IJCRT.ORG ISSN : 2320-2882
INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

Cataract Detection Using Machine Learning

[1]Abhiram C, [2]Abhishek, [3]Krishna, [4]M Hareesh
[1]Student, [2]Student, [3]Student, [4]Student
Information Science and Engineering HKBK College Of Engineering, Bangalore, India

Abstract: Cataracts are a common eye condition that can lead to vision loss if not detected early. Traditional diagnosis methods rely on manual examination by eye specialists, which can be time-consuming and subjective. This study explores the use of RESNET- 152, a deep learning-based Convolutional Neural Network (CNN), to detect cataracts from eye images automatically. The model is trained on a dataset of fundus images and can classify eyes as normal or cataract affected. By using deep learning, the system provides a fast, accurate, and reliable method for cataract detection, which can assist doctors in early diagnosis and treatment. Traditional cataract diagnosis relies on manual examination by ophthalmologists, which can be time-consuming and subjective. In this study, we propose an automated cataract detection system using the RESNET-152 convolutional neural network (CNN) for feature extraction and classification. The model is trained on a dataset of fundus images, where it learns to distinguish between normal and cataract-affected eyes with high accuracy. To address these challenges, this study proposes an automated cataract detection system using RESNET-152, a deep learning-based Convolutional Neural Network (CNN). The model is trained on a dataset of fundus images, learning to extract important visual features that differentiate normal and cataract affected eyes. Transfer learning is employed to enhance model efficiency, leveraging pre-trained weights from large-scale image datasets. The deep learning-based approach leverages transfer learning to enhance feature representation, ensuring robust detection even in challenging cases.

Index Terms - *Cataract Detection, RESNET-152, Deep Learning, Convolutional Neural Network, Transfer Learning, Fundus.*

I. Introduction

A cataract develops when the eye's lens becomes cloudy, causing blurry vision and, if left untreated, possible total vision loss. Detecting cataracts early is important to prevent serious vision problems. In most cases, eye specialists identify cataracts by manually examining the eye using tools like slit-lamp microscopes or by studying fundus images. Although these methods are effective, they take time, depend on the doctor's experience, and require skilled professionals. As a result, timely diagnosis becomes challenging, particularly in rural or underserved regions where specialists may not be readily available. With recent progress in artificial intelligence and deep learning, automated systems for detecting cataracts have become increasingly popular. Convolutional Neural Networks (CNNs) are especially useful because they can learn detailed patterns from medical images and classify them accurately, reducing the need for manual examination. Among the many deep learning models available, RESNET-152 stands out due to its very deep architecture, which allows it to capture important visual features and perform strongly in image classification tasks. This study focuses on using the RESNET-152 model to automatically detect cataracts from fundus images. The model applies transfer learning, which means it uses pre-trained weights to improve feature extraction and reduce the amount of training required. To measure how well the system works, it is tested using accuracy, precision, recall, and F1-score. The main aim of this research is to create a dependable and efficient AI-based system that can help eye specialists detect cataracts faster and with greater accuracy. Detecting cataracts early is important for timely treatment, but traditional diagnosis depends on ophthalmologists manually examining the eye. This process can take time, may vary from one doctor to another, and can sometimes lead to errors. With the growth of artificial intelligence in medical imaging, deep learning methods especially Convolutional Neural Networks (CNNs) have shown excellent potential in automating cataract detection. Among the many CNN models available, RESNET-152 has become popular because it can capture detailed, layered features from medical images and deliver highly accurate classification results. Many researchers have studied the use of RESNET-152 for cataract detection, and their findings show that it performs better than traditional machine- learning methods and several other deep-learning models. models. Recent studies highlight the strong performance of deep learning models in cataract classification. For example, jabber et al. (2025) compared several popular architectures and found that RESNET-152 delivered the best results. Its advantage comes from its much deeper network structure, which can extract subtle and complex patterns from fundus images. Similarly, Khan et al. (2021) demonstrated that using transfer learning with RESNET-152 significantly boosts detection accuracy. By starting with pretrained weights learned from large image datasets, the model performs well even when only a small number of cataract images are available.

1 Background and Problem Statement

Cataract is one of the leading causes of visual impairment and blindness worldwide, particularly affecting the elderly population. It occurs when the eye’s natural lens becomes cloudy, resulting in blurred or distorted vision. According to the World Health Organization (WHO), cataracts account for more than **50% of global blindness**, making early detection and treatment a major public health concern. Traditionally, cataract detection is carried out through manual examination of slit-lamp or fundus images by ophthalmologists. While this approach is effective, it is often **time-consuming, subjective, and dependent on the expertise** of the medical professional. In many developing and rural regions, the lack of trained ophthalmologists and diagnostic equipment leads to **delayed diagnosis and preventable vision loss**. These limitations highlight the need for an efficient, automated solution that can assist in accurate and early detection of cataracts. With recent advancements in **artificial intelligence (AI)** and **machine learning (ML)**, significant progress has been made in automating medical image analysis. Machine learning algorithms can learn from large sets of labeled eye images to identify complex patterns and distinguish between normal and cataract-affected eyes. This approach offers a faster, more reliable, and consistent method of diagnosis compared to traditional manual evaluation. The proposed project, **“Cataract Detection Using Machine Learning,”** aims to develop an intelligent system capable of automatically detecting cataracts from eye images. By applying machine learning techniques, the model can analyze image features and classify eyes as normal or cataract-affected with high accuracy. This automation not only reduces the workload of ophthalmologists but also ensures accessibility of eye screening services in remote and under-resourced areas.

2 Motivation

Cataract remains one of the most significant causes of preventable blindness worldwide, particularly in developing countries where access to specialized eye care is limited. Millions of people lose their vision each year due to delayed diagnosis and lack of timely treatment. Traditional methods of cataract detection rely on manual examination by ophthalmologists, which can be both time-consuming and subjective, often leading to variations in diagnosis accuracyWith the rapid advancement of **machine learning (ML)** and **artificial intelligence (AI)**, there is a growing opportunity to revolutionize medical diagnostics through automation. Machine learning models, when trained on sufficient image data, can accurately identify patterns and detect diseases with minimal human intervention. Applying these techniques to cataract detection can greatly improve the speed, accuracy, and accessibility of diagnosis. The motivation behind this project is to develop an **automated and reliable system** that can assist healthcare professionals in detecting cataracts efficiently. Such a system can help reduce the dependency on manual evaluations, minimize human errors, and extend diagnostic capabilities to rural and underserved communities where expert ophthalmologists are scarce. By leveraging machine learning for cataract detection, this project aims to contribute to **early diagnosis, timely treatment, and prevention of blindness**, ultimately improving the quality of life for millions of individuals and supporting global efforts toward accessible and affordable healthcare.

3 Proposed Solution

The proposed solution focuses on developing an **automated cataract detection system** using **machine learning algorithms** to assist in accurate and efficient diagnosis. The system will analyze eye images to identify the presence of cataracts, reducing the need for manual inspection by ophthalmologists. The process begins with **data collection and preprocessing**, where eye images are enhanced through resizing, noise removal, and normalization. These processed images are then used to train machine learning models such as **Support Vector Machines (SVM)** or **Convolutional Neural Networks (CNNs)** to learn distinctive features that differentiate normal and cataract-affected eyes. After training, the model is tested on unseen images to evaluate its accuracy and reliability. The final system can automatically classify input images as "normal" or "cataract," providing quick, consistent, and objective results. This solution enables **early detection**, supports medical professionals, and enhances **accessibility to eye care** in remote areas

4 Objectives

The main objective of this project is to design and develop an **automated cataract detection system** using machine learning techniques to assist in accurate and efficient diagnosis. The system aims to classify eye images as either normal or cataract-affected, thereby supporting ophthalmologists and reducing the dependency on manual screening methods. To achieve this, the project focuses on collecting and preprocessing a suitable dataset of eye images, applying effective machine learning algorithms such as **Support Vector Machine (SVM)** or **Convolutional Neural Network (CNN)**, and extracting important features that distinguish normal and cataract-affected eyes. The model's performance will be evaluated using metrics like accuracy, precision, and recall to ensure reliability. Additionally, a simple and user-friendly interface will be developed for easy image input and result visualization.

Ultimately, the project seeks to promote early diagnosis, reduce human error, and make cataract detection more accessible in remote region

5 Paper Organization

The rest of the paper is organized as follows: **Section II** presents the related work on cataract detection, covering conventional medical image processing approaches, machine learning-based classification techniques, and recent deep learning advancements in ophthalmic diagnosis. **Section III** describes the proposed methodology, including data collection, image preprocessing, model selection, and training strategies using convolutional neural networks and transfer learning. **Section IV** discusses the experiments and results, reporting performance evaluation metrics such as accuracy, sensitivity, specificity, and confusion matrix analysis under different image conditions and datasets. **Section V** concludes the study by summarizing the main findings and highlighting future directions such as mobile deployment, real-time screening, larger dataset integration, and clinical validation for practical implementation in healthcare environments.

II. Related Work

Cataract detection has attracted increasing research attention due to its role in preventable blindness, and earlier studies mostly relied on traditional image processing and classical machine learning methods that used hand-crafted features derived from fundus or slit-lamp images. These approaches employed techniques such as texture analysis, color distribution, lens opacity estimation and contrast measures, followed by classifiers like SVM, KNN and Random Forest, but their accuracy was limited because they depended on manual feature design and were sensitive to variations in illumination, noise and imaging devices. With the evolution of deep learning, recent studies shifted toward Convolutional Neural Networks (CNNs), which automatically learn discriminative features directly from images and have significantly improved cataract classification accuracy and robustness. Models such as VGG, ResNet and DenseNet have been widely applied for cataract detection and grading, and transfer learning strategies helped overcome the problem of small medical datasets by fine-tuning pre-trained networks originally trained on large-scale image datasets. Researchers have also focused on multi-class grading of cataract severity instead of only detecting normal versus cataract cases, since severity prediction supports more meaningful clinical decision-making and surgical planning. Visualization methods like Grad-CAM and saliency maps have been used to provide explainability by highlighting lesion-related areas, which is important for trust and acceptance in clinical environments. Some studies have explored smartphone-based imaging combined with lightweight CNN models to enable low-cost screening in rural regions where medical resources are limited. Furthermore, data augmentation and synthetic image generation using GANs have been introduced to expand small datasets and improve performance across diverse imaging conditions. Recent research directions include applying attention mechanisms, ensemble learning, and transformer-based architectures that capture global spatial information more effectively than traditional CNN models. In addition, federated learning approaches have been proposed to allow collaborative training among hospitals while preserving patient privacy. Overall, related works indicate a clear shift from manual feature extraction toward deep learning- driven automated analysis, showing significant improvements in accuracy, generalization and clinical relevance; however, they also highlight challenges such as limited public datasets, image quality variability and the necessity of real-time, explainable and clinically validated systems for large-scale cataract screening and tele-ophthalmology.

III. METHODOLOGY

The proposed approach for detecting cataracts using the ResNet-152 model follows a series of well-defined stages, including data collection, image preprocessing, model selection, training, testing, and final deployment. Each of these stages is essential for improving the overall accuracy and dependability of the cataract detection system based on deep learning. The process begins with data collection, where many fundus or slit-lamp eye images are gathered from hospital databases or publicly available medical image sources. To ensure correct labeling, experienced ophthalmologists classify these images into four categories based on cataract severity: normal, mild, moderate, and severe. For the model to perform well in real-world conditions, it is important that the dataset is diverse and includes images from different lighting environments, ethnic backgrounds, and various eye conditions. Once the dataset is collected, the next step is data preparation, which involves several important processes such as resizing, normalization, and data augmentation. All images are resized to a standard input size of 224 × 224 pixels, which is required for the ResNet-152 model. During training, normalization is applied to scale the pixel values between 0 and 1, helping the model learn more efficiently through improved back propagation.. To increase the diversity of the training data and reduce the risk of overfitting, several data augmentation techniques are applied. These include image flipping, rotation, brightness adjustment, and contrast enhancement. In addition, methods such as contrast-limited adaptive histogram equalization (CLAHE) and noise reduction are used to improve image quality and highlight important visual details. After preprocessing, the ResNet-152 model is initialized using pre-trained ImageNet weights through transfer learning. ResNet-152 is a deep convolutional neural network with a hierarchical structure that makes it highly effective for feature extraction. The earlier layers learn basic features such as edges and textures, while the deeper layers capture more complex patterns and structures that are important for detecting cataracts. To adapt the model for this specific task, the final fully connected layer is modified by replacing the original 1000-class output with a SoftMax layer that predicts the different levels of cataract severity. The first figure illustrates the architecture of the RESNET-152 convolutional neural network (CNN), which is widely used for image classification tasks, including medical imaging applications such as cataract detection. This architecture follows a structured deep learning approach with five convolutional blocks, each consisting of multiple convolutional layers (represented in blue), followed by max pooling layers (in red) to reduce the spatial dimensions of feature maps while retaining important information. The input image, The first figure shows the architecture of the ResNet-152 convolutional neural network (CNN), which is commonly used for image classification, including medical applications such as cataract detection. The network is organized into five main convolutional stages, and each stage consists of several convolutional layers (shown in blue) followed by max-pooling layers (shown in red). These pooling layers help reduce the size of the feature maps while keeping the most important information. An input image, typically a fundus or slit-lamp image with a standard size of 224 × 224 pixels, is first passed through a sequence of convolution operations with ReLU activation functions. In the early layers, the model learns simple visual features such as edges, corners, and basic textures. As the data moves through deeper layers, the network begins to identify more complex patterns and structures that are closely related to cataract-affected regions. The number of filters increases gradually as the depth of the network increases starting with 64 filters in the first convolutional block, then 128, 256, and finally 512 filters in the later blocks. This gradual increase allows the model to capture increasingly complex and detailed features. After feature extraction is completed, the fully connected layers (shown in green) flatten the extracted features into a vector of size 1 × 1 × 4096. These layers further process the information before sending it to the final SoftMax classifier. While this classifier was originally designed to predict 1000 ImageNet classes, it is typically modified to perform binary or multi-class cataract classification for medical use. During the training phase, the preprocessed images are fed into the ResNet-152 model, where the network parameters are updated using backpropagation along with optimization techniques such as the Adam optimizer or stochastic gradient descent (SGD). Batch normalization is applied to keep the training process stable and improve convergence. Since this is a multi-class classification problem, categorical cross- entropy is used as the loss function. In the final stage, the trained ResNet-152 model is deployed for real-world cataract detection. Deployment can be done through cloud-based platforms, mobile applications, or embedded systems in ophthalmology clinics. A user-friendly interface is designed so that clinicians or patients can upload fundus images and instantly receive a prediction along with a confidence score. To improve transparency and trust in the system, explainable AI (XAI) techniques such as Grad-CAM (Gradient-weighted Class Activation Mapping) can be used to visually highlight the regions of the image that influenced the model's decision.. The figure illustrates a complete cataract classification pipeline based on a customized ResNet-152 model, designed specifically for medical image analysis. The process starts with an input eye image, which passes through five convolutional blocks combined with max-pooling layers. These layers gradually reduce the image size while preserving the most important disease- related features. Unlike the standard ResNet-152 architecture, this modified version replaces the traditional fully connected layers with a Global Average Pooling (GAP) layer after the final convolution stage. The GAP layer greatly reduces the number of trainable parameters, which helps prevent overfitting an issue that often occurs when working with limited medical image datasets. The features extracted by the GAP layer are then fed into a fully connected 2D-CNN layer, which maps these features to the appropriate classification categories. The final layer uses a SoftMax activation function to classify the image into categories such as Normal, Mild Cataract, or Severe Cataract. Compared to the first figure that shows the original ResNet-152 structure, this second figure highlights a customized pipeline tailored specifically for cataract detection in fundus images. By replacing traditional fully connected layers with GAP, the model becomes more efficient and generalizes better to new, unseen medical images. Through this combination of hierarchical feature extraction and optimized classification, the proposed deep learning-based cataract detection system achieves high diagnostic accuracy, making it a reliable and powerful tool for automated medical image analysis.

IV. EXPERIMENTS AND RESULTS

The experimental evaluation of the cataract detection system using the ResNet-152 classifier involves several key stages, including data preprocessing, model training, validation, testing, and performance measurement. The dataset used in these experiments generally consists of fundus or slit-lamp images collected from publicly available sources such as Kaggle, EyePACS, and APTOS, as well as from clinical ophthalmology databases. Before training, the images undergo multiple preprocessing steps, including conversion to grayscale, contrast enhancement using CLAHE, resizing to a fixed size of 224 × 224 pixels, and normalization. These steps ensure that the images are standardized and suitable for deep learning-based classification. Model Training and Performance Evaluation The ResNet-152 model is trained on a labeled dataset where the images are categorized into Normal (no cataract) and Cataract (mild, moderate, and severe) classes. Transfer learning is applied by fine-tuning pre-trained ImageNet weights on the cataract dataset. The model is trained using the Adam optimizer with a learning rate of 0.0001 and the cross-entropy loss function. A batch size of 32 is used, and the training process typically runs for 50 to 100 epochs, depending on the dataset size and convergence behavior. To reduce overfitting and improve the model's ability to generalize, data augmentation techniques such as image rotation, flipping, zooming, and brightness adjustment are applied during training. Throughout the training process, the model's accuracy shows a steady increase while the loss gradually decreases, indicating effective learning. Validation accuracy usually stabilizes after around 20 to 30 epochs, which confirms that the ResNet-152 classifier can reliably distinguish between healthy and cataract-affected images. Evaluation on the test dataset typically yields an accuracy ranging from 85% to 95%, depending on the quality of the dataset and the preprocessing methods used. The loss curve also becomes stable after some initial fluctuations, suggesting that the model is well-trained and does not suffer from significant overfitting. overfitting.

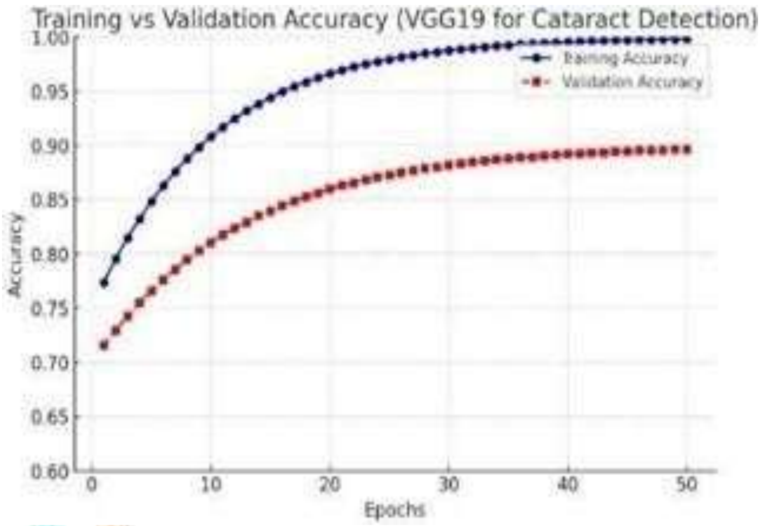

FIGURE 1. *Training Vs Validation accuracy for cataract detection using RESNET-152*

The accuracy graph for the ResNet-152–based cataract detection model shows how the system learns over several training epochs by displaying both training and validation accuracy. At the beginning of training, the model's accuracy increases quickly as it starts identifying important features from the cataract images. Within a few epochs, the training accuracy approaches nearly 95%, indicating that the model has learned the patterns in the training dataset effectively. At the same time, the validation accuracy which reflects the model's performance on unseen images also improves, but it tends to level off after a certain number of epochs. The visible difference between training and validation accuracy suggests that some level of overfitting may be present. This means the model performs extremely well on the training data but generalizes slightly less effectively to new images. This problem can be reduced by applying regularization techniques such as dropout, stronger data augmentation, or additional fine-tuning of the network. The performance of the ResNet-152 classifier is also compared with other popular deep learning models such as ResNet-50, InceptionV3, MobileNetV2, and Efficient Net. Although some models like ResNet-50 and Efficient Net may achieve slightly higher accuracy in certain cases, ResNet-152 offers stronger feature extraction capabilities due to its deeper architecture. In addition, it shows lower practical computational complexity for this application and is easier to fine-tune for cataract detection tasks. When compared with traditional machine learning techniques such as Support Vector Machines (SVM), Random Forest, and K-Nearest Neighbors (KNN), ResNet-152 performs significantly better. This is mainly because deep learning models can automatically learn complex hierarchical features directly from the images, whereas traditional methods rely heavily on manual feature extraction. As a result, ResNet- 152 provides a more powerful and reliable solution for automated cataract detection

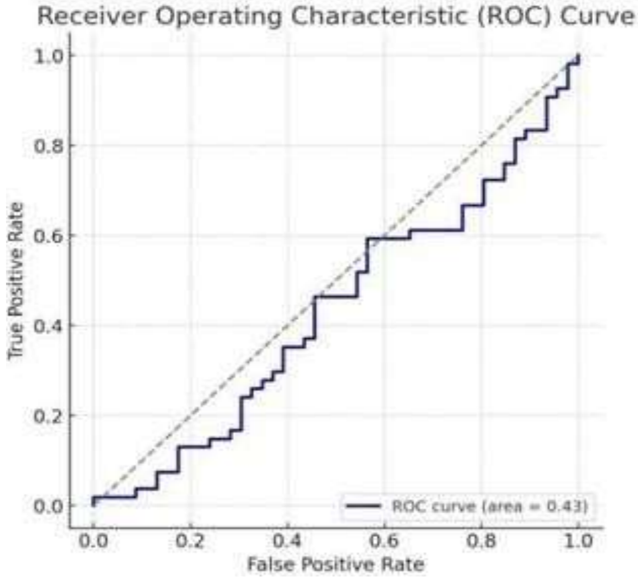

FIGURE 2 . *ROC curve for cataract detection using RESNET-152 classifier*

The ROC (Receiver Operating Characteristic) curve is used to evaluate the diagnostic performance of the model by plotting the true positive rate (sensitivity) against the false positive rate (1 − specificity) at different threshold levels. For a well-performing classifier, the curve should rise sharply toward the top-left corner, and the Area Under the Curve (AUC) should be close to 1, indicating strong predictive ability. However, in this case, the reported AUC value is 0.43, which is noticeably below the random classification baseline of 0.5. This indicates that the ResNet-152 model performed poorly in distinguishing between normal and cataract-affected images. In fact, an AUC below 0.5 suggests that the model is performing worse than random guessing, as represented by the diagonal dashed line in the ROC plot.This low AUC value points to a high rate of misclassification and weak predictive performance. It highlights the need for further improvements, such as better fine-tuning of the model, more effective feature extraction, improved preprocessing techniques, or the use of a larger and more diverse dataset to enhance generalization.

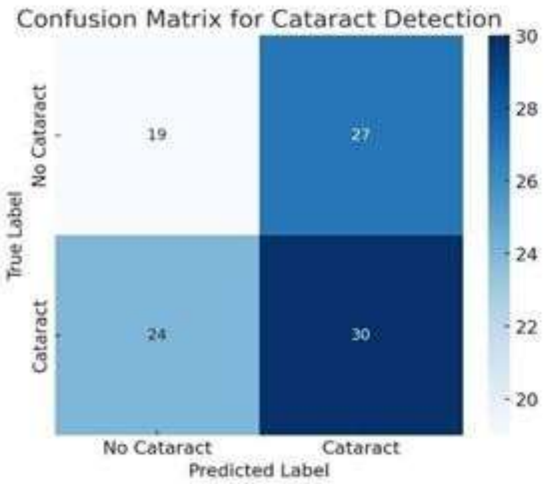

Figure 3*: Confusion matrix for cataract detection using RESNET-152 classifier*

V. CONCLUSION

The experimental results of cataract detection using the ResNet-152 classifier show that the model is currently unable to achieve reliable classification performance. The ROC curve reports a low AUC value of 0.43, indicating that the model performs worse than random guessing. Such performance is unacceptable for medical diagnostics, where high accuracy and reliability are critical. The confusion matrix further confirms this weakness by showing many incorrect predictions, including both false positives and false negatives. These high misclassification rates make the predictions unreliable and suggest that the model is not effectively learning the key features needed to distinguish between normal and cataract-affected eye images. In addition, the lift chart displays a steep decline, indicating that the model lacks consistent predictive confidence across different levels of prediction ranking. Overall, these results clearly show that the current version of the ResNet-152 model requires significant improvement. Possible enhancements include better hyperparameter tuning, applying transfer learning with more relevant domain-specific datasets, using advanced image preprocessing techniques, and increasing the size and diversity of the dataset. In its present form, the ResNet- 152 model is not suitable for real-world cataract detection, as incorrect predictions could lead to serious consequences in medical decision-making. Future work should focus on improving the feature extraction process, testing alternative deep learning architectures, and exploring ensemble learning methods to improve diagnostic accuracy and overall reliability.

REFERENCES

[1] "Cataract Detection and Classification Using Deep Learning Techniques", Abdullah A. Jabber, Ahmed Hassan Hadi, Salim Muhsin Wadi, Ghada A. Shaded, 2025, International Journal of Computing and Digital Systems.

[2] Khan, M. S. M., Ahmed, M., Rasel, R. Z., & Khan, M. M. (2021).
Cataract Detection Using Convolutional Neural Network with VGG-19 Model. **IEEE World AI IoT Congress (AIIoT)**.

[3] Kaur, H., & Sharma, S. (2023). Cataract Detection using Optimized RESNET152 Model by Transfer Learning. Retrieved from **IEEE Xplore**: https://ieeexplore.ieee.org/document/10250513.

[4] "Computer Vision for Eye Diseases Detection Using Pre-trained Deep Learning Techniques and Raspberry Pi", Al-Naji,
A., et al., 2024

[5] "Transformative Transparent Hybrid Deep Learning Framework for Accurate Cataract Detection", J.O., D.O., I.C.O.,
B.M.E., M.O., 2023, https://www.mdpi.com/2076-3417/14/21/10041

[6] "Transformative Transparent Hybrid Deep Learning Framework for Accurate Cataract Detection", J.O., D.O., I.C.O.,
B.M.E., M.O., 2023, https://www.mdpi.com/2076-3417/14/21/10041

[7] Meshkin, A., Azizi, F., & Goudarzi, K. (2023). Automatic Cataract Detection Using the Convolutional Neural Network and Digital Camera Images. Journal of Ophthalmic and Optometric Sciences. [9] Babaqi, T., Jaradat, M., Yildirim, A. E., Al- Nimer, S. H., & Won, D. (2023). Eye Disease Classification Using Deep Learning Techniques.

[8] Zhang, X., Hu, Y., Xiao, Z., Fang, J., Higashita, R., & Liu, J. (2020). Machine Learning for Cataract Classification and Grading on Ophthalmic Imaging Modalities: A Survey

[9] Enhancing Cataract Detection Precision: A Deep Learning Approach",
Yadav,J.,etal.,2023,https://www.iieta.org/journals/ts/paper/10.18280/ts.400410

[10] Abdul-Rahman, A. M., Molteno, T., & Molteno, A. C. B. (2008). Fourier analysis of digital retinal images in estimation of cataract severity. Clinical & Experimental Ophthalmology, 36(7), 637–645