Deep learning for ultra-widefield imaging: a scoping review

Nishaant Bhambra1 · Fares Antaki2,3 · Farida El Malt1 · AnQi Xu4 · Renaud Duval2,3

Abstract
Purpose This article is a scoping review of published and peer-reviewed articles using deep-learning (DL) applied to ultra-widefield (UWF) imaging. This study provides an overview of the published uses of DL and UWF imaging for the detection of ophthalmic and systemic diseases, generative image synthesis, quality assessment of images, and segmentation and localization of ophthalmic image features.

Methods A literature search was performed up to August 31st, 2021 using PubMed, Embase, Cochrane Library, and Google Scholar. The inclusion criteria were as follows: (1) deep learning, (2) ultra-widefield imaging. The exclusion criteria were as follows: (1) articles published in any language other than English, (2) articles not peer-reviewed (usually preprints), (3) no full-text availability, (4) articles using machine learning algorithms other than deep learning. No study design was excluded from consideration.

Results A total of 36 studies were included. Twenty-three studies discussed ophthalmic disease detection and classification, 5 discussed segmentation and localization of ultra-widefield images (UWFIs), 3 discussed generative image synthesis, 3 discussed ophthalmic image quality assessment, and 2 discussed detecting systemic diseases via UWF imaging.

Conclusion The application of DL to UWF imaging has demonstrated significant effectiveness in the diagnosis and detection of ophthalmic diseases including diabetic retinopathy, retinal detachment, and glaucoma. DL has also been applied in the generation of synthetic ophthalmic images. This scoping review highlights and discusses the current uses of DL with UWF imaging, and the future of DL applications in this field.

Keywords Deep learning · Ultra-widefield imaging · Artificial intelligence · Scoping review · Machine learning · Quality assessment

Key messages
What is known
- Deep learning has been applied to ophthalmic images for the diagnosis of ophthalmic diseases using fundus photography
- The utility of deep learning applied to ultra-widefield imaging has not been summarized in a scoping review

New information
- A scoping review of all published and peer-reviewed articles using deep learning with ultra-widefield imaging until August 2021 is provided here
- This article discusses both the detection of ophthalmic diseases, as well as the use of deep learning technologies for the generation of new ophthalmic images, the detection of systemic diseases, and quality assessment of ophthalmic images.

© The Author(s), under exclusive licence to Springer-Verlag GmbH Germany, part of Springer Nature 2022
Abbreviations

7SF  7 Standard Field
AMD  Age-related macular degeneration
AD   Alzheimer’s Disease
AUROC Area under ROC
AUPRC Area under precision-recall
ANN  Artificial neural network
AutoML Automated machine learning
BCVA  Best corrected visual acuity
baPWV  Brachial-Ankle Pulse-Wave Velocity
BRVO  Branch RVO
CRVO  Central RVO
CNNs  Convolutional neural networks
DL   Deep learning
DNN  Deep neural network
DME  Diabetic macular edema
DR   Diabetic retinopathy
FA   Fluorescein angiography
FI   Fundus image
GC-IPL  Ganglion cell-inner plexiform layer
GANs  Generative adversarial networks
GON  Glaucomatous optic neuropathy
Grad-CAM  Gradient-weighted class activation mapping
HIITL Human-in-the-loop
IMH  Idiopathic macular hole
ML   Machine learning
MAE  Mean absolute error
MAPE  Mean-absolute-percent error
NPDR  Non-proliferative DR
NPRLs  Notable peripheral retinal lesions
OCTA  OCT angiography
OCT  Optical coherence tomography
PDR  Proliferative diabetic retinopathy
PSR  Proliferative sickle cell retinopathy
ROC  Receiver operating characteristic
RDR  Referrable DR
ROI  Region of interest
RD   Retinal detachment
RED  Retinal exudates And/Or Drusen
RH   Retinal hemorrhage
RVO  Retinal vein occlusion
RVA  Retinal vessel areas
RP   Retinitis pigmentosa
RMSE  Root-mean-square error
SLO  Scanning laser ophthalmoscope
SCR  Sickle cell retinopathy
SSIM  Structural similarity
UWF  Ultra-widefield
UWF-FA  Ultra-widefield fluorescein angiography
UWFI  Ultra-widefield image
UWF-FAF  UWF fundus autofluorescence
UWF-ICGA  UWF indocyanine green angiography
VTDR  Vision-threatening DR

Introduction

In 1926, the first fundus camera was introduced by Zeiss and Nordensen. At that time, the camera provided only a 20-degree field of view. Shortly thereafter, an improved camera provided practitioners with a 30-degree field of view of the fundus [1]. While a major advance at the time, these cameras provided ophthalmologists with a limited view of the retinal periphery. In 1981, the Diabetic Retinopathy Study provided an objective method to visualize up to 75-degrees of the retina by combining seven conventional 30-degree fundus images (FIs) [2, 3]. This image type, known as 7 Standard Field (7SF) imaging, was the gold standard used in imaging for diagnosing diabetic retinopathy (DR). This remained the gold standard until technical developments in widefield (WF) imaging and ultra-widefield (UWF) imaging.

Widefield and ultra-widefield imaging

The International Widefield Imaging Study Group established anatomic definitions of widefield images (WFIs) as “images depicting retinal anatomic features beyond the posterior pole, but posterior to the vortex vein ampulla, in all 4 quadrants” while describing ultra-widefield images (UWFIs) as “images showing retinal anatomic features anterior to the vortex vein ampullae in all 4 quadrants [4].”

WF imaging utilizes a scanning laser ophthalmoscope (SLO), which separates the illuminating and imaging lasers used [5]. By separating the beams, WF imaging reduces artifacts produced from the interfaces in the ocular media [5].

UWF imaging can provide up to a 200-degree view of the retina, which allows for visualization of the optic disk and the peripheral retina in the same view [6]. Multiple WF imaging systems are available, with each differing in their technology and their field of view. The first UWF imaging system was introduced in 2000 by Optos [7]. Optos (Optos Inc, Dunfermline, UK) captures 200 degrees of the retina in a single image. The image provides coverage of approximately 82% of the retinal surface and does so without direct patient contact [8]. Other UWF imaging systems include the Heidelberg Spectralis Ultra-Widefield module, a noncontact removable lens which is an add-on to the Heidelberg HRA cSLO (Heidelberg Engineering, Heidelberg, Germany). This module expands the viewing range of the system from 55-degrees to a full UWF view of the retina [6]. Other UWF imaging products include the Zeiss Claurus 500 retinal camera (Carl Zeiss AG, Oberkochen Germany), which provides color and high-resolution imaging across the UWF anatomic range [9].
Clinical utility of UWF systems

As UWF imaging has provided a broader view of the retina, it has consistently been more effective at diagnosing retinal disease than previous imaging modalities. Ultra-widefield fluorescein angiography (UWF-FA), which combines UWF imaging with fluorescein angiography (FA) to visualize vessels, has been significantly more effective in diagnosing DR than previous 7SF imaging [10]. UWFIs provide 3.2 times more retinal surface area than 7SF and allow for a more comprehensive assessment of peripheral lesions and non-perfusion in DR [11, 12]. A comparison of ultra-widefield imaging to colour fundus photography is provided in Fig 1.

In 15 (RD), UWF imaging has provided improved assessment of peripheral retina breaks in comparison to indirect ophthalmoscopy [13]. In 16, UWF imaging has been shown to have a high agreement with color digital stereoscopy (CDS) in evaluating vertical cup-to-disc ratio and may be as effective in diagnosing glaucoma as CDS [14]. In patients with age-related macular degeneration (AMD), it was found that peripheral retinal changes were highly prevalent, indicating UWF imaging’s greater value in diagnosing AMD than traditional fundoscopy [15]. From these findings, UWF imaging may be a window to the retina, as well as to the brain more broadly.

The aim of this survey is to review articles that apply DL models to UWF imaging specifically. The goals of this paper are not specifically to discuss the added benefit of imaging the periphery to improve diagnosis or prognosis. Instead, our review focuses on the clinical utility of DL for UWF imaging and the current state of this field. This follows previous reviews on UWF imaging, which have described the landscape of UWF imaging and its clinical use in ophthalmology, as well as additional issues pertaining to UWF imaging and its clinical ability [6, 16].

Machine learning, deep learning, and supervision

Machine learning (ML) refers to the ability of machines to generate associations and patterns between variables, to learn in a sense similar to humans. By simulating the neural networks of human brains, ML networks generate probabilities and associations between variables to emulate human intelligence [17]. ML algorithms often can draw inferences between variables that are either imperceptible to humans or are too complex for human associations [17]. ML is divided into categories based on the approaches taken to assist computers in learning, on a spectrum between supervised learning and unsupervised learning [18].

Deep learning (DL) is a subset of machine learning that uses multiple layers of learning to identify features in data [19]. For example, in processing a fundus image, a lower layer may identify the edges of the vasculature, while higher layers may then utilize these edges of the vasculature in context to identify vessels as larger objects.

Fig. 1 Comparison of optos ultra-widefield imaging (200 degrees field of view) to color fundus photography (45 degrees field of view). A: Optos ultra-widefield optomap color image of the left fundus of a patient with diabetic retinopathy with an overlaid color fundus photograph from the same patient eye over the optic disc and macula region.
Supervised learning refers to ML from human-provided input and output pairs. For example, supervised learning for a task of classifying images would require a set of labelled images with their corresponding classification. The dataset is completely labelled, such that there is no ambiguity in the model that is training from it. For example, in a dataset of ophthalmic disease images, all the images would be labelled with the name of the disease presented in the image. By training a model on images and their corresponding classification, machines can learn to infer relationships between the two. The trained model should then be able to take unlabeled input data and determine its classification [20]. While this method is the most effective at training these associations, it also requires the most human involvement. Models trained using fully labelled datasets have higher accuracy, but also require greater human involvement for labelling the datasets.

Between supervised and unsupervised learning lies semi-supervised learning, which refers to ML training from incompletely labelled training datasets. This approach provides the machine with an initial relationship between input and output data, without a fully labelled set [21]. The dataset is incompletely labelled, which requires the ML model to then learn from the labelled images and then classify the unlabeled images. For example, in a dataset of ophthalmic disease images, only some of the images would be labelled with the name of the disease presented in the image. The remaining images would need to be classified by the model that has learned from the labelled images, providing a machine-generated label for the previously unlabeled images. Models trained using semi-supervised approaches generally have lower accuracy than supervised learning on completely labelled datasets. However, as they are not fully labelled, they can be less intensive for individuals to label an entire dataset.

Unsupervised learning uses algorithms to learn patterns from data that lacks human labels and input. The dataset contains no labels whatsoever, requiring the ML model to learn data features to classify the data first [22]. For example, in training a model to classify images using an unsupervised approach, a successful unsupervised learning algorithm would determine features that correspond to a given cluster and categorize each into separate categories without input labels from humans. In an example of a set of ophthalmic disease images, the model would learn the features that cluster between images, and then attempt to classify the images into categories based on image and disease features. The ML model would then be trained from these developed categories for classifying new images. Unsupervised learning has the lowest accuracy of the approaches described but requires no human involvement for labelling data. A figure showing the differences between the datasets and labels used in supervised, semi-supervised, and unsupervised learning is provided in Fig. 2.

Human-in-the-loop (HITL) is an example of human involvement in the training of ML models. In these models, human input is used to validate or negate the prediction.
produced by the ML model, allowing the model to learn based on the responses of humans involved. This allows humans to direct the training of models. This is also useful for the training of models on unlabeled data, and provides a method to influence the training of ML models [23]. It is important to note that HITL is different from unsupervised learning, as the terms “supervised” and “unsupervised” refer to the labelling of data used in the learning processes, rather than human involvement.

**Building a deep learning model**

DL specifically associates variables along nodes in a computational neural network. By associating data along these nodes, the artificial neural network (ANN) assigns a positive weight to variables with positive correlations and negative weights to variables with negative correlations. These weights determine the contributory strength of an input variable to the outcome of the neural network [24]. This develops a network of probabilistic associations between input variables. This is analogous to biological neurons, where associations between neurons are strengthened or weakened with excitatory and inhibitory stimuli respectively [25]. By associating data along these ANNs, machines can learn, and train models based on input data. By associating features of the data across the nodes of the ANN, correlations between the data are strengthened or weakened. These connections between the nodes, known as edges, are analogous to the synapses in biological brains [26].

In DL, nodes are associated into multiple layers. Each layer contains a set of nodes, and often perform different transformations on the input data. Each neural network contains an input layer, where the data enters untransformed, and an output layer, which produces the learned result. Between each is zero to multiple hidden layers, where further learning of data features occurs. Input data is processed forward from the input layer until it reaches the output layer [27].

Convolutional neural networks (CNNs) build on ANNs by organizing data and nodes in three dimensions. Furthermore, CNNs separate feature extraction and classification into distinct layers. CNNs rely on a convolution layer, which performs a convolution operation on the data array or tensor. The convolution operation extracts high-level features from a data source, such as the edges of an image. By doing so, it reduces the spatial size of the data and flexibly adjusts to the features of the data that are deemed more important to higher-level processing [28]. For this reason, CNNs are especially useful in image processing, where their convolutional operation allows them to ignore noise and focus on higher-order image structures, like edges. Multiple CNN models exist, including LeNet, AlexNet, VGGNet, and InceptionResNetV2 [29–32].

**Code-free and automated machine learning**

The development of machine learning models has been challenging for clinicians, many of whom lack the technical expertise to train and develop machine learning models. For this reason, the advent of code-free and automated machine learning (AutoML) systems has helped to democratize access to the development of effective ML models in medicine and ophthalmology [33–35]. These solutions provide a graphical user interface for individuals to help build ML models without code via graphical-user interfaces that are much more user-friendly [36].

AutoML can be described as “AI that can build AI” as it allows non-technical users to develop AI tools that can achieve accuracies close or equal to those of technical users developing code-based ML solutions. This has been applied to a variety of ophthalmic data including tabular data from electronic medical records, optical coherence tomography (OCT) scans, standard and UWFIs images, as well as surgical videos [35, 37–40]. Further, feasibility studies have shown that non-technical ophthalmologists have been able to develop effective ML models using AutoML tools, such as Google Cloud AutoML [34]. With the continued development and use of AutoML tools, it is expected that ML models will be used more effectively and more widely both in ophthalmology and medicine more broadly.

**Training, validation, and testing**

Datasets are split into training, validation, and testing sets such that each respective step has data that is similar to a model’s intended input data. When datasets include more than one image per patient (right eye, left eye, and steered images), it is important to maintain patient-level splits ensuring that each image could be used either for training, validation, or testing but not all. This restriction eliminates the risk of train-test data contamination that can arise when DL models use non-clinically relevant patterns to drive predictions.

The training dataset is used to train the model to learn the weights and biases between dataset variables. This training dataset is the input data for the model to learn from. For this reason, the quality and quantity of this data will greatly impact the ability of the model to learn the features of interest [41].

Validation datasets are used to tune the hyperparameters of the model. Hyperparameters are parameters that control the learning process of the model, while parameters are the node weights that are derived from training the model. The results of the validation set are used by the engineer to determine the optimal hyperparameters for the learning process. For this reason, this dataset is also known as the “development” or “tuning” dataset. The model does not learn from this dataset and does not develop weights or biases that would alter the model in an automated sense [41].
Finally, the test dataset is used as the input data for evaluating the model. This dataset is used to determine the accuracy and effectiveness of the model. This test dataset is not used to adjust the model, nor does the model learn from it [41].

The ratio of the dataset located to each subset depends on the goals of the model being trained and evaluated. In models that require many hyperparameters to be adjusted, a larger validation set is recommended. However, a validation set is optional if the user does not intend to tune these hyperparameters. Similarly, the amount of data to be allocated towards the training set depends on the complexity of the data and the amount of learning needed [41].

Another approach to dividing the dataset is k-fold cross-validation. In this process, image data is divided into k groups, while k-1 groups are used as training data and one group is used as validation data. This repeats until each dataset becomes a validation dataset.

**Image preprocessing**

Image preprocessing serves multiple purposes when training a DL model on an imaging task. The first purpose is to conserve computational resources by resizing images [42]. Often, large images (i.e., the standard 3900 × 3072 pixels of UWFIs) will be resized to significantly smaller, lower-resolution images (i.e., 227 × 227 pixels) [43].

The second purpose of image pre-processing is to increase the data available for training the model and to train the model on generalized cases. This is done via data augmentation, which increases the size of the data by performing transformations on it and producing new combinations of data to train on [44]. This provides the training set data that may match altered or changed input data. For example, one could augment an image dataset by adding noise to the images. This would help the model learn to classify noisy images with features of a given label correctly. As the goal is to train these models to be useful on real-world data, the training data should contain the same “errors” or adjustments that imperfect real-world input data does. This then allows for the model to become more robust in detecting the features of interest. Common image augmentation methods include adjusting brightness, gamma correction, histogram equalization, noise addition, and inversion [45]. Data augmentation can increase the size of the training set manyfold, often up to five or eighteen times the original dataset size [45, 46].

**Evaluating the model**

In ML classification tasks, predictions are classified into true positives, true negatives, false positives, and false negatives. From these values, the sensitivity and specificity values of the model are calculated. Sensitivity refers to the ability of the model to correctly predict positive observations, calculated as the number of true positives divided by the sum of the true positives and false negatives. Specificity refers to the ability to reject classifications for cases that do not fit the condition, calculated as the number of true negatives divided by the sum of the true negatives and number of false positives [47].

Sensitivity and specificity depend on the thresholds used for detection. As a threshold for detection increases, specificity (predicting positive outcomes) decreases while specificity (rejecting positive outcomes) increases. By plotting the sensitivity or true positive rate as a function of the specificity or false positive rate, the receiver operating characteristic (ROC) curve is produced. The integral of the entire curve is referred to as the area under ROC (AUROC) curve. The AUROC value serves as a measure of the model to correctly classify and predict based on the test data. A perfectly predictive model would score 1 while a perfectly inaccurate test would score 0 [48]. The area under the precision-recall (AUPRC) curve is occasionally used and is the integral of the plot of positive predictive value (precision) as a function of the sensitivity (recall) [49].

Another method for evaluation of some models is the Dice coefficient. It is often used in quantifying the performance of image segmentation tasks, as it quantifies the spatial overlap between the object intended to be identified and the object successfully segmented. For example, a Dice coefficient would be beneficial in quantifying how successfully a ML model has segmented and detected vasculature in an image [50]. Dice coefficient scores range from 0 indicating no spatial overlap between the target and the area segmented, to 1 indicating complete overlap between the target and the area segmented. A higher Dice coefficient score therefore indicates greater effectiveness in identifying the structure of interest.

**Deep learning in standard ophthalmic imaging**

Deep learning has been applied extensively to other modalities of ophthalmic imaging, such as OCT, standard fundus imaging, and fundus auto-fluorescence (FAF) imaging. While outside the scope of this review, it is worth noting that the application of DL to standard FIs has been successful and well-established for the detection of referable DR, diabetic macular edema (DME), and AMD. OCT imaging has been used with DL for the detection of AMD, and glaucoma extensively. While we discuss here the application of DL to UWF imaging specifically, it should be noted that other ophthalmic imaging modalities have been effectively used with DL for the detection and classification of ophthalmic disease [51].
Deep learning in ultrawide field imaging

Methodology

A literature search was performed up to August 31st, 2021, using the following online databases: PubMed, Embase, Cochrane Library, and Google Scholar. Article screening was done by the senior author (RD). The inclusion criteria were as follows: (1) Deep Learning (2) Ultra-Widefield Imaging. The exclusion criteria were as follows: (1) articles published in any language other than English, (2) articles not peer-reviewed (usually preprints), (3) no full-text availability (4) articles using machine learning algorithms other than deep learning. No study design was excluded from consideration. The detailed search methodology was as follows: (“deep learning” OR “artificial intelligence” OR “machine learning”) AND (“Ultra-Widefield” OR “UWF” OR “UWFI” OR “Optos”).

A total of 36 studies were included. A full listing of included studies, authors, and their respective digital object identifiers are listed in Table 1. A full listing of included studies and their respective architectures, datasets, and experimental results are listed in Table 2. A listing of the included studies and descriptions of their ground-truth datasets, human annotation, grader level, and description of patient-level splits is available in Table 3. A chart detailing the number of included publications by year is included in Fig. 3. A map highlighting the number of publications by country is provided in Fig. 4.

Disease detection and classification

Disease detection and classification have been the most thoroughly investigated uses for UWF imaging with DL. Specifically, DL has been used for disease detection and classification of DR, RD, glaucoma, AMD, retinitis pigmentosa (RP), pachychoroid, retinal vein occlusion (RVO), idiopathic macular hole (IMH), retinal hemorrhage (RH), and sickle cell retinopathy (SCR).

Diabetic retinopathy

At the time of writing, six published peer-reviewed articles have explored DL with UWF imaging in DR patients [45, 52–57]. Five have specifically used UWFIs for the detection and classification of DR [45, 53–55, 57].

Wang et al. first used UWFIs to train a DL model for the detection of referable DR in 2018 [55]. In this study, 754 UWFIs were acquired from patients with diabetes presenting at the Narayana Nethralaya hospital in Bangalore, India. The images were transmitted and graded by “certified DR graders” at the Doheny Eye Institute, of which 643 were gradable and inputted into the algorithm. The study set a threshold of moderate non-proliferative DR (NPDR) or higher (i.e., level 2 or higher on the International Clinical Diabetic Retinopathy scale) as sufficient to warrant a referral to an ophthalmologist. The EYEART algorithm, developed and trained using standard flash color images, was applied here to UWFIs. The study used their proprietary and closed-source algorithm to automatically detect and quantify DR lesions, such as hemorrhages, microaneurysms, lipid exudates, and cotton wool spots in UWFIs. Half of the dataset was used to train the classifiers in the previously developed EYEART algorithm, while 50% remained as a testing dataset.

The algorithm found 21.22% of the images contained referral-warranted DR while the graders determined 30.77% contained referral-warranted DR. The algorithm was used on each eye independently, as well as at the patient level, where referable DR in either eye classified the patient as having DR. At the patient level, the algorithm achieved a 91.7% sensitivity, 50.0% specificity, and 0.873 AUROC. When using individual eyes, the algorithm achieved a 90.3% sensitivity, 53.6% specificity, and 0.851 AUROC. While the authors were able to achieve high sensitivity, the low specificity indicates a high number of false positives using the EYEART algorithm. While the results were promising, a full understanding of their ML methods cannot be determined as the algorithm is closed source. Considering that EYEART was designed on FIs, it is expected that algorithms designed and trained on UWFIs would be more effective.

Nagasawa et al. published a study in 2019, which used DL for detecting treatment-naïve proliferative diabetic retinopathy (PDR) from UWFIs [56]. In this study, 378 UWFIs were graded for the presence of PDR by three retina specialists using the Early Treatment Diabetic Retinopathy Severity (ETDRS) scale, which includes disease features such as RH and neovascularization in grading presence of PDR.

The authors used the VGG-16 CNN, which automatically learns the local features of an image and generates a classification model [31]. The authors used 40 DL models from 40 learning cycles and chose the model with the highest correct answer rate from test data as the DL model for the study. The CNN selected achieved 94.7% sensitivity, 97.2% specificity, and 0.969 AUROC. Gradient-weighted class activation mapping (Grad-CAM) was utilized to visualize the image features used by the CNN to classify images as containing referable PDR.

The authors specifically used treatment-naïve PDR, which may have improved their results relative to Wang et al. Nonetheless, the authors were able to achieve high sensitivity and a high specificity, indicating that CNN approaches trained on UWFIs may be superior to applying algorithms designed for color FIs (i.e., the EyeArt algorithm) to UWFIs for DR detection.
| Author          | Title                                                                                                                                                                                                 | Country of Origin | Category                  | Subcategory                                                                 | Publication Title                          | DOI                                      | Date of Publication |
|-----------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|---------------------------|----------------------------------------------------------------------------|--------------------------------------------|------------------------------------------|---------------------|
| Antaki et al    | Accuracy of automated machine learning in classifying retinal pathologies from ultra-widefield pseudocolour fundus images                                                                                      | Canada            | Disease Detection and Classification | Retinal Detachment, Retinitis Pigmentosa, Retinal Vein Occlusion           | British Journal of Ophthalmology          | 10.1371/journal.pone.0238958           | August 3, 2021      |
| Bawany, M et al | Automated vessel density detection in fluorescein angiography images correlates with vision in proliferative diabetic retinopathy                                                                            | USA               | Disease Detection and Classification | Diabetic Retinopathy                                                      | PLOS ONE                                  | 10.1371/journal.pone.0238958           | September 11, 2020  |
| Cai, S et al    | Deep Learning Detection of Sea Fan Neo-vascularization From Ultra-Widefield Color Fundus Photographs of Patients With Sickle Cell Hemoglobinopathy                                                       | USA               | Disease Detection and Classification | Sickle Cell Retinopathy                                                   | JAMA Ophthalmology                        | 10.1001/jamaophthal-mol.2020.5900       | December 30, 2020   |
| Calderon-Auza, G et al | A Teleophthalmology Support System Based on the Visibility of Retinal Elements Using the CNNs                                                                                                          | Mexico            | Quality Assessment          | Sensors                                                                  | Sensors                                   | 10.3390/s20102838                      | January 1, 2020     |
| Ding, L et al   | A Novel Deep Learning Pipeline for Retinal Vessel Detection In Fluorescein Angiography                                                                                            | USA               | Segmentation and Localization | Vessel Segmentation                                                      | IEEE Transactions on Image Processing     | 10.1109/TIP.2020.2991530               | January 1, 2020     |
| Ding, L et al   | Weakly-Supervised Vessel Detection in Ultra-Widefield Fundus Photography Via Iterative Multi-Modal Registration and Learning                                                                               | USA               | Segmentation and Localization | Vessel Segmentation                                                      | IEEE Transactions on Medical Imaging      | 10.1109/TMI.2020.3027665               | September 29, 2020  |
| Author          | Title                                                                 | Country of Origin | Category                       | Subcategory            | Publication Title                  | DOI                                | Date of Publication |
|-----------------|-----------------------------------------------------------------------|-------------------|--------------------------------|------------------------|------------------------------------|------------------------------------|---------------------|
| Ju, L et al     | Leveraging Regular Fundus Images for Training UWF Fundus Diagnosis Models via Adversarial Learning and Pseudo-Labeling | Australia         | Generative Image Synthesis using GANs |                        | IEEE Transactions on Medical Imaging | 10.1109/TMI.2021.3056395        | November 27, 2020   |
| Kim, I et al    | Classification of pachychoroid disease on ultrawide-field indocyanine green angiography using auto-machine learning platform | South Korea       | Disease Detection and Classification | Pachychoroid          | British Journal of Ophthalmology   | 10.1136/bjophthalmol-2020–316,108 | July 3, 2020        |
| Li Z et al      | Automated detection of retinal exudates and drusen in ultrawidefield fundus images based on deep learning | USA               | Disease Detection and Classification | AMD                   | Eye                                | 10.1038/s41433-021–01,715-7       | August 3, 2021      |
| Li, H et al     | Automated Quality Assessment and Image Selection of Ultra-Widefield Fluorescein Angiography Images through Deep Learning | China             | Quality Assessment               |                        | Translational Vision Science & Technology | 10.1167/tvst.9.2.52             | January 28, 2020    |
| Li, Z et al     | Deep learning for automated glaucomatous optic neuropathy detection from ultrawidefield fundus images | China             | Disease Detection and Classification | Glaucoma              | British Journal of Ophthalmology   | 10.1136/bjophthalmol-2020–317,327 | September 16, 2020  |
| Li, Z et al     | A deep learning system for identifying lattice degeneration and retinal breaks using ultrawidefield fundus images | China             | Disease Detection and Classification | Retinal Detachment     | Annals of Translational Medicine    | 10.21037/atm.2019.11.28          | November 1, 2019    |
| Author    | Title                                                                 | Country of Origin | Category                                     | Subcategory               | Publication Title                                      | DOI                                      | Date of Publication |
|-----------|------------------------------------------------------------------------|-------------------|----------------------------------------------|---------------------------|--------------------------------------------------------|------------------------------------------|---------------------|
| Li, Z et al | Deep learning for detecting retinal detachment and discerning macular status using ultra-widefield fundus images | China             | Disease Detection and Classification          | Retinal Detachment Biology | Communications Biology                                  | 10.1038/s42003-019-0730-x               | December 1, 2020    |
| Li, Z et al | Development and Evaluation of a Deep Learning System for Screening Retinal Hemorrhage Based on Ultra-Widefield Fundus Images | China             | Disease Detection and Classification          | Retinal Hemorrhage        | Translational Vision Science & Technology               | 10.1167/tvst.9.2.3                      | January 29, 2020    |
| Li, Z et al | Deep learning from "passive feeding" to "selective eating" of real-world data | China             | Quality Assessment                            |                           | npj Digital Medicine                                   | 10.1038/s41746-020-00,350-y             | December 1, 2020    |
| Masumoto, H et al | Deep-learning Classifier With an Ultrawide-field Scanning Laser Ophthalmoscope Detects Glaucoma Visual Field Severity: | Japan             | Disease Detection and Classification          | Glaucoma                  | Journal of Glaucoma                                    | 10.1097/IJG.0000000000000988           | July 1, 2018        |
| Masumoto, H et al | Accuracy of a deep convolutional neural network in detection of retinitis pigmentosa on ultra-widefield images | Japan             | Disease Detection and Classification          | Retinitis Pigmentosa      | PeerJ                                                  | 10.7717/peerj.6900                     | May 7, 2019         |
| Matsuba, S et al | Accuracy of ultra-widefield fundus ophthalmoscopy-assisted deep learning, a machine-learning technology, for detecting age-related macular degeneration | Japan             | Disease Detection and Classification          | AMD                       | International Ophthalmology                            | 10.1007/s10792-018-0940-0             | June 1, 2019        |
| Author                | Title                                                                 | Country of Origin | Category               | Subcategory                                | Publication Title                  | DOI                                  | Date of Publication |
|-----------------------|------------------------------------------------------------------------|-------------------|------------------------|--------------------------------------------|------------------------------------|-------------------------------------|---------------------|
| Nagasato, D et al     | Prediction of age and brachial-ankle pulse-wave velocity using ultra-wide-field pseudo-color images by deep learning | Japan             | Systemic Diseases      | Age and Brachial-ANKle Pulse-Wave Velocity | Scientific Reports                 | 10.1038/s41598-020-76,513-4 | December 1, 2020    |
| Nagasato, D et al     | Deep-learning classifier with ultrawide-field fundus ophthalmoscopy for detecting branch retinal vein occlusion | Japan             | Disease Detection and Classification | BRVO                                      | International Journal of Ophthalmology | 10.18240/ijo.2019.01.15 | January 18, 2019    |
| Nagasato, D et al     | Deep Neural Network-Based Method for Detecting Central Retinal Vein Occlusion Using Ultrawide-Field Fundus Ophthalmoscopy | Japan             | Disease Detection and Classification | CRVO                                      | Journal of Ophthalmology           | 10.1155/2018/1875431 | November 1, 2018    |
| Nagasawa et al        | Retinopathy Staging with a Deep Convolutional Neural Network Using Ultra-Wide-Field Fundus Ophthalmoscopy and Optical Coherence Tomography Angiography | Japan             | Disease Detection and Classification | Diabetic Retinopathy                   | Journal of Ophthalmology           | 10.1155/2021/6651175 | April 5, 2021       |
| Nagasawa, T et al     | Accuracy of ultrawide-field fundus ophthalmoscopy-assisted deep learning for detecting treatment-naA’ve proliferative diabetic retinopathy | Japan             | Disease Detection and Classification | Diabetic Retinopathy                   | International Ophthalmology        | 10.1007/s10792-019-01,074-z | October 1, 2019     |
| Nagasawa, T et al     | Accuracy of deep learning, a machine learning technology, using ultra-wide-field fundus ophthalmoscopy for detecting idiopathic macular holes | Japan             | Disease Detection and Classification | Idiopathic Macular Hole                 | PeerJ                              | 10.7717/peerj.5696  | October 22, 2018    |

Table 1 (continued)
Table 1 (continued)

| Author           | Title                                                                 | Country of Origin | Category                          | Subcategory           | Publication Title                                         | DOI                          | Date of Publication |
|------------------|-----------------------------------------------------------------------|-------------------|-----------------------------------|-----------------------|----------------------------------------------------------|------------------------------|---------------------|
| Nunez do Rio, J  et al | Deep Learning-Based Segmentation and Quantification of Retinal Capillary Non-Perfusion on Ultra-Wide-Field Retinal Fluorescein Angiography | UK                | Segmentation and Localization     | Vessel Segmentation  | Journal of Clinical Medicine                             | 10.3390/jcm9082537           | August 6, 2020      |
| Oh, K et al      | Early detection of diabetic retinopathy based on deep learning and ultra-wide-field fundus images | South Korea       | Disease Detection and Classification | Diabetic Retinopathy  | Scientific Reports                                        | 10.1038/s41598-021-81,539-3  | January 21, 2021    |
| Ohsugi, H et al  | Accuracy of deep learning, a machine-learning technology, using ultra-wide-field fundus ophthalmoscopy for detecting rhegmatogenous retinal detachment | Japan             | Disease Detection and Classification | Retinal Detachment   | Scientific Reports                                        | 10.1038/s41598-017-09,891-x  | December 1, 2017    |
| Sevgi, D et al   | Deep learning-enabled ultra-widefield retinal vessel segmentation with an automated quality-optimized angiographic phase selection tool | USA               | Segmentation and Localization     | Vessel Segmentation  | Eye                                                      | 10.1038/s41433-021-01,661-4  | January 1, 2021     |
| Shi et al        | A method for the automatic detection of myopia in Optos fundus images based on deep learning | China             | Disease Detection and Classification | Myopia                | International Journal for Numerical Methods in Biomedical Engineering | 10.1002/cnm.3460            | March 26, 2021      |
| Tang, F et al    | Detection of Diabetic Retinopathy from Ultra-Wide Field Scanning Laser Ophthalmoscope Images: A Multi-Center Deep-Learning Analysis | China             | Disease Detection and Classification | Diabetic Retinopathy  | Ophthalmology Retina                                      | 10.1016/j.oret.202101.013    | February 1, 2021    |
| Author      | Title                                                                 | Country of Origin | Category                      | Subcategory             | Publication Title                  | DOI                          | Date of Publication |
|------------|----------------------------------------------------------------------|-------------------|-------------------------------|-------------------------|-----------------------------------|------------------------------|---------------------|
| Wang, K et al | Automated detection of diabetic retinopathy lesions on ultrawidefield pseudocolour images | USA               | Disease Detection and Classification | Diabetic Retinopathy | Acta Ophthalmologica                | 10.1111/aos.13528          | March 1, 2018       |
| Wang, Z et al  | Multi-Task Siamese Network for Retinal Artery/Vein Separation via Deep Convolution Along Vessel | China             | Segmentation and Localization | Vessel Segmentation    | IEEE Transactions on Medical Imaging | 10.1109/TMI.2020.2980117    | September 1, 2020   |
| Wisely, C et al | Convolutional neural network to identify symptomatic Alzheimer's disease using multimodal retinal imaging | USA               | Systemic Diseases             | Alzheimer's Disease    | British Journal of Ophthalmology   | 10.1136/bjophthalmol-2020–317,659 | November 26, 2020   |
| Xie, H et al   | AMD-GAN: Attention encoder and multi-branch structure based generative adversarial networks for fundus disease detection from scanning laser ophthalmoscopy images | China             | Generative Image Synthesis using GANs |                          | Neural Networks                    | 10.1016/j.neuronet.2020.09.005 | December 1, 2020    |
| Yoo, T et al    | Deep learning can generate traditional retinal fundus photographs using ultrawidefield images via generative adversarial networks | South Korea       | Generative Image Synthesis using GANs |                          | Computer Methods and Programs in Biomedicine | 10.1016/j cmpb.2020.105761 | December 1, 2020    |
| Zhang, C et al  | Development of a deep-learning system for detection of lattice degeneration, retinal breaks, and retinal detachment in tessellated eyes using ultra-wide-field fundus images: a pilot study | China             | Disease Detection and Classification | Retinal Detachment    | Graefe's Archive for Clinical and Experimental Ophthalmology | 10.1007/s00417-021-05,105-3 | February 4, 2021    |
| Author          | Title                                                                 | Architecture(s)               | Total Dataset Size | Train/Validate/Test | AUROC or AUPRC | Sensitivity | Specificity | Precision | Accuracy |
|-----------------|----------------------------------------------------------------------|-------------------------------|--------------------|---------------------|----------------|-------------|-------------|-----------|----------|
| Antaki et al    | Accuracy of automated machine learning in classifying retinal patholog-ies from ultra-widefield pseudocolor fundus images | Google Cloud AutoML          | 2137 UWFI          | 80/10/10            | AUPRC: RVO: 0.8491 RP: 0.8800 RD: 0.9210 | RVO: 0.9670 RP: 0.9420 RD: 0.9077 | RVO: 1.000 RP: 1.000 RD: 0.7872 | RVO: 1.0000 RP: 1.0000 RD: 0.8876 |
| Bawany, M et al | Automated vessel density detection in fluorescein angiography images correlates with vision in proliferative diabetic retinopathy | U-NET                        | 42 UWF-FA          | 19/0/81             | AUPRC: 0.9300 |             |             |           |          |
| Cai, S et al    | Deep Learning Detection of Sea Fan Neovascularization From Ultra-Widefield Color Fundus Photographs of Patients With Sickle Cell Hemoglobinopathy | InceptionV4 CNN               | 1182 UWFI          | 70/10/20            | AUROC: 0.988 0.9740 0.9700 0.9700 |             |             | 0.9700    |          |
| Calderon-Auza, G et al | A Teleophthalmology Support System Based on the Visibility of Retinal Elements Using the CNNs | OD Detection: FR-CNN AlexNet OD Quality Analysis: VGG-16 Obstruction Analysis: SegNet Vessel Segmentation: SegNet with VGG-16 | 1288 UWFI          | OD Detection and Quality Analysis: 700/30 Obstruction Analysis: 800/20 Vessel Segmentation: 900/10 | OD Detection: 0.9643 OD Quality Analysis: 0.9113 Obstruction Analysis: N/A Vessel Segmentation: 0.7169 | OD Detection: 0.4424 OD Quality Analysis: 0.8646 Obstruction Analysis: N/A Vessel Segmentation: 0.9816 | OD Detection: 0.9524 OD Quality Analysis: 0.8612 Obstruction Analysis: 1.0000 Vessel Segmentation: 0.9784 |
Table 2 (continued)

| Author       | Title                                                                                                                                                                                                 | Architecture(s)                              | Total Dataset Size | Train/Validate/ Test | AUROC or AUPRC | Sensitivity | Specificity | Precision | Accuracy |
|--------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|----------------------------------------------|--------------------|----------------------|----------------|--------------|-------------|-----------|----------|
| Ding, L et al | A Novel Deep Learning Pipeline for Retinal Vessel Detection In Fluorescein Angiography                                                                                                                 | GAN with UNet followed by human-in-the-loop approach | 8 UWFFA            | k-fold cross validation (k = 8) | AUROC: 0.987 AUPRC: 0.93 |              |             |           |          |
| Ding, L et al | Weakly-Supervised Vessel Detection in Ultra-Wide-field Fundus Photography Via Iterative Multi-Modal Registration and Learning                                                                          | Custom DNN                                   | PRIME-FP20 dataset: 15 concurrently captured UWFI and UWG-FA DRIVE dataset: 40 FP STARE dataset: 40 FP | PRIME-FP20: k-fold cross validation (k = 5) DRIVE dataset: trained DRIVE, tested STARE STARE dataset: trained on STARE, tested on DRIVE | AUPRC: PRIME-FP20: 0.8450 DRIVE: 0.8860 STARE: 0.8400 |              |             |           |          |
| Ju, L et al  | Leveraging Regular Fundus Images for Training UWF Fundus Diagnosis Models via Adversarial Learning and Pseudo-Labeilig                                                                                   | Transfer of FP to UWFI: modified Cycle-GAN Evaluation: Classification: ResNet-50 Detection: YOLOv3 Segmentation: U-Net with ResNet50 | Classification: 552 UWFI; 2500 FP Detection: 242 UWFI; 1169 FP Segmentation: 175 UWFI; 1000 FP | Detection: 0.9101 Detection: 0.5211 Classification: 0.7190 Segmentation: 0.6818 |              |             |           |          |
| Kim, I et al | Classification of pachychoroid disease on ultrawide-field indocyanine green angiography using auto-machine learning platform                                                                          | Google AutoML Vision                         | 783 UWF-ICGA       | 80/10/10             | Model 1: 0.8182 Model 2: 0.8636 | Model 1: 0.8519 Model 2: 0.8889 | Model 1: 0.8182 Model 2: 0.8636 | Model 1: 0.8367 Model 2: 0.8776 |
| Author | Title                                                                 | Architecture(s)                      | Total Dataset Size | Train/Validate/Test | AUROC or AUPRC | Sensitivity | Specificity | Precision | Accuracy |
|--------|----------------------------------------------------------------------|---------------------------------------|--------------------|---------------------|----------------|-------------|-------------|-----------|----------|
| Li, Z et al | Automated detection of retinal exudates and drusen in ultra-widefield fundus images based on deep learning | Inception-ResNetV2 | 19,891 UWFI | 70/15/15 | AUROC: CMAAI: 0.9420 ZOC: 0.9490 ZOC: 0.972 XOH: 0.988 | CMAAI: 0.9420 | ZOC: 0.9490 | XOH: 0.9730 | XOH: 0.9720 |
| Li, H et al | Automated Quality Assessment and Image Selection of Ultra-Widefield Fluorescein Angiography Images through Deep Learning | “U-Net style” CNN | 3935 UWF-FAF | 90/0/10 | AUROC: Test Set 1 Gradeability: 0.9200 | Test Set 1 Gradeability: 0.8700 Quality: 0.7890 Test Set 2 Gradeability: 0.9860 Quality: 1.000 | Test Set 1 Gradeability: 0.8900 Quality: 0.8930 Test Set 2 Gradeability: 0.9400 Quality: 0.9700 |
| Li, Z et al | Deep learning for automated glaucomatous optic neuropathy detection from ultra-widefield fundus images | Inception-ResNetV2 | 22,972 UWFI | 70/15/15 | AUROC: CMAAI: 0.9990 ZOC: 0.9830 XOH: 0.9900 TOPP: 0.9900 | CMAAI: 0.9750 ZOC: 0.9790 XOH: 0.9820 TOPP: 0.9790 | CMAAI: 0.9840 ZOC: 0.9430 XOH: 0.9670 TOPP: 0.9730 | CMAAI: 0.9130 ZOC: 0.7930 XOH: 0.7690 TOPP: 0.9130 |
| Li, Z et al | A deep learning system for identifying lattice degeneration and retinal breaks using ultra-widefield fundus images | Inception-ResNetV2, InceptionV3, ResNet50, VGG16 | 5005 UWFI | 70/15/15 | AUROC: InceptionV3: 0.9870 InceptionV3: 0.9870 ResNet50: 0.9680 ResNet50: 0.9760 Inception-ResNetV2: 0.9920 Inception-ResNetV2: 0.9920 VGG16: 0.9740 VGG16: 0.9880 | InceptionV3: 0.9870 ResNet50: 0.9750 Inception-ResNetV2: 0.9910 VGG16: 0.9850 |
| Li, Z et al | Deep learning for detecting retinal detachment and discerning macular status using ultra-widefield fundus images | Inception-ResNetV2 | 11,087 UWFI | 70/15/15 | AUROC: RD detection: 0.9890 Macular status classification: 0.9750 | RD detection: 0.9610 Macular status classification: 0.9380 | RD detection: 0.9600 Macular status classification: 0.9090 | RD detection: 0.9890 Macular status classification: 0.9170 |
| Author       | Title                                                                 | Architecture(s) | Total Dataset Size | Train/Validate/Test | AUROC or AUPRC | Sensitivity | Specificity | Precision | Accuracy |
|--------------|------------------------------------------------------------------------|------------------|--------------------|---------------------|-----------------|-------------|-------------|-----------|----------|
| Li, Z et al  | Development and Evaluation of a Deep Learning System for Screening Retinal Hemorrhage Based on Ultra-Widefield Fundus Images | Inception-ResNetV2 | 16,827 UWFI        | 70/15/15            | AUROC: CMAAI: 0.9990 CMAAI: 0.9940 ZOC: 0.9760 XOH: 0.9800 | CMAAI: 0.9890 ZOC: 0.9670 XOH: 0.9760 | CMAAI: 0.9940 ZOC: 0.9870 XOH: 0.9800 | CMAAI: 0.9930 ZOC: 0.9840 XOH: 0.9800 |
| Li, Z et al  | Deep learning from "passive feeding" to "selective eating" of real-world data | Inception-ResNetV2 | 40,562 UWFI        | 70/15/15            | AUROC: CMAAI: 0.9960 ZOC: 0.9940 XOH: 0.9970 | CMAAI: 0.9690 ZOC: 0.9560 XOH: 0.9660 | CMAAI: 0.9660 ZOC: 0.9790 XOH: 0.9880 |
| Masumoto, H et al | Deep-learning Classifier With an Ultrawide-field Scanning Laser Ophthalmoscope Detects Glaucoma Visual Field Severity: | Custom DCNN      | 1399 UWFI          | 80/0/20             | AUROC: Normal vs All: 0.813 Normal vs Early: 0.830 Normal vs Moderate: 0.864 Normal vs Severe: 0.934 | Normal vs All: 0.802 Normal vs Early: 0.753 Normal vs Moderate: 0.902 Normal vs Severe: 0.958 |
| Masumoto, H et al | Accuracy of a deep convolutional neural network in detection of retinitis pigmentosa on ultrawide-field images | VGG-16           | 373 UWFI           | k-fold cross validation (k = 5) | AUROC: UWFI: 0.9930 UW-FAF: 0.9990 | UWFI: 0.9930 | UW-FAF: 1.000 | UW-FAF: 0.9950 |
| Author          | Title                                                                 | Architecture(s)                  | Total Dataset Size | Train/Validate/Test | AUROC or AUPRC | Sensitivity | Specificity | Precision | Accuracy |
|-----------------|------------------------------------------------------------------------|----------------------------------|--------------------|---------------------|----------------|-------------|-------------|-----------|----------|
| Matsuba, S et al | Accuracy of ultra-wide-field fundus ophthalmoscopy-assisted deep learning, a machine-learning technology, for detecting age-related macular degeneration | Custom DCNN                      | 364 UWFI           | 70/0/30             | AUROC: 0.9976 | 1.000       | 0.9731      |           | 1.000    |
| Nagasato, D et al | Prediction of age and brachial-ankle pulse-wave velocity using ultra-wide-field pseudo-color images by deep learning | VGG-16                            | 170 UWFI (central, peripheral, total) | k-fold cross validation (k = 5) | AUROC: | 1.000       | 0.9731      |           | 1.000    |
| Nagasato, D et al | Deep-learning classifier with ultrawide-field fundus ophthalmoscopy for detecting branch retinal vein occlusion | DL Model: VGG-16 SVM: Soft Margin | 466 UWFI           | k-fold cross validation (k = 9) | AUROC: DL: 0.9760 SVM: 0.8570 | DL: 0.9400 SVM: 0.8050 | DL: 0.9700 SVM: 0.8430 | DL: 0.9650 SVM: 0.8350 |
| Nagasato, D et al | Deep Neural Network-Based Method for Detecting Central Retinal Vein Occlusion Using Ultrawide-Field Fundus Ophthalmoscopy | DL Model: VGG-16 SVM: Soft Margin | 363 UWFI           | k-fold cross validation (k = 9) | AUROC: DL: 0.9890 SVM: 0.8950 | DL: 0.9840 SVM: 0.8400 | DL: 0.9790 SVM: 0.8750 |           |          |
| Author | Title | Architecture(s) | Total Dataset Size | Train/Validate/Test | AUROC or AUPRC | Sensitivity | Specificity | Precision | Accuracy |
|--------|-------|-----------------|--------------------|---------------------|-----------------|--------------|-------------|------------|----------|
| Nagasawa et al | Accuracy of Diabetic Retinopathy Staging with a Deep Convolutional Neural Network Using Ultra-Wide-Field Fundus Ophthalmoscopy and Optical Coherence Tomography Angiography | VGG-16 | 491 UWFI and OCTA | k-fold cross validation (k = 5) | AUROC: 0.790 | DR Detection: Optos: 0.809 | OCTA: 0.839 | Optos-OCTA: 0.7860 | 0.8470 | PDR Detection: Optos: 0.981 | OCTA: 0.9280 | Optos-OCTA: 0.9640 | 0.9640 | 0.9810 | 0.9280 | 0.9640 |
| Nagasawa, T et al | Accuracy of deep learning, a machine learning technology, using ultra-wide-field fundus ophthalmoscopy for detecting idiopathic macular holes | Custom DCNN | 910 UWFI | 70/0/30 | AUROC: 0.999 | 1.000 | 0.995 | 1.000 |
| Nagasawa, T et al | Accuracy of ultrawide-field fundus ophthalmoscopy-assisted deep learning for detecting treatment-naive proliferative diabetic retinopathy | VGG-16 | 378 UWFI | k-fold cross validation (k = 9) | AUROC: 0.969 | 0.947 | 0.972 |
| Author          | Title                                                                 | Architecture(s)                           | Total Dataset Size | Train/Validate/Test | AUROC or AUPRC | Sensitivity | Specificity | Precision | Accuracy |
|-----------------|------------------------------------------------------------------------|-------------------------------------------|--------------------|---------------------|----------------|--------------|-------------|-----------|----------|
| Nunez do Rio, J et al | Deep Learning-Based Segmentation and Quantification of Retinal Capillary Non-Perfusion on Ultra-Wide-Field Retinal Fluorescein Angiography | "U-Net style" CNN                         | 75 UWF-FAF         | k-fold cross validation (k = 5) | AUROC: 0.82 AUPRC: 0.73 | 0.6609       |             | 0.7002    |          |
| Oh, K et al     | Early detection of diabetic retinopathy based on deep learning and ultra-wide-field fundus images | ResNet-34                                 | 13,271 UWFI        | k-fold cross validation (k = 10) | AUROC: ETDRS 7SF: 0.9150 ETDRS F1–F2: 0.8867 | ETDRS 7SF: 0.8338 ETDRS F1–F2: 0.8060 | ETDRS 7SF: 0.8338 ETDRS F1–F2: 0.8061 | 0.8338    | 0.8060   |
| Ohsugi, H et al | Accuracy of deep learning, a machine-learning technology, using ultra-wide-field fundus ophthalmoscopy for detecting rhegmatogenous retinal detachment | Custom CNN                                | 831 UWF            | 80/0/20             | AUROC: 0.988 | 0.9760       | 0.9650     |          |          |
| Sevgi, D et al  | Deep learning-enabled ultra-widefield retinal vessel segmentation with an automated quality-optimized angiographic phase selection tool | Custom CNN                                | RVA Extraction: 7787 UWF-FA patches Phase Selection: 13,980 UWF-FA sequences Cubic Spline Creation: 1578 UWF-FA sequences | 90/0/10             |              |              | 0.9650     |          |          |
| Author     | Title                                                                 | Architecture(s) | Total Dataset Size | Train/Validate/Test | AUROC or AUPRC | Sensitivity | Specificity | Precision | Accuracy |
|------------|-----------------------------------------------------------------------|------------------|--------------------|---------------------|-----------------|-------------|-------------|-----------|----------|
| Shi et al  | A method for the automatic detection of myopia in Optos fundus images based on deep learning | MDNet            | 7141 UWFI          | 95/2.5/2.5          |                 |             |             |           |          |
| Tang, F et al | Detection of Diabetic Retinopathy from Ultra-Wide Field Scanning Laser Ophthalmoscope Images: A Multi-Center Deep-Learning Analysis | ResNet50         | 9392 UWFI          | 80/20/0            | AUROC:          |             |             |           |          |
| Wang, K et al | Automated detection of diabetic retinopathy lesions on ultrawidefield pseudocolour images | EYEART           | 754 UWFI           | 50/0/50            | Gradeability:   |             |             |           |          |
| Wang, Z et al | Multi-Task Siamese Network for Retinal Artery/Vein Separation via Deep Convolution Along Vessel | Custom CNN: Convolution Along Vessel (CAV) utilizing ResNet-18 | WIDE: 30 UWFI, DRIVE: 40 FP, INSPIRE: 40 FP | k-fold cross validation (k = 2) | WIDE: 0.9600, DRIVE: 0.9690, INSPIRE: 0.9730 | WIDE: 0.9500, DRIVE: 0.9270, INSPIRE: 0.9660 | WIDE: 0.9450, DRIVE: 0.9470, INSPIRE: 0.9690 |
| Author     | Title                                                                 | Architecture(s)          | Total Dataset Size | Train/Validate/Test | AUROC or AUPRC | Sensitivity | Specificity | Precision | Accuracy |
|------------|------------------------------------------------------------------------|--------------------------|--------------------|---------------------|----------------|-------------|-------------|-----------|----------|
| Wisely, C et al | Convolutional neural network to identify symptomatic Alzheimer's disease using multimodal retinal imaging | ResNet18 | 284 UWF 284 UWF-FAF 284 GC-IPL 284 OCTA | AD: 35/10/55 Control: 21/11/68 | AUROC: UWFI: 0.4500 UWFFA: 0.6180 GC-IPL: 0.8090 OCTA: 0.5820 All: 0.8290 All + quantitative data: 0.8300 All + all data: 0.8360 GC-IPL + all data: 0.8410 |                        |            |            |          |          |
| Xie, H et al | AMD-GAN: Attention encoder and multi-branch structure based generative adversarial networks for fundus disease detection from scanning laser ophthalmoscopy images | Custom GAN ResNet-34 | SLO-1: 2228 UWFI SLO-2: 2362 UWFI | 80/0/20 | AUROC: SLO-1: 0.8215 SLO-2: 0.9736 | SLO-1: 0.7915 SLO-2: 0.9701 | SLO-2: 0.8475 |            |          |          |
| Yoo, T et al | Deep learning can generate traditional retinal fundus photographs using ultra-widefield images via generative adversarial networks | CycleGAN | 451 UWFI 745 FI | 90/0/10 |                        |             |            |            |          |          |
| Author        | Title                                                                 | Architecture(s) | Total Dataset Size | Train/Validate/ Test | AUROC or AUPRC | Sensitivity | Specificity | Precision | Accuracy |
|--------------|------------------------------------------------------------------------|-----------------|--------------------|----------------------|-----------------|-------------|-------------|-----------|----------|
| Zhang, C et al | Development of a deep-learning system for detection of lattice degeneration, retinal breaks, and retinal detachment in tessellated eyes using ultra-wide-field fundus images: a pilot study | seResNext50     | 911 UWFI           | 60/20/20             | AUROC: Original Resizing Method: LD: 0.7710 RB: 0.6250 RD: 0.8750 | Original Resizing Method: LD: 0.7970 RB: 0.9530 RD: 0.8120 | Original Resizing Method: LD: 0.7710 RB: 0.9530 RD: 0.8120 | Original Resizing Method: LD: 0.7970 RB: 0.9530 RD: 0.8120 | Original Resizing Method: LD: 0.7710 RB: 0.9530 RD: 0.8120 |

AUROC: area under receiver operating curve, AUPRC: area under precision recall curve, CAV: convolution along vessel, CMAAI: Chinese Medical Alliance for Artificial Intelligence, CNN: convolutional neural network, DNN: deep neural network, FI: fundus image, LD: lattice degeneration, OAG: open-angle glaucoma, OD: Optic Disc, RB: retinal breaks, RD: retinal detachment, RP: Retinitis Pigmentosa, RVO: Retinal Vein Occlusion, SLO: Scanning Laser Ophthalmoscopy, SVM: support vector machine, TOPP: Tsukazaki Optos Public Project, UWF-FA: ultra-widefield fluorescein angiography, UWF-FAF: ultrawide field fundus autofluorescence, XOH: Xiaohong Ophthalmic Hospital, ZOC: Zhongshan Ophthalmic Center., UWFI: ultrawide field image(s), UWF-ICGA: UWF indocyanine green angiography.
| Author          | Title                                                                 | Ground Truth                                                                 | Explicit Patient-Level Split | Description of Grader Level | Consensus Process                                                                 | External Validation                                                                 |
|-----------------|----------------------------------------------------------------------|-------------------------------------------------------------------------------|------------------------------|---------------------------|-----------------------------------------------------------------------------------|--------------------------------------------------------------------------------------|
| Antaki et al    | Accuracy of automated machine learning in classifying retinal pathologies from ultra-widefield pseudocolour fundus images | Publicly available image datasets reviewed for low-quality and mislabeled images by two ophthalmologists | Yes                          | Two ophthalmologist reviewed publicly available datasets                           | No information available for publicly available datasets used                        | External dataset compiled from 40 publicly available UWF images                        |
| Bawany, M et al | Automated vessel density detection in fluorescein angiography images correlates with vision in proliferative diabetic retinopathy | Two corresponding UWF-FA for each UWFI served as "ground truth" for vessel location | N/A                          | N/A                        | N/A                                                                               | No                                                                                   |
| Cai, S et al    | Deep Learning Detection of Sea Fan Neovascularization From Ultra-Widefield Color Fundus Photographs of Patients With Sickle Cell Hemoglobinopathy | UWFI with corresponding UWF-FA when available for grading by graders          | Yes                          | Two retinal specialists                                                      | Independently graded by a third masked retinal specialist                          | No                                                                                   |
| Calderon-Auza, G et al | A Teleophthalmology Support System Based on the Visibility of Retinal Elements Using the CNNs | DRIVE dataset for vessel segmentation                                            | No                           | Two expert ophthalmologists and five residents (general practitioners in ophthalmology training) | Not discussed                                                                      | No                                                                                   |
| Ding, L et al  | A Novel Deep Learning Pipeline for Retinal Vessel Detection In Fluorescein Angiography | Ground truth transfer from corresponding fundus autofluorescence images         | No                           | Human-in-the-loop process for detection of vessels. Unclear level of human involved | N/A                                                                               | No                                                                                   |
| Ding, L et al  | Weakly-Supervised Vessel Detection in Ultra-Widefield Fundus Photography Via Iterative Multi-Modal Registration and Learning | Vessel maps manually labelled by a human annotator                              | N/A                          | Two human annotators                                                       | Vessel maps from the first annotator used as ground truth, while evaluating the vessel maps made by the second annotator | No                                                                                   |
| Ju, L et al    | Leveraging Regular Fundus Images for Training UWF Fundus Diagnosis Models via Adversarial Learning and Pseudo-Labeling | Fundus images (source domain), UWF (target domain)                              | N/A                          | Images labeled by three ophthalmologists                                      | Images retained only if > 2/3 ophthalmologists agreed on the label                  | Two public fundus image datasets used as external validation                        |
| Kim, I et al   | Classification of pachychoroid disease on ultrawidefield indocyanine green angiography using auto-machine learning platform | UWF ICGA of eyes with macular disease were classified for pachychoroid by two retinal specialists | No                           | Two retinal specialists                                                    | Not discussed                                                                      | No                                                                                   |
### Table 3 (continued)

| Author | Title | Ground Truth | Explicit Patient-Level Split | Description of Grader Level | Consensus Process | External Validation |
|--------|-------|--------------|------------------------------|-----------------------------|-------------------|---------------------|
| Li, Z et al | Automated detection of retinal exudates and drusen in ultra-widefield fundus images based on deep learning | UWFA graded by two trained image analysts for quality | No | Graded by two trained image analysts | Disagreements resolved by a third independent reader | No |
| Li, H et al | Automated Quality Assessment and Image Selection of Ultra-Widefield Fluorescein Angiography Images through Deep Learning | UWFI dataset classified into retinal exudates (RED) and non-RED by retina specialists | Yes | 3 retina specialists (> 5 years experience) and 1 senior retina specialist (> 20 years experience) | Consensus annotation by 3 retina specialists. Disputed images adjudicated by senior retina specialist | ZOC (1222 UWFI) and XOH (2500 UWFI) external datasets |
| Li, Z et al | Deep learning for automated glaucomatous optic neuropathy detection from ultra-widefield fundus images | UWFI classified to glaucomatous optic neuropathy (GON) and non-GON by 3 glaucoma specialists | Yes | 3 glaucoma specialists (> 5 years experience) and 1 senior glaucoma specialist (> 20 years experience) | Consensus annotation by 3 glaucoma specialists. Disputed images adjudicated by senior glaucoma specialist | ZOC (1288 UWFI), XOH (2660), TOPP (12,968 UWFI) |
| Li, Z et al | A deep learning system for identifying lattice degeneration and retinal breaks using ultra-widefield fundus images | UWFI classified for features of NPRLs (lattice degeneration and/or retinal breaks) per Preferred Practice Pattern guidelines from the American Academy of Ophthalmology Retina/Vitreous Panel | Yes | Three retinal specialists with > 5 years of experience. Disagreements adjudicated by a retinal specialist with > 20 years of experience | The reference standard was determined when an agreement was achieved among all 3 ophthalmologists or adjudicated by another retinal specialist if disagreements existed | No |
| Li, Z et al | Deep learning for detecting retinal detachment and discerning macular status using ultra-widefield fundus images | UWFIs classified for retinal detachment by three retinal specialists | Yes | Three retinal specialists. Disagreements adjudicated by a retinal specialist with > 20 years of experience | The reference standard was determined when an agreement was achieved among all 3 ophthalmologists or adjudicated by another retinal specialist if disagreements existed | No |
| Li, Z et al | Development and Evaluation of a Deep Learning System for Screening Retinal Hemorrhage Based on Ultra-Widefield Fundus Images | UWFI categorized to retinal hemorrhage (RH) and non-RH by three retinal specialists | Yes | Three retinal specialists. Disagreements adjudicated by a retinal specialist with > 20 years of experience | The reference standard was determined when an agreement was achieved among all 3 ophthalmologists, or adjudicated by another retinal specialist if disagreements existed | CMAAI (16,827 UWFI), ZOC (905), XOH (1236) |
| Table 3 (continued) | Author | Title | Ground Truth | Explicit Patient-Level Split | Description of Grader Level Consensus Process | External Validation |
|---------------------|--------|--------------------------------|---------------|-------------------------------|---------------------------------------------|---------------------|
| Li Z et al | Deep learning from ‘passive feeding’ to ‘selective eating’ of real-world data | Patient-Level Split | Yes | None | UWFA graded by three retinal specialists for quality control | None |
| | | Overview: UWFA | Yes | N/A | No | Disagreements adjudicated by a retinal specialist with >20 years of experience |
| | | | | | The reference standard was determined when an agreement was achieved among all 3 ophthalmologists, or adjudicated by another retinal specialist if disagreements existed |
| Masumoto, H et al | Deep-learning Classifier With an Ultrawide-field Scanning Laser Ophthalmoscope Detects Glaucoma Visual Field Severity: | UWFI of patients open angle glaucoma and patients with angle closure glaucoma | Yes | N/A | UWFI categorized for RP based on corresponding clinical history, fluorescein angiography and full-field electroretinograms | None |
| | | | | | Accuracy of ultra-wide-field fundus ophthalmoscopy in detection of retinal changes and fluorescein angiography | None |
| | | | | | No | Not discussed |
| Masuda, S et al | Accuracy of ultra-wide-field fundus ophthalmoscopy for detecting age-related macular degeneration | UWFI of males >70 years and females >77 years, with wet AMD confirmed using the results of standard fundus examination, OCT imaging, and fluorescein angiography. | No | Two independent retinal specialists | N/A |
| | | | | | Accuracy of ultra-wide-field fundus ophthalmoscopy in detection of early AMD | None |
| Nagasato, D et al | Deep-learning classifier with ultrawide-field fundus ophthalmoscopy for detecting branch retinal vein occlusion | UWFI reviewed for BRVO by a single retinal specialist | Yes | Images reviewed by a retinal specialist for the presence of acute BRVO | None |
| | | | | | Accuracy of ultra-wide-field fundus ophthalmoscopy in detection of acute BRVO | None |
| Nagasato, D et al | Deep Neural Network-Based Method for Detecting Central Retinal Vein Occlusion Using Ultrawide-Field Fundus Ophthalmoscopy | UWFI reviewed for CRVO by a single retinal specialist | Yes | Images reviewed by a retinal specialist for the presence of acute CRVO | None |
| Author         | Title                                                                 | Ground Truth                                                                 | Explicit Patient-Level Split | Description of Grader Level Consensus Process | Consensus Process | External Validation |
|----------------|----------------------------------------------------------------------|------------------------------------------------------------------------------|------------------------------|-----------------------------------------------|-------------------|---------------------|
| Nagasawa et al | Accuracy of Diabetic Retinopathy Staging with a Deep Convolutional Neural Network Using Ultra-Wide-Field Fundus Ophthalmoscopy and Optical Coherence Tomography Angiography | UWFI and OCTA images that passed an image-quality review were classified by three retinal specialists | No                           | Images were reviewed by three retinal specialists | Not discussed     | None                |
| Nagasawa, T et al | Accuracy of deep learning, a machine learning technology, using ultra-wide-field fundus ophthalmoscopy for detecting idiopathic macular holes | UWFI classified by presence of macular hole. Cases were confirmed using fundus examinations via ophthalmoscope and swept-source OCT images | No                           | A single retinal specialist selected cases of macular hole | Not discussed     | DNN performance evaluated by human graders |
| Nagasawa, T et al | Accuracy of ultrawide-field fundus ophthalmoscopy-assisted deep learning for detecting treatment-naive proliferative diabetic retinopathy | UWFI classified for PDR or without PDR using the ETDRS scale by three retinal specialists | No                           | Images were reviewed by three retinal specialists | Not discussed     | None                |
| Nunez do Rio, J et al | Deep Learning-Based Segmentation and Quantification of Retinal Capillary Non-Perfusion on Ultra-Wide-Field Retinal Fluorescein Angiography | UWF and UWF-FA selected by two investigators to contain retinas affected by capillary non-perfusion and diabetic retinopathy | N/A                          | Images selected by two authors. One expert grader demarcated the outer boundary of the gradable retina | N/A               | None                |
| Oh, K et al    | Early detection of diabetic retinopathy based on deep learning and ultra-wide-field fundus images | UWFI images graded for presence and severity of diabetic retinopathy based on the ETDRS protocols by two independent graders | No                           | Ophthalmologist with > 10 years experience and certified grader with 2 years of experience graded the dataset independently | Not discussed     | REFUGE dataset used for training optic disk and macula detection |
| Ohsugi, H et al | Accuracy of deep learning, a machine-learning technology, using ultrawide-field fundus ophthalmoscopy for detecting rhegmatogenous retinal detachment | UWFI reviewed for RRD by two ophthalmologists | No                           | Reviewed by two ophthalmologists               | Not discussed     | None                |
| Author      | Title                                                                 | Ground Truth | Explicit Patient-Level Split | Description of Grader Level | Consensus Process | External Validation          |
|-------------|------------------------------------------------------------------------|--------------|------------------------------|-----------------------------|-------------------|-----------------------------|
| Sevgi, D et al | Deep learning-enabled ultra-widefield retinal vessel segmentation with an automated quality-optimized angiographic phase selection tool | UWFA sequences were used to create cubic splines. Not discussed how disease categorization was completed or by whom | No                           | Not discussed        | N/A                      | None                        |
| Shi et al   | A method for the automatic detection of myopia in Optos fundus images based on deep learning | UWFI with corresponding spherical equivalent (SE) measurements | No                           | N/A                        | N/A                      | None                        |
| Tang, F et al | Detection of Diabetic Retinopathy from Ultra-Wide Field Scanning Laser Ophthalmoscope Images: A Multi-Center Deep-Learning Analysis | UWFI of diabetic patients, with grading of DR severity and DME by retina specialists from dilated biomicroscopic fundus examination according to the International Clinical Diabetic Retinopathy Disease Severity Scale | No                           | Primary dataset: "well-trained graders" who are listed authors External-2, External-3, External-4: external graders | Discrepancies were resolved by consensus. If consensus could not be reached, the image would be excluded | External-1: Moorfield’s Biomedical Research Centre (London, UK) External-2: Sankara Nethralaya Hospital (Chennai, India) External-3: Giridhar Eye Institute (Kochi, India) External-4: Diagnostic Image Center (Buenos Aires, Argentina) |
| Wang, K et al | Automated detection of diabetic retinopathy lesions on ultrawidefield pseudocolor images | UWFI of diabetic patients categorized into no referral, referral-warranted DR, and ungradable categories | No                           | Two expert, certified reading centre DR graders (Doheny Eye Institute) | The grading was performed by two independent, masked, certified DR graders (K.W. and S.B.V.), using the ICDR scale level. Discrepancies were adjudicated by a third masked senior DIRC grading specialist (M.G.N.) to yield a single final severity grade for each case | None                        |
| Wang, Z et al | Multi-Task Siamese Network for Retinal Artery/ Vein Separation via Deep Convolution Along Vessel | Publicly available annotated datasets with vessels segmented and annotated (DRIVE, INSPIRE, WIDE) | N/A                          | DRIVE: three different human graders INSPIRE/WIDE: image analysis expert and ophthalmologist | DRIVE: consensus INSPIRE and WIDE: Image analysis expert manually classified and annotated vessels, and then corrected by an ophthalmologist | None                        |
| Author         | Title                                                                 | Ground Truth                                                                                       | Explicit Patient-Level Split | Description of Grader Level | Consensus Process | External Validation |
|---------------|----------------------------------------------------------------------|---------------------------------------------------------------------------------------------------|------------------------------|-----------------------------|-----------------|--------------------|
| Wisely, C et al | Convolutional neural network to identify symptomatic Alzheimer’s disease using multimodal retinal imaging | Data extracted from patients with AD enrolled in the Duke Memory Disorders clinic. Patients were evaluated and clinically diagnosed by experienced neurologists and medical records reviewed by one of two expert neurologists to confirm AD diagnosis | N/A                          | Experienced neurologists    | Not discussed     | None               |
| Xie, H et al   | AMD-GAN: Attention encoder and multi-branch structure based generative adversarial networks for fundus disease detection from scanning laser ophthalmoscopy images | Unannotated SLO-1 and SLO-2 datasets                                                               | N/A                          | N/A                         | Not discussed     | None               |
| Yoo, T et al   | Deep learning can generate traditional retinal fundus photographs using ultra-widefield images via generative adversarial networks | Unpaired UWFI and FI images, manually reviewed for quality by two board-certified ophthalmologists | No                           | Two board-certified ophthalmologists | Not discussed     | None               |
| Zhang, C et al | Development of a deep-learning system for detection of lattice degeneration, retinal breaks, and retinal detachment in tessellated eyes using ultra-wide-field fundus images: a pilot study | UWFI reviewed by 2 retinal specialists independently to confirm the presence of tessellated fundus | No                           | Three retinal specialists   | Three retinal specialists evaluated all images and proposed the reference standard when an agreement was achieved | None               |

AD: Alzheimer’s Disease, CMAAI: Chinese Medical Alliance for Artificial Intelligence, UWF-FA: ultra-widefield fluorescein angiography, UWF-FAF: ultrawide field fundus autofluorescence, XOH: Xudong Ophthalmic Hospital, ZOC: Zhongshan Ophthalmic Center., UWF: ultrawide field image(s), UWF-ICGA: UWF indocyanine green angiography
Fig. 3  Number of included studies by year of publication

Fig. 4  The countries of origin of all included studies. Each country with an included study is highlighted in blue, with the number of studies from each written in red.
Using UWF-FA from PDR patients, Bawany et al. utilized DL to correlate automated vessel density with visual acuity (VA) in 2020 [52]. While not focusing on detection of DR generally, the goals of the study were to use a DL-quantified measure (retinal vessel density) and determine if it correlated with an outcome (VA) known to be affected by PDR. From a dataset of 42 UWFIs from patients with PDR without significant center-involving DME, retinal blood vessels were first detected using a deep neural network (DNN) in a U-Net architecture. The study authors trained the dataset on UWF-FA images with corresponding ground-truth vessel maps trained using a HITL procedure first demonstrated by Ding et al. [58]. For each UWFI, two UWF-FA images effectively provided a “ground truth” of vessel location to evaluate the segmentation. The output of the DNN was a vessel map where pixel intensity indicated the likelihood of a pixel being a vessel. The trained DNN achieved 0.930 AUPRC. Vessel density was measured by calculating the percentage of vessel pixels in a circular area centered around the fovea. To study the correlation between vessel density and best corrected visual acuity (BCVA), UWF-FA were analyzed using the trained model. The study found a statistically significant positive correlation between vessel density and BCVA of 0.4071 (p = 0.0075), but no statistically significant correlation between vessel density and central retinal thickness.

Tang et al. published a study in 2021 which used DL to detect vision-threatening DR (VTDR) and referable DR (RDR) from UWFIs [54]. In this study, 2861 UWFIs from the primary dataset were labeled for the presence or absence of RDR and VTDR respectively by graders according to the International Clinical Diabetic Retinopathy Disease Severity Scale. A total of 9392 UWFIs were used for training, primary validation, and geographical external validation across the primary dataset and the four external datasets. The authors then trained three CNNs to develop a pipeline for disease detection from the UWFIs. The first CNN classified images as gradable or ungradable, the second for detecting VTDR, and the third for detecting RDR. The study used transfer learning and applied ResNet50 models pre-trained on ImageNet. Finally, the authors applied Class Activation Mapping heatmaps for each result (true positive, true negative, false positive, and false negative) to assess DL performance. The first CNN to determine gradeability achieved an 86.5% sensitivity, 82.1% specificity, and 0.923 AUROC on their primary dataset. The RDR detection CNN achieved 0.981 AUROC, 94.9% sensitivity, and 95.1% specificity. The VTDR CNN achieved 0.966 AUROC, 87.2% sensitivity, and 95.8% specificity. On four external datasets, the gradability CNN achieved >0.82 AUROCs, >79.6% sensitivity, and >70.4% specificity while the RDR and VTDR CNNs achieved AUROCs and accuracies of >0.9 and >80% respectively.

Oh et al. published a study investigating the early detection of DR using DL and UWFIs [53]. They compared the ability of CNNs to classify ETDRS 7SF vs. optic disk and macula-centered ETDRS F1-F2 images as containing DR. They first trained a U-Net model with ResNet-18 for optic disk detection on the publicly available REFUGE dataset of color FI [59]. The authors then used size and distance thresholds to determine macula locations. An ophthalmologist with over ten years of experience and a certified grader with two years of experience categorized a dataset of 13,271 UWFIs as healthy or containing DR. They then inputted these UWFIs into their trained model to detect the optic disk and macula center. From these detected locations, they segmented the UWFIs into ETDRS 7SF images and F1-F2 images. The 7SF ETDRS images contain 7 fields of 30 degrees each, while F1-F2 images contain only 30-degree overlapping circles centered on the optic disk and macula center. The authors then trained a ResNet-34 model pre-trained on ImageNet and optimized their model using their dataset. In doing so, they achieved an 0.915 AUROC, 83.38% sensitivity, and 83.41% specificity on 7SF images. However, they achieved a 0.8867 AUROC, 80.60% sensitivity, and 80.61% specificity on F1-F2 images. The 7SF images achieved results that were significantly greater for all three measures (p < 0.001) compared to those of F1-F2 images. While the authors demonstrate that DL classification systems are more accurate using 7SF images, the achieved AUROC, sensitivity, and specificities have been greater in previously published studies using whole UWFIs. This indicates the greater utility of UWFIs over 7SF and F1-F2 images of the fundus.

Nagasawa et al. published a second study on DR using CNNs and UWFIs in April 2021 [57]. They compared the accuracy of DL-based DR staging from UWFIs and OCT angiography (OCTA) images. UWFIs and OCT en face images of the superficial plexus, deep plexus, outer retina, choriocapillaris, and density map were extracted for 491 patients with diabetes. OCTA scans of a 6 × 6 mm region were acquired for each patient. The OCTA and UWFIs were combined into a single image file, to form a third “imaging modality.” The dataset used contained images stratified by three retinal specialists into categories of into categories of no apparent DR, mild NPDR, moderate NPDR, and severe NPDR using the ETDRS scale. The study authors then trained a VGG-16 CNN to first classify the images as containing DR, and the second to detect PDR. Each CNN was tested on UWF, OCTA, and UWF-OCTA combined datasets. In detecting DR, the first CNN achieved AUCs after training on the UWF, OCTA, and UWF-OCTA images of 0.790, 0.883, and 0.847 respectively. In detecting PDR, the second CNN achieved AUCs after training on the UWF, OCTA, and UWF-OCTA images.
of 0.981, 0.928, and 0.964 respectively. This study demonstrates the ability of DL systems to detect DR and PDR but also demonstrates no additive benefit of combining imaging modalities (UWFIs and OCTA images) to increase the accuracy of disease classification.

**Retinal detachment**

Five studies have been published on RD detection from UWFIs [39, 46, 60–62]. The first published, from Ohnuki et al., used 831 UWFIs to detect rhegmatogenous RD (RRD) [60]. The dataset contained 411 images from RRD patients and 420 images from non-RRD patients, which were then reviewed and classified by two ophthalmologists. The study used a CNN with 3 convolutional layers, of which each were followed by activation function (ReLU) layers and finished with two fully connected layers. The final output layer performed a binary classification using a softmax function. The trained model achieved an 0.988 AUROC, 97.6% sensitivity, and 96.5% specificity.

In 2019, Li et al. developed a DL system for identifying specific characteristics of RD from UWFIs [61]. They developed a DL system for detecting notable peripheral retinal lesions (NPRLs), such as lattice degeneration and retinal breaks, which can lead to RRD. Three retina specialists each with over 5 years of experience classified 5606 UWFIs with disagreements adjudicated by a retinal specialist with over 20 years of experience. They then compared the performance of 4 CNNs: InceptionResNetV2, InceptionV3, ResNet50, and VGG-16. With each CNN, the authors explored three methods for improving the DL algorithm: i) no data augmentation, ii) data augmentation with brightness shifts, 45-degree rotation, and horizontal and vertical flipping, and iii) data augmentation with histogram brightness equalizations, 45-degree rotation, horizontal flipping, and vertical flipping. This led to 12 models trained and compared. The study found that the dataset trained on InceptionResNetV2 with the second data augmentation method achieved the highest performance, with 98.7% sensitivity, and 99.2% specificity, and 99.1% total accuracy. This was significantly greater than comparisons with ophthalmologists in the study. The authors found that a general ophthalmologist with 5 years of experience had a 97.6% accuracy, 93.6% sensitivity, and a 98.7% specificity, while one with 3 years’ experience had a 94.5% accuracy, 85.9% sensitivity, and 96.8% specificity. These results are very promising for the continued use of DL in identifying NPRLs and its greater accuracy in comparison to trained ophthalmologists.

In 2020, Li et al. then applied an InceptionResNetV2-based DL model to detect RD and discern macular status using 11,087 UWFIs labelled for RD by three retinal specialists [46]. They first developed a DL system to detect RD. The model for detecting RD achieved a 96.1% sensitivity, 99.6% specificity, and 0.989 AUROC. A retina specialist with 3 years’ experience achieved 94.4% sensitivity and 99.1% specificity, while a specialist with 5 years’ experience achieved 95.4% sensitivity and 99.8% specificity. The RD images were then used as the dataset for the DL for macular status classification. This DL model achieved 93.8% sensitivity, 90.9% specificity, and 0.975 AUROC. The ophthalmologist with 3 years of training achieved sensitivities and specificities of 86.3% and 87.1%, while the more senior ophthalmologist achieved 91.3% and 92.4% respectively. The difference in discerning macular status between the DL and ophthalmologists is greater than their difference in RD detection. As macular status is an indication for emergency surgery, this difference is significant in demonstrating the utility and necessity of DL in ophthalmology [46].

Zhang et al. developed a DL system for detecting lattice degeneration, retinal breaks, and RD in tessellated eyes [62]. They then tested two image pre-processing techniques with the seResNext50 CNN on 911 UWFIs classified by three retinal specialists for disease features from tessellated eyes. The first technique resized all images to 512×512, and when applied to the DL model, the model would output a positive number for each lesion per image. The second method used the cropping of patches of labelled lesions. The DL model applied to this dataset would then assign a positive score to each lesion and output the max score of all the image’s patches. Furthermore, they trained three distinct models for detecting lattice degeneration, retinal breaks, and RD respectively for a total of 6 tested models. In detecting lattice degeneration, the resizing method achieved 0.888 AUROC while the cropping method achieved 0.841. For retinal breaks, the resizing method had 0.843 AUROC while the cropping method achieved 0.953 AUROC. In RD, the resizing and cropping methods achieved AUROCs of 1.000 and 0.979 respectively. The use of the full image led to greater accuracy in all cases except for retinal breaks, where the cropping method was found to be superior.

In 2021, Antaki et al. published a study exploring the use of AutoML technologies for classifying RD, RP, and RVO from UWFIs [39]. They trained a DL model through the Google Cloud AutoML platform using RD and normal UWFIs. The datasets used were publicly available and validated datasets, which were then reviewed by two ophthalmologists. The binary classification of RD achieved an 89.77% sensitivity and 78.72% specificity when the confidence level of the system was set to 0.8. This model also achieved an AUPRC of 0.921.

**Glaucoma**

Glaucoma, whose pathophysiology is related primarily to the optic disk and its degeneration, has been studied using DL with UWF imaging of the retina.
Two studies have investigated glaucoma with UWFIs [63, 64]. In 2018, Masumoto et al. applied a DL classifier to UWFIs to detect glaucoma in a patient dataset stratified by disease severity. The study authors first categorized glaucoma patients into early (-6 dB), moderate (-6 to -12 dB), and severe (-12 dB or worse) based on visual field damage from Humphrey Field Analyzer measurements. The ground truth dataset therefore did not require any human annotation or grading, as the UWFIs were categorized from a quantifiable visual field deficit. In classifying any glaucoma, the DL model achieved a mean of 0.872 AUROC. For early, moderate, and severe glaucoma, the DL model achieved AUROCs of 0.830, 0.864, and 0.934 respectively. The DL model was similarly most sensitive and most specific in classifying severe glaucoma vs. healthy UWFIs. While the results are promising, the AUROC does not reach the 0.9 threshold, which was an acknowledged weakness by the study authors.

Li et al. used DL for automated glaucomatous optic neuropathy (GON) detection using UWFIs in 2020 [63]. They trained a CNN based on the InceptionResNetV2 neural network. All 22,972 UWFIs were classified as containing GON or not by three glaucoma specialists, based on a vertical cup to disc ratio ≥ 0.7, rim width ≤ 0.1 of disc diameter, retinal nerve fiber layer defects, or disc splinter hemorrhages. The primary dataset achieved an AUROC, sensitivity, and specificity of 0.999, 97.5%, and 98.4% respectively. The range of AUROC, sensitivity, and specificity achieved were 0.983–0.999, 97.5–98.2%, and 94.3–98.4% across the primary and four external datasets. The methods demonstrated by Li et al. achieved significantly greater outcomes in detecting and classifying glaucoma on UWFIs than Masumoto et al. likely due to the increased primary dataset size [64].

While outside the scope of this review, other studies have used posterior-segment OCT imaging with DL for glaucoma detection [65]. It is of particular note that methods using 3D segmentation-free OCT volumetric data achieved an AUROC value of 0.940, which exceeds the AUROC of Masumoto et al., but not Li et al.'s use of UWFIs for DL-based classification of glaucoma [66]. This is of note, as it indicates some increased value of using UWFIs in the DL-based detection of glaucoma compared to using OCT alone.

**Age-related macular degeneration**

At the time of writing, two studies related to UWFIs and using DL to diagnose or detect AMD or its complications have been published [67, 68]. Matsuba et al. published a study in 2018 using DL to detect AMD on UWFIs [67]. In this study, they trained a CNN on UWFIs of healthy (no visible fundus disease) and patients with exudative AMD (wet AMD). Ground truth diagnosis was ascertained by two retinal specialists using a combination of standard fundus examination, OCT imaging and FA. The CNN achieved a 0.976 average AUROC, with 100% average sensitivity, and 97.31% sensitivity in detecting wet AMD. Six ophthalmologists yielded a correct classification 81.9% of the time, with 71.4% and 92.5% sensitivity and specificity respectively. The study ophthalmologists averaged 11 min and 23.54 s for classification, while the DL model averaged 26.29 s.

The second study, published in 2021, comes from Li et al. who used DL for the automated detection of retinal exudates and drusen from UWFIs [68]. Images were labelled as containing retinal exudates and/or drusen (RED) or non-RED by three retina specialists, with disagreements adjudicated by a retina specialist with over twenty years of experience. Two external datasets were then used for validation of the InceptionResNetV2 CNN model. On the primary dataset, 0.994 AUROC was achieved, with 94.2% sensitivity and 97.4% specificity. The external datasets achieved 0.972 and 0.988 AUROCs, with 94.9% and 95.1% sensitivities, and 96.5% and 97.3% specificities respectively.

**Retinitis pigmentosa**

Masumoto et al. trained a CNN on UWFIs of RP in 2018 [69]. Using UWF and UWF fundus autofluorescence (UWF-FAF) images, they trained a CNN (VGG-16) to classify images based on whether they contained RP. UWFIs and UWF-FAFs from RP and healthy patients respectively were used in their dataset. RP was diagnosed based on corresponding data from clinical history, UWF-FAF, and electroretinograms (ERGs) according to the International Society for Clinical Electrophysiology of Vision standards. The UWF CNN achieved 0.998 AUROC while that of the FAF achieved 1.000 AUROC. The UWF CNN achieved 99.3% and 99.1% sensitivity and specificity scores respectively, while the UWF-FAF CNN achieved 100% and 99.5% sensitivity and specificity scores. There were no statistically significant differences between the sensitivities and specificities of the UWF and the UWF-FAF CNNs.

Antaki et al. published a study exploring the use of AutoML technologies for classifying RD, RP, and RVO from UWFIs, as previously discussed in the section on RD [39]. They trained a DL model through the Google Cloud AutoML platform using RP and normal UWFIs. The binary classification of RP achieved an 88.0% sensitivity and 100% specificity when the confidence level of the system was set to 0.5. This model also achieved 0.942 AUPRC. When repeated using the data from Masumoto et al. the system achieved an AUPRC of 1, with sensitivity, specificity, and PPV all increased to 100% with no misclassifications made by the AutoML model [39, 69].
Pachychoroid

A single peer-reviewed study on using UWFIs in detecting pachychoroid disease has been published. Kim et al. used an AutoML platform to classify UWFIs based on their presence of pachychoroid disease [70]. Specifically, the authors trained the Google AutoML Vision on UWF indocyanine green angiography (UWF-ICGA) classified into categories of healthy and pachychoroid patients by two retinal specialists. Pachychoroid and non-pachychoroid UWF-ICGA images were uploaded. They trained two models, the first of which used all images in their original orientation and the second of which horizontally flipped left eye images such that all images were of the same laterality. The first model achieved precision, accuracy, sensitivity, and specificity values of 0.8182, 0.8367, 0.8182, and 0.8519 respectively, while the second model achieved 0.8636, 0.8776, 0.8636, and 0.8889 respectively. However, the mean precision, accuracy, sensitivity, and specificity scores of three retina specialists were 0.9048, 0.9388, 0.9500, and 0.9643. These results indicate that training the AutoML model with images of the same laterality led to better results, but that the current training did not reach the levels of precision or recall of retina specialists.

Retinal vein occlusion

Three peer-reviewed studies exist on using DL on UWFIs in RVO, two of which were published by Nagasato et al., and the third from Antaki et al. [39, 71, 72]. The first, published in 2018, uses UWFIs to classify and detect central RVO (CRVO). The study used UWFIs from CRVO and non-CRVO healthy subjects, which was classified by a single retina specialist. A VGG-16-based DNN was trained on the dataset, along with fine-tuning using parameters borrowed from ImageNet. After comparing 40 DL models obtained in 40 learning cycles, they used the DL model with the highest rate of correct answers for evaluation. The model achieved 0.989 AUROC, 98.4% sensitivity, and 97.9% specificity. They similarly used a support vector machine (SVM) algorithm to detect CRVO from UWFIs. The SVM achieved 0.895 AUROC, 84.0% sensitivity, and 87.5% specificity. The DL model achieved significantly greater results in all measures compared to the SVM (p < 0.001).

In 2019, the same group completed a similar study using a DL model on UWFIs of branch RVO (BRVO) patients. In this study, they used the same model (VGG-16), and DNN parameters on a BRVO dataset. Specifically, they trained the DNN on BRVO and non-BRVO healthy UWFIs classified by a single retinal specialist. They similarly tested an SVM model. In this study, the DNN achieved 0.976 AUROC, 94.0% sensitivity, and 97.0% specificity. The SVM model achieved 0.857 AUROC, 80.5% sensitivity, and 84.3% specificity. The authors demonstrated the ability of a DNN to accurately detect BRVO and the superiority of a DNN over SVMs in detecting BRVO.

In 2021, Antaki et al. published a study exploring the use of the Google Cloud AutoML platform for classifying RD, RP, and RVO from UWFIs [39]. The binary classification of RVO achieved a 84.9% sensitivity and 100% specificity when the confidence level of the system was set to 0.5. This model also achieved 0.967 AUPRC. While the sensitivity was lower than that of Nagasato et al., their model achieved comparable specificities [71, 72].

Myopia

In 2020, Shi et al. published a study where they studied the ability of a DL system to detect myopia using UWFIs [73]. For this task, they used a custom CNN, known as the Myopia Detection network (MDNet). This network combined dense connection and Residual Squeeze-and-Excitation attention for detecting myopia. The CNN combined attention dense blocks, transition blocks, convolutional layers, max-pooling layers, and a dense layer to make full use of shallow features and improve information flow.

They trained the CNN on left and right UWFIs. The study defined severe myopia as having a spherical equivalent (SE) more negative than -6 diopters (D), moderate myopia as between -6D and -3D, and mild myopia as SE between -3D and 0D. As myopia could be measured quantitatively, the ground truth dataset was categorized based on quantifiable spherical equivalent values. Images were then cropped for a region of interest around the optic disk of 400 × 400 pixels, centered on the optic disk and including the macula.

In 9, the study authors used mean absolute error (MAE) as the main evaluation index, as well as root-mean-square error (RMSE) and mean-absolute-percent error (MAPE). The CNN achieved optimal results at an MAE of 1.1150 D and RMSE and MAPE of 1.4520 D and 24.99% respectively. These results show that myopia is effectively detected within reasonable error using DL and UWFIs.

Idiopathic macular hole

A single peer-reviewed study on detecting idiopathic macular hole (IMH) using UWFIs and DL has been published. In 2018, Nagasawa et al. trained a CNN on normal and IMH images [45]. Ground truth diagnosis was ascertained by a single retinal specialist who conducted ophthalmoscopy and reviewed OCT imaging. The CNN achieved an 0.999 AUROC, 100% accuracy, 100% sensitivity, and 99.5% specificity. The CNN was able to classify images at an average speed of 32.80 ± 7.36 s for a series of 50 test images. They similarly tested the ability of human ophthalmologists to detect IMH from the same UWF test images. The
ophthalmologists were able to achieve an $80.6 \pm 5.9\%$ accuracy, $69.5 \pm 15.7\%$ sensitivity, and $95.2 \pm 4.3\%$ specificity, and required an average time of $838 \pm 199.16$ s to classify the same 50 images. From this study, it is clear that IMH is more accurately and rapidly diagnosed using CNNs than trained ophthalmologists.

**Retinal hemorrhage**

Li et al. published a study using a DL system to screen RH from a dataset of RH and non-RH UWFIs [74]. The dataset was categorized by three retinal specialists with disagreements adjudicated by a more experienced retinal specialist. The study used InceptionResNetV2, with weights pre-trained for ImageNet classification for CNN initialization. On the primary dataset, the CNN achieved an 0.999 AUROC, 98.9% sensitivity, 99.4% specificity, and 99.3% accuracy. Two external datasets were used for further testing, which achieved 0.998 and 0.997 AUROCs, 96.7% and 97.6% sensitivities, 98.7% and 98.0% specificities, and 98.4% and 98.0% accuracies respectively. On an external dataset, an ophthalmologist with five years of training achieved a 95.9% sensitivity and a 99.5% specificity, while an ophthalmologist with three years of training achieved 92.6% and 98.9% respectively. Here, the ophthalmologists scored sensitivities lower than the trained CNN, but specificities were close to the specificity of the CNN.

**Sickle cell retinopathy**

A single study, published in 2020, explores using DL with UWFIs to diagnose SCR. Specifically, the study from Cai et al. explored the detection of sea fan neovascularization (SFN) from UWFIs of patients with sickle cell hemoglobinopathy [75]. The dataset was categorized by two retinal specialists, using corresponding UWF-FA data when available. The study notes that the detection of potentially asymptomatic SFN provides the opportunity for prophylactic scatter laser photocoagulation, which can help to reduce the rates of proliferative SCR vision loss. An InceptionV4 CNN, pre-trained on the ImageNet dataset, was trained on the image set for 100 iterations. After training, the CNN achieved an 0.988 AUROC, 97.0% accuracy, 97.4% sensitivity, and 97.4% specificity. Only a single image received a false-negative classification from the CNN, due to a severe lid artifact obscuring the retinal vasculature.

**Quality assessment**

Three studies have been published on using DL methods for quality assessment of UWFIs [76–78]. The first, published in 2020 by Calderon-Auza et al., focuses on using CNNs as a teleophthalmology support system to determine the quality of images provided. Low-quality images were defined as containing factors such as an inability to distinguish the optic disc, as well as artifacts from the eyelashes and/or eyelids obstructing the image. Specifically, the system they proposed uses four steps to determine UWFIs quality. First, the system detects the optic disc (OD), performs quality analysis on the OD, determines obstruction (i.e., eyelash shadows) detection of the region of interest (ROI), and then segments the vessels of the image. For OD detection, Faster Region Based-CNN (FR-CNN) for feature extraction along with the AlexNet CNN architecture was used. On their dataset, this CNN configuration achieved an accuracy, sensitivity, and specificity of 0.9254, 0.9643, and 0.4424 respectively. Images determined as containing an OD were then used as the dataset for the OD quality analysis step. For this step, VGG-16 was used and achieved a 0.8612 accuracy, 0.9113 sensitivity, and 0.8064 specificity in detecting and classifying ODs by quality. For obstruction analysis in the ROI, centered on the optic disk and macula, a SegNet-trained CNN achieved a 1.0 accuracy due to the low number of artifacts in the training and test sets. Finally, for vessel segmentation in the ROI, a SegNet architecture with VGG-16 proposed by the authors achieved an 0.9784 accuracy, 0.7169 sensitivity, and 0.9816 specificity on their dataset.

While the study from Calderon-Auza et al. is a proof of concept of the uses of DL for detecting low-quality UWFIs, it also provides examples of these techniques in practice. For this reason, this study provides readers with a clear implementation of the uses of DL in UWFIs for a multi-step process in tele-ophthalmology.

Li et al. proposed and designed a classification system using “U-Net style” CNN and UWF-FA in 2020 [77]. UWF-FA were graded as ungradable, poor, good, or best by two image analysts with disagreements adjudicated by a third grader. The CNN achieved 90.5% sensitivity and 87.0% specificity for distinguishing between gradable and ungradable images and a sensitivity and specificity of 78.9% and 94.1% for distinguishing between optimal quality (good, best) and limited quality (poor, ungradable) images. The authors calculated the overall accuracy of the classifier as 89.0% for gradable vs. ungradable classification and 89.3% for recognizing optimal quality versus limited quality. The model also achieved an 0.920 AUROC.

In 2020, Li et al. proposed a DL-based image filtering system (DLIFS) to filter out poor-quality UWFIs in an automated fashion, such that only images of sufficient quality would be used in subsequent AI diagnostic systems [78]. Images were identified as poor-quality or good-quality images by 3 retina specialists with disagreements adjudicated by a more experienced retinal specialist. Image quality was categorized as poor if more than 1/3 of the fundus was obscured, macular vessels could not be identified or > 50% of the macular area was obscured, or if the vessels within
a 1-disc diameter of the OD margin could not be identified. From this dataset, they trained InceptionResNetV2 with weights pre-trained for ImageNet. The CNN would classify the image quality for each inputted UWFIs. The trained DLIFS achieved an 0.996 AUROC, 96.9% sensitivity, and 96.6% specificity. Two external datasets were used for testing, with which the DLIFS achieved 0.994 and 0.997 AUROCs, 95.6% and 96.6% sensitivities, and 97.9% and 98.8% specificities respectively.

### Segmentation and localization

Five peer-reviewed studies have been published on segmentation and localization using UWFIs, all of which focus on vessel segmentation [58, 79–82]. The first, from Ding et al. presented a method to detect retinal vessels in UWF-FA [58]. In this study, the authors developed a method to produce vessel segmentation maps without previously labelled ground-truth datasets. They primarily relied on cross-modality transfer and HITL learning. The HITL approach allowed the DL system to predict the vessels and respond to human feedback regarding whether it had segmented the vessels correctly and accurately. Over multiple iterations, this led to complete segmentation of the vessels. The authors were able to reduce manual annotation effort by first using morphological analysis to segment the vessels in a preliminary fashion. This was followed by a cross-modality approach that transferred vessel maps from UWFIs to UWF-FA images using robust chamfer alignment in an Expectation–Maximization framework. These were combined using the HITL iterative DL process for detection of retinal vessels.

The first step in the pipeline, relying on cross-modality transfer, trained a DNN on a dataset of ground truth color UWFIs with UWF-FA images from the same patient eye taken at the same time. Specifically, the DNN was trained on existing labelled UWFIs to extract the vessel maps from unlabeled UWFIs. These detected vessel maps were then geometrically aligned and transferred to the UWF-FA. These new vessel maps, aligned to UWF-FA, served as the approximate ground truth for training a DNN for vessel detection in UWF-FA images. From this point, the DNN for detecting vessel segmentation was continually run starting from the approximate ground truth from the UWFIs, until the DNN did not produce maps with new changes or more vessels segmented.

The process of producing vessel maps was approached as one that would be best suited for a generative adversarial network (GAN), in producing an output image of vessel segmentation from an input UWF-FA image.

The authors then evaluated their method of reducing the burden of annotation by calculating the number of pixels added and removed at each iteration. After 7 iterations, approximately 19,300 (2.0%) new pixels were added, and 14,100 (1.4%) of pixels were removed. Vessel detection on their primary dataset achieved an AUROC of 0.980 and a Dice coefficient of 0.829. In validating their approach on an external dataset, they achieved a maximal 0.987 AUROC, with significant improvements over traditional morphological techniques for vessel segmentation.

The same team of Ding et al. then published a method to segment vessels from color UWFIs via iterative multi-modal registration and learning [79]. In this project, they similarly utilized concurrently captured UWF-FA images to segment the vessels from UWFIs. The first step requires multi-modal registration of the vessels segmented first from UWF-FA using a pre-trained DNN to the UWFIs, using parametric chamfer alignment. The second step utilized a learning method to mitigate the noisy labels due to the differences in the UWF-FA and UWFIs modalities. The detected UWFIs vessel maps are then used for the registration in the following iteration, allowing for iterative improvement until the segmented vessel maps are accurate. After this training, the DNN can detect vessels from UWFIs without concurrently captured UWF-FA images. On their primary dataset, they achieved an AUROC of 0.987 and maximal Dice coefficient of 0.987. After training their DNN, they evaluated the model on an external dataset of UWFIs, achieving an AUPRC of 0.886.

Nunez do Rio et al. published a study in 2020 that explored the use of DL-based segmentation for quantification of retinal capillary non-perfusion using UWF-FA [80]. Capillary non-perfusion (CNP) is a metric that is useful in determining retinal ischemia. For this reason, they sought to use UWF-FA, which is a high-resolution image with a clearly defined retinal vasculature, to quantify this. For this process, they trained a U-Net-style CNN on 75 UWF-FA that were manually graded for CNP to segment and extract the vasculature of these images. 20 images were also segmented by an expert grader manually. To standardize the CNP measurement, a circular grid of rings of increasing radius was centered on the foveal avascular zone (FAZ). The segmentation model achieved an 0.82 AUROC. Between the manually graded images and the automatically segmented images, an inter-grader dice similarity coefficient (DSC) of 65.51 was achieved. In comparing the assessment of CNP between the CNN model and the grader, a Kappa score of 0.55 was achieved. The authors conclude that this automatic segmentation method allows for a DL-based segmentation of CNP and a quantifiable measurement of CNP from UWF-FA.

Wang, Z et al. published a study in 2020 that utilized a multi-task Siamese network for separating retinal arteries from retinal veins using deep convolution [81]. They did so on an FI dataset (DRIVE), a UWFI dataset (WIDE), and an OCT dataset (INSPIRE) [83–85]. Using these datasets, they first segmented the vessels using a CNN-based approach, followed by skeletonization of the vessels. Next, they built a graph representing the vascular network by...
finding branching and end points on the skeleton map. Next, errors such as twinborn nodes produced by overlapping vessels were removed by morphological analysis through the skeleton. This produced a refined vascular graph. They then used the Convolution Along Vessel method to extract visual features by convolving the image along the vessel segments and the geometric features of the vessels, by tracking the direction of blood flow in the vessels. Following this, the Siamese network was trained to learn to classify vessel types by visual features of vessel segments, and by estimating the similarity of every two connected segments by comparing their visual and geometric features. This was done to separate the vasculature types into individual trees of arteries and veins. On the WIDE dataset of UWFI, they were able to achieve an accuracy value of 94.5%.

In 2021, Sevgi et al. published a study that explored the ability to extract the cumulative retinal vessel areas (RVA) from UWF-FA images using CNN-based DL segmentation [82]. For this study, they extracted the RVA from the available UWF-FA image frames. Images that contained the maximum RVA were considered the optimum early phase, while a frame that was taken ≥ 4 min after that closely mirrored the RVA from the early image was considered the late phase frame. Image analysts then evaluated the selected pairs. There were 1578 UWF-FA sequences from 66 sessions used to create cubic splines and a total of 13,980 pairs. There were 1578 UWF-FA sequences from 66 sessions used for evaluation. 85.2% of the sessions had appropriate images for both phases successfully identified. 90.7% of early and 94.6% of late frames were successfully identified.

**Generative image synthesis using GANs**

At the time of writing, four studies involving generative adversarial networks (GANs) and UWFI have been published [86–89]. GANs, designed in 2014 by Goodfellow et al. are ML frameworks that utilize two competing neural networks to generate new data [90]. The first neural network, named the “generator,” generates random data. The second neural network, the “discriminator,” is trained on data that is to be modelled and produced. As the generator produces data, the discriminator will reject synthesized data from the generator that does not sufficiently represent the training data source. Through an iterative process, the generator becomes more effective at generating data that effectively represents the target data, until synthetic data that appears close to the ground-truth dataset is produced [91]. These approaches have been effective at generating data such as human faces that appear realistic [92].

Ju et al. published another study in 2020, where they utilized GANs to produce labelled datasets of UWFIIs from labelled fundus image datasets [87]. They noted that due to the differences in FIs and UWFIIs, labelled datasets of FIs could not be used for classifying UWFIIs. For this reason, they used a GAN to generate synthetic UWFIIs for training. Using a quantitative, classification-based consistency regularization method, they ensured that the pathologies present in labelled FIs were similarly present in corresponding generated UWFIIs. FIs were first labelled by three ophthalmologists, requiring 2/3 ophthalmologists to agree on the label for the image for it to be included in the dataset. The authors tested the generated images with and without the consistency regularization method and found that it improved the generation of matching features. The first step in this process required using target UWFIIs to train a target-task model, which helps to regulate the quality of generated data. Following this, pseudo-labels were generated for the generated UWFIIs. Finally, they used the original UWFI samples and the generated samples to train the target-task model together. To test that the generated UWFIIs were properly pseudo-labelled and carried the disease pathology of interest, they then classified the images that contained DR using a ResNet50 based residual neural network. They then similarly validated their synthetic UWFIIs by testing them with vessel segmentation and lesion detection tasks. This was then repeated on two public FI datasets as external validation. The study authors effectively succeeded at producing high-quality UWFIIs that mirrored FIs in pathology, and mirrored “natural” UWFIIs in image quality and complexity.

In 2020, Xie et al. published a study where they proposed a GAN which used an attention encoder (AE) and generation flow network to build a UWFI classifier for retinal pathologies found in patients under the age of eighteen (i.e., coats, familial exudative vitreoretinopathy, morning glory syndrome, RP, and DR) [88]. The two datasets used contained a total of 3072 abnormal and 1,518 normal UWFII respectively. The goal of this project was to harness the adversarial learning that occurs between the generator and the discriminator to build robustness into their classification model. Their proposed method achieved higher classification accuracy (84.75% and 97.25%) compared to classifiers based on a standard CNN architecture such as ResNet-50 (77.35% and 87.95%).

In 2020, Yoo et al. used GANs in a way opposite to Ju et al. [87, 89]. In their study, they utilized a GAN architecture to produce synthetic FIs from UWFIIs. Specifically, they used the CycleGAN architecture to translate the UWFIIs to FIs while maintaining the structure, pathology, and lesions specific to the original FIs without generating new or fake features into the FI. The authors began by using a dataset of UWFIIs and FIs, which were reviewed by two ophthalmologists for image quality. The GAN was trained on the dataset of UWFIIs and FIs, and then tested on the test dataset of UWFIIs to generate synthetic FIs. Image registration was applied to crop the region of interest on the input UWFIIs, focused on the optic disk and fovea, for conversion into
an FI. After training the CycleGAN model for 40 epochs, the model was able to successfully transfer the image from UWFIs to FIs with high fidelity to the original UWFI structure and pathologies. For example, UWFIs with DR microaneurysms and blot hemorrhages, GON, RD, CRVO, Drusen, and retinal atrophy all had their specific lesions transferred to FIs successfully. Finally, they calculated structural similarity (SSIM) indices between the generated FIs and the ground truth FIs, and achieved an average SSIM level of 0.802, indicating strong similarities between the image produced and the ground truth image.

Systemic diseases

UWF imaging is also being used in conjunction with DL for the prediction of non-ocular and neurological factors. As UWF imaging is a rich image format, exploratory studies have been conducted to determine if retinal changes can be associated with features like an individual’s age, vascular changes, and neurological status.

Age and brachial-ankle pulse-wave velocity

Nagasato et al. published a study in 2020 demonstrating an ability to predict both patient age and their brachial-ankle pulse-wave velocity (baPWV) using UWFIs and DL [93]. For each patient included in the study, they also recorded patient baPWV on the same day UWFIs were taken. They then processed these images to contain the entire image (the total image), a cropped region of the optic disk and macula (the central region), and the total image with the central region covered in black pixels (the peripheral region). Each of these processed images were used as separate datasets for the model and compared by the study authors in their performance for DL prediction of age and baPWV. They then used patient baPWV, UWFIs, and age as input data for a VGG-16 based CNN. The results showed that the total, central, and peripheral images were all able to predict the age and baPWV of a patient with statistical significance. Standardized regression coefficients of 0.833 and 0.390 were achieved for age prediction and baPWV prediction respectively. Specifically, the authors show that UWFIs can be used to make clear and specific predictions of a patient’s age and baPWV specifically, which is itself a marker of vascular health.

Alzheimer’s disease

In 2020, Wisely et al. used multiple imaging modalities to train a DL model to identify symptomatic Alzheimer’s disease (AD) [94]. In this study, the authors used UWFIs, UWF-FAF images, color maps of ganglion cell-inner plexiform layer (GC-IPL) thickness, and superficial capillary plexus en face OCTA for their training. They included these imaging modalities from eyes from cognitively healthy subjects and patients with symptomatic AD as confirmed by two expert neurologists. The DL model designed took the three imaging modalities as input, as well as OCT and OCTA numerical and patient data. The model used a shared-weight image feature extractor to extract modality-agnostic features that were then used in a modality-specific function in a fully connected layer. After training the model, they then tested the model on each imaging modality individually, as well as combinations of the data. They found UWFIs to lead to an 0.450 AUROC when inputted alone, and UWF-FAF images to achieve an 0.618 AUROC when inputted alone. On their own, OCTA achieved an 0.582 AUROC, and GC-IPL achieved an 0.809 AUROC. All images when inputted together achieved an 0.829 AUROC, while all images along with quantitative data achieved an 0.830 AUROC. All images and all data achieved an 0.836 AUROC, while GC-IPL, quantitative data, and patient data together achieved the highest AUROC of 0.841. These findings indicate that GC-IPL has the strongest individual predictive value of symptomatic AD and that the inclusion of more imaging modalities (i.e., OCTA, UWF-FAF, and UWF) do not significantly improve the predictive value in this case. As well, the predictive value of UWF alone for symptomatic AD is low.

Discussion

Across a broad range of domains and diseases, DL has been demonstrated to be useful when used in conjunction with UWF imaging. In the detection and classification of disease, DL models have been accurate, sensitive, and specific across a variety of ophthalmic disorders. In this review, we summarized the use of DL in detecting DR, RD, glaucoma, AMD, RP, pachychoroid, RVO, IMH, RH, and SCR from UWFIs. While disease detection has been the most published application of DL in UWFIs, its use in 25 of UWFIs and segmenting and localizing the structures of the retina should not be overlooked. Similarly, the high-resolution provided by each UWFI allows for novel generative uses with GANs. Finally, authors have demonstrated the novel utility of UWFI imaging’s high-resolution imaging of the eye for estimating a patient’s age and estimating vascular health via baPWV.

Benefits and risks of deep learning with UWFI

As shown by the studies discussed above, the diagnostic potential of DL when used with UWF imaging is accurate
and often exceeds the accuracy of trained ophthalmologists in many cases. The detection of ophthalmic diseases is a clear use for DL in UWF imaging. With automated DL systems for detection and diagnosis, the likelihood of detecting vision-threatening pathology early is greater. As VA often does not return as many ophthalmic disorders progress, the early and high sensitivity detection of these disorders is beneficial to patient care. In multiple studies discussed above, DL systems achieved sensitivities greater than trained ophthalmologists but in some cases, specificities that were lower than their human counterparts [46]. For this reason, while DL models have achieved impressive accuracies, a reduced specificity compared to humans may lead to greater false-positive diagnoses. From this, the risk of unnecessary medical or surgical intervention increases if DL models are followed without questioning their results. For cases like this, where DL models are more sensitive but less specific than ophthalmologists, clinical use of DL models is still beneficial. In these cases, the combined increased sensitivity of DL models paired with the greater or equal specificity of ophthalmologists may lead to a synergistic effect in accuracy. In a clinical environment, this pairing can help to reduce both false negatives via the DL model’s superior sensitivity and false positives via human specificity equal or greater than those of the DL models.

As DL models continue to learn associations between disease classifications and UWFI features, the risk of poor explainability is possible. Considering that DL models typically do not explain the associations they form, the possibility of associations being formed between unexpected image features and disease classification is possible. For example, the biases present in the training datasets may become codified in the DL models’ associations. This reproduction of human biases has been seen in AI implementations in healthcare in the past [95]. However, this can be mitigated by using large datasets with multiple graders and reviewers, to minimize individual human bias. Furthermore, to ensure that ophthalmologists can understand the image features leading to disease classification, Grad-CAM can be applied to show heatmaps of the image regions and lesions leading to stronger associations with a disease type [96, 97].

It should be noted as well, that there may be instances where the use of UWF imaging may not give added benefit to the goals of DL-based classification or disease detection. As discussed previously, in pathologies that do not use the periphery of the retina for diagnosis, the added view of the retina may not be beneficial. This is specifically important in the use of UWF imaging for glaucoma, where OCT may be a richer and more targeted data type. While the discussion of alternative modalities and their comparison is outside the scope of this review, it remains worthy of consideration for any practitioner interested in applying DL models for the detection and classification of ophthalmic disease.

**The future of ultra-widefield imaging and deep learning**

In this review, several novel methods using DL and UWFIIs have emerged. In particular, the use of GANs is useful to translate UWFIIs to other imaging modalities, such as FIs with high fidelity [89]. Furthermore, the ability of DL to translate existing FIs to UWFIIs provides the opportunity to translate decades of FIs available to ophthalmologists into a novel imaging modality, which can further strengthen DL training in the future.

UWF imaging and DL also have the potential to lead to an improved understanding of the pathophysiology of ophthalmic and non-ophthalmic disorders. While not yet demonstrated by the papers included here, using DL as an exploratory method can be proposed. For example, with a large enough dataset carefully classified, Grad-CAM interpretability models could demonstrate lesions on UWFIIs that may predispose or increase the likelihood of a specific ophthalmic disorder. When combined with other data types and other imaging modalities, as exemplified by Wisely et al. in their study for the detection of AD, the use of DL as an exploratory option for further understanding ophthalmic pathophysiology becomes possible [94].

As demonstrated by the quality assessment studies discussed in this review, DL may also be useful for expanded use of tele-ophthalmology. If general practitioners or technicians can access UWFI imaging for patients in remote areas, quality assessment methods can be used to determine if the UWFI imaging is of sufficient quality for an ophthalmologist to review remotely or at a later date. Other possibilities include the use of DL with UWFI imaging for the purpose of population-based screening of retinal and ophthalmic diseases. For this reason, DL’s utility is not simply in diagnosing and detecting diseases, but also in improving access to ophthalmic services for individuals in remote regions or with minimal access to ophthalmology expertise.

**Future directions**

For future studies, we hope to explore the differences of each study systematically and compare the accuracies of the proposed DL models quantitatively using meta-analysis methods. However, due to the heterogeneity of data and aims with UWFI imaging and DL, such analysis was not feasible at this time.

**Acknowledgements** None.
Declarations

Ethics approval This article does not contain any studies with human participants performed by any of the authors. As this is a review of previously published articles, no ethical or IRB approval was required.

Conflict of interest Fares Antaki reports research funding from Bayer and honoraria from Snell Medical Communication. Renaud Duval reports speaker honoraria from Bayer, Roche, Novartis and Alcon, a research grant and an unrestricted education grant from Bayer, and equity ownership in Optina Diagnostics. Nishaant Bhambra, Farida El-Malt, and AnQi Xu declare that they have no conflicts of interest.

Conflict of interest The authors declare no competing interests.

References

1. Agarwal A (2007) Fundus Fluorescein and Indocyanine Green Angiography: A Textbook and Atlas, 1st edn. Slack Incorporated, Thorofare, NJ
2. (1981) Diabetic retinopathy study. Report Number 6. Design, methods, and baseline results. Report Number 7. A modification of the Airlie House classification of diabetic retinopathy. Prepared by the Diabetic Retinopathy. Invest Ophthalmol Vis Sci 21:1–226
3. Kumar V, Surve A, Kumawat D et al (2021) Ultra-wide field retinal imaging: a wider clinical perspective. Indian J Ophthalmol 69:824–835. https://doi.org/10.4103/ijo.IJO_1403_20
4. Choudhry N, Duker JS, Freund KB et al (2019) Classification and guidelines for widefield imaging. Ophthalmology Retina 3:843–849. https://doi.org/10.1016/j.oret.2019.05.007
5. Kaines A, Oliver S, Reddy S, Schwartz SD (2009) Ultrawide angle angiography for the detection and management of diabetic retinopathy. Int Ophthalmol Clin 49:53–59. https://doi.org/10.1097/IIO.0b013e318191d471
6. Nagi A, Lalane RA, Sadda SR, Schwartz SD (2016) Ultra-wide-field fundus imaging: a review of clinical applications and future trends. Retina 36:660–678. https://doi.org/10.1097/IAR.0000000000000937
7. Wittmer MT, Kiss S (2012) The clinical utility of ultra-wide-field imaging. In: Review of ophthalmology. https://www.reviewofophthalmology.com/article/the-clinical-utility-of-ultra-wide-field-imaging. Accessed 1 Oct 2021
8. Optos Inc. (2021) Optos.com - Optos products. In: Optos products. https://www.optos.com/products/. Accessed 21 Aug 2021
9. Zeiss Inc. CLARUS 500. In: Clarus 500 Product Information. https://www.zeiss.com/meditec/int/product-portfolio/retinal-cameras/clarus500.html. Accessed 23 Apr 2022
10. Wessel MM, Aaker GD, Pearlitis G et al (2012) Ultra–wide-field angiography improves the detection and classification of diabetic retinopathy. Retina 32:785–791. https://doi.org/10.1097/IAR.0b013e3182278b64
11. Antaki F, Coussa RG, Mikhail M et al (2020) The prognostic value of peripheral retinal nonperfusion in diabetic retinopathy using ultra-widefield fluorescein angiography. Graefe’s Arch Clin Exp Ophthalmol 258:2681–2690
12. Liu TA, Arevalo JF (2019) Wide-field imaging in proliferative diabetic retinopathy. Int J Retin Vitreous 5:1–4
13. Fogliato G, Borrelli E, Iuliano L et al (2019) Comparison between ultra-widefield pseudocolor imaging and indirect ophthalmoscopy in the detection of peripheral retinal lesions. Ophthalmic Surg Lasers Imaging Retina 50:544–549. https://doi.org/10.3928/23258160-20190905-02
14. Quinn NB, Azuara-Blanco A, Graham K et al (2018) Can ultra-wide field retinal imaging replace colour digital stereoscopy for glaucoma detection? Ophthalmic Epidemiol 25:63–69. https://doi.org/10.1080/09286586.2017.1351998
15. Forshaw TRJ, Minór ÁS, Subhi Y, Sørensen TL (2019) Peripheral retinal lesions in eyes with age-related macular degeneration using ultra-widefield imaging: a systematic review with meta-analyses. Ophthalmol Retin 3:734–743. https://doi.org/10.1016/j.oret.2019.04.014
16. Sadda S (2019) Wide-field imaging in retina and vitreous diseases. In: International Journal of Retina and Vitreous
17. El Naqa I, Murphy MJ (2015) What Is Machine Learning? In: El Naqa I, Li R, Murphy MJ (eds) Machine learning in radiation oncology: theory and applications. Springer International Publishing, Cham, pp 3–11
18. Sathya R, Abraham A (2013) Comparison of supervised and unsupervised learning algorithms for pattern classification. International Journal of Advanced Research in Artificial Intelligence 2:34–38
19. Deng L (2014) Deep Learning: Methods and Applications. FNT in Signal Process 7:197–387. https://doi.org/10.1561/2000000039
20. Russell S, Norvig P (2020) Artificial intelligence: a modern approach, 4th edn. Pearson, Hoboken
21. Zhou X, Belkin M (2014) Chapter 22 - Semi-Supervised Learning. In: Diniz PSR, Suykens JAK, Chellappa R, Theodoridis S (eds) Academic Press Library in Signal Processing. Elsevier, pp 1239–1269
22. Hinton G (1999) Unsupervised learning: foundations of neural computation, 1st edn. Bradford Books, Cambridge, Mass
23. Monarch R (2021) Human-in-the-loop machine learning: active learning and annotation for human-centered AI. Manning
24. Dongare AD, Kharde RR, Kachare AD (2012) Introduction to artificial neural network. Int J Eng Innov Technol (IJEIT) 2:189–194
25. Zhou V (2019) Machine learning for beginners: an introduction to neural networks. In: Medium. https://towardsdatascience.com/machine-learning-for-beginners-an-introduction-to-neural-networks-d49f22d2389. Accessed 21 Aug 2021
26. Fornito A, Zalesky A, Bullmore ET (2016) Chapter 2 - Nodes and Edges. Fundamentals of brain network analysis. Academic Press, San Diego, pp 37–88
27. Ciresan DC, Meier U, Masci J et al (2011) Flexible, high performance convolutional neural networks for image classification. Proc Twenty-Second Int Joint Conf Artif Intell 2:1237–1242
28. Saha S (2018) A comprehensive guide to convolutional neural networks — the ELI5 way. In: Medium. https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3db21164a53. Accessed 21 Aug 2021
29. LeCun Y, Boser B, Denker JS et al (1989) Backpropagation applied to handwritten zip code recognition. Neural Comput 1:541–551. https://doi.org/10.1162/neco.1989.1.4.541
30. Krizhevsky A, Sutskever I, Hinton GE (2012) ImageNet Classification with Deep Convolutional Neural Networks. In: Pereira F, Burges C, Bottou L, Weinberger KQ (eds) Advances in Neural Information Processing Systems. Curran Associates Inc
31. Alippi C, Disabato S, Roversi M (2018) Moving convolutional neural networks to embedded systems: the alexnet and VGG-16 Case. In: 2018 17th ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN), pp 212–223
32. MathWorks Pretrained Inception-ResNet-v2 convolutional neural network - MATLAB inceptionresnetv2. In: MATLAB Mathworks. https://www.mathworks.com/help/deeplearning/ref/inceptionresnetv2.html?sessid=ae4b2abde0579ecab16f783b42f. Accessed 21 Aug 2021
33. Waring J, Lindvall C, Umeton R (2020) Automated machine learning: review of the state-of-the-art and opportunities for
healthcare. Artif Intell Med 104:101822. https://doi.org/10.1016/j.artmed.2020.101822
34. Faes L, Wagner SK, Fu DJ et al (2019) Automated deep learning design for medical image classification by health-care professionals with no coding experience: a feasibility study. The Lancet Digital Health 1:e232–e242. https://doi.org/10.1016/S2589-7500(19)30108-6
35. Korot E, Guan Z, Ferraz D et al (2021) Code-free deep learning for multi-modality medical image classification. Nat Mach Intell 3:288–298. https://doi.org/10.1038/s42256-021-00305-2
36. Alphabet Inc. Cloud automl custom machine learning models. In: Google cloud. https://cloud.google.com/automl. Accessed 5 May 2022
37. Touma S, Antaki F, Duval R (2022) Development of a code-free machine learning model for the classification of cataract surgery phases. Sci Rep 12:2398. https://doi.org/10.1038/s41598-022-06127-5
38. Antaki F, Coussa RG, Hamamjki K, Duval R (2021) Revisiting the problem of optic nerve detection in a retinal image using automated machine learning. Asia Pac J Ophthalmol (Philad) 10:335–336. https://doi.org/10.1097/APO.0000000000000398
39. Antaki F, Coussa RG, Kahwati G et al (2021) Accuracy of automated machine learning in classifying retinal pathologies from ultra-widefield pseudocolor fundus images. Br J Ophthalmol. https://doi.org/10.1136/bjophthalmol-2021-319030
40. Antaki F, Kahwati G, Sebag J et al (2020) Predictive modeling of proliferative vitreoretinopathy using automated machine learning by ophthalmologists without coding experience. Sci Rep 10:19528. https://doi.org/10.1038/s41598-020-76665-3
41. Shah T (2020) About train, validation and test sets in machine learning. In: Medium. https://towardsdatascience.com/train-validation-and-test-sets-72cb40cb9e7. Accessed 21 Aug 2021
42. Talebi H, Milanfar P (2021) Learning to resize images for computer vision tasks. CoRR abs/2103.09950
43. Calderon-Auzoa G, Perez A, Nakano-Miyatake M et al (2019) CNN-based quality assessment for retinal image captured by wide field of view non-mydriatic fundus camera. 2019 42nd International Conference on Telecommunications and Signal Processing (TSP) 282–285
44. Perez L, Wang J (2017) The effectiveness of data augmentation in image classification using deep learning
45. Nagasawa T, Tabuchi H, Masumoto H et al (2018) Accuracy of deep learning, a machine-learning technology, using ultra-widefield fundus ophthalmoscopy and optical coherence tomography angiography. Ophthalmology 125:1155–1165. https://doi.org/10.1016/j.ophtha.2017.12.028
46. Li Z, Guo C, Nie D et al (2020) Deep learning for detecting retinal detachment and discerning macular status using ultra-widefield fundus images. Commun Biol 3:15. https://doi.org/10.1038/s42003-019-0730-x
47. Parikh R, Mathai A, Parikh S et al (2008) Understanding and using sensitivity, specificity and predictive values. Indian J Ophthalmol 56:45–50
48. Mandrekar JN (2010) Receiver operating characteristic curve in diagnostic test assessment. J Thorac Oncol 5:1315–1316. https://doi.org/10.1097/JTO.0b013e3181e173d
49. Boyd K, Eng KH, Page CD (2013) Area under the precision-recall curve: point estimates and confidence intervals. In: Blockeel H, Kersting K, Nijsen S, Železný F (eds) Machine Learning and Knowledge Discovery in Databases. Springer, Berlin, Heidelberg, pp 451–466
50. Zou KH, Warfield SK, Bharath A et al (2004) Statistical validation of image segmentation quality based on a spatial overlap index. Acad Radiol 11:178–189. https://doi.org/10.1016/S1076-6332(03)00671-8
51. Wang Z, Keane PA, Chiang M et al (2021) Artificial intelligence and deep learning in ophthalmology. In: Lidström N, Ashrafian H (eds) Artificial Intelligence in Medicine. Springer International Publishing, Cham, pp 1–34
52. Bawany MH, Ding L, Ramachandran RS et al (2020) Automated vessel density detection in fluorescein angiography images correlates with vision in proliferative diabetic retinopathy. PLoS ONE 15:e0238958. https://doi.org/10.1371/journal.pone.0238958
53. Oh K, Kang HM, Leem D et al (2021) Early detection of diabetic retinopathy based on deep learning and ultra-wide-field fundus images. Sci Rep 11:1897. https://doi.org/10.1038/s41598-021-81539-9
54. Tang F, Luenam P, Ran AR, et al (2021) Detection of diabetic retinopathy from ultra-wide field scanning laser ophthalmoscope images: a multi-center deep-learning analysis. Ophthalmology Retina S246865302100035X. https://doi.org/10.1016/j.joret.2021.01.013
55. Wang K, Jayadev C, Nittala MG et al (2018) Automated detection of diabetic retinopathy lesions on ultrawidefield pseudocolor images. Acta Ophthalmol 96:e168–e173. https://doi.org/10.1111/aos.13528
56. Nagasawa T, Tabuchi H, Masumoto H et al (2019) Accuracy of ultrawide-field fundus ophthalmoscopy-assisted deep learning for detecting treatment-naïve proliferative diabetic retinopathy. Int Ophthalmol 39:2153–2159. https://doi.org/10.1007/s10792-019-01074-x
57. Nagasawa T, Tabuchi H, Masumoto H et al (2021) Accuracy of diabetic retinopathy staging with a deep convolutional neural network using ultra-wide-field fundus ophthalmoscopy and optical coherence tomography angiography. J Ophthalmol 2021:6651175. https://doi.org/10.1155/2021/6651175
58. Ding L, Bawany MH, Kuriyan AE et al (2020) A novel deep learning pipeline for retinal vessel detection in fluorescein angiography. IEEE Trans Image Process 29:6561–6573. https://doi.org/10.1109/TIP.2020.2991530
59. Orlando JI, Fu H, Barbosa Breda J et al (2020) REFUGE Challenge: A unified framework for evaluating automated methods for glaucoma assessment from fundus photographs. Med Image Anal 59:101570. https://doi.org/10.1016/j.media.2019.101570
60. Ohsugi H, Tabuchi H, Enno H, Ishitobi N (2017) Accuracy of deep learning, a machine-learning technology, using ultra-widefield fundus ophthalmoscopy for detecting rhegmatogenous retinal detachment. Sci Rep 7:9425. https://doi.org/10.1038/s41598-017-00891-x
61. Li Z, Guo C, Nie D et al (2019) A deep learning system for identifying lattice degeneration and retinal breaks using ultra-widefield fundus images. Annals of Translational Medicine 7:618–618. https://doi.org/10.21037/atm.2019.11.28
62. Zhang C, He F, Li B et al (2021) Development of a deep-learning system for detection of lattice degeneration, retinal breaks, and retinal detachment in tessellated eyes using ultra-widefield fundus images: a pilot study. Graefes Arch Clin Exp Ophthalmol. https://doi.org/10.1007/s00417-021-05105-3
63. Li Z, Guo C, Lin D et al (2021) Deep learning for automated glaucomatous optic neuropathy detection from ultra-widefield fundus images. Br J Ophthalmol 105:1548–1554. https://doi.org/10.1136/bjophthalmol-2020-317327
64. Masumoto H, Tabuchi H, Nakakura S et al (2018) Deep-learning classifier with an ultra-widefield scanning laser ophthalmoscope detects glaucoma visual field severity. J Glaucoma 27:647–652. https://doi.org/10.1016/j.jglau.2017.05.009
65. Ran AR, Tham CC, Chan PP et al (2021) Deep learning in glaucoma with optical coherence tomography: a review. Eye 35:188–201. https://doi.org/10.1038/s41433-020-01191-5
66. Maetschke S, Antony B, Ishikawa H et al (2019) A feature agnostic approach for glaucoma detection in OCT volumes. PLoS ONE 14:e0219126. https://doi.org/10.1371/journal.pone.0219126

Springer
67. Matsuba S, Tabuchi H, Ohsugi H et al (2019) Accuracy of ultra-wide-field fundus ophthalmoscopy-assisted deep learning, a machine-learning technology, for detecting age-related macular degeneration. Int Ophthalmol 39:1269–1275. https://doi.org/10.1007/s10972-018-0940-0

68. Li Z, Guo C, Nie D et al (2021) Automated detection of retinal exudates and drusen in ultra-widefield fundus images based on deep learning. Eye 1–6. https://doi.org/10.1038/s41433-021-01715-7

69. Masumoto H, Tabuchi H, Nakakura S et al (2019) Accuracy of a deep convolutional neural network in detection of retinitis pigmentosa on ultra-widefield images. PeerJ 7:e6900. https://doi.org/10.7717/peerj.6900

70. Kim IK, Lee K, Park JH et al (2021) Classification of pachychoroid disease on ultra-widefield indocyanine green angiography using auto-machine learning platform. Br J Ophthalmol 105:856–861. https://doi.org/10.1136/bjophthalmol-2020-316108

71. Nagasato D, Tabuchi H, Ohsugi H et al (2018) Deep neural network-based method for detecting central retinal vein occlusion using ultra-widefield fundus ophthalmoscopy. Journal of Ophthalmology 2018:1–6. https://doi.org/10.1155/2018/1875431

72. Nagasato D, Tabuchi H, Ohsugi H et al (2019) Deep-learning classifier with ultra-widefield fundus ophthalmoscopy for detecting branch retinal vein occlusion. Int J Ophthalmol 12:94–99. https://doi.org/10.18240/ijo.2019.01.15

73. Shi Z, Wang T, Huang Z et al (2021) A method for the automatic detection of myopia in Optos fundus images based on deep learning. Int J Numer Methods Biomed Eng 37:e3460. https://doi.org/10.1002/cnm.3460

74. Li Z, Guo C, Nie D et al (2020) Development and evaluation of a deep learning system for screening retinal hemorrhage based on ultra-widefield fundus images. Transl Vision Sci Technol 9:3. https://doi.org/10.1167/tvst.9.2.3

75. Dai L, Wu L, Li H et al (2021) A deep learning system for detecting diabetic retinopathy across the disease spectrum. Nat Commun 12:3242. https://doi.org/10.1038/s41467-021-23458-5

76. Calderon-Azuza G, Carrillo-Gomez C, Nakano M et al (2020) A teleophthalmology support system based on the visibility of retinal elements using the CNNs. Sensors 20:2838. https://doi.org/10.3390/s20102838

77. Li HH, Abraham JR, Sevgi DD et al (2020) Automated quality assessment and image selection of ultra-widefield fluorescein angiography images through deep learning. Trans Vis Sci Tech 9:52–52. https://doi.org/10.1167/tvst.9.2.52

78. Li Z, Guo C, Nie D et al (2020) Deep learning from “passive feeding” to “selective eating” of real-world data. NPJ Digit Med 3:143. https://doi.org/10.1038/s41746-020-00350-y

79. Ding L, Kuriyan AE, Ramchandran RS et al (2021) Weakly-supervised vessel detection in ultra-widefield fundus photography via iterative multi-modal registration and learning. IEEE Trans Med Imaging 40:2748–2758. https://doi.org/10.1109/TMI.2020.302765

80. do Nunez Rio JM, Sen P, Rasheed R et al (2020) Deep learning-based segmentation and quantification of retinal capillary non-perfusion on ultra-widefield retinal fluorescein angiography. J Clin Med 9:2537. https://doi.org/10.3390/jcm9082537

81. Wang Z, Jiang X, Liu J et al (2020) Multi-task siamese network for retinal artery/vein separation via deep convolution along vessel. IEEE Trans Med Imaging 39:2904–2919. https://doi.org/10.1109/TMI.2020.2980117

82. Sevgi DD, Srivastava SK, Wykoff C et al (2021) Deep learning-enabled ultra-widefield retinal vessel segmentation with an automated quality-optimized angiographic phase selection tool. Eye (Lond). https://doi.org/10.1038/s41433-021-01661-4

83. Niemeijer M, Xu X, Dumitrescu AV et al (2011) Automated measurement of the arteriolar-to-venular width ratio in digital color fundus photographs. IEEE Trans Med Imaging 30:1941–1950

84. Staal J, Abramoff MD, Niemeijer M et al (2004) Ridge-based vessel segmentation in color images of the retina. IEEE Trans Med Imaging 23:501–509

85. Estrada R, Tomasi C, Schmidler SC, Farsiu S (2014) Tree topology estimation. IEEE Trans Pattern Anal Mach Intell 37:1688–1701

86. Ju L, Wang X, Zhou Q et al (2020) Bridge the domain gap between ultra-widefield and traditional fundus images via adversarial domain adaptation. https://doi.org/10.48550/ARXIV.2003.10042

87. Ju L, Wang X, Zhao X et al (2021) Leveraging regular fundus images for training UWF fundus diagnosis models via adversarial learning and pseudo-labeling. IEEE Trans Imaging 40:2911–2925. https://doi.org/10.1109/TMI.2021.3056395

88. Xie H, Lei H, Zeng X et al (2020) AMD-GAN: Attention encoder and multi-branch structure based generative adversarial networks for fundus disease detection from scanning laser ophthalmoscopy images. Neural Netw 132:477–490. https://doi.org/10.1016/j.neunet.2020.09.005

89. Yoo TK, Ryu IH, Kim JK et al (2020) Deep learning can generate traditional retinal fundus photographs using ultra-widefield images via generative adversarial networks. Comput Methods Programs Biomed 197:105761. https://doi.org/10.1016/j.cmpb.2020.105761

90. Goodfellow I, Pouget-Abadie J, Mirza M et al (2014) Generative adversarial nets. In: Ghahramani Z, Welling M, Cortes C (eds) Advances in neural information processing systems. Curran Associates, Inc.

91. Karpathy A, Abbeel P, Brockman G et al (2016) Generative models. In: OpenAI. https://openai.com/blog/generative-models/. Accessed 18 Aug 2021

92. NVIDIA Research Projects (2021) StyleGAN - official tensorflow implementation. NVIDIA Research Projects

93. Nagasato D, Tabuchi H, Masumoto H et al (2020) Prediction of age and brachial-ankle pulse-wave velocity using ultra-widefield pseudo-color images by deep learning. Scientific Reports 10:19369. https://doi.org/10.1038/s41598-020-76511-4

94. Wisely CE, Wang D, Henoa R et al (2022) Convolutional neural network to identify symptomatic Alzheimer’s disease using multimodal retinal imaging. Br J Ophthalmol 106:388–395. https://doi.org/10.1136/bjophthalmol-2020-317659

95. Parikh RB, Teeple S, Navathe AS (2019) Addressing bias in artificial intelligence in health care. JAMA 322:2377–2378. https://doi.org/10.1001/jama.2019.18058

96. Selvaraju RR, Cogswell M, Das A et al (2020) Grad-CAM: visual explanations from deep networks via gradient-based localization. Int J Comput Vis 128:336–359. https://doi.org/10.1007/s11263-019-01228-7

97. Hanif AM, Beqiri S, Keane PA, Campbell JP (2021) Applications of interpretability in deep learning models for ophthalmology. Curr Opin Ophthalmol 32:452–458. https://doi.org/10.1097/ICO.0000000000000780

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