

Received August 14, 2021, accepted September 10, 2021, date of publication September 15, 2021, date of current version September 24, 2021.

Digital Object Identifier 10.1109/ACCESS.2021.3112938

CataractNet: An Automated Cataract Detection System Using Deep Learning for Fundus Images

MASUM SHAH JUNAYED^{1,2}, MD BAHARUL ISLAM^{1,2,3}, (Senior Member, IEEE), AREZOO SADEGHZADEH², AND SAIMUNUR RAHMAN⁴

¹Department of Computer Science and Engineering, Daffodil International University, Dhaka 1207, Bangladesh

²Department of Computer Engineering, Bahcesehir University, Beşiktaş, 34349 Istanbul, Turkey

³College of Data Science and Engineering, American University of Malta, 1013 Cospicua, Malta

⁴CSIRO Data61, Epping, NSW 1710, Australia

Corresponding author: Md Baharul Islam (bislam.eng@gmail.com)

This work was supported in part by the Scientific and Technological Research Council of Turkey (TUBITAK) through the 2232 Outstanding Researchers Program under Project 118C301.

ABSTRACT Cataract is one of the most common eye disorders that causes vision distortion. Accurate and timely detection of cataracts is the best way to control the risk and avoid blindness. Recently, artificial intelligence-based cataract detection systems have been received research attention. In this paper, a novel deep neural network, namely *CataractNet*, is proposed for automatic cataract detection in fundus images. The loss and activation functions are tuned to train the network with small kernels, fewer training parameters, and layers. Thus, the computational cost and average running time of *CataractNet* are significantly reduced compared to other pre-trained Convolutional Neural Network (CNN) models. The proposed network is optimized with the Adam optimizer. A total of 1130 cataract and non-cataract fundus images are collected and augmented to 4746 images to train the model. For avoiding the over-fitting problem, the dataset is extended through augmentation before model training. Experimental results prove that the proposed method outperforms the state-of-the-art cataract detection approaches with an average accuracy of 99.13%.

INDEX TERMS Cataract detection, fundus images, neural network, classification.

I. INTRODUCTION

Cataract is a lenticular opacity clouding the transparent lens in human eyes. Typically, the lens converges the light to the retina. The presence of the cataract causes this light to be blocked and not reach the lens that results in poor visual acuity. It is a worldwide leading eye disease that develops gradually and does not affect sight early. However, after a while, it can interfere with vision and even cause vision loss in people over age 40 [1]. Cataract detection in earlier stages may avoid painful and costly surgeries and prevent blindness depending on its severity [2]. The world health organization (WHO) [3] reported that about 285 million people in the world have a visual impairment. Among them, 39 million people have limited vision, and the remaining ones have impaired vision. Cataract was responsible for 33% of visual impairment, and 51% of blindness [4]. In 2020, Flaxman *et al.* [5] predicted that the number of people suf-

fering from moderate to severe vision impairment (MSVI) and blindness would be 237.1 and 38.5 million, respectively. Of them, 57.1 million (24%) and 13.4 million (35%) people would be affected by cataract. The worldwide blindness will exceed 40 million by 2025 [6].

Comparing the results of these reports prove that there was only a slight improvement in the eye care system and controlling the vision loss during the last decade. Among the leading causes of blindness such as glaucoma [7], corneal opacity, trachoma, and diabetic retinopathy [8], cataract accounts for the most significant proportion. It is considered as one of the leading causes of blindness [5]. Cataract can be categorized into three main groups based on the location and area where it develops: Nuclear Cataract [9], Cortical Cataract [10], and Posterior Sub Capsular (PSC) Cataract [11]. These three types of cataracts occur due to several common factors such as aging, diabetes, and smoking [5].

Early detection of cataracts plays a vital role in the treatment and can significantly reduce the risk of blindness. Providing the automatic system for cataract detection is a

The associate editor coordinating the review of this manuscript and approving it for publication was Junhua Li¹.

challenging issue for three reasons, including (i) the vast spectrum of cataract lesions and human eye tones, (ii) the scale, form, and location of cataracts, and (iii) the age, gender, and eye type dependence. In recent years, automatic cataract detection based on different imaging modalities has been investigated. Generally, automatic cataract detection and classification systems utilize four types of images: slit lamp, retro-illumination, ultrasound, or fundus images. Among these imaging modalities, fundus images have attracted significant attention in this field as technologists or even patients themselves can easily employ the fundus camera [12]. In contrast, slit-lamp cameras need to be operated only by well-experienced ophthalmologists. Consequently, the lack of professional ophthalmologists, especially in underdeveloped countries, results in timely treatments [4]. Thus, to simplify the process for early cataract screening, an automatic cataract detection system based on fundus images is highly required.

Artificial intelligence-based systems for cataract detection are mostly based on global features (e.g. discrete cosine transformation (DCT) [13]), local features (e.g., local standard deviation [14]) and deep features (e.g., deep CNN which have achieved higher accuracy). Although numerous deep-learning-based automatic cataract detection systems have been reported in the literature, they still suffer from limitations such as low detection accuracy, a high number of model parameters, and thus being computationally expensive.

A fundus image-based automatic cataract detection system is proposed in this paper to overcome the limitations mentioned above, which classifies the patients into two groups: cataract or non-cataract conditions. The novelties of the proposed method are as follows: 1) reducing the number of parameters in the model such as layers and weights, thus reducing the computational cost and time, and 2) increasing the detection accuracy based on the proposed deep neural network structure. Thus, the proposed preliminary cataract detection can be used for mass screening and cataract grading. The main contributions of this article are as follows:

- A cataract dataset is collected, reorganized, and pre-processed from different standard datasets of fundus images published in the last two decades, i.e., the high-resolution fundus (HRF) [15] image archive, fundus image registration (FIRE) [16] dataset, ACHIKO-I fundus image dataset [17], Indian diabetic retinopathy image dataset (IDRiD) [18], color fundus image database [19] and digital retinal images for vessel extraction (DRIVE) database [20]. Then, it is extended to a considerable number of images through data augmentation process.
- To detect cataract, a new 16-layer deep learning neural network, i.e., *CataractNet*, is proposed. The number of layers and the activation and loss functions are tuned to improve the detection accuracy significantly.
- Five CNN models, i.e., VGG-16, VGG-19, MobileNet, Inception-v3, and ResNet-50, have been implemented to compare and demonstrate the capability and accuracy of our proposed *CataractNet*.

The remaining part of this paper is organized as follows: Section II briefly discusses the related works reported in recent years. Section III explains the proposed *CataractNet* for automated cataract detection. The experimental setup is presented in section IV. The experimental results are discussed in section V. Finally, we conclude this work with future work directions in section VI.

II. RELATED WORKS

The state-of-the-art automatic cataract detection systems consist three steps: pre-processing, feature extraction, and classification [21]. These methods are categorized into two groups based on the algorithms used in either feature extraction or classification stages: Machine Learning (ML)-based and Deep Learning (DL)-based methods. These methods have been discussed in recent reports [21]–[24]. In this section, we briefly discuss some leading works in both groups.

A. MACHINE LEARNING-BASED METHODS

Gao *et al.* [25] proposed a computer-aided cataract detection system used for mass screening or as preprocessing for cataract grading. An enhanced texture feature was presented and used to train the linear discriminant analysis (LDA). The experimental results on a clinical database demonstrated an accuracy of 84.8%. Yang *et al.* [26] proposed an automatic cataract detection method that performed in three steps. A top-bottom hat transformation was utilized to improve the contrast between the foreground and background. The luminance and texture were considered as features. The classifier was constructed using a backpropagation neural network (BBNN) to classify the cataract severity into mild, medium, or severe stages.

Guo *et al.* [27] presented a computer-aided cataract classification based on fundus images. The feature extraction step was carried out using wavelet transform and sketch-based methods. Then, a multiclass discriminant analysis algorithm was used for cataract detection and grading where correct classification rates (CCRs) were respectively 90.9% and 77.1% for wavelet transform-based feature extraction and 86.1% and 74.0% for sketch-based feature extraction. In [28], Fuadah *et al.* used K-Nearest Neighbor (KNN) as the classifier and an optimal combination of texture features, i.e., dissimilarity, contrast, and uniformity. This system was implemented for smartphones and obtained a high accuracy of 97.5%. A competitive cataract detection system was proposed in [29] based on statistical texture analysis and KNN. The Gray-Level Co-occurrence Matrix (GLCM) was applied for texture feature extraction. The testing set was classified into normal or cataract conditions using KNN and received an average 94.5% accuracy. As the pupil area was cropped and extracted manually by users/experts in the processing step, it could not be considered as a fully automatic system. However, this system was performed on the eye images taken by a standard camera without using any slit lamp or fundus cameras.

Yang *et al.* [4] proposed an ensemble learning-based approach for cataract detection and grading. Three independent feature sets were extracted, and two learning models were formed for each group. The image classification was achieved by combining the multiple-based learning models based on the ensemble methods, whose CCRs were 93.2% and 84.5% for cataract detection and grading, respectively. Caixinha *et al.* [30] provided an in-vivo automatic Nuclear Cataract detection and classification system by applying machine learning and ultrasound techniques. It extracted 27 features in frequency and time domain, and support vector machine (SVM), Bayes, multilayer perceptron, and random forest classifiers were investigated in the classification phase. Although the methods based on ultrasound images achieve high accuracy in cataract screening, the imaging techniques are expensive with the complicated operation. In [31], the fundus images were classified as cataract images using an SVM classifier, and then an RBF Network graded their severity with the specificity of 93.33%. Rana and Galib [1] proposed a mobile application on a smartphone (Android, iOS, Windows) that enabled the public to carry out self-screening cataract detection. They considered the texture information and reported 85% efficiency.

Jagadale *et al.* [32] utilized Hough circle detection transform for detecting the center of the lens and their radius. Then, the statistical features were extracted and used in an SVM classifier for cataract detection with an accuracy of 90.25%. Sigit *et al.* [33] presented an android smartphone-based method for cataract detection. The classification was carried out by a single-layer perceptron method with an accuracy of 85%. Recently, a hierarchical feature extraction-based method was presented in [6] for cataract grading. The four-class classification problem of the cataract severity grading was transformed into three adjacent two-class classifications. These were carried out with three individual neural networks before integration. This system achieved the accuracies of 94.83% and 85.98% for cataract detection and grading, respectively.

B. DEEP LEARNING-BASED METHODS

Deep learning-based methods can learn the essential features and then integrate the feature learning steps into the model building process to decrease the incompleteness of the manual design features and use them in different medical imaging modalities [34], [35]. Gao *et al.* [36] investigated a deep learning-based method for grading the severity of Nuclear Cataracts from slit-lamp images. Local filters are obtained by clustering the image patches fed into a convolutional neural network (CNN). Then a set of recursive neural networks (RNNs) was used to extract more higher-order features. The cataract grading was performed using support vector regression. Zhang *et al.* [37] proposed a Deep CNN (DCNN) for cataract detection and grading that used the feature maps from the pooling layers of the architecture. This method was time-efficient and achieved 93.52% and 86.69% accuracies in cataract detection and grading, respectively.

Ran *et al.* [38] proposed a method for six-level cataract grading based on a combination of DCNN and Random Forests (RF). Three modules formed the proposed DCNN for extracting features at different levels on fundus images. On the other hand, a feature dataset was created by DCNN and used by RF to implement more intricate six-level cataract grading. On average, this method achieved an accuracy of 90.69%. This six-level grading system could help the specialists to understand the patients' condition more precisely. Pratap and Kokil [39] presented a computer-aided method for detecting cataract severity from normal to severe based on fundus images. This method utilized a pre-trained CNN as transfer learning for automatic cataract classification. The final classification was carried out using feature extraction, and an SVM classifier whose four-stage CCR was 92.91%. Jun *et al.* [40] proposed a cataract grading system based on Tournament based Ranking CNN composed of tournament structure and binary CNN models.

Hossain *et al.* [41] proposed an automatic cataract detection system using DCNNs and a trained classifier model based on Res-Net, whose accuracy was 95.77%. Recently, Zhang *et al.* [42] have provided an attention-based Multi-Model Ensemble method for automatic cataract detection on ultrasound images, which, to the best knowledge of the authors, obtained the highest accuracy (97.5%) among the other deep learning-based approaches in the literature. In this method, the whole system was composed of three main parts: an object detection network, three classification networks, and a model ensemble module. The performance was still low but satisfactory, especially in cases of inadequate training samples. However, the main limitation of this method was that it evaluated the cataract degree based on the blurriness of the retinal images, which can be caused by the cataract and the other eye diseases such as corneal edema and diabetes mellitus. Hence, different types of eye diseases may not be distinguished in this method. Almost a similar accuracy (97.47%) has been achieved recently by Khan *et al.* [43] for fundus images based on VGG-19 model with transfer learning approach on a recently published dataset in KAGGLE [44]. In another recent work by Pratap and Kokil [45], cataract diagnosis has been investigated under a noisy environment. A pre-trained CNN was applied for feature extraction formed of a set of locally- and globally-trained independent support vector networks. The obtained results proved its robustness against noise. It was the first work that investigated the robustness of the cataract detection systems.

It was observed that many works had been done based on conventional machine learning methods, while there are a few works reported on cataract detection and grading using deep learning methods. Therefore, there are still several challenges to deal with, such as improving the accuracy of the models while minimizing their complexity by reducing the number of training parameters, layers, depth, running time, and the overall model size.

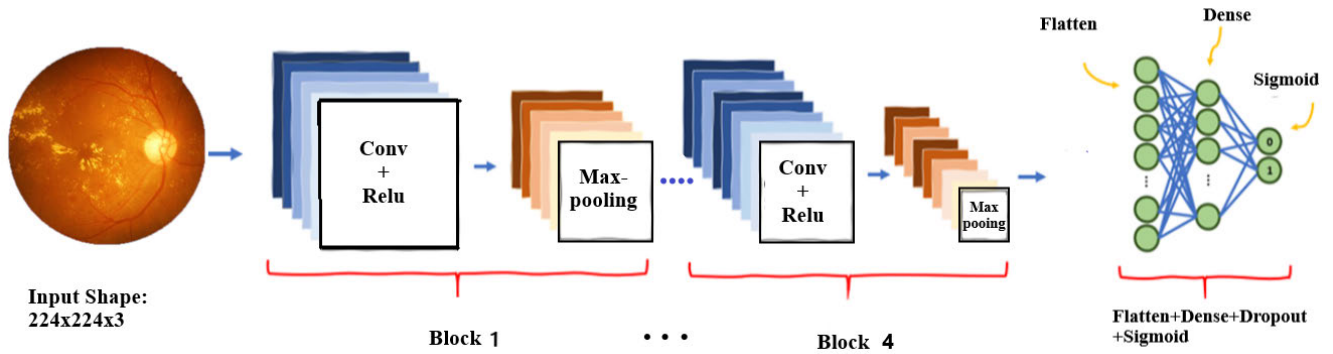


FIGURE 1. Proposed architecture of *CataractNet*. The model consists of four blocks each of which is made of convolutional and max-pooling layers.

III. PROPOSED CataractNet ARCHITECTURE

Convolutional Neural Network (CNN) is a deep neural network that acquires a complex hierarchy of features by convolutional, and pooling layers and non-linear activation functions [46], [47]. The feature extraction phase and the classification process are integrated into deep learning-based methods while these two steps are separated in the manual feature extraction methods. A novel deep learning model is proposed, namely *CataractNet*, to address the limitation of manual feature extraction and reduce the computational cost.

TABLE 1. Description of different layers of the proposed *CataractNet*.

Layer	Filter	Config.	Stride	Output
Conv2D	32	KS: 3 × 3; ReLU	-	222 × 222 × 32
Max-pooling2D	-	KS: 2 × 2	2	111 × 111 × 32
Conv2D	32	KS: 3 × 3; ReLU	-	109 × 109 × 32
Max-pooling2D	-	KS: 2 × 2	2	54 × 54 × 32
Conv2D	64	KS: 3 × 3; ReLU	-	52 × 52 × 64
Max-pooling2D	-	KS: 2 × 2;	2	26 × 26 × 64
Conv2D	128	KS: 3 × 3; ReLU	-	24 × 24 × 128
Max-pooling2D	-	KS: 2 × 2	2	12 × 12 × 128
Flatten	-	-	-	9216
Dense	-	64; ReLU	-	64
Dropout	-	0.4	-	64
Dense	-	128; ReLU	-	128
Dropout	-	0.4	-	128
Dense	-	256; ReLU	-	256
Dropout	-	0.5	-	256
Dense	-	Sigmoid	-	2

Figure 1 depicts the proposed *CataractNet* architecture. It contains 16 layers whose details are presented in Table 1. Half of these layers are placed in four blocks (each block is composed of two layers), and the rest is for classification. In the first block, the inputs are RGB (i.e., 3-input channels) images with the size of 224 × 224, and 32 filters with kernel sizes (KS) of 3 × 3 and padding as “valid” are applied. Then, a Max-Pooling (MP) layer with a stride of 2 is applied to reduce the space size of the data representation (width and height). It mainly minimizes the image size since a higher number of pixels corresponds to more parameters, requiring vast amounts of data. Finally, this block is activated by the ReLU activation function, which means the matrix’s negative values are considered 0 while the positive values are unchanged. The same block with the same values of

the parameters is used as the second block. Next, a similar block is used as the third block, but this time with 64 filters. In the fourth block, the number of filters is increased to 128. Outputs of all four blocks are combined into a feature map fed into the fully connected layers. These layers, namely flatten, dense, and dropout layers, are designed for cataract detection. Three sets of dense and dropout layers are constructed, among which dense layers are characterized with 64, 128, and 256 flattened neurons to collect the filtered cataract characteristics. Furthermore, three dropout layers are set as 0.4, 0.4, and 0.5 to prevent the model from overfitting by setting 40%, 40%, and 50% of neurons in hidden layers to 0 at each update of the training phase. As cataract detection is a binary classification, the sigmoid activation function is used in the last dense layer that is given as below:

$$\sigma(x) = \frac{1}{1 + e^{-x}} \tag{1}$$

The sigmoid function’s input is denoted by x and used as the final activation function. The total number of classes in the sigmoid layer is N , and each class represents one neuron. There are two main classes in our system: people with cataracts and normal eye conditions. In binary classification, the CNN architecture generates output at two neurons. Given a cataract image, the contribution of the first and the second neurons would be 1 or 0, or vice versa.

To investigate the effect of the block numbers on the classification accuracy and so the cataract detection, three different models based on 3, 4, and 5 blocks with three different sets of filters, i.e. (16, 32, 64), (32, 32, 64, 128) and (32, 32, 64, 96, 128), respectively, are developed and evaluated on the dataset. Table 2 represents the accuracy achieved from these models. The model’s efficiency is improved while increasing the number of blocks to 4. On the other hand, this efficiency was reduced at 5-blocks. Therefore, the model with 4-blocks and (32, 32, 64, 128) filters outperforms the others.

To explain the reason behind selecting the specific number of blocks and the convolutional layers, we go deeper into the effects of increasing the number of layers and filters on the extracted features. Patterns such as edges, dots, corners,

TABLE 2. Accuracy analysis over the number of the blocks.

No. of the Blocks	Description of Filters	Accuracy
3	16,32,64	97.84%
4	32,32,64,128	99.13%
5	32,32,64,96,128	97.17%

etc., are extracted in the first layer. Then, the following layers make more extensive patterns such as squares, circles, etc., by combining the extracted patterns from the previous layer. In our method, the required features of the cataract are extracted at 4-convolutional layers. Adding more layers does not necessarily always results in more accuracy. Increasing the number of layers helps to extract more features. Still, to a certain extent, it leads to overfitting and false positives instead of improving accuracy. It happened on *CataractNet* when the number of layers reached 5. Therefore, the optimal number of blocks for *CataractNet* is determined as 4.

As it is common in binary classification, binary cross-entropy is defined as the loss function as $Cross - entropy = -(i \log(p) + (1-i) \log(1-p))$, where i is the binary class marker predictor (0 or 1), the log is the normal logarithm, and p is the predicted probability.

IV. EXPERIMENTAL SETUP

A. DATASET

Employing a suitable dataset with many samples is critical in improving validation and training in deep learning-based classification. In this research, the high-resolution fundus (HRF) [15] image archive is used to collect all of the cataract retinal images. Additionally, some more images are employed from other datasets, i.e., fundus image registration (FIRE) [16] dataset, ACHIKO-I fundus image dataset [17], Indian diabetic retinopathy image dataset (IDRiD) [18], color fundus image database [19] and digital retinal images for vessel extraction (DRIVE) database [20], so that the total number of images in the classification phase is 1130, out of which 904 (80%) images are utilized for training and validation, and the remaining 226 (20%) images are used for testing. The model learns the patterns in the training process, while in the validation process, the weights are normalized. During the testing period, the model is evaluated for getting the accuracy and the loss.

B. PRE-PROCESSING

Since the dataset images are collected from various sources, their sizes are not identical and appropriate for the classification task. Consequently, the images are resized to a unified format with 224×224 pixels. For the RGB images, the intensities of the three channels are normalized in a range between 0 to 1. One of the essential steps before training the networks is image normalization to ensure that each input pixel includes a similar distribution, which results in faster convergence in training. The formula for

normalization is given as [48]:

$$N = \frac{p - MinV}{MaxV - MinV} \quad (2)$$

where p and N refer to the original intensity (between 0-255) and normalized intensity (between 0-1) of the cataract images, respectively, $MaxV$ and $MinV$ define the maximum and minimum intensities of the original images, respectively.

TABLE 3. Description of the dataset.

Class Name	Training Images	Testing Images	Total
Cataract	2560	119	2679
Non-cataract	1960	107	2067

C. DATA AUGMENTATION

The lack of an extensive training medical image dataset is a challenging issue that makes it hurdles to have further improvement in deep learning [49]. Hence, to deal with the insufficiency of the dataset, data augmentation is applied to training samples through four geometric transformations: re-scaling, rotation (30 degrees randomly to right or left), zooming, and horizontal flip, which results in 3616 additional training images (4 times of the original number of the training images) and prevents overfitting in the network. The numbers of the testing and training images in cataract and non-cataract classes are presented in detail in Table 3. The total numbers of non-cataract and cataract images are 2067 and 2679, respectively.

D. IMPLEMENTATION DETAILS

All the experiments are carried out in a computer with the following properties: core i9-10850K CPU, 64GB RAM, and NVIDIA Geforce RTX 2080 super GPU with 3.60 GHz. Image pre-processing, augmentation, and the CNN-based model are all implemented in Python, Keras, and Tensorflow environments. In our proposed *CataractNet*, a new model is implemented and optimized by an ADAM optimizer with a learning rate of 0.0001, whose results approve that the combination of 32 batches works satisfactorily during *CataractNet* training.

E. EVALUATION CRITERIA

Accuracy alone is not sufficient for assessing the effectiveness of the model [55]. In addition to accuracy, various evaluation metrics such as Recall/Sensitivity, Precision, Specificity, F-Score, and Matthews Correlation Coefficient (MCC) are employed to evaluate our model and five pre-trained deep learning models, as accuracy = $\frac{(TP+TN)}{(TP+TN+FP+FN)}$, sensitivity = $\frac{(TP)}{(FN+TP)}$, precision = $\frac{(TP)}{(TP+FP)}$, specificity = $\frac{(TN)}{(FP+TN)}$, F1-Score = $\frac{2*(Precision*Recall)}{(Precision+Recall)}$, MCC = $\frac{(TP*TN)-(FP*FN)}{\sqrt{(TP+FP)*(TP+FN)*(TN+FP)*(TN+FN)}}$. The TP , TN , FP , and FN state for true-positive, true-negative, false-positive, and false-negative samples, respectively. Accuracy and F1-Score are successful metrics that are calculated by averaging the

TABLE 4. Comparison between the performance of the proposed CataractNet and those of the other five pre-trained models.

Train-Test set (%)	Methods	Accuracy (%)	Precision (%)	Recall (%)	Specificity (%)	F1-Score (%)	MCC (%)
90% - 10%	Inception-v3 [50]	83.51	79.79	74.58	74.03	81.26	71.55
	MobileNet [51]	86.11	78.46	74.38	75.40	71.62	74.35
	VGG-16 [52]	91.35	86.19	91.92	95.46	89.54	83.21
	VGG-19 [53]	92.88	86.68	91.76	97.83	89.97	84.01
	ResNet-50 [54]	91.18	85.69	90.43	95.49	86.52	83.57
	CataractNet	92.92	86.37	95.43	94.75	89.51	84.17
80% - 20%	Inception-v3 [50]	87.85	85.38	86.24	81.18	88.26	88.95
	MobileNet [51]	88.62	86.43	84.39	82.47	87.95	89.47
	VGG-16 [52]	91.35	86.19	91.92	97.46	90.54	90.85
	VGG-19 [53]	92.88	86.68	91.76	97.83	89.97	92.01
	ResNet-50 [54]	97.66	96.75	97.38	98.16	97.04	98.11
	CataractNet	99.13	99.08	99.07	99.17	99.07	98.23
70% - 30%	Inception-v3 [50]	86.65	88.79	83.49	86.28	88.21	87.99
	MobileNet [51]	87.74	88.16	84.51	85.76	88.53	88.87
	VGG-16 [52]	94.89	86.19	92.86	96.49	90.54	94.34
	VGG-19 [53]	95.63	86.68	93.76	97.83	95.97	96.13
	ResNet-50 [54]	96.26	95.62	97.43	96.89	98.56	96.78
	CataractNet	98.96	98.24	99.08	98.98	98.97	98.15
60% - 40%	Inception-v3 [50]	90.16	91.47	87.64	89.81	88.78	87.67
	MobileNet [51]	90.29	88.70	89.94	91.76	87.37	88.84
	VGG-16 [52]	94.83	94.25	94.42	93.89	95.31	93.52
	VGG-19 [53]	95.95	94.42	95.96	95.58	95.72	94.23
	ResNet-50 [54]	96.63	95.58	96.71	95.84	95.80	94.46
	CataractNet	96.76	95.44	96.72	95.89	95.68	96.59
50% - 50%	Inception-v3 [50]	83.51	91.79	74.58	92.03	81.26	86.78
	MobileNet [51]	89.58	88.73	90.46	89.17	89.54	85.07
	VGG-16 [52]	91.81	90.25	92.73	94.85	91.14	90.28
	VGG-19 [53]	92.88	91.78	95.76	92.55	91.76	93.17
	ResNet-50 [54]	94.38	94.62	93.27	93.13	95.02	95.89
	CataractNet	95.02	94.86	95.68	94.79	95.05	96.34

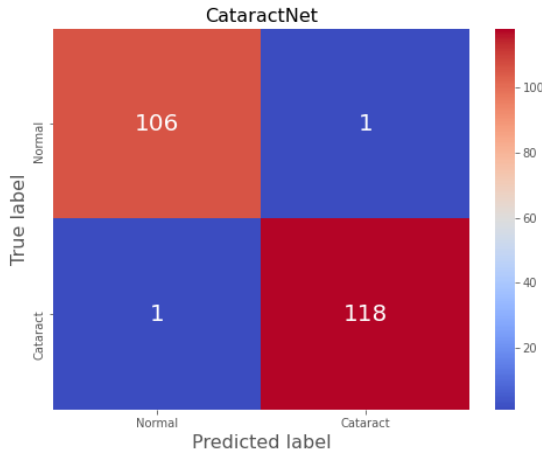


FIGURE 2. Confusion matrix of the proposed CataractNet.

results of the performance verification. The accuracy determines how the values are expected to be accurate. Precision learns how the measurement is reproducible or correctly predicted. Recall decides about the right outcome. To determine the average of all values, F1-score uses precision.

V. RESULTS AND DISCUSSIONS

In this section, we discuss the performance of the proposed CataractNet. Our model and five other pre-trained models are tested for cataract detection using the same

dataset. Experimental results are compared with state-of-the-art cataract detection methods.

A. PERFORMANCE OF CataractNet

We develop five pre-trained CNN models (with high classification results) to evaluate the performance of the proposed CataractNet on the same dataset. These models are (i) MobileNet [51], (ii) VGG-16 [52], (iii) VGG-19 [53], (iv) ResNet-50 [54], and (v) Inception-v3 [50]. We split dataset into training and testing sets as 90-10%, 80-20%, 70-30%, 60-40%, 50-50%. The performance of our CataractNet is compared with these pre-trained models in Table 4 in terms of Accuracy, Precision, Recall (Sensitivity), Specificity, F1-Score, and MCC. We mark the best performance in bold. Our proposed CataractNet ranks first and achieves the highest performance in all evaluation metrics for 80-20% and 70-30% dataset splitting conditions. CataractNet receives an accuracy of 99.13% that outperforms the other pre-trained models for 80-20% dataset splitting. The model parameters of these networks, e.g., model size, trainable parameters, total layers, depth, and average running time, are also compared in Table 5. In addition to the high performance, the proposed model has the least time complexity among others. Our model size and the trainable parameters (19.78 MB and 1.17 M, respectively) are significantly minimized compared to other pre-trained models. CataractNet has only 16 layers, and the average running time is 1035 seconds which is slightly less than that of

TABLE 5. Comparison between the proposed *CataractNet* and five pre-trained models in terms of size and other model parameters.

Model Name	Model Size	Trainable Parameters	Total Layers	Depth	Average running time
Inception-v3 [50]	92MB	23.9M	48	159	2621s
MobileNet [51]	49MB	4.2M	30	88	1296s
ResNet-50 [54]	98MB	25.6M	177	-	2765s
VGG-16 [52]	528MB	138M	16	23	3148s
VGG-19 [53]	549MB	143M	19	26	3279s
CataractNet	19.78 MB	1.17M	16	-	1035s

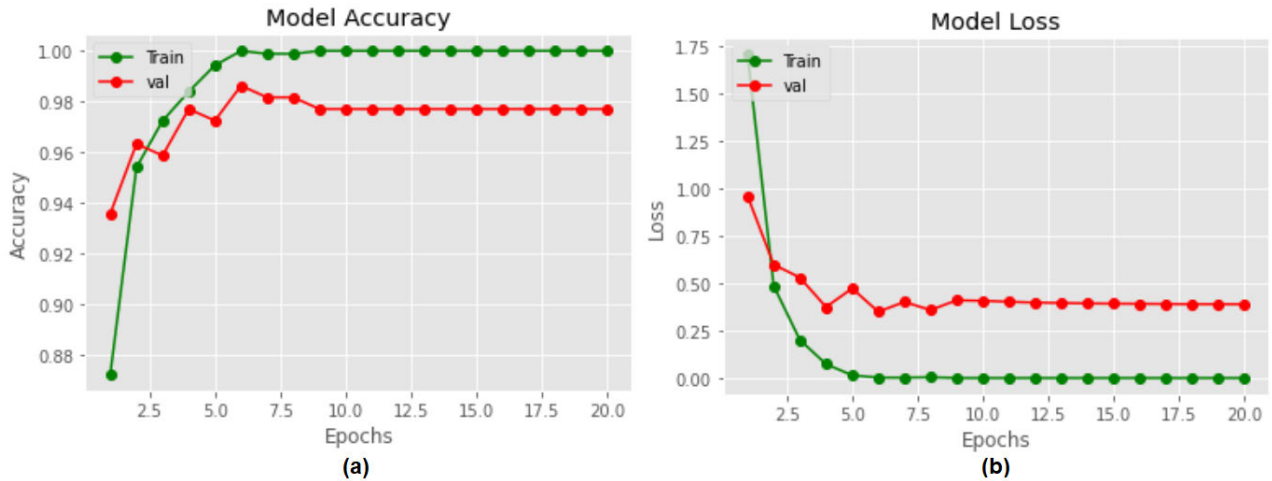


FIGURE 3. (a) Accuracy and (b) loss graphs of the proposed *CataractNet*.

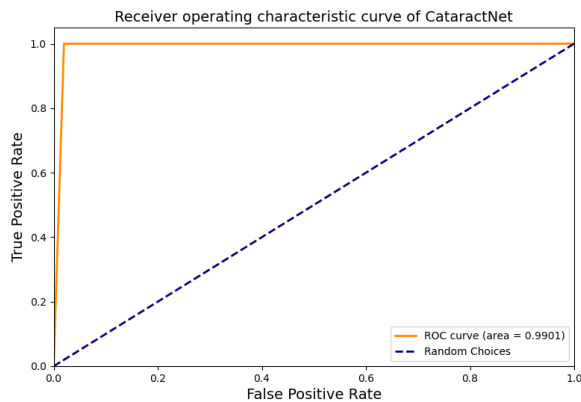


FIGURE 4. Receiver operating characteristic (ROC) curve for illustrating the relationship between the True Positive Rate (TPR) and the True Negative Rate (TNR) for showing the capability of the *CataractNet* in binary classification.

MobileNet [51] and almost three times smaller than other models [50], [52]–[54].

The confusion matrix of the *CataractNet* is depicted in Figure 2. Each column and row of the table corresponds to a predicted label and a true label, respectively. As illustrated in the matrix, there are only two wrong classification results where healthy eyes are detected as cataracts and vice versa in our test set.

In Figure 3, the (a) and (b) illustrate the training and validation accuracy and loss graphs of the *CataractNet*,

respectively. The X-axis represents the number of epochs, and Y-axis is the value of the accuracy or loss. As shown in these graphs, the validation accuracy and loss flatten out and become parallel to training after 9 epochs which is sufficient to simulate the pattern with minor anomalies.

To demonstrate the *CataractNet* efficiency, the Receiver Operating Characteristic (ROC) curve is plotted using a testing set to make the final prediction as shown in Figure 4. The X-axis and Y-axis represent the false-positive and true-positive rates. As demonstrated, the blue dash line is random choices (probability 50%), and its slope is 1 because there are two classes (cataract and normal eyes). The area under the curve (AUC) is 0.9901, which is nearly perfect.

A graphical representation of cataract detection using *CataractNet* is illustrated in Figure 5. As shown in this figure, our model can successfully detect the cataract no matter it's mild or severe. Two wrong classification results make it clear that they are classified wrongly with a slight difference in scores.

B. COMPARISON

The reported studies on cataract detection based on image processing, machine learning, and deep learning are summarized in Table 6 providing their accuracies and the size of their built datasets. Gao et al. [25], Guo et al. [27], and Fuadah et al. [29] used image processing methods for cataract detection with accuracies of 84.8%, 90.9%, and 94.5%, respectively. Yang et al. [4], Harini

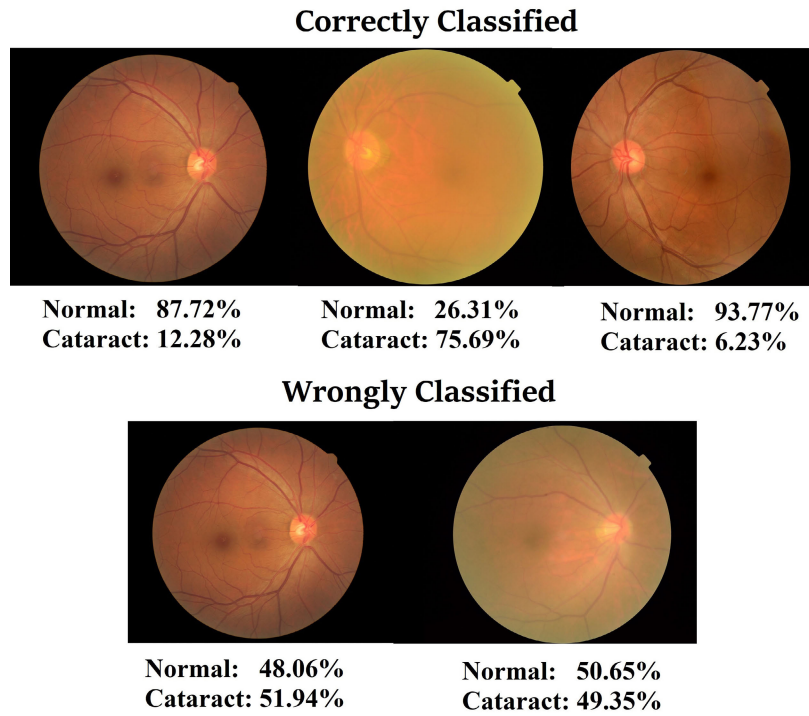


FIGURE 5. Graphical representation of cataract detection based on the proposed *CataractNet*.

TABLE 6. Accuracy comparison between the proposed method and the other state-of-the-art approaches in the available literature.

Approaches	Year	Methods	Dataset Size	Accuracy
Gao <i>et al.</i> [25]	2011	Texture analysis	4545	84.8%
Guo <i>et al.</i> [27]	2015	Wavelet transformation	445	90.9%
Fuadah <i>et al.</i> [29]	2015	Texture analysis & KNN	160	94.5%
Yang <i>et al.</i> [4]	2016	SVM & NN	1239	93.2%
Harini <i>et al.</i> [31]	2016	SVM	-	93.33%
Zhang <i>et al.</i> [37]	2017	Deep CNN	5620	93.52%
Ran <i>et al.</i> [38]	2018	Random forests & DCNN	5408	90.69%
Pratap <i>et al.</i> [39]	2019	AlexNet & SVM	800	92.91%
Sigit <i>et al.</i> [33]	2019	Single layer perceptron	50	85%
Cao <i>et al.</i> [6]	2020	NN	1355	94.83%
Hossain <i>et al.</i> [41]	2020	ResNet-50	4000	95.77%
Khan <i>et al.</i> [43]	2021	VGG-19	1400	97.47%
CataractNet	2021	Proposed CNN	4746	99.13%

and Bhanumathi [31], Sigit *et al.* [33], and Cao *et al.* [6] used machine learning-based classifiers to detect cataract with accuracies of 93.2%, 93.33%, 85%, and 94.83%, respectively. Zhang *et al.* [37], Ran *et al.* [38], Pratap and Kokil [39], Hossain *et al.* [41], and Khan *et al.* [43] proposed deep learning-based approaches with accuracies of 93.52%, 90.69%, 92.91%, 95.77%, and 97.47%, respectively.

It is worth mentioning that these methods have been implemented on their built datasets that are not publicly available. So, the accuracy can differ from different datasets. To compare our method with state-of-the-arts, it is required to implement and test them using our dataset. There are not any available source codes. We tried to contact these researchers through email to know the implementation settings, but unfortunately, we could not reach them.

TABLE 7. Comparison between the proposed method and the recent state-of-the-art approach in the Kaggle dataset.

Approaches	Dataset	Dataset Size	Accuracy
Khan <i>et al.</i> [43] (2021)	Kaggle dataset	1400	97.47%
CataractNet	Kaggle dataset	1400	98.62%

Consequently, our model is only compared with the state-of-the-art approach of Khan *et al.* [43] by implementing it on the same dataset of [43] which is accessible in Kaggle [44]. The results are presented in Table 7. Our deep learning-based *CataractNet* achieves higher performance (98.62%) than the state-of-the-art approach of [43] (with a reported accuracy of 97.47%) on the same small dataset of 1400 images without even any augmentations that indicates the capabilities of our proposed model. Additionally, our method was compared to the other pre-trained CNN models in Table 5 which proved its cost- and time-efficiency. Due to applying augmentation to the training dataset, our dataset size is more significant than most earlier works, significantly protecting the model from overfitting.

VI. CONCLUSION AND FUTURE WORK

In this paper, we presented an automated cataract detection system, namely *CataractNet*, based on lightweight deep learning. Initially, a cataract dataset of fundus images was rearranged, pre-processed, and augmented to improve the dataset to feed the deep network. The developed *CataractNet* focused on investigating different layers, activation

function, loss function, and optimization algorithms for minimizing the computational cost without sacrificing the model accuracy. Comparing with five pre-trained CNN models, i.e., MobileNet, VGG-16, VGG-19, Inception-v3, and ResNet-50, the *CataractNet* achieved competitive performance. Our model outperformed the state-of-the-art cataract detection approaches in terms of accuracy (99.13%), precision (99.08%), recall (99.07%), specificity (99.17%), MCC (98.23%), and f1-score (99.07%). Being highly accurate, cost- and time-efficient enabled the ophthalmologists to detect cataract disease timely and more precisely using *CataractNet*. However, our method can not discriminate the three types of age-related cataracts (nuclear cataracts, cortical cataracts, and PSCs). Besides, it was only proposed for cataract detection and not for grading or finding its exact location, which can be helpful for ophthalmologists. These issues need further investigation in the future.

REFERENCES

- [1] J. Rana and S. M. Galib, "Cataract detection using smartphone," in *Proc. 3rd Int. Conf. Electr. Inf. Commun. Technol. (EICT)*, Dec. 2017, pp. 1–4.
- [2] D. Pascolini and S. Mariotti, "Global estimates of visual impairment: 2010," *Brit. J. Ophthalmol.*, vol. 96, no. 5, pp. 614–618, Dec. 2012.
- [3] D. Allen and A. Vasavada, "Cataract and surgery for cataract," *BMJ*, vol. 333, no. 7559, pp. 128–132, Jul. 2006.
- [4] J.-J. Yang, J. Li, R. Shen, Y. Zeng, J. He, J. Bi, Y. Li, Q. Zhang, L. Peng, and Q. Wang, "Exploiting ensemble learning for automatic cataract detection and grading," *Comput. Methods Programs Biomed.*, vol. 124, pp. 45–57, Feb. 2016.
- [5] S. R. Flaxman, R. R. A. Bourne, S. Resnikoff, P. Ackland, T. Braithwaite, M. V. Cicinelli, A. Das, J. B. Jonas, J. Keeffe, J. H. Kempen, J. Leasher, H. Limburg, K. Naidoo, K. Pesudovs, A. Silvester, G. A. Stevens, N. Tahhan, T. Y. Wong, and H. R. Taylor, "Global causes of blindness and distance vision impairment 1990–2020: A systematic review and meta-analysis," *Lancet Global Health*, vol. 5, no. 12, pp. e1221–e1234, 2017.
- [6] L. Cao, H. Li, Y. Zhang, L. Zhang, and L. Xu, "Hierarchical method for cataract grading based on retinal images using improved Haar wavelet," *Inf. Fusion*, vol. 53, pp. 196–208, Jan. 2020.
- [7] O. J. Afolabi, G. P. Mabuza-Hocquet, F. V. Nelwamondo, and B. S. Paul, "The use of U-Net lite and extreme gradient boost (XGB) for glaucoma detection," *IEEE Access*, vol. 9, pp. 47411–47424, 2021.
- [8] L. Qiao, Y. Zhu, and H. Zhou, "Diabetic retinopathy detection using prognosis of microaneurysm and early diagnosis system for non-proliferative diabetic retinopathy based on deep learning algorithms," *IEEE Access*, vol. 8, pp. 104292–104302, 2020.
- [9] S. Hu, X. Wang, H. Wu, X. Luan, P. Qi, Y. Lin, X. He, and W. He, "Unified diagnosis framework for automated nuclear cataract grading based on smartphone slit-lamp images," *IEEE Access*, vol. 8, pp. 174169–174178, 2020.
- [10] X. Gao, D. W. K. Wong, T.-T. Ng, C. Y. L. Cheung, C.-Y. Cheng, and T. Y. Wong, "Automatic grading of cortical and PSC cataracts using retroillumination lens images," in *Proc. Asian Conf. Comput. Vis.* Berlin, Germany: Springer, 2012, pp. 256–267.
- [11] C. P. Niya and T. V. Jayakumar, "Analysis of different automatic cataract detection and classification methods," in *Proc. IEEE Int. Advance Comput. Conf. (IACC)*, Jun. 2015, pp. 696–700.
- [12] B. Raju, N. S. D. Raju, J. Akkara, and A. Pathengay, "Do it yourself smartphone fundus camera—DIYretCAM," *Indian J. Ophthalmol.*, vol. 64, no. 9, p. 663, 2016.
- [13] W. Fan, R. Shen, Q. Zhang, J.-J. Yang, and J. Li, "Principal component analysis based cataract grading and classification," in *Proc. 17th Int. Conf. E-Health Netw., Appl. Services (HealthCom)*, Oct. 2015, pp. 459–462.
- [14] L. Xiong, H. Li, and L. Xu, "An approach to evaluate blurriness in retinal images with vitreous opacity for cataract diagnosis," *J. Healthcare Eng.*, vol. 2017, pp. 1–16, Apr. 2017.
- [15] A. Budai, R. Bock, A. Maier, J. Hornegger, and G. Michelson, "Robust vessel segmentation in fundus images," *Int. J. Biomed. Imag.*, vol. 2013, pp. 1–11, Dec. 2013.
- [16] C. Hernandez-Matas, X. Zabulis, A. Triantafyllou, P. Anyfanti, S. Douma, and A. A. Argyros, "FIRE: Fundus image registration dataset," *Model. Artif. Intell. Ophthalmol.*, vol. 1, no. 4, pp. 16–28, 2017.
- [17] Z. Zhang, F. S. Yin, J. Liu, W. K. Wong, N. M. Tan, B. H. Lee, J. Cheng, and T. Y. Wong, "ORIGA-light: An online retinal fundus image database for glaucoma analysis and research," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol.*, Aug. 2010, pp. 3065–3068.
- [18] P. Porwal, S. Pachade, R. Kamble, M. Kokare, G. Deshmukh, V. Sahasrabudhe, and F. Meriaudeau, "Indian diabetic retinopathy image dataset (IDRiD): A database for diabetic retinopathy screening research," *Data*, vol. 3, no. 3, p. 25, Sep. 2018.
- [19] E. Decencière, G. Cazuguel, X. Zhang, G. Thibault, J.-C. Klein, F. Meyer, B. Marcotegui, G. Quellec, M. Lamard, R. Danno, D. Elie, P. Massin, Z. Viktor, A. Erginay, B. Laÿ, and A. Chabouis, "TeleOphta: Machine learning and image processing methods for teleophthalmology," *IRBM*, vol. 34, no. 2, pp. 196–203, Apr. 2013.
- [20] J. Staal, M. D. Abràmoff, M. Niemeijer, M. A. Viergever, and B. van Ginneken, "Ridge-based vessel segmentation in color images of the retina," *IEEE Trans. Med. Imag.*, vol. 23, no. 4, pp. 501–509, Apr. 2004.
- [21] H. Morales-Lopez, I. Cruz-Vega, and J. Rangel-Magdaleno, "Cataract detection and classification systems using computational intelligence: A survey," *Arch. Comput. Methods Eng.*, vol. 28, pp. 1–14, Jun. 2020.
- [22] H. E. Gali, R. Sella, and N. A. Afshari, "Cataract grading systems: A review of past and present," *Current Opinion Ophthalmol.*, vol. 30, no. 1, pp. 13–18, 2019.
- [23] I. Shaheen and A. Tariq, "Survey analysis of automatic detection and grading of cataract using different imaging modalities," in *Applications of Intelligent Technologies in Healthcare*. Cham, Switzerland: Springer, 2019, pp. 35–45.
- [24] X. Zhang, Y. Hu, J. Fang, Z. Xiao, R. Higashita, and J. Liu, "Machine learning for cataract classification and grading on ophthalmic imaging modalities: A survey," 2020, *arXiv:2012.04830*. [Online]. Available: <http://arxiv.org/abs/2012.04830>
- [25] X. Gao, H. Li, J. H. Lim, and T. Y. Wong, "Computer-aided cataract detection using enhanced texture features on retro-illumination lens images," in *Proc. 18th IEEE Int. Conf. Image Process.*, Sep. 2011, pp. 1565–1568.
- [26] M. Yang, J.-J. Yang, Q. Zhang, Y. Niu, and J. Li, "Classification of retinal image for automatic cataract detection," in *Proc. IEEE 15th Int. Conf. e-Health Netw., Appl. Services (Healthcom)*, Oct. 2013, pp. 674–679.
- [27] L. Guo, J.-J. Yang, L. Peng, J. Li, and Q. Liang, "A computer-aided healthcare system for cataract classification and grading based on fundus image analysis," *Comput. Ind.*, vol. 69, pp. 72–80, May 2015.
- [28] Y. N. Fuadah, A. W. Setiawan, T. L. R. Mengko, and Budiman, "Mobile cataract detection using optimal combination of statistical texture analysis," in *Proc. 4th Int. Conf. Instrum., Commun., Inf. Technol., Biomed. Eng. (ICICI-BME)*, Nov. 2015, pp. 232–236.
- [29] Y. N. Fuadah, A. W. Setiawan, and T. L. R. Mengko, "Performing high accuracy of the system for cataract detection using statistical texture analysis and K-nearest neighbor," in *Proc. Int. Seminar Intell. Technol. Appl. (ISITIA)*, May 2015, pp. 85–88.
- [30] M. Caixinha, J. Amaro, M. Santos, F. Perdigo, M. Gomes, and J. Santos, "In-vivo automatic nuclear cataract detection and classification in an animal model by ultrasounds," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 11, pp. 2326–2335, Nov. 2016.
- [31] V. Harini and V. Bhanumathi, "Automatic cataract classification system," in *Proc. Int. Conf. Commun. Signal Process. (ICCS)*, Apr. 2016, pp. 0815–0819.
- [32] A. B. Jagadale, S. S. Sonavane, and D. V. Jadav, "Computer aided system for early detection of nuclear cataract using circle Hough transform," in *Proc. 3rd Int. Conf. Trends Electron. Informat. (ICOEI)*, Apr. 2019, pp. 1009–1012.
- [33] R. Sigit, E. Triyana, and M. Rochmad, "Cataract detection using single layer perceptron based on smartphone," in *Proc. 3rd Int. Conf. Informat. Comput. Sci. (ICICoS)*, Oct. 2019, pp. 1–6.

- [34] E. Uchino, K. Suzuki, N. Sato, R. Kojima, Y. Tamada, S. Hiragi, H. Yokoi, N. Yugami, S. Minamiguchi, H. Haga, M. Yanagita, and Y. Okuno, "Classification of glomerular pathological findings using deep learning and nephrologist-AI collective intelligence approach," *Int. J. Med. Informat.*, vol. 141, Sep. 2020, Art. no. 104231.
- [35] M. S. Junayed, A. A. Jeny, S. T. Atik, N. Neehal, A. Karim, S. Azam, and B. Shanmugam, "AcneNet—A deep CNN based classification approach for acne classes," in *Proc. 12th Int. Conf. Inf. Commun. Technol. Syst. (ICTS)*, Jul. 2019, pp. 203–208.
- [36] X. Gao, S. Lin, and T. Y. Wong, "Automatic feature learning to grade nuclear cataracts based on deep learning," *IEEE Trans. Biomed. Eng.*, vol. 62, no. 11, pp. 2693–2701, Nov. 2015.
- [37] L. Zhang, J. Li, i. Zhang, H. Han, B. Liu, J. Yang, and Q. Wang, "Automatic cataract detection and grading using deep convolutional neural network," in *Proc. IEEE 14th Int. Conf. Netw., Sens. Control (ICNSC)*, May 2017, pp. 60–65.
- [38] J. Ran, K. Niu, Z. He, H. Zhang, and H. Song, "Cataract detection and grading based on combination of deep convolutional neural network and random forests," in *Proc. Int. Conf. Netw. Infrastruct. Digit. Content (IC-NIDC)*, Aug. 2018, pp. 155–159.
- [39] T. Pratap and P. Kokil, "Computer-aided diagnosis of cataract using deep transfer learning," *Biomed. Signal Process. Control*, vol. 53, Aug. 2019, Art. no. 101533.
- [40] D. Kim, T. J. Jun, Y. Eom, C. Kim, and D. Kim, "Tournament based ranking CNN for the cataract grading," in *Proc. 41st Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Jul. 2019, pp. 1630–1636.
- [41] M. R. Hossain, S. Afroze, N. Siddique, and M. M. Hoque, "Automatic detection of eye cataract using deep convolution neural networks (DCNNs)," in *Proc. IEEE Region Symp. (TENSYP)*, Jun. 2020, pp. 1333–1338.
- [42] X. Zhang, J. Lv, H. Zheng, and Y. Sang, "Attention-based multi-model ensemble for automatic cataract detection in B-scan eye ultrasound images," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2020, pp. 1–10.
- [43] M. S. M. Khan, M. Ahmed, R. Z. Rasel, and M. M. Khan, "Cataract detection using convolutional neural network with VGG-19 model," in *Proc. IEEE World AI IoT Congr. (AIoT)*, May 2021, pp. 0209–0212.
- [44] Kaggle. (2021). *Ocular Disease Recognition*. Accessed: Feb. 11, 2021. [Online]. Available: <https://www.kaggle.com/andrewmvd/ocular-disease-recognitionodir5k>
- [45] T. Pratap and P. Kokil, "Efficient network selection for computer-aided cataract diagnosis under noisy environment," *Comput. Methods Programs Biomed.*, vol. 200, Mar. 2021, Art. no. 105927.
- [46] M. S. Junayed, A. N. M. Sakib, N. Anjum, M. B. Islam, and A. A. Jeny, "EczemaNet: A deep CNN-based eczema diseases classification," in *Proc. IEEE 4th Int. Conf. Image Process., Appl. Syst. (IPAS)*, Dec. 2020, pp. 174–179.
- [47] J.-Y. Hung, C. Perera, K.-W. Chen, D. Myung, H.-K. Chiu, C.-S. Fuh, C.-R. Hsu, S.-L. Liao, and A. L. Kossler, "A deep learning approach to identify blepharoptosis by convolutional neural networks," *Int. J. Med. Informat.*, vol. 148, Apr. 2021, Art. no. 104402.
- [48] C. R. Gonzalez, "Richard Eugene woods," *Digit. Image Process.*, vol. 85, no. 3, 2007.
- [49] M. Heidari, S. Mirniaharikandehi, A. Z. Khuzani, G. Danala, Y. Qiu, and B. Zheng, "Improving the performance of CNN to predict the likelihood of COVID-19 using chest X-ray images with preprocessing algorithms," *Int. J. Med. Informat.*, vol. 144, Dec. 2020, Art. no. 104284.
- [50] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the inception architecture for computer vision," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2818–2826.
- [51] A. G. Howard, M. Zhu, B. Chen, D. Kalenichenko, W. Wang, T. Weyand, M. Andreetto, and H. Adam, "MobileNets: Efficient convolutional neural networks for mobile vision applications," 2017, *arXiv:1704.04861*. [Online]. Available: <http://arxiv.org/abs/1704.04861>
- [52] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, *arXiv:1409.1556*. [Online]. Available: <http://arxiv.org/abs/1409.1556>
- [53] M. Simon, E. Rodner, and J. Denzler, "ImageNet pre-trained models with batch normalization," 2016, *arXiv:1612.01452*. [Online]. Available: <http://arxiv.org/abs/1612.01452>
- [54] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.
- [55] A. A. Jeny, A. N. M. Sakib, M. S. Junayed, K. A. Lima, I. Ahmed, and M. B. Islam, "SkNet: A convolutional neural networks based classification approach for skin cancer classes," in *Proc. 23rd Int. Conf. Comput. Inf. Technol. (ICCIIT)*, Dec. 2020, pp. 1–6.



MASUM SHAH JUNAYED received the bachelor's degree in computer science and engineering from Daffodil International University, in 2019. He is currently pursuing the master's degree with the Department of Computer Engineering, Bahcesehir University (BAU), Istanbul, Turkey. He is also a Graduate Research Assistant with the Computer Vision Laboratory under the TUBITAK 2232 Project. Previously, he worked as a Research Assistant with the Gradient Lab for AI Research (GLAIR). He is also a member of Machine Intelligence Research Labs (MIR Labs), and the Founder and the Director of the Vision Research Lab of AI (VRLAI). He has published several peer-reviewed conferences and journal papers. His research interests include computer vision, machine learning, image and video processing, and medical image analysis.



MD BAHARUL ISLAM (Senior Member, IEEE) received the B.Sc. degree in computer science and engineering from RUET, Bangladesh, the M.Sc. degree in digital media from Nanyang Technological University, Singapore, and the Ph.D. degree in computer science from Multimedia University, Malaysia. Before becoming a Faculty Member, he was a Postdoctoral Research Fellow with the AI and Augmented Vision Laboratory, Miller School of Medicine, University of Miami, USA. He has more than 15 years of working experience in teaching and cutting-edge research in image processing and computer vision. He has authored/coauthored more than 40 international peer-reviewed research papers, including journal articles, conference proceedings, books, and book chapters. His current research interests include 3D processing and AR/VR-based vision rehabilitation. He secured several gold medals and best paper awards from national and international scientific and technological competitions and conferences. His Ph.D. thesis was selected for the Best Research Work by the IEEE SPS Research Excellence Award in 2018. He was awarded the International Fellowship and Grant for Outstanding Young Researchers from the Scientific and Technological Research Council of Turkey (TUBITAK) in 2019.



AREZOO SADEGHZADEH received the B.Sc. degree in electrical engineering from Azarbaijan Shahid Madani University, Tabriz, Iran, in 2014, and the M.S. degree in electrical engineering (telecommunication) from Sahand University of Technology, Tabriz, in 2016. She is currently pursuing the Ph.D. degree with the Computer Engineering Department, Bahcesehir University (BAU), Istanbul, Turkey. Previously, she worked as a Lecturer with the Computer Engineering Department, University College of Nabi Akram, Tabriz. She is also working on sign language recognition and translation and visual field map assessment at the Computer Vision Laboratory, BAU. Her research interests include artificial intelligence, image and video processing, computer vision, and machine learning.



SAIMUNUR RAHMAN received the M.Sc. degree (by research) from the Visual Processing Laboratory, Multimedia University, Cyberjaya, in 2017. He is currently pursuing the Ph.D. degree with the University of Wollongong, Australia, and Data61/CSIRO. He previously worked as a Lecturer with International Islamic University Chittagong and a Machine Learning Researcher with Vitrox Corporation Berhad, Malaysia. His current research interests include computer vision, medical image analysis, and machine learning. He has served on the program committees of various international conferences.