# Iris Tumour Detection Using Convolutional Neural Networks (CNN)


Gaana Shree S
*Dept. of CSE(AI-ML)*
Dayananda Sagar University
gaanashrees25@gmail.com

Fariha Naaz
*Dept. of CSE(AI-ML)*
Malla Reddy Engineering college for women
farihanaaz2004@gmail.com

Yogesh Rai
*Dept. of CSE*
Graphic Era Hill University
yogeshrai9413@gmail.com

Gayathri Tati
*Dept. of CSE*
Prasad V Potluri Siddhartha Institute of Technology
gayathriatati367@gmail.com

Charmika Mittapally
*Dept. of IT*
Malla Reddy Engineering College for women
charmikamittapally@gmail.com



***Abstract*—Iris tumors are among the smallest cystic tumors associated with cancer, making their identification particularly challenging due to their diminutive size. If not addressed in the early stages, these tumors can lead to irreversible blindness. To facilitate the diagnosis of iris tumors, convolutional neural networks (CNNs) are utilized. The model is trained with images from both healthy individuals and those with tumors, which undergo preprocessing prior to analysis. Preprocessing involves reshaping and standardizing the images into a particular format. The convolutional layers assign weights to the input pixels to determine whether the image represents a healthy or affected eye. The use of CNNs allows for the extraction of intricate features from the input images, enabling the detection of subtle abnormalities that might be overlooked by manual examination. Advanced data augmentation techniques are employed during training to improve model robustness and ensure reliable performance across diverse image variations. The dataset is classified into 2 categories (tumor or no tumor). Each category is labeled before it is fed to the model for training. Furthermore, the integration of this technology into clinical workflows can provide ophthalmologists with real-time diagnostic support, paving the way for early intervention and improved patient outcomes. This research involves the digital classification of iris tumor images and identifying the best CNN model with higher accuracy.**



***Index Terms*—CNN -Convolution Neural Network, ECG -Electrocardiogram EfficientNet-B0, MobileNetV3-Large and GoogLeNet**


## I. Introduction

The diagnosis of iris tumors is a critical aspect of ophthalmic healthcare. Any delays in the detection process could result in irreversible blindness. Our model helps in preventing such scenarios by detecting abnormalities in iris images and classifying them using various machine learning techniques. The raw iris images often contain noise due to factors such as lighting conditions, camera angles, and image quality. The dimensions of these datasets are also not standardized. In order to obtain optimal images for training the model, applying transforms plays a major role. Transforms help in resizing the images into equal dimensions, converting these images into numerical values, and normalizing them.

Post-transformation, the dataset is labeled into two categories: tumor or no tumor. Once the data is preprocessed into the required format, the ResNet-18 model is trained with this dataset. ResNet-18 is a pre-trained model based on a convolutional neural network that consists of 18 layers for image classification purposes. This trained model is then tested and evaluated based on metrics such as F1-score, recall, and accuracy to improve the model's performance. This model can further be implemented in clinical devices or applications to assist ophthalmologists in early diagnosis, which helps reduce the risk of irreversible conditions and improves patient outcomes.

## II. Related Works

Iris tumor detection has become increasingly reliant on automated systems to enhance diagnostic accuracy and efficiency, addressing the challenges posed by the small size of iris tumors. Image processing techniques, such as converting RGB images to grayscale, median filtering, and background extraction, are crucial for improving image quality and isolating regions of interest. Median filtering is particularly effective in reducing impulsive noise while preserving important features, a technique commonly used in medical imaging (Gonzalez and Woods, 2008). Once the image is enhanced, edge detection methods like the Canny algorithm are employed for segmentation, identifying tumor boundaries by detecting intensity discontinuities (Canny, 1986). These preprocessing steps ensure the accurate localization of the tumor, essential for effective classification.The final step of the process involves image fusion, where the tumor region is superimposed onto the original image to clearly mark the affected area. This approach has been successfully used in medical image analysis to combine features from different sources and highlight abnormalities. By automating this process with machine learning models, iris tumor detection systems can provide real-time diagnostic support, enabling early intervention and improved patient outcomes. Integrating such systems into clinical workflows could significantly reduce diagnostic delays, ultimately improving the prognosis for patients with iris tumors [1].

The Eye Tumour Detection System (ETDS) utilizes image processing techniques to detect tumours in eye images. The process begins by converting the original image to grayscale, reducing computational load, followed by filtering to remove noise and enhance key features. A median filter is applied to preserve edges while reducing impulsive noise. Morphological operations, such as dilation and erosion, are then used to refine the image by extracting the background and enhancing object boundaries. Image enhancement techniques like image addition and adjustment further improve the brightness and contrast, highlighting the tumour area. Edge detection is performed using the Canny edge detector, which identifies intensity discontinuities for accurate segmentation of the tumour region. To finalize the process, pixel-level image fusion combines the segmented image with the original to create a clearer, more informative output. This comprehensive approach ensures that the ETDS efficiently detects eye tumours by providing clear, enhanced images for further analysis [2].

This section highlights several key image processing techniques, focusing on wavelet transforms, denoising, and compression. Wavelet transforms, particularly the Discrete Wavelet Transform (DWT) and Discrete Wavelet Packet Transform (DWPT), are used to decompose images into frequency bands for multi-resolution analysis. DWT decomposes an image into four subbands, with further iterative filtering of the low-frequency subband, while DWPT offers enhanced time-frequency resolution through redundant basis functions. These methods are effective in applications such as structural damage assessment, where dynamic signals are decomposed and analyzed to identify damage characteristics. Wavelet-based denoising, on the other hand, transforms a noisy image using DWT, applies thresholding techniques, and then uses inverse DWT to restore the image, balancing noise removal with detail preservation. In addition to these techniques, image compression and global thresholding are discussed. Lossless compression methods are employed to retain the original data, with coding methods and transform domain techniques used for efficient compression. Global thresholding, a segmentation technique based on the image histogram, involves iterating to find an optimal threshold that divides pixels into two groups based on their grey levels. This approach works well when the histogram is suitably partitioned. Together, these methods provide a comprehensive approach for image analysis, denoising, and compression, making them valuable for a range of applications where image quality and clarity are paramount [4].

Numerous alternative models exist for similar objectives. EfficientNet represents a deep learning architecture that standardizes input dimensions for classification tasks. This model integrates an efficient net for feature extraction alongside a prediction block that categorizes images into various labels. The efficient net component is constructed from 16 iterations of MobileNet, a computer vision model recognized for its superior image classification performance. This architecture facilitates extensive feature extraction from large datasets. The prediction block subsequently classifies the extracted features into 24 distinct classes, each associated with unique labels. These labels are then utilized as training data for the model, while the original labels are applied to other datasets. Ultimately, the model's performance was assessed using the Adam optimizer, achieving an accuracy rate of 91.5 percent. Additionally, comparisons were conducted among various CNN models, including AlexNet, GoogLeNet, and SqueezeNet. AlexNet is a deep learning model designed for multi-class classification, featuring convolutional layers and rectified linear units to handle non-linear data. GoogLeNet is adept at processing data with significant scalability and incorporates a continuous wavelet transform to represent ECG signals in two dimensions. Upon evaluation, the models achieved classification accuracies of 97.8, 97.78, and 97.2 percent, respectively [5].

## III. Methodology

This process begins with gathering a dataset from Mendeley Data containing classified ECG images. Followed by data pre-processing, this involves normalising and standardising data. The dataset is labelled into four categories. After labelling, the dataset is combined and divided into training and testing datasets with a ratio of 80:20. The training data is fed into various pre-trained models to compute the best model. During the training phase, the model can improve its performance by using a cross-entropy loss function and be optimised using the Adam optimizer. Based on evaluation metrics like recall, precision, and accuracy, a model's performance is determined.

### A. Dataset

The data set consists of images categorized into two classes: Tumor and No Tumor. The data set is real-time data collected from Miles Research by Jon Miles. The images had to be segregated before preprocessing and fed into the model. A custom PyTorch dataset class is implemented to load the data efficiently. The dataset paths are defined, and images are read along with their corresponding labels (1 for Tumor and 0 for No Tumor). The dataset is augmented with preprocessing techniques to ensure consistent input for the model.

Fig. 1. Eyes with Tumor

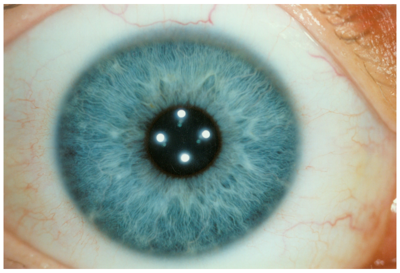

Fig. 2. Healthy eye

### B. Data pre-processing

Preprocessing is critical to enhancing image features and reducing noise. Grayscale Conversion: Reduces the image to a single intensity channel, focusing on structural and intensity-based features. Median Filtering: Removes noise while preserving edges, ensuring clearer and sharper features. Canny Edge Detection: Highlights boundaries and sharp transitions, emphasizing tumor edges or features. Resizing: All images are resized to (224, 224) to meet ResNet18's input size requirements. Normalization: Images are normalized using the mean and standard deviation of the ImageNet dataset to improve model compatibility and convergence. The processed images are converted to 3-channel format (RGB) to align with the ResNet18 input format.

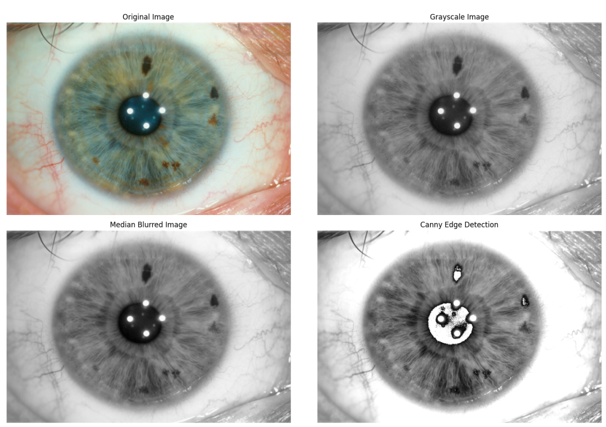


Fig. 3. Images after prepreocessing

### C. Test-Train-Validation Split

The combined dataset is divided into training, validation, and testing subsets to ensure the model's effective learning and evaluation. The training set, comprising 80 percent of the data, is used to train the model by optimizing its weights and learning patterns. The **validation set**, consisting of 20 percent of the data, is used to evaluate the model's performance after each training epoch, ensuring it generalizes well and does not overfit the training data. An optional **test set** could be employed for final evaluation on entirely unseen data to assess the model's real-world applicability.

### D. Model development

Multiple models are trained and evaluated to achieve better accuracy and performance. The following models are used to classify Iris tumor images.

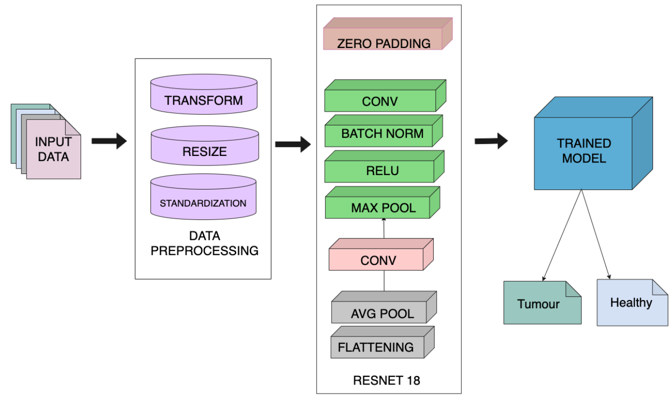


Fig. 4. Work flow diagram

*1) EfficientNet-B0:* EfficientNet is capable of providing high accuracy and better computational efficiency. This model can work on a large dataset and works well even if noise exists in the dataset. The time-sensitivity of EfficientNet leads to the analysis of micro-scale features in dataset.

*2) GoogLeNet:* The GoogLeNet model is based on inception architecture, in which hidden layers can use multiple filters without alteration in the depth of the neural network. In this model, larger layers are replaced with smaller individual layers for easy propagation of data in the hidden layer, due to this model's efficiency is increased.

*3) Resnet-18:* The ResNet-18 model architecture has 18 residual layers that can overcome vanishing gradient issues in neural networks. Residual layers help in skipping certain intermediate layers, due to which the input remains almost the same throughout the model. It is best to analyse both long-term and short-term features, as iris image contain minute variations this feature can capture it.

### E. Optimization and Entropy

To improve the accuracy and performance of our model, we have used the Adam optimizer, as this can overcome sparse gradient issues. It helps the model to converge at a faster rate. All hyperparameters are fine-tuned by using this optimizer. This study focuses on Binary classification, and predicting dissimilarities among them is important. This is achieved using a cross-entropy loss function.

## IV. Results and Discussion

*a) Result:* The performance of three deep learning models—GoogLeNet, EfficientNet-B0, and ResNet18—was assessed using key metrics: accuracy, recall, and F1-score. Among these, ResNet18 demonstrated the highest performance, achieving an accuracy of 94.44 percent a recall of 87.50 percent, and an F1-score of 93.33 percent. These results highlight ResNet18's ability to extract and leverage critical features effectively, making it highly suitable for the task of iris tumor classification. Its robustness in maintaining a balanced precision-recall tradeoff further supports its reliability in identifying tumor cases accurately while minimizing false negatives and false positives.

TABLE I
TABLE OF RESULTS

| Model | Metrics | | |
|---|---|---|---|
| | ***Accuracy*** | ***Recall*** | ***F1-score*** |
| GoogLeNet | 0.8333 | 0.8333 | 0.7694 |
| EfficientNet-B0 | 0.8888 | 0.8571 | 0.8571 |
| ResNet18 | 0.9444 | 0.8750 | 0.9333 |

EfficientNet-B0, with an accuracy of 88.88 percent, a recall of 85.71 percent, and an F1-score of 85.71 percent, performed slightly lower than ResNet18 but remains competitive due to its efficiency in computational resource usage. Conversely, GoogLeNet, achieving an accuracy of 83.33 percent, a recall of 83.33 percent, and an F1-score of 76.94 percent, showed limitations in feature extraction for this specific dataset, suggesting the need for further optimization. Overall, ResNet18 is identified as the most suitable model for this classification task, with EfficientNet-B0 serving as a viable alternative for resource-constrained scenarios.

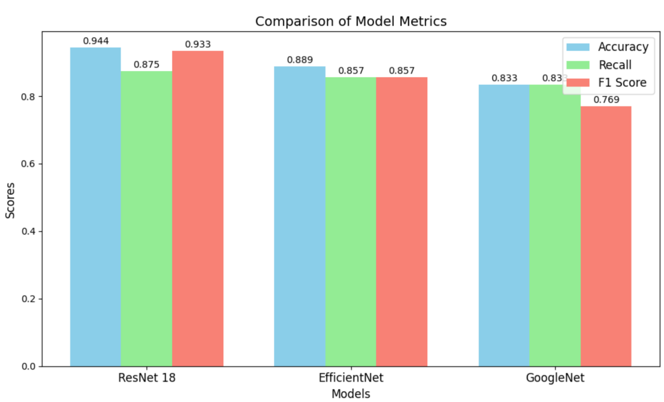


Fig. 5. Model performance

## V. CONCLUSION AND SCOPE FOR FUTURE ENHANCEMENT

*a) Conclusion:* This study demonstrated the effectiveness of deep learning models in the classification of iris tumor images. Among the evaluated models, ResNet18 proving to be the most reliable for this task. EfficientNet-B0 also exhibited competitive performance, making it a suitable alternative for environments with computational constraints. The results highlight the importance of leveraging transfer learning and robust preprocessing techniques, such as grayscale conversion, median filtering, and edge detection, to improve feature extraction and enhance model performance.

*b) Scope for Future Enhancement:* Future work can focus on improving the robustness and scalability of the model by incorporating larger and more diverse datasets. Exploring other state-of-the-art architectures, such as Vision Transformers (ViT) or advanced EfficientNet variants, could further enhance accuracy and efficiency. Additionally, integrating explainable AI techniques could provide interpretability, allowing medical practitioners to understand and trust the model's predictions. Deploying the model on real-time systems, such as edge devices or cloud-based platforms, could extend its utility in clinical settings. Finally, expanding the scope to multi-class classification for detecting other types of tumors or related conditions could significantly broaden the model's impact in medical diagnostics.

## ACKNOWLEDGMENT

For their helpful advice and assistance during this study, We would like to express our sincere gratitude to Infosys Springboard for continuous support and their helpful criticism and continuous support during this research project. We extend our heartfelt thanks to our family and friends for their unwavering encouragement, support, and patience, which were instrumental in completing this study.

## REFERENCES

[1] Abdulkader Helwan, ITDS: Iris Tumor Detection System using Image Processing Techniques, International Journal of Scientific and Engineering Research, Volume 5, Issue 11, November-2014 ISSN 2229-5518

[2] Kamil Dimililer, Yoney Kirsal Ever, Haithm Ratemi, Intelligent eye Tumour Detection System, Procedia Computer Science, Volume 102,2016, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2016.09.408.

[3] Foran DJ, Chen W, Kurc T, Gupta R, Kaczmarzyk JR, Torre-Healy LA, Bremer E, Ajjarapu S, Do N, Harris G, Stroup A, Durbin E, Saltz JH. An Intelligent Search and Retrieval System (IRIS) and Clinical and Research Repository for Decision Support Based on Machine Learning and Joint Kernel-based Supervised Hashing. Cancer Inform. 2024 Feb 4;23:11769351231223806. doi: 10.1177/11769351231223806. PMID: 38322427; PMCID: PMC10840403.

[4] M. Saravanan, D. Kalpanapriya. A Novel Iris Cancer Detection using wavelet and wavelet packet Transform. Research J. Pharm. and Tech 2018; 11(7): 2926-2931. doi: 10.5958/0974-360X.2018.00540.1

[5] Nitin Rahuja, Sudarshan K. Valluru, A Comparative Analysis of Deep Neural Network Models using Transfer Learning for Electrocardiogram Signal Classification, 2021 IEEE International Conference on Recent Trends in Electronics, Information and Communication Technology (RTEICT).

[6] Deep Learning in Opthalmology: Iris Melanocytic Tumors Intelligent Diagnosis, Abdulkader Helwan medRxiv 2021.09.14.21263573; doi: https://doi.org/10.1101/2021.09.14.21263573 Volume 93, July 2024, 106141

[7] Y. Pan, J.W. Phillips, B.D. Zhang, M. Noguchi, E. Kutschera, J. McLaughlin, P.A. Nesterenko, Z. Mao, N.J. Bangayan, R. Wang, W. Tran, H.T. Yang, Y. Wang, Y. Xu, M.B. Obusan, D. Cheng, A.H. Lee, K.E. Kadash-Edmondson, A. Champhekar, C. Puig-Saus, A. Ribas, R.M. Prins, C.S. Seet, G.M. Crooks, O.N. Witte, Y. Xing, IRIS: Discovery of cancer immunotherapy targets arising from pre-mRNA alternative splicing, Proc. Natl. Acad. Sci. U.S.A. https://doi.org/10.1073/pnas.2221116120 (2023).

[8] Mitamura M, Kase S, Suimon Y, Ito T, Ishida S. A Case of Metastatic Iris Tumor Observed With Anterior Segment Optical Coherence Tomography Before and After Radiation Therapy. In Vivo. 2020 Jul-Aug;34(4):2159-2162. doi: 10.21873/invivo.12023. PMID: 32606198; PMCID: PMC7439902.

[9] Iris Detection Using Segmentation Techniques, Dr. Mohammad Shabaz, International Journal of Computer Sciences and Engineering, 2018

[10] Sneha Jayant Patil, and Prof. R. D. Patil. (2021). DETECTION IRIS MELANOMA IN HUMAN EYE. International Journal of Innovations in Engineering Research and Technology, 1-6. https://repo.ijiert.org/index.php/ijiert/article/view/1072