Glaucoma Stage Classification Using Image Empirical Mode Decomposition (IEMD) and Deep Learning from Fundus Images

D. Shankar[1(✉)], I. Sri Harsha[1], P. Shyamala Madhuri[1], J. N. S. S. Janardhana Naidu[1], P. Krishna Madhuri[1], and Srikanth Cherukuvada[2]

[1] Department of Computer Science and Engineering, Vishnu Institute of Technology, Kovvada, India
shankar.d@vishnu.edu.in

[2] Department of Networking and Communications, School of Computing, SRM Institute of Science and Technology, Chennai, India

Abstract. Glaucoma is an ocular pathology characterized by the gradual deterioration of neural cells in the eye, which is attributed to elevated intra ocular pressure within the retina. Glaucoma takes the second spot in terms of its prevalence as a neurodegenerative eye disease, failure to diagnose glaucoma at an early stage can lead to complete blindness. This underlying issue requires a streamlined system that uses experienced medical experts, little equipment, and less time. To categories the stages of glaucoma, a Computer-Aided Diagnosis (CAD) method is used called Image Empirical Mode Decomposition (IEMD). Segmentation-based algorithms utilizing features like cup-to-disc ratio (CDR) and textural characteristics close to the optic disc region can distinguish glaucoma. To capture pixel variations for this investigation, the pre-processed fundus images are separated and transformed into a diverse range of intrinsic mode function (IMFs). The study employs deep learning-based framework for finding the optic nerve head from coloured fundus photographs, to transform the fundus images into a delimited region of interest (ROI) and utilizes multiple deep networks like RFCN and RCNN classifiers identified glaucoma stages separately. The CAD system serves as an automated tool for retinal image processing and demonstrates superior performance over the RFCN classifier, achieving an impressive accuracy of 96% with the RCNN classifier in classifying glaucoma stages.

Keywords: glaucoma · intraocular pressure · retina · computer-aided diagnosis · image decomposition · fundus images · optic nerve head · deep learning · glaucoma classification · optic disc · ensemble classifiers · RFCN · RCNN

P. Pareek et al. (Eds.): IC4S 2023, LNICST 536, pp. 385–402, 2024.
https://doi.org/10.1007/978-3-031-48888-7_33

1 Introduction

This study is centered on the application of convolutional neural networks (CNN) for the classification of glaucoma stages. After careful consideration, the researchers decided to adopt the RCNN (Region-based CNN) and RFCN (Region-based Fully Convolutional Network) approaches. We found that RCNN and RFCN provide the necessary flexibility, reliability, cost effectiveness, and efficiency required for the research.

Glaucoma is a sight-threatening disorder affecting the optic nerve and RNFL. It can be asymptomatic in early stages, leading to unnoticed vision changes. Fundus imaging helps diagnose and monitor glaucoma, capturing images of critical eye structures to enable early detection and intervention [1] (Fig. 1).

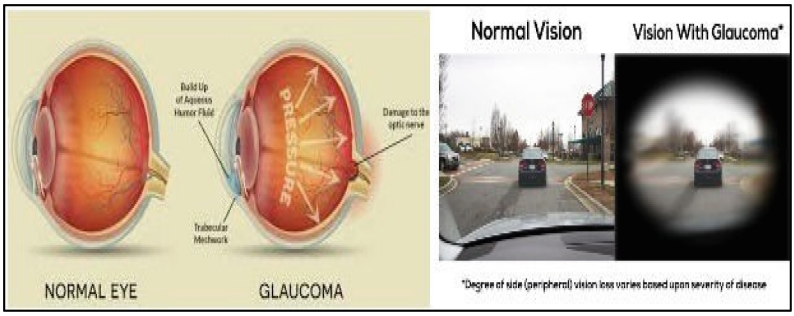

Fig. 1. Fundus Images with Normal and Glaucoma Vision

An extensive array of diverse techniques and algorithms has been developed to perform automated analysis of retinal images and extract blood vessel segments. These methods can be divided into two categories: supervised and unsupervised approaches. Blood vessels are generally visible as borders or outlines inside retinal pictures. Unsupervised methods require the use of contour models to detect and identify these vessels. This category includes techniques such as matching filter responses, morphological approaches, and deformable models. It is difficult to develop an economical and high performing segmentation approach for combined segmentation of the optic disc (OD) and optic cup (OC) [2, 3].

To solve the issues, we devised a customized Computer-Aided Diagnosis (CAD) approach utilizing retinal fundus images designed for clinical evaluation. The development of the system mainly concentrates on the use of image based empirical mode decomposition technique for image processing, RCNN and RFCN deep learning model which have low computational overhead and it is also accurate and efficient. This study aims to design a system that can accurately classify glaucoma stages based on fundus images while minimizing resource requirements. By achieving maximum accuracy, the system outcomes will assist in taking timely precautionary actions against the disease [4, 5].

2 Related Work

Segmenting the cup and disc regions in fundus images is a crucial step in detecting glaucoma. Researchers have explored various techniques to achieve accurate segmentation.

Deepak Parashar [6] a novel methodology was developed to classify glaucoma, with a focus on utilizing the 2D-TEWT for analysis. The provided images underwent preprocessing and were subsequently separated into sub-band images (SBIs) through the utilization of 2D-TEWT. These images were classified using a trained multi-class least squares support vector machine (MC-LS-SVM) classifier. Using tenfold cross-validation, the approach achieved a phenomenal classification accuracy of 93.65%.

Shubham Joshi [7] to aid in the early detection, screening, and treatment of glaucoma, the proposed CAD system makes use of three CNN models. By classifying glaucoma according to precise criteria, these models make it possible for prompt diagnosis, efficient screening, and customized therapy recommendations are used. This system identifies normal and abnormal digital fundus glaucoma images. By using the U-Net models with CNN layers it is getting accuracy of 91.11%.

Huazhu Fu [8] evaluated methods for glaucoma screening using automated processes; two deep learning-based approaches, M-Net and DENet, were examined. M-Net employs a cohesive structure to simultaneously categorize the optic disc (OD) and optic cup (OC) in a solitary stage. In contrast, DENet integrates multiple deep streams with varying levels and modules to directly predict glaucoma from fundus images.

Ratuja Shinde [9] The proposal entails the integration of a desktop application into an automated computer aided design (CAD) system that aims to facilitate detection of glaucoma through the examine of retinal fundus images. The programme incorporates the deep learning strategy known as LeNet in order to verify the input photos. An approach that relies on the brightest point is presented, and it is used to extract the area of interest (ROI). To enhance the precision of region of interest (ROI) identification, this approach uses a preprocessing stage in which an image is first gray scaled and then a Gaussian blur is applied.

Ramgopal Kashyap [10] To revolutionize glaucoma diagnosis and prediction, a cutting edge approach based on deep learning is put forward. The strategy involves the utilization of deep learning concepts to address optic cup segmentation, which entails the integration of pre trained transfer learning algorithms with the U-Net architectural style. In order to extract features, the DenseNet-201 deep convolutional neural network (DCNN) is utilized. By leveraging the power of DCNN, the presence of glaucoma can be accurately determined. The implications of this study extend to various imaging modalities, making it a promising avenue for future research.

3 Glaucoma Classification Methodology

3.1 Dataset Description

3.1.1 Training Dataset

Here, we have the complete training dataset. We can extract features and train to fit a model and so on. The findings from the learning algorithms training set gains knowledge from the experience encountered during tasks with supervised learning, where each instance includes an dependent variable in observation along with One or multiple recorded input variables. The size of the training dataset is 520.

3.1.2 Testing Dataset

It is used to test if our model is working accurately. Here, once you have obtained the model, you can use it to make predictions on the training set. The test set is an assembled group of data points used to evaluate the efficacy of the model against predefined criteria. It is vital to check that no samples from the training set made it into the exam data. It's more difficult to tell whether an algorithm has learnt to generalize from training data when the test set incorporates instances from the training set. Within our dataset, which consists of numerous records, we have allocated 80% using the data to train the mode, while the remaining 20% will be utilized for evaluating the models performance. The size of the test dataset is 130.

3.2 Pre-Processing

3.2.1 IEMD

Image Empirical Mode Decomposition is an approach to image analysis that was developed from the concept of empirical mode decomposition in Hilbert-Huang transform (HHT). This strategy was modified so that it could be applied to images. It is possible to apply IEMD to image processing jobs after increasing its scope to include two dimensions. The capacity of IEMD to efficiently separate locally overlaid frequency ranges within an image is what sets it apart as a tool with a distinctive set of capabilities (Fig. 2) and (Fig. 3).

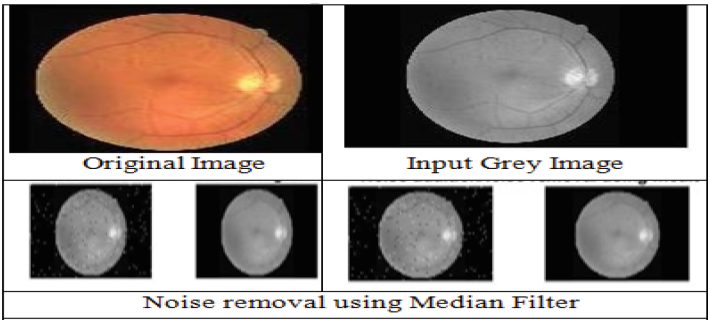

Fig. 2. Original and Grey Images

3.2.2 CLAHE

Throughout the previous few decades, there has been considerable progress in the evolution of contrast enhancement algorithms. These algorithms primarily aim to achieve two main objectives: enhancing the visual appearance of an image for improved interpretation and enhancing the efficacy of subsequent endeavors such as image segmentation, perception of object, and image analysis.

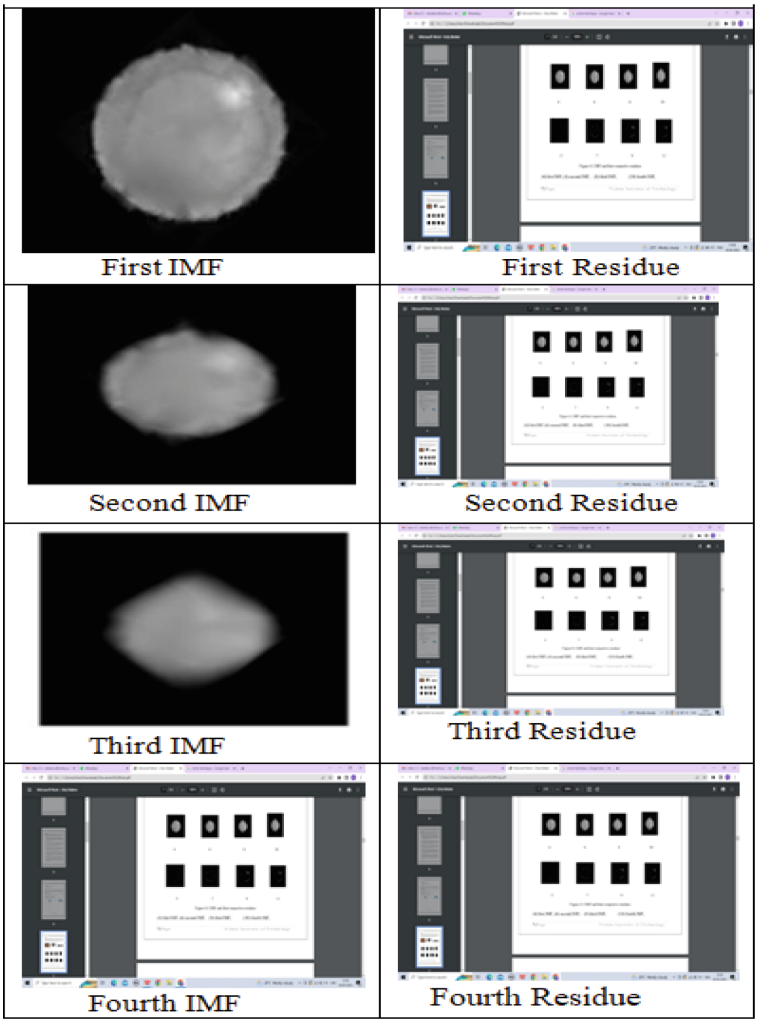

Fig. 3. IMF and their residues

Histogram adjustments, either done worldwide or regionally, are the foundation of contrast improvement approaches. To get over these worldwide restrictions, we present the Contrast Limited Adaptive Histogram Equalization (CLAHE) technique [11]. The Contrast Limited Adaptive Histogram Equalization (CLAHE) algorithm performs localized processing on image regions known as tiles, thereby mitigating the risk of excessive amplification. Additionally, the algorithm employs interpolation techniques to seamlessly merge adjacent tiles. The approach presented herein yields a notable enhancement in image contrast and is amenable to colour images, with a particular emphasis on the brightness channels in HSV images. CLAHE outperforms equalization of all channels in BGR images (Fig. 4).

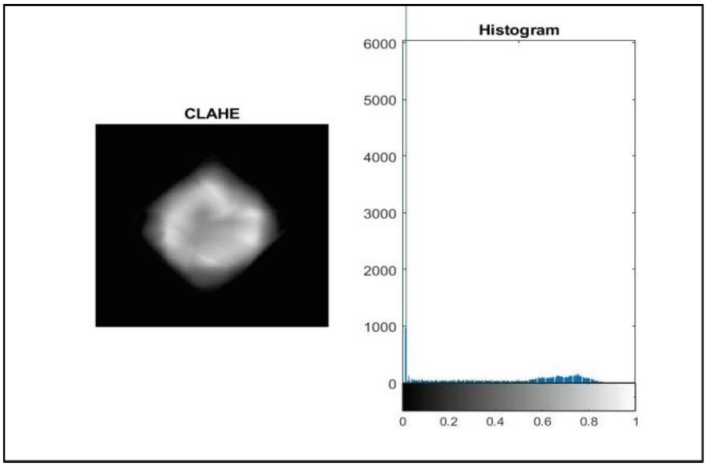

Fig. 4. CLAHE

3.2.3 Histogram

The histogram of an image is a visual representation that shows the distribution of pixel intensities. Essentially, it depicts the frequency of occurrence for each intensity value present in the image (Fig. 5) and (Fig. 6).

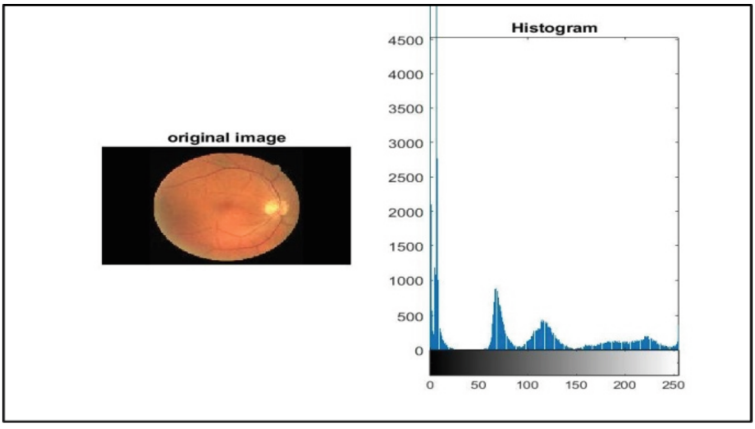

Fig. 5. Histogram of Original Eye Image

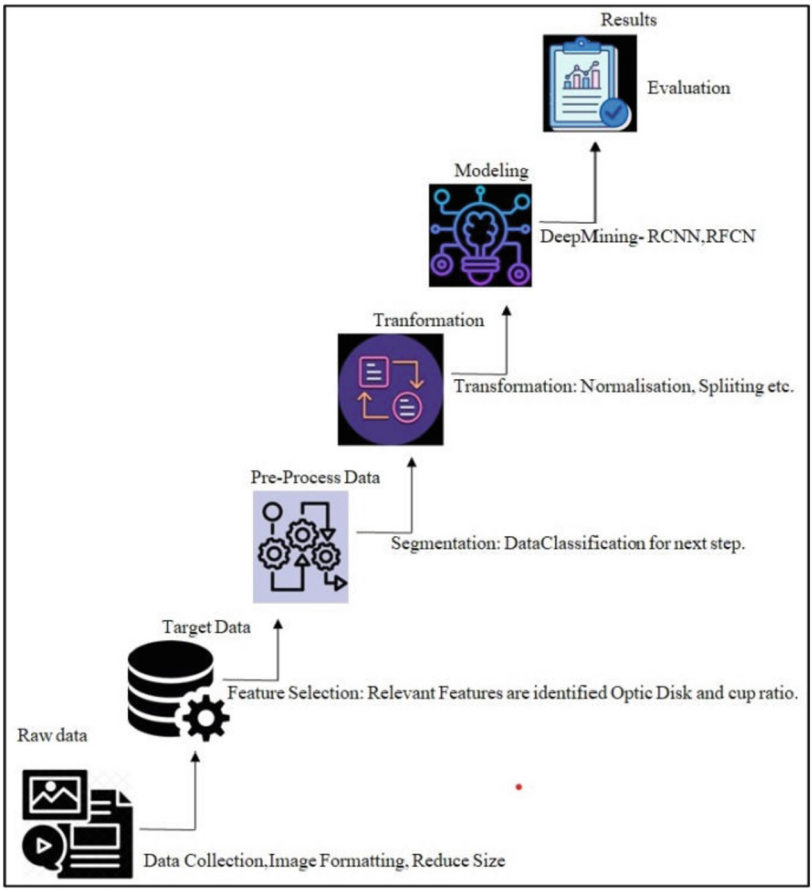

Fig. 6. Glaucoma Classification Methodology

3.2.4 RCNN

R-CNN is a cutting-edge visual object detection system that employing selective search to produce approximately 2000 region proposals per image. It then employs a pre- trained CNN to extracting 4096-dimensional feature vectors for each region of proposal. The system evaluates accuracy using the Intersection-over-Union (IoU) score, fine-tunes the CNN using selected proposals, and achieves outstanding performance without contextual rescoring or ensemble techniques [12].

In glaucoma fundus image classification, manually labeled images are used as positive examples for each class, while region proposals with IoU less than 30% act as negative examples. Bounding box regression refines proposals, and the CNN's last pooling layer features are used to train three regression models for each class. The trained R-CNN categorizes region suggestions into glaucoma, non- glaucoma, or background classes based on a 128 $\times$ 200 resized input, extracting 4096 dimensional feature vectors for each proposal. These feature vectors are then fed into a classification model to

categorize each region proposal into one of the three classes (glaucoma, non-glaucoma, background). The region proposal is assigned the class name with the highest score, representing the most probable object type. To address overlaps and redundancies among bounding boxes, a non-maximum suppression method is used. It compares class scores and the geographic overlap of bounding boxes and keeps only the most reliable and non- overlapping bounding boxes while eliminating redundant and overlapping ones [15]. In summary, glaucoma fundus image classification involves processing the data through selective search, feature extraction, bounding box regression, and non-maximum suppression to achieve accurate and efficient glaucoma classification [16, 17] (Fig. 7).

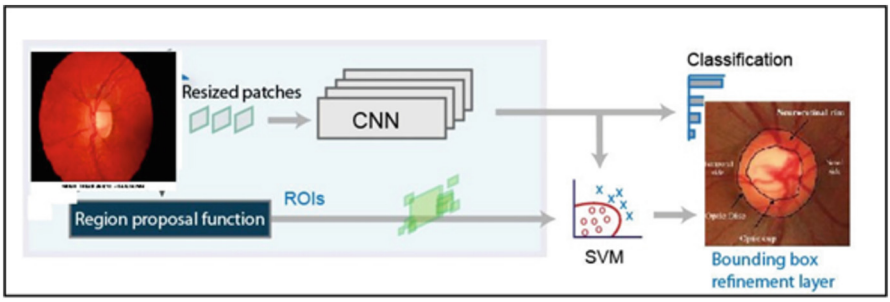

Fig. 7. RCNN Architecture

3.2.5 RFCN

A region-based detection of objects method called R-FCN makes use of fully convolutional deep networks. It offers an extremely precise and effective method of object detection. R-FCN is built as a fully CNN in contrast to earlier region-based analyzers like Fast/Faster R-CNN, which need to run a highly computational sub network for each region. As a result, there are large efficiency gains because the majority of the computation is distributed across the entire image [13, 14] (Fig. 8).

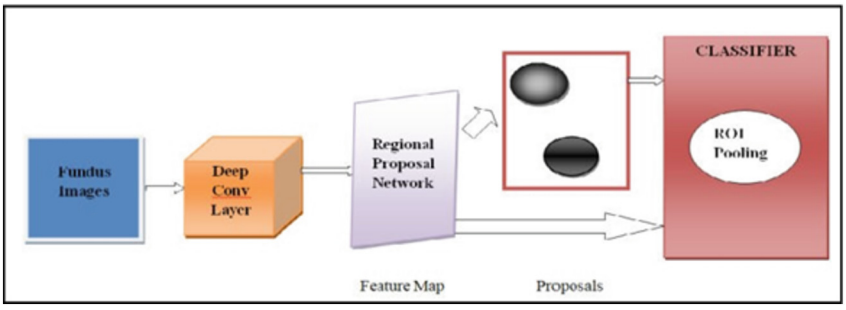

Fig. 8. RFCN Architecture

Take a look at feature map M, which has a square object and measures 5 by 5. We evenly divide the square item into three sections in order to analyze it more thoroughly. Our objective is to build a new feature map that utilizes M that only recognizes the square's top left corner. The size of the final activation map will match M's. To show that the upper left corner is present, it can only activate a subset of the grid cells. The sole activated grid cell in this instance will be the one at location [2], which is indicated in yellow. This activation clearly indicates the top left corner of a region within the map's structure (Fig. 9).

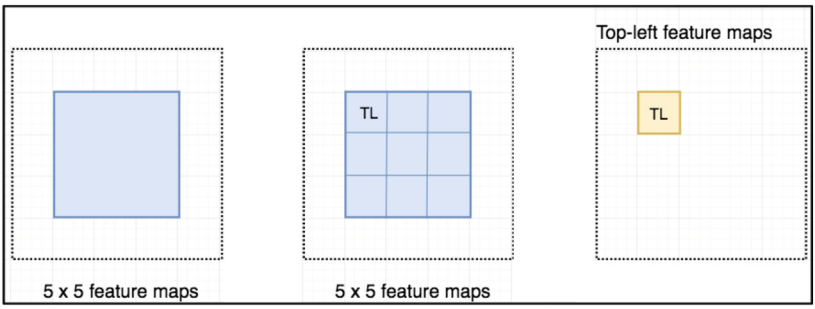

Fig. 9. Top Left Feature Maps

We can create a feature map that concentrates on the left side of the item in order to identify the top left corner of an object. We can produce nine distinct feature maps by splitting an object's area into its nine constituent pieces. Each of these maps is intended to identify a certain area of the object, such as the centre left, top left, top middle, and top right. As they are in charge of scoring and identifying sub-regions of the object according to their positions, these feature maps are known as position sensitive score maps (Fig. 10).

A proposed region of interest (ROI) is shown as a dotted red rectangle in the diagram below. We use a 3×3 grid to separate this ROI into smaller sections for analysis. The probability that each of these tiny regions contains the corresponding component of the thing we are particularly interested is subsequently evaluated.

The diagram on the far right shows how the analysis's findings are saved in a 3×3 vote array (Fig. 11).

The procedure of aligning score maps and ROIs with the vote array is commonly known as position-sensitive ROI-pooling. This technique shares similarities with the ROI pooling method employed in Fast R-CNN.

- The initial step involves the selection of the ROI region located at the top-left, which is then mapped to the corresponding top-left score map.
- The next phase involves computing the mean value of the scores obtained from the region of interest (ROI) located at the top-left quadrant. This ROI is visually depicted by the blue rectangle in the diagram.
- The findings of the research indicate that nearly 40% of the region enclosed by the blue rectangle demonstrates absence of activation, whereas the remaining 60% manifests complete activation, culminating in an average score of 0.6.

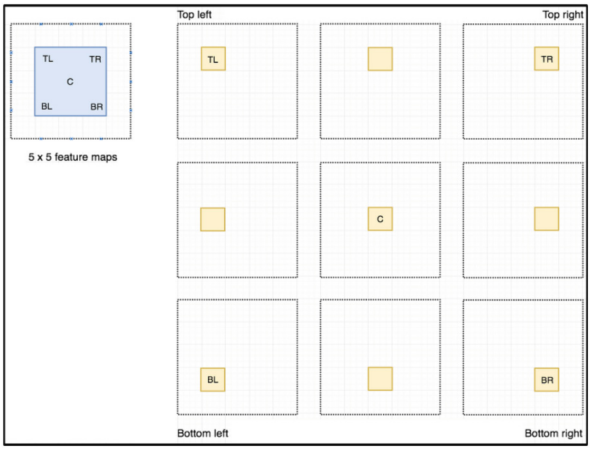

Fig. 10. Position-Sensitive Score Maps

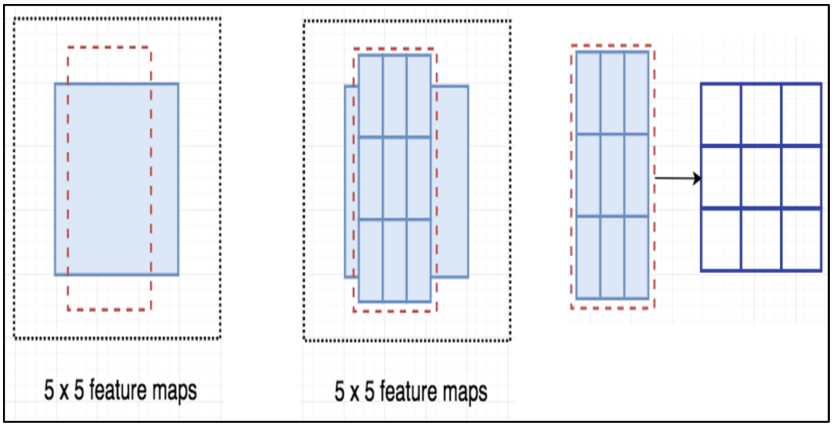

Fig. 11. Generate 9 Score Maps

- This suggests a probability of recognizing the object located in the top-left position with a level of trust of 0.6.
- The outcome of 0.6 is recorded in the array's first row and first column, denoted as [0][0].
- The aforementioned process is reiterated utilizing the top-middle Region of Interest (ROI) and its corresponding score map.
- The outcome obtained is 0.55, denoting the probability of detecting the object located at the top-middle position. This result has been saved in the array's index [0][1].

After calculating each value in the position-sensitive ROI pool, the class score is then calculated as the mean of those values (Fig. 12).

Assuming the task involves detecting C classes, we extend the number of classes to $C + 1$, which includes an additional class representing the background or non-object.

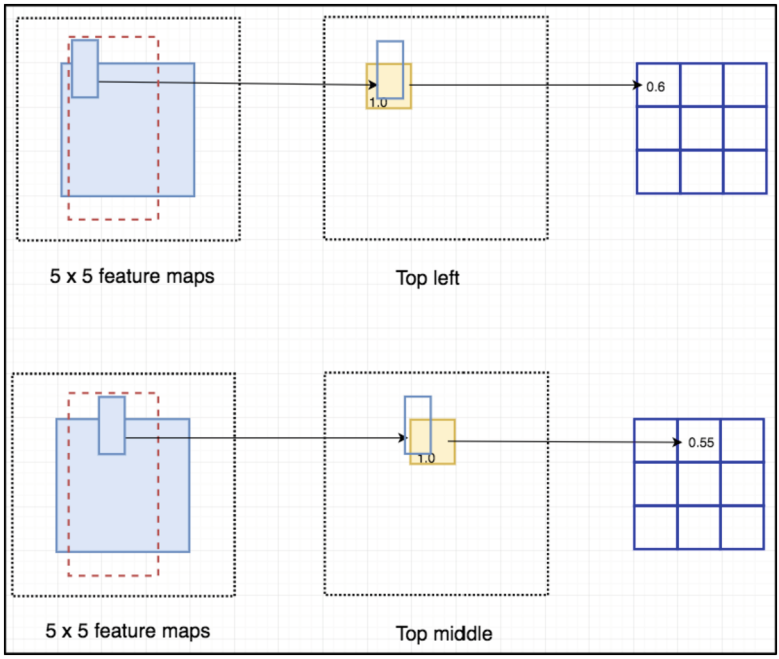

Fig. 12. Top Left and Top Middle Score Map

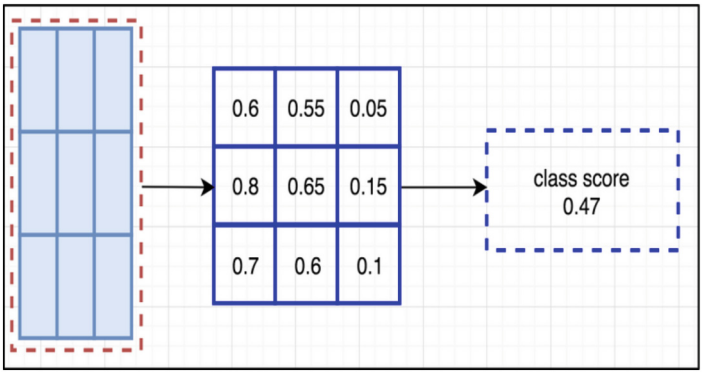

Fig. 13. Class Score

As a result, each class is associated with its own set of 3×3 score maps, resulting in a total of $(C + 1) \times 3 \times 3$ score maps. Using these individual score maps, class scores are predicted for each class. Subsequently, a softmax function is applied to these scores, yielding the probability distribution for each class (Fig. 13) and (Fig. 14).

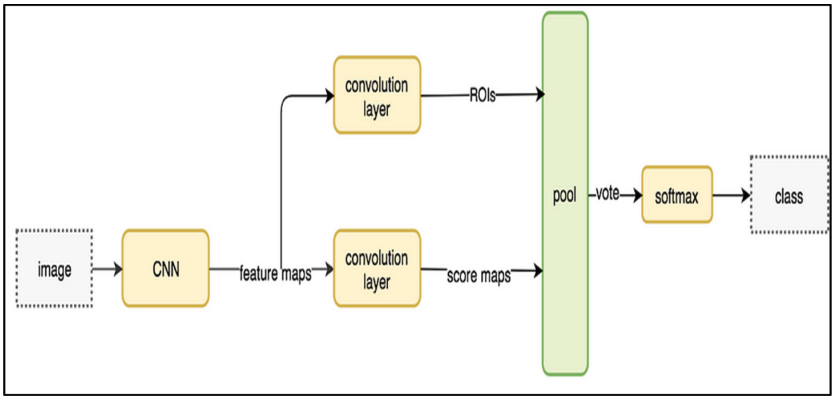

Fig. 14. Data Flow for the R-FCN

3.2.6 Training Progress

Fig. 15. Training Progress

3.2.7 Accuracy

Classification Accuracy is what we usually as commonly understood refers to the proportion of accurate predictions to the overall number of input samples However, it is important to note that this measure performs optimally only when there is an equal division of samples across each class. To illustrate, let's consider a training dataset where 98% of the samples belong to class A and only 2% belong to class B. In such a scenario, a model can easily achieve a training accuracy of 98% by simply predicting all samples as class A.

Nonetheless, if the aforementioned model were to be assessed on a test dataset comprising 60% class A instances and 40% class B instances, the accuracy of the test would diminish to 60% (Fig. 15).

Although Classification Accuracy may initially appear favorable, it can lead to a deceptive impression of achieving high accuracy. The real challenge arises when the misclassification cost of samples from the minor class is considerably high. Particularly in cases involving rare yet severe diseases, the consequences of failing to diagnose an illness in an affected individual far outweigh the cost of subjecting a healthy person to additional tests. The accuracy of test data is 93 - 96% using RCNN and 81% using RFCN (Figs. 16, 17, 18, 19, 20 and 21).

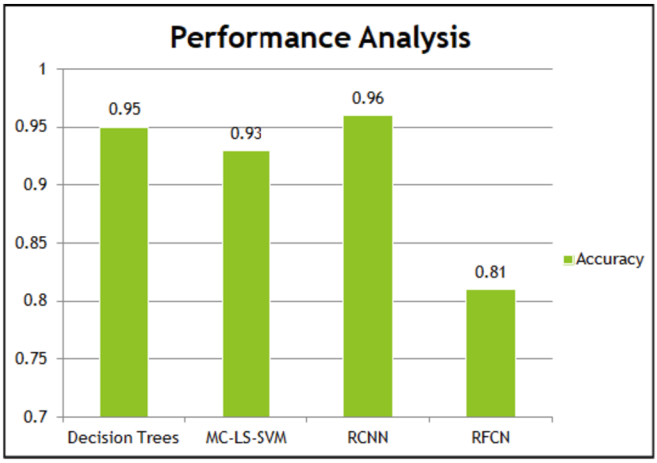

Fig. 16. Performance Analysis of Different Models

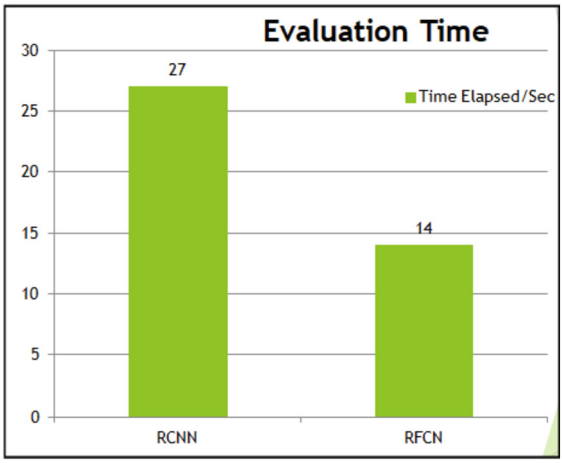

Fig. 17. Time Elapsed Comparison Analysis of Implemented Models

Epoch	Iteration	Time Elapsed (hh:mm:ss)	Mini-batch Accuracy	Mini-batch Loss	Base Learning Rate
1	1	00:00:12	74.07%	0.6322	1.0000e-04
10	10	00:00:27	96.30%	0.2039	1.0000e-04

Fig. 18. Accuracy for RFCN

Epoch	Iteration	Time Elapsed (hh:mm:ss)	Mini-batch Accuracy	Mini-batch Loss	Base Learning Rate
1	1	00:00:01	81.48%	0.9156	1.0000e-06
10	10	00:00:14	81.48%	0.8050	1.0000e-06

Fig. 19. Accuracy for RCCN

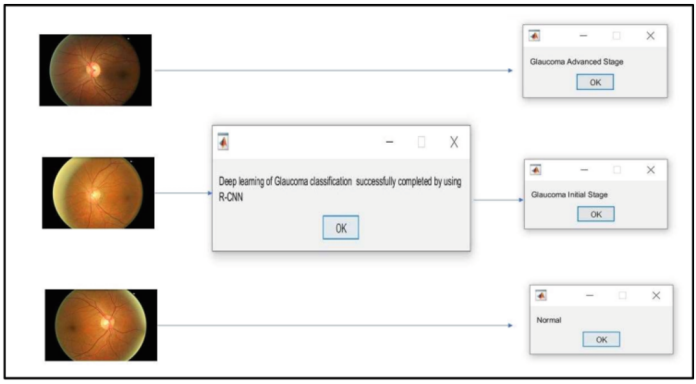

Fig. 20. Classification of glaucoma stages using RCNN

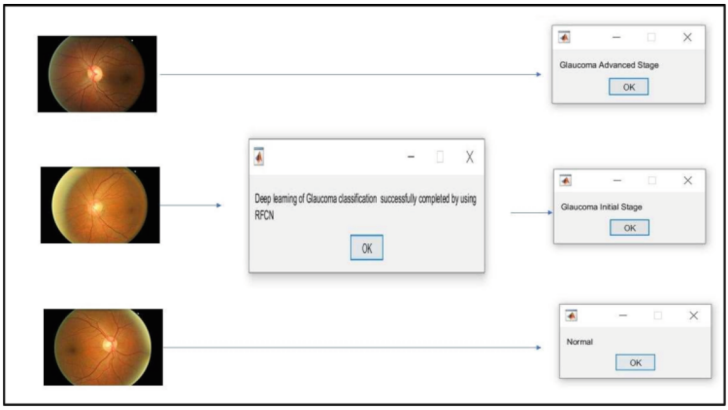

Fig. 21. Classification of glaucoma stages using RFNN

4 Conclusion

Glaucoma is an ocular condition that can lead to permanent vision loss in affected individuals. Early detection of glaucoma-related optic nerve damage is essential to prevent further deterioration. In our research, we assessed the effectiveness of R-CNN and R-FCN approaches for accurately detecting glaucoma in color fundus images. We assessed other aspects of the images, including the colour, brightness, and RGB intensity value histograms. In effect, we noticed that datasets with comparable histograms produced better results. In order to ensure its robustness over various vessel widths and luminosity circumstances, we used a supervised learning method and trained it on datasets. The accuracy curve shows that our suggested strategy outperforms existing methods. In particular, ResNet-50, a well-known deep learning architecture, performed better in our tests. By evaluating our method on the Orgia dataset, comprising 650 images, we implied that it could be a useful replacement for widespread computer-aided glaucoma screening programmes. Additionally, we assessed the performance of R-FNN and the

R-CNN. The outcomes clearly showed that R-CNN outperformed R-FCN in terms of accuracy, scoring 96.30% versus 81.48%. However, R-FCN boasted a faster processing time, taking only 14 s as opposed to R-CNN's 27 s.

5 Future Work

There is the opportunity to investigate a variety of neural network topologies in order to improve the results of vascular segmentation by lowering the number of instances of false-positive detections. In the course of our ongoing research, we are attempting to design innovative structures that are tailored to the detection of glaucoma on vast databases.

References

1. Hagiwara, Y., et al.: Computer-aided diagnosis of glaucoma using fundus images: a review. Comput. Methods Programs Biomed. **165**, 1–12 (2018)
2. Diaz-Pinto, A., Colomer, A., Naranjo, V., Morales, S., Xu, Y., Frangi, A.F.: Retinal image synthesis and semi-supervised learning for glaucoma assessment. IEEE Trans. Med. Imaging **38**, 2211–2218 (2019). https://doi.org/10.1109/tmi.2019.2903434
3. Aloudat, M., Faezipour, M., El-Sayed, A.: High intraocular pressure detection from frontal eye images: a machine learning based approach. PubMed (2018). https://doi.org/10.1109/embc.2018.8513645
4. Saha, S.K., Fernando, B., Cuadros, J., Xiao, D., Kanagasingam, Y.: Automated quality assessment of colour fundus images for diabetic retinopathy screening in telemedicine. J. Digit. Imaging **31**, 869–878 (2018). https://doi.org/10.1007/s10278-018-0084-9
5. Maheshwari, S., Pachori, R.B., Kanhangad, V., Bhandary, S.V., Acharya, U.R.: Iterative variational mode decomposition based automated detection of glaucoma using fundus images. Comput. Biol. Med. **88**, 142–149 (2017). https://doi.org/10.1016/j.compbiomed.2017.06.017
6. Parashar, D., Agrawal, D.K.: Classification of glaucoma stages using image empirical mode decomposition from fundus images. J. Digit. Imaging **35**(5), 1283–1292 (2022). https://doi.org/10.1007/s10278-022-00648-1
7. Joshi, S., Partibane, B., Hatamleh, W.A., Tarazi, H., Yadav, C.S., Krah, D.: Glaucoma detection using image processing and supervised learning for classification. J. Healthcare Eng. **2022**, 2988262 (2022). https://doi.org/10.1155/2022/2988262
8. Fu, H., Cheng, J., Xu, Y., Liu, J.: Glaucoma detection based on deep learning network in fundus image. Deep Learning and Convolutional Neural Networks for Medical Imaging and Clinical Informatics. 119–137 (2019). https://doi.org/10.1007/978-3-030-13969-8_6
9. Shinde, R.: Glaucoma detection in retinal fundus images using U-Net and supervised machine learning algorithms. Intell.-Based Med. **5**, 100038 (2021). https://doi.org/10.1016/j.ibmed.2021.100038
10. Kashyap, R., Nair, R., Gangadharan, S.M.P., Botto-Tobar, M., Farooq, S., Rizwan, A.: Glaucoma detection and classification using improved U-Net deep learning model. Healthcare **10**, 2497 (2022). https://doi.org/10.3390/healthcare10122497
11. Agarwal, T.K., Tiwari, M., Lamba, S.S.: Modified Histogram based contrast enhancement using Homomorphic Filtering for medical images. In: 2014 IEEE International Advance Computing Conference (IACC), pp. 964–968 (2014). https://doi.org/10.1109/iadcc.2014.6779453

12. Li, L., et al.: A large-scale database and a CNN model for attention-based glaucoma detection. IEEE Trans. Med. Imaging **39**(2), 413–424 (2019). https://doi.org/10.1109/tmi.2019.2927226
13. Shelhamer, E., Long, J., Darrell, T.: Fully convolutional networks for semantic segmentation. IEEE Trans. Pattern Anal. Mach. Intell. **39**, 640–651 (2017). https://doi.org/10.1109/TPAMI.2016.2572683
14. Shankar, D., George, G.V.S., JNSS, J.N., Madhuri, P.S.: Deep analysis of risks and recent trends towards network intrusion detection system. Int. J. Adv. Comput. Sci. Appl. 14, (2023). https://doi.org/10.14569/ijacsa.2023.0140129
15. Shankar, D., George, G.V.S., Kanya, N.: OptiBiNet_GRU: robust network intrusion detection system using optimum bi-directional gated recurrent unit. Int. J. Intell. Eng. Syst. **16**, 75–91 (2023). https://doi.org/10.22266/ijies2023.0630.06
16. Shanmugam, P., Raja, J., Pitchai, R.: An automatic recognition of glaucoma in fundus images using deep learning and random forest classifier. Appl. Soft Comput. **109**, 107512 (2021). https://doi.org/10.1016/j.asoc.2021.107512
17. Sandhya, M., Morampudi, M.K., Grandhe, R., Kumari, R., Banda, C., Gonthina, N.: Detection of Diabetic Retinopathy (DR) severity from fundus photographs: an ensemble approach using weighted average. Arab. J. Sci. Eng. (2022). https://doi.org/10.1007/s13369-021-06381-1