A critical review on diagnosis of diabetic retinopathy using machine learning and deep learning

Dolly Das[1] · Saroj Kr. Biswas[1] · Sivaji Bandyopadhyay[1]

Abstract

Diabetic Retinopathy (DR) is a health condition caused due to Diabetes Mellitus (DM). It causes vision problems and blindness due to disfigurement of human retina. According to statistics, 80% of diabetes patients battling from long diabetic period of 15 to 20 years, suffer from DR. Hence, it has become a dangerous threat to the health and life of people. To overcome DR, manual diagnosis of the disease is feasible but overwhelming and cumbersome at the same time and hence requires a revolutionary method. Thus, such a health condition necessitates primary recognition and diagnosis to prevent DR from developing into severe stages and prevent blindness. Innumerable Machine Learning (ML) models are proposed by researchers across the globe, to achieve this purpose. Various feature extraction techniques are proposed for extraction of DR features for early detection. However, traditional ML models have shown either meagre generalization throughout feature extraction and classification for deploying smaller datasets or consumes more of training time causing inefficiency in prediction while using larger datasets. Hence Deep Learning (DL), a new domain of ML, is introduced. DL models can handle a smaller dataset with help of efficient data processing techniques. However, they generally incorporate larger datasets for their deep architectures to enhance performance in feature extraction and image classification. This paper gives a detailed review on DR, its features, causes, ML models, state-of-the-art DL models, challenges, comparisons and future directions, for early detection of DR.

✉ Dolly Das
das.dolly0019@gmail.com

Saroj Kr. Biswas
bissarojkum@yahoo.com

Sivaji Bandyopadhyay
director@nits.ac.in

[1] National Institute of Technology Silchar, Cachar, Assam, India

1 Introduction

Diabetic Retinopathy (DR) is a health condition that arises due to Diabetes Mellitus (DM). DM is caused due to numerous micro and macrovascular abnormalities and impaired glucose metabolism leading to an enduring disease. DR is one of the most common and grave complications of DM leading to severe blindness due to disfigurement of the human retina. According to statistics, 80% of diabetes patients battling from long diabetic period of 15 to 20 years, suffer from DR [40]. DR due to diabetes is also recorded to be the chief reason of blindness amongst the working-age people, in advanced nations [74]. More than 171 million people suffer from diabetes worldwide. The World Health Organization (WHO) has surveyed that, there will be 366 million cases of diabetes in the world by 2030 [139].

The general signs and symptoms of DR are blurry vision, floaters and flashes, and loss of vision [4]. DR occurs due to metabolic fluctuations in retinal blood vessels, caused due to irregular blood flow, leakage of blood and blood constituents over the retina thereby affecting the macula. This leads to swelling of the retinal tissue, causing cloudy or blurred vision. The disorder affects both eyes, and with longer period of diabetes without treatment, DR causes blindness causing diabetic maculopathy [31, 91].

When DR remains untreated and undiagnosed, its progressive nature to serious stages worsens the vision capacity of a person. With periodic or random progress in the disease, retinal lesions are formed from the ruptured Retinal Blood Vessels (RBVs) such as Microaneurysms (MAs), Hemorrhages (HEs), Exudates (EXs), Cotton Wool Spots (CWSs), Foveal Avascular Zone (FAZ), fibrotic bands, Intra Retinal Microvascular Abnormalities (IRMAs), Neovascularization on Disc (NVD), Neovascularization Elsewhere (NVE), traction-al bands etc. [37, 39, 52, 75, 128]. These retinal lesions occur in the rear view of the human eye i.e., the fundus. The presence of these retinal lesions and abnormalities and their timely detection, helps in identifying the various stages of DR [16].

To observe the retinal anatomy such as Optic Disc (OD), RBVs, fovea and the macula, the pupil dilation takes place with the help of certain medically identified and approved contrasting agents which are injected into the retina. Such a method employs Fluorescein Angiography (FA) or a mydriatic fundus camera. This helps in acquisition of fundus images from diabetic patients which can be assessed for the effective detection and early diagnosis of DR [1]. To diagnose DR at an early stage, manual methods such as bio-microscopy, retinal imaging of the fundus, Retinal Thickness Analyzer (RTA), Scanning Laser Ophthalmoscopy (SLO), Adaptive Optics, Retinal Oximetry, Optical Coherence Tomography (OCT), OCT Angiography, Doppler OCT, and many more can be adopted [48]. However, such conventional methods for manually analyzing the disease makes it cumbersome, time consuming and highly prone to error. Besides, it demands a sophisticated task force which is sometimes not feasible w.r.t (with respect to) prevailing circumstances. Thus, it is not feasible to perform manual diagnosis for early detection of DR at any time and at any place.

The present ratio of Ophthalmologists to patient especially in India is 1:10000 [91] and in such a situation, the need of an automated intelligent detection system for primary analysis of early signs of DR is realized. Thus, a faster and a revolutionary method, proposing an intelligent system which uses a huge dataset of fundus image acquired through various sources, is essential to detect the disease at a premature stage such that lives of people suffering from prolonged diabetes can be made better through possible retainment of vision. Consequently, various intelligent and computer-assisted systems are proposed for DR detection using ML techniques such as Support Vector Machine (SVM) [10, 20, 81, 90, 108, 115],

Decision Tree [10, 28], Neural Network [17, 90], etc. However, conventional ML techniques are incompetent against real-time, large, complex, and high-dimensional data such as images. They lack domain awareness and data representation, which makes them computationally intensive and inflexible in performance.

Deep Learning (DL) is a new advent of ML which can perform automated and complex tasks, discover unseen insights, highly scalable, better domain knowledge, reliable decision making etc. and efficiently applies them upon high-dimensional data, thereby outperforming shallow ML models. DL methods such as Convolutional Neural Network (CNN) [17, 70, 115], Deep Convolutional Neural Network (DCNN), and Deep Neural Network (DNN/ DLNN) architectures such as AlexNet, Visual Geometry Group Network (VGG- 16, VGG- 19) [115], GoogLeNet [17] and its variants, Residual Network (ResNet) and its variants, Densely Connected Network (DenseNet) and its variants, Inception Convolutional Recurrent Neural Networks (IRCNN), Generative Adversarial Network (GAN), Autoencoder, Restricted Boltzmann Machine (RBM), Long Short-Term Memory (LSTM), Deep Reinforcement Learning (DRL) etc. are proposed for deep feature extraction and image classification [9].

Digital Image Processing (DIP) techniques and advancements have also an effective role to play in better image feature extraction and image classification performances, through enhancement and removal of errors [16, 41, 137]. Various pre-trained DL models also known as *Transfer Learning (TL)* techniques have found application for DR detection using smaller datasets to overcome scarcity of data, boost the classification performance and learn useful representations. Various data augmentation methods, sampling techniques, cost-sensitive algorithms, hybrid and ensemble architectures have been adopted in existing works, to overcome the constraint of imbalanced and noisy fundus image data, and improve feature extraction and prediction, for DR detection. Various dimensionality reduction techniques and attention mechanisms such as Principal Component Analysis (PCA), Linear Discriminant Analysis (LDA), Singular Value Decomposition (SVD), Fully Convolutional Network (FCN) etc. have been adopted for compact feature representation of big data and better feature discrimination.

The main objective of this manuscript is to compare the diverse studies performed earlier for early DR detection and configure their drawbacks and limitations. Earlier works and literature surveys have distinguished various conventional ML models, DL models, pre-trained TL models, hybrid ML-DL models, evolutionary models, ensemble models and comprehensive models, for DR detection. The previous works have a separate base for each of these learning algorithms or a combination and comparison of these models to a limited extent, to lay an emphasis on a particular specified task say feature extraction or segmentation or classification. On the basis of such works, it is important to compare and determine how all these models and algorithms differ from one another, when they are all capable enough to produce acceptable results in various contexts. Thus, earlier survey works have only established a one-way conclusion or have fewer perception and domain knowledge of the complete problem, which can be manipulative in the making of a reliable decision for researchers and also makes the process of study vague and time-consuming. Thus, this paper makes an effort to inculcate a comprehensive study and behavior of different learning models and advancements, in the research and development of an early detection system for DR. This paper illustrates in *detail about DR*, DR lesions and their behavior, structure, challenges in detection, and stages of occurrence and development, in a chronological order, using DR images, in contrast to previous works where the chronology cannot be identified. It establishes a comparison on different kinds of DR lesions identified on different grounds using various techniques. The

study includes a comparison on various DL models such as hybrid ML-DL models [71, 108, 142], CNN [25, 34, 44, 57, 108, 112, 117], DCNN [142], TL models [6, 70, 78], DLNN [36, 43], ensemble ML/DL models [44, 58, 80, 101, 109, 112, 134], evolutionary [43, 96] and comprehensive learning algorithms [21, 134], and their corresponding performances w.r.t early DR detection, to conclude on a firm note in identifying the best model(s), with better and *generalized predictions*. The paper lays a huge stress on encouraging DL models for high dimensional data, and their incorporation with ML-based classifiers for ensemble-based feature extraction or classification, to propose methodologies for DR detection. Besides, single-classifier systems using Neural Networks, Support Vector Machines (SVMs) or Decision Trees are highly prone to overfitting and produces ineffective and unreliable performances, due to lack of domain knowledge. Thus, this manuscript has established some critical observations based on the study and implementation of classical models in comparison to DL models and ensemble models. The paper affirms on obtaining better results using a better learning model such as DL along with hyperparameter tuning and cost-effective strategies and developments, to improve feature extraction and image classification.

This paper illustrates the different aspects of DR based on different perspectives necessary for the early diagnosis and detection of DR. In section II, it discusses the different DR lesions and features, their characteristics and stage of DR. In section III, the paper illustrates the various kinds of ML techniques adopted for the process of diagnosis of DR. In section IV, the paper entails a detailed illustration of various DL techniques. In section V, the paper illustrates various challenges related to fundus image analysis, data acquisition, feature extraction and classification for DR diagnosis and detection, and their corresponding predictable solutions. In section VI, it establishes and emphasizes on a comparative analysis upon existing techniques and experimental evaluation of some of the best performing CNNs upon an imbalanced dataset using classical, DL and ensemble methods. In section VII, the paper proposes various future directions to encourage new solutions for early DR detection. Finally, the paper concludes on a note to focus on advanced methods such as DL techniques for early DR detection and keep a foundation on ML techniques as they are conventional yet are better learning algorithms, and can be improvised with DL techniques.

2 DR features

There are various features that can be used to detect and classify DR at an initial stage, for prevention of blindness. The presence of DR features helps in identifying the stage of DR, for diagnosis and treatment. Therefore, identification of DR features is a crucial research point as good identification makes good DR detection system. In this section, some DR features and challenges associated with them are discussed. Block Diagram I depicts the various DR features which can be used to detect the disease.

2.1 Microaneurysms

Microaneurysms (MAs) are localized capillary dilations, red in color and saccular in structure [31, 94]. They may either appear in clusters or in isolation. They are 1 to 3 pixels in diameter [37] or 10 μm to 100 μm [104]. MAs are the first symptom of DR, instigated by the focal dilatation of thin blood vessels. Figure 1 depict MAs, HEs, EXs [11] in fundus image.

Diagram I DR features

2.2 Hemorrhages

Hemorrhages (HEs) are structural distortions in the walls of blood vessels with growing risk of blood leakage from the vessels, producing irregular shapes. These thin bloods vessels are sufficiently deteriorated, which may rupture and give rise to an HE. They are usually 3 to 10 pixels in diameter [37]. Sometimes, HEs and MAs occur together and are called as red lesions, based on their shape and similarity [31]. HEs may be as small as MAs and as large as Optic Disc. HE or Intraretinal HE [28, 64] may appear in wide variety of shapes such as dot, blot or flame shaped, based on its depth in the retina [30] with varying contrast. Flame HEs are elongated structures, found as blood leaking into the nerve fiber layer of the capillary network.
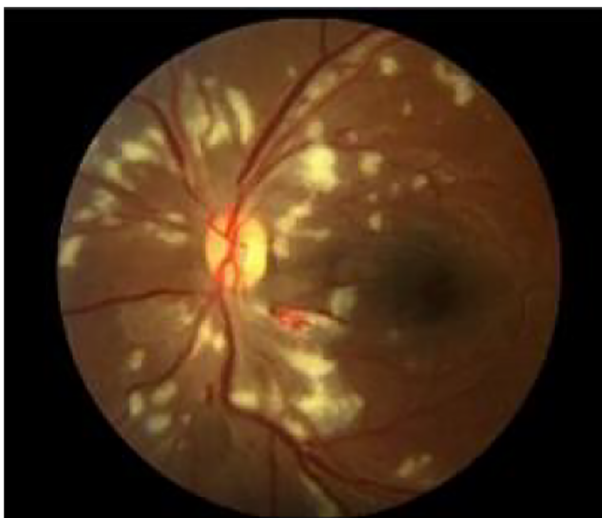

Fig. 2 MAs, HEs, EXs in fundus [11]

The appearance of HEs does not significantly affect vision. However, numerous blot HEs may infer significant ischemia, a notable characteristic of pre-proliferative retinopathy. Figure 2 depicts dot-and-blot HEs in DR retina [11]. HEs are the next sign of DR after MAs.

2.3 Exudates

Exudates (EXs) are discrete yellowish-white intraretinal deposits frequently observed with MAs [91] and contains extracellular lipids and proteins due to leakage of blood from abnormal retinal capillaries. They can vary from tiny specks to big patches and gradually evolve into ring-like constructions, within a diameter of 1 to 6 pixels [14], called circinate. They can appear with soft boundaries and cloudy structures called soft EXs, or with distinct boundaries and bright structures called hard EXs [16]. They are located in the posterior pole of the fundus, and appear as bright, well-contrasted patterns with high grey level, between the dark vessels [16]. They can also lead to vascular damage [93]. Hard EXs cause retinal thickening, which leads to malfunctioning of macula [3, 16, 61], thus causing DME or complete blindness [126]. Figure 2 also depicts EXs in DR retina [11]. HEs are the next sign of DR after MAs.

2.4 Cotton wool spots

Cotton Wool Spots (CWSs) are the largest and irregular, cloudy structures with soft boundaries, in comparison to MAs and EXs. They are retinal infarction caused by thrombosis and obstruction of blood vessels [37]. They are greyish-white patches of discoloration in the nerve fiber layer. They are a consequential of local ischemia which leads to disruption of axoplasmic flow. Multiple CWSs such as nearly 6 or more in one eye may indicate generalized retinal ischemia leading to the stage of pre-proliferative DR [29, 127]. Figure 2 depicts CWSs in DR fundus image [116].

Fig. 3 CWS in fundus image [116]

2.5 Foveal avascular zone

Foveal Avascular Zone (FAZ) is a region within the fovea in the macula which is devoid of RBVs. Its diameter is 0.5 mm [52, 73]. The fovea is of 1.5 mm [67] and is darker than the surrounding retinal tissues. FAZ exhibits a non-specific structure and differs among people due to disparities in the levels of pigment related with factors such as ethnicity, age, diet, and disease conditions. The macula is a small area approximately 5–5.5 mm in diameter, located temporal to the optic nerve head. To identify FAZ, the macula center [39, 52] is identified and then the vessel end points are localized around the macula, using the nearest distance from the center point. The FAZ area is computed by connecting these end points of the lost capillaries. The loss of capillaries enhances the rapid loss in visual acuity, thus causing of DR. The enlargement of this zone appears early in the development of the disease [15, 67, 76] and hence needs to be detected to prevent DR. Figure 3 depicts FAZ [38].

2.6 Optic disc

The Optic Disc (OD) is the brightest, homogeneous, circular structure in a normal eye fundus image and appears yellowish in color [18, 91]. The OD center and its diameter give information such as position of the origin of the blood vessels and the macula region [28]. It is important to detect any abnormality in the structure, shape, size or in the region of OD, to suspect for early changes causing visual loss. Figure 4 depicts the normal OD (bright circular structure) and RBVs of left eye of a normal patient (grade 0) acquired from Kaggle DR dataset [32].

2.7 Retinal blood vessels

The Retinal Blood Vessels (RBVs) are the central retinal artery and vein, and their branches, in the retina. The artery bifurcates into an upper and a lower branch, and each of these again divides into a medial or nasal and a lateral or temporal branch. These branches, a minute capillary plexus, which do not outspread beyond the inner nuclear layer. The macula receives two small branches, the superior and inferior macular arteries, from the temporal branches and

Fig. 4 FAZ [38]

(a)   (b)

Fig. 5 Normal OD

small twigs from the central artery, not reaching fovea centralis. Thus, the segmentation of RBVs and their branching pattern can provide plentiful information about any kind of abnormalities or disease by examining its pathological variance [130]. The unique curved shape vascular arcade arising from the OD and encircling the macula, can be exploited to know about earlier abnormalities. Prolonged diabetes in patients can damage RBVs, causing DR lesions [119]. During the detection process of RBVs, the grey level variation of vessels is high and causes high local contrast, which increases its sensitivity but decreases its predictivity. Figure 5(a) depicts normal RBVs (branches). Fig. VII depicts some of the abnormal changes in RBVs due to DR.

2.8 Neovascularization and intra retinal microvascular abnormalities

The extensive lack of oxygen in RBVs, causes diminished blood flow to ocular tissues which causes the creation of new fragile vessels, making the OD dense. These new vessels are together called as Neovascularization (NV) [110, 127, 131] which is a serious threat to eye sight. These new blood vessels have feebler walls and may break down and bleed, or cause scar tissue thus causing retinal detachment. If the retinal detachment is not treated, it can cause severe vision loss. Again, the breakage of these blood vessels at the onset of DR, increases the number of nodal points, which indicates DR severity [96]. The formation of these new vessels in the OD or within 1disc diameter of the margin of the disc, growing along the posterior hyaloid interface around the optic nerve is known as NVD and if it forms in the periphery of the retina, then it is called as NVE. NV is often confused with IRMAs. IRMAs epitomize either new vessel growth within the retina or remodeling of pre-existing vessels through endothelial cell proliferation, stimulated by hypoxia bordering areas of capillary non-perfusion [72]. They are larger in caliber with a wide-ranging arrangement and are always contained to the intraretinal layers. Conversely, NVs are fine and delicate in caliber, and more focal in location. In FA, NV often causes leakage whereas IRMAs do not leak. NV and IRMAs, both occur in response to ischemic retina at the severe NPDR or early PDR stage.
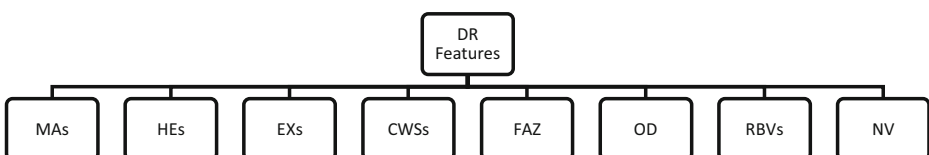

Fig. 6 (a) IRMAs in quadrant 1 [131] (b) Active NV (white arrow), hard EXs (black arrow) in early PDR [131]

Figure 5 shows the presence of (a) IRMAs in quadrant 1 [131] (b) Active NV (white arrow), hard EXs (black arrow) in early PDR [131].

3 Diagnosis of DR using ML

Different techniques are employed for the detection of DR features. The extraction of MAs, HEs, EXs, CWSs, OD, RBVs are performed based on the pre-processing operations. Image pre-processing plays an important role for better feature extraction and classification, as it enhances the properties and attributes of the raw fundus images for better interpretation by an intelligent system. It enhances the contrast of the image, reduces illumination error and blurriness, removes noise, balance intensity in structures, detects minute patterns (subtle lesions) hindered due to bright intensity structures etc. It enables an intelligent system to detect mild and intermediate stages in case of DR. It also highlights the significance of removal of artifacts and background subtraction.

Various techniques such as contrast enhancement [31, 39, 52, 109, 126, 141] for identification of green channel of the image, contrast stretching [39, 52], morphological operations [3, 16, 18, 101, 109, 113, 143], histogram thresholding and histogram equalization [93, 115], smoothing [16], post-processing [16, 88, 92, 130, 145], shade-correction [130], illumination equalization/correction [97], denoising, image restoration using Wiener filter, etc. are performed upon fundus images. MAs and HEs are distinguished using thresholding and adaptive pre-processing [118, 126], morphological image flooding, multiscale Hessian eigen value analysis for vessel enhancement [12] etc.

Various segmentation techniques such as region growing [39, 52] [88, 126, 130, 135, 145], thresholding [5, 39, 47, 51, 52, 88, 92, 109, 126, 135], bottom-hat transform [39, 52, 121, 143], unsupervised segmentation techniques [15], watershed transformation [16], active contour model [16, 67, 82], image reconstruction [16, 102], template matching [68, 111, 136, 142], ensemble-based techniques [5, 33, 105, 113, 142], a priori shape knowledge approach [95], Dynamic Decision Thresholding (DDT) [64], nature-inspired optimization techniques [87, 105], Bayesian Statistical Algorithm (BSA) are used for extraction of DR features. Additionally, similarity-based detection methods [15, 69, 120, 126], Gabor filter incorporated with Hough transform [16, 33, 67], top-hat transformation [68, 97, 126, 130, 141, 142, 145], curvelet transform and level-set-based segmentation techniques [85], canny edge enhancement for boundary detection [19], Sobel operators, Prewitt operators, signal valley analysis [99] are also used for segmentation for obtaining the Region of Interest (RoI), magnitude of intensity of the pixel, and the gradient.

Various other techniques such as sliding window technique [91], Multi Resolution Gabor Transform [115], Gaussian kernels [144], intensity-based techniques [66, 79], statistical classifier [5, 61], Principal Component Analysis (PCA) [46, 67], Singular Value Decomposition (SVD), Linear Discriminant Analysis (LDA), Semantic Image Transformation (SIT) [24], entropy-based backtracking approach [63], ON detection algorithm [133], deformable models [88, 100] and Locally Statistical Active Contour Model with the Structure Prior (LSACM-SP) approach [146] are also used to accomplish the purpose of feature segmentation and extraction, for DR detection. DR classification is performed using Clustering [5, 46, 51, 88, 91, 116, 145], ensemble techniques [13, 14, 18, 141], SVM [22, 63, 108], Sparse Representation Classifier (SRC) [71], Neural Networks [42, 68, 85, 126, 135], *Random Forest Classifier (RFC)* [61, 62, 125], SVM based hybrid classifier [5], *Majority Voting (MV)* [53] etc. Supervised

classification techniques such as KNN classification [37, 93], Extreme Learning Machine (ELM) and Naive Bayes (NB) [17], Bayesian classifier [55], *cascade Adaboost CNN classifier* [8], Naïve–Bayes and Decision Tree (DT) C4.5 enhanced with bagging techniques [46], etc. are used for DR detection.

3.1 DR detection and classification using image analysis

DR detection is highly dependent on assessment and analysis of fundus images. The presence of various DR lesions can be identified using high resolution fundus images. Based on the presence and absence of DR retinal lesions and the corresponding severity level of the disease, this paper classifies DR into five categories such as No DR-0, DR-1, mild Proliferative Diabetic Retinopathy (mPDR)-2, Non-Proliferative Diabetic Retinopathy (NPDR)-3 and Proliferative Diabetic Retinopathy (PDR)-4 [39, 52]. Grade 0 which implies No DR signifies that there is no retinal lesion in the fundus image and hence the patient is not suffering from DR. It designates the fundus is normal and there is no chance of the person losing vision. Grade 1 which implies DR signifies that there exist certain retinal lesions which may be due to foveal enlargement or presence of any early signs such as MAs and HEs. This implies that the fundus is not normal and requires immediate medical treatment. Grade 2 signifies MPDR based on the presence of a few MAs and requires adoption of immediate medical treatment for prevention of blindness. Grade 1 and grade 2 indicates low risks of blindness as MAs and HEs have hardly any impact on the vision and hence vision loss can be prevented. Grade 3 signifies NPDR based on the presence of MAs and HEs, intraretinal hemorrhaging such as venous beading, or a few IRMAs. It is the beginning of a severe stage of the disease where size of retinal lesions may be larger and the patient is at high risk of losing vision. Patient may observe dark spots which indicate a rapid progress towards blindness. Grade 4 signifies PDR, a severe stage of DR based on the rupture of retinal vessels causing multiple blot HEs, flame HEs [15, 82], retinal thickening, unhealthy macula, vascular damage, multiple CWSs, EXs, changes in width of venous caliber, Intra Retinal Microvascular Abnormalities (IRMAs) [15, 82], NVD, NVE and vitreous hemorrhage. This is a critical stage where treatment may be hopeful but the probability of cure is very unlikely.

DR is also categorized into various other categories based on the features identified such as Medhi et al. [91], Akram et al. [31], Noor-ul-huda et al. [98], ETDRS Report Number 10 [49], Hani et al. [52], Fadzil et al. [39], Ege et al. [37], Bhargavi et al. [20], Raja et al. [115], Li. et al. [82], Meshram et al. [93], ETDRS Report Number 7 [35] and Gadekallu et al. [43], have proposed various other phases for DR detection and classification. They have proposed these phases on the basis of presence of DR lesions, absence of edema, increase in retinal thickening, hard EXs [49, 140], foveal enlargement [39, 52, 73, 82, 132], intensity and contrast of features [37] and increase in retinal permeability [35, 43].

3.2 Using supervised ML

Various intelligent and computer-assisted systems are proposed for DR detection using ML techniques [15, 17, 28, 37, 39, 52, 55, 81, 86, 106, 108, 128, 129] such as Support Vector Machine (SVM) [10, 18, 81, 90, 108, 115], Naïve Bayes Classifier [10, 17], Decision Tree [10, 28], K-Nearest Neighbor Classifier [16, 76], Neural Network [17, 90], ensemble classifiers etc.

Jelinek et al. [59] have proposed a multi-lesion detection algorithm, which takes the output of fusion of visual words dictionary-based detectors formed using a set of Points of Interest

(POI) from an image region using Speeded-Up Robust Features (SURF) algorithm, to predict DR. These POIs identify anomalies such as hard EXs, deep HEs, superficial HEs, drusen and CWSs. The proposed model has evaluated three classifier fusion methods such as OR, Majority Voting (MV) and meta-SVM on an imbalanced dataset of 7137 images for single lesion and multi-lesion detection. In the single lesion detection phase, the method has classified hard EXs with an AUC of 91.6%, whereas during multi-lesion detection, it has achieved an AUC of 88.3%. The fusion strategies OR, MV and meta-SVM have achieved a sensitivity of 72%,30% and 80%, a specificity of 86%, 99% and 78%, and an accuracy of 84%, 92% and 78%, respectively.

4 DL models for early detection of DR

Researchers and Scientists all across the globe are continuously trying to make innovations for early detection of DR using ML techniques. They have switched from conventional methods of data acquisition such as Fluorescein Angiography (FA) to Digital Fundus Camera (DFC), for better image data or using a high resolution and large fundus image dataset, to carry out mass diagnosis easily and efficiently. This section discusses various intelligent systems based on Artificial Intelligence (AI) techniques. The section has been divided into applications of three subfields of AI namely supervised ML, unsupervised ML and Evolutionary Algorithms. In supervised ML, the paper discusses about hybrid ML-DL based, pure CNN based and CNN-DL based feature extraction and classification techniques. In unsupervised ML the paper discusses various unsupervised techniques incorporated with Deep Neural Networks (DNNs). In Evolutionary Algorithms, the paper discusses the applications of various nature inspired and metaheuristic optimization techniques, irrespective of supervised or unsupervised learning. All these techniques take into concern various objectives to fulfill for an image related task such as DR such as feature extraction, segmentation, object detection and localization and image classification. In the recent years, various models and intelligent systems using DL, with significant modifications in their proposed methodologies and structures, are also introduced, for early detection of DR. Block Diagram II depicts a brief sketch on various methodologies used for DR detection through use of ML and DL techniques.
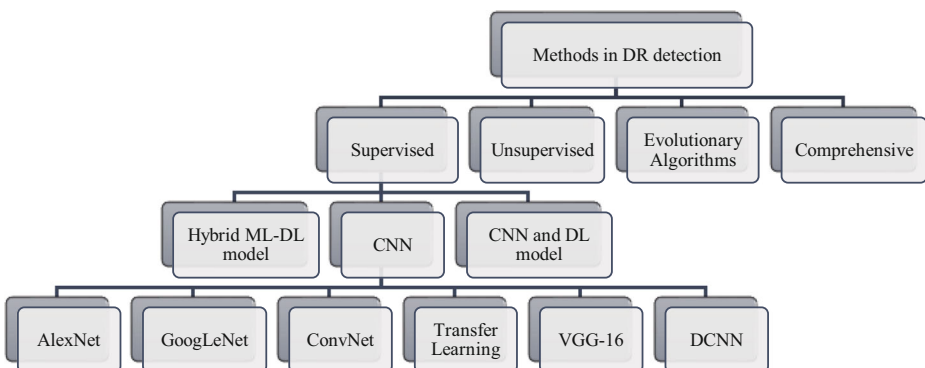

Diagram II Methodologies for DR detection

4.1 Using supervised DL

Various intelligent and computer-assisted systems are proposed for DR detection using DL techniques such as DNN, DCNN, etc. [9].

4.1.1 Using hybrid ML-DL model

This section focuses mainly on hybrid ML-DL techniques used on image and how ML classifiers are introduced as a replacement to standard classification layers of DL CNN models. A few of these methods are discussed below.

Lam et al. [77] have developed an automated system for DR detection using 243 retinal images from Kaggle's EyePACS dataset, and have generated 1324 image patches to detect HEs, MAs, EXs, NV from normal structures. The CNN model is trained on 1050 image patches and tested on 274 image patches. The model has used a $128 \times 128 \times 3$ patch-trained GoogleLeNet-v1 CNN sliding window to scan the image patches and generate a probability score for each of the five classes of DR through detection of MAs and HEs. The model has compared the performance of AlexNet, VGG16, GoogLeNet, ResNet, and Inception-v3 on the 274 test patches and has achieved five-class binary classification accuracy of 74% and 79%, 86% and 90%, 95% and 98%, 92% and 95%, and 96% and 98%, respectively.

Pratt et al. [108] have proposed a CNN architecture to extract DR features using the Kaggle dataset which comprises of 80,000 images, for the detection of DR. The model has used color normalization for preprocessing, real-time data augmentation, L2-regularization for updating weights and biases and cross-entropy loss function for optimization. The proposed CNN is trained using Stochastic Gradient Descent (SGD) and has achieved a specificity of 95% and an accuracy of 75% and sensitivity of 30%.

Xu et al. [142] have proposed a model which uses label preserving transformation for data augmentation and Deep Convolutional Neural Network (DCNN) based image classification, for the detection of DR, using Kaggle's dataset. The proposed methodology has used two classifiers in which one combines each of the extracted features with the Gradient Boosting Machines (GBM) - eXtreme Gradient Boosting method (XGBoost) and the other classifier uses CNN-based features, with and without data augmentation. The entire network is optimized using backpropagation and Stochastic Gradient Descent (SGD). The proposed methodology has detected hard EXs, red lesions, MAs and RBVs. The proposed methodology has obtained an accuracy of 89.4% for hard EXs and GBM, 88.7% for red lesions and GBM, 86.2% for MAs and GBM, 79.1% for RBVs and GBM, 91.5% for CNN without data augmentation and 94.5% for CNN with data augmentation. It is observed that DL models perform better with CNN-extracted features in comparison to conventional feature extraction methods with data augmentation.

Khojasteh et al. [71] have proposed a ten-layered CNN and employs patch-based and image-based analysis upon the fundus images, for the detection of DR. The model has used a total of 284 retinal images from DIARETDB1 and e-ophtha datasets, of which 75 images from DIARETDBI dataset are used for training, for patch-based analysis and the remaining 209 images, both from DIARETDB1 and e-ophtha, are used for testing, for image-based analysis. The model has performed contrast enhancement to extract EXs, HEs and MAs, and then segmented patches of size 50X50 to obtain rule-based probability maps. During the patch-based analysis, the model has detected EXs with sensitivity, specificity and accuracy of 0.96, 0.98 and 0.98, HEs with sensitivity, specificity and accuracy of 0.84, 0.92 and 0.90, and MAs

with sensitivity, specificity and accuracy of 0.85, 0.96 and 0.94, respectively. In the image level evaluation, the proposed method has achieved an accuracy of 0.96, 0.98 and 0.97 with error rate of 3.9%, 2.1% and 2.04% for segmentation of EXs, HEs and MAs, respectively on DIARETDB1 test set, and an accuracy of 0.88, and 3.0, and error rate of 4.2% and 3.1%, for EXs and MAs, respectively, on e-Ophtha dataset. It is observed that simultaneous detection of features can reduce potential error more accurately than individual detection, without any redundancy. It is also observed that post-processing has reduced the error rate and image patching has improved the quality of the images through consideration of the neighborhood and background of candidate lesions.

Soniya et al. [127] have proposed two CNN models 1 and 2, which consists of a single CNN and heterogeneous CNN modules, trained using gradient descent and backpropagation respectively, and are compared to evaluate the effectiveness of detecting DR, using DIARETDB0 dataset. The CNN model 1 has identified MAs, HEs, hard EXs and soft EXs and the CNN model 2 has identified NV. The proposed model has used a multilayer perceptron network classifier with 1620-10-5-6 architecture whose output corresponds to six classes as class 1, class2, class 3, class 4, class 5, and class 6 for normal images, MAs, HEs, hard EXs, soft EXs and NV, respectively. The model has used 130 color images from DIARETDB0 dataset and has performed four experiments of which three experiments have used single CNN with different filter size and receptive field and the fourth experiment has employed heterogenous CNN modules. On using single CNN for the first two experiments, with three convolutional layers each, having 10–30–30 filters for the former, and 30–30-10 filters for the latter, the model has achieved accuracies of 95%, 65%, 42.5%, 67.5% and 92.5% for the former, and accuracies of 95%, 75%, 62.5%, 65% and 95% for the latter, for detection of MAs, HEs, hard EXs, soft EXs and NV respectively. On using single CNN in the third experiment with 50–70–80-100-200 filters, the model has achieved accuracies of 75%, 77.5%, 70%, 52.5%, and 95% for MAs, HEs, hard EXs, soft EXs and NV respectively. The heterogenous CNN modules introduced have achieved 100% accuracy for extraction and detection of class specific features in comparison to the single CNN which has continuously shown low sensitivity and specificity values. It is observed that slight modifications in the filters of CNN, have enhanced the performance of detection of DR lesions. It is also observed that heterogenous CNN has performed better than single CNN.

Alghamdi et al. [8] have proposed an end-to-end supervised model for OD abnormality detection, which constitutes two successive DL architectures with integrated cascade CNN classifiers and abnormality assessment through feature learning, respectively, for the detection of DR. The model has used AdaBoost ensemble algorithm for feature selection and training of the classifier. The proposed approach has used a total of 5781 images from datasets such as DRIVE, DIARETDB1, MESSIDOR, STARE, KENYA, HAPIEE, PAMDI and KFSH. The model has used the annotated OD images in PAMDI and HAPIEE, to train and evaluate the abnormality detector. The model has achieved an accuracy of 100%, 98.88%, 99.20%, 86.71%, 99.53%, 98.36%, 98.13% and 92%, on DRIVE, DIARETDB1, MESSIDOR, STARE, KENYA, HAPIEE, PAMDI and KFSH respectively, for OD localization. The proposed OD abnormality detector has achieved a sensitivity of 96.42%, a specificity of 86% and an accuracy of 86.52% on HAPIEE dataset, and a sensitivity of 94.54%, a specificity of 98.59% and an accuracy of 97.76% on PAMDI dataset. It is observed that the use of *cascade classifiers* which is an ensemble of weak classifiers can work well with good quality images only and cannot withstand variations and hence requires learning of discriminative features using CNN.

Amongst all these, SVM, Neural Networks (NN), hybrid ML-DL models and ensemble algorithms have produced effective results upon an effective dataset. ML techniques exhibit high generalization error and presents only sub-optimal solutions, and are incompetent against any real-time, complex, and high-dimensional data such as medical images (here fundus images). Moreover, they lack domain awareness and representation, which makes them computationally intensive and inflexible to extract patterns and relationships using handcrafting rules and algorithms upon high-dimensional image data.

4.1.2 Using deep learning CNN

Deep Learning (DL) is a new advent of ML and inherits the appropriate and advantageous attributes of ML such as perform complex tasks, smart and automated, better generalization, domain knowledge, decision making etc. and efficiently applies them upon image data, thereby outperforming shallow ML algorithms. DL permits data processing of diverse types in amalgamation, known as cross-modal learning (multiple forms of representations), and can generate well-defined features through automated feature learning, unlike ML algorithms which are dependent on various feature extraction algorithms and procedures. DL models have the capability to learn and to generate new features from extracted and existing features such as points, lines, edges, gradients, vessel structure, corners, boundaries etc. using representation learning.

DL supervised networks based on Convolutional Neural Network (CNN) [17, 70, 115] are used for deep feature extraction and image classification. DL can be encouraging to find an early and new remedy for detection of DR. A few of the DL CNN based techniques for DR detection are discussed below-.

Orlando et al. [101] have proposed an *ensemble CNN approach based on LeNet architecture*, for the detection of MAs and HEs, for DR detection. The method has used r-polynomial transformation for pre-processing, Gaussian filter to reduce noise, morphological operations to avoid noise and image patching to recover candidate red lesions. The CNN is trained with Stochastic Gradient Descent (SGD), cross entropy loss function and weight decay. The model has used datasets such as e-Ophtha for training, DIARETDB1 (Standard Diabetic Retinopathy Database) for per-lesion detection, and MESSIDOR for image-level evaluation. The model has achieved per lesion sensitivity of 0.2885,0.202 and 0.2 for combined, CNN, and hand-crafted features, respectively for False Positive per Image (FPI) value of 1, in the interval [1/8,1/4,1/2,1,2,4,8], for MAs detection in DIARETDB1 dataset. On detecting HEs on the DIARETDB1 dataset, the combined approach has reported a per lesion sensitivity of 0.4907 than manually engineered descriptors achieving a per lesion sensitivity of 0.4724, at FPI value of 1. On experimenting MESSIDOR dataset and using CNN as a classifier, the model has achieved an AUC of 0.7912 with CNN features, 0.7325 with hand crafted features, and 0.8932 on combination of both features. It is observed that, the combination of both CNN and hand-crafted features is a better option for DR detection than CNN or hand-crafted features, alone.

Wang et al. [138] have proposed two CNN models namely Net-5 and Net-4 for a large-scale DR dataset, which includes a Regression Activation Map (RAM) layer to minimize the number of parameters through the inclusion of Global Average Pooling (GAP) layer instead of fully connected layers, for DR detection. The model has used 35,126 images from Kaggle dataset for training and testing in the ratio of 9:1 and has implemented a baseline model of three networks -small, medium and large using 128-pixel images, 256-pixel images and 512-pixel images, respectively. The proposed model has analyzed and compared its performance

using orthogonal initialization, data augmentation and feature blending. The model has used *Fully Convolutional Neural Network (FCNN)* upon the blended features to obtain the final predicted regression values discretized at the thresholds (0:5; 1:5; 2:5; 3:5) so as to obtain integer levels for computation of Kappa scores. On the validation set, the proposed network has achieved a Kappa score of 0.70 for 256-pixel images, 0.80 for 512-pixel images and 0.81 for 768-pixel images, on both Net-5 and Net-4 settings of the architecture, without feature blending, respectively. It is observed that as the number of pixels has increased, the performance of the proposed system has enhanced along with cost of computation but stopped beyond a threshold of 512 and above. This has widened the scope of further implementation of the system at pixel level and work on enhancing the discrete ranges of regression values.

AlexNet based feature extraction Dai et al. [27] have proposed a 5-staged image-to-text mapping expert-guided statistical model and an interleaved deep mining CNN called 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. The preprocessed fundus images are over-segmented using Simple Linear Iterative Clustering (SLIC) for AlexNet-based feature extraction and random partition of feature space through semantic mapping and random fern approach. The proposed methodology is a 'partition frequency-inverse lesion frequency' model which represents and predicts certain lesion types for each over-segmented super-pixel. The MS-CNN has taken a 'r x r' patch, where r = 64, upon which candidate selection, mean filtering, segmentation of MAs using top-hat transform and Gaussian filter, are performed and a cascaded CNN classifier is used for classification. The proposed model has used a dataset of 646 images and DIARETDB1 dataset of 89 fundus images for training and testing. The model has achieved 17.9% recall, 100% precision, 17.8% accuracy and F1 score of 30.4% in the first stage of MS-CNN, and has achieved 87.8% recall, 99.7% precision, 96.1% accuracy, and F1 score of 93.4% in the second stage of MS-CNN. It is observed that the proposed model is effective and has scope for inculcation of DL techniques in text-to-image mapping, deep feature extraction and image classification.

Abràmoff et al. [2] have compared the performance of a DL enhanced algorithm, to the earlier published non-DL algorithm, the Iowa Detection Program (IDP) based on the MESSIDOR-2 dataset consisting of 1748 augmented images, and reference standard set by retinal specialists, for automated detection of DR. The proposed model has implemented a CNN motivated IDx-DR device named IDx-DR X2.1 which is a lesion detector and has implemented AlexNet for 10,000 augmented samples and VGGNet for 1,250,000 augmented samples, for detection of referable DR (rDR) and vision-threatening DR (vtDR), so as to detect various phases of DR and Macular Edema (ME). The model has obtained feature vectors of the predicted abnormality, which are fed into two fusion classifiers implemented using *Random Forest (RF)*. The proposed CNN detectors and classifiers are trained on 25,000 complete examinations of four expert annotated photographs per subject for detecting normal anatomy such as OD and fovea and DR lesions such as HEs, EXs and NV. The model has achieved a sensitivity of 96.8%, a specificity of 87%, a negative predictive value of 99.0%, positive predictive value of 67.4% and AUC of 0.980 for rDR detection, and a sensitivity of 100%, a specificity of 90.8%, negative predictive value of 100.0%, positive predictive value of 56.4% and AUC of 0.989 for vtDR detection. It is observed that the proposed methodology has achieved an overlap in detection of PDR and ME. It is also observed that the authors have claimed that the representation of MESSIDOR-2 dataset as a reference standard is not suitable

for screening algorithms whereas various models proposed earlier have produced impactful results based on it.

GoogLeNet based feature extraction Gulshan et al. [50] have proposed a DL algorithm for computerized detection of DR and Diabetic Macular Edema (DME), using fundus images and have identified HEs and MAs using Inception-V3-architecture Neural Network, for the detection of DR. The proposed methodology has deployed the EyePACS-1 dataset consisting of 9963 images and the MESSIDOR-2 dataset having 1748 images. The model has performed preprocessing, network weight optimization using distributed Stochastic Gradient Descent (SGD), and hyperparameter optimization. The entire development set of 128,175 images, has used an *ensemble of 10 networks* and computed the final linear average prediction upon the ensemble predictions. The algorithm has detected rDR with an AUC of 0.991 for EyePACS-1 and 0.990 for MESSIDOR-2. At high specificity in the development set, EyePACS-1 has achieved a sensitivity of 90.3%, and a specificity of 98.1% whereas MESSIDOR-2 has achieved a sensitivity of 87.0% and a specificity of 98.5%. At high sensitivity in the development set, EyePACS-1 has achieved a sensitivity of 97.5% and a specificity of 93.4% whereas MESSIDOR-2 has achieved a sensitivity of 96.1% and a specificity of 93.9%. On approximation of 8% prevalence of rDR per image, the model has achieved a negative predictive value of 99.8% for EyePACS-1 and 99.6% for MESSIDOR-2. The proposed algorithm has evaluated moderate or worse DR, rDME, and ungradable images using the EyePACS-1 dataset only and has achieved an AUC of 0.974. At high specificity operating point, the algorithm has achieved a sensitivity of 90.7% and a specificity of 93.8%, and at high sensitivity operating point, the algorithm has achieved a sensitivity of 96.7% and a specificity of 84.0%. It is observed that rDR detection is adopted for detection of DR and DME and that there are higher probabilities of lesion overlap or misclassification of lesions, if the training dataset is imbalanced.

Takahashi et al. [132] have proposed a modified and randomly initialized GoogLeNet DCNN for the detection of DR, which is trained using 9443 of the 9939-posterior pole color fundus images and have used manual staging with three additional fundus images. The proposed model is a composition of AI1 model and AI2 model in which the AI1 model is trained using ResNet and both the models are trained on modified Davis grading of a concatenated figure of four photographs. The modified Davis grading includes Simple DR(SDR), Pre-proliferative DR (PPDR) and PDR. The AI2 model of the network is also trained on the same number of images using manual staging with only one original image which is used to detect retinal HEs and hard EXs and is trained on the pairs of a patient's image and its modified Davis grading. The model has achieved a Prevalence and Bias-Adjusted Fleiss' Kappa (PABAK) of 0.74 and a mean accuracy of 81%, on 496/9443 images (5%). On manual grading with one image, the model has achieved a PABAK of 0.71 with a mean accuracy of 77%, whereas on manual grading of four images for No DR(NDR), SDR, PPDR and PDR detection, the model has achieved a PABAK of 0.64 with a mean accuracy of 72%, on 496/9443 images. The proposed model has graded randomly chosen 4709 images of the total 9939 posterior pole fundus images from 0 to 14 using real prognoses, of which 95% are used for training and rest 5% are used for validation. The modified GoogLeNet has achieved a PABAK of 0.98 with mean accuracy of 96% during real prognosis, and a PABAK of 0.98 using traditional modified Davis grading with a mean accuracy of 92%, in 224 unseen images. The three retinal specialists HT, YA, and YI, who have graded the images during real prognosis have achieved a PABAK of 0.93 with mean accuracy 93%, 0.92 with mean
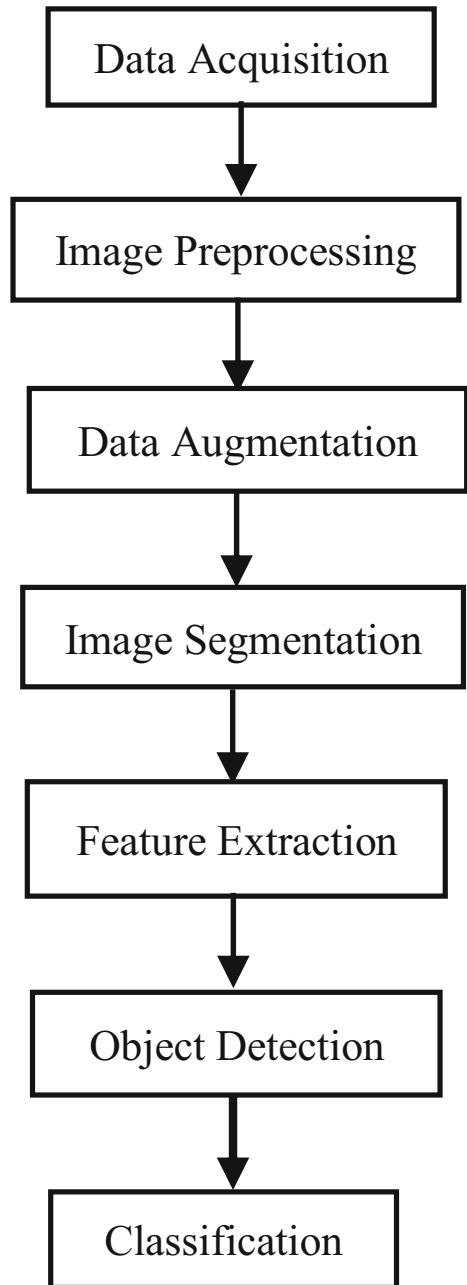
accuracy of 92%, and 0.93 with mean accuracy of 93%, respectively. It is observed that the work has identified surface reflection in retina as an abnormality seen in young people which can be helpful for DR detection in young generation. It is observed that the work has identified surface reflection in retina as an abnormality seen in young people which can be helpful for DR detection in young generation. It is observed that the four concatenated image-trained neural network is more useful and better at detecting DR than one image-trained neural network, to obtain more features and understand the behavior of DR through multiple gradings. The proposed model is a continuation of previous proposed models based on ResNet-52 and ResNet-152 having memory constraint and thereby establishes a comparison in their performance.

ConvNet based feature extraction Quellec et al. [112] have proposed a DL detector Convolutional Network (ConvNet) for detection of MAs, HEs, EXs and CWSs, and new biomarkers of DR, for rDR detection, using 88,702 fundus images from the 2015 Kaggle DR dataset, 107,799 images from e-ophtha dataset and 89 images from DiaretDB1 dataset. The proposed model has evaluated heatmap generation and an optimization solution for DR screening, using backpropagation-based ConvNets. The proposed model has adapted min-pooling based image preprocessing, data augmentation and DL fractional max-pooling based network structures (o_O solution) namely netA and netB, for visualization. It has used an ensemble of RF classifiers. The performance of the model is evaluated w.r.t lesion level and more specifically in the image level, using the DiaretDB1 dataset, where MAs, HEs, EXs and CWSs are manually segmented. The model has achieved a detection performance, area under ROC(A_z) of 0.954 in Kaggle's dataset using netB and A_z of 0.949 in e-ophtha dataset, for rDR detection and an A_z of 0.9490 upon ensemble classifier. It is observed that detection quality is not dependent on image quality and visualization of images through heatmaps using ConvNets helps in detection of subtle lesions especially using the netB. The proposed model requires more manual segmentation of lesions for advanced signs of DR such as NV.

Block Diagrams III and IV depicts two different frameworks, used for the detection of DR using conventional ML model and advanced DL model.
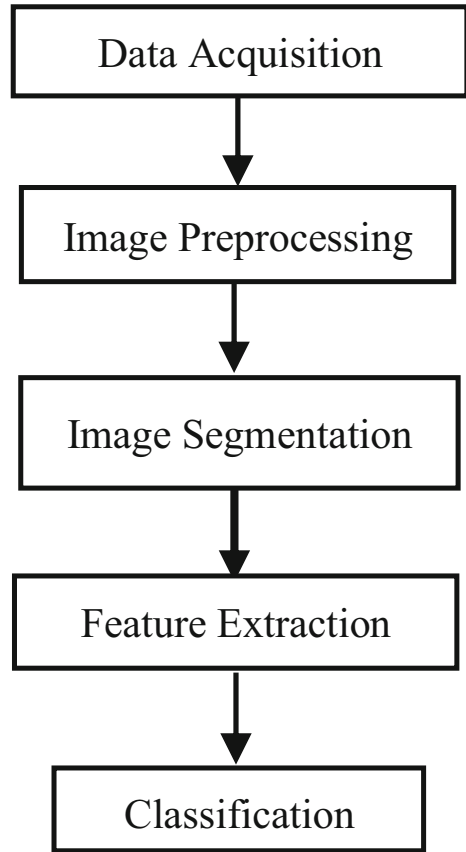
VGG-16 based feature extraction Dutta et al. [34] have proposed a DL model which is trained using backpropagation Neural Network (NN), DNN and VGG-16, for DR detection. The model has acquired the retinopathy images from Kaggle dataset and has used 2000 images in the ratio of 7:3 for training and testing the model. The model has extracted DR features such as RBVs, fluid drip, EXs, HEs and MAs and has thresholded each target lesion using Fuzzy C-Means clustering. The proposed model has performed image filtering and background subtraction, using median filter and morphological processing, respectively. The model has extracted the edge and border features using Canny edge detection. The VGG-16 model has achieved better testing accuracy, in comparison to backpropagation neural network and DNN. The backpropagation NN model has achieved a training accuracy of 45.7% and a testing accuracy of 35.6% whereas the DNN model has achieved a training accuracy of 84.7% and a testing accuracy of 82.3%, for the statistical data model. The ***VGG-16*** model has considered 1000 images for training and 300 images for testing and has achieved an accuracy of 72.5%. During the training and the testing phase, the DNN model has achieved an accuracy of 89.6% and 86.3%, respectively, on image data model than backpropagation NN which has achieved a training accuracy of 62.7% and testing accuracy of 42% and ***VGG-16*** which has achieved a training accuracy of 76.4% and testing accuracy of 78.3%, for image classification.

Diagram III General Framework
[7, 22, 34, 36, 96, 101, 112]

```
┌─────────────────────────┐
│    Data Acquisition     │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│   Image Preprocessing   │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│    Data Augmentation    │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│   Image Segmentation    │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│   Feature Extraction    │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│    Object Detection     │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│     Classification      │
└─────────────────────────┘
```

Grinsven et al. [136] have proposed a technique to improve and accelerate CNN training through dynamic sampling of misclassified negative samples at pixel-level, for detection of DR. The proposed model has acquired the image data from the Kaggle dataset which contains 35,126 training images and 53,576 test images of which 6679 training images are used. The Kaggle test set and the MESSIDOR dataset of 1200 images, are used as test images to extract

Diagram IV Supervised and Unsupervised Learning Framework using ML and DL [8, 27, 45, 57, 70, 71, 78, 142]

```
┌─────────────────────────┐
│   Data Acquisition      │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│   Image Preprocessing   │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│   Image Segmentation    │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│   Feature Extraction    │
└─────────────────────────┘
            │
            ▼
┌─────────────────────────┐
│     Classification      │
└─────────────────────────┘
```

candidate HEs. The proposed model has performed preprocessing, contrast enhancement, segmentation using circular template matching and data augmentation by spatial translation. The proposed model has compared two CNN models namely CNN iterative Selective Sampling (SeS) model having 60 epochs and iterative non-Selective Sampling (NSeS) model having 170 epochs, and at 1 FPI the models have achieved sensitivities of 78.6% and 75.3%, whereas at 0.1 FPI, both CNNs have achieved sensitivities of 51.1% and 31.6%, respectively. The model has achieved image level performance Az values of 0.919 and 0.907, on the Kaggle and MESSIDOR test sets, respectively using CNN(SeS) 60 and 0.981 and 0.967 on the Kaggle and MESSIDOR test sets, respectively using CNN (NSeS) 170. Two observers have also graded the test sets and the model has compared their performance with the two proposed models. Using Kaggle dataset, observer 1 has achieved a sensitivity and a specificity of 81.6% and 94.7%, observer 2 has achieved a sensitivity and a specificity of 80.6% and 94.2%, CNN(SeS) 60 has achieved a sensitivity and a specificity of 83.7% and 85.1% and CNN(NSeS) 170 has achieved a sensitivity and a specificity of 77.4% and 85.1%, respectively. Using Messidor dataset, observer 1 has achieved a sensitivity and a specificity of 97.6% and 89.4%, observer 2 has achieved a sensitivity and a specificity of 95.8% and 87.2%, CNN(SeS) 60 has achieved a sensitivity and a specificity of 93.1% and 91.5% and CNN(NSeS) 170 has achieved a sensitivity and a specificity of 90.3% and 93.1%, respectively. The model has also achieved an AUROC of 0.894 and 0.972 on Kaggle and MESSIDOR datasets, respectively. It

is observed that hyperparameter tuning and optimization has a significant role to play in enhancing the performance of the model.

Liskowski et al. [84] have proposed a supervised DL Retinal Blood Vessel (RBV) segmentation technique using a ***DNN,*** trained on 400,000 examples, which are enhanced, contrast normalized, and amplified using geometric transformations and gamma corrections, for the detection of DR. The model has acquired images from DRIVE, STARE and CHASE datasets. The proposed model has trained the network using backpropagation and dropout. The model has proposed two basic configurations namely PLAIN BALANCED and NO-POOL which are dependent on structure prediction for simultaneous classification of multiple pixels. The PLAIN BALANCED model has achieved highest AUC of 0.9738 on DRIVE and 0.9820 ± 0.0045 on STARE, but has achieved a better accuracy of 0.9620 ± 0.0051 in the STARE dataset. The proposed model has achieved an area under ROC curve measure of >99% and classification accuracy of >97%. The technique is resilient to the phenomenon of central vessel reflex and sensitive in recognition of fine vessels with a measure of >87%.

Deep convolutional neural network based feature extraction Islam et al. [57] have proposed a *DCNN* for early-stage detection of DR using Kaggle's EyePACS dataset of 88,702 images of which 35,126 are training images and 53,576 are testing images, through identification of MAs. The model has performed preprocessing through rescaling, followed by data augmentation to reduce imbalance data, feature blending, orthogonal weight initialization, Stochastic Gradient Descent (SGD) optimization, L2 regularization and Adam optimizer for model training. In binary classification problem, the proposed method has achieved a sensitivity of 98% and specificity of 94%, in low-high DR detection, and a sensitivity of 94.5% and specificity of 90.2% in healthy-sick DR detection. The model has used threshold coefficients of (0:5; 1:5; 2:5; 3:5) to discretize the predicted regression values and convert the class levels into integers, which has led to the achievement of a quadratic weighted kappa score of 0.851 on test set. The proposed model has achieved an AUROC and F-Score of 0.844 and 0.743, respectively on the dataset. It is observed that the proposed method is suitable only for binary classification of intermediate stages of DR.

Prentasic et al. [109] have proposed a DCNN which considers RBVs and OD in the DR detection procedure to upsurge the accuracy of EXs detection, for DR detection. The model has performed image preprocessing using Frangi vesselness filter, Total Variation (TV) regularization denoising and split Bregman algorithm for denoising, morphological operations, dynamic thresholding, pixel-wise feature extraction and classification and clustering approaches. The proposed model has combined different landmark detection algorithms, for the detection and localization of EXs. The proposed model has used an ensemble of OD detection algorithms such as entropy-based method, Laplacian of Gaussian (LoG) filtering method, brightness method, Simulated Annealing (SA) and Hough transformation of vessels, which performs various preprocessing, thresholding, localization and object detection. The proposed CNN model has classified an EX or a non-EX using DRiDB dataset of 50 images in which the model has achieved a sensitivity of 78%, a Positive Predictive Value (PPV) of 78% and an F-score of 0.78. It is observed that landmark detection of retinal features such as OD and RBVs plays an important role in identifying and analyzing abnormal retinal features either through their inclusion in severe cases or through their subtraction in normal and mild cases.

Pour et al. [107] have used EfficientNet B5 for feature extraction and classification, for DR detection using MESSIDOR, MESSIDOR-2 and IDRiD datasets, upon Contrast Limited

Adaptive Histogram Equalization (CLAHE) based preprocessed images. The model has achieved an AUC of 0.945 on MESSIDOR, and AUC 0.932 on IDRiD.

Chetoui et al. [25] have used EfficientNet B7 for feature extraction and classification, and Global Average Pooling, for the detection of Referable Diabetic Retinopathy (RDR) and vision-threatening DR. The model has used Kaggle EyePACs and Asia Pacific Tele-Ophthalmology Society (APTOS) 2019 datasets, and have extracted features such as EXs, HEs and MAs using Gradient-weighted Class Activation Mapping (Grad-CAM). The model has achieved an AUC of 0.984 for RDR and 0.990 for vision-threatening DR on EyePACS dataset, and for APTOS 2019 dataset the model has achieved an AUC of 0.966 and 0.998 for referable and vision-threatening DR, respectively.

Rakhlin [117] has proposed a *VGG-Net based modified DCNN* model for identification of DR features such as drusen, EXs, MAs, CWSs and HEs, and has used Kaggle dataset consisting of 88,696 images of which 81,670 images are used for training, and the rest 7026 images along with the entire MESSIDOR-2 dataset consisting of 1748 augmented images, are used for testing, for the detection of rDR. The diagnostic pipeline of the proposed model comprises of preprocessing, image quality assessment module, generation of randomly augmented images, localization and segmentation of features, dropout, and classification of retinal lesions. In the Kaggle dataset, the model has achieved AUROC of 0.923, sensitivity of 92%, and specificity of 72%, at high sensitivity operating point. At high specificity operating point, sensitivity and specificity of the model in Kaggle dataset is 80% and 92%, respectively. In the MESSIDOR-2 dataset, the model has achieved AUROC of 0.967, sensitivity of 99%, specificity of 71%, at high sensitivity operating point, and at high specificity operating point, the model has achieved a sensitivity and specificity of 87% and 92%, respectively. It is observed that the proposed model has outlined the importance of the image assessment module for detection of subtle lesions, and the importance of grading standard of DR, for DR detection.

Chaturvedi et al. [23] have utilized a pre-trained DenseNet121 network on 3662 fundus photography images, obtained from 5-class APTOS2019 dataset, for early detection of DR. The proposed method has achieved 96.51% validation accuracy in multi-label and multi-level DR classification and achieved 94.44% validation accuracy for single-class classification method.

Li et al. [83] have proposed a pure DCNN and a modified DCNN approach, using fractional max-pooling, for DR classification. The model has used 34,124 images from the publicly available DR Kaggle dataset for training which are preprocessed and the parameters are optimized using Teaching-Learning-Based Optimization (TLBO). The model has used 1000 validation images and 53,572 testing images and has achieved a recognition rate of 86.17%.

Transfer learning based feature extraction and classification Various pre-trained DL models also known as Transfer Learning (TL) techniques have found application for DR detection. They are basically applied when the dataset is really very small and when there are higher chances of occurrence of underfitting of data or higher generalization error. In such cases TL can be applied over standard DL techniques. TL enables combination of features consequential from various layers of pre-trained models which helps in boosting the performance of image classification. Such techniques also facilitate amalgamation of feature representations from pre-trained VGG16 and Xception [21] specifically, using a set of feature blending approaches, which boosts the classification performance and lowers high generalization error. A few of the TL-based methods are discussed below-.

Lam et al. [78] have proposed TL techniques for exploring optimal CNN models, on 35,000 fundus images from Kaggle dataset and 1200 fundus images from MESSIDOR-1 dataset, for the detection of DR. The proposed model has implemented CLAHE for preprocessing, real-time data augmentation and Otsu's thresholding for segmentation of the fundus images. The proposed model has used hyperparameter tuning techniques such as batch normalization, L2 regularization, dropout, learning rate, cross entropy loss function, weight initialization and gradient descent update. The model is trained and tested using pretrained AlexNet and GoogLeNet, as 2-ary, 3-ary and 4-ary classification models, where GoogLeNet has performed better than AlexNet. The model has performed binary classification (normal or mild vs moderate or severe) using AlexNet, VGG16 and GoogLeNet models on Kaggle dataset in which GoogLeNet has achieved better performance with a sensitivity of 95% and specificity of 96%. In the multi-class classification phase, the 3-ary classifier has achieved a sensitivity of 98%, 93% and 7% for no DR, severe DR and mild DR, respectively on Kaggle dataset, and a sensitivity of 85%, 75% and 29% for no DR, severe DR and mild DR, respectively on MESSIDOR-1 dataset. The 3-ary classifier has also achieved test accuracy of 67.2% and 71.25% on raw-data and TL data, respectively. The 4-ary classifier gets inclined towards majority classification and has failed to train GoogLeNet. Using TL, the proposed methodology has achieved a peak test set accuracy of 74.5%, 68.75%, and 51.25% on 2-ary, 3-ary, and 4-ary classification models, respectively. It is observed that the proposed method face challenges w.r.t. dataset fidelity, misclassification and inclination towards majority classification, with or without TL.

Alban et al. [6] have incorporated TL-based DL networks, Non-Local Means Denoising (NLMD) for prediction of features and noise, image restoration and data augmentation, to diagnose DR, upon 35,126 images, acquired from Kaggle's EyePACS dataset. The proposed model has addressed data imbalance using over-sampling and cost-sensitive learning. The proposed model has three different models such as the baseline model, a classifier using *pretrained AlexNet and a GoogLeNet*, in addition to two error troubleshooters namely 2-class classifier for binary classification of DR and 3-class classifier for merging classes. The baseline model has achieved an accuracy of 54.1%, recall of 0.502 and precision of 0.489 for 2-class classification, an accuracy of 35.3%, recall of 0.387 and precision of 0.301 for 3-class classification, and an accuracy of 22.7%, recall of 0.201 and precision of 0.235 for 5-class DR classification. The pretrained AlexNet model has achieved an accuracy of 66.95% for 2-class classification, an accuracy of 57.05% for 3-class classification and an accuracy of 40.73% for.

5-class classification. The pretrained GoogLeNet model has achieved an accuracy of 71.05% for 2-class classification, 58.21% for 3-class classification and 41.68% for 5-class classification. The proposed GoogLeNet for 5-class severity classification namely class 0, class 1, class 2, class 3 and class 4 has achieved an AUC of 0.79 whereas AlexNet has achieved an AUC of 0.69. It is observed that the GoogLeNet based TL model has performed better than AlexNet based TL model, using inception modules and has detected features without redundancy. On analysis of error, the model has endured factors such as black space, eye image color, low image brightness, and bad images, which have contributed to misclassification of the images and requires to be mitigated.

Kermany et al. [70] have proposed a diagnostic tool based on *Inception V3 architecture* pretrained TL model for the screening of DR patients and has used 207,130 Optical Coherence Tomography (OCT) labelled images of which 108,312 images are used for training the AI system and 1000 images are used for testing. The AI model has considered DR abnormalities

such as Choroidal Neovascularization (CNV), Diabetic Macular Edema (DME) and drusen, for DR detection. The model has extracted and localized the RoI of DR features through occlusion testing using an occlusion window of size 20 × 20 and computed the probability to designate the RoI which is responsible for the learning of the algorithm. In a multi-class comparison between CNV, DME, drusen, and normal, the proposed model has achieved an accuracy of 96.6%, with a sensitivity of 97.8%, a specificity of 97.4%, a weighted error of 6.6% and Receiver Operating Characteristic (ROC) of 99.9%. On binary classification, CNV versus normal images, the model has achieved an accuracy of 100.0%, with a sensitivity of 100.0%, specificity of 100.0% and ROC of 100.0% and for DME versus normal images, the model has achieved an accuracy of 98.2%, a sensitivity of 96.8%, specificity of 99.6% and ROC of 99.87%. Again, the classifier which distinguishes drusen has achieved an accuracy of 99.0%, with a sensitivity of 98.0%, specificity of 99.2% and ROC of 99.96%. It is observed that TL is suitable for binary classification and imaging modality has an important role to play, as for instance the proposed model concludes OCT imaging to be more reliable than fundus photography. It is also observed that pure detection of DR abnormalities is a challenge using occlusion testing as lesions tend to appear larger or smaller than the occlusion window due to random initialization on a large dataset of images.

4.1.3 DR detection using DL models

Gargeya et al. [45] have proposed a data-driven DL algorithm for deep feature extraction and image classification, using Deep Residual Learning (DRL) to develop a CNN for automated DR detection. The model is trained using 75,137 fundus images from EyePACS dataset, and tested using an augmented MESSIDOR-2 dataset and E-Ophtha dataset, containing 1748 and 463 images, respectively. The proposed model has performed preprocessing, dataset augmentation, batch normalization, ReLU activation, and categorical cross entropy loss function for class discrimination using gradient boosting classifiers. The model has extracted 1024 deep features using the convolutional method. The model has detected retinal HEs, hard EXs and NV, through visualization of heatmaps. The proposed model has achieved an AUC of 0.97 with an average sensitivity of 94% and specificity of 98% on EyePACS dataset, whereas it has achieved and AUC of 0.94, with an average sensitivity of 93% and specificity of 87% on MESSIDOR-2 dataset, and an AUC of 0.95 with an average sensitivity of 90% and specificity of 94% on E-Ophtha dataset. It is observed that the implementation of a residual network has eased and excelled the training of the proposed network with augmented data and heatmap visualization. In most of the studies, detection of NV has proved to be critical but with DRL, it was possible. In the future, DRL based deep architectures can probably excel independently without data augmentation and visualization.

Eftekhari et al. [36] have proposed a Deep Learning Neural Network (DLNN), which is a two-stage training architecture consisting of two completely different structures of CNN namely a basic CNN and a final CNN, for detection of MAs, for the diagnosis of DR. The proposed model has used images acquired from datasets such as Retinopathy Online Challenge (ROC) containing 100 images and E-Ophtha-MA containing 381 images, to train and test the model. The proposed model has performed pre-processing and has generated a probability map in the basic CNN to detect MAs and non-MAs, which has led to a balanced dataset. The model has performed backpropagation for optimization of parameters, post-processing upon the output of final CNN, and has used Stochastic Gradient Descent (SGD),

dropout and binary cross-entropy loss function for training. The proposed method is assessed on ROC and E-Ophtha-MA datasets, and has achieved a sensitivity of 0.8 for an average of >6 FPI. The proposed method has achieved sensitivities of 0.047, 0.173, 0.351, 0.552, 0.613, 0.722, and 0.769 for FPI values in the interval [1/8,1/4, 1/2, 1,2, 4, 8] respectively for the ROC, and for E-Ophtha, the proposed method has achieved sensitivities of 0.091, 0.258, 0.401, 0.534, 0.579, 0.667, and 0.771 for FPI values in the interval [1/8,1/4, 1/2, 1,2, 4, 8], respectively. The proposed model has also achieved Free-response Receiver Operating Characteristic Curves (FROC or F_{AUC}) of 0.660 for ROC dataset and 0.637 for E-Optha-MA dataset. It is observed that the evaluation of the model using parameters such as False Positive per Image (FPI) and Free-response Receiver Operating Characteristic (FROC) curve have proved to be effective in eradication of misclassification.

Al-Bander et al. [7] have proposed a multi-sequential DL technique for detecting the centers of OD and fovea, for the detection of DR, using CNNs. The model has used the MESSIDOR database of 1200 images and 10,000 images from the Kaggle dataset, for training and testing respectively. The proposed model has enhanced the contrast of the resized image using CLAHE and has obtained the ROIs using the first CNN and performed classification using the second CNN. The proposed model is trained on augmented data using Stochastic Gradient Descent (SGD). The proposed model has detected the OD and fovea, based on 1R, 0.5R and 0.25R conditions where R refers to the radius of OD. The proposed method has achieved accuracies in terms of the 1R criterion of 97% and 96.6% for detection of OD and foveal centers, respectively in MESSIDOR test set and 96.7% and 95.6% for the detection of the OD and foveal centers, respectively in the Kaggle test set. On the Kaggle test set, the model has obtained accuracies of 95.8% and 90.3% for OD detection, for 0.5R and 0.25R criterions, respectively, while 90.7% and 70.1% were achieved for fovea detection. On MESSIDOR, the model has obtained accuracies of 95% and 83.6% for 0.5R and 0.25R criterions, for localizing OD and 91.4% and 66.8%, for the foveal center detection. It is observed that the detection of OD and fovea center plays a vital role in fundus examination and abnormality detection related to DR. The standard radius of the OD highlights the geometrical and morphological properties of OD and fovea for effective segmentation and feature extraction.

4.2 DR detection using unsupervised DL

DL unsupervised networks such as Inception Convolutional Recurrent Neural Networks (IRCNN), Generative Adversarial Network (GAN), Autoencoder, Restricted Boltzmann Machine (RBM), Long Short-Term Memory (LSTM), and semi-supervised DL networks such as Deep Reinforcement Learning (DRL) are also used for deep feature extraction and image classification.

Mansour et al. [89] have proposed AlexNet-based DR model, which performs a comparative study on DL based feature extraction techniques against ML based feature extraction methods, and classifies the fundus images for the recognition of DR. The proposed methodology has applied a multi-level optimization measure that incorporates data collection from Kaggle dataset, preprocessing, adaptive learning Gaussian Mixture Model (GMM)-based region segmentation, Connected Component Analysis (CCA) based localization and DNN feature extraction. The model has segmented hard EXs, blot intraretinal HEs and MAs. The model has derived high dimensional features from the Fully Connected (FC) layers 6 and 7 of the DNN and has used PCA and Linear Discriminant Analysis (LDA) for dimensionality reduction and feature selection [58]. The proposed model has performed optimal five-class DR

classification using Radial Basis Function (RBF) kernel-based SVM. The DNN has achieved a better classification accuracy upon high dimensional FC6 features and FC7 features, for feature extraction, using PCA and LDA, in comparison to Scale Invariant Feature Transform (SIFT) features, for image classification thus producing a hybrid ML-DL model which can outperform other ML algorithms on grounds of effective and deep feature extraction. The AlexNet DNN has achieved a classification accuracy of 90.15% on FC6 features and a classification accuracy of 95.26% on FC7 features, using PCA whereas it has achieved a classification accuracy of 97.93% on FC6 features and a classification accuracy of 97.28% on FC7 features, using LDA. On using SIFT features, the proposed model has achieved a classification accuracy of 91.03% using PCA whereas it has achieved a classification accuracy of 94.40% using LDA.

Chudzik et al. [26] have proposed a *patch-based Fully CNN* (FCNN) model resembling a *convolutional Autoencoder* which performs image preprocessing, data preprocessing to mitigate data scarcity, patch extraction using a sliding window to map the corresponding annotation, and pixel-wise probabilistic classification, for the detection and segmentation of MAs, for DR detection. The model has used batch normalization layers and Dice coefficient loss function upon images from E-Ophtha, DIARETDB1, and Retinopathy Online Challenge (ROC) datasets, consisting of 381, 100 and 89 images, respectively. The model is trained on 354 images and evaluated on 27 images, from the E-Ophtha dataset. The proposed method has obtained sensitivities of 0.039, 0.067, 0.141, 0.174, 0.243, 0.306, 0.385, and a FROC score of 0.193 ± 0.116 for low FPI values in the interval [1/8,1/4,1/2,1,2,4,8] and sensitivities of 0.174, 0.243, 0.306, 0.385, 0.431, 0.461, 0.485, and a FROC score of 0.355 ± 0.109 for FPI values in the interval [4, 10, 39, 40, 48, 74, 98], on ROC training dataset. The model has achieved sensitivities of 0.187, 0.246, 0.288, 0.365, 0.449, 0.570, 0.641, and FROC score of 0.392 ± 0.157 for FPI values in the interval [1/8,1/4,1/2,1,2,4,8] for DIARETDB1 dataset, and sensitivities of 0.185, 0.313, 0.465, 0.604, 0.716, 0.801, 0.849, and FROC score of 0.562 ± 0.233 for FPI values in the interval [1/8,1/4,1/2,1,2,4,8] for E-Ophtha dataset. It is observed that the proposed model has extracted MAs for early detection of DR but it may not efficiently detect and distinguish between resembling features of MAs such as HEs, when both occurs together as red lesions.

4.3 Feature identification and classification using evolutionary algorithms

Mookiah et al. [96] have proposed a system using 156 preprocessed fundus images for extraction and classification of abnormal signs of DR such as EXs, through segmentation of RBVs using 2D Gabor matched filter, and texture extraction using Local Binary Pattern (LBP) and Laws Texture Energy, for the detection of DR. The model has used various preprocessing techniques such as gray level shading correction, contrast enhancement and image restoration. The model has proposed an OD segmentation method based on fuzzy set theoretic model A-IFS histon and region merging algorithm. The model has extracted twenty-five features of which thirteen features are extracted and fed to each Probabilistic Neural Network (PNN), Decision Tree (DT) C4.5, and SVM. The best classifier i.e., PNN is determined based on the smoothing parameter σ, which is identified using Genetic Algorithm (GA) and Particle Swarm Optimization (PSO). The model has used 104 images for training and 52 images for testing and has achieved a sensitivity of 96.27%, specificity of 96.08%, PPV of 98.20%, and an accuracy of 96.15% in PNN for s = 0.0104 determined using One-way ANOVA statistical tests, whereas DT C4.5 has achieved a sensitivity, specificity and PPV of 100% each and an

accuracy of 88.46%. The RBF SVM has achieved a sensitivity of 86.19%, a specificity of 79.60%, a PPV of 90.92%, and an accuracy of 66.02%. It is observed that very minute and intricate features can be extracted from fundus images through efficient segmentation of RBVs, and to localize and understand the behavior of DR features.

Gadekallu et al. [43] have proposed a 5-layered DNN model which performs normalization, a three-layered pre-processing for enhancement of images, PCA-based feature extraction, dimensionality reduction using Firefly algorithm and feature engineering, and DR classification for early detection of DR. The model has collected the images from DR Debrecen dataset from the UCI ML repository having 1151 instances, and 20 attributes which are features extracted from the MESSIDOR dataset, and has compared its performance to various traditional and hybrid ML models. The *DNN model* has used Adam optimizer and softsign activation function at each layer, and sigmoid function at the output layer for classification. The proposed model has adopted dimensionality reduction and feature engineering and has achieved an accuracy of 97%, precision of 96%, recall of 96%, sensitivity of 92% and specificity of 95%. It is observed that the inculcation of evolutionary algorithms with the concept of deep neural models can help in modelling of efficient expert systems.

4.4 Evaluating comprehensive CNNs and DNNs for DR detection

Sarki et al. [124] have proposed a comprehensive evaluation of 13 CNN architectures, pre-trained and fine-tuned on comprehensive ImageNet database using TL, upon MESSIDOR and Kaggle dataset, for early detection of mild stages of DR. The model has achieved 90% accuracy on classification of severe cases and 86% accuracy on mild/no DR cases, using 35,126 fundus images.

Hattiya et al. [54] have concluded AlexNet as the most appropriate CNN architecture for DR detection application when compared with different CNN architectures namely AlexNet, ResNet50, DenseNet201, InceptionV3, MobileNet, MnasNet and NASNetMobile, implemented upon 23,513 retina images, in a comprehensive process of evaluation. The model has achieved better results with AlexNet, achieving accuracy values of 98.42% and 81.32% for training and testing sets, respectively.

Bodapati et al. [21] have proposed a DR classification model upon Kaggle APTOS 2019 contest dataset using TL and deep feature aggregation from multiple convolution blocks of pretrained models such as NASNet, Xception, Inception ResNetV2 and VGG-16 to enhance feature representation, and has thereby established a comparison with handcrafted features, for assessment of DR severity. The model has compared various pooling and feature fusion strategies and have concluded that averaging pooling with simple fusion approaches upon Deep Neural Networks (DNN) performs better. The model has achieved an accuracy of 84.31% and an AUC 97.

Kamal et al. [65] have proposed a TL model for recognition of *COVID-19* using 760 Chest X-Ray (CXR) images and increased trainable parameters. The model has comprehensively evaluated CNNs such as VGG-19, InceptionV3, ResNet50, ResNet50V2, MobileNet, MobileNetV2, DenseNet121 and NasNetMobile. The fine-tuned DenseNet121 model performs better and has achieved 98.69% test accuracy and highest macro f1-score of 0.99, on CXR-B dataset (80:20 split).

Lee et al. [80] have proposed a TL NASNet-A (large) architecture to extract bottleneck features from 307 Spectral-Domain Optical Coherence Tomography (SD-OCT) image sets, and has deployed an ensemble training model for DR prediction. The proposed model

highlights the importance of CNN for OCT evaluation and how DNN's (NASNet) flexibility to evaluate and assess is greatly dependent upon the nature of imaging modality adopted. The proposed model has achieved an AUC of 0.990 with a sensitivity of 94.7% and a specificity of 100.0%.

Huang et al. [56] have proposed a pre-trained CondenseNet which combines dense connectivity with novel learned *group convolutions*, for better feature re-use, upon CIFAR-10 (C-10), CIFAR-100 (C-100) and ImageNet datasets, for image classification. The CNN architecture is cost-efficient compared to models such as DenseNet-190, MobileNets and ShuffleNets. The proposed model CondenseNet^{light}-160 (pruned) has achieved an error rate of 3.46% in C-10 and 17.55% in C-100.

Ji et al. [60] have proposed an optimized DNN model through removal of deep convolutional layers from pre-trained networks such as Inception V3, ResNet50 and DenseNet121, for DR image analysis. The model has proposed various subnetworks for each of the DNN and analysed their performance on large OCT image datasets of 83,484 images, for detection of DR lesions. The C5_b4 sub-network of DenseNet121 has achieved the highest accuracy of 99.80% amongst all others.

Samanta et al. [122] have proposed a CNN-TL DenseNet121 model, upon *3050 training images* which are fine-tuned and tested upon architectures such as Inception V1, Inception V2, Inception V3, Xception, VGG16, ResNet-50, DenseNet and AlexNet, for DR detection. The model has achieved a validation accuracy of 84.10%.

Tymchenko et al. [134] have proposed a DCNN encoder-based feature extraction for DR detection using pre-trained EfficientNet-B4, EfficientNet-B5, SE-ResNeXt50 and ensemble of 20 models (4 architectures × 5 folds). The proposed model has achieved a sensitivity and specificity of 0.99 and quadratic weighted kappa score of 0.9254 on APTOS 2019 Blindness Detection Dataset consisting of 13,000 images.

5 Challenges and predictable solutions

During fundus image analysis, various challenges are required to overcome while detecting and extracting DR features due to poor quality of images, close localization of features such as MAs towards blood vessels, presence of red lesions [24, 130], etc. which results in lower correlation coefficient. DR features such as EXs may be often confused with OD for they are bright intensity structures with non-homogeneous intensity characteristics [146], and because of presence of noise, low contrast, uneven illumination, and color variation, prominent detection might be difficult. It is also important to comprehend and perceive transformations amongst features and those having same organization or behavior such as large HEs and RBVs. Again, segmentation of retinal vessels is crucial because of the noise caused due to uneven illumination and variations in structural behavior [92]. Thus, background subtraction for elimination of anatomical structures of the retina, is required for better detection of DR lesions. However, subtraction of OD and RBVs can lead to the ignorance of abnormal behavior of OD, and NVs and IRMAs in RBVs. In such situations, two kinds of methodologies may arise as solutions - image analysis and background subtraction. This indirectly raises computational cost, and highlights the importance of preprocessing techniques.

During model training, challenges such as system specifications, operations, large-scale implementations, trade-off between accuracy and efficiency, may be encountered. Besides, incorrect and improper inferences drawn from irresponsible review and flawed facts,

prejudgments and false amplification, may also cause a biased model. Such models may also be prone to data drift [103] i.e., the relationship between input data and output target variable changes w.r.t time, thus exhibiting an unpredictable behavior, due to incremental, gradual, and reoccurring behaviors. As a remedy, such kind of problems require authentic information, well interpretation, eradication of flawed facts, for informed decision making. Additionally, automated strategies, periodic model retraining, and maintenance, elimination of data drift etc. can be of paramount help. On technical grounds, data preparation, better feature extraction, classification, compact representation of high-dimensional features, proper tuning of hyperparameters and optimization strategies can be of immense help for better and increased performance of a model.

Big data is another important challenge and DL models have effectively dealt with its volume, velocity, variety, and veracity. ML models are overwhelmed because of huge data which affects their training performance. The performance of DL model increases with increase in data, unlike ML models. However, big datasets are often imbalanced, and can degrade the performance of an efficient model. Therefore, various strategies are adopted for creation and use of a balanced dataset such as undersampling and oversampling techniques. When undersampling is adopted, data is tremendously reduced causing the formation of smaller datasets. The use of such a dataset can cause underfitting of a model. On the other hand, oversampling of a dataset may cause overfitting of the model, thus producing biased results. To overcome data imbalance, proper sampling, image/data preprocessing and data augmentation can be performed for generation of a balanced dataset. Besides, to overcome data privacy and scarcity, different repositories are available which constitutes fundus images such as the Kaggle dataset, MESSIDOR (Methods to Evaluate Segmentation and Indexing Techniques in the field of Retinal Ophthalmology) dataset, APTOS (Asia Pacific Tele-Ophthalmology Society) 2019 dataset etc. These datasets contain raw fundus images which requires efficient preprocessing for the purpose of model training. Digital equipment such as Digital Fundus Camera (DFC) can be used to capture fundus images and create image dataset(s). Besides, fundus images can also be collected from Ophthalmologists or from medical institutions and organizations. TL pre-trained models are adopted to overcome problems associated with smaller datasets. However, challenges such as adversarial attacks and biases may occur due to presence of artefacts, and which can be reduced using image preprocessing or data preprocessing techniques. Generative Adversarial Networks (GANs) can also be adopted for data generation and discrimination but it is mostly suitable in an unsupervised learning environment. All these techniques can be used to mitigate data imbalance, data acquisition, privacy concern and noisy dataset for efficient feature extraction and corresponding image classification.

The computational cost of DL networks, large ensemble networks, Neural Networks (NNs), hybrid ML-DL techniques, natural optimization techniques, evolutionary algorithms, feature concatenation [114] etc. upon image data is another important challenge. It depends upon the size of the network architecture and the number of parameters included which indicates the required processing speed and the need of a highly computational memory such as Graphical Processing Unit (GPU). *Such computational needs and resources can be requested and accessed from authorized technical institutions.*

Various challenges are also encountered when predictions obtained and derived are dependent on a single perspective i.e., on a single classifier system. Classifiers such as Neural Network (NN) undergoes a process of stochastic training algorithm. This means the network exhibits a variation in mapping and changes in weights in every iteration of the training

process, thereby producing different predictions each time, and hence exhibits an unreliable form of single classifier-based classification. This is referred to as NN executing a low bias and a high variance, in making predictions, which when incorporated with DL models, triggers its sensitivity towards weights and noise. An effective method to drop the variance of NN-based models is to train multiple ML models for classification and integrate the predictions of the ensemble to obtain an average prediction through majority voting. This leads to the consideration of various perspectives and insights for decision making with lower and better error rate and different prediction loss for establishing the final decision.

Again, DL models and shallow ML models such as SVM, RF exhibits a black-box nature unlike traditional ML algorithms which provides an informed explanation. Hence DL models are required to be explored, exploited and explained such that end users can have access to predictions, and behaviours of such systems can be understood, for better generalization and prediction. Therefore, eXplainable AI(XAI) [123] has been introduced to provide an explanation to complex, non-linear, accountable, and augmented behaviors of DL models and shallow ML models, to make them reliable, adoptable and predictable.

Thus, various challenges can occur while adopting optimal DL methodologies for DR detection. These challenges create various opportunities and highlights the importance and need of critical solutions.

6 Comparison and analysis of DR detection methods

This text highlights the comparison between different methods deployed for DR diagnosis and detection, in earlier proposed models and identifies the loopholes and gaps present in the study. The classical procedures are helpful but time consuming and inflexible, during detection. The advanced DL techniques are black-box, and may be time consuming based on data, but are highly automated, flexible and adaptable. ML techniques are manually-staged automated procedures and are better at providing an explanation than DL. On reviewing different ML and more specifically DL based techniques, for feature extraction it is found that DL incorporates representation learning to learn new patterns and features from existing and extracted features, which enables it towards better and generalized predictions. It is observed that in standard DL models, dense classification is carried out using fully connected networks which increases the number of parameters of a highly-parameterized deep architecture. This increases computational cost and model training time. To eliminate/reduce the number of parameters, dense layers of standard DL models are replaced through efficient ML classifiers and learning algorithms such as SVM, Neural Network (NN), Decision Trees etc. However, such single-classifier models have limited perception and knowledge. Thus, for further improvements in the detection process such as better prediction and correct error rate, and to reduce high bias and variance, homogeneous and heterogeneous ensemble learning and classification may be adopted for synchronization, complexity, dynamic learning, stability and better predictability, than a single classifier model.

On comparing the earlier proposed literatures and experimental models, various insights from the study are identified. ML fusion strategies [22] for classification are less efficient than boosting weak classifiers in ensembles [8]. The handcrafted ML algorithms for feature extraction are time consuming and prone to error than DL based features extractors. However, it is claimed that combination of handcrafted and CNN features [101] is the best of feature sets, it is still an inefficient approach. Again, heterogenous CNN models have performed better than

a single CNN model [127]. The dimensionality reduction techniques or compact representation techniques have the capability to effectively boost the performance of DL models, compared to ML models [89]. An ensemble of weak classifiers [8] may not be able to produce generalized predictions and requires discriminative procedures for feature learning. Image concatenation-based training of model [132] can be more efficient for reading large data and extracting more features than a single-image trained model. The use of visualization techniques [112] and image assessment modules [117] while employing DL models is completely rare or may be insignificant, however it can boost the performance of the model for detection of subtle lesions, where CNNs have often shown poor results. DRL based deep architectures [45] are capable of identifying new features with augmentation and visualization techniques. The consideration of performance metrics in DL models has an effective role to play. For e.g., accuracy is often not a suitable term to signify performance as with larger dataset it can cause overfitting, poor generalization and misclassification. Metrics such as False Positive per Image (FPI) and Free-response Receiver Operating Characteristic (FROC) curve or F1-score are more effective [36]. Evolutionary deep intelligent models may be more effective than standard or modified DL models [43]. Patch-based ML techniques [26] are conventional compared to image-based DL techniques for intermediate DR phase detection and lesion identification. DL pooling and feature fusion strategies [21] provides flexibility to models irrespective of imaging modality. However, imaging modality plays a vital role in performance.

This section of the paper discusses and analyzes the performance of some classical as well as contemporary methods using various ML and DL models such as DCNN and RF ensemble classification, for DR detection. In the process of implementation, Kaggle training dataset which consists of 35,126 fundus images, are used for DR detection. There are 5-classes of DR namely grade-0, grade-1, grade-2, grade-3 and grade-4 and each implies no DR, mild NPDR, moderate NPDR, severe NPDR and PDR, respectively. The total number of images already identified and annotated, corresponding to each of these classes is mentioned below in Table 1. The DR fundus image Kaggle dataset is huge and its distribution throughout is highly imbalanced.

Based on the implementation performed, two models are proposed namely Deep Diabetic Retinopathy Detection System (DDRDS) and Deep Diabetic Retinopathy Feature eXtraction and RF based ensemble Classification System (DDRFXRFCS), for DR detection. DDRDS performs a comprehensive evaluation and a comparative analysis of DCNN models namely VGG-16, InceptionV3, MobileNet V1 and Xception, upon 35,126 fundus images from Kaggle dataset. The DCNN performs inbuilt preprocessing upon the fundus images followed by each individual DCNN performing respective feature extraction and classification tasks on the basis of DR grades, for DR detection. The DDRDS is trained on 27,446 images, validated on 7430 images and tested on 250 images, each containing the five classes of DR, for 50 epochs. The

Table 1 Distribution in Kaggle Training dataset

Type of DR	DR Grade	Total No. of images
No DR	0	25,810
Mild NPDR	1	2443
Moderate NPDR	2	5292
Severe NPDR	3	873
PDR	4	708

entire implementation is carried on 64-bit Windows 10 platform, with python version 3.8 and TensorFlow version 2.4. Block Diagram V describes the working process of DDRDS and DDRFXRFCS, for detection of DR.

The different architectures of DCNN has achieved contrasting results. VGG-16 has achieved a training accuracy of 77.52% and validation accuracy of 75.67%. Inception V3 has achieved a training accuracy of 99.29% and validation accuracy of 81.82%. The MobileNet V1 has achieved a training accuracy of 98.90% and validation accuracy of 76.55%. Xception has achieved a training accuracy of 99.46% and validation accuracy of 75.22%. Table 2 depicts performances of different DCCN models in DDRDS.

It is observed that VGG16 is a better learning model with less overfitting and less of generalization error compared to Inception V3, MobileNet V1 and Xception architecture, which have comparatively shown higher generalization error and overfitting upon the highly skewed dataset. However, Inception V3 has achieved better results in DR detection compared to rest of the DCNN models. The proposed DDRDS differs from various other comprehensive evaluation methods in terms of dataset, architecture, number and variant of DCNN model(s) used and hyperparameter tuning.

In DDRFXRFCS, DCNN based feature extraction and RF-based ensemble classification is performed using VGG-16, Inception V3 and Xception, for DR detection. It is also implemented using 35,126 fundus images. The DCNNs of the proposed model performs individual inbuilt image preprocessing for selection of valuable features, to train the RF ensemble classifier, in a mutually exclusive manner. Block Diagram IV also depicts the working model of DDRFXRFCS. It is trained on 28,079 images and tested on 7047 images, each containing the five classes of DR. The RF in the proposed model, used for classification is an ensemble of 50 decision trees, and is trained for 50 epochs. The entire implementation is performed on 64-bit Windows 10 platform, with python version 3.8 and TensorFlow version 2.4. On
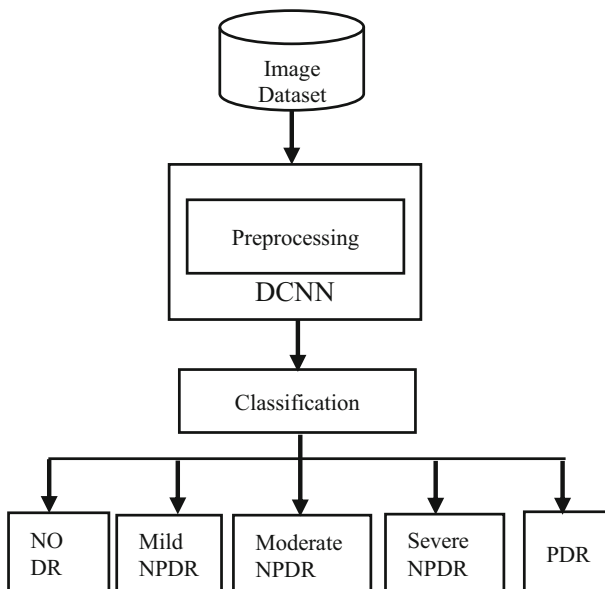

Diagram V Working process of DDRDS and DDRFXRFCS

Table 2 Performance of DCCN models in DDRDS

DCNN Model	Dataset	Training Accuracy (%)	Validation Accuracy (%)
VGG16	35,126	77.52	75.67
Inception V3	35,126	99.29	81.82
MobileNet V1	35,126	98.90	76.55
Xception	35,126	99.46	75.22

incorporating VGG-16 based deep feature extraction and RF based classification (VGG-16 RF), the proposed model has achieved an accuracy of 73.19%. On incorporating Inception V3 based deep feature extraction and RF based classification (InceptionV3-RF), the proposed model has achieved an accuracy of 73.09%. Again, on incorporating Xception based deep feature extraction and RF based classification (Xception RF), the proposed model has achieved an accuracy of 73.12%. From the above results, it is certain that VGG-16 RF has again performed better than Xception RF, followed by InceptionV3-RF. In both the proposed models of our work, it is certain that VGG-16 is a better learning model on its own as well as on incorporation with ensembles, the CNN excels with better and generalized prediction accuracy. Table 3 depicts performances of different DCCN models in DDRFXRFCS.

The proposed DCNN-ensemble model on using a comparatively larger dataset, has performed better in comparison to TL NASNet-A (large) [17] and cascaded CNN Adaboost ensemble model [8], in terms of better generalization and validation accuracy. The VGG-16 in proposed models has also achieved better and generalized predictions upon the dataset in comparison to VGG-16 [34].

Table 4 presents the literature review in a tabular form explaining in brief the various recent works performed upon DR detection using varied datasets, CNNs and classifiers and their corresponding outcomes, thereby establishing a comparison.

7 Future directions

In future, we propose to work on a high-dimensional balanced dataset using supervised learning methods and approach with the best of its kind DL model for deep feature extraction, DR lesion identification, lesion-based image classification and DR detection. We aim to work with ensemble of ML classifiers instead of dense classifiers of DL models for better error detection, multiple prediction to obtain the final average prediction thereby avoiding ambiguity and misclassification. We also aim to work with better processing of data with contemporary techniques for error retrieval and removal in DR images, for effective feature extraction and

Table 3 Performance of DCCN models in DDRDS

DCNN	Dataset	Classifier	Accuracy (%)
VGG-16	35,126	RF	73.19
Inception V3	35,126	RF	73.09
Xception	35,126	RF	73.12

Table 4 Literature Review on ML and DL models for early detection of DR

Paper Name	Images/Dataset	Features	Methodology	Classifier	Results
Lam et al. [77]	243	MAs, HEs, EXs, NV	CNN	CNN	AlexNet accuracy 74% and 79% VGG16 accuracy 86% and 90% GoogLeNet accuracy 95% and 98% ResNet accuracy 92% and 95% Inception-v3 accuracy 96% and 98%
Buades et al. [22]	7137	EXs, HEs, drusen, CWSs	SURF, OR, MV and meta-SVM	SVM	AUC 91.6% for single lesion detection (hard EXs) AUC of 88.3%multi-lesion detection
Mansour et al. [89]	Kaggle dataset		AlexNet, SIFT, LDA and PCA	SVM	Accuracy of 90.15% (PCA) and 97.23% on FC6 features (LDA) Accuracy of 95.26% (PCA) and 97.28% on FC7 features (LDA)
Orlando et al. [101]	DIARETDB1, MESSIDOR, e-ophtha	MAs, HEs	LeNet	CNN	AUC of 0.7912 with CNN features AUC of 0.7325 with hand crafted features AUC of 0.8932 on combination of both features
Pratt et al. [108]	80,000	MAs, EXs and HEs	CNN architecture		Sensitivity 30%, Specificity 95%, Accuracy 75%
Xu et al. [142]	Kaggle	EXs, red lesions, MAs RBVs	CNN	GBM and CNN	91.5% accuracy without data augmentation 94.5% accuracy with data augmentation
Khojasteh et al. [71]	DIARETDB1 and e-Ophtha	EXs, MAs, HEs	CNN	CNN	Accuracy of 0.96 (EXs), 0.98(HEs) and 0.97 (MAs) on DIARETDB1 Accuracy of 0.88(EXs), and 3.0(MAs), on e-Ophtha dataset
Soniya et al. [127]	DIARETDB0	MA, HE, hard EX, soft EX, NVE	Single CNN and heterogenous CNN	Multilayer perceptron network	Single CNN accuracy: 40%–90% Heterogenous CNN accuracy: 100%
Alghamdi et al. [8]	PAMDI and HAPIEE	OD	Cascaded CNNs	AdaBoost ensemble algorithm	Sensitivity 96.42%, Specificity 86%, Accuracy 86.52% on HAPIEE Sensitivity 94.54%, Specificity 98.59%, Accuracy 97.76% on PAMDI

Table 4 (continued)

Paper Name	Images/Dataset	Features	Methodology	Classifier	Results
Gardner et al. [44]	EyePACS-1 MESSIDO-R-2	HE and MA	Inception-V3--architecture Neural Network	Ensemble of 10 networks	EyePACS-1: AUC of 0.991 MESSIDOR-2: AUC of 0.990
Abràmoff et al. [86]	1748	OD, Fovea, HEs, EXs NV	AlexNet, VGGNet	RFC	Sensitivity of 96.8% (rDR), Specificity of 87%, AUC 0.98 (rDR) Sensitivity of 100% (vtDR), Specificity of 90.8%, AUC 0.989 (vtDR)
Lam et al. [78]	Kaggle MESSIDO-R-1		Pretrained AlexNet and GoogLeNet	GoogLeNet 2-ary, 3-ary and 4-ary	2-ary accuracy 74.5% 3-ary accuracy 68.75% 4-ary accuracy 51.25%
Takahashi et al. [132]	9939 images	HE and hard EX	GoogLeNet DCNN and ResNet		Accuracy 96% (real prognosis) Accuracy 92% (Davis grading)
Quellec et al. [112]	MESSIDOR Kaggle e-ophtha DiaretDB1	hard EXs, soft EXs, small red dots, HEs, lesions	ConvNet netB	Ensemble classifier	A_z of 0.954 in Kaggle's dataset A_z of 0.949 in e-Ophta dataset A_z of 0.9490 using ensemble classifier
Alban et al. [6]	35,126		Pre-trained AlexNet	GoogLeNet	AUC 0.79 (GoogLeNet) AUC 0.69 (AlexNet)
Kermany et al. [70]	207,130 OCT	CNV, DME Drusen	Pretrained Inception V3		Binary classification accuracy: >98% Multi-class classification accuracy of 96.6%
Dutta et al. [34]	Kaggle dataset	RBVs, fluid drip, EXs, HEs, MAs	NN, DNN, VGG-16	VGG-16	Accuracy: 72.5% for 300 test images Accuracy: 78.3% for 600 test images
Grinsven et al. [136]	Kaggle MESSIDOR		VGGNet		area under ROC of 0.894 on Kaggle dataset area under ROC of 0.972 MESSIDOR dataset
Liskowski et al. [84]	DRIVE STARE CHASE		DNN		area under ROC >99% Accuracy >97%
Islam et al. [57]	EyePACS	MAs	DCNN-18, PLAIN BALANCE, NO-POOL		area under ROC 0.844 F1-Score 0.743
Prentasic et al. [109]	DRiDB	OD, RBVs, EXs	DCNN		Sensitivity 78%, Positive Predictive Value (PPV) 78%, F-score 78%
Mookiah et al. [96]	156	RBVs, EX, OD, NVE	PNN	SVM	Sensitivity 96.27%, Specificity 96.08%, PPV 98.2%, Accuracy 96.15%
	Kaggle	drusen,	DCNN	VGGNet	

Table 4 (continued)

Paper Name	Images/Dataset	Features	Methodology	Classifier	Results
Rakhlin [117]	MESSIDOR-2	EXs, MAs, CWSs, HEs			area under ROC 0.923 in Kaggle dataset area under ROC 0.967 in MESSIDOR-2 dataset
Gargeya et al. [45]	EyePACS MESSIDOR-2 E-Ophtha	HEs, hard EXs and NVE	DRL CNN		AUC 0.94 on MESSIDOR-2 AUC 0.95 on E-Ophtha AUC 0.97 on EyePACS
Eftekhari et al. [36]	ROC E-Ophtha-MA	MAs	DLNN		FROC (F_{AUC}) of 0.660 for ROC FROC (F_{AUC}) of 0.637 for E-Optha-MA dataset
Wang et al. [138]	35,126		CNN Net-5 and Net-4		Kappa score 0.70 for 256-pixel images, Kappa score 0.80 for 512-pixel images Kappa score 0.81 for 768-pixel images
Dai et al. [27]	735	MAs	MS-CNN		Recall 87.8%, precision 99.7%, accuracy 96.1%, and F1 score of 93.4%
Al-Bander et al. [7]	MESSIDOR Kaggle	Fovea and OD	multi sequential DL technique		1R criterion: Accuracy 97% (OD) and 96.6% (foveal) in MESSIDOR test set Accuracy 96.7%(OD) and 95.6% (foveal) in Kaggle test set
Gadekallu et al. [43]	DR Debrecen		PCA, Firefly model	DNN	Accuracy of 97%, Precision of 96%, Recall of 96%, Sensitivity of 92%, Specificity of 95%.
Chudzik et al. [26]	E-Ophtha, ROC DIARETD-B1	MAs	FCNN		FROC score 0.193± 0.116 on ROC FROC score 0.392± 0.157 on DIARETDB1 FROC score 0.562± 0.233 on E-Ophtha

classification. In the future, the proposed model also plans on extending the work through introduction of feature concatenation using ensemble of DCNNs and corresponding feature descriptors and compact representation techniques. DL models are overly complex to cause overfitting and hence advanced DL algorithms with better generalization and error detection will be sought thereby concerning trade-off relationships between accuracy, computational complexity, memory constraints and processing power to increase portability, availability, and flexibility. Thus, the proposed methodology would continue in optimizing the in-depth algorithms to obtain a reliable performance for real-life applications for early DR detection.

8 Conclusion

DR is a critical medical health disorder causing blindness which is of utmost concern, and DL techniques can have an effective role in its diagnosis and early detection than traditional techniques. This paper precisely describes DR, its symptoms, features, shape, size and location of the features, and how DR causes blindness. It also describes various ML and DL techniques used for the detection of abnormal behavior of RBVs and OD to identify DR lesions such as MAs, HEs, EXs, CWS, FAZ, IRMA, Neovascularization in a chronological order. To avoid hindrance of subtle lesions and prevent misclassification, identification of a general or a strategically specialized framework is realized. Various methodologies are studied and reviewed for early detection of DR. It can be understood and realized that ML techniques are highly unscalable w.r.t high-dimensional data and takes more time in analysis and training of model in comparison to DL techniques. As the number of feature and data increases, ML models conclude with sub-optimal solutions whereas DL models strives to obtain the optimal output. Based on the immense applications of DL in recently proposed models, this paper identifies and reviews a significant number of DL models and their frameworks, for understanding of the working principle, their evolution and integration on using hybrid techniques and how such models can be transitioned on scarcity of data and resources, to produce effective models and outcomes. Besides, the text highlights some critically interesting challenges such as balanced data acquisition, data preprocessing, deep feature extraction, DL model black-box explanation, generalization upon unseen data, overfitting and underfitting of the model, constraints of single-system based classification, presence of adversarial attacks and bias in datasets and in pretrained models used for TL, dimensionality reduction, data drift, vanishing gradient problem, etc. along with suitable solutions, comparative studies, future works and directions. Thus, this paper will be helpful for aspiring, young and engaged researchers interested in the domain of DR, medical imaging and DL, and making it more approachable towards new ideas, innovations and technology.

Declarations

References

1. Abed S, Al-Roomi SA, Al-Shayeji M (2016) Effective optic disc detection method based on swarm intelligence techniques and novel pre-processing steps. Appl Soft Comput 49:146–163. https://doi.org/10.1016/j.asoc.2016.08.015
2. Abràmoff MD, Lou Y, Erginay A, Clarida W, Amelon R, Folk JC, Niemeijer M (2016) Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. Invest Ophthalmol Vis Sci 57(13):5200–5206. https://doi.org/10.1167/iovs.16-19964
3. Agurto C, Murray V, Yu H, Wigdahl J, Pattichis M, Nemeth S, Barriga ES, Soliz P (2014) A multiscale optimization approach to detect exudates in the macula. IEEE J Biomed Health Inf 18(4):1328–1336. https://doi.org/10.1109/JBHI.2013.2296399
4. Akram MU, Khalid S, Khan SA (2013) Identification and classification of microaneurysms for early detection of diabetic retinopathy. Pattern Recogn 46(1):107–116. https://doi.org/10.1016/j.patcog.2012.07.002
5. Akram MU, Khalid S, Tariq A, Khan SA, Azam F (2014) Detection and classification of retinal lesions for grading of diabetic retinopathy. Comput Biol Med 45:161–171. https://doi.org/10.1016/j.compbiomed.2013.11.014

6.	Alban M, Gilligan T (2016) Automated detection of diabetic retinopathy using fluorescein angiography photographs. Report of Standford Education.

7.	Al-Bander B, Al-Nuaimy W, Williams BM, Zheng Y (2017) Multiscale sequential convolutional neural networks for simultaneous detection of fovea and optic disc. Biomed Signal Process Control 40:91–101. https://doi.org/10.1016/j.bspc.2017.09.008

8.	Alghamdi HS, Tang HL, Waheeb SA, Peto T (2016) Automatic optic disc abnormality detection in fundus images: a DL approach. Proceedings of the ophthalmic medical image analysis international workshop. Pp. 17-24. https://doi.org/10.17077/omia.1042

9.	Alom MZ, Taha TM, Yakopcic C, Westberg S, Sidike P, Nasrin MS, Hasan M, Van Essen BC, Awwal AAS, Asari VK (2019) A State-of-the-Art Survey on Deep Learning Theory and Architectures, electronics, 8(3):1–66. https://doi.org/10.3390/electronics8030292

10.	Amin J, Sharif M, Yasmin M, Ali H, Fernandes SL (2017) A method for the detection and classification of diabetic retinopathy using structural predictors of bright lesions. J Comput Sci 19:153–164. https://doi.org/10.1016/j.jocs.2017.01.002

11.	Andrew D, MD, PhD, University of Iowa, Retina, Diabetic Retinopathy, EyeRounds.org Available at - https://webeye.ophth.uiowa.edu/eyeforum/atlas/photos/DR/DR.jpg, Accessed on – 13-10-2020

12.	Annunziata R, Garzelli A, Ballerini L, Mecocci A, Trucco E (2015) Leveraging multiscale hessian-based enhancement with a novel exudate Inpainting technique for retinal vessel segmentation. IEEE J Biomed Health Inf 20(4):1129–1138.https://doi.org/10.1109/JBHI.2015.2440091

13.	Antal B, Hajdu A (2012) An ensemble-based system for microaneurysm detection and diabetic retinopathy grading. IEEE Trans Biomed Eng 59(6):1720–1726. https://doi.org/10.1109/TBME.2012.2193126

14.	Antal B, Lázár I, Hajdu A (2012) An adaptive weighting approach for ensemble-based detection of microaneurysms in color fundus images. 2012 annual international conference of the IEEE engineering in medicine and biology society. Pp. 5955-5958. https://doi.org/10.1109/EMBC.2012.6347350

15.	Aquino A (2014) Establishing the macular grading grid by means of fovea Centre detection using anatomical-based and visual-based features. Comput Biol Med 55(1):61–73. https://doi.org/10.1016/j.compbiomed.2014.10.007

16.	Argade KS, Deshmukh KA, Narkhede MM, Sonawane NN, Jore S (2015) Automatic detection of diabetic retinopathy using image processing and data mining techniques. Proceedings of the 2015 international conference on Green computing and internet of things (ICGCIoT'15). 517-521. https://doi.org/10.1109/ICGCIoT.2015.7380519

17.	Asha PR, Karpagavalli S (2015) Diabetic retinal exudates detection using extreme learning machine. Emerging ICT for bridging the future-proceedings of the 49th annual convention of the Computer Society of India CSI. 2:573-578. https://doi.org/10.1109/ICACCS.2015.7324057

18.	Aslam T, Chua P, Richardson M, Patel P, Musadiq M (2009) A system for computerised retinal haemorrhage analysis. BMC Res Notes 2(1):1–6. https://doi.org/10.1186/1756-0500-2-196

19.	Banerjee S, Kayal D (2016) Detection of hard exudates using mean shift and normalized cut method. Biocybern Biomed Eng 36(4):679–685. https://doi.org/10.1016/j.bbe.2016.07.001

20.	Bhargavi VR, Senapati RK (2016) Bright lesion detection in color fundus images based on texture features. Bull Electric Eng Inf 5(1):92–100. https://doi.org/10.11591/eei.v5i1.553

21.	Bodapati JD, Shaik NS, Naralasetti V (2021) Deep convolution feature aggregation: an application to diabetic retinopathy severity level prediction. SIViP 15:923–930. https://doi.org/10.1007/s11760-020-01816-y

22.	Buades A, Coll B, Morel JM (2005) A non-local algorithm for image denoising. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition 2: 60–65. https://doi.org/10.1109/CVPR.2005.38

23.	Chaturvedi SS, Gupta K, Ninawe V, Prasad PS (2020) Automated diabetic retinopathy grading using deep convolutional neural network. arXiv:2004.06334:1-12

24.	Cheng X, Wong DWK, Liu J, Lee BH, Tan NM, Zhang J, Cheng CY, Cheung G, Wong TY (2012) Automatic localization of retinal landmarks. Ann Int Conf IEEE Eng Med Biol Soc 2012:4954–4957. https://doi.org/10.1109/EMBC.2012.6347104

25.	Chetoui M, Akhloufi MA (2020) Explainable diabetic retinopathy using EfficientNET. In 2020 42nd annual international conference of the IEEE engineering in Medicine & Biology Society (EMBC).1966-1969. https://doi.org/10.1109/EMBC44109.2020.9175664

26.	Chudzik P, Majumdar S, Caliváa F, Al-Diri B, Hunter A (2018) Microaneurysm detection using fully convolutional neural networks. Comput Methods Prog Biomed 158:185–192. https://doi.org/10.1016/j.cmpb.2018.02.016

27.	Dai L, Fang R, Li H, Hou X, Sheng B, Wu Q, Jia W (2018) Clinical report guided retinal microaneurysm detection with multi-sieving deep learning. IEEE Trans Med Imaging 37(5):1149–1161. https://doi.org/10.1109/TMI.2018.2794988

28. Decencière E, Cazuguel G, Zhang X, Thibault G, Klein JC, Meyer F, Marcotegui B, Quellec G, Lamard M, Danno R, Elie D, Massin P, Viktor Z, Erginay A, Lay B, Chabouis A (2013) TeleOphta: machine learning and image processing methods for teleophthalmology. IRBM. 34(2):196–203. https://doi.org/10.1016/j.irbm.2013.01.010

29. Diabetic Retinopathy – Features of Diabetes: Cotton Wool Spots, Glycosmedia, Diabetes News Service, Available at https://www.glycosmedia.com/education/diabetic-retinopathy/diabetic-retinopathy-features-of-diabetes-cotton-wool-spots, Accessed on 12-06-2020

30. Diabetic Retinopathy – Features of Diabetes : Intraretinal Haemorrhages, Glycosmedia, Diabetes News Service, Available at https://www.glycosmedia.com/education/diabetic-retinopathy/diabetic-retinopathy-features-of-diabetes-intraretinal-haemorrhages, Accessed on 12-06-2020

31. Diabetic retinopathy, American Optometric Association, https://www.aoa.org/patients-and-public/eye-and-vision-problems/glossary-of-eye-and-vision-conditions/diabetic-retinopathy-Over-time-diabetes-damages-small-condition-usually-affects-both-eyes, Accessed on 13-05-2020.

32. Diabetic Retinopathy Detection, Kaggle repository, Available at: https://www.kaggle.com/c/diabetic-retinopathy-detection/data, Accessed on 14-06-2021

33. Duanggate C, Uyyanonvara B, Makhanov SS, Barman S, Williamson T (2011) Parameter-free optic disc detection. Comput Med Imaging Graph 35(1):51–63. https://doi.org/10.1016/j.compmedimag.2010.09.004

34. Dutta S, Manideep BCS, Basha SM, Caytiles RD, Iyengar NCSN (2018) Classification of diabetic retinopathy images by using deep learning models. Int J Grid Distrib Comput 11(1):89–106. https://doi.org/10.14257/ijgdc.2018.11.1.09

35. Early Treatment Diabetic Retinopathy Study Design and Baseline Patient Characteristics: ETDRS Report Number 7 (1991) Early treatment diabetic retinopathy study research group. Ophthalmology. 98(5):741–756. https://doi.org/10.1016/S0161-6420(13)38009-9

36. Eftekhari N, Pourreza HR, Masoudi M, Ghiasi-Shirazi K, Saeedi E (2019) Microaneurysm detection in fundus images using a two-step convolutional neural network. Biomed Eng Online 18(1):1–16. https://doi.org/10.1186/s12938-019-0675-9

37. Ege BM, Hejlesen OK, Larsen OV, Møller K, Jennings B, Kerr D, Cavan DA (2000) Screening for diabetic retinopathy using computer based image analysis and statistical classification. Comput Methods Prog Biomed 62(3):165–175. https://doi.org/10.1016/S0169-2607(00)00065-1

38. Fadzil MHA, Izhar LI, Nugroho H, Nugroho HA (2010) Determination of foveal avascular zone in diabetic retinopathy digital fundus images. Comput Biol Med 40:657–664. https://doi.org/10.1016/j.compbiomed.2010.05.004

39. Fadzil MHA, Izhar LI, Nugroho H, Nugroho HA (2011) Analysis of retinal fundus images for grading of diabetic retinopathy severity. Med Biol Eng Comput 49(6):693–700. https://doi.org/10.1007/s11517-011-0734-2

40. Fong DS, Aiello L, Gardner TW, King GL, Blankenship G, Cavallerano JD, Ferris FL, Klein R (2004) Retinopathy in diabetes. Diabetes Care 27:84–87. https://doi.org/10.2337/diacare.27.2007.S84

41. Fraz MM, Jahangir W, Zahid S, Hamayun MM, Barman SA (2017) Multiscale segmentation of exudates in retinal images using contextual cues and ensemble classification. Biomed Signal Process Control 35:50–62. https://doi.org/10.1016/j.bspc.2017.02.012

42. Fundus Photographic Risk Factors for Progression of Diabetic Retinopathy: ETDRS Report Number 12 (1991) Early treatment diabetic retinopathy study research group. Ophthalmology. 98(5):823–833. https://doi.org/10.1016/S0161-6420(13)38014-2

43. Gadekallu TR, Khare N, Bhattacharya S, Singh S, Maddikunta PKR, Ra I, Alazab M (2020) Early detection of diabetic retinopathy using PCA-firefly based deep learning model. Electronics 9(2):1–16. https://doi.org/10.3390/electronics9020274

44. Gardner GG, Keating D, Williamson TH, Elliott AT (1996) Automatic detection of diabetic retinopathy using an artificial neural network: a screening tool. Br J Ophthalmol 80(11):940–944. https://doi.org/10.1136/bjo.80.11.940

45. Gargeya R, Leng T (2017) Automated identification of diabetic retinopathy using DL. Ophthalmology. 124(7):962–969. https://doi.org/10.1016/j.ophtha.2017.02.008

46. Geetharamani R, Balasubramanian L (2015) Automatic segmentation of blood vessels from retinal fundus images through image processing and data mining techniques. Sadhana. 40(6):1715–1736

47. Gegundez-Arias ME, Marin D, Bravo JM, Suero A (2013) Locating the fovea center position in digital fundus images using thresholding and feature extraction techniques. Comput Med Imaging Graph 37(5–6):386–393. https://doi.org/10.1016/j.compmedimag.2013.06.002

48. Goh JKH, Cheung CY, Sim SS, Tan PC, Tan GSW, Wong TY (2016) Retinal imaging techniques for diabetic retinopathy screening. J Diabetes Sci Technol 10(2):282–294. https://doi.org/10.1177/1932296816629491

49. Grading Diabetic Retinopathy from Stereoscopic Color Fundus Photographs—An Extension of the Modified Airlie House Classification: ETDRS Report Number 10 (1991) Early treatment diabetic retinopathy study research group. Ophthalmology. 98(5):786–806. https://doi.org/10.1016/S0161-6420(13)38012-9

50. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J, Kim R, Raman R, Nelson PC, Mega JL, Webster DR (2016) Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. Jama. 316(22):2402–2410. https://doi.org/10.1001/jama.2016.17216

51. Habib MM, Welikala RA, Hoppe A, Owen CG, Rudnicka AR, Barman SA (2017) Detection of microaneurysms in retinal images using an ensemble classifier. Inf Med Unlocked 9:44–57. https://doi.org/10.1016/j.imu.2017.05.006

52. Hani AFM, Ngah NF, George TM, Izhar LI, Nugroho H, Nugroho HA (2010) Analysis of foveal avascular zone in colour fundus images for grading of diabetic retinopathy severity. 32nd annual international conference of the IEEE EMBS Buenos Aires, Argentina. 5632-5635. https://doi.org/10.1109/IEMBS.2010.5628041

53. Harangi B, Hajdu A (2015) Detection of the optic disc in fundus images by combining probability models. Comput Biol Med 65:10–24. https://doi.org/10.1016/j.compbiomed.2015.07.002

54. Hattiya T, Dittakan K, Musikasuwan S (2021) Diabetic retinopathy detection using convolutional neural network: a comparative study on different architectures. Mahasarakham international journal of engineering. Technology 7(1):50–60. https://doi.org/10.14456/mijet.2021.8

55. Hsiao HK, Liu CC, Yu CY, Kuo SW, Yu SS (2012) A novel optic disc detection scheme on retinal images. Expert Syst Appl 39(12):10600–10606. https://doi.org/10.1016/j.eswa.2012.02.157

56. Huang G, Liu S, Maaten L, Weinberger KQ (2018) CondenseNet: an efficient DenseNet using learned group convolutions. In proceedings of the IEEE conference on computer vision and pattern recognition. 2752-2761

57. Islam SMS, Hasan MM, Abdullah S (2018) Deep learning based early detection and grading of diabetic retinopathy using retinal fundus images. arXiv preprint arXiv:1812.10595v1.

58. Iwendi C, Khan S, Anajemba JH, Mittal M, Alenezi M, Alazab M (2020) The use of ensemble models for multiple class and binary class classification for improving intrusion detection systems. Sensors 20(9):1–37. https://doi.org/10.3390/s20092559

59. Jelinek HF, Pires R, Padilha R, Goldenstein S, Wainer J, Bossomaier T, Rocha A (2012) Data fusion for multi-lesion diabetic retinopathy detection. 25th IEEE international symposium on computer-based medical systems (CBMS), pp. 1-4. https://doi.org/10.1109/CBMS.2012.6266342

60. Ji Q, Huang J, He W, Sun Y (2019) Optimized deep convolutional neural networks for identification of macular diseases from optical coherence tomography images. Algorithms 12(3):1–12. https://doi.org/10.3390/a12030051

61. Jiang X, Xiang D, Zhang B, Zhu W, Shi F and Chen X (2016) Automatic Co-segmentation of Lung Tumor based on Random forest in PET-CT Images. Medical Imaging 2016: Image processing. SPIE Medical Imaging https://doi.org/10.1117/12.2216361

62. Jin C, Shi F, Xiang D, Jiang X, Zhang B, Wang X, Zhu W, Gao E, Chen X (2016) 3D fast automatic segmentation of kidney based on modified AAM and random Forest. IEEE Trans Med Imaging 35(6):1395–1407. https://doi.org/10.1109/TMI.2015.2512606

63. Junior SB, Welfer D (2013) Automatic detection of microaneurysms and hemorrhages in color eye fundus images. Int J Comput Sci Inf Technol 5(5):21–37. https://doi.org/10.5121/ijcsit.2013.5502

64. Kale P, Janwe N (2017) Detection of retinal hemorrhage in color fundus image. Int J Adv Res Comput Commun Eng 6(3):1002–1005. https://doi.org/10.17148/IJARCCE.2017.63233

65. Kamal KC, Yin Z, Wu M, Wu Z (2021) Evaluation of deep learning-based approaches for COVID-19 classification based on chest X-ray images. SIViP 15:959–966. https://doi.org/10.1007/s11760-020-01820-2

66. Kamble R, Kokare M (2017) Detection of microaneurysms using local rank transform in color fundus images. 2017 IEEE international conference on image processing. Pp. 4442—4446. https://doi.org/10.1109/ICIP.2017.8297122

67. Kamble R, Kokare M, Deshmukh G, Hussain FA, Meriaudeau F (2017) Localization of optic disc and fovea in retinal images using intensity based line scanning analysis. Comput Biol Med 87:382–396. https://doi.org/10.1016/j.compbiomed.2017.04.016

68. Kamel M, Belkassim S, Mendonca AM, Campilho A (2001) A neural network approach for the automatic detection of microaneurysms in retinal angiograms. International joint conference on neural networks. Proceedings, 4: 2695-2699. https://doi.org/10.1109/IJCNN.2001.938798

69. Kao EF, Lin PC, Chou MC, Jaw TS, Liu GC (2014) Automated detection of fovea in fundus images based on vessel-free zone and adaptive Gaussian template. Comput Methods Prog Biomed 117(2):92–103. https://doi.org/10.1016/j.cmpb.2014.08.003

70. Kermany DS, Goldbaum M, Cai W, Valentim CCS, Liang H, Baxter SL, McKeown A, Yang G, Wu X, Yan F, Dong J, Prasadha MK, Pei J, Ting MYL, Zhu J, Li C, Hewett S, Dong J, Ziyar I, … Zhang K (2018) Identifying medical diagnoses and treatable diseases by image-based deep Learning. Cell 172:1122–1131. https://doi.org/10.1016/j.cell.2018.02.010

71. Khojasteh P, Aliahmad B, Kumar DK (2018) Fundus images analysis using deep features for detection of exudates, hemorrhages and microaneurysms. BMC Ophthalmology.1-13. 18. https://doi.org/10.1186/s12886-018-0954-4

72. Kirkpatrick C (2013) Intraretinal microvascular abnormality (IrMA). EyeRounds online atlas of ophthalmology. Available at https://webeye.ophth.uiowa.edu/eyeforum/atlas/pages/IRMA.htm, accessed on 27-05-2020

73. Kokame GT, Lai JC (2012) Intraretinal microvascular abnormalities, retina image Bank, American Society of Retina Specialists Available at: https://imagebank.asrs.org/file/1361/intraretinal-microvascular-abnormalities, accessed on 14-05-2020

74. Kulenovic I, Rasic S, Karcic S (2006) Development of microvascular complications in type 1 diabetic patients 10 years follow-up. Bosnian J Basic Med Sci 6(2):47–50. https://doi.org/10.17305/bjbms.2006.3171

75. Kumar PNS, Deepak RU, Sathar A, Sahasranamam V, Kumar RR (2016) Automated detection system for diabetic retinopathy using two field fundus photography. Procedia Comput Sci 93:486–494. https://doi.org/10.1016/j.procs.2016.07.237

76. Lachure J, Deorankar AV, Lachure S, Gupta S, Jadhav R (2015) Diabetic retinopathy using morphological operations and machine learning. In 2015 IEEE international advance computing conference (IACC). 617-622. https://doi.org/10.1109/IADCC.2015.7154781

77. Lam C, Yu C, Huang L, Rubin D (2018) Retinal lesion detection with deep learning using image patches. Multidisciplinary Ophthalmic Imaging 59(1):590–596. https://doi.org/10.1167/iovs.17-22721

78. Lam C, Yi D, Guo M, Lindsey T (2018) Automated Detection of Diabetic Retinopathy using DL. AMIA Summits on Translational Science Proceedings.147–155. PMID: 29888061; PMCID: PMC5961805.

79. Lazar I, Hajdu A (2013) Retinal microaneurysm detection through local rotating cross-section profile analysis. IEEE Trans Med Imaging 32(2):400–407. https://doi.org/10.1109/TMI.2012.2228665

80. Lee J, Kim YK, Park KH, Jeoung JW (2020) Diagnosing Glaucoma with spectral-domain optical coherence tomography using deep learning classifier. J Glaucoma 29(4):287–294. https://doi.org/10.1097/IJG.0000000000001458

81. Li B, Li HK (2013) Automated analysis of diabetic retinopathy images: principles, recent developments, and emerging trends. Curr Diab Rep 13(4):453–459. https://doi.org/10.1007/s11892-013-0393-9

82. Li C, Kao CY, Gore JC, Ding Z (2007) Implicit active contours driven by local binary fitting energy. 2007 IEEE conference on computer vision and pattern Recognition.1-7.https://doi.org/10.1109/CVPR.2007.383014

83. Li Y, Yeh N, Chen S, and Chung Y (2019) Computer-assisted diagnosis for diabetic retinopathy based on fundus images using deep convolutional neural network. Mob Inf Syst 1-14. https://doi.org/10.1155/2019/6142839.

84. Liskowski P, Krawiec K (2015) Segmenting retinal blood vessels with deep neural networks. IEEE Trans Med Imaging 35(11):2369–2380. https://doi.org/10.1109/TMI.2016.2546227

85. Ma J, Plonka G (2010) A review of Curvelets and recent applications. IEEE Signal Process Mag 27(2):118–133

86. Manjaramkar A, Kokare M (2017) Statistical Geometrical Features for Microaneurysm Detection. J Digit Imaging 31:224–234. https://doi.org/10.1007/s10278-017-0008-0

87. Mann KS, Kaur S (2017) Segmentation of retinal blood vessels using optimized features for detection of diabetic retinopathy. Int J Res Appl Sci Eng Technol 5(12):2811–2821. https://doi.org/10.1063/1.4981966

88. Mansour RF (2017) Evolutionary computing enriched computer aided diagnosis system for diabetic retinopathy: a survey. IEEE Rev Biomed Eng 10:334–349. https://doi.org/10.1109/RBME.2017.2705064

89. Mansour RF (2018) Deep-learning-based automatic computer-aided diagnosis system for diabetic retinopathy. Biomed Eng Lett 8(1):41–57. https://doi.org/10.1007/s13534-017-0047-y

90. Massey EM, Hunter A (2011) Augmenting the classification of retinal lesions using spatial distribution. 33rd annual international conference of the IEEE EMBS Boston. 3967-3970. https://doi.org/10.1109/IEMBS.2011.6090985

91. Medhi JP, Dandapat S (2016) An effective fovea detection and automatic assessment of diabetic maculopathy in color fundus images. Comput Biol Med 74:30–44. https://doi.org/10.1016/j.compbiomed.2016.04.007

92. Memari N, Ramli AR, Saripan MIB, Mashohor S, Moghbel M (2018) Retinal blood vessel segmentation by using matched filtering and fuzzy C-means clustering with integrated level set method for diabetic retinopathy assessment. J Med Biol Eng 39:713–731. https://doi.org/10.1007/s40846-018-0454-2

93. Meshram SP, Pawar MS (2013) Extraction of retinal blood vessels from diabetic retinopathy imagery using contrast limited adaptive histogram equalization. Int J Adv Comput Theory Eng 2(3):143–147

94. Microaneyrysms, The COMS Grading Scheme: Graded Features, University of IOWA Health Care, Department of Ophthalmology and Visual Sciences, Available at http://webeye.ophth.uiowa.edu/dept/coms/grading/images/11-mircoaneurysms.jpg Accessed on 13-07-2020

95. Mizutani A, Muramatsu C, Hatanaka Y, Suemori S, Hara T, Fujita H (2009) Automated microaneurysm detection method based on double-ring filter in retinal fundus images. Proceedings of SPIE 7260. Medical imaging 2009: computer-aided diagnosis. 72-78. https://doi.org/10.1117/12.813468

96. Mookiah MRK, Acharya UR, Martis RJ, Chua CK, Lim CM, Ng EYK, Laude A (2013) Evolutionary algorithm based classifier parameter tuning for automatic diabetic retinopathy grading: a hybrid feature extraction approach. Knowl-Based Syst 39:9–22. https://doi.org/10.1016/j.knosys.2012.09.008

97. Niemeijer M, Van Ginneken B, Staal J, Suttorp-Schulten MSA, Abramoff MD (2005) Automatic detection of red lesions in digital color fundus photographs. IEEE Trans Med Imaging 24(5):584–592. https://doi.org/10.1109/TMI.2005.843738

98. Noor-ul-huda M, Tehsin S, Ahmed S, Niazi FAK, Murtaza Z (2019) Retinal images benchmark for the detection of diabetic retinopathy and clinically significant macular edema (CSME). Biomed Eng 64(3):297–307. https://doi.org/10.1515/bmt-2018-0098

99. Nugroho HA, Purnamasari D, Soesanti I, Oktoeberza WKZ, Dharmawan DA (2017) Segmentation of foveal avascular zone in colour fundus images based on retinal capillary endpoints detection. J Telecom Electron Comput Eng 9(3–8):107–112

100. Oliveira WS, Teixeira JV, Ren TI, Cavalcanti GDC, Sijbers J (2016) Unsupervised retinal vessel segmentation using combined filters. PLoS One 1-21. 11. https://doi.org/10.1371/journal.pone.0149943

101. Orlando JI, Prokofyeva E, del Fresno M, Blaschko MB (2018) An ensemble deep learning based approach for red lesion detection in fundus images. Comput Methods Prog Biomed 153:115–127. https://doi.org/10.1016/j.cmpb.2017.10.017

102. Patwari MB, Manza RR, Rajput YM, Saswade M, Deshpande N (2013) Detection and counting the microaneurysms using image processing techniques. Int J Appl Inf Syst 6(5):11–17

103. Pentland BT, Liu P, Kremser W, Haerem T (2020) The dynamics of drift in digitized processes. MIS quarterly. 44(1): 19–47. https://doi.org/10.25300/MISQ/2020/14458

104. Pereira C, Veiga D, Mahdjoub J, Guessoum Z, Gonçalves L, Ferreira M, Monteiro J (2014) Using a multi-agent system approach for microaneurysm detection in fundus images. Artif Intell Med 60(3):179–188. https://doi.org/10.1016/j.artmed.2013.12.005

105. Pereira C, Gonçalves L, Ferreira M (2015) Exudate segmentation in fundus images using an ant colony optimization approach. Inf Sci 296:14–24. https://doi.org/10.1016/j.ins.2014.10.059

106. Porwal P, Pachade S, Kamble R, Kokare M, Deshmukh G, Sahasrabuddhe V, Meriaudeau F (2018) Indian diabetic retinopathy image dataset (IDRiD): a database for diabetic retinopathy screening research. Data. 3(3):25. https://doi.org/10.3390/data3030025

107. Pour AM, Seyedarabi H, Jahromi SHA, Javadzadeh A (2020) Automatic detection and monitoring of diabetic retinopathy using efficient convolutional neural networks and contrast limited adaptive histogram equalization. IEEE Access 8:136668–136673. https://doi.org/10.1109/ACCESS.2020.3005044

108. Pratt H, Coenen F, Broadbent DM, Harding SP, Zheng Y (2016) Convolutional neural networks for diabetic retinopathy. Procedia Comput Sci 90:200–205. https://doi.org/10.1016/j.procs.2016.07.014

109. Prentašić P, Lončarić S (2016) Detection of exudates in fundus photographs using deep neural networks and anatomical landmark detection fusion. Comput Methods Prog Biomed 137:281–292. https://doi.org/10.1016/j.cmpb.2016.09.018

110. Proliferative Diabetic Retinopathy - Optic Disc Neovascularization - Post Intravitreal Avastin Injections, The Retina Reference, Available at - http://www.retinareference.com/diseases/beb00894be590ec0/images/46019c8a9e/, Accessed on 23-11-2020

111. Quellec G, Lamard M, Josselin PM, Cazuguel G, Cochener B, Roux C (2008) Optimal wavelet transform for the detection of microaneurysms in retina photographs. IEEE Trans Med Imaging 27(9):1230–1241. https://doi.org/10.1109/TMI.2008.920619

112. Quellec G, Charrière K, Boudi Y, Cochener B, Lamard M (2017) Deep image Mining for Diabetic Retinopathy Screening. Med Image Anal 39:178–193. https://doi.org/10.1016/j.media.2017.04.012

113. Qureshi RJ, Kovacs L, Harangi B, Nagy B, Peto T, Hajdu A (2012) Combining algorithms for automatic detection of optic disc and macula in fundus images. Comput Vis Image Underst 116(1):138–145. https://doi.org/10.1016/j.cviu.2011.09.001

114. Rahimzadeh M, Attar A (2020) A modified deep convolutional neural network for detecting COVID-19 and pneumonia from chest X-ray images based on the concatenation of Xception and ResNet50V2. Inf Med unlocked 19:1–9. https://doi.org/10.1016/j.imu.2020.100360

115. Raja DSS, Vasuki S (2015) Screening diabetic retinopathy in developing countries using retinal images. Appl Comput Inf 36(1):13–22

116. Rajput YM, Manza RR, Patwari MB (2015) Extraction of cotton wool spot using multi resolution analysis and classification using K-means clustering. Int J Comput Appl 975:1–5

117. Rakhlin A (2017) Diabetic retinopathy detection through integration of deep learning classification framework. bioRxiv, p.225508. https://doi.org/10.1101/225508

118. Ravishankar S, Jain A, Mittal A (2009) Automated feature extraction for early detection of diabetic retinopathy in fundus images. 2009 IEEE conference on computer vision and pattern recognition. Pp. 210-217. https://doi.org/10.1109/CVPR.2009.5206763

119. Retinal blood vessels, IMAIOS, Available at - https://www.imaios.com/en/e-Anatomy/Anatomical-Parts/Retinal-blood-vessels, Accessed on 19-06-2020

120. Reza MN (2018) Automatic detection of optic disc in color fundus retinal images using circle operator. Biomed Signal Process Control 45:274–283. https://doi.org/10.1016/j.bspc.2018.05.027

121. Rosas-Romero R, Martínez-Carballido J, Hernández-Capistrán J, Uribe-Valencia LJ (2015) A method to assist in the diagnosis of early diabetic retinopathy: image processing applied to detection of microaneurysms in fundus images. Comput Med Imaging Graph 44:41–53. https://doi.org/10.1016/j.compmedimag.2015.07.001

122. Samanta A, Saha A, Satapathy SC, Fernandes SL, Zhang Y (2020) Automated detection of diabetic retinopathy using convolutional neural networks on a small dataset. Pattern Recogn Lett 135:293–298. https://doi.org/10.1016/j.patrec.2020.04.026

123. Samek W, Montavon G, Lapuschkin S, Anders CJ, Müller K (2021) Explaining deep neural networks and beyond: a review of methods and applications. Proc IEEE 109(3):247–278. https://doi.org/10.1109/JPROC.2021.3060483

124. Sarki R, Michalska S, Ahmed K, Wang H, Zhang Y (2019) Convolutional neural networks for mild diabetic retinopathy detection: an experimental study. bioRxiv. 1-18. https://doi.org/10.1101/763136

125. Seoud L, Hurtut T, Chelbi J, Cheriet F, Langlois JMP (2015) Red lesion detection using dynamic shape features for diabetic retinopathy screening. IEEE Trans Med Imaging 35(4):1116–1126. https://doi.org/10.1109/TMI.2015.2509785

126. Sinthanayothin C, Boyce JF, Williamson TH, Cook HL, Mensah E, Lal S, Usher D (2002) Automated detection of diabetic retinopathy on digital fundus images. Diabet Med 19(2):105–112. https://doi.org/10.1046/j.1464-5491.2002.00613.x

127. Soniya, Paul S, Singh L (2016) Heterogeneous modular deep neural network for diabetic retinopathy detection. In 2016 IEEE Region 10 Humanitarian Technology Conference (R10-HTC). pp. 1–6. https://doi.org/10.1109/R10-HTC.2016.7906821

128. Sopharak A, Uyyanonvara B, Barman S (2009) Automatic exudate detection from non-dilated diabetic retinopathy retinal images using fuzzy C-means clustering. Sensors. 9(3):2148–2161. https://doi.org/10.3390/s90302148

129. Sopharak A, Uyyanonvara B, Barman S (2011) Automatic microaneurysm detection from non-dilated diabetic retinopathy retinal images using mathematical morphology methods. IAENG Int J Comput Sci 38(3):295–301

130. Spencer T, Olson JA, McHardy KC, Sharp PF, Forrester JV (1996) An image-processing strategy for the segmentation and quantification of microaneurysms in fluorescein angiograms of the ocular fundus. Comput Biomed Res 29(4):284–302. https://doi.org/10.1006/cbmr.1996.0021

131. Stewart JM, Coassin M, Schwartz DM. Diabetic Retinopathy. (2017) In: Feingold KR, Anawalt B, Boyce A, et al., editors. Endotext [Internet]. South Dartmouth (MA): MDText.com, Inc.; 2000-. Figure 8, [Active neovascularization in PDR. Fibrovascular...]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK278967/figure/diab-retinopathy_f_diab-retinopathy_etx-dm-ch29-fig 8/ Accessed on 27-05-2020

132. Takahashi H, Tampo H, Arai Y, Inoue Y, Kawashima H (2017) Applying artificial intelligence to disease staging: deep learning for improved staging of diabetic retinopathy. PLoS One 12(6):1–11. https://doi.org/10.1371/journal.pone.0179790

133. Tobin KW, Chaum E, Govindasamy VP, Karnowski TP (2007) Detection of anatomic structures in human retinal imagery. IEEE Trans Med Imaging 26(12):1729–1739. https://doi.org/10.1109/TMI.2007.902801

134. Tymchenko B, Marchenko P, Spodarets D (2020) Deep Learning Approach to Diabetic Retinopathy Detection. arXiv preprint arXiv:2003.02261.1–9.

135. Usher D, Dumskyj M, Himaga M, Williamson TH, Nussey S, Boyce J (2003) Automated detection of diabetic retinopathy in digital retinal images: a tool for diabetic retinopathy screening. Diabet Med 21(1):84–90. https://doi.org/10.1046/j.1464-5491.2003.01085.x

136. Van Grinsven MJJP, Van Ginneken B, Hoyng CB, Theelen T, Sánchez CI (2016) Fast convolutional neural network training using selective data sampling: application to hemorrhage detection in color fundus images. IEEE Trans Med Imaging 35(5):1273–1284. https://doi.org/10.1109/TMI.2016.2526689

137. Walter T, Klein JC, Massin P, Erginay A (2002) A contribution of image processing to the diagnosis of diabetic retinopathy-detection of exudates in color fundus images of the human retina. IEEE Trans Med Imaging 21(10):1236–1243. https://doi.org/10.1109/TMI.2002.806290

138. Wang Z, Yang J (2018) Diabetic retinopathy detection via deep convolutional networks for discriminative localization and visual explanation. The Workshops of the Thirty-Second AAAI Conference on Artificial Intelligence:514–521

139. Wild S, Roglic G, Green A, Sicree R, King H (2004) Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. Diabetes Care 27(5):1047–1053. https://doi.org/10.2337/diacare.27.5.1047

140. Wilkinson CP, Ferris FL, Klein RE, Lee PP, Agardh CD, Davis M, Dills D, Kampik A, Pararajasegaram R, Verdaguer JT (2003) Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. Global diabetic retinopathy project group. Ophthalmology. 110(9):1677–1682. https://doi.org/10.1016/S0161-6420(03)00475-5

141. Wu B, Zhu W, Shi F, Zhu S, Chen X (2017) Automatic detection of microaneurysms in retinal fundus images. Comput Med Imaging Graph 55:106–112. https://doi.org/10.1016/j.compmedimag.2016.08.001

142. Xu K, Feng D, Mi H (2017) Deep convolutional neural network-based early automated detection of diabetic retinopathy using fundus image. Molecules 22(12):2054. https://doi.org/10.3390/molecules22122054

143. Yu F, Sun J, Li A, Cheng J, Wan C, Liu J (2017) Image quality classification for DR screening using deep learning. 2017 39th annual international conference of the IEEE engineering in medicine and biology society. 664-667. https://doi.org/10.1109/EMBC.2017.8036912

144. Zhang B, Wu X, You J, Li Q, Karray F (2010) Detection of microaneurysms using multi-scale correlation coefficients. Pattern Recogn 43(6):2237–2248. https://doi.org/10.1016/j.patcog.2009.12.017

145. Zhang B, Karray F, Li Q, Zhang L (2012) Sparse representation classifier for microaneurysm detection and retinal blood vessel extraction. Inf Sci 200:78–90. https://doi.org/10.1016/j.ins.2012.03.003

146. Zhou W, Yi Y, Gao Y, Dai J (2019) Optic disc and cup segmentation in retinal images for Glaucoma diagnosis by locally statistical active contour model with structure prior. Comput Math Methods Med 2019:1–17. https://doi.org/10.1155/2019/8973287

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.