A Survey on Medical Image Analysis in Diabetic Retinopathy

Skylar Stolte [1], Ruogu Fang [1]

1. J. Crayton Pruitt Family Department of Biomedical Engineering, University of Florida. 1275 Center Drive, Biomedical Sciences Building JG56 P.O. Box 116131 Gainesville, FL 32611-6131, USA.

Corresponding Author:
Ruogu Fang
J287, Biomedical Science Building
1275 Center Drive, APT: AR 351B
Gainesville Florida 32611, United States
Tel: (352) 294-1375
E-mail: ruogu.fang@bme.ufl.edu

A Survey on Medical Image Analysis in Diabetic Retinopathy

Skylar Stolte, Ruogu Fang

J. Crayton Pruitt Family Department of Biomedical Engineering, University of Florida

---

Abstract

Diabetic Retinopathy (DR) represents a highly-prevalent complication of diabetes in which individuals suffer from damage to the blood vessels in the retina. The disease manifests itself through lesion presence, starting with microaneurysms, at the nonproliferative stage before being characterized by neovascularization in the proliferative stage. Retinal specialists strive to detect DR early so that the disease can be treated before substantial, irreversible vision loss occurs. The level of DR severity indicates the extent of treatment necessary - vision loss may be preventable by effective diabetes management in mild (early) stages, rather than subjecting the patient to invasive laser surgery. Using artificial intelligence (AI), highly accurate and efficient systems can be developed to help assist medical professionals in screening and diagnosing DR earlier and without the full resources that are available in specialty clinics. In particular, deep learning facilitates diagnosis earlier and with higher sensitivity and specificity. Such systems make decisions based on minimally handcrafted features and pave the way for personalized therapies. Thus, this survey provides a comprehensive description of the current technology used in each step of DR diagnosis. First, it begins with an introduction to the disease and the current technologies and resources available in this space. It proceeds to discuss the frameworks that different teams have used to detect and classify DR. Ultimately, we conclude that deep learning systems offer revolutionary potential to DR identification and prevention of vision loss.

Keywords: Diabetic Retinopathy, Deep Learning, Image Mining, Lesion Detection

---

1. Introduction

Diabetic Retinopathy (DR) is a complication of diabetes that can lead to blindness if untreated Watkins (2003). Patients who have diabetes must undergo frequent screenings that are limited by resources and can put high strain on patients and ophthalmologists. DR is classified as either non-proliferative (NPDR), which is characterized by lesions such as microaneurysms (MAs) and exudates, or proliferative (PDR), in which neovascularization of weak blood vessels occurs Yun et al. (2008). More accurate systems would allow DR to be diagnosed early, such that patients can undergo treatment and avoid blindness. Manual diagnosis has difficulty meeting the eighty percent sensitivity that is recommended in Britain, and less developed areas particularly suffer from an imbalanced patient-to-specialist ratio. Studies estimate that by twenty years after diabetes onset, nearly all of patients with type I diabetes and sixty percent of those with type II diabetes will have DR.

Automated processing techniques have risen to prominence to address issues in DR classification, enabling screening to distinguish those who require further referral from those who are classified as low-risk. Machine learning (ML) algorithms extract image features which are feed into statistical classifiers Litjens et al. (2017). Pattern recognition is facilitated using training sets, which allows the algorithms to optimize decisions in the high-dimensional feature space. Systems vary in the level of hand-crafted features. Deep learning (DL) processes, most notably Convolutional Neural Networks (CNNs), determine the features and rules that optimize classification accuracy with minimum handcrafted components Krause et al. (2018a).

Treatments such as laser surgery (i.e. panretinal photocoagulation) or intravitreous injections of anti-vascular endothelial growth factor provide effective means of halting vision loss from DR or diabetic macular edema (DME) if the disease is detected sufficiently early; however, many individuals cannot undergo timely screenings due to limited resources in personnel and technology Williamson and Keating (1998); Wong and Bressler (2016). Traditional imaging methodologies rely on expensive or non-transportable technologies and thus limit a non-local approach to DR screening and diagnosis Williamson and Keating (1998). Therefore, the rise in digital cameras and scanning ophthalmoscopes have increased the feasibility of a telemedicine approach to addressing DR. Under such a system, digital images can be transported to remote or cloud-based storage systems for subsequent review. However, this increasing availability would also be likely be accompanied by reduced image quality and image field-of-view (FOV). This, along with limitations in expert personnel, increases the need for computer-aided diagnostic (CAD) systems.

The 1990's saw the first modern advances in DR diagnosis and grading with the introduction of a new clinical standard in the ETDRS scale ETDRS Research Group (1991) and the results of artificial neural networks (ANN) in DR detection Gardner et al. (1996b); Williamson and Keating (1998). Interestingly, ANNs seemed to fall in popularity in this space in the early 2000s. Rather, processing methods focused more on image processing methods and different statistical classifiers. This may have been due to a belief that the current NNs were insufficient for detecting specific features of DR, like DR lesions Lee et al. (2001). Many methods strayed away from image-level diagnosis during this time. This may perhaps be due to a fear that NNs learn risk factors of DR, or even meaningless features, rather than the underlying disease Wong and Bressler (2016). At the same time, digital imaging and other rapid screening tools have also heightened the prevalence of statistical and machine learning classifiers in this space, as in Bayesian, Mahalanobis, KNN, SVM, and other systems Ege et al. (2000).

Deep learning systems rose to prominence in the computer vision space with the success of AlexNet and the ImageNet challenge Krizhevsky et al. (2012). In particular, Google's work in DR classification and grading shifted the focus of this space to deep learning algorithms Gulshan et al. (2016a). Recent works show potential for taking DR prediction in increasingly complex and clinically relevant directions. For instance, a recent work in Nature achieved promising results to pave the way for longitudinal DR prediction and tracking Arcadu et al. (2019). Despite this advances, there has been slow progress in terms of the acceptance and implementation of these systems in widespread use. The United States Federal Drug Agency (US FDA) only approved the first AI system for DR detection, IDx-DR, in 2018 Stark. Even so, the system's approval was limited in use. For example, those with other eye diseases or severe DR were not included in its approved use cases.

Researchers must consider the advances in DR imaging methodologies when designing appropriate AI systems for DR diagnosis and grading. The golden standard for DR detection has traditionally been the ETDRS seven standard fields Aiello et al. (2019). This standard uses seven 30-degree stereoscopic images that encompass the optic disc, macula, area temporal to the macula, and the four vascular arcades Li et al. (2011). While this method correlates well to DR detection, it only covers approximately a 75-degree field-of-view (FOV) and was established before digital fundus imaging was possible Bae et al.. In fact, a recent study on ultra-wide field photography (above 200 FOV) found that at least 0.11 of eyes with DR may be at least two stages more severe than that found using the ETDRS standard Aiello et al. (2019). This is likely due to the importance of the peripheral retina in detecting retinopathy signs like microvascular abnormalities, neovascularizations, vessel leakage, and nonperfused areas Rabiolo et al. (2017). Therefore, there has been increasing evidence that diagnostic systems need to compensate for a wider view of the retina, particularly in grading tasks. Currently, retinal specialists primarily diagnose DR on the basis of color stereographic photography, fluorescein angiography, and optical coherence tomography (OCT) Cicinelli et al. (2019). However, new technologies, such as fundus autofluorescence and OCT angiography (OCTA), are emerging as interest areas in the field. These technologies, and compatible computer-aided diagnostic (CAD) systems, show great promise for capturing the intricacies of the disease. At the same time, an increasing need for DR screening in lower access areas has raised an interest in use of lower quality imaging devices. This movement has correlated to increase interest in CAD systems that compensate for lower FOV (i.e. single field, 20-50 FOV) with technologies such as image stitching and montage Palacios et al. (2019).

Our survey differs from past work, as we include novel DL pipelines, overview imaging and ML processes, and discuss all tasks for grading DR (i.e. optic disc, blood vessels, lesions, and grading). This discussion facilitates clinical implementation of state-of-the-art systems. Our inter-disciplinary material will provide one source for research teams to find and understand these systems. Ultimately, DL frameworks offer particular promise for customizable and patient-specific diagnostic and therapeutic care. Section 2 overviews current resources for studying DR. Sections 3-4 describe the imaging modalities and automated frameworks for retinal imaging and classifying DR, respectively. Finally, section 5 outlines different methodologies that address aspects of DR detection and grading, and sections 6-7 provide our insights.

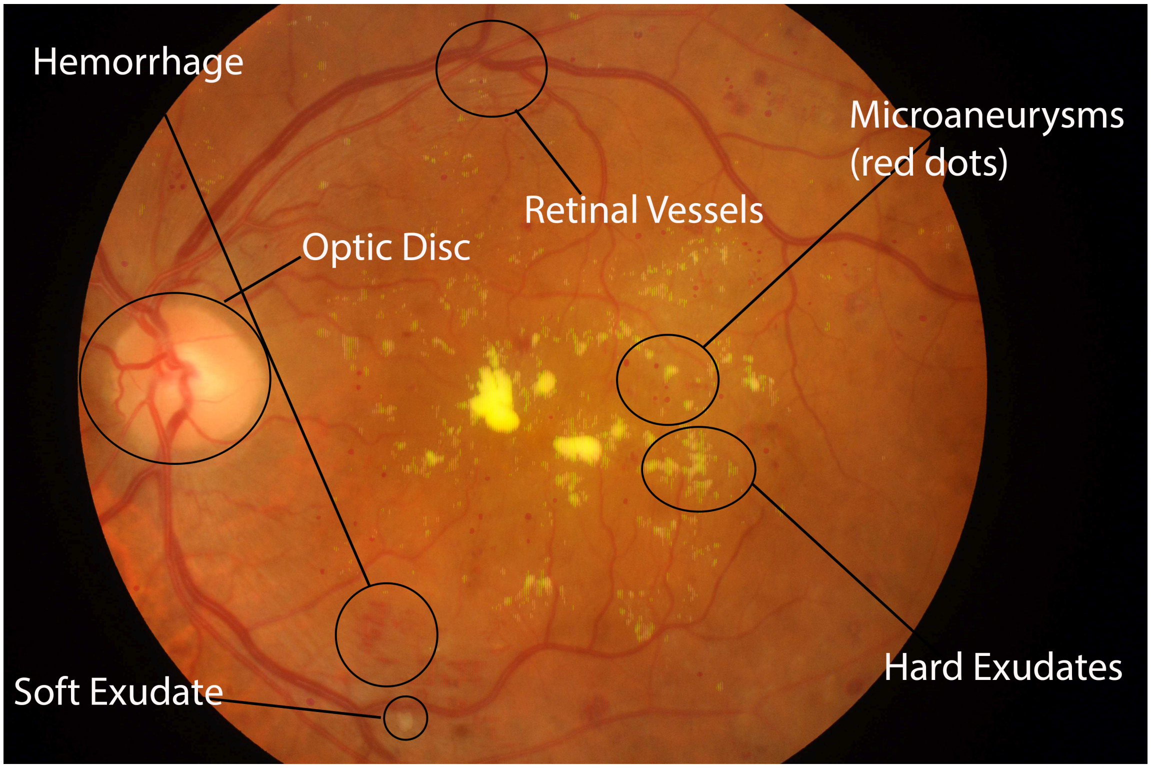

Figure 1: Examples of the different eye structures and lesion types that are important to diagnose DR

2. Overview of Literature Resources

2.1. Datasets

Dataset	Modality	Task	Year	Country	Total # Images	Labels	Camera	FOV (°)
FIRE Hernandez-Matas et al. (2017)	Color Fundus	image registration	2017	Thessaloniki	129	-	Nidek AFC-210	45
STARE Hoover and Goldbaum	Color Fundus	vessel segmentation	-	USA	400 (overall) 40 (vessels)	0-1	-	-
DRIVE Staal et al. (2004)	Color Fundus	vessel segmentation	2004	The Netherlands	40	0-1	Canon CR5 non-mydriatic 3CCD	45
REVIEW Al-Diri et al. (2008)	Color Fundus	vessel segmentation	2008	United Kingdom	16	0-1, width (m)	Cannon 60 UV, Zeiss FF 450 and JVC 3CCD	50-60
HRF Image Database Budai et al. (2013)	Color Fundus	vessel segmentation	2013	-	15	0-1	Canon CR-1	45
ROC Niemeijer et al. (2010)	Color Fundus	MA	2010	USA	100	0-1	Topcon NW 100, NW 200 or Canon CR5-45NM non-mydriatic	45
E-Ophtha Decencière et al. (2013)	Color Fundus	MA, EXU	2014	France	MA:381, EXU:82	0-1	Topcon TRC NW6 non-mydriatic	45
DIARETDB0 Kauppi et al. (2006)	Color Fundus	EXU, MA, HEM, NV	2006	Finland	130	0-1	-	50
DIARETDB1 v2.1 Kauppi et al. (2007); Tomi et al. (2009)	Color Fundus	EXU, MA, HEM	2009	Finland	89	0-1	-	50
HEI-MED (DMED) Giancardo et al. (2012)	Color Fundus	EXU, CWS, DRU, visible fluid	2012	USA	169	0-1	Zeiss Visucam PRO fundus	45
DRiDB Prentašić et al. (2013)	Color Fundus	Lesion segmentation and grading	2013	Croatia	100	0-1	VISUCAM 200	45
DDR Li et al. (2019)	Color Fundus	DR grading MA, EXU, HEM	2019	China	13,673	DR: 0-4 Lesions: 0-1	42 types, mainly Topcon, Nikon, and Canon	45
VICAVR VARPA Group (2019)	Color Fundus	DR detection	2010	Spain	58	0-1	TopCon non-mydriatic camera NW-100	-
Messidor Decencière et al. (2014)	Color Fundus	DR grading	2014	France	1200	0-3	Topcon TRC NW6 non-mydriatic	45
Kaggle Kaggle (2015a)	Color Fundus	DR grading	2015	USA	10,000+	0-4	-	-
IDRiD Porwal et al. (2018a,b)	Color Fundus	DR/DME Grading	2018	India	516	DR: 0-4, DEM: 0-1	Kowa VX 10 alpha digital fundus	50
Duke Farsiu SD-OCT Farsiu et al. (2014)	OCT	AMD detection	2013	USA	384	0-1	Bioptigen, Inc. SD-OCT	70
OCT Image Database Gholami et al. (2020a,b)	OCT	DR, AMD, MH grading	2018	India	+500	0-3 (each)	Cirrus HD-OCT machine	12mm (degree not provided)
Unnamed Noor Eye Hospital dataset Rasti et al. (2018)	OCT	Dry AMD, DME detection	2018	Iran	148	0-1 (each)	Heidelberg SD-OCT	30x30, optional 55 lens
Kaggle Retinal OCT Mooney (2018)	OCT	CNV, DME, Drusen detection	2018	USA, China	84,495	0-1 (all)	Spectralis OCT	55
Feiz Hospital FA photographs of patients with Diabetes Hajeb Mohammad Alipour et al. (2014)	FA	DR	2003	Iran	70	0-4	-	-
Fundus FA Photographs Colour Fundus Images of Patients with Diabetes Alipour et al. (2012)	FA and color fundus	DR	2012	Iran	60 per modality	0-2	-	-
Duke Farsiu FA data [?]	FA videos and late images	DME	2015	USA	24 per modality	area of leakage	Heidelberg Spectalis 6-mode HRA/OCT unit	55

Table 1: The retinal databases available for DR study. EXU: Exudates, MA: Microaneurysms, HEM: Hemorrhages, NV: Neovascularization, DME: Diabetic macular edema, CWS: Cotton wool spots, DRU: Drusen, AMD: Age-related Macular Degeneration, MH: Macular Hole, CNV: Choroidal neovascularization, FOV: field-of-view in degrees

Table 1 shows publicly-available retinal databases that are used to study DR. These databases focus primarily on DR classification and grading and on segmentation of lesions, blood vessels, and the optic disc. These databases facilitate direct comparison between competing DL frameworks and model validation across research teams. Such databases compensate for issues underlying local or private databases that may impair the model training. For instance, general clinical data may suffer from incomplete or inaccurate labeling or from small size. As such, public databases supplement this data with large image collections with consistent labeling. Commonly, research teams apply this data to pretrain their models and adjust network weights to smaller datasets through transfer learning.

Key considerations when selecting a database depend on (1) the task you are focusing on and (2) your intended subject population. Task selection involves the overall task (i.e. grading versus lesion detection) and the target results (binary versus scaled grading). Different population considerations encompass total number of images, population base (i.e. country or region, interest "group" of focus), and technical classes like camera type and FOV. Additionally, you should consider the computational strength of your operating system(s) when selecting a database, especially in terms of number of images. Rights to certain databases depend on participation in eye diagnosis competitions, or other affairs that involve petitioning membership or use on the owner's site. Keep in mind that certain publishers may restrict use of their databases; for instance, they may require data not be shared to outside parties and/or that a group cites their original works when publishing.

2.2. Literature Trends

Figs. 2-4 display literature trends for DR over time in reference to general search (fig. 2), machine learning search (fig. 3), and deep learning search (fig. 4). These figures indicate that interest surrounding DR has increased over time, likely due to increased prevalence. DR studies using AI have only emerged in recent years - DL studies in particular have increased dramatically since 2015. While AI methods for DR diagnosis and grading comprise a small percentage of the total DR publications, these articles include landmark studies published by sources like JAMA, MICCAI, Ophthalmology, and IEEE. Section 5 surveys key publications that follow each step of DR analysis for comprehensive evaluation of retinal images. These sources provide clinicians and researchers with insight into DR research to pave the way to future progress in DR diagnosis and therapies.

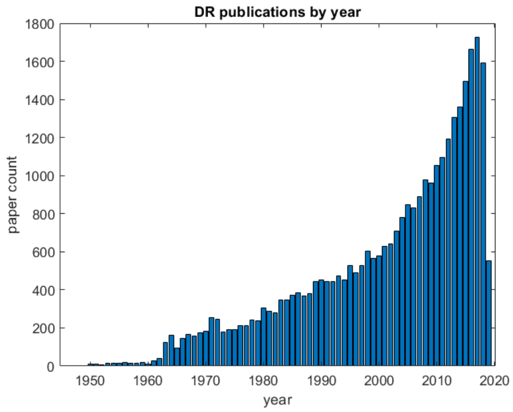

Figure 2: Publications that reference DR by year according to PubMed search.

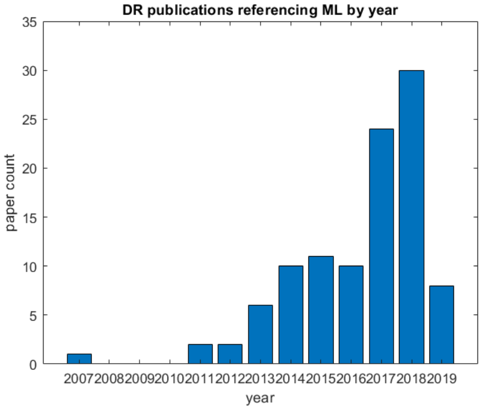

Figure 3: Publications that reference DR and ML by year according to PubMed search.

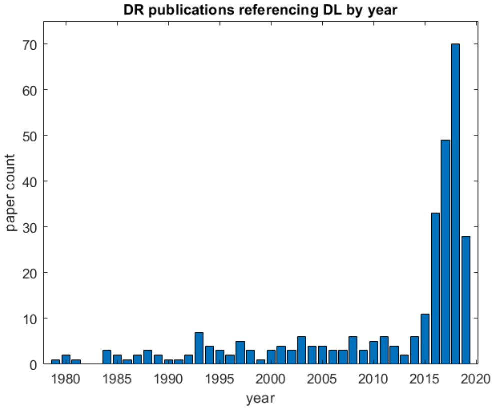

Figure 4: Publications that reference DR and DL (keyword: ”Deep”) by year according to PubMed search.

3. Overview of Imaging Modalities

3.1. Single-field imaging

3.1.1. General Single-field imaging

Single-field fundus images capture the optic nerve and macula under 20-50 degrees of FOV. Use of this method relies on the assumption that they capture the most prevalent pathologies, but this method risks missing important fundus lesions within the periphery Witmer and Kiss (2013). Traditional systems rely on single-field montages, with seven standard 30-degree images from the Early Treatment Diabetic Retinopathy Study (ETDRS) representing the standard for DR detection since 1991 ETDRS Research Group (1991). Montage images show results in DR detection and classification, but collection relies on skilled procedures that are inconvenient and time-consuming in practice.

3.1.2. Mobile Screening Systems

Despite increasing rates of DR globally, financial and practical resources impair access to treatment in developing areas Sengupta et al. (2019). Lower income areas often lack proper ophthalmic resources for active screening, therein individuals often miss adequate follow-up due to financial and travel restrictions. A portable fundus camera could mitigate this health barrier by providing a mobile screening system at low cost. Ideally, these devices would deliver high-quality images that are comparable to results using tabletop systems in terms of gradability and alignment. The Remidio Fundus on Phone (FOP) represents a current implementation that delivers smartphone-based results with a 45-degree FOV, three fields, and mydriatic or non-mydriatic modes. Further, the system offers a HIPAA-compliant, cloud-based storage system for off-site clinical viewing and ophthalmic evaluation. However, FOP still poses challenges concerning user variability and limited FOV as compared to current standards in DR screening and treatment. Thus, current studies integrate systems like FOP with web-based deep learning systems to offset user variability, specialist availability, and limitations in FOV and gradability.

3.2. Wide-field Imaging

Wide-field imaging (WFI) presents important information regarding disease manifestations in the peripheral retina Falavarjani et al. (2017). The advent of modern technologies allows field-of-view (FOV) up to 200-degrees, therefore facilitating significant advance in retinal imaging. WFI modalities include WF color imaging, autofluorescence, fluorescein angiography, and indocyanine green angiography.

WFI refer to systems with FOV capabilities above 30-degrees, based on historical standards. FOV denotes the external angle of the light source hitting the eye—a camera's FOV depends directly on lens power and indirectly on focal length. Using a confocal scanning laser ophthalmoscope, the Optos camera offers large scan angles and separable green and red laser light components (for focus on specific structures). The Optos system maximally allows a 200-degree FOV, representing eighty-two percent of the retinal surface. Despite its many advantages, the Optos system carries some limitations pertaining to ora-to-ora capabilities, imaging of anterior pathologies to the equator in all quadrants, distortion and decreased resolution of peripheral images on the two-dimensional scale, and artifacts created by anterior objects. However, current equipment addresses these issues, including combining steered-ultra field images for full retinal viewing and stereographic projection software for combat of image warp.

Retcam imaging systems maximally provide a 130-degree FOV Witmer and Kiss (2013). It consists of a portable camera system that it lit by a fiberoptic cable. Such systems yield high use with younger patients. Spectalis represents a noncontact ultra-wide-field system with a maximum 102-degree FOV.

Ultra-wide field (UWF) images present higher detection of severe non-PDR and PDR than in clinical study, with higher level classification than in ETDRS fields in ten percent of eyes. Non-mydriatic Optos UWF color images yield high agreement to mydriatic UWF color images. Further, predominantly peripheral lesions appear outside of standard ETDRS fields in over fifty percent of cases; hence lesion types that help detect DR may go undetected—these lesions include microaneurysms (MAs), hemorrhages, venous beading, and microvascular abnormalities.

3.2.1. Fluorescein Angiography

Fluorescein angiography provides ultra-wide field potential; it offers advantages surrounding diseases which require tracking and detection of peripheral neovascularization, vessel development, and/or obliterative vasculitis Falavarjani et al. (2017).

WFI initially was applied in DR screening, later Friberg proposed WF studies to predict PDR by correlating peripheral capillary nonperfusion to neovascularization (NV). Fluorescein angiography shows high potential concerning detection of microaneurysms (MAs) AND NV, to screen for initial or mild DR and PDR respectively. Ultra-wide field imaging provides 3.2 times the retinal surface coverage compared to convention seven standard fields; thus, ultra-wide-field angiography provides changes indicative of PDR with higher accuracy, including peripheral vessel leakage (PVL), NV, peripheral nonperfusion, and macular edema. Wide-field angiography facilitates superior treatment of PDR as well, via targeted laser photocoagulation.

3.2.2. Auto-Fluorescent Imaging

In auto-fluorescent imaging, 532nm (green) light excites particles and the emitted signal is detected in the 570-780nm range using a raster scan and detector Falavarjani et al. (2017). The technology scans the eye for fluorophores, which fluoresce when exposed to specific wavelengths of light. Auto-fluorescent imaging differs from fluorescein angiography, as the former records fluorescence that occurs in the eye naturally or as a by-product of disease. Natural fluorescence occurs due to optic nerve drusen, astrocytic hamartomas, lipofuscin pigments, and the aging crystalline lens. Since this imaging modality measures natural fluorescence, it does not require dye injection.

Auto-fluorescent imaging primarily centers on lipofuscin deposited in the retinal pigment epithelium (RPE). Structures like the optic nerve, retinal blood vessels, and the fovea appear dark compared to the RPE (and thus easily visualized). This is due to the absence of RPE, blocking of RPE fluorescence, or absorption of the excitation wavelengths, respectively.

3.3. Optical Coherence Tomography (OCT)

3.3.1. General OCT

Optical Coherence Topography (OCT) facilitates noninvasive examination for diabetic retinopathy with high accuracy, such that patients can be treated before irreversible vision loss Cheung and Wong (2018). OCT applies low-coherence interferometry to produce two-dimensional images of optical scattering from tissue microstructures Huang et al. (1991). It resolves structures with spatial resolutions within a micrometer range, thus detecting signals that are significantly smaller than the incident optical power. Specifically, this method measures cross-sections of tissues by applying low-coherence light and measuring the resulting reflections off tissue surfaces. Low-coherence reflectometry uses time-of-flight delay from reflection boundaries and backscattering sites to determine the longitudinal location of these sites. Multiple longitudinal scans are captured at lateral locations. OCT shows superior optical sectioning capabilities that are not limited by depth or spatial resolutions.

3.3.2. OCT-angiography (OCTA)

OCTA offers a noninvasive alternative to fluorescein angiography and has gathered substantial interest concerning detection of DR Cheung and Wong (2018). Recent trials have focused on advanced stages of DR, including leading causes of blindness such as the following: severe non-proliferative DR, PDR, diabetic macular edema (DME), and diabetic macular ischemia (DMI) without DME. In comparison to fluorescein angiography, which particularly captures changes in the superficial capillary plexus, OCTA independently captures three capillary plexuses: superficial, intermediate, and deep. Numerous OCTA metrics, such as foveal avascular zone (FAZ) size and vessel density in the different plexus (i.e deep), may influence DR severity and patient visual acuity. Therefore, many OCTA metrics may prove useful in detection and monitoring of DR. However, studies must first establish parameters to normalize these metrics such as to limit the interference of other factors like age, comorbidities, and duration of diabetes.

4. Overview of Processing Methods

4.1. Image Processing

Image processing methods address limitations in DR screening that emerge due to image quality, an imbalance in patients-to-specialists, and inter-grader inconsistencies Winder et al. (2009). Grading images for DR is performed through pattern recognition, but features such as microaneurysms, hemorrhages, and exudates can be difficult to distinguish from the background retina. Processing steps may include normalization of sharpness and illumination, contrast enhancement, and noise and artifact reduction Walter et al. (2002). Mass screening procedures and the need for consistency in the screening quality mandate use of these procedures; they may aid in human interpretation or in preprocessing stages for automatic algorithms. Additionally, image processing increases accuracy in DR tracking. For example, longitudinal registration of retinal scans facilitates more careful tracking of precise quantity and spatial changes to lesions and other relevant features.

Walter et al. (2002), Lowell et al. (2004), Mahfouz and Fahmy (2009), Youssif et al. (2008), Aquino et al. (2010), and Zheng et al. (2013) segment the Optic Disc (OD) based on the intensity and geometry of the OD and surrounding blood vessels. Common image processing strategies include morphological operations, template matching, circular hough transform, and graph cut. Traditional methods often depend on image processing to extract candidate lesions before classifying them using machine learning or rule-based systems. For example, Spencer et al. (1996) uses bilinear top-hat and matched filtering (MF) and feeds the outputs to a region-growing algorithm. Other papers transform image-based features to another space, including wavelet transform in Quellec et al. (2008a) and Abràmoff et al. (2010) and radon transform in Giancardo et al. (2011). Many such algorithms have achieved desirable results while remaining computationally efficient; however, systems that rely heavily on user specifications and simple processing methods remain functionally challenged. For instance, intensity-based algorithms may confuse structures like microaneurysms (MAs) and hemorrhages (HEMs) or bright lesions and the OD.

4.2. Traditional Machine Learning (ML) Algorithms

4.2.1. Support Vector Machine (SVM)

SVMs use pattern recognition to establish a rule to separate classes using feature expression profiles Noble (2006). For example, a line might separate DR classes by lesion presence/quantity. When many features are present, classes are separated based on maximizing distance to a high-dimensional boundary. However, this method is limited in the following respects: it assumes training and test data are derived from similar distributions, requires user-specified parameters, and may overfit with too many input variables.

Welikala et al. (2014) and Goatman et al. (2011) applied SVMs to detect blood vessels and identify cases of PDR. Akram et al. (2013) detects microaneurysms (MAs) by feeding image priors into a hybrid ensemble based on SVM, Gaussian Mixture Models (GMM), and multi-model mediod based modeling. SVM classified MA turnover Xu et al. (2018) and measured changes in red lesions to grade DR Adal et al. (2018). While this method allows complex feature representations and efficient class separation, it requires prior feature extraction and still represents a degree of user manipulation that may breed errors in complex tasks.

4.2.2. K-Nearest Neighbors (KNN)

KNN classifies objects in unseen images by distance to “K” training samples in the feature space Keller et al. (1985). Traditional KNN classifiers may suffer in cases of class imbalance due to equal weighting between labeled samples. Classification may be uncertain during ties between classes, which may be solved in “crisp” (traditional) KNN by reducing K or by minimizing sum of distances to each neighbor with a tied class. Fuzzy-KNN addresses uncertainty by assigning an input vector to multiple classes; it decides based on strongest membership.

Niemeijer et al. (2005) and Abramoff and Suttorp-Schulten (2005) apply KNN to detect red lesions for web-based DR screening. In Niemeijer et al. (2007a), KNN receives filter responses from candidate lesions and identifies exudates and cotton wool spots (CWSs). Tang et al. (2013) segments retinal images into splat partitions, selects the optimal features using a filter and wrapper sequence, then classifies candidate hemorrhages using KNN. Niemeijer et al. (2009) combined KNN with likelihood distribution normalization. KNN serves as a nonspecific classifier that works well across tasks; however, it suffers from potential issues in computational efficiency and generalizability. Thus, diverse patient populations would mandate (1) storing very large training data or (2) training on new data across different population uses.

4.2.3. Random Forest (RF)

Random forest combines decision trees in a learning ensemble, such that accuracy combines individual trees and the correlation between trees Breiman (2001). Tree ensembles are grown through random subsamples of the training set (i.e. through bagging and random split selection). A number of random attributes is selected at each node; the algorithm computes entropy of each to select the optimum attribute to classify remaining training examples Sanromà et al. (2016). Independent trees use these decisions to vote for class label for later aggregation.

Acharya et al. (2017) used a series of radon, discrete wavelet, and discrete cosine transforms to extract features from retinal images. They found that a decision tree classified DME with the highest performance based on MESSIDOR data. More complicated classification systems use random forests to make decisions; for example, Casanova et al. (2014) and Sanromà et al. (2016) detect binary DR based on RFs. RFs may also separate lesions from non-lesions, which is accomplished using candidate morphology in Zhang et al. (2014). This method can provide decisive class separations; however, it requires hand-tuning and may suffer from difficulties concerning less-clear class distinctions (i.e. mild from moderate).

4.2.4. Neural Networks (NNs)

NNs learn mathematical weights that specify the probabilities of an inputs belonging to certain output classes Gardner et al. (1996a). To accomplish this, the NNs recognize significant features in determining output class (i.e. DR severity) in input samples of known class. Some handcrafting is necessary to set parameters. Alternately, NNs may be used to extract features and fed them to a subsequent classifier.

An Artificial NN (ANN) can extract image features to identify lesion candidates, which will be labeled by a clinical rule-set or subsequent classifier. Alternatively, ANNs may predict lesions or disease labels based on input features. In Gardner et al. (1996a) and Usher et al., ANNs differentiate signs of DR like lesions and new vessels from background pixels for subsequent DR detection. The algorithm in Hatanaka et al. (2012) uses rule-based and ANN methods to predict microaneurysms based on double-ring filter responses. Marin et al. (2011) computes a seven-dimensional (7D) vector of intensity and moment invariant features. Then, it feeds this information to a NN to label pixels as blood vessel or non-blood vessel. In Osareh et al. (2009), an algorithm detects exudates by extracting candidate features using fuzzy c-means clustering, ranking feature importance under a genetic-based algorithm, and feeding the features into a multilayer NN to output the final labels. Usher et al. classifies lesions at the pixel-level using wavelet signatures and an ANN, then it feeds these results through rule-based stages at the region, area, and image levels. Herliana et al. (2018) extracted features from retinal images using particle swarm optimization (PSO), then they fed them to a NN to classify DR.

4.3. Deep Learning (DL)

DL selects increasingly higher-level features to map input data to output labels, using training data to weigh features relevance Litjens et al. (2017). The most successful DL methods are constructed using deep convolutional neural networks (CNNs), which rose to prominence with AlexNet Krizhevsky et al. (2012). CNNs use convolution filters to transform inputs and allow weights to be shared spatially. DL algorithms require less user specification than traditional ML systems, but it can be unclear as to which features influence their classifications.

Zilly et al. (2015) segments the optic disc (OD) using entropy-based sampling and a boosted CNN. The algorithms in Yu et al. (2017), Haloi (2015), Chudzik et al. (2018b), Chudzik et al. (2018a), Hatanaka et al. (2018), Dai et al. (2018), and Tan et al. (2017) use CNNs to detect exudates, MAs, and HEMs. Yu et al. (2017) focuses on an off-line classifiers to enhance computational efficiency while classifying pixels, whereas Chudzik et al. (2018b) and Chudzik et al. (2018a) stress the importance of transfer learning and layer freezing. Transfer learning allows CNNs to learn lower-level features from public datasets, thus compensating for limited data. Hatanaka et al. (2018) uses a two-step DCNN that detects MAs and filters false positives. Dai et al. (2018) compensates for weakly-supervised images by natural language processing (NLP). During training, the multi-sieving CNN relates the supervised information in clinical reports to lesion locations in retinal images. Then, the CNN uses these learned associations to detect lesions in test images, without the clinical text. Gargeya and Leng (2017) harnesses a CNN to extract features from images, then it feeds them to a tree-based model that classifies binary DR. Deep learning enhances learning by extracting higher-level features that may be missed by handcrafted systems; however, deeper systems often suffer from limitations in interpretability that may lead to clinical unacceptance. Therefore, Quellec et al. (2017) generates heatmaps during DR grading. This system involves a backpropagation method that includes three passes through the CNN to pass second-derivatives forward. The algorithm results in image and lesion-wise predictions. A heatmap represents all relevant pixels. Gulshan et al. (2016a) use inception-v3 to grade images for referable DR, DME, and image quality. Ting et al. (2017) extend this approach to multiple diseased types, applying VGGNet to grade for DR, glaucoma, and AMD. Zhou et al. (2018), Zhang et al. (2017), Ardiyanto et al. (2017), and Roy et al. (2017) also use deep learning to grade severity. Zhang et al. (2017) and Ardiyanto et al. (2017) apply Residual Networks (ResNets) to the issue. Yang et al. (2017) detects lesions and grades DR using a two-stage DCNN with imbalanced weighting. Wang et al. (2017) introduces a novel framework, Zoom-in-Net, to highlight attention maps of lesion patches and diagnose DR. ElTanboly et al. (2018) graded OCT images for NPDR using a joint segmentation model and CNN trained by stacked non-negatively constraint autoencoders. Rasti et al. (2018) detects DME and age-related macular degeneration (AMD) using a multi-scale convolutional mixture of experts (MCME) model. Shanthi and Sabeenian (2019) modified the AlexNet model to grade retinal images for DR severity. Wan et al. (2018) compared popular deep CNN frameworks (AlexNet, VGGNet, GoogleNet, and ResNet) for grading DR in retinal images.

5. Overview of Tasks

5.1. Optic Disk (OD) Segmentation

OD removal improves bright lesion detection due to the maximum intensity values that frequently occur within this region Lowell et al.

(2004). Also, the OD rim may impact blood vessel detection. Alternatively, OD localization may contribute to blood vessel segmentation, as the major retinal vessels branch off from the OD. Lastly, the OD might be located to establish distance to the fovea. Maculopathy, or retinopathy of the macula, is particularly damaging due to the high visual importance of this region. Thus, severity assessment can benefit from accurate localization of the macula, particularly when it is obscured by exudates. Interestingly, the surge in deep learning methods in the DR space has correlated to decreased use of preprocessing methods like OD segmentation and removal. Ideally, this would indicate that the less handcrafted systems can learn features that can distinguish the OD from other structures; however, the "black box" nature of most deep learning systems causes it to be difficult to distinguish whether an algorithm is learning features and feature weights correctly (i.e. OD versus lesions). This can be particularly troublesome due to the natural variety of the OD across patients. For instance, a model that learns from patients who have darker OD appearances may mistake a lighter OD as a lesion with greater ease. Ting et al. (2017) point out that they did not validate whether OD contour affected their results. On the other hand, removing the OD prevents an algorithm from concurrently learning other important eye diseases (i.e. maculopathy, glaucoma). Therefore, OD segmentation is important to DR lesion detection and severity grading, whereas its removal depends on the intended task.

Researchers have implemented machine learning in OD segmentation with high generalizability, accuracy, and computational efficiency. Mahfouz and Fahmy (2009) projected the image space into two orthogonal axes to localize the OD faster. Zheng et al. (2013) improved OD and OC segmentation by incorporating an extended general energy function into global optimization based on s-t graph cut. Then, a GMM determined the posterior probability of each pixel belonging to each section. Abramoff and Niemeijer (2006) applied a circular template and KNN regression to localize the OD. The algorithm in Niemeijer et al. (2004) supplied the vessel segmentations. Rehman et al. (2019) proposed a multi-parametric scheme for segmenting the OD using region-based statistics and texture. After preprocessing, Simple Linear Iterative Clustering (SLIC) segmented images into superpixels. Then, the algorithm discriminated important regional features according to mutual information (MI). Finally, it fed all features into one of four classifiers (SVM, RF, AdaBoost, or RusBoost). Zilly et al. (2015) segmented the OC and OD using entropy-based downsampling and a boosted CNN. A logistic regression classifier generated the final pixel probability maps using convolutional outputs, L*a*b color values, and centricity. An unsupervised graph cut smoothed the probability map, then convex hull transform connected disjointed regions.

Other methods apply image processing to detect the OD using features like high pixel intensity, blood vessel convergence, and known shape. These methods rely on more handcrafted features than necessary in deep learning; however, they are often faster and more easily interpreted. Aquino et al. (2010) segmented the OD using the weighted performance of median-filtering of the green channel, Gaussian low-pass filtering, and Otsu thresholding of the blue channel. Then, the algorithm removed blood vessels and extracted the OD boundary using morphological operators, Prewitt edge detection, and Circular Hough Transform. Youssif et al. (2008) combined orientation-based filters and maximum pixel responses to generate a vessel directional map (VDM). The minimum accumulated difference in vessel directions represented the OD center. Uribe-Valencia and Martínez-Carballido (2019) modeled column variations in pixel intensity to localize the OD. The algorithm detected promising OD regions in a top-down approach. A size and weighted centroid approach narrowed candidates. In Kar and Maity (2018) and Akram et al. (2013), morphological operators extracted and smoothed the OD, respectively. Kar and Maity (2018) and Zhang et al. (2014) incorporated vessel structures. Contrast in intensities also improved OD localization as in Zhang et al. (2014) and Usher et al.. Morphological operators were sometimes combined with other methods, as with region growing and local phase symmetry in Yu et al. (2017).

5.2. Blood Vessel (BV) Segmentation

Segmenting BVs is essential to DR screening for reducing false positives when detecting dark lesions and for tracking microvascular changes when detecting PDR. In the former, false positives could arise due to similar pixel intensities between BVs and these lesion types, as well as the proximity of most lesions to BVs. For the later, PDR is characterized by angiogenesis, or growth of, new retinal BVs that lack structural integrity. These vessels cannot support regular function, and they are more likely to rupture or otherwise cause irreversible damage.

Traditional machine learning has been highly effective for segmenting BVs. Orlando and Blaschko (2014) harnessed a discriminatively trained, fully-connected conditional random field (CRF). A 1-slack SOSVM with margin-rescaling learned the resulting pairwise potentials, as well as unary and bias features. Cai and Chung (2006) proposed an unsupervised method using a Gaussian pyramid, gradient matrix, the normalized cut criterion, and a tracking strategy. Nguyen et al. (2011) labeled pixels using Bayesian classification with bagging. The system classified BVs according to Gabor filter responses and pixel-wise gray levels. Welikala et al. (2014) used standard and modified operators to distinguish new BVs from old BVs and non-vessel edges, respectively. Then, a linear SVM integrated features from both binary vessel maps to make PDR classifications. Goatman et al. (2011) segmented BV candidates using watershed transform and ridge strength measurement. The Wilcoxon rank sum and Ansari-Bradley tests weighed feature relevance and fed the results to a radial-basis function (RBF) SVM to assess PDR. Marin et al. (2011) detected BVs using a four-layer feedforward NN to classify pixel. Xu et al. (2009) applied a boosting algorithm to 3D-OCT images to segment BVs by classifying pixels, without segmenting retinal layers. 0th - 2nd order Gaussian filter responses generated 2D features from A-scans; gradient magnitude and divergence were also added. The model incorporated 3D Haar-features based on optimization parameters and the 0th - 4th moments of the filter responses. The LogiBoost adaptive boosting algorithm trained on hand-labeled A-scans.

Paper	**Focus**	**Method**	**Image features**	**Results**
Mahfouz and Fahmy (2009)	OD localization	dimensionality reduction (1D vectors) and edge detection	retinal vessels orientation & OD brightness/shape	97% accuracy on 340 images from STARE/ DRIVE) and a processing time of 0.67s/image
Abramoff and Niemeijer (2006)	OD localization	circular template, KNN regression	number of vessels and vessel pixels, orientation of the widest vessels, average image intensity, average vessel width	0.999 accuracy on 1000 images
Uribe-Valencia and Martínez-Carballido (2019)	OD localization	local multi-level thresholds, best channel selection, size and weighted centroids, column-wise Intensity Profile Model	high pixel intensities and blood vessel geometry, Left Lobe (LL), Right Lobe (RL), intermediary Vasculature, and appropriate OD diameter	0.997 accuracy on 341 images
Youssif et al. (2008)	OD localization	morphological operators, Otsu thresholds, Gaussian filtering	directional pattern of blood vessels – minimum difference in vessel directions	0.9877 accuracy on STARE and 1.000 accuracy on DRIVE
Zilly et al. (2015)	OD segmentation	entropy downsampling, boosting CNN, logistic regression and unsupervised graph cut algorithm	centricity, brightness	F-score of 94.7 and overlap measure with manual segmentation of 89.7 on DRISHTI-GS
Aquino et al. (2010)	OD segmentation	weighted average: median filtering, Gaussian low-pass filtering, Otsu thresholds, morphological and Prewitt edge detection	pixel intensities, removed blood vessels	overlapping measure of 0.86
Zheng et al. (2013)	OD segmentation	template-matching, extended energy function into s-t graph cut, Gaussian Mixture Modeling	priors like topological information (i.e. circularity), adaptive priors (i.e. minimum rim thickness)	disc area agreement with OCT machine software between 0.86-0.88
Rehman et al. (2019)	OD segmentation	multi-parametric scheme – Simple Linear Iterative Clustering (SLIC); mutual information criterion; Texton-map histogram and Gabor filter bank; multilevel Otsu thresholds; SVM, RF, AdaBoost, RusBoost	region-based statistical and texture features	accuracies of 0.993, 0.988 and 0.993 on DRIONS, MESSIDOR and ONHSD

Table 2: Sample of methods used to localize, segment, and extract the OD in retinal images

Deep learning facilitates automatic, multi-scale learning of BVs by combining local and global image features. Many methods harness U-Net-based structures to learn from multiple receptive fields, as BVs appear at many complex patterns, shapes, and sizes. This section introduces unique CNN structures to perform this task. Jiang et al. (2019) proposed a new deep CNN (DCNN) structure, D-NET, with a multi-scale information fusion module (MSIF) during encoding. The MSIF combined parallel convolution layers with different dilation rates. The network's decoding stage incorporated skip connections. Soomro et al. (2019) approached this problem through a strided-CNN. Preprocessing improved image contrast using morphological mappings and PCA. Then, the strided-CNN passed these contrast images through encoder/decoder stages with residual connections, wherein the model replaced traditional pooling layers with strided convolutional layers. Jebaseeli et al. (2019) applied a Tandem Pulse Coupled NN (TPCNN) and DL-Based SVM (DLBSVM). The model fused data from two source images and simultaneously operated on the inter- and intra-channel linkings of the input neurons. A DLBSVM model imputed these features and classified vessels on a pixel-wise basis. Jin et al. (2019) introduced a Deformable U-Net framework, DUNet, which replaced some convolutional layers with deformable blocks. Feng et al. (2019) proposed a cross-connected CNN (CcNet). CcNet fused multi-scale features through a U-net framework with two modifications: (1) equally-sized feature maps and (2) convolution-ReLu-maxpooling (CRM) modules that linked each layer of the primary path to all layers within the secondary path. Noh et al. (2019) proposed a scale-space approximated (SSA) CNN (SSANet). SSANet combated potential aliasing by replacing decimation with upsampling by bilinear interpolation. The model adapted the DRIU network as a base Sengupta et al. (2019), added residual blocks, and combined SSA and traditional decimation.

Image processing remains prevalent in segmenting BVs. This is particularly true in papers that focus on lesion detection and require rapid BV removal. Zhang et al. (2014) and Yu et al. (2017) harnessed morphological operators followed by inpainting, whereas Hipwell et al. used linear thresholding and Niemeijer et al. (2005) performed connected component analysis. Hatanaka et al. (2012) subtracted BVs from candidate microaneurysms by combining double-ring filtering and black-top-hat transform. Aquino et al. (2010) used a linear structuring element based on rotational variance and maximum gray-levels. Shahid and Taj (2018) combined filtering operations and Otsu thresholding along a vessel location map (VLM) and Frangi binarized image. Difficulties in these methods largely stem from the geographical and intensity similarities between vessels and dark lesions; in fact, Hipwell et al. aims to remove features like hemorrhages. Such a procedure could hence subtract from classification accuracy.

Paper	**Focus**	**Method**	**Image features**	**Results**
Orlando and Blaschko (2014)	BV segmentation	conditional random fields (CRFs), 1-slack structured output support vector machine (SOSVM) of pairwise potentials	unary and bias features like 2-D Gabor wavelets, line detectors, and vessel enhancement	sensitivity of 0.776 and specificity of 0.973
Cai and Chung (2006)	BV segmentation	Gaussian pyramid, eigen-decomposition of the gradient matrix of the Lucas-Kanade equation, normalized cut criterion	vessel width, intensities	distance from ideal point of 0.1823 (ROC figure provided in paper)
Nguyen et al. (2011)	BV segmentation and extraction	Bayesian classification with bagging, Gabor filter responses at different scales/orientations	pixel-wise average local gray levels were subtracted from maximum gray levels	The AUC obtained on DRIVE was 0.9491
Welikala et al. (2014)	BV detection for PDR	line operators, linear SVM	vessel edges	sensitivity of 1.00 and specificity of 0.90 on MESSIDOR
Goatman et al. (2011)	BV detection on the OD for PDR	watershed transform and ridge strength measurement, Wilcoxon rank sum and Ansari-Bradley tests, radial-basis SVM	vessel intensities	Leave-one-out cross-validation: AUC of 0.911, a sensitivity of 0.842, and a specificity of 0.859
Marin et al. (2011)	BV detection for subsequent DR classification	four-layer feedforward NN	gray-level and moment invariants-based features	average overall accuracy of 0.9489 and AUC of 0.9678 for DR on DRIVE and STARE
Shahid and Taj (2018)	BV segmentation	multi-scale Frangi filtering, morphological top hat, lowpass, Otsu thresholding, hysteresis threholding	vessel location map (VLM), Frangi binarized image	accuracy, sensitivity, and specificity above 0.95, 0.73, and 0.95 on STARE and DRIVE
Jiang et al. (2019)	BV segmentation	DCNN structure (D-Net) with a multi-scale information fusion module (MSIF) in the encoding stage	multi-resolution image and hierarchical information	0.9864 AUC, 0.9709 accuracy, 0.8246 F-score, 0.7839 sensitivity, 0.9890 specificity, 1.5s/image
Soomro et al. (2019)	BV segmentation	morphological operators, PCA, strided-CNN, class balancing loss function	image-wide	sensitivity of 0.87, 0.808, 0.886 and 0.829 on DRIVE, STARE, CHASE_DB1 and HRF; accuracies of 0.956, 0.954, 0.976 and 0.962
Jebaseeli et al. (2019)	BV segmentation	Tandem Pulse Coupled Neural Network (TPCNN), Deep learning based SVM (DLBSVM)	unsupervised TPCNN feature vectors	0.7445 sensitivity, 0.9940 specificity, and 0.9916 accuracy
Jin et al. (2019)	BV segmentation	deformable U-Net framework (DUNet)	fundus image inputs	accuracy of 0.9566, 0.9641, 0.9610, and 0.9651 and AUC of 0.9802, 0.9832, 0.9804, and 0.9831 on DRIVE, STARE, CHASE_DB1 and HRF respectively
Feng et al. (2019)	BV segmentation	cross-connected CNN (CCNet)	fundus image inputs	sensitivity of 0.7625 and 0.7709 and accuracy of 0.9528 and 0.9633 on DRIVE and STARE, respectively
Noh et al. (2019)	BV segmentation	scale-space approximate CNN (SSANet)	fundus image inputs	AUC of 0.9916 on CHASE_DB1
Xu et al. (2009)	BV segmentation in 3D-OCT	Gaussian filtering, Haar responses, LogiBoost adaptive boosting algorithm	2D and 3D features from A-scan projections, magnitude, divergence	sensitivity of 0.85 and specificity of 0.88

Table 3: Sample of methods used to segment blood vessels

5.3. Lesion Detection

Lesion detection are essential to detecting and grading for DR. These algorithms help generate comprehensive lesion maps; thus, they enhance system credibility by allowing clinicians to observe what image regions factor into a DR grade. Further, different studies also focus on different lesions depending on their purposes. For instance, most algorithms focus on detecting microaneurysms (MAs) to aid ophthalmologists in capturing the disease before preventable vision loss. Other methods also track lesions like exudates and hemorrhages to aid in grading DR severity. Image registration allows lesion prevalence and locations to be tracked across visits as a measure of DR progression. Lesion tracking enables key advantages over end-to-end learning DR grading. Algorithms that learn directly from images in an end-to-end manner learn without user input operate like "black boxes", which means that the user may not know what features the systems are learning. Therefore, a seemingly accurate system may learn incorrect information during training and fail when provided unseen images. Further, particular lesions may provide clinicians with finer information concerning what procedures the patient might need. In fact, works such as Li et al. (2019) show accuracies for DR that surpass 0.80, despite poor identifications of lesions (i.e. precision below 0.05 for MAs). This suggests that lesion detection may be more challenging than system-wide DR detection. At the same time, systems that diagnose DR based on lesion presence may suffer from inaccuracies due to interpatient variability. Thus, superior systems combine global image information with lesion heatmaps for precise and comprehensive diagnosis and treatment. Keel et al. (2019) shows how this information offers insight into DR and how algorithms make predictions.

5.3.1. Microaneurysm (MA) Detection

Quellec et al. (2008a) matched a MA template in sub-bands of translation invariant wavelet-transformed images. A lifting scheme selected adaptive wavelets through a genetic algorithm followed by Powell's direction set descent. Akram et al. (2013) built a learning ensemble using GMMs, SVMs, and multivariate m-Mediods with Semi-Fuzzy Self-Organizing Maps (SFSOMs). Supervised Fisher discriminant analysis (LFDA) maximized class discrimination. A genetic algorithm learned relative classifier weights. Giancardo et al. (2011) applied radon transform and a RBF SVM. Dashtbozorg et al. (2018) extracted local convergence index features (LCFs) using multi-scale multi-orientation gradient weighting and iterative thresholding. RUSBoost imputed the LCFs to identify MAs. Karim et al. (2019) employed the MATLAB NN Pattern Recognition Tool (NPRTOOL). The algorithm segmented images using K-means clustering, removed the OD and BVs using a canny edge detector, and classified MAs using a NN. Selcuk and Alkan (2019) applied an ant colony algorithm. Colony filters optimized a "pheromone matrix" for each ant's ideal path and aggregated results. Niemeijer et al. (2005) used KNNs to generate pixel probability maps and classify images for red lesions.

Many works highlighted dimensionality reduction while identifying MAs. Zhou et al. (2017) harnessed sparse lasso-based PCA and multi-scale Gaussian correlation coefficients. Hatanaka et al. (2012) applied PCA, a rule-based method, and a three-layer feed-forward ANN to classify lesions. Deepa et al. (2019) harnessed discrete orthonormal Stockwell transform (DOST) and local binary pattern (LBP) histograms. The

DOST phase reduced features through Haralick feature extraction. An ANN classified the results using the concatenated features. Cao et al. (2018) ranked features by PCA and RF feature importance. The raw pixels and two reduced-feature sets were fed into three classifiers (25-tree RF, two-layer NN, and RBF SVM).

Chudzik et al. (2018a) reduced the training time for a fully convolutional NN (FCNN) with interleaved freezing. The CNN sampled annotated patches across all possible image locations. It classified pixels by the ratio of positive-to-total predictions. Haloi (2015) structured a five-layer DCNN to generate a pixel-wise probability map. Layers incorporated dropout and maxout activation. Chudzik et al. (2018b) pre-initialized a FCNN by transfer learning. Overall, the CNN included: encoder/decoder stages, eighteen alternating convolutional and batch normalization layers, three max-pooling layers, three upsampling layers, and four inter-path connections. Hatanaka et al. (2018) proposed a two-step Deep CNN based on GoogLeNet to detect MAs and three-layer, 48 feature perceptron to reduce false positives. Processing included a double-ring filter, a Gabor filter, and shape index based on the Hessian matrix. Dai et al. (2018) developed a cascaded multi-sieving CNN (MS-CNN) that utilized supervised information from clinical reports to identify MAs in images. The hybrid mining system mapped images to text to learn lesion feature subspaces. The trained model used these features to segment images without the corresponding clinical notes. Each layer fed false positives forward as negative examples.

Spencer et al. (1996) detected MAs in FAs using thresholding and region growing. Hipwell et al. detected MAs in 50-degree digital red-free fundus photographs. The algorithm generated a rule set based on thirteen clinical features like MA shape and intensity. Hatanaka et al. (2012) improved their previous method by adding double-ring filtering and PCA. Haloi (2015) improved DCNN-based classification by foveation and nonuniform sampling. Chudzik et al. (2018b) applied Otsu thresholding and morphological operators to the green channel of color fundus images to remove false negatives.

5.3.2. *Exudates (Exu)*

Kar and Maity (2018) maximized the MI of the responses to Laplacian of Gaussian (LoG) filters and MF to segment exudates. This method approximated bright lesions as edges with a zero crossing about the LoG center. Dark lesions were discerned using a MF kernel and a negative LoG kernel. A differential evolution (DE) algorithm found the fuzzy parameters that maximized MI. Zhang et al. (2014) subtracted adaptive templates containing pixel maxima from images to yield bright structures. Exudate candidates were found using MF and morphological top-hat filtering. Finally, 500-tree RF algorithm classified exudates and DR. Osareh et al. (2009) used a color-based method based on Gaussian-smoothed histogram analysis and fuzzy c-means (FCM) clustering. Coarse segmentation analyzed histogram analysis, then FCM assigned unclassified pixels. A genetic algorithm ranked features for optimal classification using a one-point crossover operator. Brandon and Hoover (2003) designed a multi-level method to tune progressively higher levels of drusen classification. Wavelet-based equalization transformed the image under an orthonormal wavelet basis with a high-order coiflet. The method classified pixels according to similarity measures between the product of four 1D wavelet responses and the Mexican hat wavelet. A feed forward NN imputed the wavelet signatures to classify the results. Finally, a rule system decided if connected pixels formed lesions. Khojasteh et al. (2019) compared a basic CNN, a pre-trained residual network, ResNet-50, and Discriminative Restricted Boltzmann Machines (DRBMs) for detecting exudates. The optimal system replaced the softmax layer of ResNet-50 with SVM and learned from PCA-reduced data inputs.

Lokuarachchi et al. (2019) and Wisaeng and Sa-Ngiamvibool (2019) detected exudates through image processing. Lokuarachchi et al. (2019) applied intensity thresholding, contrast limited adaptive histogram equalization (CLAHE), and mean filtering. The method subtracted the binary mask of the OD from lesion candidates. Wisaeng and Sa-Ngiamvibool (2019) applied a morphology mean shift algorithm (MSA) to coarsely segment candidates. This method used a kernel density estimator (KDE) with a Gaussian kernel function. Then, a mathematical morphology algorithm (MMA) finely segmented these results.

5.3.3. Multiple Lesion Types (MLT)

Lee et al. (2001) empirically trained a rule system to grade images for lesion presence. Usher et al. extracted lesions using recursive region growing and adaptive intensity thresholding; dark lesions also required a ”moat operator” Sinthanayothin et al.. An ANN classified lesions and graded images for lesion presence. Niemeijer et al. (2007b) differentiated hard exudates, CWSs, and drusen. A KNN classified pixels by thresholded filter responses from a set of second-order irreducible invariants. A second KNN suppressed spurious bright lesions based on features like contrast, size, shape, and proximity to BVs. Finally, linear discriminant analysis (LDA) separated lesion types. He et al. (2019) harnessed the spatial-spectral features of multispectral imaging (MSI). The authors first extracted LBPs across MSI channels. A generalized low-rank approximation of matrices (GLRAM) framework approximated the resulting matrix as low-rank under a supervised regularization term (SRT). Finally, a Gaussian-based SVM classified pixels.

Tan et al. (2017) distinguished exudates, HEMs, and MAs using a 9-layer CNN. The CNN consisted of five convolutional layers, max-pooling in layer six, a 256-neuron fully-connected seventh layer and 128-neuron fully-connected eighth layer, and a 4-neuron output layer. Li et al. (2019) segmented lesions in their new dataset using two semantic segmentation models, DeepLab-v3+ Chen et al. (2018) and HED Xie and Tu (2015). Keel et al. (2019) adapted two validated deep learning models for referable DR Keel et al. (2018) and glaucomatous optic neuropathy (GON) Li et al. (2018) to visualize CNN outputs and enhance model interpretability. The algorithm produced feature maps using adaptive kernel visualization, sliding windows, and thresholding of class probabilities. Saha et al. (2019) proposed a DCNN, SegNet, for semantic segmentation to detect lesions and the OD. SegNet included an initial 13-layer encoder stage based on VGGNet, a corresponding decoder stage, and a final pixel classifier. Each decoder stage harnessed memorized max-pooling indices from its respective encoder feature map. Then, the method convolved each sparse representation with a trainable decoder filter bank to produce dense feature maps. Costa et al. (2019) developed EyeWeS, a weakly supervised CNN, using multiple instance learning (MIL) and transfer learning approaches. This framework identified lesion-containing regions based on bagging and image labels. First, a pre-trained inception-v3 constrained the model's receptive field. Subsequent layers used 1 x 1 convolutions to classify pixels. Image labels were constructed by max-pooling with patch outcomes. Son et al. (2019) screened and produced lesion heatmaps for 12 major retinal abnormalities. The overall system included 12 regionally-guided CNNs that each binarily classified one abnormality. An overall lesion heatmap highlighted pixels that correlated to target findings. Each CNN consisted of residual layers, reduction layers, average pooling layers, an Atrous Pyramid pooling layer, and a final 1 x 1 convolutional layer.

5.4. DR Grading

5.4.1. Screening Systems

Abramoff and Suttorp-Schulten (2005) developed EyeCheck, an ATA category 2 web-based screening tool for DR. Measurement parameters consisted of image gradability, inter-grader agreement, need for pupil dilation, and the risk and urgency level of DR. Risk factors included neovascularization, higher age, diabetes type I, and bright and red lesions. This work was furthered by Abràmoff et al. (2008), who screened for DR by combining previously published algorithms Abramoff and Niemeijer (2006); Niemeijer et al. (2004, 2005, 2007b, 2006). First, a statistical classifier graded images for quality based on filterbank responses and RGB hisogrmas Niemeijer et al. (2006). Next, a supervised classifier segmented BVs to reduce false-positive red lesions Niemeijer et al. (2004). A KNN regression model localized the OD using prior vessel segmentation Abramoff and Niemeijer (2006). Finally, two algorithms detected red lesions and other lesions Niemeijer et al. (2005, 2007b). Abràmoff et al. (2013) assessed the Iowa Detection Program (IDP) for detecting referable DR (RDR) as compared to an adjudicated International Clinical Diabetic Retinopathy (ICDR) severity standard. American Academy of Ophthalmology (2014). IDP combined components for assessing image quality Niemeijer et al. (2006), detecting MAs and HEMs Niemeijer et al. (2005); Quellec et al. (2008a), detecting exudates and CWSs Niemeijer et al. (2007b), and detecting irregular lesions Abràmoff et al. (2008); Tang et al. (2013). Abràmoff et al. (2016) compared IDx-DR X21 to their previous system, IDP. IDx-DR X21 applies CNN detectors to detect lesions characteristic for DR. These features served as input to two Rfs to classify RDR and vision-threatening DR. The algorithm surpassed IDP due to its deep learning inclusion, improved lesion detection, and high diagnostic accuracy in DR and DME.

Keel et al. (2018) constructed their system, EyeGrader, of four inception-v3 models for classifying RDR, DME, image quality for DR, and image quality for DME. The study trained each model with 66,790 images from Healgoo Interactive (2017). González-Gonzalo et al. (2019) evaluated RetCAD v.1.3.0, for joint detection of DR and AMD. RetCAD v.1.3.0 included an initial preprocessing stage, image quality assessment, and two ensemble CNN architectures. The overall system aggregated disease-wise probabilities from each CNN. Quellec et al. (2019) introduced OphtAI, a system of CNN-ensembles that recognized eye laterality, detected RDR, assessed DR severity, and generated decision heatmaps. The work investigated several CNN frameworks: Inception-v3, Inception-v4, VGG-16, VGG-19, ResNet-50, ResNet-101, ResNet-152, and NASNet-A.

Paper	**Focus**	**Method**	**Image features**	**Results**
Spencer et al. (1996)	MAs in fluorescein angiograms	Bilinear top-hat transform, matched filtering, thresholding, region-growing	pixel intensities	a sensitivity of 0.82
Hipwell et al.	MAs in 50-degree digital red-free fundus photographs	rule-based	thirteen clinically specified features like MA shape and intensity	image-level sensitivity of 0.81 and specificity of 0.93 for MA presence
Hatanaka et al. (2012)	MAs	Double-ring filtering, principal component analysis (PCA)	126 image features, included texture	0.68 true positive rate on a lesion basis
Akram et al. (2013)	MAs	morphological opening, Gabor filter banks, learning ensemble: Gaussian mixture model (GMM), SVM, and multivariate m-Mediods with Semi-Fuzzy Self-Organizing Maps (SFSOMs), supervised Fisher discriminant analysis (LFDA), genetic algorithm	features like shape, color, intensity and statistics	image-level sensitivity of 0.9864, specificity of 0.9969, and accuracy of 0.9940 on DIARETDB0 and DIARETDB1
Haloi (2015)	MAs	DCNNs, foveation, nonuniform sampling	Intensities, convexity, and region area	AUC of 0.98 on the ROC dataset
Quellec et al. (2008a)	MAs	translation-invariant wavelet transform, genetic algorithm, Powell's direction set descent	Wavelet sub-bands	sensitivity of 0.8962 on a lesion basis in 120 expert-labeled images
Giancardo et al. (2011)	MAs	Radon transform, region growing, radial basis function SVM	Radon transformed features (modeled as Gaussian)	global score of 0.375 and sensitivity of 0.366 on ROC. AUC of 0.96 with PCA
Zhou et al. (2017)	MAs	sparse lasso-based PCA, multi-scale Gaussian correlation coefficients, false positives reduced by thresholding and region growing	thirty-four features (i.e. Gaussian filtering, shape, brightness, proximity to vasculature)	sensitivity of 0.4831
Hatanaka et al. (2012)	MAs	PCA, rule-based, three-layer feed-forward ANN	PCA-selected	true positive rate of 0.68 and 15 false positives/image on ROC
Chudzik et al. (2018a)	MAs	fully convolutional NN (FCNN) with interleaved freezing and dice loss	image patches	free-response receiver operatic characteristic (FROC) of 0.298 on ROC and 0.431 on E-Ophtha

Haloi (2015)	MAs	Five-layer DCNN with dropout and maxout	pixel-wise probabilistic maps	AUC of 0.988 on MESSIDOR and 0.98 on ROC
Chudzik et al. (2018b)	MAs	FCNN with batch normalization and dice loss	pre-annotated image patch analogues	FROC of 0.562 on E-Ophtha, 0.392 on DIARETDB1, and 0.355 on ROC
Hatanaka et al. (2018)	MAs	Two-step DCNN; three-layer, 48 feature perceptron; double-ring filter, Gabor filter, Hessian matrix; GoogLeNet	BVs extracted, green channel inputs	sensitivity of 0.84 at 8.0 false positives/image on DIARETDB1
Dai et al. (2018)	MAs	Cascaded Multi-sieving CNN (MS-CNN), natural language processing (NLP)	Hybrid text/ image features	a recall of .878, precision of .997, accuracy of .961, and F-score of .934
Deepa et al. (2019)	MAs	discrete orthonormal Stockwell transform (DOST) with short-time Fourier transform (SSFT) and wavelet transform (WT), Haralick feature extraction; Local Binary Pattern (LBP) histograms, ANN	statistical and textural features, skewness and kurtosis from the LBP	sensitivity of 0.965 and specificity of 0.953
Cao et al. (2018)	MAs	PCA, RF, NN, Radial basis function (RBF) SVM	PCA and RF feature importance reduced features	AUC of 0.985 and F-measure of 0.926 on DIARETDB1
Dashtbozorg et al. (2018)	MAs	multi-scale multi-orientation gradient weighting, iterative thresholding, gradient weighting, RUSBoost	Local convergence index features (LCF), intensity, shape	sensitivity score of 0.471 on ROC
Karim et al. (2019)	MAs	MATLAB NN Pattern Recognition Tool (NPRTOOL), K-means clustering, canny edge detector	intensity images	sensitivity of 0.616, specificity of 0.414, accuracy of 0.563
Selcuk and Alkan (2019)	MAs	Ant colony algorithm, Frangi filters	BVs removed	Dice index values of 0.81-0.90
Niemeijer et al. (2005)	All red lesions (i.e. MAs, hemorrhages (HEMs))	KNN classification, thresholding, morphological operators	removed bright lesions and BVs, probabilistic pixel-wise lesion map	image-level results with a sensitivity of 1.00 and a specificity of 0.87
Kar and Maity (2018)	Exudates	Mutual information of Laplacian of Gaussian (LoG) filter, matched filter (MF), differential evolution (DE) algorithm	LoG responses (bright lesions), MF and negative LoG responses (dark lesions)	accuracy of 0.9771 and AUC of 0.9554 on DRIVE, STARE, DIARETDB1, and MESSIDOR

Zhang et al. (2014)	Exudates	thresholding, FP, morphological top-hat filters, 500-tree RF	Adaptive template with pixel maxima	AUC of 0.95, sensitivity of 0.96, and specificity of 0.89 on e-ophtha EX
Osareh et al. (2009)	Exudates	Gaussian-smoothed histogram analysis, fuzzy c-means (FCM) clustering, genetic algorithm, multi-layer NN	LUV space features, histogram features	image-level sensitivity of 0.960 and specificity of 0.946
Lokuarachchi et al. (2019)	Exudates	intensity thresholding, contrast limited adaptive histogram equalization (CLAHE), mean filtering	hue saturation value (HSV) space images	DIARETDB1 sensitivity of 0.9459 and specificity of 0.8846
Wisaeng and Sa-Ngiamvibool (2019)	Exudates	morphology mean shift algorithm (MSA), mathematical morphology algorithm (MMA)	contrast, size	sensitivity, specificity, and accuracy of 0.9840, 0.9813, and 0.9835
Khojasteh et al. (2019)	Exudates	PCA, hybrid ResNet-50/SVM model	standard image input	accuracy and sensitivity of 0.976 and 0.99, respectively on DIARETDB1 and e-Ophtha
Brandon and Hoover (2003)	Drusen	Wavelet equalization, Mexican hat wavelet, feedforward NN, regional classification using rule system	orthonormal wavelets responses at the pixel level, area and intensity values, border intensity values, and gradient and density properties	correct detection rate of 0.87
Lee et al. (2001)	HEM, MAs, hard exudates (HE), cotton-wool spots (CWS), other lesions	rule-based	clinically specified	sensitivity and specificity above 0.90, and k-score above 0.75 (compared to experts)
Usher et al.	MAs, HEMs, exudates	ANN, locally adaptive contrast enhancement, recursive region growing, adaptive intensity thresholding	Color-enhanced intensity regions	maximum sensitivity of 0.951 at a specificity of 0.463
Tan et al. (2017)	MAs, HEMs, exudates	10-layer CNN	standard image inputs	pixel-wise sensitivity/specificity of 0.6257/0.9893 for HEMs, 0.4606/0.9799 for MAs, 0.8758/0.9873 for exudates, and 0.7158 sensitivity for dark lesions on CLEOPATRA

Niemeijer et al. (2007b)	HEs, CWS, drusen, bright lesions	KNN, filter responses, thresholding, linear discriminant analysis (LDA)	lesion probability mapping, contrast, size, shape, and proximity to vessels and red lesions, red lesions/image, bright lesions/image	AUC of 0.95 and sensitivity/specificity of 0.95/0.88, 0.95/0.86, 0.70/0.93, and 0.77/0.88 for bright lesions, exudates, cotton-wool spots, and drusen, respectively
Li et al. (2019)	MAs, soft exudates (SE), HEs, and HEMs	DeepLab-v3+, HED	standard image inputs	average precision (AP)/ intersection over union (IoU) of 0.0311/0.325 (MAs), 0.5634/0.3118 (HEs), 0.4359/0.2295 (SEs), 0.1842/0.1425 (HEMs)
Keel et al. (2019)	DR heatmaps	validated deep learning models for referable DR Keel et al. (2018)	lesion and vessel probabilistic maps	exudate, HEM, or vessel abnormality in 0.96 of true positives for DR; 0.85 of false positives displayed nontraditional fundus regions with or without retinal venules
He et al. (2019)	DR lesions such as MAs, HEMs, HEs, CWSs, or macular edema in Multispectral images (MSI)	local binary patterns (LBPs), generalized low-rank approximation of matrices (GLRAM), supervised regularization term (SRT), Gaussian kernel-based SVM	LBP feature vectors; MSI spectral bands	sensitivity of 0.925, specificity of 0.983, and accuracy of 0.981
Saha et al. (2019)	MAs, HEs, HEMs, SEs, OD	a DCNN, SegNet, based on VGG16 and Binary Cross Entropy loss	standard image inputs	Area under the positive prediction value vs. sensitivity of 0.0059 (MA), 0.5498 (HE), 0.0829 (HEM), 0.1823 (SE)
Costa et al. (2019)	unspecified DR lesions, lesion probability maps	Inception-v3 (CNN) with multi-instance learning (MIL) and a bagging scheme	image inputs	0.958 AUC while maintaining 0.05 of the Inception V3's number of parameters
Son et al. (2019)	HEMs, HEs, CWSs, drusen, membrane irregularities, macular holes, myelinated nerve fibers, chorioretinal atrophy, vascular abnormalities, retinal nerve fiber layer defects, disc changes, lesion heatmaps	twelve NNs perform binary classification using regionally guided CNN	image inputs	AUCs of 0.947 to 0.980 on IDRiD and e-ophtha

Table 4: Sample of methods used to segment lesions that are indicative of DR

Tufail et al. (2016) assessed different software for their potentials to replace manual prescreening in the National Health Service (NHS) diabetic eye screening program (DESP). The experiment compared three system: iGradingM Philip et al. (2007), Retmarker Oliveira et al. (2011); Ribeiro et al. (2013), and EyeArt Solanki et al. (2015). iGradingM Philip et al. (2007) classified disease by dot-HEM/MA presence or insufficient image quality. The system assessed quality by field definition and clarity; a following process extracted lesions by morphology. RetmarkerDR Ribeiro et al. (2013) extracted lesions using PCA and a region covariance descriptor. SVM distinguished lesions and classified DR. Further, RetmarkerDR implemented image co-registration to track disease progression. The algorithm estimated translation relative to the fovea and retinal vascular tree. EyeArt Solanki et al. (2015); Bhaskaranand et al. (2016) constituted a cloud system for rapid and large-scale screening for RDR. After quality assessment, the system identified pixels of interest using multiscale morphological filters and supervised learning classifiers. An ensemble classifier aggregated increasingly higher hierarchical levels to generate an encounter grade and produce image mappings. Further, image registration software tracked changes in MAs to identify disease progression. Overall, Retmarker and EyeArt achieved favorable grading results. EyeArt's cloud infrastructure produced the fewest technical issues and the greatest scaling to large data volumes. Walton et al. (2016) evaluated the Intelligent Retinal Imaging System (IRIS) for DR screening from teleretinal sites. The system used an NN to match images to those in a reference dataset. When compared to iGrading, Retmark, and EyeArt, the IRIS system correctly identified more cases of no disease and less cases of PDR.

5.4.2. Longitudinal Tracking

Xu et al. (2018) compared two procedures to diagnose DR progression. One method combined image registration, via Generalized Dual Bootstrap-ICP (GDBICP), with iGrading software. Differences were tracked by the angular proximity of MA coordinates. The other method used linear SVM to predict factor weights over longitudinal study. Both methods predicted DR severity and progress through classifying unchanged, new, and resolved MAs.

Adal et al. (2018) diagnosed DR by longitudinal changes to red lesions. Then, quadratic deformation registered images to account for nonlinear spatial deformation. Vessels were aligned by multi-resolution matching employed with a hierarchical registration model. A LoG operator detected MAs. Spatio-temporal retinal changes were measured by the absolute difference of the extremes of multi-scale blobness responses. RBF SVM yielded the highest lesion-level accuracy.

Arcadu et al. (2019) predicted DR progress according to 2-step or greater increase on the ETDRS DR Severity Scale (DRSS). This scheme trained field-specific Inception-v3 DCNNs as to recognize signs of DR disease and worsening. Overall, each pillar contained 25 DCNNs that were generated under a five-times fivefold cross-validation and transfer learning scheme. Then, RF aggregated the probabilities across the seven standard ETDRS image fields. The authors performed analysis via Shapley Additive Explanations (SHAP) to assess the positive or negative correlations of individual DCNN probabilities to the final RF predictions. Further, a guided back propagation method generated attribute maps to show regions that would be important to predict future progression (at 6, 12, or 24 months) from given baseline images.

5.4.3. Binary DR Classification

Giancardo et al. (2011) applied RBF SVM and PCA on features derived from Radon transform. An un-normalized Bayes method combined the SVM predictions/window and average gray level at the window center. Neighboring MAs were combined by enforcing non-maximum suppression. Tang et al. (2013) partitioned images into scale-specific splats to predict DR by retinal HEMs. Background regions were partitioned into fewer large splats by aggregating gradient magnitudes followed by the watershed method. Sequential filter and wrapper approaches detected the optimal splat features, which were fed into KNN. An image-level HEM index aggregated the posteriori probability of each splat representing DR. Niemeijer et al. (2009) combined KNN with k=281 with a posterior probability distribution normalization (PPDN)-based fusion system. Abràmoff et al. (2010) compared Challenge2009 Quellec et al. (2008b) and EyeCheck. EyeCheck clustered red lesion candidates and processed features using KNN. In contrast, Challenge2009 fit MAs to a parametric template in an adapted wavelet domain. Combining the systems performed better than either method in isolation. Bourouis et al. (2019) developed a hybrid generative/discriminative model. The authors generated a probabilistic SVM kernel from the mixture of scaled Dirichlet distributions. Overall, the Dirichlet-based Fisher model with minimum description length (MDL) criterion performed best. Frazao et al. (2019) extracted holistic texture features and local retinal features. One stream constructed a basis image set using Independent Component Analysis (ICA) and matched new images based on their projections to this subspace. The other stream extracted lesion features using image processing and DFFS measures Moghaddam and Pentland (1994). The authors compared several classifiers (KNNs, NNs, SVMs, and RFs) for classification. Gardner et al. (1996b) used an ANN to detect features of DR and identify the disease.

Gargeya and Leng (2017) constructed a CNN with six residual blocks, a stride of 2, batch normalization, and ReLU activation. This system generated heatmaps using an embedded convolutional visualization layer with a width of 1024 filters, average pooling layer, and a softmax layer. Other features included metadata related to pixel dimensions and image FOV. Overall, a second-level gradient boosting classifier harnessed 1027 features to predict DR. Quellec et al. (2017) used CNN ensembles based on the O-o framework Kaggle (2015b) to extract parameters at different iterations and optimally predict lesions. The third pass of the CNN produced heatmaps to show relevant parameters. DR grading was performed by RF following min-pooling. Hagos and Kant (2019) applied transfer learning using inception-v3 to tackle data imbalance in detecting DR. Gulshan et al. (2019) amended their deep NN Gulshan et al. (2016b) to detect RDR and DME. The improved model trained on the adjudicated data in Krause et al. (2018b). Further changes increased input resolution and replaced the framework with Inception-v4. Mobeen-ur-Rehman et al. (2019) trained a new 5-layer CNN to avoid overfitting. The structure used two initial 3x3 convolutional layers and three fully connect layers. Saranya Rubini et al. (2019) developed a new model, Deep CNN-based DR Detection (DCNN-DRD), that included five convolution and pooling layers, a dropout layer, three fully connected layers, and gradient descent graph. Liu et al. (2019) proposed a novel

105-block weighted path (WP) CNN (WP-CNN) to detect RDR using stacking WP blocks. This method employed multi-path weight coefficients and averaged the results. Pires et al. (2019) developed a data-driven CNN approach to screen for RDR in mobile devices. The network resembled VGGNet in weighted layers and o_O in fully connected layers. A multi-resolution strategy trained simplified variants of the CNN to pre-initialize progressively larger networks. Zhou et al. (2018) proposed a deep multiple instance learning (MIL) algorithm. AlexNet classified patches and fed the results into a global aggregation stage to classify images. Image outcomes averaged predictions over multiple scales to compensate for irregularities in lesion shape.

Pathological changes in DR cause conventional OCT-layer analysis to be difficult Khansari et al. (2019). Thus, Rathke et al. (2017); Khansari et al. (2019) focus on this modality, which offers unique layer information that cannot be harnessed in traditional retinal fundus photography. Rathke et al. (2017) constructed a nonspecific approach to segmentation based on local models. Sum-product networks (SPN) calculated the ideal model combination based on tractable globally optimal inferences. The model was trained on healthy OCT scans with local class-specific Gaussian models selectively modified to target pathology shapes. Markov random field regularization accounted for interactions between shape and appearance. The full model added conditional independence while minimizing Kullback-Leibler divergence. SPN combined the submodels using maximum-likelihood and dynamic programming. Translation-invariant pathology specific models were incorporated into appearance and shape modifications to maximize total log-likelihood. Optimal area factorization was computed through an iterative algorithm that utilized recursive computation and shifting windows of increasing widths. Regularization punished smaller models before the model enacted optimal MAPinference in the SPN. Subregions were smoothed into a single segmentation using a constrained least-squares problem. Khansari et al. (2019) aligned volumes of interest (VOIs) in OCT using designed masking and affine registration. The system performed stochastic gradient descent to optimize B-spline transform. Next, the authors nonlinearly registered OCT volumes to an atlas constructed from normal controls to measure pathological changes. Specifically, the Jacobian determinant of deformation indicated pathological contraction and expansion, whereas tensor-based morphometry (TBM) quantified local structural changes.

Other data sources may complement DR diagnosis. Sun and Zhang (2019) proposed an automated system to diagnose DR based on Electronic Health Records (EHR) and RFs. EHR are easily accessible and complement imaging based software for DR diagnosis. Sun (2019) proposed a 1-D CNN to detect DR from EHR. The study harnessed diabetes diagnoses, glycosylation, and biochemical tests. The approach integrated the LeNet-5 model, batch normalization, and an adaptive learning algorithm. As the input data are unrelated, the 1-D convolutional method treated each feature as an independent channel and applied deconvolution.

5.4.4. DR Severity Classification

Roychowdhury et al. (2014) proposed DR Analysis Using Machine Learning (DREAM). GMM and KNN identified bright and red lesions, respectively. Then, the model implemented feature-ranking and filtered false positives. The final stage identified DR severity based on lesion classifications. Yun et al. (2008) trained an ANN to grade DR based on veins, hemorrhages, and MAs. Costa et al. (2018) applied multiple instance learning (MIL) to leverage implicit information from clinical annotations when classifying images. A Bag of Visual Words (BoVW) described the image (bag) by local low-level visual descriptors (instances). One NN encoded instance-level feature vectors onto useful mid-level representations, whereas the second classified bags accordingly. Error from the second network backpropagated to the first to enforce useful representations. Sarwinda et al. (2018) classified NPDR stages based on Histogram of Oriented Gradients (HOG) and shallow learning. Factor analysis (FA) selected the most relevant features; then, SVM and RF independently graded for NPDR. Sil Kar and Maity (2018) graded DR by detecting lesions and neovascularization in sub-sampled retinal images following compressed sensing (CS). CS captured images using sampling rates that fell significantly below the Nyquist frequency. An initial stage maximized fuzzy entropy for MF responses to extract vessels. The algorithm detected neovascularization according to the 2D MI between vessel density and tortuosity in thin vessels. Lesion detection jointly maximized the MI between the maximum MF and Laplacian of Gaussian (LoG) filter responses. A rule-based scheme graded DR based on types and numbers of lesions (NPDR) and neovascularization (PDR). Kaya et al. (2018) introduced hybrid Hilbert-Huang Transform (HHT) and particle swarm optimization (PSO)-based ANNs (HHT-ANNPSO). The framework decomposed Video-Oculography (VOG) signals and extracted statistical features for classification. ANNPSO included 20 input nodes, one hidden layer with 16 neurons, and an output classifier. The HHT-ANNPSO model used intrinsic mode function (IMF) groups with 4, 8, 12, and 16 input neurons. PSO optimized parameter learning. Mohammadpoory et al. (2019) applied visibility graph (VG) and Radon transform (RT). The VG method represented radon projections along a graph according to signal complexity and fractality; the study used two types of VG: natural (NVG) and horizontal VG (HVG). The method extracted features from the graph topology and structure using graph theory. Then, an error-correcting output codes (ECOC) classifier used the extracted graph topologies to associate training data with defined codewords.

Gulshan et al. (2016a) applied ensemble of 10NNs based on Inception-v3. Ting et al. (2017) employed eight CNNs, each adopted from VGGNet, to classify: DR severity (two networks), referable glaucoma (two networks), referable AMD (two networks), image quality (one network), and non-retinal images (one network). Yang et al. (2017) mapped lesion presence and severity by a 10-layer CNN. Specifically, this local network included four convolution layers, max-pooling at layers three and five, two fully connected layers, and a softmax layer. Next, this stage passed two lesion maps to the global network along with the original image. The global network classified NPDR severity. It was composed of a 24-layer structure with alternating convolutional and max-pooling layers, three fully-connected layers, and a softmax layer. Zhang et al. (2017) proposed a deeply supervised residual network (ResNet). Their model included five residual units with two convolution layers each. A fusion layer combined outputs from three multi-scale side layers. Class imbalance was addressed by multi-class balanced cross-entropy loss. Ardiyanto et al. (2017) developed a DCNN, Deep-DR-Net, using an encoder-classifier with residual connections between each layer. The encoder constructed 16-D feature maps using concatenated 13-filter convolutional layers and a max pooling layer. Five cascaded convolutions blocks contained projection-receptive-projection sequences. A standard log-softmax layer classified images. Roy et al. (2017) estimated DR severity using a CNN trained on ImageNet (DR-CNN) and fed the resulting features into two visual dictionaries: a discriminative and generative pathology histograms (DPH and GPH). RF predicted DR severity using input features from DR-CNN and both dictionaries. Takahashi et al. (2017) evaluated a modified GoogLeNet using Davis grading and prevalence- and bias-adjusted kappa (PABAK) scoring.

Mansour (2018) segmented interest regions using adaptive learning-based GMM and connected component analysis. AlexNet and LDA extracted high dimensional features from these regions and optimized feature selection, respectively. Ultimately, SVM labeled inputs. Zhang et al. (2019) developed a two-part ensemble, DeepDR, to detect lesions and grade DR. Feature extraction used three CNN learners: ResNet50, DenseNet169, and DenseNet201. Three feature classifiers operated as standard deep NNs (SDNNs). Each SDNN included initial global average pooling (GAP), a 2048-unit fully connected layer, and a leaky rectified linear unit function (leaky ReLU). Sahlsten et al. (2019) modified inception-v3 by doubling the fully-connected layers, employing Adam optimization, and using dropout in the first fully-connected layer. Li et al. (2019) compared VGG-16, ResNet-18, GoogLeNet, DenseNet-121, and SE-BN-Inception. DenseNet-121 performed best on mild and severe DR. Overall, SE-BN-Inception achieved the highest accuracy and kappa score. Raumviboonsuk et al. (2019) employed a ten-model Inception-v4 ensemble to grade images for DR severity, referable DME, and image quality. Zeng et al. (2019) proposed a Siamese-like CNN framework that graded DR in binocular fundus images. The Siamese-like architecture simultaneously imputed images from both eyes into parallel inception-v3 models that shared weights. A fully-connected layer concatenated features from both streams to make final predictions. Gao et al. (2019) split each image into four 300 x 300 pixel regions and processed them through four different inception-v3 models. The overall model, Inception@4, concatenated the global average pooling layers from each model for final classification.

The following papers produced attention heatmaps to explain accompanying DR grades from retinal images. Wang et al. (2017) developed a weakly-supervised CNN, Zoom-in-Net, to cluster features which likely represented lesions within minimum bounding boxes. Zoom-in-Net consisted of a main network (M-Net) that graded images, an attention network (A-Net) that created the heatmaps, and a Crop-Network (C-Net) that concatenated optimum features from the other networks. M-Net resembled Inception-ResNet and labeled images. A-Net imputed feature maps from M-Net and processed information through two branches. In one branch, a 1x1 convolutional layer linearly classified pixels. The other branch used three convolution layers and a spatial softmax operator to generate the attention maps. C-Net greedily sampled the gated attention maps from A-Net. de la

Torre et al. (2019) proposed a 17-layer CNN with quadratic-weighted kappa (QWK) loss. This method defined pixel relevancy scores by the sum of all node contributions from the pixel space plus a constant score depending on individual layer receptive fields. The authors mapped hidden layer contributions to the input receptive field using summed gaussian distributions. Finally, they conducted PCA after the last layer. Sayres et al. (2019) employed inception-v4 to grade for quality and DR. The authors generated heatmaps for the predicted severity grade based on the integrated gradients method Sundararajan et al. (2017). Kumar et al. (2019) proposed CLass-Enhanced Attentive Response Discovery Radiomics (CLEAR-DR) to compute radiomic sequences. The last sequencer stage deconvolved each kernel response to visualize relevant pixels in two maps: a dominant class attentive map (region grade) and a dominant response map (region contribution to image grade). The final CLEAR-DR map combined these two.

Hwang et al. (2018) extracted quantified nonperfusion parameters from projection-resolved OCTA (PR-OCTA). The study quantified capillary nonperfusion in 3 vascular plexuses in the macula. Split-spectrum amplitude decorrelation angiography detected blood flow, and PR-OCTA compared this signal strength with that of the location's superficial signal. Directional graph search algorithm segmented the volumes into superficial vascular complex (SVC), intermediate capillary plexus (ICP), and deep capillary plexus (DCP). The highest sensitivity for DR was achieved using the extrafoveal avascular area (EAA) from the SVC. ElTanboly et al. (2018) graded NPDR in OCT using a two-stage deep fusion classification network (DFCN). Three global features (reflectivity, curvature, and thickness) were fed into the DFCN with a cumulative distribution function (CDF) of twelve retinal layers. Normalized reflectivity was measured from the foveal peak on the nasal and temporal sides. Huber's M-estimates detected average intensity on a given segment. For layer curvature, the surface was smoothed using a local-weighted polynomial, and the Menger curvature values were combined across each layer. Thickness was calculated by solving the planar Laplace equation for the harmonic function under the Dirichlet boundary condition. The Euclidean distances that linked vertical layers represented layer thickness. The DFCN was trained by stacked non-negativity constraint autoencoders (SNCAEs). In stage one, multiple SNCAEs per layer independently compressed their input features based on greedy unsupervised pretraining. After fine-tuning, a SNCAE was trained on the fused top activators. Stage two exclusively classified images with DR.

Paper	Focus	Method	Image features	Results
Abramoff and Suttorp-Schulten (2005)	DR screening, image gradability, pupil dilation, urgency level of treatment	EyeCheck, online database and image processing	neovascularization, red and bright lesions, and other relevant abnormalities	inter-grader agreement was 0.93
Abràmoff et al. (2008)	DR screening, image gradability, BVs, lesions, other clinical features	Ensemble of machine learning classifiers	Filter responses, color histograms,	AUC of 0.84 on EyeCheck
Abràmoff et al. (2013)	Referable DR	Iowa Detection Program (IDP)	standard image inputs	AUC of 0.937
Abràmoff et al. (2016)	Referable and vision-threatening DR	IDx-DR X21 system, AlexNet (CNN), RF	lesion characteristics	sensitivity was 0.968, specificity was 0.870, negative predictive value of 0.99, AUC was 0.980
Keel et al. (2018)	Referable DR, DME, image quality	EyeGrader system, Inception-v3 (CNN)	standard image inputs	sensitivity and specificity for correct referral was 0.923 and 0.937, respectively
Philip et al. (2007)	Binary DR detection	iGradingM, morphological operators, clinical features	Dot HEM / MA – area, mean intensity, image quality – OD localization, small vessels near fovea	0.905 sensitivity and 0.674 specificity for technical failures or any retinopathy
Ribeiro et al. (2013)	DR detection and longitudinal change	RetmarkerDR, PCA, region covariance descriptor, SVM, image co-registration	MAs, vascular lesions, area, shape, intensity	0.73 sensitivity for any retinopathy, 0.85 for referable retinopathy and 0.979 for PDR
Solanki et al. (2015); Bhaskaranand et al. (2016)	Referable DR, image quality	EyeArt, multi-scale, morphological filters, supervised learning, image registration	MAs, filter responses, aggregated level information (pixel, lesion, image, encounter)	0.947 for any retinopathy, 0.938 for referable retinopathy and 0.996 for PDR
Walton et al. (2016)	DR screening	Intelligent Retinal Imaging System (IRIS), NN classifier	highest probabilistic match to reference dataset	sensitivity for sight-threatening DR was 0.664, a false-negative rate of 0.02, specificity of 0.728, positive predictive value was 0.108, negative predictive value was 0.978
González-Gonzalo et al. (2019)	Detection of DR and age-related macular degeneration (AMD)	RetCAD v.1.3.0, ensemble CNNs	standard image inputs	AUC of 0.951, sensitivity of 0.901, specificity of 0.906
Quellec et al. (2019)	RDR and DR severity	OphAI, CNN ensembles for laterality, detection, grading	single images or full examination records	vision-threatening DR with an AUC of 0.997 and PDR with an AUC of 0.997

Xu et al. (2018)	DR tracking	Generalized Dual Bootstrap-ICP (GDBICP) and iGrading software; linear SVM	MAs, other pathological features	a sensitivity of 0.94 and specificity of 0.93 using image analysis and a sensitivity of 0.89 and specificity of 0.88 using SVM
Adal et al. (2018)	DR tracking	Quadratic deformation, multi-resolution matching strategy, LoG operator, multi-scale blobness, RBF SVM	red lesions, vessel alignment	highest accuracy of 0.80 at the lesion level
Arcadu et al. (2019)	DR tracking	Field-wise DCNN pillars aggregated in RF, Shapley Additive Explanations (SHAP) and attribute maps (map importance)	7-field color fundus photographs, 2-step or more ETDRS severity scale worsening	AUC of 0.68 (sensitivity = 0.66, specificity = 0.77) (6 months), AUC of 0.79 (sensitivity = 0.91, specificity = 0.65) (12 months), and AUC of 0.77 (sensitivity = 0.79, specificity = 0.72) (24 months)
Gargeya and Leng (2017)	DR detection	CNN, a second-level gradient boosting classifier	Standard image inputs, abnormality heatmap	AUC of 0.97, sensitivity of 0.94, and specificity of 0.98 on EyePACS
Giancardo et al. (2011)	DR detection	SVM, PCA, radon transform, Bayes, non-maximum suppression	Radon feature vector	global score of 0.375 on ROC
Tang et al. (2013)	DR detection	Splat partitioning, gradient and watershed, sequential filter and wrapper approaches, KNN	Splat features	AUC of 0.96 at the splat level and 0.87 at the image level on MESSIDOR
Niemeijer et al. (2009)	DR detection and referral	KNN, leave out classification, posterior probability distribution normalization (PPDN), model fusion	Fused image information	AUC of 0.881
Abràmoff et al. (2010)	DR detection	Challenge2009 (parametric template with wavelets) and EyeCheck (KNN classifier)	Fused features from both systems (exudates, CWSs, MAs, HEMs)	AUC of 0.86
Quellec et al. (2017)	DR detection	O-o framework: CNN ensembles, RF	attribute heatmaps	accuracy of 0.95
Rathke et al. (2017)	DR detection in OCT scans	Sum-product networks (SPN), linear Gaussian, Markov random fields, Kullback-Leibler divergence, MAP interference, translation-invariant pathology specific models recursive iteration	shape variations in layer boundaries, trained on healthy OCT scans only	unsigned error of 4.08 for the mild NPDR testing dataset

Hagos and Kant (2019)	DR detection	Inception-v3, cosine loss, transfer learning	standard image inputs	accuracy of 0.909
Gulshan et al. (2019)	RDR and diabetic macular edema (DME)	Inception-v4	standard image inputs	AUC of 0.963, sensitivity of 0.889 and specificity of 0.922
Khansari et al. (2019)	detecting NPDR or PDR in OCT images	masking and affine registration, stochastic gradient descent to optimize B-spline transform, Jacobian determinant of deformation, tensor-based morphometry (TBM)	difference from “healthy subject” atlas	not given for classification task; differences in NPDR/PDR in voxel deformation and layer surface distance provided
Bourouis et al. (2019)	DR detection	new SVM kernel: scaled Dirichlet distributions with minimum description length criterion	standard image inputs	0.98 AUC and 0.9872 accuracy on DIARETDB1
Saranya Rubini et al. (2019)	DR detection	Deep CNN-based DR Detection (DCNN-DRD)	standard image inputs	0.97 accuracy
Liu et al. (2019)	RDR detection	105-block weighted path CNN (WP-CNN)	standard image inputs	accuracy of 0.9423, sensitivity of 0.9094, specificity of 0.9574, AUC of 0.9823, and F1-score of 0.9087
Pires et al. (2019)	DR detection for mobile devices	VGGNet, O-o framework, muti-resolution learning, Nesterov momentum	standard image inputs at different resolutions	AUC of 0.982 (training on Kaggle, testing on Messidor)
Frazao et al. (2019)	DR detection	Independent Component Analysis (ICA)	holistic texture features, local retinal features (MAs and exudates), removed OD	sensitivity of 0.901, specificity of 0.561 on Messidor
Zhou et al. (2018)	DR detection	multiple instance learning (MIL), patch-level CNNs (AlexNet), image aggregation	image patch inputs	AUC of 0.925 on Kaggle and 0.960 on Messidor, F1-score of 0.924, sensitivity of 0.995, and precision of 0.863 on DIARETDB1
Gardner et al. (1996b)	DR detection in sixty degree red-free fundus images	ANN, thresholding	vessel abnormalities, exudates, and hemorrhages; image patch inputs	sensitivity of 0.884 and specificity of 0.835
Roychowdhury et al. (2014)	DR grading	DR Analysis Using Machine Learning (DREAM), GMM, KNN	removal of OD and BVs, identified bright and red lesions	1.00 sensitivity, 0.5316 specificity, and 0.904 AUC on Messidor
Yun et al. (2008)	DR grading	ANN	veins, HEMs, MAs; perimeter and area values	sensitivity of 0.91, specificity of 1.00, and accuracy of 0.84

Gulshan et al. (2016a)	DR detection and grading, diabetic macular edema (DME), image gradeability	Inception-v3, 10 CNN ensemble	standard image inputs	AUC of 0.99, sensitivity of 0.975, and specificity of 0.985
Ting et al. (2017)	DR grading, glaucoma, referable age-related macular degeneration, image quality	VGGNet, 8 CNN ensemble	standard image inputs	sensitivity of 0.905 and specificity of 0.916 for RDR; 1.0 sensitivity and 0.911 specificity for vision-threatening DR
Costa et al. (2018)	DR grading	Multi-instance learning (MIL) with Bag of Visual Words (BoVW), NNs, SURF features	weakly-labeled image inputs	AUC of 0.90 on Messidor, 0.93 on DR1, and 0.96 on DR2
Yang et al. (2017)	DR grading	local and global DCNN	standard image inputs	lesion-wise AUC of 0.9687 and image-wise AUC of 0.9590 on Kaggle
Wang et al. (2017)	DR grading	Weakly-supervised CNN (Zoom-in-Net): M-Net (inception-resnet), A-Net, C-Net	standard image inputs, attention heatmaps	AUC of 0.854 on EyePACS and of 0.921 on MESSIDOR
Zhang et al. (2017)	DR grading	deeply supervised residual network (ResNet), multi-class balanced cross entropy	standard image inputs	accuracy of 0.808, sensitivity of 0.94, specificity of 0.665, and precision of 0.54
Ardiyanto et al. (2017)	DR grading	Deep-DR-Net (cascaded encoder classifier NN with residual connections)	standard image inputs	image accuracy of 0.9571, sensitivity of 0.7692, and specificity of 1.00 on FINDeRS
Roy et al. (2017)	DR grading	CNN, discriminative pathology histogram (DPH) and generative pathology histogram (GPH), RF	image patches	quadratic kappa score of 0.86
Hwang et al. (2018)	DR grading in projection resolved OCTA (PR-OCTA)	split-spectrum amplitude decorrelation, directional graph search	capillary nonperfusion in 3 vascular plexuses in the macula	with specificity fixed at 0.95, sensitivities of 0.70 (healthy), 0.84 (DR), and 1.0 (severe DR)
ElTanboly et al. (2018)	DR grading in OCT	two-stage deep fusion classification network (DFCN) with stacked non-negativity constrained autoencoders, Huber's M-estimates, Menger curvature, Dirichlet boundary	shape, intensity, and spatial information (segmentation); reflectivity, curvature, and thickness (DFCN)	non-DR and DR with 0.93 accuracy, and early DR from mild/moderate DR with 0.98 accuracy
Takahashi et al. (2017)	DR grading	modified GoogLeNet	single image, two images, or four stitched images available in alternative models	prevalence- and bias-adjusted kappa (PABAK) to modified Davis grading = 0.64; PABAK to real prognosis grading = 0.37

de la Torre et al. (2019)	DR grading	17-layer CNN, quadratic-weighted kappa (QWK) loss, PCA	standard image inputs	QWK of 0.832, sensitivity = 0.908, specificity = 0.911, positive prediction value (PV) = 0.884, Negative PV = 0.930, Accuracy = 0.910, F1 = 0.896, MCC = 0.817 on MESSIDOR
Sayres et al. (2019)	DR grading, image quality	Inception-v4 (CNN)	standard image inputs	sensitivity of 0.9155 and specificity of 0.9469
Mansour (2018)	DR grading	GMM and connected component analysis (segmentation),AlexNet (CNN) and linear discriminant analysis (LDA) (feature extraction), SVM (grading)	standard image inputs	0.9793 accuracy on Kaggle
Zhang et al. (2019)	DR grading	ResNet50, DenseNet169, and DenseNet201 (extraction) standard deep NN (SDNN) (classify)	standard image inputs	sensitivity of 0.981 and a specificity of 0.989
Sahlsten et al. (2019)	DR grading	Inception-v3	standard image inputs	AUC of 0.962, accuracy of 0.869, and quadratic weighted kappa of 0.910
Li et al. (2019)	DR grading	VGG-16, ResNet-18, GoogLeNet, DenseNet-121, and SE-BN-Inception	standard image inputs	accuracy of 0.8284
Raumviboonsuk et al. (2019)	DR grading, referable DME, and image quality	10-NN ensemble based on Inception-v4	standard image inputs	QWK was 0.85
Zeng et al. (2019)	DR grading	“Siamese-like” parallel Inception v-3 frameworks	images from each eye used to strengthen the predictions in both	AUC of 0.951, kappa score of 0.829
Gao et al. (2019)	DR grading	“Inception@4” - Inception-v3 ensemble (4 CNNs)	images divided into four equal patches	accuracy of 0.8872
Sarwinda et al. (2018)	NPDR grading	Histogram of Oriented Gradients (HOG) and shallow learning methods, factor analysis (FA), SVM and RF	local intensity and edge information	accuracy above 0.9067, sensitivity above 0.8603, specificity above 0.9308
Mohammadpoory et al. (2019)	DR grading	Visibility graph (VG) method, radon transform (RT), graph theory, error-correcting output codes, adaboost and KNN	graph topologies and structures, radon-transformed features	accuracy of 0.9792, sensitivity of 0.9583, and a specificity of 0.9861

Table 5: Sample of methods used to detect and grade DR

5.5. Diabetic Macular Edema (DME)

Deepak and Sivaswamy (2012) evaluated a rotational asymmetry metric at the macula to grade DME in color fundus images. First, the system generated a smear pattern for the circular region of interest (ROI) that encircled the macula. Smearing was spatially dependent on maximum rotation motion. Then, radon transform (RT) responses were concatenated at different orientations. DME correlated to peaks in intensity that corresponded to hemorrhages. To assess DME severity, the approach approximated the macula to be rotationally symmetric in moderate DME (no hemorrhages in the macula) and rotationally asymmetric in severe DME (hemorrhages in the macula). The ROI was divided into eight angular patches with ten histogram bins each. Specifically, the measure defined symmetry by the second norm of the distance between histograms of diametrically opposite pair(s) of patches. Acharya et al. (2017) conducted a series of RT, discrete wavelet transform (DWT), and discrete cosine transform (DCT) to extract features for DME diagnosis in color fundus images. After these processes, locality sensitive discriminant analysis (LSDA) narrowed the features extracted from DCT coefficients. A decision tree classified DME using seven input features from LSDA.

Rasti et al. (2018) differentiated normal retina, dry AMD, and DME in OCT. First, a graph-based curvature algorithm corrected retinal distortions and extracted volumes of interest. Prior to classification, multi-scale spatial pyramid (MSSP) decomposition decomposed image slices. The multi-scale convolutional mixture of expert (MCME) model combined four scale-dependent CNNs. For the approach, a GMM modeled individual experts by independent multivariate Gaussians and expert interactions by correlated multivariate components. Each expert model individually estimated membership for each class label based on a partition of the feature space. A CNN gating network (CGN) weighed the expert predictions to predict the overall image label. During back propagation, a cross-correlation penalty term negatively correlated individual experts' errors. Sidibé et al. (2017) approached screening for DME in spectral domain OCT (SD-OCT) by an anomaly method. The method constructed a Gaussian Mixture Model (GMM) from normal OCT scans. Then, it applied a threshold number of scans/volume that significantly differed in Mahalanobis distance from this GMM as an assessment of abnormality.

6. Discussion

DR encompasses a wide range of possible lesions, defects, and defect locations that presents a unique challenge to effective diagnosis. Lesion detection algorithms particularly emphasize MA detection as a typical indication of early DR, such that the disease can be diagnosed before significant eye damage occurs. However, detection of specific lesion types sacrifices effective DR grading - such grading could heighten standard of care and provide options to individuals in later stages of disease progression. Further, individualistic differences in the number and progression of specific lesion types could risk leading individuals to being "missed" during screening that relies solely on one presentation of disease.

In addition to challenges concerning DR's diverse pathology, lesions and BV abnormalities may appear in virtually any area of the retina. Engineering approaches to DR diagnosis must consider the clinical relevance of DR geography. For instance, DR requires a wide-FOV (at least 60 degrees). Such stipulations limit use for lower FOV imaging systems and contribute to difficulties in simultaneous diagnosis of DR with conditions like glaucoma that primarily affect the Optic nerve head (ONH). DR requires a proper view of the retinal periphery, where significant BV irregularities and damage are common in the disease. Further, more central structures are also essential, as damage to areas such as the macula in DME is particularly vision threatening. Other than color fundus images, imaging modalities such as flourescein angiography and OCT provide insight into DR diagnosis. Specifically, flourescein angiography offers high contrast for defects such as MAs, whereas OCT provides differences in retinal layers that would be impossible to distinguish otherwise.

The most successful studies used deep learning for classification, particularly in DR grading. Many papers had high success with end-to-end learning; others detected lesions using less computationally expensive systems before feeding the resultant features to a CNN. Segmentation of the OD and/or BVs increased overall performance in many instances; however, such procedures were sometimes problematic. For instance, removal of BVs generally aided in MA detection, but this procedure eliminates the ability to detect the proliferative disease. Therefore, clinical relevance should be of the outmost importance when making decisions for related preprocessing stages: MA detection, the most common focus of lesion detection systems, serves as a great benefit for early DR detection. As DR progresses, BV changes and other lesion types become more important. Further, DL systems often perform at high sensitivity and specificity without requiring extensive preprocessing. This less-handcrafted approach reduces the likelihood of important parameters being removed.

Deep learning emerged for detecting DR after 2015 and have grown dramatically since. This trend represents a movement towards less handcrafted features, with the following advantages: (1) recognizing features from images that were previously unknown or manually undetectable, (2) easier adjusting to diverse cohorts or imaging modalities, and (3) rapid and accurate learning that addresses user variability and specialist availability. While numerous research groups have received relative success, clinical implementation has been limited. These limitations concern data availability and uncertainty surrounding deep learning systems. Researchers address small dataset sizes through transfer learning, in which lower-level features like image contours are learning from public datasets (Table 1). In this method, an algorithm trains select deeper layers on the target small database to learn specific information and perform optimally on the intended population. Clinical data also frequently suffers from limited or weak labels (i.e. image-wise disease severity, lesion-wise segmentations). Thus, methods as in Dai et al. (2018) harness the supervised information in clinical texts to compensate for weakly-labeled images. So-called natural language processing (NLP) uses information that is already clinically available and increase the accuracy of ground-truth labels.

Further, uncertainty surrounds DL's “black-box” nature and hinders its clinical acceptance. Therefore, systems as in Quellec et al. (2017) and Wang et al. (2017) output complementary attention maps to highlight which pixels factor into disease predictions. These attention maps provide insight into the complex systems that control AI algorithms, help validate the decisions made (i.e. prove not "randomly guessing"), and function as clinical support tools (i.e. spotlight areas or features that may have not been clinical focuses previously). Other superior methods facilitate clinical utility in other ways, like providing for multiple diseased states as in Gulshan et al. (2016a) and Ting et al. (2017). Future research should focus on enhancing clinical utility and meeting real problems surrounding retinal diagnostics and therapeutics, rather than simply achieving higher label "accuracies". For instance, DL offers great potential for clinical settings in which resources (i.e. specialists and equipment) are limited. Such systems would compensate for poor image quality to facilitate use with mobile cameras, in addition to grading image quality as in Gulshan et al. (2016a).

This survey was limited in effective comparison of the methods for DR detection and grading systems for two primary reasons: (1) narrow focus within the overall topic of DR, and (2) inconsistent use of datasets and evaluation metrics. In the first case, it was observed that the highest performances were generally notable for papers that focused on specific tasks. For instance, one study eliminated BVs and HEMs to improve MA detection, as noted above. While this eliminates false positives for MAs, the authors sacrifice a more complete picture of DR status. Superior systems used learning ensembles or deep learning procedures to incorporate more specific tasks into overall DR grading. Other papers used deep learning to achieve high performance on DR within larger eye disease diagnosis systems. Furthermore, many studies were difficult to compare to one another effectively due to inconsistent evaluation metrics. In reading scientific literature, readers of this survey should be cautioned against accepting article's claims that report one evaluation metric. For instance, accuracy can easily be manipulated by data imbalance, or less common evaluation metrics can be utilized to present the results more favorably than a more objective method (i.e. AUC). Readers should be equally skeptical of papers that exclusively use private databases, especially when the authors have not made their code available for external validation.

7. Conclusions

Diabetic Retinopathy represents a high-risk complication of diabetes that can lead to irreversible blindness without proper treatment. Proper artificial intelligence (AI) systems offer high potential in facilitating expedited detection and relevant care for individuals who suffer from this disease. Early detection systems benefit highly from a focus on mild DR symptoms, notably MAs, leading to a prevalence of such works in this area. Faster and less computationally expensive algorithms also offer high promise for application in limited-resource areas that would benefit from rapid screening. However, more complicated and diverse systems of imaging and grading are emerging that can increase the sensitivity and specificity of DR detection and grading than is possible by human observation. Overall, deep learning systems show the most significant promise in simultaneous detection of DR and evaluation of its severity level. Future implementation of such algorithms will help doctors more rapidly distinguish between different eye diseases that may be present at once or present similar symptoms, as well as diagnose disease stage so that the most relevant and helpful treatments enacted rapidly and effectively.

References

Abramoff, M.D., Niemeijer, M., 2006. The automatic detection of the optic disc location in retinal images using optic disc location regression, in: 2006 International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 4432–4435. doi:`10.1109/IEMBS.2006.259622`.

Abràmoff, M.D., Niemeijer, M., Suttorp-Schulten, M.S., Viergever, M.A., Russell, S.R., van Ginneken, B., 2008. Evaluation of a system for automatic detection of diabetic retinopathy from color fundus photographs in a large population of patients with diabetes. Diabetes Care 31, 193–198. URL: `https://care.diabetesjournals.org/content/31/2/193`, doi:`10.2337/dc07-1312`.

Abramoff, M.D., Suttorp-Schulten, M.S., 2005. Web-based screening for diabetic retinopathy in a primary care population: The eyecheck project. Telemedicine and e-Health 11, 668–674. URL: `https://doi.org/10.1089/tmj.2005.11.668`, doi:`10.1089/tmj.2005.11.668`. pMID:16430386.

Abràmoff, M.D., Folk, J.C., Han, D.P., Walker, J.D., Williams, D.F., Russell, S.R., Massin, P., Cochener, B., Gain, P., Tang, L., Lamard, M., Moga, D.C., Quellec, G., Niemeijer, M., 2013. Automated Analysis of Retinal Images for Detection of Referable Diabetic Retinopathy Automated Analysis of Retinal Images. JAMA Ophthalmology 131, 351–357. URL: `https://doi.org/10.1001/jamaophthalmol.2013.1743`, doi:`10.1001/jamaophthalmol.2013.1743`.

Abràmoff, M.D., Lou, Y., Erginay, A., Clarida, W., Amelon, R., Folk, J.C., Niemeijer, M., 2016. Improved Automated Detection of Diabetic Retinopathy on a Publicly Available Dataset Through Integration of Deep Learning. Investigative Ophthalmology Visual Science 57, 5200–5206. URL: `https://doi.org/10.1167/iovs.16-19964`, doi:`10.1167/iovs.16-19964`.

Abràmoff, M.D., Reinhardt, J.M., Russell, S.R., Folk, J.C., Mahajan, V.B., Niemeijer, M., Quellec, G., 2010. Automated early detection of diabetic retinopathy. Ophthalmology 117, 1147 – 1154. URL: `http://www.sciencedirect.com/science/article/pii/S0161642010003258`, doi:`https://doi.org/10.1016/j.ophtha.2010.03.046`.

Acharya, U.R., Mookiah, M.R.K., Koh, J.E., Tan, J.H., Bhandary, S.V., Rao, A.K., Hagiwara, Y., Chua, C.K., Laude, A., 2017. Automated diabetic macular edema (dme) grading system using dwt, dct features and maculopathy index. Computers in Biology and Medicine 84, 59 – 68. URL: `http://www.sciencedirect.com/science/article/pii/S0010482517300689`, doi:`https://doi.org/10.1016/j.compbiomed.2017.03.016`.

Adal, K.M., van Etten, P.G., Martinez, J.P., Rouwen, K.W., Vermeer, K.A., van Vliet, L.J., 2018. An automated system for the detection and classification of retinal changes due to red lesions in longitudinal fundus images. IEEE Transactions on Biomedical Engineering 65, 1382–1390. doi:`10.1109/TBME.2017.2752701`.

Aiello, L.P., Odia, I., Glassman, A.R., Melia, M., Jampol, L.M., Bressler, N.M., Kiss, S., Silva, P.S., Wykoff, C.C., Sun, J.K., for the Diabetic Retinopathy Clinical Research Network, 2019. Comparison of Early Treatment Diabetic Retinopathy Study Standard 7-Field Imaging With Ultrawide-Field Imaging for Determining Severity of Diabetic Retinopathy. JAMA Ophthalmology 137, 65–73. URL: `https://doi.org/10.1001/jamaophthalmol.2018.4982`, doi:`10.1001/jamaophthalmol.2018.4982`.

Akram, M.U., Khalid, S., Khan, S.A., 2013. Identification and classification of microaneurysms for early detection of diabetic retinopathy. Pattern Recognition 46, 107 – 116. URL: `http://www.sciencedirect.com/science/article/pii/S003132031200297X`, doi:`https://doi.org/10.1016/j.patcog.2012.07.002`.

Al-Diri, B., Hunter, A., Steel, D., Habib, M., Hudaib, T., Berry, S., 2008. Review - a reference data set for retinal vessel profiles, in: 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 2262–2265. doi:`10.1109/IEMBS.2008.4649647`.

Alipour, S.H.M., Rabbani, H., Akhlaghi, M.R., 2012. Diabetic retinopathy grading by digital curvelet transform. Computational and mathematical methods in medicine 2012.

American Academy of Ophthalmology, 2014. International clinical classification system for diabetic retinopathy and diabetic macular edema - 2012. URL: `https://www.aao.org/clinical-statement\/international-clinical-classification-system\-diabe`.

Aquino, A., Gegundez-Arias, M.E., Marin, D., 2010. Detecting the optic disc boundary in digital fundus images using morphological, edge detection, and feature extraction techniques. IEEE Transactions on Medical Imaging 29, 1860–1869. doi:`10.1109/TMI.2010.2053042`.

Arcadu, F., Benmansour, F., Maunz, A., Willis, J., Haskova, Z., Prunotto, M., 2019. Deep learning algorithm predicts diabetic retinopathy progression in individual patients. npj Digital Medicine 2. URL: `https://doi.org/10.1038/s41746-019-0172-3`, doi:`10.1038/s41746-019-0172-3`.

Ardiyanto, I., Nugroho, H.A., Buana, R.L.B., 2017. Deep learning-based diabetic retinopathy assessment on embedded system, in: 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 1760–1763. doi:`10.1109/EMBC.2017.8037184`.

Bae, K., Lee, J.Y., Kim, T.H., Cho, G.E., Ahn, J., Kim, S.J., Kim, J.H., Kang, S.W., . Anterior Diabetic Retinopathy Studied by Ultra-widefield Angiography. Korean J Ophthalmol .

Bhaskaranand, M., Ramachandra, C., Bhat, S., Cuadros, J., Nittala, M.G., Sadda, S., Solanki, K., 2016. Automated diabetic retinopathy screening and monitoring using retinal fundus image analysis. Journal of Diabetes Science and Technology 10, 254–261. URL: `https://doi.org/10.1177/1932296816628546`, doi:`10.1177/1932296816628546`. pMID: 26888972.

Bourouis, S., Zaguia, A., Bouguila, N., Alroobaea, R., 2019. Deriving probabilistic svm kernels from flexible statistical mixture models and its application to retinal images classification. IEEE Access 7, 1107–1117. doi:`10.1109/ACCESS.2018.2886315`.

Brandon, L., Hoover, A., 2003. Drusen detection in a retinal image using multi-level analysis, in: Ellis, R.E., Peters, T.M. (Eds.), Medical Image Computing and Computer-Assisted Intervention - MICCAI 2003, Springer Berlin Heidelberg, Berlin, Heidelberg. pp. 618–625.

Breiman, L., 2001. Random forests. Machine Learning 45, 5–32. URL: `https://doi.org/10.1023/A:1010933404324`, doi:`10.1023/A:1010933404324`.

Budai, A., Bock, R., Maier, A., Hornegger, J., Michelson, G., 2013. Robust vessel segmentation in fundus images. International journal of biomedical imaging 2013. doi:`https://doi.org/10.1155/2013/154860`.

Cai, W., Chung, A.C.S., 2006. Multi-resolution vessel segmentation using normalized cuts in retinal images, in: Larsen, R., Nielsen, M., Sporring, J. (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2006, Springer Berlin Heidelberg, Berlin, Heidelberg. pp. 928–936.

Cao, W., Czarnek, N., Shan, J., Li, L., 2018. Microaneurysm detection using principal component analysis and machine learning methods. IEEE Transactions on NanoBioscience 17, 191–198. doi:`10.1109/TNB.2018.2840084`.

Casanova, R., Saldana, S., Chew, E.Y., Danis, R.P., Greven, C.M., Ambrosius, W.T., 2014. Application of random forests methods to diabetic retinopathy classification analyses. PLOS ONE 9, 1–8. URL: `https://doi.org/10.1371/journal.pone.0098587`, doi:`10.1371/journal.pone.0098587`.

Chen, L.C., Zhu, Y., Papandreou, G., Schroff, F., Adam, H., 2018. Encoder-decoder with atrous separable convolution for semantic image segmentation, in: Proceedings of the European conference on computer vision (ECCV), pp. 801–818.

Cheung, C., Wong, T., 2018. Clinical use of optical coherence tomography angiography in diabetic retinopathy treatment: Ready for showtime? JAMA Ophthalmology 136, 729–730. URL: `+http://dx.doi.org/10.1001/jamaophthalmol.2018.1538`, doi:`10.1001/jamaophthalmol.2018.1538`.

Chudzik, P., Majumdar, S., Caliva, F., Al-Diri, B., Hunter, A., 2018a. Microaneurysm detection using deep learning and interleaved freezing. URL: `https://doi.org/10.1117/12.2293520`, doi:`10.1117/12.2293520`.

Chudzik, P., Majumdar, S., Calivá, F., Al-Diri, B., Hunter, A., 2018b. Microaneurysm detection using fully convolutional neural networks. Computer

Methods and Programs in Biomedicine 158, 185 – 192. URL: `http://www.sciencedirect.com/science/article/pii/S0169260717308544`, doi:`https://doi.org/10.1016/j.cmpb.2018.02.016`.

Cicinelli, M.V., Cavalleri, M., Brambati, M., Lattanzio, R., Bandello, F., 2019. New imaging systems in diabetic retinopathy. Acta diabetologica 56, 981–994.

Costa, P., Araújo, T., Aresta, G., Galdran, A., Mendonça, A.M., Smailagic, A., Campilho, A., 2019. Eyewes: Weakly supervised pre-trained convolutional neural networks for diabetic retinopathy detection, in: 2019 16th International Conference on Machine Vision Applications (MVA), pp. 1–6. doi:`10.23919/MVA.2019.8757991`.

Costa, P., Galdran, A., Smailagic, A., Campilho, A., 2018. A weakly-supervised framework for interpretable diabetic retinopathy detection on retinal images. IEEE Access 6, 18747–18758. doi:`10.1109/ACCESS.2018.2816003`.

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 Transactions on Medical Imaging 37, 1149–1161. doi:`10.1109/TMI.2018.2794988`.

Dashtbozorg, B., Zhang, J., Huang, F., ter Haar Romeny, B.M., 2018. Retinal microaneurysms detection using local convergence index features. IEEE Transactions on Image Processing 27, 3300–3315. doi:`10.1109/TIP.2018.2815345`.

Decencière, E., Cazuguel, G., Zhang, X., Thibault, G., Klein, J.C., Meyer, F., Marcotegui, B., Quellec, G., Lamard, M., Danno, R., Elie, D., Massin, P., Viktor, Z., Erginay, A., Laÿ, B., Chabouis, A., 2013. Teleophta: Machine learning and image processing methods for teleophthalmology. IRBM 34, 196 – 203. URL: `http://www.sciencedirect.com/science/article/pii/S1959031813000237`, doi:`https://doi.org/10.1016/j.irbm.2013.01.010`. special issue : ANR TECSAN : Technologies for Health and Autonomy.

Decencière, E., Zhang, X., Cazuguel, G., Lay, B., Cochener, B., Trone, C., Gain, P., Ordonez, R., Massin, P., Erginay, A., Charton, B., Klein, J.C., 2014. Feedback on a publicly distributed database: the messidor database. Image Analysis & Stereology 33, 231–234. URL: `http://www.ias-iss.org/ojs/IAS/article/view/1155`, doi:`10.5566/ias.1155`.

Deepa, V., Sathish Kumar, C., Susan Andrews, S., 2019. Automated detection of microaneurysms using stockwell transform and statistical features. IET Image Processing 13, 1341–1348. doi:`10.1049/iet-ipr.2018.5672`.

Deepak, K.S., Sivaswamy, J., 2012. Automatic assessment of macular edema from color retinal images. IEEE Transactions on Medical Imaging 31, 766–776. doi:`10.1109/TMI.2011.2178856`.

Ege, B.M., Hejlesen, O.K., Larsen, O.V., Møller, K., Jennings, B., Kerr, D., Cavan, D.A., 2000. Screening for diabetic retinopathy using computer based image analysis and statistical classification. Computer Methods and Programs in Biomedicine 62, 165 – 175. URL: `http://www.sciencedirect.com/science/article/pii/S0169260700000651`, doi:`https://doi.org/10.1016/S0169-2607(00)00065-1`.

ElTanboly, A., Ghazaf, M., Khalil, A., Shalaby, A., Mahmoud, A., Switala, A., El-Azab, M., Schaal, S., El-Baz, A., 2018. An integrated framework for automatic clinical assessment of diabetic retinopathy grade using spectral domain oct images, in: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), pp. 1431–1435. doi:`10.1109/ISBI.2018.8363841`.

ETDRS Research Group, 1991. Grading diabetic retinopathy from stereoscopic color fundus photographs—an extension of the modified airlie house classification: Etdrs report number 10. Ophthalmology 98, 786 – 806. doi:`https://doi.org/10.1016/S0161-6420(13)38012-9`.

Falavarjani, K.G., Tsui, I., Sadda, S.R., 2017. Ultra-wide-field imaging in diabetic retinopathy. Vision Research 139, 187 – 190. URL: `http://www.sciencedirect.com/science/article/pii/S0042698917301189`, doi:`https://doi.org/10.1016/j.visres.2017.02.009`. diabetic Retinopathy - an Overview.

Farsiu, S., Chiu, S.J., Oconnell, R.V., Folgar, F.A., Yuan, E., Izatt, J.A., Toth, C.A., 2014. Quantitative classification of eyes with and without intermediate age-related macular degeneration using optical coherence tomography. Ophthalmology 121, 162–172. doi:`10.1016/j.ophtha.2013.07.013`.

Feng, S., Zhuo, Z., Pan, D., Tian, Q., 2019. Ccnet: A cross-connected convolutional network for segmenting retinal vessels using multi-scale features. Neurocomputing URL: `http://www.sciencedirect.com/science/article/pii/S0925231219304655`, doi:`https://doi.org/10.1016/j.neucom.2018.10.098`.

Frazao, L.B., Theera-Umpon, N., Auephanwiriyakul, S., 2019. Diagnosis of diabetic retinopathy based on holistic texture and local retinal features. Information Sciences 475, 44 – 66. URL: `http://www.sciencedirect.com/science/article/pii/S0020025518307849`, doi:`https://doi.org/10.1016/j.ins.2018.09.064`.

Gao, Z., Li, J., Guo, J., Chen, Y., Yi, Z., Zhong, J., 2019. Diagnosis of diabetic retinopathy using deep neural networks. IEEE Access 7, 3360–3370. doi:`10.1109/ACCESS.2018.2888639`.

Gardner, G., Keating, D., Williamson, T.H., Elliott, A.T., 1996a. Automatic detection of diabetic retinopathy using an artificial neural network: a screening tool. British journal of Ophthalmology 80, 940–944. doi:`10.1136/bjo.80.11.940`.

Gardner, G.G., Keating, D., Williamson, T.H., Elliott, A.T., 1996b. Automatic detection of diabetic retinopathy using an artificial neural network: a screening tool. British Journal of Ophthalmology 80, 940–944. URL: `https://bjo.bmj.com/content/80/11/940`, doi:`10.1136/bjo.80.11.940`.

Gargeya, R., Leng, T., 2017. Automated identification of diabetic retinopathy using deep learning. Ophthalmology 124, 962 – 969. URL: `http://www.sciencedirect.com/science/article/pii/S0161642016317742`, doi:`https://doi.org/10.1016/j.ophtha.2017.02.008`.

Gholami, P., Roy, P., Parthasarathy, M.K., Lakshminarayanan, V., 2020a. Octid: Optical coherence tomography image database. Computers Electrical Engineering 81, 106532. doi:`10.1016/j.compeleceng.2019.106532`.

Gholami, P., Roy, P., Parthasarathy, M.K., Lakshminarayanan, V., 2020b. Octid: Optical coherence tomography image database. Computers Electrical Engineering 81, 106532. doi:`10.1016/j.compeleceng.2019.106532`.

Giancardo, L., Meriaudeau, F., Karnowski, T.P., Li, Y., Garg, S., Tobin, K.W., Chaum, E., 2012. Exudate-based diabetic macular edema detection in fundus images using publicly available datasets. Medical Image Analysis 16, 216 – 226. URL: `http://www.sciencedirect.com/science/article/pii/S1361841511001010`, doi:`https://doi.org/10.1016/j.media.2011.07.004`.

Giancardo, L., Meriaudeau, F., Karnowski, T.P., Li, Y., Tobin, K.W., Chaum, E., 2011. Microaneurysm detection with radon transform-based classification on retina images, in: 2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society, pp. 5939–5942. doi:`10.1109/IEMBS.2011.6091562`.

Goatman, K.A., Fleming, A.D., Philip, S., Williams, G.J., Olson, J.A., Sharp, P.F., 2011. Detection of new vessels on the optic disc using retinal photographs. IEEE Transactions on Medical Imaging 30, 972–979. doi:`10.1109/TMI.2010.2099236`.

González-Gonzalo, C., Sánchez-Gutiérrez, V., Hernández-Martínez, P., Contreras, I., Lechanteur, Y.T.E., Domanian, A., van Ginneken, B., Sánchez, C.I., 2019. Evaluation of a deep learning system for the joint automated detection of diabetic retinopathy and age-related macular degeneration. CoRR abs/1903.09555. URL: `http://arxiv.org/abs/1903.09555`.

Gulshan, V., Peng, L., Coram, M., et al, 2016a. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 316, 2402–2410. URL: `+http://dx.doi.org/10.1001/jama.2016.17216`, doi:`10.1001/jama.2016.17216`.

Gulshan, V., Peng, L., Coram, M., Stumpe, M.C., Wu, D., Narayanaswamy, A., Venugopalan, S., Widner, K., Madams, T., Cuadros, J., et al., 2016b. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. Jama 316, 2402–2410.

Gulshan, V., Rajan, R.P., Widner, K., Wu, D., Wubbels, P., Rhodes, T., Whitehouse, K., Coram, M., Corrado, G., Ramasamy, K., Raman, R., Peng, L., Webster, D.R., 2019. Performance of a Deep-Learning Algorithm vs Manual Grading for Detecting Diabetic Retinopathy in India. JAMA Ophthalmology URL: `https://doi.org/10.1001/jamaophthalmol.2019.2004`, doi:`10.1001/jamaophthalmol.2019.2004`.

Hagos, M.T., Kant, S., 2019. Transfer learning based detection of diabetic retinopathy from small dataset. CoRR abs/1905.07203. URL: `http://arxiv.org/abs/1905.07203`.

Hajeb Mohammad Alipour, S., Rabbani, H., Akhlaghi, M., 2014. A new combined method based on curvelet transform and morphological operators for automatic detection of foveal avascular zone. Signal, Image and Video Processing 8, 205–222. URL: `https://doi.org/10.1007/s11760-013-0530-6`, doi:`10.1007/s11760-013-0530-6`.

Haloi, M., 2015. Improved microaneurysm detection using deep neural networks. CoRR abs/1505.04424. URL: `http://arxiv.org/abs/1505.04424`.

Hatanaka, Y., Inoue, T., Okumura, S., Muramatsu, C., Fujita, H., 2012. Automated microaneurysm detection method based on double-ring filter and feature analysis in retinal fundus images, in: 2012 25th IEEE International Symposium on Computer-Based Medical Systems (CBMS), pp. 1–4. doi:`10.1109/CBMS.2012.6266339`.

Hatanaka, Y., Ogohara, K., Sunayama, W., Miyashita, M., Muramatsu, C., Fujita, H., 2018. Automatic microaneurysms detection on retinal images using deep convolution neural network, in: 2018 International Workshop on Advanced Image Technology (IWAIT), pp. 1–2. doi:`10.1109/IWAIT.2018.8369794`.

He, Y., Jiao, W., Shi, Y., Zhao, B., Zou, W., Lian, J., Zhu, Y., Zheng, Y., 2019. Segmenting diabetic retinopathy lesions in multispectral images using low-dimensional spatial-spectral matrix representation. IEEE Journal of Biomedical and Health Informatics , 1–1doi:`10.1109/JBHI.2019.2912668`.

Healgoo Interactive, 2017. Label me. URL: `http://www.labelme.org/`.

Herliana, A., Arifin, T., Susanti, S., Hikmah, A.B., 2018. Feature selection of diabetic retinopathy disease using particle swarm optimization and neural network, in: 2018 6th International Conference on Cyber and IT Service Management (CITSM), pp. 1–4. doi:`10.1109/CITSM.2018.8674295`.

Hernandez-Matas, C., Zabulis, X., Triantafyllou, A., Anyfanti, P., Douma, S., Argyros, A.A., 2017. Fire: fundus image registration dataset. Journal for Modeling in Ophthalmology 1, 16–28.

Hipwell, J.H., Strachan, F., Olson, J.A., McHardy, K.C., Sharp, P.F., Forrester, J.V., . Automated detection of microaneurysms in digital red-free photographs: a diabetic retinopathy screening tool. Diabetic Medicine 17, 588–594. URL: `https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1464-5491.2000.00338.x`, doi:`10.1046/j.1464-5491.2000.00338.x`.

Hoover, A., Goldbaum, M., . Structured analysis of the retina. URL: `http://cecas.clemson.edu/~ahoover/stare/`.

Huang, D., Swanson, E., Lin, C., Schuman, J., Stinson, W., Chang, W., Hee, M., Flotte, T., Gregory, K., Puliafito, C., et, a., 1991. Optical coherence tomography 254, 1178–1181. URL: `http://science.sciencemag.org/content/254/5035/1178`, doi:`10.1126/science.1957169`.

Hwang, T., Hagag, A., Wang, J., et al, 2018. Automated quantification of nonperfusion areas in 3 vascular plexuses with optical coherence tomography angiography in eyes of patients with diabetes. JAMA Ophthalmology URL: `+http://dx.doi.org/10.1001/jamaophthalmol.2018.2257`, doi:`10.1001/jamaophthalmol.2018.2257`.

Jebaseeli, T.J., Durai, C.A.D., Peter, J.D., 2019. Segmentation of retinal blood vessels from ophthalmologic diabetic retinopathy images. Computers Electrical Engineering 73, 245 – 258. URL: `http://www.sciencedirect.com/science/`

article/pii/S004579061732949X, doi:https://doi.org/10.1016/j.compeleceng.2018.11.024.

Jiang, Y., Tan, N., Peng, T., Zhang, H., 2019. Retinal vessels segmentation based on dilated multi-scale convolutional neural network. IEEE Access 7, 76342–76352. doi:10.1109/ACCESS.2019.2922365.

Jin, Q., Meng, Z., Pham, T.D., Chen, Q., Wei, L., Su, R., 2019. Dunet: A deformable network for retinal vessel segmentation. Knowledge-Based Systems 178, 149 – 162. URL: http://www.sciencedirect.com/science/article/pii/S0950705119301984, doi:https://doi.org/10.1016/j.knosys.2019.04.025.

Kaggle, 2015a. Diabetic retinopathy detection. URL: https://www.kaggle.com/c/diabetic-retinopathy-detection/overview.

Kaggle, 2015b. Diabetic retinopathy detection. URL: https://www.kaggle.com/c/diabetic-retinopathy-detection/discussion/15617.

Kar, S.S., Maity, S.P., 2018. Automatic detection of retinal lesions for screening of diabetic retinopathy. IEEE Transactions on Biomedical Engineering 65, 608–618. doi:10.1109/TBME.2017.2707578.

Karim, T., Riad, M., Kabir, R., 2019. Symptom analysis of diabetic retinopathy by micro-aneurysm detection using nprtool, in: 2019 International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST), pp. 606–610. doi:10.1109/ICREST.2019.8644439.

Kauppi, T., Kalesnykiene, V., Kamarainen, J.K., Lensu, L., Sorri, I., Raninen, A., Voutilainen, R., Uusitalo, H., Kälviäinen, H., Pietilä, J., 2007. The diaretdb1 diabetic retinopathy database and evaluation protocol., in: BMVC, pp. 1–10.

Kauppi, T., Kalesnykiene, V., Kamarainen, J.K., Lensu, L., Sorri, I., Uusitalo, H., Kälviäinen, H., Pietilä, J., 2006. Diaretdb0: Evaluation database and methodology for diabetic retinopathy algorithms. Machine Vision and Pattern Recognition Research Group, Lappeenranta University of Technology, Finland 73.

Kaya, C., Erkaymaz, O., Ayar, O., Özer, M., 2018. Impact of hybrid neural network on the early diagnosis of diabetic retinopathy disease from video-oculography signals. Chaos, Solitons Fractals 114, 164 – 174. URL: http://www.sciencedirect.com/science/article/pii/S0960077918304909, doi:https://doi.org/10.1016/j.chaos.2018.06.034.

Keel, S., Lee, P.Y., Li, Z., Kotowicz, M.A., MacIsaac, R.J., He, M., 2018. Feasibility and patient acceptability of a novel artificial intelligence-based screening model for diabetic retinopathy at endocrinology outpatient services: a pilot study. Scientific Reports 8. URL: https://doi.org/10.1038/s41598-018-22612-2.

Keel, S., Wu, J., Lee, P.Y., Scheetz, J., He, M., 2019. Visualizing Deep Learning Models for the Detection of Referable Diabetic Retinopathy and Glaucoma. JAMA Ophthalmology 137, 288–292. URL: https://doi.org/10.1001/jamaophthalmol.2018.6035, doi:10.1001/jamaophthalmol.2018.6035.

Keller, J.M., Gray, M.R., Givens, J.A., 1985. A fuzzy k-nearest neighbor algorithm. IEEE Transactions on Systems, Man, and Cybernetics SMC-15, 580–585. doi:10.1109/TSMC.1985.6313426.

Khansari, M.M., Zhang, J., Qiao, Y., Gahm, J.K., Sarabi, M.S., Kashani, A.H., Shi, Y., 2019. Automated deformation-based analysis of 3d optical coherence tomography in diabetic retinopathy. IEEE Transactions on Medical Imaging , 1–1doi:10.1109/TMI.2019.2924452.

Khojasteh, P., Júnior, L.A.P., Carvalho, T., Rezende, E., Aliahmad, B., Papa, J.P., Kumar, D.K., 2019. Exudate detection in fundus images using deeply-learnable features. Computers in Biology and Medicine 104, 62 – 69. URL: http://www.sciencedirect.com/science/article/pii/S0010482518303330, doi:https://doi.org/10.1016/j.compbiomed.2018.10.031.

Krause, J., Gulshan, V., Rahimy, E., Karth, P., Widner, K., Corrado, G.S., Peng, L., Webster, D.R., 2018a. Grader variability and the importance of reference standards for evaluating machine learning models for diabetic retinopathy. Ophthalmology 125, 1264 – 1272. URL: http://www.sciencedirect.com/science/article/pii/S0161642017326982, doi:https://doi.org/10.1016/j.ophtha.2018.01.034.

Krause, J., Gulshan, V., Rahimy, E., Karth, P., Widner, K., Corrado, G.S., Peng, L., Webster, D.R., 2018b.

Grader variability and the importance of reference standards for evaluating machine learning models for diabetic retinopathy. Ophthalmology 125, 1264–1272.

Krizhevsky, A., Sutskever, I., Hinton, G.E., 2012. Imagenet classification with deep convolutional neural networks, in: Pereira, F., Burges, C.J.C., Bottou, L., Weinberger, K.Q. (Eds.), Advances in Neural Information Processing Systems 25. Curran Associates, Inc., pp. 1097–1105. doi:`10.1145/3065386`.

Kumar, D., Taylor, G.W., Wong, A., 2019. Discovery radiomics with clear-dr: Interpretable computer aided diagnosis of diabetic retinopathy. IEEE Access 7, 25891–25896. doi:`10.1109/ACCESS.2019.2893635`.

Lee, S.C., Lee, E.T., Kingsley, R.M., Wang, Y., Russell, D., Klein, R., Warn, A., 2001. Comparison of Diagnosis of Early Retinal Lesions of Diabetic Retinopathy Between a Computer System and Human Experts. JAMA Ophthalmology 119, 509–515. URL: `https://doi.org/10.1001/archopht.119.4.509`, doi:`10.1001/archopht.119.4.509`.

Li, H.K., Esquivel, A., Hubbard, L.D., Florez-Arango, J.F., Danis, R.P., Krupinski, E.A., 2011. MOSAICS VERSUS EARLY TREATMENT DIABETIC RETINOPATHY SEVEN STANDARD FIELDS FOR EVALUATION OF DIABETIC RETINOPATHY SEVERITY. Retina 31, 1553–1563. doi:`10.1097/IAE.0b013e3182084273`.

Li, T., Gao, Y., Wang, K., Guo, S., Liu, H., Kang, H., 2019. Diagnostic assessment of deep learning algorithms for diabetic retinopathy screening. Information Sciences URL: `http://www.sciencedirect.com/science/article/pii/S0020025519305377`, doi:`https://doi.org/10.1016/j.ins.2019.06.011`.

Li, Z., He, Y., Keel, S., Meng, W., Chang, R.T., He, M., 2018. Efficacy of a deep learning system for detecting glaucomatous optic neuropathy based on color fundus photographs. Ophthalmology 125, 1199 – 1206. URL: `http://www.sciencedirect.com/science/article/pii/S0161642017335650`, doi:`https://doi.org/10.1016/j.ophtha.2018.01.023`.

Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., van der Laak, J.A., van Ginneken, B., Sánchez, C.I., 2017. A survey on deep learning in medical image analysis. Medical Image Analysis 42, 60 – 88. URL: `http://www.sciencedirect.com/science/article/pii/S1361841517301135`, doi:`https://doi.org/10.1016/j.media.2017.07.005`.

Liu, Y.P., Li, Z., Xu, C., Li, J., Liang, R., 2019. Referable diabetic retinopathy identification from eye fundus images with weighted path for convolutional neural network. Artificial Intelligence in Medicine URL: `http://www.sciencedirect.com/science/article/pii/S0933365718307747`, doi:`https://doi.org/10.1016/j.artmed.2019.07.002`.

Lokuarachchi, D., Gunarathna, K., Muthumal, L., Gamage, T., 2019. Automated detection of exudates in retinal images, in: 2019 IEEE 15th International Colloquium on Signal Processing Its Applications (CSPA), pp. 43–47. doi:`10.1109/CSPA.2019.8696052`.

Lowell, J., Hunter, A., Steel, D., Basu, A., Ryder, R., Fletcher, E., Kennedy, L., 2004. Optic nerve head segmentation. IEEE Transactions on Medical Imaging 23, 256–264. doi:`10.1109/TMI.2003.823261`.

Mahfouz, A.E., Fahmy, A.S., 2009. Ultrafast localization of the optic disc using dimensionality reduction of the search space, in: Yang, G.Z., Hawkes, D., Rueckert, D., Noble, A., Taylor, C. (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2009, Springer Berlin Heidelberg, Berlin, Heidelberg. pp. 985–992.

Mansour, R.F., 2018. Deep-learning-based automatic computer-aided diagnosis system for diabetic retinopathy. Biomedical Engineering Letters 8, 41–57. URL: `https://doi.org/10.1007/s13534-017-0047-y`, doi:`10.1007/s13534-017-0047-y`.

Marin, D., Aquino, A., Gegundez-Arias, M.E., Bravo, J.M., 2011. A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features. IEEE Transactions on Medical Imaging 30, 146–158.

Mobeen-ur-Rehman, Khan, S.H., Abbas, Z., Danish Rizvi, S.M., 2019. Classification of diabetic retinopathy images based on customised cnn architecture, in: 2019 Amity International Conference on Artificial Intelligence (AICAI), pp. 244–248. doi:`10.1109/AICAI.2019.8701231`.

Moghaddam, B., Pentland, A.P., 1994. Face recognition using view-based and modular eigenspaces, in: Automatic Systems for the Identification and Inspection of Humans, International Society for Optics and Photonics. pp. 12–21.

Mohammadpoory, Z., Nasrolahzadeh, M., Mahmoodian, N., Haddadnia, J., 2019. Automatic identification of diabetic retinopathy stages by using fundus images and visibility graph method. Measurement 140, 133 – 141. URL: `http://www.sciencedirect.com/science/article/pii/S0263224119302064`, doi:`https://doi.org/10.1016/j.measurement.2019.02.089`.

Mooney, P., 2018. Retinal oct images (optical coherence tomography). URL: `https://www.kaggle.com/paultimothymooney/kermany2018`.

Nguyen, U.T.V., Bhuiyan, A., Ramamohanarao, K., Park, L.A.F., 2011. An effective supervised framework for retinal blood vessel segmentation using local standardisation and bagging, in: Suzuki, K., Wang, F., Shen, D., Yan, P. (Eds.), Machine Learning in Medical Imaging, Springer Berlin Heidelberg, Berlin, Heidelberg. pp. 117–125.

Niemeijer, M., Abramoff, M.D., van Ginneken, B., 2009. Information fusion for diabetic retinopathy cad in digital color fundus photographs. IEEE Transactions on Medical Imaging 28, 775–785. doi:`10.1109/TMI.2008.2012029`.

Niemeijer, M., Abràmoff, M.D., van Ginneken, B., 2006. Image structure clustering for image quality verification of color retina images in diabetic retinopathy screening. Medical Image Analysis 10, 888 – 898. URL: `http://www.sciencedirect.com/science/article/pii/S1361841506000739`, doi:`https://doi.org/10.1016/j.media.2006.09.006`.

Niemeijer, M., van Ginneken, B., Russell, S.R., Suttorp-Schulten, M.S.A., Abramoff, M.D., 2007a. Automated detection and differentiation of drusen, exudates, and cotton-wool spots in digital color fundus photographs for diabetic retinopathy diagnosis. Investigative Ophthalmology Visual Science 48, 2260. URL: `+http://dx.doi.org/10.1167/iovs.06-0996`, doi:`10.1167/iovs.06-0996`.

Niemeijer, M., van Ginneken, B., Russell, S.R., Suttorp-Schulten, M.S.A., Abramoff, M.D., 2007b. Automated Detection and Differentiation of Drusen, Exudates, and Cotton-Wool Spots in Digital Color Fundus Photographs for Diabetic Retinopathy Diagnosis. Investigative Ophthalmology Visual Science 48, 2260–2267. URL: `https://doi.org/10.1167/iovs.06-0996`, doi:`10.1167/iovs.06-0996`.

Niemeijer, M., van Ginneken, B., Staal, J., Suttorp-Schulten, M.S.A., Abramoff, M.D., 2005. Automatic detection of red lesions in digital color fundus photographs. IEEE Transactions on Medical Imaging 24, 584–592. doi:`10.1109/TMI.2005.843738`.

Niemeijer, M., Staal, J., van Ginneken, B., Loog, M., Abramoff, M.D., 2004. Comparative study of retinal vessel segmentation methods on a new publicly available database. URL: `https://doi.org/10.1117/12.535349`, doi:`10.1117/12.535349`.

Niemeijer, M., van Ginneken, B., Cree, M.J., Mizutani, A., Quellec, G., Sanchez, C.I., Zhang, B., Hornero, R., Lamard, M., Muramatsu, C., Wu, X., Cazuguel, G., You, J., Mayo, A., Li, Q., Hatanaka, Y., Cochener, B., Roux, C., Karray, F., Garcia, M., Fujita, H., Abramoff, M.D., 2010. Retinopathy online challenge: Automatic detection of microaneurysms in digital color fundus photographs. IEEE Transactions on Medical Imaging 29, 185–195. doi:`10.1109/TMI.2009.2033909`.

Noble, W.S., 2006. What is a support vector machine? Nature biotechnology 24, 1565. doi:`https://doi.org/10.1038/nbt1206-1565`.

Noh, K.J., Park, S.J., Lee, S., 2019. Scale-space approximated convolutional neural networks for retinal vessel segmentation. Computer Methods and Programs in Biomedicine 178, 237 – 246. URL: `http://www.sciencedirect.com/science/article/pii/S0169260719302226`, doi:`https://doi.org/10.1016/j.cmpb.2019.06.030`.

Oliveira, C., Cristovao, L., Ribeiro, M., Abreu, J., 2011. Improved automated screening of diabetic retinopathy. Ophthalmologica 91, 191 – 197. doi:`10.1159/000330285`.

Orlando, J.I., Blaschko, M., 2014. Learning fully-connected crfs for blood vessel segmentation in retinal images, in: Golland, P., Hata, N., Barillot, C., Hornegger, J., Howe, R. (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2014, Springer International Publishing, Cham. pp. 634–641.

Osareh, A., Shadgar, B., Markham, R., 2009. A computational-intelligence-based approach for detection of exudates in diabetic retinopathy images. IEEE Transactions on Information Technology in Biomedicine 13, 535–545. doi:`10.1109/TITB.2008.2007493`.

Palacios, D., Shen, K., Baig, S., Wang, J.H., Wang, M.R., 2019. Wide field of view retinal imaging by handheld fundus camera , in: Manns, F., Söderberg, P.G., Ho, A. (Eds.), Ophthalmic Technologies XXIX, International Society for Optics and Photonics. SPIE. pp. 57 – 63. URL: `https://doi.org/10.1117/12.2510544`, doi:`10.1117/12.2510544`.

Philip, S., Fleming, A.D., Goatman, K.A., Fonseca, S., Mcnamee, P., Scotland, G.S., Prescott, G.J., Sharp, P.F., Olson, J.A., 2007. The efficacy of automated "disease/no disease" grading for diabetic retinopathy in a systematic screening programme. British Journal of Ophthalmology 91, 1512–1517. URL: `https://bjo.bmj.com/content/91/11/1512`, doi:`10.1136/bjo.2007.119453`.

Pires, R., Avila, S., Wainer, J., Valle, E., Abramoff, M.D., Rocha, A., 2019. A data-driven approach to referable diabetic retinopathy detection. Artificial Intelligence in Medicine 96, 93 – 106. URL: `http://www.sciencedirect.com/science/article/pii/S0933365718306353`, doi:`https://doi.org/10.1016/j.artmed.2019.03.009`.

Porwal, P., Pachade, S., Kamble, R., Kokare, M., Deshmukh, G., Sahasrabuddhe, V., Meriaudeau, F., 2018a. Indian diabetic retinopathy image dataset (idrid). URL: `http://dx.doi.org/10.21227/H25W98`, doi:`10.21227/H25W98`.

Porwal, P., Pachade, S., Kamble, R., Kokare, M., Deshmukh, G., Sahasrabuddhe, V., Meriaudeau, F., 2018b. Indian diabetic retinopathy image dataset (idrid): A database for diabetic retinopathy screening research. Data 3. URL: `https://www.mdpi.com/2306-5729/3/3/25`, doi:`10.3390/data3030025`.

Prentašić, P., Lončarić, S., Vatavuk, Z., Benčić, G., Subašić, M., Petković, T., Dujmović, L., Malenica-Ravlić, M., Budimlija, N., Tadić, R., 2013. Diabetic retinopathy image database(dridb): A new database for diabetic retinopathy screening programs research, in: 2013 8th International Symposium on Image and Signal Processing and Analysis (ISPA), pp. 711–716. doi:`10.1109/ISPA.2013.6703830`.

Quellec, G., Charrière, K., Boudi, Y., Cochener, B., Lamard, M., 2017. Deep image mining for diabetic retinopathy screening. Medical Image Analysis 39, 178 – 193. URL: `http://www.sciencedirect.com/science/article/pii/S136184151730066X`, doi:`https://doi.org/10.1016/j.media.2017.04.012`.

Quellec, G., Lamard, M., Josselin, P.M., Cazuguel, G., Cochener, B., Roux, C., 2008a. Optimal wavelet transform for the detection of microaneurysms in retina photographs. IEEE Transactions on Medical Imaging 27, 1230–1241. doi:`10.1109/TMI.2008.920619`.

Quellec, G., Lamard, M., Josselin, P.M., Cazuguel, G., Cochener, B., Roux, C., 2008b. Optimal wavelet transform for the detection of microaneurysms in retina photographs. IEEE transactions on medical imaging 27, 1230–1241.

Quellec, G., Lamard, M., Lay, B., Guilcher, A.L., Erginay, A., Cochener, B., Massin, P., 2019. Instant automatic diagnosis of diabetic retinopathy. Image and Video Processing URL: `https://arxiv.org/abs/1906.11875`.

Rabiolo, A., Parravano, M., Querques, L., Cicinelli, M.V., Carnevali, A., Sacconi, R., Centoducati, T., Vujosevic, S., Bandello, F., Querques, G., 2017. Ultra-wide-field fluorescein angiography in diabetic retinopathy: a narrative review. Clinical ophthalmology (Auckland, NZ) 11, 803.

Rasti, R., Rabbani, H., Mehridehnavi, A., Hajizadeh, F., 2018. Macular oct classification using a multi-scale convolutional neural network ensemble. IEEE Transactions on Medical Imaging 37, 1024–1034. doi:`10.1109/TMI.2017.2780115`.

Rathke, F., Desana, M., Schnörr, C., 2017. Locally adaptive probabilistic models for global segmentation of pathological oct scans, in: Descoteaux, M., Maier-Hein, L., Franz, A., Jannin, P., Collins, D.L., Duchesne, S. (Eds.), Medical Image Computing and Computer Assisted Intervention MICCAI 2017, Springer International Publishing, Cham. pp. 177–184.

Raumviboonsuk, P., Krause, J., Chotcomwongse, P., Sayres, R., Raman, R., Widner, K., Campana, B.J.L., Phene, S., Hemarat, K., Tadarati, M., Silpa-Archa, S., Limwattanayingyong, J., Rao, C., Kuruvilla, O., Jung, J., Tan, J., Orprayoon, S., Kangwanwongpaisan, C., Sukumalpaiboon, R., Luengchaichawang, C., Fuangkaew, J., Kongsap, P., Chualinpha, L., Saree, S., Kawinpanitan, S., Mitvongsa, K., Lawanasakol, S., Thepchatri, C., Wongpichedchai, L., Corrado, G., Peng, L., Webster, D., 2019. Deep learning versus human graders for classifying diabetic retinopathy severity in a nationwide screening program. npj Digital Medicine 2. URL: `https://doi.org/10.1038/s41746-019-0099-8`, doi:`10.1038/s41746-019-0099-8`.

Rehman, Z.U., Naqvi, S.S., Khan, T.M., Arsalan, M., Khan, M.A., Khalil, M., 2019. Multi-parametric optic disc segmentation using superpixel based feature classification. Expert Systems with Applications 120, 461 – 473. URL: `http://www.sciencedirect.com/science/article/pii/S095741741830770X`, doi:`https://doi.org/10.1016/j.eswa.2018.12.008`.

Ribeiro, M.L., Nunes, S.G., Cunha-Vaz, J.G., 2013. Microaneurysm turnover at the macula predicts risk of development of clinically significant macular edema in persons with mild nonproliferative diabetic retinopathy. Diabetes Care 36, 1254–1259. URL: `https://care.diabetesjournals.org/content/36/5/1254`, doi:`10.2337/dc12-1491`.

Roy, P., Tennakoon, R., Cao, K., Sedai, S., Mahapatra, D., Maetschke, S., Garnavi, R., 2017. A novel hybrid approach for severity assessment of diabetic retinopathy in colour fundus images, in: 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017), pp. 1078–1082. doi:`10.1109/ISBI.2017.7950703`.

Roychowdhury, S., Koozekanani, D.D., Parhi, K.K., 2014. Dream: Diabetic retinopathy analysis using machine learning. IEEE Journal of Biomedical and Health Informatics 18, 1717–1728. doi:`10.1109/JBHI.2013.2294635`.

Saha, O., Sathish, R., Sheet, D., 2019. Fully convolutional neural network for semantic segmentation of anatomical structure and pathologies in colour fundus images associated with diabetic retinopathy. CoRR abs/1902.03122. URL: `http://arxiv.org/abs/1902.03122`.

Sahlsten, J., Jaskari, J., Kivinen, J., Turunen, L., Jaanio, E., Hietala, K., Kaski, K., 2019. Deep Learning Fundus Image Analysis for Diabetic Retinopathy and Macular Edema Grading. arXiv e-prints , arXiv:1904.08764.

Sanromà, S., Moreno, A., Valls, A., Romero, P., de la Riva, S., Sagarra, R., 2016. Assessment of diabetic retinopathy risk with random forests, in: 24th European symposium on artificial neural networks, computational intelligence and machine learning (ESANN 2016). Bruges, Belgium, pp. 313–318. URL: `https://www.elen.ucl.ac.be/Proceedings/esann/esannpdf/es2016-11.pdf`.

Saranya Rubini, S., Saai Nithil, R., Kunthavai, A., Sharma, A., 2019. Deep convolutional neural network-based diabetic retinopathy detection in digital fundus images, in: Wang, J., Reddy, G.R.M., Prasad, V.K., Reddy, V.S. (Eds.), Soft Computing and Signal Processing, Springer Singapore, Singapore. pp. 201–209.

Sarwinda, D., Siswantining, T., Bustamam, A., 2018. Classification of diabetic retinopathy stages using histogram of oriented gradients and shallow learning, in: 2018 International Conference on Computer, Control, Informatics and its Applications (IC3INA), pp. 83–87. doi:`10.1109/IC3INA.2018.8629502`.

Sayres, R., Taly, A., Rahimy, E., Blumer, K., Coz, D., Hammel, N., Krause, J., Narayanaswamy, A., Rastegar, Z., Wu, D., Xu, S., Barb, S., Joseph, A., Shumski, M., Smith, J., Sood, A.B., Corrado, G.S., Peng, L., Webster, D.R., 2019. Using a deep learning algorithm and integrated gradients explanation to assist grading for diabetic retinopathy. Ophthalmology 126, 552 – 564. URL: `http://www.sciencedirect.com/science/article/pii/S0161642018315756`, doi:`https://doi.org/10.1016/j.ophtha.2018.11.016`.

Selcuk, T., Alkan, A., 2019. Detection of microaneurysms using ant colony algorithm in the early diagnosis of diabetic retinopathy. Medical Hypotheses 129, 109242. URL: `http://www.sciencedirect.com/science/article/pii/S0306987719304098`, doi:`https://doi.org/10.1016/j.mehy.2019.109242`.

Sengupta, S., Sindal, M.D., Baskaran, P., Pan, U., Venkatesh, R., 2019. Sensitivity and specificity of smartphone-based retinal imaging for diabetic retinopathy: A comparative study. Ophthalmology Retina 3, 146 – 153. URL: `http://www.sciencedirect.com/science/article/pii/S2468653018302884`, doi:`https://doi.org/10.1016/j.oret.2018.09.016`.

Shahid, M., Taj, I.A., 2018. Robust retinal vessel segmentation using vessel's location map and frangi enhancement filter. IET Image Processing 12, 494–501. doi:`10.1049/iet-ipr.2017.0457`.

Shanthi, T., Sabeenian, R., 2019. Modified alexnet architecture for classification of diabetic retinopathy images. Computers Electrical Engineering 76, 56 – 64. URL: `http://www.sciencedirect.com/science/article/pii/S0045790618334190`, doi:`https://doi.org/10.1016/j.compeleceng.2019.03.004`.

Sidibé, D., Sankar, S., Lemaître, G., Rastgoo, M., Massich, J., Cheung, C.Y., Tan, G.S., Milea, D., Lamoureux, E., Wong, T.Y., Mériaudeau, F., 2017. An anomaly detection approach for the identification of dme patients using spectral domain optical coherence tomography images. Computer Methods and Programs in Biomedicine 139, 109 – 117. URL: `http://www.sciencedirect.com/science/article/pii/S0169260716305715`, doi:`https://doi.org/10.1016/j.cmpb.2016.11.001`.

Sil Kar, S., Maity, S.P., 2018. Gradation of diabetic retinopathy on reconstructed image using compressed sensing. IET Image Processing 12, 1956–1963. doi:`10.1049/iet-ipr.2017.1013`.

Sinthanayothin, C., Boyce, J.F., Williamson, T.H., Cook, H.L., Mensah, E., Lal, S., Usher, D., . Automated detection of diabetic retinopathy on digital fundus images. Diabetic Medicine 19, 105–112. URL: `https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1464-5491.2002.00613.x`, doi:`10.1046/j.1464-5491.2002.00613.x`.

Solanki, K., Ramachandra, C., Bhat, S., Bhaskaranand, M., Nittala, M., Sadda, S., 2015. Eyeart: Automated, high-throughput, image analysis for diabetic retinopathy screening. Invest Ophthalmology Vision Sciences 56. URL: `https://iovs.arvojournals.org/article.aspx?articleid=2331130`.

Son, J., Shin, J.Y., Kim, H.D., Jung, K.H., Park, K.H., Park, S.J., 2019. Development and validation of deep learning models for screening multiple abnormal findings in retinal fundus images. Ophthalmology URL: `http://www.sciencedirect.com/science/article/pii/S0161642019303744`, doi:`https://doi.org/10.1016/j.ophtha.2019.05.029`.

Soomro, T.A., Afifi, A.J., Gao, J., Hellwich, O., Zheng, L., Paul, M., 2019. Strided fully convolutional neural network for boosting the sensitivity of retinal blood vessels segmentation. Expert Systems with Applications 134, 36 – 52. URL: `http://www.sciencedirect.com/science/article/pii/S0957417419303586`, doi:`https://doi.org/10.1016/j.eswa.2019.05.029`.

Spencer, T., Olson, J.A., McHardy, K.C., Sharp, P.F., Forrester, J.V., 1996. An image-processing strategy for the segmentation and quantification of microaneurysms in fluorescein angiograms of the ocular fundus. Computers and Biomedical Research 29, 284 – 302. URL: `http://www.sciencedirect.com/science/article/pii/S001048099690021X`, doi:`https://doi.org/10.1006/cbmr.1996.0021`.

Staal, J., Abramoff, M., Niemeijer, M., Viergever, M., van Ginneken, B., 2004. Ridge based vessel segmentation in color images of the retina. IEEE Transactions on Medical Imaging 23, 501–509.

Stark, A., . Fda permits marketing of artificial intelligence-based device to detect certain diabetes-related eye problems. URL: `https://www.fda.gov/news-events/press-announcements/fda-permits-marketing-artificial-intelligence\-based-device-detect-certain-diabetes-related\-eye`.

Sun, Y., 2019. The neural network of one-dimensional convolution-an example of the diagnosis of diabetic retinopathy. IEEE Access 7, 69657–69666. doi:`10.1109/ACCESS.2019.2916922`.

Sun, Y., Zhang, D., 2019. Diagnosis and analysis of diabetic retinopathy based on electronic health records doi:`10.1109/ACCESS.2019.2918625`.

Sundararajan, M., Taly, A., Yan, Q., 2017. Axiomatic attribution for deep networks, in: Proceedings of the 34th International Conference on Machine Learning-Volume 70, JMLR. org. pp. 3319–3328.

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, 1–11. URL: `https://doi.org/10.1371/journal.pone.0179790`, doi:`10.1371/journal.pone.0179790`.

Tan, J.H., Fujita, H., Sivaprasad, S., Bhandary, S.V., Rao, A.K., Chua, K.C., Acharya, U.R., 2017. Automated segmentation of exudates, haemorrhages, microaneurysms using single convolutional neural network. Information Sciences 420, 66 – 76. URL: `http://www.sciencedirect.com/science/article/pii/S0020025517308927`, doi:`https://doi.org/10.1016/j.ins.2017.08.050`.

Tang, L., Niemeijer, M., Reinhardt, J.M., Garvin, M.K., Abramoff, M.D., 2013. Splat feature classification with application to retinal hemorrhage detection in fundus images. IEEE Transactions on Medical Imaging 32, 364–375. doi:`10.1109/TMI.2012.2227119`.

Ting, D., Cheung, C., Lim, G., et al, 2017. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. JAMA 318, 2211–2223. URL: `+http://dx.doi.org/10.1001/jama.2017.18152`, doi:`10.1001/jama.2017.18152`.

Tomi, K., Valentina, K., Iiris, S., Asta, R., Raija, V., Joni, K., Lasse, L., Hannu, U., 2009. Diaretdb1 v2. 1-diabetic retinopathy database and evaluation protocol.

de la Torre, J., Valls, A., Puig, D., 2019. A deep learning interpretable classifier for diabetic retinopathy disease grading. Neurocomputing URL: `http://www.sciencedirect.com/science/article/pii/S0925231219304539`, doi:`https://doi.org/10.1016/j.neucom.2018.07.102`.

Tufail, A., Kapetanakis, V., Salas-Vega, S., Egan, C., Rudisill, C., Owen, C., Lee, A., Louw, V., Anderson, J., Liew, G., Bolter, L., Bailey, C., Sadda, S., Taylor, P., Rudnicka, A., 2016. An observational study to assess if automated diabetic retinopathy image assessment software can replace one or more steps of manual imaging grading and to determine their cost-effectiveness. Health Technology Assessment 20. URL: `https://doi.org/10.3310/hta20920`, doi:`10.3310/hta20920`.

Uribe-Valencia, L.J., Martínez-Carballido, J.F., 2019. Automated optic disc region location from fundus images: Using local multi-level thresholding, best channel selection, and an intensity profile model. Biomedical Signal Processing and Control 51, 148 – 161. URL: `http://www.sciencedirect.com/science/article/pii/S1746809419300424`, doi:`https://doi.org/10.1016/j.bspc.2019.02.006`.

Usher, D., Dumskyj, M., Himaga, M., Williamson, T.H., Nussey, S., Boyce, J., . Automated detection of diabetic retinopathy in digital retinal images: a tool for diabetic retinopathy screening. Diabetic Medicine 21, 84–90. URL: `https://onlinelibrary.wiley.com/doi/abs/10.1046/j.1464-5491.2003.01085.x`, doi:`10.1046/j.1464-5491.2003.01085.x`.

VARPA Group, 2019. Public databases: Vicavr database. URL: `http://www.varpa.es/research/ophtalmology.html#vicavr`.

Walter, T., Klein, J., 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 Transactions on Medical Imaging 21, 1236–1243. doi:`10.1109/TMI.2002.806290`.

Walton, O. Bennett, I., Garoon, R.B., Weng, C.Y., Gross, J., Young, A.K., Camero, K.A., Jin, H., Carvounis, P.E., Coffee, R.E., Chu, Y.I., 2016. Evaluation of Automated Teleretinal Screening Program for Diabetic Retinopathy. JAMA Ophthalmology 134, 204–209. URL: `https://doi.org/10.1001/jamaophthalmol.2015.5083`, doi:`10.1001/jamaophthalmol.2015.5083`.

Wan, S., Liang, Y., Zhang, Y., 2018. Deep convolutional neural networks for diabetic retinopathy detection by image classification. Computers Electrical Engineering 72, 274 – 282.

URL: http://www.sciencedirect.com/science/article/pii/S0045790618302556, doi:https://doi.org/10.1016/j.compeleceng.2018.07.042.

Wang, Z., Yin, Y., Shi, J., Fang, W., Li, H., Wang, X., 2017. Zoom-in-net: Deep mining lesions for diabetic retinopathy detection, in: Descoteaux, M., Maier-Hein, L., Franz, A., Jannin, P., Collins, D.L., Duchesne, S. (Eds.), Medical Image Computing and Computer-Assisted Intervention MICCAI 2017, Springer International Publishing, Cham. pp. 267–275.

Watkins, P.J., 2003. Retinopathy. BMJ 326, 924–926. URL: https://www.bmj.com/content/326/7395/924, doi:10.1136/bmj.326.7395.924.

Welikala, R., Dehmeshki, J., Hoppe, A., Tah, V., Mann, S., Williamson, T., Barman, S., 2014. Automated detection of proliferative diabetic retinopathy using a modified line operator and dual classification. Computer Methods and Programs in Biomedicine 114, 247 – 261. URL: http://www.sciencedirect.com/science/article/pii/S0169260714000650, doi:https://doi.org/10.1016/j.cmpb.2014.02.010.

Williamson, T.H., Keating, D., 1998. Telemedicine and computers in diabetic retinopathy screening. British Journal of Ophthalmology 82, 5–6. URL: https://bjo.bmj.com/content/82/1/5, doi:10.1136/bjo.82.1.5.

Winder, R., Morrow, P., McRitchie, I., Bailie, J., Hart, P., 2009. Algorithms for digital image processing in diabetic retinopathy. Computerized Medical Imaging and Graphics 33, 608 – 622. URL: http://www.sciencedirect.com/science/article/pii/S0895611109000810, doi:https://doi.org/10.1016/j.compmedimag.2009.06.003.

Wisaeng, K., Sa-Ngiamvibool, W., 2019. Exudates detection using morphology mean shift algorithm in retinal images. IEEE Access 7, 11946–11958. doi:10.1109/ACCESS.2018.2890426.

Witmer, M.T., Kiss, S., 2013. Wide-field imaging of the retina. Survey of Ophthalmology 58, 143 – 154. URL: http://www.sciencedirect.com/science/article/pii/S0039625712001713, doi:https://doi.org/10.1016/j.survophthal.2012.07.003.

Wong, T.Y., Bressler, N.M., 2016. Artificial Intelligence With Deep Learning Technology Looks Into Diabetic Retinopathy Screening. JAMA 316, 2366–2367. URL: https://doi.org/10.1001/jama.2016.17563, doi:10.1001/jama.2016.17563.

Xie, S., Tu, Z., 2015. Holistically-nested edge detection, in: Proceedings of the IEEE international conference on computer vision, pp. 1395–1403.

Xu, J., Tolliver, D.A., Ishikawa, H., Wollstein, G., Schuman, J.S., 2009. 3d oct retinal vessel segmentation based on boosting learning, in: Dössel, O., Schlegel, W.C. (Eds.), World Congress on Medical Physics and Biomedical Engineering, September 7 - 12, 2009, Munich, Germany, Springer Berlin Heidelberg, Berlin, Heidelberg. pp. 179–182.

Xu, J., Zhang, X., Chen, H., Li, J., Zhang, J., Shao, L., Wang, G., 2018. Automatic analysis of microaneurysms turnover to diagnose the progression of diabetic retinopathy. IEEE Access 6, 9632–9642. doi:10.1109/ACCESS.2018.2808160.

Yang, Y., Li, T., Li, W., Wu, H., Fan, W., Zhang, W., 2017. Lesion detection and grading of diabetic retinopathy via two-stages deep convolutional neural networks, in: Descoteaux, M., Maier-Hein, L., Franz, A., Jannin, P., Collins, D.L., Duchesne, S. (Eds.), Medical Image Computing and Computer-Assisted Intervention MICCAI 2017, Springer International Publishing, Cham. pp. 533–540.

Youssif, A., Ghalwash, A., Ghoneim, A., 2008. Optic disc detection from normalized digital fundus images by means of a vessels' direction matched filter. IEEE Transactions on Medical Imaging 27, 11–18. doi:10.1109/TMI.2007.900326.

Yu, S., Xiao, D., Kanagasingam, Y., 2017. Exudate detection for diabetic retinopathy with convolutional neural networks, in: 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 1744–1747. doi:10.1109/EMBC.2017.8037180.

Yun, W.L., Acharya, U.R., Venkatesh, Y., Chee, C., Min, L.C., Ng, E., 2008. Identification of different stages of diabetic retinopathy using retinal optical images. Information Sciences 178, 106 –

121. URL: `http://www.sciencedirect.com/science/article/pii/S0020025507003635`, doi:`https://doi.org/10.1016/j.ins.2007.07.020`.

Zeng, X., Chen, H., Luo, Y., Ye, W., 2019. Automated diabetic retinopathy detection based on binocular siamese-like convolutional neural network. IEEE Access 7, 30744–30753. doi:`10.1109/ACCESS.2019.2903171`.

Zhang, D., Bu, W., Wu, X., 2017. Diabetic retinopathy classification using deeply supervised resnet, in: 2017 IEEE SmartWorld, Ubiquitous Intelligence Computing, Advanced Trusted Computed, Scalable Computing Communications, Cloud Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI), pp. 1–6. doi:`10.1109/UIC-ATC.2017.8397469`.

Zhang, W., Zhong, J., Yang, S., Gao, Z., Hu, J., Chen, Y., Yi, Z., 2019. Automated identification and grading system of diabetic retinopathy using deep neural networks. Knowledge-Based Systems 175, 12 – 25. URL: `http://www.sciencedirect.com/science/article/pii/S0950705119301303`, doi:`https://doi.org/10.1016/j.knosys.2019.03.016`.

Zhang, X., Thibault, G., Decencière, E., Marcotegui, B., Laÿ, B., Danno, R., Cazuguel, G., Quellec, G., Lamard, M., Massin, P., Chabouis, A., Victor, Z., Erginay, A., 2014. Exudate detection in color retinal images for mass screening of diabetic retinopathy. Medical Image Analysis 18, 1026 – 1043. URL: `http://www.sciencedirect.com/science/article/pii/S1361841514000693`, doi:`https://doi.org/10.1016/j.media.2014.05.004`.

Zheng, Y., Stambolian, D., O'Brien, J., Gee, J.C., 2013. Optic disc and cup segmentation from color fundus photograph using graph cut with priors, in: Mori, K., Sakuma, I., Sato, Y., Barillot, C., Navab, N. (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013, Springer Berlin Heidelberg, Berlin, Heidelberg. pp. 75–82.

Zhou, L., Zhao, Y., Yang, J., Yu, Q., Xu, X., 2018. Deep multiple instance learning for automatic detection of diabetic retinopathy in retinal images. IET Image Processing 12, 563–571. doi:`10.1049/iet-ipr.2017.0636`.

Zhou, W., Wu, C., Chen, D., Yi, Y., Du, W., 2017. Automatic microaneurysm detection using the sparse principal component analysis-based unsupervised classification method. IEEE Access 5, 2563–2572. doi:`10.1109/ACCESS.2017.2671918`.

Zilly, J.G., Buhmann, J.M., Mahapatra, D., 2015. Boosting convolutional filters with entropy sampling for optic cup and disc image segmentation from fundus images, in: Zhou, L., Wang, L., Wang, Q., Shi, Y. (Eds.), Machine Learning in Medical Imaging, Springer International Publishing, Cham. pp. 136–143.