is then utilized to classify images. The suggested model was trained and tested using a dataset of 646 photos and 89 fundus images of the DiaretDB1 dataset. The model had a first-stage MS-CNN score of 30.4%, 100% precision, 17.8% accuracy, and 17.9% recall. It completed the second level with 87.8% recall, 99.7% precision, 96.1% accuracy, and a 93.4% F1 score. It is noted that the model is practical and offers space for incorporating DL approaches in the classification of images, extraction of deep features, and text-to-image mapping. Using 88,702 images from Kaggle's EyePACS dataset, of which 35,126 are used to train and 53,576 serve for testing, Islam et al. [48] presented a DCNN to detect DR through MAs identification. The model has undergone rescaling, data augmentation, feature blending, orthogonal weight initialization, stochastic gradient descent (SGD) optimization, L2 regularization, and Adam optimizer for the model's training. The obtained results have a 0.844 AUROC and 0.743 F-Score. MAs, HMs, HEs, SEs, and the OP were automatically segmented using a fully convolutional deep neural network that was trained end-to-end [49].

For their experiment, they employed the Drishti-GS dataset. Vertical and horizontal flipping of images were included after image resizing. The "SegNet" network has a decoder that controls pixel-by-pixel classification and a 13-layer convolutional VGGNet encoder. The "SegNet" network consists of a decoder that controls classification pixel-by-pixel and a 13-layer convolutional VGGNet encoder.

The experiment yielded sensitivity scores for the OD, MAs, HEs, HMs, and SEs of 0.8572, 0.0059, 0.5498, 0.0829, and 0.1823, respectively. To segment various forms of retinal disorders, **Ananda et al.** adjusted the U-Net deep neural network by lowering the filter's number and layers of coding. They experimented using the 81 and 1200 picture IDRiD and Messidor datasets. Through simple fundus image rotation and resizing, they improved the training data. One of DR symptoms, such as HMs, MAs, HEs, SEs, and OD, is segmented using each of the U-Net models. For bleeding segmentation, the U-Net model scored a 0.86 value of dice coefficient [50]. The DR signs such as HMs, MAs, SEs, and HEs were all detected by the L-Seg model [51]. Each image in the IDRiD, e-ophtha, and DDR datasets contains 81,463 and 13673 photos and is resized. More training samples were produced using the data augmentation method through rotating, horizontal, and vertical flipping. The suggested approach combines multiscale features to effectively address the challenge of segmenting small areas. When the IDRiD dataset is used, the L-Seg architecture yields better results; this method produced 67.34 AUC for hemorrhage. **Yan et al.** presented a model of GlobalNet and LocalNet. By combining the results of these two U-net models, they performed segmentation of HMs, MAs, HE, or EX in each image of the Segmentation subchallenge dataset. The dataset for the segmentation subchallenge consists of 245 photos. The images are first down-sampled and trimmed. HE and MA may be successfully segmented using the combination model. While segmentation rates of HM and SE in the GlobalNet model were the highest. The Fused Model produces 0.889 and 0.525 AUPR for EX and MA, respectively.

Yan et al. [52] used a Random Forest classifier in combination with a handcrafted and enhanced pre-trained LeNet architecture to detect red lesions in DR dataset. 89 photographs make up the DiaretDB1 dataset. The images were improved using CLAHE, and the green channel was clipped. The Gaussian filter also eliminated the noise and applied a morphological approach. The U-net CNN architecture was then used to segment the blood vessels from the images. Four convolutional layers, three max-pooling layers, and one Fully Connected layer comprise the upgraded LeNet design. Red lesion identification in this study has a sensitivity of 48.71% [53]. To detect MAs and diagnose DR, **Eftekhari et al.** presented a Deep Learning Neural Network (DLNN). It is a training network made up of two CNN structures that are completely different. These CNN structures are called the basic CNN and the final CNN. Retinopathy Online Challenge (ROC), which contains 100 photos, and E- Ophtha-MA, which has 381 images, were two datasets utilized to train and evaluate the proposed model. A balanced dataset resulted from pre-processing performed by the suggested model, which, in the basic CNN, also produced a probability map to distinguish between MAs and non-MAs. The model uses Stochastic Gradient Descent (SGD), dropout, and binary cross-entropy loss function for training, as well as backpropagation for parameter

optimization and post-processing on the output of the final CNN. For the ROC dataset and the E-Ophtha dataset, the suggested model has attained 0.660 and 0.637 Free-response Receiver Operating Characteristic Curves (FROC or FAUC), respectively [54]. To detect the existence of hemorrhages, **Huang et al.** used a CNN that combined RetinaNet and a bounding box refining network (BBR-Net). It increases the annotation's accuracy of data related to training. A private dataset with 80 and 590 photos was used in addition to the IDRiD. The method in question begins by preprocessing a fundus image using adaptive gamma correction and CLAHE for contrast adjustment. The technique improved manually traced hemorrhage annotations, outperforming the default RetinaNet system. On the IDRiD data set, a mean IoU (Intersection over Union) value of 0.8715 was noted [55].

By fusing the features of a custom-built and hand-made CNN with a Random Forest classifier, hard exudate lesions were identified in the E-ophtha dataset and the HEI-MED dataset [56]. 463 and 169 images are in the HEI-MED dataset and E- ophtha, respectively. Cropping, color normalization, changing the camera's aperture, and detecting the candidates through morphological construction and dynamic thresholding were all steps in processing these datasets. After that, 32*32-inch patches are gathered and enhanced. To find the patch features, the customized CNN has three convolutional layers, three pooling layers, and a fully connected layer. For the E-ophtha and HEI-MED datasets, this work achieved 0.8990 and 0.9477 sensitivity values and 0.9644 and 0.9323 of AUC. **Pour et al.** employed CLAHE-based pre-processed images in conjunction with Messidor, Messidor-2, and IDRiD datasets to perform feature extraction and classification using EfficientNet B5. The model has a 0.945 AUC on Messidor and a 0.932 AUC on IDRiD [57]. EfficientNet B7 was utilized in the study [58] to extract and classify features, while Global Average Pooling was employed to find DR. The model has been trained using the Kaggle Eye-PACS and Asia Pacific TeleOphthalmology Society (APTOS) 2019 datasets, and features including EXs, HEs, and MAs have been retrieved using Grad-CAM (Gradient-weighted Class Activation Mapping). The model has a 0.990 and 0.998 value of AUC on the EyePACS dataset and APTOS 2019 dataset, respectively. The detection and classification of diabetic retinopathy through deep learning models, emphasizing red lesion identification in retinal images, is given in [59]. Utilizing UNet for semantic segmentation and CNN for classification, the study demonstrates promising results across multiple datasets.

2) BINARY/MULTILEVEL CLASSIFICATION

Alghamdi et al. [60] proposed two consecutive DL architectures to detect DR with integrated cascade CNN classifiers and feature learning. The AdaBoost ensemble technique was employed in the model to train the classifier and select features. The suggested method has used five thousand seven hundred eight photos from datasets like Drive, DiaretDB1,

Messidor, Stare, Kenya, Hapiece, Pamdi, and KFSH. The model trained and evaluated the abnormality detector using the annotated OD photos from Pamdi and Hapiece. The model's accuracy for OD localization on Drive, DiaretDB1, Messidor, Stare, Kenya, Hapiece, Pamdi, and KFSH was 100%, 98.88%, 99.20%, 86.71%, 99.53%, 98.36%, 98.13%, and 92%. **Takahashi et al.** developed a trained, modified, and randomly initialized GoogLeNet DCNN to detect DR. This suggested model combines the two models, i.e., AI1 and AI2. ResNet was used to train the AI1 model using ResNet. These models were trained with a modified version of Davis grading on a concatenated picture of 4 images. Simple DR, Pre-proliferative DR, and PDR are all part of the modified Davis grading system. The AI2 model of the network is used to detect retinal HM and HE. On 496 photos the model scored 0.74 Prevalence and Bias Adjusted Fleiss' Kappa (PABAK) and 81% accuracy. It has been noted that the research has found the retina's reflection of the surface as an anomaly present in people of younger age that may help to detect DR earlier [61].

Wang et al. recommend Net-5 and Net-4 CNN models with larger datasets of DR to minimize the number of parameters that are needed to detect DR. These models incorporate a Regression Activation Map (RAM) layer and fully linked Global Average Pooling (GAP) layers. The model has created a baseline model of three (small, medium, and large) networks. It has used 35,126 images from the Kaggle dataset to train and test the model. The performance of the suggested model has been evaluated and compared using orthogonal initialization, data augmentation, and feature blending. This approach applied a Fully Convolutional Neural Network (FCNN) to the combined features to produce the final projected regression values discretized at the different thresholds and to get integer levels for the Kappa scores computation. Using the Net-5 and Net-4 settings of the architecture, without feature blending, the highest Kappa score of the proposed network was 0.81 [62]. To classify DR, Li et al. have developed a pure DCNN approach and a modified DCNN strategy that uses fractional max-pooling. The model was trained using 34,124 preprocessed Teaching-Learning-Based Optimization (TLBO) parameterized images from the freely accessible DR dataset obtained from Kaggle. This model has a recognition rate of 86.17% after using 1000 validation photos and 53,572 testing images [63].

For early DR identification, **Chaturvedi et al.** used 3662 images of fundus photography APTOS2019 dataset and a pre-trained DenseNet121 network. The proposed technique has achieved 96.44% and 96.51% accuracy for the classification of a single class and multi-level DR classification, respectively [64]. Hattiya et al. compared different CNN architectures upon 23,513 retinal images and concluded that AlexNet is the most suitable CNN architecture to detect DR. AlexNet has achieved 98.42% and 81.32% accuracy values for the training and testing sets [65]. The study [66] has proposed a model for the Kaggle APTOS 2019 contest dataset

using Transfer Learning and aggregation of deep features from multiple convolution blocks of pre-trained models like NASNet and VGG-16 to enhance feature representation. This has established a comparison with manually created features for determining the severity of DR. Averaging pooling with straightforward fusion methods on top of Deep Neural Networks (DNN) performs better, according to the model's comparison of various feature pooling and fusion procedures. The model's accuracy was 84.31%, and its AUC was 97. Due to the increased strength and capacity to automatically extract features compared to machine learning-based approaches, deep learning-based systems have grown in popularity, as the related work shows. Deep learning also enables precise localization of the boundaries of the retina. The only drawback is that it requires time and challenging training.

The study [67] investigates two deep learning models for Diabetic Retinopathy (DR) detection and classification: a hybrid model combining VGG16 and XGBoost and a DenseNet 121 model using the APTOS 2019 dataset. The DenseNet 121 model achieved a superior accuracy of 97.30%, demonstrating deep learning architectures' effectiveness in early DR detection. Another study [68] presents a hybrid method for Diabetic Retinopathy (DR) detection using a novel CNN model, GraphNet124, and ResNet50 for feature extraction from the Kaggle EyePACS dataset. It incorporates Local Binary Patterns (LBP) for texture features and optimizes the feature vector using the Binary Dragonfly Algorithm (BDA) and Sine Cosine Algorithm (SCA). The optimized features are classified using SVM, achieving high accuracy and showcasing the method's effectiveness in DR detection and classification. The Study [69] explores deep learning and transfer learning algorithms to detect various stages of Diabetic Retinopathy (DR) using a large dataset of approximately 60,000 images. The study evaluates the performance of four deep learning models: ResNet-101, DenseNet121, InceptionResNetV2, and EfficientNetB0. DenseNet121 was the most effective, with accuracies for the models reported as 97%, 96%, 95%, and 94%, respectively.

Upon reviewing prior methodologies, we advocate for microaneurysm detection as it offers insights into the progression of diabetic retinopathy. Enhancing this approach with the segmentation of retinal vessels could further aid ophthalmologists in assessing disease advancement, facilitating timely intervention and management for patients [70], [71]. The commonly used frameworks for detecting and classifying DR are explained in the subsequent section.

Deep learning uses neural networks with multiple layers (e.g., CNNs) to automatically learn features from retinal images for diabetic retinopathy detection. Gardner et al. utilized deep learning to detect hemorrhages and exudates in fundus images, demonstrating high sensitivity and specificity. Grinsven et al. developed a CNN architecture for hemorrhage

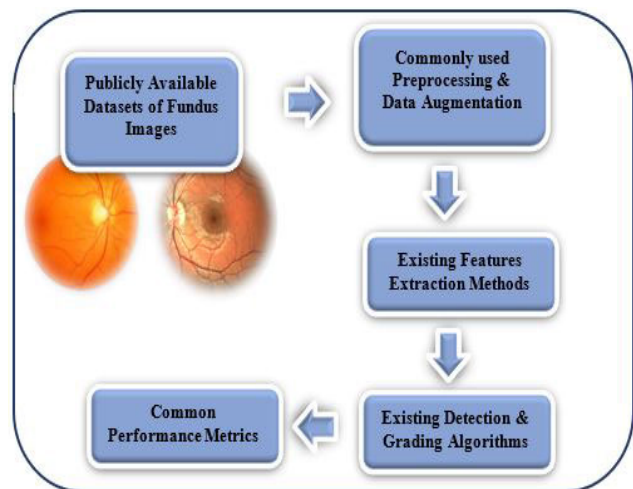


FIGURE 8. Framework for detection and grading DR.

detection, achieving high accuracy and validating the power of deep learning in medical image analysis. Deep learning models can automatically learn complex features from data, reducing the need for manual feature extraction. These models have shown superior performance in image classification tasks due to their ability to capture hierarchical features in images. Deep learning models require large amounts of data and significant computational resources for training. They are often considered “black boxes,” making it difficult to interpret their decision-making processes, which can hinder clinical acceptance. Deep learning continues transforming diabetic retinopathy detection, offering advanced capabilities in analyzing retinal images. However, to enhance their adoption in clinical practice, future research should aim at improving the interpretability of these models and developing methods to train them effectively with smaller datasets. Integrating deep learning models with clinical knowledge can provide more comprehensive and reliable diagnostic tools.

III. FRAMEWORK FOR DR DETECTION AND GRADING

Based on the literature review, most existing studies employed a common framework for detecting and classifying diabetic retinopathy. The framework comprises the following steps: (1) Retinal Datasets, (2) Preprocessing and Data Augmentation, (3) Feature Engineering, (4) Detection and Grading Algorithm, and (5) Performance Metrics. The proposed framework includes commonly used steps extracted from a detailed examination of the literature review. Numerous ML and DL techniques are used to detect and classify diabetic retinopathy. The framework for DR detection and grading is visually illustrated in Fig. 8.

A. RETINAL DATASETS

The number of public retinal datasets used for the DR detection and grading experiment. Some of the publicly available datasets are discussed below.

1) DIARETDB1

It has 89 retina fundus images that are freely available to the public. 84 images belong to DR affected; the remaining are normal images annotated by four medical experts. Quellec et al. [38], Orlando et al. [40] and Adem [44] have utilized this dataset for the experiment of red lesion detection.

2) DIABETIC RETINOPATHY DETECTION

It includes 88,702 high-quality images gathered from various cameras and range in resolution from 433 * 289 pixels to 5184 * 3456 pixels. Different phases of DR are utilized to categorize all photos. Many Photographs on Kaggle Retinopathy are not of good quality and have inaccurate labels [72]. For the retinal lesion, Grinsven et al. [33] and Lam et al. [39] have used the Kaggle dataset.

3) E-OPHTHA

Both sections of EX and MA are included in this dataset, which is openly accessible. The exudate section consists of 35 regular photos and 47 EX images. 148 photos with MA and 233 images without MA⁴ are included in MA section of E-ophtha. Mo et al. [45] and Wang et al. [56] have utilized this dataset for EX detection.

4) DOUBLE DATA RATE (DDR)

This dataset, open to the public, has 13,673 fundus images taken at a 45-degree FOV and labeled for five different DR stages: normal, mild, moderate, severe, and PDR. The dataset includes 757 pictures with DR lesions labeled. Various authors have used DDR datasets for DR detection.

5) DIGITAL RETINAL IMAGES FOR VESSEL EXTRACTION (DRIVE)

The segmentation of blood vessels is done using this publicly accessible dataset. It has 40 photos that were taken with a FOV of 45 degrees. The images are 565 by 584 pixels in size. Seven of these are small DR images, while the rest are pictures of a normal retina. Various authors have used the DRIVE dataset for DR detection.

6) HIGH RESOLUTION FUNDUS (HRF)

For blood vessel segmentation, these publicly accessible images are offered. It has 45 photos. 15 images belong to glaucomatous, 15 are healthy images, and 15 images are of DR. Naveen et al. [26] have used the HRF dataset for DR detection through image processing.

7) MESSIDOR

The 1200 fundus color images in this freely available dataset were taken at an FOV 45-degree. Images are annotated to four DR stages (Normal, Mild NDPR, Moderate NDPR, and PDR). Bhatia et al. [18], Chetoui et al. [21] and Odeh et al. [27] have used the Messidor dataset for DR detection through machine learning techniques. Its extended dataset named Messidor-2 contains 1748 images acquired at a 45-degree

TABLE 5. Systematic analysis of literature review of DL approaches.

R.No	Author /Year	Dataset	Pre-Processing	Feature Engineering	Results
1	Gardner et al./1996	Fundus images/179	Histogram Equalization, Edge Detection, and Median Filter.	A neural network was trained to recognize features of diabetic fundus images and tested for the sensitivity and specificity of recognition of these features.	Sensitivity = 88.4% and Specificity = 83.5%.
2	Sinthanayothin et al /2001	Digital retinal images/30	The RGB components of an image were transformed to an Intensity, Hue, Saturation (IHS) model. Adaptive, local, contrast enhancement was applied to the intensity band and the subsequent image was converted back to RGB for display.	An automated image analysis system with 'Moat Operator' was used to automatically detect features (HE, HM, MA) of NPDR.	HM and MA Sensitivity=77.5% specificity = 88.7EX Sensitivity =88.5 Specificity= 99.7
3	Kandala et al/2009	Stare, DiaretDB0 and DiaretDB1/ 89 (Combined 1)	Red and Green Channel Extraction, Contrast Stretching, and Median Filtering.	Pixel classification and mathematical morphology are used for feature extraction and SVM is used for classification.	Sensitivity = 100% and Specificity = 91%
4	Tang et al / 2013	Messidor / 1200	The FOV was detected automatically and the images were rescaled to 1026*681 pixels with an FOV of approximately 630 pixels in diameter across the entire dataset.	Retinal color images are partitioned into nonoverlapping splats covering the entire image. A set of features is extracted from each splat to describe its characteristics relative to its surroundings. A classifier is trained with splat-based expert annotations and evaluated on the Messidor dataset.	ROC= 0.96 is achieved at the splat level and 0.87 at the image level.
5	Liskowski et al/ 2015	Drive/40, Stare/20, Chase/28	Contrast Normalized and Enhanced.	Retinal Blood Vessels (RBV) are segmented using a Deep Neural Network.	Accuracy of classification (up to > 0.97), ROC curve measure (up to > 0.99).
6	Grinsven et al / 2016	Kaggle/ 6679, Messidor / 1200	FOV of the color fundus images are segmented, Used circular template matching, Images are cropped and resized, Contrast Enhancement	Created a CNN architecture to detect hemorrhage with nine layers formed by selective samples and 41x41 size patches labelled with or without evidence of bleeding.	Sensitivity= 84.8 and specificity 90.4 for both datasets.
7	Soniya et al / 2016	DiaretDB0/ 130	NA	Heterogenous CNN and single CNN detect and classify MAs, HEs, hard EXs, soft EXs, and NV.	Single CNN Accuracy= 95%, Heterogenous CNN Accuracy= 95%,
8	Alghamdi et al/ 2016	Drive, DiaretDB1, Messidor, Stare, Kenya, Hapiece, Pamdi, and KFSH/ 5781	NA	The model constitutes of DL architectures With CNN classifiers and optic disc abnormality assessment through feature learning, the AdaBoost ensemble algorithm has been used for feature selection and classifier training.	Accuracy of 100%, 98.88%, 99.20%, 86.71%, 99.53%, 98.36%, 98.13% and 92%, on Drive, DiaretDB1, Messidor, Stare, Kenya, Hapiece, Pamdi and KFSH.
9	Gulshan et al / 2016	EyePACS- 1 / 9963, Messidor-2/ 1748	Contrast Enhancement	Inception-V3- architecture Neural Network is trained to detect HEs and MAs for DR detection.	AUC of 0.991 for EyePACS-1 and 0.990 for Messidor-2.
10	Prentas et al /2016	DRiDB/50	Image Filtering Proposed	Model has used an ensemble of OD detection algorithms, which performs various preprocessing, thresholding, localization, and object detection. The proposed CNN model has classified an EX or a non-EX.	F-score o= 0.78.
11	Tan et al / 2017	Cleopatra/298	NA	10-layer CNN is used to automatically, simultaneously segment and discriminate EX, HM and MA. The input image is normalized before segmentation. The net is trained in two stages.	HM Sensitivity= 0.6257, MA Sensitivity= 0.4606, EX Sensitivity= 0.8758.
12	Quellec et al / 2017	Kaggle Diabetic Retinopathy/ 88,702, DiaretDB1/ 89, E- ophtha/463	Image resizing, Gaussian Filtering, Rotated, Scaled, horizontally flipped, and Contrast Enhancement.	Developed a CNN model by using the ConvNets network structure that generates heat maps to detect four forms of diabetic retinopathy lesions: MA, HM, HE, and SE.	HM AUC=0.614, HE AUC= 0.735, SE AUC=0.809, MA AUC= 0.500.

TABLE 5. (Continued.) Systematic analysis of literature review of DL approaches.

13	Lam et al / 2017	Kaggle retinopathy / 35,126, e-Ophtha/463	NA	Tested five CNN models, including AlexNet, VGG16, GoogLeNet, ResNet, and Inception-v3. The image patches are used to train convolutional neural networks to predict the existence of MA and EX. The sliding window method is used to create a probability map for the entire image.	AUC= 0.94 and ROC= 0.86 for MA, AUC=0.95 and ROC 0.64 for EX.
14	Orlando et al/2017	E-Ophtha/463, DiaretDB1/89, Messidor/1200	Green Channel Extraction, Gaussian, Filtering, Thresholding Operation	Red lesions were detected by combining both deep learning and domain knowledge. Features learned by CNN are augmented by incorporating hand-crafted features. Such an ensemble vector of descriptors identifies true lesion candidates using a Random Forest classifier.	Competition Metric is 0.4874 and 0.3683 for the DiaretDB1 and the E-ophtha datasets.
15	Takahashi et al/ 2017	Fundus Photographic set/9,939	NA	GoogLeNet DCNN was trained to classify different stages of DR.	Accuracy= 96%,
16	Rakhlin/ 2017	Kaggle DR/ 88,696, Messidor- 2/1748	Normalize, Scale, Center and Crop Images	A VGG-Net based The modified DCNN model has proposed identifying DR features such as drusen, EXs, MAs, SEs and HEs.	AUROC of Kaggle and Messidor-2 = 0.923, 0.967.
17	Gargeya /2017	Messidor- 2 /1748, E-Ophtha /463, Eye- PACS /75,137	Scaling, Cropped, Rotate, Contrast and Brightness Adjustment	Data-driven DL algorithm for deep feature extraction and image classification, using Deep Residual Learning (DRL) to develop a CNN for automated DR detection.	AUC of 0.94, 0.97 and 0.95 for Messidor-2, EyePACS and E-ophtha.
18	Chudzik et al./2018	ROC/100, E-ophtha/381, DiaretDB1 /89	Color Space Conversion, Cropping and Resizing.	Used custom CNN architecture to detect MA.	FROC (Free- response ROC) score=0.355, 0.392 and 0.562 for ROC, DiaretDB1, E-ophtha
19	Adem/2018	DiaretDB0 /130, DiaretDB1 /89, DrimDB /125	Grayscale Conversion. Then, Canny edge detection and adaptive histogram equalization functions were applied.	Detected the EX from DR images using the custom CNN with Circular Hough Transformation.	Accuracy = 99.17, 98.53, and 99.18 for DiaretDB0, DiaretDB1, and DrimDB
20	Mo et al/ 2018	E-ophtha /463, HEI- MED/ 169	NA	Detected EX lesions by segmenting and classifying the exudates using a deep residual network.	Sensitivity of 0.9227 and 0.9255 and an AUC of 0.9647 and 0.9709 for the E-ophtha and HEI-MED datasets.
21	Khojasteh et al/	DiaretDB1 and E- ophtha/463	Contrast Enhancement	A ten-layered CNN has proposed and employed patch-based and image-based analysis upon the fundus images to extract EXs, HEs, and MAs.	Sensitivity, Specificity and Accuracy EXs, HEs, and MAs are 0.96, 0.98 and 0.98, 0.84, 0.92 and 0.90, and 0.85, 0.96 and 0.94.
22	Wang et al/ 2018	Kaggle DR/ 35,126	Stretching, Rotation, Flipping and Color Augmentation	Proposed two CNN models namely Net-5 and Net-4 to detect DR.	Kappa score=0.81
23	Dai et al/ 2018	DiaretDB1/89, Local DS / 646	Resizing, Histogram Equalization	Proposed Multi Sieving CNN (MS-CNN) technique to solve an imbalanced MAs detection problem by bridging the semantic gap between fundus images and clinical data for DR detection.	Accuracy= 96%
24	Islam/2018	EyePACS /88,702	Rescaling image	Proposed a DCNN for early-stage detection of DR.	AUROC and F- Score of 0.844 and 0.743.
25	Saha et al/2019	Drishti-GS	Resizing, Horizontal and Vertical Flipping.	Automated retinal lesions and optic disk segmentation in fundus images using a deep fully CNN for semantic segmentation. This trainable segmentation pipeline consists of an encoder network, a corresponding decoder network followed by pixel-wise classification to segment MA, HM, HE, SE and OD from the background.	Sensitivity Score of OP, MA, HE, HM and SE are 0.8572, 0.0059, 0.5498, 0.0829 and 0.1823.
26	Ananda et al/2019	IDRiD/81, Messidor/1200	Rotation and Resizing.	Constructed multiple CNNs, (UNet, SegNet), and each CNN segments only a single class, i.e., MA, HM, HE, SE and OD.	Dice Coefficient= 0.86, 0.79, 0.85, 0.94 and 0.99 for HM, MA, HE, SE and OD.

TABLE 5. (Continued.) Systematic analysis of literature review of DL approaches.

27	Guo et al./2019	IDRiD/81, e-ophtha/463, DDR/13673	Resizing, Rotation, Horizontal and Vertical Flipping.	Designed a segmentation network (L-Seg) that simultaneously segments the four kinds of lesions (EX, HE, MA, SE).	AUC of IDRiD=0.65, AUC of e-ophtha=0.29, and AUC of DDR=0.32.
28	Yan et al./2019	Segmentation subchallenge/245	Cropping and Downsampling.	LocalNet and GlobalNet models are fused at the end of their decoder component. The two streams are jointly optimized, ensuring they are mutually enhanced when the lesion regions are scattered and small-scale.	Fused Model AUPR for EX and MA= 0.889 and 0.525.
29	Yan et al./2019	DiaretDB1 /89	The green channel of the images were cropped, and CLAHE enhanced them. Also, the Gaussian filter removed the noise, and a morphological method was used.	Detected Red lesions based on hybrid features, which consist of deep-learned features extracted via an improved LeNet architecture and hand-crafted features.	Sensitivity = 48.71%
30	Li et al /2019	DR Kaggle/34,124	Rescaling, Removing Color divergence and Boundary effects.	Proposed a modified DCNN with fractional max-pooling, SVM and teaching-learning-based optimization for DR classification.	Accuracy= 86%
31	Mohanty et al./2023	APTOS 2019 Blindness Detection/3662	Resizing to 224x224, Gaussian blur filter, Ben Graham procedure for cropping	Hybrid Model (VGG16 + XGBoost), DenseNet 121 used for binary classification	Hybrid Accuracy = 80% Accuracy, DenseNet 121 Accuracy = 97.30%
32	Uzair Ishtiaq et al./2023	Kaggle EyePACS/88,702	Image resizing, Data augmentation, Median filter application, Image sharpening	Hybrid approach using GraphNet124, ResNet50, and LBP; feature optimization with BDA and SCA for multi-level classification.	SVM Accuracy= 98.85%
33	P Saranya, et al./2023	IDRiD/81, DIARETDB1/130, MESSIDOR/1200 STARE/20	IBinarization, CLAHE and Morphological operations	Utilized U-Net segmentation, CNN for Lesion detection and classification	IDRiD Accuracy= 95.65%, Messidor Accuracy= 94%
34	Ahmadi et al./2024	EyePACS/88,702	Cropping, Gaussian & Sobel filtering, Rescaling, Data augmentation	DenseNet121, ResNet-101, Inception-ResNetV2, EfficientNetB0 are used for multilevel classification.	Accuracy= 97%, 96%, 95%, and 94% for DenseNet121, ResNet-101, InceptionResNetV2, and EfficientNetB0 respectively

FOV. Various authors have used the Messidor-2 dataset for DR detection.

8) STRUCTURED ANALYSIS OF THE RETINA (STARE)

Blood vessel segmentation is carried out using this freely accessible dataset. It has 20 photos that were taken with a FOV of 35 degrees. 10 of the images are normal. Several authors have used the STARE dataset for DR detection.

9) CHASE-DB1

For blood vessel segmentation, this publicly accessible dataset is offered. It has 28 photographs at a resolution of 1280 * 960 pixels that were taken with a 30-degree FOV. Several authors have used the CHASE DB1 dataset for DR detection.

10) INDIAN DIABETIC RETINOPATHY IMAGE DATASET (IDRID)

The 516 fundus images in this publicly available dataset were taken at a 50-degree FOV and labeled to five DR phases.

Guo et al. [51] and Huang et al. [55] have used the IDRiD dataset for automated DR detection.

11) RETINOPATHY ONLINE CHALLENGE (ROC) DATASET

It includes 100 retinal images taken at a 45-degree FOV and is publicly available. Its pixels range in size from 768 by 576 to 1389 by 1383. The photos were annotated to find MA. There are just training ground realities. Chudzik et al. [43] have used a dataset for microaneurysms detection.

B. PREPROCESSING

The dataset's blurred or less-than-clear photos cause issues. By minimizing unexpected distortions or increasing specific features that are essential for further processing, pre-processing seeks to enhance the image data. Different steps for pre-processing images are performed; some common steps are mentioned below.

1) GRAYSCALE CONVERSION AND IMAGE ENHANCEMENT

As processing grayscale images is more pleasant than processing color images, the images are transformed to

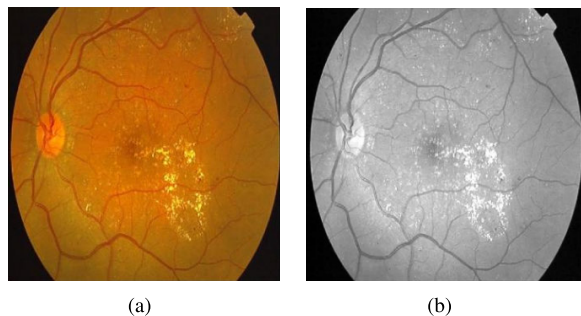


FIGURE 9. Visualization of original Vs processed fundus image.

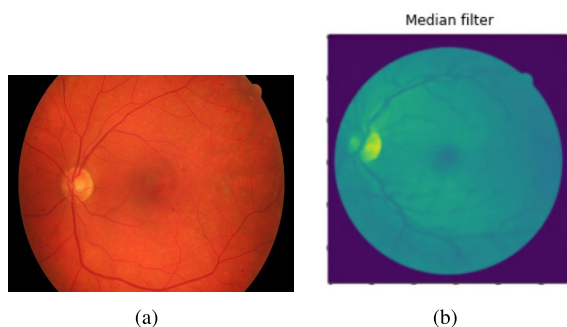


FIGURE 10. Original image Vs filtered image.

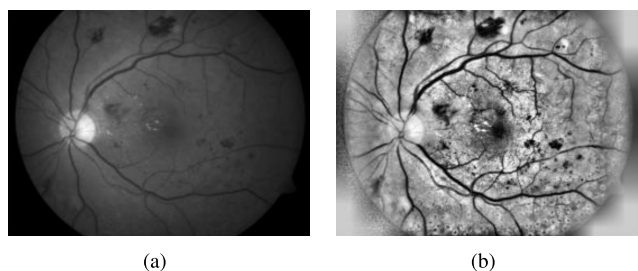


FIGURE 11. (a) Original Fundus image (b) the results after applying CLAHE technique on the original fundus image.

grayscale (Fig. 9). The Histogram equalization adds contrast to the filtered image by doing the contrast enhancement as applied in [15].

2) MEDIAN FILTERING

The adaptive median filter’s primary goals are to eliminate salt and pepper noise and smooth and minimize image distortions (Fig. 10) as applied in [15].

3) CLAHE

Contrast Limited Adaptive Histogram Equalization (CLAHE) is a type of adaptive histogram equalization. The CLAHE (Fig. 11) was created to stop the excessive noise amplification that might result from adaptive Histogram Equalization (AHE). Limiting AHE’s contrast enhancement helps achieve this, as shown in [26] and [27].

4) GREEN CHANNEL EXTRACTION

Because of the highest intensity of the color images and superior contrast than the other color channels (Red and

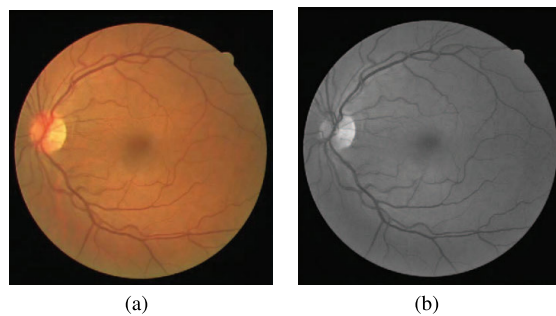


FIGURE 12. (a) Original Fundus image (b) Green Extraction of the original fundus image.

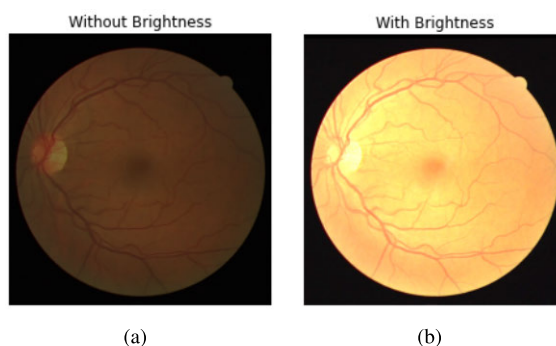


FIGURE 13. (a) Without Brightness Fundus image (b) With Brightness Fundus image.

Blue), the green level extraction of the RGB fundus images (Fig. 12) is usually performed during DR process as applied in [8].

5) BRIGHTNESS

The fundus images are often dark, so we can add brightness (Fig 13).

C. DATA AUGMENTATION

Generating new images from ones already in the dataset is known as augmentation. Data augmentation techniques would increase the number of photos (Fig 14) i.e., vertically and horizontally rotating, flipping, cropping, and resizing.

D. FEATURE ENGINEERING

The different feature extraction methods applied to fundus images are discussed here.

1) TEXTURE FEATURES

The texture is a feature used to divide and classify regions of interest in photographs. The way that colors or intensities are distributed spatially in an image is revealed by texture. The spatial distribution of intensity levels within a neighbourhood defines the texture. Texture features may be beneficial for illustrating certain local patterns that repeat themselves and arranging regularity in particular areas of photos. It could provide defining metrics like smoothness, roughness, and

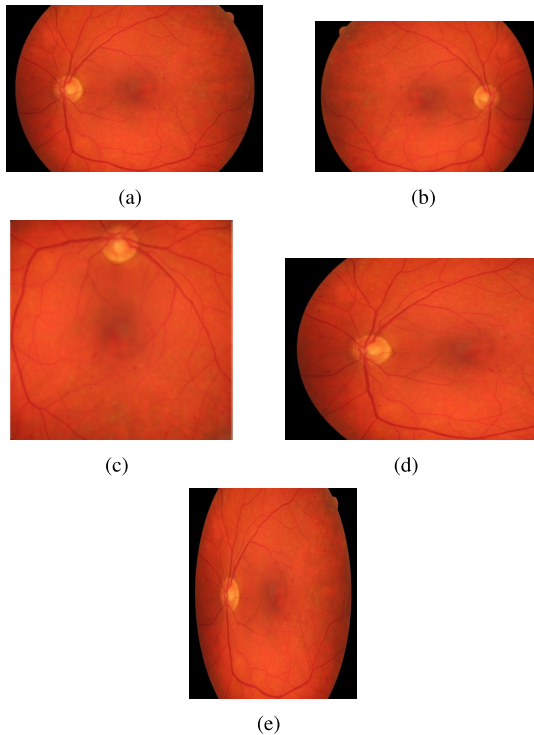


FIGURE 14. (a) Original Fundus image (b) Vertically Flipped fundus image (c) Rotate fundus image (d) Cropped fundus image (e) Resized fundus image.

regularity, which could describe the local properties of images [21], [73].

2) GLCM FEATURES

Considering the spatial relationship between pixels, the grey-level co-occurrence matrix (GLCM), also known as the grey-level spatial dependence matrix, is a statistical method for analyzing texture. The GLCM functions calculate the frequency of pairs of values and spatial relationships in an image, creating a GLCM. From this matrix, statistical measures are then extracted to describe an image's texture [74], [75].

3) WAVELET FEATURES

A wavelet mathematical function is utilized in digital signal processing and image compression. Its primary goal is to enhance image quality. Wavelets can also separate signals into their temporal and frequency components. A signal is broken down into its frequency components using a wavelet transform. One of the fundamental advantages of wavelets is the simultaneous localization in the domain of time and frequency. The second important advantage of the wavelet transform is its extraordinarily rapid computation speed. One of the main advantages of wavelets is their capacity to identify the smallest details in a signal [76].

4) COLOUR FEATURES

Color is the most useful visual element and has been extensively utilized in picture retrieval systems. Three-

dimensional color spaces are typically used to define colors. Red, Green, and Blue (RGB), Hue, Saturation, and Value (HSV), or Hue, Saturation, and Brightness (HSB) are some examples of these. The most popular method for displaying color features is using color histograms. The RGB color area is used to hold image data in the majority of image formats, including Joint Photographic Experts Group (JPEG), Bitmap (BMP), and Graphics Interchange Format (GIF) [77].

5) HISTOGRAM FEATURES

The original image can be utilized to retrieve the features of a histogram. The histograms are processed to produce the meta-features that the semantic mapper uses as an input to create the semantic features [78], [79]. The image histogram depicts the grey level distribution in a photo in two dimensions. A histogram is merely a graphic representation. It reveals the image's optical composition. The amount of light and darkness in an image is what is meant by optical content. Using statistical torques associated with the histogram of the intensity of an image or region is one of the simplest ways to characterize a texture, and numerous features can be retrieved from it [80].

6) EDGE FEATURES

The margins of the two consecutive grey levels or the brightness values of the two pixels, which occur at a specific point in the image, serve as the boundaries between an object and its background. Edge detection aims to find the areas of an image where the light intensity abruptly changes. Vision may be possible at the edges. They can alter depending on the point of view and, typically, the scene's geometry, the items that cross paths with one another, and so on. They can also be influenced by perspective, which typically depicts the characteristics of the things seen, such as marks and surface shape. This action can be done by several operators like Sobel [81], Pewit et al. [82], Armi and Fekri-Ershad [80]

E. DETECTION AND GRADING ALGORITHMS

The number of algorithms used to detect and classify DR. Some commonly used algorithms are discussed below.

1) ARTIFICIAL NEURAL NETWORK

By having the ability to recognize patterns in data, an artificial neural network (ANN) computer program can be said to replicate the analytical functions of the human brain. An input layer, an output layer, and a hidden layer comprise the three layers of the ANN. The hidden layer comprises several nodes coupled to the input and output nodes by mathematical formulas (weights). The network is trained by repeatedly exposing it to data instances with known results, known as an iteration [83]. After training has changed the hidden layer's weights, the network can be presented with an unknown input and classified into the proper output. It can identify handwriting or speech patterns and forecast financial or

economic trends [28].

$$Y = W1X1 + W2 \times 2 + b \quad (1)$$

In equation (1), X1 and X2 feature, W1 and W2 are weights. The weights of the features are multiplied by one another, and the bias is added up. Following the application of this summation function over an activation function, the output of this neuron is multiplied by the weight W3 and sent as input to the output layer. ANN-based models have been integrated with deep learning layers to improve pattern recognition in fundus images, as seen in the work by Gardner et al., where preprocessing steps like histogram equalization and edge detection helped train ANNs for haemorrhage and exudate detection.

2) SUPPORT VECTOR MACHINE

Support Vector Machines (SVMs) are supervised learning methods for detecting classification, regression, and outliers [84]. SVMs aim to find the optimal separating hyperplane, maximising the margin between different classes. The data points that lie closest to the hyperplane and which determine its position are known as support vectors. This reliance on only the closest points makes SVMs robust to outliers and capable of producing generalized models that avoid overfitting. The decision function of an SVM is defined by the inner products of the input features and the support vectors, weighted by the support vectors' coefficients. The kernel function, which can be linear, polynomial, radial basis function (RBF), or sigmoid, plays a critical role in this process. It enables the SVM to access higher-dimensional space without directly computing its dimensions, allowing it to fit the maximum-margin hyperplane in a transformed feature space. The choice of kernel and its parameters can significantly influence the performance of an SVM classifier, making model selection and tuning an important part of using SVMs effectively [18], [24]. Recent advancements in SVM for diabetic retinopathy detection have explored advanced feature selection techniques and kernel optimization. An enhanced SVM model, leveraging feature extraction from retinal images, showed improved performance in classifying DR stages, as evidenced by Acharya et al., who achieved 82% sensitivity and 88% specificity.

$$Y = w \cdot x + b \quad (2)$$

In equation (2), W is the normal direction of the plane and b represents the threshold.

3) K-NEAREST NEIGHBORS

The k-nearest neighbours (KNN) is a straightforward and simple-to-implement algorithm that is employed to address classification and regression issues [85]. The function generates a suitable result when fresh, unlabeled input is provided to the algorithm. KNN quantifies similarity, which is frequently referred to as distance/proximity/closeness [22]. Recent advancements in k-NN for diabetic retinopathy

detection have incorporated feature engineering to improve classification accuracy, as demonstrated by Niemeijer et al., who used k-NN in conjunction with other machine learning techniques to classify retinal images. Finding which data points are closest to a certain query point requires figuring out the distance between the query point and the other data points. This can be calculated by Euclidean distance eq (3).

$$d(p, q) = \sqrt{\sum_{i=1}^n (q_i - p_i)^2} \quad (3)$$

4) RANDOM FOREST

Random Forest is an ensemble learning method that operates by constructing a multitude of decision trees at training time and outputting the class that is the mode of the classes (classification) or mean prediction (regression) of the individual trees [86]. It combines the simplicity of decision trees with flexibility, resulting in a vast improvement in accuracy. Random Forests create a forest of decision trees, usually through the bagging method. Instead, the best split among a random subset of features is chosen. This strategy of selecting a subset of features at each split point adds an extra layer of randomness to the model, which helps in increasing the diversity among the trees in the model, leading to more robust overall predictions. Random Forests are known for their high accuracy, robustness, and ease of use. By looking at how much the accuracy decreases when a feature is excluded, one can gauge the significance of each feature in the prediction process [22]. Advanced implementations of Random Forest algorithms have been applied to diabetic retinopathy classification, utilizing ensemble learning to enhance predictive accuracy. Studies have focused on optimizing tree structures and integrating RF with deep learning frameworks for better retinal image classification.

5) NAIVE BAYES

The "naive" assumption that each pair of features is conditionally independent given the value of the class variable underlies a class of supervised learning algorithms collectively referred to as "naive Bayes methods." Both binary and multiclass classification benefit from this method of categorization [87]. Naive Bayes performs better when dealing with categorical input variables than it does when dealing with numerical input variables. Making forecasts based on historical data and anticipating data are both aided by it. Enhanced Naive Bayes classifiers, integrating image preprocessing and feature extraction methods, have shown improved performance in detecting diabetic retinopathy, with studies indicating the use of NB in conjunction with neural network frameworks to leverage probabilistic classification alongside deep learning's feature learning capabilities.

6) K-MEAN CLUSTERING

The K-Means algorithm extracts the characteristics from the retinal images. K-Means determines which cluster employs the feature extraction technique [88]. When this algorithm

is used to process the information gathered about blood vessels, it can benefit grading DR severity [16]. In diabetic retinopathy detection, k-means Clustering has been utilized for segmenting retinal images, aiding in identifying pathological features. Recent innovations include adaptive clustering algorithms that dynamically adjust to the specific characteristics of retinal photos, improving lesion detection.

7) LOGISTIC REGRESSION

The statistical regression model known as logistic regression uses a categorical output variable rather than a continuous one. The output variable's categorical nature may represent separate classes. One vs. all is the algorithm used for multi-class categorization. The outcome of logistic regression is the probability of whether the example falls into category 0 (No DR) or 1 (DR), with the features being fed into the trained hypothesis [18]. Logistic Regression has been refined with machine learning techniques to better model the probability of occurrence of diabetic retinopathy, with innovations including the use of regularized Logistic Regression to handle high-dimensional data from retinal images, preventing overfitting and improving model generalizability.

8) CONVOLUTIONAL NEURAL NETWORK

Convolutional Neural Networks (CNNs) excel in capturing hierarchical patterns in data, making them especially suited for image recognition and processing tasks. Central to CNNs, the convolution operation involves sliding a filter or kernel over the input data to produce a feature map that summarizes the presence of detected features in the input. This process allows CNNs to automatically and adaptively learn spatial hierarchies of features from input images, from simple edges to more complex patterns, by stacking multiple convolutional layers, each learning to recognize increasingly complex features. After several convolutional and pooling layers, the high-level reasoning in the neural network is done through fully connected layers, where all neurons from the previous layer are connected to each neuron [89]. The ability of CNNs to learn from vast amounts of data and their efficiency in recognizing patterns across different contexts and domains make them a cornerstone of modern AI applications [90]. CNNs have been at the forefront of recent advances in diabetic retinopathy detection. Studies have shown that deep CNN models, pre-trained on large datasets and fine-tuned for specific DR features, offer superior accuracy in detecting and grading diabetic retinopathy.

9) AUTOENCODERS

Autoencoders are used to learn efficient coding of unlabeled data, typically for dimensionality reduction or feature learning. An autoencoder learns to compress (encode) the input into a lower-dimensional latent space and then reconstruct (decode) the input from this latent representation, effectively learning a compact representation of the input data [91]. The key idea is to train the network to minimize the reconstruction

error, encouraging the autoencoder to capture the most salient features of the data in the encoding. Autoencoders consist of an encoder, a decoder, and a loss function that measures the difference between the input and its reconstruction [92]. Autoencoders are used for unsupervised feature learning in diabetic retinopathy diagnosis. They effectively reduce dimensionality and extract meaningful features from fundus images, enhancing DR detection and grading, as demonstrated in the research, combining them with other deep learning models for improved accuracy.

10) RESTRICTED BOLTZMANN MACHINES (RBMS)

RBMs are a class of generative stochastic artificial neural networks that can learn a probability distribution over its set of inputs. RBMs are composed of visible and hidden layers with bidirectional connections between them but no connections within a layer, making them a type of bipartite graph. Unlike general Boltzmann machines, they are "restricted" because they do not allow intra-layer communication, simplifying the learning algorithm [93]. Their ability to represent complex distributions and capture correlations between variables makes them valuable for deep learning and unsupervised learning tasks [94]. RBMs are explored for their potential in unsupervised feature learning in diabetic retinopathy detection, with recent approaches using RBMs to pre-train layers of deep neural networks, enhancing the feature extraction process and leading to more accurate classification of disease stages.

F. PERFORMANCE METRICS

The usefulness of the suggested methods can be assessed through the various metrics. Measures employ a variety of common words, including True positive, True negative, False positive, and False negative. True positive (TP) refers to the circumstance in which a test is positive, and an individual can identify the disease. When the test is negative and an individual is not given a disease diagnosis, the condition is known as a true negative (TN). False positive (FP) is the circumstance in which a test result is positive, but an individual cannot show it. False negatives (FN) occur when a result is negative, yet an individual can have it. The used performance metrics are mentioned below.

1) ACCURACY

Accuracy refers to the degree of agreement between a quantity value that has been measured and the actual quantity of a measurand (i.e., the quantity that is being measured). Accuracy can be measured by equation (iv).

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

2) RECALL

A recall metric counts the percentage of correct positive predictions among all potential positive guesses. A recall is also known as sensitivity. The recall is

measured by equation (v).

$$Recall = \frac{TP}{TP + FN} \quad (5)$$

3) PRECISION

A measure of precision counts how many correctly positive forecasts were made. Precision is measured by equation (vi).

$$Precision = \frac{TP}{TP + FP} \quad (6)$$

4) F-MEASURE

The F-score or F-measure tells you how accurate a test is. It is determined using the test's recall and precision. It can be measured by equation (vii).

$$F - Measure = 2 * \frac{Precision * Recall}{Precision + Recall} \quad (7)$$

5) SPECIFICITY

The ability of the proposed model to estimate true negatives for each accessible category. It can be measured by equation (viii).

$$Specificity = \frac{TN}{TN + FP} \quad (8)$$

6) AREA UNDER CURVE

AUC stands for the area under the curve. Classification analysis is done to identify the model that most accurately predicts the classes. ROC (Receiver Operating Characteristics) curves are one application of it.

The entire two-dimensional region beneath the full ROC curve from (0,0) to (1,1) is measured by AUC. References [38] and [39] used AUC to evaluate retinal lesion detection. Gulshan et al. [35] used the F measure as an evaluation metric with Accuracy, Precision, and Recall. Kande et al. [30] and Grinsven et al. [33] used Sensitivity and Specificity to assess proposed DR detection and grading methods.

IV. OPEN PROBLEMS AND DISCUSSION

In this section, we have conducted an in-depth analysis of current methodologies, identifying their limitations and discussing the unresolved challenges in existing approaches to diabetic retinopathy detection. The two central classification schemes employed in the most recent studies on identifying diabetic retinopathy are binary and multi-level classifications. Binary classification refers to classifying a retinal image as either DR Present or DR Absent, but multi-level classification may contain several labellings, such as mild, severe, and others. Another well-known classification method is lesion-based classification, which identifies Diabetic Retinopathy lesions (MA, HM, HE, or SE). Machine learning and deep learning methods are effective in multiple applications [95], [96], [97].

A. EXISTING METHODS

1) BINARY CLASSIFICATION

Fundus images of the retina were classified into DR or No DR using CNN by K. Xu et al. The study employed about 1000 photographs, which were rescaled before being used as input to the CNN. A CNN model with 8 convolutional layers and 4 maximum pooling layers was used. The final layer for classification was a softmax layer. A 94.5% accuracy rate was attained [97]. Another [98] classified photographs as usual or DR using the ResNet 34. The image quality was enhanced using methods such as the Gaussian filter, weighted addition, and image normalization, and their dimensions were set at 512*512. The accuracy rate was 85%. Two CNN models were utilized by Zago et al., one of which was built from scratch and the other of which was trained using VGG16. These models distinguished between non-red and red lesions in photos. AUC was evaluated with a result of 0.912 [99].

2) MULTI-LEVEL CLASSIFICATION

SVM for Multi-Level Diabetic Retinopathy Classification was proposed by Kandhasamy et al. utilizing local binary patterns; image features are retrieved and sent as input to the SVM. The average accuracy was 99.3% [99]. A softmax layer was suggested [100] as the classifier in a CNN-based technique. To feed the CNN, images were downsized to 512*512. Several regularization techniques were applied to decrease overfitting in the CNN. A 75% accuracy rate was attained. Wang et al. employed only 166 images to perform multi-level DR classification utilizing three available CNN models: Inception Net V3, Alexnet, and VGG16. Accuracy levels of 63.23%, 37.43%, and 50.03%, respectively, were reached [101]. For the multi-label stage-wise categorization of retinal pictures, a different method based on a bag of words was presented [102]. The multi-level classification was then done using three different classifiers: SVM, Random Forest, and Multinomial Logistic Regression. A 72%, 73%, and 68% accuracy rate was attained, respectively. Instead of categorizing retinal images according to the severity of the disease, Sadek et al. categorized retinal images according to lesions (normal, exudates, and drusen). BoVW, VGG, VGG-VD, GoogLeNet, and ResNet were applied for classification. The achieved accuracy was 77.76%, 91.83%, 90.76%, 92.00%, and 91.23% [103]. Harangi et al. pre-trained AlexNet with several new features for multilevel DR classification. A 90.07% accuracy was attained [104]. Another study in [105] produced a data set of 13,767 pictures for DR categorization. On a scale of 1 to 4, each image was categorized. The photographs were cropped and scaled before being provided to the models. Four CNNs were used, and further FC layers were placed on top.

B. LIMITATIONS

Detection and classification methods of DR can help identify and treat the severity level of DR, but there are also limitations to these methods. Here are some of the main limitations:

TABLE 6. Limitations from the literature review.

Ref#	Limitations
[9]	Limited dataset size (5 images) might not be sufficient to generalize the method's effectiveness across diverse retinal conditions.
[10]	While high sensitivity was achieved, the specificity varied greatly, indicating potential issues with false positives in retinopathy detection.
[11]	The method relies heavily on the quality of image preprocessing steps, which might affect performance in diverse real-world scenarios.
[12]	The study used SIFT for feature extraction, which may not capture all relevant features of retinal images for DR detection effectively.
[13]	The study's approach might be limited by the image quality and the effectiveness of CLAHE in enhancing the contrast of retinal blood vessels.
[14]	The dataset used (40 images from the DRIVE database) is quite small, which may not provide a comprehensive validation of the algorithm.
[15]	The method might struggle with differentiating between similar intensity levels of different retinal features, leading to potential misclassification.
[16]	The varying accuracy rates across different classifiers suggest inconsistency in feature extraction and classification robustness.
[17]	The study's reliance on lesion detection algorithms might limit its ability to accurately classify DR stages in less clear or more complex cases.
[18]	The method's performance is highly dependent on the success of individual classifiers, which may not consistently handle diverse image conditions.
[19]	The use of only 100 images for training may not be sufficient for an SVM to effectively learn the variability of DR features effectively.
[20]	The approach achieved relatively low FROC scores, indicating room for improvement in accurately identifying microaneurysms across datasets.
[21]	The method's accuracy and AUC varied significantly between the textural features used, suggesting inconsistency in feature detection efficiency.
[22]	The maximum accuracy of 82% indicates potential limitations in the hybrid model's ability to accurately identify and classify DR features.
[16]	The study indicates SVM as the best method but does not provide a comprehensive comparison or validation across a sufficiently large and varied dataset.
[23]	The study may have a limited dataset size (124 images), which could affect the generalizability and robustness of the findings.
[24]	Sensitivity is relatively lower at 82%, indicating potential room for improvement in detecting all stages of DR effectively.
[25]	The method relies on manual selection of regions, which may introduce bias and limit scalability for larger datasets.
[26]	The study uses a very small dataset (10 images), raising concerns about the statistical significance and generalizability of the results.
[27]	Achieved an accuracy of 75.1%, which is relatively low, indicating that the model might struggle with the complexity of DR classification or the diversity of the dataset used.
[28]	The study used a limited number of images (179), which may affect the model's ability to generalize across diverse DR conditions.
[29]	Sensitivity and specificity metrics show a large variance, indicating potential inconsistency in lesion segmentation.
[30]	Although the method showed high sensitivity, it was tested on a relatively small composite dataset, questioning its scalability.
[31]	The method's reliance on specific image preprocessing and resizing may limit its applicability to images with different characteristics.

TABLE 6. (Continued.) Limitations from the literature review.

[32]	The study focuses on blood vessel segmentation without directly addressing the broader spectrum of DR symptoms.
[33]	Hemorrhage detection sensitivity varied significantly between datasets, suggesting potential model overfitting or underfitting.
[34]	The study compares two CNN models on a small dataset, which may not provide a comprehensive view of their effectiveness.
[35]	Despite high AUC scores, the method's effectiveness in real-world clinical settings remains to be fully proven.
[36]	Sensitivity and Positive Predictive Value are moderate, indicating room for improvement in EX detection accuracy.
[37]	The model shows variable sensitivity across different lesions, with particularly low sensitivity for detecting microaneurysms.
[38]	Some AUC values for lesion detection are low, indicating challenges in reliably identifying specific DR features.
[39]	While the model tested multiple CNN architectures, the performance varied, highlighting challenges in model selection for lesion detection.
[40]	The approach shows moderate Competition Metrics (CPM), suggesting potential limitations in the segmentation and feature classification.
[41]	Although the model shows promising AUROC, sensitivity and specificity rates vary, suggesting inconsistencies in lesion detection.
[42]	The method's deep feature retrieval and classification process requires validation across more diverse and extensive datasets.
[43]	The study presents a unique CNN architecture but achieves only moderate FROC scores, indicating potential for improvement in MA detection.
[44]	While the model shows high accuracy in EX detection, the complexity and scalability of the CNN approach need further examination.
[45]	Despite high sensitivity and AUC, the method's performance in distinguishing fine vessels and dealing with central vessel reflex is limited.
[46]	The study suggests simultaneous feature detection but requires further validation to confirm its efficiency and reduce error rates.
[47]	The proposed model shows significant variability in performance across stages, indicating challenges in consistent lesion detection.
[48]	Despite a comprehensive dataset, the model's AUROC and F-Score suggest that there is room for improvement in MAs detection.
[49]	The method achieves varied sensitivity scores for different lesions, indicating a need for improved detection accuracy.
[50]	The adjusted U-Net model shows promise, but the dice coefficient values indicate potential for enhanced segmentation accuracy.
[51]	Although the method addresses small area segmentation, the overall AUC for hemorrhage detection suggests improvement is needed.
[52]	The study combines different approaches for red lesion detection but shows low sensitivity, highlighting the need for optimization.
[53]	Utilizing a modified LeNet architecture, the study indicates moderate sensitivity, suggesting potential limitations in lesion detection.
[54]	The method's FROC scores are moderate, pointing to the need for enhanced accuracy in MA detection and DR diagnosis.
[55]	While innovative in annotation accuracy, the method's IoU value suggests that there is potential for improving hemorrhage detection.
[56]	Despite achieving high sensitivity and AUC, the method's reliance on specific CNN and classifier combinations may limit generalizability.

TABLE 6. (Continued.) Limitations from the literature review.

[57]	The model shows high AUC on specific datasets, but the overall effectiveness in diverse clinical scenarios is not fully established.
[58]	High AUC values are achieved, yet the model's performance across varied datasets and in practical clinical applications needs validation.
[59]	The study demonstrates the effectiveness of UNet and CNN for lesion detection, yet comprehensive performance metrics across datasets are not detailed.
[60]	The use of multiple datasets for training might introduce variability in image quality and annotations, potentially affecting the model's consistency across different data sources.
[61]	The modified and randomly initialized GoogLeNet DCNN might face challenges in generalizing across different stages of DR due to the complex nature of the disease's progression.
[62]	Despite using large datasets, the model may require further optimization to improve the Kappa score and ensure more accurate DR detection.
[63]	The pure DCNN approach might struggle with overfitting, given the large number of training images and the complexity of the task.
[64]	While achieving high accuracy, the model's performance in real-world clinical settings needs validation to confirm its effectiveness in multi-level DR classification.
[65]	The suitability of AlexNet for DR detection may not hold across all datasets or in comparison with more advanced CNN architectures.
[66]	The model relies on transfer learning and deep feature aggregation, which might limit its ability to identify novel or subtle DR features not present in the training data.
[67]	While DenseNet 121 showed high accuracy, the study's comparison between different models and their training complexity requires deeper investigation.
[68]	The hybrid method's reliance on a novel CNN model and feature optimization algorithms may introduce complexity and challenges in model interpretation and clinical application.
[69]	The study, despite evaluating multiple models, may still face challenges in real-world application, particularly in consistently detecting all stages of DR across diverse patient populations.

- 1) **FOV:** Most datasets comprise one- or three-field colored images. Therefore, The lesions outside these fields are not seen [106].
- 2) **Grading:** Numerous studies concentrated on the binary grading of DR. However, this is not helpful because it cannot provide information related to the severity level of diabetic retinopathy. It is needed to allow an ophthalmologist to intervene in cases of high severity of DR.
- 3) **Dataset:** Since the datasets were so small, the recommended approaches would not perform well in practical applications.
- 4) **Manual Annotation:** Manual annotations of features in handcrafted-based image classification methods produce false results.
- 5) **Availability of trained professionals:** The accuracy of detection and classification of diabetic retinopathy relies on the expertise of medical professionals trained

to identify the condition. In areas with a shortage of trained professionals, accurate diagnosis and treatment may be limited.

- 6) **Reliance on imaging:** Imaging methods like fundus photography and OCT are frequently used to detect diabetic retinopathy. However, elements including patient compliance, eye movements, and other eye problems might impact the quality of these photographs.
- 7) **False positives and false negatives:** Detecting and classifying DR can sometimes result in false positives, where a patient is incorrectly diagnosed with the condition, or false negatives, where a patient with the condition is not identified.
- 8) **Cost:** Some imaging techniques used to detect and classify diabetic retinopathy can be costly, which can limit access to diagnosis and treatment for patients who cannot afford them.
- 9) **Invasive procedures:** In some cases, more invasive procedures, such as retinal fluorescein angiography may be required for accurate diagnosis, which can be uncomfortable and carry some risks.

In Table 6, the limitations have been identified by using a literature review that is mentioned with cited references.

C. ADDRESSING ETHICAL CONSIDERATIONS

Ethical considerations entail the actions and decisions prioritizing fairness, transparency, and respect for individuals and communities. To uphold ethical standards, practices are typically reviewed and reassessed regularly. This includes seeking informed consent, maintaining confidentiality, and implementing clear guidelines. In this section, we address crucial ethical considerations related to the use of publicly available diabetic retinopathy datasets. Ensuring patient privacy is paramount; datasets are typically anonymized to protect individuals' identities. Consent is another cornerstone, often obtained before data collection, ensuring patients are informed about the research purposes. Acknowledging potential dataset biases related to age, ethnicity, or disease stage is essential, as they could affect the model's performance and generalizability. Researchers should strive for diverse data representation and be transparent about limitations, fostering trust and ethical integrity in medical AI research.

D. FUTURE TRENDS

- 1) Expanding the collection of high-resolution retinal images will significantly aid in training more accurate and sophisticated machine learning models. Higher-resolution images provide more detailed information about the retina, allowing for better feature extraction and improved diagnosis accuracy.
- 2) Enhancing classification performance by integrating dynamic features (like temporal changes in the retina) with traditional hand-engineered features and modern

deep learning-based non-hand-engineered features will provide a more comprehensive analysis, improving diagnostic precision [107].

- 3) Given the rapid impact of DME on vision, future research should prioritize its early detection. Advanced imaging techniques and predictive analytics can be crucial in identifying patients at risk of DME and facilitating timely intervention and treatment.
- 4) Incorporating clinical diagnostic measures such as STARD (Standard for the Reporting of Diagnostic Accuracy Studies) and DRRI (Diabetic Retinopathy Risk Index) in the evaluation of DR classification models can provide clinicians with numerical assessments that reflect a range of disease detection thresholds, thereby enhancing clinical decision-making [107].
- 5) Developing new color space and appearance models will improve the characterization of complex visual patterns in retinal images, aiding in detecting subtle signs of diabetic retinopathy and enhancing the overall diagnostic process.
- 6) An entirely automated, real-time, standardized method that can be used with retinal fundus image databases of any size is required to identify DR in the eye with a higher level of precision and a reduced false positive rate [108].
- 7) Advancements in image processing techniques that can effectively handle low-resolution and blurry images will make it possible to utilize a broader range of retinal images for DR detection, including those not of optimal quality, thus extending the reach of DR screening programs [109].

E. DISCUSSION

Detecting and classifying retinopathy requires timely intervention and treatment, as this condition can cause blindness if not treated on time. Different methods for detecting and classifying DR have been proposed, including fundus photography, optical coherence tomography (OCT), and fluorescein angiography. These techniques are non-invasive and relatively quick to perform, making them suitable for screening large populations. We have discussed different DR pipelines used in ML and DL approaches in systematic literature reviews for detecting and grading DR. In the current review, 60 publications were reviewed. Machine learning and deep learning approaches were used in all of the research discussed in the current paper to manipulate the detection of diabetic retinopathy. Due to the rise in diabetic patients, the necessity for trustworthy diabetic retinopathy screening technologies has recently become a significant concern. The issue of choosing reliable features for ML is solved by using DL in DR detection and classification, but it requires a large amount of training data. Most research used data augmentation to enhance the quantity of photos and avoid overfitting during the training phase. This study discussed in this work utilized public datasets for 72% of the cases, private

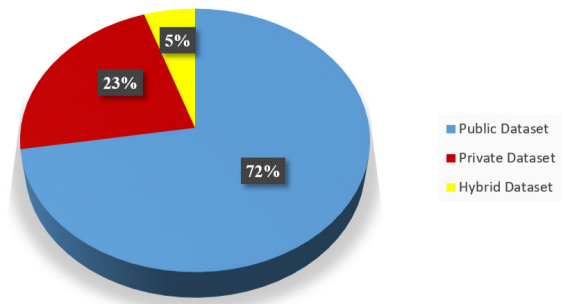


FIGURE 15. The percentage of studies that used public and private datasets.

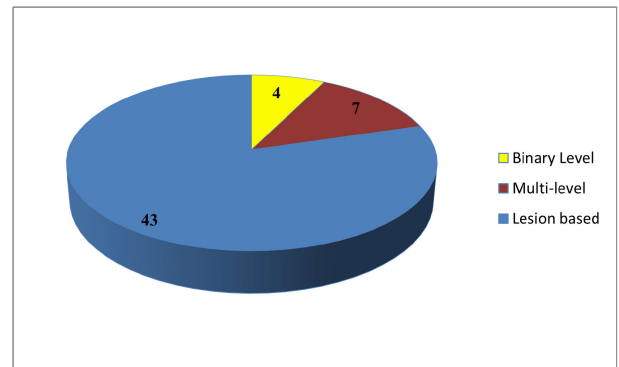


FIGURE 16. The percentage of studies based on classification methods.

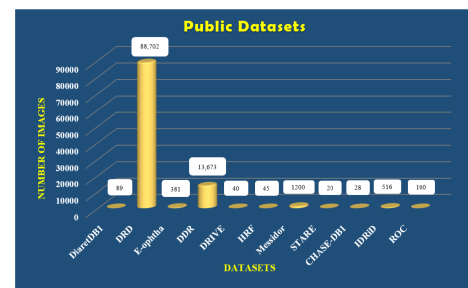


FIGURE 17. Number of images in public datasets.

datasets for 23%, and a combination of private and public datasets for the remaining 5% (Fig 15).

The DR classification is broadly classified into binary, multilevel, and lesion-based classifications. We split the selected studies as per their classification criteria, as depicted in Fig 16.

Further, the public datasets used in these approaches consist of a different number of images (refer to Fig 17)

The different techniques of ML and DL for DR detection, such as SVM, KNN, CNN, etc., have been discussed. Various features of DR have been extracted, including Microaneurysms, Haemorrhages, Exudates, Optic disc, etc. Different metrics have been used to evaluate methods like Accuracy, Specificity, Sensitivity, F-Score, and AUC (Fig 18).

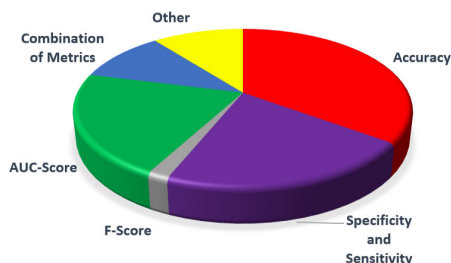


FIGURE 18. Percentage of used metrics in studies.

35% of the studies used accuracy, 21% used specificity and sensitivity, 21% used AUC -Score, and 2% used F-score for the evaluation of their methods while 10% of them used a combination of these metrics. The remaining 11% used other metrics like Dice-coefficient, Kappa, and AUPR. Advances in artificial intelligence (AI) and machine learning have shown promise in improving the detection and classification of diabetic retinopathy. AI algorithms can be trained on large datasets of retinal images to automatically detect and classify signs of diabetic retinopathy, reducing the reliance on human interpretation and potentially improving accuracy and reliability. However, further research and validation are needed before these AI techniques can be widely adopted in clinical practice.

The transition from traditional ML techniques to advanced DL models marks a significant advancement in DR detection, offering automated feature learning from retinal images. Integrating these technologies into clinical practice remains challenging, necessitating further research to ensure their efficacy and reliability in real-world settings. Future work should enhance AI interpretability, explore multimodal data integration, and validate these tools through extensive clinical trials to establish their utility in ophthalmology. In generic, addressing the limitations of existing studies and exploring these future trends, research in diabetic retinopathy detection and grading can move towards more effective, efficient, and patient-centred solutions.

V. CONCLUSION

A significant complication of diabetes mellitus, diabetic retinopathy causes progressive retinal deterioration and has the potential to cause blindness. Early detection and treatment are essential to prevent it from deteriorating and causing retinal damage. Automated systems significantly shorten the time needed to make diagnoses, saving ophthalmologists time and money and enabling prompt patient treatment. The stages of DR are determined by the kind of lesions that develop on the retina. Computer-aided diagnosis methods based on machine learning and deep learning have been created over the past few years to identify and categorize diabetic retinopathy. We systematically reviewed these algorithms. We have described the publicly accessible common fundus DR datasets. The pipelines used, including pre-processing and data augmentation steps, feature extraction methods, detection and grading algorithms for DR, and performance

metrics, are discussed in detail. The main benefit of this SLR is that it helps the scientific community build a strong, completely automated framework for the early identification of DR disease by highlighting the advantages and disadvantages of various machine learning and deep learning-based existing techniques. We have discussed some limitations to these approaches, including inter-observer and intra-observer variability, reliance on imaging, costly techniques, and invasive procedures. These limitations can affect the accuracy and reliability of the diagnosis and classification of diabetic retinopathy and may lead to missed early intervention and treatment opportunities. Finally, we conclude with the future trends of DR. In conclusion, it's essential to emphasize the transformative potential of ML and DL in DR detection, which not only streamlines the diagnostic process but also opens avenues for more personalized patient care. Future endeavours should pivot towards integrating these systems within clinical workflows, enhancing their accessibility and utility. The anticipated advancements in AI could significantly democratize eye care, making early DR detection more ubiquitous and impactful across diverse healthcare settings.

ACKNOWLEDGMENT

The authors like to thank the anonymous reviewers for their constructive comments on the original manuscript.

REFERENCES

- [1] (2023). *Human Eye Structure*. BYJU, India. Accessed: Nov. 11, 2023. [Online]. Available: <https://byjus.com/physics/structure-human-eye-functioning/>
- [2] D. S Fong, "Retinopathy in diabetes," *Diabetes Care*, vol. 27, no. 1, pp. s84–s87, 2004.
- [3] *Global Report on Diabetes*, World Health Organization, Geneva, Switzerland, 2016, pp. 1–88.
- [4] T. Aslam, P. Chua, M. Richardson, P. Patel, and M. Musadiq, "A system for computerised retinal haemorrhage analysis," *BMC Res. Notes*, vol. 2, no. 1, p. 196, 2009.
- [5] L. Guariguata, D. R. Whiting, I. Hambleton, J. Beagley, U. Linnenkamp, and J. E. Shaw, "Global estimates of diabetes prevalence for 2013 and projections for 2035," *Diabetes Res. Clin. Pract.*, vol. 103, no. 2, pp. 137–149, Feb. 2014.
- [6] I. Kononenko, "Machine learning for medical diagnosis: History, state of the art and perspective," *Artif. Intell. Med.*, vol. 23, no. 1, pp. 89–109, Aug. 2001.
- [7] X. Zhang, G. Thibault, E. Decencière, B. Marcotegui, B. Lay, R. Danno, G. Cazuguel, G. Quellec, M. Lamard, P. Massin, A. Chabouis, Z. Victor, and A. Erginay, "Exudate detection in color retinal images for mass screening of diabetic retinopathy," *Med. Image Anal.*, vol. 18, no. 7, pp. 1026–1043, Oct. 2014.
- [8] S. D. Kasurde and S. N. Randive, "An automatic detection of proliferative diabetic retinopathy," in *Proc. Int. Conf. Energy Syst. Appl.*, Oct. 2015, pp. 86–90.
- [9] T. Walter and J.-C. Klein, "Automatic detection of microaneurysms in color fundus images of the human retina by means of the bounding box closing," in *Proc. ISMDA*, 2002, pp. 210–220.
- [10] D. Usher, M. Dumskyj, M. Himaga, T. H. Williamson, S. Nussey, and J. Boyce, "Automated detection of diabetic retinopathy in digital retinal images: A tool for diabetic retinopathy screening," *Diabetic Med.*, vol. 21, no. 1, pp. 84–90, Jan. 2004.
- [11] M. Niemeijer, B. van Ginneken, S. R. Russell, M. S. A. Suttorp-Schulten, and M. D. Abramoff, "Automated detection and differentiation of drusen, exudates, and cotton-wool spots in digital color fundus photographs for diabetic retinopathy diagnosis," *Investigative Ophthalmol. Vis. Sci.*, vol. 48, no. 5, p. 2260, May 2007.

- [12] N. Silberman, K. Ahrlich, R. Fergus, and L. Subramanian, "Case for automated detection of diabetic retinopathy," in *Proc. AAAI Spring Symp. Ser.*, 2010, pp. 85–90.
- [13] M. H. A. Fadzil, L. I. Izhar, H. Nugroho, and H. A. Nugroho, "Analysis of retinal fundus images for grading of diabetic retinopathy severity," *Med. Biol. Eng. Comput.*, vol. 49, no. 6, pp. 693–700, Jun. 2011.
- [14] M. D. Saleh, C. Eswaran, and A. Mueen, "An automated blood vessel segmentation algorithm using histogram equalization and automatic threshold selection," *J. Digit. Imag.*, vol. 24, no. 4, pp. 564–572, Aug. 2011.
- [15] M. Gandhi and R. Dhanasekaran, "Diagnosis of diabetic retinopathy using morphological process and SVM classifier," in *Proc. Int. Conf. Commun. Signal Process.*, Apr. 2013, pp. 873–877.
- [16] R. Priya and P. Aruna, "Diagnosis of diabetic retinopathy using machine learning techniques," *ICTACT J. Soft Comput.*, vol. 3, no. 4, pp. 563–575, 2013.
- [17] B. Antal and A. Hajdu, "An ensemble-based system for automatic screening of diabetic retinopathy," *Knowl.-Based Syst.*, vol. 60, pp. 20–27, Apr. 2014.
- [18] K. Bhatia, S. Arora, and R. Tomar, "Diagnosis of diabetic retinopathy using machine learning classification algorithm," in *Proc. 2nd Int. Conf. Next Gener. Comput. Technol. (NGCT)*, Oct. 2016, pp. 347–351.
- [19] B. S. Mankar and N. Rout, "Automatic detection of diabetic retinopathy using morphological operation and machine learning," *ABHIYANTRIKI Int. J. Eng. Technol.*, vol. 3, no. 5, pp. 12–19, 2016.
- [20] B. Wu, W. Zhu, F. Shi, S. Zhu, and X. Chen, "Automatic detection of microaneurysms in retinal fundus images," *Computerized Med. Imag. Graph.*, vol. 55, pp. 106–112, Jan. 2017.
- [21] M. Chetoui, M. A. Akhloufi, and M. Kardouchi, "Diabetic retinopathy detection using machine learning and texture features," in *Proc. IEEE Can. Conf. Electr. Comput. Eng. (CCECE)*, May 2018, pp. 1–4.
- [22] T. Balogun, R. Saliu, S. Faluyi, and K. Fapohunda, "Comparative analysis of deep learning models for the detection and classification of diabetes retinopathy," in *Proc. 5th Inf. Technol. Educ. Develop. (ITED)*, Nov. 2022, pp. 1–6.
- [23] W. L. Yun, U. R. Acharya, Y. V. Venkatesh, C. Chee, L. C. Min, and E. Y. K. Ng, "Identification of different stages of diabetic retinopathy using retinal optical images," *Inf. Sci.*, vol. 178, no. 1, pp. 106–121, Jan. 2008.
- [24] R. Acharya U, C. K. Chua, E. Y. K. Ng, W. Yu, and C. Chee, "Application of higher order spectra for the identification of diabetes retinopathy stages," *J. Med. Syst.*, vol. 32, no. 6, pp. 481–488, Dec. 2008.
- [25] C. Agurto, V. Murray, E. Barriga, S. Murillo, M. Pattichis, H. Davis, S. Russell, M. Abramoff, and P. Soliz, "Multiscale AM-FM methods for diabetic retinopathy lesion detection," *IEEE Trans. Med. Imag.*, vol. 29, no. 2, pp. 502–512, Feb. 2010.
- [26] R. Naveen, S. A. Sivakumar, B. M. Shankar, and A. K. Priyaa, "Diabetic retinopathy detection using image processing," *Int. J. Eng. Adv. Technol.*, vol. 8, no. 6S, pp. 937–941, 2019.
- [27] I. Odeh, M. Alkasasbeh, and M. Alauthman, "Diabetic retinopathy detection using ensemble machine learning," in *Proc. Int. Conf. Inf. Technol. (ICIT)*, Jul. 2021, pp. 173–178.
- [28] G. G. Gardner, D. Keating, T. H. Williamson, and A. T. Elliott, "Automatic detection of diabetic retinopathy using an artificial neural network: A screening tool," *Brit. J. Ophthalmology*, vol. 80, no. 11, pp. 940–944, Nov. 1996.
- [29] C. Sinthanayothin, J. F. Boyce, T. H. Williamson, H. L. Cook, E. Mensah, S. Lal, and D. Usher, "Automated detection of diabetic retinopathy on digital fundus images," *Diabetic Med.*, vol. 19, no. 2, pp. 105–112, Feb. 2002.
- [30] G. B. Kande, T. S. Savithri, and P. V. Subbaiah, "Automatic detection of microaneurysms and hemorrhages in digital fundus images," *J. Digit. Imag.*, vol. 23, no. 4, pp. 430–437, Aug. 2010.
- [31] L. Tang, M. Niemeijer, J. M. Reinhardt, M. K. Garvin, and M. D. Abramoff, "Splat feature classification with application to retinal hemorrhage detection in fundus images," *IEEE Trans. Med. Imag.*, vol. 32, no. 2, pp. 364–375, Feb. 2013.
- [32] P. Liskowski and K. Krawiec, "Segmenting retinal blood vessels with deep neural networks," *IEEE Trans. Med. Imag.*, vol. 35, no. 11, pp. 2369–2380, Nov. 2016.
- [33] M. J. J. P. van Grinsven, B. van Ginneken, C. B. Hoyng, T. Theelen, and C. I. Sánchez, "Fast convolutional neural network training using selective data sampling: Application to hemorrhage detection in color fundus images," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1273–1284, May 2016.
- [34] Soniya, S. Paul, and L. Singh, "Heterogeneous modular deep neural network for diabetic retinopathy detection," in *Proc. IEEE Region 10 Humanitarian Technol. Conf. (R10-HTC)*, Dec. 2016, pp. 1–6.
- [35] V. Gulshan, L. Peng, M. Coram, M. C. Stumpe, D. Wu, A. Narayanaswamy, S. Venugopalan, K. Widner, T. Madams, J. Cuadros, R. Kim, R. Raman, P. C. Nelson, J. L. Mega, and D. R. Webster, "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," *J. Amer. Med. Assoc.*, vol. 316, no. 22, p. 2402, Dec. 2016.
- [36] P. Prentašić and S. Lončarić, "Detection of exudates in fundus photographs using deep neural networks and anatomical landmark detection fusion," *Comput. Methods Programs Biomed.*, vol. 137, pp. 281–292, Dec. 2016.
- [37] J. H. Tan, H. Fujita, S. Sivaprasad, S. V. Bhandary, A. K. Rao, K. C. Chua, and U. R. Acharya, "Automated segmentation of exudates, haemorrhages, microaneurysms using single convolutional neural network," *Inf. Sci.*, vol. 420, pp. 66–76, Dec. 2017.
- [38] G. Quellec, K. Charrière, Y. Boudi, B. Cochener, and M. Lamard, "Deep image mining for diabetic retinopathy screening," *Med. Image Anal.*, vol. 39, pp. 178–193, Jul. 2017.
- [39] C. Lam, C. Yu, L. Huang, and D. Rubin, "Retinal lesion detection with deep learning using image patches," *Investigative Ophthalmol. Vis. Sci.*, vol. 59, no. 1, p. 590, Jan. 2018.
- [40] J. I. Orlando, E. Prokofyeva, M. del Fresno, and M. B. Blaschko, "An ensemble deep learning based approach for red lesion detection in fundus images," *Comput. Methods Programs Biomed.*, vol. 153, pp. 115–127, Jan. 2018.
- [41] A. Rakhlin, "Diabetic retinopathy detection through integration of deep learning classification framework," *BioRxiv*, vol. 1, Jan. 2017, Art. no. 225508.
- [42] R. Gargeya and T. Leng, "Automated identification of diabetic retinopathy using deep learning," *Ophthalmology*, vol. 124, no. 7, pp. 962–969, Jul. 2017.
- [43] P. Chudzik, S. Majumdar, F. Calivá, B. Al-Diri, and A. Hunter, "Microaneurysm detection using fully convolutional neural networks," *Comput. Methods Programs Biomed.*, vol. 158, pp. 185–192, May 2018.
- [44] K. Adem, "Exudate detection for diabetic retinopathy with circular Hough transformation and convolutional neural networks," *Expert Syst. Appl.*, vol. 114, pp. 289–295, Dec. 2018.
- [45] J. Mo, L. Zhang, and Y. Feng, "Exudate-based diabetic macular edema recognition in retinal images using cascaded deep residual networks," *Neurocomputing*, vol. 290, pp. 161–171, May 2018.
- [46] P. Khojasteh, B. Aliahmad, and D. K. Kumar, "Fundus images analysis using deep features for detection of exudates, hemorrhages and microaneurysms," *BMC Ophthalmol.*, vol. 18, no. 1, pp. 1–13, Dec. 2018.
- [47] L. Dai, R. Fang, H. Li, X. Hou, B. Sheng, Q. Wu, and W. Jia, "Clinical report guided retinal microaneurysm detection with multi-sieving deep learning," *IEEE Trans. Med. Imag.*, vol. 37, no. 5, pp. 1149–1161, May 2018.
- [48] S. M. S. Islam, M. M. Hasan, and S. Abdullah, "Deep learning based early detection and grading of diabetic retinopathy using retinal fundus images," 2018, *arXiv:1812.10595*.
- [49] P. Porwal, S. Pachade, M. Kokare, G. Deshmukh, J. Son, W. Bae, L. Liu, J. Wang, X. Liu, and L. Gao, "IDRiD: Diabetic retinopathy–segmentation and grading challenge," *Med. Image Anal.*, vol. 59, Jan. 2020, Art. no. 101561.
- [50] S. Ananda, D. Kitahara, A. Hirabayashi, and K. R. U. K. Reddy, "Automated fundus image segmentation for diabetic retinopathy diagnosis by multiple modified U-Nets and SegNets," in *Proc. Asia-Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. (APSIPA ASC)*, Nov. 2019, pp. 1582–1588.
- [51] 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.
- [52] Z. Yan, X. Han, C. Wang, Y. Qiu, Z. Xiong, and S. Cui, "Learning mutually local-global U-Nets for high-resolution retinal lesion segmentation in fundus images," in *Proc. IEEE 16th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2019, pp. 597–600.

- [53] Y. Yan, J. Gong, and Y. Liu, "A novel deep learning method for red lesions detection using hybrid feature," in *Proc. Chin. Control Decis. Conf. (CCDC)*, Jun. 2019, pp. 2287–2292.
- [54] N. Eftekhari, H.-R. Pourreza, M. Masoudi, K. Ghiasi-Shirazi, and E. Saeedi, "Microaneurysm detection in fundus images using a two-step convolutional neural network," *Biomed. Eng. OnLine*, vol. 18, no. 1, pp. 1–16, Dec. 2019.
- [55] Y. Huang, L. Lin, M. Li, J. Wu, P. Cheng, K. Wang, J. Yuan, and X. Tang, "Automated hemorrhage detection from coarsely annotated fundus images in diabetic retinopathy," in *Proc. IEEE 17th Int. Symp. Biomed. Imag. (ISBI)*, Apr. 2020, pp. 1369–1372.
- [56] H. Wang, G. Yuan, X. Zhao, L. Peng, Z. Wang, Y. He, C. Qu, and Z. Peng, "Hard exudate detection based on deep model learned information and multi-feature joint representation for diabetic retinopathy screening," *Comput. Methods Programs Biomed.*, vol. 191, Jul. 2020, Art. no. 105398.
- [57] A. M. Pour, H. Seyedarabi, S. H. A. Jahromi, and A. Javadzadeh, "Automatic detection and monitoring of diabetic retinopathy using efficient convolutional neural networks and contrast limited adaptive histogram equalization," *IEEE Access*, vol. 8, pp. 136668–136673, 2020.
- [58] M. Chetoui and M. A. Akhloufi, "Explainable diabetic retinopathy using EfficientNET," in *Proc. 42nd Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2020, pp. 1966–1969.
- [59] P. Saranya, R. Pranati, and S. S. Patro, "Detection and classification of red lesions from retinal images for diabetic retinopathy detection using deep learning models," *Multimedia Tools Appl.*, vol. 82, no. 25, pp. 39327–39347, Oct. 2023.
- [60] H. S. Alghamdi, H. L. Tang, S. A. Waheeb, and T. Peto, "Automatic optic disc abnormality detection in fundus images: A deep learning approach," in *Proc. Ophthalmic Med. Image Anal. 3rd Int. Workshop*, Oct. 2016, pp. 17–24.
- [61] H. Takahashi, H. Tampo, Y. Arai, Y. Inoue, and H. Kawashima, "Applying artificial intelligence to disease staging: Deep learning for improved staging of diabetic retinopathy," *PLoS ONE*, vol. 12, no. 6, Jun. 2017, Art. no. e0179790.
- [62] Z. Wang and J. Yang, "Diabetic retinopathy detection via deep convolutional networks for discriminative localization and visual explanation," 2017, *arXiv:1703.10757*.
- [63] Y.-H. Li, N.-N. Yeh, S.-J. Chen, and Y.-C. Chung, "Computer-assisted diagnosis for diabetic retinopathy based on fundus images using deep convolutional neural network," *Mobile Inf. Syst.*, vol. 2019, pp. 1–14, Jan. 2019.
- [64] S. S. Chaturvedi, K. Gupta, V. Ninawe, and P. S. Prasad, "Automated diabetic retinopathy grading using deep convolutional neural network," 2020, *arXiv:2004.06334*.
- [65] T. Hattiya, K. Dittakan, and S. Musikasawan, "Diabetic retinopathy detection using convolutional neural network: A comparative study on different architectures," *Eng. Access*, vol. 7, no. 1, pp. 50–60, 2021.
- [66] J. D. Bodapati, N. S. Shaik, and V. Naralasetti, "Deep convolution feature aggregation: An application to diabetic retinopathy severity level prediction," *Signal, Image Video Process.*, vol. 15, no. 5, pp. 923–930, Jul. 2021.
- [67] C. Mohanty, S. Mahapatra, B. Acharya, F. Kokkoras, V. C. Gerogiannis, I. Karamitsos, and A. Kanavos, "Using deep learning architectures for detection and classification of diabetic retinopathy," *Sensors*, vol. 23, no. 12, p. 5726, Jun. 2023.
- [68] U. Ishtiaq, E. R. M. F. Abdullah, and Z. Ishtiaque, "A hybrid technique for diabetic retinopathy detection based on ensemble-optimized CNN and texture features," *Diagnostics*, vol. 13, no. 10, p. 1816, May 2023.
- [69] R. Al-Ahmadi, H. Al-Ghamdi, and L. Hsairi, "Classification of diabetic retinopathy by deep learning," *Int. J. Online Biomed. Eng.*, vol. 20, no. 1, pp. 74–88, 2024.
- [70] T. A. Soomro, J. Gao, T. Khan, A. F. M. Hani, M. A. U. Khan, and M. Paul, "Computerised approaches for the detection of diabetic retinopathy using retinal fundus images: A survey," *Pattern Anal. Appl.*, vol. 20, no. 4, pp. 927–961, Nov. 2017.
- [71] T. A. Soomro, A. J. Afifi, L. Zheng, S. Soomro, J. Gao, O. Hellwich, and M. Paul, "Deep learning models for retinal blood vessels segmentation: A review," *IEEE Access*, vol. 7, pp. 71696–71717, 2019.
- [72] W. L. Alyoubi, W. M. Shalash, and M. F. Abulkhair, "Diabetic retinopathy detection through deep learning techniques: A review," *Informat. Med. Unlocked*, vol. 20, Jan. 2020, Art. no. 100377.
- [73] M. N. Ashraf, Z. Habib, and M. Hussain, "Texture feature analysis of digital fundus images for early detection of diabetic retinopathy," in *Proc. 11th Int. Conf. Comput. Graph., Imag. Visualizat.*, Aug. 2014, pp. 57–62.
- [74] S. Giraddi, J. Pujari, and S. Seeri, "Role of GLCM features in identifying abnormalities in the retinal images," *Int. J. Image, Graph. Signal Process.*, vol. 7, no. 6, pp. 45–51, May 2015.
- [75] A. Z. Foeady, D. C. R. Novitasari, A. H. Asyhar, and M. Firmansjah, "Automated diagnosis system of diabetic retinopathy using GLCM method and SVM classifier," in *Proc. 5th Int. Conf. Electr. Eng., Comput. Sci. Informat. (EECSI)*, Oct. 2018, pp. 154–160.
- [76] M. ur Rehman, Z. Abbas, S. H. Khan, S. H. Ghani, and Najam, "Diabetic retinopathy fundus image classification using discrete wavelet transform," in *Proc. 2nd Int. Conf. Eng. Innov. (ICEI)*, Jul. 2018, pp. 75–80.
- [77] H. H. Vo and A. Verma, "Discriminant color texture descriptors for diabetic retinopathy recognition," in *Proc. IEEE 12th Int. Conf. Intell. Comput. Commun. Process. (ICCP)*, Sep. 2016, pp. 309–315.
- [78] T. Vijayan, M. Sangeetha, A. Kumaravel, and B. Karthik, "Feature selection for simple color histogram filter based on retinal fundus images for diabetic retinopathy recognition," *IETE J. Res.*, vol. 69, no. 2, pp. 987–994, Feb. 2023.
- [79] D. Sarwinda, T. Siswantining, and A. Bustamam, "Classification of diabetic retinopathy stages using histogram of oriented gradients and shallow learning," in *Proc. Int. Conf. Comput., Control, Informat. Appl. (IC3INA)*, Nov. 2018, pp. 83–87.
- [80] L. Armi and S. Fekri-Ershad, "Texture image analysis and texture classification methods—A review," 2019, *arXiv:1904.06554*.
- [81] I. Sobel, "An isotropic 3×3 image gradient operator," in *Machine Vision for Three-Dimensional Scenes*. Sweden: Stanford Artificial Intelligence Project, 1990, pp. 376–379.
- [82] J. M. Prewitt, "Object enhancement and extraction," *Picture Process. Psychopictorics*, vol. 10, no. 1, pp. 15–19, 1970.
- [83] N. H. Harun, Y. Yusof, F. Hassan, and Z. Embong, "Classification of fundus images for diabetic retinopathy using artificial neural network," in *Proc. IEEE Jordan Int. Joint Conf. Electr. Eng. Inf. Technol. (JEEIT)*, Apr. 2019, pp. 498–501.
- [84] M. Hardas, S. Mathur, A. Bhaskar, and M. Kalla, "Retinal fundus image classification for diabetic retinopathy using SVM predictions," *Phys. Eng. Sci. Med.*, vol. 45, no. 3, pp. 781–791, Sep. 2022.
- [85] J. Kaur and P. Kaur, "Automated computer-aided diagnosis of diabetic retinopathy based on segmentation and classification using K-nearest neighbor algorithm in retinal images," *Comput. J.*, vol. 66, no. 8, pp. 2011–2032, Aug. 2023.
- [86] N. Zaaboub and A. Douik, "Early diagnosis of diabetic retinopathy using random forest algorithm," in *Proc. 5th Int. Conf. Adv. Technol. Signal Image Process. (ATSIP)*, Sep. 2020, pp. 1–5.
- [87] Y. Kang, Y. Fang, and X. Lai, "Automatic detection of diabetic retinopathy with statistical method and Bayesian classifier," *J. Med. Imag. Health Informat.*, vol. 10, no. 5, pp. 1225–1233, May 2020.
- [88] N. Arora, A. Singh, M. Z. N. Al-Dabagh, and S. K. Maitra, "A novel architecture for diabetes patients' prediction using k-means clustering and SVM," *Math. Problems Eng.*, vol. 2022, no. 1, 2022, Art. no. 4815521.
- [89] P. Macsik, J. Pavlovicova, J. Goga, and S. Kajan, "Local binary CNN for diabetic retinopathy classification on fundus images," *Acta Polytechnica Hungarica*, vol. 19, no. 7, pp. 27–45, 2022.
- [90] K. O'Shea and R. Nash, "An introduction to convolutional neural networks," 2015, *arXiv:1511.08458*.
- [91] Y. Li, Z. Song, S. Kang, S. Jung, and W. Kang, "Semi-supervised auto-encoder graph network for diabetic retinopathy grading," *IEEE Access*, vol. 9, pp. 140759–140767, 2021.
- [92] D. Bank, N. Koenigstein, and R. Giryes, "Autoencoders," in *Machine Learning for Data Science Handbook: Data Mining and Knowledge Discovery Handbook*. New York, NY, USA: Cornell Univ., 2023, pp. 353–374.
- [93] V. R. Naramala, B. A. Kumar, V. S. Rao, A. Mishra, S. A. Hannan, Y. A. B. El-Ebiary, and R. Manikandan, "Enhancing diabetic retinopathy detection through machine learning with restricted Boltzmann machines," *Int. J. Adv. Comput. Sci. Appl.*, vol. 14, no. 9, pp. 573–585, 2023.
- [94] A. Fischer and C. Igel, "An introduction to restricted Boltzmann machines," in *Proc. Iberoamerican Congr. Pattern Recognit.*, Buenos Aires, Argentina. Cham, Switzerland: Springer, Sep. 2012, pp. 14–36.

- [95] A. Reghukumar, L. J. Anbarasi, J. Prassanna, R. Manikandan, and F. Al-Turjman, "Vision based segmentation and classification of cracks using deep neural networks," *Int. J. Uncertainty, Fuzziness Knowl.-Based Syst.*, vol. 29, no. 1, pp. 141–156, Apr. 2021.
- [96] L. J. Anbarasi, M. J. Vincent, and G. S. A. Mala, "A novel visual secret sharing scheme for multiple secrets via error diffusion in halftone visual cryptography," in *Proc. Int. Conf. Recent Trends Inf. Technol. (ICRITT)*, Jun. 2011, pp. 129–133.
- [97] M. Jawahar, L. J. Anbarasi, S. G. Jasmine, and M. Narendra, "Diabetic foot ulcer segmentation using color space models," in *Proc. 5th Int. Conf. Commun. Electron. Syst. (ICCES)*, Jun. 2020, pp. 742–747.
- [98] K. Xu, D. Feng, and H. Mi, "Deep convolutional neural network-based early automated detection of diabetic retinopathy using fundus image," *Molecules*, vol. 22, no. 12, p. 2054, Nov. 2017.
- [99] M. T. Esfahani, M. Ghaderi, and R. Kafiyeh, "Classification of diabetic and normal fundus images using new deep learning method," *Leonardo Electron. J. Pract. Technol.*, vol. 17, no. 32, pp. 233–248, 2018.
- [100] G. T. Zago, R. V. Andreão, B. Dorizzi, and E. O. Teatini Salles, "Diabetic retinopathy detection using red lesion localization and convolutional neural networks," *Comput. Biol. Med.*, vol. 116, Jan. 2020, Art. no. 103537.
- [101] J. P. Kandhasamy, S. Balamurali, S. Kadry, and L. K. Ramasamy, "Diagnosis of diabetic retinopathy using multi level set segmentation algorithm with feature extraction using SVM with selective features," *Multimedia Tools Appl.*, vol. 79, nos. 15–16, pp. 10581–10596, Apr. 2020.
- [102] H. Pratt, F. Coenen, D. M. Broadbent, S. P. Harding, and Y. Zheng, "Convolutional neural networks for diabetic retinopathy," *Proc. Comput. Sci.*, vol. 90, pp. 200–205, Jan. 2016.
- [103] X. Wang, Y. Lu, Y. Wang, and W.-B. Chen, "Diabetic retinopathy stage classification using convolutional neural networks," in *Proc. IEEE Int. Conf. Inf. Reuse Integr. (IRI)*, Jul. 2018, pp. 465–471.
- [104] S. Honnunar, S. Mehra, and S. Joseph, "Diabetic retinopathy identification and severity classification," in *Proc. Fall*, 2016, pp. 1–5.
- [105] I. Sadek, M. Elawady, and A. El Rahman Shabayek, "Automatic classification of bright retinal lesions via deep network features," 2017, *arXiv:1707.02022*.
- [106] A. V. Prasad, A. Gupta, L. J. Anbarasi, and M. Jawahar, "Current trends, challenges, and future prospects for automated detection of diabetic retinopathy," in *Proc. 3rd Int. Conf. Signal Process. Commun. (ICSPSC)*, May 2021, pp. 340–343.
- [107] I. Qureshi, J. Ma, and Q. Abbas, "Recent development on detection methods for the diagnosis of diabetic retinopathy," *Symmetry*, vol. 11, no. 6, p. 749, Jun. 2019.
- [108] A. Mishra, L. Singh, M. Pandey, and S. Lakra, "Image based early detection of diabetic retinopathy: A systematic review on artificial intelligence (AI) based recent trends and approaches," *J. Intell. Fuzzy Syst.*, vol. 43, no. 5, pp. 6709–6741, Sep. 2022.
- [109] G. Selvachandran, S. G. Quek, R. Paramesran, W. Ding, and L. H. Son, "Developments in the detection of diabetic retinopathy: A state-of-the-art review of computer-aided diagnosis and machine learning methods," *Artif. Intell. Rev.*, vol. 56, no. 2, pp. 915–964, Feb. 2023.



AMNA IKRAM received the master's degree in computer science from the Shaheed Zulfiqar Ali Bhutto Institute of Science and Technology (SZABIST), Pakistan. During her pursuit of the master's degree, she undertook extensive research focused on the software requirement traceability matrix (SRTM) as a part of her thesis project. She was a Visiting Lecturer with the Department of Computer Science, in different institutions in Pakistan. She is currently a Lecturer with the Department of Computer Science, Faculty of Software Engineering, National University of Technology, Islamabad, Pakistan. Her research interests include image processing, medical image diagnosis, machine learning, deep learning, and software requirements engineering. She aims to contribute to interdisciplinary research of computer science and human-related disciplines.



AZHAR IMRAN (Senior Member, IEEE) received the master's degree in computer science from the University of Sargodha, Pakistan, and the Ph.D. degree in software engineering from Beijing University of Technology, China. He was a Senior Lecturer with the Department of Computer Science, University of Sargodha, from 2012 to 2017. He is currently an Assistant Professor with the Department of Creative Technologies, Faculty of Computing and Artificial Intelligence, Air University, Islamabad, Pakistan. He is a Renowned Expert of image processing, healthcare informatics, and social media analysis. He has over 11.5 years of national and international academic experience as a full-time Faculty Member and teaching software engineering and core computing courses. His research interests include image processing, social media analysis, medical image diagnosis, machine learning, and data mining. He aims to contribute to interdisciplinary research of computer science and human-related disciplines. He is a Regular Member of IEEE and has contributed more than 60 research papers in well-reputed international journals and conferences. He has delivered guest talks and conducted seminars and training at numerous national and international forums in the past. He has contributed to multiple international conferences in diverse roles (keynote speaker, technical/committee member, registration, and speaker). He is an Editorial Member and a Reviewer of various journals, including IEEE Access, *Cancers* (MDPI), *Applied Sciences*, *Mathematics*, *The Visual Computer* (Springer), *Biomedical Imaging* (Taylor and Francis), *Journal of Visualization*, *Multimedia Tools and Applications*, IGI Global, and *Journal of Imaging*.



JIANQIANG LI (Senior Member, IEEE) received the B.S. degree in mechatronics from Beijing Institute of Technology, Beijing, China, in 1996, and the M.S. and Ph.D. degrees in control science and engineering from Tsinghua University, Beijing, in 2001 and 2004, respectively. He was a Researcher with the Digital Enterprise Research Institute, National University of Ireland, Galway, from 2004 to 2005. From 2005 to 2013, he was with NEC Laboratories China, as a Researcher. He was with the Department of Computer Science, Stanford University, as a Visiting Scholar, from 2009 to 2010. He joined Beijing University of Technology, Beijing, in 2013, as a Beijing Distinguished Professor. He has more than 200 publications, including one book, more than 100 journal articles, and 58 international patent applications (27 of them have been granted in China, USA, or Japan). His research interests include petri nets, enterprise information systems, business processes, data mining, information retrieval, semantic web, privacy protection, and big data. He was a PC Member of the multiple international conferences and organized the IEEE workshop on medical computing. He served as the Guest Editor to organize a Special Issue on information technology for enhanced healthcare service in computer and industry.



ABDULAZIZ ALZUBAIDI received the bachelor's degree in computer science from the University College, King Abdulaziz University, Saudi Arabia, in 2001, the Master of Science degree from Jordan University, and the Ph.D. degree from the University of Colorado at Colorado Springs, USA, in 2017. His research interests include intrusion detection, computer vision, human activity recognition, cyber-security awareness, and smart-devices security.



SAFA FAHIM received the bachelor's degree in computer science from Air University, Islamabad, Pakistan, in 2021. She is currently pursuing the master's degree in data science with the Department of Creative Technologies, Faculty of Computing and Artificial Intelligence, Air University, Islamabad. She is an Active Researcher who has published more than four research papers in reputable journals and conferences. She has also presented multiple research papers in conferences, showcasing her contributions to the academic community. Her research interests include image processing, social media analysis, medical image diagnosis, machine learning, and data mining. She is driven to contribute to interdisciplinary research in computer science and related fields.



AMANULLAH YASIN received the master's degree in data science and the Ph.D. degree in data science and artificial intelligence from the University of Nantes, France. He has supervised more than 20 M.S. research thesis. Currently, three Ph.D. students are working under his supervision. He has multiple international publications in his research areas. His current research interests include big data analytics, machine learning, and Bayesian networks to solve the social problems of Pakistan. He is a reviewer and program committee member of various journals and conferences, respectively.



HANAA FATHI received the Ph.D. degree in computer science from the Faculty of Science, Menoufia University, Egypt, in 2022. She is currently an Assistant Professor with the Faculty of Information Technology, Applied Science Private University. She has involved on several research topics. She has contributed more than ten technical articles in feature selection, classification, optimization, machine learning, and web service in international journals. She attended the 2019 International Conference on Intelligent Systems and Advanced Computing Sciences (ISACS), in December 2019, Taza, Morocco, and the fourth edition of the International Conference on Intelligent Systems and Computer Vision (ISCV 2020). She majors in machine learning, optimization, and medical application.

...