Open Access ISSN: 3007-1437 (Online), ISSN: 3007-1429 (Print)

IRAQI Academic Scientific Journals

Alkadhim Journal for Computer Science (KJCS)

Journal Homepage: https://alkadhum-col.edu.iq/JKCEAS

KJCS ALKADHIM JOURNAL FOR COMPUTER SCIENCE

A Comprehensive Review: AI Techniques in Ophthalmic Imaging Analysis

[1]Hayder Rahman Abdulmohsin, [2] Saif Ali Abd Alradha Alsaidi

[1, 2] College of Computer Science and Information Technology, Wasit University, Kut – Iraq

Article information

Article history:
Received: April, 27, 2026
Accepted: May, 23, 2026
Available online: September, 25, 2026

Keywords:
Computer-Aided Diagnosis (CAD),
Computer Vision,
Ophthalmic Image Analysis,
Medical Image Classification,
Automated Diagnosis,
Ocular Diseases

***Corresponding Author:**
habdulmohsin@uowasit.edu.iq

DOI:
https://doi.org/10.61710/sm8wp171

Abstract

In recent years, there has been significant progress in Artificial Intelligence (AI)-based technologies that can facilitate the accurate analysis of ophthalmic images. With the help of computer vision, the early diagnosis of eye diseases such as Diabetic Retinopathy (DR), cataracts, glaucoma, and Age-Related Macular Degeneration (AMD) has become possible. Nevertheless, overlapping symptoms and signs appear in different clinical cases, making it difficult for health care providers to make an accurate diagnosis. As a result, patients might fail to get early treatment, resulting in worse outcomes and the rise in the incidence of these diseases. Moreover, conventional methods are usually costly, l abor-intensive, and require the involvement of a professional ophthalmologist who has sufficient knowledge and skills to diagnose the condition. The objective of this paper is to conduct a literature review on the current status of AI technology-based approaches used to differentiate among the most prevalent types of ocular diseases by exploring relevant literature published between 2020 and 2025, analyzing the benefits and drawbacks of proposed methods and comparing them to other existing tools, and describing popular databases that researchers use to test the performance of their models. An important finding of this study is the ability of modern computer-aided diagnosis (CAD) systems to accurately identify vision-threatening diseases, which may outperform expert ophthalmologists. On the other hand, there are several challenges that need to be addressed in future studies, such as data imbalance, generalizability, validation, and expensive computation.

1 Introduction

Vision is an important component of people's overall health status and quality of life; however, ocular diseases that endanger vision are a pressing problem globally. The World Health Organization (WHO) estimates that approximately 2.2 billion people worldwide suffer from visual impairment or blindness, with a minimum of a billion preventable cases [1]. It is therefore crucial that early and correct diagnoses be made, as they have proven to effectively minimize the risk of visual loss [2]. Traditionally, the diagnosis of such diseases is performed manually by qualified ophthalmologists based on the analysis of retinal fundus photographs and optical coherence tomography (OCT) [3]. Although human analysis remains the clinical standard, it is highly prone to inter-rater variability and is time-consuming to perform. Moreover, it is hampered by an extreme shortage of professionals in the field, especially in low- and middle-income countries [4].

Therefore, ophthalmology is currently experiencing a revolution in AI-based CAD systems [5]. Previous ML techniques require handcrafted features and classical classifiers that rely on expert knowledge and have issues with generalization owing to the laborious process of designing their features [6]. This has led to the emergence of DL techniques, specifically CNNs. They allow the automatic extraction of features from medical images and better diagnostic results [7], thus providing improved support for clinical decision-making [8]. Still, DL models require massive annotated datasets, which are usually hard to gather because of clinical and privacy restrictions. As a result, researchers are increasingly turning to transfer learning (TL) to adapt pretrained models for smaller ophthalmic datasets [9]. Apart from DL architectures, more recent developments include the introduction of attention-based algorithms, vision transformers (ViT), hybrid CNN-transformer models, and ensemble models that offer distinct advantages over classic CNN models [10], [11]. Despite these advances, several challenges remain in applying AI to clinical practice, including limited model interpretability ("black box" problem), dataset bias due to unrepresentative patient cohorts, and the absence of comprehensive guidelines for AI-assisted diagnostic devices [12], [13].

In this review, we aim to comprehensively and concurrently examine the role of ML and DL techniques used in ophthalmic image analysis across glaucoma (G), diabetic retinopathy (DR), cataract (C), and age-related macular degeneration (AMD), focusing on studies published between 2020 and 2025. We also systematically appraise the datasets employed and critically address challenges and limitations related to clinical translation. Finally, we review current developments and highlight future trends in the field of ophthalmology. The organization of this review is shown in **Figure (1)**.

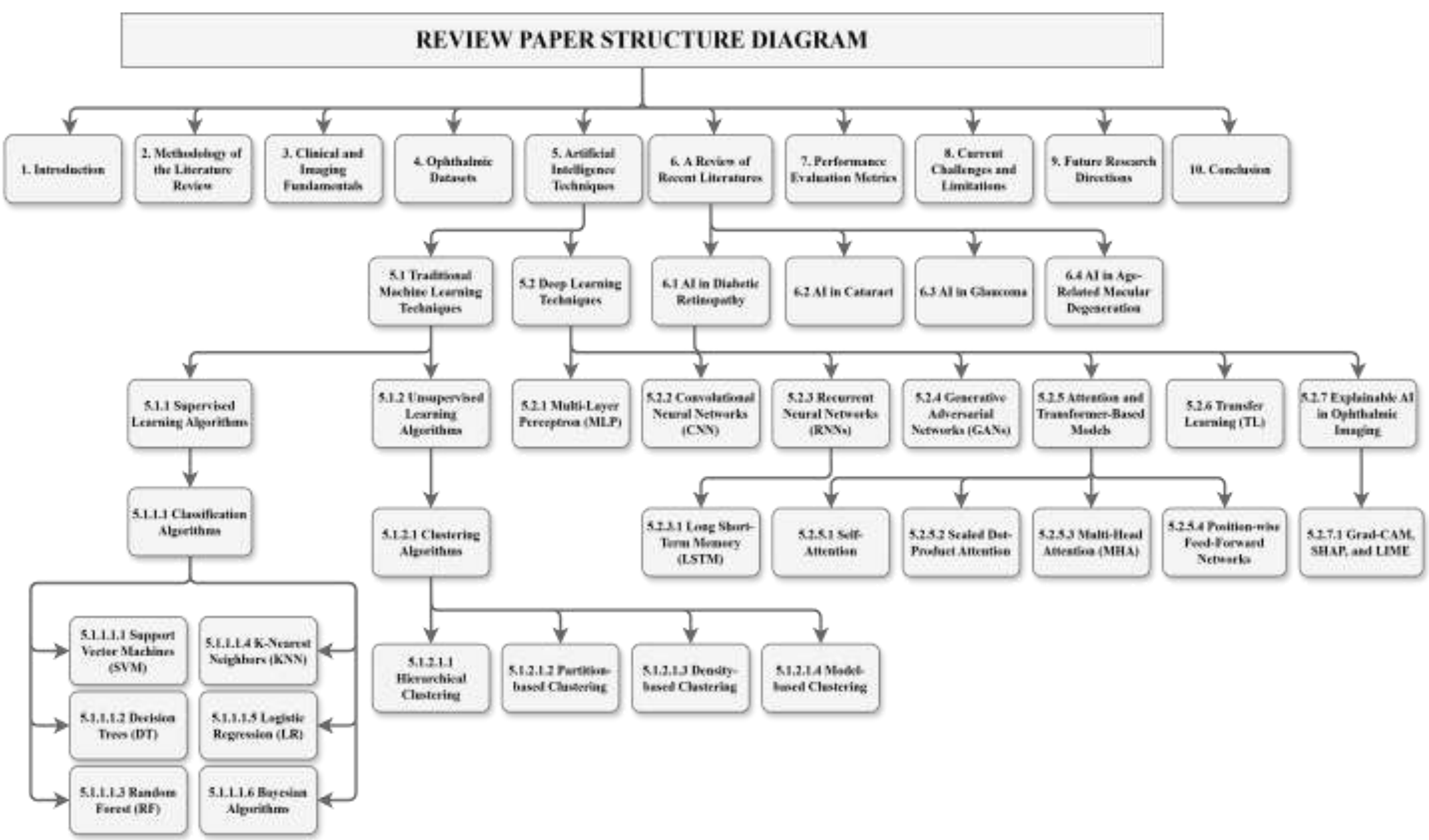

Figure (1): The Paper Structure Diagram.

2 Methodology of the Literature Review

This study adopts a structured narrative literature review approach to examine the most recent advances in AI applications across eye health. Relevant scientific studies were obtained from various scientific databases from 2020 to 2025, such as Web of Science, IEEE Xplore, Science-Direct, PubMed, Springer-Link, MDPI, and arXiv. Google Scholar was used as a supplementary search tool. Search terms related to AI, deep learning, retinal imaging, OCT classification, fundus image analysis, and AI in ophthalmology were utilized in combination with Boolean operators (AND/OR). Selection criteria included relevance, methodological clarity, and contribution to AI-based ophthalmic diagnosis. Articles chosen for further investigation were peer-reviewed empirical studies, systematic reviews, and methodological articles. Additional representative ML/DL studies were included to provide a broader methodological perspective on medical image analysis. Non-English, duplicate, non-peer-reviewed, and methodologically unclear studies were excluded. All references were organized with Zotero in order to keep references consistent and accurate.

3 Clinical and Imaging Fundamentals

The retina is a vital part of the eye involved in the conversion of light stimuli to neural signals with the help of photoreceptor cells (rods and cones) to allow visual perception through the optic nerve that sends impulses to the brain's visual cortex [14] (see **Figure (2)**). Structurally, the retina measures an approximate length of 32 mm and has a surface area of 1094 mm² based on typical measurements [15]. Important anatomical structures of the retina include the optic disc, which has an average vertical measurement of 1.86 mm and an average horizontal measurement of 1.75 mm [16], [17]. The fovea is a depression within the macula measuring 1.5 mm across and contains a dense concentration of photoreceptors for high-resolution and color vision [18]. Additionally, macular pigmentation increases the retina's ability to absorb light and improve acuity [19]. Fundus imaging is a process by which artificial intelligence algorithms detect and diagnose conditions like diabetic retinopathy, macular edema, and glaucoma among others [20].

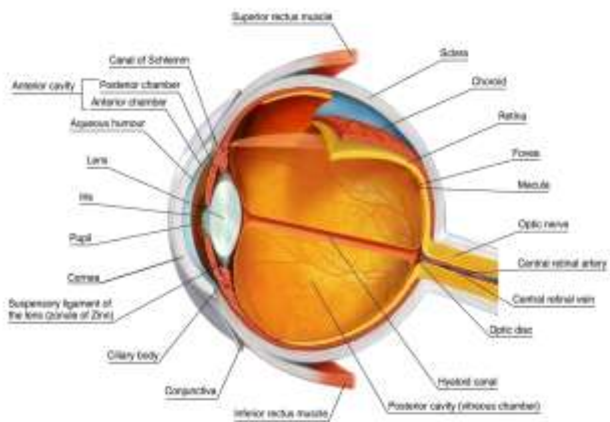

Figure (2): Anatomy of the human eye [21].

Diagnosing and monitoring eye diseases rely heavily on advanced ophthalmic imaging modalities. To evaluate ocular conditions in a clinical practice, professionals employ a variety of methods, such as Optical Coherence Tomography (OCT) [22], Fundus Fluorescein Angiography (FFA) [23], and 2D Color Fundus Photography (CFP) [24], in addition to other imaging techniques. OCT and CFP seem to be the most often utilized of them (a comparison is presented in **Figure (3)**). OCT has become a vital tool in ophthalmology as well as, in certain situations, in imaging other organs like the brain. It gives cross-sectional images of the retina, enabling doctors to quantify retinal thickness and assess structural changes. In contrast, CFP takes photographs of the inner surface of the eye and is frequently used to record and track retinal abnormalities throughout time, both modalities have proven useful for early detection of ocular diseases, although CFP is generally considered a more cost-effective and accessible option in routine clinical settings [25].

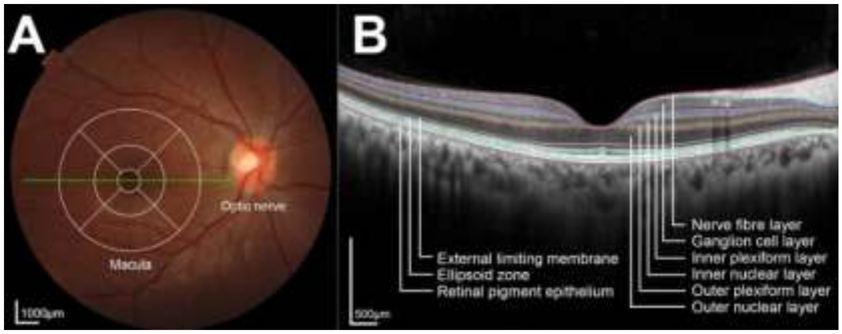

Figure (3): Comparison of retinal imaging modalities: (**A**) Fundus photograph, (**B**) Optical Coherence Tomography (OCT) [26].

Such imaging technologies play an important role in the diagnosis, monitoring, and treatment of the most common ocular diseases that cause visual impairment. The clinical features, pathology, and imaging aspects of four common ocular diseases have been presented below.

- **Diabetic Retinopathy (DR):** A diabetic micro-vascular complication associated with retinal damage in the form of micro-aneurysm (MA), hemorrhage (H), hard exudates (EX), and cotton-wool spot (CWS). The disease starts from non-proliferative (NPDR) to proliferative stages (PDR) [27], [28]. It is typically asymptomatic until advanced stages; hence, imaging is crucial for early diagnosis to avoid loss of vision [29]. As illustrated in **Figure (4)**.

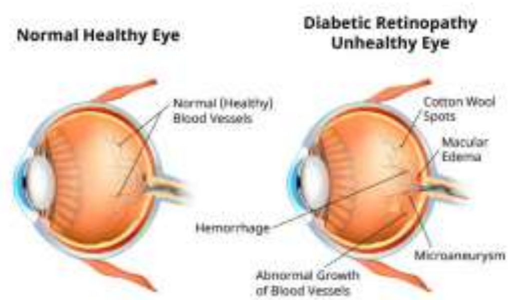

Figure (4): A normal healthy eye and an eye affected by DR [30].

- **Glaucoma:** This is a progressive optic neuropathy that occurs due to high intraocular pressure (IOP) levels resulting in optic nerve damage [5]. Clinically, it is identified through structural changes in the optic nerve head, commonly evaluated using the cup-to-disc ratio (with values exceeding 0.5 often considered suspicious) [18], as illustrated in **Figure (5)**. Glaucoma has an early asymptomatic stage, thus requiring image-based diagnosis [32].

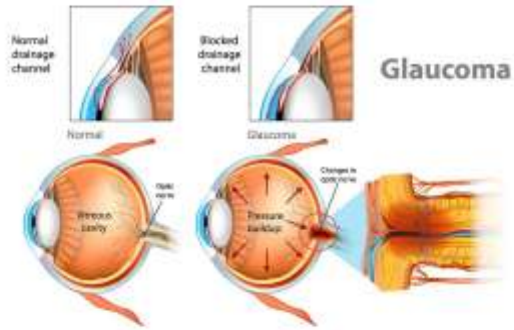

Figure (5): Glaucoma and its effect on the normal eye [33].

- **Cataract:** The development of a disease where there is a progressive opacity of the lens. This causes poor transmission of light and vision [34], [35]. It is a disease linked to old age, among others, and can be detected using slit lamp imagery [29], [36]. Different from other diseases, vision impairment caused by cataracts is generally reversible with surgery [29], **Figure (6)**.

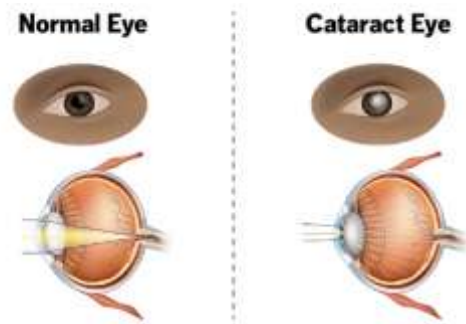

Figure (6): Difference between the normal eye and cataract [37].

- **Age-related Macular Degeneration (AMD):** A progressive retinal disease affecting the macula and central vision, including dry (drusen accumulation) and wet (choroidal neovascularization, CNV) forms [38]. Early stages are asymptomatic, requiring imaging techniques such as fundus photography and OCT for detection [39]. Early diagnosis is critical to prevent severe vision loss [29], [40], **Figure (7)**.

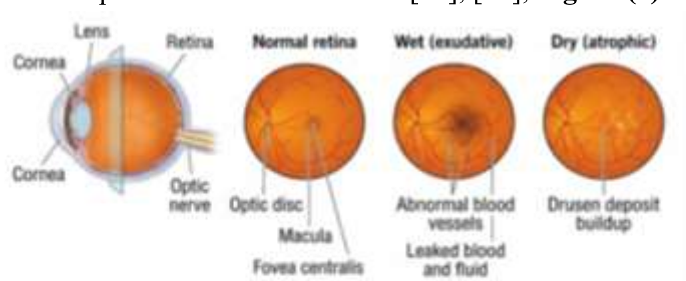

Figure (7): Both the wet (exudative) and dry (atrophic) forms of macular degeneration affect the macula [40].

4 Ophthalmic Datasets

Datasets employed in ophthalmology are fundamental components of AI-assisted diagnostic tools for identifying and diagnosing diseases. Nonetheless, those ophthalmology data are associated with challenges that impact the success rate of these models in the clinic.

To begin with, problems associated with poor image quality such as blurriness, noisiness, lack of illumination, motion artifacts, and acquisition issues occur frequently in ophthalmological datasets [41], [42]. In addition, low-quality scans are rejected before being used in several benchmark datasets, resulting in high performance at the expense of the model's robustness. In terms of annotation, labeling ophthalmological images according to standardized grading scales is an arduous task due to the expensive cost and time-consuming process involved [41]. Hence, noisy labels and weak annotations could negatively impact generalizability, increasing the over-fitting risk even when state-of-the-art architectures are utilized [43]. The issue of class imbalance, in which the prevalence of one category can bias the model, is another significant challenge for reliable clinical use, as the model may become biased towards majority classes and thus fail to detect minority conditions correctly [44]. Datasets including APTOS 2019 are examples of class imbalance due to imbalances between different degrees of severity [45]. In turn, multi-disease datasets like ODIR contain diseases in long-tailed distributions, where specific pathologies are underrepresented compared to other dominant classes [46].

A review of available ophthalmic imaging datasets further revealed that retinal fundus images constitute the majority (57%), followed by optical coherence tomography (OCT) and OCT angiography (19%), whereas other imaging modalities are less represented [47]. These distributional differences reflect the broader challenges of dataset diversity discussed previously. **Table (1)** presents the most widely used datasets for ophthalmic image analysis, categorized by imaging modality, target disease, and key characteristics.

Table (1): Summary of ophthalmic imaging datasets used in AI-based analysis.

Source Type	Disease or Task	Dataset Name	Number of images	Country of Origin
Color Fundus Photography (CFP)	Vessel Segmentation Datasets	DRIVE [48]	40	Netherlands
		STARE [49]	397	USA
		CHASE_DB1 [50], [51]	28	UK
		HRF [52]	45	Germany and Czech Republic
		IOSTAR [53]	30	---
	Diabetic Retinopathy Datasets	MESSIDOR [54]	1,200	France
		MESSIDOR-2 [55]	1,748	France
		EPACS [56]	88,702	USA
		DDR [57]	12,524	China
		IDRiD [58]	512	India
		APTOS2019 [59]	5590	India
		DiaRetDB1 V2.1 [60]	89	Finland
	Glaucoma Datasets	ORIGA [61], [62]	650	Singapore
		REFUGE Challenge 2020 [63]	1200	China
		ACRIMA [64]	705	Spain
		DRISHTI-GS [65], [66]	101	India
		RIM-ONE R1 [67]	169	Spain
		RIM-ONE R2 [67]	455	Spain
		RIM-ONE R3 [67]	159	Spain
		RIM-ONE DL [67]	485	Spain
		OFOAOS [68]	3909	China
		G1020 [69]	1020	Germany
	Myopia	PALM [70]	1200	China
		HPMI [71]	4011	China
	Multi-Diseases	RFMiD [72], [73]	3200	India
		JSIEC [74]	1000	China
		ODIR [75], [76]	10,000	China
Optical Coherence Tomography (OCT)		OCT2017 [77]	109,309	USA
		Duke's OCT [78]	---	USA
		OCTID [79], [80]	572	---
		OCTDL [81]	2,064	Russia
		2015_BOE_Chiu2 [82]	110	USA
		NEH_UT_2021RetinalOCT [83]	16822	Iran

Glossary: DRIVE: Digital Retinal Images for Vessel Extraction, STARE: STructured Analysis of the REtina, CHASE_DB1: Child Heart and Health Study in England Database 1, HRF: High-Resolution Fundus image database, IOSTAR: IOSTAR Vessel Segmentation, MESSIDOR: Methods to Evaluate Segmentation and Indexing Techniques in the field of Retinal Ophthalmology, MESSIDOR-2: Methods to Evaluate Segmentation and Indexing Techniques in the field of Retinal Ophthalmology version 2, EPACS: EyePACS Diabetic Retinopathy, DDR: Diabetic retinopathy Detection and Reporting, IDRiD: Indian Diabetic Retinopathy Image, APTOS2019: Asia Pacific Tele-Ophthalmology Society 2019 Blindness Detection, DiaRetDB1 V2.1: Diabetic Retinopathy Database and Evaluation Protocol version 2.1, ORIGA: Online Retinal Fundus Image Database for Glaucoma Analysis and Research, REFUGE: Retinal Fundus Glaucoma Challenge, ACRIMA: ACRIMA, DRISHTI-GS: DRISHTI Glaucoma Screening, RIM-ONE R1: Retinal Image database for Optic Nerve Evaluation Release 1, RIM-ONE R2: Retinal Image database for Optic Nerve Evaluation Release 2, RIM-ONE R3: Retinal Image database for Optic Nerve Evaluation Release 3, RIM-ONE DL: Retinal Image database for Optic Nerve Evaluation Deep Learning edition, OFOAOS: Optic Fundus image for Ocular disease AI screening and Optic disc Segmentation, G1020: Glaucoma 1020 fundus images, PALM: PAthologic Myopia challenge, HPMI: High-resolution Pathological Myopia Image, RFMiD: Retinal Fundus Multi-Disease image, JSIEC: Joint Shantou International Eye Center, ODIR: Ocular Disease Intelligent Recognition, OCT2017: Kermany OCT 2017, Duke's OCT: Duke University OCT Retinal Layer Segmentation, OCTID: Optical Coherence Tomography Image Database, OCTDL: Optical Coherence Tomography Dataset for Image-Based Deep Learning Methods, 2015_BOE_Chiu2: Biomedical Optics Express 2015 Chiu et al. OCT Layer Segmentation version 2, NEH_UT_2021RetinalOCT: Noor Eye Hospital and University of Tehran 2021 Retinal OCT

5 Artificial Intelligence Techniques

Ophthalmology is one of the most AI-ready medical specialties, and the field has seen rapid, transformative advances. Here's an introduction to the key techniques.

5.1 Traditional Machine Learning Techniques

Machine Learning (ML), as a sub-discipline of artificial intelligence, has its basis in the recognition of patterns in data. The repeated patterns in the data are recognized by employing statistical and computational techniques [84]. The machine learning algorithms handle the complete range of data processing activities such as clustering, regression analysis, anomaly detection, classification, prediction, and decision-making in accordance with the recognized patterns. The basic concepts of machine learning can be described in terms of the following three paradigms of learning. The supervised learning approach involves the training of the model on the data set to recognize the relationship between the input and the corresponding output. The unsupervised learning approach involves the training of the model on the data set without the corresponding outputs. The model then attempts to recognize the underlying structure or pattern in the data. The reinforcement learning approach involves the use of trial and error in the decision-making process. The model attempts to optimize the decision-making process by the use of reward or a penalty in the feedback [85], [86]. Apart from the above three paradigms of machine learning, the semi-supervised approach has been proposed to use the combination of both labeled and unlabeled data [87].

5.1.1 Supervised learning: Classification Algorithms

Classification algorithms are designed to assign input data into predefined categories or classes based on learned patterns from labeled training data. These algorithms can be categorized according to their underlying mathematical principles and computational strategies.

5.1.1.1 Support Vector Machines (SVM)

Support Vector Machines (SVMs) are machine learning techniques that utilize the representation of data in terms of vectors within a high-dimensional feature space. In such a setting, an SVM algorithm aims to identify the best possible hyperplane that separates distinct classes and maximizes the margin between them [88]. The margin represents the shortest distance between the dividing line between classes and the closest points belonging to each class. Such points are known as support vectors. This point needs to be emphasized because these vectors play a crucial role in the creation of hyper-planes [89]. SVM theory is based on structural risk minimization. Unlike traditional methods that focus solely on the empirical error minimization, this method attempts to find an upper bound of the generalization error. Because of that, SVMs demonstrate very good performance in terms of generalization capabilities [88]. These basic ideas can be demonstrated using **Figure (8)**.

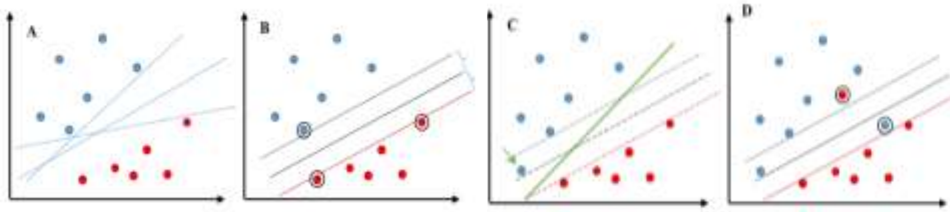

Figure (8): The maximum margin classifier is illustrated in figures **A** and **B**, whereas the use of a slack variable in figures **C** and **D** helps the support vector classifier to obtain maximum margin regardless of the influence of nearby points even when the data set is not separable [90].

5.1.1.2 Decision Trees (DT)

Decision Trees (DTs) are hierarchical models with multiple components working in concert to make interpretable predictions. The DT model has a root, internal nodes, edges, and leaves. The internal nodes make decisions based on particular features, and the edges show the results of these decisions [91]. The leaves contain the final predicted results, which can be class labels for classification models and numerical values for regression models [92]. DTs are built through an algorithm called Top-Down Induction of Decision Trees (TDIDT), which follows a divide-and-conquer approach. The process begins at the root, where all instances are present, and recursively selects features to partition the data into more homogeneous sets [93]. The process continues until specific termination criteria are met, such as maximum depth, minimum number of instances at nodes, and when all instances in a set belong to the same class [91]. To make predictions, DTs work by traversing from the root to the leaves based on the decision process determined by input features. The final prediction is made when all decisions are traversed to reach a leaf [94]. As illustrated in **Figure (9)** the basic architecture of a Decision Tree.

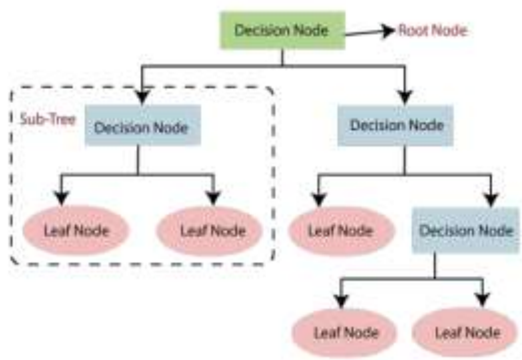

Figure (9): Structure of the decision tree (DT) algorithm [95].

5.1.1.3 Random Forest (RF)

Random Forest (RF) generates multiple decision trees and combines their output to generate an outcome, as illustrated in **Figure (10)**. The first step is bootstrap sampling, whereby N samples are randomly drawn with replacement from the initial dataset to form a new training set for each tree [96]. This technique, termed bagging, injects randomness, making the training dataset different for each tree [97]. Secondly, random feature selection is another critical aspect that sets apart RF from ordinary decision trees. Instead of selecting all features at every node, a small number of features are chosen randomly to perform the splitting operation in RF. This number is usually p, where p denotes the square root of the overall number of features in a classification problem. This feature reduces the correlation between different trees and increases the generalizability of the model [91]. The third component involves full-depth tree construction, where RF trees are grown to their maximum size without pruning. These trees operate independently and perform binary splitting using specific threshold values [98]. Finally, predictions from all trees are combined using majority voting for classification and averaging for regression tasks [91].

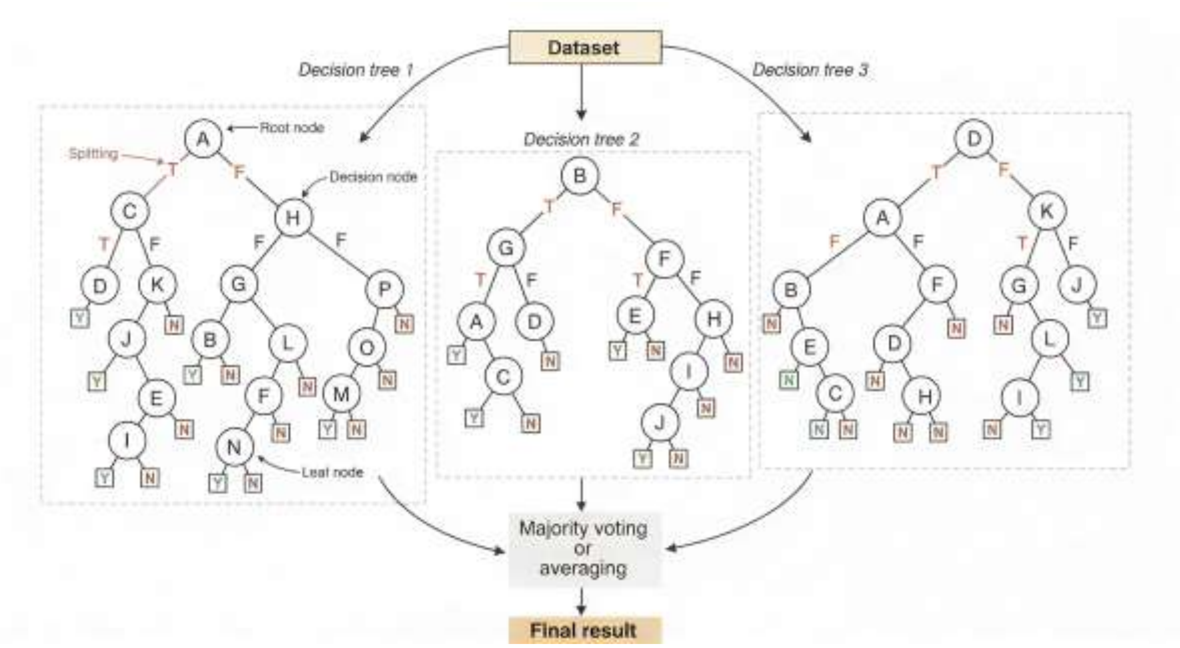

Figure (10): An example of a random forest method using three decision trees. Either majority voting or the average is used to summarize results across decision trees [97].

5.1.1.4 K-Nearest Neighbors (KNN)

One of the simplest models of machine learning includes K-Nearest Neighbors (KNN), which can be used for classification and regression. The procedure requires several steps that include the definition of the number of nearest neighbors, determination of distances between the input point and all points, finding points with the nearest distances, and finally predicting based on such points [99]. In this approach, an input point will be classified based on the comparison between the point and all points from the list of data that have been stored. In the case of classification, prediction is made through the application of majority voting. Prediction in the case of regression will depend on the averaging of all points. The value of k determines the flexibility and bias of the model. With small k, the model becomes very flexible, while large values make the decision boundaries smooth [100]. **Figure (11)** visually outlines the fundamental steps of the KNN classification process, from plotting the initial data to the final voting mechanism.

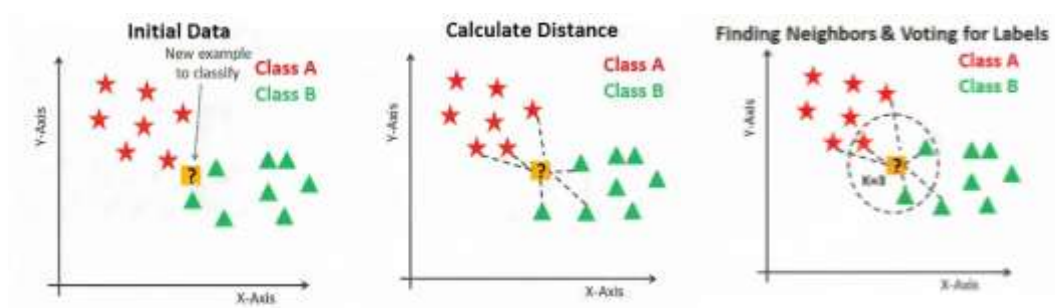

Figure (11): K-nearest neighbors [101].

5.1.1.5 Logistic Regression (LR)

The logistic regression (LR) is a generalized linear model, estimating the probability of an outcome using a logit link function that relates a linear combination to its output, as shown in **Figure (12)**. Practically, the probability of occurrence of a particular class label is estimated using the sigmoid activation function that maps the linear output into the range [0,1] [102]. The estimation of such probabilities proves helpful for decision-making processes. The logistic regression can work with numerical and categorical input data, where the predicted labels result from the estimation of the probability of belonging to each class using the logistic function [103]. According to the task type, different versions of the logistic regression can be utilized. Binary LR is the common version for the problems that have two class labels where the inputs are mapped into the interval (0,1) using the logistic function [104]. MLR is the extension of binary LR for multi-class problems, using different logistic functions for the calculation of the probabilities of occurrence of different class labels [103].

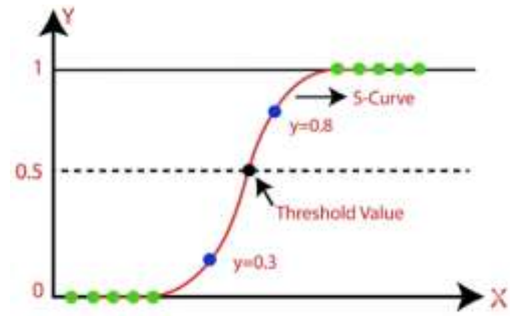

Figure (12): Logistic Regression [105].

5.1.1.6 Bayesian Algorithms

A Bayesian classifier operates within a probabilistic model of decision-making wherein the result is derived from hypothesis probability estimation. Contrary to other techniques that use handpicked features, Bayesian classification makes use of all the features available, thus improving predictive ability [106]. Within the Bayesian classification family, the Naïve Bayes Classifier technique is arguably the most commonly used algorithm for classification, as illustrated in **Figure (13**). It uses labeled data to classify data instances into classes. The joint probability of features and labels is estimated during training. While making predictions, Bayes' theorem is used to estimate the posterior probability, after which the class with the greatest probability is selected [107]. Naïve Bayes' algorithm possesses several benefits, including great performance in higher dimensional spaces and efficiency when there is insufficient training data [108]. In some situations, Naïve Bayes performs equally well or even better than more complicated models, such as decision trees and neural networks [106].

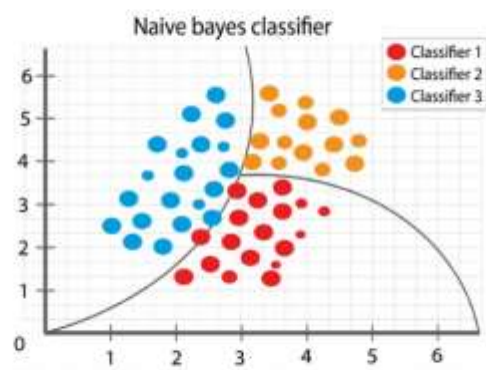

Figure (13): Naive Bayes Classifier [105].

5.1.2 Unsupervised Learning: Clustering Algorithms

The main aim of clustering algorithms is to partition a dataset into a number of clusters based on the intrinsic similarities of the data objects. The data samples inside one cluster should be similar to each other, while those belonging to different clusters are dissimilar from one another. Various categories of clustering algorithms have been proposed, which differ in terms of their approaches to identifying these clusters.

5.1.2.1 Hierarchical Clustering

Hierarchical clustering is one of the most common unsupervised machine learning techniques that demonstrates relationships between data elements with the help of tree structures, allowing for analysis on different scales [109]. Hierarchical clustering provides a lot of advantages compared to other unsupervised learning techniques, such as k-means clustering, since it does not require any prior knowledge about the required number of clusters [110].

Hierarchical clustering has two algorithms namely agglomerative and divisive. The agglomerative algorithm is a clustering algorithm that starts with each data element belonging to its own cluster and merges similar clusters one after another into a single cluster [111]. Meanwhile, the divisive algorithm takes the assumption that at the beginning, there is a single cluster containing all the data elements and then recursively partitions the clusters till every data element belongs to its own cluster [112]. Dendrograms can be used to visualize the hierarchy created with the help of agglomerative and divisive algorithms [113].

5.1.2.2 Partition-based Clustering

These algorithms are mainly characterized by partitioning of the dataset into various clusters through the use of an objective function that mainly involves the minimization of intra-cluster variation and maximization of inter-cluster differences [114]. Traditionally, there were two main types of partition-based data clustering methodologies, namely hard data clustering and soft data clustering. The hard data clustering approach involves partitioning the data points into clusters where no overlaps between clusters exists. Soft data clustering involves partitioning of data points into clusters such that the data points belong to different clusters [115]. The most commonly applied data clustering technique among the hard data clustering approaches is the K-means approach since its implementation is relatively easy [116]. This technique involves partitioning of the dataset into different clusters referred to as centroids or clusters. The algorithm mainly seeks to minimize the variances within clusters through partitioning based on the Euclidean distance metric [117]. Unlike other data clustering approaches such as hierarchical data clustering, partition-based data clustering approaches are favored when dealing with very large datasets due to their computationally efficient nature [115]. However, partition-based data clustering algorithms suffer from some limitations, including the challenge of dealing with non-spherical clusters, sensitivity to initialization of clusters, and the existence of different data points referred to as outliers [118].

5.1.2.3 Density-based Clustering

The density-based clustering method is one of the essential techniques in cluster analysis that defines clusters based on high-density regions that are surrounded by regions of low density. In contrast to other clustering algorithms, it does not rely on the number of clusters as input, is capable of capturing clusters of different sizes and arbitrary shapes, and is robust to noise while simultaneously detecting outliers [119]. The two widely applied density-based clustering algorithms are Density-Based Spatial Clustering of Applications with Noise (DBSCAN) and Ordering Points to Identify the Clustering Structure (OPTICS).

- **DBSCAN**: The algorithm is based on recursively labeling points as either core points, border points, or noise points (outliers) [120]. It allows for detecting clusters of arbitrary shapes and sizes without specifying the number of clusters beforehand [121]. However, it shows high dependence on pre-defined parameters [122].
- **OPTICS**: The algorithm was created to address some of DBSCAN's weaknesses, namely its parameter sensitivity and difficulty handling varying densities. It employs core and reachability distances to assign ranks to the data points, resulting in a reachability graph with valleys denoting possible clusters [121]. This technique enables finding non-spherical and nested clusters without knowing the number of clusters beforehand. Yet, OPTICS has several downsides such as complicated results interpretation, high memory usage for nearest-neighbor calculations, and extra effort for creating reachability graphs [123].

5.1.2.4 Model-based Clustering

Model-based clustering approaches rely on the assumption that the data set used for clustering consists of a combination of various probability distributions, where each distribution corresponds to a distinct cluster [124]. The most widely employed clustering approach from among the model-based clustering approaches is the Gaussian Mixture Model (GMM) [121]. The fundamental difference between the model-based and non-model-based clustering approaches is that while the former allows for the determination of the posterior probability of the clusters, which gives a better representation of the uncertainty associated with the clusters, the latter does not provide this information. Nonetheless, the primary premise behind clustering in both cases is similar: determining the parameters corresponding to the clusters and to which cluster an element belongs using maximum likelihood estimates [124].

5.2 Deep Learning Techniques

Deep Learning (DL) is considered an emerging technology, and it is a core technique behind the current CAD systems [27], [125]. While other machine learning techniques require hand-crafted feature extraction, the deep learning algorithm allows itself to automatically discover the discriminative features to be used for predictions [126], as illustrated in **Figure (14a)**. The performance of conventional machine learning and deep learning techniques will improve as the amount of training data increases. Conventional learning techniques will show improvements in performance during the initial stages as more training data samples become available; yet, the performance stops improving as the techniques hit the limit in capturing complex features because of their limitations in capacity. On the other hand, deep learning techniques show continuous improvements as the number of data samples used in training increases, **Figure (14b)** [127].

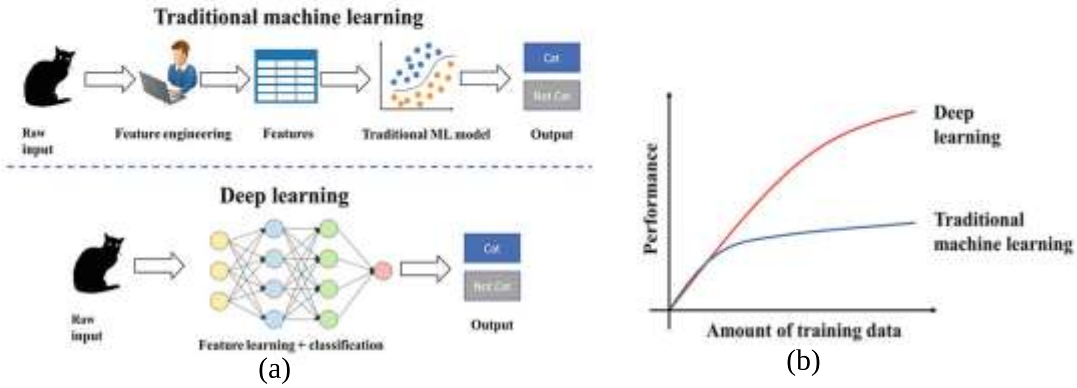

Figure (14): (**a**) ML vs DL, (**b**) Performance curves of DL and ML with increasing amount of training data [127].

5.2.1 Multi-Layer Perceptron (MLP)

The basic multi-layer perceptron is an essential component of numerous types of artificial neural networks used in deep learning applications [128]. The multi-layer perceptron architecture involves three layers: input, hidden, and output layers. The initial input layer includes the inputs to the network. The inputs propagate forward to the hidden layer where computations take place. The output layer computes the outcome or decision based on the computations in the hidden layers [129].

The essential property of the basic multi-layer perceptron is its full connection structure. In the fully connected architecture, all neurons of one layer are connected to all neurons of the adjacent layer [129]. All connections have their weights, determining how much impact signals transmitted between the neurons have. In the context of a single neuron, the inputs are being weighted and aggregated into one value. The bias is calculated from the aggregation of weighted values. Then, the aggregation and the bias are used in combination with an activation function for output computation [130]. Different activation functions like sigmoid, tanh, and rectified linear units (ReLU) provide the nonlinear activation necessary for learning complex patterns [131]. Due to the hierarchical architecture, features are progressively extracted by each consecutive layer of a perceptron, thus making the process of learning more effective. An example of the single neuron perceptron is depicted in **Figure (15a)**; the activation function (σ) is a nonlinear function of the weighted inputs y, which are calculated based on Eq. (1) [129].

$$y = \sigma(xw + b) \qquad \text{Eq. (1)}$$

The input vector, weighting vector, bias, and output value are denoted by the terms x, w, b, and y, respectively. The Multi-Layer Perceptron (MLP) model's structure is shown in **Figure (15b)**.

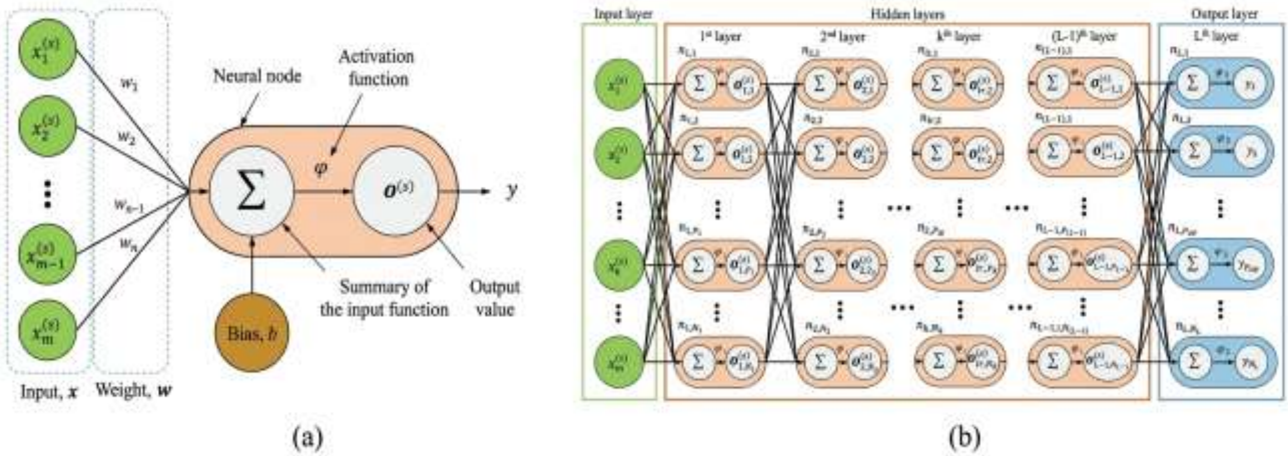

Figure (15): (**a**) Perceptron model with a single neuron. (**b**) MLP Structure [129].

5.2.2 Conventional Neural Networks (CNN)

The fundamental architecture of a CNN is based on its hierarchical nature, with the building blocks arranged in layers primarily comprising convolutional, pooling, and fully-connected (FC) layers [132] (as shown in **Figure (16)**). The convolutional layer acts as the backbone of a CNN, performing feature extraction from input raw images. A set of trainable filters (or kernels) is used to convolve over the image pixels to obtain a set of feature maps that allow the model to detect features from simple edges to complex structures [133]. To increase the capacity of representation, the convolution operation is supplemented by an activation function, which most often is the ReLU activation function, making use of non-linear functions by assigning zero to all negative activations [133], [134]. Spatial dimensionality reduction of feature maps is done using a pooling layer. This helps in reducing computational cost, memory consumption, and over-fitting concerns while maintaining important features in the data. Finally, the Fully Connected (FC) layer performs high-level reasoning. Before this stage, feature maps are flattened into a one-dimensional vector, which is then used by the FC layer to compute final classification probabilities [132].

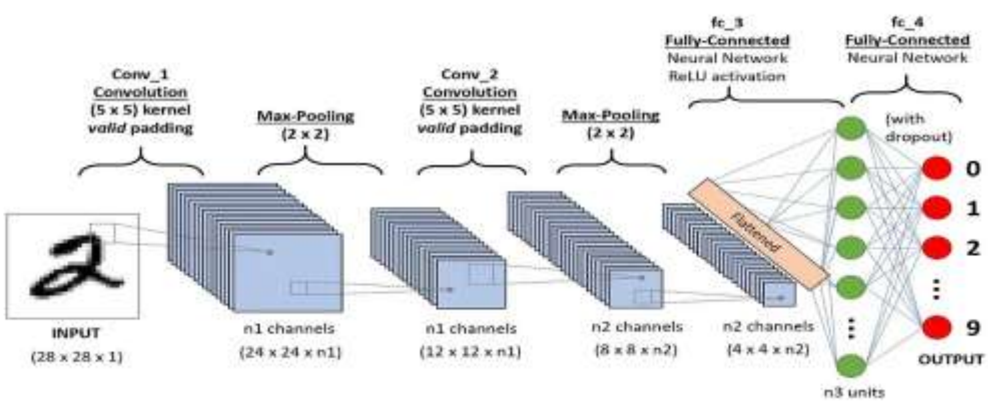

Figure (16): An example of CNN architecture [135].

5.2.3 Recurrent Neural Networks (RNNs)

Recurrent Neural Networks (RNNs) are another category of neural networks that have been specifically designed to deal with sequences of data and time-dependent data [136]. Unlike feed-forward neural networks, RNNs use connections between neurons where the information obtained at the previous stages is taken into consideration at the present stage of the network. This will make it easy to maintain a hidden state of the network, which consequently makes the network capable of recognizing dependencies within sequences of data. In summary, RNNs are quite useful for applications such as speech recognition, sentiment analysis, machine translation, time series prediction, and many more [137]. Despite being quite successful in dealing with sequences of data, RNNs face several issues. For example, one of the issues associated with RNN is that the network tends to suffer from the problem of vanishing gradients when trying to retain memory over a longer duration of time [136]. The **Figure (17)** below shows an illustration of a simple RNN network where the hidden memory of the network (h_t) is defined using **equation (2)** below [129].

$$h_t = g\,(W_{xt} + U_{ht} + b) \quad \text{Eq. (2)}$$

In the above equation, (g) refers to the activation function (usually Tanh), while (U) and (W) are trainable weights of the hidden state (h). (b) on the other hand, represents the bias of the neuron while (x) stands for the input vector. One of the advanced RNN architectures used to overcome the basic limitations includes the Long Short-Term Memory (LSTM).

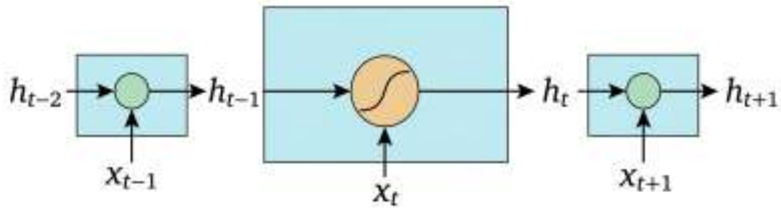

Figure (17): Simple RNN internal operation [129].

5.2.3.1 Long short-term memory (LSTM)

An important extension of RNN is known as Long Short-Term Memory (LSTM) network. The distinctive feature of the LSTM network is that the network has gates in its memory cell. More specifically, there are three gates: the forget gate decides which information should be forgotten and which retained, the input gate controls the amount of the new information that should be entered into the memory cell, and the output gate decides what should be done with the information that has been retained [128]. The memory cell can be seen as the key component of the LSTM network. Its purpose is to preserve information in a pathway, with as little modification of information over time as possible [138]. That means that the information that has been stored will be preserved through time. Therefore, preserving information for a longer period of time is the distinctive feature of LSTM since it cannot be done with traditional RNN. In addition to the memory cell, there is the hidden state that preserves the information, filtered depending on the needs of the particular step [139]. The general architecture of the LSTM network is illustrated in **Figure (18a)**.

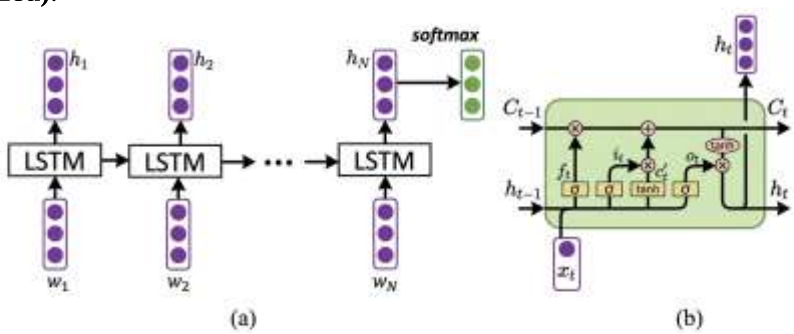

Figure (18): (**a**) LSTM's high-level architecture. (**b**) The LSTM unit's internal structure [140].

Figure (18b) illustrates the procedure of the information update in the internal architecture of the LSTM. The process of the information update in the LSTM unit is expressed in **equation (3)** presented below [129].

$$\begin{cases} h^{(t)} = g_o^{(t)} f_h\left(s^{(t)}\right) \\ s^{(t-1)} = g_f^{(t)} s^{(t-1)} + g_i^{(t)} f_s\left(wh^{(t-1)}\right) + uX^{(t)} + b \\ g_i^{(t)} = \text{sigmoid}\left(w_i h^{(t-1)} + u_i X^{(t)} + b_i\right) \\ g_f^{(t)} = \text{sigmoid}\left(w_f h^{(t-1)} + u_f X^{(t)} + b_f\right) \\ g_o^{(t)} = \text{sigmoid}\left(w_o h^{(t-1)} + u_o X^{(t)} + b_o\right) \end{cases} \quad \text{Eq. (3)}$$

Where (f_h) and (f_s) are the activation functions for the hidden state and cell state. They are usually implemented using the hyperbolic tangent function. The gating function, (g) is implemented using a feed-forward neural network with a sigmoid activation function resulting in the outputs in the range of [0, 1] and serving as weighting coefficients. While (i), (o), and (f) are the subscripts of the input, output, and forget gates [129]. LSTM networks have proven to perform well in a variety of areas including speech recognition, natural language processing, machine translation, time series prediction, and sentiment analysis [137]. Despite being more complex to compute compared to the basic RNN, the ability of the network to model long dependencies has turned out to be its advantage.

5.2.4 Generative Adversarial Networks (GANs)

GANs refer to one of the major advancements in the field of deep learning that enable the creation of new data samples that mirror the features of an original dataset. Two neural networks, including a generator and a discriminator, form the basic architecture of GANs. The generator network is developed to be fed with noise, whose output is to mimic original data samples, while the discriminator network is trained using a binary classification problem that involves original and synthetic data samples [141]. The interaction between the two networks forms the basis of the GAN framework as illustrated in **Figure (19)**.

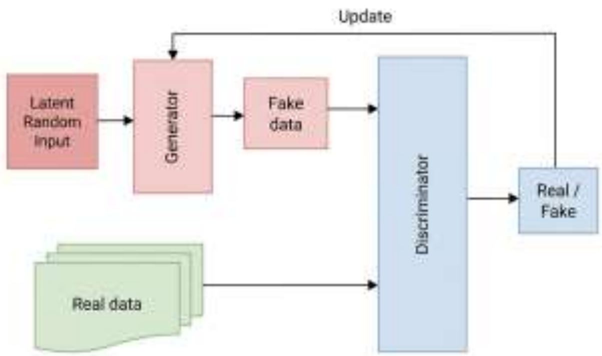

Figure (19): The framework of a GAN [129].

However, some notable drawbacks can be experienced when dealing with GANs. For example, training of GAN models is known to be complicated due to the numerous hyper-parameters required [142]. Thus, the implementation of GAN models may be complicated. On the other hand, another problem associated with GANs is that of control, meaning that controlling a GAN model's output is difficult [143]. One of the most serious drawbacks of GANs is mode collapse occurring when a generator converges to a limited area of a given data distribution [142].

5.2.5 Attention and Transformer-Based Models

The transformer introduced by Ashish Vaswani et al. in 2017 marked another significant improvement in sequence modeling by abandoning recurrence and convolutional approaches and fully relying on attention mechanisms. While the architecture has greatly outperformed existing models for machine translation, improving the efficiency of parallelization and training, the BERT model (Bidirectional Encoder Representations from Transformers), developed by Jacob Devlin et al. in 2018, has become one of the key developments in NLP [144]. Thus, there are many studies devoted to the analysis of transformers in various NLP tasks [145]. Transformers proved effective for computer vision as well, showing great results in tasks such as image classification, object detection, and segmentation while ensuring computational efficiency that grows linearly as the image size increases [146]. **Figure (20)** shows the structure of the transformer, including several essential components that are detailed in the following subsections.

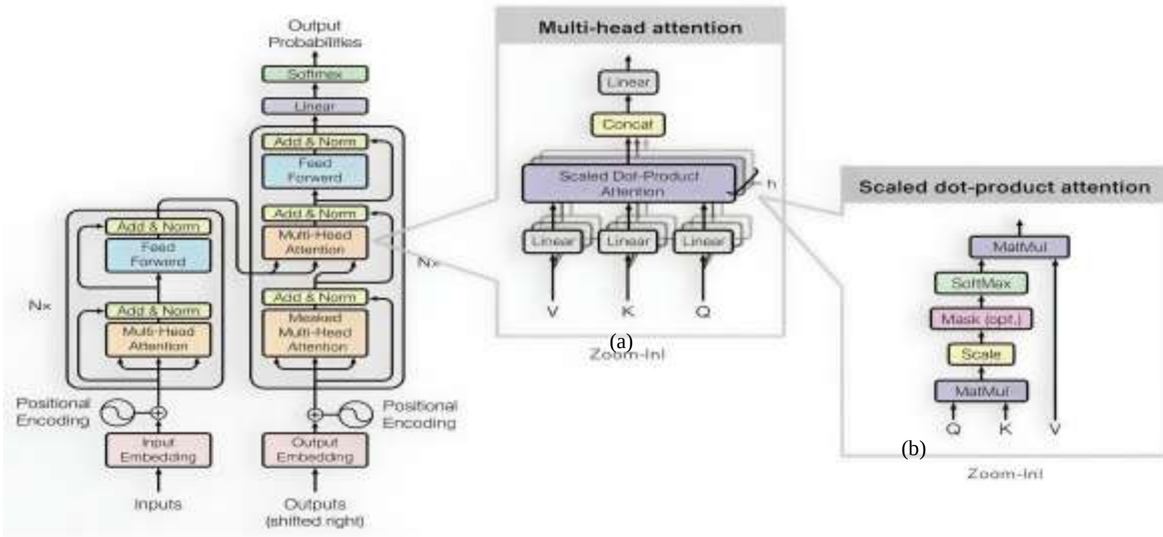

Figure (20): Transformer architecture, (**a**) MHA, (**b**) Scaled dot-product attention [147], [148].

5.2.5.1 Self-attention

Intra-attention (self-attention) provides an opportunity for a neural network to consider context by examining relations between all elements of input data. In contrast to sequential/local approaches, it gives a chance to consider the weights of all elements in order to enable direct interaction. This leads to a constant maximum distance between elements ($O(1)$)[148]. As opposed to CNN where the length of interaction is limited to kernel size, self-attention is able to detect long-range dependencies [149]. Its computational cost equals ($O(N^2)$) per pixel. In order to resolve the problem, Vision Transformers [150] treat images as sequences of patches/apply attention to downscaled feature maps [151].

5.2.5.2 Scaled Dot-Product Attention

Scaled Dot-Product Attention is one of the basic operations used in self-attention mechanisms (**Figure (20b)**). The attention score of queries (Q), keys (K), and values (V) is determined through dot products and scaled by a factor of $\frac{1}{\sqrt{d_k}}$. Then, a similarity is obtained using a softmax function. The attention matrix is calculated as follows **equation (4)** [148].

$$\text{Attention}\,(Q, K, V) = \text{softmax}\left(\frac{QK^T}{\sqrt{d_k}}\right)V \qquad \text{Eq. (4)}$$

Where, (d_k) denotes the dimension of keys/queries, and (softmax) converts similarity scores into probability distributions [148].

5.2.5.3 Multi-head Attention (MHA)

In order to be capable of capturing various types of dependencies, multi-head attention decomposes Q, K, and V in multiple projection spaces, as shown in (**Figure (20a)**). Then, self-attention is independently applied within those spaces. The MHA is calculated using equation (5), where each attention head computes its value according to **equation (6)** [148].

$$\text{MultiHead}(Q, K, V) = \text{Concat}(\text{head}_1, \dots \dots, \text{head}_h)W^O \qquad \text{Eq. (5)}$$

$$\text{head}_i = \text{Attention}(QW_i^Q, KW_i^K, VW_i^V) \qquad \text{Eq. (6)}$$

Here, QW_i^Q, KW_i^K, VW_i^V are the parameter matrices for the projections of the $i-$th head, and W^O is the final output projection matrix. This architecture allows the model to jointly attend to information from different representation subspaces, which is crucial for identifying complex patterns in medical images [148].

5.2.5.4 Position-wise Feed-Forward Network (FFN)

Position-wise Feed Forward Networks (FFNN) are an integral part of the Transformer architecture, as described in [148]. It has been included in each of the Transformer layers, which work in conjunction with the multi-head self-attention mechanism, typically after the attention component, with residual connections and normalization following this [152]. This component plays an important role in the computation process by providing non-linearity to the model, which helps in learning non-linear functions. In the self-attention layers, the information is disseminated across various positions in the input sequence, which are then processed individually by the feed-forward network, providing an additional processing component to the model, which enhances the flexibility of the model to work with various kinds of data [153]. The FFN can be defined by the **equation (7)** below [154].

$$\mathrm{FFN}(x_i) = \sigma(x_i W_1 + b_1) W_2 + b_2 \qquad \text{Eq. (7)}$$

where (W_1), (W_2) are the weight matrices for the two linear transformation operations, (b_1) and (b_2) are the corresponding bias vectors, with (σ) representing the non-linear activation function.

5.2.6 Transfer Learning (TL)

With transfer learning, knowledge accumulated from data-abundant tasks can be exploited to improve performance on data-limited problems, as shown in **Figure (21)**. Rather than designing models from scratch, pre-trained models are initialized [155], assuming the presence of common low-level representations between tasks, such as edges and textures in computer vision. There are different techniques involved in transferring knowledge. The instance-based technique adjusts the weights associated with samples within the source domain to prioritize those most useful for the target task. The mapping-based technique projects the features from the source and target domains into one common space [156]. In the network-based method, which is commonly applied in deep learning, certain components of a pre-trained model and its parameters are reused for new tasks [157]. This technique shows remarkable performance with convolutional and recurrent neural networks because the hierarchical representations can be tuned to the application at hand.

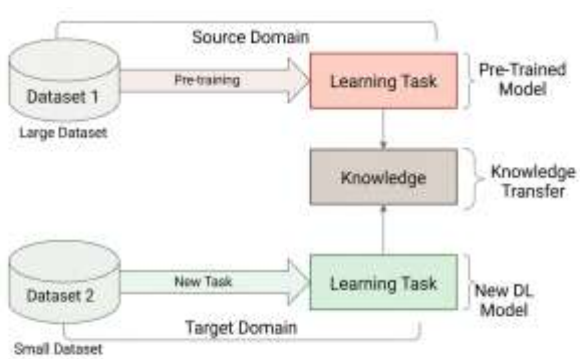

Figure (21): A generic framework for the transfer learning process, in which a new DL model incorporates knowledge from a pre-trained model [128].

In practice, transfer learning is normally done using two steps: pre-training and fine-tuning. In the first step, the models are initially trained using large datasets. The models are then used for the target task using a smaller dataset. This has been proven to work well in scenarios where there is limited data. The existing models provide robust initializations for the models, thus improving performance [155].

5.2.7 Explainable AI in Ophthalmic Imaging

Explainable AI (XAI) has become increasingly important in ophthalmic imaging, as high predictive accuracy alone is insufficient for safe clinical deployment. Although deep learning systems can achieve expert-level performance in retinal image analysis, their inherent opacity limits clinical acceptance, since clinicians cannot readily determine how predictions are generated [158]. In ophthalmology, the diagnosis is based on the actual pathological changes observed, such as hemorrhages, exudates, changes in the optic disc, defects of the retinal nerve fiber layer, abnormalities of the retinal layers, etc., and in this field, interpretability is especially important for clinical validation. XAI is therefore seen as a tool for increasing transparency, making for safer decisions, building clinician trust, and meeting regulatory requirements for medical AI systems [159].

5.2.7.1 Explain-ability Methods: Grad-CAM, SHAP, and LIME

Grad-CAM is one of the most popular post-hoc explain-ability techniques for retinal imaging and OCT analysis, generating heat-maps that highlight the areas of an image that are most likely to contribute to a particular prediction. The method uses the last convolutional layer to calculate gradients and defines class-specific localization maps, which help answer a clinically relevant question of what area of the retina influenced the diagnosis [160].
During glaucoma detection, for example, models trained with Grad-CAM visualization generally target optic disc and optic cup regions, while those for diabetic retinopathy tend to target areas of the retina with lesions, showing clinically significant attention patterns. However, saliency maps are still qualitative and can be misleading even when models are based on non-generalizable and biased features, highlighting the need for external validation and expert review [161].
In contrast to Grad-CAM which emphasizes spatial attention, SHAP and LIME provide complementary explanations with feature attribution. LIME approximates local model behavior for individual cases, enabling the retinal regions that influence a specific prediction to be identified for clinician review [162]. By contrast, SHAP quantifies the contribution of image features or retinal regions to the final prediction and is often used to evaluate the interpretability and robustness of the model used in diabetic retinopathy screening systems [163]. Nevertheless, these two approaches have crucial drawbacks in ophthalmic imaging. Typical approaches for LIME use super pixel segmentation, but this may not align with clinically relevant retinal features like lesions or vessel boundaries, which limits the medical interpretability of the explanations produced by LIME [164]. Moreover, LIME offers explanations by using a local surrogate model which can be too simple to encompass complex feature interactions commonly found in deep neural networks [165]. SHAP explanations, conversely, can be hard to interpret at the pixel level and can point to any part of the image within a broad region [166].
They are designed to provide responses to different clinical questions. Grad-CAM is generally preferred when strong spatial localization on the image itself is required, SHAP is employed when feature contributions must be quantified, and LIME is applied when a local explanation for an individual prediction is needed [167]. Their collective value in ophthalmology is therefore pragmatic: they provide clinicians with an additional layer of evidence for fundus and OCT predictions, help identify implausible or misleading cues, and can support clinical trust, safety review, and regulatory acceptance when used in conjunction with validation studies and expert inspection [162] [167].

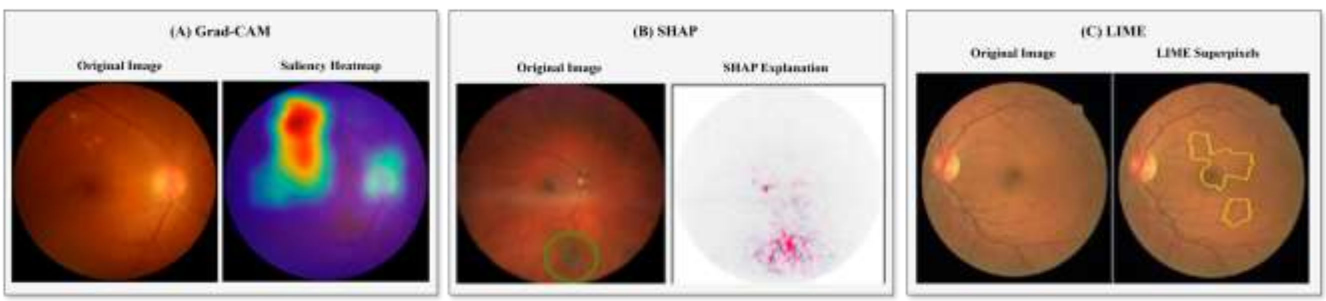

Figure (22): Comparison of XAI methods in ophthalmic imaging.

6 AI-Driven Methods for Ocular Diseases Analysis: A Review of Recent Literatures

This section offers a comprehensive overview of the latest advances in the field of Artificial Intelligence for ophthalmic image analysis. It focuses on the rapid evolution and usage of AI techniques in the context of the four major ocular diseases. **Table (2)** provides a summary of various comprehensive and state-of-the-art studies that have been discussed.

6.1 AI-Based Approaches for Diabetic Retinopathy Analysis

Butt et al. (2022) [27] used pre-trained models (Google Net and ResNet-18) to obtain a 2000-feature vector which was further classified using machine learning algorithms such as SVM, Random Forest, and Naïve Bayes. This method demonstrated 97.8% accuracy in binary classification and 89.29% in multi-class classification on the APTOS dataset. Nevertheless, relying on only one dataset limited the generalizability of the findings; moreover, the significant disparity between binary and multi-class accuracy indicates problems with the precise grading of disease severity.

Jabbar et al. (2022) [9] proposed a model based on VGG Net architecture tested on the Eye PACS dataset, where comprehensive preprocessing combined with grade-specific augmentation and manual feature engineering greatly improved classification results. The authors reported 96.6% accuracy with an area under the curve of 0.97. Despite the encouraging results, the reliance on handcrafted features alongside neural networks is methodologically inconsistent since the automatic learning of representation layers should be the main advantage of deep learning techniques, and the manual augmentation of features may artificially inflate performance without contributing to genuine representational learning.

Mutawa et al. (2023) [168] evaluated four CNN architectures, concluding that a combination of three datasets (APTOS, Eye PACS, and ODIR) positively affected the overall performance of models, with DenseNet121 demonstrating the best accuracy at 98.97%. Nonetheless, merging datasets implies discrepancies among image resolutions, acquisition tools, colors, and even grading criteria. The authors did not provide any information concerning whether the researchers applied methods of domain adaptation or normalization to deal with those inconsistencies. Consequently, it remains unclear whether the result reflects genuine model robustness or simply the dominance of one dataset's characteristics over others during training.

Krishna and Rao (2023) [169] developed a model incorporating the Flower Pollination Optimization algorithm combined with the transfer learning-based DeepID3 network to detect multiple eye diseases, including diabetic retinopathy, with 99.23% accuracy on the Messidor and DRISHTI-GS datasets. The proposed approach is innovative, but the extremely high accuracy achieved on two relatively small benchmark datasets causes suspicion concerning the possibility of over-fitting.

Jabbar et al. (2024) [170] designed a deep transfer learning model for real-time detection of diabetic retinopathy for telemedicine purposes using a pre-trained VGG Net architecture on the Eye PACS, Messidor, and IDRiD datasets, resulting in 97.6% accuracy. The telemedicine focus adds practical clinical value; nevertheless, the study does not discuss model latency, hardware requirements, or performance under low-quality image conditions, which are critical factors for real-world deployment in resource-limited settings.

Bhoopalan et al. (2225) [171] designed an optimal task-based vision transformer (TOViT) to diagnose and assess the severity of diabetic retinopathy and reported 99% accuracy across the EyePACS, Messidor-2, and APTOS datasets. Additionally, the authors presented the results of model implementation on a Raspberry Pi platform as a part of edge-computing applications. Although the involvement of the proposed technology in edge computing is a practical contribution, the high accuracy achieved on different datasets casts doubts on over-fitting or potential data leakage.

6.2 AI-Based Approaches for Cataract Analysis

Varma et al. (2023) [172] suggested a deep convolutional neural network (DCNN) model for identifying four grades of cataracts according to disease severity, attaining 92.7% classification accuracy through green channel contrast enhancement and data augmentation for softmax classification. Although the proposed four-tier classification is practical in a clinical setting, the absence of AUC and sensitivity statistics makes it difficult to assess the model's reliability for detecting early-stage cataracts.

Saqib et al. (2024) [173] employed lightweight transfer learning using MobileNetV1 and MobileNetV2 to detect both cataracts and glaucoma, reaching 90% and 87% accuracy, respectively. While the lightweight nature of the proposed architecture renders it suitable for limited resource settings, the accuracy rate remains relatively mediocre when compared to recent state-of-the-art techniques, in addition the study does not address class imbalance.

Mahmood et al. (2024) [174] applied transfer learning based on the ResNet50 architecture, attaining an excellent 96.63% accuracy on the publicly available ODIR-5K data set, in the binary classification of cataracts. Although the presented performance figure is very promising, focusing only on the binary classification of the disease and not taking into consideration variations in their severity limits the effectiveness. Furthermore, reliance on a single publicly available data set without any cross-validation and only ResNet50 as a baseline without comparison to more recent models may affect the generalizability and robustness of the findings.

Feng et al. (2024) [175] proposed a hybrid deep learning technique comprising ResNet50, Squeeze-and-Excitation blocks, and a prototypical network classifier. The combination resulted in 98.75% and 99.84% accuracy and AUC, respectively. However, the assessment of this model was confined to internal validation, which raises questions about whether such performance is reproducible across external testing, different imaging devices.

Jabber et al. (2025) [176] performed an extensive comparison among DenseNet-201, ResNet-101, and Google Net based on a Kaggle dataset, employing image preprocessing via histogram equalization, contrast enhancement, and segmentation techniques. The results attained 98.33% accuracy with DenseNet-201 and 90% with Google Net for three-tier cataract grading. Although the application of sophisticated image preprocessing techniques is considered a significant strength of the study, the usage of a single Kaggle dataset and the large performance gap between architectures suggest sensitivity to model selection that warrants further investigation.

6.3 AI-Based Approaches for Glaucoma Analysis

Singh et al. (2022) [177] used HOG features to detect and classify glaucomatous eyes based on KNN and ANN classifiers, obtaining 96.90% and 96.0% accuracy, respectively, demonstrating that handcrafted features remain competitive with deep learning on smaller datasets. However, the reliance on handcrafted features introduces a degree of manual engineering that may not scale effectively to larger or more heterogeneous datasets.

Kashyap et al. (2022) [31] introduced a methodology involving transfer learning with a U-Net architecture, where DenseNet-201 was applied for feature extraction and classification tasks, obtaining 98.82% and 96.90% accuracy on training and test sets, respectively. While the integration of segmentation and classification is well-grounded, there is a significant difference between training and testing accuracy rates. The high performance on the training dataset implies the possibility of over-fitting, limiting the model's generalizability.

Velpula and Sharma (2023) [178] presented a multi-stage classification system that employs five pre-trained CNNs to diagnose glaucoma. The authors obtained excellent accuracy on four different datasets, i.e., 99.57% on the ACRIMA, 94.95% on RIM-ONE, 90.55% on Drishti, and 85.43% on the HVD dataset. Testing the proposed model on multiple datasets is a great advantage since it demonstrates its true generalizability. However, the wide gap in performance on different datasets highlights the sensitivity of the model to differences in image quality, acquisition, and demographic characteristics.

Patil and Sharma (2024) [125] proposed a framework based on Deep Learning technology to minimize diagnostic time and automate the glaucoma diagnostics process. The experiments were performed using a Kaggle dataset with 1,022 fundus images. Several pretrained transfer learning models were tested, such as DenseNet169, MobileNet, InceptionV3, Xception, ResNet152V2, and VGG19. In particular, DenseNet169 showed the highest accuracy equal to 0.9935 with precision and recall rates of 0.9936 and 0.9935, respectively. Nevertheless, almost perfect results obtained on a small single-source dataset raise concerns about overfitting, which reduces the practical utility of the model.

Chincholi and Koestler (2024) [179] have tried to apply ViT and DETR approaches for detecting and classifying glaucoma on fundus images with the help of vision transformers for extracting features and detection transformers for recognizing the optic disc and cup and calculating cup-to-disc ratios. As a result, DETR outperforms ViT, showing an accuracy rate up to 90.48% and an AUC of 0.95. The use of transformers is innovative, and it enables better generalization. Nevertheless, the accuracy has not yet reached full maturity for ophthalmic classification tasks without substantially larger training datasets.

Edan and Feizi-Derakhshi (2025) [180] overcome the problem of using a separate dataset for testing and training because the authors experiment with an ensemble-based network combining ResNet-101, ResNet-18, and Darknet-

19 models using an accuracy-weighted voting strategy. The resulting ensemble-based CNN obtained accuracies of 90% on the Drishti-GS dataset, 95.6% on the RIM-ONE v2 dataset, and 99.3% on the ACRIMA dataset. The ensemble strategy effectively mitigates individual model limitations and the multi-dataset evaluation strengthens generalizability claims; however, the wide variation in accuracy across datasets underscores the persistent challenge of domain shift, and the computational overhead of ensemble inference may limit deployment in real-time clinical systems.

6.4 AI-Based Approaches for Age-Related Macular Degeneration Analysis

Wang et al. (2022) [181] suggested a self-attention-based neural network with an encoder-decoder architecture resembling U-Net for segmenting and analyzing AMD and Stargardt disease from OCT volumes. The proposed approach has shown excellent segmentation capacity and interpretability and provided 98% of accuracy for AMD segmentation. At the same time, the model was validated on a relatively small number of images, suggesting its possible over-fitting limits the model's generalizability across different imaging conditions.

Mishra et al. (2022) [182] presented MacularNet, which is an end-to-end CNN including pretrained models and deformation-aware attention modules for classifying macular diseases without preprocessing stages. It provided 99.94% and 99.79% accuracies for five-fold cross-validation on the Duke and NEH datasets, correspondingly. On the one hand, the use of cross-validation and deformable attention mechanisms may be considered as advantages. On the other hand, near-perfect results on two relatively small benchmark sets may suggest a risk of the model being overfit.

Han et al. (2023) [183] proposed a CNN-based algorithm for the differential diagnosis of neovascular AMD (nAMD), central serous chorioretinopathy (CSC), and healthy subjects using spectral-domain OCT. Applying transfer learning and mixup augmentation on VGG-16, VGG-19, and ResNet architectures, they obtained 99.7% and 91.1% accuracies for 3-class classification and 6-class classification, respectively, on a dataset consisting of 6,063 images. The multi-class approach and big dataset may be viewed as relative strengths of this research. At the same time, a noticeable difference between 3-class and 6-class classification suggests that the model may struggle to reliably distinguish between clinically similar subtypes, which poses a problem for clinical practice.

Ali et al. (2024) [184] developed a CNN-LSTM hybrid model aimed at capturing temporal relations between OCT images with LSTM layers. The proposed approach resulted in a classification accuracy of 96.5% based on 2,000 images from six public sources. Although multi-source data aggregation and temporal analysis of images a contribution point, the use of a relatively small dataset and a rather complex hybrid architecture can increase the chances of overfitting of the model.

Azizi et al. (2024) [185] introduced a Medical Vision Transformer (MedViT)-based framework for AMD detection from OCT images. A “micro” MedViT variant was proposed, as well as stitchable neural networks to be built with linear insertion of pretrained layers and iterative optimization with SGD and AdamW. The approach achieved 97.9% accuracy in an in-distribution evaluation, incorporating extra out-of-distribution robustness testing. Although the inclusion of out-of-distribution evaluation is a distinguishing strength of this work, the stitchable neural network design, while technically innovative, introduces considerable architectural complexity, as such designs typically require careful hyperparameter tuning and are sensitive to the specific combination of pretrained components used during integration.

Yusufoğlu et al. (2024) [186] proposed a hybrid CNN structure that incorporates ConvMixer, Depthwise Squeeze-and-Excitation blocks and a modified version of the Inception modules. The model achieved 97.98% accuracy, 97.95% precision, 97.77% recall, and 97.86% F1 score on a private clinical dataset and 100% across all the metrics on the public Noor dataset, highlighting the effectiveness of hybrid and attention-enhanced CNN designs. While the hybrid design shows strong performance, the perfect scores on the Noor dataset raise concerns about potential overfitting of the model. The results could be representative of the dataset and not representative of the overall model generalizability, which means that independent validation is needed.

Table (2): Comparative Analysis of AI Models in Diagnosing Ocular Diseases.

Ref.	Disease	Dataset	Methodology	Key Results	Limitation
Butt et al. (2022) [27]	DR	APTOS (Kaggle)	Hybrid (GoogleNet, ResNet18 + ML classifiers (RF, SVM, RBF, NB))	Binary Accuracy: 97.80%, Multi-class Accuracy: 89.29%	Small dataset. Class imbalance. Limited generalization. No clinical validation.
Jabbar et al. (2022) [9]	DR	EyePACS (Kaggle)	VGGNet-16	Accuracy: 96.6%, AUC: 0.971	Strong dependence on augmentation and potential overfitting.
Krishna & Rao (2023) [169]	DR	DRISHTI-GS, Messidor-2, Retina datasets	DeepID3 + FPOA optimization	Accuracy: 99.23%, F1-score: 98.3%	Class imbalance and bias risk.
Mutawa et al. (2023) [168]	DR	APTOS, EyePACS, ODIR	Ensemble Transfer Learning models (VGG-16, InceptionV3, DenseNet-121, MobileNetV2)	Accuracy: 98.97%, AUC: 0.9947	Use of only a small dataset size, evaluation on balanced data
Jabbar et al. (2024) [170]	DR	IDRiD, Messidor, EyePACS	TAlexNet, ResNet, VGGNet	Accuracy: 97.6%, F1-score: 0.987, AUC: 0.981	The accuracy of the model depends on image quality.
Bhoopalan et al. (2025) [171]	DR	EyePACS, Messidor-2, APTOS	Task-Optimized Vision Transformer (TOViT)	Accuracy: 99%, F1-scores: 93%	Very large model with (267 million) parameters.
Varma et al. (2023) [172]	Cataract	HRF, STARE, MESSIDOR, DRIVE, DRIONS-DB, IDRiD, Internet images	Custom CNN	Overall Accuracy: 92.7% (Per-class: Mild 87.5%, Mod 93.8%, Severe 96.3%, Normal 93.1%)	Weak performance in mild cataract. Early cases missed.
Saqib et al. (2024) [173]	Cataract	Not specified (N/A)	MobileNetV1 & MobileNetV2 (Transfer Learning with custom head)	Cataract vs Normal Accuracy: 90% (V1), 87% (V2)	Weak generalization.
Feng et al. (2024) [175]	Cataract	ODIR-5K, Retina (GitHub), Kaggle Cataract dataset	Hybrid Transfer Learning (ResNet50 + SE module + Prototypical Network)	Accuracy: 98.75%, AUC: 0.9984, F1: 0.9855	Low demographic diversity. Low interpretability.
Mahmood et al. (2024) [174]	Cataract	ODIR-5K (Normal vs. Cataract only)	ResNet50	Test Accuracy: 96.63% Prec: 0.9715, AUC: 0.9661	Only used one dataset, and Poor explainability.
Jabber et al. (2025) [176]	Cataract	Kaggle	DenseNet-201, ResNet-101, GoogLeNet	Detection Accuracy: 98.33% Classification Accuracy: 90% (3-tier)	Only used one dataset.
Singh et al. (2022) [177]	Glaucoma	Drishti-GS, Feiz Hospital	Histogram of Oriented Gradients (HOG) + classifiers (KNN, SVM, LDA, Naïve Bayes, and ANN)	Accuracy on ANN: 96.90%, Naïve Bayes: 96%, SVM: 90%, KNN: 86%	Very small dataset.
Kashyap et al. (2022) [31]	Glaucoma	Kaggle, SiMES	U-Net + DenseNet-201 + DCNN	Test Accuracy: 96.90%, Test Recall: 97.03%, Spec: 98.15%	Small dataset. Limited generalization.
Velpula & Sharma (2023) [178]	Glaucoma	ACRIMA, RIM-ONE, HVD, Drishti	Pre-trained CNN ensemble + Maximum voting-based fusion	Accuracy on ACRIMA: 99.57% (AUC 1.00), RIM-ONE: 94.95%, Drishti 90.55%, HVD: 85.43%	High computation cost
Patil & Sharma (2024) [125]	Glaucoma	Kaggle	DenseNet169, MobileNet, InceptionV3, Xception, ResNet152V2, VGG19, ResNet50	Accuracy: 99.35%, Precision: 99.36%, Recall: 99.35%	Small dataset. Limited generalization.
Chincholi and Koestler (2024) [179]	Glaucoma	SMDG	Vision Transformer (ViT) and Detection Transformer (DETR)	DETR accuracy: 90.48%, ViT accuracy: 87.87%.	Data scarcity. Class imbalance. Dataset bias.
Edan & Feizi-Derakhshi (2025) [180]	Glaucoma	ACRIMA, RIM-ONE v2, Drishti-GS	Pre-trained CNNs (ResNet-101, ResNet-18, and DarkNet-19) + voting classifier fusion technique.	Fusion Accuracy: ACRIMA 99.3%, RIM-ONE v2 95.6%, Drishti-GS 90.0%	Dataset-dependent performance. Inconsistent results between the datasets.

Wang et al. (2022) [181]	AMD	FAF images (Spectralis HRA+OCT)	U-Net + Self-attention	AMD Segmentation: Accuracy 98%, Dice 0.8512. Month Progression: Accuracy 95%, Dice 0.78	Limited ability to diagnose and predict the slow progression of Stargardt AMD.
Mishra et al. (2022) [182]	AMD	Duke, NEH, UCSD, OCTID	MacularNet (CNN / VGG)	Acc (5-fold): Duke 99.94%, NEH 99.79%, Accuracy: UCSD 92.60%, OCTID 93.12%	lack of evaluation on early-stage macular diseases
Han et al. (2023) [183]	AMD	Hangil Eye Hospital SD-OCT	VGG-16/19, ResNet + custom heads	Test Accuracy: 99.7% (3-class), 91.1% (6-class), Ext. Test Accuracy: 100% (3-class), 92.3% (6-class)	Limited diversity. Single-device data affects the models' generalizability.
Yusufoğlu et al. (2024) [186]	AMD	Private OCT, Public OCT (Noor Eye Hospital)	PM (Inception + SE Blocks + ConvMixer)	(Private Dataset) Accuracy: 97.98%, F1: 97.86% (Noor Dataset) Accuracy: 100%, F1: 100%	Limited dataset diversity may affect generalization; real-world challenges were not addressed.
Azizi et al. (2024) [185]	AMD	NEH	Medical Vision Transformer (MedViT)	Accuracy: 94.9%	Class mismatch. Needs fine-tuning.
Ali et al. (2024) [184]	AMD	ODIR, DR-200, RFMiD, ARIA, Kaggle	AMDNet23 (Hybrid CNN-LSTM)	Overall Accuracy: 96.50%, Sens: 96.50%, Spec: 99.32%, F1: 96.49%	Class imbalance. Overfitting risk.

7 Performance Evaluation Metrics

The assessment of the performance of DL and ML models is achieved through the application of various metrics. The application of one metric might be insufficient since there is a variety of metrics that evaluate different facets of models' performance. Thus, the employment of different evaluation criteria allows comprehending model's performance in a more complete manner and estimating its predictive potential. The following performance metrics are typically used: accuracy (**Equation (8)**), sensitivity (recall) (**Equation (9)**), specificity (**Equation (10)**), precision (**Equation (11)**), F1-score (**Equation (12)**), and the area under the curve (AUC) (**Equation (13)**). All these metrics are calculated based on basic classification results and, therefore, reflect the model performance differently. The accuracy, sensitivity, specificity, precision, and F1-score are calculated as follows: These evaluation measures depend on four numbers: True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). These four numbers describe classification results produced by the model. The true positive rate (TPR) and false positive rate (FPR) are calculated based on the previously described values and are needed for estimating AUC. This particular metric shows how well the model can separate between two classes depending on the decision threshold. All these metrics are easy to calculate with the help of the scikit-learn library.

$$\text{Accuracy (ACC)} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})} \quad \text{Eq. (8)}$$

$$\text{Recall (Sensitivity)} = \frac{\text{TP}}{(\text{TP}+\text{FN})} \quad \text{Eq. (9)}$$

$$\text{Specificity (SPC)} = \frac{\text{TN}}{(\text{TN}+\text{FP})} \quad \text{Eq. (10)}$$

$$\text{Precision (PPV)} = \frac{\text{TP}}{(\text{TP}+\text{FP})} \quad \text{Eq. (11)}$$

$$\text{F1} - \text{score} = \frac{2*\text{TP}}{(2*\text{TP}+\text{FP}+\text{FN})} \quad \text{Eq. (12)}$$

$$\text{AUC} = \int_{x=0}^{1} \text{TPR}\left(\text{FPR}^{-1}(x)\right)dx, \ \ \text{TPR} = \frac{\text{TP}}{\text{TP}+\text{FN}}, \quad \text{FPR} = \frac{\text{FP}}{\text{FP}+\text{TN}} \quad \text{Eq. (13)}$$

8 Current Challenges and Limitations

The use of artificial intelligence for analyzing ophthalmological data is a complicated and intricate issue. It entails a variety of aspects that need to be properly considered in order to achieve positive results. Despite the fact that considerable advances have been made in both artificial intelligence and medical imaging, there are still some problems that hinder the efficiency of existing methods. In particular, the problems of model accuracy and generalization still play an important role. The following is a list of the major challenges and limitations of AI in ophthalmology.

- **Limited datasets:** One of the major drawbacks is the scarcity of extensive and quality ophthalmological datasets for training AI algorithms. Poor-quality images or datasets with bias can result in low model performance and limited generalization to different patient groups. Most of the existing databases have relatively small sample sizes, inconsistencies in image annotation, and alignment problems across various imaging modalities.
- **No standardization:** The absence of common standards for some imaging modalities (e.g., fundus photography and OCT) and the representation of clinical metadata. Consequently, heterogeneous data integration is hampered by the lack of appropriate AI-ready frameworks for imaging and metadata processing.
- **Data imbalance:** Along with data heterogeneity, class imbalance is yet another problem in the development of ophthalmic AI systems. The lack of balance between the underrepresented diseases and predominant categories results in the model being biased towards the majority class and leading to a reduced sensitivity of the model to the detection of the less frequent cases.
- **Single-disease focus:** Most existing AI models are designed to identify only a single disease. However, in real life, people suffer from more than one disease and there is more complexity in the relationships of these conditions.
- **Privacy issues:** Serious regulatory and privacy-related issues are associated with the sensitive nature of both clinical data and imaging results.
- **Low interpretability:** The poor interpretability of neural network architectures commonly used in ophthalmic AI makes models difficult to understand and trust.

9 Future Research Directions

Recent advances in ophthalmic artificial intelligence have demonstrated strong potential for retinal disease detection, segmentation, and clinical decision support. However, several previously discussed limitations continue to restrict real-world deployment. Consequently, future ophthalmic AI research is increasingly shifting toward several emerging directions, including federated learning, explainable AI, foundation models, multimodal learning, and real-time clinical deployment.

9.1 Federated Learning and Privacy-Preserving AI

One of the major challenges in ophthalmic imaging is the limited availability of large and diverse multicenter datasets. Privacy regulations and institutional restrictions often prevent direct sharing of retinal images across hospitals. Consequently, federated learning (FL) has become a promising approach for collaborative development of ophthalmic AI, which allows multiple institutions to train shared models without transferring patient data between institutions [187], [188]. This approach can enhance the protection of privacy and the generalizability of the model in other patient populations and imaging modalities. FL can deliver performance similar to centralized training in retinal imaging applications, including diabetic retinopathy classification and OCT/OCTA analysis, without compromising data privacy [189]. However, domain heterogeneity is still a big challenge due to the differences in scanners, image quality, and acquisition protocols, which can reduce the robustness of the model across institutions [158]. Thus, future research should work on integrating FL with domain adaptation, self-supervised pertaining, and utilizing transformer-based encoders to boost the cross-site generalization and fairness.

9.2 Self-Supervised Learning

Lack of expert-labeled ophthalmic datasets is still a key bottleneck in the field of medical image analysis. As retinal image annotation is often a specialist task, self-supervised and semi-supervised learning approaches which can learn meaningful visual representations from unlabeled data [190], which could have a great impact in reducing the associated annotation burden. Taken together, the main future research question is no longer whether ophthalmic AI can work with labels alone, but how to build reusable pertained representations that transfer across tasks, modalities, populations, and care settings with minimal annotation burden.

9.3 Vision Transformers (ViTs)

Recently, vision Transformer (ViT) has become a popular technique in retinal image analysis, as it has the ability to simultaneously extract the local lesions and global spatial relationships via self-attention mechanisms. This offers a natural fit for ophthalmology applications where pathological findings (e.g., vascular pathology, changes in the optic disc and retinal lesions) may often also be confined to localized image regions. Also, the transformer may be more robust than standard CNN systems in the face of domain shift and image artifacts [191]. However, ViTs take larger quantities of data and have higher computational requirements, making a case for larger metacentric datasets, more computationally efficient pretraining approaches, as well as lightweight hybrid networks suitable for clinical deployment.

9.4 Foundation Segmentation Models and Segment Anything Model (SAM)

In ophthalmic imaging, segmentation is a crucial aspect as clinicians often need to analyze retinal layers, vessels, lesions, and fluid region structure levels. Recently, the Segment Anything Model (SAM) introduced a foundation-style promotable segmentation framework trained on extremely large-scale image-mask datasets [192]. This has led to a lot of research on adapting foundation segmentation models to retinal imaging applications. Looking ahead, the real question is not if SAM is interesting but how it can be used in ophthalmology to its benefit. The major focus areas should include tuning of the ophthalmic-specific algorithms, evaluation of existing OCT and fundus data from multiple devices, improved clinician interaction prompts and user interfaces, and integration with clinical workflows where segmentation can support measurement and monitoring rather than only visual demos [193].

9.5 Edge AI and Real-Time Clinical Deployment

Another area of development in the future is AI systems for the ophthalmic field that are lightweight and real-time, appropriate for teleophthalmology, mobile screening, and low-resource clinical settings [194]. Numerous models have been developed that reach high benchmark accuracy but are not easily deployable due to their computational cost. Future work should therefore emphasize model compression, efficient inference, and hardware-aware optimization for edge devices such as portable retinal cameras and embedded diagnostic systems. In addition, integration with hospital workflows, picture archiving systems (PACS), and clinical decision-support platforms will be critical for large-scale adoption [195]. Importantly, future edge AI systems need to be robust and calibrated despite having less computational complexity.

9.6 Multimodal Learning

Future ophthalmic AI systems will also shift from single modality analysis to combining multiple imaging modalities and clinical data sources. A recent study indicates that integration of fundus photography, OCT, OCTA, angiography, clinical metadata, and genetics could enhance the diagnostic value and deepen the understanding of disease mechanisms [196]. But there are still significant challenges in a multimodal approach to data collection, including missing modalities, fusion strategy, privacy protection, and standardization of datasets. Consequently, future research should prioritize large paired multimodal datasets, robust multimodal fusion architectures, and clinically validated multimodal foundation models capable of supporting real-world ophthalmic decision-making.

10 Conclusion

This review demonstrates that artificial intelligence has fundamentally transformed the landscape of ophthalmic imaging analysis, enabling automated diagnostic systems that approach or match expert-level performance across major ocular diseases: Diabetic Retinopathy, Cataract, Glaucoma, and Age-Related Macular Degeneration. The convergence of deep learning architectures, transfer learning strategies, attention mechanisms, and transformer-based models has substantially advanced the capacity of computer-aided diagnosis systems to detect subtle pathological features, stratify disease severity, and support clinical decision-making in settings where specialist access remains limited.

Despite these technical advances, the transition from benchmark performance to reliable clinical deployment remains constrained by a set of persistent and interconnected challenges. Data-related limitations, including class imbalance, inconsistent annotation protocols, imaging device heterogeneity, and domain shift across institutions and populations, continue to undermine model generalizability and external validity. Furthermore, the absence of clinical validation across the reviewed literature suggests that reported accuracy metrics may not accurately reflect real-world diagnostic reliability. In addition, the issues of patient data privacy, algorithmic transparency, and the inherent opacity of deep learning models remain unresolved and represent significant barriers to regulatory approval and clinician trust.

It is therefore essential to emphasize that AI systems in ophthalmology are intended to function as decision-support tools that augment clinical expertise rather than replace it. The most meaningful advances in the field will likely emerge not from incremental accuracy improvements on existing benchmarks, but from rigorous multi-center validation studies, standardized evaluation frameworks, demographically diverse datasets, and collaborative development processes that incorporate the clinical knowledge of ophthalmologists alongside the technical contributions of AI researchers. Addressing these gaps systematically will be critical to realizing the full potential of AI in reducing the global burden of preventable visual impairment.

References

[1] R. Alsohemi and S. Dardouri, “Fundus Image-Based Eye Disease Detection Using EfficientNetB3 Architecture,” J. Imaging, vol. 11, no. 8, p. 279, Aug. 2025, doi: 10.3390/jimaging11080279.

[2] P. Zang, T. T. Hormel, T. S. Hwang, S. T. Bailey, D. Huang, and Y. Jia, “Deep-Learning–Aided Diagnosis of Diabetic Retinopathy, Age-Related Macular Degeneration, and Glaucoma Based on Structural and Angiographic OCT,” Ophthalmol. Sci., vol. 3, no. 1, p. 100245, Mar. 2023, doi: 10.1016/j.xops.2022.100245.

[3] S. Srinivasan et al., “Inter-observer agreement in grading severity of diabetic retinopathy in wide-field fundus photographs,” Eye, vol. 37, no. 6, pp. 1231–1235, Apr. 2023, doi: 10.1038/s41433-022-02107-1.

[4] F. Rizvi, A. Rizvi, and A. V. Das, “Can artificial intelligence improve accessibility to ophthalmic image screening and diagnosis in low- and middle-income countries: a review,” Expert Rev. Ophthalmol., vol. 19, no. 3, pp. 159–161, May 2024, doi: 10.1080/17469899.2024.2321936.

[5] R. Mahum, S. U. Rehman, O. D. Okon, A. Alabrah, T. Meraj, and H. T. Rauf, “A Novel Hybrid Approach Based on Deep CNN to Detect Glaucoma Using Fundus Imaging,” Electronics, vol. 11, no. 1, p. 26, Dec. 2021, doi: 10.3390/electronics11010026.

[6] N. Kumar, M. Sharma, V. P. Singh, C. Madan, and S. Mehandia, “An empirical study of handcrafted and dense feature extraction techniques for lung and colon cancer classification from histopathological images,” Biomed. Signal Process. Control, vol. 75, p. 103596, May 2022, doi: 10.1016/j.bspc.2022.103596.

[7] Puneet, R. Kumar, and M. Gupta, “Optical coherence tomography image based eye disease detection using deep convolutional neural network,” Health Inf. Sci. Syst., vol. 10, no. 1, p. 13, Jun. 2022, doi: 10.1007/s13755-022-00182-y.

[8] M. Alghamdi and M. Abdel-Mottaleb, “A Comparative Study of Deep Learning Models for Diagnosing Glaucoma From Fundus Images,” IEEE Access, vol. 9, pp. 23894–23906, 2021, doi: 10.1109/ACCESS.2021.3056641.

[9] M. K. Jabbar, J. Yan, H. Xu, Z. Ur Rehman, and A. Jabbar, “Transfer Learning-Based Model for Diabetic Retinopathy Diagnosis Using Retinal Images,” Brain Sci., vol. 12, no. 5, p. 535, Apr. 2022, doi: 10.3390/brainsci12050535.

[10] J.-H. Wu, N. D. Koseoglu, C. Jones, and T. Y. A. Liu, “Vision transformers: The next frontier for deep learning-based ophthalmic image analysis,” Saudi J. Ophthalmol., vol. 37, no. 3, pp. 173–178, Jul. 2023, doi: 10.4103/sjopt.sjopt_91_23.

[11] R. Azad et al., “Advances in medical image analysis with vision Transformers: A comprehensive review,” Med. Image Anal., vol. 91, p. 103000, Jan. 2024, doi: 10.1016/j.media.2023.103000.

[12] S. C. Sonmez, M. Sevgi, F. Antaki, J. Huemer, and P. A. Keane, “Generative artificial intelligence in ophthalmology: current innovations, future applications and challenges,” Br. J. Ophthalmol., vol. 108, no. 10, pp. 1335–1340, Oct. 2024, doi: 10.1136/bjo-2024-325458.

[13] K. Kazemzadeh, “Artificial intelligence in ophthalmology: opportunities, challenges, and ethical considerations,” Med. Hypothesis Discov. Innov. Ophthalmol., vol. 14, no. 1, pp. 255–272, May 2025, doi: 10.51329/mehdiophthal1517.

[14] D. Shu et al., “Role of Oxidative Stress in Ocular Diseases: A Balancing Act,” Metabolites, vol. 13, no. 2, p. 187, Jan. 2023, doi: 10.3390/metabo13020187.

[15] P. Kumar, R. Kumar, and M. Gupta, “Deep Learning Based Analysis of Ophthalmology: A Systematic Review,” EAI Endorsed Trans. Pervasive Health Technol., vol. 7, no. 29, p. e4, Sep. 2021, doi: 10.4108/eai.10-9-2021.170950.

[16] A. Das, R. Giri, G. Chourasia, and A. A. Bala, "Classification of Retinal Diseases Using Transfer Learning Approach," in 2019 International Conference on Communication and Electronics Systems (ICCES), Coimbatore, India: IEEE, Jul. 2019, pp. 2080–2084. doi: 10.1109/ICCES45898.2019.9002415.

[17] B. Hassan et al., "A comprehensive review of artificial intelligence models for screening major retinal diseases," Artif. Intell. Rev., vol. 57, no. 5, p. 111, Apr. 2024, doi: 10.1007/s10462-024-10736-z.

[18] K. Delaunay et al., "Glial cells of the human fovea," Mol. Vis., vol. 26, pp. 235–245, 2020.

[19] C. M. Travieso-Gonzalez, Applications of Pattern Recognition. IntechOpen, 2021. doi: 10.5772/intechopen.80151.

[20] L. Zhang, "Transfer learning approach for diabetic retinopathy detection using residual network," Appl. Comput. Eng., vol. 49, pp. 324–331, Mar. 2024, doi: 10.54254/2755-2721/49/20241698.

[21] K. Jain and M. Bharkatiya, "Nanoparticulate based ocular drug delivery technology: 'Nanoemulsion a novel approach,'" Indian J. Pharm. Pharmacol., vol. 7, no. 3, pp. 168–176, Oct. 2020, doi: 10.18231/j.ijpp.2020.029.

[22] M. Zeppieri et al., "Optical Coherence Tomography (OCT): A Brief Look at the Uses and Technological Evolution of Ophthalmology," Medicina (Mex.), vol. 59, no. 12, p. 2114, Dec. 2023, doi: 10.3390/medicina59122114.

[23] T. Yu et al., "A Systematic Review of Advances in AI-Assisted Analysis of Fundus Fluorescein Angiography (FFA) Images: From Detection to Report Generation," Ophthalmol. Ther., vol. 14, no. 4, pp. 599–619, Apr. 2025, doi: 10.1007/s40123-025-01109-y.

[24] N. Unno et al., "Comparison of Multicolor Scanning Laser Imaging and Color Fundus Photography in Evaluating Vessel Whitening in Branch Retinal Vein Occlusion," Ophthalmic Res., vol. 66, no. 1, pp. 413–420, 2023, doi: 10.1159/000528251.

[25] J. He, C. Li, J. Ye, Y. Qiao, and L. Gu, "Multi-label ocular disease classification with a dense correlation deep neural network," Biomed. Signal Process. Control, vol. 63, p. 102167, Jan. 2021, doi: 10.1016/j.bspc.2020.102167.

[26] H. Trinh et al., "Artificial intelligence techniques in inherited retinal diseases: a review," Biomed. Phys. Eng. Express, vol. 11, no. 4, p. 042004, Jul. 2025, doi: 10.1088/2057-1976/ade9c7.

[27] M. M. Butt, D. N. F. A. Iskandar, S. E. Abdelhamid, G. Latif, and R. Alghazo, "Diabetic Retinopathy Detection from Fundus Images of the Eye Using Hybrid Deep Learning Features," Diagnostics, vol. 12, no. 7, p. 1607, Jul. 2022, doi: 10.3390/diagnostics12071607.

[28] I. Qureshi, J. Ma, and Q. Abbas, "Diabetic retinopathy detection and stage classification in eye fundus images using active deep learning," Multimed. Tools Appl., vol. 80, no. 8, pp. 11691–11721, Mar. 2021, doi: 10.1007/s11042-020-10238-4.

[29] "World report on vision." Accessed: Sep. 17, 2025. [Online]. Available: https://www.who.int/publications/i/item/9789241516570

[30] M. Bangar and P. Chaudhary, "A novel approach for the classification of diabetic maculopathy using discrete wavelet transforms and a support vector machine," AIMS Electron. Electr. Eng., vol. 7, no. 1, pp. 1–13, 2022, doi: 10.3934/electreng.2023001.

[31] R. Kashyap, R. Nair, S. M. P. Gangadharan, M. Botto-Tobar, S. Farooq, and A. Rizwan, "Glaucoma Detection and Classification Using Improved U-Net Deep Learning Model," Healthcare, vol. 10, no. 12, p. 2497, Dec. 2022, doi: 10.3390/healthcare10122497.

[32] S.-H. Hung, W.-T. Yen, and D.-W. Lu, "Advances in Glaucoma Diagnosis and Treatment: Integrating Innovations for Enhanced Patient Outcomes," Biomedicines, vol. 13, no. 4, p. 850, Apr. 2025, doi: 10.3390/biomedicines13040850.

[33] H. A. Albarqi, A. Garg, M. Z. Ahmad, A. A. Alqahtani, I. A. Walbi, and J. Ahmad, "Recent Progress in Chitosan-Based Nanomedicine for Its Ocular Application in Glaucoma," Pharmaceutics, vol. 15, no. 2, p. 681, Feb. 2023, doi: 10.3390/pharmaceutics15020681.

[34] S. M. Saqib et al., "Cataract and glaucoma detection based on Transfer Learning using MobileNet," Heliyon, vol. 10, no. 17, p. e36759, Sep. 2024, doi: 10.1016/j.heliyon.2024.e36759.

[35] H. ÇetiNer, "Cataract disease classification from fundus images with transfer learning based deep learning model on two ocular disease datasets," Gümüşhane Üniversitesi Fen Bilim. Enstitüsü Derg., Jan. 2023, doi: 10.17714/gumusfenbil.1168842.

[36] K. Pesudovs et al., "Cataract-related blindness and vision impairment in 2020 and trends over time in relation to VISION 2020: the Right to Sight: an analysis for the Global Burden of Disease Study.," Invest. Ophthalmol. Vis. Sci., vol. 62, no. 8, p. 3523, Jun. 2021.

[37] A. Sirajudeen, A. Balasubramaniam, and S. Karthikeyan, "Novel angular binary pattern (NABP) and kernel based convolutional neural networks classifier for cataract detection," Multimed. Tools Appl., vol. 81, no. 27, pp. 38485–38512, Nov. 2022, doi: 10.1007/s11042-022-13092-8.

[38] Y.-M. Chen, W.-T. Huang, W.-H. Ho, and J.-T. Tsai, "Classification of age-related macular degeneration using convolutional-neural-network-based transfer learning," BMC Bioinformatics, vol. 22, no. S5, p. 99, Nov. 2021, doi: 10.1186/s12859-021-04001-1.

[39] A. Arbab, A. Habibi, H. Rabbani, and M. Tajmirriahi, "From Image to Sequence: Exploring Vision Transformers for

Optical Coherence Tomography Classification," J. Med. Signals Sens., vol. 15, no. 6, Jun. 2025, doi: 10.4103/jmss.jmss_58_24.

[40] "What Is Macular Degeneration?," Cleveland Clinic. Accessed: Mar. 17, 2026. [Online]. Available: https://my.clevelandclinic.org/health/diseases/15246-macular-degeneration

[41] Y. Tong, W. Lu, Y. Yu, and Y. Shen, "Application of machine learning in ophthalmic imaging modalities," Eye Vis., vol. 7, no. 1, p. 22, Dec. 2020, doi: 10.1186/s40662-020-00183-6.

[42] P. K. Karn and W. H. Abdulla, "On Machine Learning in Clinical Interpretation of Retinal Diseases Using OCT Images," Bioengineering, vol. 10, no. 4, p. 407, Mar. 2023, doi: 10.3390/bioengineering10040407.

[43] E. Tonti et al., "Artificial Intelligence and Advanced Technology in Glaucoma: A Review," J. Pers. Med., vol. 14, no. 10, p. 1062, Oct. 2024, doi: 10.3390/jpm14101062.

[44] M. Ashtari-Majlan, M. M. Dehshibi, and D. Masip, "Deep Learning and Computer Vision for Glaucoma Detection: A Review," Jul. 31, 2023, arXiv: arXiv:2307.16528. doi: 10.48550/arXiv.2307.16528.

[45] C. Mohanty et al., "Using Deep Learning Architectures for Detection and Classification of Diabetic Retinopathy," Sensors, vol. 23, no. 12, p. 5726, Jun. 2023, doi: 10.3390/s23125726.

[46] M. A. Al Jbaar and S. A. Dawwd, "DCNN-based embedded models for parallel diagnosis of ocular diseases," East.-Eur. J. Enterp. Technol., vol. 4, no. 2 (124), pp. 53–69, Aug. 2023, doi: 10.15587/1729-4061.2023.281790.

[47] S. M. Khan et al., "A global review of publicly available datasets for ophthalmological imaging: barriers to access, usability, and generalisability," Lancet Digit. Health, vol. 3, no. 1, pp. e51–e66, Jan. 2021, doi: 10.1016/S2589-7500(20)30240-5.

[48] "DRIVE - Grand Challenge," grand-challenge.org. Accessed: Mar. 28, 2026. [Online]. Available: https://drive.grand-challenge.org/

[49] "The STARE Project." Accessed: Mar. 28, 2026. [Online]. Available: https://cecas.clemson.edu/~ahoover/stare/

[50] "CHASE_DB1." [Online]. Available: https://blogs.kingston.ac.uk/retinal/chasedb1/

[51] "CHASEDB1." Accessed: Mar. 28, 2026. [Online]. Available: https://www.kaggle.com/datasets/khoongweihao/chasedb1

[52] "High-Resolution Fundus (HRF) Image Database." [Online]. Available: https://www5.cs.fau.de/research/data/fundus-images/

[53] "Download IOSTAR Retinal Vessel Segmentation Dataset « RetinaCheck." Accessed: Mar. 28, 2026. [Online]. Available: http://www.retinacheck.org/download-iostar-retinal-vessel-segmentation-dataset

[54] G. P. MAFFRE Gervais GAUTHIER, Bruno LAY, Julien ROGER, Damien ELIE, Mélanie FOLTETE, Arthur DONJON, Hugo, "Messidor," ADCIS. Accessed: Mar. 28, 2026. [Online]. Available: https://www.adcis.net/en/third-party/messidor/

[55] G. P. MAFFRE Gervais GAUTHIER, Bruno LAY, Julien ROGER, Damien ELIE, Mélanie FOLTETE, Arthur DONJON, Hugo, "Messidor-2," ADCIS. Accessed: Mar. 28, 2026. [Online]. Available: https://www.adcis.net/en/third-party/messidor2/

[56] "Diabetic Retinopathy Detection." Accessed: Mar. 28, 2026. [Online]. Available: https://kaggle.com/diabetic-retinopathy-detection

[57] "DDR dataset." Accessed: Mar. 28, 2026. [Online]. Available: https://www.kaggle.com/datasets/mariaherrerot/ddrdataset

[58] "IDRiD - Grand Challenge," grand-challenge.org. Accessed: Mar. 28, 2026. [Online]. Available: https://idrid.grand-challenge.org/Rules/

[59] "APTOS 2019 Blindness Detection." Accessed: Mar. 28, 2026. [Online]. Available: https://kaggle.com/aptos2019-blindness-detection

[60] "DiaRetDB1 V2.1." Accessed: Mar. 28, 2026. [Online]. Available: https://www.kaggle.com/datasets/nguyenhung1903/diaretdb1-v21

[61] J. Ye, "GRAPE: A multi-modal glaucoma dataset of follow-up visual field and fundus images for glaucoma management," Aug. 2023, doi: 10.6084/m9.figshare.c.6406319.v1.

[62] "Glaucoma Fundus Imaging Datasets." Accessed: Mar. 29, 2026. [Online]. Available: https://www.kaggle.com/datasets/arnavjain1/glaucoma-datasets

[63] "REFUGE - Grand Challenge," grand-challenge.org. Accessed: Mar. 29, 2026. [Online]. Available: https://refuge.grand-challenge.org/

[64] "CNNs for Automatic Glaucoma Assessment using Fundus Images: An Extensive Validation." figshare. Accessed: Mar. 29, 2026. [Online]. Available: https://figshare.com/s/c2d31f850af14c5b5232

[65] "Drishti-GS - RETINA DATASET FOR ONH SEGMENTATION." Accessed: Mar. 29, 2026. [Online]. Available: https://www.kaggle.com/datasets/lokeshsaipureddi/drishtigs-retina-dataset-for-onh-segmentation

[66] "Drishti-GS Dataset Webpage." Accessed: Mar. 29, 2026. [Online]. Available: https://cvit.iiit.ac.in/projects/mip/drishti-gs/mip-dataset2/Home.php

[67] "MIAG ULL." Accessed: Mar. 29, 2026. [Online]. Available: https://medimrg.webs.ull.es/

[68] “ORIGA for OD and OC segmentation_数据集-飞桨AI Studio星河社区.” Accessed: Mar. 29, 2026. [Online]. Available: https://aistudio.baidu.com/datasetdetail/122940

[69] “Glaucoma Fundus Imaging Datasets.” Accessed: Mar. 29, 2026. [Online]. Available: https://www.kaggle.com/datasets/arnavjain1/glaucoma-datasets

[70] “PALM - Grand Challenge,” grand-challenge.org. Accessed: Mar. 29, 2026. [Online]. Available: https://palm.grand-challenge.org/

[71] “HPMI: A retinal fundus image dataset for identification of high and pathological myopia based on deep learning.” figshare, Dec. 13, 2023. doi: 10.6084/m9.figshare.24800232.v2.

[72] “RIADD (ISBI-2021) - Grand Challenge,” grand-challenge.org. Accessed: Mar. 29, 2026. [Online]. Available: https://riadd.grand-challenge.org/Home/

[73] “Retinal Disease Classification.” Accessed: Mar. 29, 2026. [Online]. Available: https://www.kaggle.com/datasets/andrewmvd/retinal-disease-classification

[74] “1000 Fundus images with 39 categories.” Accessed: Mar. 29, 2026. [Online]. Available: https://www.kaggle.com/datasets/linchundan/fundusimage1000

[75] “Ocular Disease Recognition.” Accessed: Mar. 29, 2026. [Online]. Available: https://www.kaggle.com/datasets/andrewmvd/ocular-disease-recognition-odir5k

[76] “ODIR-2019 - Grand Challenge,” grand-challenge.org. Accessed: Mar. 29, 2026. [Online]. Available: https://odir2019.grand-challenge.org/dataset/

[77] D. Kermany, K. Zhang, and M. Goldbaum, “Labeled Optical Coherence Tomography (OCT) and Chest X-Ray Images for Classification,” vol. 2, Jan. 2018, doi: 10.17632/rscbjbr9sj.2.

[78] “Farsiu Ophthalmology 2013.” Accessed: Mar. 29, 2026. [Online]. Available: https://people.duke.edu/~sf59/RPEDC_Ophth_2013_dataset.htm

[79] “octid-dataset.” Accessed: Mar. 29, 2026. [Online]. Available: https://www.kaggle.com/datasets/kawtarnaim/octid-dataset

[80] U. of W. Peyman Gholami and U. of W. Vasudevan Lakshminarayanan, “Optical Coherence Tomography Image Retinal Database.” Inter-university Consortium for Political and Social Research (ICPSR), Feb. 20, 2019. doi: 10.3886/E108503V1.

[81] “OCTDL: Retinal OCT Images dataset.” Accessed: Mar. 29, 2026. [Online]. Available: https://www.kaggle.com/datasets/shakilrana/octdl-retinal-oct-images-dataset

[82] “Chiu_BOE_2015.” Accessed: Mar. 29, 2026. [Online]. Available: https://people.duke.edu/~sf59/Chiu_BOE_2014_dataset.htm

[83] S. Sotoudeh-Paima, F. Hajizadeh, and H. Soltanian-Zadeh, “Labeled Retinal Optical Coherence Tomography Dataset for Classification of Normal, Drusen, and CNV Cases,” vol. 1, Oct. 2021, doi: 10.17632/8kt969dhx6.1.

[84] A. Montoro-Montarroso et al., “Fighting disinformation with artificial intelligence: fundamentals, advances and challenges,” El Prof. Inf., p. e320322, Jun. 2023, doi: 10.3145/epi.2023.may.22.

[85] A. F. A. H. Alnuaimi and T. H. K. Albaldawi, “An overview of machine learning classification techniques,” BIO Web Conf., vol. 97, p. 00133, 2024, doi: 10.1051/bioconf/20249700133.

[86] R. Seccia, S. Romano, M. Salvetti, A. Crisanti, L. Palagi, and F. Grassi, “Machine Learning Use for Prognostic Purposes in Multiple Sclerosis,” Life, vol. 11, no. 2, p. 122, Feb. 2021, doi: 10.3390/life11020122.

[87] P. Keramida, N. K. Syrigos, M. Kouvela, G. Poulakou, A. Charpidou, and O. Fiste, “AI-Driven Drug Repurposing: Applications and Challenges,” Medicines, vol. 12, no. 4, p. 28, Nov. 2025, doi: 10.3390/medicines12040028.

[88] Y. El Ouahabi, M. H. E. Y. Alaoui, B. Nsiri, A. Soulaymani, A. Mokhtari, and B. Benaji, “Vigilance towards the use of artificial intelligence applications for breast cancer screening and early diagnosis,” E3S Web Conf., vol. 319, p. 01099, 2021, doi: 10.1051/e3sconf/202131901099.

[89] A. Inglis, A. C. Parnell, N. Subramani, and F. M. Doohan, “Machine Learning Applied to the Detection of Mycotoxin in Food: A Systematic Review,” Toxins, vol. 16, no. 6, p. 268, Jun. 2024, doi: 10.3390/toxins16060268.

[90] D. Valkenborg, A.-J. Rousseau, M. Geubbelmans, and T. Burzykowski, “Support vector machines,” Am. J. Orthod. Dentofacial Orthop., vol. 164, no. 5, pp. 754–757, Nov. 2023, doi: 10.1016/j.ajodo.2023.08.003.

[91] I. Karamanlis et al., “Deep Learning-Based Black Spot Identification on Greek Road Networks,” Data, vol. 8, no. 6, p. 110, Jun. 2023, doi: 10.3390/data8060110.

[92] A. B. Silva, E. D. S. Rocha, J. F. Lorenzato, and P. T. Endo, “Evaluating how different balancing data techniques impact on prediction of premature birth using machine learning models,” PLOS ONE, vol. 20, no. 3, p. e0316574, Apr. 2025, doi: 10.1371/journal.pone.0316574.

[93] W. Wongkaew, W. Muanyoksakul, C. Ngamkhanong, J. Sresakoolchai, and S. Kaewunruen, “Data driven machine learning prognostics of buckling failure modes in ballasted railway track,” Discov. Appl. Sci., vol. 6, no. 4, p. 212, Apr. 2024, doi: 10.1007/s42452-024-05885-3.

[94] C. Singh, J. Morris, A. Rush, J. Gao, and Y. Deng, “Tree Prompting: Efficient Task Adaptation without Fine-Tuning,”

in Proceedings of the 2023 Conference on Empirical Methods in Natural Language Processing, Singapore: Association for Computational Linguistics, 2023, pp. 6253–6267. doi: 10.18653/v1/2023.emnlp-main.384.
[95] M. A. Hafeez, M. Rashid, H. Tariq, Z. U. Abideen, S. S. Alotaibi, and M. H. Sinky, “Performance Improvement of Decision Tree: A Robust Classifier Using Tabu Search Algorithm,” Appl. Sci., vol. 11, no. 15, p. 6728, Jul. 2021, doi: 10.3390/app11156728.
[96] M. Georgioudakis and V. Plevris, “Response Spectrum Analysis of Multi-Story Shear Buildings Using Machine Learning Techniques,” Computation, vol. 11, no. 7, p. 126, Jun. 2023, doi: 10.3390/computation11070126.
[97] M. N. U. Alam et al., “Machine learning in biological research: key algorithms, applications, and future directions,” BMC Biol., vol. 23, no. 1, p. 324, Oct. 2025, doi: 10.1186/s12915-025-02424-3.
[98] F. Prezja et al., “Improving performance in colorectal cancer histology decomposition using deep and ensemble machine learning,” Heliyon, vol. 10, no. 18, p. e37561, Sep. 2024, doi: 10.1016/j.heliyon.2024.e37561.
[99] Y. Feng and J. Park, “Using machine learning-based binary classifiers for predicting organizational members’ user satisfaction with collaboration software,” PeerJ Comput. Sci., vol. 9, p. e1481, Jul. 2023, doi: 10.7717/peerj-cs.1481.
[100] F. D. E. Baran and M. Cetin, “AI-driven early diagnosis of specific mental disorders: a comprehensive study,” Cogn. Neurodyn., vol. 19, no. 1, p. 70, Dec. 2025, doi: 10.1007/s11571-025-10253-x.
[101] S. Nyawa, D. Tchuente, and S. Fosso-Wamba, “COVID-19 vaccine hesitancy: a social media analysis using deep learning,” Ann. Oper. Res., vol. 339, no. 1–2, pp. 477–515, Aug. 2024, doi: 10.1007/s10479-022-04792-3.
[102] M. Dirik, “Application of machine learning techniques for obesity prediction: a comparative study,” J. Complex. Health Sci., vol. 6, no. 2, pp. 16–34, Dec. 2023, doi: 10.21595/chs.2023.23193.
[103] M. Pal et al., “Early detection of human Mpox: A comparative study by using machine learning and deep learning models with ensemble approach,” Digit. Health, vol. 11, p. 20552076251344135, May 2025, doi: 10.1177/20552076251344135.
[104] N. A. Saran and F. Nar, “Fast binary logistic regression,” PeerJ Comput. Sci., vol. 11, p. e2579, Jan. 2025, doi: 10.7717/peerj-cs.2579.
[105] J. Musa and A. M. Abdulazeez, “A Review on Diabetes Classification Based on Machine Learning Algorithms,” Indones. J. Comput. Sci., vol. 13, no. 2, Apr. 2024, doi: 10.33022/ijcs.v13i2.3886.
[106] W. Zhou, “A Random Traffic Assignment Model for Networks Based on Discrete Dynamic Bayesian Algorithms,” Discrete Dyn. Nat. Soc., vol. 2022, no. 1, p. 8998352, Jan. 2022, doi: 10.1155/2022/8998352.
[107] Y. Wei and M. Shangguan, “A review of deep learning based intrusion detection systems,” vol. 56, 2023, doi: 10.54097/hset.v56i.10104.
[108] A. E. Ezugwu et al., “Classical Machine Learning: Seventy Years of Algorithmic Learning Evolution,” Data Intell., vol. 7, no. 4, pp. 947–996, Dec. 2025, doi: 10.3724/2096-7004.di.2024.0051.
[109] R. Priyadarshi, R. Ranjan, A. Kumar Vishwakarma, T. Yang, and R. Singh Rathore, “Exploring the Frontiers of Unsupervised Learning Techniques for Diagnosis of Cardiovascular Disorder: A Systematic Review,” IEEE Access, vol. 12, pp. 139253–139272, 2024, doi: 10.1109/ACCESS.2024.3468163.
[110] L. Manduchi, M. Vandenhirtz, A. Ryser, and J. Vogt, “Tree Variational Autoencoders,” Nov. 17, 2023, arXiv: arXiv:2306.08984. doi: 10.48550/arXiv.2306.08984.
[111] A. Yaqoob, R. Musheer Aziz, and N. K. Verma, “Applications and Techniques of Machine Learning in Cancer Classification: A Systematic Review,” Hum.-Centric Intell. Syst., vol. 3, no. 4, pp. 588–615, Sep. 2023, doi: 10.1007/s44230-023-00041-3.
[112] Z. Lu, “Clustering Longitudinal Data: A Review of Methods and Software Packages,” Int. Stat. Rev., vol. 93, no. 3, pp. 425–458, Dec. 2025, doi: 10.1111/insr.12588.
[113] K. Kim, J. Kim, L. A. Wasserman, and E. H. Kennedy, “Hierarchical and Density-based Causal Clustering,” Nov. 02, 2024, arXiv: arXiv:2411.01250. doi: 10.48550/arXiv.2411.01250.
[114] H. Mittal, A. C. Pandey, M. Saraswat, S. Kumar, R. Pal, and G. Modwel, “A comprehensive survey of image segmentation: clustering methods, performance parameters, and benchmark datasets,” Multimed. Tools Appl., vol. 81, no. 24, pp. 35001–35026, Oct. 2022, doi: 10.1007/s11042-021-10594-9.
[115] B. Sasmal and K. G. Dhal, “A survey on the utilization of Superpixel image for clustering based image segmentation,” Multimed. Tools Appl., vol. 82, no. 23, pp. 35493–35555, Sep. 2023, doi: 10.1007/s11042-023-14861-9.
[116] A. Pegado-Bardayo, A. Lorenzo-Espejo, J. Muñuzuri, and A. Escudero-Santana, “A review of unsupervised k-value selection techniques in clustering algorithms,” J. Ind. Eng. Manag., vol. 17, no. 3, p. 641, Aug. 2024, doi: 10.3926/jiem.6791.
[117] N. Pillai, A. R. Das, M. Ayoola, G. Gireesan, B. Nanduri, and M. Ramkumar, “EndToEndML: An Open-Source End-to-End Pipeline for Machine Learning Applications,” Mar. 27, 2024, arXiv: arXiv:2403.18203. doi: 10.48550/arXiv.2403.18203.
[118] A. T. Abdel-Wahed, Z. Ullah, A. S. Abdel-Khalik, M. S. Hamad, S. Ahmed, and N. Elmalhy, “Drive Cycle-Based Design With the Aid of Data Mining Methods: A Review on Clustering Techniques of Electric Vehicle Motor Design With a Case Study,” IEEE Access, vol. 11, pp. 115775–115797, 2023, doi: 10.1109/ACCESS.2023.3325400.

[119] M. Drogkoula, K. Kokkinos, and N. Samaras, "A Comprehensive Survey of Machine Learning Methodologies with Emphasis in Water Resources Management," Appl. Sci., vol. 13, no. 22, p. 12147, Nov. 2023, doi: 10.3390/app132212147.

[120] M. H. A. Khan, "STICA: Spanning-Tree-Based Identification of Clusters With Agglomeration for Numerical Data," IEEE Access, vol. 12, pp. 150430–150450, 2024, doi: 10.1109/ACCESS.2024.3480460.

[121] D. Papakyriakou and I. S. Barbounakis, "Data Mining Methods: A Review," Int. J. Comput. Appl., vol. 183, no. 48, pp. 5–19, Jan. 2022, doi: 10.5120/ijca2022921884.

[122] Q. Zheng, K. Zhang, Q. Chen, J. Wu, and J. Lin, "DKCDC: A clustering algorithm focusing on genuine boundary search for regional division," PLOS One, vol. 20, no. 9, p. e0331555, Sep. 2025, doi: 10.1371/journal.pone.0331555.

[123] K. Aggarwal et al., "Robust Assessment of Clustering Methods for Fast Radio Transient Candidates," Astrophys. J., vol. 914, no. 1, p. 53, Jun. 2021, doi: 10.3847/1538-4357/abf92b.

[124] T. Dinh et al., "Data clustering: a fundamental method in data science and management," Data Sci. Manag., p. S2666764925000360, Aug. 2025, doi: 10.1016/j.dsm.2025.08.001.

[125] R. Patil and S. Sharma, "Automatic glaucoma detection from fundus images using transfer learning," Multimed. Tools Appl., vol. 83, no. 32, pp. 78207–78226, Feb. 2024, doi: 10.1007/s11042-024-18242-8.

[126] A. García-Floriano and E. Ventura-Molina, "Age-Related Macular Degeneration Detection in Retinal Fundus Images by a Deep Convolutional Neural Network," Mathematics, vol. 12, no. 10, p. 1445, May 2024, doi: 10.3390/math12101445.

[127] L. Cheng, "The Practice and Application of Machine Learning in Data Analysis," J. Artif. Intell. Pract., vol. 8, no. 1, pp. 79–85, 2025, doi: 10.23977/jaip.2025.080111.

[128] I. H. Sarker, "Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions," SN Comput. Sci., vol. 2, no. 6, p. 420, Nov. 2021, doi: 10.1007/s42979-021-00815-1.

[129] T. Perumal, N. Mustapha, R. Mohamed, and F. M. Shiri, "A Comprehensive Overview and Comparative Analysis on Deep Learning Models," J. Artif. Intell., vol. 6, no. 1, pp. 301–360, 2024, doi: 10.32604/jai.2024.054314.

[130] E. R. AlBasiouny, A.-F. A. Heliel, H. E. Abdelmunim, and H. M. Abbas, "Multilayer Perceptron Generative Model via Adversarial Learning for Robust Visual Tracking," IEEE Access, vol. 10, pp. 121230–121248, 2022, doi: 10.1109/ACCESS.2022.3222867.

[131] A. Salhi, R. Alshamrani, A. Althbiti, A. Ismail, M. Abd-ElRahman, and B. M. Hassan, "Optimizing high dimensional data classification with a hybrid AI driven feature selection framework and machine learning schema," Sci. Rep., vol. 15, no. 1, p. 35038, Oct. 2025, doi: 10.1038/s41598-025-08699-4.

[132] D. Bhatt et al., "CNN Variants for Computer Vision: History, Architecture, Application, Challenges and Future Scope," Electronics, vol. 10, no. 20, p. 2470, Oct. 2021, doi: 10.3390/electronics10202470.

[133] L. Alzubaidi et al., "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," J. Big Data, vol. 8, no. 1, p. 53, Mar. 2021, doi: 10.1186/s40537-021-00444-8.

[134] A. W. Salehi et al., "A Study of CNN and Transfer Learning in Medical Imaging: Advantages, Challenges, Future Scope," Sustainability, vol. 15, no. 7, p. 5930, Mar. 2023, doi: 10.3390/su15075930.

[135] M.E student from Madhav Institute of Technology and Science, Gwalior, Madhya Pradesh, India., A. Gupta, Dr. R. P. Narwaria, and Assistant Professor in MITS, Gwalior, Madhya Pradesh, India., "Review on Deep Learning Handwritten Digit Recognition using Convolutional Neural Network," Int. J. Recent Technol. Eng. IJRTE, vol. 9, no. 5, pp. 245–247, Jan. 2021, doi: 10.35940/ijrte.E5287.019521.

[136] M. Z. Khan, M. K. Gajendran, Y. Lee, and M. A. Khan, "Deep Neural Architectures for Medical Image Semantic Segmentation: Review," IEEE Access, vol. 9, pp. 83002–83024, 2021, doi: 10.1109/ACCESS.2021.3086530.

[137] M. B. Shaikh and D. Chai, "RGB-D Data-Based Action Recognition: A Review," Sensors, vol. 21, no. 12, p. 4246, Jun. 2021, doi: 10.3390/s21124246.

[138] Y. Liu et al., "Brain model state space reconstruction using an LSTM neural network," J. Neural Eng., vol. 20, no. 3, p. 036024, Jun. 2023, doi: 10.1088/1741-2552/acd871.

[139] A. Gutiérrez, J. Bobadilla, and S. Alons, "Comparison of Models for Predicting the Number of Calls Received in a Call Center through Time Series Analysis," J. Adv. Inf. Technol., vol. 15, no. 11, pp. 1273–1282, 2024, doi: 10.12720/jait.15.11.1273-1282.

[140] W. Fang, Y. Chen, and Q. Xue, "Survey on Research of RNN-Based Spatio-Temporal Sequence Prediction Algorithms," J. Big Data, vol. 3, no. 3, pp. 97–110, 2021, doi: 10.32604/jbd.2021.016993.

[141] S.-W. Park, J.-S. Ko, J.-H. Huh, and J.-C. Kim, "Review on Generative Adversarial Networks: Focusing on Computer Vision and Its Applications," Electronics, vol. 10, no. 10, p. 1216, May 2021, doi: 10.3390/electronics10101216.

[142] Y. Zhao, "Application of deep learning for prediction of alzheimer's disease in PET/MR imaging," Bioengineering, vol. 10, no. 10, p. 1120, Sep. 2023, doi: 10.3390/bioengineering10101120.

[143] Y. Zhao et al., "Application of Deep Learning for Prediction of Alzheimer's Disease in PET/MR Imaging," Bioengineering, vol. 10, no. 10, p. 1120, Sep. 2023, doi: 10.3390/bioengineering10101120.

[144] V. Machaca et al., "Transformers meets neoantigen detection: a systematic literature review," J. Integr. Bioinforma.,

vol. 21, no. 2, p. 20230043, Aug. 2024, doi: 10.1515/jib-2023-0043.
[145] A. Rahali and M. A. Akhloufi, “End-to-End Transformer-Based Models in Textual-Based NLP,” AI, vol. 4, no. 1, pp. 54–110, Jan. 2023, doi: 10.3390/ai4010004.
[146] Z. Liu et al., “Swin Transformer: Hierarchical Vision Transformer using Shifted Windows,” in 2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, QC, Canada: IEEE, Oct. 2021, pp. 9992–10002. doi: 10.1109/ICCV48922.2021.00986.
[147] A. Bennetot et al., “A Practical Tutorial on Explainable AI Techniques,” ACM Comput. Surv., vol. 57, no. 2, pp. 1–44, Feb. 2025, doi: 10.1145/3670685.
[148] A. Vaswani et al., “Attention is All you Need,” NeurIPS, 2017.
[149] G. Papanastasiou, N. Dikaios, J. Huang, C. Wang, and G. Yang, “Is Attention all You Need in Medical Image Analysis? A Review,” IEEE J. Biomed. Health Inform., vol. 28, no. 3, pp. 1398–1411, Mar. 2024, doi: 10.1109/JBHI.2023.3348436.
[150] A. Dosovitskiy et al., “An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale,” Jun. 03, 2021, arXiv: arXiv:2010.11929. doi: 10.48550/arXiv.2010.11929.
[151] X. Yang, “An Overview of the Attention Mechanisms in Computer Vision,” J. Phys. Conf. Ser., vol. 1693, no. 1, p. 012173, Dec. 2020, doi: 10.1088/1742-6596/1693/1/012173.
[152] P. Jiang, X. Hou, and Y. Zhou, “Device Tuning for Multi-Task Large Model,” Feb. 21, 2023, arXiv: arXiv:2302.10820. doi: 10.48550/arXiv.2302.10820.
[153] K. Lu et al., “DataSculpt: Crafting Data Landscapes for Long-Context LLMs through Multi-Objective Partitioning,” Oct. 02, 2024, arXiv: arXiv:2409.00997. doi: 10.48550/arXiv.2409.00997.
[154] P. Xu, X. Zhu, and D. A. Clifton, “Multimodal Learning with Transformers: A Survey,” May 10, 2023, arXiv: arXiv:2206.06488. doi: 10.48550/arXiv.2206.06488.
[155] W. L. Alyoubi, M. F. Abulkhair, and W. M. Shalash, “Diabetic Retinopathy Fundus Image Classification and Lesions Localization System Using Deep Learning,” Sensors, vol. 21, no. 11, p. 3704, May 2021, doi: 10.3390/s21113704.
[156] J. Li et al., “An Interpretable High-Accuracy Method for Rice Disease Detection Based on Multisource Data and Transfer Learning,” Plants, vol. 12, no. 18, p. 3273, Sep. 2023, doi: 10.3390/plants12183273.
[157] Y. Hagiwara, K. Taguchi, S. Ishibushi, A. Taniguchi, and T. Taniguchi, “Hierarchical Bayesian model for the transfer of knowledge on spatial concepts based on multimodal information,” Adv. Robot., vol. 36, no. 1–2, pp. 33–53, Jan. 2022, doi: 10.1080/01691864.2021.2004224.
[158] Q. Chen, T. Leng, S. Niu, and E. Trucco, “Editorial: Generalizable and explainable artificial intelligence methods for retinal disease analysis: challenges and future trends,” Front. Med., vol. 11, p. 1465369, Jul. 2024, doi: 10.3389/fmed.2024.1465369.
[159] S. Nazir, D. M. Dickson, and M. U. Akram, “Survey of explainable artificial intelligence techniques for biomedical imaging with deep neural networks,” Comput. Biol. Med., vol. 156, p. 106668, Apr. 2023, doi: 10.1016/j.compbiomed.2023.106668.
[160] Y. Sabiri, W. Houmaidi, and A. Abouaomar, “EYE-DEX: Eye Disease Detection and EXplanation System,” Sep. 29, 2025, arXiv: arXiv:2509.24136. doi: 10.48550/arXiv.2509.24136.
[161] V. Sarao, D. Veritti, A. De Nardin, M. Misciagna, G. Foresti, and P. Lanzetta, “Explainable artificial intelligence model for the detection of geographic atrophy using colour retinal photographs,” BMJ Open Ophthalmol., vol. 8, no. 1, p. e001411, Dec. 2023, doi: 10.1136/bmjophth-2023-001411.
[162] Y. A. Fahim, I. W. Hasani, S. Kabba, and W. M. Ragab, “Artificial intelligence in healthcare and medicine: clinical applications, therapeutic advances, and future perspectives,” Eur. J. Med. Res., vol. 30, no. 1, p. 848, Sep. 2025, doi: 10.1186/s40001-025-03196-w.
[163] S. Acosta-Jiménez et al., “Assessing ResNeXt and RegNet Models for Diabetic Retinopathy Classification: A Comprehensive Comparative Study,” Diagnostics, vol. 15, no. 15, p. 1966, Aug. 2025, doi: 10.3390/diagnostics15151966.
[164] M. I. Nazir, A. Akter, M. A. Hussen Wadud, and M. A. Uddin, “Utilizing customized CNN for brain tumor prediction with explainable AI,” Heliyon, vol. 10, no. 20, p. e38997, Oct. 2024, doi: 10.1016/j.heliyon.2024.e38997.
[165] Y. Li, X. Yi, J. Fu, Y. Yang, C. Duan, and J. Wang, “Reducing misdiagnosis in AI-driven medical diagnostics: a multidimensional framework for technical, ethical, and policy solutions,” Front. Med., vol. 12, p. 1594450, Oct. 2025, doi: 10.3389/fmed.2025.1594450.
[166] M. H. Udin et al., “Investigating methods to enhance interpretability and performance in cardiac MRI for myocardial scarring diagnosis using convolutional neural network classification and One Match,” PLOS One, vol. 20, no. 6, p. e0313971, Jun. 2025, doi: 10.1371/journal.pone.0313971.
[167] S. Iftikhar, N. Anjum, A. B. Siddiqui, M. Ur Rehman, and N. Ramzan, “Explainable CNN for brain tumor detection and classification through XAI based key features identification,” Brain Inform., vol. 12, no. 1, p. 10, Dec. 2025, doi: 10.1186/s40708-025-00257-y.
[168] A. M. Mutawa, S. Alnajdi, and S. Sruthi, “Transfer Learning for Diabetic Retinopathy Detection: A Study of Dataset Combination and Model Performance,” Appl. Sci., vol. 13, no. 9, p. 5685, May 2023, doi: 10.3390/app13095685.

[169] "Detection and Diagnosis of Diabetic Retinopathy Using Transfer Learning Approach," Int. J. Intell. Eng. Syst., vol. 16, no. 3, pp. 62–74, Jun. 2023, doi: 10.22266/ijies2023.0630.05.
[170] A. Jabbar et al., "Deep Transfer Learning-Based Automated Diabetic Retinopathy Detection Using Retinal Fundus Images in Remote Areas," Int. J. Comput. Intell. Syst., vol. 17, no. 1, p. 135, May 2024, doi: 10.1007/s44196-024-00520-w.
[171] R. Bhoopalan, P. Sekar, N. Nagaprasad, T. R. Mamo, and R. Krishnaraj, "Task optimized vision transformer for diabetic retinopathy detection and classification in resource constrained early diagnosis settings," Sci. Rep., vol. 15, no. 1, p. 39047, Nov. 2025, doi: 10.1038/s41598-025-25399-1.
[172] N. Varma, S. Yadav, and J. K. P. S. Yadav, "A reliable automatic cataract detection using deep learning," Int. J. Syst. Assur. Eng. Manag., vol. 14, no. 3, pp. 1089–1102, Jun. 2023, doi: 10.1007/s13198-023-01923-2.
[173] S. M. Saqib et al., "Cataract and glaucoma detection based on Transfer Learning using MobileNet," Heliyon, vol. 10, no. 17, p. e36759, Sep. 2024, doi: 10.1016/j.heliyon.2024.e36759.
[174] S. S. Mahmood, S. Chaabouni, and A. Fakhfakh, "Improving Automated Detection of Cataract Disease through Transfer Learning using ResNet50," Eng. Technol. Appl. Sci. Res., vol. 14, no. 5, pp. 17541–17547, Oct. 2024, doi: 10.48084/etasr.8530.
[175] Z. Feng, K. Xu, L. Li, and Y. Wang, "Hybrid Deep Learning Model for Cataract Diagnosis Assistance," Appl. Sci., vol. 14, no. 23, p. 11314, Dec. 2024, doi: 10.3390/app142311314.
[176] A. A. Jabber, A. Hadi, S. Wadi, and G. Ali, "Cataract Detection and Classification Using Deep Learning Techniques," Int. J. Comput. Digit. Syst., vol. 17, no. 1, pp. 1–10, Jan. 2025, doi: 10.12785/ijcds/1571037888.
[177] L. K. Singh, Pooja, H. Garg, and M. Khanna, "Histogram of Oriented Gradients (HOG)-Based Artificial Neural Network (ANN) Classifier for Glaucoma Detection:," Int. J. Swarm Intell. Res., vol. 13, no. 1, pp. 1–32, Oct. 2022, doi: 10.4018/IJSIR.309940.
[178] V. K. Velpula and L. D. Sharma, "Multi-stage glaucoma classification using pre-trained convolutional neural networks and voting-based classifier fusion," Front. Physiol., vol. 14, p. 1175881, Jun. 2023, doi: 10.3389/fphys.2023.1175881.
[179] F. Chincholi and H. Koestler, "Transforming glaucoma diagnosis: transformers at the forefront," Front. Artif. Intell., vol. 7, p. 1324109, Jan. 2024, doi: 10.3389/frai.2024.1324109.
[180] S. J. Edan and M.-R. Feizi-Derakhshi, "An Efficient Glaucoma Classification for Fundus Images by Pre-trained Deep CNNs and Accuracy-Weighted Voting Classifier Fusion," Baghdad Sci. J., vol. 22, no. 8, pp. 2813–2823, Aug. 2025, doi: 10.21123/2411-7986.5040.
[181] Z. Wang, S. R. Sadda, A. Lee, and Z. J. Hu, "Automated segmentation and feature discovery of age-related macular degeneration and Stargardt disease via self-attended neural networks," Sci. Rep., vol. 12, no. 1, p. 14565, Aug. 2022, doi: 10.1038/s41598-022-18785-6.
[182] S. S. Mishra, B. Mandal, and N. B. Puhan, "MacularNet: Towards Fully Automated Attention-Based Deep CNN for Macular Disease Classification," SN Comput. Sci., vol. 3, no. 2, p. 142, Mar. 2022, doi: 10.1007/s42979-022-01024-0.
[183] J. Han et al., "Detecting Macular Disease Based on Optical Coherence Tomography Using a Deep Convolutional Network," J. Clin. Med., vol. 12, no. 3, p. 1005, Jan. 2023, doi: 10.3390/jcm12031005.
[184] Md. A. Ali, Md. S. Hossain, Md. K. Hossain, S. S. Sikder, S. A. Khushbu, and M. Islam, "AMDNet23: Hybrid CNN-LSTM deep learning approach with enhanced preprocessing for age-related macular degeneration (AMD) detection," Intell. Syst. Appl., vol. 21, p. 200334, Mar. 2024, doi: 10.1016/j.iswa.2024.200334.
[185] M. M. Azizi, S. Abhari, and H. Sajedi, "Stitched vision transformer for age-related macular degeneration detection using retinal optical coherence tomography images," PLOS ONE, vol. 19, no. 6, p. e0304943, Jun. 2024, doi: 10.1371/journal.pone.0304943.
[186] E. Yusufoğlu et al., "A Comprehensive CNN Model for Age-Related Macular Degeneration Classification Using OCT: Integrating Inception Modules, SE Blocks, and ConvMixer," Diagnostics, vol. 14, no. 24, p. 2836, Dec. 2024, doi: 10.3390/diagnostics14242836.
[187] I. D. Mienye, T. G. Swart, G. Obaido, M. Jordan, and P. Ilono, "Deep Convolutional Neural Networks in Medical Image Analysis: A Review," Information, vol. 16, no. 3, p. 195, Mar. 2025, doi: 10.3390/info16030195.
[188] L. Alsadoun et al., "Artificial Intelligence (AI)-Enhanced Detection of Diabetic Retinopathy From Fundus Images: The Current Landscape and Future Directions," Cureus, Aug. 2024, doi: 10.7759/cureus.67844.
[189] J. Lo et al., "Federated Learning for Microvasculature Segmentation and Diabetic Retinopathy Classification of OCT Data," Ophthalmol. Sci., vol. 1, no. 4, p. 100069, Dec. 2021, doi: 10.1016/j.xops.2021.100069.
[190] T. D. Nguyen, D.-T. Le, J. Bum, S. Kim, S. J. Song, and H. Choo, "Retinal Disease Diagnosis Using Deep Learning on Ultra-Wide-Field Fundus Images," Diagnostics, vol. 14, no. 1, p. 105, Jan. 2024, doi: 10.3390/diagnostics14010105.
[191] S. Paul and P.-Y. Chen, "Vision Transformers Are Robust Learners," Proc. AAAI Conf. Artif. Intell., vol. 36, no. 2, pp. 2071–2081, Jun. 2022, doi: 10.1609/aaai.v36i2.20103.

[192] A. Kirillov et al., “Segment Anything,” in 2023 IEEE/CVF International Conference on Computer Vision (ICCV), Paris, France: IEEE, Oct. 2023, pp. 3992–4003. doi: 10.1109/ICCV51070.2023.00371.
[193] M. Kulyabin et al., “Segment Anything in Optical Coherence Tomography: SAM 2 for Volumetric Segmentation of Retinal Biomarkers,” Bioengineering, vol. 11, no. 9, p. 940, Sep. 2024, doi: 10.3390/bioengineering11090940.
[194] P. Li, J. Tao, Q. Yuan, R. Zhang, and P. Gao, “Research progress in deep learning-based fundus image analysis for the diagnosis and prediction of hypertension-related diseases,” Front. Cell Dev. Biol., vol. 13, p. 1608994, Aug. 2025, doi: 10.3389/fcell.2025.1608994.
[195] B. Guo et al., “WaveAttention-ResNet: a deep learning-based intelligent diagnostic model for the auxiliary diagnosis of multiple retinal diseases,” Front. Radiol., vol. 5, p. 1608052, Jul. 2025, doi: 10.3389/fradi.2025.1608052.
[196] Z. Li, S. Yin, S. Wang, Y. Wang, W. Qiang, and J. Jiang, “Transformative applications of oculomics-based AI approaches in the management of systemic diseases: A systematic review,” J. Adv. Res., vol. 75, pp. 679–696, Sep. 2025, doi: 10.1016/j.jare.2024.11.018.