Journal of Al-Qadisiyah for Computer Science and Mathematics Vol.17.(4) 2025,pp.Comp 174–183

Available online at www.qu.edu.iq/journalcm
JOURNAL OF AL-QADISIYAH FOR COMPUTER SCIENCE AND MATHEMATICS
ISSN:2521-3504(online) ISSN:2074-0204(print)

# Deep Learning Models for Ocular Disease Detection

**Bahaa Salih Mandeel [a]*, Manar Raad Abdalhassan [b]*, Diyaa Kareem Abd-Alhmza [c]*, Muhtadi Thamer Mousa Raad [d] ***

[a]College of Technical Engineering, Islamic University, Dewania, Iraq.

[bcd]University of Al-Qadisiyah, College of Computer Science and Information Technology, Dewania, Iraq.

[a]Email:bahaqq645@gmail.com

[b]Email: malmuhja@gmail.com

[c]Email: diyaakareem68@gmail.com

[d]Email: muhtadi201348@gmail.com

ARTICLE INFO

*Article history:*
Received: 08/08/2025
Rrevised form: 06 /10/2025
Accepted : 12/10/2025
Available online: 30/12/2025

*Keywords:*

ODIR-2019 dataset, InceptionV3, GoogLeNet, Deep Learning, Binary classification.

ABSTRACT

Early and precise identification of ocular diseases using fundus imaging is essential for effective ophthalmic care, yet it remains a non-trivial challenge due to the complexity of visual patterns and inter-observer variability. Conventional manual diagnosis, while widely practiced, is labor-intensive and susceptible to subjective bias, thereby limiting scalability and consistency in clinical settings. To address these limitations, automated diagnostic systems based on deep learning have emerged as a promising alternative, offering improved efficiency, reproducibility, and diagnostic accuracy. In this work, we propose a robust deep learning framework for automated detection of ocular diseases, with a specific focus on distinguishing between normal and cataract cases. This was performed through two deep learning models, an InceptionV3 structure and a pre-trained GoogLeNet model, which was developed and extensively validated. Both models were trained and evaluated on the Ocular Disease Intelligent Recognition (ODIR-2019) dataset in case of binary classification tasks. The InceptionV3 model had a relatively high accuracy of 97.9\% and the GoogLeNet model also achieved a high accuracy of 97.6\%. These results demonstrate the possibility of professional DL algorithms to provide comprehensible, precise, and efficient solutions for automation diagnosing ocular diseases, and to be a significantly innovative breakthrough to ophthalmic disease detection and clinical decision supporting.

MSC..

https://doi.org/ 10.29304/jqcsm.2025.17.42550

## 1. Introduction

If left unattended, some eye problems can result in loss of vision and in extreme cases, total blindness. Visual impairment significantly impairs the quality of life and affects approximately 2.2 billion individuals worldwide, as reported recently [2]. Early detection is crucial for maintaining their vision, as many ocular diseases can be slowed

*Corresponding author Bahaa Salih Mandeel

Email addresses: bahaqq645@gmail.com

Communicated by 'sub etitor'

with prompt treatment [3].The retina, an elegant sheet of neural tissue at the back of the eye, transduces light into neural signals [4]. Accurate disease diagnosis is necessary for early state intervention in retinal diseases such as diabetic retinopathy (DR), glaucoma, cataract, and age-related macula degeneration to avoid severe visual consequences [5]. The ability to rapidly and accurately evaluate large sets of technical data using artificial intelligence (AI) has transformed the practice of medicine, particularly ophthalmology [6].In this study we investigate the applicability of deep learning techniques on efficientl y detection and classification of eye diseases. We concentrate our attention on two models, InceptionV3 and the pre-trained GoogLeNet. We employ here the Ocular Disease Intelligent Recognition (ODIR-2019) dataset, which consists of images associated to several ocular conditions, to a binary classification task that aims to discern healthy eyes and those afflicted by cataracts. In this study, we would like to verify the diagnostic performance and efficiency in identifying eye disease and increasing opportunities for early diagnosis in the clinical setting with automated systems"In this study, we aim to assess the performance of the model in diagnosing the eye condition, and clinical performance of the surrogate and in the early diagnosis of the eye condition.

## *2. Literature Review*

The development of deep learning technique has greatly improved the accuracy and efficiency in detecting ocular diseases. Many researches have attempted to use CNNs and a diverse of machine learning algorithms to classify and diagnose eye diseases from fundus images.

Afsana Ahsan Jeny et al [7] proposed a deep CNN-based ensemble method which consists of 20 layers with activation, boosting, and loss functions. They used pre-processing techniques such as CLAHE (Contrast-Limited Adaptive Histogram Equalization) and a Gaussian filter to ensure the sharpness of the images and to reduce the noise, respectively. The authors have made a comparison of their CNN ensemble with pre-trained CNNs such as VGG16, DenseNet201, and ResNet50 and were able to demonstrate that their ensemble approach has outperformed these networks with accuracy of 95.95%.

Ari Taha Mustafa et al. proposed a method of ODM detection based on deep learning in [8] using the Ocular Disease Intelligent Recognition (ODIR) dataset. The difficulty of the classification of eye diseases by using simpler models demonstrated here by using all features and SVM with the accuracy of 71.3%.

Bajwa MN et al. [9] proposed a two-staged approach to the localization and classification of optic disc from retinal fundus images. They classified glaucomatous and normal discs using a deep CNN and used a Regions with Convolutional Neural Networks (RCNN) for disc localization. It achieved an AUC of 0.874 for glaucoma classification on the public datasets and had 100% localization accuracy on six datasets. The performance of the model has been demonstrated, indeed, it is difficult for practical use due to the high computational costs.

Zahraa Najm AbedAbbas and M. Al-Bakry [10] also worked on the ODIR dataset where it was divided into 70% training and 30% testing dataset. They pre-processed the images by converting images to grayscale, histogram equalization, blurring and scaling. Feature extraction with Scale-Invariant Feature Transform (SIFT) and GLCM was then performed. We performed classification, using a number of machine learning methods, including Naive Bayes, Decision Tree, Random Forest and K-Nearest Neighbor. The best accuracy achieved in binary classification was 75% by Naïve Bayes while Random Forest was the best in multiclass with 88% accuracy. This paper also highlights the good performance of preprocessing methods with conventional machine learning techniques in detecting ocular diseases as well as the difficulty in obtaining higher accuracy.

Kai et al. [13] used loss function of a CNN model to produce probability map. A convolutional neural network (CNN) was applied to the feature maps to recover accurate information about the blood vessels in the retina.

Clement et al. [14] presented a convolutional architecture which was\ enhanced through supervised learning, trained on the Messidor dataset. The system generated and examined pixel-level slices after the detection of the lesions, which were furthered separated into red and bright elements. This method achieved excellent performance with AUC 83.9%.

Rahul et al. [15] proposed a hybrid system of CNN and SVM to detect cataracts in fundus images. After extraction of the features, the dataset was augmented resulting in a CNN accuracy of 87.08% and SVM accuracy of 87.5%.

*The latter point is to draw attention that we are responding the ongoing development of strong, efficient, and accurate models for ocular disease detection. Despite significant progress, challenges including the computational cost or complex model structures, and the need for a high diagnostic accuracy, remain. The work presented in this paper builds upon these findings and further investigates the performance of InceptionV3 as well as a pre-trained GoogLeNet model in an attempt to enhance the diagnostic accuracy and computational efficiency of automated ocular disease detection systems.*

## 3. Methodology

The primary objective of the research is to create efficient and accurate computer models to identify and diagnose eye ailments. The overwhelming number of eye images seen in clinical environment present particular issues that our models attempt to tackle. We performed a detailed assessment on two deep learning models:InceptionV3 and GoogLeNet. To fine-tune the InceptionV3 model, on top of the final output layer, several additional layers were included. Layers were made dense to reduce data and increase speed. The final layer with softmax activation greatly improves the classification accuracy by making use of effectively classifying the images. The best results were achieved with InceptionV3 using the SGDM optimizer.

Our approach involves data process, architecture design, training schedules and evaluation protocol. The approach starts with the acquisition of the ocular images and ends with the analysis of the train models. This is a consolidated approach comprising state-of-the-art techniques and best practises in deep learning and medical image analysis techniques. FIGURE 1 Figure 1 Sequence of steps in the proposed models.

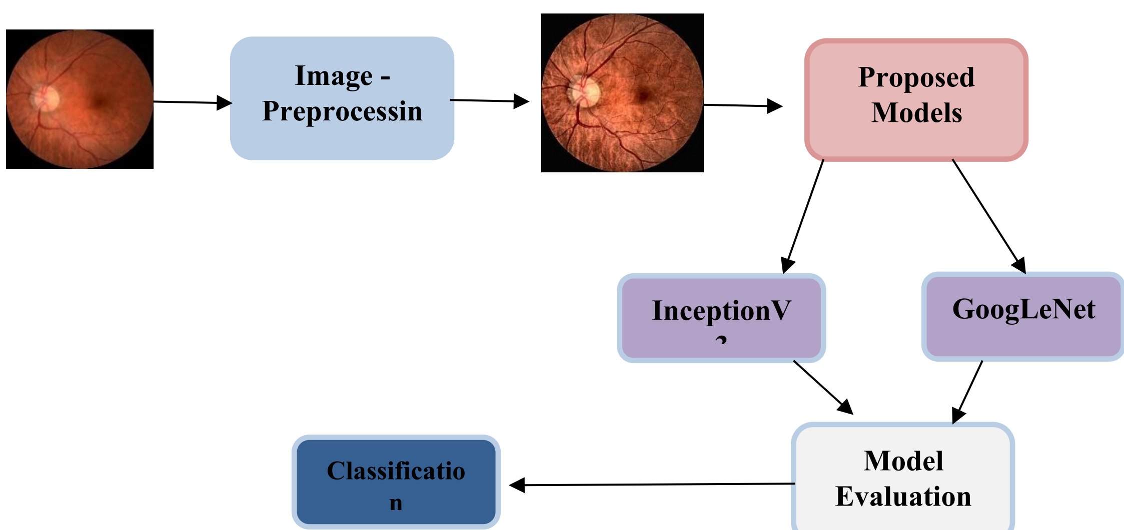



**Fig. 1 - The overall flow chart of the proposed models.**

### 3.1. Dataset

In this study, the Ocular illness Intelligent Recognition (ODIR) dataset [11], one of the largest public datasets for multi-class ocular disease diagnosis was utilized. This datasetsb consists of fundus images from several hospitals in China which was collected by Shanggong Medical Technology Co., Ltd. The ODIR dataset split the images into 7 out of 8 diseases/abnormality classes including diabetes (D), cataract (C), glaucoma (G), age-related macular degeneration (A), myopia (M), hypertension (H) and other diseases/abnormalities (O). METHODS: A data set comprises 5000 CFPs, which is organized into the training (and test) sets in an even portion. Roughly 3,500 instances are used for training, and the rest for testing. The ODIR dataset with some example photos is presented in Figure 2.

(a)Normal (b) Glaucoma

**Fig. 2 - Sample images from the ODIR dataset, Source:ODIR (11)**

### *3.2 Preprocessing Techniques*

The Ocular disease Intelligent Recognition (ODIR) dataset contains fundus images diagnosed with different eye diseases such as diabetes (D), cataract (C), glaucoma (G), age-related macular degeneration (A), myopia (M), hypertension (H), and other abnormalities or diseases (O). A detailed image preparation method was established to reduce noise and enhance the clarity of essential features. Such tasks include edge refinement, noise reduction and contrast enhancement.

Fundus images were improved using Contrast Limited Adaptive Histogram Equalisation (CLAHE) that increase image contrast while keeping the edge details intact. This is advantageous particularly when the lighting and texturing conditions differ between examples. The clarity of the image was further improved by removing noise using Gaussian Blur with 5x5 pixel kernel. We applied a 3x3 kernel sharpening filter to further sharpen the clarity/definition of the image edges by enhancing high-frequency edge components. Upon application of calibrated preprocessing steps, detection of lesions by machine learning algorithms improved substantially by increasing feature clarity and reducing retained noise.

### *3.3 Data Augmentation*

During the network training, data augmentation methods can be used to enhance the dataset by modifying the original data and creating new training examples which were not previously seen. There it is a common practice to use an additive generator whose variedes instances are the different versions of an image where one applies all sort of modification to the original image before giving it to the network. The typical data augmentation techniques are scaling, rotating, flipping, zooming, and injecting Gaussian noise [12]. Data augmentation enhances a model when it is trained with few data and thereby reduces overfitting, especially in high stakes real life applications such as medical datasets. This is very important for the performance of machine learning algorithms when the data amounts are small.

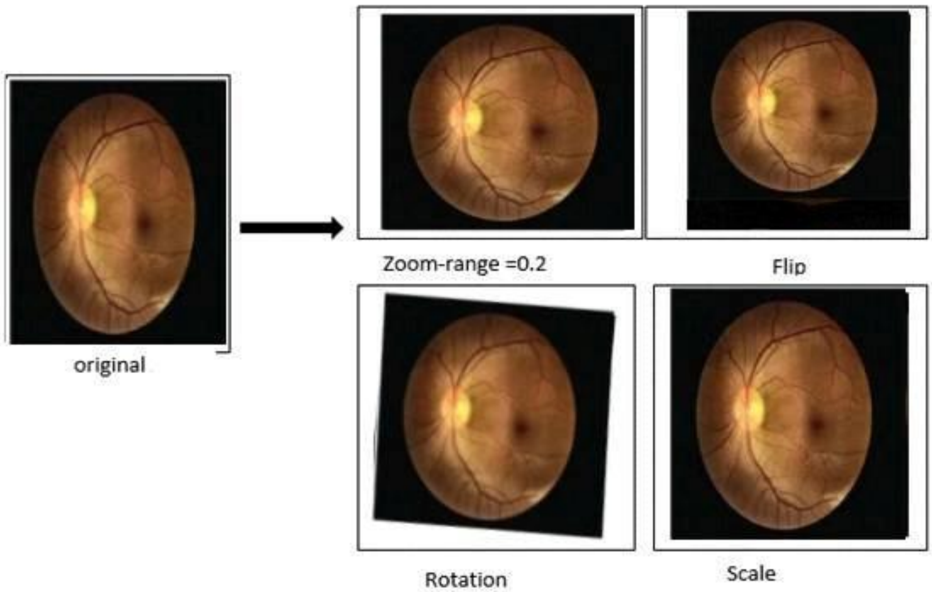



**Fig. 3 - Data Augmentation.**

### *3.4 InceptionV3 model architecture*

In this work, InceptionV3 architecture [ 17 ] is used, which has 23.8 million parameters, and it has a balance between model complexity and computational expenses. InceptionV3 is suitable for learning complex patterns from image data through performing multi-scale feature extraction that is realized by integrating different filter sizes across per Inception module. InceptionV3 has fewer parameters than its predecessors, such as AlexNet and VGG, while offering richer feature representation. It has three sets of Inception layers and two grid-reduction modules and has a total of 350 unique pathways; therefore, it provides strong performance in image classification tasks.

In our execution, the input size 250 × 250 × 3 was adapted in the input layer to match the exact dimensions of the dataset. The global average is used for the last classfication and follows by a fully connected layer. The architecture uses a softmax activation function for multi-class classification. The Stochastic Gradient Descent with Momentum (SGDM) optimizer is also employed for the acceleration of convergence and the co-ordination of the overall training procedure. These modifications along with the optimized architectures have shown the enhanced feature extraction and classification accuracy.

The full training and compilation parameters for InceptionV3 model is provided in Table 1. This architecture works well in challenging machine learning tasks, which involves deep analysis of images such as ocular disease detection by using ODIR dataset.

**Table 1 - Training and Compilation Parameters.**

| Parameter | Typical Value/Setting |
|---|---|
| Batch Size | 32 |
| Optimizer | Sgdm |
| Learning Rate | 0.001 |
| Loss Function | Cross-entropy |
| Epochs | 50 |
| Metric(s) for Evaluation | Accuracy, Precision, Recall, F1 Scor. |
| Validation Split | Typically 0.2 (20% for validation) |

### *3.5 Googlenet model architecture*

We adopt the GoogLeNet architecture to classify eye images in the Ocular Disease Intelligent Recognition (ODIR) dataset. The relatively efficient architecture of GoogLeNet is derived from the "Inception Module" that simultaneously performs parallel convolutions and the corresponding max-pooling to capture both local and global information from input images. The architecture consists of 9 Inception Modules which are able to simultaneously perform the 1×1, 3×3 and 5×5 convolutions and max-pooling in a single layer on the same input. This multi-scale strategy enable GoogLeNet to effectively learn detailed patterns in ocular images with the computational budget.

The 1×1 convolutions utilized in the Inception Modules are especially useful in reducing the dimensionality of input features such that the computational cost is reduced without losing feature richness. Meanwhile, the 3×3 and 5×5 convolutions also squeezer bigger scale patterns in the fundus images, which are useful to recognize the disease-related signs. This structure makes GoogLeNet a good choice for ocular disease classification, as it can effectively handle high-dimension medical images with far fewer parameters than the classical methods like VGG.

We resize each input image to 1\(\times \)224\(\times \)224\(\times \)3 since GoogLeNet requires input of this dimension. The network was adapted to the number of classes in our dataset by substituting the final fully connected layer with one specially designed for our classification problem. Furthermore, the softmax activation layer and classification layer were changed to classify the images adequately. The model convergence was improved using the Stochastic Gradient Descent with Momentum (SGDM) optimizer.

The performance of the model, trained on over-sampled data, was tested by the standard measures such as accuracy, precision, recall and F1 score for overall performance of the model in classification.

## 4. Results and Discussion

We train the proposed models on the Ocular Disease Intelligent Recognition (ODIR) dataset training set and evaluate it on the testing set. The test dataset consists of a wide variety of fundus images from different ocular diseases. Performance metrics e.g.AUC and F1 score derived from the confusion matrix are employed to measure the agreement between the predicted labels and the true labels. AUC measures the balance between sensitivity and specificity of models, and the ROC curve displays the sensitivity versus specificity at all classification thresholds.

Accuracy (1), precision (2), and recall (3) are derived from TP, FN, FP and TN. The F1 Score (4) is the harmonic mean of precision and recall and offers a balanced measurement of models handling performance of false positive and false negative predictions.

$$Acc. = \frac{correct\ predictions\ result\ in\ the}{whole\ number\ of\ results} * 100\% \quad (1)$$

$$\text{Precision} = \frac{TP}{FP+TP} \quad (2)$$

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

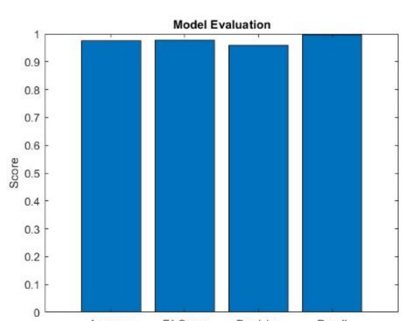
$$\text{F1} = \frac{2(\text{Precision}*\text{recall})}{\text{Precision} + recall} \quad (4)$$

The GoogLeNet and InceptionV3 networks were tested over the Ocular Disease Intelligent Recognition (ODIR) dataset, which includes fundus images of different ocular cases. On examining the results more, it was found that the best model can still be improved by adding more layers directly above the output layer. We have added these dense layers to reduce the dimensionality of the data and added performance. The output layer uses softmax activation function to get the accurate image categories.

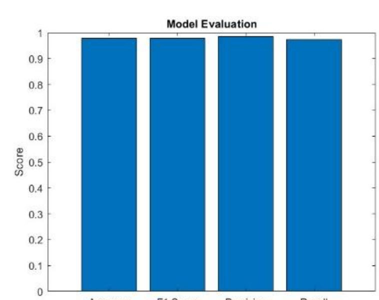
Models were evaluated twice on the ODIR dataset, the original and hacked version of the models were tested. The results in Table 2 and Figure 4 offer an all-round assessment on model performances and have shown the efficacy of the modifications in enhancing the classification accuracy.

**Table 2 -Performance Metrics.**

| Model | Accuracy | Precision | Recall | F1-Score | AUC |
|---|---|---|---|---|---|
| GoogLeNet | 97.64% | 95.96% | 99.53% | 97.72% | 99.33% |
| InceptionV3 | 97.87% | 98.58% | 97.21% | 97.89% | 99.59% |



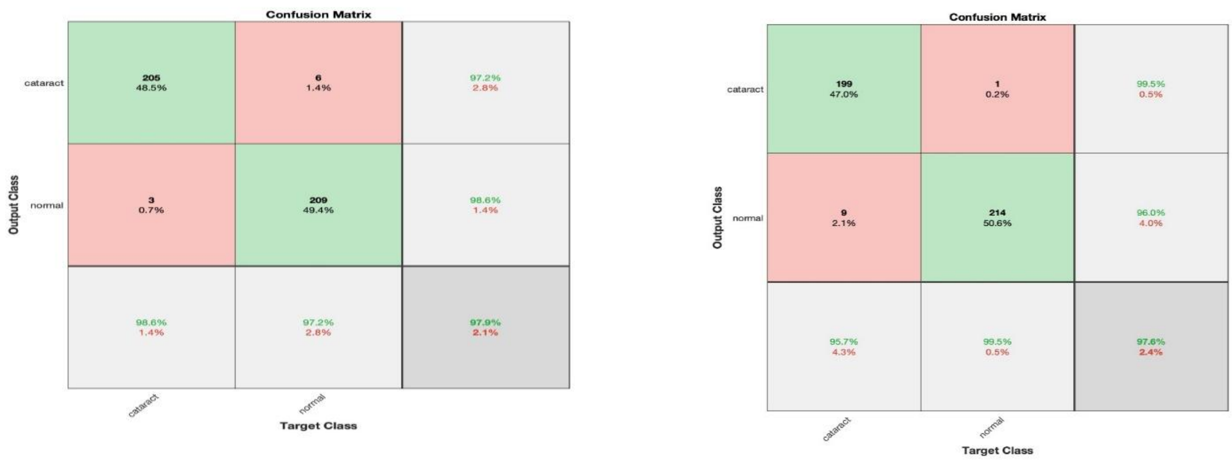

(a) GoogLeNet Performance




(b) InceptionV3 Performance

**Fig. 4 - Custom CNN and modified InceptionV3 Performance**

The Confusion Matrix is a very useful way to evaluate the performance of classification models, but is even more important for pre-trained models such as GoogLeNet and InceptionV3. It provides a breakdown of the models predictions by categories. Comparison of confusion matrices for the proposed models, especially GoogLeNet, InceptionV3 is illustrated in Figure 5.




(a) GoogLeNet (b) InceptionV3

**Fig. 5- Confusion matrix**

### 4.1 Comparison with Previous Works

We compared our results in details with results of state-of-the-art methods, applied on the same dataset utilized in our work for ocular diseases classification. Table 3 presents the results corresponding to injury severity prediction. As it can be seen, our network has achieved 97.64% with the a GoogLeNet and 97.87% with the InceptionV3. This comparison with existing work shows the advances of the efficacy of our models:

**Table 3 - Comparison with Previous Works**

| References | Method | Dataset | ACC |
|---|---|---|---|
| [10] | NB | ODIR | 75% |
| This study | Modified GoogLeNet model | same as used | 97.64 |
| This study | Modified InceptionV3 model | same as used | 97.87 |

## 5. Conclusion

Ocular diseases (such as cataract and glaucoma) are one of the most significant threats towards vision and quality of life, and their prevalence is increasing worldwide at a rapid pace. While some recent works have made progress in the computer-aided detection of these conditions from fundus images, several challenges remain, such as lighting condition variations, color cast differences, and diverse morphologies of ocular abnormalities. In this paper, we examine the use of deep learning based techniques for improving the diagnosis and classification of ocular diseases.

We implemented two deep learning architectures: GoogLeNet and InceptionV3 pre-trained and fine-tuned using X-ray images. We used the Ocular Disease Intelligent Recognition (ODIR) data set on binary classification between cataract vs. healthy eye. Our analysis showed the high potential of these models in reliable detection of the disease.

The accuracy of GoogLeNet model is 97.64%, which demonstrates good precision for its compact structure. The InceptionV3 model, a little more complex, performed better, reaching an accuracy of 97.87%. These results demonstrate the great potential of both hand-crafted and pre-trained deep models, for accurately detecting the ocular diseases.

Our study demonstrates the potential of such models for significant improvement of accuracy and efficiency of ocular disease diagnosis. GoogLeNet also offers an efficient solution that can be implemented even with limited computational resources. On the other hand, the InceptionV3 model delivers the higher accuracy which is also very important for applications that require high accuracy. The current study highlights that deep learning methods have great promise for the diagnosis of ocular diseases, revealing the development and application of clinical diagnosis techniques that are more accurate and effective.

## References

[1] Dipu, N. M., Shohan, S. A., & Salam, K. M. A. (2021). Ocular disease detection using advanced neural network based classification algorithms. *Asian Journal For Convergence In Technology (AJCT) ISSN-2350-1146*, *7*(2), 91-99.

[2] Wahab Sait, A. R. (2023). Artificial Intelligence-Driven Eye Disease Classification Model. *Applied Sciences*, *13*(20), 11437.

[3] Srinivasulu, A., Varaprasad Karu, C., & Sreenivasulu, G. (2024). Automatic Detection and Classification of Eye Diseases from Retinal Images Using Deep Learning: A Comprehensive Research on the ODIR Dataset. *Advances in Engineering and Intelligence Systems*, *3*(01), 45-55.

[4] Kumar, E. S., & Bindu, C. S. (2021). Mdcf: Multi-disease classification framework on fundus image using ensemble cnn models. *Journal of Jilin University*, *40*(09), 35-45.

[5] Khan, M. S., Tafshir, N., Alam, K. N., Dhruba, A. R., Khan, M. M., Albraikan, A. A., & Almalki, F. A. (2022). [Retracted] Deep Learning for Ocular Disease Recognition: An Inner-Class Balance. *Computational Intelligence and Neuroscience*, *2022*(1), 5007111.

[6] Anwer, M. A., Qattan, G. A., & Ali, A. M. (2024). Ocular Disease Classification Using Different Kinds of Machine Learning Algorithms. *Zanco Journal of Pure and Applied Sciences*, *36*(2), 25-34.

[7] Jeny, A. A., Junayed, M. S., & Islam, M. B. (2023). Deep Neural Network-Based Ensemble Model for Eye Diseases Detection and Classification. *Image Analysis and Stereology*, *42*(2), 77-91.

[8] Guron, A. T., Anwer, M. A., Sulaiman, S. K., & AbdulSamad, S. J. (2023). Classification of the cause of eye impairment using different kinds of machine learning algorithms. *Passer Journal of Basic and Applied Sciences*, *5*(2), 410-416.

[9] Bajwa, M. N., Malik, M. I., Siddiqui, S. A., Dengel, A., Shafait, F., Neumeier, W., & Ahmed, S. (2019). Two-stage framework for optic disc localization and glaucoma classification in retinal fundus images using deep learning. *BMC medical informatics and decision making*, *19*, 1-16.

[10] Abed, Z. N., & Al-Bakry, A. M. (2023). DIAGNOSE EYES DISEASES USING VARIOUS FEATURES EXTRACTION APPROACHES AND MACHINE LEARNING ALGORITHMS. *Iraqi Journal for Computers and Informatics*, *49*(2), 44-52.

[11] Li, N., Li, T., Hu, C., Wang, K., & Kang, H. (2021). A benchmark of ocular disease intelligent recognition: One shot for multi-disease detection. In *Benchmarking, Measuring, and Optimizing: Third BenchCouncil International Symposium, Bench 2020, Virtual Event, November 15–16, 2020, Revised Selected Papers 3* (pp. 177-193). Springer International Publishing.

[12] Perez, H., Tah, J. H., & Mosavi, A. (2019). Deep learning for detecting building defects using convolutional neural networks. *Sensors*, *19*(16), 3556.

[13] Hu, K., Zhang, Z., Niu, X., Zhang, Y., Cao, C., Xiao, F., & Gao, X. (2018). Retinal vessel segmentation of color fundus images using multiscale convolutional neural network with an improved cross-entropy loss function. *Neurocomputing*, *309*, 179-191.

[14] Playout, C., Duval, R., & Cheriet, F. (2019). A novel weakly supervised multitask architecture for retinal lesions segmentation on fundus images. *IEEE transactions on medical imaging*, *38*(10), 2434-2444.

[15] Pahuja, R., Sisodia, U., Tiwari, A., Sharma, S., & Nagrath, P. (2022). A Dynamic approach of eye disease classification using deep learning and machine learning model. In *Proceedings of Data Analytics and Management: ICDAM 2021, Volume 1* (pp. 719-736). Springer Singapore.