



VICTORIA UNIVERSITY
MELBOURNE AUSTRALIA

Convolutional neural network for multi-class classification of diabetic eye disease

This is the Published version of the following publication

Sarki, Rubina, Ahmed, Khandakar, Wang, Hua, Zhang, Yanchun and Wang, Kate (2021) Convolutional neural network for multi-class classification of diabetic eye disease. *EAI Endorsed Transactions on Scalable Information Systems*, 9 (4). ISSN 2032-9407

The publisher's official version can be found at
<https://eudl.eu/doi/10.4108/eai.16-12-2021.172436>
Note that access to this version may require subscription.

Downloaded from VU Research Repository <https://vuir.vu.edu.au/47084/>

Convolutional Neural Network for Multi-class Classification of Diabetic Eye Disease

Rubina Sarki¹, Khandakar Ahmed¹, Hua Wang¹, Yanchun Zhang¹, Kate Wang²

¹ISILC, Victoria University, Melbourne, VIC 3011, Australia, ²School of Health and Biomedical Sciences, RMIT University, Bundoora, VIC 3083, Australia

Abstract

Prompt examination increases the chances of effective treatment of Diabetic Eye Disease (DED) and reduces the likelihood of permanent deterioration of vision. A key tool commonly used for the initial diagnosis of patients with DED or other eye disorders is the screening of retinal fundus images. Manual detection with these images is, however, labour intensive and time consuming. As deep learning (DL) has recently been demonstrated to provide impressive benefits to clinical practice, researchers have attempted to use DL method to detect retinal eye diseases from retinal fundus photographs. DL techniques in machine learning (ML) have achieved state-of-the-art performance in the binary classification of healthy and diseased retinal fundus images while the classification of multi-class retinal eye diseases remains an open challenge. Multi-class DED is therefore considered in this study seeking to develop an automated classification framework for DED. Detecting multiple DEDs from retinal fundus images is an important research topic with practical consequences. Our proposed model was tested on various retinal fundus images gathered from the publicly available dataset and annotated by an ophthalmologist. This experiment was conducted employing a new convolutional neural network (CNN) model. Our proposed model for multi-class classification achieved a maximum accuracy of 81.33%, sensitivity of 100%, and specificity of 100%.

Received on 09 November 2021; accepted on 12 December 2021; published on 16 December 2021

Keywords: Diabetic Eye Disease, Deep Learning, Multi-class Classification, Image Processing

Copyright © 2021 Rubina Sarki *et al.*, licensed to EAI. This is an open access article distributed under the terms of the [Creative Commons Attribution license](#), which permits unlimited use, distribution and reproduction in any medium so long as the original work is properly cited.

doi:10.4108/eai.16-12-2021.172436

1. Introduction

DED is a category of eye disorders that can affect people with diabetes. Diabetic retinopathy (DR), diabetic macular edema (DME), glaucoma (GI), and cataract (Ca) are among these diseases (see Fig. 1). Diabetes can cause eye damage over time, resulting in blurred vision or even vision impairment. However, by keeping track of diabetes, one can avoid DED or keep it from worsening [1]. Around one-third of people with diabetes are likely to be diagnosed with DED. 2.2 billion people worldwide are confirmed by the World Health Organization (WHO) to have blindness or vision loss, and at least 1 billion have a vision impairment which

could have been reversed¹. Identifying and diagnosing these diseases for treatment is critical.

Motivated by the need for successful development of detection and preventive measures to implement the wide range of lifespan needs associated with retinal conditions and visual impairments. Deep learning (DL) has recently been demonstrated to provide impressive benefits to many real applications [2–5], researchers have attempted to use DL method to detect retinal eye diseases from retinal fundus photographs [6, 7]. DL techniques in machine learning (ML) have achieved state-of-the-art performance in the binary classification of healthy and diseased retinal fundus images while the classification of multi-class retinal eye diseases remains an open challenge.

*Corresponding author. Email: rubina.sarki@live.vu.edu.au

¹<https://www.who.int/en/news-room/fact-sheets/detail/blindness-and-visual-impairment>

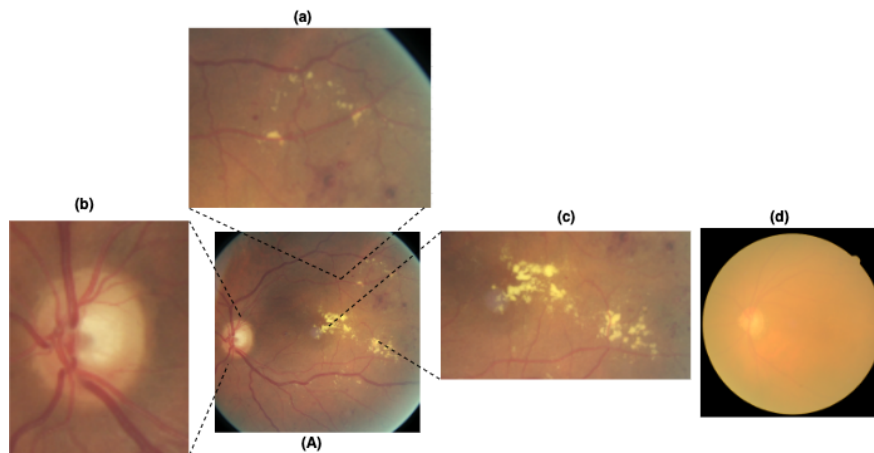


Figure 1. Complications of DED in fundus images (A); (a). Microaneurysms; narrow bulges in the side of the blood vessel (Diabetic Retinopathy) (b). Progressive damage of optic nerve damage (Glaucoma) (c). Exudates formation in macular region and thickening of macula (Diabetic Macular Edema) (d). Degeneration of lens (Cataract).

To solve these problems, automated DED diagnostic techniques using DL are vital [8, 9]. The time and labour intensive [10] nature of screening of DED make it a crucial inaccurate prognosis based on a competent ophthalmologist. While DL has generally achieved high accuracy of validity for healthy and diseased (binary) classification, the results of multi-class classification, particularly for early-stage disability, are less impressive [11, 12]. We, therefore, present in this paper an automated DED classification model based on a deep convolutional neural network (DCNN) that can distinguish healthy images from disease pathology.

Initially, various DCNN architectures were evaluated to determine the best performing system for the mild and multi-class DED classification tasks [13]. We aspired to achieve the highest standards of output that previous works have recorded. The following is a list of the research's main contributions.

- We trained the proposed CNN model and evaluated a multi-class classification model to improve sensitivities for the various DED levels.
- Incorporating various pre-processing methods.
- Augmentation approaches to improve the accuracy of the results and sample volume adequacy for the dataset.

Early amplification of retinal diseases is important however, it takes a tremendous amount of time and memory to diagnose these diseases using neural networks. To improve accuracy, additional data must be given, but this requires high computing capacity and a large amount of time spent. Therefore, the method will benefit from a pre-trained model as it adapts the design to minimize losses. Pretrained

models or models for transfer learning (TL)[14] have already shown promising results in the classification and detection of medical images [15–18]. In this phase of the research, following TL theory, we used state-of-the-art CNN models, pre-trained on the sizeable public image repository, ImageNet, also used in our previous study [13, 19]. The deep neural network's top layers were trained with the publicly available fundus image corpora using initialized weights for mild and multi-class classification. We proposed a new CNN model to solve the problem of multi-class classification. Initially, the highest-performing CNN model was developed based on a comprehensive experiment conducted in previous studies [20, 21]. Then we evaluated the set of performance improvements, including image-processing and optimizer selection. Finally, an ideal specificity and sensitivity were achieved by the proposed model, thus facilitating an efficient and effective fully automated DL-based system improving the outcomes of mass screening services for the at-risk population.

The arrangement of this article is as follows. Section 3 This section covers study design based on datasets, image processing, and DL algorithms. Section 4 addresses the experiments performed in two scenarios. Section 5 compared the results based on the classification evaluation. Section 6 discussion of the article. Finally, section 7 concludes the paper.

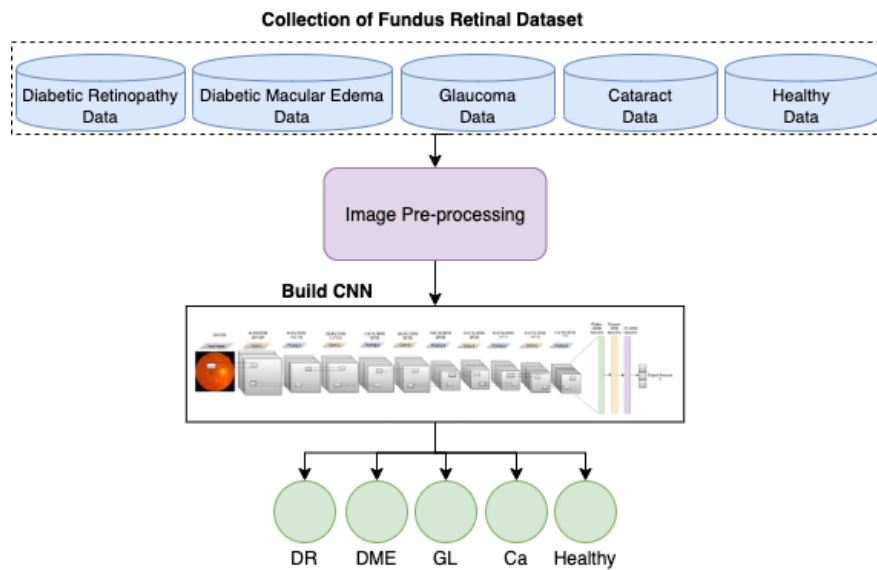


Figure 2. The computational steps

2. Literature Review

By using machine learning algorithms[22–25] to identify a huge proportion of fundus images captured from ocular scanning programs[26, 27], several previous researches focused on automatic retinal disease detection. Various machine learning techniques have been introduced for the automatic detection of retinal diseases[15]: Artificial Neural Network(ANN), K-nearest neighbor algorithm, Support Vector Machine(SVM), and Naive Bayes classifier. To define the difference between glaucoma and non-glaucoma[28, 29], multiple studies have introduced ANN models. By finding preperimetric glaucoma using DL feedforward neural networks(FNN)[30], a glaucoma research team observed visual field assessment. An artificial DCNN was applied for the grading sensitivity of nuclear cataracts[31]. In the field of ML techniques, DL emerges as a common solution for different classification issues [32–34]. An advanced DL model capable of diagnosing Diabetes Mellitus Retinopathy(DMR)[8] has been developed by the Google research community. Using identical DL techniques, fundus photographs and optical coherence tomography[35, 36], the Age-related Macular Degeneration(AMD) research was performed. However, binary classification was chosen by all retinal image classification studies from which problems with normal versus one disease were resolved[37]. Besides, Lam et al.[38] used pre-trained networks (GoogleNet and AlexNet) in the Messidor dataset to distinguish mild and multi-class diabetic retinopathy, researchers developed selective contrast-limited adaptive histogram equalization (CLAHE) and reported enhancement in recognition of subtle characteristics. Multi-class DL architectures have been tested by Choi et

al.[39] for automated detection of ten retinal disorders. This work’s findings have shown that existing DL algorithms have failed to distinguish retinal images from small multi-group datasets. While the outcome of high classification success in structured experimental settings has been published in research conducted in this field, it is fundamentally difficult to apply the binary classification model in real medical practice where visiting patients suffer from different DEDs. Indeed, there have been minimal studies on mild and multi-class groups aimed at recognizing DED. This study used a state-of-the-art CNN for fundus image analysis to adapt TL in mild and multi-class DED environments. This paper articulates a pilot study intended to evaluate TL on mild and multi-class classification using the small open-source fundus retinal image database.

3. Study Design

The real aim of this study is to increase performance in automated multi-class DED classification and detection through an experimental assessment of different classification improvement techniques. The following are key definitions of associated priorities and goals.

1. Performance analysis of proposed CNN framework with multiple datasets collected from multiple sources;
2. Impact of an optimizer on the performance of proposed frameworks;
3. Visual representation of the performance of the model using heat-map;
4. Analysis of image quality improvement using contrast enhancement techniques for further

classification improvements on multi-class DED detection task.

The DL process is illustrated (see Fig. 2), to demonstrate the steps that followed.

Data Collection. Open-source data was collected, including publicly available Messidor, Messidor-2, DRISHTI-GS, and kaggle cataract datasets. Despite its relatively limited size, Messidor’s dataset contains high-resolution photos with accurate labeling. Similarly, Messidor-2 is a public dataset for analyzing DED algorithm output used by other individuals. The data comprises 1,748 photographs of 874 topics. Messidor-2 varies from the primary 1200 image data set of Messidor, and for each object has two images, one for each eye. They acquired the cataract and glaucoma dataset from the retina dataset obtained by Kaggle. This dataset consists of 100 images of a cataract and 100 images of glaucoma.

Image Processing. To accomplish contrast enhancement in the retinal fundus images, mathematical morphology has been used. Mathematical morphology approaches work hinged on the structural values of objects. To pull out the elements of an image, these methods use relationships between classes and mathematical fundamentals, which help explain areas. In morphological operators, the input consists of two data sets. The original image is included in the first set, and the second one illustrates the structural element (SE), which is also called a mask. The original image is in grey level or binary, and the mask is a 0s and 1s value matrix[40]. In morphological operators, if the gray-level image matrix id represented by $I(x, y)$ and the SE by $S(u, v)$, the erosion and dilation operators are defined as equation 1 and 2 [41].

$$I \ominus S = \min_{u,v} \{I(x + u, y + v) - S(u, v)\} \quad (1)$$

$$I \oplus S = \max_{u,v} \{I(x - u, y - v) + S(u, v)\} \quad (2)$$

The erosion operator decreases the objects’ size and increases the size of an image’s holes and eliminates very tiny information from that image. It makes the final image appear darker than the original image by removing bright areas under the SE. The dilation operator operates in reverse; in other words, the size of objects increases and holes in the image decreases, respectively. Therefore, the opening operator is similar to implementing the dilation and erosion operations on the same image equation 3, while the closing operator acts in reverse equation 4.

$$I \circ S = (I \ominus S) \oplus S \quad (3)$$

$$I \bullet S = (I \oplus S) \ominus S \quad (4)$$

The opening operator eliminates poor relations between artifacts and small information, while small gaps are eliminated, and the closing operator fills cracks. The size and shape of a SE are usually chosen arbitrarily; however, disk-shaped SE is used more frequently than other masks for medical images.

Proposed Convolutional Neural Network. This layer comprises a filter set (kernel). Each filter is convoluted against the input image and then extract features by creating a new layer. Each layer signifies some of the important features or characteristics of the input image. The * symbol identifies the operation of the convolution. The output (or function map) $K(t)$ where t is time, defined below when input $M_n(t)$ where t is time is convoluted with a filter or $k(a)$ kernel where a is the measurement. We obtained a function new function K to estimate the position. This operation is called convolution.

$$K(t) = (M_n * k)(t). \quad (5)$$

If t can only accept integer values, the following discrete convolution is provided by the following equation:

$$K(t) = \sum_a M_n(a) \cdot k(t - a). \quad (6)$$

The above assumes a one-dimensional convolutional operation. A two dimension convolution operation with input $M_n(m, n)$ and a kernel $k(a, b)$ is defined as:

$$K(t) = \sum_a \sum_b M_n(m, n) \cdot k(m - a, n - b). \quad (7)$$

By the commutative law, the kernel is flipped and the above is equivalent to:

$$K(t) = \sum_a \sum_b M_n(m - a, n - b) \cdot k(a, b). \quad (8)$$

Neural networks implement the cross-correlation function, which is the same as convolution but without flipping the kernel.

$$K(t) = \sum_a \sum_b M_n(m + a, n + b) \cdot k(a, b). \quad (9)$$

So there is less variation in the range of valid values of m and n , this last perspective is usually easier to implement in a machine learning framework.

Rectified Linear Unit (ReLU) Layer. This layer is an activation function that sets the negative input value to zero, which optimizes and speeds up analyses and training, and helps prevent the gradient from disappearing. Mathematically, this described as:

$$P(x) = \max(0, x). \quad (10)$$

In which x is input to the neuron.

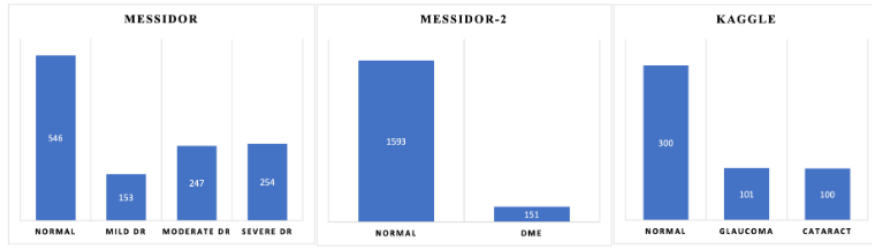


Figure 3. Data Distribution

Maxpooling Layer. This Layer is a sample-based discretization method. It is employed to down-sample an input design (input image, hidden-layers, output matrix, etc.), its dimensionality is compressed, allowing assumptions about the components available in the binned sub-regions. This will decrease the size of learning parameters and provide fundamental interpretation invariance to internal depiction, thus further reducing the cost of computation. Our model adopted the kernel size of 3×3 during the Maxpooling process. After the final convolution block, the network flattened to one dimension.

Batch Normalization. Batch normalization enables every layer of the network to learn a little more independently of the other layers. It also normalizes the output from the previous activation layer by subtracting the batch mean and dividing the batch standard deviation [42] to improve the steadiness of the neural network.

Fully Connected Layer. This layer takes the output of the previous layer (Convolutional, ReLU, or Pooling) as its input and calculates the probability values for classification into the various groups.

Loss Function. This layer applies a soft-max function to the input data sample. This layer is used for the final prediction. Therefore, our loss function is given as:

$$L_i = -\log\left(\frac{e^{\beta_y}}{\sum_j^c e^{\beta_j}}\right) \quad (11)$$

Where β_j is the j th element of the vector of class scores β , β_y is the CNN score for the positive class and c is classes for each image. The softmax ensures a proper prediction probability in the log of the equation.

Regularization. An efficient regularization method named as a dropout is employed. This strategy was being proposed by Srivastava et al. [43]. During the training process, the dropout is conducted by maintaining the neuron active with a certain probability P or by setting it to 0. In our study, we set hyperparameter to 0.50 because it outputs in the maximum amount of regularization[44].

Classification Description. Training and testing performed by new CNN architecture, as mentioned above, in Scenario I classification models.

- *Scenario I:* In this scenario, we will classify five classes of DED, such as healthy, DR, DME, Gl, and Ca retinal fundus images (multi-class classification)

Classification Performance Analysis. Retinal fundus images are evaluated as follows, the efficacy of each CNN is calculated by various metrics implemented to measure the true and false category for the diagnosed DED. Next, the cross-validation estimator[45] is used, resulting in a Table.2 confusion matrix. If the classification algorithm accurately categorizes samples based on a specific category, they are included in the TP indices. The TN indices of the confusion matrix contain the other samples which relate correctly to some other classes defined. Similarly, in the confusion matrix, the FP and FN indices represent the number of samples that the classifier wrongly estimates.

Accuracy: For experimental analytics, output parameters derived from the confusion matrix are used. Accuracy in eq (5) is essentially the fraction of the uncertainty matrix's total values between true positive and true negative. Therefore, the above elements of the confusion matrix should determine the classifier's efficacy in our present framework.

$$Accuracy(\%) = \frac{TP + TN}{TP + FN + TN + FP}. \quad (12)$$

Similarly, **Sensitivity (Recall):** A test's sensitivity is often called the true positive rate (TPR), which is the percentage of truly positive tests that use the test in question to give a positive result.

$$Sensitivity(Recall) = \frac{TP}{TP + FN} \quad (13)$$

Specificity: The specificity also refers to it as the true negative rate (TNR), which is the percentage of genuinely negative samples testing negative using the test in question.

$$Specificity = \frac{TN}{TN + FP} \quad (14)$$

Table 1. Hyper-parameters of the build CNN model and preferred weights in this study.

R1	R2	R3	R4	R5	R6	R7	R8
CNN	224*224	RMSprop	32	10-fold	3e-4	BCE	50

Legend: R1 - Model, R2 - Image Size, R3 - Optimizers, R4 - Mini Batch Size, R5 - cross validation, R6 - Initial Learning Rate, R7 - Loss function, R8 - Epoch, BCE - Binary cross-entropy.

Table 2. Confusion Matrix

	Predictive Positive	Predictive Negative	Total
Actual Positive	TP	FN	TP + FN
Actual Negative	FP	TN	FP + TN
Total	TP + FP	FN + TN	

Legend: TP = True Positive, FN = False Negative, FP = False Positive, TN = True Negative.

Precision Call: Precision demonstrates the ratio of the real true positive test samples are flattering for all test samples classified by the model. The formula will be seen as follows:

$$Precision = \frac{TP}{TP + FP} \quad (15)$$

ROC curve: A ROC (receiver operating characteristic curve) curve is a graphical representation that demonstrates how a test’s specificity and sensitivity differ concerning each other. Using the test, samples considered to be true or false are calculated to create a ROC curve. To offer a graph identical to the one below, the TPR (sensitivity) is mapped against the FPR (1 - specificity) for specified cut-off values. Preferably, a point is selected across the curve’s shoulder, reducing false positives while optimizing real positives.

4. Experiment Setup

All the studies conducted used Python, Keras library, TensorFlow as a back-end. The resolution of the images has been standardized to a uniform size, following each model’s input requirements. The epoch number set at 50 because of the use of new CNN trained on weights in our experiments. The distribution of training/testing dataset at 80/20. Stratified standard preference made to ensure a nearly similar dispersion of the class. Mini-batch size set to 32, and the categorical cross-entropy loss function was selected due to its suitability for multi-class classification tasks. The default RMSProp was the Optimiser. The primary assessment metric for Accuracy, Specificity, and Sensitivity of test data was used for final scores validation.

4.1. Performance enhancement

In this article, CNN recommended that the classification of the DED dataset be included. A deep convolutionary neural network (CNN) converts a function vector with a fixed weight matrix to obtain specific representations of features without losing spatial arrangement information[46]. Optimizer selection is a vital component of the neural network, helping to pick which one to use for the model by understanding how they function. Several hyperparameters could be modified to increase the efficiency of the neural network. Not all of them, however, have a significant influence on the efficiency of the network. Not all of them, however, have a significant influence on the efficiency of the network. The optimizer is one of the parameters that could allow the adjustment between the algorithm assemblies or set-off. There are different optimizers we have picked from, the most widely used ones.

RMSprop Optimizer The RMSprop is an unpublished method proposed by Geoff Hinton based on adaptive learning rate [47]. The RMSprop optimizer is equivalent with momentum to the gradient descent algorithm. In the vertical direction, this optimizer limits the oscillations. Thus, we can maximize our learning rate, and our algorithm will take larger steps to converge faster in the horizontal direction. The discrepancy in how the gradients are measured is between RMSprop. For the RMSprop with momentum, the following equations illustrate how the gradients are determined. Hinton suggested the momentum value is normally set to 0.9 and a good default value for the learning rate η is 0.001.

$$D[f^2]_t = 0.9D[f^2]_{t-1} + 0.1f_t^2 \quad (16)$$

$$\beta_{t+1} = \beta_t - \frac{\eta}{\sqrt{D[f^2]_t + \epsilon}} f_t \quad (17)$$

RMSprop's gist is to maintain a moving average of the gradient square and divide the gradient by the root of this average ².

5. Results

The proposed CNN model has obtained accuracy in the test dataset. The efficiency acquired by the model has been used to demonstrate That multi-class classification in DED can be improved by improving the quality of the training images and using the right parameters for the model. The highest sensitivity and specificity was identified by proposed model in RMSprop optimizer. Finally, the accuracy obtained is shown in Table.3.

This analysis is a study of multi-class DED classification using the DL algorithm. A minimum of 80 percent sensitivity and 95 percent precision for sight-threatening DED identification must be obtained by any method[48] under the British Diabetic Association (BDA) guidelines. In scenario I, we achieved a maximum sensitivity of 100 percent and a maximum specificity of 100 percent, respectively, after checking our strategy in DED detection tasks. Thus, according to the BDA criteria, multi-class DED detection is adequate for its sensitivity and specificity.

5.1. Visualizing feature map

The feature maps, or activation maps, record the input applied with filters, such as the source images or other feature maps. The purpose of visualizing a feature map for particular source images would explain what attributes in the feature maps are observed or retained. The idea would be that the feature maps near the input detect fine-grained or small information while featuring maps near the model output to capture more distinctive characteristics. The first layer of CNN always learns features like edges, line patterns, color, and deeper layer network to identify more complex features like pathological lesions. Later layers construct their features by merging features from previous layers. To analyze the visualization of feature maps, we used the highest performed model with fundus retinal images, i.e., proposed CNN model, and used to create activations. The activations for CNN network models shown in Fig. 5.

5.2. Explaining Proposed Model using Grad-CAM

To make deep learning more practical and explainable, a range of work was performed [49–51]. It is also essential to make the deep neural network more interpretative in various deep learning applications linked to medical imaging. A technique of Gradient

Weighted Class Activation Mapping (Grad-CAM) is developed by Selvaraju et al. [52], which provides an illustrative view of deep learning techniques. The technique of Grad-CAM offers a visual description for any deeply related neural network. This helps to decide more about the model when conducting identification or prediction tasks. The simple retinal fundus image is given as input and uses the proposed model as a detection method. After calculating the predicted label using the full model, Grad-CAM is applied to the last Convolution layer. Fig. 6 shows the heatmap visualization on various retinal fundus images by the proposed model.

6. Discussion

This research focuses on DL algorithms to classify multi-class DED automatically. Previous studies in this field found that the new DL algorithms are ineffective in classifying multi-class classification of fundus images from small datasets. For a computer-aided medical application, it failed to produce realistic and efficient results. Consequently, because of each disease's significance, this article adapted the optimized DL architectures for the automated classification of regular, DR, DME, GL, and Ca to create an automatic framework for the classification.

In the proposed CNN model multi-class DED classification, the performance was reduced by 3 percent [53]. As DED fundus images consist of subtle features that can be crucial for diagnosis, this finding is prevalent. Interestingly, the most widely implemented architectures were designed to classify features based on artifacts such as those found in the ImageNet dataset. However, a new mechanism such as lesion-based (e.g., exudates) may be required to diagnose diseases using CNN models. Our future goals include multi-class classification of segmented DED lesions (region of interest)[54, 55][56] for improving the accuracy of multi-class DED identification and progressing to more complicated and beneficial multi-grade disease identification.

Automated DED recognition has been the topic of several studies in the past, with the main emphasis on healthy/unhealthy binary retinal classification[57]. In the case of multi-class classification, the performance of DL models has been reduced as categories have multiplied. When categories increased, the predicted precision of the random distribution decreased. This finding corresponded to the previous studies[58]. Recent research using the GoogLeNet model to identify skin cancer has shown that increase in the number of classes has underperformed (with an accuracy of 72.1% for a three class problem and 55.4% for a nine class problem)[59]. Thus it is essential to create disease-specific strategies to differentiate between DED

²<https://keras.io/api/optimizers/rmsprop/>

Table 3. Average performance of the models in multi-class DED classification.

Model	Optimiser	Learning Rate	Accuracy	Sensitivity	Specificity
CNN	<i>RMSprop</i>	1e-3	81.33%	100%	100%

Legend:

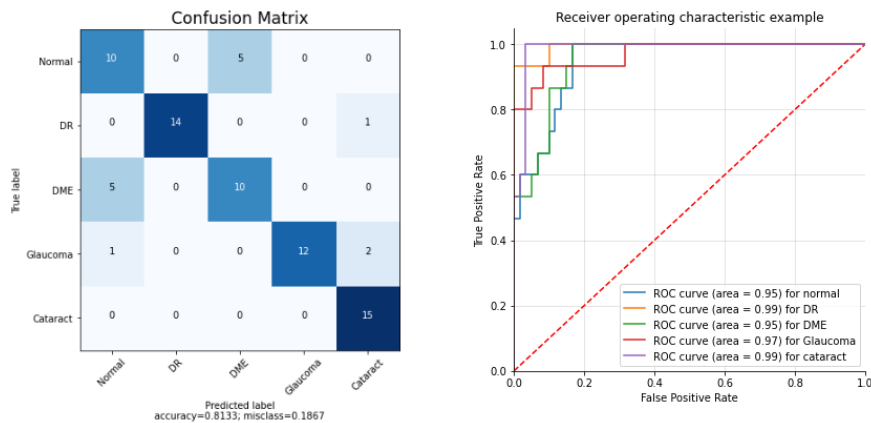


Figure 4. Confusion Matrix and ROC Curve obtained by proposed CNN

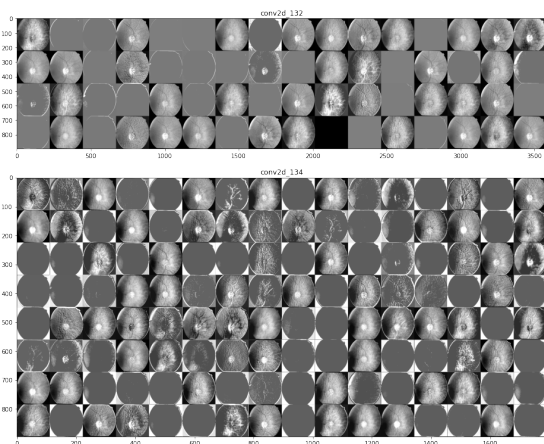


Figure 5. Visual feature maps in first layer and deep layer

to enhance the efficiency of multi-class classification. Therefore, the research proposed a system that focuses entirely on the identification of multi-class DED among healthy instances, as discussed in previous studies. According to the empiric nature of DL, a variety of performance optimization techniques have been applied (i) optimizer choice, (ii) data augmentation, and (iii) contrast enhancement. Besides, the study used combined datasets from different sources to evaluate the system’s robustness in its flexibility to cope with real world scenarios. As Wan et al.[60] have pointed out, the single data collection environment presents difficulties in the validation of accurate models[61, 62].

7. Conclusions

Images of the retinal fundus are a popular and useful instrument and are used to provide accurate DED details. Such photographs can easily show injuries, anomalies and help to prevent permanent loss of vision. However, it is challenging to identify DEDs through retinal fundus images accurately, and even highly experienced ophthalmologists are prone to misdiagnose eye lesions. Severe DEDs, which require rapid diagnosis and treatment, cause irreversible vision loss, visual imparities, and vision distortion disorders. Consequently, to assist in the diagnosis, it is essential to use DL methods. Multiple diseases may represent one collection of fundus photos. Using a single picture for diagnosis with a fundus image, which is well examined by traditional methods, is inaccurate. Furthermore, it requires a large amount of time to mark a series of retinal fundus images one by one. For this analysis, we, therefore, used publicly accessible and annotated photographs of the fundus. This research suggested a model that learns the characteristics of fundus images in retinal fundus photography and their feature dependencies for multi-class classification. We grouped images of the retinal fundus into five types of DED. The findings presented in this work show that DL algorithms can automatically identify the form of DED. This technology may have possible clinical applications, and may enhance healthcare delivery by identifying different acute DED diseases.

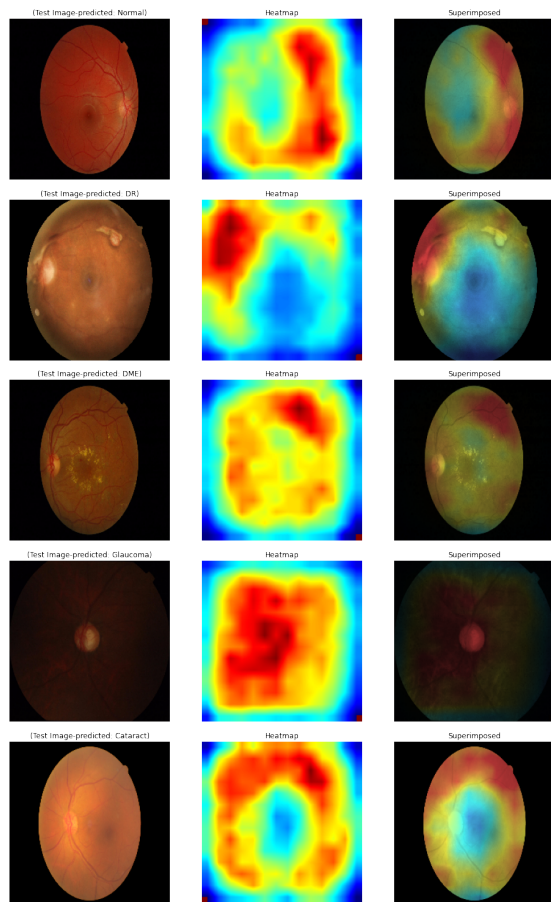


Figure 6. Visualisation on fundus retinal images of Normal/DR/DME/Glaucoma/Cataract infected using Grad-CAM on the proposed model.

References

- [1] CONGDON, N.G., FRIEDMAN, D.S. and LIETMAN, T. (2003) Important causes of visual impairment in the world today. *Jama* **290**(15): 2057–2060.
- [2] CHEN, Z.G., HUI ZHAN, Z., WANG, H. and ZHANG, J. (2020) Distributed individuals for multiple peaks: A novel differential evolution for multimodal optimization problems. *IEEE Transactions on Evolutionary Computation* **24**: 708–719.
- [3] WANG, H. and SUN, L. (2010) Trust-involved access control in collaborative open social networks. *2010 Fourth International Conference on Network and System Security* : 239–246.
- [4] GAO, W., PENG, M., WANG, H., ZHANG, Y., XIE, Q. and TIAN, G. (2018) Incorporating word embeddings into topic modeling of short text. *Knowledge and Information Systems* : 1–23.
- [5] WANG, H., SUN, L. and BERTINO, E. (2014) Building access control policy model for privacy preserving and testing policy conflicting problems. *J. Comput. Syst. Sci.* **80**: 1493–1503.
- [6] SIULY, S., KHARE, S.K., BAJAJ, V., WANG, H. and ZHANG, Y. (2020) A computerized method for automatic detection of schizophrenia using eeg signals. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **28**: 2390–2400.
- [7] SIULY, S., ALÇIN, Ö.F., KABIR, E., ŞENGÜR, A., WANG, H., ZHANG, Y. and WHITTAKER, F. (2020) A new framework for automatic detection of patients with mild cognitive impairment using resting-state eeg signals. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **28**: 1966–1976.
- [8] GULSHAN, V., PENG, L., CORAM, M., STUMPE, M.C., WU, D., NARAYANASWAMY, A., VENUGOPALAN, S. *et al.* (2016) Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *Jama* **316**(22): 2402–2410.
- [9] ABRÀMOFF, M.D., REINHARDT, J.M., RUSSELL, S.R., FOLK, J.C., MAHAJAN, V.B., NIEMEIJER, M. and QUELLEC, G. (2010) Automated early detection of diabetic retinopathy. *Ophthalmology* **117**(6): 1147–1154.
- [10] GARGEYA, R. and LENG, T. (2017) Automated identification of diabetic retinopathy using deep learning. *Ophthalmology* **124**(7): 962–969.
- [11] HU, H., LI, J., WANG, H. and DAGGARD, G. (2006) Combined gene selection methods for microarray data analysis. In *International conference on knowledge-based and intelligent information and engineering systems*: 976–983.
- [12] KHALIL, F., WANG, H. and LI, J. (2007) Integrating markov model with clustering for predicting web page accesses. In *13th Australasian world wide web conference (AusWeb07)*: 63–74.
- [13] SARKI, R., AHMED, K., WANG, H. and ZHANG, Y. (2020) Automated detection of mild and multi-class diabetic eye diseases using deep learning. *Health Information Science and Systems* **8**(1): 1–9.
- [14] TORREY, L. and SHAVLIK, J. (2010) Transfer learning. In *Handbook of research on machine learning applications and trends: algorithms, methods, and techniques* (IGI Global), 242–264.
- [15] MOOKIAH, M.R.K., ACHARYA, U.R., CHUA, C.K., LIM, C.M., NG, E. and LAUDE, A. (2013) Computer-aided diagnosis of diabetic retinopathy: A review. *Computers in biology and medicine* **43**(12): 2136–2155.
- [16] BARGSHADY, G., ZHOU, X., DEO, R.C., SOAR, J., WHITTAKER, F. and WANG, H. (2020) Enhanced deep learning algorithm development to detect pain intensity from facial expression images. *Expert Systems with Applications* **149**: 113305.
- [17] LI, X., PANG, T., XIONG, B., LIU, W., LIANG, P. and WANG, T. (2017) Convolutional neural networks based transfer learning for diabetic retinopathy fundus image classification. In *2017 10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)* (IEEE): 1–11.
- [18] KARRI, S.P.K., CHAKRABORTY, D. and CHATTERJEE, J. (2017) Transfer learning based classification of optical coherence tomography images with diabetic macular edema and dry age-related macular degeneration. *Biomedical optics express* **8**(2): 579–592.
- [19] HU, H., LI, J., WANG, H., DAGGARD, G. and SHI, M. (2006) A maximally diversified multiple decision tree algorithm for microarray data classification. In *Intelligent Systems*

- For *Bioinformatics*: 35–38.
- [20] KHALIL, F., LI, J. and WANG, H. (2009) An integrated model for next page access prediction. *Int. J. Knowl. Web Intell.* 1: 48–80.
- [21] KHALIL, F., LI, J. and WANG, H. (2006) A framework of combining markov model with association rules for predicting web page accesses. In *AusDM*: 177–184.
- [22] DE FAUW, J., LEDSAM, J.R., ROMERA-PAREDES, B., NIKOLOV, S., TOMASEV, N., BLACKWELL, S., ASKHAM, H. et al. (2018) Clinically applicable deep learning for diagnosis and referral in retinal disease. *Nature medicine* 24(9): 1342–1350.
- [23] ARUNKUMAR, R. and KARTHIGAIKUMAR, P. (2017) Multi-retinal disease classification by reduced deep learning features. *Neural Computing and Applications* 28(2): 329–334.
- [24] WANG, Y., ZHANG, Y., YAO, Z., ZHAO, R. and ZHOU, F. (2016) Machine learning based detection of age-related macular degeneration (amd) and diabetic macular edema (dme) from optical coherence tomography (oct) images. *Biomedical optics express* 7(12): 4928–4940.
- [25] DE FAUW, J., KEANE, P., TOMASEV, N., VISENTIN, D., VAN DEN DRIESSCHE, G., JOHNSON, M., HUGHES, C.O. et al. (2016) Automated analysis of retinal imaging using machine learning techniques for computer vision. *F1000Research* 5.
- [26] CAIXINHA, M. and NUNES, S. (2017) Machine learning techniques in clinical vision sciences. *Current eye research* 42(1): 1–15.
- [27] YOO, T.K. and PARK, E.C. (2013) Diabetic retinopathy risk prediction for fundus examination using sparse learning: a cross-sectional study. *BMC medical informatics and decision making* 13(1): 106.
- [28] CHEN, X., XU, Y., WONG, D.W.K., WONG, T.Y. and LIU, J. (2015) Glaucoma detection based on deep convolutional neural network. In *2015 37th annual international conference of the IEEE engineering in medicine and biology society (EMBC) (IEEE)*: 715–718.
- [29] YOO, T.K. and HONG, S. (2015) Artificial neural network approach for differentiating open-angle glaucoma from glaucoma suspect without a visual field test. *Investigative Ophthalmology & Visual Science* 56(6): 3957–3966.
- [30] ASAOKA, R., MURATA, H., IWASE, A. and ARAIE, M. (2016) Detecting preperimetric glaucoma with standard automated perimetry using a deep learning classifier. *Ophthalmology* 123(9): 1974–1980.
- [31] GAO, X., LIN, S. and WONG, T.Y. (2015) Automatic feature learning to grade nuclear cataracts based on deep learning. *IEEE Transactions on Biomedical Engineering* 62(11): 2693–2701.
- [32] LECUN, Y., BENGIO, Y. and HINTON, G. (2015) Deep learning. *nature* 521(7553): 436–444.
- [33] HE, J., RONG, J., SUN, L., WANG, H., ZHANG, Y. and MA, J. (2020) A framework for cardiac arrhythmia detection from iot-based ecgs. *World Wide Web* : 1–16.
- [34] MA, J., SUN, L., WANG, H., ZHANG, Y. and AICKELIN, U. (2016) Supervised anomaly detection in uncertain pseudoperiodic data streams. *ACM Transactions on Internet Technology (TOIT)* 16(1): 1–20.
- [35] BURLINA, P., FREUND, D.E., JOSHI, N., WOLFSON, Y. and BRESSLER, N.M. (2016) Detection of age-related macular degeneration via deep learning. In *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI) (IEEE)*: 184–188.
- [36] LEE, C.S., BAUGHMAN, D.M. and LEE, A.Y. (2017) Deep learning is effective for classifying normal versus age-related macular degeneration oct images. *Ophthalmology Retina* 1(4): 322–327.
- [37] KWASIGROCH, A., JARZEMBINSKI, B. and GROCHOWSKI, M. (2018) Deep cnn based decision support system for detection and assessing the stage of diabetic retinopathy. In *2018 International Interdisciplinary PhD Workshop (IIPhDW) (IEEE)*: 111–116.
- [38] LAM, C., YI, D., GUO, M. and LINDSEY, T. (2018) Automated detection of diabetic retinopathy using deep learning. *AMIA Summits on Translational Science Proceedings* 2018: 147.
- [39] CHOI, J.Y., YOO, T.K., SEO, J.G., KWAK, J., UM, T.T. and RIM, T.H. (2017) Multi-categorical deep learning neural network to classify retinal images: a pilot study employing small database. *PloS one* 12(11).
- [40] SUN, K. and SANG, N. (2008) Morphological enhancement of vascular angiogram with multiscale detected by gabor filters. *Electronics Letters* 44(2): 86–87.
- [41] HASSANPOUR, H., SAMADIANI, N. and SALEHI, S.M. (2015) Using morphological transforms to enhance the contrast of medical images. *The Egyptian Journal of Radiology and Nuclear Medicine* 46(2): 481–489.
- [42] IOFFE, S. and SZEGEDY, C. (2015) Batch normalization: Accelerating deep network training by reducing internal covariate shift. *arXiv preprint arXiv:1502.03167* .
- [43] SRIVASTAVA, N., HINTON, G., KRIZHEVSKY, A., SUTSKEVER, I. and SALAKHUTDINOV, R. (2014) Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research* 15(1): 1929–1958.
- [44] BALDI, P. and SADOWSKI, P.J. (2013) Understanding dropout. In *Advances in neural information processing systems*: 2814–2822.
- [45] SOKOLOVA, M. and LAPALME, G. (2009) A systematic analysis of performance measures for classification tasks. *Information processing & management* 45(4): 427–437.
- [46] YAMASHITA, R., NISHIO, M., DO, R.K.G. and TOGASHI, K. (2018) Convolutional neural networks: an overview and application in radiology. *Insights into imaging* 9(4): 611–629.
- [47] HINTON, G., SRIVASTAVA, N. and SWERSKY, K. (2012) Neural networks for machine learning lecture 6a overview of mini-batch gradient descent. *Cited on* 14(8).
- [48] ASSOCIATION, B.D. et al. (1997) Retinal photography screening for diabetic eye disease. *London: BDA* .
- [49] ZHANG, J., LI, H., LIU, X., LUO, Y., CHEN, F., WANG, H. and CHANG, L. (2015) On efficient and robust anonymization for privacy protection on massive streaming categorical information. *IEEE Transactions on Dependable and Secure Computing* 14(5): 507–520.
- [50] SUN, X., LI, M., WANG, H. and PLANK, A. (2008) An efficient hash-based algorithm for minimal k-anonymity. In *Conferences in Research and Practice in Information Technology (CRPIT) (Australian Computer Society Inc.)*, 74: 101–107.

- [51] RASOOL, R.U., ASHRAF, U., AHMED, K., WANG, H., RAFIQUE, W. and ANWAR, Z. (2019) Cyberpulse: a machine learning based link flooding attack mitigation system for software defined networks. *IEEE Access* 7: 34885–34899.
- [52] SELVARAJU, R.R., COGSWELL, M., DAS, A., VEDANTAM, R., PARIKH, D. and BATRA, D. (2017) Grad-cam: Visual explanations from deep networks via gradient-based localization. In *Proceedings of the IEEE international conference on computer vision*: 618–626.
- [53] SARKI, R., AHMED, K., WANG, H. and ZHANG, Y. (2020) Automatic detection of diabetic eye disease through deep learning using fundus images: A survey. *IEEE Access* 8: 151133–151149.
- [54] PANDEY, D., YIN, X., WANG, H., SU, M.Y., CHEN, J.H., WU, J. and ZHANG, Y. (2018) Automatic and fast segmentation of breast region-of-interest (roi) and density in mris. *Heliyon* 4(12): e01042.
- [55] PANDEY, D., YIN, X., WANG, H. and ZHANG, Y. (2017) Accurate vessel segmentation using maximum entropy incorporating line detection and phase-preserving denoising. *Computer Vision and Image Understanding* 155: 162–172.
- [56] YIN, X., NG, B.W., HE, J., ZHANG, Y. and ABBOTT, D. (2014) Accurate image analysis of the retina using hessian matrix and binarisation of thresholded entropy with application of texture mapping. *PLoS One* 9(4): e95943.
- [57] PRATT, H., COENEN, F., BROADBENT, D.M., HARDING, S.P. and ZHENG, Y. (2016) Convolutional neural networks for diabetic retinopathy. *Procedia Computer Science* 90: 200–205.
- [58] DENG, J., BERG, A.C., LI, K. and FEI-FEI, L. (2010) What does classifying more than 10,000 image categories tell us? In *European conference on computer vision* (Springer): 71–84.
- [59] ESTEVA, A., KUPREL, B., NOVOA, R.A., KO, J., SWETTER, S.M., BLAU, H.M. and THRUN, S. (2017) Dermatologist-level classification of skin cancer with deep neural networks. *Nature* 542(7639): 115–118.
- [60] WAN, S., LIANG, Y. and ZHANG, Y. (2018) Deep convolutional neural networks for diabetic retinopathy detection by image classification. *Computers & Electrical Engineering* 72: 274–282.
- [61] SARKI, R., AHMED, K. and ZHANG, Y. (2020) Early detection of diabetic eye disease through deep learning using fundus images. *EAI Endorsed Transactions on Pervasive Health and Technology* 6(22). doi:10.4108/eai.13-7-2018.164588.
- [62] SUN, X., WANG, H., LI, J. and PEI, J. (2011) Publishing anonymous survey rating data. *Data Mining and Knowledge Discovery* 23(3): 379–406.