




Eye diseases detection using deep learning with BAM attention module

Amna Zia¹ · Rabbia Mahum²  · Nabeel Ahmad¹ · Muhammad Awais^{3,4} · Ahmad M. Alshamrani⁵

Received: 12 September 2022 / Revised: 16 September 2023 / Accepted: 8 December 2023 /
Published online: 27 December 2023

© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2023

Abstract

With the changing lifestyle, a large population suffers from eye diseases such as glaucoma, cataract, and diabetic retinopathy. Therefore, timely detection and classification of the disease are necessary to minimize vision loss, however, it is time taking task and requires various tests and physicians' in-depth analysis. Thus, an accurate automated technique, timely detection, and classification are needed to cope with the aforementioned challenges. Therefore, this study proposes a technique based on an improved deep learning algorithm i.e., SqueezeNet that uses the eye image' features to detect various diseases such as cataract, glaucoma, and diabetic retinopathy simultaneously. In our proposed model, we employed Bottleneck Attention Module (BAM) with SqueezeNet having an additional layer. Our proposed attention module utilizes two different ways and effectively extracts the most representative features and drops the image's background features of eyes which don't take part in the detection of diseases. Moreover, the algorithm is a pre-trained network that doesn't require a huge training set, therefore, the existing dataset i.e., ODIR, cataract, ORIGA, and glaucoma datasets have been utilized for the training and testing. Additionally, cross-validation has been employed using the cataract dataset to assess the performance of the proposed model. The squeezed connections with regularization power help to minimize the overfitting during the training of eye samples training sets. The proposed algorithm is a novel and effective technique to report the successful implementation for the early detection and classification of eye disease images. The algorithm achieved 98.9% accuracy over the testing dataset and 98.1% accuracy over cross-validation. Various experiments have been performed to confirm that our proposed algorithm performs significantly to detect and classify eye diseases than existing state-of-the-art.

Keywords Eye Disease Detection · Deep Learning · Cataract · Glaucoma · Diabetic Retinopathy · Improved SqueezeNet

1 Introduction

Eye diseases may cause partial or complete vision loss if a proper observation is not performed at an initial stage. Early diagnosis of eye disease can halt vision disability [1]. Vision impairment, called vision loss is usually defined based on visual acuity [2]. The normal value of visual acuity is 20/20 which less than 20/40 or 20/60 is in a visionary-impaired person. In 2015, 1.5 billion out of the 7.33 billion world population is found as visionary impaired [2]. The main causes of vision loss are cataracts, glaucoma, and diabetic retinopathy (DR) [2]. People residing in remote or rural areas are affected by eye diseases due to the inaccessibility of ophthalmologists and health services. In developing countries, manual eye assessments are also less accurate due to untrained ophthalmologists and cause a delay in timely treatment [3]. Therefore, it is need time to develop an intelligent eye disease detection system that can be used for the early detection of eye diseases to prevent vision loss [4]. Machine learning techniques have been widely used for medical applications [5]. Conventional machine learning algorithms that have been used for the classification include decision trees, random forest, naive Bayes algorithm, support vector machine, and k-nearest neighbor [6]. However, traditional machine learning methods are not effective for a large amount of data [7]. More precisely, in decision trees, there are chances of overfitting if repeated tree construction is performed whereas large trees are hard to comprehend. Random forest takes more time in training and its complexity is high. Naive Bayes is also computationally expensive specifically for models having more variables. Support vector machines cannot work well if the dataset is noisy and an appropriate kernel function with SVM is challenging. K-nearest neighbor is computationally high and unknown data is comparatively expensive to distinguish [8]. Furthermore, machine traditional learning algorithms have shown less accuracy in the classification of eye disease [9]. Deep learning models perform an important role in various domains such as Facebook for speech identification [3], performs image exploration in Google [3], medical diagnosis [3, 10, 11], and plant disease detection [12]. Representative deep learning algorithms include convolutional neural networks, deep belief networks, and recurrent neural networks [6]. Convolutional neural networks are the widely used algorithms for image identification and classification [13], speech analysis [14], and natural language processing [15].

It is evident from the research that a method for increasing the resolution of input images can enhance the performance of classifiers [16]. Various DL-based classifiers are based on transfer learning, however, they may cause the challenge of negative transfer (NTs). Seyyed et al. [17] introduced an approach to minimize the NT outcomes. Furthermore, iris segmentation techniques have been proposed to aid the better classification [18].

Existing techniques used for eye disease detection face some limitations. Small datasets of retinal images have been utilized for eye disease detection in most of the techniques employing existing machine learning algorithms (SVM [19], Neural Networks, Decision Trees [19], Naïve Bayes [20], Random Forest [21]). Therefore, the performance of the models has been degraded over large-size unseen test samples [22]. In [23], authors achieved 85.5% accuracy approximately by applying VGG16 for the diagnosis of diabetic retinopathy using retinal fundus images. Despite of the promising results of neural networks, its modulating complexity is not as anticipated and is found as a black box [24]. Moreover, Recent works related to eye disease detection deal with one or two diseases therefore an automated model should be developed that could detect various eye diseases with maximum accuracy [1].

To solve the aforementioned problems of high computation time, complexity, inadequate size of training data [22] to increase the performance of the neural network, we propose an improved SqueezeNet model [25] for the detection of cataracts, glaucoma, and diabetic retinopathy simultaneously. SqueezeNet increases the accuracy level with fewer parameters than some pre-trained models such as AlexNet on the ImageNet dataset. In our improved SqueezeNet model, we improved classification accuracy by adding a BAM attention block that allows the classifier to focus on the most discriminative features while suppressing less relevant or noisy information. This enhances the overall feature representation capability of the model, leading to better classification accuracy.

The main contributions of the paper are below:

- Efficient performance of improved SqueezeNet algorithm due to the addition of a attention module i.e. BAM for enhanced feature extraction.
- Accurate classification of eye diseases on unseen datasets.
- Computationally efficient framework, due to the lightweight and simple architecture of improved SqueezeNet with insignificant additional computational cost.
- We compared the proposed method with other revolutionary techniques and attained high-performance outcomes.
- A single technique is used for the detection of multiple eye diseases.

2 Acronyms

Diabetic Retinopathy (DR), Computer-Aided Design (CAD), Machine Learning (ML), Support Vector Machine (SVM), Convolutional Neural Network (CNN), Support Vector Machine (SVM), Neural Network (NN), That is (i.e.)

3 Literature review

Cataract eye disease may cause loss of vision and is included in severe diseases of the eyes. Therefore, its timely detection and prevention is the key aspect to overcoming serious loss. Rahul et al. [26] introduced a combined approach using Convolutional Neural Networks (CNN) and Support Vector Machine (SVM) for detecting cataracts in micro-read fundus images. They expanded the dataset by including additional images and then extracted relevant features. The CNN achieved an accuracy of 87.08%, while the SVM achieved a slightly higher accuracy of 87.5%. In a study by Rohit [27], an intelligent system was introduced for analyzing color fundus images and detecting glaucoma. The system utilized a combination of Convolutional Neural Networks (CNN) and machine learning techniques. The process involved two stages: first, CNN was employed for feature extraction, and then machine learning algorithms were used for classification. Interestingly, the logistic regression network demonstrated superior performance compared to other machine learning algorithms in this hybrid system. In [2], an efficient method based on network selection is proposed for the detection of cataract disease known as uses additive white Gaussian noise (AWGN). The method involved features selection, training of the network, and efficient network selection. For the evaluation of the proposed method, the EyePACS dataset was employed. The performance of the proposed CACD method outperformed the existing CCN-based CACD models. In [28], a deep learning-based model has been utilized for

cataract detection based on binary classification. Firstly, they developed a dataset of 1335 images from China Hospital, including 433 images of the healthy eye and 922 of cataract disease. They used discrete state transition (DST) with ResNet for the classification. A hybrid global–local-based representation model has been proposed for cataract detection [29]. They employed a dataset of 8030 images and employed AlexNet for global features extraction and then employed deconvolutional architecture among CNN layers to extract key features. In the end, they used a hybrid model of various AlexNet with a combination of global and local features.

For the detection of diabetic retinopathy (DR) eye disease, [30] uses the EyePACS dataset and Meena's model [31] as a baseline. Data is preprocessed by removing noise, and outlier images, normalization of color, and augmentation of data. Then, pre-trained neural network models, VGG-16 and VGG-19 were employed for the classification of normal and DR images. The model provided 71% and 73% accuracy for VGG-16 and VGG-19 respectively. The model was improved to give 83% accuracy with the Keras CNN model having a customized layer and dense layer. In [32], a first model was proposed using CNN for the DR classification. They employed a class-weighted approach while updating the parameters in each batch which in turn reduced the class imbalance and overfitting. Islam et al. [33] converted the first approach of the classification problem to regression to get better results of detection. Furthermore, Torre et al. [34] employed a CNN-based technique combining both eyes images and feature representations for the classification. Raju et al. [35] stated that due to small filters in the convolutional layers, the classification performance of DR has been improved due to the small lesion size. Inception models have been employed in ([36–38]) to extract features from relevant lesion marks in images to detect DR. Attention module has been employed to improve the classification accuracy of CNNs. Zhao et al. [39], employed an attention-based technique and a bilinear approach to train a CNN and enhanced the performance of the classification. Li et al. [40], developed a model for the detection of DR and DMR employing an attention mechanism to discover the inter-disease correlations.

Key factor for disease is computed as the fraction of vertical cup diameter (VCD) and vertical disc diameter (VDD) [41]. The two-step model for glaucoma disease classification has been used in [42]. The first step is the localization and extraction of the optic disc from retinal fundus images and the second one involves the classification of this extracted disc into normal and glaucomatous. The localization was achieved through "Regions with Convolutional Neural Network (RCNN)" and the classification stage used deep CNN to label the localized disc. The proposed methodology has been evaluated on seven public datasets for localization and the ORIGA dataset for the classification of glaucoma. Localization results show 100% accuracy on six datasets and the classification of glaucoma shows an "area under the curve (AUC)" of 0.874. The model performed significantly however it required a high computational cost. Almazro et al. [36], has developed a model to train for glaucoma detection using the RIGA dataset that is publicly available and attained considerable accuracy. R Mahum [43], employed CNN, LBP, SURF, and HOG for feature extraction from fundus images and then utilized machine learning-based classifiers such as KNN, a SVM, and RF for the classification of Glaucoma. Sucheet et al. [44] introduced three CNN models designed for predicting retinal diseases. To enhance the images, they utilized a Gaussian kernel and employed entropy-based thresholding to extract blood vessels from the retinal images. In [45], path classification has been employed using the deep learning mechanism to detect the retinal nerves and their defects. In [46], local and holistic features have been used with deep learning techniques to detect and classify glaucoma. Moreover, a deep residual technique has been utilized using fundus images to detect glaucoma and the

obtained performance results are compared with ophthalmologists [47]. In [48], the objective measure for the digital image inpainting method has been proposed to refill the damage in images, that can be used for processing of fundus images to be more visible. The summary of the existing works is reported in Table 1.

4 Proposed methodology

In this study, we propose a novel and improved SqueezeNet architecture that preserves accuracy with fewer parameters for the multiple eye disease detection based on BAM attention module. It adopts three intrinsic strategies to enhance the performance of state-of-the-art convolutional neural networks. First, most of the filters of the network are 1×1 rather than 3×3 which decreases network weights. Second, it reduces input channels up to 3×3 filters which also lessens the weight of the network. Third, it performs late down-sampling to get large activation maps. Furthermore, the enhanced features set is formed due to the attention mechanism attached to the fire module of the SqueezeNet. In the end, the soft-max function is employed for the classification of eye diseases. The details of the proposed architecture is given below.

4.1 SqueezeNet Design Strategies

SqueezeNet carries three key strategies to increase the performance of conventional CNN architectures.

Strategy 1: Replacement of 3×3 filters with 1×1 filters

In a specific no of convolution filters, the majority of the 1×1 filters are chosen, since the 1×1 filter has 9X fewer parameters than the 3×3 filter.

Strategy 2. Decreasing the number of input channels to 3×3 filters

The convolution layer comprising 3×3 filters contains a total amount of parameters equal to no of input channels * no of filters * 3×3 . To maintain a small total no of parameters in CNN, it is essential to reduce the no of input channels to 3×3 filters with the reduction of 3×3 filters i.e., strategy 1.

Strategy 3. Late downsampling in the network to convolution layers has large activation maps

In a CNN, each layer produces an output activation map with a spatial resolution of 1×1 or greater than it. The width and height of such activation maps are managed by 1) input data size ($256 * 256$ images) and 2) the option of layers which is used for downsampling in CNN architecture. Most often, downsampling is set up in CNN architectures by putting the stride greater than one in some convolutional or pooling layers [61–63]. If the initial layers of the network have massive strides, then many of the layers will possess small activation maps. Contrarily, if many layers have a stride of 1 and the strides with values greater than 1 are centralized towards the end of the network, then many network layers will own massive activation maps. Our instinct is that massive activation maps (due to late down sampling)

Table 1 The summary of some existing works

Reference	Datasets	Method	Outcomes	Observations
[49]	SCES and ORIGA	Classification	83% and 88%	Overlapping pooling layers have been used for binary classification in CNN
[50]	499 Images	Classification	97%	DT, KNN, SVM, and RF have been compared using various eye measurements
[51]	279 Images	Classification	92%	Pre-parametric glaucoma detection is performed using a feed-forward neural network
[52]	48116 Fundus Images	Classification	98%	Binary classification has been performed on suspected and exact glaucoma images
[53]	RIM-ONE, Messidor, DRISHTI-GS	OC and OD Segmentation	80%, and 89%	Entropy sampling based on convolutional filters has been utilized to extract key features
[54]	DRIONS-DB, RIM-ONE	OD and Vessels Segmentation	95% recall and 95% precision	Transfer learning concept has been employed using VGG network then images have been classified after supervised learning
[55]	DRIVE	Disc, vessels, and fovea Segmentation	IOU: 62%	33 × 33 central pixels patch has been categorized employing 7 layered CNN
[56]	SiMES	Disc Segmentation	IOU: 90%	Handcrafted features have been extracted to segment the images pixel-wise and CNN with 7 layers has been used to assess the segmentation of OD
[57]	28 fundus images	Classification	Not Available	Texture features have been mined in nerve fibers layers
[58]	71 fundus images	Not Reported	97.2%	OC detection has been performed employing Neuro Retinal Images
[59]	1866 images	Not Reported	Not Available	Retcam images have been utilized having close angle for Glaucoma Detection
[60]	60 fundus images	Classification	91%	High-order spectra and texture features have been utilized to detect glaucoma

can induce better classification accuracy, with everything else held identically. Indeed, J. Sun and K. He performed delayed down-sampling in four distinct CNN frameworks and attained higher classification accuracy in each setup [49].

The first two design strategies relate to the wise reduction of parameters amount in a CNN, preserving the network accuracy. Strategy 3 relates to maximizing the accuracy of the network on the limited forecast of parameters. Next, we explain the Fire module, our constituent for CNN architectures that empowers us to effectively employ these three strategies.

4.2 The fire module

The Fig. 1 exhibits the Fire module's architecture in SqueezeNet.

A Fire module consists of a squeeze convolution layer (with 1×1 filters) that feeds into an expanded layer having a mix of 1×1 and 3×3 convolutional filters. The prolific utilization of 1×1 filters in the Fire modules is the prosecution of the first design strategy of SqueezeNet. Three adjustable hyper-parameters in the Fire module are $s_{1 \times 1} = 3$, $e_{1 \times 1} = 4$, and $e_{3 \times 3} = 4$. In the Fire module, $s_{1 \times 1}$ is the value of squeeze layer filters (overall 1×1), $e_{1 \times 1}$ is the value of 1×1 expand layer filters and $e_{3 \times 3}$ is the value of 3×3 expand layer filters. While using Fire modules, we keep $s_{1 \times 1}$ to be fewer than $(e_{1 \times 1} + e_{3 \times 3})$, so that the squeeze layer help in limiting the value of input channels to 3×3 filters (as per the second design strategy of SqueezeNet).

4.3 Architecture of SqueezeNet

The architecture of SqueezeNet CNN is illustrated in Fig. 2. SqueezeNet starts with a discrete convolution layer i.e., conv1, ensued by 8 Fire modules i.e., fire2-9, closing with the last convolution layer i.e., conv10. We uniformly increase the value of a fire module filter from the start to the last of the network. SqueezeNet executes max-pooling with the stride of 2 following layers conv1, fire4, fire8, and fire10. These tolerably late positionings of pooling are according to strategy 3 of SqueezeNet design. Table 2 put forward the entire SqueezeNet architecture.

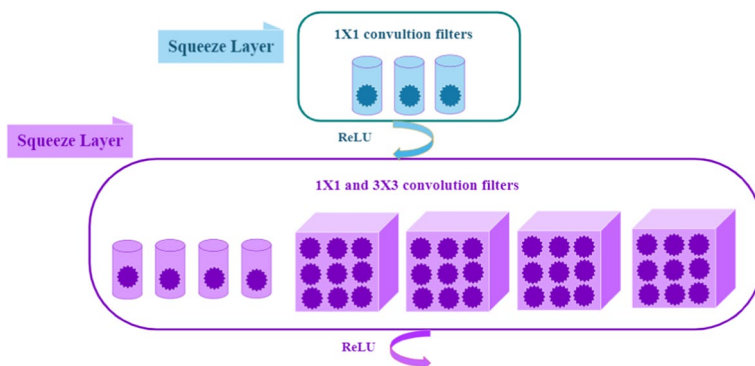


Fig. 1 Micro-design view: Convolution Filters in Fire Module ($s_{1 \times 1} = 3$, $e_{1 \times 1} = 4$, and $e_{3 \times 3} = 4$)

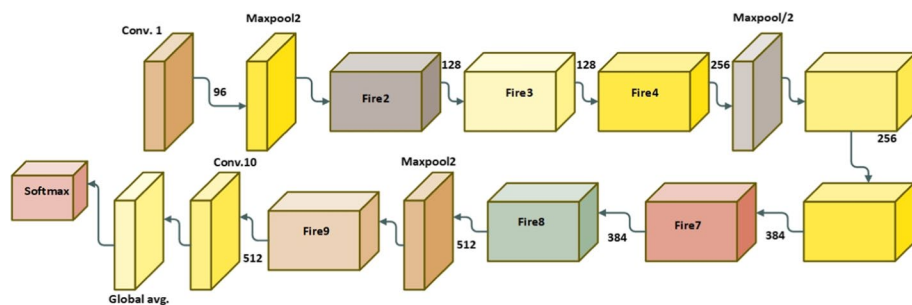


Fig. 2 SqueezeNet architecture

4.4 SqueezeNet characteristics

Many SqueezeNet design choices are provided below. The instinct beyond these choices can be perceived in the papers referenced below.

- To obtain the same width and height of output activations from 1×1 and 3×3 filters, a 1-pixel border having zero padding is added in input data to the 3×3 filters of expand module.
- Relu [64] is prosecuted to activations via squeeze layers and expand layers.
- After the fire9 module, dropout [65] is applied with a 50% ratio.
- Shortness in fully connected layers to SqueezeNet. This architecture choice was motivated by NiN [66].
- While training SqueezeNet, a learning rate of 0.04 is chosen which is decreased linearly throughout the training, as explained in [67].
- The Café configuration does not naturally endure a convolutional layer containing many filter resolutions (e.g., 1×1 & 3×3) ([68]). To get over this, expand layer is implemented with two discrete convolution layers: one with 1×1 filters and the other with 3×3 filters. Then, the outcomes of both layers are concatenated altogether in the channel dimension. This is computationally equivalent to executing a single layer that carries both (1×1 & 3×3) filters.

SqueezeNet CNN architecture is ported with other CNN software architectures for compatibility:

- MXNet [69] SqueezeNet port: [70]
- Chainer [71] SqueezeNet port: [72]
- Keras [73] SqueezeNet port: [74]
- Torch [75] SqueezeNet Fire Modules port: [76]

4.5 Improved SqueezeNet

We implement our proposed network by adding a batch normalization layer in SqueezeNet architecture. Batch normalization is a technique for training very deep neural networks that normalizes the contributions to a layer for every mini-batch. This has the impact of settling

Table 2 SqueezeNet design dimensions

Name / Type of layer	Size of output	Size of filter size/stride (for non-fire layer)	depth	s1 × 1 (#1 × 1 squeeze)	e1 × 1 (#1 × 1 expand)	e3 × 3 (#3 × 3 expand)	s1 × 1 sparsity	e1 × 1 sparsity	e3 × 3 sparsity	# Bits	#Parameter before pruning	#Parameter after pruning
input image	224 × 224 × 3											
conv1	111 × 111 × 96	7 × 7/2 (× 96)	1				100% (7 × 7)			6bit	14,208	14,208
maxpool1	55 × 55 × 96	3 × 3/2	0									
fire2	55 × 55 × 128		2	16	64	64	100%	100%	33%	6bit	11,920	5,746
fire3	55 × 55 × 128		2	16	64	64	100%	100%	33%	6bit	12,432	6,258
fire4	55 × 55 × 256		2	32	128	128	100%	100%	33%	6bit	45,344	20,646
maxpool4	27 × 27 × 256	3 × 3/2	0									
fire5	27 × 27 × 256		2	32	128	128	100%	100%	33%	6bit	49,440	24,742
fire6	27 × 27 × 384		2	48	192	192	100%	50%	33%	6bit	104,880	44,700
fire7	27 × 27 × 384		2	48	192	192	50%	100%	33%	6bit	111,024	46,236
fire8	27 × 27 × 512		2	64	256	256	100%	50%	33%	6bit	188,992	77,581
maxpool8	13 × 12 × 512	3 × 3/2	0									
fire9	13 × 13 × 512		2	64	256	256	50%	100%	30%	6bit	197,184	77,581
conv10	13 × 13 × 1000	1 × 1/1 (× 1000)	1				20% (3 × 3)			6bit	513,000	103,400
avgpool10	1 × 1 × 1000	13 × 13/1	0								1,248,424 (total)	421,098 (total)

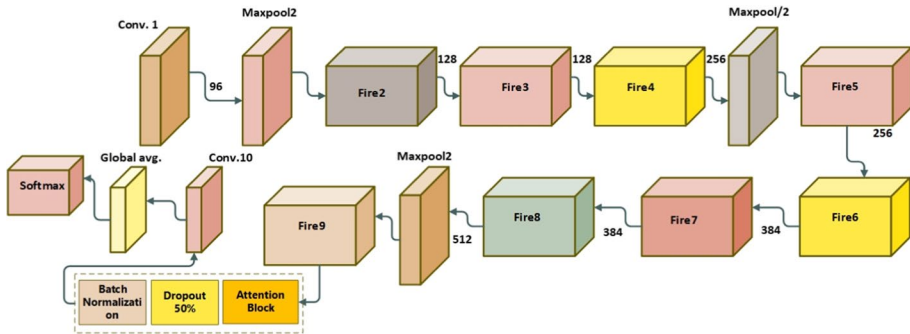


Fig. 3 Improved SqueezeNet Architecture

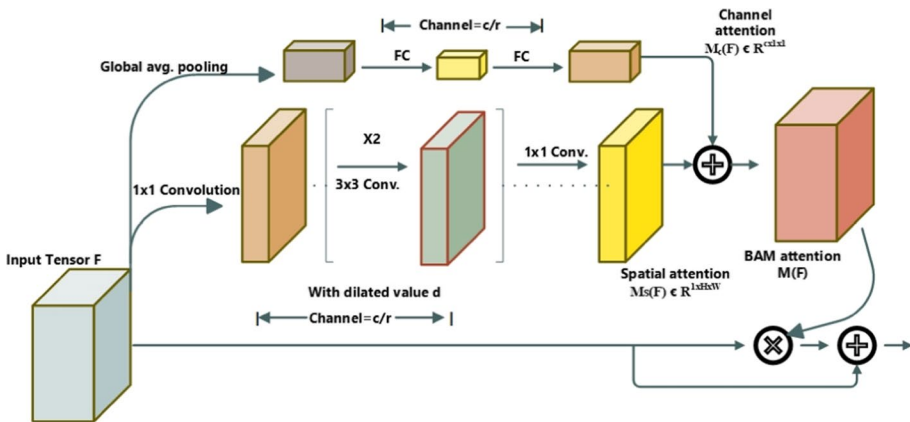


Fig. 4 Architecture of BAM attention module

the learning process drastically decreasing the number of training epochs required to train deep neural networks. Figure 3 shows an improved architecture of the proposed model.

4.6 Attention module

Various disease lesions can be there in the human eye which are normally of various shapes. Channel dependency usage is a significant way of improving CNN model execution. To increase the performance of state-of-the-art models with negligible computational cost, we used attention block i.e. BottleNeck Attention Module (BAM) [77] in our improved SqueezeNet model as shown in Fig. 4. BAM's architecture depends on two pathways such as spatial and channel. It gets training in an end-to-end way with our proposed SqueezeNet. Different channel weights are trained using the cost function and weight coefficients of the feature channel are obtained automatically. The attention module assists the model in attaining intermediate features more effectively. The proposed attention module's architecture is shown in Fig. 4. The F represents the feature map, whereas $M(F)$ is an attention map computed by the module using two attention methods such as spatial as M_s and channel as M_c . There exist 2 hyperparameters i.e. r as reduction ratio, and d as dilation

value. More specifically, r controls the overhead among both attention methods, and d spatially helps in contextual information using the receptive field size.

5 Experimental evaluation

In this section, we describe the experimental environment, metrics, comparative analysis with various existing methods, discussion describing the limitations and challenges, and conclusion of our proposed study.

5.1 Environment

We performed the experiments using a GPU NVIDIA card i.e., GEFORCE GTX with 4 GB memory. The details of the employed hardware are shown in Table 3. The operating system was Windows 10 having a RAM of 16 GB. The experiment was performed on the Anaconda framework.

5.2 Dataset

First, we gathered a dataset namely ODIR [78] version-7 that comprises 4166 JPG images of the eye. The dataset contains 1293 cataracts and 2873 normal eye images. We used the Glaucoma dataset [30] version- 1 which contains 4854 JPG images of glaucoma and normal eye. For DR (Diabetic Retinopathy), we gathered the diabetic retinopathy dataset [79] version-1 which contains 13.3 k JPEG images of the eye. We combined these three datasets to train our model on cataracts, glaucoma, DR, and normal eye. Some training samples of the dataset are shown in Fig. 5. The details of the dataset are displayed in Table 4. The plot for the dataset is presented in Fig. 6. We split the overall dataset approximately 83% for training set and 17% for testing respectively. The split ratio of approximately 83% for training and 17% for testing represents a balanced 5:1 ratio. This balance ensures a substantial training set to facilitate the model's learning process while also reserving a significant portion of the data for testing and evaluation purposes.

Second, we applied the Kaggle dataset for cataracts [80], Glaucoma Detection [81], and Diabetic Retinopathy (resized version) [82] for the cross-validation of our proposed model. The cataract dataset comprises 602 PNG images of cataracts and normal eyes. The glaucoma dataset consists of 2005 JPG images of glaucoma and normal eye and the DR dataset contains 70.2 k JPEG images of DR and normal eye.

Table 3 System Specifications of the Employed Model

Hardware	Specifications
Computer	GPU Server
CPU	Intel Core i5
RAM	16 GB
GPU	NVIDIA GEFORCE GTX×4

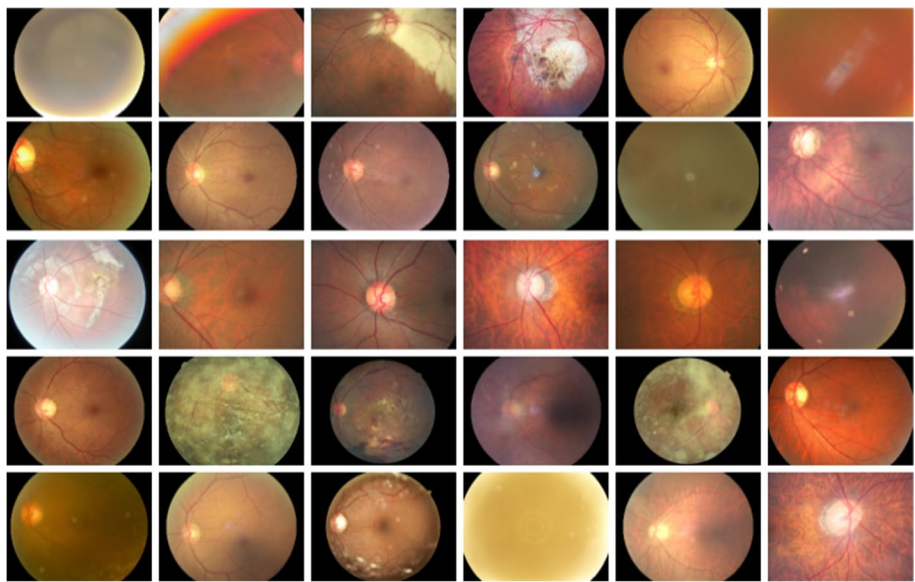


Fig. 5 Training Samples

Table 4 Dataset Details for Training and Testing

Class Label	Training Data Samples	Testing Data Samples
Cataract	1500	400
Glaucoma	1200	200
DR	1200	200
Neutral	1000	150
Total	4900	950

5.3 Results and analysis

During the training phase, the parameters were set as shown in Table 5. The initial learning rate was 0.1, which was divided by 5 when the epoch finished its 59 and 89 iterations. The maximum number of epochs was set to 180, the momentum was set 0.9, and the batch size was set 16. At every 10 epochs, a sample from the training set was picked for the continuous training of the proposed network till the last sample.

5.4 Metrics

For the performance evaluation of the proposed model, we have utilized various metrics such as Precision, Recall, and Accuracy. Moreover, these metrics are relied on true positive (TP), false positive (FP), true negative (TN), and false negative (FN). The TP refers to the correctly classified diseases by our proposed model, FP refers to the number of images that

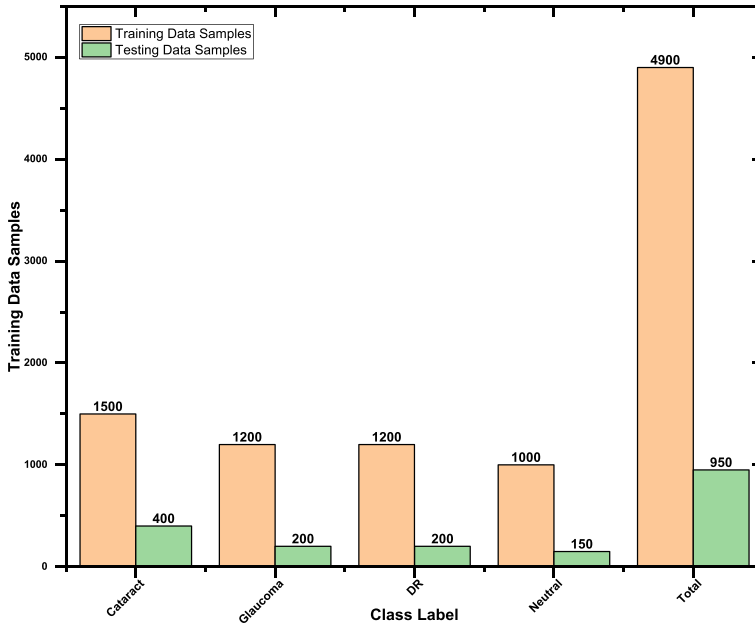


Fig. 6 Plot for Training Dataset

Table 5 Details of Hyper-Parameters of the Proposed Network

Parameters	Value
learning rate	0.001
No of epochs	180
batch size	16
Model's Confidence Value	0.25

were incorrectly classified as another disease than the actual one, and FN denotes the number of diseases that were incorrectly classified as negative i.e., neutral class, and TN refers the number of images that were correctly classified as a negative class such as neutral. Furthermore, precision refers to the fraction of TP over the total images classified as positive. The mathematical equation is given below.

$$\text{Precision} = \frac{TP}{TP + FP} \quad (1)$$

The accuracy of the system indicates the correctly classified images by the proposed system. The equation is presented below.

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (2)$$

The recall is the fraction of the classified positive class images to all images of positive class whether they were classified as a negative class by the system. The recall value closer to 1 refers to the better model. The equation of Recall is given below.

Table 6 Class-wise performance of the Proposed Model

Class	Precision (%)	Recall (%)	Accuracy (%)
Cataract	98.2	98.8	99.2
Glaucoma	99.2	98.5	98.9
DR	99.2	98.4	98.3
Neutral	98.9	99.1	99.2
Average	98.875	98.7	98.9

Table 7 Comparison with the original SqueezeNet

Model	Precision (%)	Recall (%)	Accuracy (%)	Mean Training Time(minutes)
Original SqueezeNet	94	95.4	96	14
Our Proposed Model	98.875	98.7	98.9	13.7

$$Recall = \frac{TP}{TP + FN} \quad (3)$$

In this section, we evaluated the metrics defined above for each category of disease. We observe from the results as shown in Table 6 that we have attained a significant average value for precision, recall, and accuracy. In cataract identification, we got an accuracy of up to 99.2%. Similarly, we are more precise in the case of glaucoma and DR identification i.e., 99.2%. In the case of neutral eye images, we got an accuracy value of 99.2%.

5.5 Results of the proposed model

In this section, we analyzed the performance of our proposed model with the existing original SqueezeNet. We observed that three out of four metrics (precision, recall, accuracy) have improved values for our proposed model than the original SqueezeNet i.e., 98% whereas the mean training time for our proposed model is 13.7 which is less than the original model. The outcomes comparatively with the base network are shown in Table 7.

5.6 Comparison with existing DL models

We employed our model using ODIR and Kaggle datasets and attained better results as compared to existing methods for eye disease detection. In [23], the VGG19 model is applied to STARE and ImageNet databases which give 87.4% accuracy for retinal image classification. [83] gives 97.93% & 98.10% accuracies for the CenterNet model for the datasets APTOS-2019 and IDRiD. [4] applied RESNET-50 on hospitals and IDRiD datasets. It gives 81% accuracy with hospital datasets and 86% accuracy with the IDRiD dataset. [84] implied fast region-based convolutional neural network (FRCNN) on the datasets HRF and ORIGA. Classification accuracies for these datasets are found to be 95.8% and 95.7%. In [85], optimal residual deep neural networks (RESNETs) are applied for the

classification of eye diseases from normal eye images. Results show the accuracy values on datasets IDRiD and MESSIDOR as 73.12% and 69.65%. [86] applied deep convolutional neural network model for the classification of multiple eye diseases i.e., diabetic retinopathy (DR) and glaucoma. Datasets used for analysis were Kaggle dataset and Medimrg. The accurate value for classification was found to be 80%. [30] uses the VGG16 model for the detection of DR disease on the Eyepacs dataset and attained an 82% accurate value. [87] has used the MESSIDOR dataset and deep CNN model and got 97% accuracy in DR identification. Similarly, [3] applied the deep CNN model on HRF and ORIGA datasets and attained 93.86% accuracy. Our proposed model depicts a 98.9% accuracy value on ODIR and Kaggle datasets containing Cataract, Glaucoma, and DR images. On the dataset, it gives a 97.8% accurate value. As compared to other models, the proposed model shows better results. The comparison with existing deep learning methods is shown in Table 8.

5.7 Comparison with existing ML models

We compared our proposed model with machine learning algorithms and found better results. [20] implied Naïve Bayes algorithm for the detection of cataract eye disease. The dataset was collected from the interviews conducted with specialists/experts. It shows 86% accuracy. [89] classified DR disease, using k nearest neighbor, Naïve Bayes, polynomial kernel SVM, and RBF kernel SVM. Eye images were gathered from the ophthalmology department, at Melaka hospital, Malaysia. It shows 93%, 75%, 70%, and 93% accuracies for each ML technique respectively. [90] attained 88% accuracy in cataract detection, using the CNN model and Kambang Eye Hospital dataset. [9] applied probabilistic neural network (PNN), Bayesian classification, and support vector machine. Database DIARETDB0 (Evaluation Database and Methodology for Diabetic Retinopathy) was used for classification purposes. It showed 87.69%, 90.76%, and 95.38% accuracies for each ML model respectively. [55] gained 88.85% to 94.54% accuracy in classifying eye diseases using the CNN model and DRIVE dataset. [19] applied

Table 8 Comparison with Existing Deep Learning Models

Model	Dataset1 name	Accuracy (%)	Dataset2 name	Accuracy (%)
VGG-19 [23]	STARE database	87.4	ImageNet database	-
CenterNet [83]	APTOS-2019	97.93	IDRiD	98.10
RESNET-50 [4]	Hospitals dataset	81	IDRiD	86
FRCNN [84]	HRF	95.8	ORIGA	95.7
RESNETs [85]	IDRiD	73.12	MESSIDOR	69.65
DCNN [86]	Kaggle dataset	80	Medimrg	80
VGG16 [30]	Eyepacs dataset	82	-	-
DCNN [87]	MESSIDOR	97	-	-
DCCN [3]	HRF	93.86	ORIGA	-
DCNN [88]	Beijing Tongren Eye Center Database	93.52% (detection accu- racy), 86.69 (grading accuracy)	-	-
Our proposed Model	ODIR-IMAGE	98.9	Kaggle Dataset	97.8

the SVM technique for the classification of the DR eye from the normal eye. It used FLIR T400 thermal camera images and got 86.22% accuracy. [21] used random forest classifier and the databases DiaretDB0, DiaretDB1, and Messidor for classifying retinal diseases. It gained a 93.58% accurate value. Our proposed model outperformed the other techniques, showing an accuracy value of 98.9%. The comparative analysis with existing machine learning models is reported in Table 9.

5.8 Cross-Validation

In this section, we performed an evaluation using the Cataract dataset [80] for eye diseases. The Cataract dataset contained four classes such as cataract, glaucoma, DR, and normal. It consists of images having a size of 256×256 . Some samples from the dataset have been shown in Fig. 7. The cataract class samples were 100, glaucoma class samples were 101, DR class samples were 100, and normal class samples were 300. The purpose of this experiment was to analyze the generalization of the proposed model. Moreover, we have already trained our proposed model using mentioned datasets in the dataset section, therefore in this section we only tested the proposed model employing the cataract dataset. The results are exhibited in the form of a ROC curve in Fig. 8. It is depicted that our proposed model performs significantly over the cross-validation. Our proposed technique is very robust due to the squeezed layers of the proposed network which allow us to reduce the number of parameters and attain maximum accuracy. More precisely, the bypass connections allow to extract the most representative features from samples and each next layer gets the most meaningful features from the previous layer and transfers these features to the subsequent layers. The squeezed architecture may have less size than ResNet, and due to bypass connections, it has the minimum number of parameters. Therefore, as in the form of the ROC curve, the classification results into four classes using the cataract dataset have been presented. The proposed model attained 97.4% accuracy on cross-validation. From the results, it can be said that our proposed model is robust for eye disease detection on various unseen datasets.

Table 9 Comparison with Existing Machine Learning Models

Model	Dataset name	Accuracy
Naïve Bayes [20]	Specialists/experts interviews	86%
k-nearest neighbor, Naïve Bayes, polynomial kernel SVM, RBF kernel SVM [89]	Images from eye clinic, Department of Ophthalmology, Hospital Melaka, Malaysia	93%, 75%, 70%, 93% for each model respectively
Convolutional Neural Network [90]	Kambang Eye Hospital	88%
Probabilistic Neural network (PNN), Bayesian Classification, and Support vector machine (SVM) [9]	DIARETDB0: Evaluation Database and Methodology for Diabetic Retinopathy"	87.69%, 90.76%, and 95.38% respectively for each ML model
CNN [55]	DRIVE dataset	88.85% to 94.54%
SVM [19]	FLIR T400 thermal Camera images	86.22%
Random Forest classifier [21]	DiaretDB0, DiaretDB1, Messidor	93.58%
Our proposed Model	Kaggle datasets	98.9%

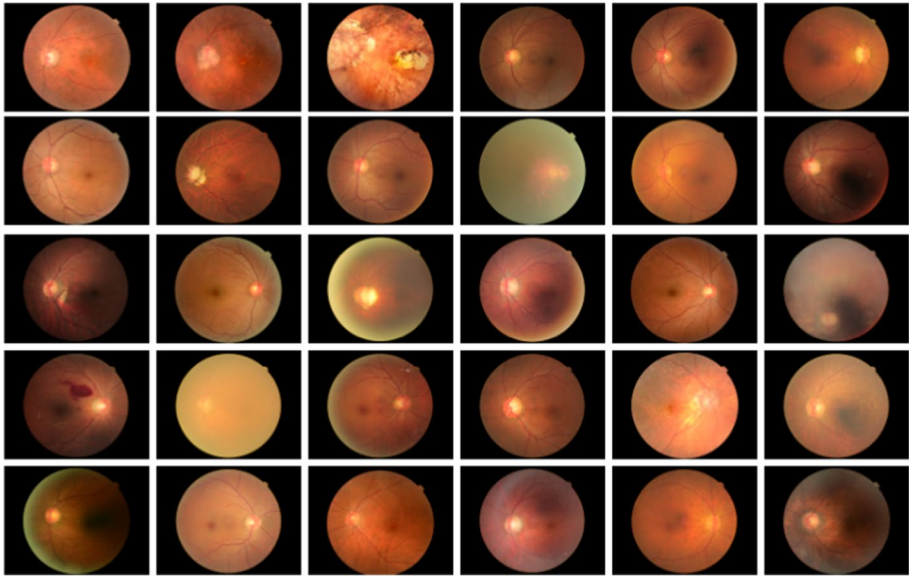


Fig. 7 Samples from Cataract Dataset

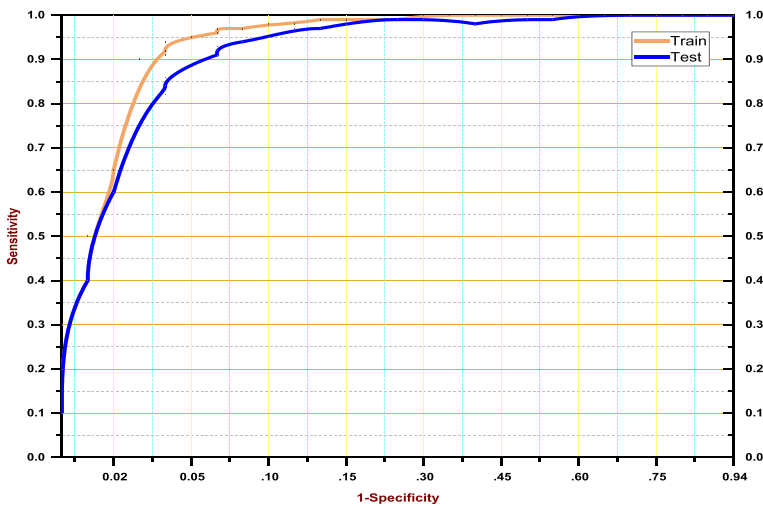


Fig. 8 ROC curve for Cross-validation

5.9 Comparison on same dataset

The performance of the proposed model is assessed by comparing it with 6 state-of-the-art deep learning models, namely VGG16, VGG19 [91], AlexNet [92], ResNet [93], InceptionNet-v4 [94], and Vision Transformer [95]. These models are renowned for their robustness in classification tasks and have demonstrated significant performance in various works. To ensure a fair and proper assessment, we used the same hyper-parameters and

Table 10 The comparison with existing DL models on same dataset

Model	Accuracy(%)	Precision(%)	Recall(%)
VGG16	76.3	72.2	74.2
VGG19	78.9	75.7	76.5
AlexNet	63.4	64.3	61.3
ResNet	87.2	88.8	86.1
InceptionNet-v4	84.2	85.2	82.4
VisionTransformer	89.1	87.7	86.3
Our proposed approach	98.9	98.87	98.8

dataset for training and testing as described in dataset section. The outcomes are reported in Table 10, and it is clearly visible that our proposed model achieved considerable performance for eye disease detection.

6 Discussion

In this study, we have proposed a model based on SqueezeNet and BAM attention mechanism to classify eye diseases. SqueezeNet is designed to have a significantly smaller model size compared to traditional deep neural networks like AlexNet or VGG, making it more memory-efficient and easier to deploy on resource-constrained devices such as mobile phones or embedded systems. Moreover, the reduced number of parameters and computations in SqueezeNet results in faster inference speed. However, the compactness of SqueezeNet comes at the cost of a reduced representation capacity compared to deeper architectures like ResNet or InceptionNet. This may lead to less capability in capturing intricate features or patterns from complex datasets. Therefore, we attached the BAM attention mechanism to enhance the feature representation.

7 Conclusion

In this research, we propose an advanced deep learning model with the addition of a batch normalization layer, classification layer, and BAM attention module to detect eye diseases i.e., cataract, glaucoma, and diabetic retinopathy. Moreover, the proposed system is based on an efficient pre-trained SqueezeNet architecture that is lighter and computes fast as compared to other deep neural networks. It contains five modules that consist of convolutional, downsampling, and global average pooling layers. It gives AlexNet-level accuracy with 50% fewer parameters. Our improved SqueezeNet model outperforms other deep learning and machine learning models. In our proposed model, we employed Bottleneck Attention Module (BAM) with SqueezeNet having an additional layer. Our proposed attention module utilizes two different ways and effectively extracts the most representative features and drops the image's background features of eyes which don't take part in the detection of diseases. Moreover, we utilized three datasets in the proposed study such as ORID and Kaggle Datasets used for training and testing, and the cataract Dataset used for the cross-validation. Multiple experiments have been carried out to assess the performance of the proposed model. The model attained 98.9% accuracy for the overall detection and classification of

eye diseases. More precisely, 99.2%, 98.9%, 98.3%, 99.2%, and 98.9% accuracies have been attained for the cataract, glaucoma, DR, Neutral classes, and average over the ODIR dataset. In addition, 98.08% accuracy has been achieved over cross-validation on the cataract dataset, which exhibits the robustness of the proposed algorithm. We also evaluated our improved SqueezeNet with the comparison of the original SqueezeNet and found our proposed model better than the other. Our improved model shows 98.9% accuracy in comparison to the original SqueezeNet i.e., 96%. The most useful key aspect of our research is to use it for the timely and accurate detection of eye diseases to minimize the time and cost needed for additional examination techniques. The model efficiently identifies the diseases in eye images due to the pre-trained proposed architecture over the large dataset.

The proposed technique can also be employed for other disease detection and classification i.e., knee diseases. In the future, we aim to improve our proposed method by employing auto-fine-tuning techniques to minimize training time. Moreover, we will also explore the combination of our proposed classifier with other imaging modalities or patient data, such as genetic information or electronic health records, to enhance diagnostic accuracy and personalized medicine.

Acknowledgements The authors present their appreciation to King Saud University for funding this research through Researchers Supporting Program number (RSPD2023R533), King Saud University, Riyadh, Saudi Arabia.

Data availability The datasets generated during and/or analysed during the current study are available in the Kaggle repository, [<https://www.kaggle.com/datasets/annisarizki/odir-preprocessing-augmentation>].

Declarations

Conflict of interest The authors have no conflicts of interest to declare that are relevant to the content of this article.

References

1. Akram A, Debnath R (2020) An automated eye disease recognition system from visual content of facial images using machine learning techniques. *Turkish J Electr Eng Comput Scis* 28(2):917–932
2. Pratap T, Kokil P (2021) Efficient network selection for computer-aided cataract diagnosis under noisy environment. *Comput Methods Programs Biomed* 200:105927
3. Mahum R, Irtaza A, Javed A (2023) EDL-Det: A Robust TTS Synthesis Detector Using VGG19-Based YAMNet and Ensemble Learning Block. *IEEE Access* 11:134701–134716. <https://doi.org/10.1109/ACCESS.2023.3332561>
4. Hossain MR, Afroze S, Siddique N, Hoque MM (2020) Automatic detection of eye cataract using deep convolution neural networks (DCNNs). In 2020 IEEE region 10 symposium (TENSYP). IEEE, pp 1333–1338
5. An G, Omodaka K, Tsuda S, Shiga Y, Takada N, Kikawa T, Nakazawa T, Yokot H, Akiba M (2019) Comparison of Machine-Learning Classification Models for Glaucoma Management. *J Healthc Eng* 2018:8. <https://doi.org/10.1155/2018/6874765>
6. Tong Y et al (2020) Application of machine learning in ophthalmic imaging modalities. *Eye and Vision* 7(1):1–15
7. Syarifah MA, Bustamam A, Tampubolon PP (2020) Cataract classification based on fundus image using an optimized convolution neural network with lookahead optimizer. In AIP Conference Proceedings 2296(1)
8. Ibrahim I, Abdulazeez A (2021) The role of machine learning algorithms for diagnosing diseases. *J Appl Sci Technol Trends* 2(01):10–19
9. Priya R, Aruna P (2013) Diagnosis of diabetic retinopathy using machine learning techniques. *ICTACT J Soft Comput* 3(4):563–575

10. Mahum R et al (2021) A novel hybrid approach based on deep cnn features to detect knee osteoarthritis. *Sensors* 21(18):6189
11. Munir MH, Mahum R, Nafees M, Aitazaz M, Irtaza A (2022) An automated framework for Corona virus severity detection using combination of AlexNet and faster RCNN. *International Journal of Innovations in Science and Technology* 3:197–209
12. Mahum Rabbia, Munir Haris, Mughal Zaib-Un-Nisa, Awais Muhammad, Khan Falak Sher, Saqlain Muhammad, Mahamad Saipunidzam, Thili Iskander (2023) A novel framework for potato leaf disease detection using an efficient deep learning model. *Hum Ecol Risk Assess Int J* 29(2):303–326
13. Zhao Y et al (2017) Automatic recognition of fMRI-derived functional networks using 3-D convolutional neural networks. *IEEE Trans Biomed Eng* 65(9):1975–1984
14. Abdel-Hamid O et al (2014) Convolutional neural networks for speech recognition. *IEEE/ACM Trans Audio Speech Language Process* 22(10):1533–1545
15. Korpusik M, Collins Z, Glass J (2017) Semantic mapping of natural language input to database entries via convolutional neural networks. In 2017 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, pp 5685–5689
16. Bastanfard A, Amirkhani D, Mohammadi M (2022) Toward image super-resolution based on local regression and nonlocal means. *Multimed Tools Appl* 81(16):23473–23492
17. Minoofam SAH, Bastanfard A, Keyvanpour MR (2023) TRCLA: A Transfer Learning Approach to Reduce Negative Transfer for Cellular Learning Automata. *IEEE Trans Neural Netw Learn Syst* 34(5):2480–2489. <https://doi.org/10.1109/TNNLS.2021.3106705>
18. Harifi S, Bastanfard A (2015) Efficient iris segmentation based on converting iris images to high dynamic range images. In 2015 Second International Conference on Computing Technology and Information Management (ICCTIM). IEEE, pp 115–119
19. Selvathi D, Suganya K (2019) Support vector machine based method for automatic detection of diabetic eye disease using thermal images. In 2019 1st International Conference on Innovations in Information and Communication Technology (ICICT). IEEE, pp 1–6
20. Darussalam U, Benrahman B (2020) Web-Based Expert System for Diagnosing Human Eye Disease Using the Naïve Bayes Method. *J Teknik Informatika CIT Medicom* 12(1):16–25
21. Chowdhury AR, Chatterjee T, Banerjee S (2019) A Random Forest classifier-based approach in the detection of abnormalities in the retina. *Med Biol Eng Compu* 57(1):193–203
22. Rajyaguru V, Vithalani C, Thanki R (2022) A literature review: various learning techniques and its applications for eye disease identification using retinal images. *Int J Inf Technol* 14:713–724. <https://doi.org/10.1007/s41870-020-00442-8>
23. Choi JY et al (2017) Multi-categorical deep learning neural network to classify retinal images: A pilot study employing small database. *PLoS ONE* 12(11):e0187336
24. Sarki R et al (2020) Automatic detection of diabetic eye disease through deep learning using fundus images: A survey. *IEEE Access* 8:151133–151149
25. Iandola FN, Han S, Moskewicz MW, Ashraf K, Dally WJ, Keutzer K (2016) SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5 MB model size. *arXiv preprint arXiv:1602.07360*
26. Pahuja Rahul (2022) Udit Sisodia, Abhishek Tiwari, Siddharth Sharma, and Preeti Nagrath (2021) "A Dynamic approach of eye disease classification using deep learning and machine learning model. *Proc Data Analytics Manag: ICDAM 2021* 1:719–736 (Springer Singapore)
27. Thanki R (2023) A deep neural network and machine learning approach for retinal fundus image classification. *Healthcare Analytics* 3:100140
28. Zhou Y, Li G, Li H (2019) Automatic cataract classification using deep neural network with discrete state transition. *IEEE Trans Med Imaging* 39(2):436–446
29. Xu X et al (2019) A hybrid global-local representation CNN model for automatic cataract grading. *IEEE J Biomed Health Inform* 24(2):556–567
30. Nguyen QH, Muthuraman R, Singh L, Sen G, Tran AC, Nguyen BP, Chua M (2020) Diabetic retinopathy detection using deep learning. In *Proceedings of the 4th international conference on machine learning and soft computing*, pp. 103–107
31. Vyas M (2015) Kaggle diabetic retinopathy detection competition report. <https://www.kaggle.com/meenavyas/diabetic-retinopathy-detection>
32. Pratt H et al (2016) Convolutional neural networks for diabetic retinopathy. *Procedia Comp Sci* 90:200–205
33. Islam SM, Hasan MM, Abdullah S (2018) Deep learning based early detection and grading of diabetic retinopathy using retinal fundus images. *arXiv preprint arXiv:1812.10595*
34. de La Torre J, Valls A, Puig D (2020) A deep learning interpretable classifier for diabetic retinopathy disease grading. *Neurocomputing* 396:465–476

35. Raju M et al (2017) Development of a deep learning algorithm for automatic diagnosis of diabetic retinopathy. MEDINFO 2017: Precision Healthcare through Informatics. IOS Press, pp 559–563
36. Almazroa A, Alodhayb S, Osman E, Ramadan E, Hummadi M, Dlaim M, Alkatee M, Raahemifar K, Lakshminarayanan V (2018) Retinal fundus images for glaucoma analysis: the RIGA dataset. In Medical Imaging 2018: Imaging Informatics for Healthcare, Research, and Applications. SPIE. 10579:55–62
37. Hagos MT, Kant S (2019) Transfer learning based detection of diabetic retinopathy from small dataset. arXiv preprint arXiv:1905.07203
38. Gulshan V 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
39. Wang Z, Yin Y, Shi J, Fang W, Li H, Wang X (2017) Zoom-in-net: Deep mining lesions for diabetic retinopathy detection. In Medical Image Computing and Computer Assisted Intervention– MICCAI 2017: 20th International Conference, Quebec City, QC, Canada, September 11–13, 2017, Proceedings, Part III 20 (pp. 267–275). Springer International Publishing
40. Li X et al (2019) CANet: cross-disease attention network for joint diabetic retinopathy and diabetic macular edema grading. IEEE Trans Med Imaging 39(5):1483–1493
41. Shankaranarayana SM et al (2017) Joint optic disc and cup segmentation using fully convolutional and adversarial networks. Fetal, infant and ophthalmic medical image analysis. Springer, pp 168–176
42. Bajwa MN et al (2019) Two-stage framework for optic disc localization and glaucoma classification in retinal fundus images using deep learning. BMC Med Inform Decis Mak 19(1):1–16
43. Mahum R, Rehman SU, Okon OD, Alabrah A, Meraj T, Rauf HT (2021) A novel hybrid approach based on deep CNN to detect glaucoma using fundus imaging. Electronics 11(1):26
44. Kumar KS, Singh NP (2023) Retinal disease prediction through blood vessel segmentation and classification using ensemble-based deep learning approaches. Neural Comput Appls 35(17):12495–12511
45. Panda R et al (2018) Deep convolutional neural network-based patch classification for retinal nerve fiber layer defect detection in early glaucoma. J Med Imaging 5(4):044003
46. Li A, Cheng J, Wong DWK, Liu J (2016) Integrating holistic and local deep features for glaucoma classification. In 2016 38th annual international conference of the IEEE engineering in medicine and biology society (EMBC). IEEE, pp 1328–1331
47. Shibata N et al (2018) Development of a deep residual learning algorithm to screen for glaucoma from fundus photography. Sci Rep 8(1):1–9
48. Amirkhani D, Bastanfard A (2019) Impainted Image Quality Evaluation Based on Saliency Map Features. 2019 5th Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS). Shahrood, Iran. <https://doi.org/10.1109/ICSPIS48872.2019.9066140>
49. Chen X, Xu Y, Wong DWK, Wong TY, 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, pp 715–718
50. Kim SJ, Cho KJ, Oh S (2017) Development of machine learning models for diagnosis of glaucoma. PLoS ONE 12(5):e0177726
51. Asaoka R et al (2016) Detecting preperimetric glaucoma with standard automated perimetry using a deep learning classifier. Ophthalmology 123(9):1974–1980
52. Li Z et al (2018) Efficacy of a deep learning system for detecting glaucomatous optic neuropathy based on color fundus photographs. Ophthalmology 125(8):1199–1206
53. Zilly J, Buhmann JM, Mahapatra D (2017) Glaucoma detection using entropy sampling and ensemble learning for automatic optic cup and disc segmentation. Comput Med Imaging Graph 55:28–41
54. Maninis KK, Pont-Tuset J, Arbeláez P, Van Gool L (2016) Deep retinal image understanding. In Medical Image Computing and Computer-Assisted Intervention–MICCAI 2016: 19th International Conference, Athens, Greece, October 17–21, 2016, Proceedings, Part II 19 (pp. 140–148). Springer International Publishing
55. Tan JH, Acharya UR, Bhandary SV, Chua KC, Sivaprasad S (2017) Segmentation of optic disc, fovea and retinal vasculature using a single convolutional neural network. J Comput Sci 20:70–79
56. Srivastava R, Cheng J, Wong DWK, Liu J (2015) Using deep learning for robustness to parapapillary atrophy in optic disc segmentation. In 2015 IEEE 12th international symposium on biomedical imaging (ISBI). IEEE pp. 768–771
57. Novotny A, Odstreilik Jan, Kolar Radim, Jan Jiří (2010) Textural analysis of nerve fibre layer in retinal images via local binary patterns and Gaussian Markov random fields. Anal Biomed Signals Images 20:308–315
58. Zhang Z, Liu J, Wong WK, Tan NM, Lim JH, Lu S, Li H, Liang Z, Wong TY (2009) Neuro-retinal optic cup detection in glaucoma diagnosis. In 2009 2nd International Conference on Biomedical Engineering and Informatics. IEEE pp. 1–4


59. Qureshi I (2015) Glaucoma detection in retinal images using image processing techniques: a survey. *Int J Adv Networking Appl* 7(2):2705
60. Acharya UR et al (2011) Automated diagnosis of glaucoma using texture and higher order spectra features. *IEEE Trans Inf Technol Biomed* 15(3):449–455
61. Reed S, Lee H, Anguelov D, Szegedy C, Erhan D, Rabinovich A (2014) Training deep neural networks on noisy labels with bootstrapping. *arXiv preprint arXiv:1412.6596*
62. Simonyan K, A Zisserman (2014) Very deep convolutional networks for large-scale image recognition. In *ICLR* May 7 - 9, 2015, *arXiv preprint arXiv:1409.1556*
63. Krizhevsky A, I Sutskever, GE Hinton (2012) Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems*, 2012
64. Nair V, Hinton GE (2010) Rectified linear units improve restricted boltzmann machines." In *Proceedings of the 27th international conference on machine learning (ICML-10)*, pp. 807–814
65. Srivastava N et al (2014) Dropout: a simple way to prevent neural networks from overfitting. *J Mach Learn Res* 15(1):1929–1958
66. Lin M, Chen Q, Yan S (2013) Network in network. *arXiv preprint arXiv:1312.4400*
67. Mishkin D, Sergievskiy N, Matas J (2017) Systematic evaluation of convolution neural network advances on the imagenet. *Comput Vis Image Underst* 161:11–19
68. Jia Y, Shelhamer E, Donahue J, Karayev S, Long J, Girshick R, Guadarrama S, Darrell T (2014) Caffe: Convolutional architecture for fast feature embedding. In *Proceedings of the 22nd ACM international conference on Multimedia*, pp. 675–678
69. Chen T, Li M, Li Y, Lin M, Wang N, Wang M, Xiao T, Xu B, Zhang C, Zhang Z (2015) Mxnet: A flexible and efficient machine learning library for heterogeneous distributed systems. *arXiv preprint arXiv:1512.01274*
70. Yang H, Fritzsche M, Bartz C, Meinel C (2017) Bmxnet: An open-source binary neural network implementation based on mxnet. In *Proceedings of the 25th ACM international conference on Multimedia*, pp. 1209–1212
71. Tokui S, Oono K, Hido S, Clayton J (2015) Chainer: a next-generation open source framework for deep learning. In *Proceedings of workshop on machine learning systems (LearningSys) in the twenty-ninth annual conference on neural information processing systems (NIPS)*, 5:1–6
72. Kora P, Ooi CP, Faust O, Raghavendra U, Gudigar A, Chan WY, Meenakshi K, Swaraja K, Plawiak P, Rajendra Acharya U (2022) Transfer learning techniques for medical image analysis: A review. *Biocybernetics and Biomedical Engineering* 42(1):79–107
73. Brownlee Jason (2019) Develop deep learning models on theano and TensorFlow using keras. *J Chem Inf Model* 53(9):1689–1699
74. DT42 (2016) Squeezenet keras implementation. https://github.com/DT42/squeezenet_demo
75. Collobert R, Kavukcuoglu K, Farabet C (2011) Torch7: A matlab-like environment for machine learning. In *BigLearn, NIPS workshop*, no. CONF
76. Waghmare SM (2016) FireModule.lua. <https://github.com/Element-Research/dpnn/blob/master/FireModule.lua>
77. Park J, Woo S, Lee J-Y, Kweon IS (2018) Bam: Bottleneck attention module. *arXiv preprint arXiv:1807.06514*
78. Dataset: Peking University International Competition on Ocular Disease Intelligent Recognition (ODIR-2019). <https://www.kaggle.com/andrewmvd/ocular-disease-recognition-odir5k>
79. Chaudhary PK, Pachori RB (2022) Automatic diagnosis of different grades of diabetic retinopathy and diabetic macular edema using 2-D-FBSE-FAWT. *IEEE Trans Instrum Meas* 71:1–9
80. Anonymous, Cataract Dataset (2019) <https://www.kaggle.com/datasets/jr2ngb/cataractdataset?select=dataset>
81. Edward Zhang and Sauman Das *Glaucoma Detection* (2021) <https://www.kaggle.com/datasets/sshikamaru/glaucoma-detection>
82. Sovit Ranjan Rath, *Diabetic Retinopathy* (2019) <https://www.kaggle.com/datasets/sovirath/diabetic-retinopathy-224x224-2019-data>
83. Nazir T et al (2021) Detection of diabetic eye disease from retinal images using a deep learning based CenterNet model. *Sensors* 21(16):5283
84. Nazir T et al (2020) Retinal image analysis for diabetes-based eye disease detection using deep learning. *Appl Sci* 10(18):6185
85. Chea N, Nam Y (2021) Classification of Fundus Images Based on Deep Learning for Detecting Eye Diseases. *Computers, Materials & Continua* 67(1)

86. Prasad K, Sajith PS, Neema M, Madhu L, Priya PN (2019) Multiple eye disease detection using Deep Neural Network. In TENCON 2019-2019 IEEE Region 10 Conference (TENCON). IEEE pp. 2148–2153
87. Hemanth DJ, Deperlioglu O, Kose U (2020) An enhanced diabetic retinopathy detection and classification approach using deep convolutional neural network. *Neural Comput Appl* 32(3):707–721
88. Zhang L, Li J, Han H, Liu B, Yang J, Wang Q (2017) Automatic cataract detection and grading using deep convolutional neural network. In 2017 IEEE 14th international conference on networking, sensing and control (ICNSC). IEEE pp. 60–65
89. Rahim SS et al (2016) Automatic screening and classification of diabetic retinopathy and maculopathy using fuzzy image processing. *Brain informatics* 3(4):249–267
90. Weni I et al (2021) Detection of Cataract Based on Image Features Using Convolutional Neural Networks. *Indones J Comput Cybern Syst* 15(1):75–86
91. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. <http://arxiv.org/abs/1409.1556>
92. Krizhevsky A, Sutskever I, Hinton GE (2017) ImageNet classification with deep convolutional neural networks. *Commun ACM* 60:84–90. <https://doi.org/10.1145/3065386>
93. Yu X, Kang C, Guttery DS, Kadry S, Chen Y, Zhang Y-D (2020) ResNet-SCDA-50 for breast abnormality classification. *IEEE/ 10570 Neural Computing and Applications* (2023) 35:10551–10571123 *ACM Trans Comput Biol Bioinforma.* 18:94–102. <https://doi.org/10.1109/tcbb.2020.2986544>
94. Szegedy C, Ioffe S, Vanhoucke V, Alemi A (2016) Inception-v4, inception-ResNet and the impact of residual connections on learning. <http://arxiv.org/abs/1602.07261>
95. Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, Dehghani M, Minderer M, Heigold G, Gelly S, Uszkoreit J (2020) An image is worth 16x16 words: transformers for image recognition at scale. <http://arxiv.org/abs/2010.11929>

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

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.

Authors and Affiliations

Amna Zia¹ · Rabbia Mahum²  · Nabeel Ahmad¹ · Muhammad Awais^{3,4} · Ahmad M. Alshamrani⁵

✉ Rabbia Mahum
rabbia.mahum@uettaxila.edu.pk

Amna Zia
amnazia113@gmail.com

Nabeel Ahmad
nabeel.ahmed@cydea.tech

Muhammad Awais
dr.muhammad.awais@henau.edu.cn

Ahmad M. Alshamrani
ahmadm@ksu.edu.sa

¹ Department of Computer Science, Sir Syed CASE Institute of Information Technology, Islamabad, Pakistan

² Department of Computer Science, University of Engineering and Technology, Taxila, Pakistan

³ Department of Electrical Engineering, Henan Agricultural University, Zhengzhou 450002, China

⁴ Henan International Joint Laboratory of Laser Technology in Agriculture Sciences, Zhengzhou 450002, China

⁵ Statistics and Operations Research Department, College of Science, King Saud University, Riyadh 11451, Kingdom of Saudi Arabia