Eye-Vision Net: Cataract Detection and Classification in Retinal and Slit Lamp Images using Deep Network

Binju Saju[1]*, Rajesh R[2]
Research Scholar[1], Associate Professor[2]
Department of Computer Science, Christ (Deemed to be University), Bengaluru, Karnataka 560029, India

Abstract*—In the modern world, cataracts are the predominant cause of blindness. Early treatment and detection can reduce the number of cataract patients and prevent surgery. However, cataract grade classification is necessary to control risk and avoid blindness. Previously, various studies focused on developing a system to detect cataract type and grade. However, the existing works on cataract detection does not provide optimal results because of high detection error, lack of learning ability, computational complexity issues, etc. Therefore, the proposed work aims to develop an effective deep learning techniques for detecting and classifying cataracts from the given input samples. Here, the cataract detection and classification are performed using two phases. In order to provide an accurate cataract detection, the proposed study introduced Deep Optimized Convolutional Recurrent Network_Improved Aquila Optimization (Deep OCRN_IAO) model in phase I. Here, both retinal and slit lamp images are utilized for cataract detection. Then, the performance of these two image datasets are analysed, and the best one is chosen for cataract type and grade classification. By analysing the performance, the slit lamp images attain higher results. Therefore, phase II uses slit lamp images and detects the type and grade of cataracts through the proposed Batch Equivalence ResNet-101 (BE_ResNet101) model. The proposed classification model is highly efficient to classify the type and grades of cataracts. The experimental setup is done using MATLAB software, and the datasets used for simulation purposes are DRIMDB (Diabetic Retinopathy Images Database) and real-time slit lamp images. The proposed type and grade detection model has an accuracy of 98.87%, specificity of 99.66%, the sensitivity of 98.28%, Youden index of 95.04%, Kappa of 97.83%, and F1-score is 95.68%. The obtained results and comparative analysis proves that the proposed model is highly suitable for cataract detection and classification.

Keywords*—Cataract detection; grade classification; CRNN; dense CNN; Aquila optimization; BE-ResNet101

I. INTRODUCTION

The eyes are the major organs that offer a clear vision to humans and living organisms. A cataract is a severe eye disorder that affects the original vision of the eyes and produces vision distortion. A cataract is the major cause of vision impairment or blindness worldwide [1-2]. Vision or Visual Impairment (VI) in humans refers to the inability to perceive clear vision. A global survey by the WHO (World Health Organization) estimated that among 285 million people, 39 million are blind, and the remaining are reported to other VIs [3]. Some of the retinal disorders are Cataract, DR (Diabetic Retinopathy), Glaucoma, AMD (Aged Macular Degeneration), RP (Retinitis Pigmentosa), [4-5] etc. Cataracts are the leading cause of vision loss in aged and younger individuals [6-7]. The risk factors related to cataracts are people over 40 years, uncontrolled diabetes, steroid usage, family history, trauma in the eye, UV (Ultra-Violet) light exposure, [8] etc.

Early cataract diagnosis can control global vision loss. Based on the location and development, the cataract is divided into different types, namely NC (nuclear cataract), CC (cortical cataract), and PSC (posterior subcapsular cataract) [9]. The severity of the cataract can be identified with the grading procedure. Eye diseases are often diagnosed using several methods such as slit-lamp images [10], VA (visual acuity) examination [11], digital photography [12], retinal images [13] and ultrasonic images [14-15]. Ophthalmologist utilizes retinal and Slit-lamp images to detect the presence and absence of cataracts [16-17]. In ophthalmology, the screening and detection of cataracts have produced robust outcomes with deep learning (DL) models [18-20]. DL offers significant advantages by using several HL (Hidden Layer) sequences to extract useful image features. Moreover, in the medical field, DL learns the image features effectively and automatically. The advent of computer infrastructure has contributed to the faster adoption of DL due to its high level of feature extraction, processing of huge amounts of data and accurate classification.

Motivation: The eyes are the sensory organs that view the elegance of our surrounding environment. In recent years, the research organization are focussed on developing a robust cataract detection and classification methodology to assist affected people. From the recent studies, it is observed that deep learning techniques are attaining great attention for cataract detection. By using different types of datasets, the deep learning approaches gain better results in detecting and classifying cataract diseases. Many existing studies exhibit that using slit-lamp images produces high performance because of reduced cost, easy of maintain, flexibility etc. The neural network based feature extraction techniques can detect the features automatically. Also, such kind of feature extraction methods directly enhances the classification accuracy. Thus, it motivates the author to establish a deep learning based method for cataract detection and grade classification using slit lamp images. The major contribution of proposed work is given as,

- To develop an effective cataract detection and classification in retinal and slit-lamp images by

*Corresponding Author

proposing a new Deep Optimized Convolutional Recurrent Network (Deep OCRN).

- To reduce the detection error rate, the presented loss in the proposed Deep OCRN model is minimized by updating the weight parameters through Improved Aquila Optimization (IAO) algorithm.
- To extract the required features and to effectively identify the cataract type by proposing Dense CNN.
- To develop Batch Equivalence ResNet-101 method for classifying the grades of cataracts.
- To evaluate the performance of the proposed DL model by comparing it with other state-of-art classifiers.

The rest of this paper is structured as follows: Section II discussed recent existing studies on cataract detection, and Section III presents the proposed cataract detection classification model grade classification model. Section IV represents the results and discussion of the proposed study, and Section V deals with the overall conclusion and future scope for further research.

II. Related Works

Some recent research studies which are carried out on cataract detection and classification are listed as,

Pratap and Priyanka [21] presented a computer-aided robust cataract diagnosis using DNN (Deep Neural Network) in retinal fundus images. This diagnosis model involves CFE (Combined Feature Extraction) using dual fine-tuned independent DCNNs (Deep Convolutional Neural Networks). Moreover, multi-class SVM (Support Vector Machine) classifiers were trained on the noise levels from 0–25. The developed model obtained higher accuracy of (93.49%). However, the performance is affected due to noise problems.

Akil et al. [22] introduced the deep learning based CNN model for detecting retinal abnormalities in retinal fundus images. The ensemble learning based DR (Diabetic retinopathy) detection attained 97.7% accuracy and, with the grading achieved 98.5% accuracy. The main drawbacks were caused due to the limited number of training images and the lack of earlier screening.

Hu et al. [23] developed a unified automated NC (Nuclear Cataract) grading in slit-lamp smartphone images. The DL YOLOv3 model was used to localize the ocular lens nuclear region. Moreover, the combination of ShuffleNet and SVM was used in grading cataract severity by measuring the gray conjugate features of the core region. The overall performance obtained was accuracy (93.5%), F1 (92.3%), AUC (0.9198), and Kappa (95.4%). The drawbacks were high computational complexity and convergence loss.

Liu et al. [24] presented the diagnosis framework based on the localization of paediatric cataracts in slit-lamp images using deep feature modelling of CNN. The classification was completed with the combination of SVM and Softmax classifier. The overall classification accuracy was 97.07%. The Performance is not as much effective due to an inappropriate extraction of features.

Zhang et al. [25] developed automatic detection and grading of cataracts using DCNN (Deep Convolutional Neural Network). This DL framework DCNN was used to automatically examine cataracts by visualizing the feature maps with empirical higher-order semantic meaning at the pool5 layer. The overall accuracy performance obtained with cataract detection was (93.52%), and the grading of cataracts was (86.69%). The major limitation was classification and grading were analysed only on the fundus retinal images.

Problem statement: Earlier and accurate diagnosis of cataracts is more essential to limit the risk categories and prevent blindness. In existing, different techniques are developed to afford accurate cataract detection. However, they face several challenges while detecting cataracts from the given input samples. The detection accuracy of existing methods are reduced due to an enhanced computational complexity problem and large number of parameters in the detection models. By concerning this, the proposed study optimally selects the needed parameters using an effective algorithm. In addition, to make effective detection, extraction and selection of optimal features are more important. But, the existing detection techniques cannot extract the relevant features because of reduced efficiency. So, that the existing methods are not as suitable for detecting cataracts from the provided inputs. Also, many of the existing studies prefer retinal images for detecting cataracts and only few of them uses slit lamp images. To fulfil this gap, the proposed study used both retinal and slit lamp images for detecting cataracts. Moreover, robust DL models are utilized in the proposed study to learn appropriate features and make the system to attain higher detection accuracy with reduced complexity.

III. Proposed Methodology

The proposed detection and classification of cataract images involves two phases. In phase 1, the detection of cataracts using both retinal and slit-lamp images are performed. Initially, image acquisition is enabled for collecting retinal and slit-lamp images from the DRIMDB data source and real-time slit lamp images. Next, pre-processing is enabled to refine the quality of images, which involves image resizing, color conversion, removal of uneven illumination, quality selection and image enhancement. Here, score based image quality selection is performed using Hybrid NIQE-PIQE, and image enhancement is processed with Improved Mean Adjustment (IMA). To reduce the complexity and processing time, the proposed study extracts different features like shape, wavelet and texture. The relevant shape features are initially extracted through Pyramidal HOG (Histogram Oriented Gradient). Then, wavelet features are extracted using Haar Wavelet Transform (HWT) and the texture features are extracted using GLCM (Gray Level Co-occurrence Matrix). In order to reduce the high feature dimensionality, selecting an optimal feature set is essential. Thus, Relief Neighbourhood Component Analysis (RNCA) approach is proposed. Finally, the cataract detection is performed using Deep Optimized Convolutional Recurrent Network_Improved Aquila Optimization (Deep OCRN_IAO). Here, the loss in the network is optimized using Improved Aquila Optimization (IAO) and through this learning ability is increased. Moreover, the detection performance can be compared with other DL

models to validate the proposed performance. Fig. 1 represents the proposed methodology of the proposed work in both phases.

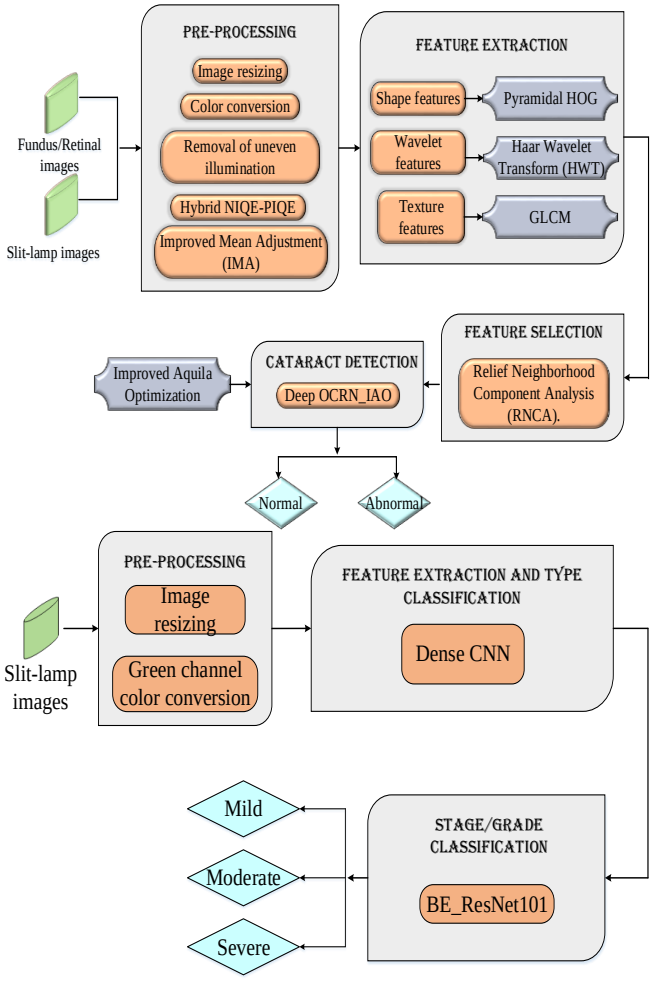


Fig. 1. Block diagram of proposed methodology

In phase 2, the slit-lamp cataract type and stage classification are done. Initially, image resizing and green channel color conversion are performed to enhance the image quality. Next, feature extraction and type classification can be performed using the Dense Convolutional Neural Network (Dense CNN) based deep learning framework. Finally, the classification of stages or grading as mild, moderate and severe can be attained using Batch Equivalence ResNet-101 (BE_ResNet101).

A. Image Acquisition

Image acquisition is the initial step for cataract detection, where various images are collected from different sources. The proposed work collects the retinal and slit-lamp images from popular datasets like DRIMDB and real-time slit lamp images. Fig. 2 shows some of the input images acquired from the dataset.

Fig. 2. Original input images

B. Pre-processing

Pre-processing an image is a significant step because it affords better classification results. The proposed pre-processing stage aims to refine the image quality to obtain effective performance. The images are pre-processed by,

- Image resizing
- Color conversion
- Uneven illumination removal
- Image enhancement
- Quality selection

In the proposed work, the high resolution images of the varied pixel are resized into 224*224 pixels for easy processing. The proposed work used green channel conversion in the pre-processing stage. Due to the presented uneven illumination and the available reflection of the eyes, the accurate classification of the different grades of cataracts is more difficult. Thus, the uneven illumination is removed in the pre-processing stage to enhance the performance of grade classification.

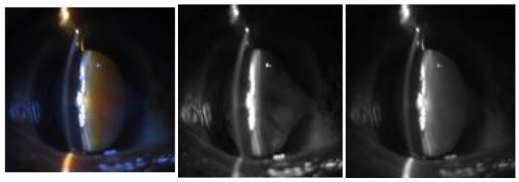
Image enhancement using Improved Mean Adjustment (IMA): Image quality enhancement plays a vital role in each algorithm. The poor quality of images lead to performance degradation of a specific classifier. Therefore, it is necessary to enhance the image quality during the phase of training and testing. This method is performed to elevate the image contrast. Here, the original input image I^O is utilized to generate the enhanced I^E image. The original image is split into 3×3 blocks of non-overlapping, and for each block mean is calculated. The highest mean value is mentioned as $\mu_{\max}$ and the lowest mean is indicated as $\mu_{\min}$. When the variations among $\mu_{\max}$ and $\mu_{\min}$ is small, then it is represented that the image contains a poor contrast. Hence, it can be enhanced by,

$$I^E(k,l) = T\left(I^O(k,l)\right) \tag{1}$$

Where, $T(f) = \frac{f \times s}{\mu_{\max} - \mu_{\min}} - \frac{\mu_{\min} \times s}{\mu_{\max} - \mu_{\min}}$ (2)

Here, the gray scale value at $I^{O}(k,l)$ is mentioned as f and the maximum gray scale value in the given image is represented as S.

Quality selection using hybrid NIQE-PIQE: The module of image quality selection is performed to filter out the improved quality of images for cataract diagnosis. The proposed work selects the image quality through the NIQE-PIQE model [21]. In NIQE-PIQE, the score value is evaluated for each image in which a low score is considered the best quality of image and the high score of images is assumed to be the worst quality. By [26] and [27], the higher values of NIQE-PIQE for the best image quality are observed to be five and 50 correspondingly. NIQE-PIQE is evaluated on the presented dataset, and the quality selection is performed based on the threshold. Fig. 3 shows the pre-processed images in the proposed study.



(a) (b) (c)

Fig. 3. Pre-processed images (a) Resized image (b) Output image of color conversion (c) Enhanced image

C. Feature Extraction

When the pre-processed images are directly subjected to the classification stage, the system faces several problems in diagnosing cataract disease and also it degrades the performance. Thus, extracting suitable features is essential to attain optimal detection results. The features that extracted by proposed feature extraction stage is described as,

1) Extraction of shape feature through pyramidal HOG: Pyramidal HOG is utilized for the object retrieval process and is highly suitable for extracting shape features from the input samples. In existing several methods are used to extract the shape features but they are not as much effective. The proposed method separates a localized and tracked region into several cells at varying pyramid levels. Depending the direction of the image gradient, the pyramidal HOG is termed an image descriptor. The gradient direction $\theta(a,b)$ and gradient magnitude $m(a,b)$ is evaluated in each cell on every pixel. It is mentioned as,

$$m = \sqrt{g_a(a,b)^2 + g_b(a,b)^2} \quad (3)$$

$$\theta = \arctan\frac{g_a(a,b)}{g_b(a,b)} \quad (4)$$

Where, $g_a(a,b)$ and $g_b(a,b)$ are indicated by the image gradients with the directions a and b correspondingly. The histogram vector is computed based on the direction and size values in all the cells. A large feature vector is generated by concatenating the histogram vectors of cells together, and the last feature vector is the outcome of the pyramidal HOG approach.

2) Extraction of wavelet features through Haar Wavelet Transform (HWT): The Haar transform is one of the simplest methods in wavelet transforms. This HWT method cross multiplies a function over the Haar wavelet with several stretches and shifts. In various existing studies, the wavelet transform is used in the pre-processed retinal image to contrast with the background and blood vessels. Still, they face some troubles while extracting wavelet features. Thus, the proposed study uses a robust HWT to extract the wavelet features in which the retinal images are dissolved into three levels: reduced contrast vessels, the background, and the vessel mentioning the background. Then, it crosses over the sum and continues to present the upcoming scale. The HWT is expressed in the form of a matrix.

$$B = HIH^{T} \quad (5)$$

Where, I states the images of $M \times M$ matrix, H mentions the Haar transformation of $M \times M$ matrix and B denotes the transform of $M \times M$ matrix that involves the Haar basis functions. Thus, the wavelet features are extracted via HWT.

3) Text feature extraction through GLCM (gray Level co-occurrence Matrix): The texture features correspond to the image spatial organization of pixel values. The proposed model prefers the GLCM method to attain text-based features. Because GLCM is one of the powerful method to extract texture based features. The GLCM approach acquires the specific relationship among two pixels spaced by a certain distance in an image. The GLCM techniques contain several features, where the proposed work extracts the statistical features like Angular Second Moment (ASM), correlation, contrast, entropy (ENT_y), energy (ENE_y), homogeneity (HOM_y), inverse difference moment (IDM) and variance (VAR).

The proposed feature extraction methods extract numerous features, which has the chance to enhance the computational complexity. Thus, to minimize the computational complexity, selecting an optimal feature set is essential.

D. Relief Neighbourhood Component Analysis (RNCA) based Feature Selection

The feature selection process reduces the high data dimensionality by choosing the most effective features. In the proposed cataract detection system, RNCA is utilized for feature selection. This method can make optimal feature predictions through feature weights. The aim of this stage is to diminish the computational complexity, and also it helps to boost the classification accuracy. The proposed RNCA is one of the multi-level feature selection approaches which incorporates both ReliefF and NCA algorithms. The mentioned NCA [28] is a feature selection and dimensionality minimization method and is mainly utilized in varied classification studies. The NCA utilizes training data with varied class labels when selecting the projection that will better isolate classes in the prescribed region. The objective function used in the NCA method is expressed as,

$$\Gamma(w) = \frac{1}{m}\sum_{i=1}^{m} Q_i - \psi \sum_{r=1}^{q} w_r^2 \qquad (6)$$

Where, ψ represents the regularization parameter, feature weight is indicated as w_r, total amount of images are given as m, the probability score of the image i is mentioned as Q_i, q signifies dimensionality. When the parameter ψ is randomly selected, each of the feature weights can acquire values that are nearer to 0. The ReliefF approach is a method which can generate an accurate prediction of features. With the aid of feature weights, the features are estimated. Initially, the feature weights are set as zero then it randomly chooses the images from the input at every step. After that, the nearest k images from a similar class and the nearest from a varied class are identified. Then, weights are updated for each feature, and the feature that does not reach the particular condition is ignored from the input. The formulation of the ReliefF approach is represented as,

$$\omega(z^b) = \omega(z^b) - \frac{\sum_{j=1}^{k} diff(F, C_i, Y_j)}{nzk} + \sum_{d \neq class(C_i)} \left[\frac{Q(D)}{1 - Q(class(C_i))} z \sum_{j=1}^{k} \frac{diff(F, C_{i,N_j})}{nzk} \right] \qquad (7)$$

Where, z^b mentions the b^{th} feature, the feature set is represented as F, C_i and Y_j indicates the feature set instances and the user chosen parameter is signified as η and k. By this, the optimal set of features is selected using the proposed RNCA, and the selected reduced feature sets are subjected to the Deep OCRN_IAO model for cataract detection. By generating low dimensionality features, the process becomes smooth and reduces computational time.

E. Cataract Detection using Deep OCRN_IAO

This stage detects the cataract from the input images using a hybrid mechanism. Recently, recurrent neural network gained more attention for detecting diseases. In existing, several deep learning models are used to detect cataracts. But they cannot achieve higher detection accuracy due to the reduced learning capability. In order to obtain better detection results, the proposed work hybridizes the convolutional recurrent neural network (CRNN) with IAO algorithm. The major module of CRNN is the recurrent convolutional layer (RCL). The parameter setting of CRNN is given as, total number of epochs-10, maximum iteration-500, iteration per epoch-50 and learning rate-0.001. The RCL units develop against discrete time steps. For a unit placed at (i, j) on p^{th} feature map in RCL, its input is represented as,

$$x_{ijp}(t) = \left(\omega_p^h\right)^T v^{(i,j)}(t) + \left(\omega_p^\beta\right) z^{(t-1)} + b_p \qquad (8)$$

Where, $v^{(i,j)}(t)$ indicate the feed forward input and the recurrent input is mentioned as $z^{(t-1)}$, ω_p^h signifies as feed forward weights, ω_p^β mentions the recurrent weights and bias is represented as b_p. The initial term in equation (19) is utilized in conventional CNN, and the second term represents the recurrent connections. The activity of this unit is mentioned as the function of its input and is given as,

$$z_{ijp}(t) = g\left(r\left(x_{ijp}(t)\right)\right) \qquad (9)$$

Where, r specifies the rectified linear activation function. The function of local response normalization is represented as g.

$$r\left(x_{ijp}(t)\right) = \max\left(x_{ijp}(t), 0\right) \qquad (10)$$

$$g\left(r_{ijp}(t)\right) = \frac{r_{ijp}(t)}{\left(1 + \frac{\alpha}{M} \sum_{p'=\max(0,\, p-M/2)}^{\min(P,\, p+M/2)} \left(r_{ijp'}\right)^2\right)^\eta} \qquad (11)$$

Where, P mentions the overall feature maps in the present layer. It is observed that in equation (23), the sum process over M feature maps at a similar region (i, j), the constant that maintains the normalization's amplitude is mentioned as α and η. The exploding state is avoided by using local response normalization. The dynamic characteristics of RCL are described in equations (20) and (21). The CRNN holds a stack of RCLs with max pooling layers. The initial layer is the feed-forward convolutional layer without the connections of recurrent layers. This convolutional layer generates an activation map and transforms all the pixel values into a single value. The second layer is the max pooling layer in which the feature dimensionality is reduced and minimizes the number of parameters to learn. The third is stacked with recurrent layers, and the next layer is fully connected. In the fully connected layer, each input from one layer is linked to each activation function of the upcoming layer. Finally, the cataract is detected in the softmax layer, which is given as,

$$o_p = \frac{\exp\left(\omega_p^T z\right)}{\sum_{p'} \exp\left(\omega_{p'}^T z\right)} \qquad (p = 1, 2, \ldots\ldots, D) \qquad (12)$$

Where, o_p represents the predicted probability from the p^{th} category and feature vector created by max pooling layer is represented as z. However, the available loss function influences the detection performance of CRNN. Thus, the loss function is minimized by updating the weight parameters using the IAO approach. The available loss function is described as,

$$L_F = \frac{1}{T}\sum_{i=1}^{T} \left(a_i - k_i^{'}\right)^2 \qquad (13)$$

Where, T represents the total number of iterations, the actual value is signified as a_i and $k_i^{'}$ mentions the predicted value. Cataracts in the retinal images are detected based on the objective function and are expressed as,

$$Objective\ function = Min[L_F] \quad (14)$$

The proposed IAO method aims to minimize the loss function through weight parameter updation. This is one of the population based approaches and is motivated by the hunting behaviour of Aquila. In order to update the weight parameters, Aquila algorithm used the below formulation.

$$L_1(h+1) = L_{best}(h) \times \gamma(P) + L_r(h) + (x-y) \times rand \quad (15)$$

Where, $L_1(h+1)$ represents the upcoming iteration of the solution h, P mentions the position of the feature, levy flight distribution function is denoted as $\gamma(P)$ and $L_r(h)$ represents the random solution. To enhance the efficiency of Aquila algorithm, weight strategy is introduced.

Weight strategy: According to the maximum iteration, the feature's distribution scope is compressed, generating the search space as small. The weight factor is applied to enhance the different characteristics of features and recover the optimization algorithm from the local optima. This is represented as,

$$\varpi_j = \exp(n\ iter_j / n\ iter_{\max} - 1) \quad (16)$$

Where, ϖ_j denotes the weight at j^{th} iteration, the maximum iteration number is specified as $n\ iter_{\max}$ and the current iteration number is signified as $n\ iter_j$ where, $n\ iter_j \le n\ iter_{\max}$. In this, the objective function is evaluated for each iteration, and the attained result from every iteration is compared. By analysing the best objective function, the optimal solution is obtained. Therefore, the proposed Deep OCRN_IAO model identifies the cataract from the given dataset image. However, determining the grade and severity of the cataract is highly important to cure the disease. Thus, the proposed work used the cataract type classification and grade categorization in phase 2 using effective deep learning techniques.

F. Cataract Type and Grade Classification using Slit Lamp Images through Dense CNN+BE_ResNet101 Model

The automatic detection of grading and type is challenging for several existing studies. Thus, the proposed work focussed on developing an automatic cataract grade and type classification with the support of deep learning techniques. Here, the Dense CNN is developed to detect the type of cataracts and BE-ResNet101 is performed to classify the cataract grades. The hyperparameters of the proposed type and grade detection model is described as, maximum number of epochs-10, maximum iteration-500, iteration per epoch-50 and learning rate-0.001.

1) Pre-processing of slit lamp images: Initially, the slip lamp images are pre-processed to enhance the classification performance. The pre-processing steps that carried out on slit lamp images are,

- Resizing of slit lamp images
- Green channel color conversion

After pre-processing the input images, the image quality is enhanced and fed to the Dense CNN method for feature extraction and cataract type categorization.

2) Feature extraction and detection of cataract type using dense CNN: The proposed work is done on feature extraction from the pre-processed images to minimize the classification error. For this purpose, Dense CNN is preferred in this study for feature extraction. Because CNN has the ability to extract features automatically, it helps to produce higher classification results. After feature extraction, the Dense CNN model's final layer classifies the cataract type from the provided slit lamp images. The proposed approach contains a set of dense layers that evaluate each input's weight average and is transmitted to the activation function. The input layer is the initial layer of CNN, where the images from the dataset are set as input. Next, the convolutional layer is the second layer designed by various convolutional kernels. The parameters in the convolutional layer are optimized via a back propagation approach. The major intention of this layer is to extract the most needed features from the slit lamp image, which helps to maintain the complexity. Next, a max-pooling layer is presented which generates feature mapping, and it reduces the feature dimensionality. The information from the pooling layer is passed into the set of dense layers. Here, the neurons of the layer are linked to each neuron of its previous layer. The CNN functionality mainly confides in the activation function, a systematic layer. It is noted that the rate of training can prevent over-fitting problems.

$$h(z)\ K \max(z;0) \quad (17)$$

The functionality of the convolutional layer is described as follows,

$$z_p^l = h\left(\sum_{i \in Np} z_i^{l-1} * K_{ip}^{l-1} + b_p^l\right) \quad (18)$$

Where, p mentions feature mapping, Np represents input mapping, K_{ip} is denoted as filter and bias of feature mapping is specified as b_p. The training process of the network is accelerated using the activation layer, and batch normalization is performed among the activation function and convolutional layer. The cross entropy is analysed in the output layer and is defined as,

$$Loss_{class} = -\sum_{i=1}^{m} 1\{f = i\} \log \frac{e^{xi}}{\sum_{l=1}^{f} e^{xi}} \quad (19)$$

Where, q is denoted as the number of classes, x_i is the final layer and f represents the information of each feature from the input image. During the process of CNN training, the feature extraction is performed, and also it diminishes the

cross entropy loss and reconstruction error. In the training process, each kernel is trained to acquire cataract types from the input source. The confidence value is enhanced and fine-tuned by minimizing the cross entropy loss function. Thus, the proposed Dense CNN classifies the types of cataracts precisely.

3) Grade or stage classification of cataracts using BE_ResNet101 framework: The proposed study used the BE-ResNet101 approach for cataract grade categorization. The fundamental ResNet is also a type of CNN model, and the structure of ResNet was inspired by the traditional VGG-19 method. The proposed BE-ResNet101 detects the grade by avoiding the issues of attaining information from unbalanced instances. The presented batch equivalence in ResNet101 helps to improve the cataract grade detection results. Also, the proposed model helps to reduce the degradation issue. The basicResNet-101 is a residual network that contains 101 layers and is the advanced version of the ResNet-50 model. The layers available in the ResNet contain a similar amount of filters for a similar size of output feature map. The amount of filter is twice if the feature map's size is minimized to reduce the time complexity. In the ResNet model, down sampling is directly performed by convolving layers with two strides. This ResNet structure ends with a pooling layer and a fully connected layer with a softmax layer. The objective of distributive layers is to avoid the vanishing gradients issues through reutilizing activation from the previous layer till the closest layer of the current layer has learned its weights.

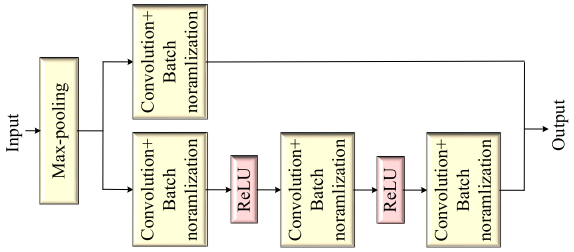


Fig. 4. ResNet-101 structure

The ResNet-101 involves a total of 104 convolutional layers. It contains layers of 33 blocks, and 29 of such blocks utilize the output of the preceding block. It is shown above as remainder connections, and such remainders are used as the initial operand of the addition operator used at the end of all blocks to get the input of the coming blocks. The rested four blocks acquire the output of the preceding block and utilize it in a convolutional layer with 1×1 of filter size and one stride, followed by a batch normalization layer. This layer enables the normalization process, and the attained outcome is transmitted to the addition operator at the output of the block. The structure of ResNet is illustrated in Fig. 4.

Batch equivalence: In the proposed experimental data, the number of images graded from 3.0 to 4.0 is considered the highest. When a mini-batch is designed randomly, the grading model obtains the information from unbalanced occurrences and becomes biased. This issue is solved by applying a strategy of batch balancing. Here, the batch is separated in to five varied groups like [1, 2], [2, 3], [3, 4], [4, 5] and [5, 6]. A similar number of occurrences are randomly chosen during the construction of a mini-batch, i.e., 25 from all the groups and generate the batch fully balanced. Thus, the proposed BE_ResNet101 model classifies and detects the grade of cataracts effectively. Fig. 5 illustrates the cataract grade classified images from the proposed Dense CNN+BE_ResNet101 model.

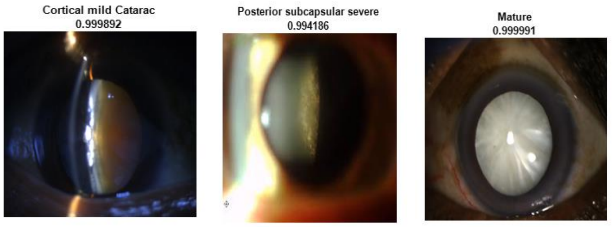


Fig. 5. Classified images

IV. Results and Discussions

This section presents the simulation results of the proposed cataract detection and classification system with clear analysis. The experimental setup was executed through MATLAB software.

A. Dataset Description

The input images utilized in this work are chosen from two different datasets like DRIMDB (Diabetic Retinopathy Images Database)https://academictorrents.com/details/99811ba62918f8e73791d21be29dcc372d660305b and real time slit lamp images. The DRIMDB dataset is utilized in several retinal image analysis studies. In this dataset, a total of 216 consecutive images are taken in the proposed experimental analysis. The real-time slit lamp images are collected from various hospitals containing five types of cataracts: cortical, hyper mature, mature, nuclear and posterior.

B. Performance Metrics

The performance metrics are utilized to evaluate the efficacy of the proposed model. The metrics utilized in the proposed study are expressed below.

$$Accuracy:\ A_y = \frac{Tp+Tn}{Tp+Tn+Fp+Fn} \quad (20)$$

$$Sensitivity:\ S_y = \frac{Tp}{Tp+Fn} \quad (21)$$

$$Specificity:\ Sp_y = \frac{Tn}{Tn+Fp} \quad (22)$$

$$Youden\ Index = S_y + Sp_y - 1 = TPR - FPR \quad (23)$$

$$F_1 - score = 2\times\frac{\Pr ecision\times Sensitivity}{\Pr ecision + Sensitivity} \quad (24)$$

$$Kappa = \frac{P_{observed} - P_{chance}}{1 - P_{chance}} \quad (25)$$

C. Performance Analysis of Proposed Classification

The performance of the proposed BE_ResNet101 model is analysed by measuring the performance metrics and comparing the attained results with some other approaches. The confusion matrix of the proposed classification is shown in Fig. 6.

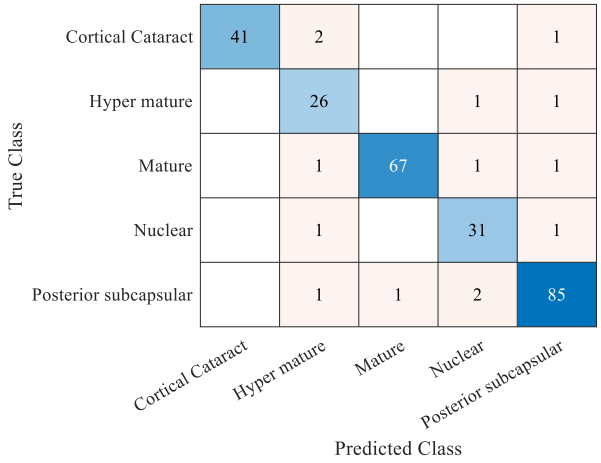


Fig. 6. Confusion matrix of proposed Dense CNN+BE_ResNet101 classification model

The above confusion matrix reveals the efficiency of the proposed model also, it mentions the prediction ability. At each testing phase, different cataract types are detected. During classification, only few classes are misclassified as others. Thus, the confusion matrix exhibits that the proposed model provides accurate classification. Fig. 7 shows the obtained accuracy and loss during training and testing.

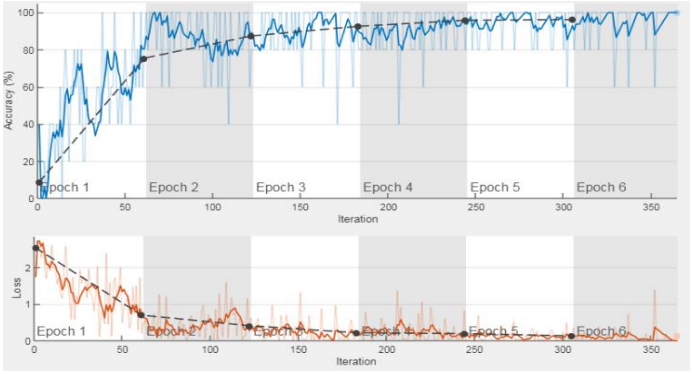


Fig. 7. Accuracy and loss graph of proposed Dense CNN+BB_ResNet101model

The accuracy and loss are measured at each iteration by varying the number of epochs. By analysing the results, it is clear that the accuracy and loss is well maintained in the proposed model. This proves that the proposed classifier is powerful for cataract type and grade classification. The results obtained from phase I by using retinal and slit lamp images are shown in Table I.

TABLE I. PERFORMANCE ANALYSIS IN PHASE I USING RETINAL AND SLIT LAMP IMAGES

Datasets	*Accuracy (%)*	*Sensitivity (%)*	*Specificity (%)*	*Precision (%)*	*F1 score (%)*	*Youden Index (%)*	*Kappa (%)*
(DRIMDB) Retinal images	92.36	94.58	96.57	91.06	91.47	89.65	89.32
Slit lamp images	94.67	96.68	98.29	93.03	93.81	91.23	90.28

In order to determine the strength of datasets, the performance metrics are evaluated, and the obtained results reveals that the slit lamp images obtain a better outcome than retinal images from the DRIMDB dataset. Thus, slit lamp images are preferred in phase II to detect the type and grade of the cataract. The accuracy measure comparison is illustrated in Fig. 8.

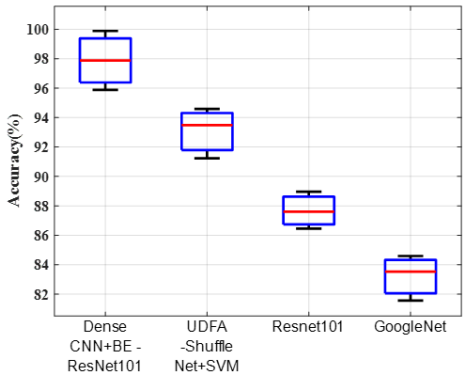


Fig. 8. Accuracy performance comparison

The above analysis clearly shows that, the accuracy of the proposed classifier is enhanced because of the higher efficiency. Besides that, the proposed BE-ResNet101 model diminishes the training error and also it avoids the vanishing gradient issue. But, the existing approaches generates higher classification errors and leads to low accuracy. The accuracy of proposed model is 98.87%, ResNet-101 is 87.6%, GoogleNet is 83.53% and UDFA is 93.48%. The comparative analysis of proposed and existing works in terms of sensitivity is shown in Fig. 9.

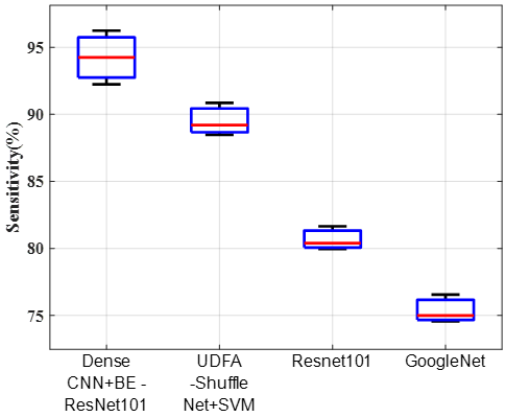


Fig. 9. Sensitivity performance comparison

The above analysis shows that the proposed Dense CNN +BE_ResNet101 model obtains enhanced sensitivity to existing approaches. The average value of sensitivity attained from the proposed model is98.28%, ResNet101 is 80.39%, GoogleNet is 75%, and UDFA is 89.2%. The specificity comparison analysis is illustrated in Fig. 10.

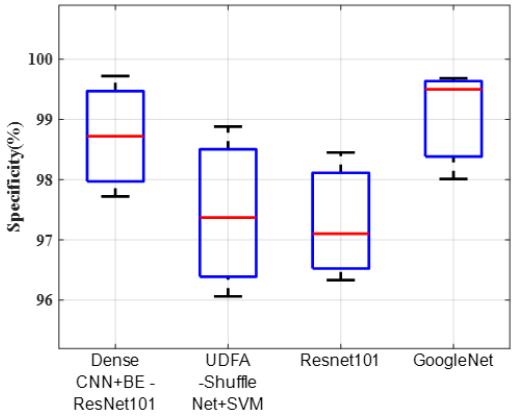


Fig. 10. Specificity performance comparison

The existing approaches face several issues when detecting cataracts from the given image. Thus, the previous studies are unable to provide improved specificity results. The average specificity value attained from the proposed Dense CNN+BE_ResNet101 model is 99.66%, ResNet 101 is 97.1%, UDFA is 97.37% and GoogleNet is 99.5%. The F1-score comparison is depicted in Fig. 11.

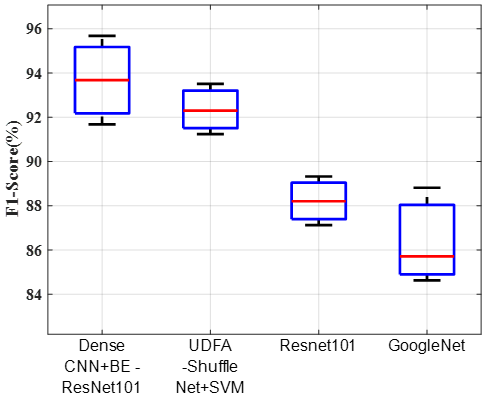


Fig. 11. F1 score performance comparison

The average F1-score of proposed classifier is 95.68%, UDFA is 92.3%, and ResNet101 is 88.2% and GoogleNet 85.71%. Compared with other approaches, the F1score of the proposed classifier is superior. This states that the developed Dense CNN+BE ResNet101 is more suitable for detecting cataract type and grade. The comparative analysis of the kappa coefficient is shown in Fig. 12.

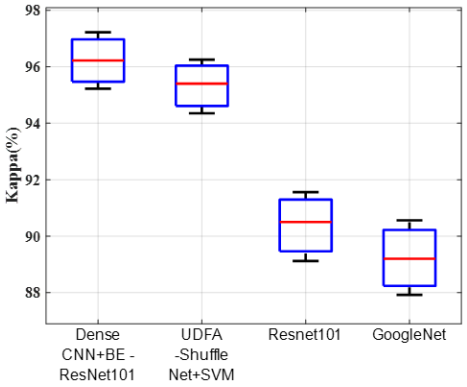


Fig. 12. Kappa performance comparison

The comparison analysis using the kappa measure shows that the proposed techniques are more potent than the others. The average kappa value obtains from the proposed Dense CNN+BE-ResNet101 is 97.83%, UDFA is 95.4%, ResNet101 is 90.5% and GoogleNet is 89.2%. The Youden index comparison of proposed and existing techniques is shown in Fig. 13.

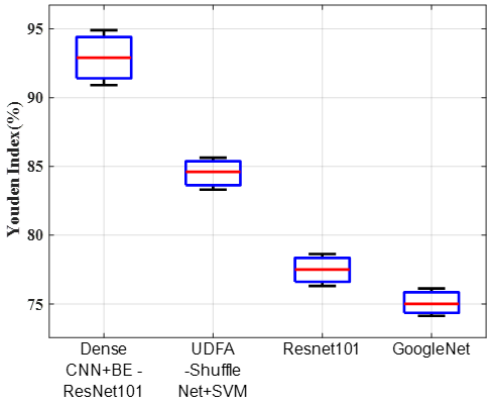


Fig. 13. Performance comparison in terms of Youden index

The average Youden index attained from the proposed model is 95.04, UDFA is 84.6%, ResNet101 is 77.5%, and GoogleNet is 75%. Compared with others, the proposed method is the highly enlarged range in the Youden index. The comparison of the ROC curve and AUC values of different techniques is illustrated in Fig. 14.

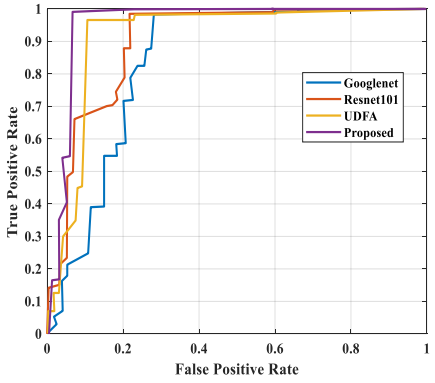


Fig. 14. ROC curve and AUC values of proposed and existing approaches

The ROC curve comparison represents that the proposed model is better than the other techniques. Analysing the ROC curve, the obtained AUC values are superior in the proposed cataract grade detection model. By analysing the proposed work with different approaches, the proposed model obtains improved outcomes in terms of accuracy, sensitivity, specificity, Youden index, Kappa and F1 score.

V. CONCLUSION

This research presents cataract detection and classification using deep learning approaches through retinal and slit lamp images. Here, the proposed work is carried out in two phases. In the first phase, the cataract is identified, and in the second phase, the type and grade of the cataract is categorized. In phase I, the images are gathered from DRIMDB and slit lamp images. By using these two varied dataset images, the cataract is detected through proposed Deep OCRN_IAO model. The experimental results prove that the slit lamp images are obtained improved performance than the retinal images.

Thus, in phase II, the slit lamp images are preferred for cataract type and grade classification. In the proposed work, the type of cataract is detected using Dense CNN model, where the convolutional layer of CNN reduces the feature dimensionality. Depending on the outcome of convolutional layer, the dense layer categorizes the features. Because of the presence of a dense layer, the CNN network accurately detects the type of cataracts. Also, the proposed BE-ResNet101 model detects the grade of cataracts by reducing the degradation issue. This model classifies the grades as mild, severe and moderate. Therefore, the system obtains enhanced results than the existing ones. The experimental results show that the proposed Dense CNN+BE-ResNet101 model attains good performance results in an accuracy of 98.87%, sensitivity 98.28%, specificity 99.66%, kappa coefficient 97.83%, F1 score 95.68% and Youden index 95.04%. In future studies, the proposed work will be applied to different biomedical images and detect their performance and effectiveness through various datasets.

REFERENCES

[1] D. Lam, S.K. Rao, V. Ratra, Y. Liu, P. Mitchell, J. King, M.J. Tassignon, J. Jonas, C.P. Pang, D.F. Chang, "Cataract," Nature reviews Disease primers, Vol. 1, no. 1, pp. 1-15, 2015.

[2] R.J. Olson, R. Braga-Mele, S.H. Chen, K.M. Miller, R. Pineda, J.P. Tweeten, D.C. Musch, "Cataract in the adult eye preferred practice pattern®." Ophthalmology, Vol. 124, no. 2, pp. P1-P119, 2017.

[3] S.R. Flaxman, R.R. Bourne, S. Resnikoff, P. Ackland, T. Braithwaite, M.V. Cicinelli, A. Das, J.B. Jonas, J. Keeffe, J.H. Kempen, J. Leasher, "Global causes of blindness and distance vision impairment 1990–2020: a systematic review and meta-analysis," The Lancet Global Health, Vol. 5, no. 12, pp. e1221-e1234, 2017.

[4] K. Landau, M. Levin, "Retinal disorders," Handbook of clinical neurology, pp. 97-116, Vol. 102, 2011.

[5] I.K. Muftuoglu, M. Al-Sheikh, M.A. Rasheed, S.R. Singh, J. Chhablani, Imaging "Imaging in inherited retinal disorders," European Journal of Ophthalmology, Vol. 31, no. 4, pp. 1656-1676, 2021.

[6] C. Karakosta, A. Tzamalis, M. Aivaliotis, I. Tsinopoulos, "Pathogenesis of Age-Related Cataract: A Systematic Review of Proteomic Studies," Current Proteomics, Vol. 18, No. 4, pp. 458-466, 2021.

[7] Ç. Öktem, F. Aslan, "Vitamin D Levels in Young Adult Cataract Patients: A Case-Control Study," Ophthalmic research, Vol. 64, no. 1, pp. 116-120, 2021.

[8] Q. Fan, X. Han, J. Luo, L. Cai, X. Qiu, Y. Lu, J. Yang, "Risk factors of intraocular lens dislocation following routine cataract surgery: a case–control study," Clinical and Experimental Optometry, Vol. 104, no. 4, pp. 510-517, 2021.

[9] R.E. Ruiz-Lozano, J.C. Hernandez-Camarena, M. Roman-Zamudio, R.J. Alcazar-Felix, O. Davila-Cavazos, J.A. Cardenas-de la Garza, "Three types of cataract associated with atopic dermatitis and chronic topical corticosteroid use: A case report," Dermatologic therapy, Vol. 34, no. 1, pp. e14600, 2021.

[10] S. Hu, H. Wu, X. Luan, Z. Wang, M. Adu, X. Wang, C. Yan, B. Li, K. Li, Y. Zou, X. Yu, "Portable handheld slit-lamp based on a smartphone camera for cataract screening," Journal of Ophthalmology, Vol. 2020, 2020.

[11] J. Chen, Y. Chen, Y. Zhong, J. Li, "Comparison of visual acuity and complications between primary IOL implantation and aphakia in patients with congenital cataract younger than 2 years: a meta-analysis," Journal of Cataract & Refractive Surgery, Vol. 46, no. 3, pp. 465-473, 2020.

[12] S. Kumar, S. Pathak, B. Kumar, "Automated detection of eye related diseases using digital image processing," In Handbook of multimedia information security: techniques and applications, pp. 513-544, 2019. Springer, Cham,

[13] Y. Luo, K. Chen, L. Liu, J. Liu, J. Mao, G. Ke, M. Sun, "Dehaze of cataractous retinal images using an unpaired generative adversarial network," IEEE Journal of Biomedical and Health Informatics, Vol. 24, no. 12, pp. 3374-3383, 2020.

[14] Y. Wang, C. Tang, J. Wang, Y. Sang, J. Lv, "Cataract detection based on ocular B-ultrasound images by collaborative monitoring deep learning." Knowledge-Based Systems, Vol. 231, pp. 107442, 2021.

[15] X. Zhang, J. Lv, H. Zheng, Y. Sang, "Attention-based multi-model ensemble for automatic cataract detection in b-scan eye ultrasound images," In 2020 international joint conference on neural networks (IJCNN), pp. 1-10, 2020. IEEE,

[16] A. Koyama, D. Miyazaki, Y. Nakagawa, Y. Ayatsuka, H. Miyake, F. Ehara, S.I. Sasaki, Y.Shimizu, Y. Inoue, "Determination of probability of causative pathogen in infectious keratitis using deep learning algorithm of slit-lamp images," Scientific reports, Vol. 11, no. 1, pp. 1-13, 2021.

[17] A. Imran, J. Li, Y. Pei, F. Akhtar, J.J. Yang, Q. Wang, "Cataract detection and grading with retinal images using SOM-RBF neural network," In 2019 IEEE Symposium Series on Computational Intelligence (SSCI), pp. 2626-2632, 2019. IEEE,

[18] C. P. Niya, T. V. Jayakumar, "Analysis of different automatic cataract detection and classification methods," In 2015 IEEE International Advance Computing Conference (IACC), pp. 696-700, 2015. IEEE,

[19] L. Guo, J.J. Yang, L. Peng, J. Li, Q. Liang, "A computer-aided healthcare system for cataract classification and grading based on fundus image analysis," Computers in Industry, Vol. 69, pp. 72-80, 2015.

[20] X. Gao, S. Lin, T.Y. Wong, "Automatic feature learning to grade nuclear cataracts based on deep learning," IEEE Transactions on Biomedical Engineering, Vol. 62, no. 11, pp. 2693-2701, 2015.

[21] T. Pratap, P. Kokil, "Deep neural network based robust computer-aided cataract diagnosis system using fundus retinal images," Biomedical Signal Processing and Control, Vol. 70, pp. 102985, 2021.

[22] M. Akil, Y. Elloumi, R. Kachouri, "Detection of retinal abnormalities in fundus image using CNN deep learning networks," In State of the Art in Neural Networks and their Applications, pp. 19-61, 2021. Academic Press,

[23] S. Hu, X. Wang, H. Wu, X. Luan, P. Qi, Y. Lin, X. He, W.He, "Unified diagnosis framework for automated nuclear cataract grading based on smartphone slit-lamp images," IEEE Access, Vol. 8, pp. 174169-174178, 2020.

[24] X. Liu, J. Jiang, K. Zhang, E. Long, J. Cui, M. Zhu, Y. An, "Localization and diagnosis framework for pediatric cataracts based on slit-lamp images using deep features of a convolutional neural network," PloS one, Vol. 12, no. 3, pp. e0168606, 2017.

[25] L. Zhang, J. Li, H. Han, B. Liu, J. Yang, Q. Wang, "Automatic cataract detection and grading using deep convolutional neural network," In

2017 IEEE 14th international conference on networking, sensing and control (ICNSC), pp. 60-65, 2017. IEEE,

[26] A. Mittal, R. Soundararajan, C. B. Alan, "Making a "completely blind" image quality analyser," IEEE Signal processing letters, Vol. 20, no. 3, pp. 209-212, 2012.

[27] N. Venkatanath, D. Praneeth, M.C. Bh, S.S. Channappayya, S.S. Medasani, "Blind image quality evaluation using perception based features," In 2015 Twenty First National Conference on Communications (NCC), pp. 1-6, 2015. IEEE,

[28] E. Başaran, Z. Cömert, Y. Çelik, "Neighbourhood component analysis and deep feature-based diagnosis model for middle ear otoscope images," Neural Computing and Applications, Vol. 34, no. 8, pp. 6027-6038, 2022.