BIOMEDICAL ENGINEERING | REVIEW ARTICLE

Current state of artificial intelligence applications in ophthalmology and their potential to influence clinical practice

Dasharathraj K Shetty¹, Abhiroop Talasila², Swapna Shanbhag³, Vathsala Patil⁴, B.M Zeeshan Hameed⁵, Nithesh Naik⁶* and Adithya Raju⁷

Abstract: Artificial intelligence (AI) has emerged as a major frontier in healthcare and finds broad range of applications. It has the potential to revolutionize current procedures of disease diagnosis and treatment, thus influencing the clinical practice. Artificial intelligence (AI) in ophthalmology, primarily concentrates on diagnostic and treatment pathways for eye conditions such as cataract, glaucoma, age-related macular degeneration (MDA) and diabetic retinopathy (DR). The purpose of this article is to systematically review the existing state of literature on the various AI techniques and its applications in the diagnosis and treatment of eye diseases and conduct an in-depth enquiry to identify the challenges in accurate detection, pre-processing of data, monitoring and assessment through various AI algorithms. The results suggest that all AI models proposed reduce the detection time considerably. The potential limitations and challenges in the development and application play a significant role in clinical practice. There is a need for the development of AI-assisted technologies that shall consider the clinical implications based on experience and guided by patient-centred healthcare principles. The diagnostic models should assist ophthalmologists on making quick and accurate decisions in determining the progression of various ocular diseases.

ABOUT THE AUTHOR

Dasharathraj K Shetty is a faculty member of Department of Humanities and Management, Manipal Institute of Technology (MIT), Manipal Academy of Higher Education (MAHE), Manipal. He is an Author, Columnist, Engineer and Social Entrepreneur. He is also the Secretary-General of Indian Bureau of Administrators and Technocrats and the Director of Micro Souharda Credit Cooperative Ltd. He has recently authored the book “Learning like a Lion”. Dasharathraj is a B.E. (Computer Science and Engineering) and has three Post-graduation Degrees - MBA (Finance), MPhil (Management) and M.Tech (Computer Science and Engineering). He was awarded a PhD by MAHE, Manipal. He is also a Certified Microsoft Certified Technology Specialist, Dale Carnegie High Impact Teaching Skills, AIMA Certified Management Trainer and RBNQA Examiner.

PUBLIC INTEREST STATEMENT

Artificial Intelligence (AI), a technology that enables machines and equipment to “learn” and adapt from their experience. AI-based platforms have obtained clinically acceptable diagnostic efficiency in the automated diagnosis of many retinal diseases. AI in ophthalmology concentrates primarily on illnesses like diabetic retinopathy (DR), glaucoma and macular degeneration related with age (MDA). Artificial intelligence (AI) is a technology that can significantly influence the field of ophthalmology in the coming decades. In this review, an in-depth enquiry to identify the challenges in accurate detection, pre-processing of data, monitoring and assessment through various AI algorithms is conducted. This study contributes to the literature on AI models by thorough methodical processes for applying diagnostic AI systems and its application for the diagnosis of retinal diseases using evidence-based systematic approach.
1. Introduction

Artificial Intelligence (AI) is the reproduction of the human thinking process by computers. It has a wide array of applications including but not limited to, law enforcement, finance, education and healthcare. AI platforms are used to replicate human intelligence and then made to complete tasks such as image and speech recognition, sentiment analysis and problem solving (Heidary & Gharebaghi, 2012; Hogarty et al., 2018). Traditionally, a set of instructions is given to computer to complete a task and allow machines to make independent decisions, i.e. without any programming, they undergo a task called Machine Learning (ML).

Machine learning is a process where a computer trains itself based on sample labelled data with a validation dataset, and a basic learning structure for chosen algorithms. There are three types of machine learning algorithms:

1. Supervised machine-learning: Data sets are labelled and fed into the classifier for training.

2. Unsupervised machine-learning: Data sets are not labelled initially, but sorted according to their differences or similarities and then trained.

3. Reinforcement learning: Memory networking differentiates the connections between supervised and unsupervised learned information (Lee et al., 2017).

Algorithms that a computer uses to learn and classify data are termed as classifiers. Some examples of classifiers of artificial intelligence that include support vector machines (SVM) and neural networks are shown in Table 1.

SVM is a popular supervised learning classifier (Figure 1). In this classifier, we first plot each data value as a coordinate in n-dimensional space. Then, classification finds the hyper-plane differentiating different classes. SVMs are memory efficient and work well with a clear margin of separation. However, training time is proportional to input size and less accurate when different classes' overlap (Sevik et al., 2014).

Traditional supervised training algorithms suffer when training on huge amounts of data. To solve this problem, Deep learning (DL), a subset of AI is used. DL algorithms take inspiration from the functioning and structure of the human brain. Using DL algorithms, performance of neural networks increases rapidly. DL is based on Artificial Neural Networks (ANN), which are inspired from biological neural networks. The unique feature of ANNs is that each layer learns different features with different weights for different stimuli. When multiple layers are used, it is termed as deep learning. A Convolutional Neural Network (CNN) is a collection of deep neural networks, most commonly applied to analysing visual imagery.

At present, AI has applications in fields such as cancer, cardiology and neurology. Given deep learning's practical implications in image processing, our technology needs to improve to be able to process various other imaging modalities such as computed tomography scans and magnetic resonance imaging. While AI tools vary in functionality, their use raises questions in terms of their reliability because humans select the data used to train the AI program and that the potential for
human bias may undermine the platform. This review directly explores the current state of artificial intelligence applications in ophthalmology and discusses about the current limitations and the future challenges in clinical practice.

2. Methods

2.1. Method of literature search
Detailed and systematic literature review was conducted according to the Preferred Reporting Items for Systematic Reviews and Meta-analysis (PRISMA) guidelines, to identify all potential relevant studies using the PubMed®, Springer®, Medline®, Scopus®, Web of knowledge®, and Inspec® databases for reporting artificial intelligence in ophthalmology. The articles were searched with the search terms: “Artificial intelligence” and “ophthalmology” “Artificial intelligence” and “glaucoma”, “Artificial intelligence” and “diabetic retinopathy”, “Artificial intelligence” and “macular degeneration”, “Artificial intelligence” and “cataract”.

2.2. Inclusion and exclusion criteria
The primary goal of this systematic review was to include all published articles, which utilized AI platforms for automatic detection of retinal diseases. Figure 2 shows a flow chart of study selection process for the articles considered in the study for analysis. A search was performed using the keywords: Ophthalmology, artificial intelligence, CNN, machine learning, and deep learning. Inclusion and exclusion criteria of the articles are as follows: a) Exposure of interest: Only patients with confirmed ocular diseases were included in the study. b) Language: Only studies written in English were included. c) Reported outcomes: Only outcomes obtained using objective measures were included. d) Type of publication: Review articles were excluded.

2.3. Data extraction
Titles and abstracts have been screened for all identified studies. Irrelevant and duplicate papers were excluded, and a full text review was conducted to examine the remaining articles for
compliance with requirements for inclusion and exclusion. The data extracted at this stage included the title, year of publication, authors, study objective, type of study, diagnostic criteria and selection criteria of participants, method, algorithm, results and performance metrics. The recent articles were handpicked, curated for the current year and subject to the same inclusion criteria. For articles cited in the results bibliographies, the same search strategy was used.

3. Results
A total of 773 articles were identified in the initial search performed. After screening the titles 454 article and abstract 296 articles were excluded, as per inclusion and exclusion criteria. Based on full text accessibility from the mentioned databases in method of literature 23 articles, which met all inclusion criteria were included for the review. Other 32 articles were included after searching bibliographies that highlighted the methods and algorithms of application, challenges and future perspectives of artificial intelligence in healthcare.

3.1. Building artificial intelligence models
Various imaging methods used in healthcare support AI platforms in diagnosis, such as Computed Tomography (CT), X-rays, fundus images, etc. Fundus images and Optical Coherence Tomography (OCT) scans are the most prevalent and widely available methods used. In order to create an AI model, picture information are first processed, trained, validated and assessed with corresponding classifiers as shown in Figure 3.
3.2. Data pre-processing
The images used are often obtained from multiple sources and to provide accurate guidelines for the classifier, some pre-processing steps need to be carried out on the data sets. Many pre-processing steps often involve converting the image data sets into greyscale or extracting a particular RGB channel to obtain a more detailed image. The green channel is preferred generally as it appears more contrasted compared to blue and red channels (Faroq & Sattar, 2015).

3.3. Training and validation
After pre-processing the images, the classifier is trained with labelled data (supervised-learning). To ensure better accuracy, datasets are divided into training and testing sets. While the classifier defines parameters for the model using the training data, a validation set (partitioned from the training set) is used to verify the model's performance and tune the parameters if necessary. To evaluate the trained model, the testing set is used.

3.4. Evaluation
A Receiver Operating Characteristic (ROC) curve shows the variation of the diagnostic ability of a binary classifier system as its discrimination threshold is varied (Hajian-Tilaki, 2013). It is created by plotting probability of detection against probability of false alarm. To assess a model in AI diagnosis, area under receiver operating characteristic curves (AUC) is widely used (Paulus et al., 2010). Effective models have an AUC value in the range of 0.5 to 1.

4. Applications in ophthalmology
Many studies have shown that deep learning models can nearly achieve and sometimes even exceed human performance (Al-Fadhili et al., 2017). In view of recent publications, almost 85%-90% of them focus on DR, MDA, glaucoma and cataract. The probability deviation map acts as a virtual environment filter that helps us know the problems of a patient while walking by demonstrating which regions of his vision have been affected. Incorrect identification of intra-retinal cystoid fluid regions in OCT images might result in the removal of existing cysts or ignore false positives, hindering the performance of the AI system (Al-Fadhili et al., 2017; Moura et al., 2017).

4.1. Diabetic retinopathy
A pooled assessment of 22,896 diabetes sufferers from 35 population-based studies (from 1980 to 2008), carried out in the USA, Australia, Europe, and Asia, showed that the prevalence of DR (in type 1 and type 2 diabetes) reached up to 34.6% with 7% vision-related retinopathy (DR) (Yau et al., 2012). Various health care professionals including ophthalmologists, optometrists, general practitioners and clinical photographers can perform DR screening. Screening procedures include direct ophthalmoscopic screening, dilated bio microscopy with hand-held lenses (90 D or 78 D), mydriatic or non-mydriatic screening, tele retinal screening and video recording. However, the problems of
execution, the availability of human assessors and long-term economic sustainability are questioned in DR screening programmes.

DR is one of the most commonly researched fields of AI applications globally for visual impairment. The ability to recognize images of NN models allows users to use widespread fundus images for early identification. Gulshan et al. were the first to identify the applications of DL models for DR identification (Gulshan et al., 2016). They used supervised learning and very large fundus data sets to train a deep CNN (DCNN). The high number of input data sets resulted in a very good statistical performance with an AUC of 0.99. Chowdhury et al. developed a Random Forest classifier with five performance measures, which obtained an accuracy of 93.58% and an AUC of 0.893, which significantly outperforms the Naive Bayes classifier (NBC) (Chowdhury et al., 2018).

To counter the shortage of specialists in the field, Hnoohom et al. used a local thresholding method to separate foreground region from background region to clearly identify the Optical Disc (OD) and exudates that in turn facilitated in making preliminary decisions, reducing the detection time (Hnoohom & Tanthuwapathom, 2017). Exudates in the layers of the retina obstruct vision and early detection helps to estimate the severity of the condition. Ravindraiah et al. (Krishnapuram & Keller, 1993) presented a framework for exudate detection using a Spatial Possibilistic C-means Clustering (SPCM) algorithm proposed by Krishnapuram et al. (Ravindraiah & Chandra Mohan Reddy, 2018) to avoid the problem of outliers and noise in the traditional Fuzzy c-means clustering algorithm (FCM). Takahashi et al. tried a different approach towards DR staging by using nonmydriatic posterior pole photographs to train a modified GoogLeNet DCNN (Takahashi et al., 2017).

4.2. Age-related macular degeneration

AMD is an important cause for impaired vision in the world’s elderly population. The Age-Related Eye Disease Study (AREDS) classified no, early, intermediate and late AMD stages of AMD. The American Academy of Ophthalmology recommends that intermediate AMD patients be seen at least once every 2 years. 288 million patients are expected to be able to experience some types of AMD until 2040, with an intermediate AMD or worse of about 10%. The elderly population is in need of an urgent clinical system to screen these patients in tertiary eye care centres (Bogunovic et al., 2017; Garcia-Floriano et al., 2019). AMD is a retinal disease that occurs when the macula is damaged and hence, central vision is lost. This occurs secondary to deposition of lipid in the macular region; these depositions are known as drusens. A gradient segmentation-based algorithm to identify the borders of the drusen is used in the detection and segmentation process. Floriano et al. developed a feature vector from the processed fundus image and then trained an SVM to automatically detect the drusen (Garcia-Floriano et al., 2019).

Anti-vascular endothelial growth factor therapy (anti-VEGF) prevents vascular endothelial growth, thus reducing the growth of new vessels in the macular region of the retina. Using ML to predict anti-VEGF injection requirements for AMD can reduce economic burden on patients. Bogunovic et al. used a random forest classifier to train Optical coherence tomography (OCT) images of patients on anti-VEGF medication to predict future requirements and obtained very solid AUC between 70% and 80% for the predictive model (Bogunovic et al., 2017). However, when a DCNN was trained with OCT scans regarding anti-VEGF injection requirements, Prahs et al. obtained better accuracy than using a RF classifier (Prahs et al., 2017). These studies proved to be an important step towards using image modalities to predict treatment intervals for medication for AMD.

Typically, stages of AMD were classified as none, early-stage, intermediate-stage and advanced (Philippe Burlina et al., 2017). Nevertheless, when a 2-stage classification system (none or early-stage and intermediate or advanced AMD) was used, the diagnostic accuracy was better than that of the 4-stage classification system (Schmidt-Erfurth et al., 2018) Various DL platforms have been used for the automatic detection of various abnormalities such as drusen and exudates.
4.3. Glaucoma

Glaucoma is a condition that is due to increase in intraocular pressure, which subsequently affects the optic nerve. When it is not detected early, glaucoma can lead to permanent visual loss (Wang et al., 2019). For individuals 40–80 years old, the global prevalence of glaucoma is 3.4%, and it is forecast that around 112 million people are impacted globally by 2040. The improvements of disease detection, assessing the progressive structural and functional harm, optimizing therapy for visual impairment, and precise long-term forecasts would be both welcome by clinicians and patients (Galilea et al., 2007; Wang et al., 2019). Glaucoma is an optic nerve disorder that is clinically manifested by enhanced optic nerve head (ONH) classified by excavation and neuroretinal edge erosion. However, because the ONH region differs by five, almost no Cup to Disk Ratio (CDR) describes pathological cupping, preventing the detection of a disease. The most common type of glaucoma is Open-angle glaucoma, which occurs when fluid does not flow normally out of the trabecular meshwork. Most glaucoma patients suffer from high intraocular pressure (IOP) which leads to and retinal nerve fiber layer defects with concomitant damage to the optic nerve, which leads to visual loss. Early diagnosis and automatic detection in older patients has proven to be highly beneficial. Apreutesei et al. (Apreutesei et al., 2018) attempted to establish a relation between open-angle glaucoma and diabetes in patients. Fundus images were pre-processed and trained using a back-propagation algorithm. The feed forward neural model with parameters like cup-disc ratio, intraocular pressure, glycosylated haemoglobin levels attained an accuracy of 95%.

The second type of glaucoma is Angle-closure glaucoma (ACG) which might occur when the drainage space between the iris and cornea becomes too narrow (Apreutesei et al., 2018). This form of glaucoma causes a sudden increase in the intraocular pressure and is an ocular emergency. Niwas et al. (2016) (Galilea et al., 2007) proposed segmenting AS-OCT images based on the four major mechanisms causing ACG as the required medication differs for each mechanism. Morphological features were extracted from the AS-OCT images, features with minimum redundancy were trained using Neighborhood-Based Clustering (NBC), and 89.2% accuracy was obtained using a leave-one-out cross-validation method. However, Galilea et al. who implemented a multilayer perceptron ANN with backpropagation and analyzed nerve fibers using laser polarimetry obtained the most promising results (Li et al., 2018). This allowed them to analyze more parameters and the final neural network model had an accuracy of 100%.

Diagnostic models built using both fundus images and OCT scans. Primary cause for inaccurate results is misalignment of optic disc on the fundus images. Li et al. in their study presented that DL can be applied to identify referable glucomatous optic neuropathy with high sensitivity and specificity (Ran et al., 2018). These results demonstrate potential of clinical decision support software compared to human clinician accuracy tests, particularly given the practicality of the ability to recognize many specific referred diseases.

4.4. Cataract

A cataract is the excessive build-up of protein in the natural lens of the eye that leads to the opacification of the eye lens. Since this is an age-related phenomenon, it is common in older adults. Early detection and treatment at the appropriate stage can help in reducing the blindness caused by this condition. AI platforms using random forests and SVMs can accurately detect cataracts using fundus images. Ran et al. used an elaborate six-level cataract grading random forest based on the feature datasets generated by a DCNN (Long et al., 2017). The results show that random forests reduce the concusion of DCNN on smaller datasets and with an average accuracy of 90.69%. Long et al. classified and graded patients, with pediatric cataract using a CNN-based CAD framework (Almeida & De et al., 2015).

Other than the above-mentioned retinal diseases, AI systems can detect keratoconus (Almeida & De et al., 2015), to formulate plans for horizontal strabismus (Koprowski et al., 2016), to evaluate corneal power after myopic corneal refractive surgery (Xu et al., 2017), and to detect pigment
epithelial detachment in polypoidal choroidal vasculopathy (Hassan et al., 2018). In this review, we outlined studies on DR, AMD, glaucoma and cataracts using various DL techniques in Table 2.

### 5. Discussions

Automatic screening and diagnostics with AI assistance for common eye diseases may ultimately contribute towards maximizing the role of doctors in the clinic. Outside the clinic, AI platforms offer more health opportunities to patients and reduce barriers in eye care where an ophthalmologist is not available. To a certain extent, new AI-based technologies can reduce social inequalities (Khalid et al., 2018). AI-assisted systems will demonstrate the potential to alleviate the problems of the overcharged healthcare system.

In particular, there are primarily three steps in the method of automatically detecting a disease (Sengur et al., 2017; Yau et al., 2012). Firstly, a large amount of image collection is necessary and relative experts must label the characteristic lesions. Secondly, computers extract disease characteristics from the input of labelled pictures via a specific program. Lastly, the statistical characteristic of the target lesions can distinguish a specified picture from other types of disease.

Bourne et al. (Fleming et al., 2006) detailed upon the global prevalence of vision impairment. As of 2015, more than 200 million people worldwide were suffering from moderate to severe vision impairment. The age-standardized prevalence was highest among developing regions of the world like South Asia, western sub-Saharan Africa and North Africa Bengio et al. (2013), Ting Daniel et al. (2011). Visual defects and retinal diseases for people in these regions affect quality of life and economic opportunities.

The retina has a complex morphology and is among the most complex parts in the human body. AI systems are more reliable than a human counterpart in the long run with the accessibility of high-resolution scanners. Many retinal conditions must be critically evaluated because they are highly subjective. Some of the characteristics of diabetic eyes are so minute and rare that specialists will not check until and without reason (Bellemo et al., 2019). However, an AI platform can reliably and without partiality look at all these features. As far as the concerns that doctors might lose their employment owing to automation are concerned, automation will allow more

### Table 1. Classifiers of artificial intelligence presently used in healthcare

| Classifiers         | Description                                                                                                                                                                                                 |
|---------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Bayesian Classifiers| Every pair of features is classified independent of each other                                                                                                                                              |
| Decision Trees      | Uses tree-like graphs to solve classification problems where internal nodes represent a test on an attribute and terminal nodes contain a class label                                                                 |
| k-means clustering  | Divides observations into clusters, with individual observations classified into cluster with nearest mean                                                                                                   |
| k-nearest neighbors | Without making any underlying assumptions about the data, coordinates are classified into groups identified by an attribute                                                                                   |
| Neural Networks     | A circuit of nodes with each connection being modelled as weights. Used for predictive modeling when trained with a dataset and to discover any complex correlations between datasets                                      |
| Random Forests      | Create a set of decision trees from random training sets and decide class of test object from sum of votes from different decision trees                                                                     |
| Support Vector Machines | Outputs optimal hyperplane which categorizes new examples using a supervised-learning algorithm                                                                                              |
| Study group         | Aim                              | Dataset size                      | Method                          | Results                                      | Conclusions                                                  |
|---------------------|----------------------------------|-----------------------------------|---------------------------------|----------------------------------------------|--------------------------------------------------------------|
| Abbas et al.        | DR grading                       | MESSidor, Diaretdb, FAZ 750 fundus images | DNN                             | AUC: 0.924  
Sensitivity: 92.18%  
Specificity: 94.50% | Suitable for early detection and for treatment prediction of diabetes |
| Apreutesei et al.   | Prove relation between           | SF, Spiridon hospital, Iasi       | Feed Forward Neural Network     | Accuracy: 95%  
Confidence Interval: ± 15%                  | Model was successful in predicting the relationship between glaucoma and diabetes. The most accurate results predicted the presence or absence of changes in DR |
|                     | glaucoma and diabetes             | 101 fundus images                 |                                 |                               |                                                             |
| Burlina et al.      | AMD detection                     | AREDS 130,000 fundus images       | DCNN (AlexNet)                  | Accuracy: 88.4% ~ 91.6%  
AUC: 0.94 ~ 0.96 | CNN model produces results similar to specialists |
| Chen et al.         | Glaucoma detection                | Origo, Sces 2326 fundus images    | DCNN                            | AUC: 0.887 for Sces 0.831 for Origo          | Accurate DCNN for glaucoma detection                         |
| Chowdhury et al.    | Detection of abnormalities in     | Diaret, Tele-ophtha, MESSidor, University of Lincoln, HRF database and others 405 fundus images | K-means clustering              | Accuracy: 93.58%  
AUC: 0.893  
Sensitivity: 73.1%  
Specificity: 95.2% | The performance analysis of Random Forest classifier was significantly higher compared to Naive Bayes classifier |
|                     | the retina                        |                                   |                                 |                               |                                                             |
| Farooq et al.       | Automatic Localization of Optic   | Messidor, STARE 100 fundus images | SVM                             | Accuracy: 90%                        | Accurately able to locate optic disc using SVM               |
| Floriano et al.     | AMD detection                     | STARE 51 fundus images            | SVM                             | Accuracy: 92.15%  
Precision: 0.93 | Using feature selection and Leave-One-Out Cross Validation, SVM doesn’t exhibit false positives |
| Galilea et al.      | Identification of Glaucoma Stages | 106 fundus images                 | ANN                             | Accuracy: 100%  
Sensitivity: 100%  
Specificity: 100% | Neuronal networks and use of laser polarimetry increased available parameters to aid diagnosis |
| Gulshan et al.      | DR detection                      | Messidor, EyePACS 128,175 fundus images | DCNN                            | AUC: 0.990 for Messidor 0.991 for EyePACS | The DCNN demonstrated high sensitivity and specificity in detecting DR |
|                     |                                   |                                   |                                 |                               |                                                             |

(Continued)
| Study group          | Aim                                      | Dataset size          | Method     | Results                     | Conclusions                                                                 |
|---------------------|------------------------------------------|-----------------------|------------|-----------------------------|----------------------------------------------------------------------------|
| Hassan et al.⁵³      | CSCR diagnosis                           | 80 OCT scans          | k-NN       | Accuracy: 100%              | Efficiency of the classifier highly depended on the number of nearest neighbors |
| Hnoohom et al.⁵⁴     | DR classification                        | Institute of Medical Research and Technology 100 fundus images | ANN        | Accuracy: 96%               | Local thresholding is applied to better separate the OD and exudate regions |
| Kolaie et al.⁶²      | Retinal blood vessel classification and detection | Review, Drive 105 fundus images | PGM        | Precision: 88.67%           | Probabilistic Graphical Models with a Maximum Likelihood solution based on Laplace approximation found to be more accurate and have higher recall rates than vessel centerline tacking method |
| Khalid et al.⁵⁶      | Drusen segmentation and quantification   | AFIO 6800 OCT scans 100 fundus images | SVM        | Accuracy: 98% Specificity: 97.14% | Accurately distinguish between early, suspected and confirmed AMD stages by correlating OCT scans with fundus images |
| Long et al.²⁸        | Pediatric cataract detection             | 886 Slit-lamp images  | DCNN       | Accuracy: 98.87% for detection 97.56% for treatment suggestion | Platform accurately diagnoses and suggests treatment for cataracts |
| Moura et al.¹⁰       | Automatic Identification of Intraretinal Cystoid Regions | 50 OCT scans         | Linear Discriminant Classifier | Accuracy: 94.61%           | Among LDC, kNN and SVM, the LDC achieved best performance |
| Niwas et al.²⁵       | OCT image analysis for ACG                | NUHS 148 AS-OCT images | NBC        | Accuracy: 89.2%             | Major mechanisms of ACG are classified successfully using parameters from OCT images |
| Ran et al.²⁷         | Cataract detection                       | 5408 fundus images    | DCNN       | Accuracy: 90.69%            | Six-level cataract grading allows specialists to more accurately diagnose than a four-level grading |

(Continued)
| Study group          | Aim                        | Dataset size                           | Method   | Results                      | Conclusions                                                                 |
|---------------------|----------------------------|----------------------------------------|----------|------------------------------|-----------------------------------------------------------------------------|
| Ravindraiah et al.  | Exudates detection         | Diaretdb, American Society of Retina Specialists 25 fundus images | SPCM     | Sensitivity: 89.86% Specificity: 99.44% | Using PCM over traditional FCM proved to be beneficial without the problem of outliers and noise clutter |
| Sengür et al.       | Retinal Vessel detection   | DRIVE 40 fundus images                  | CNN      | Accuracy: 91.78% AUC: 0.96    | Detection task is formulated as a classification task. Proposed method demonstrates second best AUC score among other methods |
| Stevenson et al.    | Classify pathology and clinical features | 4435 fundus images                        | CNN      | Accuracy: 89% AUC: 0.58 Specificity: 75% Specificity: 89% | Proof-of-concept AI platform that could be implemented within a diabetic photo-screening pathway. Statistical performance was limited by the small sample size |
| Takahashi et al.    | DR staging                 | Jichi Medical University 9939 fundus images | DCNN (GoogleNet) | Accuracy: 64% ~ 82%          | The proposed AI platform is able to grade retinal areas typically difficult to visualize on fundoscopy |
| Treder et al.       | AMD detection              | 1112 SD-OCT images                      | DCNN     | Accuracy: 96% Sensitivity: 100% Specificity: 92% | Used DL approach can achieve more practical value with more input data |
| Yadav et al.        | Retinal Image classification | HRF, STARE 110 fundus images           | SVM      | Accuracy: 77.3%              | SVM classifier was found to be more accurate compared to other classifiers like k-NN, decision tree, and linear discriminant |
individuals to take cheaper health care and thus become an original screening test that allows experts to concentrate and become more effective on key instances.

Most automated diagnostic studies on retinal diseases concentrate on one problem at a time. However, this is perhaps not the case in current clinical cases where patients may have several retinal diseases. In order to enhance the application of various AI platforms, multiple conditions with high precision need to be detected (Elze et al., 2015). Use of multiple image methods confirm a given illness rather than just a single data source to further provide a concrete diagnosis. The classifiers are also dependent on the image quality (Abrâmoff et al., 2016). Studies have shown that some algorithms such as DR, AMD, glaucoma, and cataract have been preliminarily created. However, with so many current reports, 100% accuracy and sensitivity are seldom achieved. In other words, not every image can be accurately identified or missed. The accuracy of the results obtained depends not only on computer technology but also on the quality of input pictures (Christopher et al., 2018; Elze et al., 2015). The primary variables leading to bad quality in segmentation of the images includes head and eyeball movement, undilated pupil, frequent blinking, opaque refractive medium and poor fixation. It is the basis of computer education. The annotators must therefore be trained to achieve a uniform standard Bengio et al. (2013), Ting Daniel et al. (2011).

6. Future challenges
Although the AI-based models are highly accurate in many ophthalmic diseases, the clinical and technological difficulties and the real-time use of these models in clinical practice remain numerous. In research and clinical settings, these challenges could arise at different stage. Many of the research have used comparatively homogenous populations of training sets. Retinal images for AI practice and testing often come under various variables, including field width, field of perspective, picture magnification, image quality and the ethnicity of participants (Garigea & Leng, 2017; Mayro et al., 2019). Diversification of the information set could assist to tackle this issue in terms of ethnicity and imaging hardware. The restricted accessibility of high-level information for both unusual conditions (e.g., ocular tumours) and prevalent diseases that are not regularly imaged in clinical exercise is another challenge for the growth of AI models of ophthalmology (Zheng et al., 2012).

Furthermore, the formation of an algorithm requires a lot of computational expense and preparation. This means that AI can only be useful for highly morbid illnesses. It might not be accessible for rare diseases. Secondly, the computer mechanically recognizes a structure or function so AI cannot recognize a disease that was totally separated from our procedure. There will be a small portion of characteristics and variety, which are uncommon (Nguyen et al., 2015). We conclude that AI can select vast majority, not all, people with a disease. Thirdly, this job is somewhat complex. The features of a disorder and the algorithm parameters vary from task to task. Finally, the machine may not construct a model if the connection between input and anticipated output is complicated. Most importantly, it can lead to an error. Nguyen et al. described the process of wrong classification of data by the neural networks. Even though AI can effectively perform a task, it is essential to have a certain level of human intervention during the process [55].

7. Research limitations
Current limitations of accurate diagnosis of retinal diseases are:

a) Quality of the training sets: The labelled set of images tend to have low accuracy if the training set images do not have strong reference standards.

b) Black Box dilemma: Most of the image recognition models use Convolutional Neural Network (CNN) based systems. Wherever a CNN analyses data, it follows some self-generated rules and is difficult to interpret decisions made by the algorithms.
c) Incorrect diagnosis: CNN systems are very sensitive to even minor pixel-level changes in the images leading to inaccurate diagnosis.

d) Image Quality: Current state-of-the-art models are accurate when detecting retinal diseases. However, they cannot recognize when and if an image does not contain retinal diseases. For example, they might confuse a central retinal vein occlusion with DR. Blurry or partial images might hinder the accuracy of model.

8. Conclusions

Artificial intelligence techniques like deep learning and machine learning have dramatically altered the healthcare sector. AI-based platforms have obtained clinically acceptable diagnostic efficiency in the automated diagnosis of many retinal diseases. In order to assess the clinical deployment and cost-effectiveness of various AI systems in clinical practice, future research is crucial. It is important to reuse existing and future methodologies to improve clinical acceptance of AI systems. Although challenges lie ahead, AI-based automated detection platforms are likely to influence the field of medicine and ophthalmology in the coming decades.

Funding
The authors received no direct funding for this research.

Disclosure Statement
The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this article.

Author details
Dasharathra K Shetty1
ORCID ID: http://orcid.org/0000-0002-5021-4029
Abhiroop Talasila2
Swapna Shanbhag3
Vathsala Patil4
ORCID ID: http://orcid.org/0000-0002-8656-8080
B.M Zeeshan Hameed5
Nithesh Naik6
E-mail: nithesh.naik@manipal.edu
ORCID ID: http://orcid.org/0000-0003-0356-7697
Aditya Raju7
1 Department of Humanities and Management, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, Karnataka, India.
2 Department of Computer Science Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, Karnataka, India.
3 Cornea and Ocular Surface, L.V. Prasad Eye Institute, Hyderabad, Telangana, India.
4 Department of Oral Medicine and Radiology, Manipal College of Dental Sciences, Manipal, Manipal Academy of Higher Education, Manipal, Karnataka, India.
5 Department of Urology, Kasturba Medical College, Manipal Academy of Higher Education, Manipal, Karnataka, India.
6 Department of Mechanical and Manufacturing Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, Karnataka, India.
7 Medical Engineering, KTH Royal Institute of Technology, University in Stockholm, Brinellvägen 8, 114 28 Stockholm, Sweden.

Abbreviation

| Abbreviation | Description                  |
|--------------|------------------------------|
| AI           | Artificial Intelligence      |
| MDA          | Macular Degeneration         |
| DR           | Diabetic Retinopathy         |
| ML           | Machine Learning             |
| SVM          | Support Vector Machines      |
| ANN          | Artificial Neural Network    |
| CNN          | Convolutional Neural Network |
| DL           | Deep Learning                |
| CT           | Computed Tomography          |
| OCT          | Optical Coherence Tomography |
| ROC          | Receiver Operating Characteristic |
| AUC          | Area Under ROC               |
| SPCM         | Spatial Possibilistic C-means Clustering |
| FCM          | Fuzzy c-means clustering algorithm |
| AREDS        | Age-Related Eye Disease Study |
| CDR          | Cup to Disk Ratio            |
| ONH          | optic nerve head             |
| IOP          | intraocular pressure        |
| ACG          | Angle-closure glaucoma       |
| NBC          | Neighborhood-Based Clustering |
| anti-VEGF    | Anti-vascular endothelial growth factor therapy |
| NBC          | Naïve Bayes classifier       |
| OD           | Optical Disc                 |
| AS-OCT       | Anterior segment OCT         |
Funding
The authors received no direct funding for this research.

Citation information
Cite this article as: Current state of artificial intelligence applications in ophthalmology and their potential to influence clinical practice, Dasharatrach K Shetty, Abhiroop Talasila, Swapna Shanbhag, Vathsala Patil, B.M. Zeeshan Hameed, Nithesh Naik & Aditya Raju, Cogent Engineering (2021), 8: 1920707.

References
Abramoff, M. D., Lou, Y., Erginay, A., Clarida, W., Amelunxen, R., Folk, J. C., & Niemeijer, M. (2016). Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of Deep Learning. Investigative Ophthalmology & Visual Science, 57(19), 13950. https://doi.org/10.1167/iovs.16-19966

Al-Fadhl, Y. Q., Chung, P. W., Li, B., & Bowman, R. (2017). 3D simulation of navigation problem of people with cerebral visual impairment. Advances in Intelligent Systems and Computing Advances in Computational Intelligence Systems, 265–275. https://doi.org/10.1007/978-3-319-66939-7_23

Almeida, J. D., Silva, A. C., Teixeira, J. A., J. A., Poiva, A. C., & Gattass, M. (2015). Surgical planning for horizontal strabismus using support vector regression. Computers in Biology and Medicine, 63(2015), 178–186. https://doi.org/10.1016/j.compbiomed.2015.05.025

Apreutesei, N. A., Tircoveanu, F., Cantemir, A., Bogdanici, C., Lisa, C., Curteanu, S., & Chiseliţă, D. (2018). Predictions of ocular changes caused by diabetes in glaucoma patients. Computer Methods and Programs in Biomedicine, 154, 183–190. https://doi.org/10.1016/j.cmpb.2018.06.027

Bellomo, V., Yip, M. Y., Xie, Y., Lee, X. Q., Nguyen, Q. D., Hamzah, H., Ting, D. S. (2019). Artificial Intelligence using Deep Learning in classifying side of the eyes and width of field for retinal fundus photographs. Computer Vision – ACCV 2018 Workshops Lecture Notes in Computer Science, 309–315. https://doi.org/10.1007/978-3-030-21074-8_26

Bengio, Y., Courville, A., & Vincent, P. (2013). Representation learning: A review and new perspectives. IEEE Transactions on Pattern Analysis and Machine Intelligence, 35(8), 1798–1828. https://doi.org/10.1109/tpami.2013.50

Bogunovic, H., Waldstein, S. M., Schlegl, T., Langs, G., Sadeghipour, A., Liu, X., Gerendas, B. S., ... & Schmidt-Erfurth, U. (2017). Prediction of anti-VEGF treatment requirements in neovascular AMD using a Machine Learning approach. Investigative Ophthalmology & Visual Science, 58(7), 3240. https://doi.org/10.1167/iovs.16-21053

Burlina, P., Pacheco, K. D., Joshi, N., Freund, D. E., & Bressler, N. M. (2017). Comparing humans and Deep Learning performance for grading AMD: A study in using universal deep features and transfer learning for automated AMD analysis. Computers in Biology and Medicine, 82, 80–86. https://doi.org/10.1016/j.compbiomed.2017.01.018

Chowdhury, A. R., Chatterjee, T., & Banerjee, S. (2018). A random forest classifier-based approach in the detection of abnormalities in the retina. Medical & Biological Engineering & Computing, 57(1), 193–203. https://doi.org/10.1007/s11517-018-1878-0

Christopher, M., Belghith, A., Weinreb, R. N., Bowd, C., Goldbaum, M. H., Saunders, L. J., & Zangwill, L. M. (2018). Retinal nerve fiber layer features identified by unsupervised machine learning on optical coherence tomography scans predict glaucoma progression. Investigative Ophthalmology & Visual Science, 59(7), 2748. https://doi.org/10.1167/iovs.17-23387

Elze, T., Pasqualet, L. R., Shen, L. Q., Chen, T. C., Wiggs, J. L., & Bex, P. J. (2015). Patterns of functional vision loss in glaucoma determined with orthetypal analysis. Journal of the Royal Society Interface, 12(103), 20141118. https://doi.org/10.1098/rsif.2014.1118

Farooq, U., & Sattar, N. Y. (2015). Improved automatic localization of optic disc in retinal fundus using image enhancement techniques and SVM. 2015 IEEE International Conference on Control System, Computing and Engineering (ICCSCE). Batu Ferringhi, Penang, Malaysia. https://doi.org/10.1109/iccsce.2015.7482242

Fleming, A. D., Philip, S., Goatman, K. A., Olson, J. A., & Sharan, P. F. (2016). Automated assessment of diabetic retinal image quality based on clarity and field definition. Investigative Ophthalmology & Visual Science, 47(3), 1120–1125. https://doi.org/10.1167/iovs.05-1155

Gallego, E. H., Santos-Garcia, G., & Sudrez-Bárcena, I. F. (2020). Identification of glaucoma stage with artificial neural networks using retinal nerve fibre layer analysis and visual field parameters. In Innovations in hybrid intelligent systems (pp. 418–424). Berlin, Springer. https://doi.org/10.1007/978-3-540-74972-1_54

García-Fioriano, A., Ferreiro-Santiago, Á., Camacho-Nieto, O., & Yáñez-Marquez, C. (2019). A machine learning approach to medical image classification: Detecting age-related macular degeneration in fundus images. Computers & Electrical Engineering, 75, 218–229. https://doi.org/10.1016/j.compeleceng.2017.11.008

Gorgaya, R., & Leng, T. (2017). Automated identification of diabetic retinopathy using Deep Learning. Investigative Ophthalmology, 124(7), 962–969. https://doi.org/10.1016/j.ophtha.2017.02.008

Gulshan, V., Peng, L., Coram, M., Stumpe, M. C., Wu, D., Narayanaswamy, A., & Webster, D. R. (2016). Development and validation of a Deep Learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. Jama, 316(22), 2402. https://doi.org/10.1001/jama.2016.17216

Hojian-Tiloki, K. (2013). Receiver Operating Characteristic (ROC) curve analysis for medical diagnostic test evaluation. Caspian Journal of Internal Medicine, 4(2), 627–635. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3755824/

Hassan, B., Ahmed, R., & Li, B. (2018). “Computer aided diagnosis of idiopathic central serous chorioretinopathy.” In 2018 2nd IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC) of 2018, (pp. 824–828). Xi’an, China: IEEE. https://doi.org/10.1109/imcec.2018.8469292.

Heidary, F., & Gharebaghi, R. (2012). Ideas to assist the underprivileged dispossessed individuals. Medical Hypothesis, Discovery and Innovation in Ophthalmology, 3(3), 43–44. https://iranjournals.nalid.ir/handle/123456789/71694

Hnoohom, N., & Tanthuwapathom, R. (2017). Classification of diabetic retinopathy stages using image segmentation and an artificial neural network. Lecture Notes in Computer Science Trends in Artificial Intelligence: PRICAI, 2016(Workshops), 51–62. https://doi.org/10.1007/978-3-319-60675-0_5

Hogarty, D. T., Mackey, D. A., & Hewitt, A. W. (2018). Current state and future prospects of artificial intelligence in
ophthalmology: A review. Clinical & Experimental Ophthalmology, 47(1), 128–139. https://doi.org/10.1111/ceo.13381

Khalid, S., Akram, M. U., Hassan, T., et al. (2018). Automated segmentation and quantification of drusen in fundus and optical coherence tomography images for detection of ARMD. Journal of Digital Imaging, 31, 464–476. https://doi.org/10.1007/s10278-017-0038-7

Koprowski, R., Lanza, M., & Irregerole, C. (2016). Corneal power evaluation after myopic corneal refractive surgery using artificial neural networks. Biomedical Engineering OnLine, 15(1), 121. https://doi.org/10.1186/s12938-016-0243-5.

Krishnapuram, R., & Keller, J. (1993). A possibilistic approach to clustering. IEEE Transactions on Fuzzy Systems, 1(2), 98–110. https://doi.org/10.1109/91.227387

Lee, J., Jun, S., Cho, Y., Lee, H., Kim, G. B., Seo, J. B., & Kim, N. (2017). Deep Learning in medical imaging: General overview. Korean Journal of Radiology, 18(4), 570. https://doi.org/10.3348/kjr.2017.18.4.570

Li, Z., He, Y., Keel, S., et al. (2018). Efficacy of a Deep Learning system for detecting glaucomatous optic neuropathy based on color fundus photographs. Ophthalmology, 125(8), 1199–1206. https://doi.org/10.1016/j.ophtha.2018.01.023

Long, E., Lin, H., Liu, Z., Wu, X., Wang, L., Jiang, J., ... & Liu, Y. (2017). An artificial intelligence platform for the multihospital collaborative management of congenital cataracts. Nature Biomedical Engineering, 1(2), 1–8. https://doi.org/10.1038/s41551-016-0024.

Mayo, E. L., Wang, M., Elze, T., & Pasquale, L. R. (2019). The impact of artificial intelligence in the diagnosis and management of glaucoma. Eye, 34(1), 1–11. https://doi.org/10.1038/s41433-019-0577-x

Moura, J. D., Novo, J., Rouco, J., Penedo, M. G., & Ortega, M. (2017). Automatic identification of intraretinal cystoid regions in optical coherence tomography. Artificial Intelligence in Medicine Lecture Notes in Computer Science, 1035–1045. https://doi.org/10.1007/978-3-319-59758-4_35

Nguyen, A., Yosinski, J., & Clune, J. (2015). Deep neural networks are easily fooled: High confidence predictions for unrecognizable images. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, MA, USA. https://doi.org/10.1109/cvpr.2015.7298640

Niwas, S. I., Lin, W., Bai, X., Kwoh, C. K., Jay Kuo, C. C., Sng, C. C., Aquino, M. C., ... & Chew, P. T. (2016). Automated anterior segment OCT image analysis for Angle Closure Glaucoma mechanisms classification. Computer methods and programs in biomedicine, 130, 65–75. https://doi.org/10.1016/j.cmpb.2016.03.018

Paulus, J., Meier, J., Bock, R., Hornegger, J., & Michelsson, G. (2010). Automated quality assessment of retinal fundus photos. International Journal of Computer Assisted Radiology and Surgery, 5(6), 557–566. https://doi.org/10.1007/s11548-010-0479-7

Pras, P., Radeck, V., Mayer, C., Cvetkov, Y., Cvetkova, N., Helbig, H., & Märker, D. (2017). OCT-based Deep Learning algorithm for the evaluation of treatment indication with anti-vascular endothelial growth factor medications. Graefe’s Archive for Clinical and Experimental Ophthalmology = Albrecht von Graefes Arch f klinische und experimentelle Ophthalmologie, 256(1), 91–98. https://doi.org/10.1007/s00417-017-3839-y

Ran, J., et al. “Cataract detection and grading based on combination of deep convolutional neural network and random forests.” In 2018 International Conference on Network Infrastructure and Digital Content (IC-NIDC) (pp. 155–159). Guiyang,China: IEEE. https://doi.org/10.1109/nidc.2018.8525852

RaoMohanreddy, R., & Chandra Mohan Reddy, S. (2018). Exudates detection in diabetic retinopathy images using possibilistic C-means clustering algorithm with induced spatial constraint. Advances in Intelligent Systems and Computing Artificial Intelligence and Evolutionary Computations in Engineering Systems, 455–463. https://doi.org/10.1007/978-981-10-7868-2_44

Schmidt-Erfurth, U., Waldstein, S. M., Klümper, S., Sadeghipour, A., Hu, X., Gerendas, B. S., Osborne, A., & Bogunovic, H. (2018). Prediction of individual disease conversion in early AMD using artificial intelligence. Investigative Ophthalmology & Visual Science, 59(8), 3199–3208. Malatya, Turkey, https://doi.org/10.1167/iovs.18-24106

Sengur, A., Gun, Y., Budak, A., & Vespa, L. J. “A retinal vessel detection approach using convolution neural network.” In 2017 International Artificial Intelligence and Data Processing Symposium (IDAP) (pp. 1-4). Malatya: IEEE. https://doi.org/10.1109/idap.2017.8090331

Sievik, U., Köse, C., Berber, T., & Erdö, H. (2014). Identification of suitable fundus images using automated quality assessment methods. Journal of Biomedical Optics, 19(4), 046006. https://doi.org/10.1117/1.jbo.19.4.046006

Tekahashi, H., Tampo, H., Arai, Y., Inoue, Y., & Kawashima, H. (2017). Applying Artificial Intelligence to disease staging: Deep Learning for improved staging of diabetic retinopathy. Plos One, 12(6), e0179790. https://doi.org/10.1371/journal.pone.0179790

Ting Daniel, S. W., Tay-Kearney, M. L., Constable, I., Lim, L., Preen, D. B., & Kanagasigam, Y. (2011). Retinal video recording a new way to image and diagnose diabetic retinopathy. Ophthalmology, 118(8), 1588–1593. https://doi.org/10.1016/j.ophtha.2011.04.009

Treder, M., Lauermann, J. L., & Eter, N. (2018). Automated detection of exudative age-related macular degeneration in spectral domain optical coherence tomography using deep learning. Graefe’s archive for clinical and experimental ophthalmology = Albrecht von Graefes Arch f klinische und experimentelle Ophthalmologie, 256(2), 259–265. https://doi.org/10.1007/s00417-017-3850-3

Wang, M., Shen, L. Q., Pasquale, L. R., Petrokas, P., Formico, S., Boland, M. V., Wellik, S. R., ... & Elze, T. (2019). An Artificial Intelligence approach to detect visual field progression in glaucoma based on spatial pattern analysis. Investigative Ophthalmology & Visual Science, 60(1), 365–375. https://doi.org/10.1167/iovs.18-25568

Xu, Y., Yan, K., Kim, J., Wang, X., Li, C., Su, L., Yu, S., ... & Feng, D. D. (2017). Dual-stage Deep Learning framework for pigment epithelium detachment segmentation in polypoidal choroidal vasculopathy. Biomedical Optics Express, 8(9), 4061–4076. https://doi.org/10.1364/boe.8.004061

You, J., Rogers, S., Kawai, R., Lamoureux, E., Kowalski, J., Bek, T., et al. (2011). Global prevalence and major risk factors of diabetic retinopathy. Diabetes Care, 34(6), 556–564. https://doi.org/10.2337/dc11-1909

Zheng, Y., Hijazi, M. H., & Coenen, F. (2012). Automated “disease/no disease” grading of age-related macular degeneration by an image mining approach. Investigative Ophthalmology & Visual Science, 53(13), 8310. https://doi.org/10.1167/iovs.12-9576
