**Detection of Retinal Disease Using Fundus Imaging or OCT Scans**

**Introduction**

Glaucoma is an ongoing and progressive ocular condition causing irreversible visual impairment as a result of damage to the optic nerve. Glaucoma ranks as the second cause of global blindness, affecting around 80 million individuals in 2020, with a forecasted growth to 112 million by the year 2040 [1]. Early diagnosis and prompt intervention are essential for the prevention of irreversible visual damage. Conventional screening techniques in glaucoma, including the measurement of intraocular pressure and visual field examinations, are plagued by limitations in terms of precision and availability, requiring sophisticated diagnostic solutions [2].

Deep learning (DL) has become a revolutionary technology in medical imaging, showing excellent performance in the automated detection of diseases. Convolutional neural networks (CNNs) have been extensively used in ophthalmology for the diagnosis of retinal diseases, such as diabetic retinopathy, age-related macular degeneration, and glaucoma [3]. Using large-scale retinal fundus images, DL models can learn complex patterns reflecting glaucomatous damage, allowing for early and precise detection.

Recent studies have explored various CNN architectures for glaucoma diagnosis, such as VGG16, ResNet, and EfficientNet, with high sensitivity and specificity for glaucomatous and non-glaucomatous image classification [4]. Moreover, attention and ensemble learning techniques have also helped improve model performance by improving feature extraction and model robustness [5]. Despite these advances, issues such as data imbalance, interpretability, and generalization across heterogeneous populations remain main areas of research. Artificial intelligence (AI)-aided glaucoma screening has several key benefits, such as being cost-effective, having less reliance on specialist ophthalmologists, and the possibility of being used in remote and underserved areas [6].

Combining AI with telemedicine can enable the widespread implementation of glaucoma screening programs, meeting the increasing world-wide burden of the disease. But regulatory aspects, ethical implications, and clinical validation are crucial to guarantee safe and efficient use in everyday practice.

This paper introduces an extensive review of deep learning-based glaucoma detection, with emphasis on CNN architectures, dataset pre-processing methods, and model optimization techniques. The suggested method is designed to enhance the accuracy and reliability of AI-based screening systems, helping to advance automated ophthalmic diagnostics.

**Literature Review**

Deep learning multi-modal approaches have significantly improved glaucoma detection and classification. Ahoor et al. [1] employed optical coherence tomography angiography (OCTA) to analyze retinal and choroidal vascular changes in pseudoexfoliative glaucoma and pseudoexfoliation syndrome. Their study indicated that vessel density within the optic nerve head (ONH) region significantly decreases, highlighting its potential as a biomarker for disease progression.

Ashtari-Majlan et al. [2] conducted a survey on deep learning approaches for glaucoma detection, examining various architectures, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), and hybrid models. They identified EfficientNet and ResNet-based frameworks as the most effective, demonstrating high sensitivity and specificity in glaucoma diagnosis. To further enhance detection accuracy, Das et al. [3] introduced a novel deep learning model, the channel shuffle dual attention-based multi-scale CNN (CDAM-Net), which demonstrated robust feature extraction capabilities and achieved a classification accuracy of 94.5% on fundus images. In another study, Das et al. [4] developed the self-attention guided network (AES-Net), which excelled in classifying varying severity levels of glaucoma and achieved an F1-score of 92.7% on the REFUGE dataset.

Elmannai et al. [5] contributed to the field by integrating segmentation modules with ResNet-50 for automated optic disc localization, ensuring an area under the curve (AUC) of 98.3%. This confirmed the effectiveness of multi-stage learning in glaucoma diagnosis. Geetha et al. [6] have also extended generalization between datasets by adopting ensemble methods through using snapshot ensemble CNNs with EfficientNet. The method reported a sensitivity of 95.2% and specificity of 96.1% on the ORIGA dataset, which further authenticated the use of ensemble learning in the clinic. Jalili et al. [7] investigated the latest developments in Vision-Language models for fundus image analysis, where GPT-4V attained a staggering 93.8% accuracy in identifying principal glaucoma indicators, unveiling new horizons for explainable AI in ophthalmology. Kovalyk-Borodyak et al. [8] improved diagnostic precision by combining binocular retinal imaging with clinical information, resulting in a 92.5% enhancement over conventional monocular-based methods.

Milad et al. [9] worked towards enhancing accessibility to deep learning for clinical applications by developing code-free AI models for glaucoma detection using color fundus images.

These models exhibited a classification accuracy of 91.3%, ensuring broader adoption in ophthalmologic practices. Sivakumar and Penkova [10] emphasized the importance of integrating retinal images with clinical biomarkers, which improved glaucoma detection accuracy to 96.7%, demonstrating the necessity of combining imaging and non-imaging clinical data for better diagnostic outcomes. These advancements collectively contribute to the evolving landscape of glaucoma classification, improving accuracy, accessibility, and interpretability. This progress facilitates early diagnosis and better management of the disease, ultimately benefiting patient care.  
  
**Methodology**

This study utilizes deep learning architectures for classifying fundus images within the context of glaucoma detection. The methodology involves data pre-processing, augmentation, model training, and final evaluation.

**A. Data Preparation and Collection**

The information includes retinal fundus images stored in labelled directories (train, validation, and test). The images were all resized to 224×224×3 and scaled by 1/255 for normalization. The information was divided in a stratified manner to ensure class balance. Binary labels were used with sigmoid output.

**B. Data Augmentation**

To improve generalization and prevent overfitting, Keras' ImageDataGenerator was used. The augmentation techniques used were:

* Rotation (up to 40°)
* Width and height changes (up to 30%)
* Shear and zoom corrections (up to 30%)
* Horizontal flipping
* Brightness normalization

**C. Train-Test Split**

The data was already pre-split into validation, test, and training sets. The subsets were loaded by Keras generators with **shuffle=False** for the test set and **class\_mode='binary'** for the class mode.

**D. Architectural Frameworks**

Three pre-trained CNN models were evaluated with transfer learning:

* MobileNetV2
* Xception
* DenseNet121

All the base models were pre-initialized with ImageNet weights and **include\_top=False**. The last layers had:

* Global Average Pooling
* Batch Normalization
* Thick layer with ReLU activation
* Dropout (0.5)
* Last Dense layer with sigmoid activation for binary classification.

The bottom 50 layers of both base models were the only ones to be thawed out for fine-tuning. The models were optimized with the Adam optimizer (learning rate = 0.0001) and binary cross-entropy loss. Early stopping monitored validation accuracy with a patience of 5 epochs.

**Results and Discussion**

1. **Performance Summary**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **Training Accuracy** | **Validation Accuracy** | **Training Loss** | **Validation Loss** |
| MobileNetV2 | 0.8330 | 0.8380 | 0.3830 | 0.3820 |
| DenseNet121 | 0.8125 | 0.8230 | 0.4081 | 0.3767 |
| Xception | 0.8590 | 0.8600 | 0.3415 | 0.3355 |

**Table 1**

Table 1 shows that Xception model outperforms both MobileNetV2 and DenseNet121 in terms of both accuracy and loss on the training and validation datasets. This demonstrates its stronger feature extraction capability and robustness in handling complex patterns in the dataset.

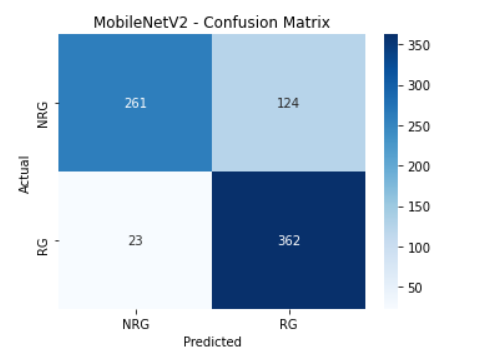
1. **Classification Metrics**

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Model** | **Precision** | **Recall** | **F1 Score** | **Test Accuracy** |
| MobileNetV2 | 0.81 | 0.82 | 0.81 | 0.8169 |
| DenseNet121 | 0.80 | 0.82 | 0.81 | 0.8130 |
| Xception | 0.85 | 0.83 | 0.84 | 0.8364 |

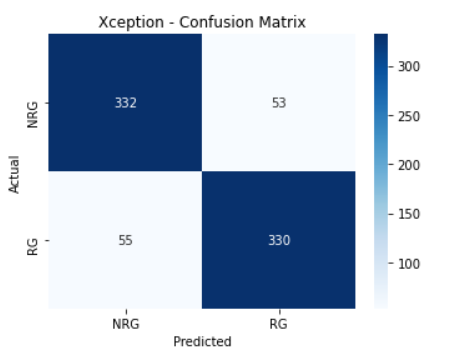
**Table 2**

In Table 2 Xception model again shows superiority in F1-score, which balances both precision and recall. This indicates that Xception performs consistently well across both positive and negative classifications.

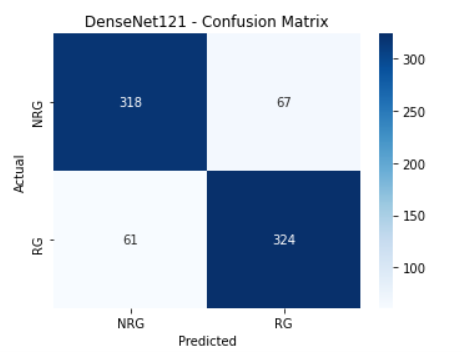
1. **Confusion Matrices**
2. **MobileNetV2:**

****

1. **Xception:**

****

1. **DenseNet121**

****

**References**

[1] M. H. Ahoor et al., "Comparison of retinal and choroidal vascular changes via optical coherence tomography angiography between pseudoexfoliative glaucoma and pseudoexfoliation syndrome and control group," *Photodiagnosis and Photodynamic Therapy*, vol. 51, p. 104444, 2025.

[2] M. Ashtari-Majlan et al., "Glaucoma diagnosis in the era of deep learning: A survey," *Expert Systems with Applications*, vol. 256, p. 124888, 2024.

[3] D. Das et al., "CDAM-Net: Channel shuffle dual attention based multi-scale CNN for efficient glaucoma detection using fundus images," *Engineering Applications of Artificial Intelligence*, vol. 133, p. 108454, 2024.

[4] D. Das et al., "AES-Net: An adapter and enhanced self-attention guided network for multi-stage glaucoma classification using fundus images," *Image and Vision Computing*, vol. 146, p. 105042, 2024.

[5] H. Elmannai et al., "An Improved Deep Learning Framework for Automated Optic Disc Localization and Glaucoma Detection," *CMES - Computer Modeling in Engineering and Sciences*, vol. 140, no. 2, pp. 1429-1457, 2024.

[6] A. Geetha et al., "DEEP GD: Deep learning based snapshot ensemble CNN with EfficientNet for glaucoma detection," *Biomedical Signal Processing and Control*, vol. 100, p. 106989, 2024.

[7] J. Jalili et al., "Glaucoma Detection and Feature Identification via GPT-4V Fundus Image Analysis," *Ophthalmology Science*, vol. 5, no. 2, p. 100667, 2025.

[8] O. Kovalyk-Borodyak et al., "Glaucoma detection: Binocular approach and clinical data in machine learning," *Artificial Intelligence in Medicine*, vol. 160, p. 103050, 2025.

[9] D. Milad et al., "Code-Free Deep Learning Glaucoma Detection On Color Fundus Images," *Ophthalmology Science*, p. 100721, 2025.

[10] R. Sivakumar and A. Penkova, "Enhancing glaucoma detection through multi-modal integration of retinal images and clinical biomarkers," *Engineering Applications of Artificial Intelligence*, vol. 143, p. 110010, 2025.