


# Machine Learning Techniques for Ophthalmic Data Processing: A Review

Mhd Hasan Sarhan , M. Ali Nasser, Daniel Zapp, Mathias Maier, Chris P. Lohmann, Nassir Navab, and Abouzar Eslami

**Abstract**—Machine learning and especially deep learning techniques are dominating medical image and data analysis. This article reviews machine learning approaches proposed for diagnosing ophthalmic diseases during the last four years. Three diseases are addressed in this survey, namely diabetic retinopathy, age-related macular degeneration, and glaucoma. The review covers over 60 publications and 25 public datasets and challenges related to the detection, grading, and lesion segmentation of the three considered diseases. Each section provides a summary of the public datasets and challenges related to each pathology and the current methods that have been applied to the problem. Furthermore, the recent machine learning approaches used for retinal vessels segmentation, and methods of retinal layers and fluid segmentation are reviewed. Two main imaging modalities are considered in this survey, namely color fundus imaging, and optical coherence tomography. Machine learning approaches that use eye measurements and visual field data for glaucoma detection are also included in the survey. Finally, the authors provide their views, expectations and the limitations of the future of these techniques in the clinical practice.

**Index Terms**—Ophthalmic diagnostics, deep learning, diabetic retinopathy, age-related macular degeneration, glaucoma.

## I. INTRODUCTION

**D**URING the past years, machine learning has been widely employed in both ophthalmic data analysis [1], [2], and ophthalmic image analysis [3]–[5]. The application of machine learning for ophthalmic image analysis covered in this survey

broadly falls into two categories, 1) Detection and grading of diseases where the objective is to classify the disease found in an image or a set of images [6]. Further works use multi-class classification to grade the level of severity of a given pathology rather than merely detecting it [7]; 2) Segmentation of anatomical structures from pixel-wise labels, such as segmentation of retinal vessels [3], haemorrhages [8], drusen [9] or optic disk [10]. Analyzing images with machine learning systems involves extraction of features that represent the input data in a way suitable for the classification or segmentation task. Features are either hand-crafted, i.e. defined manually for example by an algorithm developer, or they are extracted with a predefined filter banks, or ultimately learned during training to perform suitable task-specific processing of the input data. The latter is referred to as automatic feature extraction. Recently, after the success of AlexNet in image classification [11], automatic feature extraction using Convolutional Neural Networks (CNNs) surfaced as the preferred feature extraction method in medical image analysis [12] and it is widely used for ophthalmic image analysis [13]–[15]. In this review, publications about both feature extraction methods (handcrafted and automatic) are discussed. Methods of Ophthalmic data Analysis for glaucoma analysis based on multiple anatomical and functional measurement are also reviewed in this paper [1], [2].

Classification and segmentation works related to three diseases namely diabetic retinopathy, age-related macular degeneration, and glaucoma are considered in the this review. Furthermore, the recent machine learning approaches used for retinal vessels segmentation, and methods of retinal layers and fluid segmentation are reviewed. Two main image modalities that are widely used to investigate ophthalmic related pathologies are considered in this survey. (1) **Fundus Color** images that show the retinal vasculature, optic disk, macula, and various abnormalities related to ophthalmic diseases. (2) **Optical Coherence Tomography (OCT)** imaging is used to acquire high-resolution cross-sectional scans of the retinal layers in the eye's posterior segment or the anterior segment structure. Only the former is considered in this review. As a subcategory of OCT, few works on **OCT Angiography (OCTA)** are considered in the survey. OCTA is used to examine the microvasculature of the retina by differentiating blood flow from static tissue in OCT over time. Fig. 1 shows fundus and OCT example images with pathologies.

Various works in the literature have been proposed to review deep learning for ophthalmology or a specific ophthalmic disease [16]–[19]. To our knowledge, none of the existing works

Manuscript received December 9, 2019; revised May 7, 2020 and June 20, 2020; accepted July 18, 2020. Date of publication July 28, 2020; date of current version December 4, 2020. The work of Mhd Hasan Sarhan was supported by Carl Zeiss Meditec. (Corresponding author: Mhd Hasan Sarhan.)

Mhd Hasan Sarhan is with the Translational Research Lab, Carl Zeiss Meditec, 85748 Munich, Germany.

Computer Aided Medical Procedures Chair, Technical University of Munich, 80333 Munich, Germany (e-mail: hasan.sarhan@tum.de).

M. Ali Nasser, Daniel Zapp, Mathias Maier, and Chris P. Lohmann are with the Department of Ophthalmology, Klinikum rechts der Isar, Technical University of Munich, 80333 Munich, Germany (e-mail: ali.nasser@mri.tum.de; daniel.zapp@mri.tum.de; mathias.maier@mri.tum.de; chris.lohmann@mri.tum.de).

Nassir Navab is with the Computer Aided Medical Procedures chair, Technical University of Munich, 80333 Munich, Germany.

Whiting School of Engineering, Johns Hopkins University, Baltimore, MD 21218 USA (e-mail: nassir.navab@tum.de).

Abouzar Eslami is with the Translational Research Lab, Carl Zeiss Meditec, 80333 Munich, Germany (e-mail: abouzar.eslami@zeiss.com).

Digital Object Identifier 10.1109/JBHI.2020.3012134

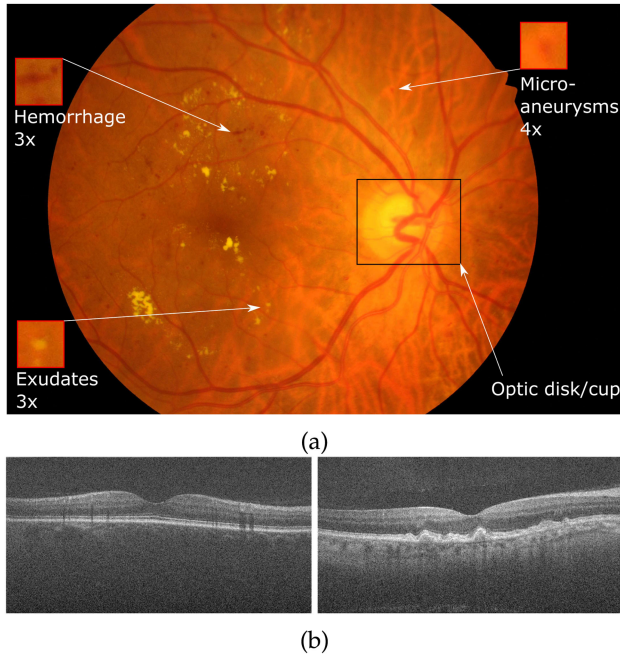


Fig. 1. (a) is an example fundus image from IDRiD dataset with retinal lesions and optic disk/cup highlighted, (b) left is a healthy OCT B-scan cross-section that shows retinal layers, and right is an OCT B-scan of patient with AMD signs [20].

combine technical details with disease/application categorization and list all datasets used in the works to make the reader more aware of the limitations and strengths of the discussed works. This is beneficial for readers from technical backgrounds who are interested in the ophthalmic data analysis domain and require an overview of the advances, the common diseases and applications, the available datasets, and the current limitations. To select papers in this review, we have looked into the most impactful papers in the past 4 years that appeared in ophthalmology/medical high impact journals/conferences. From these works, we branched to other works and studied them as well and kept branching in the publications graph. We have reached over 100 submissions in total and filtered some of them for the sake of capacity and impact. We included papers that have used a new methodology that has not been seen before to our knowledge. We eventually settled on more than 60 works that are presented in the manuscript.

In each of the following sections, we focus on an ophthalmic pathology and list the major machine learning publications relevant to it with a summary of the main contributions. We then select a number of these methods and discuss them in more details. The results of each method are presented alongside the used datasets in a summary table. Moreover, the public datasets and challenges relevant to the presented publications are presented in tabular format. The works since 2015 are considered in these sections except for the used public datasets which could be older.

For reporting the results of works included in the review, we have used the metrics reported by the included papers. Whenever it is feasible, we report the area under the curve

(AUC) of the receiver operating characteristic (ROC) curve because it does not require a cut-off point which gives a more general idea about the classifier performance. For segmentation approaches, we report dice or intersection over union (IOU) whenever feasible. Dice and IOU measure the overlap area between the gold standard segmentation and predicted one. Accuracy, sensitivity, specificity, precision, recall or kappa scores are reported otherwise depending on the paper content. Section II focuses on diabetic retinopathy, detection of the disease and segmentation of various lesions are covered. Section III deals with Age-related Macular Degeneration (AMD) detection and segmentation tasks. Section IV presents glaucoma detection and segmentation. Section V looks at the task of retinal vessel segmentation in fundus and OCT-A images, and in Section VI we focus on retinal layer and fluid segmentation in OCT scans. At the end, Section VII discusses the limitations and future of machine learning in ophthalmology. A summary of the paper is finally provided in Section VIII.

## II. DIABETIC RETINOPATHY

Diabetic Retinopathy (DR) is the leading cause of visual impairment in middle-aged groups [21]. Screening for early detection and referral is the widely adopted strategy in practice for blindness prevention. It is still challenging to implement screening programs as needed because of financial issues and the availability of human monitoring to all sites. Hence, the need for reliable automatic systems for detecting DR emerges. The capillaries start to weaken even at the early stages of DR. This causes abnormalities in the microvascular structure leading to microaneurysms (MA), which appear as small red dots on the retina. These weak capillaries may rupture causing hemorrhages (HA), which appear as larger darker red areas. Moreover, DR makes capillaries more permeable causing hard exudates (HE). When the capillary support fails, pale areas with soft edges appear as soft exudates (SE). In more severe cases, dilation appears in the intraretinal microvascular structure. New fragile small vessels start growing in response to the signals from tissues that lack nutrition which could lead to a sudden loss of vision when growing into the vitreous humour. This process is known as neovascularisation [22]. In the following subsections, literature related to the tasks of detection and grading of diabetic retinopathy, segmentation of lesions that characterize diabetic retinopathy, and attempts to use the detection and grading results for weakly supervised lesion localization are discussed. Table I lists the publicly available datasets for DR detection and segmentation.

### A. Detection and Grading of Diabetic Retinopathy

Recent approaches have applied machine learning techniques and convolutional neural networks for the task of Diabetic Retinopathy grading based on image data. Typically, in studies presented in the literature, datasets are annotated for five grades of DR according to the International Clinical Diabetic Retinopathy Disease Severity Scale (ICDRD) [31] namely healthy, mild DR, moderate DR, severe DR, and proliferative DR. In this scale, DR is defined as the signs of any DR related lesions which is mild

**TABLE I**  
PUBLICLY AVAILABLE FUNDUS IMAGE DIABETIC RETINOPATHY DATASETS

Dataset	Count	Modality	Task	Notes
DiaretDB0 [23]	130: 22 normal, 108 DR	Fundus	Segmentation	Exudates, microaneurysms, haemorrhages segmentation
DiaretDB1 [24]	89: 5 normal, 84 mild signs of DR	Fundus	Segmentation	microaneurysms, haemorrhages, hard and soft exudates
Retinopathy Online Challenge [25]	100	Fundus	Segmentation	Annotations for microaneurysms
E-Ophtha [26]	381: 233 normal, 148 with red lesions	Fundus	Segmentation, Classification	Segmented small red lesions. Images with any red lesion is labeled as DR
Messidor [27]	1200: 4 DR stages	Fundus	Classification	Images are graded into 4 DR stages: healthy, mild, moderate and severe
Messidor-2 [27]	1748: 5 DR stages	Fundus	Classification	Images are graded into 5 ICDRD stages
DR2 [28]	529	Fundus	Classification	Only the presence (not the location) of different types of lesions is provided
Kaggle DR challenge [29]	88702: 5 DR stages split into train, validation, and test sets	Fundus	Classification	Images are graded into 5 ICDRD stages
IDRiD [30]	516 images split into train and test sets	Fundus	Segmentation, Classification	DR grading, pixel level annotation for four lesions and optic disc.

or more DR. Patients with moderate or more severe retinopathy are referred to ophthalmologists for examination and treatment. This level of DR is called referable DR.

In [32], the authors use multiple deep feature extractors for DR lesions and anatomies in fundus (e.g. optic disk). These features are then fused together to get good referable DR classification. The method is trained on a private dataset and evaluated on Messidor-2 dataset. This approach, while performing well, requires multiple steps of segmentation and feature extraction to detect DR. To mitigate the effect of sub-optimal solutions of multi-stage models, Google team [4] have used an end-to-end model consisting of an ensemble of 10 inception networks [33]. The training is done on a private dataset containing 128,175 fundus images provided by EyePACS. The images are graded by 3–7 annotators according to the ICDRD severity scale. The network is trained for a binary classification problem of detecting referable Diabetic Retinopathy. The model is evaluated on the publicly available Messidor-2 dataset and another private dataset of 9963 images provided by EyePACS that does not overlap with the training dataset. To have a more robust detection model that could generalize better to different types of retinas, the authors in [5] use a dataset of multi-ethnic populations with diabetes to develop and evaluate the model. In order to focus the training of lesions and informative areas of the retina, Wang *et al.* [34] utilize attention mechanisms for DR grading. This enforces the network to be more attentive to discriminative areas of the image for making grading decisions. Furthermore, the areas of focus are cropped and used in higher resolution alongside the full image to make a more informed decision. Ultra-widefield (UWF) fundus imaging is a newer more advanced technology with the ability to show a wider area of the retina to examine the periphery in addition to the foveal area. This, in turn, allows for better staging of diabetic retinopathy [35]. Since large datasets of UWF fundus images are not yet available, utilization of narrow field fundus large datasets through transfer learning is done to improve the grading of UWF fundus images [36]. Table II summarizes DR related works by stating the main task, the datasets used, and

the results of each work. Results are not directly comparable because of utilizing different datasets and splits.

In a step to make the models more interpretable visually and have some insights about the classification results, weak localization is utilized to generate heatmaps similar to the input image size. In these heatmaps, areas that have more effect on the classification have higher values [14], [37]–[39]. In early works, Class Activation Map (CAM)s [40] are used to visualize the prominent areas of the image. CAMs are easy to calculate and widely used in the literature to localize areas of the image that are responsible for specific pathology (e.g. DR) [37]–[39]. In [38], the authors use a classification network to differentiate normal fundus images from ones with DR. This networks is then used to localize salient features in the image guiding the DR prediction. CAMs are not only used for qualitative analysis of trained models but are also evaluated quantitatively as in [39] where localization results are evaluated for sensitivity and false positives per image using an annotated dataset on pixel level. To calculate CAMs, specific network design guidelines should be taken into consideration which limits the applicability of the method. Some authors have analyzed the gradients flowing back in the neural network as in [14] where the significance of each pixel is calculated by computing the q-norm of the last layer gradient with respect to the pixel of the channel. Methods for detection and grading of DR have gone a long way and are now performing in competitive levels compared with human annotators. However, there are still issues with trust, domain shifts and explainability of these models. Research directions into utilizing other modalities such as UWF or studying the progression of the disease over time could be promising fields for better understanding DR.

## B. Diabetic Retinopathy Segmentation

Further assessment of Diabetic Retinopathy characteristics is done by the identification and segmentation of lesions that characterize DR such as microaneurysms, haemorrhages, hard



TABLE II

SUMMARY: DIABETIC RETINOPATHY. AUC IS AREA UNDER THE CURVE OF ROC CURVE. ACC IS THE ACCURACY SCORE. SEN IS THE SENSITIVITY SCORE. FPs/I IS FALSE POSITIVES PER IMAGE

Reference	Task	Summary	Results	Dataset(s)
Gulshan <i>et al.</i> [4]	Classification	Ensemble of 10 inception neural networks for referable DR classification	AUC: 0.99	Messidor-2, private
Abramoff <i>et al.</i>	Classification	Deep learning used for feature extraction and features are used as input to IDx-DR X2.1	AUC: 0.98	Messidor-2, private
Ting <i>et al.</i> [5]	Classification	Learning referable diabetic retinopathy classification from multi-ethnic population dataset	AUC: 0.93	Private >76k fundus images
Wang <i>et al.</i> [34]	Classification	Fusion of multiple attention maps to focus the training of DR grading on pathological areas	AUC: 0.95	Kaggle, Messidor
Lin <i>et al.</i> [41]	Classification	Using a classification model with extra bounding box annotations for some lesions to enhance the detection of DR	AUC: 0.96	Kaggle, Messidor
Sarhan <i>et al.</i> [36]	Classification	Grading of UWF fundus images by utilizing a large dataset of narrow field fundus and transfer learning	AUC: 0.99	Kaggle, private 154 UWF
Gargeya & Leng [38]	Classification & Weak Localization	Two level classification: GAP for visualization, decision tree ensemble for GAP features + image metadata	AUC: 0.94	Messidor-2, E-Ophtha
Yang <i>et al.</i> [7]	Classification & Localization	Two-stage CNN with local lesion detection CNN and a global image level grading deeper CNN	AUC: 0.95	Kaggle
Quellec <i>et al.</i> [14]	Classification & weak localization	Generalization of backpropagation with a sparsity constraint for training using second-order gradients	AUC: 0.95, Lesion: 0.95	Kaggle, E-Ophtha, DiaretDB1
Orlando <i>et al.</i> [15]	Red lesion segmentation	Use a combination of deep and hand crafted features to train a random forest on pixel level segmentation and image level classification	AUC: 0.93	E-Ophtha, DiaretDB1, Messidor
Van Grinsven <i>et al.</i> [8]	Haemorrhages segmentation	Introduce selective sampling for training datasets with unbalanced labels and apply it to haemorrhages segmentation	AUC: 0.97	Kaggle, Messidor-2
Abbasi-Sureshjani <i>et al.</i> [42]	Exudates segmentation	Residual networks with selective sampling for full image segmentation	Dice: 0.83	E-Ophtha-EX, DiaretDB1, DR2
Huang <i>et al.</i> [43]	Neovascularization segmentation	Extreme Learning Machine with pixel-wise filter banks as input features for the model	Acc: 0.89	Messidor-2, DiaretDB0, DiaretDB1
Levenkova <i>et al.</i> [44]	Segmentation in UWF	Lesion segmentation in ultra wide-field fundus imaging using a deep feature extractor and an SVM	AUC: 0.86	Private 146 UWF
Lam <i>et al.</i> [45]	Segmentation	Segmentation of multiple DR lesions in fundus images and comparing various neural networks architectures	Acc: 0.98	Kaggle extra annotation
Guo <i>et al.</i> [46]	Avascular area segmentation	Automatic quantification of retinal non-perfusion area based on OCT angiography $6 \times 6 \text{ mm}^2$ en face angiograms in DR patients	Average dice: 0.87	Private 76 healthy and 104 DR OCTA scans
Sarhan <i>et al.</i> [47]	Microaneurysms segmentation	Multi-scale segmentation of microaneurysms using selective sampling and embedding triplet loss	Dice: 0.43	IDRiD
Wang & Yang [48]	Weak localization	CNN with multiple max-poolings and a GAP layer to visualize class activation maps	Kappa: 0.85	Kaggle
Gondal <i>et al.</i> [49]	Weak localization	Weakly supervised learning to visualize lesions using image-level classification information with increased GAP size and batch normalization	Sen: 0.72 @ 2.25 FPs/I	Kaggle, DiaretDB1

exudates, soft exudates and neovascularization. Most current approaches for the segmentation of lesions in fundus images are not holistic and partition the image into patches for processing. These approaches incorporate high-resolution patches of the image to classify each patch into the studied lesions [8], [15], [42]. In such scenarios, the patches coming from healthy samples are plenty compared to patches containing lesions which creates an unbalanced dataset. This could be handled with selective sampling [8] where a weighting scheme is applied to the selection of negative (over-represented) class samples for training the next

epoch. This approach is combined with multi-scale segmentation and embedding triplet loss to preserve spatial and semantic information on one hand and learn a robust representation on the other [47]. In [15], the authors use a hybrid feature vector (a combination of features from a deep neural network model and features extracted in a handcrafted fashion) to detect red lesions (microaneurysms and haemorrhages) in fundus images using random forest classifier. Learning-based methods are also used for the segmentation of lesions characterizing later stages of diabetic retinopathy such as neovascularization [43]. The authors

**TABLE III**  
DATASETS FOR AMD, RETINAL LAYERS AND FLUIDS SEGMENTATION

Reference	Count	Modality	Task	Notes
Duke AMD OCT [55]	384: 115 no AMD, 269 intermediate AMD	OCT volumes	Classification	100 B-scans per volume (1000 × 500). Volumes are centered at the foveal pit
Duke AMD-DME [56]	45: 15 normal, 15 AMD, 15 Diabetic Macular Edema	OCT volumes	Classification	31-97 B-scans per volume and 512-1024 A-scans per B-scan
Zhang Lab Data	109k: 51k Normal, 37k CNV, 8k drusen, 11k DME	OCT B-Scans	Classification	OCT B-scans acquired with Spectralis OCT, Heidelberg Engineering, Germany
Duke SD-OCT DME Dataset [57]	110 B-scans from 10 patients	OCT B-scans	Retinal layers and fluid segmentation	Images of subjects with Diabetic Macular Edema. B-scans are centered around the fovea and annotated for retinal layers and fluid regions
OPTIMA MICCAI 2015 Challeng [58]	30: 15 training, 15 testing	OCT volumes	Retinal layers and fluid segmentation	Two manually annotated segmentation maps of retinal cyst areas are provided
RETOUCH [59]	112: 70 training, 42 testing	OCT volumes	Retinal layers and fluid segmentation	3 fluid types segmentation maps for OCT volumes from 3 different devices

use Extreme Learning Machine (ELM) [48] for segmenting neovascularization in color fundus images and report an accuracy of 89.2%. Ultra-widefield fundus images are used in [44] to detect DR lesions automatically using a deep neural network. Detection and segmentation of DR lesions are challenging tasks especially in the case of small lesions such as microaneurysms. Models that could handle high-resolution images or work in multiple scales of the data are viable solutions to the problem.

### III. AGE-RELATED MACULAR DEGENERATION

Age-related Macular Degeneration (AMD) is one of the major causes of vision impairment for the elderly. The choroid and outer layers of retina such as retinal pigment epithelium (RPE) are particularly affected. Damages to the retinal layers that contain photoreceptors could lead to vision loss. There are multiple scales to grade AMD, one commonly used is the Age-Related Eye Disease Study (AREDS) simplified scale [49]. This scale grades AMD into four stages (none, early, intermediate, and advanced/late AMD). Another fine-grained scale from AREDS provides 9 increasing grades of nonadvanced AMD. AMD could be dry or wet. In the wet form of AMD, choroidal neovascularisations (CNV) could grow into the RPE. These new vessels can leak which may cause sudden vision loss. The dry AMD represents atrophy affecting the RPE layer of the retina. In the following subsections, two tasks that have been addressed by researchers will be presented: detection of AMD and segmentation of AMD related structures. Table III lists the publicly available datasets for AMD detection, retinal layers, and retinal fluids segmentation.

#### A. Detection of Age-Related Macular Degeneration

The automatic detection of AMD is done either with fundus imaging or with OCT volumes and scans. Burlina *et al.* [50] use fundus images around the macula to extract features using a pre-trained neural network on ImageNet. The features are extracted on multiple scales and concatenated to be used as an input to a support vector machine (SVM). The model is

developed for detecting intermediate AMD (level 3 in a 4-level grading scheme [49]) as it is the asymptomatic stage after which vision loss is a serious risk. A 0.96 AUC of the ROC curve is reported in the study. To have a more fine-grained grading of AMD, an ensemble of deep neural networks is used in [51] to detect 12 grades of AMD. These grades include 9 classes based on the ARED [49] 9-step severity scale, 3 late-stage classes, and 1 additional class for ungradable images. The method achieves a quadratic weighted kappa of 0.92. OCT scans of the eye posterior segment show the anatomical structure of the retinal layers and biomarkers related to AMD are clear in the scans (e.g. drusen) making it possible to detect AMD through OCT. Other approaches use OCT imaging to detect AMD. In [52], transfer learning is used in 3-class classification task (normal, AMD, DME). The authors used the Duke AMD-DME dataset and a pre-trained version of the Google Inception network [53]. The OCT B-scans are pre-processed to have a similar spatial resolution as the pre-trained network input with three channels by replicating the grayscale image. The results show better performance when using the pre-trained network rather than a random initialization of the network. The achieved mean accuracy for each class is 0.99, 0.89, and 0.86 for normal, AMD, and DME respectively. To solve the problem of exhaustively annotating each B-scan separately, Apostolopoulos *et al.* [20] propose a model that works on the volume level first to infer classification on the B-scan level. The model trained to classify healthy vs. intermediate AMD patients using OCT volumes from the Duke AMD dataset. This approach shows good performance compared to multiple baselines with a 0.997 AUC or ROC curve. For better detection of AMD, other approaches detect the presence of AMD-related abnormalities rather than directly detecting AMD. This could also contribute to the interpretability of such models. This is done in [54] where the authors use a dataset of 109 k OCT B-scans to detect the presence of CNV, drusen, and DME. CNV and drusen are indicators of wet and dry AMD respectively and the detection of their presence is correlated with AMD. Transfer learning is used in this case with a pre-trained model on ImageNet. The last layers are optimized

TABLE IV

SUMMARY: AGE-RELATED MACULAR DEGENERATION AND MACULAR PATHOLOGIES. AUC IS THE AREA UNDER THE ROC CURVE. ACC IS THE ACCURACY SCORE. IOU IS THE INTERSECTION OVER UNION SCORE

Reference	Task	Summary	Results	Dataset(s)
Apostolopoulos <i>et al.</i> [20]	Classification	Full OCT volume classification for AMD screening using CNNs and pre training on single B-scans with volume annotation	AUC: 0.99	Duke AMD
Deng <i>et al.</i> [65]	Classification	Distinguish between normal, dry AMD, and wet AMD OCT scans using Gabor filtering with multiple machine learning methods (random forests, SVM, and ANN)	Acc: 0.94	Private 21 patients
Lee <i>et al.</i> [66]	Classification & weak localization	Use OCT scans and VGG network to classify AMD vs. normal patients with occlusion tests to localize salient areas	AUC: 0.93	Private: 48,312 normal, 52,690 AMD
Venhuizen <i>et al.</i> [67]	Classification	Use visual bag of words dictionary and random forest classifier to classify five severity levels of AMD using OCT scans. Levels are No AMD, Early AMD, Intermediate AMD, Advanced AMD with geographic atrophy, and Advanced AMD with choroidal neovascularization	AUC: 0.97	EUGENDA 3265 OCT volumes, Duke AMD-DME
Karri <i>et al.</i> [52]	Classification	Detect AMD and DME with transfer learning and multiple pre-processing steps for saturation, flattening, resizing, and noise reduction	Acc: 0.89	Duke AMD-DME
Burlina <i>et al.</i> [50]	Classification	A pre-trained network is used to extract features around the macula in multiple scales of fundus images. Then, these features are used as input to an SVM	AUC: 0.96	AREDS >130,000 color fundus images
Grassmann <i>et al.</i> [51]	Classification	An ensemble of deep neural networks is used to detect AMD fine-grained stages and image gradability using fundus images	Kappa: 0.92	AREDS >120,000 color fundus images
Kermany <i>et al.</i> [54]	Classification	Detection of CNV and drusen from OCT B-scans as an indicator for AMD. Transfer learning is used for this task.	AUC: 0.99	Zhang Lab Data
Treder <i>et al.</i> [68]	Classification	Use transfer learning to detect wet AMD in OCT B-scans	Acc: 0.96	1,112 OCT B-scans
De Fauw <i>et al.</i> [63]	Classification & segmentation	Detect 10 retinal pathologies and referral scores of OCT volumes based on a segmentation step of 15 retinal structures	AUC: 0.99 urgent referral	Moorfield hospital 14,884 OCT volumes
Schlegl <i>et al.</i> [69]	Segmentation	Segmentation model to quantify intraretinal cystoid fluid (IRC) and subretinal fluid (SRF) in 3 different pathologies	AUC IRC: 0.94, AUC SRF: 0.92	Private 1200 OCT volumes
Feeny <i>et al.</i> [60]	Geographic atrophy segmentation	Random forest model for pixel-wise segmentation of geographic atrophy using handcrafted features extracted from colored fundus images	Dice: 0.68	Private 143 fundus images
Zedeh <i>et al.</i> [9]	Drusen segmentation	Drusen segmentation using U-Net architecture. Ground truth generation from OCT volumes of segmented retinal layers	IOU: 0.82	Private >50k images
Lee <i>et al.</i> [70]	IRF segmentation	Segmentation of intra-retinal fluids in OCT B-scans	Dice: 0.91	Private 1,289 OCT B-scans
Ji <i>et al.</i> [62]	Geographic atrophy segmentation	Voting scheme for combining geographic atrophy segmentation of multiple models from OCT volumes	IOU: 0.86	Private 51 OCT volumes, 54 OCT volumes
Xue <i>et al.</i> [71]	Choroidal neovascularization segmentation	Unsupervised clustering-based segmentation of choroidal neovascularization based on OCT angiography $3 \times 3 \text{ mm}^2$ scans	IOU: 0.87	Private 22 OCTA scans

for the new task and a separate test set annotated by 6 experts is used for evaluation. The proposed network reaches a 0.99 AUC of ROC curve for the task of urgent referral. Urgent referral is defined in the study as the presence of CNV or DME. Approaches for detecting AMD perform very well in later stages but suffer in the earlier stages where abnormalities are not yet prominent. Moreover, domain shift problems may occur since many studies rely on data from AREDS. Table IV summarizes AMD related works by stating the main task, the datasets used, and the results of each work. Results are not directly comparable because of utilizing different datasets and splits.

## B. Segmentation for Age-Related Macular Degeneration

Segmentation of AMD related lesions is done on both fundus images variants and OCT scans. Fundus imaging is used for

segmentation of AMD-related lesion as in [60] where a random forest is utilized to segment Geographic Atrophy (GA) in AMD patients. A hand-crafted feature vector is extracted for each pixel neighborhood and a model is trained to predict if a pixel belongs to a GA. The model is tested on the full dataset and a subset that contains images with low ambiguity and the achieved dice coefficients are  $0.68 \pm 0.25$  and  $0.70 \pm 0.21$ , respectively. OCT is becoming more accessible and OCT cross-sections show information about retinal layers pathologies, OCT is widely used in AMD lesion segmentation. In [9], the authors use U-Net [61] architecture to segment drusen in OCT retinal images. A drusen segmentation dataset is created by utilizing a pre-existing dataset that contains segmentation maps for Bruch's membrane (BM) and the retinal pigment epithelium (RPE) layers. It is also possible to segment GA from OCT volumes without going through a retinal layer segmentation step [62]. A

**TABLE V**  
DATASETS FOR GLAUCOMA DETECTION AND OPTIC DISC/CUP SEGMENTATION

Reference	Count	Modality	Task	Notes
DRISHTI-GS [73]	50	Fundus	Segmentation	Annotation for optic disc and optic cup
DRIONS-DB [74]	110	Fundus	Segmentation	Annotation for optic disc
RIM-ONE [75]	159	Fundus	Segmentation	Annotation for optic disc
SiMES [76]	230	Fundus	Segmentation	Optic disc segmentation for subjects with parapapillary atrophy
ORIGA [77]	650: 482 No glaucoma, 168 with glaucoma	Fundus	Classification	Binary classification
SCES [78]	1676: 1630 No glaucoma, 46 with glaucoma	Fundus	Classification	Binary classification
MICCAI 2018 REFUGE [79]	1200: 1080 No glaucoma, 120 with glaucoma	Fundus	Classification and segmentation	Binary classification, optic disc and cup segmentation and fovea localization

voting strategy is suggested to combine the segmentation from multiple networks into the final segmentation result. The method achieves an intersection over union (IoU) of  $86.94\% \pm 8.75\%$ . Most of the methods in the literature apply machine learning techniques on a limited dataset where the comparison is between normal and pathological cases. However, pathological cases can interfere with each other and affect the robustness of neural networks on unseen pathologies. This effect is not well studied and investigated. To account for that, DeepMind's work [63] includes 10 different retinal pathologies to be detected. Some of these pathologies are AMD-related such as CNV, GA and drusen. A dataset from Moorfield eye hospital is utilized to create a two-stage deep learning network for detecting 10 OCT conditions and give a referral decision. The first stage model is trained to segment 15 different structures and pathologies in OCT volumes. The segmentation results are then used as input to a classification network. The classification network is trained for 1) referral suggestion (urgent, semi-urgent, routine, observation), 2) diagnosis probability for multiple characteristics (normal, CNV, macular retinal edema, full macula hole, partial macula hole, central serous retinopathy, Vitreomacular traction, GA), and 3) tissue volume of drusen and the epiretinal membrane (ERM). The reported results are comparable with expert results on these tasks with 0.99 AUC of ROC curve for the Task of referring urgent cases. The intermediate segmentation step enables easier adoption of the model in case it is needed to change OCT vendors as the second model is independent of the OCT source. Detecting lesions for AMD is an important task for detecting early signs and tracking the progression of the disease. However, there is high variability among experts when it comes to pixel-level annotation. To partly overcome this difficulty, approaches that model uncertainty are desired to model this variability.

#### IV. GLAUCOMA

Glaucoma affects the optic nerve progressively and it is detected commonly in three approaches [64] (1) detection of increased intraocular pressure, (2) identifying the field of abnormal vision, or (3) assessing the damage of the optic nerve by calculating the cup-to-disc ratio (CDR). Optic Disc (OD) is the area where the optic nerve leaves the eye and it can be divided into two parts, the optic cup which is in the center of

the Optic Disc as a bright circular area and the peripheral region around the cup which is the neuroretinal rim. The Optic Disc changes visually when the optic nerve fibers are damaged due to glaucoma, this leads to enlargement of the cup region called cupping and it is an indicator for glaucoma suspect detection. Detection of glaucoma and segmentation of areas that characterize glaucoma are presented in the following subsections. Table V lists the publicly available datasets for glaucoma detection, and optic disk and cup segmentation.

##### A. Detection of Glaucoma

Multiple measurements are used for detecting glaucoma. In [2], the authors used Visual Fields measurements to differentiate open-angle glaucoma patients from healthy patients using deep feedforward neural networks. Visual Fields using Humphrey Field Analyzer from 171 pre-perimetric glaucoma patients and 108 healthy cases are acquired for analysis. The input to the neural network model is the 52 visual field values generated by the measuring machine with Pattern Standard Deviation (PSD) and Mean Deviation (MD) values. The neural network consists of two hidden layers and classifies pre-perimetric glaucoma from normal visual field. The method is compared with several other machine learning models including SVM, Random Forest, and K-Nearest Neighbors. It achieved the highest AUC of 0.926. To go beyond visual field measurements, some works use anatomical features alongside the visual field data. One work [1] selects the best features from various measurements for 297 glaucoma eyes and 202 non-glaucoma eyes. The measurements included ocular pressure and corneal thickness measurement as general exams, OCT scan for retinal nerve fiber layer (RNFL) thickness, and visual field (VF) exam. The model achieves 0.979 AUC of ROC curve for detecting if a subject has glaucoma or not. The RNFL thickness is the anatomical information most used in the detection process. Other works have attempted to utilize fundus imagery for detecting glaucoma. Li *et al.* [72] have used fundus images for detecting referable glaucomatous optic neuropathy. The images are annotated by 21 trained ophthalmologists into unlikely suspects of glaucoma and certain glaucoma depending on the vertical CDR, rim width, RNFL defect, and disc hemorrhage. The deep model registers a 0.98 AUC of ROC for referable glaucomatous optic neuropathy. Detection of glaucoma is a challenging task because the early



TABLE VI

SUMMARY OF WORKS ON GLAUCOMA DETECTION AND GLAUCOMA-RELATED ANATOMIES SEGMENTATION. AUC IS THE AREA UNDER THE ROC CURVE. IOU IS INTERSECTION OVER UNION

Reference	Task	Summary	Results	Dataset(s)
Cheng <i>et al.</i> [80]	Classification	CNN with overlapping pooling layers for binary classification of glaucoma in fundus images	AUC: 0.83, 0.88	ORIGA, SCES
kim <i>et al.</i> [1]	Classification	Use multiple eye measurements as input for decision trees, random forests, SVMs, and KNNs for comparison	AUC: 0.97	Private 499 Cases
Asaoka <i>et al.</i> [2]	Classification	Use Visual Field measurements for preperimetric glaucoma detection using feedforward neural networks	AUC: 0.92	Private 279 Visual Fields
Li <i>et al.</i> [72]	Classification	Glaucoma suspect and glaucoma certain classification based on fundus imaging	AUC: 0.98	Private 48,116 fundus images
Zilly <i>et al.</i> [10]	Disk & cup segmentation	Entropy based sampling for informative points selection with convolutional filters optimization using greedy boosting	IOU: disk 0.89, cup 0.80 on RIM-One	DRISHTI-GS, RIM-ONE, Messidor
Maninis <i>et al.</i> [81]	Disk & vessels segmentation	Fully convolutional approach with pre-trained VGG network and supervision on multiple internal layers	0.95 precision @ 0.95 recall	DRIONS-DB, RIM-ONE
Tan <i>et al.</i> [82]	Disc, fovea, vessels segmentation	7-layer CNN to classify central pixel of $33 \times 33$ patches into disc, fovea, vessel, or background	IOU: 0.62	DRIVE
Srivastava <i>et al.</i> [83]	Disc segmentation	Pixel-wise classification of OD with distance and intensity handcrafted features and a 7-layer neural network to test optic disc segmentation robustness against parapapillary atrophy	IOU: 0.90	SiMES

signs of the disease could be either functional or anatomical. Approaches that aim for the early detection of glaucoma should consider both factors. Table VI summarizes glaucoma-related approaches and states the main task, the used datasets, and the results of each approach. Results are not directly comparable because of different datasets and splits utilization.

### B. Segmentation for Glaucoma

The authors in [10] used an entropy-based sampling scheme to select points of interest for classification. The idea behind entropy based sampling is to pick the most informative patches for the training process. The informative points are selected to have the highest entropy values where the entropy of each pixel is calculated from its  $N \times N$  neighborhood. A convex hull transform is applied to the initial segmentation assuming that the disk and cup shapes are convex. The method is compared with state-of-the-art methods and achieves intersection over union (IOU) values for optic disc and cup segmentation of 91.4, 85.0 on DRISHTI-GS dataset and 89.0, 80.2 on the RIM-ONE dataset. To have more global information about the image, the authors in [81] have used the full fundus image rather than image patches to segment retinal vessels and optic disk simultaneously. The authors used a pre-trained VGG network [84] and inspired by GoogLeNet inception network [53], supervision to multiple internal layers is added. The authors further added two sets of specialized layers connected to the internal layers of the network. One set is used to do full image segmentation and uses the first four blocks of VGG to have higher spatial resolution. The second set of added layers is used for optic disk segmentation by connecting to the last four layers of VGG. This approach is fully convolutional without any fully connected layers and a probability map of the same size as the original image is produced for each task. For optic disk segmentation, the evaluation is done on DRIONS-DB and RIM-ONE datasets and the model achieves dice coefficient values of 0.971 and 0.959 for

optic disk segmentation on the two datasets respectively. While the segmentation of glaucoma related anatomies (such as optic disc/cup) is achievable with good performance. It is important to note that it is clinically challenging to decide for glaucoma based on data acquired at one instance. It would be more beneficial to utilize such quantifiable segmentation systems for the tracking of glaucoma suspects to monitor any glaucoma progression symptoms.

### V. RETINAL VESSELS SEGMENTATION

Retinal vessel segmentation is widely researched in the literature. Retinal vessel networks contain rich information about various retinal pathologies. Early signs of diabetic retinopathy and vascular burden from hypertension could be detected. Moreover, vision threatening diseases such as Retinal Vein Occlusion (RVO) and Retinal Artery Occlusion (RAO) are detected from abnormalities in microvascular structure. The exclusion of retinal vessels network is a preprocessing step used in various automatic detection algorithms. Therefore, reliable methods for automatic retinal vessel segmentation are needed. Retinal vessel segmentation is a challenging task for multiple reasons. Retinal vessels vary in thickness and could reach 1-pixel thickness (depending on resolution and anatomical characteristics) and vessels could have intersections and sometimes thicker vessels have centerline reflex effect where it appears as two vessels. Table VII lists the publicly available datasets for retinal vessels segmentation.

Using Machine Learning to label central pixel of small extracted patches is one approach frequently used in the literature [85], [86], [87], [88]. Liskowski *et al.* [85] use patches of size  $27 \times 27$  and applied a preprocessing for global contrast normalization on image level and zero-phase component analysis whitening on patch level. A deep convolutional network is then trained to classify the center pixel of the pre-processed patches. As for end-to-end direct image segmentation, holistic



TABLE VII  
PUBLICLY AVAILABLE DATASETS FOR RETINAL VESSELS SEGMENTATION

Reference	Count	Modality	Notes
DRIVE [94]	40: 7 mild DR signs	Fundus	Manual segmentation of the blood vessels is provided
STARE [95]	20: 10 cases with pathologies that overlap with blood vessels	Fundus	Two manually annotated segmentation maps are provided
CHASE_DB [96]	28 images	Fundus	Children Fundus images. The images have worse contrast for blood vessels, nonuniform background illumination and wide vessels that have the central vessel reflex

TABLE VIII  
SUMMARY OF RETINAL VESSELS SEGMENTATION WORKS

Reference	Summary	Dataset(s)
Liskowski <i>et al.</i> [85]	CNNs with and without max-pooling applied for central pixel classification of $27 \times 27$ patches	DRIVE, STARE, CHASE_DB
Wang <i>et al.</i> [87]	Ensemble learning with random forests trained on feature maps extracted with a CNN ( $25 \times 25$ patches)	DRIVE, STARE
Maji <i>et al.</i> [88]	Ensemble of multiple CNNs trained on binary classification of central pixel for $31 \times 31$ patches	DRIVE
Li [3]	Use artificial neural networks and denoising autoencoders to learn full patch segmentation map	DRIVE, STARE
Fu <i>et al.</i> [13]	A CNN with sided output and conditional random field to capture non-local correlations	DRIVE, STARE, CHASE_DB
Prentavsis <i>et al.</i> [91]	OCT angiography microvasculature segmentation using a 3-layer CNN to classify center pixel of $61 \times 61$ patches	Private 80 OCTA images
Leopold <i>et al.</i> [86]	Gabor filter preprocessing to train a CNN with batch normalization for central pixel classification of $65 \times 65$ patches	DRIVE
Son <i>et al.</i> [97]	Retinal vessels segmentation using generative adversarial networks where the generator has an architecture similar to U-Net and the discriminator classifies human and non-human segmentation	DRIVE, STARE
Costa <i>et al.</i> [90]	Automatic generation of fundus images and vessels segmentation maps pairs to account for data scarcity problems using generative adversarial neural networks	DRIVE, Messidor

view of the image structure is considered which allows for a more context-aware segmentation. Moreover, a faster segmentation is possible since there is no need for a sliding window patch-wise classification. In [13] the vessel segmentation is done on image level using a convolutional neural network with sided output and a conditional random field. The sided output is used as a classifier for shallow layers outputs. Each sided output has its loss that is weighted and added to the final loss as in [89]. A Conditional Random Field is used to capture non-local correlations between pixels and takes the image and the sided outputs from the convolution layers as input. The final loss of the network is the weighted sum of the sided outputs and the output of the Conditional Random Field. The time needed for full segmentation of one image of size  $565 \times 584$  is 1.3 seconds. The emergence of generative adversarial neural networks is also affecting research in ophthalmic image analysis. In order to counter the problems of data scarcity and high cost of annotations, Costa *et al.* [90] proposed a framework to synthesize fundus images with their corresponding vessel map segmentation. Two segmentation networks are trained on the real and generated images. The performance of the two networks is close which indicates good generation. However, the generated images are of relatively low resolution ( $256 \times 256$ ). Hence, the segmentation results of both models are much lower than state-of-the-art. Machine learning is also used to segment microvasculature

in OCTA images [91]. This work includes 80 OCTA foveal region images of 12 eyes from 6 volunteers with no sign of pathologies. The network is composed of 3 convolutional layers and 3 max-pooling layers. The input is  $61 \times 61$  pixels patch of the image and the network is trained to classify the central pixel of each patch. The number of sampled patches with and without vessels is equal to avoid bias towards the over-represented class. The method achieves an accuracy of 0.83 and a dice score of 0.7 and it is compared with the human performance by asking a second annotator to segment 10 of the images. Full segmentation of OCT-A image microvasculature takes 2 minutes using the proposed method. While various approaches perform well on the retinal vessels segmentation task, it is still challenging to detect smaller vessels with low contrast and to detect vessels in cases of severe pathologies. Using OCTA is a good option for monitoring microvascular abnormalities in limited areas. Table VIII summarizes retinal vessels segmentation works by stating the main task and the datasets used. The results of each work on retinal vessels segmentation are shown in Table IX.

## VI. RETINAL LAYERS SEGMENTATION IN OCT IMAGES

Retinal layers segmentation is valuable for analyzing OCT scans to diagnose or monitor retinal diseases. This task can be challenging due to pathologies such as cysts, subretinal

**TABLE IX**  
RETINAL VESSELS SEGMENTATION RESULTS IN TERMS OF SENSITIVITY (SEN), SPECIFICITY (SPE), ACCURACY (ACC), AND AREA UNDER THE CURVE OF ROC (AUC)

Reference	DRIVE dataset				STARE dataset				CHASE-DB dataset			
	Sen	Spc	Acc	AUC	Sen	Spc	Acc	AUC	Sen	Spc	Acc	AUC
Li <i>et al.</i> [3]	75.69	<b>98.16</b>	95.27	97.38	77.26	<b>98.44</b>	96.28	<b>98.79</b>	<b>75.07</b>	<b>97.93</b>	<b>95.81</b>	<b>97.16</b>
Fu <i>et al.</i> [13]	76.03	-	95.23	-	74.12	-	95.85	-	71.30	-	94.89	-
Liskowski <i>et al.</i> [85]	77.63	97.68	94.95	97.20	79.67	97.54	95.66	97.85	-	-	-	-
Wang <i>et al.</i> [87]	<b>81.73</b>	97.33	<b>97.67</b>	94.75	<b>81.04</b>	97.91	<b>98.13</b>	97.51	-	-	-	-
Leopold <i>et al.</i> [86]	78.00	97.27	94.78	-	-	-	-	-	-	-	-	-
Maji <i>et al.</i> [88]	-	-	94.7	92.83	-	-	-	-	-	-	-	-
Son <i>et al.</i> [97]	-	-	-	<b>98.03</b>	-	-	-	98.38	-	-	-	-

**TABLE X**  
SUMMARY OF RETINAL LAYERS SEGMENTATION WORKS. ACC IS THE ACCURACY SCORE. AVERAGE DICE IS THE AVERAGE OF DICE OF MULTIPLE STRUCTURES REPORTED IN THE PAPER

Reference	Summary	Results	Dataset(s)
Schlegl <i>et al.</i> [98]	Cystoid structures weak segmentation in OCT images using semantic descriptors extracted from clinical reports without pixel-wise labels for training	Acc: 0.81	Private 157 OCT Volumes
Venhuizen <i>et al.</i> [92]	Retinal layers segmentation by using a generalized form of U-Net to have a larger receptive field and deeper architecture	Dice: 0.95	EUGENDA, Duke_DME
Roy <i>et al.</i> [93]	Retinal layers segmentation by using U-Net like structure with batch normalization to segment retinal layers and fluid using a tailored loss function with a differentiable approximation of dice loss	Average dice: 0.90	Duke_DME
Apostolopoulos <i>et al.</i> [99]	Retinal layers segmentation by using dilated convolution for bigger receptive field and residual connections with MSE loss on U-Net like architecture	Dice: 0.99	Private 20 OCT volumes
Gopinath <i>et al.</i> [100]	Cystoid structures segmentation in OCT images using selective enhancement with generalized motion pattern on 3D OCT data	Dice: 0.71	OPTIMA, Duke_DME
Schlegl <i>et al.</i> [101]	Using adversarial neural networks to mark biomarkers candidates such as retinal fluids in OCT B-scans without the need for annotated data	Dice: 0.79	Private 280 healthy volumes + 10 volumes with fluids
venhuizen <i>et al.</i> [102]	Deep learning for segmentation and quantification of intraretinal cystoid fluid in OCT volumes on B-scan level	Dice: 0.75	EUGENDA 221 OCT volumes

fluids or drusen where the structure of the retinal layers varies significantly.

In [92], the authors use a generalized form of U-Net [61] to segment retinal layers in OCT images with AMD. AMD introduces difficulties in capturing the structure of the retina. The generalized U-Net used in this work is a deeper network than original U-Net with more downsampling layers and more feature maps per layer. The receptive field becomes  $572 \times 572$  pixels rather than  $140 \times 140$  pixels in the original U-Net. This is important for the network to be able to capture larger structures in the image such as fluid-filled spaces. The model achieves a central macular thickness estimation error of  $14.0 \pm 22.1 \mu m$ . It takes the system 5 seconds to segment a volume of 38 B-scans. In [93], authors use an encoder-decoder architecture to segment retinal layers and fluid into 10 classes (7 anatomical layers, 1 fluid, 2 above and under retina classes). The loss function uses the weighted logistic regression loss where the weights are used to balance imbalanced classes and to give higher importance to segmentation of boundary pixels. Moreover, a differentiable approximation of dice loss is added to the loss function making the total loss as a weighted sum of the two losses. The network employs convolutional kernels of size  $7 \times 3$  to keep consistency with OCT spatial resolution, this kernel size ensures the

receptive field of the network to cover the whole retinal layers. The images are sliced with none overlapping width-wise stripes. This allows training a larger batch size with GPU limitations. At test time the whole B-Scan is given to the network as input and it takes 10 milliseconds to segment one B-scan. The achieved dice score for fluid segmentation which is the most challenging part in the segmentation task is 0.77. Segmenting retinal layers is a challenging task when retinal fluids and abnormalities occur. For such cases, it could be beneficial to use certainty aware approaches to model the inter-rater variability. Table III lists the publicly available datasets for retinal layers and fluids segmentation. Table X summarizes retinal layers segmentation works by stating the main task, the datasets used, and the results of each work. Results are not directly comparable because of different datasets and splits utilization.

## VII. LIMITATIONS AND FUTURE

Machine learning techniques are being widely researched for their good performance on challenging tasks of fine-grained grading and pixel-level segmentation. However, these techniques are hindered by various limitations. Some important limitations are the availability of data and the regulations for

data usage. Unlike most computer vision applications, medical applications require datasets annotated by medical experts. This limits the amount of available annotated data because of the high cost of annotation. One way to resolve this in the literature is data augmentation which could be done more realistically with generative models [103] that learn the data distribution and possibly generate unseen examples [90]. Another way is utilizing semi-supervised learning techniques [104] which allow the incorporation of unlabeled data alongside labeled data into the model to enhance the overall performance on the labeled data. This is beneficial for medical applications where a large amount of unlabeled data is not used in developing the models. Another data-related limitation is the country-specific regulations on data collection and patient consent which mostly results in homogeneous datasets in terms of ethnicity and/or hardware settings for development. Models developed on such datasets are prone to generalization problems as one population data might have different characteristics that introduce a bias in the model. Hence, domain adaptation techniques should be kept in mind while developing these models. One solution to the problem is to decompose it into multiple parts as in first segment then classify [63] to allow easier domain adaptation when needed. An alternative solution is to separate shared content information from attribute (domain-specific) information using image-to-image translation via disentangled representations [105]. Although deep models perform well on quantitative measures, trust is still an issue for the adoption of these models into clinical scenarios. A deep model is seen as a black box where results are not interpretable and failure cases cannot be anticipated nor explained. Approaches such as [14], [38] visualize heat maps associated with the classification to show attention areas of the model that lead the classification. Another possibility to explore is disentangling image generative factors in image representation latent space. Disentangling allows a better understanding of the image representation in high dimensional space where each component of representation holds information about one trait in the data [106], [107].

### VIII. SUMMARY

Machine learning and deep convolutional neural networks have been the focus of most recent publications on ophthalmic data analysis outperforming legacy algorithms on different applications from detection and grading diseases such as AMD, DR, and glaucoma to the segmentation of anatomical structures such as vessels and retinal layers or intraretinal fluids. In this paper, we reviewed the recently published machine learning methods in Ophthalmology and listed the datasets publicly available for research. Classification and segmentation are the two most addressed problems in these publications. Detection and grading of eye diseases through classification leads to more efficient screening programs enabling the healthcare organizations to cover a larger population. Segmentation of vessels and retinal layer or lesions such as intra-retinal fluid, on the other hand, can be used for improving the performance in detecting diseases or as a crucial intermediate process of the clinical decision support systems in the future. Together with the rise of these methods

came the concerns about patient privacy and the rights on the use of data. Interpretability is one requirement in medical applications as more than sensitivity and specificity is needed to assess the performance of machine learning methods. There is no doubt that machine learning came a long way fast and future researches and developments will make it one of the most powerful tools in the hands of experts to diagnose and treat eye diseases.

### REFERENCES

- [1] S. J. Kim, K. J. Cho, and S. Oh, "Development of machine learning models for diagnosis of glaucoma," *PLoS ONE*, vol. 12, no. 5, 2017, Paper e0177726.
- [2] R. Asaoka, H. Murata, A. Iwase, and M. Araie, "Detecting preperimetric glaucoma with standard automated perimetry using a deep learning classifier," *Ophthalmology*, vol. 123, no. 9, pp. 1974–1980, 2016.
- [3] Q. Li, B. Feng, L. Xie, P. Liang, H. Zhang, and T. Wang, "A cross-modality learning approach for vessel segmentation in retinal images," *IEEE Trans. Med. Imag.*, vol. 35, no. 1, pp. 109–118, Jan. 2015.
- [4] V. Gulshan *et al.*, "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," *JAMA*, vol. 316, no. 22, pp. 2402–2410, 2016.
- [5] D. S. W. Ting *et al.*, "Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes," *JAMA*, vol. 318, no. 22, pp. 2211–2223, 2017.
- [6] H. Pratt, F. Coenen, D. M. Broadbent, S. P. Harding, and Y. Zheng, "Convolutional neural networks for diabetic retinopathy," *Procedia Comput. Sci.*, vol. 90, pp. 200–205, 2016.
- [7] Y. Yang, T. Li, W. Li, H. Wu, W. Fan, and W. Zhang, "Lesion detection and grading of diabetic retinopathy via two-stages deep convolutional neural networks," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervention*, 2017, pp. 533–540.
- [8] M. J. van Grinsven, B. van Ginneken, C. B. Hoyng, T. Theelen, and C. I. Sánchez, "Fast convolutional neural network training using selective data sampling: Application to hemorrhage detection in color fundus images," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1273–1284, May 2016.
- [9] S. G. Zadeh *et al.*, "CNNs enable accurate and fast segmentation of drusen in optical coherence tomography," in *Deep. Learn. in Medical. Image. Analysis. and Multimodal. Learn. for Clinical. Decision. Support.* Berlin, Germany: Springer, 2017, pp. 65–73.
- [10] J. Zilly, J. M. Buhmann, and D. Mahapatra, "Glaucoma detection using entropy sampling and ensemble learning for automatic optic cup and disc segmentation," *Computerized Med. Imag. Graph.*, vol. 55, pp. 28–41, 2017.
- [11] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Proc. Advances Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.
- [12] H. Greenspan, B. van Ginneken, and R. M. Summers, "Guest editorial deep learning in medical imaging: Overview and future promise of an exciting new technique," *IEEE Trans. Med. Imag.*, vol. 35, no. 5, pp. 1153–1159, May 2016.
- [13] H. Fu, Y. Xu, S. Lin, D. W. K. Wong, and J. Liu, "Deepvessel: Retinal vessel segmentation via deep learning and conditional random field," in *Proc. Int. Conf. Med. Image Comput. Comput.-assisted Intervention*, 2016, pp. 132–139.
- [14] G. Quellec, K. Charrière, Y. Boudi, B. Cochener, and M. Lamard, "Deep image mining for diabetic retinopathy screening," *Med. Image Anal.*, vol. 39, pp. 178–193, 2017.
- [15] J. I. Orlando, E. Prokofyeva, M. del Fresno, and M. B. Blaschko, "An ensemble deep learning based approach for red lesion detection in fundus images," *Comput. Methods Programs Biomed.*, vol. 153, pp. 115–127, 2018.
- [16] D. S. Ting *et al.*, "Deep learning in ophthalmology: the technical and clinical considerations," *Progress. Retinal. Eye. Res.*, vol. 72, 2019, Art. no. 100759.
- [17] D. S. W. Ting *et al.*, "Artificial intelligence and deep learning in ophthalmology," *Brit. J. Ophthalmol.*, vol. 103, no. 2, pp. 167–175, 2019.
- [18] U. Schmidt-Erfurth, A. Sadeghipour, B. S. Gerendas, S. M. Waldstein, and H. Bogunović, "Artificial intelligence in retina," *Progress. Retinal. Eye. Res.*, vol. 67, pp. 1–29, 2018.



- [19] K. B. Nielsen, M. L. Lautrup, J. K. Andersen, T. R. Savarimuthu, and J. Grauslund, "Deep learning-based algorithms in screening of diabetic retinopathy: A systematic review of diagnostic performance," *Ophthalmol. Retina*, vol. 3, no. 4, pp. 294–304, 2019.
- [20] S. Apostolopoulos, C. Ciller, S. De Zanet, S. Wolf, and R. Sznitman, "Retinet: Automatic amd identification in OCT volumetric data," *Investigative Ophthalmol. Visual. Sci.*, vol. 58, no. 8, pp. 387–387, 2017.
- [21] R. Lee, T. Y. Wong, and C. Sabanayagam, "Epidemiology of diabetic retinopathy, diabetic macular edema and related vision loss," *Eye Vis.*, vol. 2, no. 1, pp. 1–25, 2015.
- [22] T. Kauppi *et al.*, *Eye Fundus Image Analysis For Automatic Detection of Diabetic Retinopathy*. Lappeenranta, Finland: Lappeenranta University of Technology, 2010.
- [23] T. Kauppi *et al.*, "Diaretddb0: Evaluation database and methodology for diabetic retinopathy algorithms," *Mach. Vis. Pattern Recognit. Res. Group, Lappeenranta University Technol., Finland.*, vol. 73, pp. 1–17, 2006.
- [24] T. Kauppi *et al.*, "Diaretdb1 diabetic retinopathy database and evaluation protocol," in *Med. Image Understanding Anal.*, 2007, pp. 61–65.
- [25] M. Niemeijer *et al.*, "Retinopathy online challenge: automatic detection of microaneurysms in digital color fundus photographs," *IEEE Trans. Med. Imag.*, vol. 29, no. 1, pp. 185–195, Jan. 2010.
- [26] E. Decencière *et al.*, "Teleophtha: Machine learning and image processing methods for teleophthalmology," *IRBM*, vol. 34, no. 2, pp. 196–203, 2013.
- [27] E. Decencière *et al.*, "Feedback on a publicly distributed database: the messidor database," *Image Anal. Stereol.*, vol. 33, no. 3, pp. 231–234, Aug. 2014.
- [28] R. Pires, H. F. Jelinek, J. Wainer, E. Valle, and A. Rocha, "Advancing bag-of-visual-words representations for lesion classification in retinal images," *PLoS ONE*, vol. 9, no. 6, 2014, Paper e96814.
- [29] Kaggle, "Kaggle diabetic retinopathy detection challenge," 2015. Accessed: Sep. 26, 2018. [Online]. Available: <https://www.kaggle.com/c/diabetic-retinopathy-detection>
- [30] P. Porwal *et al.*, "Indian diabetic retinopathy image dataset (idrid)," 2018. [Online]. Available: <http://dx.doi.org/10.21227/H25W98>
- [31] "American Academy of Ophthalmology. International Clinical Diabetic Retinopathy Disease Severity Scale Detailed Table," Oct. 2002, Accessed: Jun. 11, 2019. [Online]. Available: <http://www.icoph.org/downloads/Diabetic-Retinopathy-Detail.pdf>
- [32] M. D. Abramoff *et al.*, "Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning," *Investigative Ophthalmol. Vis. Sci.*, vol. 57, no. 13, pp. 5200–5206, 2016.
- [33] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 2818–2826.
- [34] Z. Wang, Y. Yin, J. Shi, W. Fang, H. Li, and X. Wang, "Zoom-in-net: Deep mining lesions for diabetic retinopathy detection," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervention*, 2017, pp. 267–275.
- [35] P. S. Silva, J. D. Cavallerano, J. K. Sun, J. Noble, L. M. Aiello, and L. P. Aiello, "Nonmydriatic ultrawide field retinal imaging compared with dilated standard 7-field 35-mm photography and retinal specialist examination for evaluation of diabetic retinopathy," *Amer. J. Ophthalmol.*, vol. 154, no. 3, pp. 549–559, 2012.
- [36] M. H. Sarhan, P. Sha, M. Chen, M. K. Durbin, M. Yigitsoy, and A. Eslami, "Deep learning for automatic diabetic retinopathy grading of ultra-widefield fundus images," *Investigative Ophthalmol. Vis. Sci.*, vol. 60, no. 9, pp. 1094–1094, 2019.
- [37] Z. Wang and J. Yang, "Diabetic retinopathy detection via deep convolutional networks for discriminative localization and visual explanation," in *Proc. Workshops 32nd AAAI Conf. Artif. Intell.*, 2018. [Online]. Available: <https://www.aaai.org/ocs/index.php/WS/AAAIW18/paper/viewPaper/16668>
- [38] R. Gargya and T. Leng, "Automated identification of diabetic retinopathy using deep learning," *Ophthalmology*, vol. 124, pp. 962–969, 2017.
- [39] W. M. Gondal, J. M. Köhler, R. Grzeszick, G. A. Fink, and M. Hirsch, "Weakly-supervised localization of diabetic retinopathy lesions in retinal fundus images," in *Proc. IEEE Int. Conf. Image Process.*, 2017, pp. 2069–2073.
- [40] B. Zhou, A. Khosla, A. Lapedriza, A. Oliva, and A. Torralba, "Learning deep features for discriminative localization," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 2921–2929.
- [41] Z. Lin *et al.*, "A framework for identifying diabetic retinopathy based on anti-noise detection and attention-based fusion," in *Proc. Int. Conf. Med. Image Comput. Comput.-assisted Intervention*, 2018, pp. 74–82.
- [42] S. Abbasi-Sureshjani, B. Dashtbozorg, B. M. ter Haar Romeny, and F. Fleuret, "Boosted exudate segmentation in retinal images using residual nets," in *Fetal, Infant. and Ophthalmic. Medical. Image. Analysis.*. Berlin, Germany: Springer, 2017, pp. 210–218.
- [43] H. Huang, H. Ma, H. J. van Triest, Y. Wei, and W. Qian, "Automatic detection of neovascularization in retinal images using extreme learning machine," *Neurocomputing*, 2017, pp. 210–218.
- [44] A. Levenkova, A. Sowmya, M. Kalloniatis, A. Ly, and A. Ho, "Lesion detection in ultra-wide field retinal images for diabetic retinopathy diagnosis," in *Proc. Med. Imag.: Comput.-Aided Diagnosis*, 2018, vol. 10575, Art. no. 1057531.
- [45] C. Lam, C. Yu, L. Huang, and D. Rubin, "Retinal lesion detection with deep learning using image patches," *Investigative Ophthalmol. Vis. Sci.*, vol. 59, no. 1, pp. 590–596, 2018.
- [46] Y. Guo, A. Camino, J. Wang, D. Huang, T. S. Hwang, and Y. Jia, "MedNet, a neural network for automated detection of avascular area in OCT angiography," *Biomed. Opt. Exp.*, vol. 9, no. 11, pp. 5147–5158, 2018.
- [47] M. H. Sarhan, S. Albarqouni, M. Yigitsoy, N. Navab, and A. Eslami, "Multi-scale microaneurysms segmentation using embedding triplet loss," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervention*, 2019, pp. 174–182.
- [48] G.-B. Huang, Q.-Y. Zhu, and C.-K. Siew, "Extreme learning machine: theory and applications," *Neurocomputing*, vol. 70, no. 1, pp. 489–501, 2006.
- [49] F. L. Ferris *et al.*, "A simplified severity scale for age-related macular degeneration: Areds report no. 18," *Archives ophthalmol. (Chicago, Ill.: 1960)*, vol. 123, no. 11, pp. 1570–1574, 2005.
- [50] P. M. Burlina, N. Joshi, M. Pekala, K. D. Pacheco, D. E. Freund, and N. M. Bressler, "Automated grading of age-related macular degeneration from color fundus images using deep convolutional neural networks," *JAMA Ophthalmol.*, vol. 135, no. 11, pp. 1170–1176, 2017.
- [51] F. Grassmann *et al.*, "A deep learning algorithm for prediction of age-related eye disease study severity scale for age-related macular degeneration from color fundus photography," *Ophthalmology*, vol. 125, no. 9, pp. 1410–1420, 2018.
- [52] S. P. K. Karri, D. Chakraborty, and J. Chatterjee, "Transfer learning based classification of optical coherence tomography images with diabetic macular edema and dry age-related macular degeneration," *Biomed. Opt. Exp.*, vol. 8, no. 2, pp. 579–592, 2017.
- [53] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2015, pp. 1–9.
- [54] D. S. Kermany *et al.*, "Identifying medical diagnoses and treatable diseases by image-based deep learning," *Cell*, vol. 172, no. 5, pp. 1122–1131, 2018.
- [55] S. Farsiu *et al.*, "Quantitative classification of eyes with and without intermediate age-related macular degeneration using optical coherence tomography," *Ophthalmology*, vol. 121, no. 1, pp. 162–172, 2014.
- [56] P. P. Srinivasan *et al.*, "Fully automated detection of diabetic macular edema and dry age-related macular degeneration from optical coherence tomography images," *Biomed. Opt. Exp.*, vol. 5, no. 10, pp. 3568–3577, 2014.
- [57] S. J. Chiu, M. J. Allingham, P. S. Mettu, S. W. Cousins, J. A. Izatt, and S. Farsiu, "Kernel regression based segmentation of optical coherence tomography images with diabetic macular edema," *Biomed. Opt. Exp.*, vol. 6, no. 4, pp. 1172–1194, 2015.
- [58] OPTIMA, "Proceeding of the miccai 2015 optima cysts segmentation challenge," 2015. Accessed: Apr. 22, 2020. [Online]. Available: <https://optima.meduniwien.ac.at/research/challenges/>. 2015
- [59] H. Bogunović *et al.*, "Retouch-the retinal OCT fluid detection and segmentation benchmark and challenge," *IEEE Trans. Med. Imag.*, vol. 38, no. 8, pp. 1858–1874, Aug. 2019.
- [60] A. K. Feeny, M. Tadarati, D. E. Freund, N. M. Bressler, and P. Burlina, "Automated segmentation of geographic atrophy of the retinal epithelium via random forests in areds color fundus images," *Comput. Biol. Med.*, vol. 65, pp. 124–136, 2015.
- [61] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Proc. Int. Conf. Med. Image Comput. Comput.-assisted Intervention*, 2015, pp. 234–241.
- [62] Z. Ji, Q. Chen, S. Niu, T. Leng, and D. L. Rubin, "Beyond retinal layers: A deep voting model for automated geographic atrophy segmentation in SD-OCT images," *Translational Vis. Sci. Technol.*, vol. 7, no. 1, p. 1, 2018.

- [63] J. De Fauw *et al.*, "Clinically applicable deep learning for diagnosis and referral in retinal disease," *Nature Med.*, vol. 24, no. 9, pp. 1342–1350, 2018.
- [64] J. Cheng *et al.*, "Superpixel classification based optic disc and optic cup segmentation for glaucoma screening," *IEEE Trans. Med. Imag.*, vol. 32, no. 6, pp. 1019–1032, Jun. 2013.
- [65] J. Deng *et al.*, "Age-related macular degeneration detection and stage classification using choroidal OCT images," in *Proc. Int. Conf. Image Anal. Recognition.*, 2016, pp. 707–715.
- [66] C. S. Lee, D. M. Baughman, and A. Y. Lee, "Deep learning is effective for classifying normal versus age-related macular degeneration optical coherence tomography images," *Ophthalmol. Retina*, vol. 1, pp. 322–327, 2017.
- [67] F. G. Venhuizen *et al.*, "Automated staging of age-related macular degeneration using optical coherence tomography," *Investigative Ophthalmol. Vis. Sci.*, vol. 58, no. 4, pp. 2318–2328, 2017.
- [68] M. Treder, J. L. Lauer, and N. Eter, "Automated detection of exudative age-related macular degeneration in spectral domain optical coherence tomography using deep learning," *Graefes Arch. Clin. Exp. Ophthalmol.*, vol. 256, no. 2, pp. 259–265, Feb. 2018.
- [69] T. Schlegl *et al.*, "Fully automated detection and quantification of macular fluid in OCT using deep learning," *Ophthalmology*, vol. 125, no. 4, pp. 549–558, 2018.
- [70] C. S. Lee, A. J. Tying, N. P. Deruyter, Y. Wu, A. Rokem, and A. Y. Lee, "Deep-learning based, automated segmentation of macular edema in optical coherence tomography," *Biomed. Opt. Exp.*, vol. 8, no. 7, pp. 3440–3448, 2017.
- [71] J. Xue, A. Camino, S. T. Bailey, X. Liu, D. Li, and Y. Jia, "Automatic quantification of choroidal neovascularization lesion area on OCT angiography based on density cell-like p systems with active membranes," *Biomed. Opt. Exp.*, vol. 9, no. 7, pp. 3208–3219, 2018.
- [72] Z. Li, Y. He, S. Keel, W. Meng, R. T. Chang, and M. He, "Efficacy of a deep learning system for detecting glaucomatous optic neuropathy based on color fundus photographs," *Ophthalmology*, vol. 125, no. 8, pp. 1199–1206, 2018.
- [73] J. Sivaswamy, S. Krishnadas, G. D. Joshi, M. Jain, and A. U. S. Tabish, "Drishti-gs: Retinal image dataset for optic nerve head (onh) segmentation," in *Proc. IEEE 11th Int. Symp. Biomed. Imag.*, 2014, pp. 53–56.
- [74] E. J. Carmona, M. Rincón, J. García-Feijóo, and J. M. Martínez-de-la Casa, "Identification of the optic nerve head with genetic algorithms," *Artificial. Intell. Med.*, vol. 43, no. 3, pp. 243–259, 2008.
- [75] F. Fumero, S. Alayón, J. Sanchez, J. Sigut, and M. Gonzalez-Hernandez, "Rim-one: An open retinal image database for optic nerve evaluation," in *Proc. Int. Symp. Comput.-Based Med. Syst.*, 2011, pp. 1–6.
- [76] A. W. Foong *et al.*, "Rationale and methodology for a population-based study of eye diseases in malay people: The singapore malay eye study (simes)," *Ophthalmic Epidemiol.*, vol. 14, no. 1, pp. 25–35, 2007.
- [77] Z. Zhang *et al.*, "Origa-light: An online retinal fundus image database for glaucoma analysis and research," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2010, pp. 3065–3068.
- [78] C. C. Sng *et al.*, "Determinants of anterior chamber depth: the singapore chinese eye study," *Ophthalmology*, vol. 119, no. 6, pp. 1143–1150, 2012.
- [79] J. I. Orlando *et al.*, "Refuge challenge: A unified framework for evaluating automated methods for glaucoma assessment from fundus photographs," *Med. Image Anal.*, vol. 59, 2020, Art. no. 101570.
- [80] X. Chen, Y. Xu, D. W. K. Wong, T. Y. Wong, and J. Liu, "Glaucoma detection based on deep convolutional neural network," in *Proc. Annu. Int. Conf. IEEE Eng. Med. and Biol. Soc.*, 2015, pp. 715–718.
- [81] K.-K. Maninis, J. Pont-Tuset, P. Arbeláez, and L. Van Gool, "Deep retinal image understanding," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervention*, 2016, pp. 140–148.
- [82] J. H. Tan, U. R. Acharya, S. V. Bhandary, K. C. Chua, and S. Sivaprasad, "Segmentation of optic disc, fovea and retinal vasculature using a single convolutional neural network," *J. Comput. Sci.*, vol. 20, pp. 70–79, 2017.
- [83] R. Srivastava, J. Cheng, D. W. Wong, and J. Liu, "Using deep learning for robustness to parapapillary atrophy in optic disc segmentation," in *Proc. IEEE 12th Int. Symp. Biomed. Imag.*, 2015, pp. 768–771.
- [84] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. Int. Conf. Learn. Represent.*, 2015. [Online]. Available: <https://dblp.org/rec/bibtex/journals/corr/SimonyanZ14a>
- [85] P. Liskowski and K. Krawiec, "Segmenting retinal blood vessels with deep neural networks," *IEEE Trans. Med. Imag.*, vol. 35, no. 11, pp. 2369–2380, Nov. 2016.
- [86] H. A. Leopold, J. Orchard, J. Zelek, and V. Lakshminarayanan, "Use of gabor filters and deep networks in the segmentation of retinal vessel morphology," in *Proc. Imag., Manipulation, Anal. Biomolecules, Cells, Tissues XV*, 2017, vol. 10068, Paper 100680R.
- [87] S. Wang, Y. Yin, G. Cao, B. Wei, Y. Zheng, and G. Yang, "Hierarchical retinal blood vessel segmentation based on feature and ensemble learning," *Neurocomputing*, vol. 149, pp. 708–717, 2015.
- [88] D. Maji, A. Santara, P. Mitra, and D. Sheet, "Ensemble of deep convolutional neural networks for learning to detect retinal vessels in fundus images," 2016, *arXiv:1603.04833*.
- [89] S. Xie and Z. Tu, "Holistically-nested edge detection," in *Proc. IEEE Int. Conf. Comput. Vis.*, 2015, pp. 1395–1403.
- [90] P. Costa *et al.*, "End-to-end adversarial retinal image synthesis," *IEEE Trans. Med. Imag.*, vol. 37, no. 3, pp. 781–791, Mar. 2017.
- [91] P. Prentašić *et al.*, "Segmentation of the foveal microvasculature using deep learning networks," *J. Biomed. Opt.*, vol. 21, no. 7, pp. 075 008–075 008, 2016.
- [92] F. G. Venhuizen *et al.*, "Robust total retina thickness segmentation in optical coherence tomography images using convolutional neural networks," *Biomed. Opt. Exp.*, vol. 8, no. 7, pp. 3292–3316, 2017.
- [93] A. G. Roy *et al.*, "Relaynet: retinal layer and fluid segmentation of macular optical coherence tomography using fully convolutional networks," *Biomed. Opt. Exp.*, vol. 8, no. 8, pp. 3627–3642, 2017.
- [94] J. Staal, M. D. Abramoff, M. Niemeijer, M. A. Viergever, and B. Van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE Trans. Med. Imag.*, vol. 23, no. 4, pp. 501–509, Apr. 2004.
- [95] A. Hoover, V. Kouznetsova, and M. Goldbaum, "Locating blood vessels in retinal images by piecewise threshold probing of a matched filter response," *IEEE Trans. Med. Imag.*, vol. 19, no. 3, pp. 203–210, Mar. 2000.
- [96] C. G. Owen *et al.*, "Measuring retinal vessel tortuosity in 10-year-old children: validation of the computer-assisted image analysis of the retina (caiar) program," *Investigative Ophthalmol. Vis. Sci.*, vol. 50, no. 5, pp. 2004–2010, 2009.
- [97] J. Son, S. J. Park, and K.-H. Jung, "Retinal vessel segmentation in fundoscopic images with generative adversarial networks," 2017, *arXiv:1706.09318*.
- [98] T. Schlegl, S. M. Waldstein, W.-D. Vogl, U. Schmidt-Erfurth, and G. Langs, "Predicting semantic descriptions from medical images with convolutional neural networks," in *Proc. Int. Conf. Inf. Process. Med. Imag.*, 2015, pp. 437–448.
- [99] S. Apostolopoulos, S. De Zanet, C. Ciller, S. Wolf, and R. Sznitman, "Pathological OCT retinal layer segmentation using branch residual u-shape networks," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assisted Intervention*, 2017, pp. 294–301.
- [100] K. Gopinath and J. Sivaswamy, "Segmentation of retinal cysts from optical coherence tomography volumes via selective enhancement," *IEEE J. Biomed. Health Informat.*, vol. 23, no. 1, pp. 273–282, Jan. 2018.
- [101] T. Schlegl, P. Seeböck, S. M. Waldstein, G. Langs, and U. Schmidt-Erfurth, "f-anogan: Fast unsupervised anomaly detection with generative adversarial networks," *Med. Image Anal.*, vol. 54, pp. 30–44, 2019.
- [102] F. G. Venhuizen *et al.*, "Deep learning approach for the detection and quantification of intraretinal cystoid fluid in multivendor optical coherence tomography," *Biomed. Opt. Exp.*, vol. 9, no. 4, pp. 1545–1569, 2018.
- [103] I. Goodfellow *et al.*, "Generative adversarial nets," in *Proc. Advances Neural Inf. Process. Syst.*, 2014, pp. 2672–2680.
- [104] V. Cheplygina, M. de Bruijne, and J. P. Pluim, "Not-so-supervised: A survey of semi-supervised, multi-instance, and transfer learning in medical image analysis," *Med. Image Anal.*, vol. 54, pp. 280–296, 2019.
- [105] H.-Y. Lee, H.-Y. Tseng, J.-B. Huang, M. Singh, and M.-H. Yang, "Diverse image-to-image translation via disentangled representations," in *Proc. Eur. Conf. Comput. Vis.*, 2018, pp. 35–51.
- [106] I. Higgins *et al.*, "beta-vae: Learning basic visual concepts with a constrained variational framework," *Proc. Int. Conf. Learn. Representations*, 2017. [Online]. Available: <https://openreview.net/forum?id=Sy2fzU9gl>
- [107] M. H. Sarhan, A. Eslami, N. Navab, and S. Albarqouni, "Learning interpretable disentangled representations using adversarial vae's," in *Domain. Adaptation. and Represent. Transfer. and Medical. Image. Learning. with Less. Labels. and Imperfect. Data.*. Berlin, Germany: Springer International Publishing, 2019, pp. 37–44.