# A hybrid convolutional neural network model for detection of diabetic retinopathy

Alshawabkeh, Musa; Ryalat, Mohammad Hashem; Dorgham, Osama M.; Alkharabsheh, Khalid; Btoush, Mohammad Hjouj; Alazab, Mamoun

*Published in:*
International Journal of Computer Applications in Technology

*DOI:*
10.1504/IJCAT.2022.130886

Published: 13/05/2023

*Document Version*
Peer reviewed version

# A hybrid convolutional neural network model for detection of diabetic retinopathy

Musa Alshawabkeh,
Mohammad Hashem Ryalat*
and Osama M. Dorgham

Computer Science Department,
Al-Balqa Applied University,
Salt, Jordan
Email: musa.ams@outlook.com
Email: ryalat@bau.edu.jo
Email: o.dorgham@bau.edu.jo

Khalid Alkharabsheh*

Software Engineering Department,
Al-Balqa Applied University,
Salt, Jordan
Email: khalidkh@bau.edu.jo
*Corresponding authors

Mohammad Hjouj Btoush

Computer Science Department,
Al-Balqa Applied University,
Salt, Jordan
Email: hujooj@bau.edu.jo

Mamoun Alazab

College of Engineering, IT and Environment,
Charles Darwin University,
Darwin, Northen Territory, Australia
Email: alazab.m@ieee.org

**Abstract:** Diabetic retinopathy causes vision loss. Regular eye screening has to be done to provide the appropriate treatment and for vision loss prevention. Globally, patients with DR are increasing, which leads to work pressure on specialists and equipment. Fundus images are a key factor in effective retinal diagnosis. In this paper, a deep-learning approach is proposed to detect DR from retinal images. The proposed approach involves a combination of four effective techniques: image augmentation, contrast limited adaptive histogram equalisation, CNN and transfer learning and ensemble classification. The results show the proposed approach obtained high values of accuracy (93%), precision (95%) and recall (96%), and more stability compared with other approaches.

**Keywords:** deep learning; diabetic retinopathy; eye diseases; retinal diagnosis; retinal images; convolutional neural networks; medical applications; ensemble classification.

**Reference** to this paper should be made as follows: Alshawabkeh, A., Ryalat, M.H., Dorgham, O.M., Alkharabsheh, K., Btoush, M.H. and Alazab, M. (xxxx) 'A hybrid convolutional neural network model for detection of diabetic retinopathy', *Int. J. Computer Applications in Technology*, Vol. XX, No. X, pp.xxx–xxx.

**Biographical notes:** Musa Alshawabkeh received his MSc degree in Computer Science from Al-Balqa Applied University in 2021. His research interests include deep learning, artificial intelligence, machine learning and medical images. Currently, he is a Computer Science Lecturer and working with different research teams.

Mohammad Hashem Ryalat was appointed as a Lecturer at Al-Balqa Applied University (BAU) in September 2005 and won a British Council competition to pursue his PhD in 2014 at UEA in

the UK. He received the PhD degree in 2017. He was appointed as the Head of Computer Information Systems Department from 2018 to 2020 at BAU. His research interests include biomedical engineering, machine learning, optimisation, computer vision and computerised medical applications. Currently, he is an Assistant Professor and working with different research teams and committees.

Osama M. Dorgham received his BSc degree in Computer Science from Princess Sumaya University for Technology, Jordan, the MSc degree in Computer Science from Al-Balqa Applied University, Jordan and the PhD degree in Computing Sciences from the University of East Anglia, Norwich, UK. His research interests include artificial intelligence, image processing, parallel processing, computer security and computer graphics. He is an Active Member in many academic and industrial organisations; in addition, he serves as a Member in many international scientific journals. He has been awarded Erasmus grants (Post-Doc and Research Staff) and international awards during the past few years. He has appointed as the Head of Computer Information Systems science 2015–2018.

Khalid Alkharabsheh received his BSc and MSc degrees in Computer Science from Yarmouk University and Al-Balqa Applied University, Jordan in 2002 and 2005, respectively. He was appointed as a Lecturer at Al-Balqa Applied University (BAU) in September 2006 and won an Erasmus Mundus grant to pursue his PhD degree in 2014 at Santiago de Compostela University in Spain and was awarded his PhD degree in 2019. He was appointed as an Assistant Professor in the Software Engineering Department in 2019 and the Head of the Department from 2021 to date. His research interests include machine learning, big data, software quality, empirical software engineering, software validation and verification and design smell detection. Currently, he is an Assistant Professor and working with different research teams and committees.

Mohammad Hjouj Btoush have joined Al-Balqa Applied University-Jordan as a Lecturer of Computer Science in 1992. During that time, he co-authored a book: *Computer Skills*. In 2006, he joined the Informatics Research Group at Sheffield Hallam University-UK. His PhD thesis investigates the perception of users and providers of e-services in *Jordan/Artificial Intelligence & Software Engineering*. His current research interests include: information security, information systems evaluation, specifically, within the context of electronic government, operating system and software engineering.

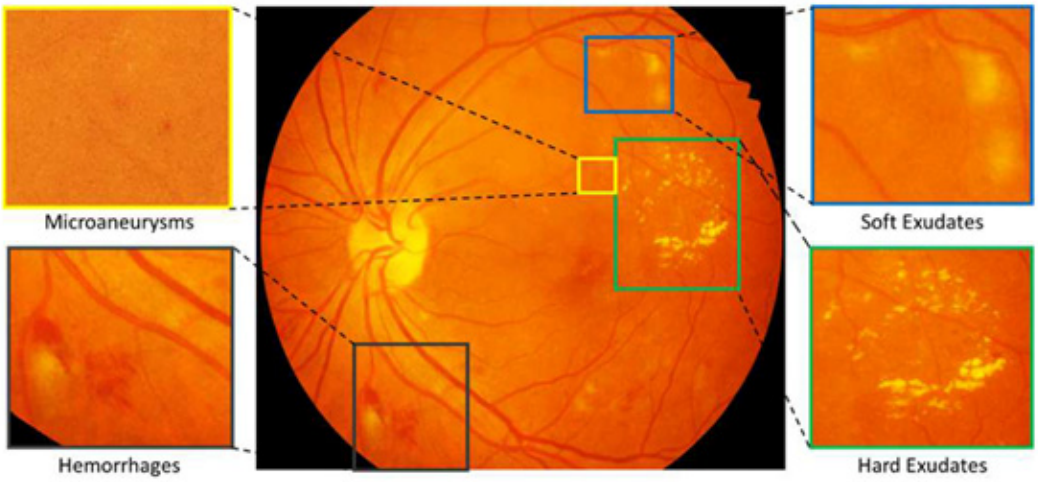
Mamoun Alazab is an Associate Professor in the College of Engineering, IT and Environment at Charles Darwin University, Australia. He received his PhD degree in Computer Science from the Federation University of Australia, School of Science, Information Technology and Engineering. He is a Cyber Security Researcher and Practitioner with Industry and Academic Experience. He is multidisciplinary that focuses on cyber security with a focus on cybercrime detection and prevention. He has more than 300 research papers and 9 authored/edited books. He is a Senior Member of the IEEE. He is the Founding Chair of the IEEE Northern Territory (NT) Subsection.

## 1 Introduction

There are many diseases that can infect the human eye and negatively affect vision capabilities, such as macular degeneration, cataracts, Diabetic Retinopathy (DR) (Litjens et al., 2017; Reyana et al., 2020), glaucoma, and other common eye disorders including amblyopia and strabismus.

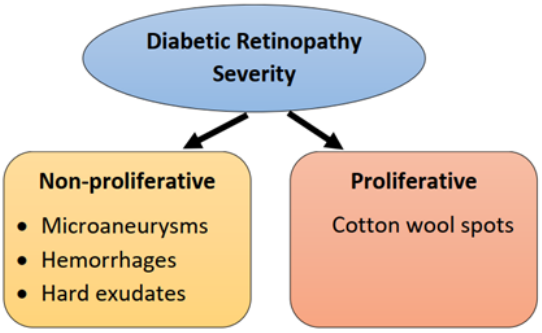
Diabetes is a health condition that is caused by a high level of blood sugar (glucose). Diabetes retinopathy can occur in individuals who have diabetes and is a complication of both types of diabetes, types 1 and 2 (Ting et al., 2016). The high level of blood pressure and high level of glucose damage blood vessels in the retina. This eventually causes blood and fat fluids to leak out into the retina, which leads to a decline in vision. Microaneurysms, haemorrhages, hard exudates and so-called 'cotton wool spots' are the main symptoms of DR. The disease development process is as follows: First, the retina receives its food needs from blood. However, when the amounts of glucose and blood pressure change to abnormal levels due to diabetes, tiny capillaries dilate to form red spots on the surface of the retina, and these red spots are called microaneurysms (Qiao et al., 2020). Over a period of time and the presence of continued high-blood pressure, the risk of damage increases and haemorrhages appear when those capillaries burst. The next symptom that appears is hard exudates. These occur when blood and protein leak out of the haemorrhages onto the retinal surface, this reduces the amount of food received by the retina. Hence, the retina needs to grow new tiny capillaries to obtain food, and these capillaries appear as cotton wool spots. Figure 1 illustrates the symptoms of DR in a series of retinal images.

**Figure 1** Diabetic retinopathy symptoms




The severity of DR determines the level of risk of loss of vision. Diabetes retinopathy is divided into two main types: Non-Proliferative Diabetic Retinopathy (NPDR) and Proliferative Diabetic Retinopathy (PDR) where the severity is more advance (Kaur et al., 2021). Each type has is indicated by the presence of particular symptoms. Microaneurysms, haemorrhages and hard exudates are the symptoms of NPDR, while cotton wool spots is the symptom of PDR, as illustrated in Figure 2.

**Figure 2** Diabetic retinopathy types and symptoms




Hence the symptoms of DR are not noticeable in the early stages, which is a barrier to treatment interventions as DR is a progressive disease that worsens with time. Thus, the early detection of DR could provide an advantage by enabling the more effective treatment of infected retinas. The treatment of DR requires the control of the level of diabetes as keeping blood sugar levels close to normal can slow down disease progression. Therefore, it is recommended that people undergo regular examinations for DR so that they can receive the appropriate medical care and treatment, and the severity level can be controlled.

Various manual methods are used to detect and diagnosis DR, such as the Dilated Eye Examination (DEE), fundus photography, fluorescence angiography and optical coherence tomography (Flaxel et al., 2020). These methods must be done by professionals, and they need to be scheduled into far apart appointments. Since time is a major factor in the progress of DR, the early detection of DR in an eye allows a patient to have effective treatments for the prevention of sight loss. Therefore, an automated screening system could help to shorten the time taken to diagnose DR, as well as lower the financial cost and to decrease the need for specialists who can free up their time to undertake other work.

Optimisation algorithms were heavily employed in medical field (Ryalat et al., 2016, 2017b). On the other hand, Deep Learning (DL) is a subcategory of Machine Learning (ML) which relies on a sequence of layers, where each layer extracts information from input data, then feeds that information as input data to the next layer. The Convolutional Neural Network (CNN) is a class of DL that is used in various applications such as in natural language processing technology, finance (Chen et al., 2016), satellite imagery and medical images (Kannan et al., 2021; Fisher et al., 2013). The CNN is a smart way to build a deep neural network that can take an image as input data and then assign a weight and bias to the features in the image (Dorgham et al., 2021; Ryalat et al., 2017a). Through this technique, the CNN can recognise objects in images and at the same time reduce the image pre-processing time. In addition, it has been shown that CNNs provide accurate results (Litjens et al., 2017).

The proposed approach in this study comprises DL, image processing, image augmentation, Transfer Learning (TL), a CNN and ensemble classification techniques. It is expected through the application of the proposed method over actual retinas that we can guarantee early detection of DR. Consequently, it will possible for people with retinal problems to obtain appropriate medical treatment in a timely manner. Furthermore, the proposed method will help to alleviate or prevent DR progression and blindness.

Regarding that and in attempts to employ DL for the diagnosis of DR, some previous studies focused on the features of DL and image processing to benefit from image characteristics. The following paragraphs present some of the recent studies that are related to this study, and how this study presents a contribution that was not discussed by the related studies

A set of studies addressed the detection of DR without separating images into three colour channels. A CNN of 13

layers was utilised in Saha et al. (2019), while pixels were normalised in Zeng et al. (2019). In a similar vein, pixels were subtracted from the mean of the surroundings, then normalised in Gharaibeh et al. (2017). In Li et al. (2019a), the colour was subtracted, while Reddy et al. (2019) used RGB images.

On the other hand, some related studies have not used deep learning technique whereas this study presents a framework which is mainly based on deep learning techniques. The author of Habib et al. (2017) proposed a tree ensemble classifier. Also, in Qummar et al. (2019), an ensemble DL model using five pre-trained CNNs was employed. In Kassani et al. (2019), an ensemble TL method with colour normalised images was proposed. In Li et al. (2019b), two independent neural networks were used. The TL method was also used in Zago (2019).

One of the main contributions in this study is the employment of colour channels. The images are split into three-colour channels. This will lead to three images each of has only one-colour channel. The green channel of RGB images has a higher contrast than the other channels. Therefore, Manjesh and Nagachandra (2016) depended on the green channel to calculate the value of (g) as

$$g = \frac{G}{R + G + B} \tag{1}$$

On the other hand, Jaya et al. (2019) omitted the green channel from pre-processing because it shows high contrast intensity compared to the red and blue channels. Meanwhile, Foeady et al. (2018) extracted the red, green and blue channels from RGB images. The green channel was used for classification, and the red and blue channels were used to filter out noise for contrast enhancement.

In RGB fundus images, the red channel has higher numeric values than the other two channels, which gives it an influence in classification, while the green channel has higher contrast. On the other hand, pixel normalisation is used to reduce the computation cost in the training and testing phases.

As presented in the previous paragraphs, the uniqueness of the work in this study that It used CLAHE on each colour channel separately, unlike the other studies in the literature that combine the colours after the application of CLAHE to one image.

This paper is divided into six sections. Section 2 illustrates the significance of this study. Section 3 presents an overview of the problem and discusses the various recent research studies in the field of DR detection. Section 4 describes and explains the methodology of the proposed framework developed in the study. Section 5 details the experimental results and evaluation of the developed framework. Finally, Section 6 concludes the paper and suggests some directions for future work.

## 2 Significance of the study

One way in which an eye can be checked for DR is through the use of the DEE, which is a simple examination method without pain. Eye drops are administered into the eye to widen and increase the size of the pupil. Then the inner eye is checked for DR and for other various eye problems. Even though the DEE is painless, it can cause vision to become blurry. Also, the eye becomes more sensitive to light for few hours after the examination. Therefore, patients may feel uncomfortable. The risk of vision loss due to DR becomes higher with the passage of time, and DR itself becomes more difficult to treat. However, DR risk can be reduced through blood sugar management, blood pressure management, and through lowering the amount of lipids in the blood. Therefore, the early diagnosis of DR and the appropriate treatment of DR could reduce the risk of vision loss. An automated system would be helpful in examining retinal images to aid DR diagnosis. Such a system would receive retinal images in digital form. It would interpret these images to analyse their content and come up with relevant features. Then these extracted features would be used to identify the presence of the symptoms of DR in other images. Thus, the system would be able to predict the presence or otherwise of DR in the retina. It has been shown that automated eye screening systems are cost-effective (Fenner et al., 2018). In addition, this type of system can help to reduce the workload of ophthalmologists because its operation does not require the involvement of ophthalmologists and specialist graders; rather, a trained non-ophthalmologist would be capable of operating the system effectively.

The number of DR cases is growing and, consequently, the load on medical specialists is increasing due to the manual effort required. Hence it is difficult for medical specialists to provide optimal medical care (Bhatt et al., 2020). Thus, an alternative reliable method to diagnose eye disease and predict DR is needed in order to reduce the demand on human resources. Therefore, in this paper, we propose a DL framework for DR detection.

In this study, we used an ensemble TL classification model of four pretrained CNNs. These CNNs are well known for their high classification accuracy (Al-Hiary et al., 2011). Three of the CNNs were trained on the colour channels, where each CNN was trained on a separate colour channel. Therefore, each colour component in the image has the opportunity to play an independent role in training and testing, and to enable the achievement of a highly accurate classification result. The fourth CNN was trained on gray-scale in order to balance the classification results of the RGB channels, and to strengthen the final classification result.

As an impact of the above-stated motivation, the purpose of this study is to provide a tested, effective, appropriate and applicable DL framework that has the capability to investigate DR in a given colour fundus image. This will be done by utilising DL and image processing techniques to diagnose and analyse retinal images and to extract the desired information. The information will then be processed to enable the prediction of a result of classification probability, and thus determine whether an eye has DR or not.

A colour fundus image contains information that is integral to the diagnosis of problems in the retina (Wedyan et al., 2020). This information provides clues as to the condition of the retina and whether a retina has DR. This motivates the researchers to find answers to the following research questions:

- Does the use of image augmentation techniques enhance the accuracy of the detection and classification of DR?
- Is it beneficial to employ TL techniques to reuse a previously trained model for use in DR detection and classification?
- Does the use of image processing techniques lead to better performance in terms of the accuracy of the detection and classification of DR? g
- What is the impact of applying ensemble classification approaches to DR detection systems?

## 3 Background and literature review

In DR detection, images are classified into two classes. The first class denotes that an image of retina that has no DR symptoms, while the other class indicates that the image contains at least one of the DR symptoms.

The CNN can be used to analyse and classify fundus images. The CNN takes the input image and reads the numerical values of the pixels. It then applies mathematical calculations in its convolutional layers to generate new values. The new values are used to create a feature map at a higher level of abstraction. The extracted features are used to estimate the probability of the symptoms being present in the image, and thus to derive a correct prediction. This section presents an overview of recent research works of relevance to this study, which were done in the field of automated DR detection.

The author of Lands et al. (2020) developed a DL model to predict the level of DR in an input image. The images were gathered by a merge of two data sets to obtain more enough fundus images for training. The total number of images was of 23,303 images. The images were graded from 1 to 4 as each grade indicates the type of DR, where 0 for No DR, 1 Mild, 2 Moderate, 3 Severe and 4 for Proliferative DR. In order to improve the model capabilities, the author applied four modifications on images for image augmentation. Those modifications were Flipping which creates a mirror image, rotation to generate images of different angles, brightness and contrast. The model pre-processed the images by resizing them into 256×256 to reduce training time, then performed Gaussian blur subtraction to enhance the features of images. After pre-processing the images, the model trained on four CNNs; ResNet50, DenseNet121, DenseNet169 and Densenet256. Each CNN trained separately for three epochs. DenseNet121 gave the best results on Kaggle test set as 80.02 kappa score, 0.89 validation accuracy, 0.93 training accuracy and 0.78 accuracy.

Huang et al. (2020) proposed a method to detect haemorrhages in retina fundus images. Detection was done by a TL CNN with bounding box refining network. The CNN was based on ResNet50. Furthermore, the network of bounding box refining was trained and validated on 80 images. The images contained haemorrhages from IDRiD data set of fundus images. Later, the entire method was trained on a private data set of 590 fundus images, as 500 images for training and 90 manually bounding box refined images for testing. The CNN with pretrained parameters in was used for detecting objects in retinal images. The parameters were trained on COCO data set. These trained parameters were employed in the CNN with label smoothing to correct miss-detected and wrong labelled haemorrhages. All images were pre-processed by Contrast Limited Adaptive Histogram Equalisation (CLAHE) to enhance poor contrast. Also, images were pre-processed by adaptive gamma correction with weighting distribution for illumination. The author used mean average precision as a measure, and scored (52.20). A CNN model of modified pretrained VGG-16 was determined to predict DR severity level in Thota and Reddy (2020). The last block of VGG-16 was made trainable, and the dense layers were replaced with 128 neurons and ReLU. In addition, batch normalisation, dropout, learning rate and image augmentation were utilised for overfitting prevention. The image data set was obtained from Kaggle data set, which were classified into five classes refer to DR severity, i.e., No DR, Mild DR, Moderate DR, Sever DR and Proliferative DR. The images are a total of 35,126, and were pre-processed by resizing into 512 × 512 for better quality and higher accuracy of classification. The data set was split into training set of 1000 images and 34,126 images for evaluating. Training images were augmented randomly by rotating from 0 to 360, flipping horizontally or vertically, and brightness, hue, contrast and saturation change. The presented CNN achieved 0.74 average class accuracy, 0.8 sensitivity, 0.65 specificity and 0.8 Area Under the Curve (AUC).

Pour et al. (2020) proposed an EfficientNet-B5 architecture CNN with fundus images. The CNN trained on data set of 2948 images, these images were gained from three data sets; Messidor, Messidor-2 and IDRiD. CLAHE was applied to make vessels being displayed better in images, by improving contrast and pixels' intensities equalisation. Because of the different grading of the images in the data sets, the author divided the images in two groups of severity, then the EfficientNet-B5 CNN was trained and evaluated twice. At first, the CNN trained on Messidor-2 and IDRiD and evaluated on Messidor data set. The yelled results were as 0.92 sensitivity and 0.945 AUC. Then, the CNN trained on Messidor and Messidor-2 and evaluated on IDRiD. The obtained was 0.932 and AUC 0.796.

Kajan et al. (2020) compared the accuracy of an implemented method of TL from three independent pretrained CNNs, ResNet-50, Inceptionv3 and VGG-16. The CNNs were tested on a data set of fundus images based on EyePACS Diabetic Retinopathy data set. The data set of EyePACS contains 35,126 images graded into five grades of DR damage. Images were divided into two sets, training and testing. Training set had 25,790 images, the author found that some classes don't have enough image for suitable training, therefore, mirroring was applied on right and left eye retina images to increase the size of the data set to 49,262 images, then divided for 20% test and 80% train data sets.

The presented approach in Qummar et al. (2019) to detect DR, was based on ensemble DL using five pre-trained CNNs (Xception, Resnet50, Inceptionv3, Dense121 and Dense169). Kaggle training data set was demanded to train and test the approach. The data set contains 35,126 images of size 3888 × 2951. The images are classified to normal, mild, moderate,

sever and PDR for DR severity. In preprocessing, the input images were resized to 786 × 512. Then, mean normalisation was performed to reduce training time. Image augmentation was performed on images of minor classes to increase the number of images. While for major classes, images have been removed to make a balance in the data set of classes representation. The data set was divided into three subsets, 64% for training, 20% for testing and 16% for validation. The achieved accuracy was 80.8%, recall 51.5%, specificity 86.72%, precision 63.85% and F1-score was 53.74%. LeNet as a simple structured CNN which has few parameters and is built of eight layers, i.e., an input layer, three convolution layers, two pooling layers, a fully connected layer and an output layer. In (Fan et al., 2019) the LeNet was modified to improve the detection of red lesions (Microaneurysms and haemorrhages) using CNN.

Verbraak et al. (2019) a study to determine diagnostic accuracy of detection of DR, for a device enhanced with DL. Individuals with diabetes type 2 were eye screened under certain strict conditions. Four pairs of images of the eye with different sites and camera settings were taken, each pair is of a macula centred image and an optic disc centred image. The camera was operated by experienced specialists, and eye dilation was performed upon the specialist recommendation. The results from the device were compared with a human reference standard grading, where images were graded into vision-threatening DR, more than mild DR and mild or more DR. Each image had to be graded independently by two experienced retina readers, and a retina experienced specialist to adjudicate disagreement. The DL algorithm was used with the device is based on lesion detection. The output was a numerical value from 0 to 1 to indicate the likelihood of having DR. The study was performed on 1,616 patients and showed the device achieved 100% sensitivity for visual threatening, 79.4% for more than mild, 97.8% specificity for visual threatening and 93.8% for more than mild.

Two images were the input of the model of Li et al. (2019a). Each image was pre-processed as rescaling to 540 × 540, then colour subtracting, and 10% of border was clipped. Then the image enters two CNNs that vary in the number of the layers. In addition of employing fractional max-pooling instead of the common max-pooling. Each CNN results 1 of 5 classess, both results are combined by Support Vector Machine (SVM) classifier.

In De La Torre et al. (2020), a DL method to grade the level of DR severity has been presented, based on scores assigned for pixels of the input image, and their distribution in the hidden layers. The scores were obtained using propagation through an activation function node, a batch normalisation node, a fully connected layer and a dropout layer. The obtained scores of the feature extracted of both eyes were linearly combined, then classified and normalised with SoftMax function. The presented method trained on EyePACS data set from Kaggle, where each input image was pre-processed by trimming the border of the background, then resized into 3 × 640 × 640. The input images passed through 7 blocks for feature extraction, each block contains 2 layers, each layer consists of a convolution of 3 × 3 with 1 × 1 stride and padding, batch normalisation and ReLU activation function. Also, a 2 × 2 max-pooling with 2 × 2 stride was applied between every two blocks. Then a 2 × 2 convolution and a 4 × 4 average-pooling. The presented model has been trained for 300 epochs, and gained +90% of sensitivity and specificity.

Zago (2019) proposed a DL model called it Multi Self-attention Network. The model generated maps of the features, which were extracted from the image using InceptionV3 model, then calculates the multi self-attention factor for weight update. At last, the model classifies the retina image with a convolution layer and a fully connected layer.

A DL model to diagnose fundus images and classify them was proposed in Zeng et al. (2019). The model was a CNN based on the similarity of one person's eyes. It took images of both right and left eyes as input, then pre-processed the images by applying down scaling to 299 × 299, subtract pixels from mean of surroundings, boundaries remove, pixels normalisation (–1, 1) and random transformation. After pre-processing, the images were fed to InceptionV3 CNN to extract features. The result for both eyes was the input to a fully connected layer to compare the similarity between both eyes.

An aggregated various well-known CNNs (Xception, ResNet50, InceptionV3 and MobileNet) method was proposed in Kassani et al. (2019), to improve the training efficiency and the classification of the severity of DR. Pre-processing was performed on images by image resize, min-pooling to enhance contrast and colour normalisation. Xception was employed to extract features from the input image. The aggregation of layers was for DR severity classification. The method trained and tested on Kaggle APTOS and gained 83% accuracy, 88% sensitivity and 87% specificity

Zeng et al. (2019) proposed a method detect and grade the severity of DR in fundus images. The method includes a CNN, image pre-processing and image augmentation. The CNN to detect DR and to grade DR severity in scale of 0 – 4, while pre-processing and augmentation were to enhance the ate set. The proposed CNN relied on 4 × 4 kernel and 18 layers. Pre-processing was performed to resize images and to remove boundaries. Image augmentation included rotation, shearing, flipping, zooming, cropping, translation and colour augmentation. The proposed CNN was trained on Kaggle EyePACS data set with sigmoid optimiser, and yield 98% sensitivity and 94% specificity.

Chakrabarty (2018) applied a DL approach of a CNN on a data set of 30 high resolution fundus images, to detect DR automatically. The data set was divided into two sub sets as 15 images for each, Healthy and Diabetic. The images were pre-processed in three phases. Initially, images were converted from RGB to gray-scale with respect to weight given to each colour channel as:

$$i = R*0.299 + G*0.587 + B*0.114 \quad (2)$$

The calculated (i) is the value of the desired pixel in the gray-scale image, and R, G and B are the values of the colours. Then, the images were resized to 1000 × 1000. Lastly, pixels

were divided by 255 to produce values between 1 and 0, for easier computation. The CNN has an input layer that receives gray-scale, then three sets of convolution layer with ReLU and max-pooling, a flattening layer and a fully connected layer. The data set was divided as 80% for training set, and 20% for testing set, with equal numbers of images of each class in the sets. Then, all of the training images were augmented to increase the number of training images by zoom in range of 0.2. Training was for 28 epochs and achieved 91.67% training accuracy, 100% validation accuracy, 100% sensitivity and 100% F1-score.

## 4 Proposed approach

In this section, the proposed approach for the detection of DR is presented. Basically, as illustrated in Figure 3, our approach reads an image of retina, then performs DR detection to give final result regarding the medical condition of the retina, i.e., whether DR is present or not.

The proposed model that was built for this study is shown in Figure 4, where each level in the diagram represents a phase of the model, and each phase goes through various steps.

**Figure 3** Detection of diabetic retinopathy in a fundus image

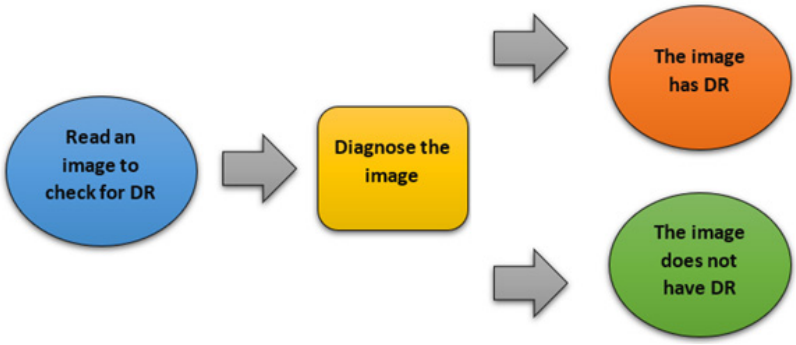
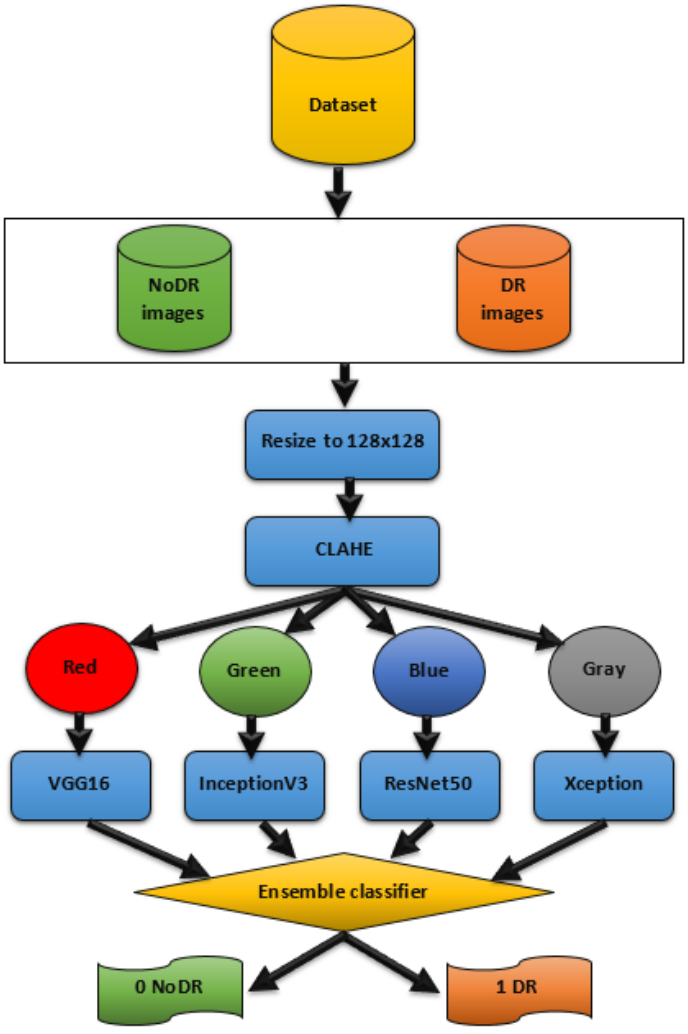



**Figure 4** Proposed model for diabetic retinopathy detection



### 4.1 Data set

The image data set we used for the training and testing of the model was gathered from the Standard Diabetic Retinopathy Database Calibration level 0 (DIARETDB0) data set. DIARETDB0 is a public database for benchmarking DR from digital images. The target is to define a database and a protocol which can be utilised to benchmark DR detection approaches. The data set can be downloaded freely (https://www.it.lut.fi/project/imageret/diaretdb0/) and can be used for scientific research purposes. DIARETDB0, which is widely used in the literature as a benchmark data set, contains a total of 130 retinal images. Figure 5 provides a sample of those images. The images were captured with a 50° field-of-view digital fundus camera, and unknown camera settings. The 130 images consist of 20 colour fundus images that are normal and 110 colour fundus images that contain symptoms of DR. These images are in png format and each image is of 1500 × 1152 resolution.

#### 4.1.1 Image pre-processing

The images we used for training and testing were of high resolution. However, the use of such images implies that there is a very high processing cost in terms of time. For this reason, we had to resize the images to an appropriate size to reduce the time cost while at the same time ensuring that it was still feasible to perform feature extraction on the resized images. Therefore, we resized the images into 128 pixels in width and 128 pixels in height.

To increase image quality, we used Contrast Limited Adaptive Histogram Equalisation (CLAHE) to improve the contrast and the visible surface of the dark parts of the image. The CLAHE method divides an image into small regions called tiles and then operates on them. Neighbouring tiles are combined by bilinear interpolation for intensity to improve image contrast (Pour et al., 2020). In order to apply CLAHE on the fundus images, we had to split the three-colour channels. This produced three images each of only one-colour channel, either red or green or blue. The next step was to apply CLAHE on the extracted one-colour channel images. Moreover, we recombined the resultant images into one main image to produce gray-scale images. Figure 6 shows the red, green, blue and gray-scale images after CLAHE.

In colour fundus images of retinas, the red channel has higher numeric values than the green and blue channels. Owing the higher red values in the pixels, red has an influence on the pixel computation. Therefore, to give each pixel the opportunity to play a role in DR detection, we employed an image for each colour channel in the classification method. We also included a gray-scale image to strengthen the roles of all the pixels and to balance classification.

The model proposed in this paper gives a pixel-based prediction. To avoid combining the processed values of the colour pixels, and to keep the values of each channel independent, we made various CNNs learn from each colour independently. We made use of a different CNN for each different colour channel. Moreover, the gray-scale images produced from the colour images were subjected to another different CNN. This allowed each pixel to affect the predicted results in respect to the gray-values.

**Figure 5** Sample from DIARETDB0 data set

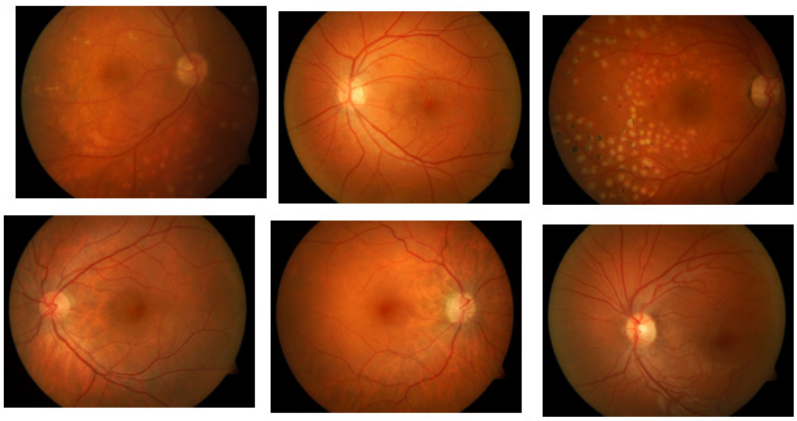
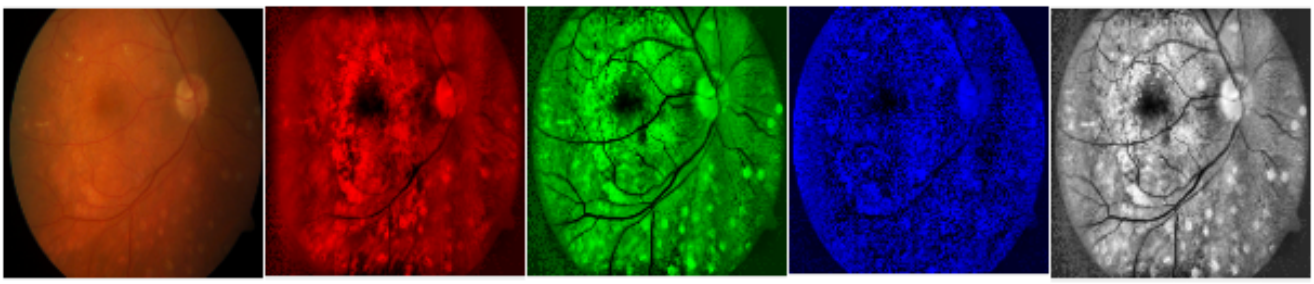


**Figure 6** CLAHE images



### 4.2 Transfer learning

Transfer learning is a technique used in ML to reuse a previously trained model on a task as a starting point for training another model on a related task. The aim of using TL is to improve the performance of the target model by transferring the learned knowledge to enable rapid learning (Zhuang et al., 2020). For our problem of DR detection in colour fundus images, we used TL from well-known CNNs. These CNNs have been trained on large data sets, and can take many days or weeks to train on powerful hardware. The TL method in our model consisted of the following three steps:

- *Step 1 (Selection of CNNs)*: We selected four CNN models – VGG16, ResNet50, InceptionV3 and Xception – all of which have proven high performance in feature extraction, image classification and object prediction. These CNNs were downloaded and incorporated directly into our model.
- *Step 2 (Reuse of CNNs)*: The four CNNs were employed as starting points in our model for the task of interest. It should be noted that this process did not involve the layers of the chosen CNNs being trained in our model.
- *Step 3 (Tuning of CNNs)*: We adapted the input layer of the four selected CNNs to match with the size and type of the input images in our model. We also refined the output by adding three fully connected layers, with SoftMax, ReLU and Sigmoid activations, one for each layer, respectively.

*VGG16*: VGG16 was proposed by Simonyan and Zisserman (2014) at the University of Oxford and published in Simonyan and Zisserman (2014) VGG16 is a well-known CNN model that achieved 92.7% test accuracy in ImageNet, which is a data set that contains more than 14 million images that belong to 1000 classes. VGG16 improved on AlexNet (Wang et al., 2018) by replacing the 11 × 11 kernel in the first convolutional layer of AlexNet with a 3 × 3 kernel, and the 5 × 5 kernel in the second layer with a 3 × 3 kernel. VGG16 consists of 16 weight layers, including 13 convolutional layers with filter size of 3 × 3 and 3 fully connected layers. The configurations of the fully connected layers in VGG16 are the same as those in AlexNet. The stride and padding are fixed to one pixel for all the convolutional layers. All of the convolutional layers are split into 5 groups, each of which is followed by one max-pooling layer.

*ResNet50*: ResNet50 is a powerful CNN that has shown high-performance results in image classification, and won many competitions in 2015. It took first place in the ImageNet detection competition, the ImageNet localisation competition, as well as the COCO detection and COCO segmentation competitions (He et al., 2016). The name ResNet is a short for Residual Networks, and the name is followed by a number that refers to the number of neural network layers used to build the architecture of the ResNet. There are various ResNet architectures based on the number of layers, and in this study, we utilised on one of the most vibrant ResNet architectures, ResNet50. The addition of more layers to the network is supposed to lead to higher accuracy values, and the ability to represent complex features. However, the stacking up of many more layers in a network load to the vanishing gradient problem. This adversely affects accuracy and accuracy degradation can be observed. The strength of ResNet lies in its ability to solve the vanishing gradient problem that occurs while training in deeper networks. ResNet introduced the concept of identity shortcuts between layers. Identity shortcuts connect the input directly to the output when they are of the same dimension, which transforms the network into parallel residual blocks. This mitigates the vanishing gradient problem by allowing the identity shortcut to flow the gradient unimpeded. Also, it ensures that the performance of the higher layers is as good if not better than that of the lower layers.

*InceptionV3*: The first Inception CNN was introduced by Szegedy et al. (2016) as a GoogLeNet deep CNN of 42 layers. In comparison to VGG networks, Inception has fewer parameters, better computation efficiency, lower error rate and similar complexity (Ryalat, 2022). There are various versions of Inception each of which represents a refinement of the architecture of the original Inception model. In InceptionV2, batch normalisation was introduced. In InceptionV3, factorisation was added in the third iteration with efficient grid size reduction. Factorisation aims to reduce the number of parameters for the whole network, without decreasing the efficiency of the network. With factorisation, the network can go deeper, and is less likely to suffer the problem of overfitting. To explain the factorisation effect, suppose we have a layer consisting of a 5 × 5 filter, then the number of parameters will be 5 × 5 = 25. However, with factorisation of the layer, if we use two layers of 3 × 3 filters, the number of parameters will be 3 × 3 + 3 × 3 = 18. Hence, the number of the parameters is reduced from 25 to 18, or by 28%. Efficient grid size reduction makes the network learn effectively with less cost. A CNN uses max-pooling to down-sample feature maps. However, the two main drawbacks of using max-pooling after a convolutional layer are that (1) it makes the network too greedy and (2) the network becomes more expensive. With effective grid size reduction, the max-pooling layer produces a set of (n) number of feature maps, and the convolutional layer produces a set of the same (n) number of feature maps. Then both are concatenated as 2 × n feature maps. After that, the feature maps pass to the next level of the network.

*Xception*: Xception was proposed in 2017 by Google as a modified version of InceptionV3 (Chollet, 2017). Xception stands for Extreme version of Inception. In Xception the order of the depthwise convolution and the pointwise convolution is modified. Depthwise convolution uses a number of convolutional layers that are equal to the convolution dimension. For example, suppose we have a 5 × 5 convolution, this means that there are five convolutional layers of 5 × 5. On the other hand, pointwise convolution involves a convolutional layer of 1 × 1. In Xception, depthwise convolution is used after pointwise convolution, instead of using depthwise convolution before pointwise convolution. Therefore, Xception is lighter than Inception in

term of computation cost. Another modification of Xception is that it is built without intermediate ReLU activation, thus Xception has high accuracy.

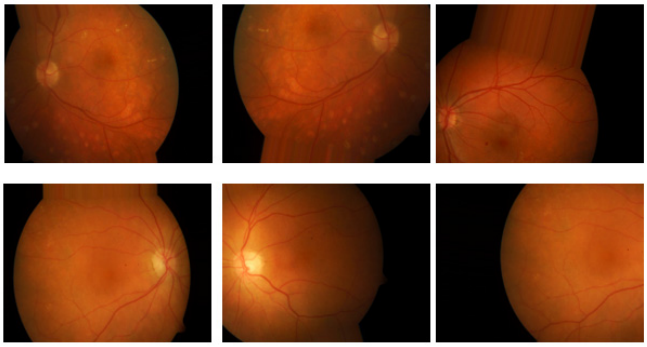
### 4.3 Image augmentation

Image augmentation is a powerful technique that creates variations of existing images in order to enlarge the data set. It modifies the images by using different processing methods, or a combination of multiple processing methods (Pour et al., 2020). In order to train a CNN, we need a large number of images to obtain good performance.

The data set for the model in this study contains 130 images. In order to build a powerful image classifier by using these images for training, image augmentation was required to boost the performance of deep networks. Therefore, we randomly applied various processing methods for image augmentation, including rotation in a range of 0 to 20, zoom to 0.15, width shift to 0.3, height shift to 0.3, as well as horizontal flip and nearest fill mode.

Using the pre-processing augmentation method, we created 32 augmented images of each original image. Thus, the number of images increased from 130 images to 4160 images of both classes. This method required a large space to store the augmented images, and more time to load the images into the model. A sample of the augmented images is provided in Figure 7.

**Figure 7** Augmented images



### 4.4 K-fold cross-validation

In order to train a model, the data set is split randomly into two parts, one for training and one for testing. This means that the model views the sets only once, and the model is not trained on all of the images in the data set. This means that the images of the test set are unseen by the model for training. *K*-fold cross-validation divides the total data set into *k* subsets in order to distribute the data in the main data set between the training and testing sets. The model is trained for *k* iterations, where on each iteration the model is trained on *k*-1 subsets and the last subset is used for testing. In the last iteration, the trained model is tested on all of the images in the main data set (Juda et al., 2020).

In our model, we divided the main data set into 10 folds. Next, we let the model train for 10 iterations. Then we computed the accuracy for each iteration. After the last iteration, we computed the net accuracy by calculating the mean accuracy of the 10 iterations.

### 4.5 Ensemble classification

Ensemble classification in DL helps to improve the result of a model by combining several models together as a set (Reddy et al., 2020). The classification results of the models in the ensemble set are combined by using a specific method, with the guarantee that no single model will play a superior role in giving the final classification results. The specified method to combine the results of the set determines the type of ensemble classification. Ensemble classification can be categorised into two types: bagging and random forest. The latter was 5.

## 5 Experimental results

In this section, we present the experimental results that we have carefully designed to evaluate our proposed framework. A comparison between our model and other selected techniques is presented as well. All the phases, presented in Section 4, of our methodology have been tested and illustrated in this section.

Our model was built using Python 3.8, and tested on Microsoft Windows 7 64-bit, Intel(R) Core (TM) i5-4200U CPU @ 1.60GHz, Intel(R) HD Graphics Family and 16 GB of memory. In order to evaluate the performance of our proposed framework, the experiments were performed on DIARETDB0 data set, which is one of the widely accepted benchmark data sets used normally to evaluate the accuracy for detection of DR.

In order to facilitate the tracing of work in this section, we will present our experimental work in three parts. In the first part, we have built a basic deep learning architecture which employs CNN to detect and classify the images into infected and non-infected images. The second part presents the outcomes of applying our proposed framework and discusses the impact and the significance of each component of it. In the last part, we show a comparison between the outcomes of our framework and other selected classifiers.

Accuracy, Precision, Recall, F1-score and Area Under Curve (AUC) are the metrics used in this study to evaluate and test the reliability of the results. These metrics are the most common metrics used in literature to test and measure the accuracy and fitness of classification (Alkharabsheh et al., 2022, 2021, 2016). Accuracy, Precision, and Recall are calculated from the confusion matrix, which as shown in Figure 8 describes the result of prediction.

**Figure 8** Confusion matrix

| | Predicted (DR) | Predicted (NoDR) |
|---|---|---|
| Actual (DR) | TP | FN |
| Actual (NODR) | FP | TN |

In the context of DR disease, the True Positive (TP), False Positive (FP), True Negative (TN) and False Negative (FN) are interpreted as: *TP*: Observation is DR, and the predicted is DR. *FN*: Observation is DR, but the predicted is NoDR. *TN*: Observation is NoDR, and the predicted is NoDR. *FP*: Observation is NoDR, but the predicted is DR. Accuracy is the proportion of the truly predicted results for both classes, DR and NoDR, results among the total number of examined images. It is calculated by the equation:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{3}$$

Precision describes the proportion of the correctly predicted DR, among the total of predicted DR images.

$$Precision = \frac{TP}{TP + FP} \tag{4}$$

Recall is the total number of actual DR images, divided by the total number of the actual DR images, which is calculated by the equation:

$$Recall = \frac{TP}{TP + FN} \tag{5}$$

F1-Score is a value between 0 and 1, calculated as the harmonic mean of precision and recall, to maintain a balance between them. Therefore, F1 score is always closer to the lower value of recall and precision. F1 score is calculated by the equation:

$$F1score = 2 * \frac{Recall * Precision}{Recall + Precision} \tag{6}$$

## 5.1 Part 1: basic deep learning architecture

In this part, we have applied the basic structure of DL in which CNNs were employed to classify retina images. Utilising the TL techniques is one of the main features of our proposed framework. The other main features of our proposed framework are: applying CLAHE as a pre-processing step, utilising image augmentation techniques, and employing ensemble classification approach. In order to evaluate the significance of these three features, we have removed each one of these features once and present the results for each case. The confusion matrix, ROC curve, AUC score, accuracy, precision, recall and F1-score are presented in Figure 9 to demonstrate the outcomes when CLAHE is not activated, Figure 10 to demonstrate the results when data augmentation is not utilized and Figure 11 to demonstrate the outputs when ensemble methods are not employed. Table 1 presents those evaluation metrics in a tabulated form. It can be clearly seen from the outcomes in Table 1, Figures 9, 10 and 11 that the performance of the basic structure of CNN to classify retina images is not good enough in the medical treatment. Although the values of precision are fairly good in the table, but the values of accuracy, recall, F1-score and AUC are unsatisfactory. These results forced us to reinforce the basic architecture of DL with other techniques in order to enhance the quality of results. The enhancements that we have adopted are presented in the next subsection.

**Table 1** Outcomes of experimental works in part 1

| | *Metric* | *Without CLAHE* | *Without augmentation* | *Without ensemble classifier* |
|---|---|---|---|---|
| Part 1 | Accuracy | 0.53 | 0.67 | 0.44 |
| | Precision | 0.86 | 0.82 | 0.92 |
| | Recall | 0.51 | 0.77 | 0.34 |
| | F1-score | 0.64 | 0.79 | 0.50 |
| | AUC | 0.61 | 0.52 | 0.58 |

## 5.2 Part 2: the outcomes of the proposed framework

In this part of experimental work, we applied all the features of our proposed framework. Firstly, the images have been pre-processed by applying CLAHE on the extracted colours and the gray-scale images. These resulted images had good patterns of retina components, which led later to better classification. However, each of the produced images went through a different architecture of CNN for training and testing, which implied the need for ensemble classification. Figure 12 shows the effect of applying CLAHE on a selected image from DIARETDB0.

**Figure 9** Classification without CLAHE

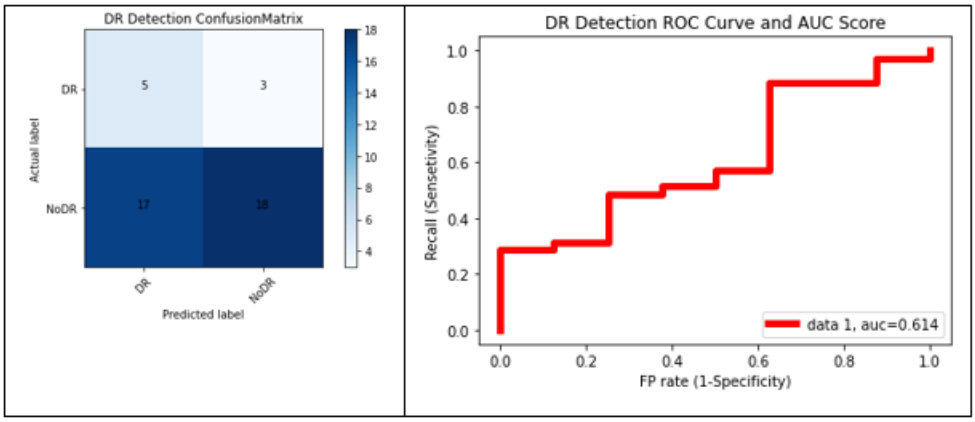


**Figure 10** Classification without image augmentation

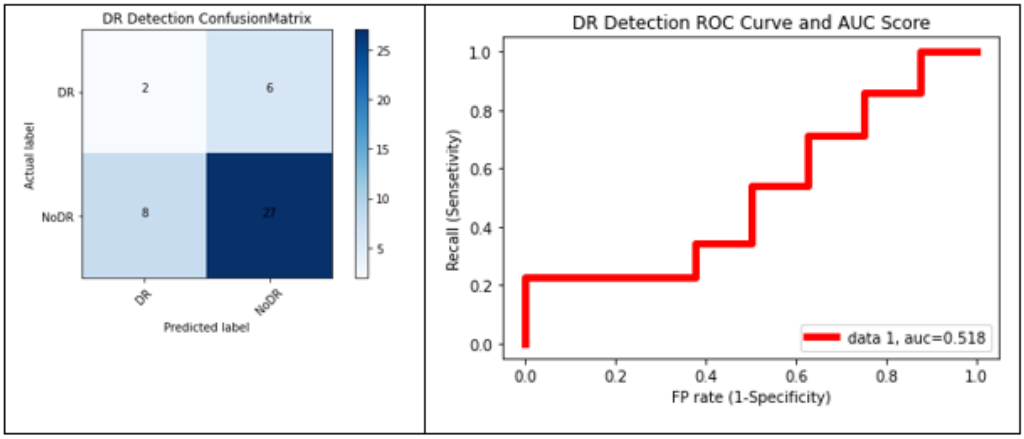



**Figure 11** Classification without ensemble classifier

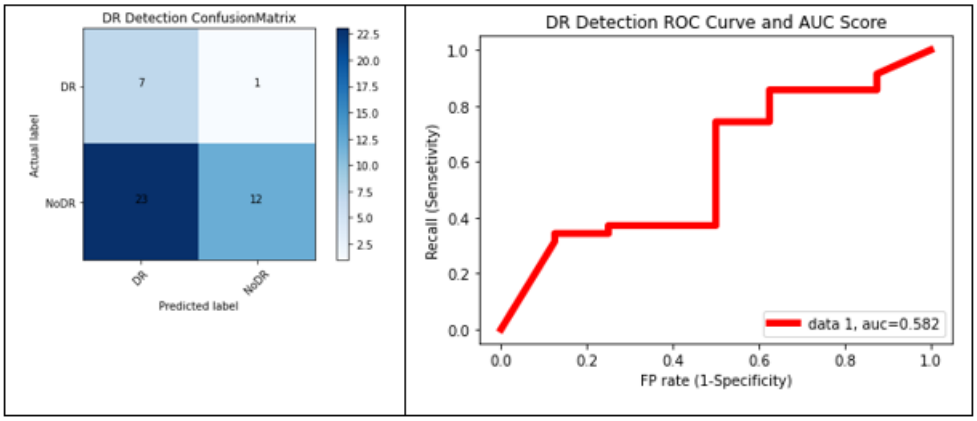



**Figure 12** CLAHE images

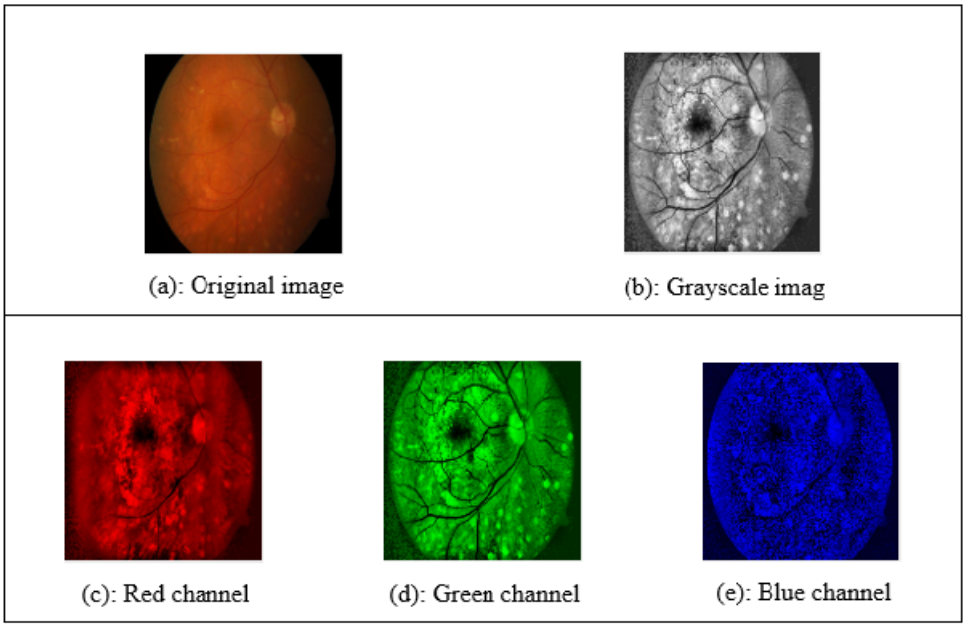


**Figure 13** The outcomes of our proposed framework

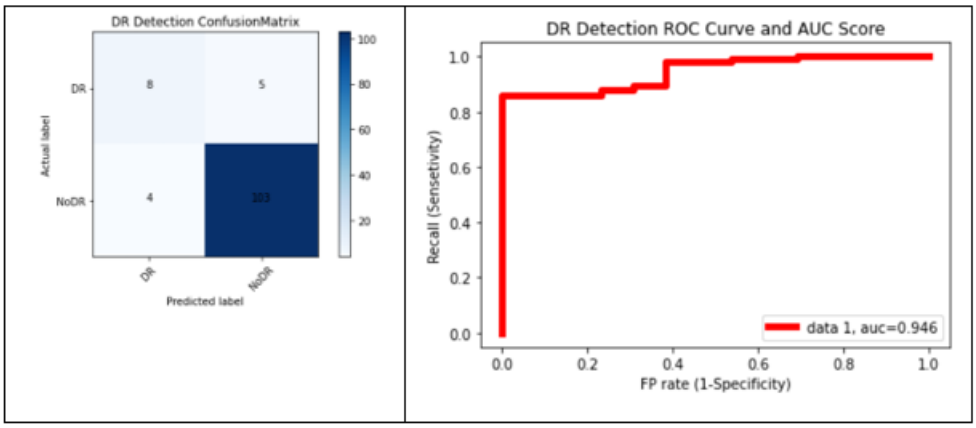



Image augmentation is then performed to enlarge the data set and to investigate the impact of using image augmentation in DR detection systems. The classification of images into infected and noninfected is then generated by the ensemble classification which is considered the last stage in our proposed framework. It is worth mentioning here that we used cross-validation statistical approach for evaluation in order to get accurate results and fair comparison between the evaluated algorithms. Figure 13 presents the confusion matrix, ROC curve, AUC score, accuracy, precision, recall and F1-score that are generated when our proposed framework is applied over the benchmark data set. The implementation of our proposed framework includes applying TL techniques to improve the performance of the target model by transferring the learned knowledge. VGG16, ResNet50, InceptionV3 and Xception are the pretrained models used in our proposed framework to enhance the learning process.

Table 3 presents the evaluation metrics values for our proposed model. The evaluation metrics of the cases that were mentioned in the previous subsection were tabulated as well in this table to facilitate the comparison. It is obvious from Table 3 that the use of CLAHE, image augmentation, ensemble classifiers in addition to TL techniques make a significant difference in terms of accuracy, precision, recall, F1-scare and AUC score. Since we presented in this part how our proposed framework enhanced the performance of basic architecture of CNN, we present in the next subsection a comparison between our classifier and other selected classifiers.

### *5.3 Part 3: comparison between our framework and other selected classifiers*

In order to compare our framework as a classifier for retina images, we compare it with the other widely accepted classifiers. The selected classifiers for comparison include Decision Tree (DT), *k*-neighbours (K-N), Logistic Regression (LR), Naive Bayes (NB), Random Forest (RF) and SVM. This set of selected classifiers were run over the same benchmark data set (i.e., DIARETDB0) and the cross-validation method was used to compute the results. Table 2 presents a comparison between our framework and other selected techniques, and Figures 14, 15, 16, 17, 18, and 19 represent the results generated by DT, K-N, LR, NB, RF and SVM, respectively. The measurements displayed in Table 2 proves that the proposed framework exceeds the other selected classifiers in terms of accuracy, precision, F1-score and area under curve whereas RF classifier achieves the best result in terms of recall. The results in this table confirms again that deep learning-based techniques can easily outperform traditional machine learning classifiers.

**Table 2** Comparison between our framework and other selected classifiers.

| | *Metric* | *DT* | *K-N* | *LR* | *NB* | *RF* | *SVM* | *Proposed* |
|---|---|---|---|---|---|---|---|---|
| Part 3 | Accuracy | 0.75 | 0.83 | 0.77 | 0.65 | 0.84 | 0.74 | **0.93** |
| | Precision | 0.86 | 0.84 | 0.85 | 0.86 | 0.85 | 0.83 | **0.95** |
| | Recall | 0.84 | 0.97 | 0.88 | 0.69 | **0.98** | 0.86 | 0.96 |
| | F1-score | 0.85 | 0.90 | 0.86 | 0.76 | 0.91 | 0.85 | **0.96** |
| | AUC | 0.57 | 0.53 | 0.55 | 0.57 | 0.55 | 0.46 | **0.95** |

**Table** 3 Outcomes of experimental works in part 1 and part 2

| | *Metric Without CLAHE* | *Without augmentation* | *Without ensembles* | | *Proposed framework* |
|---|---|---|---|---|---|
| Part 1 vs. Part 2 | Accuracy | 0.53 | 0.67 | 0.44 | 0.93 |
| | Precision | 0.86 | 0.82 | 0.92 | 0.95 |
| | Recall | 0.51 | 0.77 | 0.34 | 0.96 |
| | F1 score | 0.64 | 0.79 | 0.50 | 0.96 |
| | AUC | 0.61 | 0.52 | 0.58 | 0.95 |

It can be obviously concluded that our proposed classifier outperforms significantly the other classifiers in terms of accuracy, precision, F1-score and AUC. Regarding the recall metric, the RF classifier was the one which achieved the highest score. However, the difference between our proposed framework and the RF classifier in terms of recall metric is about 0.02 which is a small value.

**Figure 14** Classification by decision tree

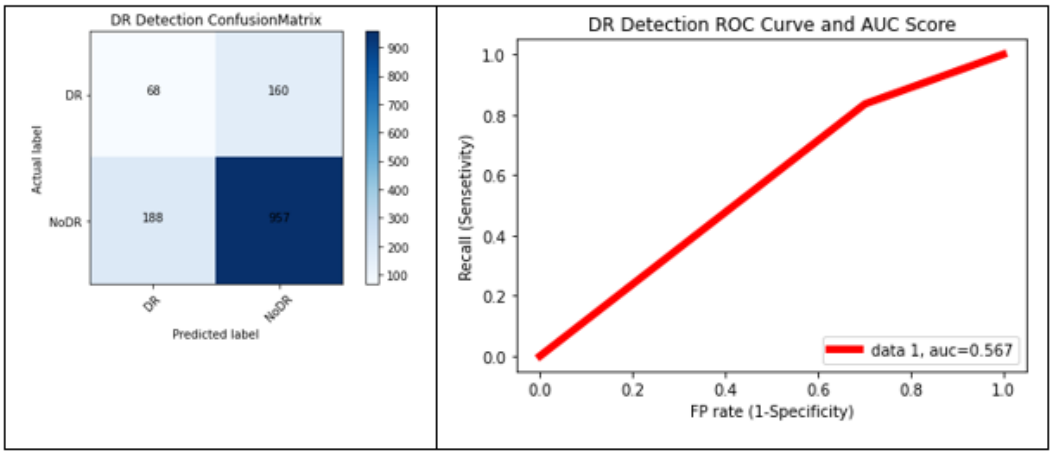



**Figure 15** Classification by *k*-neighbours

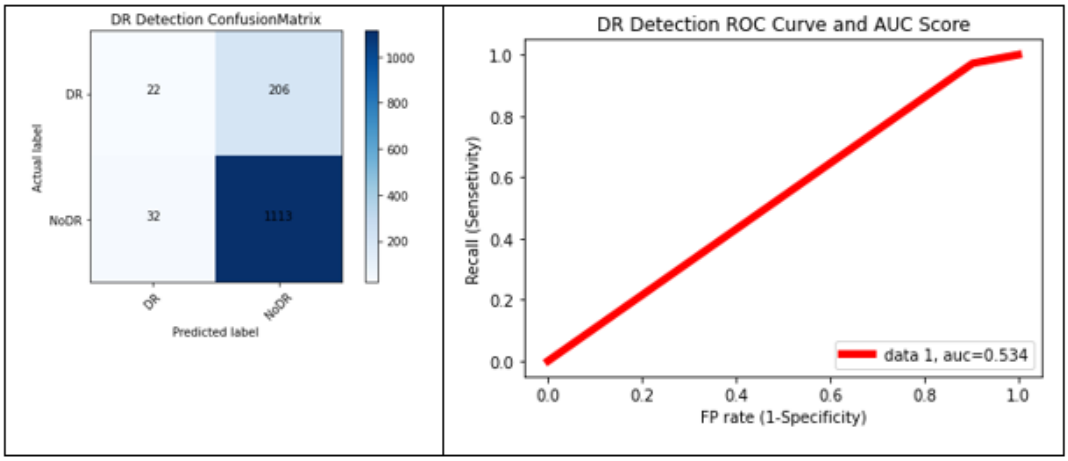



**Figure 16** Classification by logistic regression

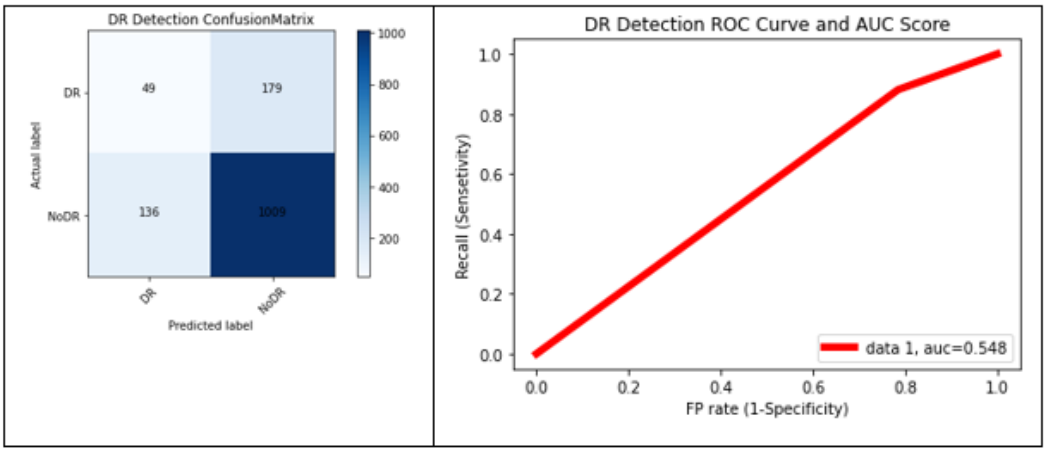


**Figure 17** Classification by naive Bayes

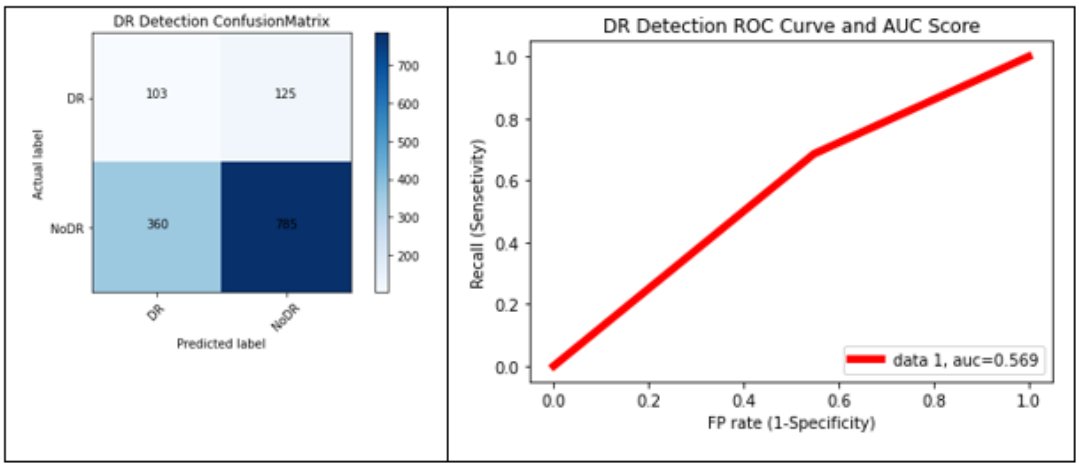



**Figure 18** Classification by random forest

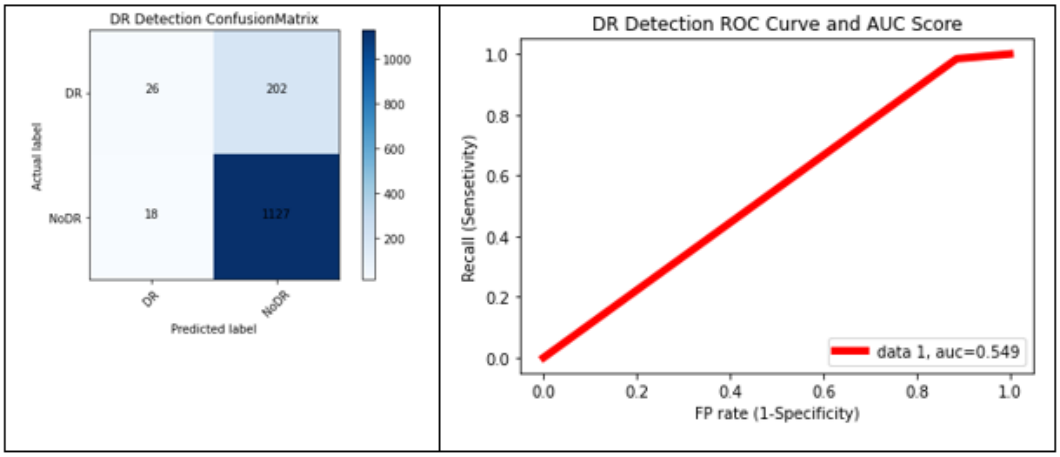
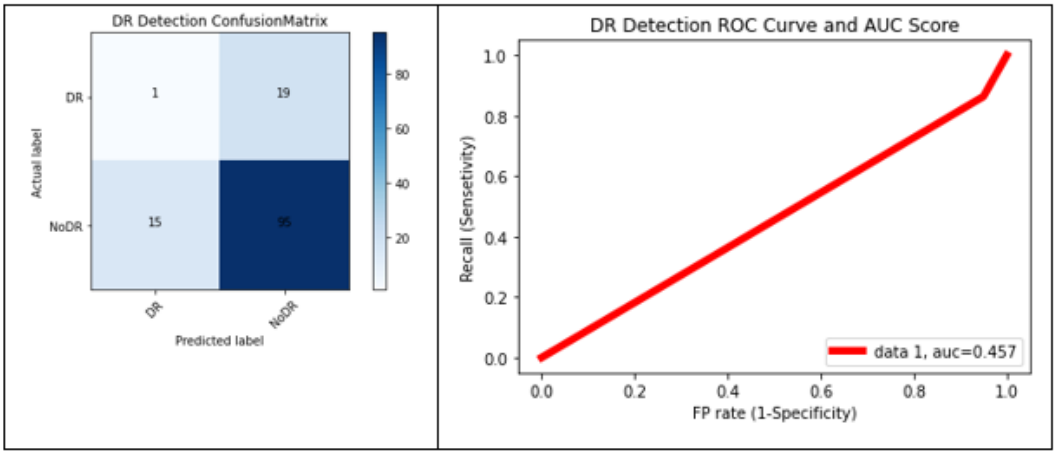



**Figure 19** Classification by support vector machine




In order to evaluate the stability and reliability of the results generated by our proposed framework, we performed the cross-validation statistical method, and so we have computed the standard deviation of the different runs of the evaluated algorithms. Table 4 presents the average accuracy and standard deviation of the proposed model and the other selected classifiers. It can be seen from Table 4 that all the selected classifiers, including ours, had (1) small values of standard deviation and (2) minor differences in values. Our classifier had the lowest value of standard deviation. This in turn confirms the stability of our proposed algorithm and the reliability of results generated by it.

**Table 4** The average accuracy and standard deviation of the proposed model and the other selected classifiers

| *Technique* | *Average accuracy* | *Standard deviation* |
|---|---|---|
| Proposed model | **0.93** | **0.0192** |
| Decision tree | 0.75 | 0.0236 |
| K-neighbours | 0.83 | 0.0226 |
| Logistic regression | 0.77 | 0.0218 |
| Naive Bayes | 0.65 | 0.0353 |
| Random forest | 0.84 | 0.0236 |
| Support vector machine | 0.74 | 0.0215 |

## 6 Conclusions

This paper presented a DL framework for the detection of DR disease in retinal images. The proposed framework includes a pipeline of four main stages which make it as a unique one. These stages are: image augmentation, contrast limited adaptive histogram equalisation, CNN and transfer learning and ensemble classification.

The experimental work demonstrated that the proposed framework achieved 93% overall accuracy, 95% precision and 96% recall. Table 4 explains how the research questions raised in Section 2 were answered by this study. Several algorithms have been proposed for automated DR detection using DL in the literature. Although our method also detects DR by using DL, it utilises a unique combination of various techniques, which is an innovative approach. The framework proposed in this study presented several contributions; It used CLAHE on each colour channel separately, unlike the other algorithms in the literature that combine the colours after the application of CLAHE to one image. Moreover, it employed four different types of CNN in TL where each one of the four CNNs (VGG16, InceptionV3, ResNet50 and Xception) is trained on a different colour channel image (i.e., red, green, blue and gray-scale), whereas the different CNNs in the algorithms in the literature were trained on one-colour images, RGB images, gray-scale images, or binary images. This unique mechanism led the framework to generate a significant enhancement. The proposed approach also applied ensemble methods in the classification of DR, which in turn enhanced the performance of the approach.

As part of the future work, further research is important to identify the feasibility of utilising our framework in ophthalmology clinical settings. Besides employing the proposed framework in smartphone-based retinal imaging systems, our proposed framework could be employed in other medical applications that require the use of colour images for classification.

## References


Al-Hiary, H., Bani-Ah Mad, S., Reyalat, M., Braik, M. and Lrahamneh, Z.A. (2011) 'Fast and accurate detection and classification of plant diseases', *International Journal of Computers and Applications*, IEEE, Vol. 17, No. 1, pp.31–38.

Alkharabsheh, K., Alawadi, S., Kebande, V.R., Crespo, Y., Delgado, M.F. and Taboada, J.A. (2022) 'A comparison of machine learning algorithms on design smell detection using balanced and imbalanced dataset: a study of god class', *Information and Software Technology*, Vol. 143. Doi: 10.1016/j.infsof.2021.106736.

Alkharabsheh, K., Almobydeen, S., Taboada, J.A. and Crespo, Y. (2016) 'Influence of nominal project knowledge in the detection of design smells: an exploratory study with God class', *International Journal of Advanced Studies in Computers, Science and Engineering*, Vol. 5, No. 11, pp.120–127.

Alkharabsheh, K., Crespo, Y., Delgado, M.F., Viqueira, J.R.R. and Taboada, J.A. (2021) 'Exploratory study of the impact of project domain and size category on the detection of the god class design smell', *Software Quality Journal*, Vol. 29, No. 2, pp.197–237.

Bhatt, T.V., Patel, R.K., Chitara, H.B., Marques, G. and Bhoi, A.K. (2020) 'Fuzzy logic system for diabetic eye morbidity prediction', *International Journal of Computer Applications in Technology*, Vol. 64, No. 4, pp.339–348.

Chakrabarty, N. (2018) 'A deep learning method for the detection of diabetic retinopathy', *Proceedings of the 5th IEEE Uttar Pradesh Section International Conference on Electrical, Electronics and Computer Engineering (UPCON)*, IEEE, pp.1–5.

Chen, J-F., Chen, W-L., Huang, C-P., Huang, S-H. and Chen, A-P. (2016) 'Financial time-series data analysis using deep convolutional neural networks', *Proceedings of the 7th International Conference on Cloud Computing and Big Data (CCBD)*, IEEE, pp.87–92.

Chollet, F> (2017) 'Xception: deep learning with depthwise separable convolutions', *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.1251–1258.

De La Torre, J., Valls, A. and Puig, D. (2020) 'A deep learning interpretable classifier for diabetic retinopathy disease grading', *Neurocomputing*, Vol. 396, pp.465–476.

Dorgham, O.M., Alweshah, M., Ryalat, M.H., Alshaer, J., Khader, M. and Alkhalaileh, S. (2021) 'Monarch butterfly optimization algorithm for computed tomography image segmentation', *Multimedia Tools and Applications*, Vol. 80, No. 20, pp.1–34.

Fan, X-N., Gong, J. and Yan, Y. (2019) 'Red lesion detection in fundus images based on convolution neural network', *Chinese Control And Decision Conference (CCDC)*, IEEE, pp.5661–5666.

Fenner, B-J., Wong, R.L.M., Lam, W-C., Tan, G.S.W. and Cheung, G.C.M. (2018) 'Advances in retinal imaging and applications in diabetic retinopathy screening: a review', *Ophthalmology and Therapy*, Vol. 7, No. 2, pp.333–346.

Fisher, M., Dorgham, O. and Laycock, S.D. (2013) 'Fast reconstructed radiographs from octree-compressed volumetric data', *International Journal of Computer Assisted Radiology and Surgery*, Vol. 8, No. 2, pp.313–322.

Flaxel, C.J., Adelman, R.A., Bailey, S.T., Fawzi, A., Lim, J.I., Vemulakonda, G.A. and Ying, G.S. (2020) 'Diabetic retinopathy preferred practice pattern', *Ophthalmology*, Vol. 127, No. 1, pp.P66–P145.

Foeady, A.Z., Novitasari, D.C.R., Asyhar, A.H. and Firmansjah, M. (2018) 'Automated diagnosis system of diabetic retinopathy using GLCM method and SVM classifier', *Proceedings pf the 5th International Conference on Electrical Engineering, Computer Science and Informatics (EECSI)*, IEEE, pp.154–160.

Gharaibeh, N.Y. (2017) 'A novel approach for detection of microaneurysms in diabetic retinopathy disease from retinal fundus images', *Computing and Information Sciences*, Vol. 10, No. 1, pp.1–15.

Habib, M.M., Welikala, R.A., Hoppe, A., Owen, C.G., Rudnicka, A.R. and Barman, S.A. (2017) 'Detection of microaneurysms in retinal images using an ensemble classifier', *Informatics in Medicine Unlocked*, Vol. 9, pp.44–57.

He, K., Zhang, X., Ren, S. and Sun, J. (2016) 'Deep residual learning for image recognition', *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.770–778.

Huang, Y., Lin, L., Li, M., Wu, J., Cheng, P., Wang, K., Yuan, J. and Tang, X. (2020) 'Automated hemorrhage detection from coarsely annotated fundus images in diabetic retinopathy', *IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, IEEE, pp.1369–1372.

Jaya, I., Andayani, U., Siregar, B., Febri, T. and Arisandi, D. (2019) 'Identification of retinoblastoma using the extreme learning machine', *Journal of Physics: Conference Series*, IOP Publishing. Doi: 10.1088/1742-6596/1235/1/012057.

Juda, P., Renard, P. and Straubhaar, J. (2020) 'A framework for the cross-validation of categorical geostatistical simulations', *Earth and Space Science*, Vol. 7, No. 8, pp.e2020EA001152, 2020.

Kajan, S., Goga, J., Lacko, K. and Pavlovičová, J. (2020) 'Detection of diabetic retinopathy using pretrained deep neural networks', *Cybernetics and Informatics (K&I)*, IEEE, pp.1–5.

Kannan, S., Yuvaraj, N., Idrees, B.A., Arulprakash, P., Ranganathan, V., Udayakumar, E. and Dhinakar, P. (2021) 'Analysis of convolutional recurrent neural network classifier for covid-19 symptoms over computerised tomography images', *International Journal of Computer Applications in Technology*, Vol. 66, Nos. 3/4, pp.427–432.

Kassani, S.H., Kassani, P.H., Khazaeinezhad, R., Wesolowski, M.J., Schneider, K.A. and Deters, R. (2019) 'Diabetic retinopathy classification using a modified xception architecture', *IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)*, IEEE, pp.1–6.

Kaur, M., Rani, S., Gupta, D. and Manocha, A.K. (2021) 'Prediction of diabetic patients using various machine learning techniques', *International Journal of Computer Applications in Technology*, Vol. 66, No. 2, pp.100–106.

Lands, A., Kottarathil, A.J., Biju, A., Jacob, E.M. and Thomas, S. (2020) 'Implementation of deep learning based algorithms for diabetic retinopathy classification from fundus images', *Proceedings of the 4th International Conference on Trends in Electronics and Informatics (ICOEI)(48184)*, IEEE, pp.1028–1032.

Li, X., Shen, L., Shen, M., Tan, F. and Qiu, C.S. (2019b) 'Deep learning based early stage diabetic retinopathy detection using optical coherence tomography', *Neurocomputing*, Vol. 369, pp.134–144.

Li, Y-H., Yeh, N-N., Chen, S-J. and Chung, Y-C. (2019a) 'Computer-assisted diagnosis for diabetic retinopathy based on fundus images using deep convolutional neural network', *Mobile Information Systems*. Doi: 10.1155/2019/6142839.

Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., Van Der Laak, J.A., Van Ginneken, B. and Sánchez, C.I. (2017) 'A survey on deep learning in medical image analysis', *Medical Image Analysis*, Vol. 42, pp.60–88.

Manjesh, P. and Nagachandra, M.K. (2016) 'Secondary glaucoma diagnosis technique using retinal layers', *International Journal of Exploring Emerging Trends in Engineering*, Vol. 3, pp.228–232.

Pour, A.M., Seyedarabi, H., Jahromi, S.H.A. and Javadzadeh, A. (2020) 'Automatic detection and monitoring of diabetic retinopathy using efficient convolutional neural networks and contrast limited adaptive histogram equalization', *IEEE Access*, Vol. 8, pp.136668–136673.

Qiao, L., Zhu, Y. and 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*, Vol. 8, pp.104292–104302.

Qummar, S., Khan, F.G., Shah, S., Khan, A., Shamshirband, S., Ur Rehman, Z., Khan, I.A. and Jadoon, W. (2019) 'A deep learning ensemble approach for diabetic retinopathy detection', *IEEE Access*, Vol. 7, pp.150530–150539.

Reddy, G.T., Bhattacharya, S., Ramakrishnan, S.S., Chowdhary, C.L., Hakak, S., Kaluri, R. and Reddy, M.P.K. (2020) 'An ensemble based machine learning model for diabetic retinopathy classification', *Proceedings of the International Conference on Emerging Trends in Information Technology and Engineering (IC-ETITE)*, IEEE, pp.1–6.

Reddy, L., Raval, M., Patel, J. and Redkar, M. (2019) 'Diabetic retinopathy detection from retinal images 2019', *International Research Journal of Engineering and Technology*, Vol. 6, No. 2.

Reyana, A., Krishnaprasath, V.T., Kautish, S., Panigrahi, R. and Shaik, M. (2020) 'Decision-making on the existence of soft exudates in diabetic retinopathy', *International Journal of Computer Applications in Technology*, Vol. 64, No. 4, pp.375–381.

Ryalat, M.H. (2022) 'A new algorithm to find the k-th smallest element in an unordered list (efficient for big data)', *Proceedings of the 2nd International Conference on Computing and Information Technology (ICCIT)*, IEEE, pp.51–56.

Ryalat, M.H., Emmens, D., Hulse, M., Bell, D., Al-Rahamneh, Z., Laycock, S. and Fisher, M. (2016) 'Evaluation of particle swarm optimisation for medical image segmentation', *Proceedings of the International Conference on Systems Science*, Springer, pp.61–72.

Ryalat, M.H., Laycock, S. and Fisher, M. (2017a) 'Automatic removal of mechanical fixations from CT imagery with particle swarm optimisation', *Proceedings of the International Conference on Bioinformatics and Biomedical Engineering*, Springer, pp.419–431.

Ryalat, M.H., Laycock, S. and Fisher, M. (2017b) 'A fast and automatic approach for removing artefacts due to immobilisation masks in x-ray CT', *Proceedings of the IEEE EMBS International Conference on Biomedical and Health Informatics (BHI)*, IEEE, pp.33–36.

Saha, O., Sathish, R. and Sheet, D. (2019) 'Fully convolutional neural network for semantic segmentation of anatomical structure and pathologies in colour fundus images associated with diabetic retinopathy', *arXiv preprint arXiv:1902.03122*.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. and Wojna, Z. (2016) 'Rethinking the inception architecture for computer vision', *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.2818–2826.

Thota, N.B. and Reddy, D.U. (2020) 'Improving the accuracy of diabetic retinopathy severity classification with transfer learning', *Proceedings of the IEEE 63rd International Midwest Symposium on Circuits and Systems (MWSCAS)*, IEEE, pp.1003–1006.

Ting, D.S.W., Cheung, G.C.M. and Wong, T.Y. (2016) 'Diabetic retinopathy: global prevalence, major risk factors, screening practices and public health challenges: a review', *Clinical and Experimental Ophthalmology*, Vol. 44, No. 4, pp.260–277.

Verbraak, F.D., Abramoff, M.D., Bausch, G.C.F., Klaver, C., Nijpels, G., Schlingemann, R.O. and Van der Heijden, A.A. (2019) 'Diagnostic accuracy of a device for the automated detection of diabetic retinopathy in a primary care setting', *Diabetes Care*, Vol. 42, No. 4, pp.651–656.

Wang, X., Lu, Y., Wang, Y. and Chen, W-B. (2018) 'Diabetic retinopathy stage classification using convolutional neural networks', *IEEE International Conference on Information Reuse and Integration (IRI)*, IEEE, pp.465–471.

Wedyan, M., Al-Jumaily, A. and Dorgham, O. (2020) 'The use of augmented reality in the diagnosis and treatment of autistic children: a review and a new system', *Multimedia Tools and Applications*, 79, No. 25, pp.18245–18291.

Zago, G.T. (2019) 'Diabetic retinopathy detection based on deep learning', (Doutorado em Engenharia Elétrica) - Centro Tecnológico, Universidade Federal do Espírito Santo, Vitória.

Zeng, X., Chen, H., Luo, Y. and Ye, W. (2019) 'Automated diabetic retinopathy detection based on binocular siamese-like convolutional neural network', *IEEE Access*, Vol. 7, pp.30744–30753.

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H. and He, Q. (2020) 'A comprehensive survey on transfer learning', *Proceedings of the IEEE*, Vol. 109, No. 1, pp.43–76.