



AOCT-NET: a convolutional network automated classification of multiclass retinal diseases using spectral-domain optical coherence tomography images

Ali Mohammad Alqudah, MSc¹

Received: 11 May 2019 / Accepted: 2 November 2019 / Published online: 14 November 2019
© International Federation for Medical and Biological Engineering 2019

Abstract

Since introducing optical coherence tomography (OCT) technology for 2D eye imaging, it has become one of the most important and widely used imaging modalities for the noninvasive assessment of retinal eye diseases. Age-related macular degeneration (AMD) and diabetic macular edema eye disease are the leading causes of blindness being diagnosed using OCT. Recently, by developing machine learning and deep learning techniques, the classification of eye retina diseases using OCT images has become quite a challenge. In this paper, a novel automated convolutional neural network (CNN) architecture for a multiclass classification system based on spectral-domain optical coherence tomography (SD-OCT) has been proposed. The system used to classify five types of retinal diseases (age-related macular degeneration (AMD), choroidal neovascularization (CNV), diabetic macular edema (DME), and drusen) in addition to normal cases. The proposed CNN architecture with a softmax classifier overall correctly identified 100% of cases with AMD, 98.86% of cases with CNV, 99.17% cases with DME, 98.97% cases with drusen, and 99.15% cases of normal with an overall accuracy of 95.30%. This architecture is a potentially impactful tool for the diagnosis of retinal diseases using SD-OCT images.

Keywords Retina · Optical coherence tomography · Spectral domain · Classification · Deep learning

1 Introduction

The retina in the human eye consists of a photosensitive layer of optic nerve tissue, which is lining the inner surface of the eyeball. This layer receives the focused light by the lens and converts it into neural signals [1]. The macula is the main region for sensing purpose, which is in the central part of the retina and it contains special layers of photoreceptor nerve cells that are responsible for color, light-intensity, and fine visual details detecting. Finally, the retina processes the information acquired by the macula and sends them to the brain via the optic nerve for visual recognition [1, 2].

There are many eye diseases that can affect retina such as age-related macular degeneration (AMD), choroidal Neovascularization (CNV), drusen, and diabetic macular edema (DME). These eye diseases are among the most common causes

of vision loss in modern communities, recent predictions of exponentially growing DME cases to affect over 300 million people worldwide in the next few years [3–5]. These diseases can cause blindness and affect the lives of patients. Based on the mentioned predictions scientists were forced to provide a new and effective tool for the diagnosis of such eye diseases [6, 7].

One of the most common imaging modalities for eye diagnosis is optical coherence tomography (OCT) with around 5.35 million OCTs performed in 2014 in the USA only. OCT is a noninvasive imaging technique which is useful to make a diagnosis of eye retina diseases [8]. Using OCT imaging, ophthalmologists see distinctive retinal layers that allow them to measure and map these layers and provide guidance for treatment [9, 10]. OCT was able to capture and generate images in two different domains: the time domain and the spectral domain. In general, spectral-domain OCT has led to a significant improvement in imaging performance and provides more accurate information about intraretinal morphology of retaining layers that can be used for disease detection and classification [11].

✉ Ali Mohammad Alqudah, MSc
ali_qudah@hotmail.com

¹ Department of Biomedical Systems and Informatics Engineering, Yarmouk University, Irbid, Jordan

Recently, with the development of artificial intelligence (AI) especially machine learning (ML) and deep learning (DL) algorithms, many researchers focus on applying them to medical data to enhance the medical decision-making systems [12–14]. As one of the largest medical fields, researchers focused on eye diagnosis using ML and DL algorithms. One of the main diseases that affect the eye are retina diseases, and many diagnosis systems and methods were proposed [15–17]. With ML and DL techniques, we can implement systems that can be used for identification, segmentation, and classification of eye retina diseases, but the main need for such research is the implementation, collection, and usage of a huge retina OCT images either in time or in spectral domain, mainly in spectral-domain OCT such as proposed by [1] and by [2].

Deep learning is the state of the art AI technologies used to implement computational models consisting of multiple layers used for data processing to learn how to represent data at different levels of abstraction [12, 15]. The convolutional neural network (CNN) technique is a state of the art DL technique for medical disease diagnosis and especially for data based on images such as X-ray images and OCT images in cases of eye disease diagnosis [12]. CNN, in general, is designed to minimize data preprocessing and to deal with raw images. It consists of many layers as follows: the input layer, convolution layer, RELU layer, fully connected layer, classification layer, and output layer [18]. CNN is based mainly on two processes: the first one is convolution, and the second one is downsampling. The convolution processes are performed using trainable filters with pre-specified size, and weights are adjusting during training phase [19]. Two main techniques are used in medical diagnosis using CNN, transferee learning and network designing. The first one is most widely used and based on transferring pretrained CNN models, and making them able to classify

new categories instead of their original categories as presented by [16]. The second technique is to design a specialized CNN for a certain application such as classifying eye retinal diseases as proposed by [15]. Fig. 1 shows the path to developing a smart system for SD-OCT classification.

In this work, a deep learning-based system for automatic multiclass classification of retinal diseases based on SD-OCT raw images that meet the higher accuracy was developed. The system is based on a novel CNN architecture for classifying five types of retinal states (four diseases and normal cases) with high accuracy and a huge SD-OCT image dataset. Five folders are created, each one consists of the collected OCT images for specific classes AMD, CNV, DME, drusen, and normal; the data is partitioned into training and testing data, 70% of data is utilized in the training stage and the rest is used in the test stage, then, the designed CNN is applied to this paper. The results show that the proposed system can achieve 97.78% accuracy, 97.78% sensitivity, and 97.78% specificity rates. The contributions of this work can be summarized as follows:

- A new CNN architecture has been proposed and trained.
- First CNN-based system to classify five classes of retina states.
- Using a huge dataset of SD-OCT images collected from two different datasets available online.

The paper is organized as follows: Section II provides details about recent related works. Section III presents with a detailed explanation of the used dataset and the proposed architecture. Section IV is the results, including the performance of the proposed CNN classifier. Section V is the discussion about the method results. And finally, section IV represents the conclusion of the work.

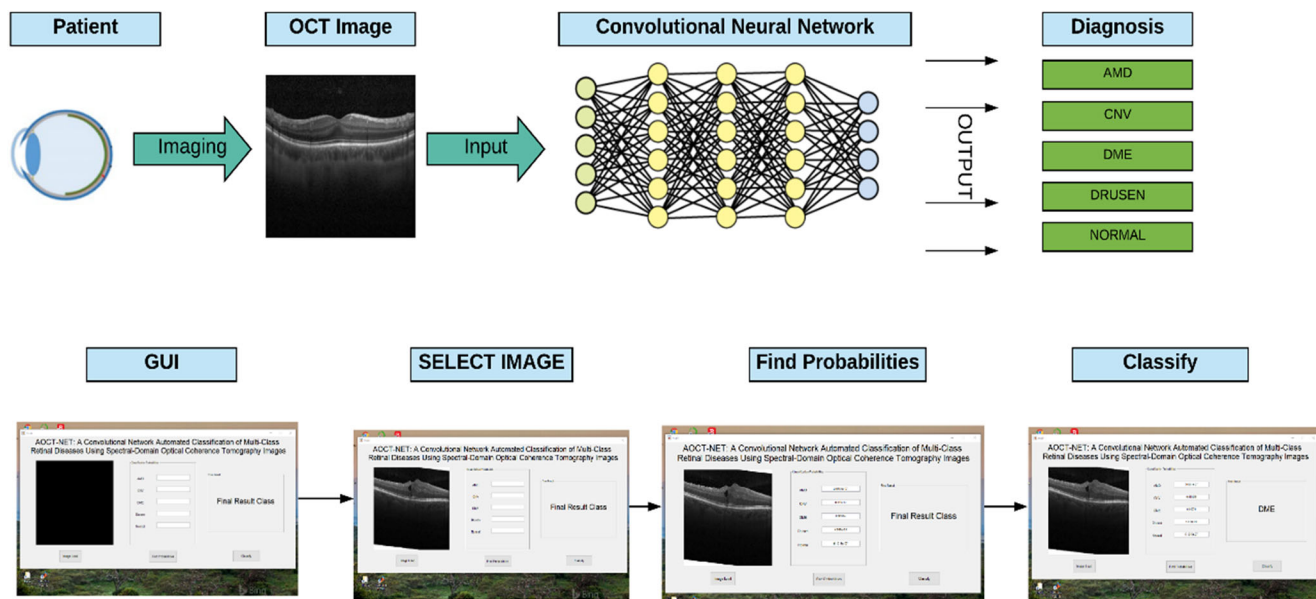


Fig. 1 Approach to developing a deep learning-based system for eye diagnosing using SD-OCT

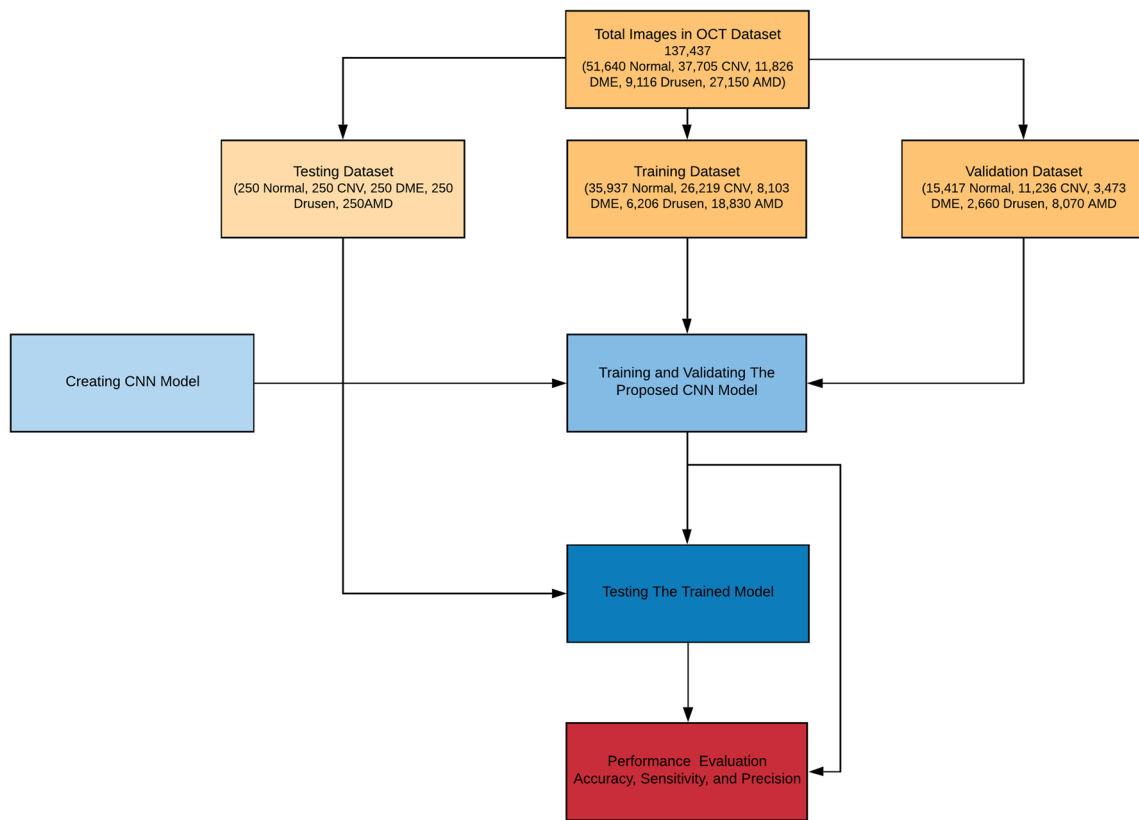


Fig. 2 Representative AOCT-NET workflow diagram

2 Related works

Machine learning and deep learning methods are more widely used to identify and classify retinal eye diseases. In this section, the most recent and related works to the paper’s topic are discussed. Karri et al. [16] proposed a GoogleNet transferee learning technique–based system, the system is used to classify three cases: normal, AMD, and DME. The system accuracy was computed for 10 experiments which are repeated to get the best model, and the highest accuracy gained was 96%. A CNN-based OCT classification system is proposed by Rong et al. [17], at first, image denoising is performed to reduce the image noise, then thresholding and morphological dilation are applied consequently to extract the masks and generate a dataset of surrogate images. The results show that the area

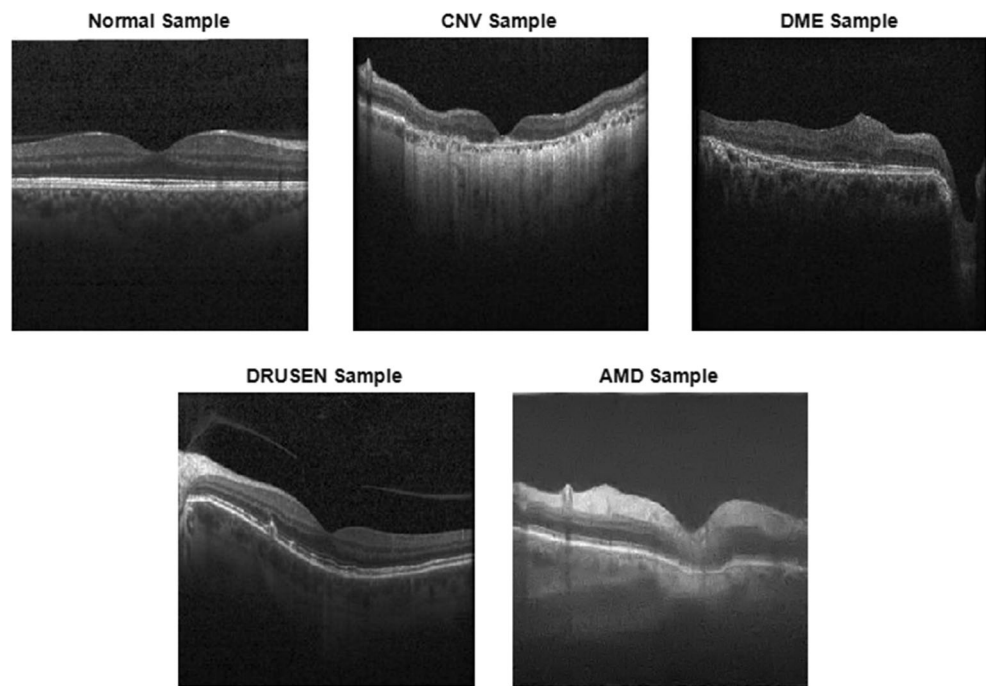
under the ROC curve (AUC) of the system is 0.9783 and 0.9856 for the local and Duke databases respectively.

Fang et al. [18] proposed an iterative fusion CNN system for the classification of optical coherence tomography. The proposed methodology will adapt the network using iterative fusion strategy by using and combining all features in current and previous convolutional layers. The system was used to classify four different classes (normal, DME, drusen, and CNV); the results show that the IFCNN overall accuracy was $87.3 \pm 2.2\%$. A macular OCT classification system by implementing an ensemble of a multiscale CNN was proposed by Rasti et al. [19]. The proposed system is applied to identify normal retina and two common types of macular: dry age-related macular degeneration (AMD) and diabetic macular edema (DME). The proposed AUC system was 0.9985.

Table 1 The distribution of images used in the system

Case	Number of training images	Number of validation images	Number of testing images	Total number of total images
Normal	35,973	15,417	250	51,640
CNV	26,219	11,236	250	37,705
DME	8103	3473	250	11,826
Drusen	6206	2660	250	9116
AMD	18,830	8070	250	27,150
Total	95,331	40,856	1250	137,437

Fig. 3 Samples of the images in the system for five classes



Amil et al. [20] proposed an unsupervised feature extraction method for anterior chamber OCT images for purposes of ordering and classification. The unsupervised method consists of three main steps: (1) images preprocess, (2) distance measurement between every pair of images is computed, (3) applying a machine learning algorithm that uses the measured distance for ordering and classifying the images. The system is applied to a large dataset consisting of 1000 images.

A fully automated system for the detection of diabetic macular edema and dry age-related macular degeneration using OCT images was proposed by Srinivasan et al. [21]. The proposed system uses multiscale histograms of oriented gradients (HOG) as features set to be fed to support vector machine (SVM) classifier. The system passes through three different stages before extracting features defined as follows: image denoising using BM3D, flattening retina curvature, and finally cropping the ROI. Then, cross-validation was applied on volumetric scans acquired from 45 subjects distributed equally between three cases; the proposed classification system was able to correctly identify 100% of cases with AMD and DME, while for normal cases, the system was able to identify only 86.67%.

A system that augmented OCT images with patient information such as age, gender, and visual acuity data for multiclass multilabel classification of OCT images was proposed by Mehta et al. [22]. The proposed system was tested to classify four different pathologies as follows epiretinal membrane, diabetic macular edema, dry age-related macular degeneration, and neovascular age-related macular degeneration. The system achieves an overall accuracy of 86% and an exact match of 66.7%. Hussain et al. [23] proposed a novel classification model for identifying patients with age-related

macular degeneration (AMD) or diabetic macular edema (DME). The model uses retinal-based features extracted from SD-OCT images. Features include the thickness of the retina, the thickness of individual retinal layers, the volume of the detected pathologies, and hyper-reflective intra-retinal spots. The model was tested on 251 (59 normal, 177 AMD, and 15 DME) subjects using 15-fold cross-validation tests. The results show that the accuracy was more than 96%.

Ji et al. [24] proposed a transfer learning based on the Inception V3 pretrained CNN system. The system is used for the classification of two types of diseases (dry AMD and DME) in addition to normal cases from retinal OCT images; the overall system accuracy was above 98% using Beijing clinic datasets. While, Perdomo et al. [25] designed a CNN (OCT-NET) consisting of 12 layers. The proposed CNN was used for the classification of DME using SD-OCT volumes.

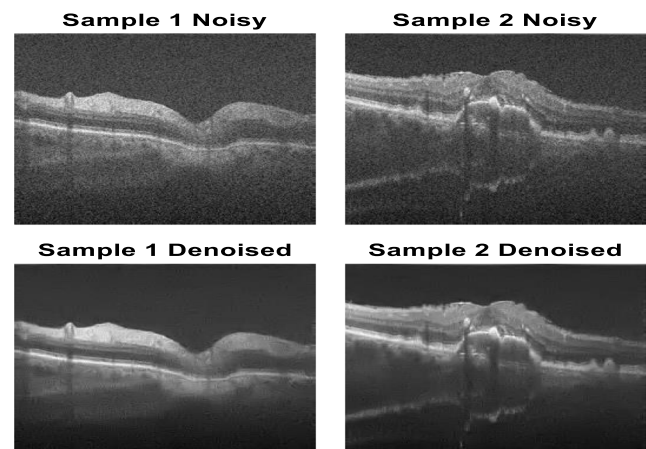


Fig. 4. Sample of noisy and denoised images

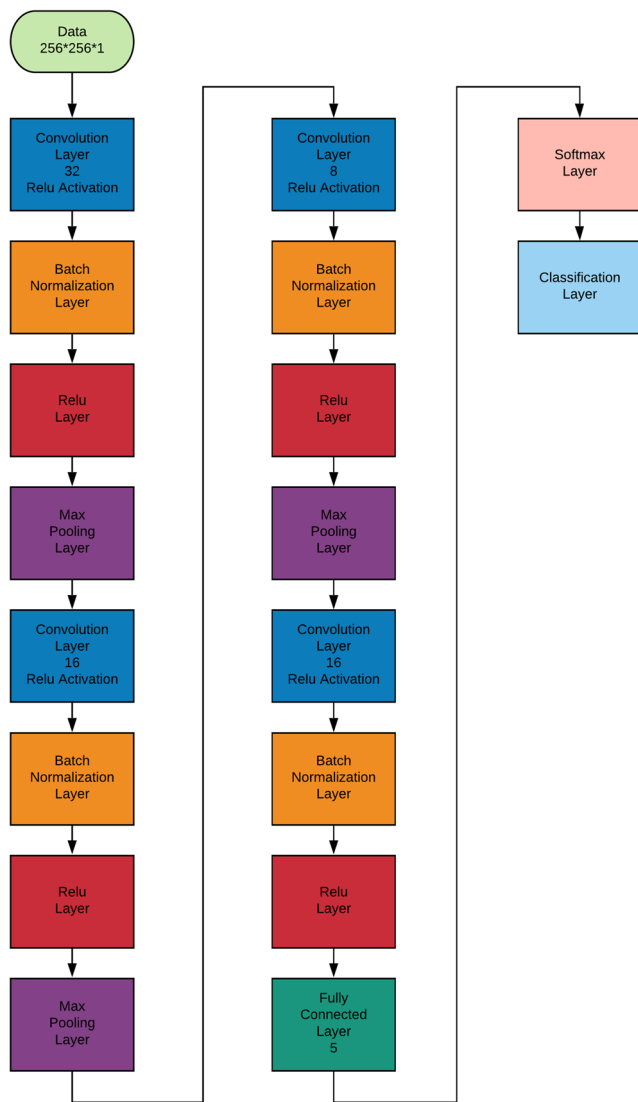


Fig. 5 Proposed CNN Architecture

The proposed CNN was tested on the Singapore Eye Research Institute (SERI) database which contains 32 SD-OCT volumes, and the overall accuracy was $93.75 \pm 3.125\%$.

Hwang et al. [26] proposed a cloud-based artificial intelligence telemedicine web tool for decision-making about the diagnosis and treatment of AMD. They used 35,900 labeled OCT images obtained from AMD patients to train three different pre-trained architectures of CNNs to build the tool. The CNN pre-trained architectures are VGG16, InceptionV3, and ResNet50. The results show that the used architectures achieved an accuracy rate of 91.20%, 96.93%, and 95.87% for VGG16, InceptionV3, and ResNet50, respectively, in addition to high specificity and sensitivity rates with values not less than 90%.

Nugroho [27] proposed a comparison between two methods of feature extractions, handcrafted and deep learning-based, to select the best features and model that can be used to classify the SD-OCT images. The handcrafted features are histogram of oriented gradient (HOG) and local binary pattern (LBP), while, the deep learning features were extracted using DenseNet-169 and ResNet50. The result shows that the deep learning-based features outperformed the handcrafted based features with an accuracy of 88% and 89% for DenseNet and ResNet respectively while it only was 50% and 42% for HOG and LBP respectively.

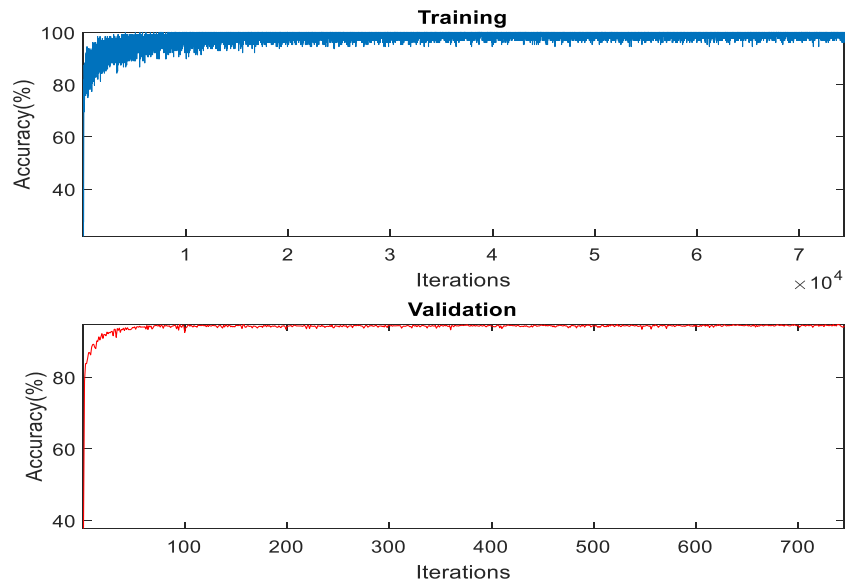
3 Methodology

In this section, the proposed methodology is discussed; the discussion will include used datasets in the system implementation and image denoising and resizing, and the proposed CNN architecture is discussed in detail. Figure 2 shows a flow chart diagram of the proposed methodology.

Table 2 Layers information for proposed CNN architecture

#	Layer	Information	#	Layer	Information		
1	Input layer	Size	256 × 256	9	Maxpol_2	Kernel size 2 × 2 Stride 2 × 2	
2	Conv_1	Number of filters Kernel size Activation	32 3 × 3 RELU	10	Conv_3	Number of filters Kernel size Activation	8 3 × 3 RELU
3	Batch_Norm_1	Number of channels	32	11	Batch_Norm_3	Number of channels	8
5	Maxpol_1	Kernel size Stride	2 × 2 2 × 2	13	Maxpol_3	Kernel size Stride	2 × 2 2 × 2
6	Conv_2	Number of filters Kernel size Activation	16 3 × 3 RELU	14	Conv_4	Number of filters Kernel size Activation	16 3 × 3 RELU
7	Batch_Norm_2	Number of channels	16	15	Batch_Norm_4	Number of channels	16

Fig. 6 Accuracy is plotted against the iterations for both training and validation



3.1 Dataset

In this research, the OCT dataset created by Zhang Lab at the University of California at San Diego (UCSD) which is available for free online at Mendeley (<https://data.mendeley.com/datasets/rscbjbr9sj/3>) is being used [1]. The dataset consists of 108,312 retina OCT images (37,206 with choroidal neovascularization, 11,349 with diabetic macular edema, 8617 with drusen, and 51,140 normal), while for the age-related macular degeneration (AMD) dataset, it was downloaded from Farsiu Ophthalmology 2013 AMD dataset which is available for free online at Duke University (https://people.duke.edu/~sf59/RPEDC_Ophth_2013_dataset.htm); the dataset contains 26,900 OCT images acquired from 269 AMD patients [2], in addition to 1250 images (250 for each

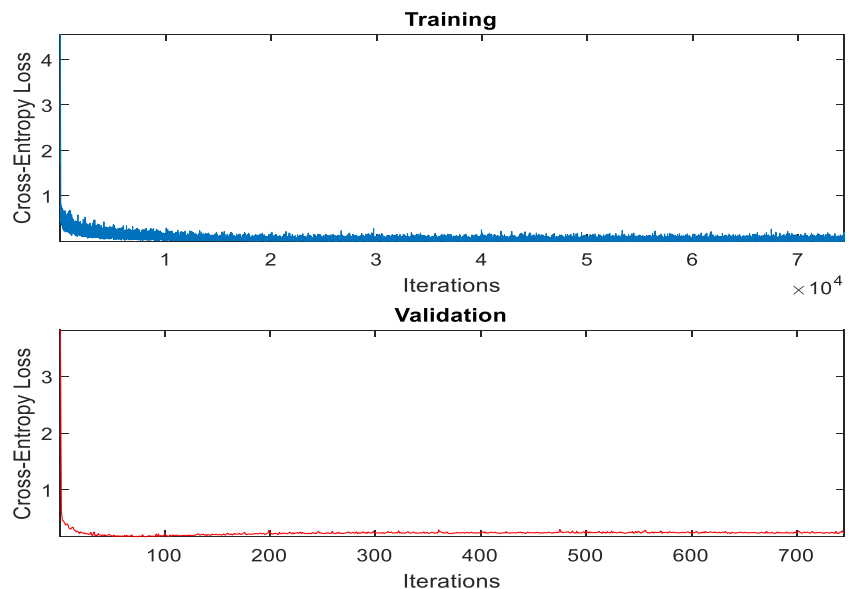
category) collected from five different hospitals and private clinics for testing purposes.

The images are split into 5 directories: AMD, CNV, DME, drusen, and normal [1, 2]. In this research, 136,187 SD-OCT images were used in the implementation of the system. Table 1 shows details about the used images from the used datasets. Figure 3 shows examples of images from the used datasets.

3.2 Image denoising and resizing

The used images from the first dataset (Zhang Dataset) are already denoised, while the second dataset (Farsiu Dataset) is noisy and can affect the classifier performance and needs to be denoised. To denoise the second dataset image, a double-

Fig. 7 Cross-entropy loss is plotted against the iterations for both training and validation



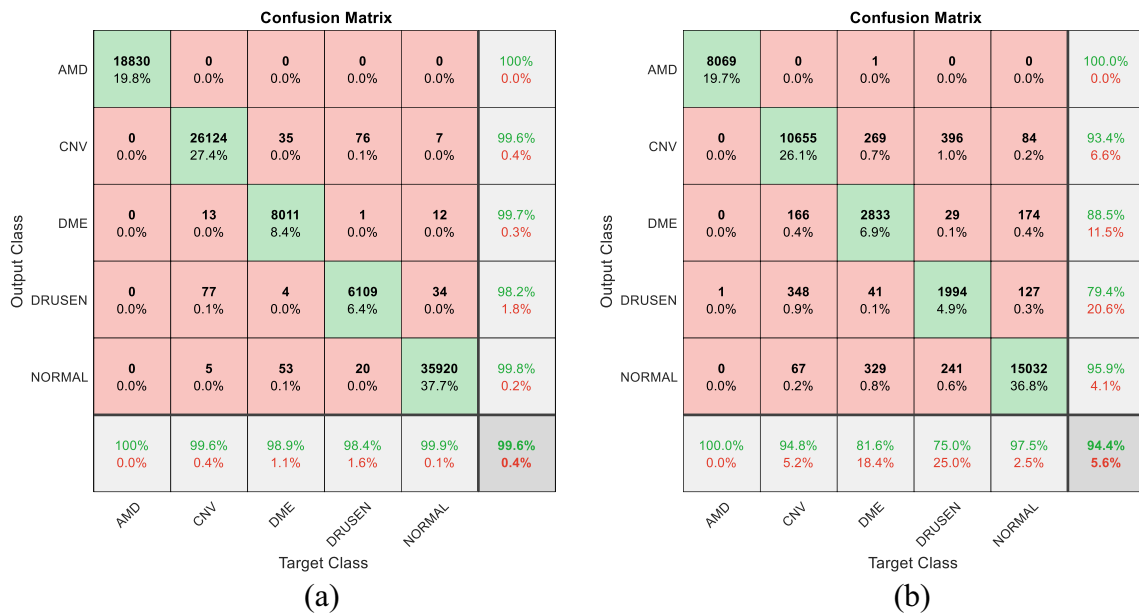


Fig. 8 Checking the performance of proposed system during training and validation. a training confusion matrix. b validation confusion matrix

density wavelet transform-based adaptive thresholding technique has been applied to all dataset images [28]. Figure 4 shows samples of noisy and denoised images from Farsiu dataset. After that, the final step in this stage is to resize all images to make them compatible with the input size of the proposed architecture that is discussed in the next section.

3.3 Adaptive moment learning rate

The adaptive moment learning rate (ADAM) is a type of stochastic gradient descent with adaptive learning rate for each parameter with momentum [29]. Stochastic is one of the most-

known optimization algorithms which is designed to deal with very complex optimization problem and it is used widely in deep learning for updating the weights based on a subset of training samples as shown in Eq. 1 below.

$$L_t(W) = \frac{1}{b} \sum_{j=1}^b l(W; x_{ij}; y_{ij}) + \gamma r(W) \tag{1}$$

where $\left\{ (x_{ij}; y_{ij})_{j=1}^b \right\}$ is the random mini-batch size chosen at iteration t , γ is the forgetting factor, l is the loss function, b is the number of training samples, W is the weights, and r is the convex regularized.

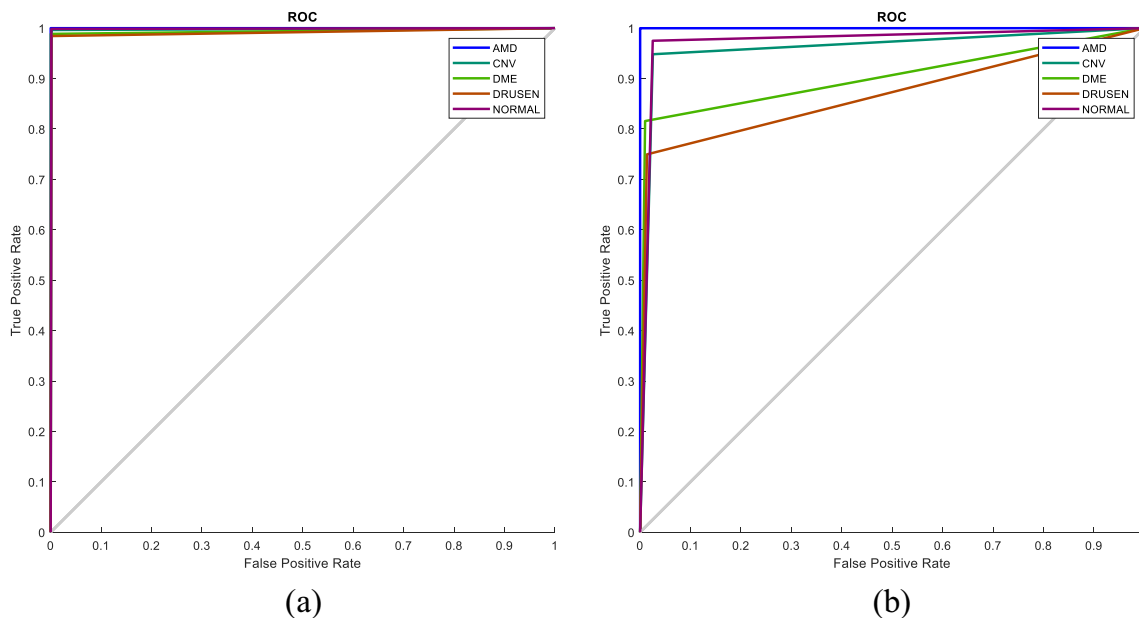


Fig. 9 Checking the performance of proposed system during training and validation. a training ROC. b validation ROC

For updating the weights in the ADAM algorithm, the illustration is in the corresponding explanations:

Step #	Equation	Explanation
1	$M_o = 0, R_o = 0$	(Initialization) For $t = 1, \dots, T$
2	$M_t = \beta_1 M_{t-1} + (1 - \beta_1) \nabla l_t(W_{t-1})$	(1st moment estimate)
3	$R_t = \beta_2 R_{t-1} + (1 - \beta_2) \nabla l_t(W_{t-1})^2$	(2nd moment estimate)
4	$\hat{M}_t = M_t / (1 - \beta_1^t)$	1st moment bias correction
5	$\hat{R}_t = R_t / (1 - \beta_2^t)$	2nd moment bias correction
6	$W_t = W_{t-1} - \alpha \hat{M}_t / \sqrt{\hat{R}_t + \epsilon}$	Update
7	Return W	Returning Value

where M is the 1st momentum estimate, R is the 2nd momentum estimate, \hat{M} is the 1st momentum correction, \hat{R} is the 2nd momentum correction, W is the weights, α is the learning rate, β_1 and β_2 are the hyperparameters, and ∇l_t is the gradient evaluated at timestep t .

In this research, the ADAM optimization method is used because it is providing a combined heuristics optimization of both momentum and root mean square propagation (RMSProp). It can build a gradient descent that combats the pathological curvature of the problem and speeds up the search at the same time.

3.4 Proposed CNN architecture

Deep learning is one of the most powerful and state of the art technology that is inspired by the deep structure of a human brain [12]. Human brain deep structures are characterized by numerous hidden layers that allow the extraction and abstraction of the features at different levels. Deep learning starts with the method proposed by [30] in 2006; the paper proposed a newly developed algorithm (greedy layer-wise training) which is used to train the neuron layers of deep network architecture. This algorithm is one of the unsupervised learning algorithms in which the deep network is trained layer by layer.

Since this method is very effective and powerful, it has been chosen to be used as a training algorithm for many deep networks. The most powerful, efficient, and widely used deep networks is the CNN; this network includes multiple hidden layers that are performing convolution and subsampling methods in order to extract low to high levels of features of the input data either single dimension or two dimensions [31].

In general, CNN showed a great performance in different areas especially in the medical field and computer vision and in all other fields, and recently, researchers devoted their effort to employing CNN on all fields [32, 33]. Basically, as mentioned before, CNN consists of six types of layers (input, convolution, RELU, fully connected, classification, and

output); arranging and ordering of these layers is crucial and must take into consideration that they must extract fine details from input data [18]. In this paper, the proposed CNN architecture which is used to classify the input images into 5 classes consists of 19 layers as shown in Fig. 5.

The proposed CNN architecture decreases the number of the layers when compared with pretrained network that is usually used with transferee learning technique from 25 layers in AlexNet, 201 layers in Densnet201, 48 layers in InceptionV3, 101 layers in ResNet-101, and 144 layers in GoogleNet to only 19 layers; these pretrained networks are usually used in the medical field. Reducing the number of layers will shorten the time required for training and for finding probabilities of new input images in addition to reducing the resources required to run the system. Table 2 shows detailed information about layers in the proposed CNN architecture.

3.5 Performance evaluation

To evaluate the performance of the proposed CNN architecture in classifying the eye retina diseases, a comparison between the diagnosis outputs with its corresponding label of the image was done. These image labels were done by the eye specialist to provide ground truth to evaluate classification outputs [34]. The performance of the proposed CNN architecture was evaluated by computing the sensitivity and specificity, where this evaluation parameter measures how precisely the eye retina is diagnosed. To compute the sensitivity, specificity, and accuracy of this approach, four statistical indices,

Output Class	AMD	CNV	DME	DRUSEN	NORMAL	
AMD	250 20.0%	0 0.0%	0 0.0%	0 0.0%	0 0.0%	100% 0.0%
CNV	0 0.0%	250 20.0%	2 0.2%	15 1.2%	0 0.0%	93.6% 6.4%
DME	0 0.0%	0 0.0%	241 19.3%	0 0.0%	1 0.1%	99.6% 0.4%
DRUSEN	0 0.0%	0 0.0%	1 0.1%	231 18.5%	7 0.6%	96.7% 3.3%
NORMAL	0 0.0%	0 0.0%	6 0.5%	4 0.3%	242 19.4%	96.0% 4.0%
	100% 0.0%	100% 0.0%	96.4% 3.6%	92.4% 7.6%	96.8% 3.2%	97.1% 2.9%
	AMD	CNV	DME	DRUSEN	NORMAL	

Fig. 10 Testing set confusion matrix to check the performance of proposed system

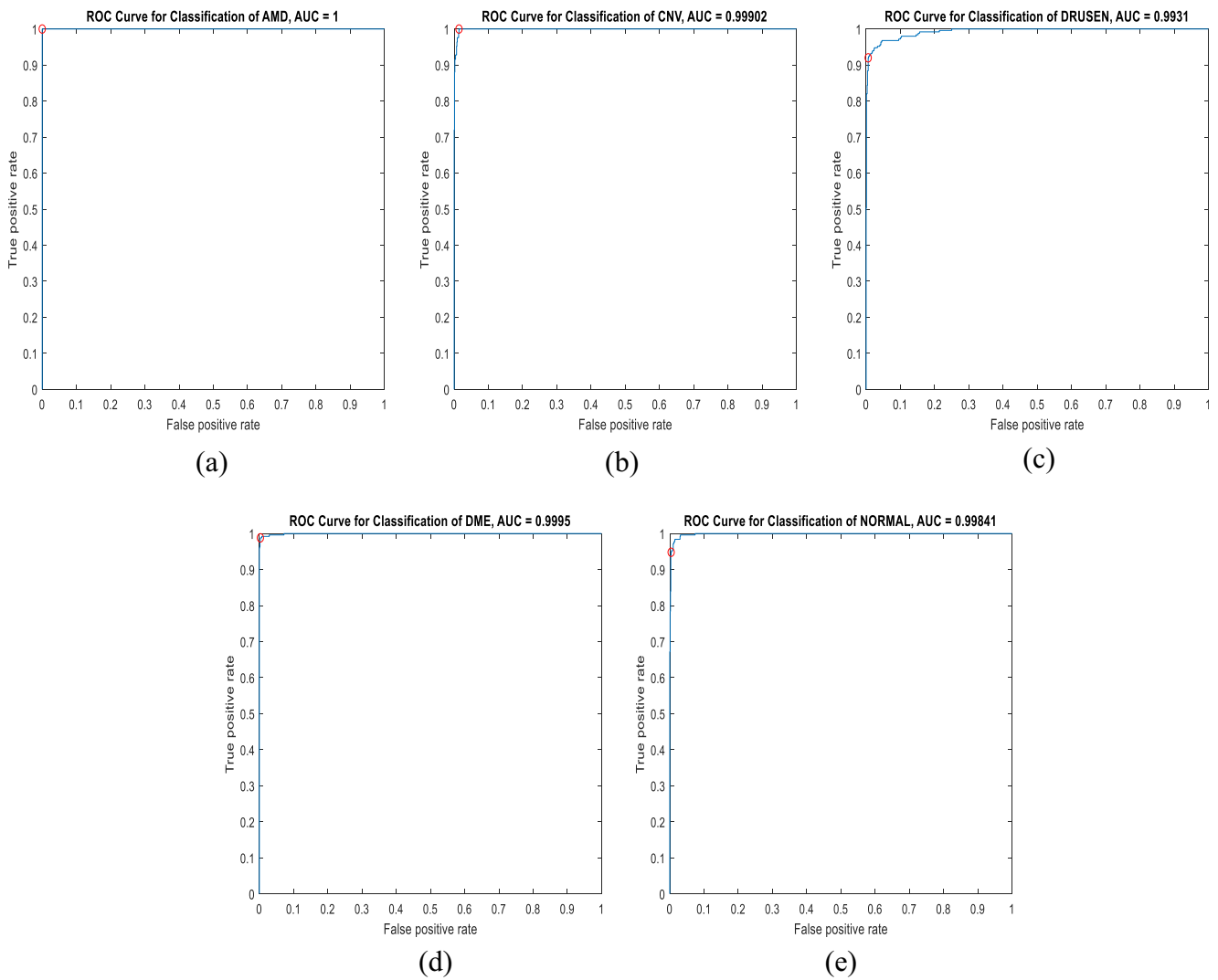


Fig. 11 Checking the performance of proposed system using testing set ROCS

namely true positive (TP), false positive (FP), false negative (FN), and true negative (TN), were calculated using confusion matrix, and consequently, the accuracy, sensitivity, and specificity were computed as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \tag{2}$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \tag{3}$$

Table 3 Testing dataset performance evaluation results

Class	Accuracy (%)	Sensitivity
AMD	100	1.0
CNV	98.64	1.0
DME	99.2	0.96
Drusen	97.84	0.92
Normal	98.56	0.97

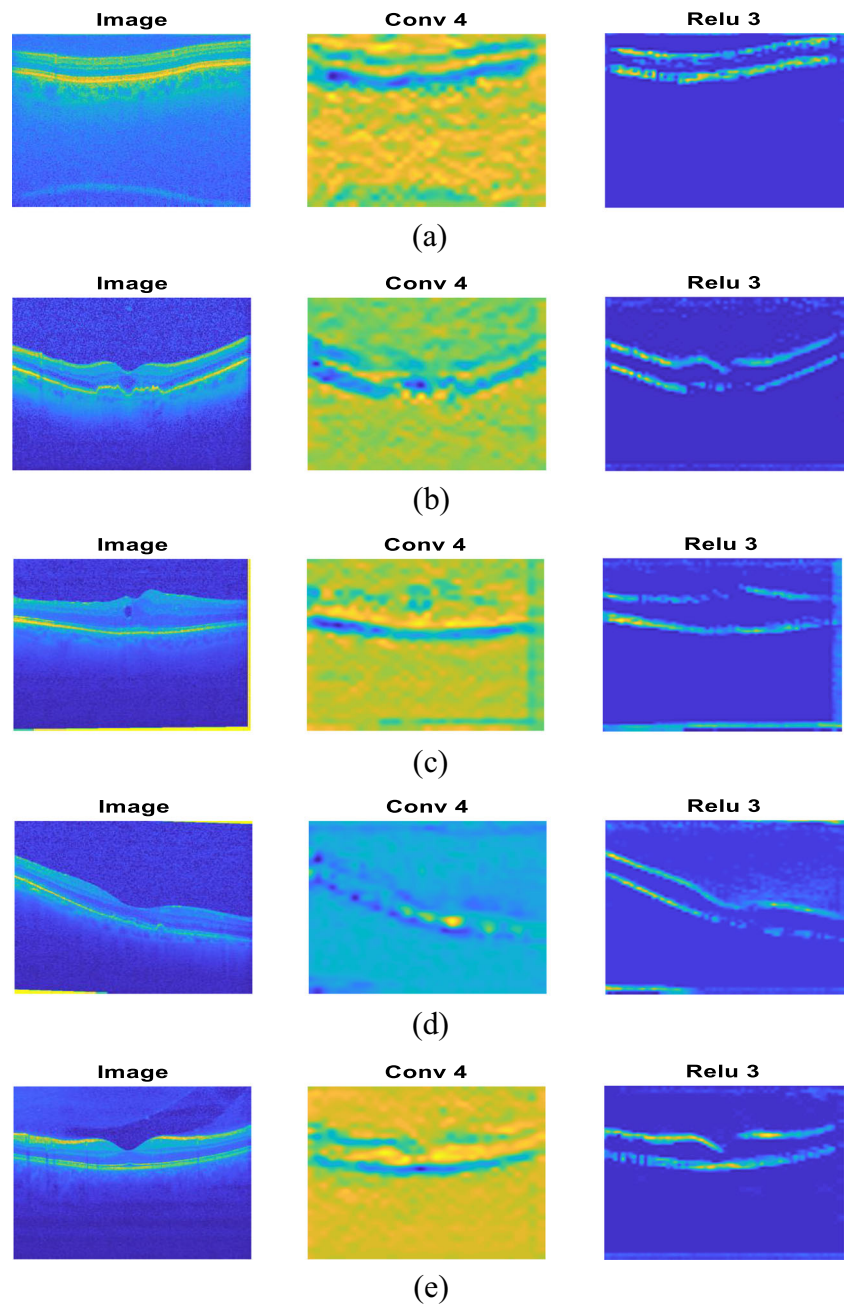
$$\text{Specificity} = \frac{TN}{TN + FP} \tag{4}$$

Also, a receiver operating characteristic (ROC) curve was plotted to present and check the performance of the proposed CNN model, with its X and Y axes defined as the false positive rate (FPR) and true positive rate (TPR), respectively, with values between 0 and 1. The TPR values resulted from the sensitivity equation, while the FPR results from subtracting the specificity value from 1. The closer the ROC curve to the upper left corner, the more satisfactorily the model was performed. While, the AUC of ROC was also used; the higher the AUC value, the more correct predictions using the model.

4 Results

The whole dataset mentioned before was used and divided into two subsets (training 70% and validation 30%) in

Fig. 12 OCT features images extracted using proposed CNN. a AMD. b CNV. c DME. d drusen. e normal



addition to 1250 images (250 images for each class) as a testing set, then fed to the CNN. The code run is conducted on a workstation with Intel Core i7-6700 at 3.4 GHz, and 16 GB of RAM with Nvidia Tesla k40 GPU is employed for training and testing of the proposed CNN architecture. The proposed CNN architecture was trained with ADAM (adaptive moment estimation) optimizer; the hyperparameters have been chosen carefully, and by using the trial and error, the following hyperparameters give the best results; initial learning rate is 0.001, the mini-batch size is 64, maximum number of epochs is 100, gradient decay factor is 0.9000, squared gradient decay factor is 0.9900, Epsilon is $1.0000e^{-08}$, L2

regularization is $1.0000e^{-04}$, gradient threshold is infinity, and the total number of training iterations was 74,400, while the validation frequency was 100 and the total number of validation iterations was 744. Figure 6 shows the performance in terms of accuracy for both the training and validation datasets. While Fig. 7 shows the performance in terms of cross-entropy loss for both the training and validation datasets. We can notice that both training and validation achieved high accuracy rates, it was 99.60% for training and 94.40% for validation respectively, with a sensitivity of 99.36% and 89.77% for validation respectively, a specificity of 99.91% and 98.52% for validation respectively. Figure 8 and Fig. 9

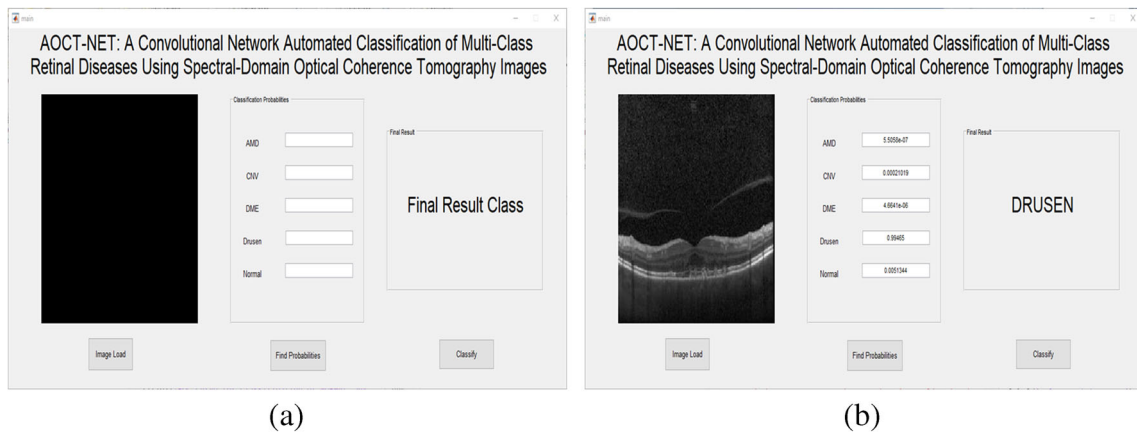


Fig. 13 The developed GUI. a Main interface. b Running example using GUI

show the confusion matrix and ROC curves for training and validation respectively.

To evaluate the proposed CNN system (AOCT-NET) in diagnosing the five classes retinal eye diseases, this system was tested using the mentioned testing; using testing set, the model achieved an accuracy of 97.12%, a sensitivity of 97.12%, and a specificity of 99.28%. The confusion matrix of the testing set is shown in Fig. 10. While the ROC curves for each class using the test set are shown in Fig. 11.

a AMD ROC, b CNV ROC, c drusen ROC, d DME ROC, e normal ROC.

The ROC curves and AUC were shown above of the proposed CNN model; the AUCs of the different classes were above 0.98 and have the following values 1, 0.99902, 0.9931, 0.9995, and 0.99841 for AMD, CNV, DME, drusen, and normal respectively. As shown in Table 3, all classes demonstrated high accuracy and sensitivity, and all were above 90%. To show the accuracy of the features extracted using the proposed CNN architecture, Fig. 12 shows the

extracted features using two layers for the five classes in the testing dataset.

To make the system more applicable and make it available to doctors and be applicable to a graphical user interface (GUI) was developed. Through this interface, the proposed AI is based on the CNN algorithm; the accuracy of eye disease diagnosis using SD-OCT is improved. Figure 13 shows the developed GUI.

5 Discussion

Comparison of the results of the proposed CNN architecture with other pre-trained and designed CNN in the literature are shown in Table 3. All research listed compared in Table 2 uses different SD-OCT databases or a combination of this dataset to form a huge one, and it is noted that they use a different number of classes, usually two or three classes. These factors can affect the performance of the used classification methods

Table 4 Comparison between the related works and the proposed method

Reference	Methodology	Number of images	Classes	Accuracy (%)
[16]	Transfer learning with GoogleNet CNN	10,000	3	96
[17]	CNN	3231	2 and 3	91.77
[18]	IFCNN	84,484	4	87.3 ± 2.2
[19]	CNN ensemble	7389	3	
[21]	SVM with HOG features	3247	3	95.56
[21]	Transfer learning with Inception Resnet V2 CNN	113,397	4	86
[23]	Random forest	1183	3	> 96
[24]	Transfer learning with Inception V3 CNN	1680	3	> 98
[25]	CNN	4096	2	93
[26]	VGG16, InceptionV3, and ResNet50	35,900	4	> 90
[27]	DenseNet and ResNet50	84,484	4	88 and 89
Proposed	CNN	137,437	5	97.1

significantly. In addition, it is noted that the researchers focused on preprocessing techniques and used transferee learning techniques with pre-trained CNN rather than designing a CNN architecture that can take raw images without preprocessing.

We can notice from Table 3 that different pre-trained CNN architectures have been used and the transferee learning technique, GoogleNet, Inception V2, and Inception V3 networks were used and also, ensemble models of these pre-trained CNN was used. Also, some researchers used machine learning techniques with extracted features from the segmented region of interest of SD-OCT as a method of classification. However, most of the listed methods in the literature have achieved high recognition rates, greater than 85%.

It is worth noting that most of these classify the OCT images into three classes, while two methods used four classes, and no method to classify images into five classes. Also, the largest dataset of images used was 113,397 with four classes. Finally, based on Table 4, the proposed system used the largest dataset of SD-OCT images that ever been used in such a system with 136,187 images, and it is the only system that classifies the SD-OCT images into five classes with one of the highest performances among other systems. The proposed system shows that the proposed CNN architecture has high classification rate compared with other methods, while it's the only system that provides a classification of five eye retina diseases based on SD-OCT with the largest dataset of images used ever.

6 Conclusion

In this paper, a CNN architecture is designed for the diagnosis of eye retina diseases using SD-OCT raw images. CNN was trained and tested using SD-OCT images containing five different cases (four diseases and a normal case). Furthermore, the proposed CNN was also trained and tested using a huge dataset of images (136,187 images). In addition, by tuning the proposed CNN network structure and using ADAM optimizing method, the more accuracy and lower time cost was achieved. Comparing the proposed system with the other methods, the experimental results show that the proposed method achieves better overall performance and a better accuracy rate of 98.1%. In the future, the work will focus on reducing the error rate and adding new eye retina diseases to make the system more generalized.

Acknowledgments The author would like to thank the Nvidia Learning Center at Yarmouk University (YU), represented by Dr. Ahmad Alomari, for providing access to the GPU Unit for training and testing the CNN architecture on the whole image dataset and getting the results. Also, the author would like to thank Zain Innovation Campus (ZINC) at Yarmouk University (YU) represented by Eng. Nour Al-Ajlouni, for providing access to the GPU Unit for initial results and testing of the CNN architecture on part of the dataset.

Availability of data and materials The image dataset that used to test and support the findings of this research is available from different hospitals in Jordan, but restrictions apply to the availability of these data, which were used under license for the current research only, and so are not publicly available. Data are however may be available from the author upon reasonable request and with permission of all hospitals.

Compliance with ethical standards

Conflict of interest The author declares that he has no conflict of interest.

Human and animal rights This article does not contain any studies with human participants or animals performed by the author.

Informed consent Informed consent was obtained from all individual participants included in the study.

References

1. Kermany DS, Goldbaum M, Cai W, Valentim CCS, Liang H, Baxter SL, McKeown A, Yang G, Wu X, Yan F, Dong J, Prasadha MK, Pei J, Ting MYL, Zhu J, Li C, Hewett S, Dong J, Ziyar I, Shi A, Zhang R, Zheng L, Hou R, Shi W, Fu X, Duan Y, Huu VAN, Wen C, Zhang ED, Zhang CL, Li O, Wang X, Singer MA, Sun X, Xu J, Tafreshi A, Lewis MA, Xia H, Zhang K (2018) Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell* 172(5):1122–1131 e9
2. Farsiu S, Chiu SJ, O'Connell RV, Folgar FA, Yuan E, Izatt JA, Toth CA, Age-Related Eye Disease Study 2 Ancillary Spectral Domain Optical Coherence Tomography Study Group (2014) Quantitative classification of eyes with and without intermediate age-related macular degeneration using optical coherence tomography. *Ophthalmology* 121(1):162–172
3. Kanagasigam Y, Bhuiyan A, Abramoff MD, Smith RT, Goldschmidt L, Wong TY (2014) Progress on retinal image analysis for age related macular degeneration. *Prog Retin Eye Res* 38: 20–42
4. Lemaitre G et al (2016) Classification of SD-OCT volumes using local binary patterns: experimental validation for DME detection. *J Ophthalmol* 2016:3298606
5. Reis AS et al (2012) Influence of clinically invisible, but optical coherence tomography detected, optic disc margin anatomy on neuroretinal rim evaluation. *Invest Ophthalmol Vis Sci* 53(4): 1852–1860
6. Schmidt-Erfurth U, Sadeghipour A, Gerendas BS, Waldstein SM, Bogunovic H (2018) Artificial intelligence in retina. *Prog Retin Eye Res* 67:1–29
7. Esteva A et al (2019) A guide to deep learning in healthcare. *Nat Med* 25(1):24–29
8. Ting DSW et al (2019) Artificial intelligence and deep learning in ophthalmology. *Br J Ophthalmol* 103(2):167–175
9. Ting DSW, Wu WC, Toth C (2018) Deep learning for retinopathy of prematurity screening. *Br J Ophthalmol*
10. Sajda P (2006) Machine learning for detection and diagnosis of disease. *Annu Rev Biomed Eng* 8:537–565
11. Pierro L, Zampedi E, Milani P, Gagliardi M, Isola V, Pece A (2012) Spectral domain OCT versus time domain OCT in the evaluation of macular features related to wet age-related macular degeneration. *Clin Ophthalmol* 6:219–223
12. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. *Nature* 521(7553):436–444

13. Alsaih K, Lemaitre G, Rastgoo M, Massich J, Sidibe D, Meriaudeau F (2017) Machine learning techniques for diabetic macular edema (DME) classification on SD-OCT images. *Biomed Eng Online* 16(1):68
14. Awais M, Müller H, Tang TB, Meriaudeau F (2017) Classification of sd-oct images using a deep learning approach. In: 2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA), IEEE, pp 489–492
15. Lee CS, Baughman DM, Lee AY (2017) Deep learning is effective for the classification of OCT images of normal versus age-related macular degeneration. *Ophthalmol Retina* 1(4):322–327
16. Karri SP, Chakraborty D, Chatterjee J (2017) Transfer learning based classification of optical coherence tomography images with diabetic macular edema and dry age-related macular degeneration. *Biomed Opt Express* 8(2):579–592
17. Rong Y, Xiang D, Zhu W, Yu K, Shi F, Fan Z, Chen X (2019) Surrogate-assisted retinal OCT image classification based on convolutional neural networks. *IEEE J Biomed Health Inform* 23(1):253–263
18. Fang L, Jin Y, Huang L, Guo S, Zhao G, Chen X (2019) Iterative fusion convolutional neural networks for classification of optical coherence tomography images. *J Vis Commun Image Represent* 59:327–333
19. Rasti R, Rabbani H, Mehridehnavi A, Hajizadeh F (2018) Macular OCT classification using a multi-scale convolutional neural network ensemble. *IEEE Trans Med Imaging* 37(4):1024–1034
20. Amil P et al (2019) Unsupervised feature extraction of anterior chamber OCT images for ordering and classification. *Sci Rep* 9(1):1157
21. Srinivasan PP et al (2014) Fully automated detection of diabetic macular edema and dry age-related macular degeneration from optical coherence tomography images. *Biomed Opt Express* 5(10):3568–3577
22. Mehta P, Lee AY, Lee C, Balazinska M, Rokem A (2018) Multilabel multiclass classification of OCT images augmented with age, gender and visual acuity data. *bioRxiv*:316349
23. Hussain MA et al (2018) Classification of healthy and diseased retina using SD-OCT imaging and random forest algorithm. *PLoS One* 13(6):e0198281
24. Ji Q, He W, Huang J, Sun Y (2018) Efficient deep learning-based automated pathology identification in retinal optical coherence tomography images. *Algorithms* 11(6)
25. Perdomo O, Otálora S, González FA, Meriaudeau F, Müller H (2018) Oct-net: a convolutional network for automatic classification of normal and diabetic macular edema using SD-OCT volumes. In: 2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018), IEEE, pp 1423–1426
26. Hwang DK, Hsu CC, Chang KJ, Chao D, Sun CH, Jheng YC, Yarmishyn AA, Wu JC, Tsai CY, Wang ML, Peng CH, Chien KH, Kao CL, Lin TC, Woung LC, Chen SJ, Chiou SH (2019) Artificial intelligence-based decision-making for age-related macular degeneration. *Theranostics* 9(1):232–245
27. Nugroho KA (2018) A comparison of handcrafted and deep neural network feature extraction for classifying optical coherence tomography (OCT) images. In: 2018 2nd International Conference on Informatics and Computational Sciences (ICICoS), IEEE, pp 1–6
28. Gnanadurai D, Sadasivam V (2005) Image de-noising using double density wavelet transform based adaptive thresholding technique. *International Journal of Wavelets, Multiresolution and Information Processing* 03(01):141–152
29. Kingma DP, Ba J (2014) Adam: a method for stochastic optimization, *arXiv preprint arXiv:1412.6980*
30. Hinton GE, Osindero S, Teh YW (2006) A fast learning algorithm for deep belief nets. *Neural Comput* 18(7):1527–1554
31. Gao F, Yue Z, Wang J, Sun J, Yang E, Zhou H (2017) A novel active semisupervised convolutional neural network algorithm for SAR image recognition. *Comput Intell Neurosci* 2017:3105053
32. Bakator M, Radosav D (2018) Deep learning and medical diagnosis: a review of literature. *Multimod Techn and Interact* 2(3)
33. Gu J et al (2018) Recent advances in convolutional neural networks. *Pattern Recog* 77:354–377
34. Alqudah AM, Alquraan H, Abu-Qasmieh I, Al-Badarnah A (2018) Employing image processing techniques and artificial intelligence for automated eye diagnosis using digital eye fundus images. *JBBBE* 39:40–56

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Ali Mohammad Alqudah received his BSc and MSc from Yarmouk University in 2015 and 2018 respectively. His research area is in the field of Biomedical Signal Processing, Image Processing and Analysis, Deep Learning, and Machine Learning. Currently, he is a Teaching and Research Assistant at Department of Biomedical Systems and Informatics Engineering, Yarmouk University, Irbid, Jordan.