



Developments in the detection of diabetic retinopathy: a state-of-the-art review of computer-aided diagnosis and machine learning methods

Ganeshsree Selvachandran¹ · Shio Gai Quek¹ · Raveendran Paramesran² · Weiping Ding³ · Le Hoang Son⁴ 

Accepted: 4 April 2022

© The Author(s), under exclusive licence to Springer Nature B.V. 2022

Abstract

The exponential increase in the number of diabetics around the world has led to an equally large increase in the number of diabetic retinopathy (DR) cases which is one of the major complications caused by diabetes. Left unattended, DR worsens the vision and would lead to partial or complete blindness. As the number of diabetics continue to increase exponentially in the coming years, the number of qualified ophthalmologists need to increase in tandem in order to meet the demand for screening of the growing number of diabetic patients. This makes it pertinent to develop ways to automate the detection process of DR. A computer aided diagnosis system has the potential to significantly reduce the burden currently placed on the ophthalmologists. Hence, this review paper is presented with the aim of summarizing, classifying, and analyzing all the recent development on automated DR detection using fundus images from 2015 up to this date. Such work offers an unprecedentedly thorough review of all the recent works on DR, which will potentially increase the understanding of all the recent studies on automated DR detection, particularly on those that deploys machine learning algorithms. Firstly, in this paper, a comprehensive state-of-the-art review of the methods that have been introduced in the detection of DR is presented, with a focus on machine learning models such as convolutional neural networks (CNN) and artificial neural networks (ANN) and various hybrid models. Each AI will then be classified according to its type (e.g. CNN, ANN, SVM), its specific task(s) in performing DR detection. In particular, the models that deploy CNN will be further analyzed and classified according to some important properties of the respective CNN architectures of each model. A total of 150 research articles related to the aforementioned areas that were published in the recent 5 years have been utilized in this review to provide a comprehensive overview of the latest developments in the detection of DR.

Keywords Diabetic retinopathy · Retina · Retinopathy · Computer-aided diagnosis · Machine learning · Fuzzy logic · Image processing · Fuzzy sets

✉ Weiping Ding
dwp9988@163.com

✉ Le Hoang Son
sonlh@vnu.edu.vn

Extended author information available on the last page of the article

1 Introduction

Diabetes mellitus is becoming an ever-common disease with Type II diabetes reaching epidemic levels and Type I diabetes on a steady increase all around the world. With diabetes comes various health complications and life-limiting diseases including but not limited to macrovascular diseases such as ischemic heart disease, peripheral artery diseases and stroke. Additionally, it may also result in various microvascular diseases such as diabetic neuropathy, retinopathy and nephropathy. Diabetic retinopathy (DR) is a health complication that occurs as a result of prolonged diabetes with poor blood sugar management. As the name suggests, DR is a health concern pertaining to the eyes and ultimately results in vision loss. Diabetes affects insulin production and sensitivity, and therefore the ability of the body to absorb glucose, leading to high blood sugar levels. Left unmanaged, high blood sugar levels may result in blood vessel damage to the retinal vasculature of the eye. Blockage of the capillaries responsible for transporting blood to and from the eye results in decreased blood flow to the retina. The continuous need for oxygen sources naturally stimulates angiogenesis to occur, generating new blood vessels to carry out blood transport which the ruptured vessels have failed to do. This, in theory, should resolve the oxygen deprivation the eye has previously faced but instead leads to further complications as for some reason, instead of reparative angiogenesis, retinal neovascularization occurs. These complicates matters as the new outburst of unexplained vessel growth compromises vision and might lead to blindness in diabetic patients.

DR affects up to 80% of people who have had diabetes for longer than 20 years and is responsible for 12% of all new cases of blindness. Globally, DR has always been the leading cause of vision loss and even blindness among those aged between 20 and 74 years old (Vujosevic et al. 2020), and it affects an estimated 285 million people worldwide. As the number of diabetic patients continue to increase all around the world, we are thus looking at the possibility that the DR arising from diabetes may prove to be a growing burden to ophthalmologists in terms of the amount of time and effort needed to treat these growing numbers. Automation of the most tedious processes may be the solution towards alleviating some of this taxing burden caused by this disease on the economy and healthcare system.

A large part of our understanding revolves around the cause of DR to tie to systemic risk factors. Studies have shown that good glycemic management substantially hinders the development of DR. On the other hand, the accumulation of advanced glycaemic end product (AGE)—a by-product of poor glycemic management—is known to correlate with retinopathy risk (Genuth et al. 2015). Dyslipidemia, when lipid amounts in blood reaches abnormal levels and hypertension, also known as high blood pressure, may also influence DR. However, studies suggest that intensive blood pressure control has little effect in hindering the severity progression of DR (The ACCORD Study Group and ACCORD Eye Study Group 2010). The systemic risk factors, however, only contribute collectively to about less than 20% of all DR risks. This suggests additional factors playing vital roles in DR initiation and progression continue to remain undiscovered. A recent discovery shows that diabetes also has an impact on the retinal neovascular unit and its interdependent neuronal, vascular, glial and immune cells (Duh et al. 2017). This opens up many new possibilities to discover and explore in the pursuit of identifying DR initiation factors. Whilst medical experts continue to seek more ways to detect DR through other factors previously gone unnoticed, the development of convolutional neural networks are showing promise in the field of early detection for DR.

DR is generally classified into two stages: non-proliferative DR and proliferative DR. A major criterion in DR evaluation and analysis is the evident development of microvascular lesions such as microaneurysms, intraretinal hemorrhages, exudates, capillary nonperfusions, intraretinal microvascular abnormalities and venous caliber abnormalities accompanied with neuronal infarcts and retinal neovascularization.

Other hallmark features foreboding the development of early DR usually accompanying and/or preceding lesion development are the thickening of the vascular basement membrane, premature death of pericytes and eventual endothelial cell death (Duh et al. 2017; Antonetti et al. 2016). Several retinal vascular pathological processes in DR have a direct impact on vision.

One such example is vessel integrity loss, which implies gradual non perfusion of the retinal vasculature. This progresses into the eventual occlusion or degeneration of capillaries (Stitt et al. 2016) which in turn results in regions of ischemia and impaired oxygenation of retinal neurons. If the condition persists and capillary nonperfusion becomes progressive, hypoxia and proangiogenic growth factors stimulate abnormal formation of blood vessels in the retina. This is an indication that NPDR has worsened into a state of DR—known for pathologic preretinal neovascularization. Vision loss is highly likely at this point. Otherwise, in the NPDR stage, things can also go sideways if retinal neovascularization occurs. Retinal neovascularization leads to vitreous hemorrhage or tractional retinal detachment (Stitt et al. 2016), both bearing high risks of visual impairment.

Although there are only two stages of DR, it is worth mentioning diabetic macular edema at this point. This is an important categorization since DME can occur at any of the two stages of DR and is also the most common cause of vision loss in DR patients. Simply put, edema is when the blood-retinal barrier breaks down and causes vascular leakage of circulated fluids and protein into the neural retina—yet another cause for vision loss.

As of current knowledge, DR is not a curable disease. One characteristic of DR that continues to baffle researchers is the how reparative angiogenesis almost never occurs despite the strong hypoxia stimulus and enhanced proangiogenic growth factor production. A recent study suggests that semaphorins—a protein class that controls axonal growth core guidance—may be the root cause of this phenomenon. Specifically, semaphorins 3A, 3F and 6A (Wei et al. 2015) are implicated to suppress the revascularization response from the ischemic retina, steering the neo-vessels to the vitreous instead. Since this proangiogenic growth factor almost always results in pathologic preretinal neovascularization rather than beneficial, reparative angiogenesis. Extensive research has been done over the decades to identify growth factors such as VEGF (Aiello et al. 1994). Along the lines of this understanding, the implementation of anti-VEGF treatments has been quite effective as a form of treatment for this condition (Rami et al. 2017; Gross et al. 2015). We will mention more on existing treatment methods but for now, knowing that for patients with no DR are mild NPDR, simple screening and glycemic control is often enough; patients experiencing moderate or higher levels of DR will have various treatment options depending on the level of severity ranging from laser scatter treatment to vitrectomy.

The study on the automated detection of DR has been gathering more and more attention worldwide since 2015. Among all search engines, it can be observed that the number of articles produced has been increasing rapidly since 2015. In examining the current research trend on the automated detection of DR, the following syntax was keyed into a research database, such as Web of Science Database:

("Diabetic Retinopathy" OR ("Fundus" AND ("exudate" OR "microaneurysm" OR "hemorrhages" OR "red lesions" OR "dark lesions" OR "cotton wool spots" OR "macular edema" OR "vessel network" OR "blood vessel" OR "optic disk")))

AND ("Convolutional" OR "Automated" OR "machine learning" OR "artificial intelligence" OR "neural network" OR "deep learning" OR "fuzzy logic" OR "random forest" OR "support vector machine")

The period of searching was restricted to 2015 or later, and only published journal articles were the target of searching. The search result on all the published journal articles in <https://app.dimensions.ai/> was subsequently exported to a bibliometric program, yielding the author co-citation network shown in Fig. 1.

Among all the 1575 publications in the list obtained, it was found that Wong T.Y. was co-cited 1262 times, which is the most number of co-citation among the thousands of authors, displayed by the bibliometric program as the largest circle. It is worth noting that Wong T.Y. is also a co-author for seven of the journal articles that were reviewed in this paper namely, references no. (Ting et al. 2017, 2018; Bellemo et al. 2019; Xie et al. 2020; Wong and Bressler 2016; Sabanayagam et al. 2018; Yip et al. 2020).

1.1 The structure of this review article

This review article consists of a total of 5 sections. The tree diagram in Fig. 2 provides an overview for the main contents of this article.

In Chapter 1, a comprehensive overview of DR which includes its complications, detection and treatment methods are presented.

Chapter 2 presents the main content of this article which is a review on the automated detection of DR using fundus images across the recent years is presented.

In Sect. 2.1, thorough general descriptions and analysis will be done on the model which an AI is built upon for the automated detection of DR using fundus images. Substantial graphical presentation is provided, showing how the choices of the model to build the AI has changed across the recent years.

Sections 2.2 and 2.3 follows immediately after Sect. 2.1 to serve as more detailed discussions to the model used for DR detection which forms the most crucial part of this article.

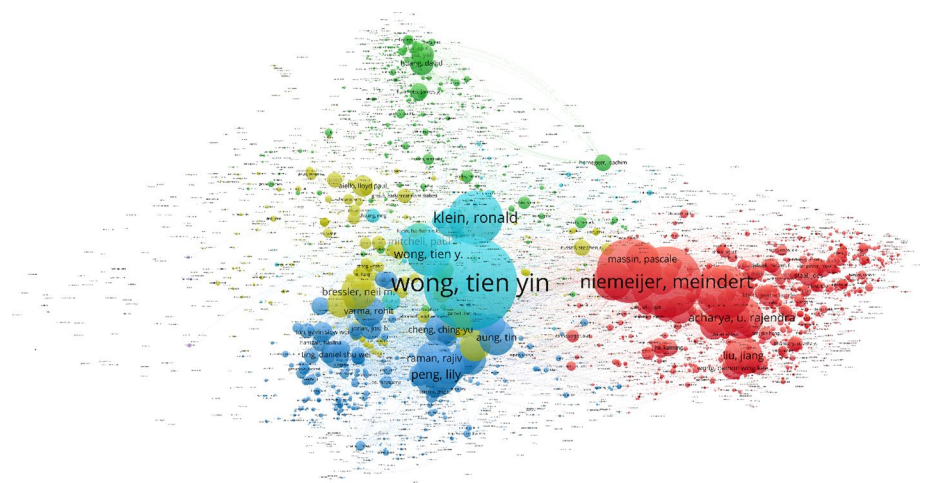


Fig. 1 The author co-citation network showing all the authors who had made publications in the field of automated DR detection since 2015

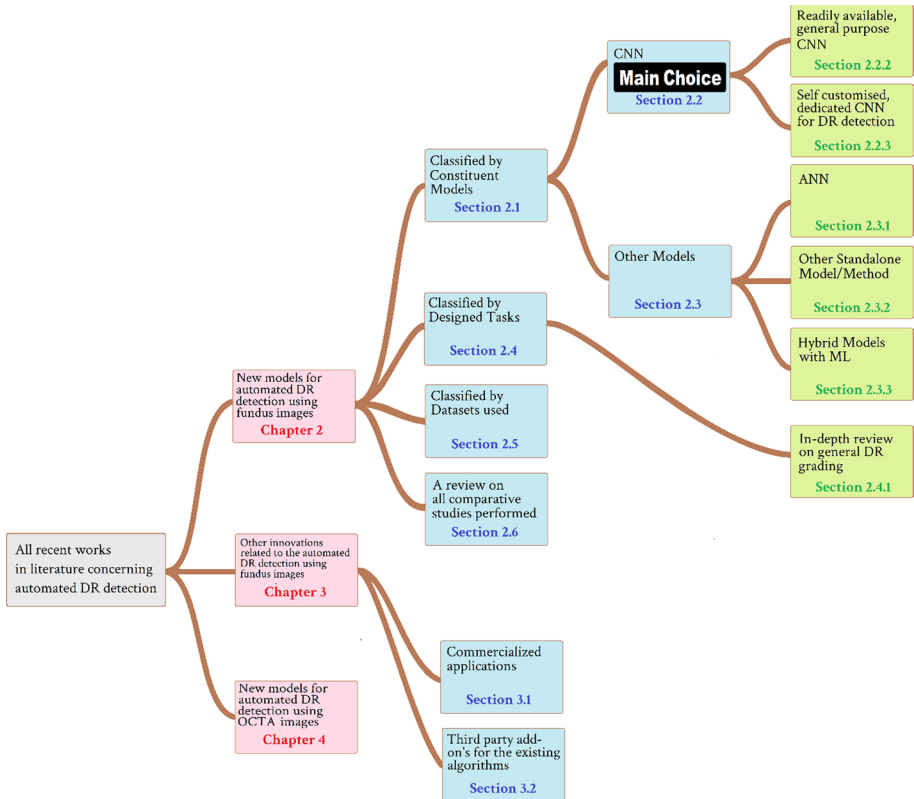


Fig. 2 An overview of the main contents of this article, presented as a tree diagram for visualization

In Sect. 2.2, an even more in-depth study is conducted specifically for CNN-based models, which were found to be the main choice among all the models to build an algorithm. In fact, there are two major groups: Those who deploys readily available general purpose CNN structure, and those who improvise their own CNN structure dedicated for the purpose of the automated detection of DR.

A thorough discussion of the other models used in the detection of DR using fundus photography is presented in Sect. 2.3. Firstly, all the AI which deploys standalone models other than CNN were reviewed. Then, the attention is given to all the AI which deploys hybrid models, for which the type of the hybrid model’s constituents, as well as the trending of choices, are reviewed and presented.

Then, in Sect. 2.4, thorough general descriptions and analysis will be done. Different tasks an AI are built to diagnose, such as the detection of various diseased tissue resulted from complications of DR. Substantial graphical presentation are likewise provided. As the majority of the AI are designed for the general grading of DR according to their levels of severity, Sect. 2.4.1 is provided for a more in-depth analysis among those articles who performed general DR gradings.

The major datasets that have been used in fundus photography in the studies that have been reviewed in this paper are presented in Sect. 2.5. This is then followed by a review on

all the comparative studies performed by all the authors, as well as our own commentaries as which models seems to perform best for the detection of DR based on our observations.

In Sect. 3, other recent developments in the field of automated fundus-based DR detection are presented. First, in Sect. 3.1, a brief review on all the recently trademarked applications for automated DR detection using fundus images is given, in which the authors had also published articles. Then, in Sect. 3.2, some newly developed performance enhancing add-ons are presented. Such add-ons were developed by some dedicated research groups to provide a performance boost when used alongside with the existing ML algorithms for automated DR detection using fundus images.

In Sect. 4, a review on the recent development of DR detection algorithm using the OCTA (Optical Coherence Tomography Angiography) images is presented. OCTA imaging is an emerging new way of retinal imaging that provide unprecedented image details over the conventional fundus imaging reviewed in Sects. 2 and 3.

Concluding remarks are presented in Sect. 5. Moreover, potential future directions of research will also be highlighted in this section, such as a shift of attention from using the conventional fundus images to using the recently developed OCTA image for automated DR detection. This is then followed by acknowledgements and the list of references.

2 Recent development on the automated detection of DR using fundus images

Altogether, more than 150 articles published since 2015 were collected from reputable journals across the World Wide Web, out of which 81 of those research articles [2015: (T̄alu et al. 2015; Mahendran and Dhanasekaran 2015a, b; Ibrahim et al. 2015), 2016: (Banerjee and Kayal 2016; Partovi et al. 2016; Wardoyo et al. 2016; Abr̄amoff et al. 2016; Prentašić and Lončarić 2016; Gulshan et al. 2016; Akyol et al. 2016; Gharaibeh 2016; Gharaibeh and Alshorman 2016), 2017: (Das and Puhan 2017; Dutta et al. 2015; Tan et al. 2017a, b; Xu et al. 2017; Quellec et al. 2017; Raju et al. 2017; Ting et al. 2017; Al-Jarrah and Shatnawi 2017; Abbas et al. 2017; Javidi et al. 2017; Zhang et al. 2017; Barkana et al. 2017), 2018: (Kaur and Mittal 2018; Mansour 2018; Brown et al. 2018; Randive et al. 2018; Ramachandran et al. 2018; Hemanth et al. 2018; Saleh et al. 2018; Seth and Agarwal 2018; Al-Hazaimeh et al. 2018), 2019: (Gao et al. 2018; Gonzalez-Gonzalo et al. 2020; Sahlsten et al. 2019; Liu et al. 2019; Hemanth et al. 2019; Sun 2019; Zhang et al. 2019; Li et al. 2019a, b; Eftekhari et al. 2019; Bellemo et al. 2019; Pires et al. 2019; Qummar et al. 2019; Zeng et al. 2019; Son et al. 2019; Gulshan et al. 2019; Nazir et al. 2019; Nair and Muthuvel 2019; Chowdhury et al. 2019; Pratheeba and Singh 2019; Memari et al. 2019; Jebaseeli et al. 2019a, b), 2020: (Colomer et al. 2020; Mateen et al. 2020; Wu et al. 2020; Shaban et al. 2020; Pao et al. 2020; Torre et al. 2020; Shah et al. 2020; Zago et al. 2020; Qiao et al. 2020; Srivastava and Purwar 2020; Shankar et al. 2020a, b; Samanta et al. 2020; Xie et al. 2020; Ayhan et al. 2020; Ali et al. 2020; Derwin et al. 2020a, b; Jadhav et al. 2020; Usman and Almejalli 2020; Luo et al. 2020; Gayathri et al. 2020; Bhardwaj et al. 2021)] are observed to be research articles in which new AI algorithms for DR detection using fundus images were innovated. In fact, from all the literature published in this area from January 2015 until September 2020, a total of 84 newly developed algorithms was observed from that 81 research articles.

Among most articles encountered, the norm has always been to focus on developing a single algorithm using a single type of model at a time. There were nonetheless 3

noticeable exceptions found during 2015–2020, where each of the 3 exceptions produced 2 innovated algorithms using a distinct type model from one another.

- (i) In Mahendran et al. (2015a): One using probabilistic neural network, thus classified as Other-NN in Sect. 2.1.2; the other using support vector machine, thus classified as SVM in Sect. 2.1.2.
- (ii) In Mahendran et al. (2015b): One using k-means clustering, thus classified as Other ML in Sect. 2.1.2; the other using fuzzy c-means, thus classified as Fuzzy-ML in Sect. 2.1.2.
- (iii) In Saleh et al. (2018): One using random forest, thus classified as RF in Sect. 2.1.2; the other using dominance-based rough set approach, thus classified as Other ML in Sect. 2.1.2.

Moreover, there also exist few articles, each produced multiple algorithms which however remained very close in structure to each other, and those algorithms in an article were all created using one same type of model, such as CNN. For clarity reasons, all algorithms of such kind from one same article using one same type of model will be counted as one in our review.

Hence, there are altogether 84 distinctive entries of newly developed AI algorithms for DR detection based on fundus images found from all the literature published in this area from January 2015 until September 2020. Those 84 distinctive entries shall be collectively referred as the *New Fundus Algorithms* throughout the remaining sections of this review article.

On the other hand, we have also gone through many review articles published previously (Amin et al. 2016; Lu et al. 2018; Ting et al. 2018; Sorour et al. 2018; Raman et al. 2019; Grzybowski et al. 2019; Nagendran et al. 2020; Yip et al. 2020; Chaki et al. 2020; Alyoubi et al. 2020; Stolte and Fang 2020; Kandel and Castelli 2020; Ishtiaq et al. 2020; Islam et al. 2020). We found that all of these review articles did not manage to provide a comprehensive comparison between the methods used by each of the authors for the detection of DR. This is especially prevalent today where most of the recent studies have deployed convolutional neural network (CNN) but with differences in its internal architecture. Moreover, though many of the review articles did provide a table mentioning the choices of datasets across the years, none of them was found to be a reliable overview of the research trend due to the fact that they reviewed only very few articles (often less than 20 articles even for a period of 10 years or more).

In three of the succeeding Sects. 2.1, 2.4 and 2.5, all the New Fundus Algorithms will be studied by examining their classifications under three different approaches:

- (i) According to the choices of their constituent models, which can be one (as in Pure CNN), many (as in a hybrid model), or none (as in a conventional morphological method), in Sect. 2.1.
- (ii) According to the specific tasks they are designed to perform, in Sect. 2.4.
- (iii) According to the datasets they have used to train their ML-based algorithms, in Sect. 2.5.

3 The study of the new fundus algorithms by their overall model groups, choices of constituent models and methods

3.1 Observation on the constituent models among all entries of the New Fundus Algorithms

Among the 84 entries of the New Fundus Algorithms, most of them were observed to be built upon one or several of the following 6 constituent models of AI which are all ML based:

- (i) *CNN* Convolutional neural network.
- (ii) *ANN* Artificial neural network.
- (iii) *Other-NN* Other derivatives of ANN than CNN, such as recurrent neural network (RNN) and long-short-term memory (LSTM).
- (iv) *Fuzzy-ML* Fuzzy logic-based methods, such as fuzzy inference system and/or fuzzy time series among others, and with ML implemented.
- (v) *SVM* Support vector machine.
- (vi) *RF* Random forest.

It is worth noting that all neural network-based models (ANN, CNN, Other-NN) as well as SVM and RF, inherently incorporates ML in their own structures.

As the entire structure of an ANN is itself a component of a CNN structure (the classification section, see Fig. 7), any algorithm which said to combines some CNN structures with some ANN structures will be regarded as a Pure CNN model instead of hybrid model. On the other hand, the structure of RNN and LSTM are entirely different from that of a conventional ANN, so an algorithm which incorporates features of both ANN and Other-NN will be considered as a hybrid of ANN with Other-NN.

With regards to the procedure of repeated iteration during the ML procedure, it was further observed that some entries had even deployed genetic algorithm to facilitate their machine learning procedures.

3.2 Classification of the New Fundus Algorithms into 10 mutually exclusive model groups

All 84 entries of the New Fundus Algorithms are found to be fully classifiable under all the 10 mutually exclusive *model groups* as described below:

- (i) *Conventional* Entries that are built solely upon the conventional methods which does not incorporate any recent AI models, with or without ML. Most of these methods are morphological based.
- (ii) *Pure CNN* Entries that are built solely using CNN, without deploying genetic algorithm.
- (iii) *Pure ANN* Entries that are built solely using ANN, without deploying genetic algorithm. In this review, if any entry is known to be NN-based, but it was not mentioned anywhere in its article about which particular NN model it belongs to, then that entry will be classified under this model group
- (iv) *Pure Other-NN* Entries that are built solely using Other-NN, without deploying genetic algorithm.

- (v) *Pure Fuzzy-ML* Entries that are built solely using Fuzzy-ML, without deploying genetic algorithm.
- (vi) *Pure SVM* Entries that are built solely using SVM, without deploying genetic algorithm.
- (vii) *Pure RF* Entries that are built solely using RF, without deploying genetic algorithm.
- (viii) *Other ML* Entries that are built using ML method(s) other than those among the CNN, ANN, Other-NN, Fuzzy-ML, SVM and RF, without deploying genetic algorithm.
- (ix) *Only Genetic* Entries which deployed genetic algorithm onto an unknown host model, i.e. such host model was not mentioned anywhere in the article.
- (x) *Hybrid ML* Entries which satisfy at least one of the following conditions below:
 - (a) Entries which incorporate 2 or more constituent models among {CNN, ANN, Other-NN, Fuzzy-ML, SVM, RF}.
 - (b) Entries which incorporate 1 constituent model among {CNN, ANN, Other-NN, Fuzzy-ML, SVM, RF} together with some other ML methods.
 - (c) Entries which incorporate 1 constituent model among {CNN, ANN, Other-NN, Fuzzy-ML, SVM, RF} together with genetic algorithm.

Due diligence had been carried out in studying all the articles containing entries of the New Fundus Algorithms. The proportion of all 84 entries among the New Fundus Algorithms that belongs to each of the 10 model groups is thus as shown in Fig. 3.

Now we present Fig. 4 which shows the 10 model groups containing these 84 entries of the New Fundus Algorithms arranged by year of publication. This is then followed by Fig. 5 which gives the other perspective around: The year of publication of these 84 entries arranged by the model groups they belong.

Thus, from Figs. 4 and 5, it is clearly visible that a lot of research work on the New Fundus Algorithms seemed to converge into the sole use of convolutional neural networks (Pure CNN) since 2016 (Abràmoff et al. 2016; Prentašić and Lončarić 2016; Gulshan et al.

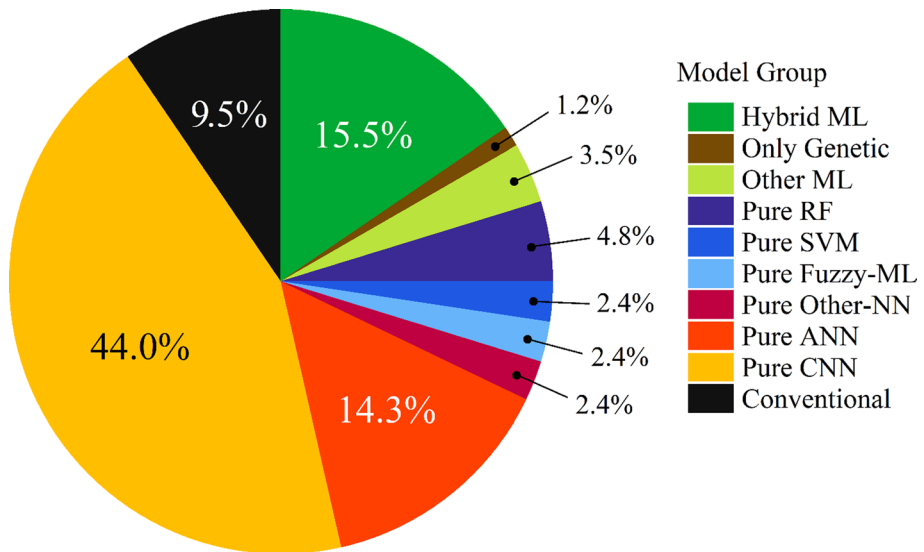


Fig. 3 Proportions among all the 84 entries among the New Fundus Algorithms that belong to each of the 10 model groups

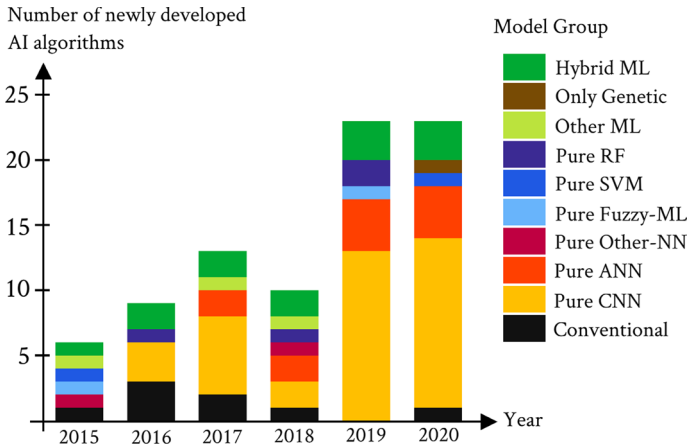


Fig. 4 The model groups of the 84 entries among the New Fundus Algorithms, arranged by their year of publication

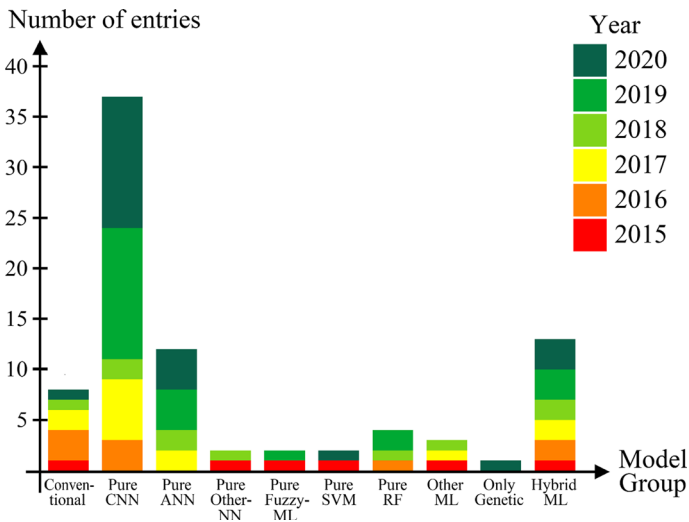


Fig. 5 The year of publications of the 84 entries among the New Fundus Algorithms, arranged by the model groups they belong

2016; Tan et al. 2017a, b; Xu et al. 2017; Quelled et al. 2017; Raju et al. 2017; Ting et al. 2017; Mansour 2018; Brown et al. 2018; Gao et al. 2018; Gonzalez-Gonzalo et al. 2020; Sahlsten et al. 2019; Liu et al. 2019; Hemanth et al. 2019; Sun 2019; Zhang et al. 2019; Li et al. 2019a; Eftekhari et al. 2019; Bellemo et al. 2019; Pires et al. 2019; Qummar et al. 2019; Zeng et al. 2019; Mateen et al. 2020; Wu et al. 2020; Shaban et al. 2020; Pao et al. 2020; Torre et al. 2020; Shah et al. 2020; Zago et al. 2020; Qiao et al. 2020; Srivastava and Purwar 2020; Shankar et al. 2020a; Samanta et al. 2020; Xie et al. 2020; Ayhan et al. 2020). Particularly since 2019, a tremendous increase was observed in the annual number of entries among the New Fundus Algorithms which adopted CNN as their main model

of development (Gao et al. 2018; Gonzalez-Gonzalo et al. 2020; Sahlsten et al. 2019; Liu et al. 2019; Hemanth et al. 2019; Sun 2019; Zhang et al. 2019; Li et al. 2019a; Eftekhari et al. 2019; Bellemo et al. 2019; Pires et al. 2019; Qummar et al. 2019; Zeng et al. 2019; Mateen et al. 2020; Wu et al. 2020; Shaban et al. 2020; Pao et al. 2020; Torre et al. 2020; Shah et al. 2020; Zago et al. 2020; Qiao et al. 2020; Srivastava and Purwar 2020; Shankar et al. 2020a; Samanta et al. 2020; Xie et al. 2020; Ayhan et al. 2020). Meanwhile, the computational power of consumer-based PC had increased tremendously since 2016 for both CPU and GPU. Most of the software needed to run the programming are now free (gratis) and provides plenty of freedom of use, even for commercial use. This grants all the researchers unprecedented access to parallel GPU computing necessary to carry out the resource intensive ML procedures to train the AI built using pure CNN. In the near future, it is therefore very likely to have an even higher number of new pure CNN based algorithms appearing every year.

3.3 An even more detailed classification of the entries in accordance with their constituent models and methods

To provide an even more detailed insight to the model groups used among the New Fundus Algorithms, Table 1 and Fig. 6 are presented to show the general distribution of all the 84 entries according to their constituent models and methods which they are based upon. In particular, the six constituent models are as previously mentioned in Sect. 2.1.1, and they are reviewed together with another 2 methods, namely: Other ML (i.e. those who deployed other ML models as their constituents, but not using genetic algorithm) and Genetic (i.e. those who deployed genetic algorithm).

The areas of all regions in Fig. 6 are shown to be directly proportional to the number of entries that fall within each combination. Moreover, in Fig. 6, all the cardinalities representing the entries under hybrid ML are displayed in green. The remaining 9 black numbers once again correspond to the number of entries under model group number 1 to 9 as defined under Sect. 2.1.2.

It is therefore clearly observable in Fig. 6, that 37 out of the 84 entries belongs to the model group of pure CNN. Section 2.2 is therefore dedicated for an in-depth study of all the pure CNN-based algorithms grouped among the 37 entries (see the beginning of Sect. 2 for more details), revealing all the intricate details and the minute differences among all entries belonging to pure CNN. Entries in other model groups will be discussed under Sect. 2.3.

3.4 Pure CNN: the main model group for the new fundus algorithms: an even more in-depth review

With the abundance of 37 entries of the New Fundus Algorithm that are purely built upon CNN, this section is dedicated to provide an in-depth studies on all the CNN based algorithms grouped among these 37 entries. Firstly in Sect. 2.2.1, the fundamental design of the CNN architecture is reviewed. This is then followed by a thorough study on the recently introduced new general purpose CNN blueprints for the detection of DR in Sect. 2.2.2. Then, in Sect. 2.2.3, several self-customized CNN architectures that are dedicated for DR detection are also discussed.

Table 1 A table showing the distribution of all the 84 entries of the New Fundus Algorithm according to their constituent models and method

Model group	Constituent model/method	Number of entries	Source articles of the entries
Pure CNN	CNN	37	Abràmoff et al. (2016), Prentasić and Lončarić (2016), Gulshan et al. (2016), Tan et al. (2017a, b), Xu et al. (2017), Quellec et al. (2017), Raju et al. (2017), Ting et al. (2017), Mansour (2018), Brown et al. (2018), Gao et al. (2018), Gonzalez-Gonzalo et al. (2020), Sahlsten et al. (2019), Liu et al. (2019), Hemanth et al. (2019), Sun (2019), Zhang et al. (2019), Li et al. (2019a), Eftekhari et al. (2019), Bellemo et al. (2019), Pires et al. (2019), Qummar et al. (2019), Zeng et al. (2019), Mateen et al. (2020), Wu et al. (2020), Shaban et al. (2020), Pao et al. (2020), Torre et al. (2020), Shah et al. (2020), Zago et al. (2020), Qiao et al. (2020), Srivastava and Purwar (2020), Shankar et al. (2020a), Samanta et al. (2020), Xie et al. (2020) and Ayhan et al. (2020)
Pure ANN	ANN	12	Al-Jarrah and Shatmawi (2017), Abbas et al. (2017), Randive et al. (2018), Ramachandran et al. (2018), Son et al. (2019), Gulshan et al. (2019), Nazir et al. (2019), Nair and Muthuvel (2019), Ali et al. (2020), Derwin et al. (2020a), Jadhav et al. (2020), and Shankar et al. (2020b)
Conventional	–	8	Tálu et al. (2015), Banerjee and Koyal (2016), Partovi et al. (2016), Wardoyo et al. (2016), Das and Puhian (2017), Dutta et al. (2015), Kaur and Mittal (2018) and Colomer et al. (2020)
Pure RF	RF	4	Akyol et al. (2016), Saleh et al. (2018), Chowdhury et al. (2019), and Pratheeba and Singh (2019)
Other ML	Other ML	3	Mahendran and Dhanasekaran (2015b), Javidi et al. (2017) and Saleh et al. (2018)
Pure fuzzy-ML	Fuzzy-ML	2	Mahendran and Dhanasekaran (2015b) and Memari et al. (2019)
Pure other-NN	Other-NN	2	Mahendran and Dhanasekaran (2015a) and Hemanth et al. (2018)
Pure SVM	SVM	2	Mahendran and Dhanasekaran (2015a) and Derwin et al. (2020b)
Only genetic	Genetic	1	Usman and Almejalli (2020)
Hybrid ML	ANN, fuzzy-ML	3	Ibrahim et al. (2015), Gharaibeh (2016), and Gharaibeh and Alshorman (2016)
	CNN, SVM	2	Seth and Agarwal (2018) and Li et al. (2019b)
	ANN, SVM	2	Jebaseeli et al. (2019a, b)
	Fuzzy-ML, CNN	1	Luo et al. (2020)
	Genetic, SVM	1	Al-Hazaimeh et al. (2018)
	Other ML, CNN	1	Gayathri et al. (2020)
	Other ML, SVM	1	Bhardwaj et al. (2021)
	Other ML, RF	1	Zhang et al. (2017)
	Fuzzy-ML, ANN, SVM	1	Barkana et al. (2017)

3.4.1 The fundamental CNN design and the needs of preprocessing

Figure 7 below shows the essential structure of a typical CNN architecture for general DR grading.

One of the most obvious methods to further improve the outcome of a CNN based algorithm is to increase the “clarity” of the image right before it is fed into the system. This was usually accomplished by the algorithms by raising the contrast and sharpness of the image with photo editing software or similar programs, to accentuate all the relevant features on the fundus image (Tan et al. 2017a, b; Eftekhari et al. 2019; Zago et al. 2020; Samanta et al. 2020) and to suppress all the unwanted artifacts and noises of the image (Prentašić and Lončarić 2016), before feeding the data into the CNN. This is also to ensure that all the images fed into the CNN possess uniform amount of overall brightness and color (Tan et al. 2017a, b; Eftekhari et al. 2019; Zago et al. 2020; Samanta et al. 2020). Moreover, in a dataset of fundus images where the resolutions vary among the images, resizing was usually done to standardize the image size before fed into the CNN (Sahlsten et al. 2019; Qummar et al. 2019). Such procedure is thus unanimously referred to as preprocessing (Fig. 8).

Among all entries of the New Fundus Algorithm built using pure CNN, we tabulated in Fig. 9 the number of algorithms where significant preprocessing was observed before feeding the input into CNN [2016: (Prentašić and Lončarić 2016; Gulshan et al. 2016), 2017: (Tan et al. 2017a, 2017b; Xu et al. 2017; Raju et al. 2017), 2018: (Mansour 2018), 2019: (Gao et al. 2018; Gonzalez-Gonzalo et al. 2020; Sahlsten et al. 2019; Liu et al. 2019; Hemanth et al. 2019; Zhang et al. 2019; Li et al. 2019a; Eftekhari et al. 2019; Pires et al. 2019; Qummar et al. 2019; Zeng et al. 2019), 2020: (Mateen et al. 2020; Wu et al. 2020; Pao et al. 2020; Zago et al. 2020; Srivastava and Purwar 2020; Samanta et al. 2020)]. Those articles which promotes pure CNN had also given elaboration on the method of preprocessing. This is then compared against those other algorithms who was not found to perform preprocessing [2016: (Abràmoff et al. 2016), 2017: (Quelleg et al. 2017; Ting et al. 2017), 2018: (Brown et al. 2018), 2019: (Sun 2019; Bellemo et al. 2019), 2020: (Shaban et al. 2020; Torre et al. 2020; Shah et al. 2020; Qiao et al. 2020; Shankar et al. 2020a; Xie et al. 2020; Ayhan et al. 2020)], either the articles mention that themselves, or there is nowhere in their article which suggest any preprocessing had been done.

It can be observed that there is a significant drop in the number of entries (it is worth reminding that each entries can either be one single algorithm or several algorithms that are structurally very close to one another, as mentioned in Sect. 1.1) who were observed to perform standalone preprocessing in 2020 (Mateen et al. 2020; Wu et al. 2020; Pao et al. 2020; Zago et al. 2020; Srivastava and Purwar 2020; Samanta et al. 2020), compared to the year 2019 (Gao et al. 2018; Gonzalez-Gonzalo et al. 2020; Sahlsten et al. 2019; Liu et al. 2019; Hemanth et al. 2019; Zhang et al. 2019; Li et al. 2019a; Eftekhari et al. 2019; Pires et al. 2019; Qummar et al. 2019; Zeng et al. 2019). Through the observation on the articles in 2020 (Mateen et al. 2020; Wu et al. 2020; Shaban et al. 2020; Pao et al. 2020; Torre et al. 2020; Shah et al. 2020; Zago et al. 2020; Qiao et al. 2020; Srivastava and Purwar 2020; Shankar et al. 2020a; Samanta et al. 2020; Xie et al. 2020; Ayhan et al. 2020), it is concluded that preprocessing no longer needed due to the newly developed CNN architectures (since 2020 and beyond) which by itself is already capable of tracing even the tiniest difference among the colors or textures, without the needs to be provided with a manually enhanced image.

Another obvious way to further improve the outcome of a CNN based algorithm, is to increase the complexity of the CNN itself. It is worth noting that the most legacy

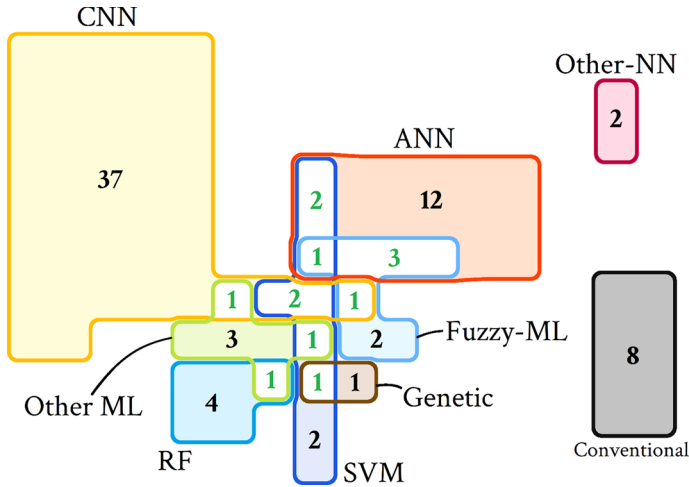


Fig. 6 An Euler diagram showing the distribution of all the 84 entries of the New Fundus Algorithm according to their constituent models and method

architecture of CNN, LeNet (Lecun et al. 1998), had remained the blueprint by most authors for innovating new CNN based algorithms for automated DR detection until 2017. Many groups of authors had thus improvised on LeNet by introducing additional number of layers in the feature extraction or the classification or both, accompanying by adding more neurons and the more feature maps in each layer. One notable example is done by Tan et al. (2017a) which they developed an AI algorithm that detects EX, MA and HM using a CNN enlarged from LeNet.

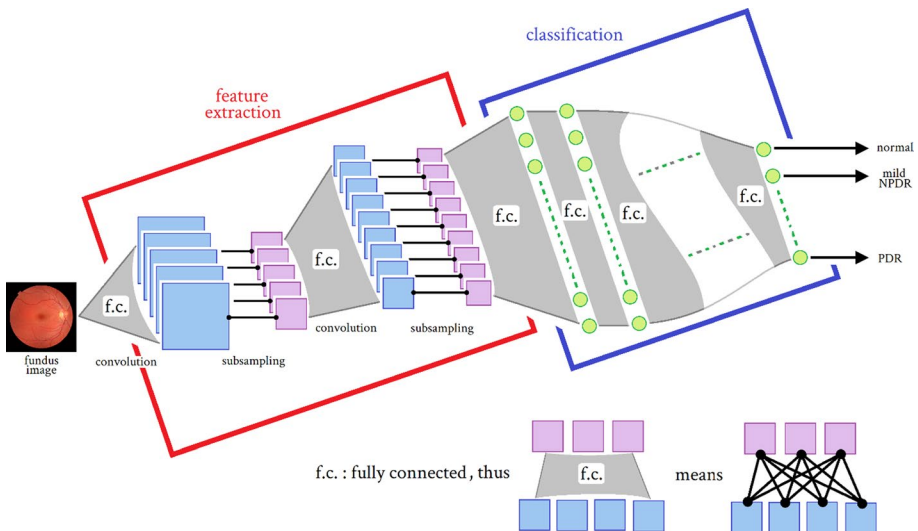


Fig. 7 The essential structure of a typical CNN architecture for general DR grading

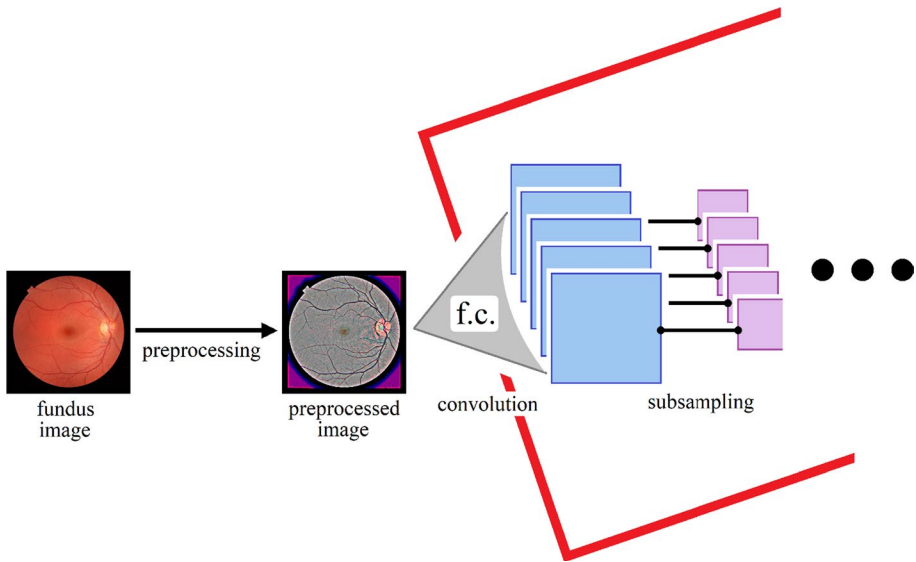


Fig. 8 The implementation of preprocessing of fundus image in prior of feeding into CNN

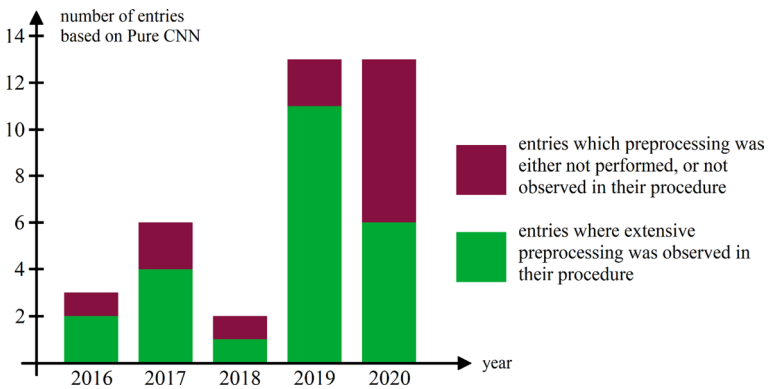


Fig. 9 The implementation of preprocessing among the entries of the New Fundus Algorithm built using pure CNN

3.4.2 The recent general-purpose CNN (RGP-CNN): new blueprints for the mainstream CNN developers for DR detection

It was not until the year 2018 that a large number of more recent, profoundly complex architectures of CNN began to get attention among AI developers for DR detection, such as AlexNet (Krizhevsky et al. 2012), VGGNet (Simonyan and Zisserman 2015), GoogleNet/Inception (Szegedy et al. 2015), ResNet (He et al. 2015), DenseNet (Huang et al. 2017) and SeNet (Hu et al. 2018). Such newly developed architectures were however general-purpose in nature and thus their structures were not specifically tuned for any practical purpose, let alone for DR detection. For the remaining part of this article, those architectures shall

be collectively referred to as *Recent General-Purpose CNN* (abbr. RGP-CNN). It is worth mentioning that an RGP-CNN tends to be extremely complex in structure, except some very early versions like AlexNet. For example, InceptionV3 consists of more than 100 crisscrossing layers and 7 different operations including convolution itself (<https://software.intel.com/content/www/us/en/develop/articles/inception-v3-deep-convolutional-architecture-for-classifying-acute-myeloidlymphoblastic.html>). Hence, throughout this review, any RGP-CNN deployed in a given AI algorithm will be represented by a single icon as shown in Fig. 10.

It is worth reminding that an RGP-CNN were not tuned for any dedicated practical applications and thus it can still be possibly outperformed by a smaller but dedicated CNN architecture customized specially for DR detection.

Following the commonly accepted notation in literature, if a number is mention after the name of an RGP-CNN, that number specifies the number of layers presented in that particular subclass of RGP-CNN. Thus, VGGNet19 specifies that it is a particular subclass of VGGNet that is 19 layers deep. Some author groups in literature however did not mention such number of layers in a particular subclass of an RGP-CNN they used in their works, in such a case, only the name of that RGP-CNN will be presented in this review.

In 2018, Mansour et al. (2018) deployed AlexNet, which is just newer than LeNet, for General DR Grading. Also in the same year, Wan et. al. (2018) were observed to be among the pioneers to perform a comparative study on the performance of six RGP-CNN when adopted for DR detection. The architecture studied by Wan et. al. (2018) were: AlexNet, 3 derivatives of VGGNet (VGGNet-s, VGGNet16, VGGNet19), GoogleNet(a.k.a. Inception), and ResNet. Then, following the footsteps of such comparative studies, most algorithms from the year 2018 had begun to discard the conventional LeNet and began to adopt one or several RGP-CNN as their blueprints to develop their own CNN based algorithm for DR detection. Among the RGP-CNN commonly studied by those inventing authors were ResNet (Liu et al. 2019; Zhang et al. 2019; Mateen et al. 2020; Shankar et al. 2020a; Wan et al. 2018), GoogleNet (Mateen et al. 2020; Shankar et al. 2020a; Wan et al. 2018), and VGGNet (Mateen et al. 2020; Shankar et al. 2020a; Wan et al. 2018).

With such RGP-CNN as the new blueprints, some CNN based algorithms innovated since 2019 (Gao et al. 2018; Gonzalez-Gonzalo et al. 2020; Zhang et al. 2019; Bellemo et al. 2019; Qummar et al. 2019; Zeng et al. 2019; Mateen et al. 2020; Shankar et al. 2020a; Xie et al. 2020) contained multiple RGP-CNN architectures (not to be confused with a hybrid AI model, because it is still a CNN based model).

Among those algorithms with multiple RGP-CNN architectures, some of them deployed mutually distinct RGP-CNN architectures (Zhang et al. 2019; Bellemo et al. 2019; Qummar et al. 2019; Mateen et al. 2020; Xie et al. 2020), all working concurrently on an identical preprocessed data of a fundus image. The final output was then computed from the outcomes yielded by all the constituent RGP-CNN using certain formula or even another ANN-like network. There were six such algorithms observed among the literature published since 2019, all of which are presented in Table 2.

On the other hand, those algorithms with multiple RGP-CNN architectures deployed mutually identical RGP-CNN (Gao et al. 2018; Gonzalez-Gonzalo et al. 2020; Zeng et al. 2019; Shankar et al. 2020a), which all working concurrently on distinctively preprocessed data of a fundus image. Likewise, the final output was then computed from the outcomes yielded by all the constituent RGP-CNN using certain formula or even another ANN-like network. It is worth noting that two mutually identical RGP-CNN used in a given innovated algorithm will still have their network's weights combination different with each

other because each of them will be trained for a different purpose, say examining the left eye and the right eye as encountered in Zeng et al. (2019).

The algorithm developed by Zheng et al. (2019) for General DR Detection takes both eyes as inputs, analyses, and even compares them simultaneously using InceptionV3 for both eyes. Hence the name “Binocular Siamese-like” called by the authors themselves (Fig. 11).

Later in 2020, Shankar et al. (2020a) proposed another algorithm for General DR Grading which deploys two concurrent ResNet50, said to be connected consecutively to an ANN (“synergic”) (Fig. 12).

Besides, there were another two invented algorithms for DR that are rather complex than the rest of the hybrid CNN architectures aforementioned: Firstly, Gonzalez-Gonzalo et al., in their recent article published in JAMA in 2019 (Gonzalez-Gonzalo et al. 2020), incorporates six DenseNet joined concurrently in their algorithm for General DR Grading, each of them took on one of the three RGB channels for each of the two distinctively preprocessed images, which were done using different methods from the raw fundus image. The final result is simply the mean (“average DR score”) of the six outputs from each of the six DenseNet architectures. Secondly, Gao et al. (2018) proposed an algorithm for General DR Detection which consists of four concurrent InceptionV3, each of them work on a quarterly section of the image and produces a tiny monochrome image instead a mere numerical output like from 0 to 4. The final answer was concatenated from that 4 monochrome images produced by that four InceptionV3. That yielded a final monochrome image which, in addition to General DR Grading, also capable of giving the approximate (because that final image is still very small) location of the diseased tissues.

3.4.3 Self-customized CNN architecture dedicated for DR detection

Nevertheless, there were some authors who chose to innovate their own customized architecture of CNN dedicated for DR detection without relying to any RGP-CNN in literature as their blueprints. These self-customized CNN architectures tend to be just as complex as RGP-CNN developed on the same period. Moreover, there is a much greater degree of diversity among the algorithms which uses customized CNN architectures than those who simply deploy RGP-CNN’s. This is because each of the internal constituents can potentially be fully customized by the authors, unlike using a package of RGP-CNN where each of the internal components can hardly be changed (or even impossible, as in the case of cloud computing). As a result, the internal constituents for each of these individually customized CNN architectures aforementioned (Sun 2019; Li et al. 2019a; Eftekhari et al. 2019; Pires et al. 2019; Wu et al. 2020; Pao et al. 2020) will likewise not mentioned in this review.

Among all of the groups of authors observed since 2018, Liu et al. (2019) was among the few who proves that an algorithm built using self-customized CNN dedicated for a



Fig. 10 A representation of the direct adaptation of a single RGP-CNN for general DR grading

certain task of DR screening, can potentially outperform one that is built using an RGP-CNN of much greater complexity. Liu et al. (2019) then further compared their invented algorithm against 5 RGP-CNN's (ResNet50, ResNet101, DenseNet121, SeNet50, SeNet101), and using only 32 layers, it outperformed all of them, achieving an accuracy of 93.80%. The algorithm built by Liu et al. (2019) was for General DR Detection.

Eftekhari et al. (2019) constructed two mutually distinct customized CNN architectures joined in series for EX detection. In contrast, Wu et al. (2020) constructed two mutually distinct customized CNN architectures joined concurrently to perform General DR Grading, involving two distinctively preprocessed data from a fundus image. One of the CNN performs General DR Detection (Normal/DR), the other performs classification on the severity of DR in 4 stages (Mild NPDR/Moderate MPDR/Severe NPDR/PDR). On the other hand, Pao et al. (2020) constructed two mutually identical customized CNN architecture which are joined concurrently to perform General DR Grading using, also involving distinctively preprocessed data of a fundus image. Both of the CNN however are capable to perform General DR Grading on their own. Instead, one of the CNN worked on the preprocessed red channel of the fundus image, the other worked on the preprocessed blue channel of the fundus image. The blue channel of the fundus image were not used in their CNN.

Pires et al. (2019) constructed three customized CNN architecture which are all joined concurrently in their models, each of them is in charge of a preprocessed input image of a different resolution. The customized CNN architecture constructed by Sun (2019), also for General DR Detection, has the all data normalized between each layer. Moreover, the algorithm developed by Li et al. (2019a) use a U-Net structure (Cao et al. 2018) for the feature extraction segment.

3.5 A review on the other minority models groups used by the entries of the new fundus algorithms

In this section, some of the other machine learning models that have been used for the detection of DR using conventional fundus photography are reviewed.

3.5.1 ANN: a tabular representation

Through observation, there was also a slight but observable increase in the annual number of articles which developed an ANN based model. Nevertheless, the number of new ANN algorithms has always been much fewer than that of CNN, this is because CNN by its fundamental structure and purpose was inherently created for analyzing graphical images (Lecun et al. 1998), thus they are much more suited to analyze fundus images than ANN which was built to analyze verbal or numerical data.

Altogether, 12 recent articles utilizing ANN models were observed during the period of 2016–2020. Their particulars are summarized in Table 3 below.

It is worth noting that most algorithms utilizing ANN are designed to perform general classification of DR's presence or severity. This is mainly because the general classification of DR does not require a large number of pixels to be used as inputs from the raw data. In such cases, the severity will mainly depend on how many inputs are with colors unusually differs from that of a retina.

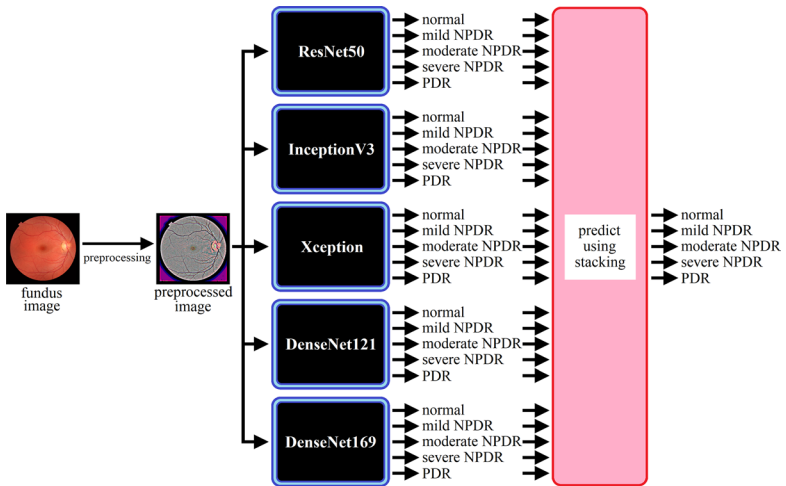
Table 2 Hybrid CNN architecture constructed using several mutually distinct RGP-CNN

Article	Overall structure of the innovated algorithm using mutually distinct RGP-CNN's (Aim)
Zhang et al. (2019)	<p>InceptionV3 + Xception + InceptionResNetV2 (general DR detection)</p>
Zhang et al. (2019)	<p>ResNet50 + DenseNet169 + DenseNet201 (general DR grading)</p>
Bellemo et al. (2019)	<p>VGGNet + ResNet (general DR grading)</p>

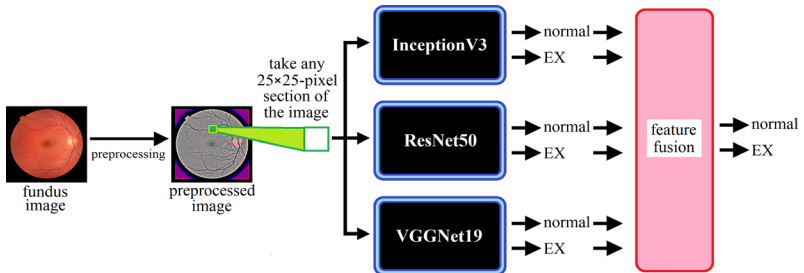
Table 2 (continued)

Article Overall structure of the innovated algorithm using mutually distinct RGP-CNN's (Aim)

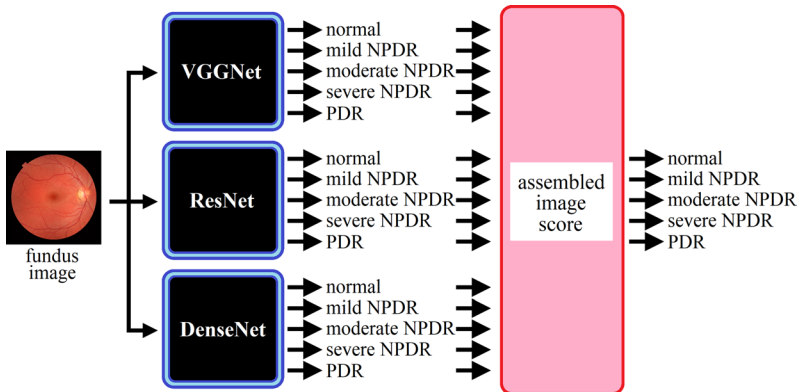
Qummar et al. (2019) ResNet50 + InceptionV3 + Xception + DenseNet121 + DenseNet169 (general DR grading)



Mateen et al. (2020) VGGNet19 + ResNet50 + InceptionV3 (EX segmentation)



Xie et al. (2020) VGGNet + ResNet + DenseNet (general DR grading)



3.5.2 Other standalone algorithms: a tabular representation

In this section we will discuss the author groups from the year 2015 who were found using models other than CNN or ANN to develop their algorithm for DR detection. Table 4 shows the list of authors who used a standalone, ML algorithm which was neither ANN nor CNN.

As shown in Table 4, FCM, RF and SVM were among the ML models right after CNN and ANN in terms of the frequency of appearance in literature.

On the other hand, in 2020, Usman and Almejalli (2020) was the first author who was observed innovating an algorithm using a standalone Genetic Algorithm. In that article, they proposed using genetic algorithms (GA) to do General DR Grading.

There were 8 entries from the New Fundus Algorithms, among seven articles (Țălu et al. 2015; Banerjee and Kayal 2016; Partovi et al. 2016; Wardoyo et al. 2016; Das and Puhan 2017; Dutta et al. 2015; Kaur and Mittal 2018), who used conventional methods till 2019. During the year 2015 to 2019, it is evident that the percentage of algorithms which deployed conventional methods generally were on decline, reaching 0 counts by the year 2019. However, in the year 2020, Colomer et al. (2020) was the only observed group of authors who revived the conventional morphological method again, this time by inventing a morphological algorithm that can be combined with some varieties of other non-NN ML algorithms, such as SVM or genetic algorithm. This was said to produce hybrids that surpass each of the standalone constituents.

3.5.3 Hybrid models with ML

In every year from 2015 to 2020, there were around 2 full-length journal articles per year that were found deploying hybrid models to develop their algorithm. In total, 13 articles were found altogether (Ibrahim et al. 2015; Gharaibeh 2016; Gharaibeh and Alshorman 2016; Zhang et al. 2017; Barkana et al. 2017; Seth and Agarwal 2018; Al-Hazaimeh et al. 2018; Li et al. 2019b; Jebaseeli et al. 2019a, b; Luo et al. 2020; Gayathri et al. 2020; Bhardwaj et al. 2021) during this period. The constituents of their algorithms are as shown in Table 5.

During 2015–2019, as shown in Table 5, 8 of the entries (Ibrahim et al. 2015; Gharaibeh 2016; Gharaibeh and Alshorman 2016; Barkana et al. 2017; Seth and Agarwal 2018; Li et al. 2019b; Jebaseeli et al. 2019a, b) used one NN-based model as one of the constituents, and some non-NN-based models as the other constituents. This was to compensate the various shortcomings of NN-based models, such as the sheer amount of time and computation needed on the ML process. There was indeed a pattern observed, among the 8 entries, on the choices of both their NN-based constituent and their non-NN-based constituent as time goes by, as shown in Table 5, where authors had begun to use CNN in place of ANN as the NN-based constituent (Seth and Agarwal 2018; Li et al. 2019b), and to use SVM in place of Fuzzy-ML as the non-NN-based constituent (Seth and Agarwal 2018; Li et al. 2019b; Jebaseeli et al. 2019a, b).

On the other hand, all three author groups in the year 2020 uses combinations that were never deployed during 2015–2019.

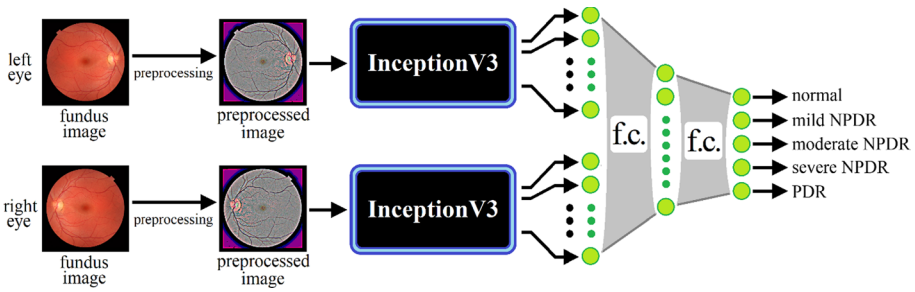


Fig. 11 The algorithm developed by Zeng et al. (2019) for General DR Detection using two mutually identical CNN architecture

3.6 The study of the New Fundus Algorithms by their designed tasks

Through due diligence of studying all the articles containing the entries of the New Fundus Algorithms, it was observed that most of these entries were built for one or several specifically designed tasks, classifiable into three main task groups as shown in Table 6.

In analogy with Table 1 and Fig. 6 under Sect. 2.1.3, we now present Table 7 and Fig. 13 showing the general distribution of all the 84 entries of the New Fundus Algorithm according to their designed tasks, shown in abbreviations referable to Table 6.

In Fig. 13, the areas of all regions are likewise shown directly proportional to the number of entries, similar to Fig. 6 from Sect. 2.1.3.

As it can be observed that the majority of the models are dedicated to general DR grading, Sect. 2.4.1 is presented to provide an in-depth review concerning about those works.

3.6.1 An in-depth review of works who performed general DR grading

In 2015, Mahendran and Dhanasekaran (2015a) accomplished the DR grading through various parameters, each summarized from the data contained in an image, such as the number of pixels as well as the grey levels of the pixels. Besides, Ibrahim et al. (2015) even considered Fourier spectrum of the textures, in addition to some parameters similar to those used by Mahendran and Dhanasekaran (2015a). On the other hand, Țălu et al. (2015) solely looked at the fractals pattern obtained from a fundus image through certain algorithms, which may or may not be resulted from blood vessels. In such a case, the grading made will depend on how much the fractal patterns truly resemble blood vessels instead of other diseased tissues.

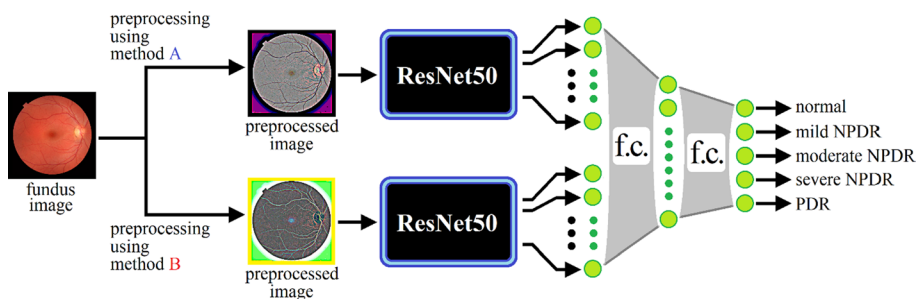


Fig. 12 The CNN architecture developed by Shankar et al. (2020a) for general DR grading

Table 3 List of observed works since 2015 on innovation on automated DR detection using purely ANN

Year	Article	Designed tasks (abbreviation's meanings refer to Table 1)
2017	Al-Jarrah and Shatnawi (2017)	G, EX and HM
	Abbas et al. (2017)	G only
2018	Randive et al. (2018)	G only
	Ramachandran et al. (2018)	G only
2019	Son et al. (2019)	CWS, EX and HM
	Gulshan et al. (2019)	ME and G
	Nazir et al. (2019)	G only
	Nair and Muthuvel (2019)	RVN
2020	Ali et al. (2020)	G only
	Derwin et al. (2020a)	MA only
	Jadhav et al. (2020)	G only
	Shankar et al. (2020b)	G only

In 2016, Gharaibeh and Alshorman (2016) separates the optical disk and blood vessels in their process of general DR grading.

In 2017, Xu et al. (2017) deployed a CNN which works on 3 input images of 224×224 , such 3 input images were obtained from the raw image using different preprocessing techniques. Quellec et al. (2017) and Raju et al. (Raju et al. 2017) deployed a CNN which works on 1 input images of 448×448 . These three CNN networks deployed in Xu et al. (2017), Quellec et al. (2017) and Raju et al. (2017) were given very detailed descriptions. As a comparison, the CNN deployed in Xu et al. (2017) contains 8 convolution layers, the CNN deployed in Raju et al. (2017) contains 11 convolution layers, whereas the CNN deployed in Quellec et al. (2017) contains 13 convolution layers. Moreover, the CNN deployed in Xu et al. (2017) deploys 2×2 max pooling, whereas the CNN deployed in Quellec et al. (2017) and Raju et al. (2017) deployed 3×3 max pooling.

Meanwhile in 2017, Al-Jarrah and Shatnawi (2017) deployed a conventional ANN which various parameters, such as the area of exudates and hemorrhages. Abbas et al. (Abbas et al. 2017) also deployed conventional ANN, however the structure of their inputs were not clearly mentioned in Abbas et al. (2017). Based on the common practice of conventional ANN, it can be expected that Abbas et al. (2017) could have deployed similar methods like Al-Jarrah and Shatnawi (2017) or Mahendran and Dhanasekaran (2015a) to obtain the appropriate input.

In 2018, Mansour (2018) feed 512×512 images as inputs into his CNN. As a trade-off due to the larger image size, his CNN consists of 5 convolution layers, less than that of Xu et al. (2017), Quellec et al. (2017) and Raju et al. (2017). Seth and Agarwal (2018) deployed a hybrid of CNN and SVM, which also worked on 512×512 images as the input. In particular, Seth and Agarwal (2018) trained their model on various orientations of a given image, such as mirror and rotations.

Meanwhile in 2018, Hemanth et al. (2018) used 6 types of textural features (mean, variance, standard deviation, contrast, entropy and energy) as inputs of his neural network-based algorithm. Such methodology is deemed similar to that deployed by Mahendran and Dhanasekaran (2015a) did in 2015.

Table 4 List of observed works since 2015 on innovation on automated DR detection using other standalone algorithms with ML, besides ANN and CNN

Model group	Model name	Article	Year	Designed tasks (refer to Table 1 for the description of the abbreviations)
Other-NN	Probabilistic neural network	Mahendran and Dhanasekaran (2015a)	2015	G only
	Modified hopfield neural network	Hemanth et al. (2018)	2018	G only
Fuzzy logic with ML	FCM (fuzzy c-means clustering)	Mahendran and Dhanasekaran (2015b)	2015	EX only
		Memari et al. (2019)	2019	RVN only
Other standalone ML	RF (random forest classifier)	Akyol et al. (2016)	2016	EX only
		Saleh et al. (2018)	2018	G only
		Chowdhury et al. (2019)	2019	CWS, MA, EX and HM
		Pratheeba and Singh (2019)	2019	EX only
	SVM (support vector machine)	Mahendran and Dhanasekaran (2015a)	2015	G only
		Derwin et al. (2020b)	2020	MA only
	k-means clustering	Mahendran and Dhanasekaran (2015b)	2015	EX only
	Discriminative dictionary learning	Javidi et al. (2017)	2017	RVN only
Dominance-based rough set balanced rule ensemble	Saleh et al. (2018)	2018	G only	

In 2019, RGP-CNN had become the norm for general DR grading. Such ML process are usually done on external server platform, such as Google Colab, and are likewise trained from ground zero without human intervention. Among the vast varieties of RGP-CNN used are InceptionV3 (Gao et al. 2018; Sahlsten et al. 2019; Zeng et al. 2019) and VGGNet (Bellemo et al. 2019). On the other hand, Saleh et al. (2018) used a unique ML algorithm, whereas Randive et al. (2018) deployed a proprietary software (MATLAB) and hence the detail of their structure remains unknown.

In 2020, the choices of model once again shifted from RGP-CNN to self-customized CNN structure. For instance, Wu et al. (2020) deployed two customized CNN networks, once worked on 256×256 preprocessed image as its input, the other worked on 512×512 preprocessed image as its input. On the other hand, Gayathri et al. (2020) deployed a customized CNN networks which worked on 227×227 preprocessed image.

Moreover, Shaban et al. (2020) deployed a customized CNN with increasingly many numbers of filters as the computation progresses. Pao et al. (2020) deployed a customized CNN with 2 channels, one accepted the red channel of the input image, the other accept the green channel of the input image. Shankar et al. (2020a) also deployed a customized CNN where all 3 channels (Red, Green, Blue) were used. Torre et al. (2020) deployed a dedicated CNN which accepts $3 \times 640 \times 640$ input images. Shah et al. (2020) deployed a simplistic CNN with only two convolutional layers. Zago et al. (2020) deployed a CNN with 5 convolution layers which accepts $3 \times 65 \times 65$ images.

Moreover, Jadhav et al. (2020) devised their unique method of general DR grading. In such grading, the blood vessels are first separated by a grey level threshold, leaving the other features to have their entropy (e.g. variance) measured. Luo et al. (2020), deployed a unique Self-supervised Fuzzy Clustering Network. Ali et al. (2020) customized a totally unique grading method in which they provided a detailed flowchart for their unique model. In addition, Bhardwaj et al. (2021).

There remain some authors in 2020 who still used RGP-CNN. For example, Shankar et al. (2020a) deployed InceptionV4 in their model for general DR grading. Samanta et al.

Table 5 The 13 entries from New Fundus Algorithms belonging to Hybrid ML, alongside with their respective constituent models

Year	Source article of the entry	Constituent models of the entry						
		CNN	ANN	Fuzzy	SVM	RF	Other ML	Genetic
2015	Ibrahim et al. (2015)		✓	✓				
2016	Gharaibeh (2016)		✓	✓				
	Gharaibeh and Alshorman (2016)		✓	✓				
2017	Zhang et al. (2017)					✓	✓	
	Barkana et al. (2017)		✓	✓	✓			
2018	Seth and Agarwal (2018)	✓			✓			
	Al-Hazaimeh et al. (2018)				✓			✓
2019	Li et al. (2019b)	✓			✓			
	Jebaseeli et al. (2019a)		✓		✓			
	Jebaseeli et al. (2019b)		✓		✓			
2020	Luo et al. (2020)	✓		✓				
	Gayathri et al. (2020)	✓					✓	
	Bhardwaj et al. (2021)				✓		✓	

Table 6 List of designed tasks among all the 84 entries of the New Fundus Algorithms, together with their respective abbreviations used in this article

Task group	Designed task	Abbreviation
General DR grading into several classes, but without pinpointing the exact locations of the diseased tissue. (G)	2-class (or Boolean) classification: The AI classifies the given fundus images into 2 classes—"normal" and "DR"	G2
	3-class classification: The AI classifies the given fundus images into 3 classes—usually "normal", "mild DR" and "severe DR"	G3
	4-class classification: The AI classifies the given fundus images into 4 classes—usually "normal", "mild NPDR", "severe NPDR" and "PDR"	G4
	5-class classification: The AI classifies the given fundus images into 5 classes—usually "normal", "mild NPDR", "moderate NPDR", "severe NPDR" and "PDR"	G5
Segmentation of DR-related diseased tissue from the normal tissue	Exudates segmentation: (whether hard, soft, or not mentioned)	EX
The AI pinpoints the pixels in the image where it is deemed to a diseased tissue of a particular type, producing a black & white two-color image	Microaneurysms segmentation	MA
	Segmentation of hemorrhages (or: red lesions/dark lesions)	HM
	Segmentation of cotton wool spots	CWS
	Segmentation of macular edema	ME
Segmentation of healthy tissue, to be used along with other algorithms	Segmentation of retinal vessel network (or: blood vessels)	RVN
The AI pinpoints the pixels in the image where it is deemed to a particular type of healthy tissue, producing a black and white two-color image	Segmentation of optic disk	OD

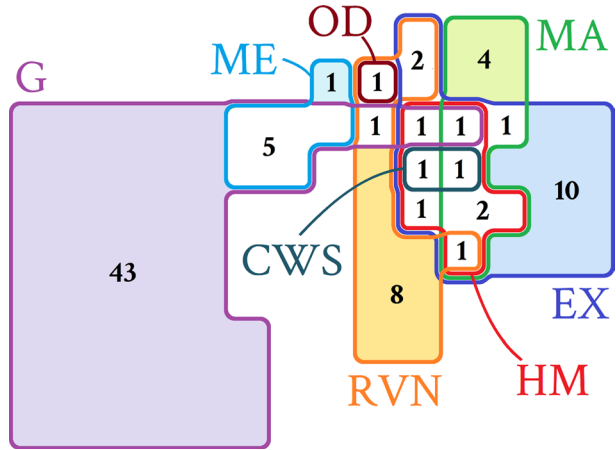
Table 7 A table showing the distribution of all the 84 entries among the New Fundus Algorithms according to their designed tasks

Designed tasks	Number of entries	Source articles of the entries
G only	43	Articles with 2 entries each: Mahendran and Dhanasekaran (2015a), Seth and Agarwal (2018) Articles with 1 entry each: (Gharaibeh and Alshorman (2016) Raju et al. (2017) Ting et al. (2017) Abbas et al. (2017) Mansour (2018) Brown et al. (2018) Randive et al. (2018) Ramachandran et al. (2018) Hemanth et al. (2018) (2019) Saleh et al. (2018) Gao et al. (2018) Gonzalez-Gonzalo et al. (2020) Sahlsten et al. (2019) Liu et al. (2019) Sun (2019) Zhang et al. (2019) Phres et al. (2019) Qummar et al. (2019) Zeng et al. (2019) Nazir et al. (2019) Li et al. (2019b) Wu et al. (2020) Shaban et al. (2020) Pao et al. (2020) Torre et al. (2020) Shah et al. (2020) Zago et al. (2020) Srivasiava and Purwar (2020) Shankar et al. (2020a) (b) Samanta et al. (2020) Xie et al. (2020) Ayhan et al. (2020) Ali et al. (2020) Jadhav et al. (2020) Luo et al. (2020) Gayathri et al. (2020) Bhardwaj et al. (2021)
EX only	10	Articles with 2 entries each: (Mahendran and Dhanasekaran (2015b) Articles with 1 entry each: (Banerjee and Kayal (2016) Partovi et al. (2016) Prentasić and Lončarić (2016) Akyol et al. (2016) Das and Puhan (2017) Kaur and Mittal (2018) Pratheeba and Singh (2019) Mateen et al. (2020)
RVN only	8	Javidi et al. (2017) Zhang et al. 2017; Barkana et al. (2017) Li et al. (2019a) Nair and Muthuvel (2019) Memari et al. (2019) Jebaseeli et al. (2019a) (2019b)
MA only	4	Gharaibeh (2016) Eftekhari et al. (2019) Derwin et al. (2020a) Derwin et al. (2020b)
ME and G	5	Abramoff et al. (2016) Gulshan et al. (2016) (2019) Bellemo et al. (2019) Usman and Almejalli (2020)
EX and RVN	2	Wardoyo et al. (2016) Dutta et al. (2015)
EX, MA and HM	2	Tan et al. (2017a) Al-Hazaimeh et al. (2018)
ME only	1	Ibrahim et al. (2015)
EX and MA	1	Qiao et al. (2020)
RVN and G	1	Țălu et al. (2015)
RVN and OD	1	Tan et al. (2017b)
EX and HM	1	Colomer et al. (2020)
G, EX and HM	1	Al-Jarrah and Shatnawi (2017)
CWS, EX and HM	1	Son et al. (2019)
G, MA, EX and HM	1	Quellec et al. 2017)(
CWS, MA, EX and HM	1	Chowdhury et al. (2019)

Table 7 (continued)

Designed tasks	Number of entries	Source articles of the entries
RVN, MA, EX and HM	1	Xu et al. 2017)

Fig. 13 An Euler diagram showing the distribution of all the 84 entries of the new fundus algorithm according to their designed tasks



(2020) deployed Densenet121. Xie et al. (2020) deployed VGGNet, ResNet and Densenet in their model.

In most articles observed which deployed preprocessing, such preprocessing was often done by nullifying the local color and luminosity, thus exposing all the textures at the same grey tone. However, the details of such formulations were not mentioned.

3.7 Study of the New Fundus Algorithms by their datasets used in the ML process

It is worth noting that the development of an ML algorithm necessitates training and testing with real life datasets. Hence, during the period of 2015–2018, most entries among the New Fundus Algorithms were observed to use *one* particular publicly available dataset, though there is nevertheless a drastic increase in the average size of the dataset used, as more and more of the authors chose to use Kaggle-EyePACS dataset (<https://www.kaggle.com/c/diabetic-retinopathy-detection/dat>), the largest dataset available publicly even as of 2020.

In 2019, there was observed an increase in the number of entries who solely rely on actual clinical data, rather than the publicly available datasets. In 2020, the Covid-19 pandemic was deemed to have severely hindered the procedure of obtaining clinical data through various medical institutions. Nevertheless, despite the Covid-19 pandemic, 13 entries among the New Fundus Algorithms were actually published just during the first 8 months 2020. 5 of those 13 entries even used multiple collection of datasets (instead of relying on a single dataset) obtained through various independent websites, to make the best out of what is available during the Covid-19 pandemic.

Just like the Euler diagrams in Figs. 6 and 13, we would like to provide similar illustration on how the authors picked their datasets. Due to the immense complexity involved, even an Euler diagram proves impractical for this case.

Thus, we will solely present Table 8 which shows the list of the most used datasets by the New Fundus Algorithms. Some of the most used datasets include Kaggle-EyePACS and Messidor. Also in Table 8, we also included references (Huang et al. 2017; Cao et al. 2018) in which the authors had done their comparative studies on several existing RGP-CNN architectures. It is worth noting that a single algorithm may had been tested on multiple datasets. In addition, there are two articles (Abbas et al. 2017; Chowdhury et al. 2019)

which only mentioned that “various publicly available datasets” were used, without pinpointing which datasets they used, whether partly or entirely. There are also another three articles observed (Țălu et al. 2015; Wardoyo et al. 2016; Wu et al. 2020) where no descriptions were given by these authors on their choices of datasets.

It is worth reminding that the three articles (Mahendran and Dhanasekaran 2015a, b; Seth and Agarwal 2018) each contributed to two entries among the New Fundus Algorithms, whereas in Huang et al. (2017) and Cao et al. (2018) the authors merely did comparative studies on the existing algorithms instead of innovating new one(s) of their own.

Thus, it is observed that more and more research teams resorted to use larger datasets, such as Kaggle-EyePACS with 88,704 samples, as computing speed progressed. The deployments of clinical datasets were also on the rise till 2019, for which 2020 marks a very difficult year to obtain such non-public datasets due to the Covid-19 pandemic.

3.8 A presentation of all the observed comparative studies performed by all the authors who innovates entries from the New Fundus Algorithms

As the entries of the New Fundus Algorithms were constructed, numerous comparative studies were done by the authors against some previous models, which was presented in Table 9. It is worth noting that for all such cases presented in Table 9, each source article yields exactly one New Fundus Algorithm.

For each New Fundus Algorithm, their claimed performances are as presented in Table 10. Due to the limitation of scope, the performances among the competitors of a given New Fundus Algorithm are presented as a range. Unless mentioned otherwise in Table 10, the percentage represents accuracies, calculated by the formula:

$$\text{Accuracy} = \frac{\text{True Positive} + \text{True Negative}}{\text{All Positive} + \text{All Negative}}$$

Hence, based on the comparative studies done across all the relevant articles, CNN was deemed to be the best overall performing models for DR detection through fundus images. This is consistent with the very nature of CNN which directly deal with pixels on an image, rather than relying on the numerical data which are themselves only representatives from the raw data.

On a quantitative perspective, however, it is worth noting that the quantitative comparisons of accuracy is only applicable to all the models involved within one single article. For example, the New Fundus Algorithm proposed in Al-Hazaimah et al. (2018), which yielded an accuracy of 98.4%, is deemed better than all the competitor tested by the same authors within (Al-Hazaimah et al. 2018) itself, which yielded accuracies ranges in 89.6% to 97.6%. Still, it *cannot* be concluded that the New Fundus Algorithm proposed in Al-Hazaimah et al. (2018), which yielded an accuracy of 98.4%, is better than the New Fundus Algorithm proposed in Liu et al. (2019), which yielded an accuracy of 90.84% by a different team of authors. This is because there remain many degrees of freedom involved, such as the duration of the machine learning process, the size of the dataset used, and the amount of computational resource available (e.g., computer hardware). Such parameters may differ significantly among different authors. Moreover, on all articles observed, there remains a great deal of detail of computations which are not given by the authors, such as the exact preprocessing formula.

4 A brief review on some other developmental progress related to automated fundus based DR detection

4.1 Some recently commercialized applications for fundus-based automated DR detections

In addition to the 81 articles which produced 84 distinctive New Fundus Algorithms. Another 5 articles (Rajalakshmi et al. 2018; Natarajan et al. 2019; He et al. 2019; Rogers et al. 2021; Wang et al. 2020) gives elaborations on the recently commercialized applications by the authors of those journals, which in turn are prominent staffs of their respective companies. An overview is provided in Table 11 below.

In particular, the authors of Rajalakshmi et al. (2018) had actually compared the results given by their algorithms with that of actual medical experts. Subsequently, their automatic DR performance is concluded to be at least on par with the manual grading by those medical experts.

In Natarajan et al. (2019) the performance of an AI was examined in actual context. The AI algorithm was designed to perform on a mobile device (with the application installed) attached to a fundus camera. It is also worth mentioning that most countries in the world do not approve apps ran on mobile devices for medical use.

The authors of He et al. (2019) only mentioned that they used a “deep learning algorithm”, however the internal mechanism of the algorithm was not discussed. It is however deemed likely to be CNN by this review article. Based on their testing, the authors in He et al. (2019) had thus concluded that it is now feasible to deploy AI for DR in China.

4.2 Some recently developed standalone add-ons to improve existing ML-based algorithms for fundus-based automated DR detections

There are altogether 4 articles (Gupta et al. 2017; Nilashi et al. 2019; Gadekallu et al. 2020a, b) found to have worked on developing standalone add-one to further enhance pre-existing ML based algorithms.

In Gupta et al. (2017), an extension is recommended to several existing MF-based algorithms. Such extension works by dividing the fundus images into several segments. Thus, when coupled with an algorithm that can only detect the presence/absence of diseased tissue but cannot pinpoint the location (e.g. G, G2, G3, G4, G5), such patching can somewhat help in pointing out the location of such diseased tissue.

In Nilashi et al. (2019) and Gadekallu et al. (2020a) The datasets consist of lists of 20 attributes derived from MASSIDOR, thus no images were involved. Instead, the add-on works in improving the efficiency of internal computations undertaken by some ML-based algorithms.

In Gadekallu et al. (2020b), another add-on of computations as also introduced in which, combined with some existing NN based algorithms (ANN, CNN, Other-NN) claimed to be able to increase their accuracies even further.

Table 8 The summary of the datasets used by the New Fundus Algorithms

Datasets number of samples)	Year					
	2015	2016	2017	2018	2019	2020
Publicly available datasets						
Kaggle-EyePACS (88,704)	–	Gulshan et al. (2016)*	Quellec et al. (2017); Raju et al. (2017) Xu et al. (2017)*	Mansour (2018) (Seth and Agarwal (2018) Huang et al. (2017)*	(Pires et al. 2019; Nazir et al. 2019; Li et al. 2019b) (Qummar et al. 2019; Zeng et al. 2019)*	(Pao et al. 2020; Torre et al. 2020; Zago et al. 2020; Aghan et al. 2020) (Shaban et al. 2020; Srivastava and Purwar 2020; Samanta et al. 2020; Gayathri et al. 2020)*
Messidor (1200)	(Mahendran and Dhasekaran 2015a)	(Gharraibeh 2016)	(Dutta et al. 2015)	(Kaur and Mittal 2018; Ramachandran et al. 2018)	(Gonzalez-Gonzalo et al. 2020)	(Shah et al. 2020; Shankar et al. 2020a, 2020b; Ali et al. 2020; Jadhav et al. 2020; Usman and Almejalli 2020; Luo et al. 2020; Gayathri et al. 2020; Bhardwaj et al. 2021)
Messidor-2 (1748)	(Abramoff et al. 2016; Gulshan et al. 2016)	–	–	–	–	Torre et al. 2020)
DiaretDB0 (200)	–	(Banerjee and Kayal 2016)	–	–	(Pratheeba and Singh 2019)	(Zago et al. 2020)
DiaretDB1 (89)	–	(Banerjee and Kayal 2016; Akyol et al. 2016)	(Das and Puhon 2017; Al-Jarrah and Shatnawi 2017)	(Kaur and Mittal 2018; Randive et al. 2018; Al-Hazaimeh et al. 2018; Cao et al. 2018)	–	(Colomer et al. 2020; Usman and Almejalli 2020; Luo et al. 2020)

Table 8 (continued)

Datasets number of samples	Year	2015	2016	2017	2018	2019	2020
STARE (400)	-	-	(Gharaibeh and Alshorman 2016)	-	(Kaur and Mittal 2018)	(Liu et al. 2019; Hemanth et al. 2019; Nazir et al. 2019; Nair and Muthuvel 2019; Memari et al. 2019; Jebaseeli et al. 2019a, 2019b)	-
DRIVE (40)	-	-	(Gharaibeh and Alshorman 2016)	(Dutta et al. 2015; Tan et al. 2017b; Javidi et al. 2017; Zhang et al. 2017; Barkana et al. 2017)	-	(Nazir et al. 2019; Nair and Muthuvel 2019; Memari et al. 2019; Jebaseeli et al. 2019a, 2019b)	(Luo et al. 2020)
IDRiD (82)	-	-	-	-	-	-	(Zago et al. 2020; Qiao et al. 2020; Ayhan et al. 2020) (Gayathri et al. 2020)*
HRF (45)	-	-	-	-	-	(Jebaseeli et al. 2019a, 2019b) (Nair and Muthuvel 2019)*	(Srivastava and Purwar 2020; Ali et al. 2020; Jadhav et al. 2020)
Others	-	-	-	(Das and Puhon 2017; Tan et al. 2017a)	(Ramachandran et al. 2018; Cao et al. 2018)	(Gao et al. 2018; Eftekhari et al. 2019; Nazir et al. 2019; Memari et al. 2019; Jebaseeli et al. 2019a, 2019b)	(Mateen et al. 2020; Zago et al. 2020; Derwin et al. 2020a, 2020b)

Table 8 (continued)

Datasets number of samples	Year	2016	2017	2018	2019	2020	
Clinical datasets		(Mahendran and Dhasekaran 2015b; Ibrahim et al. 2015)	(Partovi et al. 2016; Prentašić and Lončarić 2016; Gulshan et al. 2016)	(Ting et al. 2017)	(Brown et al. 2018; Hemanth et al. 2018; Saleh et al. 2018)	(Gonzalez-Gonzalo et al. 2020; Sahlsten et al. 2019; Sun 2019; Zhang et al. 2019; Li et al. 2019a; Bellemo et al. 2019; Son et al. 2019; Gulshan et al. 2019)	(Shah et al. 2020; Xie et al. 2020; Ayhan et al. 2020)

Citation collections marked with an asterisk (*) means only a portion of the dataset was used by the entries in those articles.

Table 9 List of New Fundus Algorithms where comparative studies have performed, together with their type of models

Year	Source article	Model group of the entry	Particulars of the innovated model(s)	Previous model(s) compared against
2016	Akyol et al. (2016)	Pure RF	RF	ANN Decision tree
	Gharaibeh and Alshorman (2016)	Hybrid ML	ANN + Fuzzy operating system	SVM PNN FCM
2017	Xu et al. (2017)	Pure CNN	CNN of self-designed structure	Gradient Boosting Machine
	Barkana et al. (2017)	Hybrid ML	Type 1 FIS + ANN + SVM	Type 1 FIS ANN SVM
2018	Al-Hazaimeh et al. (2018)	Hybrid ML	SVM + genetic algorithm	SVM PNN

Table 9 (continued)

Year	Source article	Model group of the entry	Particulars of the innovated model(s)	Previous model(s) compared against
2019	Liu et al. (2019)	Pure CNN	CNN with weighted path	ResNet DenseNet SeNet
	Sun (2019)	Pure CNN	CNN that is batch normalized	Multi layer perceptron Logistic regression
	Zhang et al. (2019)	Pure CNN	CNN incorporating Inception + Xception + another hybrid (Inception + ResNet)	Inception Xception Hybrid(Inception + ResNet) ResNet DenseNet
	Li et al. (2019a)	Pure CNN	CNN of U-Net Structure	CNN (traditional model) CNN (fully convolutional) U-Net
	Pires et al. (2019)	Pure CNN	CNN with 3-channel for 3 different resolutions	The structure used by o_O team who obtained 2 nd in the competition by Kaggle
	Zeng et al. (2019)	Pure CNN	CNN incorporating the structure of Inception V3, Binocular Siamese-like	The structure used by o_O team who obtained 2 nd in the competition by Kaggle
	Chowdhury et al. (2019)	Pure RF	RF	Naive Bayes classifiers
	Li et al. (2019b)	Hybrid ML	CNN + SVM	Random forest

Table 9 (continued)

Year	Source article	Model group of the entry	Particulars of the innovated model(s)	Previous model(s) compared against
2020	Mateen et al. (2020)	Pure CNN	CNN incorporating Inception + Resnet + VGG	Inception ResNet VGG
	Srivastava and Purwar (2020)	Pure CNN	CNN of self-designed structure	GoogleNet
	Jadhav et al. (2020)	Pure ANN	ANN with modified levy updated-dragonfly algorithm-neural network (MLUDA-NN)	SVM + KNN SVM ANN
	Shankar et al. (2020b)	Pure ANN	ANN with extension: Inception + MLP	M-AlexNet AlexNet VggNet-s VggNet-16 VggNet-19 GoogleNet ResNet
	Luo et al. (2020)	Hybrid ML	CNN + fuzzy c-means	k-means clustering hierarchical clustering fuzzy c-means

Table 10 List of New Fundus Algorithms where comparative studies have performed, together with their claimed performances

Year	Source article	Claimed accuracies (unless mentioned otherwise) of the innovated New Fundus Algorithm	Claimed accuracies (unless mentioned otherwise) ranges among all the competitors
2016	Akyol et al. (2016)	3 datasets, 92.98% 85.96% 87.50% 99%	3 datasets, 80.70–85.96% 78.94–85.96% 69.64–78.57% 89.6–97.6% 79.1–89.4%
2017	Gharaibeh and Alshorman (2016) Xu et al. (2017)	91.5% without data augmentation 94.5% with data augmentation	
2018	Barkana et al. (2017)	94.01–96.86%	89.66–96.4%
2019	Al-Hazaimeh et al. (2018) Liu et al. (2019) Sun (2019) Zhang et al. (2019) Li et al. (2019a) Pires et al. (2019) Zeng et al. (2019) Chowdhury et al. (2019) Li et al. (2019b)	98.4% 90.84% 99.85% 95.46–96.50% CR index: 0.700–0.712 AUC: 98.2% Kappa score: 0.829 93.58% 86.17%	89.6–97.6% 81.68–85.50% 96.857–97.57% 92.09–94.94% CR index: 0.552–0.683 AUC: 97.9% Kappa score: 0.800 83.64% 86.01%
2020	Mateen et al. (2020) Srivastava and Purwar (2020) Jadhav et al. (2020) Shankar et al. (2020b) Luo et al. (2020)	98.43% 2 datasets, Precision: 100.00%, 95.16% 93.33% 99.49% 97%	93.67–97.80% 2 datasets, Precision: 86.29%, 95.16% 79.55–90.00% 89.75–93.73% 82–90%

5 Recent development on the automated detection of DR using OCTA (optical coherence tomography angiography) images

OCTA (Optical Coherence Tomography Angiography) is a newly developed retinal scanning technology that can extract much more information of the retina as compared to conventional fundus images. In particular, OCTA is capable of resolving even the tiniest retinal blood vessels, causing all MA to show up with much more clarity, as compared to a conventional fundus image. However, up until the end of 2020, almost all the research done on automated DR detection has remain focused only on utilizing conventional fundus images, and all DR detection algorithms created for OCTA were only at very primitive stages of development. As technology advances, it is expected that the use of OCTA will soon replace that of conventional fundus imaging. Therefore, it is crucial to develop new algorithms that are able to utilize this frontier technology to enable early detection of DR, especially at the very early stage where only MA is often present.

There are altogether 13 articles (Srinivasan et al. 2014; Carlo et al. 2015; Sandhu et al. 2018a, b, 2020; Guo et al. 2019; Lee et al. 2019; Cano et al. 2020; Heisler et al. 2020; Zahran et al. 2020; Abdelsalam 2020; Le et al. 2020; Wang et al. 2017), each of them was found to have proposed some new algorithms that works on interpreting OCTA images to a certain extent. Thus, throughout all the 13 aforementioned articles, there are altogether 13 algorithms. These 13 algorithms shall therefore be collectively referred as the new OCTA algorithms throughout the remaining passages of this article.

Both the constituent model groups and the designed tasks, among all entries of the new OCTA algorithms, are thus summarized in Tables 12 and 13. In particular, most of the entries designed for G (from Sandhu et al. 2018a, b, 2020; Guo et al. 2019; Cano et al. 2020; Heisler et al. 2020; Zahran et al. 2020; Abdelsalam 2020; Le et al. 2020) are claimed to be capable of distinguishing the following 3 different categories of DR severity, namely.

- (i) Normal.
- (ii) Diabetic without DR (a feature that is detectable only by OCTA imaging).
- (iii) DR.

Moreover, all the datasets involved in training and testing the new OCTA algorithms are obtained clinically and therefore not publicly available.

Table 14 is presented to show the sizes of the clinical datasets obtained for training and testing all the new OCTA algorithms. Although the number of samples may seem like only a few, the OCTA image of a single eye already carries a tremendous amount of information showing every intricate detail of the blood vessels and other type of tissue.

6 Conclusion

Due to the exponential growth in numbers of predicted future diabetes mellitus patients and inadvertently the victims subject to DR, the number of ophthalmologists worldwide may not be able to match regulatory screening demands. A safe clinically available automated detection algorithm has been a strongly sought-after option in tackling the problem.

It was found that most entries from the New Fundus Algorithms, including those deploying the advanced ML models, still work on the principal of maximum likelihood.

Table 11 Particulars for the 5 recently commercialized applications for automated fundus-based DR detection

Product name	EyeArt®	Medios AI®	Airdoc®	Pegasus®	DeepDR®
Company	Eyenuk, Inc., USA	Remidio Innovative Solutions Pvt. Ltd., India	Airdoc LLC, China	Visulytix Ltd., UK	Shanghai Jiao Tong University, China
Article	Rajalakshmi et al. (2018)	Natarajan et al. (2019)	He et al. (2019)	Rogers et al. (2021)	Wang et al. (2020)
Publication year	2018	2019	2019	2020	2020
Model group (see Sect. 2.1.2)	Pure CNN (likely)	Pure ANN (unidentifiable type of neural network structure)	Pure CNN (likely)	Pure CNN (likely)	Pure CNN
Designed tasks (see Sect. 2.1.3)	G	G	G	G	G,EX,MA,HM

Table 12 The distribution of all the 13 entries of the new OCTA algorithms according to their constituent models and method

Model group	Constituent model/method	Number of entries	Source articles of the entries
Pure CNN	CNN	4	Guo et al. (2019), Lee et al. (2019), Heisler et al. (2020) and Le et al. (2020)
Pure ANN	ANN	2	Zahran et al. (2020) and Abdelsalam (2020)
Conventional	–	2	Carlo et al. (2015) and Wang et al. (2017)
Pure RF	RF	1	Sandhu et al. (2020)
Pure fuzzy-ML	Fuzzy-ML	1	Wang et al. (2017)
Pure other-NN	Other-NN	1	Sandhu et al. (2018b) (*Deep Fusion Classification Network)
Pure SVM	SVM	2	Srinivasan et al. (2014) and Sandhu et al. (2018a)

Therefore, they can only detect the disease when it is at an advanced stage. Those entries from the New Fundus Algorithms that use ML algorithms (including those using deep learning algorithms) only look at the degree of brightness to detect the optical disc and exudates in the retina based on the common assumption that the brightest area in the retina must be the optical disc. Therefore, most systems are not even able to detect the optical disc correctly. This hampers the ability of the system to do early detection which is pertinent to prevent vision loss.

It was also found that most entries from New Fundus Algorithms, even those designed to pinpoint diseased or healthy tissues (e.g., EX and HM), can only work with images with fixed and very low pixel count of length and width. For example, even algorithms designed to “pinpoint” exudates are usually designed to work on 160 by 240 pixels, manually pre-processed images. Such common practice of shrinking down the images to predetermined pixel sizes has rendered early detection almost impossible. This is because in the early stages of DR, the diseased tissues (e.g., microaneurysms) are usually too small to be visible on such a low-resolution image.

Among the entries from the New Fundus Algorithms that uses Pure CNN, many have resorted to use RGP-CNN (e.g., ResNet50) as there has been many readily made programming packages for these RGP-CNN. With this mere adaptation of the existing, general-purpose structure, the ML models deployed by most of the models are not optimized specifically for the task of automated detection of DR.

As OCTA imaging is a very recent technology that is very resource demanding, the research on the automated DR detection using OCTA imaging is still at the very early stages of development.

There are still several challenges to the detection of DR. One of them is to develop image processing techniques/methods that are able to process retinal images with low resolution and blurred retinal images. Another is to develop accurate detection methods for the detection of borderline cases. This can be done using a combination of advanced mathematical tools such as complex fuzzy logic and machine learning algorithms. Finally, by acquiring as many clinical datasets will lead to the development of a novel and fully functional DR detection algorithm for OCTA images. Then, a dedicated ML algorithm not based on any pre-existing models can be used for the purpose of automated DR detection. As every single part of the programming code is designed specifically for the task of DR

Table 13 The distribution of all the 13 entries among the new OCTA algorithm according to their designed tasks

Designed tasks (abbreviation's meanings refer to Table 1)	Number of entries	Source articles of the entries
G only	9	Sandhu et al. (2018a, b), Guo et al. (2019), Cano et al. (2020), Heisler et al. (2020), Zahran et al. (2020), Abdelsalam (2020), Le et al. (2020), and Sandhu et al. (2020)
ME only	1	Srinivasan et al. (2014)
MA only	1	Carlo et al. (2015)
RVN only	1	Lee et al. (2019)
Photoreceptor disruption (a feature that is detectable only by OCTA imaging)	1	Wang et al. (2017)

Table 14 The sizes of the clinical datasets used to train and test the entries among the new OCTA algorithm

Dataset size	Number of entries	Source articles of the entries
160 eyes (80 are healthy, 40 are diabetic but without DR, 40 are with DR)	2	Sandhu et al. (2018a, b)
89 eyes (28 are healthy, 61 are with symptoms of DR)	1	Carlo et al. (2015)
84 eyes (44 are healthy, 40 are with DR)	1	Guo et al. (2019)
35 eyes (22 are healthy, 13 are with DR)	1	Cano et al. (2020)
45 “objects”	1	Srinivasan et al. (2014)
36 “objects”	1	Wang et al. (2017)
Not mentioned	6	Lee et al. (2019), Heisler et al. (2020), Zahran et al. (2020), Abdelsalam (2020), Le et al. (2020) and Sandhu et al. (2020)

detection, the performance of our dedicated ML algorithm is thus optimized and will even be capable of being run on any electronic devices.

Moreover, it has also come to our attention that Mongan et al. (2020) had outlined a checklist for artificial intelligence in medical imaging (abbr. CLAIM) in 2020. Although the article was published in the field of radiology, it serves as a new benchmark for all upcoming AI in medical imaging to comply or even surpass. Such very recent and unprecedentedly comprehensive checklist shall prove very useful to all our future undertaking in developing a fully functional DR detection algorithm, because a full compliance with the checklist guarantees a major advantage over the previous New Fundus Algorithms, which were innovated mostly before 2020 and therefore severely fall short of the checklist suggested by Mongan et al. (2020).

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s10462-022-10185-6>.

Acknowledgements The authors would like to thank the Editor-in-Chief and the anonymous reviewers for their valuable comments and suggestions.

Funding The authors (Ganeshsree Selvachandran and Shio Gai Quek) would like to thank the Ministry of Higher Education, Malaysia for funding this research under Grant No. FRGS/1/2020/STG06/UCSI/02/1.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

References

- Abbas Q, Fondon I, Sarmiento A, Jiménez S, Alemany P (2017) Automatic recognition of severity level for diagnosis of diabetic retinopathy using deep visual features. *Med Biol Eng Comput* 55:1959–1974
- Abdelsalam MM (2020) Effective blood vessels reconstruction methodology for early detection and classification of diabetic retinopathy using OCTA images by artificial neural network. *Inf Med Unlocked* 20:100390
- Abràmoff MD, Lou Y, Erginay A, Clarida W, Amelon R, Folk JC, Niemeijer M (2016) Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. *Invest Ophthalmol vis Sci* 57(13):5200–5206

- Aiello LP, Avery RL, Arrigg PG, Keyt BA, Jampel HD, Shah ST, Pasquale LR, Thieme H, Iwamoto MA, Park JE, Nguyen HV, Aiello LM, Ferrara N, King GL (1994) Vascular endothelial growth factor in ocular fluid of patients with diabetic retinopathy and other retinal disorders. *N Engl J Med* 331(22):1480–1487
- Akyol K, Bayir Ş, Şen B (2016) Detection of hard exudates in diabetic retinopathy retinal images by utilizing visual dictionary and classifier approaches. *Mugla J Sci Technol* 2(1):1–6
- Al-Hazaimeh OM, Nahar KMO, Al-Naami B, Gharaibeh N (2018) An effective image processing method for detection of diabetic retinopathy diseases from retinal fundus images. *Int J Signal Imaging Syst Eng* 11(4):206–216
- Ali A, Qadri S, Mashwani WK, Kumam W, Kumam P, Naeem S, Goktas A, Jamal F, Chesneau C, Anam S, Sulaiman M (2020) Machine learning based automated segmentation and hybrid feature analysis for diabetic retinopathy classification using fundus image. *Entropy* 22(5):567
- Al-Jarrah MA, Shatnawi H (2017) Non-proliferative diabetic retinopathy symptoms detection and classification using neural network. *J Med Eng Technol* 41(6):498–505
- Alyoubi WL, Shalash WM, Abulkhair MF (2020) Diabetic retinopathy detection through deep learning techniques: a review. *Inf Med Unlocked* 20:100377
- Amin J, Sharif M, Yasmin M (2016) A review on recent developments for detection of diabetic retinopathy. *Scientifica* 2016:6838976
- Antonetti DA, Barber AJ, Bronson SK, Freeman WM, Gardner TW, Jefferson LS, Kester M, Kimball SR, Krady JK, LaNoue KF, Norbury CC, Quinn PG, Sandirasegarane L, Simpson IA, JDRF Diabetic Retinopathy Center Group (2016) Diabetic retinopathy: seeing beyond glucose-induced microvascular disease. *Diabetes* 55(9):2401–2411
- Ayhan MS, Kühlewein L, Aliyeva G, Inhoffen W, Ziemssen F, Berens P (2020) Expert-validated estimation of diagnostic uncertainty for deep neural networks in diabetic retinopathy detection. *Med Image Anal* 64:101724
- Banerjee S, Kayal D (2016) Detection of hard exudates using mean shift and normalized cut method. *Biocybern Biomed Eng* 36(4):679–685
- Barkana BD, Saricicek I, Yildirim B (2017) Performance analysis of descriptive statistical features in retinal vessel segmentation via fuzzy logic, ANN, SVM, and classifier fusion. *Knowl Based Syst* 118:165–176
- Bellemo V, Lim ZW, Lim G, Nguyen GD, Xie Y, Yip MYT, Hamzah H, Ho J, Lee XQ, Hsu W, Lee ML, Musonda L, Chandran M, Chipalo-Mutati G, Muma M, Tan GSW, Sivaprasad S, Menon G, Wong TY, Ting DSW (2019) Artificial intelligence using deep learning to screen for referable and vision-threatening diabetic retinopathy in Africa: a clinical validation study. *Lancet Digital Health* 1(1):e35–e44
- Bhardwaj C, Jain S, Sood M (2021) Hierarchical severity grade classification of non-proliferative diabetic retinopathy. *J Ambient Intell Humaniz Comput* 12:2649–2670
- Brown JM, Campbell JP, Beers A, Chang K, Ostmo S, Chan RVP, Dy J, Erdogmus D, Ioannidis S, Kalpathy-Cramer J, Chiang MF (2018) Automated diagnosis of plus disease in retinopathy of prematurity using deep convolutional neural networks. *JAMA Ophthalmol* 136(7):803–810
- Cano J, O'neill WD, Penn RD, Blair NP, Kashani AH, Ameri H, Kaloostian CL, Shahidi M (2020) Classification of advanced and early stages of diabetic retinopathy from non-diabetic subjects by an ordinary least squares modeling method applied to OCTA images. *Biomed Opt Express* 11(8):4666–4678
- Cao W, Czarnek N, Shan J, Li L (2018) Microaneurysm detection using principal component analysis and machine learning methods. *IEEE Trans Nanobiosci* 17(3):191–198
- Carlo TE, Chin AT, Filho MAB, Adhi M, Branchini L, Salz DA, Baumal CR, Crawford C, Reichel E, Witkin AJ, Duker JS, Waheed NK (2015) Detection of microvascular changes in eyes of patients with diabetes but not clinical diabetic retinopathy using optical coherence tomography angiography. *Retina* 35(11):2364–2370
- Chaki J, Ganesh ST, Cidham SK, Theertan SA (2020) Machine learning and artificial intelligence based diabetes mellitus detection and self-management: a systematic review. *J King Saud Univ Comput Inf Sci*. <https://doi.org/10.1016/j.jksuci.2020.06.013>
- Chowdhury AR, Chatterjee T, Banerjee S (2019) A random forest classifier-based approach in the detection of abnormalities in the retina. *Med Biol Eng Comput* 57:193–203
- Colomer A, Igual J, Naranjo V (2020) Detection of early signs of diabetic retinopathy based on textural and morphological information in fundus images. *Sensors* 20(4):1005
- Das V, Puhani NB (2017) Tsallis entropy and sparse reconstructive dictionary learning for exudate detection in diabetic retinopathy. *J Med Imaging* 4(2):024002
- Derwin DJ, Selvi ST, Singh OJ, Shan BP (2020a) A novel automated system of discriminating microaneurysms in fundus images. *Biomed Signal Process Control* 58:101839

- Derwin DJ, Selvi ST, Singh OJ (2020b) Secondary observer system for detection of microaneurysms in fundus images using texture descriptors. *J Digit Imaging* 33:159–167
- Duh EJ, Sun JK, Stitt AW (2017) Diabetic retinopathy: current understanding, mechanisms, and treatment strategies. *JCI Insight* 2(14):e93751
- Dutta MK, Parthasarathi M, Ganguly S, Ganguly S, Srivastava K (2015) An efficient image processing based technique for comprehensive detection and grading of nonproliferative diabetic retinopathy from fundus images. *Comput Methods Biomech Biomed Eng* 5(3):195–207
- Eftekhari N, Pourreza HR, Masoudi M, Ghiasi-Shirazi K, Sae E (2019) Microaneurysm detection in fundus images using a two-step convolutional neural network. *Biomed Eng Online* 18:67
- Gadekallu TR, Khare N, Bhattacharya S, Singh S, Maddikunta PKR, Srivastava G (2020a) Deep neural networks to predict diabetic retinopathy. *J Ambient Intell Humaniz Comput*. <https://doi.org/10.1007/s12652-020-01963-7>
- Gadekallu TR, Khare N, Bhattacharya S, Singh S, Maddikunta PKR, Ra IH, Alazab M (2020b) Early detection of diabetic retinopathy using PCA-Firefly based deep learning model. *Electronics* 9(2):274
- Gao Z, Li J, Guo J, Chen Y, Yi Z, Zhong J (2018) Diagnosis of diabetic retinopathy using deep neural networks. *IEEE Access* 7:3360–3370
- Gayathri S, Varun PG, Palanisamy P (2020) A lightweight CNN for diabetic retinopathy classification from fundus images. *Biomed Signal Process Control* 62:102115
- Genuth S, Sun W, Cleary P, Gao X, Sell DR, Lachin J, Monnier VM (2015) Skin advanced glycation end products glucosepane and methylglyoxal hydroimidazolone are independently associated with long-term microvascular complication progression of type 1 diabetes. *Diabetes* 64(1):266–278
- Gharaibeh NY (2016) A novel approach for detection of microaneurysms in diabetic retinopathy disease from retinal fundus images. *Comput Inf Sci* 10(1):1–15
- Gharaibeh NY, Alshorman AA (2016) An effective diagnosis of diabetic retinopathy with aid of soft computing approaches. *J Power Energy Eng* 10(8):474–485
- Gonzalez-Gonzalo C, Sanchez-Gutierrez V, Hernandez-Martinez P, Contreras I, Lechanteur YT, Domninan A, Ginneken BV, Sanchez CC (2020) Evaluation of a deep learning system for the joint automated detection of diabetic retinopathy and age-related macular degeneration. *Acta Ophthalmol* 98(4):368–377
- Gross JG, Glassman AR, Jampol LM, Inusah S, Aiello LP, Antoszyk AN, Baker CW, Berger BB, Bressler NM, Browning D, Elman MJ, Ferris FL 3rd, Friedman SM, Marcus DM, Melia M, Stockdale CR, Sun JK, Beck RW (2015) Panretinal photocoagulation vs intravitreal ranibizumab for proliferative diabetic retinopathy: a randomized clinical trial. *JAMA* 314(20):2137–2146
- Grzybowski A, Brona P, Lim G, Ruamviboonsuk P, Tan GSW, Abramoff M, Ting DSW (2019) Artificial intelligence for diabetic retinopathy screening: a review. *Eye* 34(3):451–460
- Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams T, Cuadros J, Kim R, Raman R, Nelson PC, Mega JL, Webster DR (2016) Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* 316(22):2402–2410
- Gulshan V, Rajan RP, Widner K, Wu D, Wubbels P, Rhodes T, Whitehouse K, Coram M, Corrado G, Ramasamy K, Raman R, Peng L, Webster DR (2019) Performance of a deep-learning algorithm vs manual grading for detecting diabetic retinopathy in India. *JAMA Ophthalmol* 137(9):987–993
- Guo YK, Hormel TT, Xiong HL, Wang BJ, Camino A, Wang J, Huang D, Hwang TS, Jia YL (2019) Development and validation of a deep learning algorithm for distinguishing the nonperfusion area from signal reduction artifacts on OCT angiography. *Biomed Opt Express* 10(7):3257–3268
- Gupta G, Kulasekaran S, Ram K, Joshi N, Sivaprakasam M, Gandhi R (2017) Local characterization of neovascularization and identification of proliferative diabetic retinopathy in retinal fundus images. *Comput Med Imaging Graph* 55:124–132
- He J, Cao T, Xu F, Wang S, Tao H, Wu T, Sun L, Chen J (2019) Artificial intelligence-based screening for diabetic retinopathy at community hospital. *Eye* 34(3):572–576
- He K, Zhang X, Ren S, Sun J (2015) Deep residual learning for image recognition. In: *Proceedings of the 2016 IEEE conference on computer vision and pattern recognition (CVPR)*, Las Vegas
- Heisler M, Karst S, Lo J, Mammo Z, Yu T, Warner S, Maberley D, Beg MF, Navajas EV, Sarunic MV (2020) Ensemble deep learning for diabetic retinopathy detection using optical coherence tomography angiography. *Transl vis Sci Technol* 9(2):20
- Hemanth DJ, Anitha J, Son LH, Mittal M (2018) Diabetic retinopathy diagnosis from retinal images using modified hopfield neural network. *J Med Syst* 42(12):247
- Hemanth DJ, Deperlioglu O, Kose U (2019) An enhanced diabetic retinopathy detection and classification approach using deep convolutional neural network. *Neural Comput Appl* 32:707–721

- Hu J, Shen L, Sun G (2018) Squeeze-and-excitation networks. In: Proceedings of the 2018 IEEE conference on computer vision and pattern recognition (CVPR), Salt Lake City, pp 7132–7141
- Huang G, Liu Z, Maaten L, Weinberger KQ (2017) Densely connected convolutional networks. In: Proceedings of the 2017 IEEE conference on computer vision and pattern recognition (CVPR), Honolulu, pp 4700–4708
- Ibrahim S, Chowriappa P, Dua S, Acharya UR, Noronha K, Bhandary S, Mugasa H (2015) Classification of diabetes maculopathy images using data-adaptive neuro-fuzzy inference classifier. *Med Biol Eng Comput* 53:1345–1360
- Ishtiaq U, Kareem SA, Abdullah ERMF, Mujtaba G, Jahangir R, Ghafoor HY (2020) Diabetic retinopathy detection through artificial intelligent techniques: a review and open issues. *Multimed Tools Appl* 79:15209–15252
- Islam MM, Yang HC, Poly TN, Jian WS, Li YC (2020) Deep learning algorithms for detection of diabetic retinopathy in retinal fundus photographs: a systematic review and meta-analysis. *Comput Methods Prog Biomed* 191:105320
- Jadhav AS, Patil PB, Biradar S (2020) Computer-aided diabetic retinopathy diagnostic model using optimal thresholding merged with neural network. *Int J Intell Comput Cybern* 13(3):283–310
- Javidi M, Pourreza HR, Harati A (2017) Vessel segmentation and microaneurysm detection using discriminative dictionary learning and sparse representation. *Comput Methods Programs Biomed* 139:93–108
- Jebaseeli TJ, Durai CAD, Peter JD (2019a) Segmentation of retinal blood vessels from ophthalmologic diabetic retinopathy images. *Comput Electr Eng* 73:245–258
- Jebaseeli TJ, Durai CAD, Peter JD (2019b) Retinal blood vessel segmentation from diabetic retinopathy images using tandem PCNN model and deep learning based SVM. *Optik* 199:163328
- Kandel I, Castelli M (2020) Transfer learning with convolutional neural networks for diabetic retinopathy image classification: a review. *Appl Sci* 10:2021
- Kaur J, Mittal D (2018) A generalized method for the segmentation of exudates from pathological retinal fundus images. *Biocybern Biomed Eng* 38(1):27–53
- Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. *Adv Neural Inf Process Syst* 25:1097–1105
- Le D, Alam M, Yao CK, Lim JI, Hsieh YT, Chan RVP, Toslak D, Yao X (2020) Transfer learning for automated OCTA detection of diabetic retinopathy. *Transl vis Sci Technol* 9(2):35
- Lecun Y, Bottou L, Bengio Y, Haffner P (1998) Gradient-based learning applied to document recognition. *Proc IEEE* 86(11):2278–2324
- Lee CS, Tyring AJ, Wu Y, Xiao S, Rokem AS, DeRuyter NP, Zhang Q, Tufail A, Wang RK, Lee AY (2019) Generating retinal flow maps from structural optical coherence tomography with artificial intelligence. *Sci Rep* 9:5694
- Li Q, Fan S, Chen C (2019a) An intelligent segmentation and diagnosis method for diabetic retinopathy based on improved U-NET network. *J Med Syst* 43:304
- Li YH, Yeh NN, Chen SJ, Chung YC (2019b) Computer-assisted diagnosis for diabetic retinopathy based on fundus images using deep convolutional neural network. *Mob Inf Syst* 2019:6142839
- Liu YP, Li Z, Xu C, Li J, Liang R (2019) Referable diabetic retinopathy identification from eye fundus images with weighted path for convolutional neural network. *Artif Intell Med* 99:101694
- Lu W, Tong Y, Yu Y, Xing Y, Chen C, Shen Y (2018) Applications of artificial intelligence in ophthalmology: general overview. *J Ophthalmol* 2018:527819
- Luo YG, Pan J, Fan SS, Du ZY, Zhan GH (2020) Retinal image classification by self-supervised fuzzy clustering network. *IEEE Access* 8:92352–92362
- Mahendran G, Dhanasekaran R (2015a) Investigation of the severity level of diabetic retinopathy using supervised classifier algorithms. *Comput Electr Eng* 45:312–323
- Mahendran G, Dhanasekaran R (2015b) Detection and localization of retinal exudates for diabetic retinopathy. *J Biol Syst* 23(2):195–212
- Mansour RF (2018) Deep-learning-based automatic computer-aided diagnosis system for diabetic retinopathy. *Biomed Eng Lett* 8:41–57
- Mateen M, Wen J, Nasrullah N, Sun S, Hayat S (2020) Exudate detection for diabetic retinopathy using pre-trained convolutional neural networks. *Complexity* 2020:5801870
- Memari N, Ramli AR, Saripan MI, Mashohor S, Moghbel M (2019) Retinal blood vessel segmentation by using matched filtering and fuzzy c-means clustering with integrated level set method for diabetic retinopathy assessment. *J Med Biol Eng* 39:713–731
- Mongan J, Moy L, Kahn CE (2020) Checklist for artificial intelligence in medical imaging (CLAIM): a guide for authors and reviewers. *Radiology* 2:2

- Nagendran M, Chen Y, Lovejoy CA, Gordon AC, Komorowski M, Harvey H, Topol EJ, Ioannidis JPA, Collins GS, Maruthappu M (2020) Artificial intelligence versus clinicians: systematic review of design, reporting standards, and claims of deep learning studies. *BMJ* 368:689
- Nair AT, Muthuvel K (2019) Blood vessel segmentation and diabetic retinopathy recognition: an intelligent approach. *Comput Methods Biomech Biomed Eng* 8(2):169–181
- Natarajan S, Jain A, Krishnan R, Rogye A, Sivaprasad S (2019) Diagnostic accuracy of community-based diabetic retinopathy screening with an offline artificial intelligence system on a smartphone. *JAMA Ophthalmol* 137(10):1182–1188
- Nazir T, Irtaza A, Shabbir Z, Javed A, Akram U, Mahmood MT (2019) Diabetic retinopathy detection through novel tetragonal local OCTA patterns and extreme learning machines. *Artif Intell Med* 99:101695
- Nilashi M, Samad S, Yadegaridehkordi E, Alizadeh A, Akbari E, Ibrahim O (2019) Early detection of diabetic retinopathy using ensemble learning approach. *J Soft Comput Decis Support Syst* 6(2):12–17
- Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R, Glanville J, Grimshaw JM, Hróbjartsson A, Lalu MM, Li T, Loder E, Mayo-Wilson E, McDonald S, McGuinness LA, Stewart LA, Thomas J, Tricco AC, Welch VA, Whiting P, Moher D (2021) The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *BMJ* 372:n71. <https://doi.org/10.1136/bmj.n71>
- Pao SI, Lin HZ, Chien KH, Tai MC, Chen JT, Lin GM (2020) Detection of diabetic retinopathy using bichannel convolutional neural network. *J Ophthalmol* 2020:9139713
- Partovi M, Rasta SH, Javadzadeh A (2016) Automatic detection of retinal exudates in fundus images of diabetic retinopathy patients. *J Anal Res Clin Med* 4(2):104–109
- Pires R, Avila S, Wainer J, Valle E, Abramoff MD, Rocha A (2019) A data-driven approach to referable diabetic retinopathy detection. *Artif Intell Med* 96:93–106
- Pratheeba C, Singh NN (2019) A novel approach for detection of hard exudates using random forest classifier. *J Med Syst* 43(7):180
- Prentašić P, Lončarić S (2016) Detection of exudates in fundus photographs using deep neural networks and anatomical landmark detection fusion. *Comput Methods Programs Biomed* 137:281–292
- Qiao L, Zhu Y, Zhou H (2020) Diabetic retinopathy detection using prognosis of microaneurysm and early diagnosis system for non-proliferative diabetic retinopathy based on deep learning algorithms. *IEEE Access* 8:104292–104302
- Quellec G, Charriere K, Boudi Y, Cochener B, Lamard M (2017) Deep image mining for diabetic retinopathy screening. *Med Image Anal* 39:178–193
- Qummar S, Khan FG, Shah S, Khan A, Shamshirband S, Rehman ZU, Khan IA, Jadoon W (2019) A deep learning ensemble approach for diabetic retinopathy detection. *IEEE Access* 7:150530–150539
- Rajalakshmi R, Subashini R, Anjana RM, Mohan V (2018) Automated diabetic retinopathy detection in smartphone-based fundus photography using artificial intelligence. *Eye* 32(6):1138–1144
- Raju M, Pagidimarri V, Barreto R, Kadam A, Kasivajjala V, Aswath A (2017) Development of a deep learning algorithm for automatic diagnosis of diabetic retinopathy. *Stud Health Technol Inf* 245:559–563
- Ramachandran N, Chiong HS, Sime MJ, Wilson GA (2018) Diabetic retinopathy screening using deep neural network. *Clin Exp Ophthalmol* 46(4):412–416
- Raman R, Srinivasan S, Virmani S, Sivaprasad S, Rao C, Rajalakshmi R (2019) Fundus photograph-based deep learning algorithms in detecting diabetic retinopathy. *Eye* 33:97–109
- Rami HE, Barham R, Sun JK, Silva PS (2017) Evidence-based treatment of diabetic retinopathy. *Semin Ophthalmol* 32(1):67–74
- Randive SN, Rahulkar AD, Senapati RK (2018) LVP extraction and triplet-based segmentation for diabetic retinopathy recognition. *Evol Intel* 11(1–2):117–129
- Rogers TW, Gonzalez-Bueno J, Franco RG, Star EL, Marín DM, Vassallo J, Lansingh VC, Trikha S, Jaccard N (2021) Evaluation of an AI system for the detection of diabetic retinopathy from images captured with a handheld portable fundus camera: the MAILOR AI study. *Eye* 35:632–638
- Sabanayagam C, Banu R, Chee ML, Lee R, Wang YX, Tan G, Jonas JB, Lamoureux EL, Cheng CY, Klein BEK, Mitchell P, Klein R, Cheung CMG, Wong TY (2018) Incidence and progression of diabetic retinopathy: a systematic review. *Lancet Diabetes Endocrinol* 7(2):140–149
- Sahlsten J, Jaskari J, Kivinen J, Turunen L, Jaanio E, Hietala K, Kaski K (2019) Deep learning fundus image analysis for diabetic retinopathy and macular edema grading. *Sci Rep* 9:10750
- Saleh E, Błaszczczyński J, Moreno A, Valls A, Romero-Aroca P, Riva-Fernández SDL, Słowiński R (2018) Learning ensemble classifiers for diabetic retinopathy assessment. *Artif Intell Med* 85:50–63
- Samanta A, Saha A, Satapathy SC, Fernandes SL, Zhang YD (2020) Automated detection of diabetic retinopathy using convolutional neural networks on a small dataset. *Pattern Recogn Lett* 135:293–298

- Sandhu HS, Eladawi N, Elmogy M, Keynton R, Helmy O, Schaal S, El-Baz A (2018a) Automated diabetic retinopathy detection using optical coherence tomography angiography: a pilot study. *Br J Ophthalmol* 102(11):1564–1569
- Sandhu HS, Eltanboly A, Shalaby A, Keynton RS, Schaal S, El-Baz A (2018b) Automated diagnosis and grading of diabetic retinopathy using optical coherence tomography. *Invest Ophthalmol vis Sci* 59:3155–3160
- Sandhu HS, Elmogy M, Sharafeldeen AT, Elsharkawy M, El-Adawy N, Eltanboly A, Shalaby A, Keynton R, El-Baz A (2020) Automated diagnosis of diabetic retinopathy using clinical biomarkers, optical coherence tomography, and optical coherence tomography angiography. *Am J Ophthalmol* 216:201–206
- Seth S, Agarwal B (2018) A hybrid deep learning model for detecting diabetic retinopathy. *J Stat Manag Syst* 21(4):569–574
- Shaban M, Ogur Z, Mahmoud A, Switala A, Shalaby A, Khalifeh HA, Ghazal M, Fraiwan L, Giridharan G, Sandhu H, El-Baz AS (2020) A convolutional neural network for the screening and staging of diabetic retinopathy. *PLoS ONE* 15(6):e0233514
- Shah P, Mishra DK, Shanmugam MP, Doshi B, Jayaraj H, Ramanjulu R (2020) Validation of deep convolutional neural network-based algorithm for detection of diabetic retinopathy: artificial intelligence versus clinician for screening. *Indian J Ophthalmol* 68(2):398–405
- Shankar K, Sait ARW, Guptac D, Lakshmanaprabu SK, Khanna A, Pandey HM (2020a) Automated detection and classification of fundus diabetic retinopathy images using synergic deep learning model. *Pattern Recogn Lett* 133:210–216
- Shankar K, Zhang Y, Liu Y, Wu L, Chen CH (2020b) Hyperparameter tuning deep learning for diabetic retinopathy fundus image classification. *IEEE Access* 8:118164–118173
- Simonyan K, Zisserman A (2015) Very deep convolutional networks for large-scale image recognition. In *Proceedings of the 3rd international conference on learning representations (ICLR) 2015, San Diego*
- Son J, Shin JY, Kim HD, Jung KH, Park KH, Park SJ (2019) Development and validation of deep learning models for screening multiple abnormal findings in retinal fundus images. *Ophthalmology* 127(1):85–94
- Sorour O, Arya M, Waheed N (2018) New findings and challenges in OCT angiography for diabetic retinopathy. *Ann Eye Sci* 3(8):44
- Srinivasan PP, Kim LA, Mettu PS, Cousins SW, Comer GM, Izatt JA, Farsiu S (2014) Fully automated detection of diabetic macular edema and dry age-related macular degeneration from optical coherence tomography images. *Biomed Opt Express* 5(10):3568–3577
- Srivastava V, Purwar RK (2020) Classification of eye-fundus images with diabetic retinopathy using shape based features integrated into a convolutional neural network. *J Inf Optim Sci* 41(1):217–227
- Stitt AW, Curtis TM, Chen M, Medina RJ, McKay GJ, Jenkins A, Gardiner TA, Lyons TJ, Hammes HP, Simó R, Lois N (2016) The progress in understanding and treatment of diabetic retinopathy. *Prog Retin Eye Res* 51:156–186
- Stolte S, Fang R (2020) A survey on medical image analysis in diabetic retinopathy. *Med Image Anal* 64:101742
- Sun Y (2019) The neural network of one-dimensional convolution: an example of the diagnosis of diabetic retinopathy. *IEEE Access* 7:69657–69666
- Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A (2015) Going deeper with convolutions. In: *Proceedings of the 2015 IEEE conference on computer vision and pattern recognition, Boston*
- Țălu Ș, Călugăru DM, Lupașcu CA (2015) Characterization of human non-proliferative diabetic retinopathy using the fractal analysis. *Int J Ophthalmol* 8(4):770–776
- Tan JH, Fujita H, Sivaprasad S, Bhandary SV, Rao AK, Chua KC, Acharya UR (2017a) Automated segmentation of exudates, haemorrhages, microaneurysms using single convolutional neural network. *Inf Sci* 420:66–76
- Tan JH, Acharya UR, Bhandary SV, Chua KC, Sivaprasad S (2017b) Segmentation of optic disc, fovea and retinal vasculature using a single convolutional neural network. *J Comput Sci* 20:70–79
- The ACCORD Study Group and ACCORD Eye Study Group (2010) Effects of medical therapies on retinopathy progression in type 2 diabetes. *N Engl J Med* 363(3):233–244
- Ting DSW, Cheung CY, Lim G, Tan GSW, Quang ND, Gan A, Hamzah H, Garcia-Franco R, Yeo IYS, Lee SY, Wong EYM, Sabanayagam C, Baskaran M, Ibrahim F, Tan NC, Finkelstein EA, Lamoureux EL, Wong IY, Bressler NM, Sivaprasad S, Varma R, Jonas JB, He MG, Cheng CY, Cheung GCM, Aung T, Hsu W, Lee ML, Wong TY (2017) Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. *JAMA* 318(22):2211–2223

- Ting DSW, Pasquale LR, Peng L, Campbell JP, Lee AY, Raman R, Tan GSW, Schmetterer L, Keane PA, Wong TY (2018) Artificial intelligence and deep learning in ophthalmology. *Br J Ophthalmol* 103(2):167–175
- Torre J, Valls A, Puig D (2020) A deep learning interpretable classifier for diabetic retinopathy disease grading. *Neurocomputing* 396:465–476
- Usman I, Almejalli KA (2020) Intelligent automated detection of microaneurysms in fundus images using feature-set tuning. *IEEE Access* 8:65187–65196
- Vujosevic S, Aldington SJ, Silva P, Hernández C, Scanlon P, Peto T, Simó PR (2020) Screening for diabetic retinopathy: new perspectives and challenges. *Lancet Diabetes Endocrinol* 8(4):337–347
- Wan S, Liang Y, Zhang Y (2018) Deep convolutional neural networks for diabetic retinopathy detection by image classification. *Comput Electr Eng* 72:274–282
- Wang Z, Camino A, Zhang M, Wang J, Hwang TS, Wilson DJ, Huang D, Li D, Jia Y (2017) Automated detection of photoreceptor disruption in mild diabetic retinopathy on volumetric optical coherence tomography. *Biomed Opt Express* 8(12):5384–5398
- Wang XN, Dai L, Li ST, Kong HY, Sheng B, Wu Q (2020) Automatic grading system for diabetic retinopathy diagnosis using deep learning artificial intelligence software. *Curr Eye Res* 45(12):1550–1555
- Wardoyo S, Pramudyo AS, Rizanti ED, Muttakin I (2016) Exudate and blood vessel feature extraction in diabetic retinopathy patients using morphology operation. *Telkomnika* 14(4):1493–1501
- Wei Y, Gong J, Xu Z, Thimmulappa RK, Mitchell KL, Welsbie DS, Biswal S, Duh EJ (2015) Nrf2 in ischemic neurons promotes retinal vascular regeneration through regulation of semaphorin 6A. *Proc Natl Acad Sci USA* 112(50):e6927–e6936
- Wong TY, Bressler NM (2016) Artificial intelligence with deep learning technology looks into diabetic retinopathy screening. *JAMA Ophthalmol* 316(22):2366–2367
- Wu Z, Shi G, Chen Y, Shi F, Chen X, Li S, Coatrieux G, Yang J, Luo L (2020) Coarse-to-fine classification for diabetic retinopathy grading using convolutional neural network. *Artif Intell Med* 108:101936
- Xie Y, Nguyen QD, Hamzah H, Lim G, Bellemo V, Gunasekeran DV, Yip MYT, Lee XQ, Hsu W, Lee ML, Tan CS, Wong HT, Lamoureux EL, Tan GSW, Wong TY, Finkelstein EA, Ting DSW (2020) Artificial intelligence for teleophthalmology-based diabetic retinopathy screening in a national programme: an economic analysis modelling study. *Lancet Dig Health* 2(5):e240–e249
- Xu K, Feng D, Mi H (2017) Deep convolutional neural network-based early automated detection of diabetic retinopathy using fundus image. *Molecules* 22(12):2054
- Yip MYT, Lim G, Lim ZW, Nguyen QD, Chong CCY, Yu M, Bellemo V, Xie Y, Lee XQ, Hamzah H, Ho J, Tan TE, Sabanayagam C, Grzybowski A, Tan GSW, Hsu W, Lee ML, Wong TY, Ting DSW (2020) Technical and imaging factors influencing performance of deep learning systems for diabetic retinopathy. *NPJ Dig Med* 3(40):1–12
- Zago GT, Andraeo RV, Dorizzi B, Salles EOT (2020) Diabetic retinopathy detection using red lesion localization and convolutional neural networks. *Comput Biol Med* 116:103537
- Zahran MA, Damrawi GE, Amin E, Abdelsalam MM (2020) Enforcing artificial neural network in the early detection of diabetic retinopathy OCTA images analysed by multifractal geometry. *J Taibah Univ Sci* 14(1):1067–1076
- Zeng XL, Chen HQ, Lou Y, Ye WB (2019) Automated diabetic retinopathy detection based on binocular Siamese-like convolutional neural network. *IEEE Access* 7:30744–30753
- Zhang J, Chen Y, Bekkers E, Wang M, Dashtbozorg B, Romeny BMH (2017) Retinal vessel delineation using a brain-inspired wavelet transform and random forest. *Pattern Recogn* 69:107–123
- Zhang W, Zhong J, Yang S, Gao Z, Hu J, Chen Y, Yi Z (2019) Automated identification and grading system of diabetic retinopathy using deep neural networks. *Knowl Based Syst* 175:12–25

Authors and Affiliations

Ganeshsree Selvachandran¹ · Shio Gai Quek¹ · Raveendran Paramesran² · Weiping Ding³ · Le Hoang Son⁴ 

Ganeshsree Selvachandran
Ganeshsree@ucsiuniversity.edu.my

Shio Gai Quek
queksg@ucsiuniversity.edu.my

Raveendran Paramesran
ravee@ucsiuniversity.edu.my

¹ Department of Actuarial Science and Applied Statistics, Faculty of Business & Management, UCSI University, Jalan Menara Gading, Cheras, 56000 Kuala Lumpur, Malaysia

² Institute of Computer Science and Digital Innovation, UCSI University, Jalan Menara Gading, Cheras, 56000 Kuala Lumpur, Malaysia

³ School of Information Science and Technology, Nantong University, Nantong 226019, People's Republic of China

⁴ VNU Information Technology Institute, Vietnam National University, Hanoi, Vietnam