



Machine Learning Methods for Diagnosis of Eye-Related Diseases: A Systematic Review Study Based on Ophthalmic Imaging Modalities

Qaisar Abbas¹ · Imran Qureshi² · Junhua Yan² · Kashif Shaheed³

Received: 9 July 2021 / Accepted: 23 January 2022 / Published online: 18 February 2022

© The Author(s) under exclusive licence to International Center for Numerical Methods in Engineering (CIMNE) 2022

Abstract

Glaucoma, diabetic retinopathy, diabetic hypertension (DHR), Cataract, and age-related macular degeneration are some of the most common and important retinal diseases. A permanent vision loss occurs if these diseases are not discovered at an early stage. It is illustrated by numerous abnormalities in the retina such as microaneurysms (MA), hard exudates, soft exudates, or cotton wool spots, hemorrhages (HEM), neovascularization (NV), and macular edema (DME). An analysis of Ophthalmic imaging modalities is used as computerized tools by ophthalmologists for faster screening and diagnosis of these diseases. Compared to traditional-machine learning, multilayer deep learning (MDL), speeds up diagnosis and improves precision and accuracy. As a result, this survey offers a thorough overview of the new technologies used in each phase of the retinal diagnosis process concerning various image modalities. Among different imaging modalities, this paper is primarily focused on the retinal fundus photograph and optical coherence tomography images. This review aims to provide a more in-depth survey of the most significant components, types, and network architectures of MDL algorithms. This paper also describes the advantage of using other different MDL algorithms that are not being used in the past. To assist the researcher, the challenges and potential developments of current traditional and deep learning methods are described in depth. At last, the challenges and recommended solutions are also presented to support scientists in terms of research gaps. We have measured the influence of benchmarks (GPU, and CPU) on MDL algorithms in this paper. Moreover, the limitations and future direction are also mentioned to assist other experts. The comparative results and discussion section of this article suggest that a rigorous MDL model is still required compared to traditional machine-learning to effectively diagnose eye-related diseases in several modalities.

1 Introduction

✉ Qaisar Abbas
qaabbas@imamu.edu.sa; qaisarabbasphd@gmail.com
http://www.imamu.edu.sa

Imran Qureshi
imarwat11@gmail.com

Junhua Yan
yjh9758@126.com

Kashif Shaheed
Kashifshaheed1@gmail.com

¹ College of Computer and Information Sciences, Imam Mohammad Ibn Saud Islamic University (IMSIU), P.O. Box 5701, Riyadh 11432, Saudi Arabia

² Key Laboratory of Space Photoelectric Detection and Perception, Ministry of Industry and Information Technology and College of Astronautics, Nanjing University of Aeronautics and Astronautics, Nanjing 211106, Jiangsu, China

³ School of Computer Science and Engineering, South China University of Technology, Guangzhou 510006, China

Diabetes mellitus (DM) to be 463 million in the year 2019 and increased to 700 million by 2045 [1]. As the most common and specific complication of DM, diabetic retinopathy (DR) is also one of the leading causes of preventable blindness if detected at an early stage. In particular, the retina is not only involved in the effect of DR but several other eye-related diseases such as glaucoma, diabetic hypertension, and age-related macular degeneration (AMD) [2, 3]. In practice, diabetes is quite frequent in those over 30 years old, and uncontrolled diabetes can lead to these various eye-related diseases [4]. The early stages of DR are less severe and can be effectively treated. Due to DR, many DR-related lesions are detected to diagnoses such as Microaneurysms (MA), Hard exudates, soft exudates, or cotton wool spots (CWS), hemorrhages (HEM), neovascularization (NV), and macular edema (ME) [5]. A visual example of these DR-related lesions with numerous components and disease symptoms

is shown in Fig. 1. Ophthalmologists are mainly diagnosed eye-related diseases such as DR, DHR, Glaucoma, AMD, Anterior segment lesions, and Cataracts through various imaging modalities. Table 1 shows the various eye diseases with respect to imaging modalities. Specifically, the 80 million people affected by 2020 by Glaucoma, which is a leading cause of blindness. Another prevalent cause of vision loss in humans is diabetic retinopathy (DR). The global proportion of diabetes patients is anticipated to rise from 2.8% in 2000 to 4.4% in 2030. Another prevalent eyesight condition is age-related macular degeneration (AMD) [6]. It can cause vision loss in the center of the visual field in

the human eye, leading to a full loss of central vision over time. Most countries' healthcare systems have a low doctor-to-patient ratio. Diagnosis and correct treatment become error-prone and time-consuming because of an overwhelmed patient-care system.

Several Imaging modalities [8] were used in the past to diagnose eye-related diseases such as Retinal Fundus photograph (RFP), Optical Coherence Tomography (OCT), Heidelberg Retina Tomography (HRT), RetCam, and SLIT Lamp. A visual example of these modalities is displayed in Fig. 2. Moreover, in Table 1, we summarize these different types of imaging modalities utilized in the past to diagnose

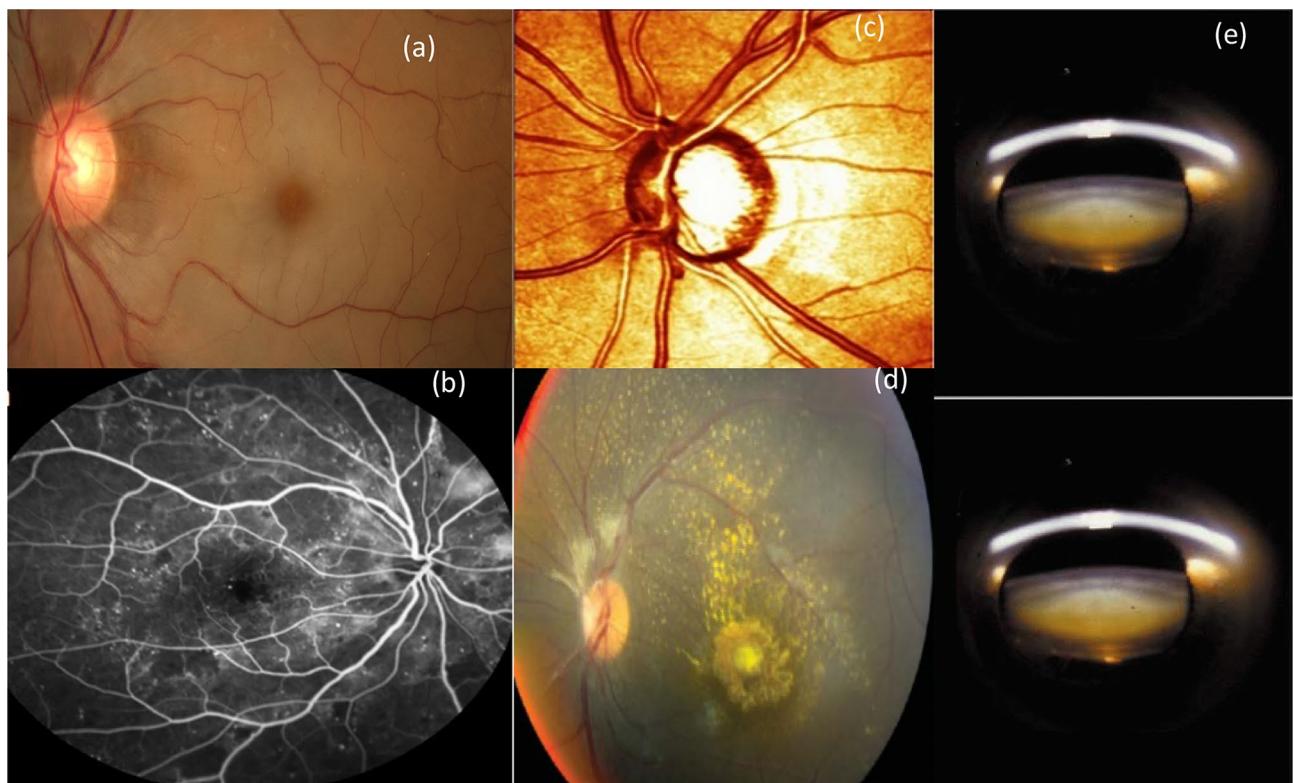


Fig. 1 Multimodalities retinal pictures to diagnosis eye-related diseases [7] where figure **a** color fundus image, **b** enhanced OCT image, **c** HRT, **d** RetCam and **e** SLIT Lamp

Table 1 Most popular image modalities used to diagnosis eye-related diseases

Eye-related disease modalities	Technology	Diseased diagnosis
Retinal Fundus photograph (RFP)	2D; It extracts large area of fundus that can provide one-time using handset ophthalmoscopes	DR, DHR, Glaucoma, AMD
Optical coherence tomography (OCT)	3D; high-resolution cross-sectional imaging	Glaucoma, AMD
Heidelberg retina tomography (HRT)	2D; confocal scanning laser ophthalmoscopies	Glaucoma
RetCam	2D; wide-angle imaging	Segment anterior lesions and diagnosis of retinopathy of prematurity
SLIT Lamp	2D; high-intensity light source stereoscopic magnified view of eye structure	Cataract

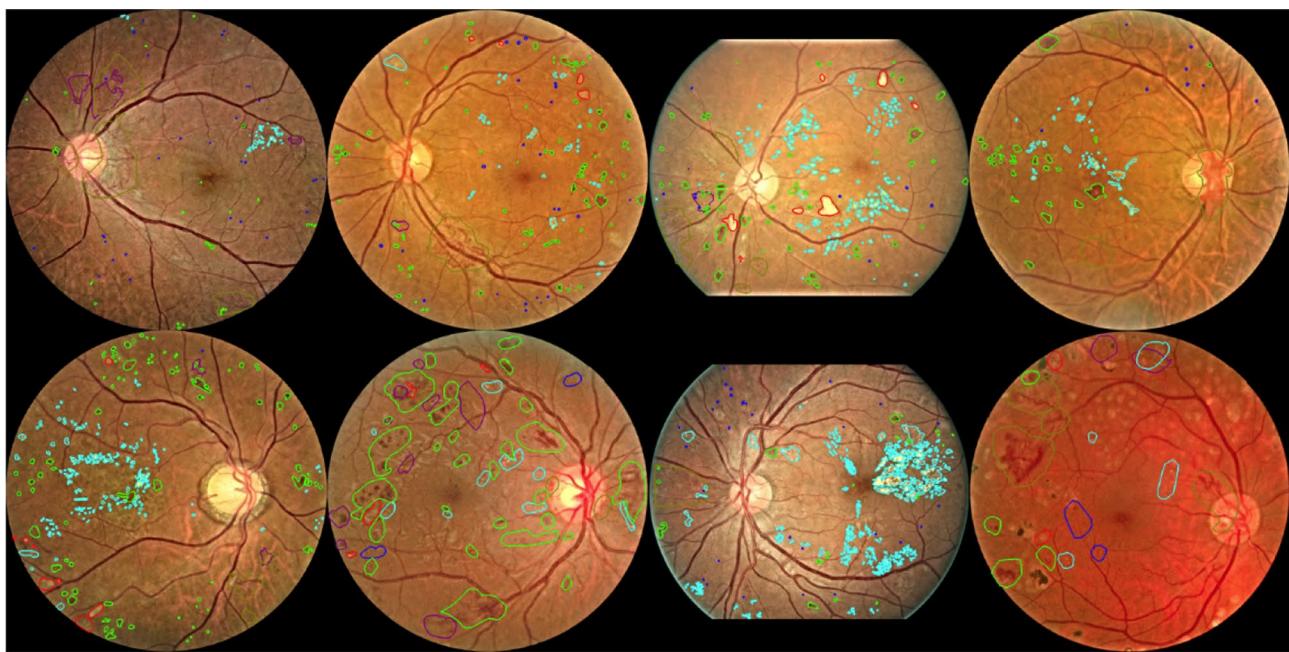


Fig. 2 A visual example of fine-grained annotated diabetic retinopathy (FGADR) dataset [5] to recognize eye-related diseases

eye-related diseases. In this study, our primary focus is based on the retinal fundus photograph (RFP) and optical coherence tomography (OCT) images. that is mostly utilized in past studies. Compared to other imaging modalities, the retinal fundus photograph (RFP) is one of the most common types of retinal imaging to diagnosis DR, DHR, Glaucoma, and AMD. The RFP is a typical ophthalmic imaging technique in which optical cameras are employed to obtain enlarged images of retinal tissues. In the past, there have been a few studies comparing to medical experts and automatic algorithm performance in diagnosing various imaging modalities [7]. Optical coherence tomography (OCT) is another optical image technology that is used to provide in-depth patterns to diagnose mainly Glaucoma and AMD eye-related diseases. The Heidelberg Retina Tomography (HRT) is also used in the past to diagnose mainly Glaucoma. This HRT takes 3-dimensional images of the optic nerve and surrounding retina with a specific laser to get the surface of the optic nerve. Also, the RetCam used the 2D-wide angle for images to segment anterior lesions. To diagnosis Cataract eye-related diseases, the scientists used mainly 2-D image technology captured by SLIT Lamp. This eye disease is one of the most common causes of reversible vision loss and blindness in the world.

Over the years, researchers have achieved significant progress in developing state-of-the-art automatic techniques [9] to diagnose eye-related diseases. In this paper, a comprehensive survey of recent advances in traditional-machine learning (TML) versus advanced multilayer deep-learning (MDL) algorithms is presented for the classification of eye-related

diseases based on multimodal ophthalmic imaging modality. Those automatic systems were developed in the past known as computer-assisted diagnostics (CADx). Several computer-assisted diagnostics (CADx) [10–12] systems were developed in the past to help ophthalmologists for screening eye-related diseases. However, those CADx systems focused on a single image modality to diagnosis mostly single eye-related diseases. To develop those CADx systems, the authors utilized in the current year many image-processing, traditional-machine learning, and advanced deep-learning approaches. Traditional CAD system architectures in retinal image processing used numerous predetermined templates and kernels to compare with manually annotated and segmented regions of these eye-related images. Compared to this, the deep learning models are exceptionally strong architectures for detecting patterns in nonlinear combinations of many data kinds. Without the need for manual feature extraction, it extracts relevant and essential representations from the data. Deep learning algorithms [13–17] have largely replaced classic machine learning algorithms in recent years, exceeding traditional classifiers in most cases.

To analyze eye-related diseases, the authors used a pre-processing step to enhance the image and then perform the segmentation step, which is a very crucial step to detect accurate lesions as shown in Fig. 2. It noticed that a pre-processing step by following the machine-learning and or deep-learning algorithms is outperformed to recognize eye-related diseases through images. The pre-processing step is very important because images contain noises and ununiform light illumination. In this article,

we comprehensively discussed and compared different machine-learning algorithms on different image modalities for the diagnosis of various eye-related diseases. According to our limited knowledge, we did not find a single article to describe this kind of comparison where different deep learning architectures have been implemented

for applications in clinical ophthalmology. Moreover, we have reviewed available data sources and computational tools to build a platform for the diagnosis of all eye-related diseases. Figure 3 shows year-wise trends in the published literature and the number of papers for different application areas. Papers are classified into subsections according

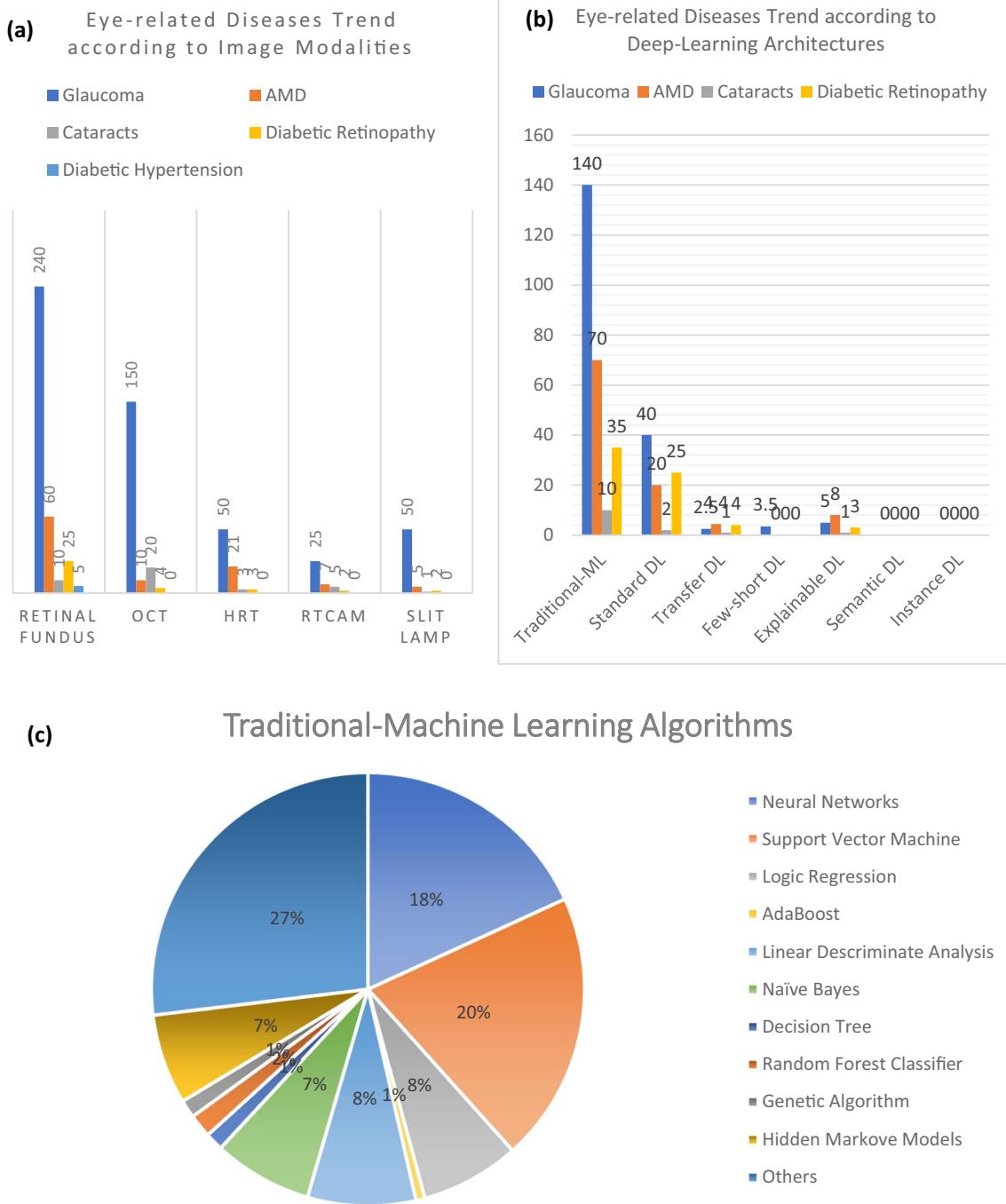


Fig. 3 A visual trend of different imaging modalities and various types of traditional-machine and deep learning algorithms to diagnosis eye-related diseases where Figure **a** shows the trend of various

imaging modalities, **b** represents eye-related diseases with respect to deep-learning and **c** figure indicate the traditional machine learning algorithms for diagnosis of eye-related diseases

to the traditional and advanced machine-learning (TML and MDL) algorithms with respect to imaging modalities. In this article, we have explained the importance of image modalities to diagnose eye-related diseases by using MDL algorithms. Several TML and MDL algorithms are evaluated to get performance measures on the CPU and GPU-based benchmarks. To further highlight the research gap in this field, the new MDL algorithms are also described that can be used in the future for the diagnosis of eye-related diseases.

1.1 Paper Organization

The remainder of this paper's structure is as follows. Figure 4 shows a flow diagram of the paper organizing process. The research protocol, which includes research questions, comparisons to previous survey studies, and the selection of publications for this review article, is discussed in Sect. 2. The framework of several deep-learning-based models that have been utilized in the past is described in Sect. 3. Section 4 examines recent trends in the state-of-the-art in the areas of machine and deep learnings for the detection of

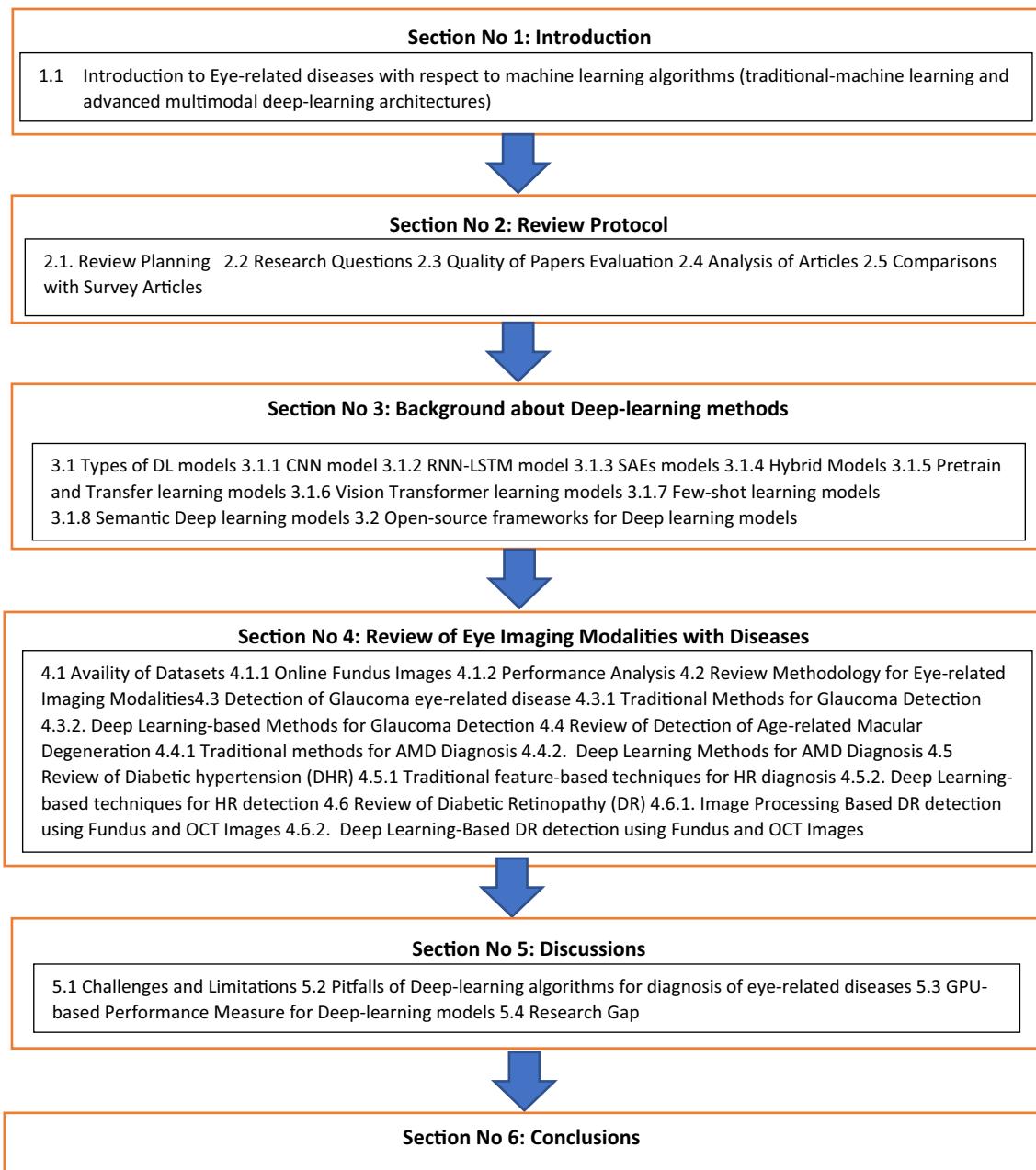


Fig. 4 Systematic flow chart of review paper for diagnosis of eye-related diseases with respect to imaging modalities

eye-related disorders using ophthalmologic imaging modalities. Moreover, we offered various open-source datasets and libraries in this section to create a system for recognizing eye-related disorders. To aid other researchers, this section discusses the current limitations and future directions in this topic. In addition, Sect. 5 discusses the problems and future directions to assist other researchers in developing up-to-date eye-related disease recognition systems. Finally, this paper concludes in Sect. 6.

2 Review Protocol

A research methodology must be understood to work in ophthalmology in any type of image modality study or survey to diagnose eye-related diseases. In this article, we have described the different research approaches that let other authors assist in a wide domain. The review protocol outlined the specific layout and methods used throughout the evaluation. The protocol also specified the guidelines and instructions for conducting a survey on current work based on current machine-learning (ML) and deep learning (DL) popular algorithms to diagnosis different eye-related diseases such as glaucoma, diabetic hypertension, and age-related macular degeneration (AMD) in different imaging modalities such as Retinal Fundus (RF), Optical Coherence Tomography (OCT), Heidelberg Retina Tomography (HRT), RetCam and SLIT Lamp. This review article is unique compared to other survey papers because we did not find a single review article that describes a systematic review about eye-related diseases related to different imaging modalities. This paper is focused on the retinal fundus photograph (RFP) and optical coherence tomography (OCT) images to diagnose eye-related diseases. that is mostly utilized in past studies. Therefore, this article is provided useful and technical information for other researchers who are just starting in this field. Moreover, we have also reviewed the latest machine-learning algorithms that help other researchers to perfectly build models to recognize eye-related diseases. In this domain, we offered sufficient information on classical and deep learning-based techniques. A review planning, research questions, study sources, data collecting sources, criteria for inclusion and exclusion, and data extraction are all examples of sub-concepts in the review procedure. The survey also helps users compared the performance of traditional and deep learning-based systems, as well as the impact of deep learning technology on the recent development of computerizing systems for recognition of different eye-related diseases, the types of deep models that have already been implemented, and how those system's performance can be improved. In addition, the survey may assist with cost, data analysis and management, and finding difficulties relating to both classical and deep learning-based eye-diagnostic tools.

2.1 Review Planning

The purpose and review planning for studying the most modern and widely utilized machine-learning algorithms to diagnosis eye-related diseases in various imaging modalities is the starting point for writing a review paper. Our target audience should be able to understand it rather easily. The approach used to communicate the objective of the review, the goals, the questions generated, and other relevant aspects determine the effectiveness of the review article. The technique utilized to attain the objective, the criterion for inclusion and exclusion, and the outcome comparison are all clearly shown in the review planning.

3 Research Questions

The research questions are at the heart of every study whether it is a literature review or a research project. This is the methodological starting point for all scientific studies, regardless of specialty. The study questions associated with the eye-related diseases characteristic in various imaging modalities and motivation are listed in Table 2.

3.1 Quality of Papers Evaluation

Several articles are studied and described to complete this article. To search articles, we used different keywords that are mentioned in Table 3. The research boundaries will be defined by a set of rules that determine inclusion and exclusion. In general, standards are evaluated after the research topic has been defined and before the research process has begun. Irrelevant and out-of-concerned documents, for example, were refused in this case. We would examine the research if the wide field of study is related to our subject. Top journals, such as IEEE, Springer, ScienceDirect, Science Wiley, Top conferences (SCI and ESCI), and Top conferences must all be mentioned. In addition, the indicated review effort must contain the research activity of competent scientists. The focus should also be on well-presented qualitative and quantitative research, which should include studies over the last five to three years. Old studies and research, on the other hand, must be excused because they do not contribute to our objectives. Following the inclusion and exclusion criterion, a quality evaluation procedure on nominated articles is used to choose acceptable and shortlisted articles for investigation. Because our study area is so large, it includes a wide range of sub-areas of interest, as well as many reputable SCI and ESCI publications. As a result, several methods must be followed to choose study quality. As a result, our paper passes the quality evaluation rule.

Table 2 Research questions and motivations involves in this study to recognize eye-related diseases

S. No.	Research questions	Motivations
RQ1	What major type of research is being done for recognition of eye-related diseases?	To understand recent studies, which are being conducted in this domain
RQ2	What kind of ocular modalities is being used to diagnosis the disease?	To know about recent popular ocular modalities
RQ3	What is the classification of ocular diseases related to modalities?	To have in-depth knowledge about eye-related diseases
RQ4	What is the classification of traditional-machine learning (TML) being used to diagnosis this disease?	To know and study the current TML approaches in this domain
RQ5	What is the technical background of deep-learning architectures being used?	To do critical study for understand current deep-learning algorithms
RQ6	What are the latest trends in recognition of eye-related diseases?	To know latest trend to recognize eye-related diseases
RQ7	To know and study the state-of-the-art approaches in this domain	To know various methodologies applied to recognize eye-related diseases
RQ8	What are the latest deep-learning architectures that are not still being used to diagnosis this disease?	To understand the models that are still being used in this domain
RQ9	What are the challenges to use DL models for recognition of eye-related diseases?	To know the current challenges of DL algorithms in this domain
RQ10	What are the latest survey articles to recognition of eye-related diseases?	To know and study the state-of-the-art approaches in this domain
RQ11	What are the online available datasets being used by other studies?	To know about the availability of datasets
RQ12	What is the favorable data size available for machine-learning algorithms?	To study and know about the size of dataset is really used
RQ13	What are the main performance measures used for validating the eye-related diseases?	To know about the statistical metrices used in the past
RQ14	What is the role of ML and deep learning architectures to recognize diabetic retinopathy and hypertension?	To study various latest-machine-learning approaches
RQ15	What are the limitations and advantages of benchmarks to predict eye-related diseases?	It assists the other researchers to select a best model for recognition
RQ16	What is the effect of imaging modalities to eye-related diseases?	To inform other researchers for selecting best modalities
RQ17	What are the latest-machine-learning architectures that must be considered to recognize eye disease?	To identify research gap in this domain
RQ18	What are the future directions in recognition of eye-related diseases by using deep-learning architectures?	To know the current and future prospective of this domain
RQ19	What is the role of latest machine learning approaches in recognition of eye-related diseases?	Learning the importance and usefulness of latest
RQ20	What is the current trend of Deep learning?	Learning the importance and usefulness of latest

3.2 Analysis of Articles

Current trend of deep-learning algorithms for diagnosis of eye-related diseases are described visually by a Fig. 5. This information is collected from different platforms such as PubMed and top journals in the years of 2006 to 2021. Moreover, some papers are also included in the year of 2022. It noticed that the latest trend is to use deep-learning architectures compared to traditional-machine learning algorithms.

3.3 Comparisons with Survey Articles

In terms of methodological comparisons of different features and machine-learning algorithms against state-of-the-art survey publications in this domain, we have conducted

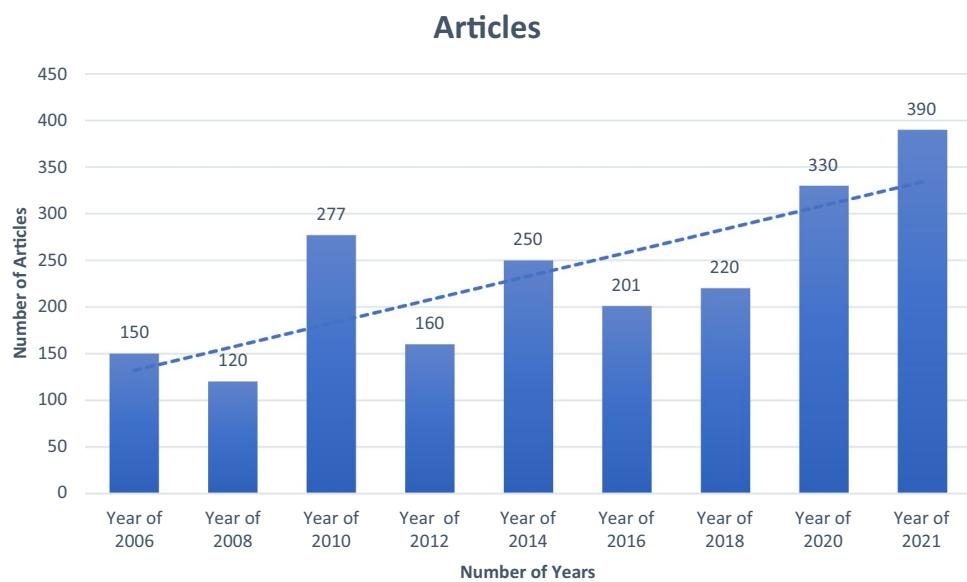
substantial research. The differences between our study and the current surveyed papers are shown in Table 4. Compared to existing survey articles, we have covered many eye-related diseases in multimodal image modalities.

4 Background of Deep Learning Architectures

Machine learning models have traditionally been trained to perform useful tasks using manually created features retrieved from raw data or features learned by other simple machine learning models. Deep learning bypasses this tedious and difficult phase by allowing computers to learn meaningful representations and features straight from raw data. Artificial neural networks and their derivatives are by far the

Table 3 Research keywords and selected papers involves in this study

S. No.	Search keywords	IEEE	Springer	ACM	Elsevier	Science Wiley	Top conferences
1	Retinopathy + deep learning	40	60	10	150	60	10
2	Retinopathy + machine Learning	15	20	15	12	8	20
3	Retinal Fundus + deep learning	40	60	10	200	100	15
4	Retinal Fundus + machine learning	5	10	5	13	10	20
5	Diabetic retinopathy + deep learning	40	60	10	200	100	20
6	Diabetic retinopathy + machine learning	14	16	15	13	7	34
7	Diabetic hypertension + deep learning	1	2	0	2	0	0
8	Diabetic hypertension + machine learning	5	10	5	13	10	20
9	Diabetic macular edema + deep learning	20	40	30	50	200	10
10	Diabetic macular edema + machine learning	5	10	5	13	10	20
11	Computerize retinal Fundus images	5	50	30	50	150	12
12	Classification + OCT	8	20	10	11	14	5
13	Classification + HRT	20	10	30	20	12	10
14	Classification + SLIT Lamp	5	10	5	13	10	20

Fig. 5 Current trend for deep-learning algorithms to recognize eye-related diseases

most prevalent models in deep learning, but there are others. The focus on feature learning, or automatically learning data representations, is a frequent aspect of deep learning approaches. This is the key distinction between deep learning and more “classical” techniques. Finding features and completing a task are combined into a single challenge, and both are improved throughout the same training phase. For general overviews of the field, see [15, 16]. Deep learning in medical imaging has inspired interest in convolutional neural networks (CNNs) [17], a powerful methodology for learning useful representations of photos and other structured data. Before CNN’s could be employed efficiently, these qualities had to be manually developed or produced by less

powerful machine learning models. Many handcrafted image features were abandoned when it became more practical to use data-driven features, as they proved to be almost useless when compared to feature detectors discovered by CNN. The CNNs have strong biases ingrained in them because of the way they are designed, which helps us understand them better.

4.1 Types of DL Architectures

Multilayer Deep learning (MDL) models for classification [34, 35] are Deep Belief Networks (DBNs), Recurrent Neural Networks (RNNs), Convolutional Neural Networks

Table 4 Summary of survey papers on eye-related diseases using latest-machine learning approaches

No	Survey articles	Reviewed year	Research questions																			
			RQ1	RQ2	RQ3	RQ4	RQ5	RQ6	RQ7	RQ8	RQ9	RQ10	RQ11	RQ12	RQ13	RQ14	RQ15	RQ16	RQ17	RQ18	RQ19	RQ20
1	Cheung et al. [2]	2021																				✓
2	Tong et al. [8]	2020																				✓
3	Saeed et al. [7]	2016	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	
4	Zhang et al. [9]	2020																				✓
5	Mayyat et al. [11]	2021																				✓
6	Ponni et al. [18]	2021																				✓
7	Ahmed et al. [19]	2014																				✓
8	Mirzania et al. [20]	2020																				✓
9	Abdullah et al. [21]	2021																				✓
10	Stefan et al. [22]	2020																				✓
11	Sengupta et al. [23]	2018																				✓
12	Usman et al. [24]	2017																				✓
13	Ran et al. [25]	2021																				✓
14	Priyanka et al. [26]	2021																				✓
15	Pead et al. [27] and Papadopoulos [28]	2019																				✓
16	Papadopoulos Z [29]	2019																				✓
17	Alyoubi et al. [30]	2020																				✓
18	Stolte et al. [31]	2020																				✓
19	Imran et al. [32]	2019																				✓
20	Imran et al. [33]	2016																				✓
21	Our review	2021																				✓

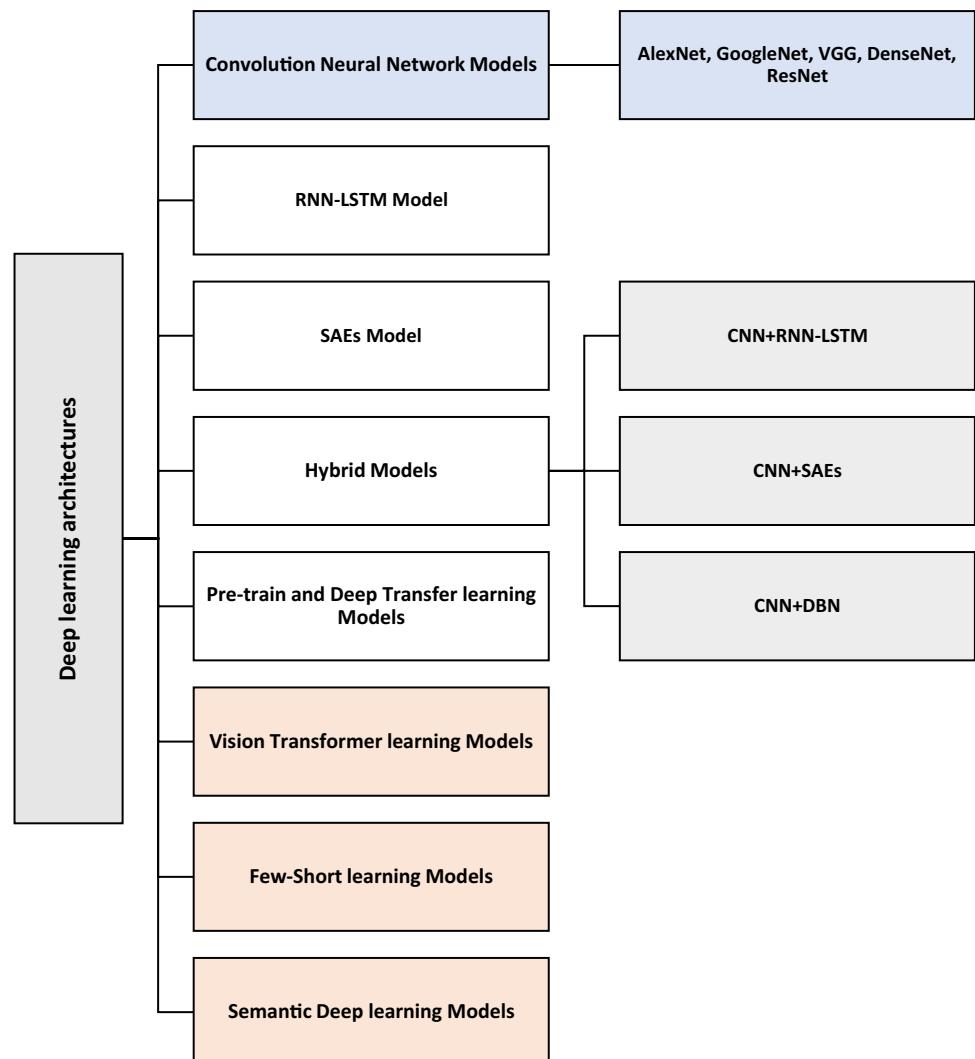
(CNNs), Stack-based Autoencoders (SAEs), Restricted Boltzmann Machines (RBMs), and Recurrent Neural Networks (RNNs) (DBN). In previous driver fatigue systems, SAEs, RNN, and CNN were used for feature engineering, but they failed to classify data. RBM and DBN, on the other hand, are the greatest deep learning algorithms for data representation, but they have not been tested in previous eye-disease detection systems. We chose CNN, RNN-LSTM, stack-based autoencoders (SAEs), and pre-train transfer learning models for our comparison as shown in Fig. 6. In the following subsections, we will go through those ML models in more detail.

4.1.1 CNN Model

Convolutional Neural Networks (CNNs) [36] models are utilized in many different applications because of their ability to distinguish spatial data from digital images. These CNNs models are the most popular variant of deep learning

algorithms and applied on many different fields such as image recognition, ubiquitous, and object searching [37, 38]. Salient features of these models include regularized structure, good spatial locality, and translation invariance. In fact, a typical CNN template was used to detect facial features. Since, the CNN model is widely utilized to various features from images. A typical CNN model is visually displayed in Fig. 7a. This type of CNN model includes one input layer, two convolutional layers (Con-L), one fully connected layer (FC-L), and one output predicted layer (Pred-L). Each of the layers are followed by pooling (Pool-L) or dropout (DRP-L) layer. The main trick to CNN is to simplify input data to a type that is easier to understand, with as little loss of information as possible in multiple layers. In practice, the convolutional layers (Con-L) in CNN models are the main blocks, which include a series of filters to convert the input data to represent the features into the different size of features maps. To design Con-L, there are several parameters that should be defined such as the number of filters, the size of each filter,

Fig. 6 Types of advanced machine-learning models that are briefly described in this survey article



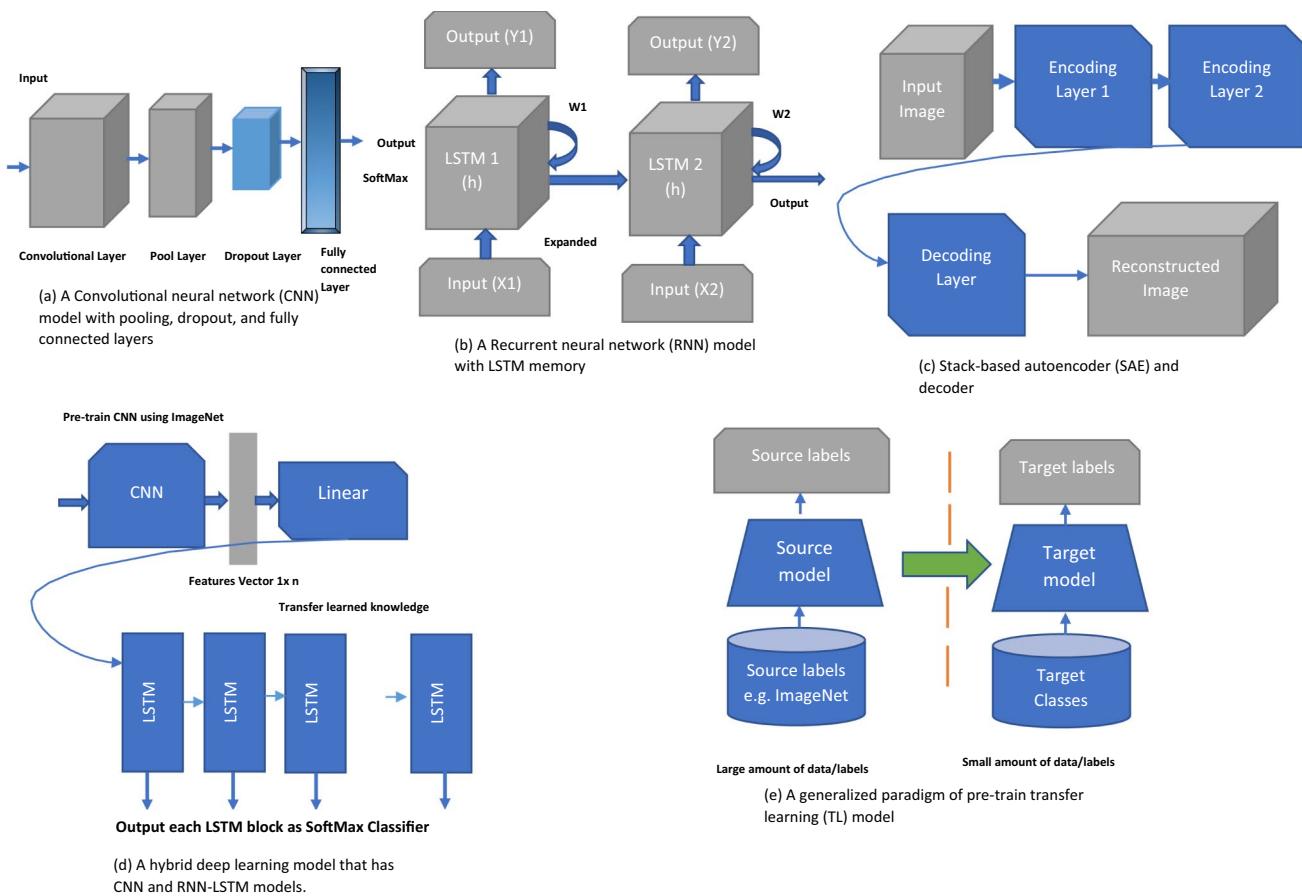


Fig. 7 Types of deep-learning models developed in the past to recognize class

etc. In fact, the Con-L is followed by Pool-L layer. The Pool-L layer is normally implemented to decrease the spatial size of the feature maps. As a result, the number of parameters (weights and bases) and the computational burden can be reduced into some extend. To develop Pool-L layer, it is needed to use three kinds such as max, min or average type of pooling layer. In the pooling layer, the hyper-parameters include the pooling activity, the size of the pooling field, the strides, etc. As in the simple neural network, in the fully connected layer (FC-L), the nodes have complete connections to all activations in the previous layer. Finally, some authors used dropout layers (Drop-L) to optimize features, but domain-expert knowledge is required to introduce this layer in the network. For the development of the feature map, a Con-L layer uses a linear filter and a nonlinear activation function for the scan of the 2D input data when it is applied to image processing. The nonlinear activation function of the linear convolution layer feature map is calculated as:

$$f_{p,q,r}^i = \max(W_T \cdot I_{P,Q} + b_r, 0) \quad (1)$$

where W_T is the weighted vector, $I_{P,Q}$ is the input patch from the input image and b_r is the parameter to include basis terms that is used in Eq. (3). Also, the \max is the non-linear activation function of the r feature map and n is the parameter indicates the number of multi-layers and calculated by Eq. (4).

$$f_{p,q,r}^n = \max(W^r \cdot n^T f_{p,q,r}^i - 1 + b_k n, 0) \quad (2)$$

Pooling layer is another significant layer, which is used in the architecture of CNN model. It was noticed that max pool layer is normally utilized in the past compared to min or avg pooling layers. So, non-linear activation function was used to implement max pool layer. This layer is gradually decreasing the spatial size of the input image and decrease the parameters. This Pool-L layer is definitely helpful for reducing the computational burden on the CNN's network. Hence, the Pool-L is used to control network overfitting. In many architectures, the Pool-L layer is placed between successive Con-L layers (followed by ReLU layer) in CNN architecture. The most common form is a pooling layer is calculated by Eq. (5)

$$f_{x,y}^n(S) = \max_{a,b=0} (S_{2x+a, 2y+b}) \quad (3)$$

Each max operation in this case is over 4 numbers. The dimension of the depth in the architecture remains unaltered. Pooling units may use other features in addition to max pooling, such as average pooling or ℓ_2 -norm pooling. Average pooling has always been used traditionally, but compared to max pooling, which performs better in practice, it has recently collapsed. The rectified linear unit (ReLU) is another important activation function $f(x) = \max(0, x)$. Other functions are also used to increase nonlinearity, for example the saturating hyperbolic tangent $f(x) = \tanh(x)$, and the sigmoid function $\sigma(x) = (1 + e^{-x})^{-1}$. In practice, the ReLU function is very important to use compared to other functions as it is used to train the network many times. The ReLU is often preferred to other functions because it trains the network several times to improve accuracy of the classifier. Afterwards, the high-level information gain is obtained in more general form by using fully connected layers (FC-L) after many convolutional and max pooling layers.

Since most of the parameters are occupied by a FC-L layer, it is vulnerable to overfitting. As a result, the Dropout layer (Drop-L) is an important layer to reduce overfitting in the network architecture. Individual nodes are either "dropped out" of the network at each training stage with probability $\{1-p\}$ or retained with probability $\{p\}$, so that a reduced network is left. To develop Drop-L layer, it is important to separate incoming and outgoing edges from the dropped-out node. Only the reduced network in that process is trained on the results. Subsequently, the removed nodes with their original weights are reinserted into the network. The probability of losing a hidden node is generally 0.5 in the training phases; however, this probability is normally even lower for input nodes. Finally, the prediction layer (Pred-L) is implemented through SoftMax function to recognize many classes in the classification problem. In the CNN network model, the SoftMax function transforms the values into 0 to 1 and this function is calculated by Eq. (6) as:

$$S(x_i) = \frac{e^{x_i}}{\sum_{j=1}^n e^{x_j}} \quad (4)$$

where the $x_1, x_2, x_3, \dots, x_n$ are the input vectors that are translated into the range of 0 to 1.

4.1.2 RNN-LSTM Model

Recurrent neural networks (RNNs) [39–41] are used in few state-of-the-art detection algorithms for diagnosis of eye-related diseases. In case of hybrid driver fatigue systems, the RNNS have been used in many studies and a simple RNN architecture is visually displayed in Fig. 7b. Although, it has capability to process data in real-time and generate

new sequence based on training dataset. In general, to avoid long sequences, the RNN models are best utilized to generate optimized sequence on hidden layers. Several researchers have used RNN model without training on eye open or close dataset. In practice, the RNN models are performing any task based on few input values classified by the network layers. However, the RNN models have limited flexibility to recover the past mistakes because it has a recurrent nature. As a result, the RNN model was totally performed unexpected prediction results. Those results can be improved by adding some noise signals to the RNN structure and if we added then it can go back and forth to recognize more distinct features based on long short-term memory (LSTM). Below we presented the mathematical representation of recurrent neural network (RNN) deep-learning model. Given a training features $(x^1, x^2, x^3, \dots, x^n)$ with the label $(y^1, y^2, y^3, \dots, y^k)$, the mapping of input layer, hidden layer and output layer is defined as:

$$Out^t = f(v s^t + b) \quad (5)$$

$$s^t = g(u x^t + w s^{t-1} + b) \quad (6)$$

where $f(\cdot)$ and $g(\cdot)$ are nonlinear function, b is the bias, and V , U and W represent the weight parameters.

4.1.3 SAEs Model

The stack-based autoencoder (SAEs) [42–44] are mostly utilized in the past to detect driver fatigue. In general, SAEs models are based on auto-encoder that represent the data model. For construction of SAEs models, an encoder mapping-function is required to transform an input vector into hidden-layer representation alongside weighted-matrix and biases. Next component, the decoder, maps the hidden units to the reconstruct input. During its operation, the auto encoder is trying to compare the reconstructed input with the original input data by optimizing the error. In fact, the reconstructed the value must as close as possible to the original input value. Later, the SAEs used trained denoising layer to recover data through trained autoencoder. The output from this autoencoder can be used in many supervised learning algorithms such as SVM, AdaBoost or ANN after all the layers of the network has been trained. A main advantage of SAEs is to train a layer-wise network and it has capability to combine with many machine learning algorithms. If one can change activation function then it can be used with real, binary, or probabilistic distribution. The SAEs is also contained self-fine tunable process through back-propagation algorithm to reduce the total reconstruction loss. In Fig. 7c, it is visually depicted. Autoencoder has an input layer, hidden layer and output layer as shown in Fig. 1. For training it has two components: an encoder and a decoder. The encoder

maps the hidden representation of the input data, and the decoder recreates it. Because of the unlabeled input dataset, where the measured hidden encoder vector is represented, and this is the output layer decoder vector. The encoding method, therefore, is as follows:

$$h_n = f(W_1 X_n + b_1) \quad (7)$$

where the function $f(\cdot)$ is the encoding function, W_1 is the weight matrix of the encoder, and b_1 is the bias vector. As shown in Eq. (8). Hence, the decoder process of Eq. (8) is defined as follows:

$$x_n = g(W_2 h_n + b_2) \quad (8)$$

where $g(\cdot)$ is the decoding function, W_2 is the weight matrix of the decoder, and b_2 is the bias vector in the Eq. (9). In stack-based autoencoders (SAEs), it contains number of autoencoder parameters that can be optimized by using Eq. (10) as follows.

$$\text{Recons - Error}(\emptyset) = \arg \min \frac{1}{n} \sum_{i=1}^n L(x, x^{i+1}) \quad (9)$$

where the parameter L represents a lost function $(x, x^{i+1}) = \|x - x^{i+1}\|^2$. The autoencoder structure uses an unsupervised layer-wise learning algorithm to stack n autoencoders into n hidden layers and then fine-tuned by a supervised process. So, it is possible to divide the SAEs-based approach into three stages. Train the first autoencoder with input data and obtain the vector of the learned feature; the former layer feature vector is used as the next layer input, and this process is repeated until the training is completed.

4.1.4 Hybrid Models

Recent trends in driver fatigue detection are to developed hybrid deep learning models [45–47] to learn features and classify them. The standard hybrid deep learning model, for instance, uses a feature extraction representation algorithm and classification discriminative algorithms [45]. The SoftMax layer, which will not be treated as an algorithmic component in this survey, implements almost all the classification functions in neural networks. In previous systems, the researchers have developed hybrid deep-learning models to provide the recognition tasks with more descriptive and low-dimensional features, whereas these models are more generative to support the training process. The CNN and RNN-LSTM [46, 47] models are often mixed to recognize features in a robust way, according to the literature as visually displayed in Fig. 7d. The combinations of RNN-LSTM and CNN models provide excellent performance in terms of temporal and spatial features that make help in the training process. It is very common in many driver fatigue

recognition systems to combine them for both temporal and spatial feature learning. Our research shows that several studies have also combined SAEs with CNN or MLP algorithms. Deep Belief Networks (DBN), Convolutional Neural Networks (CNN), and Long Short-Term Memory (LSTM) networks are among the common techniques used for these models. For pattern analysis, DBN is often used, and they practice faster than other methods of deep learning. In image processing applications, CNN is often used and has greater discriminative capacity than DBN. LSTM networks are a type of Recurrent Neural Network (RNN) that can better learn dependencies than other types of RNN between extracted features. The weight-dropped long short-term memory (WDLSTM) network is another variant of LSTM. The drop-connect technique uses this regularized RNN on the repeated hidden-to-hidden weight.

4.1.5 Pre-train and Deep Transfer Learning (DTL) Models

Pre-train CNN models [48] have been sued in recent times to extract high-level features from video sequence and still images. In pre-trained CNN model, the network is trained on ImageNet dataset or trained from the scratch and then fine-tuned on VGG16. This transfer learning method is successfully outperformed on many driver fatigue detection systems [49] to reduce the training cost in terms of uni- or multi-modal features learning. Although in the literature, we have several standard CNN pre-trained models available such as AlexNet, GoogLeNet, DenseNet and ResNet50 but for computer vision point-of-view, we can have some other popular models such as VGG-16, VGG-19, Inception V3, Xception and ResNet-50 [48–52] etc. An example of this transfer learning using CNN and RNN-LSTM models are represented in Fig. 7e. Pre-trained models are nowadays utilized in many computer vision applications. The main advantage of using a pre-trained model is to let the system for automatically features extraction by training processing. As a result, the manual features extraction step became automatic by the help of CNN training. This process is known as transfer learning. Moreover, this transfer learning is performed on the pre-trained CNN models to do fine-tuning. So, those pre-trained CNN models are best candidate for classification tasks for the development of HDx systems. Moreover, in this paper, we have applied different variation of transfer learning algorithms (TLAs) on large dataset to detect driver drowsiness as a small-scale dataset.

In case of TLAs, you can use information from the last trained model (weights and features, etc.) and manage problems such as having less information for the new assignment. The fundamental concept of transfer learning is demonstrated in Fig. 12. As shown in this figure, many deep learning-based frameworks and models are represented through by using multiple layers architecture.

Those multilayers are integrated together to learn distinct features and the last layer is always implemented in the form of FC-L layer, which contains SoftMax function at the end for the classification decision. If the last layer is removed, then this architecture can be used to extract distinct features by using a pre-trained strategy. For any pre-trained network, the extraction of visually distinguishable features is achieved by using fine-tuning. Hence, the TLAs provided very important architectures to automatically learn features from any dataset with pre-train strategy. Therefore, it is important to describe TALs architectures in a brief manner. Several transfer learning (TL) models were developed in the past for automatic recognition of different tasks. One basic TL model was developed in [51] known as VGG-NET. In VGG-NET TL model, there are 2 to 3 times in-depth Con-L layers compare to CNN model. In one variant of VGG-Net, there were around 138 million of parameters that were calculated in this model [52]. In practice, the VGG-NET model has a good representation of features for more than a million images (ImageNet dataset) from 1000 different categories. There are many variants of VGG-NET TL model but VGG-166 is mostly utilized in the past. The architecture of VGG-16 utilized 3 convolutional filters with 13 Con-L layers. Where each Con-L layer is followed by a ReLU layer and has max-pool layers. To perform classification task, this model has 3 FC-L layers in the network. This VGG-16 model can be used as best candidate to automatically extract features. The VGG16 is suitable well [53] for both object recognition and edge detection problems. To fine-tune any network as a model, we can also use the VGG16 network model. Suppose, we have a dataset with m samples $\{(x(1), y(1)), \dots, (x(m), y(m))\}$ for training. The network overall cost function can be defined as follow:

$$J(W, b) = \left[\frac{1}{m} \sum_{i=1}^m \left(\frac{1}{2} \|K_{w,b}(x^i) - y^i\|^2 \right) \right] + \frac{\lambda}{2} \sum_{k=1}^n \sum_{i=1}^{sl} \sum_{j=1}^{sj} (W_{j,i}^l)^2 \quad (10)$$

where the K_w , $(x(i))$ is the neural network model, W_{ji} (l) is the connection weight between the j th element of layer 1 and the i th element of layer $l+1$, and b is the bias term of the hidden layer neuron. There is an important TL algorithm known as residual network (ResNet). In comparison to VGG-Net, the architecture of ResNet TL algorithm is more organized. The most important point in ResNet is that it has blocks that feed the values to the next layer and we can define skip connections to skip the blocks. ResNet50 architecture consists of replacing both layer blocks in a 34-layer network with a 3-layer bottleneck layer. The architectures of ResNet101 and ResNet152 use more layers than 3-layer blocks. 11.3 billion FLOPs are found in ResNet152. It is less complex than the networks of VGG16/19 [48]. In a research

on the CIFAR10 dataset with 10 schools, 50,000 educational images and 10,000 test images, ResNet architecture was the winner. As the depth of CNN models increases, the convergence of CNN models is affected by the issue of increased gradients that disappear and explode. The ResNet architecture is defined based on multiple residual blocks in the network. In each residual block, the Con-L layer with input along with ReLU layers, there is Drop-L and skip connection layer as well. It depends on specific problems to design each residual block in the network. There are many different forms of RNN model, but the ResNet50 is the mostly utilized ResNet model [54]. Apart from VGG-NET TL algorithm, another model known as DenseNet was proposed by Huang et al. [55]. This model uses image information more effectively. In the scheme, the other layers are fed forward to each layer. In this way, the property information of all layers before it can be accessed by any layer l . The general network architecture of densely connected neural networks can be calculated by the following equation.

$$X_l(\emptyset) = H_l(X_0, X_1, X_2, \dots, X_{l-1}) \quad (11)$$

where the $H_l(\cdot)$ is the transfer function that is used to process property information and it is calculated by Eq. (11). To decrease the parameters, the authors in [53] developed another architecture known as InceptionV3 in 2013. To implement InceptionV3 TL model, we require 1×1 convolutional size in calculation compared to all other TL models thereby helping to reduce number of parameters. In the network, there are many variants of transfer learning algorithms as mentioned before. However, InceptionV3 is the TLAs model that contains many layers in the network of different sizes. In the past studies, it is also mentioned that the GoogLeNet is a type of InceptionV1 type of model. Moreover, in the past, there are many variants that depends on different applications.

The MDL networks can be expressed by a broad collection of models due to the neural network's many flexibilities. These architectures are known as deep models, and they are made up of different layers [56]. The AlexNet was the first deep learning architecture, created by Alex Krizhevsky, Geoffrey Hinton, and their collaborators, who pioneered deep learning research. The design is made up of convolutional and pooling layers that are piled on top of each other, with entirely interconnected layers on top. The advantages and supremacy come from the fact that scalability and GPU utilization are unrivaled. Because of the GPU, AlexNet has a fast processing and training speed. Visual Graphic Group Net is a website dedicated to visual graphics. This pyramid-shaped net was created by technologists from the Visual Graphics Group at Oxford University. The model is made up of bottom layers that are wide and top layers that are deep. To narrow the layers,

VGG accommodates consecutive convolutional layers and finally the pooling layers.

The GoogleNet [57] design, and therefore the name of the Net, was introduced by Google researchers. It has 22 layers, whereas VGG only had 19. Google Net is built on a cutting-edge approach known as the inception module. A single layer has a variety of feature extractors that aid the network's performance. When several of these inception modules are placed on top of each other, the final one emerges. Because of the joint and simultaneous training, the model converges faster. With a modest pre-trained GoogleNet, GoogleNet training is faster than VGG. The ResNet Residual Network is made up of a series of residual modules that are referred to as “successive residual modules”. The leftover modules are stacked one on top of the other to build a fully functional node-to-node network. The fundamental advantage of ResNet is that it can train a network with many residual layers.

The ResNeXt [58] is built using ResNet ideas, but with a new and improved architecture that improves performance. The RCNN (Regions with Convolutional Neural Network) works by drawing a bounding box around the objects in the image and identifying the object within it. The YoLo is used to overcome image detection issues. To determine the object's class, the image is segmented into bounding boxes, and then a common recognition method is applied to all the boxes. After the classes have been identified, the boxes are carefully blended to form the optimal bounding box around the objects. It is utilized in real time to deal with day-to-day issues. With little bandwidth, the SqueezeNet [59] design is the most powerful architecture to choose. The inception

procedure will take 100 MB and the network architecture will require 4.9 MB. The extreme change is handled by a fire module. SegNet is widely regarded as the most effective model for picture segmentation. SegNet is a deep neural network that is used to handle the challenges of image segmentation. It is made up of an interconnected network of encoders for pixel-wise categorization and an arrangement of processing layers termed encoders. The capacity of SegNet to maintain very high frequency features in the segmented image is a key feature. Pooling indices are connected between the encoder and decoder networks. The information is also flowing in a straight line. GAN (Generative Adversarial Networks) is an innovative network design that generates wholly new and distinct images that are not found in the available training dataset.

4.1.6 Vision Transformers Learning (VTL) Models

In recent studies [60–64], the transformers have been successfully applied to various image recognition, natural language processing (NLP) and vision tasks, with comparable performance to convolutional neural networks (CNNs). A visual architecture diagram of vision transformers learning (VTL) model is displayed in Fig. 8. In practice, the transformers used self-attention method to capture spatial patterns and global interactions of pixels [60] compared to stacking of convolutional layers in CNN models. This permits transformers to accumulate rich global data without needing to use CNNs to generate layer-wise local feature extractions, resulting in better performance. A 12-block transformers model with 22 M learnable parameters, for example,

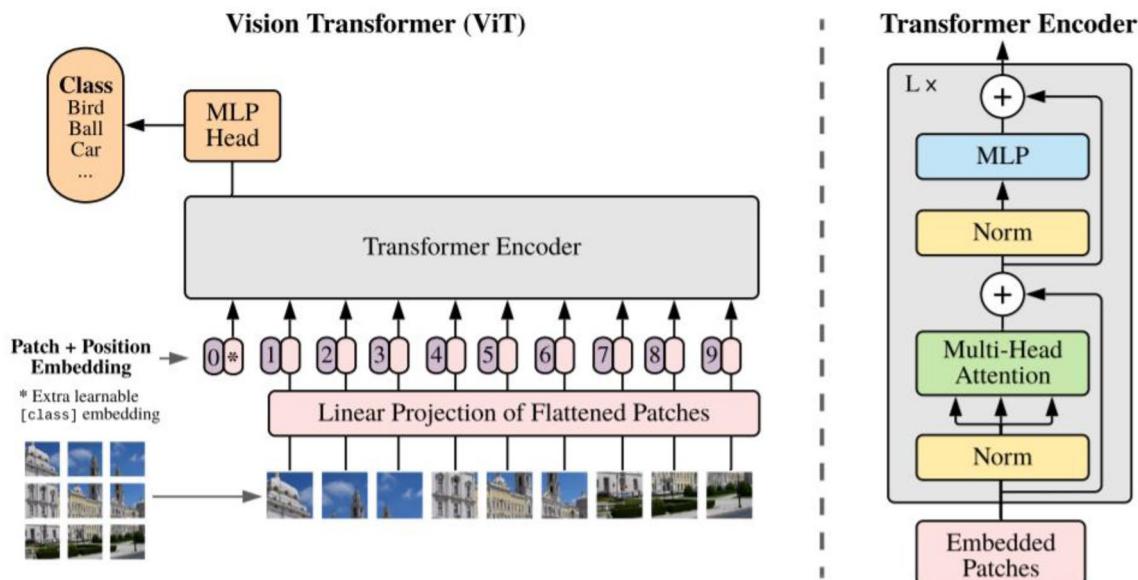


Fig. 8 The Transformer architecture [61] used in many Computer Vision Applications and medical image classification tasks

delivers better results compared to standard CNN model [61]. This transformer models are used to totally replaced recurrence and convolutions states of MDL algorithms by using self-attention process in the network. In addition, the transformers eventually become the most popular models for a variety of NLP [62] problems. Recent studies attempted to merge the self-attention mechanism into CNN models. These accomplishments further allow the researcher's interest in developing purely transformer-based models for vision tasks without worrying about convolutions or inductive bias.

The transformers [63] is one of the earliest attempts to attain competitive performance with CNNs on the image classification challenge. Due to high model complexity, the transformers required to be pre-trained on large-scale datasets to perform effectively on any image recognition tasks. With big pre-trained model, the transformers used knowledge of distillation to train the model and it allows to the user to use the architecture without doing pre-train on large datasets. On the other hand, it focuses on a distinct challenge namely, how to successfully scale transformers to be deeper for solving problems of image recognition. Several authors suggested a new self-attention mechanism that can perform well on vision tasks without the use of supplementary data, pre-train networks, or domain-specific inductive bias. Furthermore, they suggested that the stacking of more transformers blocks is ineffective as compared to CNNs [64] to improve the performance. They noticed that when the depth of transformers grows, the attention maps used for aggregating the features for each transformer block become extremely similar after a certain number of layers, causing the representations to cease evolving. As a result, the attention collapse is the term we use to describe this specific

problem. However, if one can do contribution on self-attention mechanism, the transformers can be applied to not only vision or natural language processing (NLP) tasks but also suitable for image classification.

4.1.7 Few-Shot Learning (FSL) Models

Clinical experts are needed an automatic way to find certain from medical images. Last decades, there are several machine-learning based methods are developed to recognize medical diseases through deep learning, transfer learning, semantic learning, and vision transformers. However, these models are required huge parameter and labelled training datasets to build the model architecture. While it is not always a case that big medical datasets are always available to train and test the model. To overcome these concerns, the concept of few-shot learning (FSL) [65–69] is developed in the literature. A visual architecture diagram of FSL model is shown in Fig. 9. This FSL system is used basic architecture, which required few labeled examples to train the model and they are capable to generalize on huge and unseen datasets. As a result, the FSL algorithms can be effectively utilized to recognize eye-related diseases when diagnosis by fundus or OCT image modalities.

The few-shot learning (FSL) algorithm is a well-known technique to solve different problems in practice, but it is not used in eye-related disease diagnostic tasks according to our limited knowledge. Since the creation of the first convolutional neural network (CNN) algorithms, deep learning performance on computer vision (CV) tasks has improved dramatically [69]. Microsoft announced in 2015 that their algorithm could classify photos from the ImageNet dataset

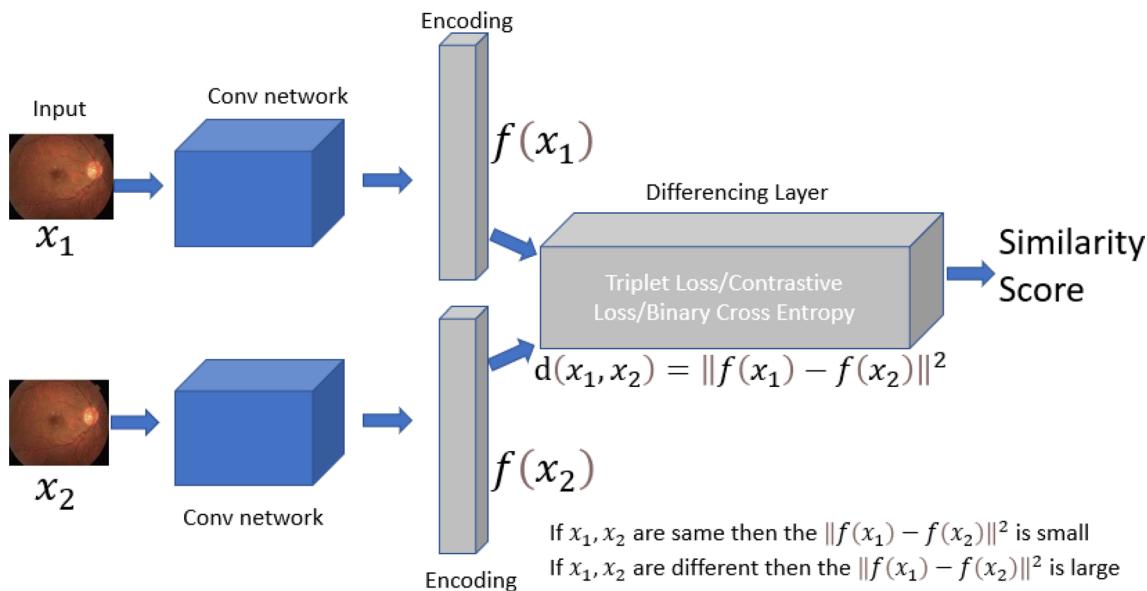


Fig. 9 One-shot learning with siamese network with prototypical network to classify retinograph images

better than humans. When it comes to leveraging billions of photos to tackle a single task, computers today are unrivaled. Even still, in the actual world, finding or creating a dataset with that many samples is rare. What are our options for dealing with this issue? We can employ data augmentation (DA) or collect and label more data if we are talking about a CV task. Labeling tasks in CVs is a time-consuming and costly operation, but it yields superior results. Both strategies may not be useful if the dataset is tiny. Consider an assignment in which we must create a classification with only one or two examples per class, each of which is extremely difficult to locate. This would necessitate novel approaches. One of these is Few-Shot Learning (FSL). This would call for innovative approaches. Few-Shot Learning (FSL) is one of them. There are different variations of Few-Shot Learning variations such as N-Shot Learning and Few-Shot Learning.

4.1.8 Semantic Deep Learning (STDL) Models

Semantic image segmentation (SIS) is known as pixel-level categorization [70], which is a wide operation that creates a label for each pixel. Currently, the SIS approach is broadly utilized in image processing and computer vision applications. In contrast to producing an image prediction, this SIS method creates a detection of objects with embedded spatial information. Deep learning algorithms have outperformed standard computer vision approaches in a variety of applications over the last decade, including object categorization, detection, and semantic segmentation [71]. An example of STDL model is visually represented in Fig. 10 to recognize medical images. Deep learning approaches generate features that are automatically adjusted for the specific classification tasks, making them superior choices for dealing with

complex scenarios. The adoption and extension of deep learning approaches were sparked by their significant success in other fields. Even after decades of work, providing meaningful names to the parts in a remote sensing image remains a difficult task. To work on a variety of imaging modalities, the SIS approach is capable to detect and learn effective features. In addition, the SIS is also facilitated by the MDL algorithm that provided benefits to researchers for better image analysis.

In recent years, reviews of deep learning algorithms [72] for remote sensing challenges have been conducted. Convolutional neural networks (CNNs), autoencoders, and other fundamental deep learning approaches were reviewed by certain writers. A follow-up survey looked at other deep neural network topologies, such as recurrent neural networks, as well as a variety of distant sensing applications. Deep learning approaches are reviewed and applied to applications such as anomaly detection, disaster analysis and assessment, land cover categorization, and segmentation. Reviews on semantic [74] image segmentation using deep learning algorithms have recently been published. One study looked at deep learning architectures for semantic segmentation of optical images and categorizes the methods that are currently available. The authors covered deep neural network methods from architectural perspectives, such as up-sampling techniques, convolutional approaches, weakly supervised and unsupervised methods, and so on. Other researchers looked examined the methodologies for semantic segmentation and the data sets used for evaluation, in addition to the five deep learning architectures. This study supplied performance analysis error measures and provided a quantitative comparison in terms of time, memory, and effort. In terms of venues and years, the review presents an overview of publications

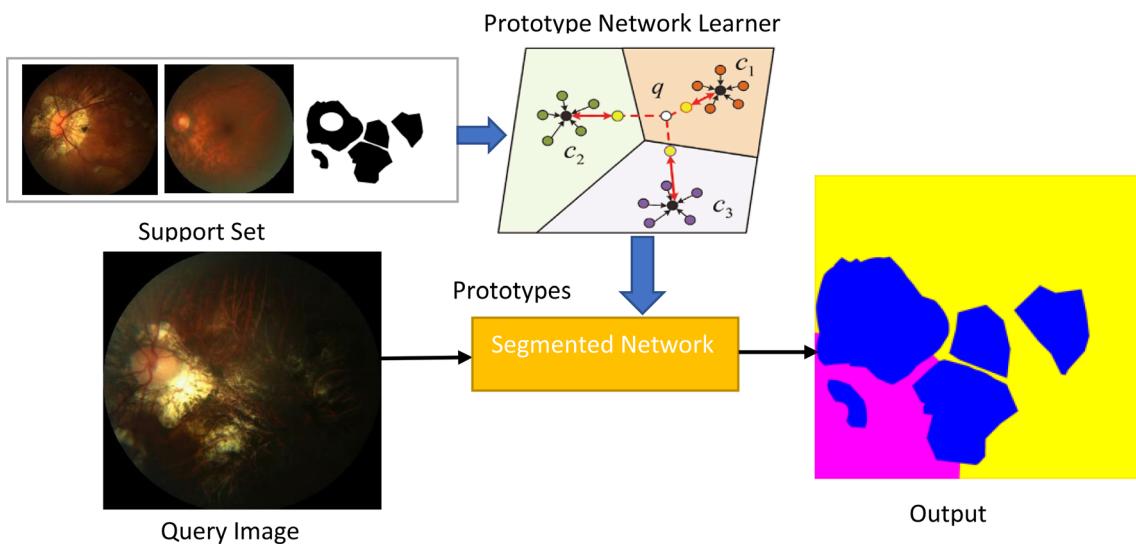


Fig. 10 An example of prototypical network based on few-short learning architecture for semantic based image segmentation

and applications. The focus is on object detection, with picture segmentation covered lightly. The evaluations are primarily concerned with current advances in deep neural networks and their applicability to remote sensing imaging in general. There is no comprehensive survey of deep learning algorithms for semantic segmentation of remote sensing data that we are aware of. Due to the rapid development of new deep learning approaches in recent years, it is vital to summarize the evolution and provide a complete evaluation to scholars and practitioners, as well as identify unresolved difficulties in the semantic segmentation of remote sensing.

4.2 Open-Source Frameworks

Several open-source frameworks were developed in the past to support multilayer deep-learning (MDL) algorithms [74]. Table 5 contains many open-source libraries with different attributes that can help other researchers to choose specific library. The MDL models have a significantly different architecture from the CPU. It was primarily intended for the rendering of graphic pictures at the time of its inception. The GPU, like the CPU, is a graphics card processor that performs sophisticated mathematical and geometric calculations necessary for graphics rendering. The CPU does not need to undertake graphics processing work with the GPU, allowing it to focus on other system activities, considerably

improving the computer's overall performance. Those frameworks are comparatively analyzed in Table 5 and described in the subsequent paragraphs.

Convolutional architecture for fast feature embedding (Caffe): BVLC now hosts Caffe, which was born in Berkeley, California. Caffe has fast performance, smooth switching between CPU and GPU modes, and Windows, Linux, and Mac cross-platform support. Blobs, Layers, and Nets are the three primary atomic structures that Caffe uses to create its programming framework. It extensively abstracts the structure of the deep neural network in terms of the "Layer," greatly improves execution performance by some intricate design, and maintains flexibility based on efficient implementation.

TensorFlow: TensorFlow is an open-source software library that performs numerical calculations using data flow diagrams. On November 9, 2015, Google formally introduced the TensorFlow computing framework, and in 2017, Google TensorFlow version 1.0 was released, signifying the framework's official use in the production environment. The TensorFlow calculation framework can handle a variety of deep learning algorithms such as CNN, RNN, and LSTM, but its use isn't restricted to deep learning; it can also be used to build general machine learning. The components of TensorFlow are fantastic, and it offers a lot of capability. TensorBoard provides visualization

Table 5 The software development framework based on deep learning theory

No.	Library name	Purpose	Deep learning methods	GPU support	Cross platform
1	Caffe	It provides the high-performance features to work on both CPU and GPU based benchmarks	CNN	Yes	Yes
2	TensorFlow	Google officially opened the computing framework TensorFlow for numerical calculations and visualization	CNN, RNN, LSTM and others	Yes	Yes
3	PyTorch	It provides more clear libraries compare to TensorFlow and it is used to create complex deep learning architectures	CNN, RNN, LSTM and others	Yes	Yes
4	DL4j	It supports images and text, and it has same functionality as tensors	CNN, RNN, LSTM and others	Yes	Yes
5	Keras	It enables fast experiments and run-on top-of TensorFlow	CNN, RNN, LSTM and others	Yes	Yes
6	Chainer	It was written in Python, and it supports NVIDIA CUDA computation. It supports different network architectures	CNN and RNN	Yes	Yes
7	Microsoft CNTK	It is a MDL architecture to build a neural network as a series of computational steps via a direct graph	CNN and RNN	Yes	No
8	Scikit-learn	It focused on data mining and analysis domain	DL models	Yes	Yes
9	Amazon SageMaker Neo	It is a python library to train machine learning models and run them anywhere in the cloud	DL models	Yes	Yes
10	ConvNetJS	This library that supports training various MDL architectures through web browser without using any expensive server machine	DL models	Yes	Yes

features, allowing users to create very strong visual representations of real-world network topologies and performance. It also enables heterogeneous distributed computing, which can run on numerous GPUs at the same time and can run the model across many platforms automatically. TensorFlow has good performance because it was designed in C++.

PyTorch: PyTorch is the Python version of the torch, a Facebook-developed neural network framework that is optimized for GPU-accelerated deep neural network development. Unlike TensorFlow's static calculation graph, PyTorch's calculation graph is dynamic, and it may be altered in real-time to meet the needs of the calculation. The Facebook Artificial Intelligence Institute (FAIR) team launched PyTorch on GitHub in January 2017 and soon rose to the top of the GitHub hotlist. PyTorch drew a lot of attention right away after it was released, and it quickly became popular among researchers.

DL4j: Deeplearnig4j (DL4j) is implemented in Java, and hence, it is more efficient when compared to Python. The ND4J tensor library used by DL4j provides the capability to work with multi-dimensional arrays or tensors. This framework supports CPUs and GPUs. Deeplearnig4j works with images, CSV as well as plaintext.

Keras is an API, written in Python and run-on top of TensorFlow. It enables fast experimentation. It supports both CNNs and RNNs and runs on CPUs and GPUs.

5 Review of Eye Imaging Modalities with Diseases

5.1 Availability of Datasets

This section discusses a variety of retinal image datasets employed to develop and benchmark machine learning systems (MLS) in the context of diagnosing ophthalmic diabetes diseases. Table 6 shows the online available datasets for the recognition of eye-related diseases. Two-dimensional digital retinal pictures are captured using a fundus camera [11]. Many early-stage DR detection methods rely on databases with images captured by dilation of the pupil. Although several retinal image datasets are publicly available to benefit the research community in the development of early screening of DR systems, such as STARE [34], DRIVE [35], MESSIDOR [36, 37], HRF [38], Kaggle DR [39, 40], ODIR [41], DeepDR [42], UoA-DR [43], etc. The STARE [34] benchmark provides the structural details of the retina, comprised of 400 samples. A.txt file is available separately for the associated diagnoses. These text files contain lesion and regional annotations of all images.

The segmentation of blood vessels includes 40 vessel maps that have been manually labeled. Optic nerve head detection and ground truth are shown in detail in 80 images. The overall dataset contains a diverse range of disease and lesion types; however, the annotation system is a little difficult to navigate. Overall, the blood vessel segmentation is

Table 6 Publicly available database for early diagnosis of eye-related diseases

Cited references	Dataset name	Image size	Total images	Ground truth	Properties	Numerous experts
[34]	STARE	(700×960) pixels	2000	Image level	DR, glaucoma and AMD grading	Yes
[35]	DRIVE	(1100×1200) pixels	4500	Image level	Glaucoma grading Retinal vessel masks	Yes
[36, 37]	MESSIDOR	(850×900) pixels	2400	Image level	DR grading Lesion segmentation	Yes
[38]	HRF	(1000×1200) pixels	2300	Image level	DR and glaucoma grading	–
[39, 40]	Kaggle DR	(900×800) pixels	6000	Image level Lesion mask	DR and AMD grading	Yes
[41]	ODIR	(1200×1100) pixels	3500	Image level	DR and hypertension grading	–
[42]	DeepDR	(1450×2000) pixels	4000	Image level	DR grading	–
[43]	UoA-DR	(800×700) pixels	3300	Image level	DR grading	Yes
[44]	DIARETDB2	(1500×1152) pixels	240	Image level	DR and glaucoma grading	Yes
[45]	e-Ophtha	Variable size	500	Image level Lesion mask	DR grading	Yes
[46]	ROC	Variable size	100	Image level	DR and glaucoma grading	Yes
[47]	DDR	Variable size	1100	Image level Lesion segmentation	DR grading	Yes
[5]	FGADR	(900×800) pixels	2842	Image level	DR grading	–

superior because it detects smaller vessels more precisely. DRIVE [35] illustrates the retinal vessel extraction details. It contains 40 photographs in which 33 are identified as normal and 7 are abnormal. In addition, all the image pixels are labeled for vessels by the clinical expert. All retinographs are divided into two sets named train and test. MESSIDOR [36, 37] consists of 1200 images organized in 400-image sets. Each set is broken into 4 sets of 100 images and matching Excel spreadsheets are created for each subset. The images are classified as having a DR of 0 to 3 based on the presence and number of MAs, HEMs, and micro vascularization. DME grades are assigned on a scale of 0 to 3 based on the presence of hard exudates and their distance(s) from the macula. 800 images depicted dilated eyes and 400 images depicted undilated eyes. HRF [38] stands for high-resolution fundus, containing 45 samples, categorized in 3 sets of 15 samples for healthy, DR, and glaucoma. These samples are accompanied by binary vessel segmentation maps, which function as the golden truth segmentations. Kaggle DR [39] comprised over 10,000 retinal samples that are rated for DR on a scale of 0 to 4. The images are not in a standard format, having noise, non-uniformity, distortions, and exposure issues, and thus occupied 82 Gigabytes size. The training labels are listed in a.csv file.

Only databases that support early DR diagnosis, i.e., those that include MA masks and allow for DR grading, were examined for this investigation. ImageRet [44, 45] provides three versions of the DR Database (DIARETDB) collected from Kuopio university hospital, the most recent of which contains 89 color fundus images and ground-truth MA masks. The V0 version, on the other hand, has 130 photos accessible for grading. Color fundus photos captured with a Topcon NW 100, a Topcon NW200, or a Canon CR5-45NM camera are included in the Retinopathy Online Challenge (ROC) [46]. The photos have three resolutions: 786,576, 10,581,061, and 13,891,383. An e-Optha [47] database contains 148 color fundus photographs of various sizes gathered

over the OPHDIAT network in 2008 and 2009. It also provides 233 photos of a healthy fundus. Retina Check (RC) [48] includes 250 annotated photographs with a resolution of 25,951,944 that were collected during the study, which was handled by the Eindhoven University of Technology in the Netherlands. The Indian DR Image Dataset (IDRiD) [50], AGAR300 [49], and Database for DR (DDR) [51], all of which incorporate pixel-level MA annotation masks, has just been released. There are about 1000 MA segmentation masks in all that would benefit the research community.

Numerous open-source OCT benchmarks are available for AMD, DME and DR. Kaggle provides 84,495nretinal OCT images, rated as normal, DME, drusen, and choroidal neovascularization. All available samples are classified into training, testing, and validation sets. OCT dataset contains 500 images that are categorized into DR, normal, AMD, macular hole, and central serous retinopathy. All labels are given in JPEG format. Duke Farsiu SD-OCT benchmark is having 384 photographs in a size of 20 Gigabytes from AMD and central control subjects. These samples are accessible using MATLAB format.

5.1.1 Performance Analysis

Many statistical metrics were utilized in the past to detect eye-related diseases using imaging modalities as described in Table 7. Performance analysis indicators such as accuracy, sensitivity, and specificity are compared after the construction machine-learning model has been built (Table 7). These statistical metrics are just described in terms of definition as they have been used in the review process. Among other metrics, the area under the receiver operating characteristic curve (AUC) is also utilized in many studies for indicating the performance of classification systems. To highlight the performance, the AUC indicators are also suggestive of important objective evaluation. The AUC may quantify both the positive and negative sample accuracies at the same

Table 7 Performance metrics used in the past to recognize eye-related diseases, where TPR: true positive rate is use, false negative rate (FNR), false positive rate (FPR), true negative rate (TNR)

Performance measures	Description
True positive rate (TPR)	The features-class identified as a true class by the classification method
False positive rate (FPR)	The features-class as wrongly identified by the classification algorithm
Accuracy (Acc)	The ratio of classified features to the total features plus ratio of features classified as non-class
Sensitivity (SN)	The ratio of classified class to the actual class in ground truth. The same is called TPR metric
Specificity (SP)	The ratio of features class, which is classified as no accurate class in the ground truth (1-SP). It is also known as FPR metric
Receiver operating curve (ROC)	Mapping of fraction of the TPR to the FPR metrices or fraction of sensitivity to specificity on 2-dimensional plane
Area under the curve (AUC)	Covered area of the receiver operating characteristic curve for optimality
PPV	Likelihood of a feature classified as true features belong to actual class
F1-measure	Likelihood of a feature recognized as true feature belong to actual class

time. The higher the value of AUC and the better the model's performance, the closer the ROC curve is to the upper-left-hand corner.

5.2 Review Methodology for Eye-Related Imaging Modalities

Traditional machine learning (TML) and multilayer deep learning (MDL) allow computers to learn from a set of data and then make prediction decisions; there are two types of supervised and unsupervised learning procedures. In supervised learning, a machine-learning algorithm is taught to predict the intended outcome using input data previously categorized by humans, allowing it to tackle classification and regression issues. This method, however, is time-consuming because it necessitates manually labeling a large amount of data. Unsupervised learning, on the other hand, involves giving a machine unlabeled input data and allowing it to find structures and patterns from the set of objects without human intervention. The TML algorithms include decision tree, support vector machine (SVM), AdaBoost, k-nearest neighbor (KNN), and many more as visually in Fig. 3c. Despite their ability to perform well on small datasets, the TML network architecture makes them more prone to failing to attain convergence and overfitting the training dataset due to the manual feature selection procedure, which limits their application. Hence, the variants of deep-learning algorithms are developed to enhance the results of the prediction and

it performs automatically feature selections without manual intervention. In this paper, an analysis of Ophthalmic imaging modalities is used in many computerized tools by ophthalmologists for faster screening and diagnosis of eye-related diseases. There are numerous abnormalities caused by eye-related diseases as shown in Fig. 11 in the retina such as microaneurysms (MA), hard exudates, soft exudates, or cotton wool spots (CWS), hemorrhages (HEM), neovascularization (NV), and macular edema (DME). To detect and classify these lesions with respect to eye-related diseases, several TML and MDL algorithms are developed. Compared to traditional-machine learning TML, multilayer deep learning (MDL), speeds up diagnosis and improves precision and accuracy. As a result, this survey offers a thorough overview of the new technologies used in each phase of the retinal diagnosis process concerning various image modalities. This review aims to provide a more in-depth survey of the most significant components, types, and network architectures of MDL algorithms. The subsequent paragraphs have described those studies in detail to show the importance of MDL over TML algorithms by using various imaging modalities.

5.3 Review of Detection of Glaucoma

Glaucoma is a vision-threatening ailment diagnosed in a large population of adults worldwide, and it is expected to affect more than 118.2 million people by 2040 [18–20]. Glaucoma is the second most common cause of vision loss

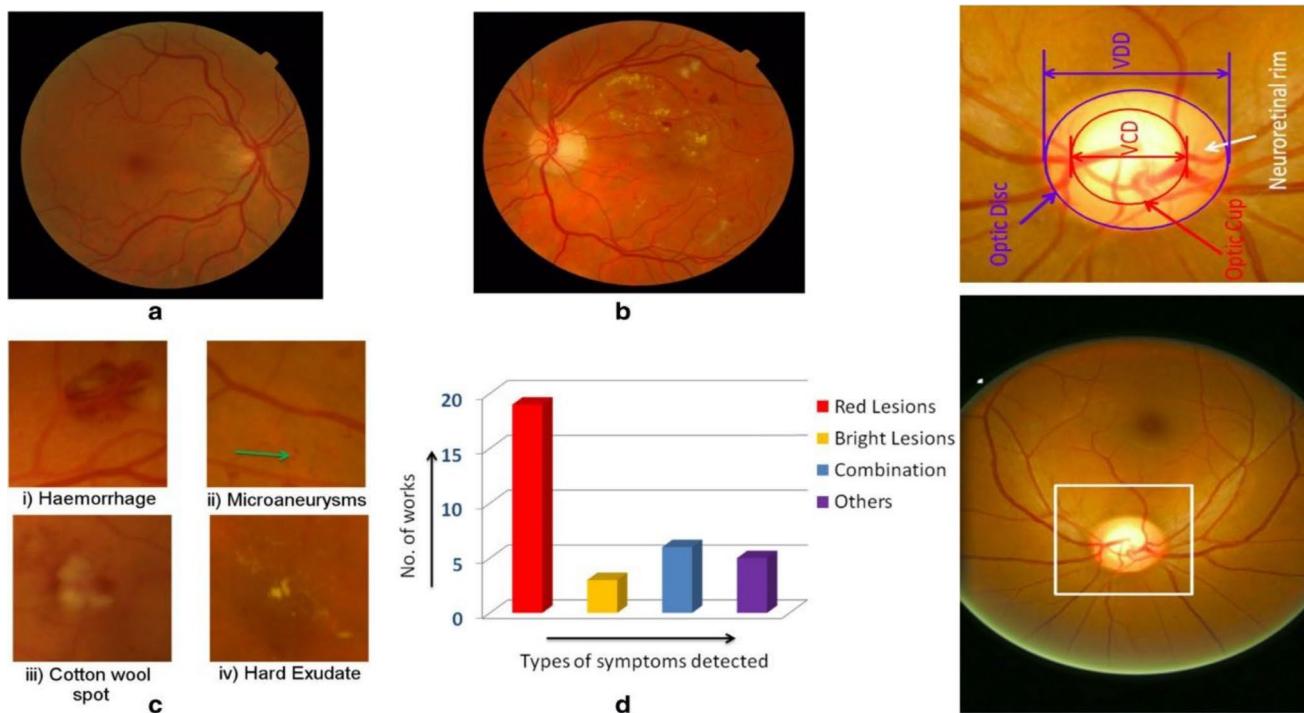


Fig. 11 Anatomical structure of different lesions presented in fundus images that visually represented different eye-related diseases

[75, 76]. Several studies have been conducted to determine the number of people affected by this chronic illness [21–23, 77–80]. Figure 1 shows the anatomy of the glaucoma eye. The main reason for glaucoma is the frequent thinning of the Retinal Nerve Fiber Layer (RNFL) [81]. Fundus images, OCT modality, and Angio-OCT images are commonly employed modalities for examining the optic nerve. The principal measure used to evaluate the Optic Nerve Head (ONH) is the ratio of the optic cup to the optic disc. When the optic nerve fibers gradually disappear, it can lead to glaucoma. Fundus images are intended to have features like the Cup to Disk Ratio for classifying the severity level of Glaucoma [82–84]. The thickness parameter measured by OCT was used to classify normal eyes and eyes with mild to severe glaucoma [85, 86]. This section discusses state-of-the-art Glaucoma traditional-based and deep learning-based methods using Color Fundus Image and Optical Coherence Tomography (OCT) images. A visual example of Glaucoma detection using fundus images is displayed in Fig. 12.

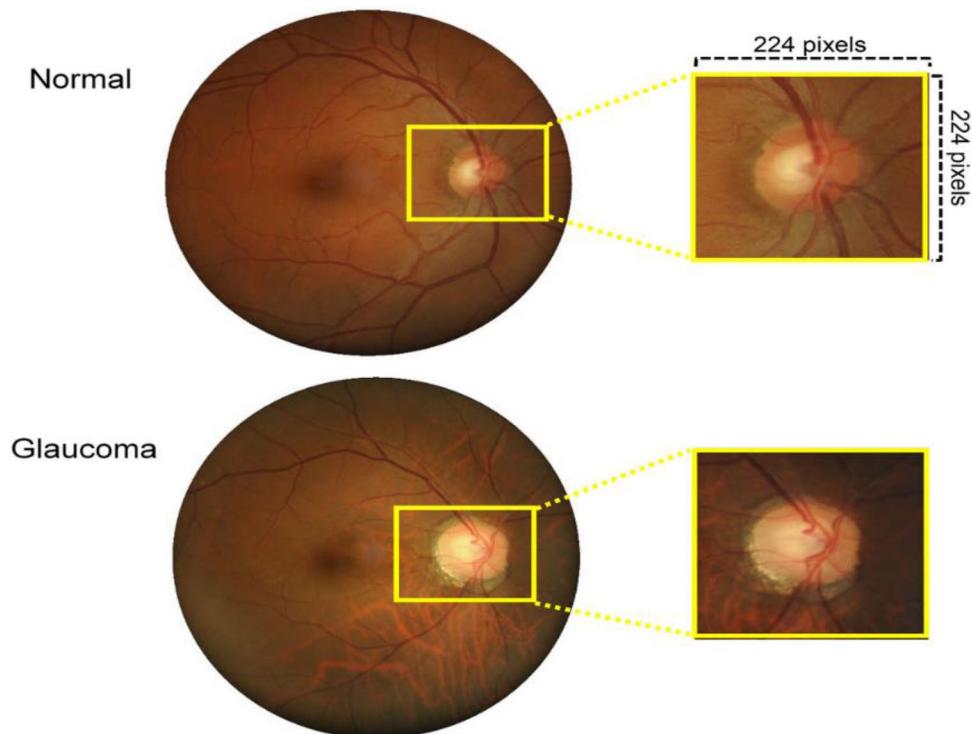
5.3.1 Traditional Methods for Glaucoma Detection

This subsection will study various handcrafted methods based on color fundus and OCT images for Glaucoma Detection. Traditional approaches are broadly divided into three main categories namely, (1) Segmentation approaches [19, 88–94], (2) Texture based approaches [91, 95–102], (3) Machine learning-based methods [103–107]. Retinal vessel identification and segmentation are a basic phase in

automated diagnosis and exciting medical image processing topics [88]. Many methods have been developed, but none is fast enough to be used in real automatic detection. Therefore, Bibiloni et al. [88] proposed a segmentation algorithm for real-time retinal segmentation based on the fuzzy morphological operation. They used the top-hat transformation method to enhance low contrast, noisy, blurred retinal image, and fuzzy mathematical morphology to handle the uncertainty in the images. The results reported from this work show that the proposed work shows better segmentation and enhances performance than other real-time systems. Another work in [89] proposed an automatic method for optic disc (OD) segmentation. For correct segmentation of OD and exact boundary detection, they employed morphology-based and snake-based active contour fitting. This work reported a 90% of the overlapping score, which shows the accuracy of the proposed art and is considered very helpful for automatic patient screening.

Pal et al. [92] proposed a method for OD segmentation based on morphological operation applied to a different channel of an image. They used the RIM-ONE dataset of 290 images and evaluated ground truth. This work reported average accuracy of 94%. Ahmad et al. [19] presented an image processing technique for glaucoma detection. Two major features are extracted from the fundus image: Cup to Disc Ratio (CDR) and Neuroretinal Rim (NRR) for glaucoma detection. The proposed work achieved an average accuracy of 97.5% with an average computational cost of 0.81 s. Finding the cup region helps find the cup-to-disk (CDR), which is an

Fig. 12 A visual example of difference between normal and Glaucoma retinal fundus image [87]



important feature of detecting disease. The study in [90] proposed an automatic cup region segmentation scheme based on gradient approach. The result of proposed work shows the effectiveness in segmentation of cup in the process of glaucoma assessment. The study in [95] proposed a method for detecting retinal nerve fiber layer (RNFL) based on texture features by creating a co-occurrence matrix and back propagation neural network as a classifier. This work achieved accuracy of 94.5% and indicate good performance. According to Arwa et al. [97], the low contrast of RNFL in fundus image lead to errors in glaucoma early detection. Therefore, they proposed a method [97] computer aided algorithm based on RNFL texture features that accurately detect the defect in very early stage of disease. Texture features are extracted and selected for classification from the segmented area after OD segmentation. They reported high accuracy using RIM-One datasets. Kirar et al. [96] proposed a hybrid and concatenating approach for computer aided diagnosis (CAD) of glaucoma using discrete wavelet transform (DWT) and empirical wavelet transform (EWT) from fundus image. Feature obtained using these methods are concatenated and fed to singular value decomposition to find out robust features. Afterward, SVM classifier is used using these robust features. The obtained results suggest that proposed method outperform the existing detection method.

Glaucoma causes vision loss, according to study [98], so Nirmala et al. [98] proposed a wavelet based contourlet transformation (WBCT) approach for glaucoma detection. They used adaptive gamma correction method with weighted distribution to enhance the image contrast and OD are segmented using Gabor filter. Coefficients are extracted from the segmented OD area using a WBCT, and then features are extracted from the computed coefficients. Naïve Bayes classifier is used to diagnose glaucoma. To enhance the performance of image classification system some studies focused on fusion based approaches [101]. Yin et.al [99] proposed a method for OD and OC segmentation using statistical model based on fusion of Hough transform and optimal channels selection for segmentation of the OD. The proposed work obtained mean absolute error of 0.10%. According to Khalid et al. [91], CDR is important feature for glaucoma detection in fundus images and its calculation is difficult because of ambiguous color pattern in optic cup and optic disk. Therefore, in [91] Fuzzy c-Mean (FCM) methods was proposed for the segmentation of optic cup and optic disk to measure the CDR. Several parameters are used to evaluate the performance of proposed work. This work obtained high value of sensitivity, specificity and accuracy using FCM with morphological operation. In [102] two segmenting schemes based on morphological reconstruction was proposed for optic disc and optic cup segmentation. This work achieved accuracy of 96.3% for optic disc segmentation and 99.1% accuracy for optic cup segmentation.

Machine learning has been used for the detection of glaucoma from last two decades [18]. In study [104] proposed machine learning approach to detect glaucoma in fundus image. First preprocessing method were employed to remove noise and enhance contrast. Then, the features are extracted using principal component analysis and then SVM are used for classification. This work reported accuracy of 96%, sensitivity of 100% and specificity of 92%. The study in [106] presents a new method for accurate detection of glaucoma from fundus image. They extract several features such as cup to disc ratio, rim to disc ratio, spatial and spectral features from fundus image. A multivariate medioids classifier is used for detection of glaucoma. The results showed that proposed work outperformed the existing work. Acharya et al. [103] proposed a method to extract feature from whole image for glaucoma classification. First, Adaptive histogram equalization was used to convert images to grayscale. Then, four filter banks were used to obtain textons. Afterwards, the features were extracted using local configuration. They considered KNN classifier the best one and obtained accuracy of 95.8%, specificity of 92.3% and sensitivity of 96.7%.

OCT is the most recent technique for measuring in-depth and multidimensional information about the retina. OCT is utilized to quantify the thickness and depth of the Retinal Nerve Fiber Layer (RNFL), the depth of the cup, and the macula, all of which are significant characteristics in the diagnosis of glaucoma [19]. Some studies [24, 108–112] focus on glaucoma detection using traditional image procession techniques. The study [108] proposed a layers segmentation algorithm to calculate CDR in OCT images. A comparison is made with clinical values and reported minor performance errors with OCT image analysis. In another study [110], image classification techniques are used for early glaucoma diagnosis based on Discrete wavelet transform (DWT) using OCT images, SVM algorithm and achieved an accuracy of 90.7%. Ramzan et al. [111] proposed a method to segment inner limiting membrane (ILM) and retinal pigment epithelium (RPE), aiming to diagnose glaucoma in OCT images. These ILM and RPE layers were used to compute cup and disc diameters and classify glaucomatous and normal from OCT images. This work reported an average error of mean Euclidean distance of 7.03%. The study in [112] used to extract CDR and two new features called cup depth and retinal thickness from OCT images. BPN and SVM are used as a classifier in this work and reported an accuracy of 96.9%. Table 8 demonstrates the traditional method to detect glaucoma using color fundus images and OCT images.

5.3.2 Deep Learning-Based Methods for Glaucoma Detection

This subsection focused on a deep learning-based study for glaucoma diagnosis schemes using fundus and OCT images.

Table 8 Traditional arts for glaucoma diagnosis using Color Fundus and OCT imaging

Cited	Methodology	Modality	Database	Performance measure
Bibiloni et al. [88]	A real-time method based on fuzzy morphological is introduced to segment the retinal images	Fundus image	DRIVE (20) STARE (20)	Accuracy = 93%, SE = 72%, SP = 97%, Precision = 78%, Accuracy = 95%, SE = 64%, SP = 97%, Precision = 73% Accuracy = 90%
Agarwal et al. [89]	Image Processing technique was used in this work such as smoothing filter for removal of blood vessels A morphological operation is used to segment the optic disc Active contour snake-based model utilized for optic disc boundary smoothing	Fundus image	Local Dataset (60)	
Pal et al. [92]	Morphological operations have been applied to combine different color channels, and the canny edge detection method was used to segment the optic disc	Fundus image	RIM-ONE (290)	Average accuracy = 94% and ratio = 83%
Ahmad et al. [19]	Glaucoma detection was presented based on extracting two features	Fundus image	DMED, FAU and MISSIDOR	Average accuracy = 97.5%
Khalid et al. [91]	Fuzzy c-Mean (FCM) and erosion and dilation used for the accurate segmentation of optic cup and disc	Fundus image	Dataset (27)	Optic Cup SE = 80%, SP = 99% and accuracy = 90% and Optic Disc SE = 87%, SP = 99% and accuracy = 95% Accuracy = 94.5%
Septiarini et al. [95]	A method is proposed to detect RNFL using texture features by making co-occurrence matrix and backpropagation classifiers	Fundus image	Dataset (40)	
Kirar et al. [96]	A hybrid and concatenating method for CAD of glaucoma are used using DWT and EWT	Fundus image	MIAG (505)	Accuracy = 88.7%, SE = 85.3%, SP = 92.1%
Nirmala et al. [98]	WBCT method for detection of glaucoma	Fundus image	FAU (NIL)	Accuracy = 89%
Arwa et al. [97]	Computer aided algorithm based on RNFL texture features	Fundus image	RIM-ONE (169)	Accuracy = 89.5%
Yin et al. [48]	Statistical model-based segmentation	Fundus image	ORIGA (650)	Mean Absolute Error = 0.10
Nugroho et al. [102]	Morphological reconstruction-based segmentation for OC and OD	Fundus image	Drishi-GS	Accuracy for optic disc = 96.3%, Accuracy for optic cup = 99.1%
Dey et al. [104]	Glaucoma detection was done using a machine learning algorithm called SVM	Fundus image	Database (100)	Accuracy = 96%, SE = 100%, SP = 92%
Akram et al. [106]	Glaucoma detection using optic disc localization, hybrid feature set, and classification methods	Fundus image	DRIVE (40) TARE (400) LARETDB (89) DRIONS-DB (110) HEI-MED (169) MESSIDOR (1200) HRF (45)	Accuracy = 100%, Accuracy = 97.5%, Accuracy = 100%, Accuracy = 98.9%, Accuracy = 100%, Accuracy = 100%, Accuracy = 96.4%, Accuracy = 95.8%, Accuracy = 95.8%
Acharya et al. [103]	A glaucoma detection method was proposed using texton and local configuration pattern features	Fundus image	Local Database (462)	Accuracy = 95.8%
Nithaya et al. [108]	The disc and cup diameter is measured by segmenting the retinal fiber and retinal epithelium layers	OCT image	Dataset (NIL)	OCT Diameter = 5.01 %

Table 8 (continued)

Cited	Methodology	Modality	Database	Performance measure
Ganesh et al. [112]	Two novel features, such as cup depth and retinal thickness and CDR, are extracted from OCT Image. Back Propagation and SVM are used as a classifier	OCT image	Dataset (100)	Accuracy = 96.9%
Ramzan et al. [111]	A segmentation algorithm is used to segment the inner limiting membrane	OCT image	Dataset (50)	Mean Euclidean Distance = 7.03
Rajan et al. [110]	An image classification system was introduced based on discrete wavelet transform (DWT) using OCT images for early-stage glaucoma detection	OCT image	Dataset (200)	Accuracy = 90.7%, Specificity = 89.7%, Sensitivity = 91.7%
Meng et al. [109]	This work uses local phase quantization (LPQ) to extract features. Information fusion and PCA are used to reduce the feature dimension	OCT image	Dataset (660)	Accuracy = 94.3%

Deep learning has shown impressive performance in medical imaging, especially in image classification and pattern classification tasks [25]. Chen used a deep learning framework for the first time to diagnosis glaucoma using fundus images [113]. In this study, a deep learning method is based on a convolutional neural network (CNN) to discriminate between glaucomatous and non-glaucomatous images. The proposed network architecture consists of 6 layers: four convolutional layers and two fully connected layers and used data augmentation and drop out a technique to improve the performance of glaucoma diagnosis. This work reported an AUC of 0.83 and 0.88 using two datasets. However, the proposed method has a false-positive detection rate.

Tan et al. [114] developed a method based on a Feed-Forward neural network (FNN) to reduce the False positive rate and improve the system's performance. Their work is based on two parts: background normalization and a 7-layer convolutional neural network. The proposed method obtained an average accuracy of 92.3%. However, usually, deep learning required a large dataset, which is challenging to design a deep learning model for glaucoma diagnosis with fewer data. Therefore, Chai et al. [115] proposed a Multi-branch neural network (MB-NN) combined with domain to extract essential features from retinal images for automatic glaucoma detection. The proposed work achieved an accuracy of 91%, sensitivity of 92%, and specificity of 90%. Another Study proposed a deep learning method based on a multi-model network termed G-EyeNet for glaucoma detection from retinal fundus image [116]. The framework consists of three parts which are (1) Encoder, (2) Decoder, and (3) Classification. They developed this method to minimize image reconstruction error and classification error. The proposed network is considered adequate on a small dataset, usually in the case of medical imaging. The proposed work obtained a ROC curve of 0.92%. The study in [117] used a deep residual network model named ResNet for glaucoma detection. They apply image augmentation to increase the data size and to improve the diagnostic system performance. High accuracy of 99% was reported using a dataset of 1364 images. Another work [118] proposed a feature learning method for glaucoma detection based on a deep convolutional neural network. They used multilayer perceptron instead of convolution layer for better extraction of robust features. This work reported ROC of 0.83% and 0.89% on two datasets.

The study [120] proposed CNN based transfer learning technique (AlexNet architecture) with a local binary pattern (LBP) data augmentation approach to detect glaucoma in fundus images. The image was separated into Red, Green, and Blue Channels in training and testing datasets. The proposed work obtained an accuracy of 98%, a sensitivity of 100%, and a specificity of 97% using RIM-ONE datasets. However, the method was evaluated on small datasets. Raghavendra et al. [120] proposed eighteen layer CNN

model to efficiently extract robust features from digital fundus images. These extracted features are used to classify into normal and glaucoma categories during testing. The proposed work achieved high accuracy of 98% on 1426 images. Few methods [93, 121] focus on optic cup and optic disc segmentation using deep learning method. The study in [93] proposed a segmentation method technique based on R-CNN for OD and OC segmentation from retinal fundus image. The proposed work obtained overlapping ratio of 93.1%. In reference [121] an ensemble based CNN architectures was proposed to segment OD and OC from fundus image. This work is reported the best method using small dataset of 50 images compared to other handcrafted feature methods. Experimental results shows that proposed work achieved superior performance to discriminate.

Several studies [26, 87, 120, 122–138, 140, 140–143] focus on deep learning-based glaucoma detection. The study in [124] proposed an algorithm for glaucoma detection based on transfer learning of convolution neural networks through fundus and OCT images. This work focusses on several findings such as disc RNFL deviation map, macular ganglion cell complex (GCC) thickness map, macular GCC deviation map, and RNFL thickness map. The study reported the improvement in diagnostic accuracy in glaucoma by achieving 0.96% AUC. Another study [122] focuses on classifying unsegmented OCT volumes into glaucomatous and normal using a 3D CNN network. They employed 3D convolution to extract depth information for glaucoma detection. The architecture of this work is composed of five 3D convolutional layers with a ReLU activation function. The proposed study obtained a higher AUC of 0.94%. Raja et al. [123] proposed a convolutional neural network (CNN) to automatically segmentation retinal layers. They used VGG-16 network architecture to extract features and classification of retinal layers pixels. The output map is integrated with the SoftMax layer for classification. The proposed work reported an accuracy of 94.6%. Table 9 shows the deep learning-based method to detect glaucoma using color fundus image and OCT. Recently, many other authors deployed [26, 87, 120, 124–138, 140, 140–143] deep learning models to detect glaucoma but still, those models required huge domain-expert knowledge to train and test the architecture. Table 9 shows the some of the deep learning schemes to detect glaucoma using color Fundus and OCT images.

5.4 Review of Detection of Age-Related Macular Degeneration (AMD)

Age-related Macular Degeneration (AMD) is an eye disease that can blur the sharp, central vision you need for activities such as reading and driving. The term "Age-related" means that it usually happens in older people. The term "Macular" means that it affects a part of your eye called the macula. The

macula is an essential part of the human eye responsible for the center loss to see the fine details. Figure 13 denotes the normal retina and macular degenerated retina with respect to different imaging modalities. AMD currently affects more than 6 million people alone in the UK. By 2040, the number of people in the world would have doubled. The disease is estimated to affect 288 million people worldwide [27, 28]. It can be mainly classified into dry or non-neovascular and wet neovascular AMD [29]. This section discusses state-of-the-art AMD traditional feature-based and deep learning-based methods using Color Fundus Image and Optical Coherence Tomography (OCT) images.

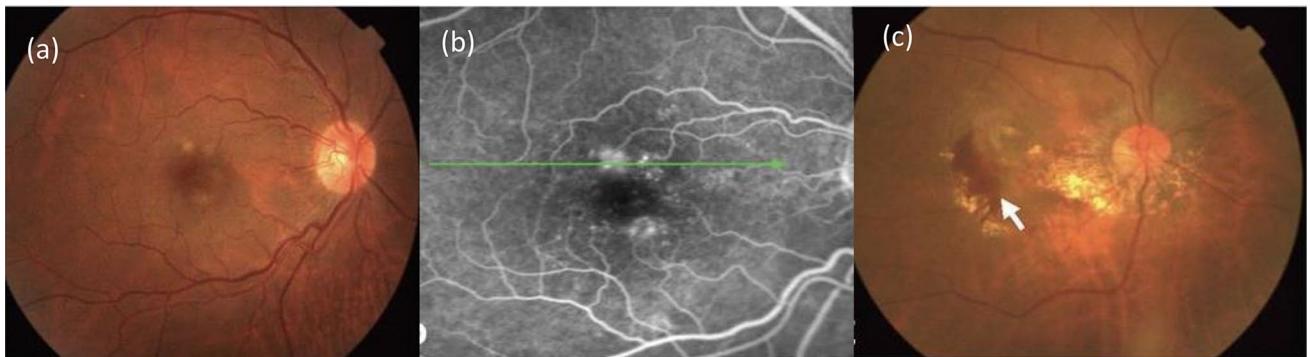
5.4.1 Traditional Methods for AMD Diagnosis

This subsection presents a comprehensive study of hand-designed feature approaches based on the AMD color fundus and OCT photographs. Automatic detection of AMD using color fundus images and OCT has been an active research area over the last decade. We will briefly review the recently traditional feature-based algorithms used to detect AMD using color fundus images [54, 55, 144–156] and OCT images [157–164]. Initially, the researchers proposed AMD diagnosis based on drusen in colored fundus imaging. For instance, a segmentation algorithm in [154] was proposed using the histogram-based adaptive local thresholding (HALT) to detect drusen within the fundus image. This work reported sensitivity (SE) and specificity (SP) of 96% for 22 samples. In that line, a mathematical model named amplitude modulation and frequency modulation (AM-FM) was implemented in [155] for the identification of pathological (drusen) structures. Their work obtained an SP and SE of 100% on different color spaces for structure classification. Another study [156] detects the drusen pathological features using fundus images. They used maximal region-based pixel intensity approach via HSV and RGB channels by achieving an SP and SE of 75%, respectively. Sultan et al. in [55] designed a method to automatically detect and segment drusens in color fundus images for AMD diagnosis. They used a color-based segmentation method based on color normalization to overcome the uneven/poor illumination and minimize color variability in a color image. The method obtained an accuracy of 96%. Drusen are of two types such as hard and soft drusen.

The work in [145] proposed a scheme to detect soft and hard drusen. They used the color, and Gabor features to extract and classify possible drusen pixels and obtained an accuracy of 96% using the STARE database. However, the methods were considered computationally expensive. Therefore. Kajal et al. in [146] introduced an automatic drusen detection method based on statistical analysis to improve the image quality and pixel-wise feature extraction. Recently, [148] proposed an automatic approach for drusen regions

Table 9 Deep learning schemes to detect glaucoma using Color Fundus and OCT

Cited	Methodology	Modality	Database	Performance measure
Chen et al. [113]	CNN for glaucoma	Fundus image	ORIGA (650) SCES (1722)	AUC=0.83, AUC=0.88
Tan et al. [114]	Feed-Forward Neural Network was used to detect glaucoma	Fundus image	Drive (40)	Average Accuracy=92.3%
Chai et al. [115]	MB-NN integrated with domain knowledge to extract robust features from fundus images	Fundus image	Dataset (2554)	Accuracy=91%, SE=92% SP=90%
Pal et al. [116]	A multi-model network called G-EyeNet was developed for glaucoma detection	Fundus image	Combined Dataset (400)	ROC of 0.92%
Asaoka et al. [117]	ResNet Model was introduced to diagnose glaucoma	Fundus image	Dataset (1364)	Accuracy=99%
Chen et al. [118]	DCNN used to diagnosis glaucoma	Fundus image	ORIGA (650), SCES (1676)	AUC=0.83%, AUC=0.89%
Maheshwari et al. [120]	AlexNet and LBP-based data augmentation	Fundus image	RIM-ONE	Accuracy=98%, SE=100%, SP=97%
Raghvendra et al. [120]	CNN for efficient feature extraction	Fundus image	Dataset (1426)	Accuracy=98.1%, SE=98.0%, SP=98.3%
Sun et al. [93]	A deep object detection network called Faster R-CNN for segmentation of OD and OC	Fundus image	ORIGA (650)	Overlapping ratio=R=93.1%
Zilly et al. [121]	Ensemble-based CNN models are used to segment OD and OC	Fundus image	DRISHTI-GS (50)	Jaccard Index of OD=0.042%, Jaccard Index of OC=0.039%
An et al. [124]	Transfer learning method was to classify glaucomatous and normal image	OCT image	Database (208)	AUC=0.96%
Maetschke et al. [122]	3D CNN network architecture used for glaucoma detection	OCT image	Database (1110)	AUC=0.94%
Raja et al. [123]	They used CNN (VGG-16 Net) to extract retinal layers from OCT images for glaucoma diagnosis	OCT image	Database (196)	Accuracy=94.6%, SE=94.6%, SP=94.1%

**Fig. 13** A visual example of Age-related Macular Degeneration (AMD), where Figure **a** represents fundus images, figure **b** shows OCT image, **c** AMD fundus image

segmentation to detect AMD using fundus images. Their work achieved an average SE of 93 and a dice classification coefficient of 73%. Acharaya et al. [147] presented a screening approach to detect dry and wet AMD based on the Pyramid of Histogram of oriented gradient (PHOG) and

non-linear feature. The study achieved an accuracy of 85%, SE of 87%, and SP of 80%. A comparison study of fundus image captured using multi-color (MC) and traditional color photography(CFP) were presented which focus on the feature ability to detect the early and late AMD [149].

The proposed study recommends that the performance of MC is effective compared to CFP by detecting early and late features for AMD diagnosis. Another research in [54] showed a conventional approach to automatically detect macular regions from fundus images using hybrid feature sets including shape\structure, color, and texture features to detect AMD accurately at an early stage. The proposed work achieved an accuracy of 95.4% and 92.3% on two available datasets. Several studies confirmed the detection and evaluation of atrophy of eyes, such as post hoc analysis of the comparison of age-related macular degeneration treatment (CATT) trials [165] and macular evaluation in the HARBOR Study for nvAMD [166]. The study in [152] evaluates the variability of mesopic and two-color dark-adapted (DA) fundus controlled perimetry (FCP) to assess the degree of impairment of rod and cone function in nvAMD. The mean sensitivity loss was more significant for DA cyan compared to DA red testing. To determine the classification system performance, Thee et al. in [153] compared the commonly used classification method for AMD and their ability to predict late AMD.

Imaging the eye using optical coherence tomography (OCT) is widely available to a large population due to its low cost and no side effects [167]. The first OCT image was published in 1993 and soon became a useful imaging modality in AMD [168]. We have briefly studied recent handcrafted-based methods to detect AMD using the OCT imaging tool. Tan et al. in [158] used traditional approaches such as histological analysis and nano-analytical measurement to confirm that calcified nodules in the retina were associated with AMD progression using multimodal imaging. It demonstrates that might help guide therapeutic method and outcome measures in patients at risk of progression to advanced AMD. Sirinavan et al. [169] proposed a fully automatic algorithm to detect AMD and diabetic macular edema (DME) using OCT image mode. They used multi-scale histograms of oriented gradient descriptors to extract, and SVM as a classifier with an accuracy of 86% was reported. Another classification-based method was presented in [170] using HOG and LBP to extract features and SVM to classify each image. Their work obtained a SE of 75% and SP of 87%. The study in [171] proposed a method to classify AMD and DME in OCT using image segmentation. The first retinal pigment epithelium (RPE) and drusen in the RPE layer were segmented with the segmentation approach. Afterward, RPE was used to find the retinal nerve fiber layer and detect the bubble blood area in the RFNL complex. Finally, binary classification was performed to classify both the disease and achieved an accuracy of 87%. In reference [172], an identification method was proposed to identify AMD using OCT images. Their methods include filtering segmentation method based on watershed algorithm and classification of a retinal image using SVM. This work achieved an accuracy

of 94%, sensitivity of 90%, and specificity of 93% and considering the effecting method compared to other existing segmentation techniques. Another study [167] focuses on extracting the RPE layer from SD-OCT images, which helps find the height of drusen in the RPE layer and classify it into normal and AMD. The proposed work achieved high accuracy of 96% using a large dataset of 2130 images. Table 10 highlights the salient features with a performance comparison of the developed traditional feature-based AMD detection approaches using fundus and OCT retinographs.

5.4.2 Deep Learning Methods for AMD Diagnosis

Due to the popularity of deep learning arts in a diverse range of image classification and computer vision tasks, the authors have developed deep learning algorithms [173–183] to screen AMD disease using fundus imaging. Felix et al. [173] proposed a deep learning method to study AMD severity using fundus images and predict age-related eye disease. They trained several deep learning models, but ensemble network architecture showed excellent performance in all 13 classes. Another study [174] also proposed an ensembling technique to combine two pre-trained models such as Inception-ResNet-V2 and Xception architecture to screen AMD using color fundus images. The proposed method obtained an accuracy of 95% and 86% on binary and multiclassification screening of AMD. The study in [178] also presents a transfer learning method based on pre-trained VGG-16 network architecture to classify AMD in two and four different classes using color fundus images. The proposed schemes reported accuracy of 92% with two levels and 83% with four classes. Peng et al. [179] proposed the DeepSeeNet deep learning model for automatic classification of AMD severity from color fundus images. This work introduces to computer severity score (0–5) for AMD. The proposed work shows better performance than retinal specialists, which enhances clinical decision-making in patients with AMD. To improve AMD patients' clinical decision-making, Peng et al. [183] introduced another research study based on the prediction risk of late AMD detection using deep learning. The proposed work report higher predictive accuracy than existing clinical standards. Bhuiyan et al. [177] presented a method that predicts late dry and wet AMD based on the artificial intelligence algorithm. The proposed model obtained a remarkable performance to identify early and late AMD with 99% and 84% classification on two different datasets. Some research studies focus on the prediction of neovascular AMD.

The study in [182] introduced a method to predict neovascular AMD (nvAMD) from color fundus images using transfer learning architecture based on Inception-V3. The proposed work shows high sensitivity compared to manual grades from 9-step scales and 4-steps scales for the

Table 10 A summary of handcrafted approaches based on fundus and OCT imaging analysis for AMD

Cited	Methodology	Image modality	Database	Performance measure
Rapantzikos et al. [154]	A segmentation method based on the HALT approach to detect drusen from fundus image	Color Fundus image	Own (23)	SE = 96%, SP = 96%
Barriga et al. [155]	A mathematical method based on AM-FM were employed to generate multi-feature for pathological classification	Color Fundus image	AREDS (Not Reported)	SP = 100%, SE = 100%
Liang et al. [156]	This work used a maximal region-based pixel intensity approach via HSV and RGB channels to detect drusen	Color Fundus Image	Own (16)	SP = 75%, SE = 75%
Waseem et al. [145]	Background illumination and illumination correction was performed at the preprocessing stage Color and Gabor feature used to extract feature of drusen	Color Fundus image	STARE (40)	Accuracy = 96%
Mittal et al. [146]	KNN classifier is used for classification, and optical disk remove to get drusen image	Color Fundus image	STARE and ARIA (36)	Visual Results were presented
Acharya et al. [147]	Drusen is detected automatically to reduce the computational cost Statistical analysis is employed to preprocess the image Pixel-based feature extraction to extract feature	Color Fundus image	Own (945)	Accuracy = 85%, SE = 87%, SP = 80%
Kim et al. [148]	PHOG and the non-linear feature was used to detect dry and wet AMD using fundus image First, fundus image change to the green channel ROI extract for macular region based on optical disk along with another preprocessing step to remove the vessel and detect candidate group Finally, Renyi entropy threshold to detect drusen	Color Fundus image	Own (30)	The average SE was 93.37% (80.95%~100%), and the average Dice Similarity Coefficient (DSC) was 0.73 (0.3~0.98)
Khalid et al. [54]	First, the Macula region was detected from input images Feature extraction was perform based on shape, structure, color, and texture An SVM classification method is used to classify abnormal macula from the normal macula	Color Fundus image	STARE (462) and AFIO (65)	SP = 97%, SE = 95%, Accuracy = 95%, SP = 93%, SE = 92%, Accuracy = 92%
Mohaimin et al. [55]	Color normalization technique was presented to remove non-uniform illumination and inter-subject variability	Color Fundus image	STARE (60) and ARIA (12)	SP = 97% SE = 96%, Accuracy = 96%, SP = 97%, SE = 87%, Accuracy = 94%
Graham et al. [149]	A Comparison study was presented for fundus image obtained by MC and CFP	Color Fundus image	Own (105)	The MC method shows effective performance compared to CFP for SP, SE, and accuracy

Table 10 (continued)

Cited	Methodology	Image modality	Database	Performance measure
Pfau et al. [152]	To determine the retest variability of mesopic and dark-adapted two-color FCP in CNV secondary to AMD to define the clinical significance Investigated the predictive value of “patient-reliability indices” To describe the functional loss for choroid neovascularization	Color Fundus image	Own (120)	Mean sensitivity loss [mean \pm SD] – 2.63 \pm 3.87 dB P, 0.001
Sirinavan et al. [169]	A new method was developed to detect AMD and DME using OCT images They used the multi-scale HOG method to extract features and SVM to classify each image	OCT	SD-OCT (36)	Accuracy = 87%
Alsaih et al. [170]	Automatic detection of DME on OCT method was proposed Features were extracted using the HOG and LBP method and SVM classifier	OCT	Dataset (32)	SE = 75%, SP = 87%
Sugrnk et al. [171]	A classification method was proposed two disease DME and AMD using image segmentation method	OCT	Dataset (16)	Accuracy = 87%
Hani et al. [172]	The computerized method based on morphological parameters and a classification scheme was proposed to diagnosis AMD. Watershed algorithm is used to segment the images and SVM used to classify the images	OCT	Dataset (15)	Accuracy = 94%, SE = 90%, SP = 93%
Thomas et al. [167]	Frist, contrast enhancement based on adaptive noise reduction method use to remove speckle elimination from image Pixel grouping and iterative elimination are used to obtain RPE layer followed by polynomial fitting and drusen removal to achieve baseline estimate Finally, based on drusen height images were classified normal and AMD	OCT	Dataset (2130)	Accuracy = 96%

progression of none/early/intermediate AMD to nvAMD at 1-year. In another study, Keel et al. [180] proposed deep learning architecture for the detection of nvAMD from color fundus images. The work includes three deep learning models based on inceptionV3 architecture to classify AMD and assess the image quality and visual macular region assessment. All the above-mentioned deep learning methods are supervised neural, requiring labor-intensive annotation and training to specific tasks. Moreover, most classification approaches are defined based on human rubrics, which are prone to bias. Therefore, to address these problems, Yellapragada et al. [175] developed an unsupervised network with Non-Parametric Instance Discrimination (NPID) to grade AMD using color fundus image from Age-Related Eyes Disease Study(AREDS). The proposed work outperformed the supervised network that was trained with both fundus images and genotype data. Also, achieve good predictive performance without human-generative labels.

Few recent deep learning studies [184–188] were published to automatically detect AMD using OCT images. Kaymak et al. [185] proposed a deep learning approach based on AlexNet architecture to automatically detect AMD from OCT images. They classified the AMD into three categories such as healthy, dry AMD, and wet AMD. The study reported accuracy of 97%, Sensitivity of 99%, and specificity of 98% on multi-classification tasks. Serener et al. [186] used ResNet architecture to classify dry and wet AMD on OCT images. The proposed work obtained an accuracy of 98%, a sensitivity of 95%, and a specificity of 99%. Motozawa et al. [187] present a CNN-based deep learning scheme to classify normal and AMD from OCT images. They achieved an accuracy of 99%, specificity of 91%, and sensitivity of 100%. Das et al. [188] introduced a denoising method based on a generative adversarial network to improve OCT images' quality for AMD detection. The results show that the proposed method effectively computes cost and enhances the system performance by achieving an accuracy of 96%. A summary of the employed deep learning approaches for AMD diagnosis is mentioned in Table 11.

5.5 Review of Diabetic Hypertension Retinopathy (DHR)

Hypertension is a common global disease that affects 9.4 million people each year [189] and is expected to increase in the US. It is caused by high blood pressure [190], and it is the most common type of eye disease that has recently spread across the world. Because of elevated arterial pressure, diabetic hypertensive retinopathy (DHR) harms the retinal and blood vessels. Many human organs, such as retinopathy, heart disease, and renal insufficiency, are impaired by hypertension. Retinopathy is the most common cause of cardiovascular disease among all hypertension, leading to

death [191]. Figure 11 shows the significant signs of HR disease, such as retinal hemorrhage, Cotton wool spots, hemorrhages, and microaneurysms (HM). Their several systems were developed to detect DHR, which is related to high blood pressure, automatically. A retinal fundus image and OCT image analysis were used in the recent past by research scientists to detect DHR disease automatically. This section surveyed several HR studies and divided them into two categories: standard feature and deep learning-based methods.

Categorizing HR into different grades is a difficult task for computerized diagnostic systems, and to the best of our knowledge, no research has assessed HR stages such as mild (MID-HR), moderate (MOD-HR), and malignant (MAL-HR) (MLG-HR). Table 12 lists the stages with diagnostic signs, while Fig. 14 displays the stages visually. The mild stage of HR shows the early stage of HR, as seen Table 1. Narrowing of arteries, copper, and silver wirings (Congealing of walls of retinal arterioles), arteries /veins nipping, and increased vascular tree tortuosity are all symptoms of mild stage HR. Many retinopathy symptoms, including mild stage defects, occur on the fundus retinal images at the intermediate stage of HR, which ophthalmologists can easily see. While Cotton wool spots, rough exudates, and flame-shaped hemorrhages are some of the anomalies seen in moderate HR stage. Beside the mild and moderate stage anomalies, the last stage of HR (malignant) is characterized by the appearance of Papilledema (swelling of optic disc). Figure 15 gives visual examples of the four different HR levels.

5.5.1 Traditional Feature-Based Techniques for DHR or HR Diagnosis

The recent discussion about the hand-engineered feature-based arts for HR detection using the fundus and OCT images is highlighted here. These methods were initially focused on recognizing exciting features, and then a machine learning classifier was proposed. In [193] a tortuosity index feature evaluation study was presented to detect HR using fundus images. The study in [194] suggested an automatic method based on the arteriolar-to-venular diameter ratio measurement (AVR). A sensitivity of 87% was recorded, and 93% were reported to be correctly classified into arteries or veins. The authors in [195] used the arteriovenous ratio and papilledema signs to detect and grade HR using retinal fundus images. Their work obtained an average accuracy of 95%, 95%, and 98% for AVR using three datasets and obtained an average accuracy of 95% and 97% for possible signs of papilledema. Manikis et al. [196] provided an AVR computation technique focused on area-based classification and multi-scale filtering, showing 93.70% and 93.10 accuracies using DRIVE and STARE databases.

A supervised classifier [197] performs initial segmentation and refinement measures to detect hemorrhage

Table 11 Deep learning methods for AMD using color fundus and OCT imaging modalities

Cited	Methodology	Modality	Database	Performance measure
Grassmann et al. [173]	First, the color fundus image was preprocessed The preprocessed image was used to train multiple convolutional neural nets The ensemble-based method was developed to train a random forest algorithm using the result of the CNN model Finally, a final model used to predict age-related disease	Color Fundus image	AREDS (120,656) KORA (5555)	Accuracy = 92.1% and 94%
Govindaiah et al. [174]	The Ensembling learning approach was employed with two pre-trained model called Inception-ResNetV2 and Xception model to detect AMD in color fundus images	Color Fundus image	AREDS (150,000)	Accuracy = 95% and 86% for binary and multi-class classification
Govindaiah et al. [178]	A Pre-trained deep learning model called VGG16 Neural network model to detect AMD in the color fundus image	Color Fundus image	AREDS (150,000)	Accuracy = 92% and 83% on binary and multi-class classification
Peng et al. [179]	DeepSeeNet deep learning model was designed to classify AMD patient using bilateral CFP	Color Fundus image	AREDS (59,302)	Accuracy = 67%, Kappa = 55%. Detection of Drusen = 94%, Late AMD = 97%
Peng et al. [183]	A deep CNN network model was used to extract multiple highly discriminative deep features A Cox, the proportional hazards model, was adopted to predict the probability of progression to late AMD using deep feature	Color Fundus image	AREDS (80,000)	5 Years C-Statistic 86.4 (95% confidence interval 86.2–86.6)
Bhuiyan et al. [177]	An ensembling deep learning method to classify late AMD and early AMD	Color Fundus image	AREDS (116,875) NAT-2	Accuracy = 99% SE = 95%, SP = 95%, Precision = 95% and Accuracy = 84% SE = 90%, SP = 81%, Precision = 71%
Babenko et al. [182]	A logistic tree model was used to predict the risk of progression for late AMD	Color Fundus image	AREDS (12,000)	SE = 80% SP = 80% AUC = 88%
Keel et al. [180]	First, data preprocessing was employed to ensure that the image is de-identified Inception V3 model was used to predict the neovascular AMD A tenfold cross-validation approach was employed to evaluate the deep learning model	Color Fundus image	Own (86,162)	AUC = 96%, SE = 100%, SP = 93%
Yellapragada et al. [175]	An automatic image processing method was used to normalize the image InceptionV3 Architecture was employed to classify AMD images, assess image quality, and assessment of the visual macular region	Color Fundus image	AREDS (100,848)	Unbalanced Data Accuracy = 94%, Balanced Data Accuracy = 82%

Table 11 (continued)

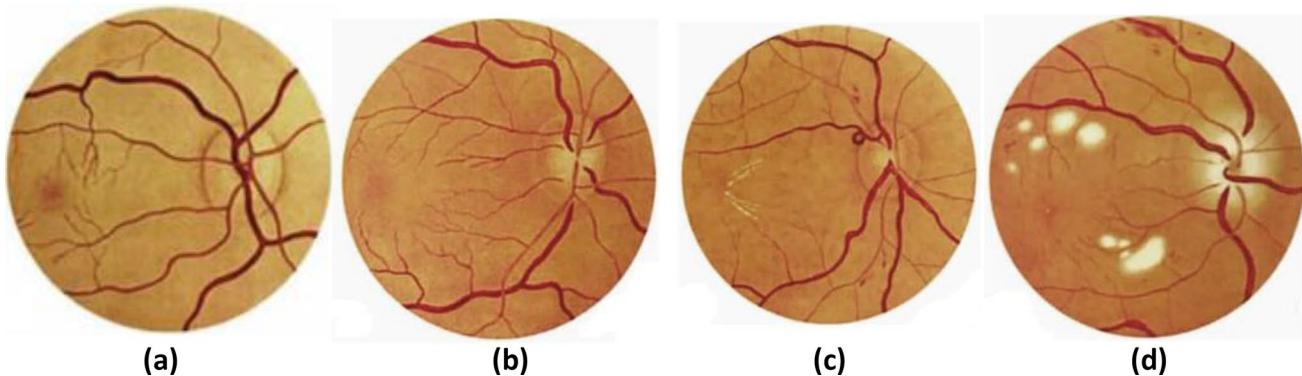
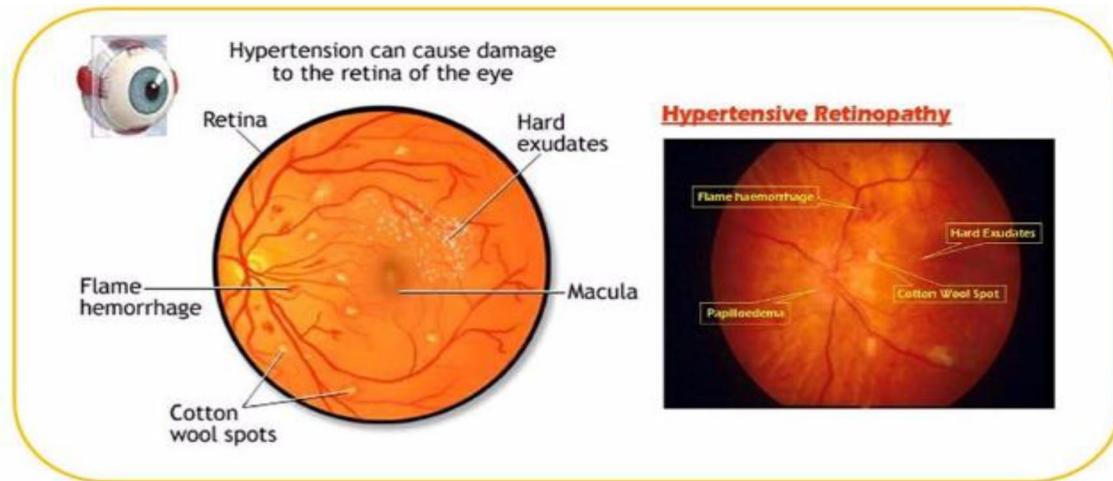
Cited	Methodology	Modality	Database	Performance measure
Kaymak et al. [185]	AlexNet deep learning architecture was used to detect the AMD	OCT	Kermany Database (83,848)	Accuracy = 97%, SE = 99%, SP = 99%
Serener et al. [186]	ResNet deep learning architecture was employed to detect the dry and wet AMD	OCT	Kermany Database (83,848)	Accuracy = 98%, SE = 95%, SP = 99%
Motozawa et al. [187]	A CNN deep learning model was used to classify AMD from healthy OCT images	OCT	Own (1616)	Accuracy = 99%, SE = 91%, SP = 100%
Das et al. [188]	Unsupervised generative adversarial network-based on super-resolution framework was proposed to improve the low-quality OCT images	OCT	Own (384 noisy OCT images)	Accuracy = 96%, SE = 96%, SP = 96%

lesions and AVR. Neha et al. [198] proposed the method detection based on AVR estimation using fundus images. They extract gray level and moment-based features from a preprocessed image to detect pixels belong to blood vessel class or not. Afterward, Intensity variation and color information are used to evaluate and classify vessels as arteries or vein vessel width estimation methods used to measure AVR. Cavallari et al. [199] proposed a method for analyzing two conditions known to be associated with retinal vessel change, namely HR Cerebral Autosomal Dominant Arteriopathy with Subcortical Infarcts and Leukoencephalopathy (CADASIL). This technique is used to detect and quantify retinal vessel abnormalities. The methodology represented a straightforward method for distinguishing disease conditions based on morphological changes in retinal vessels. The authors used a Gabor filter bank with a threshold mechanism in [200] for segmenting HR lesions. They recorded that the cotton wool spots (CWS) were the most significant clinical indicators for determining hypertensive retinopathy. Their method reported 82.21% of SE and 82.38% of PPV on the local dataset. The study in [203] presents an HR detection scheme based on vessel measurement and textural features. Their result showed a sensitivity of 90% and specificity of 67%. Another study in [195] based on diagnosis and grading of HR is presented using AVR. The proposed work is based on three modules, i.e., central component extraction, artery/vein classification, and finally, AVR classification and grading of HR. This work achieved an average accuracy of 95%, 96%, and 98% for three dataset images: INSPIRE-avr, VICA, and AVRDB local database.

Few traditional feature extraction-based studies are found to detect HR using OCT images. The research in [201] noticed that hypertension affects both the structure and function of the microvasculature. They proposed a study about the impact of hypertension on retinal capillary microvasculature using optical tomography angiography (OCTA). This work estimates the glomerular filtration rate using the creatinine equation and retinal capillary density, computed with OCT-A (AngioVue). Linear regression was used to associate the risk factor with capillary density. The result indicates that retinal capillary density decreases with high blood pressure. Another scientific report [202] focuses on the impact of risk on the choriocapillaris using OCTA with systematic hypertension. They used the same method of [201] to investigate the risk factor. The result concludes that the high blood pressure appeared to increase the choriocapillaris' blood flow, which should be considered to study eye disease in hypertension using OCTA images. The study in [203] presents the Utility of OCTA in detecting retinal vascular damage caused by blood pressure. They reported a sensitivity of 78% and specificity of 66, respectively. Dihao et al. [204] proposed an observational study to

Table 12 Definition of diabetic hypertensive retinopathy (DHR) stages along with their diagnostic symbols

No.	Stages of retinographics	HR-related lesions features	Diagnostic
1	Normal retinograph (NOM-RETINO)	Absence of HR-related lesions	None
2	MILD-Hypertensive Retinopathy (MID-HR)	This MID-HR is associated with one or more DR-related lesions such as hemorrhage, microaneurysms, cotton wool spot and hard exudates	The MID-HR eye disease is linked with heart disease and mortality of cardiovascular
3	Moderate-Hypertensive Retinopathy (MOD-HR)	This MOD-HR indicate the macular edema and swelling of optic disc	The moderate HR is related to heart failure, renal dysfunction and mortality of cardiovascular
4	Malignant- Hypertensive Retinopathy (MLG-HR)	This MLG-HR stage is having high rate of mortality with arteriolar narrowing, focal arteriolar narrowing, arteriovenous nicking, arteriolar wall opacity (silver wiring)	The malignant HR is associated with mortality

**Fig. 14** A visual example of mild (a), moderate (b), severe (c), and malignant (d) hypertensive retinopathy fundus**Fig. 15** Effect of hypertension on the retina of eye [192]

assess vascular changes in the macular and optic nerve head in hypertensive patients. Their study concluded that superficial plexus, FAZ area, capillary, and inner retinal thickness changed in patients without hypertension

retinopathy. Table 13 illustrates the traditional feature-based HR detection methods using color fundus and OCT images.

5.5.2 Deep Learning-Based Techniques for HR Detection

This subsection has discussed several recent deep learning-based methods to diagnose HR using fundus and OCT images. In [207], a multilayer deep learning model (MDL) for HR recognition was presented. They used 32 resized batches of the grayscale converted fundus to train a CNN. The CNN performance was either HR or non-HR. The accuracy of the identification was 98.6%. The study in [208] developed a deep neural network and Random Boltzmann Machine (RBM) algorithm to compute arterial blood vessels' change. They defined features for the deep-learning algorithm using the AVR and the OD area and found significantly higher accuracy. In [114], a deep learning algorithm based on CNN automatically segmented the optic disc, fovea, and blood vessel simultaneously. The architecture of the proposed network was made up of 7-layers. This work reported an accuracy of 92% on the DRIVE dataset. AlBadawi et al. [209] presented an encoder-decoder-based, fully connected deep neural network for pixel-level classification of retinal vascular into arterioles and venules by producing a detection rate of 93%. Two other studies [210, 211] also show retinal vascular classification into arterioles and venules. The study in [212] used a deep learning model to detect exudates in retinal images. The CNN architecture is composed of 4 convolutional layers and pooling layers. Their work reported the same sensitivity, positive predictive value (PPV), and F-score of 77%.

There is no deep learning-based detection method found for HR using OCT imaging modality in the literature. Table 14 shows the comparative analysis of the deep learning-based HR detection schemes using color fundus images.

5.6 Review of Diabetic Retinopathy (DR)

Diabetes mellitus (DM) has reached epidemic proportions globally, affecting 425 million people, and is anticipated to affect 642 million people by 2040 [30]. DM's leading microvascular complication is diabetic retinopathy (DR), an ultimate cause of vision loss in adulthood [213]. According to the clinical study [31], DR is mainly categorized into two types named nonproliferative NPDR and Proliferative PDR. NPDR is a painless type of DR and is often undertaken by researchers to develop automated screening systems for DR diagnosis. In contrast, PDR is a painful and severe form of DR [214]. Microaneurysm (MA), hemorrhage (HEM), exudate (EX), and cotton wool spot (CWS) are the essential clinical lesions of DR in the retinal photograph. The development of automated arts for early screening of DR relies entirely on the correct localization and segmentation of DR lesions [32, 33, 215]. Accordingly, a typical DR eye structure with lesions is shown in Fig. 16.

The purpose of this section is to survey the most recent Computer-Aided Diagnosis (CAD) studies developed for the early detection of DR and severity classification of DR using computational image processing and deep learning methods. Additionally, the authors focused on the Color Fundus (CF) and Optical Coherence Tomography (OCT) imaging modalities, clinically adopted by both ophthalmologists and scientific researchers for the early analysis of abnormal (DR) features.

5.6.1 Image Processing Based DR Detection Using Fundus and OCT Images

This section will focus on traditional approaches used to detect DR using Fundus Images and OCT Images. Walter et al. [217] proposed an image processing scheme to diagnosis DR using fundus images. They detect exudates using high grey level variation and their contour using morphological reconstruction techniques. And the optic disc was detected using the morphological filtering method and watershed transformation. This work obtained mean sensitivity of 92.8% and a mean predictive value of 92.4%. However, the proposed work was not efficient in determining all the exudates. It was also evaluated on small datasets. Sopharak et al. [218] used a set of optimally adjusted morphological operators to detect exudates from fundus images of DR patients, which can help ophthalmologists. This work reported sensitivity and specificity of 80% and 99% for exudate detection. Their proposed method still needs to improve the proposed work performance in terms of sensitivity results. Also, the proposed work could not detect soft and hard exudate, and some essential features were missing in the proposed work. Quellec et al. [219] proposed a multi-instance learning framework for automatic screening of DR. Clinicians provide the datasets of relevant and irrelevant images for image classifier to train. Similar patterns are searched in new images after training to identify them as relevant or irrelevant. Hence, no formal segmentation method was required in the proposed work. Two datasets MESSIDOR and e-OPHTHA were used for the proposed framework and achieved high performance on both datasets.

However, the training procedure of the proposed work was prolonged because of many parameters that need to be trained. Amel et al. [220] proposed an improved method for hard exudates detection approach for DR diagnosis. They used a k-means clustering algorithm along with mathematical morphology to detect hard exudates from retinal fundus images. Accuracy of 99.7% and Sensitivity of 95.95 was achieved using the proposed method. Michael et al. [221] studied the retinal fundus images for detection of DR. This work reported confidence Interval for referable DR is 95%. Another study [222] suggested a method to differentiate DR samples from standard fundus images. Fundus images

Table 13 Traditional approaches to detecting HR using color fundus and OCT images

Cited	Methodology	Modality	Datasets	Performance measure
Akbar et al. [195]	An automatic detection and grading extraction was proposed using fundus image They used AVR and papilledema as a feature and SVM as an extraction	Color Fundus image	INSPIRE-AVR (40), VICAVR (58), AVRDB (100), STARE (100), AVRDB (100),	Average Accuracy of 95% and 98% and Average Accuracy of 95% and 97%
Narasimhan et al. [198]	The method proposed to detect HR based on AVR estimation using fundus image Gabor filter bank to enhance the candidate regions and then thresholding technique	Color Fundus image	VICAVR (101)	AVR ratio=0.2 to 0.49 for HR patient and AVR ratio=0.6–0.7 in normal case
Goswami et al. [200]	Image preprocessing and segmentation algorithms utilized to measure the AVR and the OD region	Color Fundus image	Dataset (15)	Accuracy=97%
Holm et al. [205]	Image normalization and enhancement Region of interest detection of the optic disc Retinal vascular segmentation and measurement of vessels and tortuosity Color feature extraction and AVR computation Calculation of mean red intensity and saturation values for all arteries Computation of AM-FM feature for images Feature classification using HR and non-HR using linear regression	Color Fundus	Dataset (39)	Accuracy = 95%
Agurto et al. [206]	Automatic detection and grading of HR were proposed Vascular map and set of hybrid features for Artery/Vein classification	Color Fundus image	Dataset (74)	AUC = 84%, SE = 90%, SP = 67%
Akbar et al. [195]	This work estimates the glomerular filtration rate using the creatinine equation, and retinal capillary density was computed with OCT-A (AngioVue). Linear regression used to associate the risk factor with capillary density	OCT	INSPIRE-AVR, VICAVR, AVRDB	Average Accuracy = 95%, Average Accuracy = 96%, Average Accuracy = 98%
Chua et al. [201]	This study also estimates the glomerular filtration rate using creatinine equation, and retinal capillary density was computed with OCT-A (AngioVue). Linear regression used to associate the risk factor with capillary density	OCT	Dataset (163)	Probability value $p > 0.05$
Chua et al. [202]	The study was conducted to examine macular and peripapillary perfusion density and patient using OCTA	OCT	Dataset (164)	Probability value $p \leq 0.10$
Pascual-Prieto et al. [203]	The study was conducted to examine macular and peripapillary perfusion density and patient using OCTA	OCT	Dataset (73)	Probability value $p \leq 0.047$

Table 13 (continued)

Cited	Methodology	Modality	Datasets	Performance measure
Hua et al. [204]	Fifty-seven chronic hypertensive patients without hypertensive retinopathy (22 men and 35 women) and 40 healthy volunteers (17 men and 23 women), ranging in age from 60 to 70 years, were included in this study Vessel density (VD), foveal avascular zone (FAZ) area, choriocapillaris flow area, ONH capillary density, retinal nerve fiber layer (RNFL) thickness, and demographic information were compared among the groups	OCT	Dataset (97)	Probability value $p=0.96, 0.76$ and 0.64 for different age group

were processed to extract abnormal signs, such as hard exudates, area of blood vessels, bifurcation points, textures, and entropies.

In this work, 13 features were fed to the machine learning classifier algorithm named Probabilistic Neural Network (PNN). The parameter of PNN is tuned using a genetic algorithm and particle swarm optimization, and an average accuracy of 96.1% was achieved. The study reports that the proposed work was a faster DR diagnosis scheme during a mass screening of typical/DR images. However, the detection of several other significant lesions such as microaneurysms, cotton-wool spots, and hemorrhages were missing, which can improve the performance of the method. Zhang et al. [223] proposed a detection method of exudates from color retinal images for mass screening of DR. They pre-processed the image to normalize, denoise, and detect the artifact and reflections in the processed image. A segmentation algorithm based on mathematical morphology was proposed to characterize the classical and contextual features. Afterward, random forest algorithms were employed to detect the exudates among candidates and obtained an AUC of 95% on e-ophtha EX database. However, there are still some parameters left, mostly corresponding to thresholds. A hybrid classifier developed by Akram et al. [224] to detect retinal lesions by following the pre-processing step and achieved 98.1% of AUC value. However, the method is considered complex because of using a hybrid classifier for detection.

Several traditional methods [224–233] were recently developed to detect DR using OCT imaging modality. Colomer et al. [225] used local binary patterns and granulometric profiles to extract morphological and texture information from the retinal image for DR diagnosis. This information then fed into a classifier to classify into bright and dark lesions from healthy tissue. The proposed work was validated on various public databases. Ahmad et al. [226] proposed a CAD method for detecting DR using OCT images. This work is three steps process: (1) On the OCT image, locating and segmenting 12 different retinal layers using Markov-Gibbs random field of intensity and model descriptor. (2) determining the characteristics of the segmented layers with cumulative probability distributive, and (3) learning discriminating features and classify them into a diabetic and normal person. The proposed art obtained promising performance to detect early DR using OCT images. Another study [227] presents a diagnostic approach to grade and detect NP diabetic retinopathy from OCT images automatically. The images in the proposed work have segmented the retina into 12 layers: quantifies the reflectivity, curvature, and thickness of each layer, and ultimately used this information to feed into a classifier. The proposed work reported 92% accuracy on test datasets of 40 images. However, the study includes a limited number of subjects and a lack of higher DR.

Table 14 Deep learning methods to detect HR using color fundus

Cited	Methodology	Modality	Datasets	Performance measure
Triwijoyo et al. [207]	A CNN model was used to detect HR using fundus images	Color Fundus image	DRIVE	Accuracy = 98%
Triwijoyo et al. [208]	Deep neural network and RBM algorithm was used to compute the change of blood vessel and determine the AVR and OD area	Color Fundus image	DRIVE (20)	Not Reported
Tan et al. [114]	CNN method to segment the OD, fovea, and blood vessel	Color Fundus image	DRIVE (40)	Average accuracy = 92%
AlBadawi et al. [209]	Fully convolutional deep network for Retinal vascular into A/V	Color Fundus image	Dataset (100)	Accuracy = 93%
Welikala et al. [210]	CNN model to classify retinal vascular into A/V	Color Fundus image	Dataset (100)	Accuracy = 86%
Yao et al. [211]	CNN model used to extract blood vessels from fundus image	Color Fundus image	DRIVE (40)	SP = 96%, SE = 77%
Prentašić et al. [212]	DNN model used for detecting exudates	Color Fundus image	DriDB (50)	F1 score = 77%, PPV = 77%, SE = 77%

In [228], CAD was proposed to grade the DR using spectral domain OCT images automatically. First, with the segmentation method, 12 distinct layers are localized. Then, three features such as curvature, thickness, and curvature were computed from segmented layers. Finally, a classifier is used to classify the subject standard and DR and then grade the DR. The proposed work obtained an accuracy of 93% on distinguishing between normal and DR. Also, it achieved a correct classification of 98% between early and mild DR. In a previous study in [229] using spectral-domain OCT images to hard exudates in DR. This work reported 97% accuracy. However, this work was misidentified in 4 cases out of 150 images. Maziyar et al. [230] presented a study to analyze the OCT volumes in DR research automatically. In this work, First, using carefully designed masking and affine registration, create an anatomically consistent volume of interest (VOI) in distinct OCT images. Then, using stochastic gradient descent optimization, efficient B-spline transformations are computed. The experimental result of the proposed work showed successful retinal structure detection due to DR using 105 subjects. In reference [231], multi imaging modalities were used to automatically diagnose DR, such as OCT, OCTA, and clinical biomarkers. Multiple pathophysiological important features were extracted from each layer of OCT and each OCTA and combined with clinical data using a machine learning classifier called random forest. This work reported an accuracy of 96%. Another study [232] also evaluates and demonstrates the utility of multimodal imaging tools, namely OCT and OCTA, for objective diagnosis of proliferative diabetic retinopathy and achieved high detection of 100% on all modalities. However, these findings need to be further validated to alternate the current practice. Recently, ElTanboly et al. [233] proposed a CAD system for early detection of DR using OCT B-scans. From automatically segmented retinal layers, the method calculates distinct discriminant morphological and reflectivity

markers. Current, state-of-the-art machine learning classifiers are utilized to fuse and classify these descriptors. The best accuracy was achieved of 97.6% with the integration of this descriptor. Table 15 provides a summary of image processing-based DR detection arts using color fundus and OCT modalities.

5.6.2 Deep Learning-Based DR Detection Using Fundus and OCT Images

Computer vision, speech recognition and image analysis [234, 235] are the main fields where multilayer deep learning (MDL) algorithms have been outperformed compared to traditional machine-learning (TML) methods. In this subsection, we will focus on a deep learning-based method using fundus images and OCT images. Sahlsten et al. [236] built a deep learning system capable of detecting referable diabetic retinopathy using fundus images. This method shows good performance by designing a cost-efficient system. However, a small number of images with high resolutions were used as the training set. A MDL algorithm was developed by Gulshan et al. [237] to detect diabetic macular edema (DME) and retinopathy (DR). The authors used retinal fundus images. However, this research did not determine the algorithm feasibility for clinical settings. Abramoff et al. [238] developed an automatic detection algorithm for diabetic retinopathy on a public dataset through the integration of deep learning. It has been demonstrated that when the algorithm is combined with Deep learning, the efficacy of DR screening improves and thus could be utilized as a prediction tool to minimize sight loss and blindness due to this devastating disease. However, it was reported that this work might not generalize to all kinds of fundus images from different sources. Swapna et al. [239] used MDL architecture to classify diabetes patient data into diabetic or normal heart rate variability (HRV) signals. They used RNN-LSTM

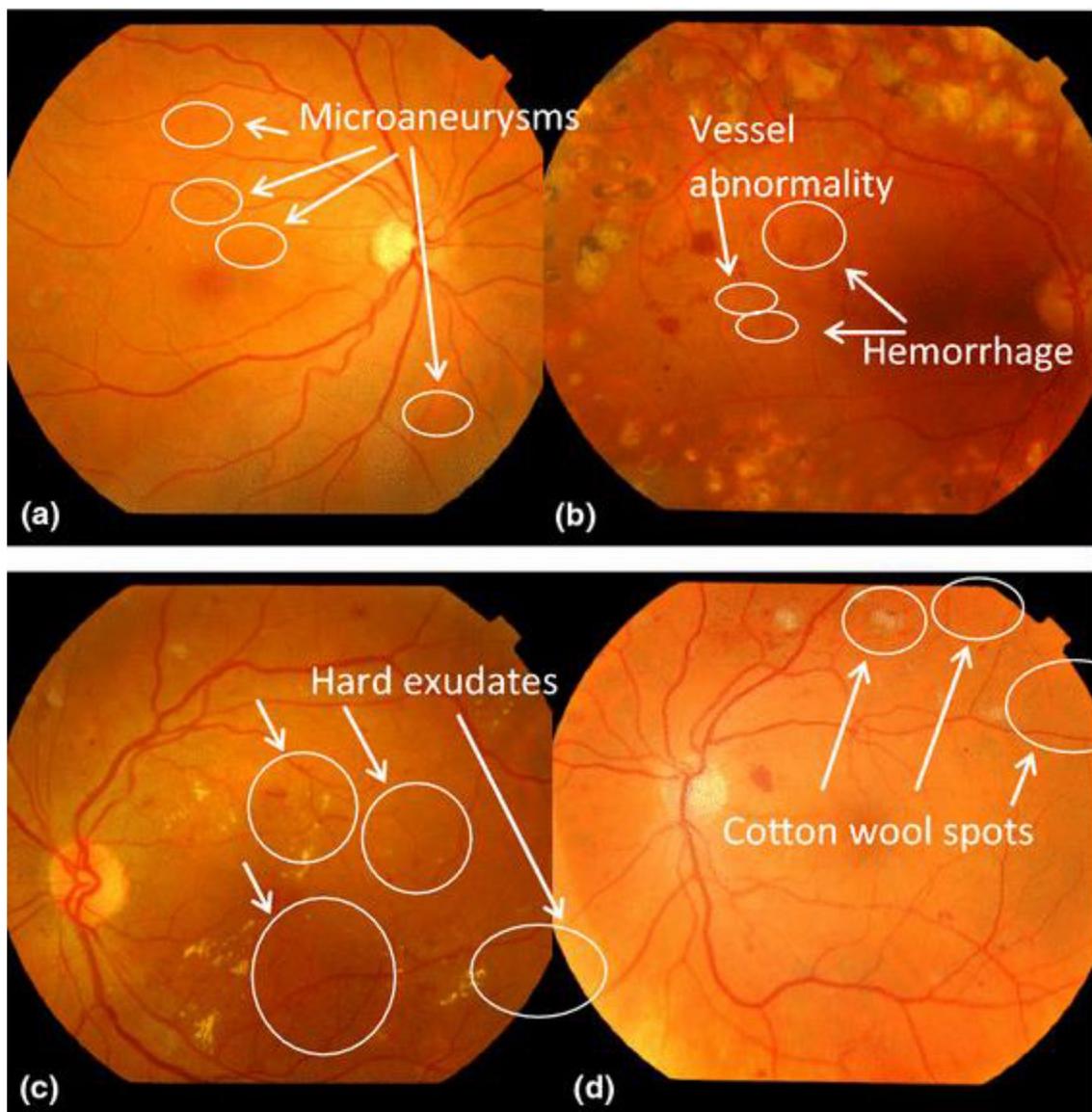


Fig. 16 Example of DR-related lesions of diabetic retinopathy where each fundus image represents **a** normal, **b** microaneurysms, **c** hemorrhages and vessel abnormality, **d** hard exudates and **e** cotton wool spots [216]

with CNN models to extract complex features and classified those features by SVM. This study obtained an accuracy of 97%. Hemanth et al. [240] proposed an improved diabetic retinopathy detection and classification approach based on a deep convolutional neural network; for improved results, the hybrid technique combined image processing and deep learning algorithm. The method was validated using 400 retinal fundus images from the MESSIDOR database, with 97% accuracy. However, detecting lesions in color fundus images for DR is considered time-consuming. In [241] classification of DR is based on modified Alexnet architecture to improve the performance of diagnostic systems. This work achieved good accuracy of 96%. Whereas in paper [242],

the authors simply deploy CNN MDL algorithm to extract features from RF images.

Several TML algorithms is used for classification and obtained 99.89% average accuracy using three different datasets. A deep network path-based approach was used with a variety of datasets, including Messidor, Kaggle, IDRiD, DDR, and DIARETDB0 [244]. The authors achieved an accuracy of 0.912. Victor et al. [244] proposed a bio-inspired scheme of synaptic meta plasticity in the CNN model to detect the DR using fundus images early. The result reported from the proposed work outperforms previous works, even using a small dataset from Kaggle for model training. This work obtained an accuracy of 95.5%, an F1score of 94%,

Table 15 Performance comparison of traditional arts for DR detection using Fundus and OCT

Cited	Methodology	Modality	Database	Performance measure
Walter et al. [217]	Detection of Exudates for DR diagnosis using image processing techniques such as morphological technique and watershed transformation	Fundus image	Database (30)	SE = 92.8% Predictive value = 92.4%
Akara et al. [218]	A set of an optimal adjusted morphological operator to be used exudate detection on DR patients	Fundus image	Database (60)	SE = 80% SP = 90%
Quellec et al. [219]	The image classifier is taught to recognize patterns of arbitrary size that only appear in relevant photos, using reference images designated as relevant or irrelevant by physicians. Similar patterns are searched in new images after training to classify them as either relevant or irrelevant	Fundus image	MESSIDOR (1200), e-phtha (273)	Accuracy = 86%, Accuracy = 76%
Amel et al. [220]	To detect hard exudates (HEs) in retinal images, we used a combination of the k-means clustering technique and mathematical morphology	Fundus image	Database (50)	Accuracy = 99.7%, SE = 95.9% Predictive value = 92.2%
Mookiah et al. [222]	Fundus images were processed to extract abnormal signs, such as hard exudates, area of blood vessels, bifurcation points, textures, and entropies. In this work, 13 features were fed to the machine learning classifier algorithm Probabilistic Neural Network (PNN)	Fundus image	Database (156)	Average Accuracy = 96.1% SE = 96.2% SP = 96.1%
Zhang et al. [223]	They preprocessed the image to normalize, denoise, and detect the artifact and reflections in the processed image. A segmentation algorithm based on mathematical morphology was proposed to characterize the classical and contextual features. Afterward, a random forest algorithm was employed to detect the exudates	Fundus image	e-ophtha EX database	AUC = 95%
Akram et al. [224]	A hybrid classifier detecting the retinal lesions by preprocessing, extracting lesions from a candidate, formulation of feature followed by the classification	Fundus image	DRIVE, STARE, DIARETDB, and MESSIDOR	AUC = 98.1%
Colomer et al. [225]	Local binary pattern and granulometric profile locally to extract morphological and texture information from the retinal image. This information then fed into a classifier to classify into bright and dark lesion from healthy tissue	Fundus image	E-OPHTHHA (47), DIARETDB1 (89)	SE = 81% SP = 81.3%
Ahmad et al. [226]	Proposed CAD method for detecting DR using OCT images	OCT image	Database (52)	Accuracy = 92% SE = 83% SP = 100%

Table 15 (continued)

Cited	Methodology	Modality	Database	Performance measure
Szymkowiak et al. [228]	An enhanced algorithm for hard exudates detection in Optical Coherence Tomography images of DR	OCT image	Database (150)	Accuracy = 97%
Khansari et al. [229]	CAD was proposed to grade the DR using spectral-domain OCT images automatically	OCT image	Database (74)	Accuracy = 93% SE = 91% SP = 97%
Sandhu et al. [230]	Presents a diagnostic approach to grade and detect NP diabetic retinopathy from OCT	OCT image	Database (80)	Accuracy = 92.5%
Schwartz et al. [231]	Multi imaging modality was used to automatically diagnosis DR such as OCT, OCTA, and clinical biomarkers	OCT image	Database (111)	Accuracy = 76% SE = 85% SP = 87%
Sharafelddeen et al. [232]	Multimodal imaging tool, namely OCT and OCTA for objective diagnosis of proliferative diabetic retinopathy	OCT image	Database (383)	Detection rate = 100%
E[Tanboly et al. [233]	Proposed CAD system for early detection of DR using OCT B-scans	OCT image	Database (260)	Accuracy = 97.6% SE = 96% SP = 99% F1-score = 97%

and 98.9% precision. Some studies [244, 246] referred to DR stage classification as multi-class classification.

We have studied some recent deep learning DR detection approaches using OCT images. The study in [247] proposed a deep learning framework called MEDNet for automatic detection of an avascular area in OCT angiography of DR. This work reported excellent performance on the ultra-wide field of OCTA, which show the potential of clinical to early detect and processions of DR. Li et al. [248] proposed deep learning network called, OCTDNet for early DR detection scheme using OCT images. One network extract feature from the original OCT image and another network extracted retinal information in this work. The proposed work obtained an accuracy of 92%, sensitivity 90%, and specificity of 95%. However, the proposed network was not tested and optimized on clinical application. Ghazal et al. [249] proposed a diagnostic system to detect NPDR using the CNN model from the OCT imaging modality. The CNN model to extract feature and SVM for classifying image into DR and normal. This work obtained promising accuracy of 94%. However, revealing two pre-trained models and the preprocessing step requirement before model training made the proposed work computationally complex. Morgan et al. [250] proposed an ensemble technique for DR detection using OCT angiography. They designed a neural network by combining pre-trained networks such as VGG19, ResNet50, and DenseNet using ImageNet weights. They used soft voting and stacking techniques for ensembled networks. The proposed network obtained 92% and 90% accuracy for the majority voting and stacking technique, respectively. However, the small dataset used in the proposed work limits network performance.

Kermany et al. [251] created a transfer learning algorithm based on OCT images for diabetic retinopathy classification. As a pre-trained model, they used the Inception-v3 deep learning model. To train, test, and validate the technique, they used four types of OCT images such as CNV, DME, DRUSEN, and NORMAL are all terms for choroidal neovascularization. To recognize DR, the authors showed that they obtained classification accuracy comparable to human expert. In another study [252], the authors used OCT imaging modality and transfer learning (TL) algorithm to recognize DR. In case of TL algorithm, they utilized ResNet-101 model as a pre-trained step to classify five classes of DR. This ensemble was utilized to execute four binary classification tasks (one vs. all) from OCT images to identify anomalies. However, various parameters such as pre-trained model, size of the image, number of channels affect the performance of the pre-trained model. Therefore, Tohidul et al. [253] used an optimized pre-trained model to extract features and feed a conventional classifier for DR diagnosis using OCT images. The result of the proposed work showed better performance with low computation cost to train the model. Another research study [254] proposed a deep convolutional

neural network for DR screening and staging. They used 18 convolutional layers and three fully connected layers of CNN architecture to analyze images. The study enables early detection and tracks the disease progression, which may help medical applications reduce vision loss. Table 16 illustrates the recent deep learning-based DR detection using color fundus and OCR modalities.

6 Discussions

We have searched and described many retinal image analysis studies related to traditional machine learning (TML) and multilayer deep learning (MDL) algorithms in various fundus imaging modalities. This review article is unique because we did not find any review article, which focused on this broader sense for classifying eye-related diseases. In particular, we have selected a set of publications by using a technique. In this paper selection process, we have gathered 3,650 scientific papers as a preliminary study. We have selected in total 270 papers by removing duplicate and irrelevant studies. As a result, we have selected 270 papers and summarized them in brief and quantity analysis. In addition, many other papers are included in terms of obtaining the distribution of a trend, which are not described in the paper. Another one of the study's major limitations was the requirement to exclude publications published in JCR impact quartiles lower than Q3. We have concluded that the MDL architectures are widely used nowadays compared to TML algorithms. In addition, the fundus images and OCT are also utilized in many studies compare to other imaging modalities for the detection of eye-related diseases. As a result, this paper is primarily focused on the retinal fundus photograph (RFP) and optical coherence tomography (OCT) images. that is mostly utilized in past studies. Because of the huge number of references to be included, a complete description of all investigations was impossible. Therefore, the research team's goals for this review article include finishing a review in which multiple disciplines converge such as machine learning, medical image analysis, and various imaging modalities.

Glaucoma, diabetic retinopathy (DR), diabetic hypertension (DHR), Cataract, and age-related macular degeneration (AMD) are some of the most common and important retinal diseases as described in Table 1. A permanent vision loss occurs if these diseases are not discovered at an early stage. To screen this eye-related disease, many studies used different imaging modalities as visually described in Fig. 1. It is illustrated by numerous abnormalities in the retina such as microaneurysms (MA), hard exudates, soft exudates, or cotton wool spots (CWS), hemorrhages (HEM), neovascularization (NV), and age-related macular edema (AME) (see Fig. 2). An analysis of Ophthalmic imaging modalities is

used as computerized tools by ophthalmologists for faster screening and diagnosis of these diseases. Among other eye-related diseases, diabetic retinopathy (DR) comprises a broad spectrum of potential lesions, abnormalities, and defect locations, posing a unique diagnostic problem. Identification of MA is emphasized in lesion detection algorithms as a characteristic indicator of early DR, allowing the illness to be recognized before. Subsequently, if it is not detected at an early stage then it could cause severe eye damage. Also, detecting lesion types precludes accurate DR grading, which could improve the screening process for patients in later stages. Additionally, a distinguishing feature of distinct lesion types could cause patients to be "missed" during screening which depends exclusively on a single disease presentation. Along with the difficulties associated with the diverse pathology of DR, lesions and blood vessel abnormalities can manifest in virtually any area of the retina. Approaches to DR diagnosis must consider the clinical significance of DR geography. For example, DR necessitates a large field of view (at least 60 degrees). These provisions reduce the use of low-field-of-view imaging systems and cause problems in the simultaneous diagnosis of DR under conditions such as glaucoma, which mainly affect the head of the Optic Nerve (ONH). DR necessitates a clear view of the retinal periphery, which is prone to severe blood vessel abnormalities and damage. Well, several other lesions are damaged due to eye-related diseases such as the macula in AMD poses a particular threat to vision. Apart from color fundus images, imaging modalities such as fluorescein angiography and OCT can aid in the diagnosis of DR. Fluorescein angiography gives great contrast for defects like MA lesions. Still, OCT reveals changes in retinal layers that would otherwise be impossible to detect.

Regardless of imaging modalities, machine learning algorithms are used in the past to recognize the eye-related disease. A visual trend of different imaging modalities and various types of traditional-machine and deep learning algorithms to diagnosis eye-related diseases, where Fig. 3a shows the trend of various imaging modalities, Fig. 3b represents eye-related diseases concerning deep-learning and Fig. 3c figure indicate the traditional machine learning algorithms for the diagnosis of eye-related diseases. Nowadays, multilayer deep learning (MDL) algorithms are mostly utilized compared to traditional machine learning (TML) methods. The MDL algorithms, particularly in DR grading, were the most successful study. The end-to-end learning scheme adopted by numerous papers was successful; others have found lesions using lower computing costs before transferring the resulting functionality to the CNN to achieve high success. In many cases, segmentation of the optic disc and/or blood vessels resulted in improved overall performance; however, such procedures were occasionally fraught with difficulty and complications. For example, removing blood

Table 16 A summary of deep learning arts for DR using fundus and OCT images

Cited	Methodology	Modality	Database	Performance measure
Sahlten et al. [236]	A deep learning system capable of detecting referable diabetic retinopathy using fundus images	Fundus image	Database (41,122)	Accuracy=97.3%
Gulshan et al. [237]	Deep learning network to create an algorithm for detecting diabetic retinopathy and diabetic macular edema in retinal fundus images	Fundus image	EyePACS-1 (9963), Messidor-2 (1748)	SE=97%, SP=93% and SE=96% SP=93%
Abramoff et al. [238]	Developed an automatic detection algorithm for diabetic retinopathy on a public dataset through the integration of deep learning	Fundus image	Messidor-2 (1748)	AUC=98%
Swapna et al. [239]	They developed deep learning architectures to classify diabetes patient data into diabetic or average Heart Rate Variability (HRV) signals	Fundus image	Database (1000)	Accuracy=95.7%
Hemanth et al. [240]	An improved diabetic retinopathy detection and classification approach based on a deep convolutional neural network	Fundus image	MESSIDOR (400)	Accuracy=97%, SP=98%, Precision=94%, FScore=94%, GMean=95%
Gayathri et al. [242]	They used the CNN model to extract feature from retinal fundus image for better image classification	Fundus image	IDRiD, MESSIDOR, and KAGGLE	Average Accuracy=99.8%, K-score=94%
Zago et al. [244]	A lesion localization model was designed using a deep network patch-based approach to reduce model complexity and improve performance	Fundus image	DIARETDB1 (84)	Sensitivity=94%, ROC=91%
Vives-Boix et al. [244]	A bio-inspired synaptic scheme meta plasticity in CNN model to early detect the DR using fundus images	Fundus image	Kaggle (3662)	Accuracy=95.5%, F1-score=94.4%, Precision=98.9%, Recall=90%
Li et al. [245]	A Dense Correlation Network (DCNet) is proposed	Fundus image	Database (1167)	AUC=93%
Chen et al. [246]	An approach to retinal image classification is proposed based on the integration of multi-scale shallow CNNs	Fundus image	Kaggle (35,126)	Accuracy=86%
Guo et al. [247]	A deep learning framework called MEDNet for automatic detection of an avascular area in OCT angiography of DR	OCT image	Database (40)	Accuracy=89%, F1-Score=91%
Li et al. [248]	Deep learning network called OCTDNet for early DR detection scheme using OCT images	OCT image	Database (4168)	Accuracy=92%, SE=90%, SP=95%
Ghazal et al. [249]	A diagnostic system to detect NPDR using CNN model from OCT imaging modality	OCT image	Cirrus HD-OCT	Accuracy=94%,
Heisler et al. [250]	Ensemble technique for DR detection using OCT angiography	OCT image	Database (463)	Accuracy=92%, SE=90%, SP=93%
Kermany et al. [251]	A transfer learning algorithm based on OCT images for diabetic retinopathy classification	OCT image	Database (1000)	Accuracy=92.3%, SE=93.0%, SP=90.1%

Table 16 (continued)

Cited	Methodology	Modality	Database	Performance measure
Lu et al. [252]	An OCT-based transfer learning approach for diagnosing diabetic retinopathy	OCT image	Database (25,134)	Kappa value = 92%, Accuracy = 95%
Tohidul et al. [253]	The optimized pre-trained model was used to extract features and fed to a conventional classifier for DR diagnosis using OCT images	OCT image	Database (109,309)	Accuracy = 97%, SP = 99%, Precision = 97%
Shaban et al. [254]	In this work, Deep Convolutional Neural Network is used to analyze OCT and extract unique features for each stage of DR and therefore classify images and stage the disease	OCT image	Kaggle (3648)	Accuracy = 89%, SE = 89%, SP = 95%

vessels benefited with MA diagnosis in the past, but this operation now makes it impossible to detect the proliferative disease. Consequently, the clinical relevance is becoming paramount when selecting associated preprocessing stages such as the detection of MA. For early detection of DR, the detection of MA or other lesions are an important concern. Also, the other lesion types become more prominent as DR advances. Furthermore, without considerable preprocessing, DL systems can often achieve excellent sensitivity and specificity. In practice, many MDL algorithms are based on less handcrafted, which ultimately minimized the possibility of omitting critical parameters.

The stage and duration of hypertension are connected to the occurrence of Hypertensive Retinopathy (HR) of eye disease. Pathological eye injuries such as arteriolar narrowing, retinal hemorrhage, macular edema, spots of cotton wool, and blood vessels are impaired by the HR. Currently, a few computerized systems developed to recognize HR by using 2-stage (HR versus non-HR) as described in subsection 4.5. Those systems, however, focused on extracting features through techniques based on handcraft and deep-learning models (DLMs). Moreover, it is difficult for DLMs to define specialized features to recognize the multistage of HR. In addition, the deep features are used in the past to detect 2-stage of HR, but the classification accuracy is not up-to-the-mark. A new hypertensive retinopathy framework is required to grade the HR based on 4-stages [255] such as normal (NR), mild (MLD), moderate (MOD), and malignant (MLG) (see Fig. 16). To define HR characteristics, it is arduous to specify and recognize relevant HR lesion properties from fundus images. There are no datasets available with a medical specialist to identify the HR severity level directly from retina fundus images to train and evaluate the network. Several variant models automatically learn characteristics using the deep-learning technique, but at each layer, they utilize the same weighted scheme. For accurate decisions,

it is hard for layers to transfer weights to deeper network levels. There is a necessity to develop an automated solution for recognizing the four stages of HR, and to the best of our knowledge, no study has addressed this problem.

6.1 Critical Analysis and Challenges

A visual example is shown in Fig. 17 to explain the distribution of different deep-learning models that were utilized in the past for the development of detection of eye-related diseases. Despite the positive results so far, there are still obstacles and restrictions to applying traditional machine learning (TML) and multilayer deep-learning (MDL) algorithms [130] in a wide variety of fields. First, the quality of input images is intrinsically variable, owing to a lack of standardized imaging annotation and patient-to-patient variation in eye-related disease features. Furthermore, inter-expert heterogeneity in clinical decision-making is a serious problem that has been well-documented [131]. Experts' high variability in interpreting ophthalmic modalities images may add bias into model training. Second, the quantity of photos with clinical annotations is severely limited due to the tremendous effort of manual annotation. As a result, improved picture annotation techniques for collecting clinical annotations should be created such as localization of exudates, AME, and retinal hemorrhages. The semi-supervised learning approach tries to make full use of unlabeled samples to improve model generalization performance. Third, due to the complexity of diseases, enough data is required to create high-accuracy models; yet data for more severe stages of the disease and rare diseases is frequently insufficient. Fourth, the current TML and MDL algorithms are developed based on a single modality of eye-related diseases in the domain of ophthalmology, but integrated diagnosis employing several imaging techniques is required.

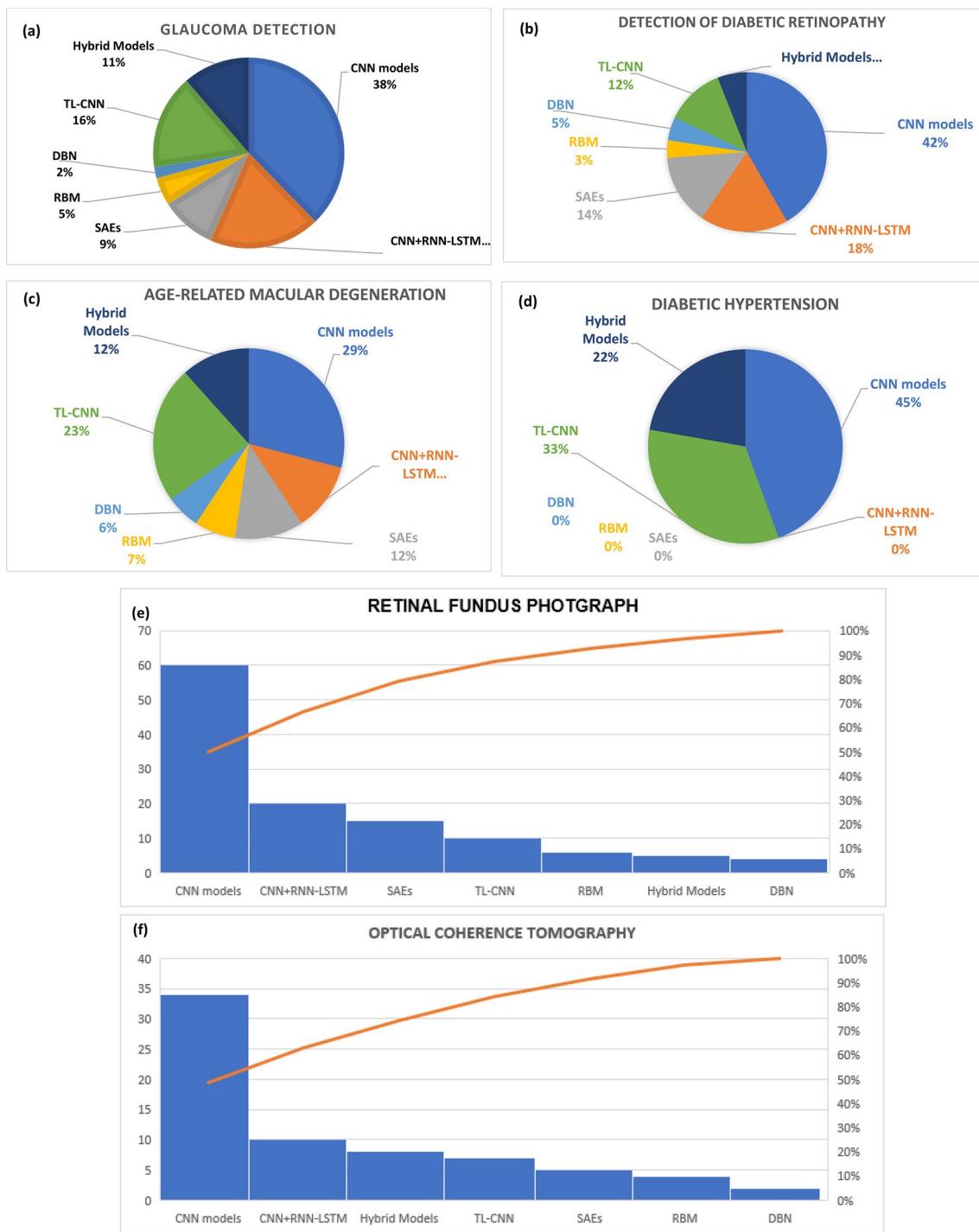


Fig. 17 Distribution of different deep-learning models that utilized in the past for the development of detection eye-related diseases using different deep-learning algorithms

Modern automated imaging may aid healthcare systems with a few staff in the future. Healthcare workers may be able to deliver better patient care if ophthalmic gadgets have intelligence built in. Furthermore, the ophthalmology-based computer systems may be implemented within ocular imaging devices (e.g., portable fundus cameras and smartphones),

enabling real-time picture diagnosis with minimum operator skill. New multimodal imaging techniques and better intelligent algorithms allow for cooperative training using complementary modalities with different strengths. Improved hardware performance and lower costs will enable this integrated computerize domain. Patients could self-screen without

supervision before seeing an ophthalmologist, thanks to the growing use of AI in medical treatment. Deep neural networks are beneficial when it comes to making precise decisions based on large amounts of data. They do, however, come with several important challenges and limits that you must either accept or attempt to overcome. In general, the use of TML and MDL models is quite challenging due to the lack of mathematics and theoretical background of machine-learning algorithms. As a result, it provides a basic challenge to the selection of hyperparameters, which is required in MDL architectures.

Others are more focused on a single domain. Over the last few years, sophisticated models, and a set of best practices for standard computer vision tasks, such as object detection and localization, have been established. Although the rate of development continues to be quite rapid, several issues appear to have been resolved, at least for the time being. Using the basic building blocks outlined above and arranging them according to the principles of, say, ResNet and SENet, you may achieve near-state-of-the-art performance on two-dimensional object identification, picture classification, and segmentation tasks. The tale for deep learning in medical imaging, on the other hand, is less clear. The fact that medical images are frequently three-dimensional, and three-dimensional convolutional neural networks are as well-developed as their two-dimensional equivalents, is one difficulty. When utilizing CNNs with higher-dimensional picture data, problems with memory and computation usage arise quickly. Researchers are experimenting with a variety of techniques to deal with the problem.

Multilayer Deep learning (MDL) is here to stay in medical data analysis. Despite the numerous hurdles connected with implementing deep learning in healthcare settings, the technologies yield far too useful outcomes to be overlooked. Many high-impact research publications are utilized the MDL algorithms in medical imaging in top journals demonstrate this (for example [59–72, 74, 74, 256–266]), all published from 2015 to 2021). Figure 17 represents the distribution trend of MDL algorithms used to detect and classification eye-related diseases. Moreover, this figure shows the imaging modalities (retinal fundus photograph (RFP), optical coherence tomography (OCT) trend that used many MDL algorithms in the last decades. To detect eye-related diseases, the practitioners and experts are getting more experience to identify the problems through traditional-machine learning (TDL), MDL solutions, or a mixture of both techniques as a hybrid solution. However, it noticed that the computational complexity related to MDL algorithms will be increased the attention due to increase training and testing size of data. Also, the GPU and cloud-computing are provided less computational burden on the machine-learning architecture but still, they suffered many difficulties as detailed described in Sect. 5.3.

The DR comprises a wide spectrum of potential lesions, abnormalities, and defect locations, making accurate diagnosis difficult. MA identification is emphasized in lesion detection algorithms as a characteristic indicator of early DR, allowing the illness to be recognized before severe eye damage occurs. However, detecting lesion types at the expense of effective DR grading, which could improve the standard of care and provide options to people in later stages of the disease. Furthermore, individual variances in the number and course of specific lesion types may cause individuals to be “missed” during screening if only one pre-screening criterion is used. Lesions and blood vessels (BVs) anomalies can arise in nearly any location of the retina, adding to the complexities of DR's complex pathophysiology. The clinical importance of DR geography must be considered in engineering methods to DR diagnosis. For example, DR necessitates a large field of view (at least 60 degrees). Such restrictions limit the use of lower FOV imaging techniques and make it more difficult to diagnose DR alongside disorders like glaucoma, which primarily affect the Optic nerve head (ONH). DR necessitates a clear view of the retinal periphery, which is prone to substantial BVs abnormalities and damage.

Damage to places such as the macula in AME is particularly vision-threatening, thus additional central structures are also necessary. Imaging methods such as fluorescein angiography and OCT, in addition to color fundus pictures, provide insight into DR diagnosis. Fluorescein angiography provides strong contrast for defects like MAs, but OCT reveals changes in retinal layers that would otherwise be impossible to detect. Deep learning was utilized in the most successful experiments for categorization, particularly in DR grading. End-to-end learning was successful in much research; others discovered lesions with less computationally expensive algorithms before feeding the resulting features to a CNN. In many cases, segmenting the OD and/or BVs improved overall performance; nonetheless, such techniques were occasionally problematic. For example, eliminating BVs benefited in MA diagnosis in the past, but this approach now makes it impossible to detect the proliferative illness. As a result, while making decisions on relevant preprocessing phases, clinical relevance should be paramount: Early DR detection is greatly aided by MA detection, which is the most prevalent target of lesion detection systems. BVs changes as DR advances and other types of lesions become increasingly essential. Furthermore, without considerable preprocessing, DL systems can often achieve excellent sensitivity and specificity. Important parameters are less likely to be omitted with this less-handcrafted method.

After the year 2015, multilayer deep learning (MDL) became popular for identifying Glaucoma, diabetic retinopathy (DR), and hypertension retinopathy (HR). Many variants of MDL models are rapidly developed and evolved. This

MDL's latest trend is moved towards automatic extraction of lesions from retinal imaging modalities without using handcrafted features. In practice, the automatic detection of features is having many benefits such as unknown features are possible to extract from different imaging modalities, rapid and accurate features learning. Still, the progress of MDL algorithms has several problems that do not allow the clinical studies to be conducted further. These drawbacks are related to data availability and the uncertainty that surrounds deep learning algorithms. In addition, the size of the dataset is small, and an imaging modality related to hypertension is not also available online to perform experiments. The DenseHyper [159] is developed by us to address the hypertension issue in diabetes-related to fundus photograph, which is used to address the two-class problem. However, hypertension [255] is a four-class category problem. Despite these facts, transfer learning (TL) is relatively used by many authors to address the issue of online available dataset size as presented in Table 6. By using TL models, the pre-train architecture is used to pick deep layers on the small dataset to learn specific features from the image. In addition to TL models, the few-short learning architectures are the best architecture to handle the issue of smaller datasets. The few short learning algorithms are not used in the past to categorize DR or HR. By integrating vision transformers or semantic-based learning to detect eye-related diseases, the classification accuracy will be improved. Rather than simply improving label "accuracies", future research should focus on improving clinical value and solving real-world difficulties in retinal diagnostics and therapies. For example, MDL has a lot of promise in clinical settings where resources (such as doctors and equipment) are scarce. It noticed that such systems would correct for poor image quality to make using mobile cameras and grading image quality easier.

For two key reasons, this survey was limited in its ability to effectively compare DR detection and grading system methods: (1) a narrow emphasis within the broad topic of DR, and (2) inconsistent usage of datasets and evaluation criteria. In the first scenario, it was discovered that the best results were found in articles that focused on specific goals. For example, as previously mentioned, one study deleted BVs and HEMs to increase MA detection. While this minimizes MA false positives, the authors give up a complete picture of DR status in the process. To incorporate more particular tasks into overall DR grading, superior systems used learning ensembles or deep-learning methods. Other sheets utilized a lot of ink. Furthermore, due to inconsistent evaluation metrics, many studies were difficult to evaluate adequately. Readers of this survey should be cautious about adopting the article's assertions that it only reports one evaluation metric when reading scientific research. For example, data imbalance can readily skew accuracy, and less prevalent assessment metrics might be used to display outcomes in

a more favorable light than a more objective manner (i.e., AUC). Readers should be wary of publications that solely rely on private databases, especially if the authors have not made their code public for external verification.

Patients could self-screen without supervision before seeing an ophthalmologist, thanks to the growing use of AI in medical treatment. Furthermore, patients in rural places could receive routine eye exams and disease progression tracking without the need for highly qualified operators. Another major research path will be to improve the interpretability of networks. The problem of the "black box" has been noted as a barrier to the use of DL in healthcare. Instead of simply obtaining a diagnosis, existing research has built unique algorithms that allow clinicians to examine and visualize the decision-making process (e.g., OCT tissue-segmentation). In terms of treatment, more study on ophthalmic robotics is needed; experiments on robotic intraretinal vascular injection and anterior macular surgery have been conducted.

6.2 Current and Future Trends of MDL Models

Multilayer deep learning (MDL) models are widely utilized to diagnose eye-related diseases through various imaging modalities. Although, these MDL models are also used in many different other tasks such as object recognition and segmentation in computer vision applications. According to a literature review, we noticed that the MDL models such as CNN, RBM, DBN, SAE, transfer learning, and hybrid learning [256] are mostly utilized. In general, Table 17 represents the most common variant of DML models and algorithms. However, these MDL algorithms are depended on certain hyper-parameters and training sizes. Compared to these MDL models, there are the latest trend in MDL algorithms that are related to few-short learning, transformers, and semantic deep learning methods to perform similar tasks with less training size [256]. Those models are still not utilized in the domain of detection of eye diseases. Later, these enhanced MDL models will be used to increase the accuracy and efficiency of MDL algorithms.

Despite these models, the CNN [257, 258] is mostly utilized to automatically extract features from retinograph images without manually learn distinct features. However, the selection of the CNN model and features map size, epochs, and hyper-parameter selection for the CNN model are major problems related to this architecture. Let us consider a simple CNN model that consists of Con-L, Dens-L, Pool-L, and Soft-max layers for classification tasks. Moreover, the activation functions are needed to add non-linearity to the model by simulating neurons, and sometimes, the backpropagation step is required for fine-tuning weights and learning a model. The most frequent methods for achieving faster convergence and avoiding overfitting are batch

Table 17 Algorithm-wise categorization of deep learning, deep transfer learning, few-short learning, explainable deep-learning and semantic-based deep learning

Multilayer-deep-learning methods	Classification	Algorithms
Deep learning	Stack-based autoencoders (SAEs)	Variant of SAEs
	Convolutional neural network (CNN)	Variant of CNN model
	Deep-belief network (DBN)	Convolutional deep belief network, Conditional restricted Boltzmann machine
	Recurrent neural network (RNN)	Bidirectional RNN, long short-term memory
Deep transfer learning (DTL)	CNN-based transfer learning	AlexNet, GoogleNet, Visual geometry group network (VGG), Deep Residual Learning, Inception v4 (v2, v3), Restnet-152 (34,50,101), LeNet
Few-short learning (FSL)	k-short learning algorithms	0-Short,1-short,2-short and n-short learning
Semantic-based deep learning (SDL)	Instance-based segmentation	U-Net

RNN Bidirectional RNN, LSTM Long short-term memory, DBN Deep belief network, CNN Convolution neural network, RNN Recurrent neural network

normalization and dropout. We have also a fully connected layer that can be added to the architecture. It is the same as a feed-forward neural network, and it is employed as the classification's final layer that is the same as the SoftMax layer. If the fully connected layer is not added, then the network architecture can be utilized for object segmentation tasks instead of classification problems. As the name suggested that the fully convolutional models are made up of locally connected convolutional layers such as downsampling (convolution), upsampling (deconvolution), and pooling, as the name suggests. The lack of a dense layer results in faster computing and a smaller number of parameters [259].

Typically, these have a convolutional layer downsampling path and a deconvolutional layer upsampling path. Skip connections, which bypass multiple layers and transport information across layers with various resolutions for better learning, are examples of these. Autoencoders are a type of encoder that automatically encodes data. It is an unsupervised neural network [260] that compresses and encodes data efficiently before reconstructing it from the reduced encoded representation. Initializing random weights and tweaking them via stochastic gradient descent and back-propagation are common methods for training neural networks. The strategies for training a neural network used in the literature on ocular diagnostics utilizing fundus images are listed below.

Deep learning methods typically necessitate huge datasets for optimal training. The datasets available in ophthalmic diagnostics are typically small, resulting in model overfitting. Transfer learning offers a solution to this problem by fitting the weights on a big dataset, often from a different domain. The weights of some or all the layers of the model are then fine-tuned on the target dataset. Ensemble learning entails independently training several deep learning models and then polling their findings for a given data sample to obtain a prediction. The most typical method is to use

majority voting, which entails selecting the most common result as the final prediction. It is based on the idea that errors from models that have been trained independently are unlikely to overlap.

6.3 GPU-Based Performance Evaluation of MDL Models

Several non-invasive imaging modalities are routinely utilized to perform patient screening, and clinical decisions are progressively providing more benefits from the supplemental data, which are given by modern medical devices. Ophthalmology [262] is an example of a discipline that makes extensive use of in-office by using various imaging modalities. When diagnosis eye-related by ophthalmologists, the Fundus photography and optical coherence tomography (OCT) are the most used imaging techniques in ophthalmology (OCT). To recognize eye-related diseases using imaging modalities, numerous machine-learning algorithms have been developed in the past 20 years. Compared to traditional machine learning (TDL) approaches, the advanced multilayer deep learning (MDL) architectures achieved encouraging results by employing many hidden layers of nodes, whose weights are derived by a training process from input data. Moreover, the CNN model in MDL models achieved better performance when diagnosis eye-related diseases by fundus images in diabetic retinopathy programs. With limited resources, there is several platforms developed in the past to perform automation by the selection of optimal network architectures, pre-processing methods, and hyperparameter optimization. Due to the use of these platforms, the clinician or researcher's effort has been reduced for automation of medical problems.

Over the past twenty years, the MDL methods have proved effective compared to TDL systems in all fields, and especially in medical image classification [263] remain prominent in all areas. This review article focuses

on important deep learning structures utilized in recognition tasks of eye-related diseases by using various imaging modalities. The Graphics Processing Unit (GPU) introduction is vital for general-purpose issues, and there is devotion to GPU exploitation for deep learning algorithms. A Graphics Processing Unit (GPU) is a dedicated, electronic circuit planned to quickly operate and modify memory to hasten the image creation in a frame buffer envisioned for output to a display device. On the other hand, computer vision trains computers to understand and interpret the visual world while utilizing digital images from videos, cameras, and deep learning models. The learning models are many, each having benefits and disadvantages, and have been checked in previous sections. In summary, the CNNs perform better than DBNs, especially when benchmarking medical image classification datasets like fundus images and OCT. The DBNs are best and outperform other models in cases non-visual is input and have difficulty checking and estimating joint probabilities. Their computational costs in developing it result in drawbacks. Feature learning is the advantage of CNNs and can bypass handcrafted structures important for other network types, and features are learned automatically. On the other hand, they depend on ground truth, including labeling the data, while DBNs/DBMS and the SAs do not have this restriction and work in an unsupervised manner. The autoencoders are ineffective if errors are in the first layer, making learning to rebuild the typical training data. Notably, CNNs and DBNs/DBMS training processes are time-consuming [264–266], but real-time training is possible in SAS. The CNNs are good in computer vision problems, like object detection in medical images, since it permits abstracting an object's uniqueness or category from the particulars of the visual input. It is possible since they are invariant to transformations like translation, scale, and rotation. The greater application is targeted in medicine, where GPU, deep learning, and computer vision are employed in genome sequencing technology. High throughput sequencing (HTS), cancer composition, drug synergy prediction, drug-target interaction, and drug response predictions play a role in suppressing cancer by generating effective drugs with fewer side effects. The MDL models to recognize eye-related diseases using OCT have shown performance comparable to retinal specialists with decades of experience. However, the development of such MDL algorithms demands substantial resources.

7 Research Gap

The published survey papers on this topic were focused on clinically detecting specific diseases using the deployment of high-level ophthalmic equipment in hospitals. These articles were failed to address the thorough analysis of

the advanced machine learning approaches for diabetes assessment. Thus, to our knowledge, this is the first review article examining the performance of deep learning algorithms and architectures for diabetes diagnosis utilizing ophthalmic imaging modalities. The clinical findings have presented that the application of retinal imaging analysis using deep learning strategies is highly beneficial and effective. Furthermore, they eliminate the need for manual feature extraction due to the methodologies' data-driven nature. However, there are several limitations of deep learning arts that are highlighted below with the potential solutions:

- Without extensive datasets, computer vision problems are much more challenging to solve. In addition, there is a scarcity of data that has been manually annotated. Deep learning is associated with massive amounts of data because the model primarily learns from the data's inherent pattern. This, in turn, creates a critical issue for this field. The valuable solution to this crucial problem is utilizing the Generative models proposed by Goodfellow et al. [267]. To date, very few efforts [261, 268] have been made to explore the possibilities of generative modeling to synthesize new fundus images with annotations and with clinical relevance. The application of Generative Adversarial Network and Variational Autoencoders are observed beneficial for the image generation. Clinical-relevant synthetic data can be generated by effectively applying these strategies. It will assist in increasing the amount of available data, but it will also help avoid privacy issues in general.
- There is a lack of standard statistical metrics used for the estimation of the deep learning model performance. Each researcher has recorded a performance index after several trials to grade their scheme accurately. Thus, it is challenging to present the comparative analysis on the deep learning performance for the disease diagnosis.
- There is a possibility of a domain shift problem occurring due to the differences in camera settings. Training and test data are typically distributed in the same manner. This is not the rule in real life. This domain shift can cause significant damage in real-world applications if it is not addressed properly beforehand. Therefore, the model should deal with data from different distributions when used for training and testing purposes. It is frequently noticed that accuracy suffers because of this domain shift problem. Deep domain adaption techniques should be given more attention to developing robust models that may be used in real-world ophthalmic diagnosis. The ophthalmic diagnostic of the future will likely look in this direction.
- Few-short learning and Semantic-based deep-learning algorithms are not still explored in the domain of detect-

- ing and classifying eye-related diseases through various imaging modalities.
- In the case of huge patients in the coming years, the computational complexity in terms of the selection of benchmarks is an important concern for many deep-learning architectures.
 - In the future, it might be possible to include more fundus imaging modalities compared to retinal fundus photographs and OCT images.
 - Instance-based segmentation is also an important field for detecting different objects in an image and achieved higher performance in other domains such as computer vision. It should be tested in the domain of detecting many lesions, which are presented in the fundus images.
 - The vision transformers also provided the best performance in many medical image segmentation and classification tasks especially in computer vision applications so they will be provided the better performance compared to many other deep-learning algorithms. In future work, it should be tested.
 - Computerized methods for the classification of many other eye-related diseases are not still explored even in this review article that should be considered in future work.
 - Although few approaches have been established for classifying retinal fundus images into two-category based HR and non-HR, and none-of-them focused on four categories of HR. There are no datasets (see Table 6) available with a medical specialist to identify the HR severity level directly from retina fundus images to train and evaluate the network. There are several variant models that automatically learn characteristics using deep-learning technique, but at each layer they utilize the same weighted scheme. For accurate decisions, it is hard for layers to transfer weights to deeper network levels. There is a necessity to develop an automated solution for recognizing the four stages of HR, and to the best of our knowledge, no study has addressed this problem.

8 Conclusions

This work presents a survey of recent advances in the machine learning arts (multi-layered deep learning architectures) to diagnose the retinal diseases caused by diabetes using various imaging modalities. The presented review starts with the essential background of diabetes and its different micro-vascular complications. Then, the comprehensive details of the retinal imaging analysis using colored fundus image, OCT, etc., and different types of deep learning frameworks were discussed to develop the fast and reliable automated retinal screening systems. Accordingly, the theoretical and practical aspects of every conventional

and advanced machine learning method were analyzed and demonstrated. Thus, the obtained observation in this review endorses the dominancy of the deep learning scheme to the traditional hand-engineered approaches for each phase of the retinal processing to assess ocular diseases.

Moreover, the appropriate machine-learning technologies have a lot of promise to facilitate early detection and proper care for people who have retinal irregularities. The clinician and researchers are using various imaging modalities to automate this screening process. Faster and less computationally intensive algorithms have a lot of potential for use in low-resource locations where rapid retinal screening is needed. Since more complex and diverse imaging modalities are emerging that can improve the sensitivity and specificity of eye-related disease detection beyond what human observation can manually achieve. Overall, the multilayer deep learning (MDL) architectures compared to traditional-machine learning (TML) systems show the most potential in detecting eye-related diseases. Future implementation of such algorithms will aid doctors in more quickly distinguishing between multiple eye diseases that may be present at the same time or have similar symptoms and diagnosing disease stages so that the most appropriate and effective treatments can be implemented quickly and effectively.

Acknowledgements This research was supported by the Deanship of Scientific Research, Imam Mohammad Ibn Saud Islamic University (IMSIU), Saudi Arabia, Grant No. (20-13-09-002).

Declarations

Conflict of interest All authors declared no conflict of interest.

References

1. Teo ZL, Tham YC, Yu MCY, Chee ML, Rim TH, Cheung N, Cheng CY (2021) Global prevalence of diabetic retinopathy and projection of burden through 2045: systematic review and meta-analysis. *Ophthalmology* 128(11):1580–1591
2. Cheung R, Chun J, Sheidow T, Motolko M, Malvankar-Mehta MS (2021) Diagnostic accuracy of current machine learning classifiers for age-related macular degeneration: a systematic review and meta-analysis. *Eye* 1–11
3. Perepelkina T, Fulton AB (2021) Artificial intelligence (AI) applications for age-related macular degeneration (AMD) and other retinal dystrophies. In: Seminars in ophthalmology, pp 1–6. Taylor & Francis
4. Bhardwaj C, Jain S, Sood M (2021) Hierarchical severity grade classification of non-proliferative diabetic retinopathy. *J Ambient Intell Humaniz Comput* 12(2):2649–2670
5. Zhou Y, Wang B, Huang L, Cui S, Shao L (2020) A benchmark for studying diabetic retinopathy: segmentation, grading, and transferability. *IEEE Trans Med Imaging* 40:818–828
6. Yim J, Chopra R, Spitz T, Winkens J, Obika A, Kelly C, De Fauw J (2020) Predicting conversion to wet age-related macular degeneration using deep learning. *Nat Med* 26(6):892–899

7. Saeed MU, Oleszczuk JD (2016) Advances in retinal imaging modalities: challenges and opportunities. *World J Ophthalmol* 6(2):10–19. <https://doi.org/10.5318/wjo.v6.i2.10>
8. Tong Y, Lu W, Yu Y, Shen Y (2020) Application of machine learning in ophthalmic imaging modalities. *Eye and Vis* 7:1–15
9. Zhang X, Fang J, Hu Y, Xu Y, Higashita R, Liu J (2020) Machine learning for cataract classification and grading on ophthalmic imaging modalities: a survey. arXiv preprint <http://arxiv.org/abs/2012.04830>
10. Abdelmaksoud E, El-Sappagh S, Barakat S, Abuhmed T, Elmogy M (2021) Automatic diabetic retinopathy grading system based on detecting multiple retinal lesions. *IEEE Access* 9:15939–15960
11. Mayya V, Kamath SS, Kulkarni U (2021) Automated microaneurysms detection for early diagnosis of diabetic retinopathy: a comprehensive review. *Comput Methods Programs Biomed Update* 1:1–15
12. Shanthini A, Manogaran G, Vadivu G, Kottilingam K, Nithyakan P, Fancy C (2021) Threshold segmentation based multi-layer analysis for detecting diabetic retinopathy using convolution neural network. *J Ambient Intell Humaniz Comput* 136:1–15
13. Sharma A, Rani R (2021) A systematic review of applications of machine learning in cancer prediction and diagnosis. *Arch Comput Methods Eng* 28:4875–4896
14. Dargan S, Kumar M, Ayyagari MR, Kumar G (2019) A survey of deep learning and its applications: a new paradigm to machine learning. *Arch Comput Methods Eng* 27:1071–1092
15. Mittal R, Arora S, Bansal V, Bhatia MPS (2021) An extensive study on deep learning: techniques, applications. *Arch Comput Methods Eng* 28:4471–4485
16. Abbas Q, Ibrahim ME, Jaffar MA (2018) Video scene analysis: an overview and challenges on deep learning algorithms. *Multimedia Tools Appl* 77(16):20415–20453
17. Abbas Q, Ibrahim ME, Jaffar MA (2019) A comprehensive review of recent advances on deep vision systems. *Artif Intell Rev* 52(1):39–76
18. Bala MP, Rajalakshmi P, Sindhuja AM, Naganandhini S (2021) A review on recent development for diagnosis of glaucoma. *Ann Rom Soc Cell Biol* 25(3):2723–2736
19. Ahmad H, Yamin A, Shakeel A, Gillani SO, Ansari U (2014) Detection of glaucoma using retinal fundus images. In: 2014 International conference on robotics and emerging allied technologies in engineering (iCREATE), pp 321–324. IEEE
20. Mirzania D, Thompson AC, Muir KW (2020) Applications of deep learning in detection of glaucoma: a systematic review. *Eur J Ophthalmol* 31:1618–1642
21. Abdullah F, Imitiaz R, Madni HA, Khan HA, Khan TM, Khan MA, Naqvi SS (2021) A review on glaucoma disease detection using computerized techniques. *IEEE Access* 9:37311–37333
22. Stefan AM, Paraschiv EA, Ovreiu S, Ovreiu E (2020) A review of glaucoma detection from digital fundus images using machine learning techniques. In: 2020 International conference on e-health and bioengineering (EHB), pp 1–4. IEEE
23. Sengupta S, Singh A, Leopold HA, Gulati T, Lakshminarayanan V (2018) Application of deep learning in fundus image processing for ophthalmic diagnosis—a review. arXiv preprint <http://arxiv.org/abs/1812.07101>
24. Usman M, Fraz MM, Barman SA (2017) Computer vision techniques applied for diagnostic analysis of retinal OCT images: a review. *Arch Comput Methods Eng* 24(3):449–465
25. Ran AR, Tham CC, Chan PP, Cheng CY, Tham YC, Rim TH, Cheung CY (2021) Deep learning in glaucoma with optical coherence tomography: a review. *Eye* 35(1):188–201
26. Mahum R, Rehman SU, Okon OD, Alabrah A, Meraj T, Rauf HT (2022) A novel hybrid approach based on deep CNN to detect glaucoma using fundus imaging. *Electronics* 11(1):26
27. Pead E, Megaw R, Cameron J et al (2019) Afibercept: a review of its effect on the treatment of exudative age-related macular degeneration. *Eur J Ophthalmol* 29(4):368–378
28. Papadopoulos Z (2019) Automated detection of age-related macular degeneration in color fundus photography: a systematic review. *Surv Ophthalmol* 64:498–511
29. Papadopoulos Z (2019) Afibercept: a review of its effect on the treatment of exudative age-related macular degeneration. *Eur J Ophthalmol* 29:368–378
30. Alyoubi WL, Shalash WM, Abulkhair MF (2020) Diabetic retinopathy detection through deep learning techniques: a review. *Inform Med Unlocked* 20:100377
31. Stolte S, Fang R (2020) A survey on medical image analysis in diabetic retinopathy. *Med Image Anal* 64:101742
32. Qureshi I, Ma J, Abbas Q (2019) Recent development on detection methods for the diagnosis of diabetic retinopathy. *Symmetry* 11(6):749
33. Qureshi I, Sharif M, Yasmin M, Raza M, Javed YM (2016) Computer aided systems for diabetic retinopathy detection using digital fundus images: a survey. *Curr Med Imaging* 12(4):234–241
34. Abbas Q, Alsheddy A (2021) A methodological review on prediction of multi-stage hypovigilance detection systems using multimodal features. *IEEE Access* 9:47530–47564
35. Shaheed K, Mao A, Qureshi I, Kumar M, Abbas Q, Ullah I, Zhang X (2021) A systematic review on physiological-based biometric recognition systems: current and future trends. *Arch Comput Methods Eng* 28:4917–4960
36. Anwar SM, Majid M, Qayyum A, Awais M, Alnowami M, Khan MK (2018) Medical image analysis using convolutional neural networks: a review. *J Med Syst* 42(11):1–13
37. Khatami A, Nazari A, Khosravi A, Lim CP, Nahavandi S (2020) A weight perturbation-based regularisation technique for convolutional neural networks and the application in medical imaging. *Expert Syst Appl* 149:113196
38. He T, Liu Y, Yu Y, Zhao Q, Hu Z (2020) Application of deep convolutional neural network on feature extraction and detection of wood defects. *Measurement* 152:107357
39. Golestan N, Moghaddam M (2020) Human activity recognition using magnetic induction-based motion signals and deep recurrent neural networks. *Nat Commun* 11(1):1–11
40. Dong S, Wang P, Abbas K (2021) A survey on deep learning and its applications. *Comput Sci Rev* 40:100379
41. Hati AS, Chakrabarti P, Abawajy JH, Keong NW (2021) Development of energy efficient drive for ventilation system using recurrent neural network. *Neural Comput Appl* 33:8659–8668
42. Drif A, Zerrad HE, Cherifi H (2020) EnsVAE: ensemble variational autoencoders for recommendations. *IEEE Access* 8:188335–188351
43. Sivaramakrishnan N, Subramanyaswamy V, Viloria A, Vijayakumar V, Senthilselvan N (2020) A deep learning-based hybrid model for recommendation generation and ranking. *Neural Comput Appl* 33:10719–10736
44. Shahamiri SR, Thabtah F (2020) An investigation towards speaker identification using a single-sound-frame. *Multimedia Tools Appl* 79(41):31265–31281
45. Hassan MM, Gumaei A, Alsanad A, Alrubaian M, Fortino G (2020) A hybrid deep learning model for efficient intrusion detection in big data environment. *Inf Sci* 513:386–396
46. Bharati S, Podder P, Mondal MRH (2020) Hybrid deep learning for detecting lung diseases from X-ray images. *Inform Med Unlocked* 20:100391
47. Sindi H, Nour M, Rawa M, Öztürk S, Polat K (2021) A novel hybrid deep learning approach including combination of 1D power signals and 2D signal images for power quality disturbance classification. *Expert Syst Appl* 174:114785

48. Wan Z, Yang R, Huang M, Zeng N, Liu X (2021) A review on transfer learning in EEG signal analysis. *Neurocomputing* 421:1–14
49. Abbas Q, Ibrahim ME, Khan S, Baig AR (2022) Hypo-driver: a multiview driver fatigue and distraction level detection system. *CMC-computers Mater Contin* 71(1):1999–2017
50. Niu S, Liu Y, Wang J, Song H (2020) A decade survey of transfer learning (2010–2020). *IEEE Trans Artif Intell* 1(2):151–166
51. Salaken SM, Khosravi A, Nguyen T, Nahavandi S (2017) Extreme learning machine based transfer learning algorithms: a survey. *Neurocomputing* 267:516–524
52. Day O, Khoshgoftaar TM (2017) A survey on heterogeneous transfer learning. *J Big Data* 4(1):1–42
53. Medicine JH Age-Related Macular Degeneration (AMD). <https://www.hopkinsmedicine.org/health/conditions-and-diseases/age-related-macular-degeneration-amd>. Accessed 19 Mar 2021
54. Khalid S, Akram MU, Shehryar T, Ahmed W, Sadiq M, Manzoor M, Nosheen N (2021) Automated diagnosis system for age-related macular degeneration using hybrid features set from fundus images. *Int J Imaging Syst Technol* 31(1):236–252
55. Mohaimin SM, Saha SK, Khan AM, Arif ASM, Kanagasingam Y (2018) Automated method for the detection and segmentation of drusen in colour fundus image for the diagnosis of age-related macular degeneration. *IET Image Proc* 12(6):919–927
56. Alom MZ, Taha TM, Yakopcic C, Westberg S, Sidike P, Nasrin MS, Asari VK (2018) The history began from alexnet: a comprehensive survey on deep learning approaches. arXiv preprint <http://arxiv.org/abs/1803.01164>
57. Tang P, Wang H, Kwong S (2017) G-MS2F: GoogLeNet based multi-stage feature fusion of deep CNN for scene recognition. *Neurocomputing* 225:188–197
58. Nan F, Zeng Q, Xing Y, Qian Y (2020) Single image super-resolution reconstruction based on the ResNeXt network. *Multimedia Tools Appl* 79(45):34459–34470
59. Ucar F, Korkmaz D (2020) COVIDagnosis-Net: deep Bayes-SqueezeNet based diagnosis of the coronavirus disease 2019 (COVID-19) from X-ray images. *Med Hypotheses* 140:109761
60. Khan S, Naseer M, Hayat M, Zamir SW, Khan FS, Shah M (2021) Transformers in vision: a survey. arXiv preprint <http://arxiv.org/abs/2101.01169>
61. Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, Houlsby, N (2020) An image is worth 16×16 words: transformers for image recognition at scale. arXiv preprint <http://arxiv.org/abs/2010.11929>
62. Chen CF, Fan Q, Panda R (2021) Crossvit: cross-attention multi-scale vision transformer for image classification. arXiv preprint <http://arxiv.org/abs/2103.14899>
63. Wu H, Xiao B, Codella N, Liu M, Dai X, Yuan L, Zhang L (2021) CvT: introducing convolutions to vision transformers. arXiv preprint <http://arxiv.org/abs/2103.15808>
64. Minaee S, Boykov YY, Porikli F, Plaza AJ, Kehtarnavaz N, Terzopoulos D (2021) Image segmentation using deep learning: a survey. *IEEE Trans Pattern Anal Mach Intell*. [arXiv:2001.05566](https://arxiv.org/abs/2001.05566)
65. Wang Y, Yao Q, Kwok JT, Ni LM (2020) Generalizing from a few examples: a survey on few-shot learning. *ACM Comput Surv (CSUR)* 53(3):1–34
66. Kotia J, Kotwal A, Bharti R, Mangrulkar R (2021) Few shot learning for medical imaging. In: Machine learning algorithms for industrial applications, pp 107–132. Springer, Cham
67. Wang T, Lei Y, Fu Y, Curran WJ, Liu T, Yang X (2020) Medical imaging synthesis using deep learning and its clinical applications: a review. arXiv preprint <http://arxiv.org/abs/2004.10322>
68. Zhou SK, Greenspan H, Davatzikos C, Duncan JS, Van Ginneken B, Madabhushi A, Summers RM (2021) A review of deep learning in medical imaging: imaging traits, technology trends, case studies with progress highlights, and future promises. In: Proceedings of the IEEE
69. Al Chanti D, Duque VG, Crouzier M, Nordez A, Lacourpaille L, Mateus D (2021) Ifss-net: interactive few-shot siamese network for faster muscle segmentation and propagation in volumetric ultrasound. *IEEE Trans Med Imaging* 40:1–14
70. Huang B, Tian J, Zhang H, Luo Z, Qin J, Huang C, Yuan C (2020) Deep semantic segmentation feature-based radiomics for the classification tasks in medical image analysis. *IEEE J Biomed Health Inform* 7:2655–2664
71. Taghanaki SA, Abhishek K, Cohen JP, Cohen-Adad J, Hamarneh G (2021) Deep semantic segmentation of natural and medical images: a review. *Artif Intell Rev* 54(1):137–178
72. Wang EK, Chen CM, Hassan MM, Almogren A (2020) A deep learning based medical image segmentation technique in Internet-of-Medical-Things domain. *Future Gener Comput Syst* 108:135–144
73. Rezaei M, Yang H, Meinel C (2020) Recurrent generative adversarial network for learning imbalanced medical image semantic segmentation. *Multimedia Tools Appl* 79(21):15329–15348
74. Mathew A, Amudha P, Sivakumari S (2020) Deep learning techniques: an overview. In: International conference on advanced machine learning technologies and applications, pp 599–608. Springer, Singapore
75. Khalil T, Akram MU, Khalid S, Dar SH, Ali N (2021) A study to identify limitations of existing automated systems to detect glaucoma at initial and curable stage. *Int J Imaging Syst Technol* 31(3):1155–1173
76. Quigley HA, Broman AT (2006) The number of people with glaucoma worldwide in 2010 and 2020. *Br J Ophthalmol* 90(3):262–267
77. Jamous KF, Kalloniatis M, Hennessy MP, Agar A, Hayen A, Zangerl B (2015) Clinical model assisting with the collaborative care of glaucoma patients and suspects. *Clin Exp Ophthalmol* 43(4):308–319
78. Shingleton BJ, Gamell LS, O'Donoghue MW, Baylus SL, King R (1999) Long-term changes in intraocular pressure after clear corneal phacoemulsification: normal patients versus glaucoma suspect and glaucoma patients. *J Cataract Refract Surg* 25(7):885–890
79. Li L, Xu M, Wang X, Jiang L, Liu H (2019) Attention based glaucoma detection: a large-scale database and CNN model. In: Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp 10571–10580
80. Verma OP, Roy S, Pandey SC, Mittal M (eds) (2019) Advancement of machine intelligence in interactive medical image analysis. Springer
81. Kim YK, Jeoung JW, Park KH (2017) Inferior macular damage in glaucoma: its relationship to retinal nerve fiber layer defect in macular vulnerability zone. *J Glaucoma* 26(2):126–132
82. Khan MW, Sharif M, Yasmin M, Fernandes SL (2016) A new approach of cup to disk ratio based glaucoma detection using fundus images. *J Integr Des Process Sci* 20(1):77–94
83. Soltani A, Battikh T, Jabri I, Lakhoud N (2018) A new expert system based on fuzzy logic and image processing algorithms for early glaucoma diagnosis. *Biomed Signal Process Control* 40:366–377
84. Keerthanashree T, Bala MP (2016) Extraction of retinal features in fundus images for glaucoma diagnosis. *Curr Trends Inf Technol* 6(1):21–28
85. Bowd C, Belghith A, Proudfoot JA, Zangwill LM, Christopher M, Goldbaum MH, Weinreb RN (2020) Gradient-boosting classifiers combining vessel density and tissue thickness measurements for classifying early to moderate glaucoma. *Am J Ophthalmol* 217:131–139

86. Moghimi S, Bowd C, Zangwill LM, Penteado RC, Hasenstab K, Hou H, Weinreb RN (2019) Measurement floors and dynamic ranges of OCT and OCT angiography in glaucoma. *Ophthalmology* 126(7):980–988
87. Mangipudi PS, Pandey HM, Choudhary A (2021) Improved optic disc and cup segmentation in Glaucomatous images using deep learning architecture. *Multimedia Tools Appl* 80:30143–30163
88. Bibiloni P, González-Hidalgo M, Massanet S (2019) A real-time fuzzy morphological algorithm for retinal vessel segmentation. *J Real-Time Image Proc* 16(6):2337–2350
89. Agarwal A, Issac A, Singh A, Dutta MK (2016) Automatic imaging method for optic disc segmentation using morphological techniques and active contour fitting. In: 2016 Ninth international conference on contemporary computing (IC3), pp 1–5. IEEE
90. Ingle R, Mishra P (2013) Cup segmentation by gradient method for the assessment of glaucoma from retinal image. *Int J Eng Trends Technol* 4(6):2540–2543
91. Khalid NEA, Noor NM, Ariff NM (2014) Fuzzy c-means (FCM) for optic cup and disc segmentation with morphological operation. *Procedia Comput Sci* 42:255–262
92. Pal S, Chatterjee S (2017) Mathematical morphology aided optic disk segmentation from retinal images. In: 2017 3rd International conference on condition assessment techniques in electrical systems (CATCON). pp 380–385. IEEE
93. Sun X, Xu Y, Zhao W, You T, Liu J (2018) Optic disc segmentation from retinal fundus images via deep object detection networks. In: 2018 40th annual international conference of the IEEE engineering in medicine and biology society (EMBC), pp 5954–5957. IEEE
94. Zhang L, Fisher M, Wang W (2015) Retinal vessel segmentation using multi-scale textons derived from keypoints. *Comput Med Imaging Graph* 45:47–56
95. Septiarini A, Harjoko A, Pulungan R, Ekantini R (2018) Automated detection of retinal nerve fiber layer by texture-based analysis for glaucoma evaluation. *Healthcare Inform Res* 24(4):335–345
96. Kirar BS, Agrawal DK (2018) Computer aided diagnosis of glaucoma using discrete and empirical wavelet transform from fundus images. *IET Image Proc* 13(1):73–82
97. Elseid AAG, Hamza AO (2019) Glaucoma detection using retinal nerve fiber layer texture features. *J Clin Eng* 44(4):180–185
98. Nirmala K, Venkateswaran N, Kumar CV, Christobel JS (2017) Glaucoma detection using wavelet based contourlet transform. In: 2017 International conference on intelligent computing and control (I2C2), pp 1–5. IEEE
99. Yin F, Liu J, Wong DWK, Tan NM, Cheung C, Baskaran M, Wong TY (2012) Automated segmentation of optic disc and optic cup in fundus images for glaucoma diagnosis. In: 2012 25th IEEE international symposium on computer-based medical systems (CBMS), pp 1–6. IEEE
100. Chang HT, Liu CH, Pai TW (2008) Estimation and extraction of B-cell linear epitopes predicted by mathematical morphology approaches. *J Mol Recognit Interdiscip J* 21(6):431–441
101. Aslam MA, Salik MN, Chughtai F, Ali N, Dar SH, Khalil T (2019) Image classification based on mid-level feature fusion. In: 2019 15th International conference on emerging technologies (ICET), pp 1–6. IEEE
102. Nugroho HA, Oktoeberza WK, Erasari A, Utami A, Cahyono C (2017) Segmentation of optic disc and optic cup in colour fundus images based on morphological reconstruction. In: 2017 9th International conference on information technology and electrical engineering (ICITEE), pp 1–5. IEEE
103. Acharya UR, Bhat S, Koh JE, Bhandary SV, Adeli H (2017) A novel algorithm to detect glaucoma risk using texton and local configuration pattern features extracted from fundus images. *Comput Biol Med* 88:72–83
104. Dey A, Bandyopadhyay SK (2016) Automated glaucoma detection using support vector machine classification method. *J Adv Med Med Res* 1–12
105. Mookiah MRK, Acharya UR, Lim CM, Petznick A, Suri JS (2012) Data mining technique for automated diagnosis of glaucoma using higher order spectra and wavelet energy features. *Knowl Based Syst* 33:73–82
106. Akram MU, Tariq A, Khalid S, Javed MY, Abbas S, Yasin UU (2015) Glaucoma detection using novel optic disc localization, hybrid feature set and classification techniques. *Aust Phys Eng Sci Med* 38(4):643–655
107. Vidotti VG, Costa VP, Silva FR, Resende GM, Cremasco F, Dias M, Gomi ES (2013) Sensitivity and specificity of machine learning classifiers and spectral domain OCT for the diagnosis of glaucoma. *Eur J Ophthalmol* 23(1):61–69
108. Nithya R, Venkateswaran N (2015) Analysis of segmentation algorithms in colour fundus and OCT images for glaucoma detection. *Indian J Sci Technol* 8(24):1
109. Chan YM, Ng EYK, Jahmunah V, Koh JEW, Lih OS, Leon LYW, Acharya UR (2019) Automated detection of glaucoma using optical coherence tomography angiogram images. *Comput Biol Med* 115:103483
110. Rajan A, Ramesh GP (2015) Automated early detection of glaucoma in wavelet domain using optical coherence tomography images. *Biosci Biotechnol Res Asia* 12(3):2821–2828
111. Ramzan A, Akram MU, Ramzan J, Salam AA, Yasin UU (2018) Automated inner limiting membrane segmentation in OCT retinal images for glaucoma detection. In: Science and information conference, pp 1278–1291. Springer, Cham
112. Babu TR, Devi S, Venkatesh R (2015) Optic nerve head segmentation using fundus images and optical coherence tomography images for glaucoma detection. *Biomed Pap* 159(4):607–615
113. Chen X, Xu Y, Wong DWK, Wong TY, Liu J (2015) Glaucoma detection based on deep convolutional neural network. In: 2015 37th Annual international conference of the IEEE engineering in medicine and biology society (EMBC), pp 715–718. IEEE
114. Tan JH, Acharya UR, Bhandary SV, Chua KC, Sivaprasad S (2017) Segmentation of optic disc, fovea and retinal vasculature using a single convolutional neural network. *J Comput Sci* 20:70–79
115. Chai Y, Liu H, Xu J (2018) Glaucoma diagnosis based on both hidden features and domain knowledge through deep learning models. *Knowl Based Syst* 161:147–156
116. Pal A, Moorthy MR, Shahina A (2018) G-eyenet: a convolutional autoencoding classifier framework for the detection of glaucoma from retinal fundus images. In: 2018 25th IEEE international conference on image processing (ICIP), pp 2775–2779. IEEE
117. Asaoka R, Tanito M, Shibata N, Mitsuhashi K, Nakahara K, Fujino Y, Kiuchi Y (2019) Validation of a deep learning model to screen for glaucoma using images from different fundus cameras and data augmentation. *Ophthalmol Glaucoma* 2(4):224–231
118. Chen X, Xu Y, Yan S, Wong DWK, Wong TY, Liu J (2015). Automatic feature learning for glaucoma detection based on deep learning. In: International conference on medical image computing and computer-assisted intervention, pp 669–677. Springer, Cham
119. Maheshwari S, Kanhangad V, Pachori RB (2020) CNN-based approach for glaucoma diagnosis using transfer learning and LBP-based data augmentation. *arXiv preprint* <http://arxiv.org/abs/2002.08013>
120. Hemelings R, Elen B, Breda JB, Blaschko MB, De Boever P, Stalmans I (2021) Glaucoma detection beyond the optic disc: The importance of the peripapillary region using explainable deep learning. *arXiv preprint* <http://arxiv.org/abs/2103.11895>
121. Zilly J, Buhmann JM, Mahapatra D (2017) Glaucoma detection using entropy sampling and ensemble learning for automatic

- optic cup and disc segmentation. *Comput Med Imaging Graph* 55:28–41
122. Maetschke S, Antony B, Ishikawa H, Wollstein G, Schuman J, Garnavi R (2019) A feature agnostic approach for glaucoma detection in OCT volumes. *PLoS ONE* 14(7):e0219126
 123. Raja H, Akram MU, Shaukat A, Khan SA, Alghamdi N, Khawaja SG, Nazir N (2020) Extraction of retinal layers through convolution neural network (CNN) in an OCT image for glaucoma diagnosis. *J Digit Imaging* 33(6):1428–1442
 124. An G, Omodaka K, Hashimoto K, Tsuda S, Shiga Y, Takada N, Nakazawa T (2019) Glaucoma diagnosis with machine learning based on optical coherence tomography and color fundus images. *J Healthcare Eng* 2019:1–9
 125. Gheisari S, Shariflou S, Phu J, Kennedy PJ, Agar A, Kalloniatis M, Golzan SM (2021) A combined convolutional and recurrent neural network for enhanced glaucoma detection. *Sci Rep* 11(1):1–11
 126. García G, Colomer A, Naranjo V (2021) Glaucoma detection from raw SD-OCT volumes: a novel approach focused on spatial dependencies. *Comput Methods Programs Biomed* 200:105855
 127. Nayak DR, Das D, Majhi B, Bhandary SV, Acharya UR (2021) ECNet: an evolutionary convolutional network for automated glaucoma detection using fundus images. *Biomed Signal Process Control* 67:102559
 128. Lee T, Jammal AA, Mariottini EB, Medeiros FA (2021) Predicting glaucoma development with longitudinal deep learning predictions from fundus photographs. *Am J Ophthalmol* 225:86–94
 129. Cho H, Hwang YH, Chung JK, Lee KB, Park JS, Kim HG, Jeong JH (2021) Deep learning ensemble method for classifying glaucoma stages using fundus photographs and convolutional neural networks. *Curr Eye Res* 46(10):1516–1524
 130. Medeiros FA, Jammal AA, Mariottini EB (2021) Detection of progressive glaucomatous optic nerve damage on fundus photographs with deep learning. *Ophthalmology* 128(3):383–392
 131. Singh LK, Garg H, Khanna M, Bhaduria RS (2021) An enhanced deep image model for glaucoma diagnosis using feature-based detection in retinal fundus. *Med Biol Eng Compu* 59(2):333–353
 132. Asano S, Asaoka R, Murata H, Hashimoto Y, Miki A, Mori K, Inoue K (2021) Predicting the central 10 degrees visual field in glaucoma by applying a deep learning algorithm to optical coherence tomography images. *Sci Rep* 11(1):1–10
 133. Chang J, Lee J, Ha A, Han YS, Bak E, Choi S, Park SM (2021) Explaining the rationale of deep learning glaucoma decisions with adversarial examples. *Ophthalmology* 128(1):78–88
 134. Priyanka V, Vaishnavi D (2021) An automated glaucoma detection in fundus images—a survey. In: Intelligent system design, pp 347–359. Springer, Singapore
 135. Nawaz M, Nazir T, Javed A, Tariq U, Yong HS, Khan MA, Cha J (2022) An efficient deep learning approach to automatic glaucoma detection using optic disc and optic cup localization. *Sensors* 22(2):434
 136. Saravanan V, Samuel R, Krishnamoorthy S, Manickam A (2022) Deep learning assisted convolutional auto-encoders framework for glaucoma detection and anterior visual pathway recognition from retinal fundus images. *J Ambient Intell Humaniz Comput* 1–11. <https://doi.org/10.1007/s12652-021-02928-0>
 137. Natarajan D, Sankaralingam E, Balraj K, Karuppusamy S (2022) A deep learning framework for glaucoma detection based on robust optic disc segmentation and transfer learning. *Int J Imaging Syst Technol* 32(1):230–250
 138. Ibrahim MH, Hacibeyoglu M, Agaoglu A, Ucar F (2022) Glaucoma disease diagnosis with an artificial algae-based deep learning algorithm. *Med Biol Eng Comput* 1–12. <https://doi.org/10.1007/s11517-022-02510-6>
 139. Abdel-Hamid L (2022) TWEEC: computer-aided glaucoma diagnosis from retinal images using deep learning techniques. *Int J Imaging Syst Technol* 32(1):387–401
 140. Deperlioglu O, Kose U, Gupta D, Khanna A, Giampaolo F, Fortino G (2022) Explainable framework for Glaucoma diagnosis by image processing and convolutional neural network synergy: analysis with doctor evaluation. *Future Gener Comput Syst* 129:152–169
 141. Tékouabou SCK, Chabbar I, Toulni H, Cherif W, Silkan H (2022) Optimizing the early glaucoma detection from visual fields by combining preprocessing techniques and ensemble classifier with selection strategies. *Expert Syst Appl* 189:115975
 142. Jain S, Indora S, Atal DK (2022) Rider Manta ray foraging optimization-based generative adversarial network and CNN feature for detecting glaucoma. *Biomed Signal Process Control* 73:103425
 143. Singh LK, Khanna M (2022) A novel multimodality based dual fusion integrated approach for efficient and early prediction of glaucoma. *Biomed Signal Process Control* 73:103468
 144. Bhuiyan A, Kawasaki R, Sasaki M, Lamoureux E, Ramamohanarao K, Guymer R, Kanagasingam Y (2013) Drusen detection and quantification for early identification of age related macular degeneration using color fundus imaging. *J Clin Exp Ophthalmol* 4(305):2
 145. Waseem S, Akram MU, Ahmed BA (2014) Drusen detection from colored fundus images for diagnosis of age related Macular degeneration. In: 7th International conference on information and automation for sustainability, pp 1–5. IEEE
 146. Mittal D, Kumari K (2015) Automated detection and segmentation of drusen in retinal fundus images. *Comput Electr Eng* 47:82–95
 147. Acharya UR, Hagiwara Y, Koh JE, Tan JH, Bhandary SV, Rao AK, Raghavendra U (2017) Automated screening tool for dry and wet age-related macular degeneration (ARMD) using pyramid of histogram of oriented gradients (PHOG) and nonlinear features. *J Comput Sci* 20:41–51
 148. Kim YJ, Kim KG (2018) Automated segmentation methods of drusen to diagnose age-related macular degeneration screening in retinal images. *Comput Math Methods Med* 2018:1–8
 149. Graham KW, Chakravarthy U, Hogg RE, Muldrew KA, Young IS, Kee F (2018) Identifying features of early and late age-related macular degeneration: a comparison of multicolor versus traditional color fundus photography. *Retina* 38(9):1751–1758
 150. Jacintha V, Simon J, Tamilarasu S, Thamizhmani R, Nagarajan J (2019) A review on facial emotion recognition techniques. In: 2019 International conference on communication and signal processing (ICCP), pp 0517–0521. IEEE
 151. von der Emde L, Pfau M, Thiele S, Möller PT, Hassenrik R, Fleckenstein M, Schmitz-Valckenberg S (2019) Mesopic and dark-adapted two-color fundus-controlled perimetry in choroidal neovascularization secondary to age-related macular degeneration. *Transl Vis Sci Technol* 8(1):7–7
 152. Pfau M, Lindner M, Gliem M, Steinberg JS, Thiele S, Finger RP, Schmitz-Valckenberg S (2018) Mesopic and dark-adapted two-color fundus-controlled perimetry in patients with cuticular, reticular, and soft drusen. *Eye* 32(12):1819–1830
 153. Thee EF, Meester-Smoor MA, Luttkhuizen DT, Colijn JM, Enthoven CA, Haarman AE, Klaver CC (2020) Performance of classification systems for age-related macular degeneration in the Rotterdam study. *Transl Vis Sci Technol* 9(2):26–26
 154. Rapantzikos K, Zervakis M, Balas K (2003) Detection and segmentation of drusen deposits on human retina: Potential in the diagnosis of age-related macular degeneration. *Med Image Anal* 7(1):95–108
 155. Barriga ES, Murray V, Agurto C, Pattichis MS, Russell S, Abramoff MD, Soliz P (2009) Multi-scale AM-FM for lesion

- phenotyping on age-related macular degeneration. In: 2009 22nd IEEE international symposium on computer-based medical systems, pp 1–5. IEEE
- 156. Liang Z, Wong DW, Liu J, Chan KL, Wong TY (2010) Towards automatic detection of age-related macular degeneration in retinal fundus images. In: 2010 Annual international conference of the IEEE engineering in medicine and biology, pp 4100–4103. IEEE
 - 157. Seddon JM, Dossett J, Widjajahakim R, Rosner B (2019) Association between drusen burden determined by OCT and genetic risk in early and intermediate age-related macular degeneration. *bioRxiv* 743633. <https://doi.org/10.1101/743633>
 - 158. Tan AC, Pilgrim MG, Fearn S, Bertazzo S, Tsolaki E, Morell AP, Curcio CA (2018) Calcified nodules in retinal drusen are associated with disease progression in age-related macular degeneration. *Sci Transl Med* 10(466):1–49
 - 159. Sadda SR, Guymer R, Holz FG, Schmitz-Valckenberg S, Curcio CA, Bird AC, Staurenghi G (2018) Consensus definition for atrophy associated with age-related macular degeneration on OCT: classification of atrophy report 3. *Ophthalmology* 125(4):537–548
 - 160. Lek JJ, Caruso E, Baglin EK, Sharangan P, Hodgson LA, Harper CA, Guymer RH (2018) Interpretation of subretinal fluid using OCT in intermediate age-related macular degeneration. *Ophthalmol Retina* 2(8):792–802
 - 161. Shi Y, Motulsky EH, Goldhardt R, Zohar Y, Thulliez M, Feuer W, Rosenfeld PJ (2019) Predictive value of the OCT double-layer sign for identifying subclinical neovascularization in age-related macular degeneration. *Ophthalmol Retina* 3(3):211–219
 - 162. Vujošević S, Toma C, Villani E, Muraca A, Torti E, Florimbi G, De Cillà S (2019) Quantitative choriocapillaris evaluation in intermediate age-related macular degeneration by swept-source optical coherence tomography angiography. *Acta Ophthalmol* 97(6):e919–e926
 - 163. Corvi F, Srinivas S, Nittala MG, Corradetti G, Velaga SB, Stambolian D, Sadda SR (2020) Reproducibility of qualitative assessment of drusen volume in eyes with age related macular degeneration. *Eye* 35:2594–2600
 - 164. Sadda SR, Abdelfattah NS, Lei J, Shi Y, Marion KM, Mengenien E, Balasubramanian S (2020) Spectral-domain OCT analysis of risk factors for macular atrophy development in the HARBOR study for neovascular age-related macular degeneration. *Ophthalmology* 127(10):1360–1370
 - 165. Jaffe GJ, Martin DF, Toth CA, Daniel E, Maguire MG, Ying GS, Huang J (2013) Comparison of age-related macular degeneration treatments trials research group. Macular morphology and visual acuity in the comparison of age-related macular degeneration treatments trials. *Ophthalmology* 120(9):1860–1870
 - 166. Sadda SR, Tuomi LL, Ding B, Fung AE, Hopkins JJ (2018) Macular atrophy in the HARBOR study for neovascular age-related macular degeneration. *Ophthalmology* 125(6):878–886
 - 167. Thomas A, Sunija AP, Manoj R, Ramachandran R, Ramachandran S, Varun PG, Palanisamy P (2021) RPE layer detection and baseline estimation using statistical methods and randomization for classification of AMD from retinal OCT. *Comput Methods Programs Biomed* 200:105822
 - 168. Guymer R, Wu Z (2020) Age-related macular degeneration (AMD): more than meets the eye. The role of multimodal imaging in today's management of AMD—a review. *Clin Exp Ophthalmol* 48(7):983–995
 - 169. Srinivasan PP, Kim LA, Mettu PS, Cousins SW, Comer GM, Izatt JA, Farsiu S (2014) Fully automated detection of diabetic macular edema and dry age-related macular degeneration from optical coherence tomography images. *Biomed Opt Express* 5(10):3568–3577
 - 170. Alsaih K, Lemaître G, Vall JM, Rastgoo M, Sidibé D, Wong TY, Mériadeau F (2016) Classification of SD-OCT volumes with multi pyramids, LBP and HOG descriptors: application to DME detections. In: 2016 38th Annual international conference of the IEEE engineering in medicine and biology society (EMBC), pp 1344–1347. IEEE
 - 171. Sugmik J, Kiattisin S, Leelasantitham A (2014) Automated classification between age-related macular degeneration and diabetic macular edema in OCT image using image segmentation. In: The 7th 2014 biomedical engineering international conference, pp 1–4. IEEE
 - 172. Hani M, Ben Slama A, Zghal I, Trabelsi H (2021) Appropriate identification of age-related macular degeneration using OCT images. *Comput Methods Biomech Biomed Eng Imaging Vis* 9(2):146–156
 - 173. Grassmann F, Mengelkamp J, Brandl C, Harsch S, Zimmermann ME, Linkohr B, Weber BH (2018) A deep learning algorithm for prediction of age-related eye disease study severity scale for age-related macular degeneration from color fundus photography. *Ophthalmology* 125(9):1410–1420
 - 174. Govindaiah A, Smith RT, Bhuiyan A (2018) A new and improved method for automated screening of age-related macular degeneration using ensemble deep neural networks. In: 2018 40th Annual international conference of the IEEE engineering in medicine and biology society (EMBC), pp 702–705. IEEE
 - 175. Yellapragada B, Hornhauer S, Snyder K, Yu S, Yiu G (2020) Unsupervised deep learning for grading of age-related macular degeneration using retinal fundus images. *arXiv preprint* <http://arxiv.org/abs/2010.11993>
 - 176. von der Emde L, Pfau M, Dysli C, Thiele S, Möller PT, Lindner M, Schmitz-Valckenberg S (2019) Artificial intelligence for morphology-based function prediction in neovascular age-related macular degeneration. *Sci Rep* 9(1):1–12
 - 177. Bhuiyan A, Wong TY, Ting DSW, Govindaiah A, Souied EH, Smith RT (2020) Artificial intelligence to stratify severity of age-related macular degeneration (AMD) and predict risk of progression to late AMD. *Transl Vis Sci Technol* 9(2):25–25
 - 178. Govindaiah A, Hussain MA, Smith RT, Bhuiyan A (2018) Deep convolutional neural network based screening and assessment of age-related macular degeneration from fundus images. In: 2018 IEEE 15th International symposium on biomedical imaging (ISBI 2018), pp 1525–1528. IEEE
 - 179. Peng Y, Dharsi S, Chen Q, Keenan TD, Agrón E, Wong WT, Lu Z (2019) DeepSeeNet: a deep learning model for automated classification of patient-based age-related macular degeneration severity from color fundus photographs. *Ophthalmology* 126(4):565–575
 - 180. Keel S, Li Z, Scheetz J, Robman L, Phung J, Makeyeva G, He M (2019) Development and validation of a deep-learning algorithm for the detection of neovascular age-related macular degeneration from colour fundus photographs. *Clin Exp Ophthalmol* 47(8):1009–1018
 - 181. González-Gonzalo C, Sánchez-Gutiérrez V, Hernández-Martínez P, Contreras I, Lechanteur YT, Domanian A, Sánchez CI (2020) Evaluation of a deep learning system for the joint automated detection of diabetic retinopathy and age-related macular degeneration. *Acta Ophthalmol* 98(4):368–377
 - 182. Babenko B, Balasubramanian S, Blumer KE, Corrado GS, Peng L, Webster DR, Varadarajan AV (2019) Predicting progression of age-related macular degeneration from fundus images using deep learning. *arXiv preprint* <http://arxiv.org/abs/1904.05478>
 - 183. Peng Y, Keenan TD, Chen Q, Agrón E, Allot A, Wong WT, Lu Z (2020) Predicting risk of late age-related macular degeneration using deep learning. *NPJ Digit Med* 3(1):1–10
 - 184. Zapata MA, Royo-Fibla D, Font O, Vela JI, Marcantonio I, Moya-Sánchez EU, Labarta J (2020) Artificial intelligence to

- identify retinal fundus images, quality validation, laterality evaluation, macular degeneration, and suspected glaucoma. *Clin Ophthalmol (Auckland, NZ)* 14:419
185. Kaymak S, Serener A (2018) Automated age-related macular degeneration and diabetic macular edema detection on OCT images using deep learning. In: 2018 IEEE 14th international conference on intelligent computer communication and processing (ICCP), pp 265–269. IEEE
 186. Serener A, Serte S (2019) Dry and wet age-related macular degeneration classification using OCT images and deep learning. In: 2019 Scientific meeting on electrical-electronics & biomedical engineering and computer science (EBBT), pp 1–4. IEEE
 187. Motozawa N, An G, Takagi S, Kitahata S, Mandai M, Hirami Y, Kurimoto Y (2019) Optical coherence tomography-based deep-learning models for classifying normal and age-related macular degeneration and exudative and non-exudative age-related macular degeneration changes. *Ophthalmol Therapy* 8(4):527–539
 188. Das V, Dandapat S, Bora PK (2020) Unsupervised super-resolution of OCT images using generative adversarial network for improved age-related macular degeneration diagnosis. *IEEE Sens J* 20(15):8746–8756
 189. Abbas Q, Ibrahim ME (2020) DenseHyper: an automatic recognition system for detection of hypertensive retinopathy using dense features transform and deep-residual learning. *Multimedia Tools Appl* 79(41):31595–31623
 190. Akagi S, Matsubara H, Nakamura K, Ito H (2018) Modern treatment to reduce pulmonary arterial pressure in pulmonary arterial hypertension. *J Cardiol* 72(6):466–472
 191. Suryani E (2019) The review of computer aided diagnostic hypertensive retinopathy based on the retinal image processing. In: IOP conference series: materials science and engineering, vol 620, no 1, p 012099. IOP Publishing
 192. Abraha B HYPERTENSIVE RETINOPATHY. <https://www.sunopticaltechnologies.com/hypertensive-retinopathy>. Accessed 19 Mar 2021
 193. Grisan E, Foracchia M, Ruggeri A (2008) A novel method for the automatic grading of retinal vessel tortuosity. *IEEE Trans Med Imaging* 27(3):310–319
 194. Muramatsu C, Hatanaka Y, Iwase T, Hara T, Fujita H (2011) Automated selection of major arteries and veins for measurement of arteriolar-to-venular diameter ratio on retinal fundus images. *Comput Med Imaging Graph* 35(6):472–480
 195. Akbar S, Akram MU, Sharif M, Tariq A, ullah Yasin, U. (2018) Arteriovenous ratio and papilledema based hybrid decision support system for detection and grading of hypertensive retinopathy. *Comput Methods Programs Biomed* 154:123–141
 196. Manikis GC, Sakkalis V, Zabulis X, Karamaounas P, Triantafylou A, Douma S, Marias K (2011) An image analysis framework for the early assessment of hypertensive retinopathy signs. In: 2011 E-health and bioengineering conference (EHB), pp 1–6. IEEE
 197. Tramontan L, Ruggeri A (2009) Computer estimation of the AVR parameter in diabetic retinopathy. In: World congress on medical physics and biomedical engineering, September 7–12, 2009, Munich, Germany, pp 141–144. Springer, Berlin
 198. Narasimhan K, Neha VC, Vijayarekha K (2012) Hypertensive retinopathy diagnosis from fundus images by estimation of Avr. *Procedia Eng* 38:980–993
 199. Cavallari M, Stamile C, Umeton R, Calimeri F, Orzi F (2015) Novel method for automated analysis of retinal images: results in subjects with hypertensive retinopathy and CADASIL. *BioMed Res Int* 2015
 200. Goswami S, Goswami S, De S (2017) Automatic measurement and analysis of vessel width in retinal fundus image. In: Proceedings of the first international conference on intelligent computing and communication, pp 451–458. Springer, Singapore
 201. Chua J, Chin CWL, Hong J, Chee ML, Le TT, Ting DSW, Schmetterer L (2019) Impact of hypertension on retinal capillary microvasculature using optical coherence tomographic angiography. *J Hypertens* 37(3):572
 202. Chua J, Chin CWL, Tan B, Wong SH, Devarajan K, Le TT, Schmetterer L (2019) Impact of systemic vascular risk factors on the choriocapillaris using optical coherence tomography angiography in patients with systemic hypertension. *Sci Rep* 9(1):1–11
 203. Pascual-Prieto J, Burgos-Blasco B, Avila Sanchez-Torija M, Fernández-Vigo JI, Arriola-Villalobos P, Barbero Pedraz MA, Martínez-de-la-Casa JM (2020) Utility of optical coherence tomography angiography in detecting vascular retinal damage caused by arterial hypertension. *Eur J Ophthalmol* 30(3):579–585
 204. Hua D, Xu Y, Zeng X, Yang N, Jiang M, Zhang X, Xing Y (2020) Use of optical coherence tomography angiography for assessment of microvascular changes in the macula and optic nerve head in hypertensive patients without hypertensive retinopathy. *Microvasc Res* 129:103969
 205. Holm S, Russell G, Nourrit V, McLoughlin N (2017) DR HAGIS—a fundus image database for the automatic extraction of retinal surface vessels from diabetic patients. *J Med Imaging* 4(1):014503
 206. Agurto C, Joshi V, Nemeth S, Soliz P, Barriga S (2014) Detection of hypertensive retinopathy using vessel measurements and textural features. In: 2014 36th Annual international conference of the IEEE engineering in medicine and biology society, pp 5406–5409. IEEE
 207. Triwijoyo BK, Budiharto W, Abdurachman E (2017) The classification of hypertensive retinopathy using convolutional neural network. *Procedia Comput Sci* 116:166–173
 208. Triwijoyo BK, Pradipto YD (2017) Detection of hypertension retinopathy using deep learning and Boltzmann machines. In: Journal of physics: conference series, vol 801, no 1, p 012039. IOP Publishing
 209. AlBadawi S, Fraz MM (2018) Arterioles and venules classification in retinal images using fully convolutional deep neural network. In: International conference image analysis and recognition, pp 659–668. Springer, Cham
 210. Welikala RA, Foster PJ, Whincup PH, Rudnicka AR, Owen CG, Strachan DP, Barman SA (2017) Automated arteriole and venule classification using deep learning for retinal images from the UK Biobank cohort. *Comput Biol Med* 90:23–32
 211. Yao Z, Zhang Z, Xu LQ (2016) Convolutional neural network for retinal blood vessel segmentation. In: 2016 9th International symposium on computational intelligence and design (ISCID), vol 1, pp 406–409. IEEE
 212. Prentašić P, Lončarić S (2016) Detection of exudates in fundus photographs using deep neural networks and anatomical landmark detection fusion. *Comput Methods Programs Biomed* 137:281–292
 213. Sridhar S, Pradeep Kandasamy J, Sinthuja M, Minish TS (2021) Diabetic retinopathy detection using convolutional neural networks algorithm. In: Materials today: proceedings
 214. Qureshi I, Ma J, Abbas Q (2021) Diabetic retinopathy detection and stage classification in eye fundus images using active deep learning. *Multimedia Tools Appl* 80(8):11691–11721
 215. Qureshi I, Ma J, Shaheed K (2019) A hybrid proposed fundus image enhancement framework for diabetic retinopathy. *Algorithms* 12(1):14
 216. Diabetic Retinopathy Detection—DEV Community, (n.d.). <https://dev.to/erol/diabetic-retinopathy-detection-5ha7>. Accessed 3 June 2021
 217. Walter T, Klein JC, Massin P, Erginay A (2002) A contribution of image processing to the diagnosis of diabetic retinopathy—detection of exudates in color fundus images of the human retina. *IEEE Trans Med Imaging* 21(10):1236–1243

218. Sopharak A, Uyyanonvara B, Barman S, Williamson TH (2008) Automatic detection of diabetic retinopathy exudates from non-dilated retinal images using mathematical morphology methods. *Comput Med Imaging Graph* 32(8):720–727
219. Quellec G, Lamard M, Abràmoff MD, Decencière E, Lay B, Erginay A, Cazuguel G (2012) A multiple-instance learning framework for diabetic retinopathy screening. *Med Image Anal* 16(6):1228–1240
220. Amel F, Mohammed M, Abdelhafid B (2012) Improvement of the hard exudates detection method used for computer-aided diagnosis of diabetic retinopathy. *Int J Image Graph Signal Process* 4(4):19
221. Abràmoff MD, Folk JC, Han DP, Walker JD, Williams DF, Russell SR, Niemeijer M (2013) Automated analysis of retinal images for detection of referable diabetic retinopathy. *JAMA Ophthalmol* 131(3):351–357
222. Mookiah MRK, Acharya UR, Martis RJ, Chua CK, Lim CM, Ng EYK, Laude A (2013) Evolutionary algorithm based classifier parameter tuning for automatic diabetic retinopathy grading: A hybrid feature extraction approach. *Knowl Based Syst* 39:9–22
223. Zhang X, Thibault G, Decencière E, Marcotegui B, Laÿ B, Danno R, Erginay A (2014) Exudate detection in color retinal images for mass screening of diabetic retinopathy. *Med Image Anal* 18(7):1026–1043
224. Akram MU, Khalid S, Tariq A, Khan SA, Azam F (2014) Detection and classification of retinal lesions for grading of diabetic retinopathy. *Comput Biol Med* 45:161–171
225. Colomer A, Igual J, Naranjo V (2020) Detection of early signs of diabetic retinopathy based on textural and morphological information in fundus images. *Sensors* 20(4):1005
226. ElTanboly A, Ismail M, Shalaby A, Switala A, El-Baz A, Schaal S, El-Azab M (2017) A computer-aided diagnostic system for detecting diabetic retinopathy in optical coherence tomography images. *Med Phys* 44(3):914–923
227. Sandhu HS, Eltanboly A, Shalaby A, Keynton RS, Schaal S, El-Baz A (2018) Automated diagnosis and grading of diabetic retinopathy using optical coherence tomography. *Invest Ophthalmol Vis Sci* 59(7):3155–3160
228. Szymkowski M, Saeed E, Saeed K, Mariak Z (2019) A simple algorithm for hard exudate detection in diabetic retinopathy using spectral-domain optical coherence tomography. In: Computer graphics international conference, pp 179–189. Springer, Cham
229. Khansari MM, Zhang J, Qiao Y, Gahm JK, Sarabi MS, Kashani AH, Shi Y (2019) Automated deformation-based analysis of 3D optical coherence tomography in diabetic retinopathy. *IEEE Trans Med Imaging* 39(1):236–245
230. Sandhu HS, Elmogy M, Sharafeldeen AT, Elsharkawy M, El-Adawy N, Eltanboly A, El-Baz A (2020) Automated diagnosis of diabetic retinopathy using clinical biomarkers, optical coherence tomography, and optical coherence tomography angiography. *Am J Ophthalmol* 216:201–206
231. Schwartz R, Khalid H, Sivaprasad S, Nicholson L, Anikina E, Sullivan P, Keane PA (2020) Objective evaluation of proliferative diabetic retinopathy using OCT. *Ophthalmol Retina* 4(2):164–174
232. Sharafeldeen A, Elsharkawy M, Khalifa F, Soliman A, Ghazal M, AlHalabi M, El-Baz A (2021) Precise higher-order reflectivity and morphology models for early diagnosis of diabetic retinopathy using OCT images. *Sci Rep* 11(1):1–16
233. ElTanjoly A, Ghazal M, Khalil A, Shalaby A, Mahmoud A, Switala A, El-Baz A (2018) An integrated framework for automatic clinical assessment of diabetic retinopathy grade using spectral domain OCT images. In: 2018 IEEE 15th international symposium on biomedical imaging (ISBI 2018), pp 1431–1435. IEEE
234. Gadekallu TR, Khare N, Bhattacharya S, Singh S, Maddikunta PKR, Srivastava G (2020) Deep neural networks to predict diabetic retinopathy. *J Ambient Intell Humaniz Comput* 1–14. <https://doi.org/10.1007/s12652-020-01963-7>
235. Katada Y, Ozawa N, Masayoshi K, Ofiji Y, Tsubota K, Kurihara T (2020) Automatic screening for diabetic retinopathy in inter-racial fundus images using artificial intelligence. *Intell Based Med* 3:100024
236. Sahlsten J, Jaskari J, Kivinen J, Turunen L, Jaanio E, Hietala K, Kaski K (2019) Deep learning fundus image analysis for diabetic retinopathy and macular edema grading. *Sci Rep* 9(1):1–11
237. Gulshan V, Peng L, Coram M, Stumpé MC, Wu D, Narayanaswamy A, Webster DR (2016) Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* 316(22):2402–2410
238. Abràmoff MD, Lou Y, Erginay A, Clarida W, Amelon R, Folk JC, Niemeijer M (2016) Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. *Invest Ophthalmol Vis Sci* 57(13):5200–5206
239. Swapna G, Vinayakumar R, Soman KP (2018) Diabetes detection using deep learning algorithms. *ICT Express* 4(4):243–246
240. Hemanth DJ, Deperlioglu O, Kose U (2020) An enhanced diabetic retinopathy detection and classification approach using deep convolutional neural network. *Neural Comput Appl* 32(3):707–721
241. Shanthi T, Sabreenian RS (2019) Modified Alexnet architecture for classification of diabetic retinopathy images. *Comput Electr Eng* 76:56–64
242. Gayathri S, Gopi VP, Palanisamy P (2020) A lightweight CNN for diabetic retinopathy classification from fundus images. *Biomed Signal Process Control* 62:102115
243. Zago GT, Andreão RV, Dorizzi B, Salles EOT (2020) Diabetic retinopathy detection using red lesion localization and convolutional neural networks. *Comput Biol Med* 116:103537
244. Vives-Boix V, Ruiz-Fernández D (2021) Diabetic retinopathy detection through convolutional neural networks with synaptic metaplasticity. *Comput Methods Programs Biomed* 206:106094
245. Li C, Ye J, He J, Wang S, Qiao Y, Gu L (2020) Dense correlation network for automated multi-label ocular disease detection with paired color fundus photographs. In: 2020 IEEE 17th international symposium on biomedical imaging (ISBI), pp 1–4. IEEE
246. Chen W, Yang B, Li J, Wang J (2020) An approach to detecting diabetic retinopathy based on integrated shallow convolutional neural networks. *IEEE Access* 8:178552–178562
247. Guo Y, Camino A, Wang J, Huang D, Hwang TS, Jia Y (2018) MEDnet, a neural network for automated detection of avascular area in OCT angiography. *Biomed Opt Express* 9(11):5147–5158
248. Li X, Shen L, Shen M, Tan F, Qiu CS (2019) Deep learning based early stage diabetic retinopathy detection using optical coherence tomography. *Neurocomputing* 369:134–144
249. Ghazal M, Ali SS, Mahmoud AH, Shalaby AM, El-Baz A (2020) Accurate detection of non-proliferative diabetic retinopathy in optical coherence tomography images using convolutional neural networks. *IEEE Access* 8:34387–34397
250. Heisler M, Karst S, Lo J, Mammo Z, Yu T, Warner S, Sarunic MV (2020) Ensemble deep learning for diabetic retinopathy detection using optical coherence tomography angiography. *Transl Vis Sci Technol* 9(2):20–20
251. Kermany DS, Goldbaum M, Cai W, Valentim CC, Liang H, Baxter SL, Zhang K (2018) Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell* 172(5):1122–1131
252. Lu W, Tong Y, Yu Y, Xing Y, Chen C, Shen Y (2018) Deep learning-based automated classification of multi-categorical abnormalities from optical coherence tomography images. *Transl Vis Sci Technol* 7(6):41–41

253. Islam KT, Wijewickrema S, O'Leary S (2019) Identifying diabetic retinopathy from oct images using deep transfer learning with artificial neural networks. In: 2019 IEEE 32nd international symposium on computer-based medical systems (CBMS), pp 281–286. IEEE
254. Shaban M, Ogur Z, Mahmoud A, Switala A, Shalaby A, Abu Khalifeh H, El-Baz AS (2020) A convolutional neural network for the screening and staging of diabetic retinopathy. PLoS ONE 15(6):e0233514
255. Abbas Q, Qureshi I, Ibrahim ME (2021) An automatic detection and classification system of five stages for hypertensive retinopathy using semantic and instance segmentation in DenseNet architecture. Sensors 21(20):6936
256. Chan HP, Hadjiiski LM, Samala RK (2020) Computer-aided diagnosis in the era of deep learning. Med Phys 47(5):e218–e227
257. Georgiou T, Liu Y, Chen W, Lew M (2020) A survey of traditional and deep learning-based feature descriptors for high dimensional data in computer vision. Int J Multimedia Inf Retr 9(3):135–170
258. Nguyen ND, Do T, Ngo TD, Le DD (2020) An evaluation of deep learning methods for small object detection. J Electr Comput Eng 2020
259. Wong KK, Fortino G, Abbott D (2020) Deep learning-based cardiovascular image diagnosis: a promising challenge. Future Gener Comput Syst 110:802–811
260. Hong S, Zhou Y, Shang J, Xiao C, Sun J (2020) Opportunities and challenges of deep learning methods for electrocardiogram data: a systematic review. Comput Biol Med 122:1–15
261. Costa P, Galdran A, Meyer MI, Abramoff MD, Niemeijer M, Mendonça AM, Campilho A (2017) Towards adversarial retinal image synthesis. arXiv preprint <http://arxiv.org/abs/1701.08974>
262. Korot E, Guan Z, Ferraz D, Wagner SK, Zhang G, Liu X, Keane PA (2021) Code-free deep learning for multi-modality medical image classification. Nat Mach Intell 3(4):288–298
263. Zhang Q, Bai C, Liu Z, Yang LT, Yu H, Zhao J, Yuan H (2020) A GPU-based residual network for medical image classification in smart medicine. Inf Sci 536:91–100
264. Yadav SS, Jadhav SM (2019) Deep convolutional neural network based medical image classification for disease diagnosis. J Big Data 6(1):1–18
265. Zhang J, Xie Y, Wu Q, Xia Y (2019) Medical image classification using synergic deep learning. Med Image Anal 54:10–19
266. Shams R, Sadeghi P, Kennedy RA, Hartley RI (2010) A survey of medical image registration on multicore and the GPU. IEEE Signal Process Mag 27(2):50–60
267. Goodfellow I, Bengio Y, Courville A (2016) Deep learning. MIT Press
268. Diaz-Pinto A, Colomer A, Naranjo V, Morales S, Xu Y, Frangi AF (2019) Retinal image synthesis and semi-supervised learning for glaucoma assessment. IEEE Trans Med Imaging 38(9):2211–2218

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