International Journal of Engineering Trends and Technology
ISSN: 2231–5381 / https://doi.org/10.14445/22315381/IJETT-V72I5P113
Volume 72 Issue 5, 120-130, May 2024

Original Article

An Intelligent Ensembling of Fine-Tuned Transfer Learning-Based Model for Cataract Diagnosis from Fundus Images

Mahdi Falah Mahdi Alyami[1], Meenu Garg[2], Vatsala Anand[3], Sheifali Gupta[4], Mana Saleh Al Reshan[1*], Saleh Hamad Sajaan Almansour[1], Shoug Salem Hasan Alyami[1], Asadullah Shaikh[1]

[1]*Department of Information Systems, College of Computer Science and Information Systems, Najran University, Najran, Saudi Arabia.*
[2]*Department of Electronics and Communication Engineering, Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India.*
[3,4]*Chitkara University Institute of Engineering and Technology, Chitkara University, Punjab, India.*

[*]*Corresponding Author : msalreshan@nu.edu.sa*

Received: 26 December 2023 Revised: 01 March 2024 Accepted: 04 May 2024 Published: 26 May 2024

Abstract - *Blindness and Visual Impairment are serious and pervasive health issues in the current situation. Although new innovative technologies are developing quickly, vision impairment continues to be a significant issue for global healthcare systems. Cataracts are among the most common causes of vision impairment. Accurate and rapid diagnosis of cataracts is the greatest way to prevent or treat it in its early stages. In this research, a weighted ensemble transfer learning based model is proposed for the prediction of cataracts from the fundus images. The weighted average ensemble is performed using three transfer learning models i.e. VGG16, VGG19 and ResNet50 model. Here, weight1, weight2 and weight3 are evaluated through optimisation for VGG16, VGG19 and ResNet50 models, respectively, to take the weighted average of three models. Different experiments based on different optimisers, different batch sizes and different numbers of epochs have been performed for the proposed ensemble model. Different transfer learning models, i.e. VGG16, VGG19 and ResNet50, are used individually also for training of the datasets and their performance is compared with the proposed ensemble model. The proposed weighted ensemble model is performing better as compared to the three transfer learning models with 99% sensitivity, 100% specificity, 100% precision, 64.5% recall and 99.37% accuracy.*

Keywords - *Artificial Intelligence, Cataract, Ensemble Mode, Ocular Disease Intelligent Recognition (ODIR), Transfer Learning (TF), Validation (Val).*

1. Introduction

Millions of individuals around the world, especially those over the age of 60, suffer from cataracts. A cataract is a vision loss that occurs due to a gradual clouding of the eye's natural lens. The lens, which sits behind the iris and pupil, focuses light onto the retina for sharp vision. Cataracts obscure the ordinarily transparent lens, reducing or obstructing vision. Cataracts develop as a result of the lens proteins degrading and clumping together due to age. One finds it challenging to read, drive, and even recognise faces when they have cataracts in their eyes [1].

Both crystalline lenses and ageing bring on cataracts. The lens' microscopic structure and chemical composition, among other interrelated factors, help to maintain its transparency and optical homogeneity. A yellow-brown pigment that darkens with age is present in the lens as a progressive deposit. Additionally, this lessens the amount of light that reaches the eyes. The terms "extracapsular cataract extraction" and "intracapsular cataract extraction" are used synonymously. The entire lens is removed during an operation called intracapsular extraction, leaving the capsule in place. In the industrialised world, this form of therapy is rarely employed. It is not dependent on an extremely reliable electricity supply.

Additionally, it only requires a brief period of training to perform. Extracapsular extraction is an additional strategy. A relatively large incision is needed in order to remove the lens's nucleus in one piece [5]. Patients who have early-stage cataracts often respond well to refractive lenses. Surgery for the removal of the cataract and intraocular lens implantation should be done in the hospital if outpatient treatment with corrective lenses and pupillary dilatation fails to enhance the patient's vision [6]. According to the World Health Organization (WHO) [2], there are 285 million visually impaired people in the world, including 246 million people who are moderately to severely blind and 39 million people who are entirely blind. Age-related cataracts cause 19.34 million people to lose vision in both eyes (less than 3/60 in the better eye), according to the 1998 World Health Report. 43% of all cases of blindness were attributable to this [3]. Recent cataract cases increased by 43.6%, and only 26.8% of those cases had cataract surgery.

In addition, the number of cataract surgeries of all kinds has increased recently. Studies show that patients are more likely to be women than men. This covers both nuclear and cortical cataracts as well as cataract surgery. Additionally, non-whites are more likely to experience it [4]. Ophthalmologists put a premium on the early discovery of cataracts since this permits provoked medication and treatment. Cataract-recognisable proof has truly included manual examination by ophthalmologists, which may be both time-consuming and error-prone. In any case, later progresses in deep learning have changed therapeutic imaging in general and the distinguishing proof of cataracts in particular.

Convolutional Neural Systems (CNNs), in particular, have been demonstrated to be a capable apparatus for handling and assessing medical images. These algorithms are able to accurately detect the existence and severity of cataracts by using massive datasets of annotated cataract photos to learn to recognise patterns and features indicative of cataracts.

There may be a shortage of research, particularly centering on utilising fundus images for cataract determination compared to other eye illnesses, such as diabetic retinopathy or age-related macular degeneration. Existing studies might need a comprehensive examination of the different highlights or designs in fundus images that are demonstrative of cataracts, possibly overlooking critical diagnostic markers. While deep learning has shown promise in medical image analysis, there might be a gap in its application for cataract diagnosis from fundus images, particularly in the context of transfer learning and ensemble methods. By identifying and addressing these research gaps, this study could contribute significantly to the advancement of cataract diagnosis from fundus images by applying an ensemble model, ultimately leading to improved patient outcomes and healthcare practices.

The following are the work's primary contributions:

- A weighted ensemble transfer learning-based model is proposed for the prediction of cataracts from the fundus images.
- The weighted ensemble is performed using three transfer learning models i.e., VGG16, VGG19 and ResNet50 models using weight1, weight2 and weight3 that are evaluated through optimisation, respectively.
- Different simulations based on different optimisers, different batch sizes and different numbers of epochs have been performed for the optimisation of the proposed ensemble model.

Utilising a weighted average ensemble of three transfer learning models (VGG16, VGG19, and ResNet50) for cataract prediction from fundus images is innovative. Ensembling allows for the combination of different models' strengths, potentially improving overall performance. The evaluation and optimisation of weights (weight1, weight2, weight3) for each model (VGG16, VGG19, ResNet50) in the ensemble is a novel approach. This optimisation likely enhances the ensemble's performance by giving more weight to the models that are more effective for cataract prediction. The rest of the paper is structured as follows: In Section 2, many methods that have been tried to diagnose and forecast cataracts in the past are shown. As described in Section 3, the proposed ensemble model is presented. Section 4 contains the findings and discussion, while Section 5 summarises the whole attempt.

2. Related Work

For the purpose of predicting cataracts in fundus pictures, numerous methods have been put forth by various authors [7-12]. A method for categorising cataracts was published and primarily consisted of four steps: pre-processing, extraction of features, selection of features, and classifier or model. Qian et al. 2018 [13] developed a five-convolution layer based on deep learning to segregate the characteristics in fundus images. A number of properties, including texture, drawing, colour, wavelet, acoustical, spectral parameters, and others, have been retrieved from retinal fundus images. The Squeeze Net model was used by Dong et al. [14] to examine techniques for cataract categorisation and detection. The pre-processed dataset is initially used to eliminate the noise from the photos. A few other models use the maximum entropy method for image pre-processing.

A model for classification utilising Softmax and Support Vector Machine was proposed by Divya et al. [15]. This research also conducted a comparative examination of many extracted features to demonstrate the superior accuracy of the softmax function. A CNN model for object recognition from a dataset of diabetic retinopathy was presented by Lia et al. utilising a pre-trained model. A CNN model was suggested by Sertkaya et al. 2019 [16] to study retinal disorders. The model's total accuracy after training on the AlexNet, LeNet, and VGG16 architectures was 82.9%. A dataset of cases of glaucoma, retinal illness, and common cataracts in the eyes was recognised with an accuracy of 82% in research by Gosh et al. 2020 [17] using CNN. Qiao et al. 2017 [18] used the approach that used the 17 components that make up the whole photo as input to the SVM algorithm. This approach has an accuracy rate of 87.52%. Li et al. had attained an accuracy of 95.00%. Using an active shape model that had been trained on the dataset of more than 5000 images. Hossain et al. 2020 [19] used ResNet50, a deep CNN module, to identify cataracts and other fundus conditions in his research. A system was introduced by Zhou et al. 2019 [20] for detection of cataracts. This system attained an accuracy of 94.00% accurate in detecting cataracts, and was also able to resolve the vanishing gradient problems. Li Xiong et al. 2017 [21] presented a study on the prediction of cataracts using VGG-19 and CNN, achieving an overall accuracy of 97.47%, with 97.47% precision and 5.27% loss. Other researchers [22-25] have also worked using different diseases with the help of deep learning.

A glaucoma classification approach using nine deep-learning models was presented by Prananda et al. [26]. They made use of the 650 glaucomatous images in the publicly accessible ORIGA-light collection. They were able to determine that the accuracy was 92.88% and the AUC was 89.34%. Thanki et al. [27] classified cataracts from the DRISTHI-GS dataset, which included 101 images, by combining deep characteristics with machine learning-based classifiers.

The categorisation was carried out by Kumar et al. [28] using three models: EfficientNetB0, VGG 16, and ResNet-152. They measured the precision, recall, accuracy, and F1-score, and the results were 99.71%, 98.63%, and 99.22%, respectively. Nine distinct machine-learning models were used by Marouf et al. [29] to classify cataracts. They had calculated the accuracy to be 99.11%. Four distinct datasets were employed in the red lesion identification algorithm that Saranya et al. [30] provided. They determined that the accuracy value for each of the four datasets was 95.65%.

3. Proposed Weighted Ensembled Transfer Learning-Based Model

Early diagnosis of cataract enables ophthalmologists to assess its progression and plan accordingly for appropriate treatment. Cataract diagnosis using deep learning is an emerging and promising application of Artificial Intelligence. Here, an ensemble deep learning model is proposed for the diagnosis of cataract at an early stage. The flow chart of the proposed model used for the prediction of cataract from fundus image is shown in Figure 1. The main modules of the methodology are pre-processing, weighted averaging of the three TF models, training of the proposed model and prediction of cataract using the proposed model.

3.1. Ocular Disease Intelligent Recognition (ODIR) Dataset

Ocular Disease Intelligent Recognition (ODIR) is a standardised ophthalmology database compiled from numerous Chinese hospitals and clinics. It includes color fundus images of the left and right eyes, age, sex, and diagnostic keywords from doctors. Several cameras with varying image resolutions are employed to acquire fundus images. Qualified human readers labeled the annotations as part of a quality control process. Patients are grouped into eight categories, including "normal," "diabetes," "glaucoma," "cataracts," "age-related macular degeneration," "hypertension," "pathological myopia," and "other diseases/abnormalities". Figure 2 represents some of the normal and cataract images of the ODIR dataset.

The cataract dataset contains left cataract images, right cataract images and normal eye images. Along with the number of images, this dataset contains patient information in terms of patient ID, age, sex, left fundus image, right fundus image, left-diagnostic keywords, right-diagnostic keywords, eight labels of patients, filename, label assigned to that patient, target and filename as represented by Figure 3.

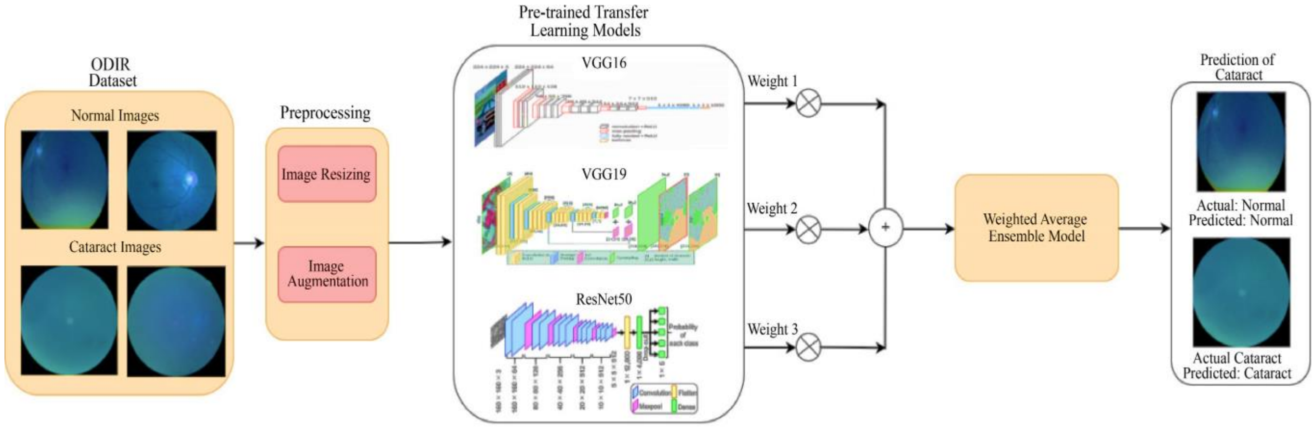


Fig. 1 Algorithm of proposed ensemble model

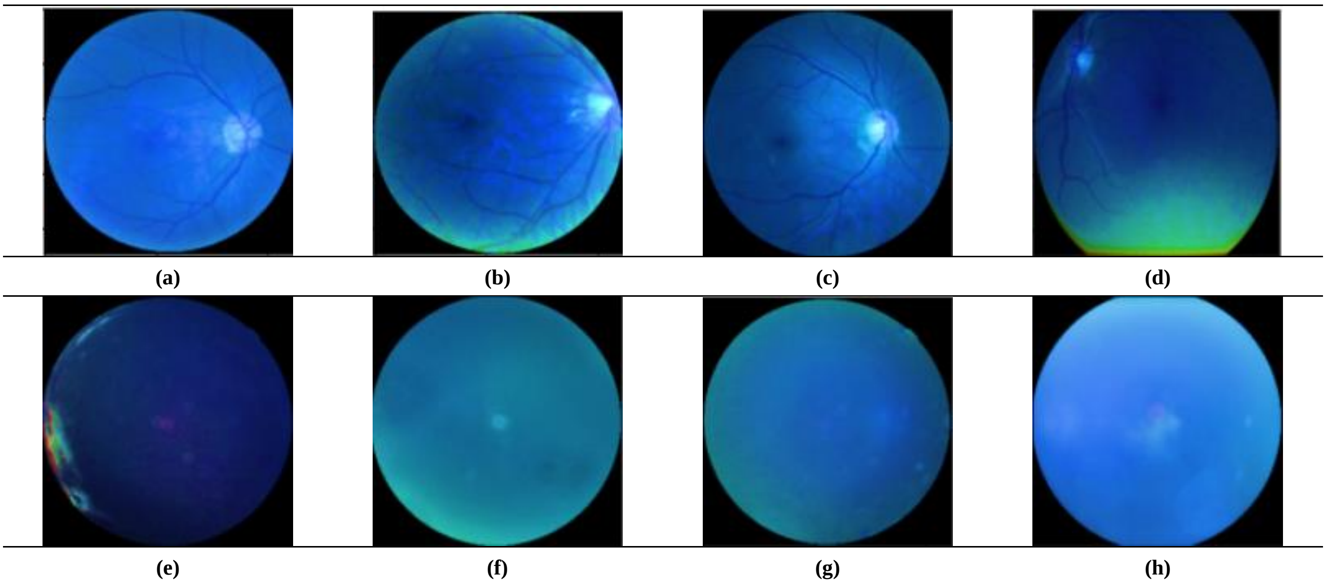


Fig. 2 Dataset of normal and cataract images (a-d) Normal images (e-h) Cataract Images

	ID	Patient Age	Patient Sex	Left-Fundus	Right-Fundus	Left-Diagnostic Keywords	Right-Diagnostic Keywords	N	D	G	C	A	H	M	O	filepath	labels	target	filename
0	0	69	Female	0_left.jpg	0_right.jpg	cataract	normal fundus	0	0	0	1	0	0	0	0	../input/ocular-disease-recognition-odir5k/ODI...	['N']	[1, 0, 0, 0, 0, 0, 0, 0]	0_right.jpg
1	1	57	Male	1_left.jpg	1_right.jpg	normal fundus	normal fundus	1	0	0	0	0	0	0	0	../input/ocular-disease-recognition-odir5k/ODI...	['N']	[1, 0, 0, 0, 0, 0, 0, 0]	1_right.jpg
2	2	42	Male	2_left.jpg	2_right.jpg	laser spot, moderate non proliferative retinopathy	moderate non proliferative retinopathy	0	1	0	0	0	0	0	1	../input/ocular-disease-recognition-odir5k/ODI...	['D']	[0, 1, 0, 0, 0, 0, 0, 0]	2_right.jpg

Fig. 3 Detailed information on patients

3.2. Pre-processing

After the extraction of the dataset into cataract and normal images, the next step is to pre-process the images. The pre-processing stage includes image resizing and image augmentation techniques.

3.2.1. Image Resizing

For pre-processing, initially image resizing operation has been performed. Each cataract and normal image are resized to 224 x 224. After that, the final dataset is created by appending the labels with each image depending upon the type of image. If the image has having cataract, then label 1 is appended to that image, and if the image is normal, then label 0 is appended to that image.

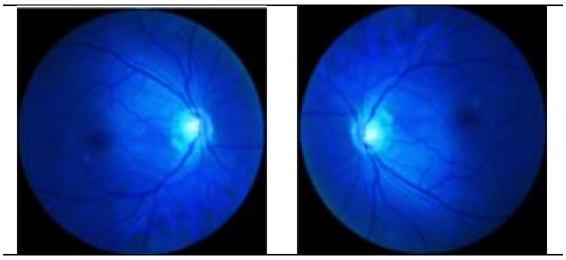
3.2.2. Image Augmentation

To increase the number of images, image augmentation techniques are also used. The use of augmentation also introduces variability in the dataset. In this paper, only a horizontal flipped operation is used for the augmentation of images. Figure 4 (a) signifies the original fundus image of the normal eye, and Figure 4 (b) characterises the corresponding horizontal flipped image.

3.2.3. Ensembling of Transfer Learning Model

In this section, the ensembling of different TF models, such as VGG16, VGG19 and Resnet50, is performed. Figure 5 shows the architecture of different TF models, where Figure 5 (a) illustrates the architecture of VGG16, Figure 5 (b) demonstrates the architecture of VGG19, and Figure 5 (c) displays the architecture of the ResNet50 model. Using weighted sum ensemble, which is also known as a weighted average ensemble, combines the predictions of numerous models while giving each model credit for its contribution in proportion to its competence or skill. The weighted average ensembling is performed by assigning three different weights, i.e. w1, w2 and w3, to three different TF models, VGG16, VGG19 and ResNet50, respectively. The respective weights are multiplied with the features extracted from different TF models as wnTn, where wn represents the weight assigned to the nth model, and Tn represents the features extracted from the nth model.

Now, these weights are optimised at the level where the best accuracy of the model is achieved. The weighted averaging of three models is shown in Figure 6. The model optimisation is performed using different values of weight for which different values of accuracy are obtained. The best weight value of accuracy is obtained using w1, w2 and w3 values as 0.5, and the value of accuracy obtained is 99.8%, as shown in Table 1.

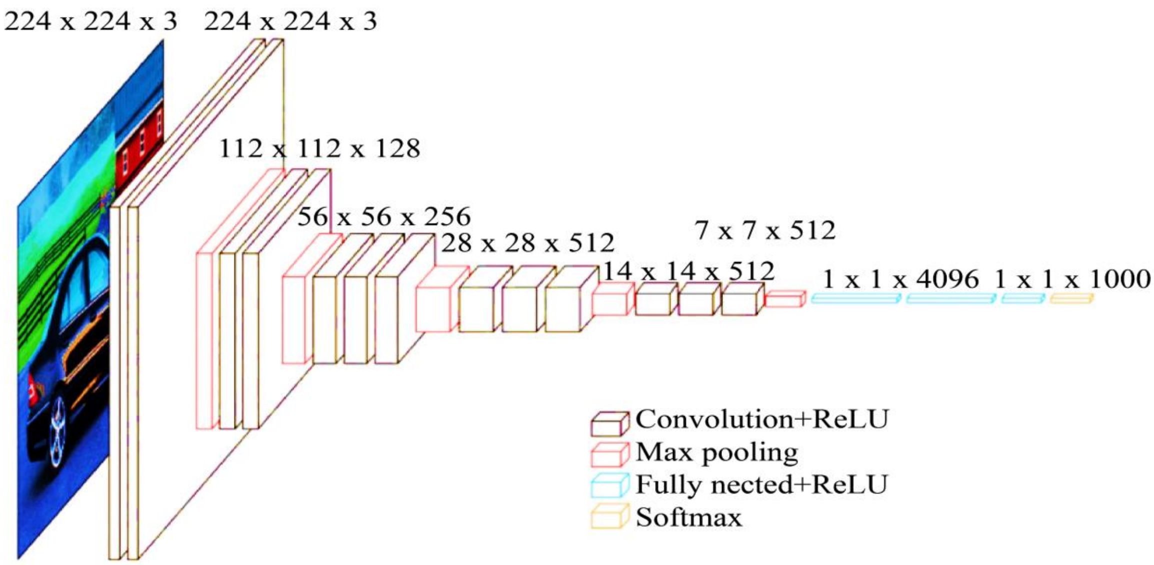


(a) **(b)**

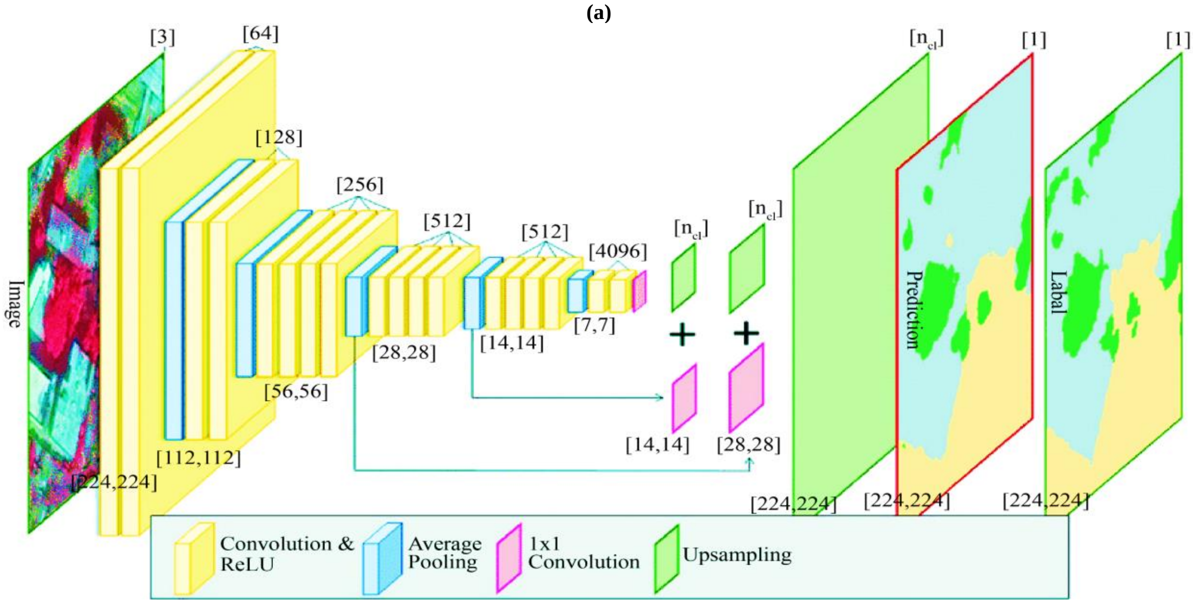
Fig. 4 (a) Original Image (b) Horizontal Flipped Image

Table 1. Accuracy of weighted average ensemble model with different weights

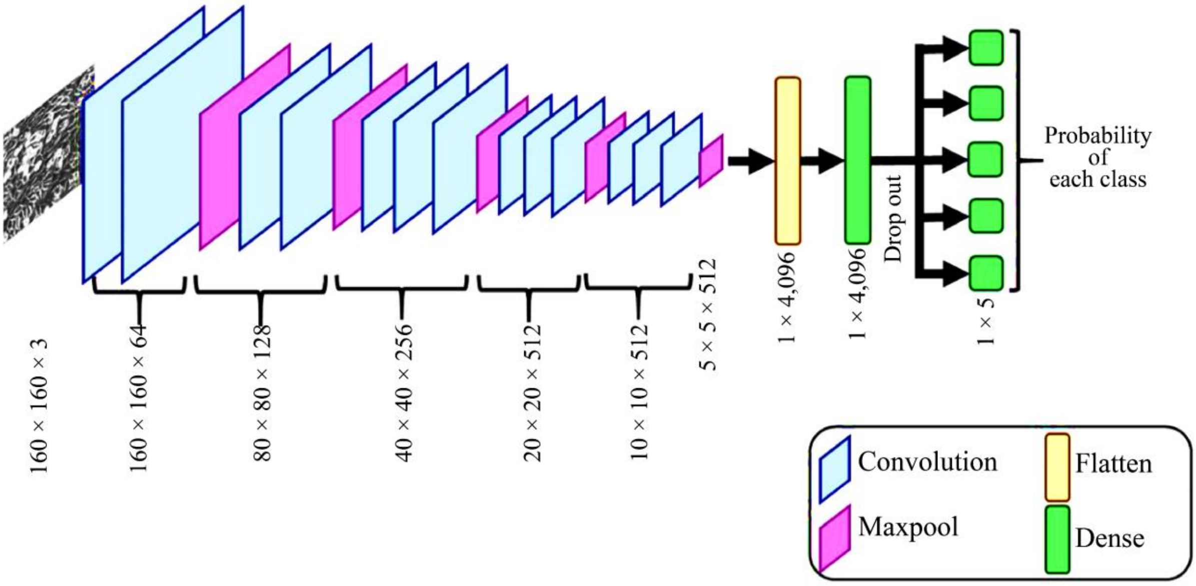
VGG16 w1	VGG19 w2	ResNet50 w3	Accuracy
0.0	0.0	0.0	53.47
0.0	0.0	0.1	63.38
0.0	0.0	0.2	65.77
0.0	0.0	0.3	69.16
0.0	0.0	0.4	71.49
…	…	…	…
0.5	0.5	0.0	97.15
0.5	0.5	0.1	97.43
0.5	0.5	0.2	97.93
0.5	0.5	0.3	98.45
0.5	0.5	0.4	98.84
0.5	0.5	0.5	99.81
0.5	0.5	0.6	98.94
0.5	0.5	0.7	97.34
…	…	…	…
0.9	0.9	0.9	62.48



(a)



(b)



(c)

Fig. 5 Architecture of pre-trained TF models (a) VGG16 (b) VGG19 (c) Resnet50

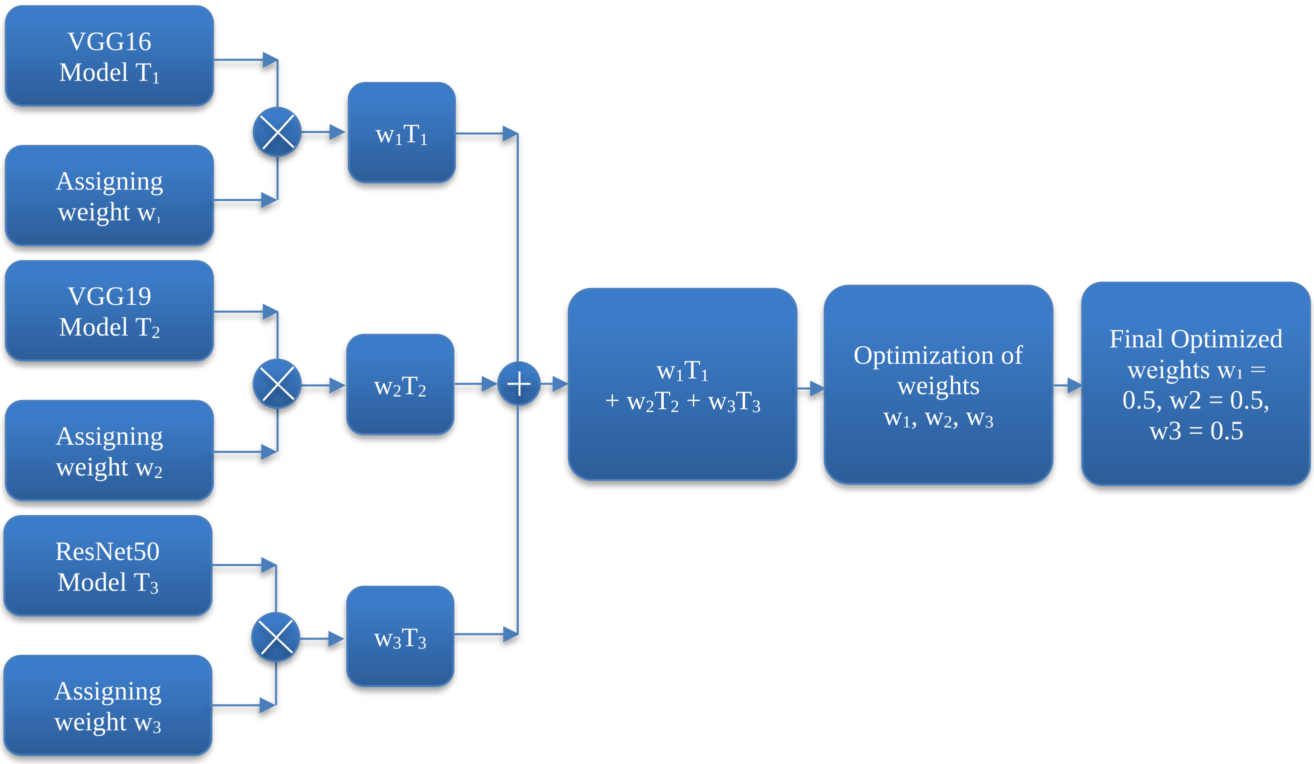


Fig. 6 Weighted average ensemble of three TF models

4. Results and Discussion

In this work, a weighted ensemble TF-based model is proposed for the prediction of cataract from the fundus images. The work on the proposed model is performed on the ODIR dataset using different batch sizes, number of epochs and different optimisers.

4.1. Analysis based on Different Optimisers

In this experiment, simulation is performed using different optimisers like Adadelta, Adam, RMSprop and SGD on batch size 8 at 10 epochs. The sigmoid function is used for the dense layer as an activation function. Each optimiser has its own set of hyperparameters that can be tuned to enhance model performance further. By experimenting with different optimisers, you can gain insights into how these hyperparameters affect your model. Different optimisers have different behaviors in terms of how they update model parameters. By observing how each optimiser affects the training process, you can gain a better understanding of your model's behavior and performance characteristics. Table 2 presents a comparison of diverse optimisation calculations utilised during the preparation of an ensemble model, besides their respective Val loss and Val accuracy. The Adadelta optimiser accomplished a val loss of 0.5538 and a Val precision of 0.8486. In this case, the model utilising the Adadelta optimiser had conventional precision but a moderately higher loss compared to a few other optimisers. The Adam optimiser outperformed the other optimisers with a Val loss of 0.5307 and a Val accuracy of 0.9725. This shows that the model utilising the Adam optimiser achieved way better, because it had a lower loss and higher accuracy on the validation set. The RMSprop optimiser achieved a Val loss of 1.0669 and a Val accuracy of 0.9587. At the same time, it performed well in terms of accuracy.

Table 2. Val loss and Val accuracy for different optimisers

Optimisers	Val Loss	Val Accuracy
Adadelta	0.5538	0.8486
Adam	0.5307	0.9725
RMSprop	1.0669	0.9587
SGD	12.0510	0.9633

The SGD optimiser had the most elevated Val loss of 12.0510 among all the optimisers, demonstrating that the model's expectations had a noteworthy mismatch with the genuine target values during validation. In summary, the Adam optimiser produced the best results, achieving the lowest validation loss and highest validation accuracy. Based on this examination, it is evident that the choice of optimiser can have a critical effect on the execution of the model. Adam shows up to be the foremost effective optimiser for this dataset, accomplishing the most reduced loss and most elevated accuracy.

4.2. Analysis based on Different Batch Size

The batch size can essentially affect the execution and merging of the ensemble model. By testing with diverse batch sizes, we will recognise the batch size that leads to ideal execution for our particular task. Additionally, larger batch sizes regularly lead to quicker training times due to more proficient utilise of hardware assets (e.g., GPUs). By finding the ideal batch estimate, we will accomplish speedier training without compromising model execution. Table 3 presents a comparison of different Batch Sizes (BS) used during the training of an ensemble model, along with their respective validation loss and validation accuracy. The model prepared with a BS of 8 accomplished a Val loss of 0.5925 and an accuracy of 0.9811. This demonstrates that employing a smaller batch size allowed the model to attain great precision on the validation set.

Table 3. Val loss and Val accuracy for different batch sizes

Batch Size	Val Loss	Val Accuracy
8	0.5925	0.9811
16	0.5692	0.972
32	0.3580	0.9774
64	0.3166	0.9837
128	0.2217	0.9858
256	0.2888	0.9774

The model prepared with a BS of 32 accomplished a Val loss of 0.3580 and a Val accuracy of 0.9774. The model prepared with a BS of 64 achieved a Val loss of 0.3166 and a Val accuracy of 0.9837. The model prepared with a batch measure of 256 achieved a Val loss of 0.2888 and a Val accuracy of 0.9774. Whereas the precision is still great, it is marginally lower compared to the batch estimate of 128. In summary, the results demonstrate that increasing the batch size generally leads to improved Val accuracy and reduced Val loss. It is essential to choose the batch size that balances computational resources and training performance for a given deep learning task. Based on this analysis, using a batch size of 128 seems to be optimal for this model, as it achieves both a low validation loss and a high validation accuracy. Using smaller batch sizes may result in slower convergence and higher loss values, while using larger batch sizes may lead to a slight decrease in accuracy.

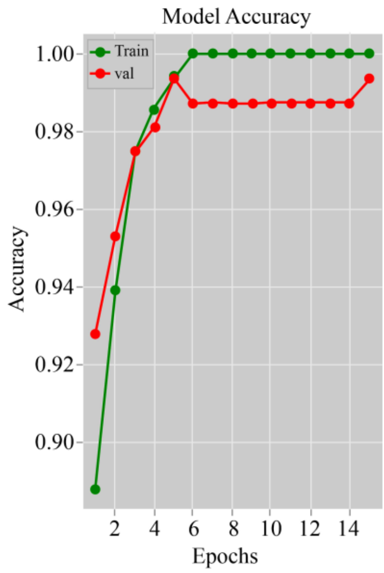
4.3. Analysis based on Different Number of Epochs

Training for too few epochs may lead to underfitting, where the model fails to capture the underlying patterns in the data. On the other hand, training for as well numerous epochs may lead to overfitting, where the model learns noise within the prepared information. By recreating with different numbers of epochs, we can analyse the nearness of overfitting and decide the point at which the model begins to overfit. Watching the model's execution over distinctive numbers of epochs can offer assistance in understanding its joining behavior. This may give experiences into how rapidly the model learns and whether further training would lead to noteworthy advancements. Table 4 presents a comparison of different numbers of epochs used during the training of the ensemble model, along with their respective Val loss and Val accuracy. The model trained for five epochs attained a Val loss of 0.425 and a high Val accuracy of 0.9811. The model trained for 10 epochs attained a Val loss of 0.2217 and an impressive Val accuracy of 0.9858. Doubling the number of epochs allowed the model to improve its accuracy and reduce the loss significantly.

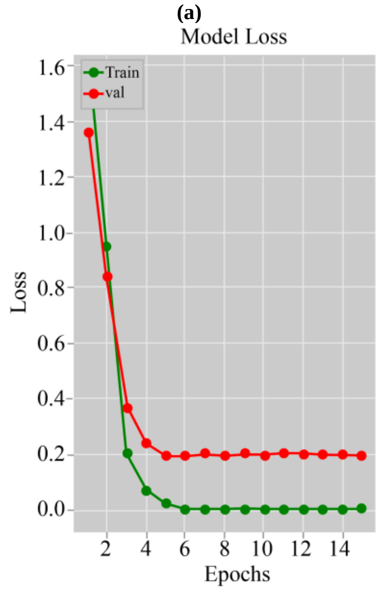
The model trained for 20 epochs attained a Val loss of 0.2266 and a Val accuracy of 0.9874. Although the accuracy is still relatively high, it appears that the model may have started to overfit the data as the Val loss increased slightly compared to the model trained for 15 epochs. In summary, the results demonstrate the impact of the number of epochs on the proposed model's performance. Increasing the number of epochs can lead to improved accuracy and reduced loss, but there is a point of diminishing returns where the model may start to overfit the data. In this specific case, training the model for 15 epochs produced the best results, achieving the highest Val accuracy and the lowest Val loss. It is essential to monitor the Val metrics during training and select the number of epochs that strike the right balance between training performance and generalisation on unseen data. Early stopping techniques can also be applied to prevent overfitting and select the optimal number of epochs for the given deep-learning task. Based on this analysis, it appears that training the model for around 15 epochs may be optimal, as it achieves a good balance between minimising loss and maximising accuracy. Further increasing the number of epochs beyond this point may lead to overfitting, as indicated by the increase in validation loss at 20 epochs. Model accuracy and model loss attained using the proposed model for Adam optimiser, batch 128 and at different numbers of epochs are represented by Figure 7 (a) and (b), respectively.

Table 4. Val loss and Val accuracy at different numbers of epochs

Number of Epochs	Val Loss	Val Accuracy
5	0.425	0.9811
10	0.2217	0.9858
15	0.1817	0.9937
20	0.2266	0.9874



(a)



(b)

Fig. 7 Analysis at different numbers of epochs (a) Model Accuracy (b) Model Loss

4.4. Analysis based on Different Performance Metrics and Different Transfer Learning Models

Table 5 presents a comparison of different TF models with the proposed model for cataract detection, along with their respective evaluation metrics. The TF models evaluated are ResNet50, VGG16, VGG19, and a proposed ensemble model. ResNet50 model attained a high sensitivity (recall) of 0.98, indicating it can correctly identify 98% of positive cases (cataracts). It also showed a good specificity of 0.97, implying that it can correctly identify 97% of negative cases (non-cataract). The precision, which measures the proportion of true positive predictions among all positive predictions, was 0.96. However, the recall (44%) was relatively low, suggesting that the model had some difficulty in correctly identifying all cataract cases. Overall accuracy was high at 0.97.

VGG16 model performed well with a sensitivity of 0.95, indicating it can correctly identify 95% of cataract cases. It showed perfect specificity (1), meaning it can correctly identify all non-cataract cases. The precision was 0.95, indicating that 95% of the positive predictions were correct. However, the recall (49.5%) was relatively low, indicating that the model may miss some cataract cases. The overall accuracy was 0.97.

VGG19 model attained a very high sensitivity (recall) of 0.989, indicating it can correctly identify 98.9% of cataract cases. It also showed a good specificity of 0.983, meaning it can correctly identify 98.3% of non-cataract cases. The precision was 0.989, indicating that 98.9% of the positive predictions were correct. The recall (75.6%) was higher than the previous models, suggesting that it can better identify cataract cases. The overall accuracy was 0.987.

The proposed ensemble model outperformed the others in terms of sensitivity (recall) with a value of 0.990, indicating it can correctly identify 99% of cataract cases. It showed perfect specificity (1), correctly identifying all non-cataract cases. The precision was 1, meaning all positive predictions were correct. The recall (64.5%) was higher than some models but still indicates some missed cataract cases. The overall accuracy was high at 0.9937. In summary, the proposed ensemble model attained the highest sensitivity, specificity, precision, and accuracy among the evaluated models, making it a promising candidate for cataract detection.

4.5. Visual Analysis of Classification and Misclassification Results Using Ensembled Model

After compiling and fitting of model, the labels of each image are predicted. Suppose the predicted labels of images are 0. In that case, the images are classified into normal images, and if the predicted labels of images are 1, then the images are classified into cataract images. The images classified using the proposed ensemble model are shown in Table 6. The first row shows predictions for normal images, and the second row shows predictions for cataract images.

4.6. Comparison with Previous State-of-Art Methods

Comparative analysis of proposed work is also a mode with state-of-art methods, as shown in Table 7. The Table presents a diverse set of techniques and models used for cataract classification tasks, showcasing varying levels of accuracy attained on different datasets. The traditional image processing methods like top-bottom hat transformation combined with neural networks also exhibit promising results, particularly on smaller datasets. Sertkaya et al. [16] used LeNet, AlexNet, and Vgg16 models for cataract classification and attained an accuracy of 82.9% on a dataset containing 26,581 images. Yang et al. [18] worked on top-bottom hat transformation and a neural network classifier, resulting in an accuracy of 90.86% on a relatively small dataset of 504 images. Hossain et al. [19] employed a Deep Convolution Neural Network and attained a high accuracy of 97.38% on a dataset of 4,000 images for cataract classification. Zhou et al. [20] used DST-ResNet and Vanilla-ResNet models separately for cataract classification. DST-ResNet attained an accuracy of 94.0%, and Vanilla-ResNet attained an accuracy of 92.5%. Both models were evaluated on the same dataset containing 1,355 images. Xiong et al. [21] used a Decision Tree model for cataract classification, which attained an accuracy of 92.8% on the same dataset of 1,355 images. The proposed Ensemble model stands out as the most accurate approach in this comparison, achieving an accuracy of 99.81% on a larger dataset having 8000 cataract fundus images.

The ensemble of three TF models (VGG16, VGG19, and ResNet50) in our investigation is likely able to attain superior results compared to state-of-the-art methods for cataract detection in fundus images. Each of the three models (VGG16, VGG19, and ResNet50) has diverse models and learns diverse highlights from the information. By combining them in an ensemble, we are viably leveraging the complementary highlights learned by each model, which can lead to a more comprehensive representation of the information. Ensembles are regularly more robust to noise and exceptions within the information compared to personal models. This is because the predictions of multiple models are averaged, which can help smooth out any inconsistencies or errors in individual predictions. Moreover, by using weighted averaging of the predictions from the three models, we are effectively fusing the information from each model in a way that maximises performance. The optimisation of weights for each model further enhances the ensemble's ability to make accurate predictions.

Table 5. Analysis of different performance metrics for different pre-trained models and proposed ensemble model

Model	Sensitivity	Specificity	Precision	Recall	Accuracy
Resnet50	0.98	0.97	0.96	0.44	0.97
VGG16	0.95	1	0.95	0.495	0.97
VGG19	0.989	0.983	0.989	0.756	0.987
Proposed Ensemble model	0.990	1	1	0.645	0.9937

Table 6. Prediction of cataract using ensemble model

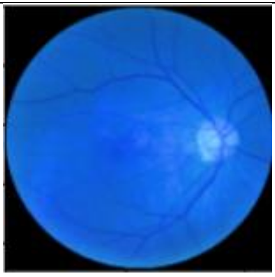
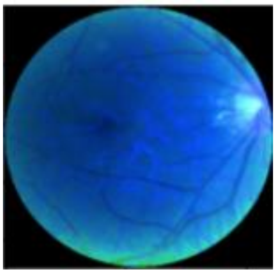
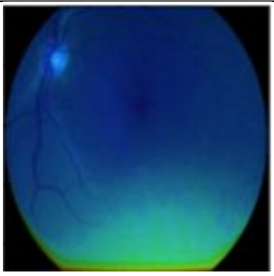
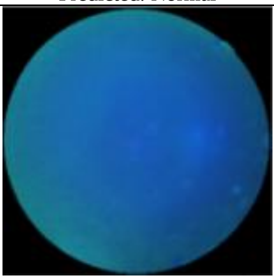
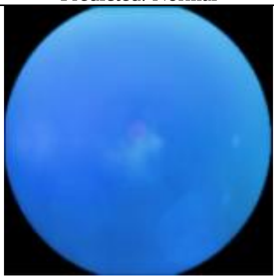
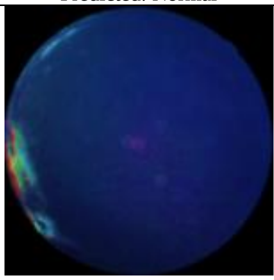
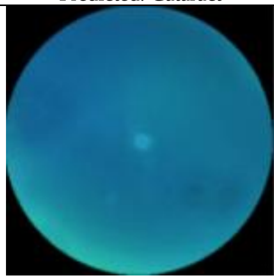
 Actual: Normal Predicted: Normal	 Actual: Normal Predicted: Normal	 Actual: Normal Predicted: Normal	 Actual: Normal Predicted: Cataract
 Actual: Cataract Predicted: Cataract	 Actual: Cataract Predicted: Cataract	 Actual: Cataract Predicted: Cataract	 Actual: Cataract Predicted: Normal

Table 7. Comparative analysis of the proposed ensemble model with previous state-of-art methods

Authors	Technique Used	Number of Images	Accuracy (%)
Sertkaya et al. [16]	LeNet, AlexNet, Vgg16	26,581	82.9
Yang et al. [18]	Top-bottom hat transformation and neural network classifier	504	90.86
Hossain et al. [19]	Deep Convolution Neural Network	4000	97.38
Zhou et al. [20]	DST-ResNet Vanilla-ResNet	1355	94.0 92.5
Xiong et al. [21]	Decision Tree	1355	92.8
Kumar et al. [31]	Hybrid Deep Learning	-	99.00
Shamsan et al. [32]	Artificial Neural Network	4217	98.5
Proposed Ensemble model	weighted average ensemble of VGG16, VGG19 and ResNet50	8000	99.81

5. Conclusion

In this research, images are separated into a cataract and non-cataract category using a pre-trained ensemble model. In order to use the dataset effectively, it was first assembled and then split into two parts: the training set and the test set. 20% of the images in the dataset are used for testing, while 80% are used for training new images to classify the cataract in the fresh input image effectively. The proposed model is performing well for Adam optimiser, for batch 128 and at 15 number of epochs. Comparative analysis of the proposed ensemble model is also made with pre-trained models.

When compared to existing pre-trained models, the proposed ensemble model outperforms in terms of all performance metrics. It is important to note that the proposed deep learning model has shown significant potential in cataract diagnosis, it should be considered as an aid to clinicians rather than a replacement for professional medical judgment. It can assist in the screening process and provide quick initial assessments, but qualified healthcare professionals should always assist with any diagnosis or treatment decisions.

Funding Statement

The authors are thankful to the Deanship of Scientific Research at Najran University for funding this work under the Future Funding program grant code (NU/SRP/SERC/12/2).

References

[1] Dishant Padalia, Abhishek Mazumdar, and Bharati Singh, “A CNN-LSTM Combination Network for Cataract Detection Using Eye Fundus Images,” *Arxiv*, pp. 1-8, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[2] Seth R. Flaxman et al., "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. [CrossRef] [Google Scholar] [Publisher Link]
[3] A. Foster, "Vision 2020: The Cataract Challenge," *Community Eye Health*, vol. 13, no. 34, pp. 17-19, 2000. [Google Scholar] [Publisher Link]
[4] Euna Koo et al., "Ten-Year Incidence Rates of Age-Related Cataract in the Age-Related Eye Disease Study (AREDS): AREDS Report No. 33," *Ophthalmic Epidemiology*, vol. 20, no. 2, pp. 71-81, 2013. [CrossRef] [Google Scholar] [Publisher Link]
[5] David Allen, and Abhay Vasavada, "Cataract and Surgery for Cataract," *Bmj Journals*, vol. 333, pp. 128-132, 2006. [CrossRef] [Google Scholar] [Publisher Link]
[6] Md Kamrul Hasan et al., "Cataract Disease Detection by Using Transfer Learning-Based Intelligent Methods," *Computational and Mathematical Methods in Medicine*, vol. 2021, pp. 1-12, 2021. [CrossRef] [Google Scholar] [Publisher Link]
[7] Ji-Jiang Yang et al., "Exploiting Ensemble Learning For Automatic Cataract Detection and Grading," *Computer Methods and Programs in Biomedicine*, vol. 124, pp. 45-57, 2016. [CrossRef] [Google Scholar] [Publisher Link]
[8] Huiqi Li et al., "A Computer-Aided Diagnosis System of Nuclear Cataract," *IEEE Transactions on Biomedical Engineering*, vol. 57, no. 7, pp. 1690-1698, 2010. [CrossRef] [Google Scholar] [Publisher Link]
[9] Miguel Caixinha et al., "New Approach for Objective Cataract Classification Based on Ultrasound Techniques Using Multiclass SVM Classifiers," *2014 IEEE International Ultrasonics Symposium*, Chicago, IL, USA, pp. 2402-2405, 2014. [CrossRef] [Google Scholar] [Publisher Link]
[10] Linglin Zhang et al., "Automatic Cataract Detection and Grading Using Deep Convolutional Neural Network," *2017 IEEE 14th International Conference on Networking, Sensing and Control (ICNSC)*, Calabria, pp. 60-65, 2017. [CrossRef] [Google Scholar] [Publisher Link]
[11] Weiming Fan et al., "Principal Component Analysis Based Cataract Grading and Classification," *2015 17th International Conference on E-Health Networking, Application & Services (HealthCom)*, Boston, MA, pp. 459-462, 2015. [CrossRef] [Google Scholar] [Publisher Link]
[12] Liye Guo et al., "A Computer-Aided Healthcare System for Cataract Classification and Grading Based on Fundus Image Analysis," *Computers in Industry*, vol. 69, pp. 72-80, 2015. [CrossRef] [Google Scholar] [Publisher Link]
[13] Xingzhi Qian et al., "Machine Learning on Cataracts Classification using SqueezeNet," *2018 4th International Conference on Universal Village (UV)*, Boston, MA, USA, pp. 1-3, 2018. [CrossRef] [Google Scholar] [Publisher Link]
[14] Yanyan Dong et al., "Classification of Cataract Fundus Image Based on Deep Learning," *2017 IEEE International Conference on Imaging Systems and Techniques (IST)*, Beijing, China, pp. 1-5, 2017. [CrossRef] [Google Scholar] [Publisher Link]
[15] Divya, and C. Senbagavalli, "Blood Vessel Segmentation and Classification Based On ANFIS Method for Retinal Images," *International Journal of Computer Science Trends and Technology*, vol. 6, no. 4, pp. 123-130, 2018. [Google Scholar] [Publisher Link]
[16] Mehmet Emre Sertkaya, Burhan Ergen, and Mesut Togacar, "Diagnosis of Eye Retinal Diseases Based on Convolutional Neural Networks Using Optical Coherence Images," *2019 23rd International Conference Electronics*, Palanga, Lithuania, pp. 1-5, 2019. [CrossRef] [Google Scholar] [Publisher Link]
[17] Chandra Lekha Dondapati, Ashutosh Ghosh, and TYJ. Naga Malleswari, "Classification of Eye Disorders Based on Deep Convolutional Neural Network," *International Journal of Innovative Technology and Exploring Engineering*, vol. 9, no. 6, pp. 1388-1393, 2020. [CrossRef] [Google Scholar] [Publisher Link]
[18] Meimei Yang et al., "Classification of Retinal Image for Automatic Cataract Detection," *2013 IEEE 15th International Conference on E-Health Networking, Applications and Services (Healthcom 2013)*, Lisbon, pp. 674-679, 2013. [CrossRef] [Google Scholar] [Publisher Link]
[19] Rajib Hossain et al., "Automatic Detection of Eye Cataract Using Deep Convolution Neural Networks (DCNNs)," *2020 IEEE Region 10 Symposium (TENSYMP)*, Dhaka, Bangladesh, pp. 1333-1338, 2020. [CrossRef] [Google Scholar] [Publisher Link]
[20] Yue Zhou, Guoqi Li, and Huiqi Li, "Automatic Cataract Classification Using Deep Neural Network with Discrete State Transition," *IEEE Transactions on Medical Imaging*, vol. 39, no. 2, pp. 436-446, 2020. [CrossRef] [Google Scholar] [Publisher Link]
[21] Li Xiong, Huiqi Li, and Liang Xu, "An Approach to Evaluate Blurriness in Retinal Images with Vitreous Opacity for Cataract Diagnosis," *Journal of Healthcare Engineering*, vol. 2017, pp. 1-17, 2017. [CrossRef] [Google Scholar] [Publisher Link]
[22] T.R Soumya et al., "Profile Similarity Recognition in Online Social Network using Machine Learning Approach," *2022 4th International Conference on Inventive Research in Computing Applications (ICIRCA)*, Coimbatore, India, pp. 805-809, 2022. [CrossRef] [Google Scholar] [Publisher Link]
[23] Sanjay Singh et al., "COVID 19: Identification of Masked Face Using CNN Architecture," *2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC)*, Coimbatore, India, pp. 1045-1051, 2022. [CrossRef] [Google Scholar] [Publisher Link]
[24] Tanima Bhattacharya et al., "Applications of Phyto-Nanotechnology for the Treatment of Neurodegenerative Disorders," *Materials*, vol. 15, no. 3, pp. 1-32, 2022. [CrossRef] [Google Scholar] [Publisher Link]
[25] Bhisham Sharma, and Deepika Koundal, "Cattle Health Monitoring System Using Wireless Sensor Network: A Survey From Innovation Perspective," *IET Wireless Sensor Systems*, vol. 8, no. 4, pp.143-151, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[26] Alifia Revan Prananda et al., “Retinal Nerve Fiber Layer Analysis Using Deep Learning to Improve Glaucoma Detection in Eye Disease Assessment,” *Applied Sciences*, vol. 13, no. 1, pp. 1-21, 2023. [CrossRef] [Google Scholar] [Publisher Link]
[27] Rohit Thanki, “A Deep Neural Network and Machine Learning Approach for Retinal Fundus Image Classification,” *Healthcare Analytics*, vol. 3, pp. 1-9, 2023. [CrossRef] [Google Scholar] [Publisher Link]
[28] K. Susheel Kumar, and Nagendra Pratap Singh, “Retinal Disease Prediction through Blood Vessel Segmentation and Classification Using Ensemble-Based Deep Learning Approaches,” *Neural Computing and Applications*, vol. 35, pp. 12495-12511, 2023. [CrossRef] [Google Scholar] [Publisher Link]
[29] Ahmed Al Marouf et al., “An Efficient Approach to Predict Eye Diseases from Symptoms Using Machine Learning and Ranker-Based Feature Selection Methods,” *Bioengineering*, vol. 10, no. 1, pp. 1-20, 2023. [CrossRef] [Google Scholar] [Publisher Link]
[30] P. Saranya, R. Pranati, and Sneha Shruti Patro, “Detection and Classification of Red Lesions from Retinal Images for Diabetic Retinopathy Detection Using Deep Learning Models,” *Multimedia Tools and Applications*, vol. 82, pp. 39327-39347, 2023. [CrossRef] [Google Scholar] [Publisher Link]
[31] Yogesh Kumar, and Bharat Gupta, “Retinal Image Blood Vessel Classification using Hybrid Deep Learning in Cataract Diseased Fundus Images,” *Biomedical Signal Processing and Control*, vol. 84, 2023. [CrossRef] [Google Scholar] [Publisher Link]
[32] Ahlam Shamsan, Ebrahim Mohammed Senan, and Hamzeh Salameh Ahmad Shatnawi, “Automatic Classification of Colour Fundus Images for Prediction Eye Disease Types Based on Hybrid Features,” *Diagnostics*, vol. 13, no. 10, pp. 1-23, 2023. [CrossRef] [Google Scholar] [Publisher Link]