



ENHANCING GLAUCOMA DIAGNOSIS USING DEEP LEARNING TECHNIQUE

Thesis-AI900: Dr. Hager

Abdullah Abdulaziz
242000982

Contents

Abstract	2
2. Related Work.....	3
2.1 Deep Learning for Glaucoma Detection.....	3
2.2 Dimensionality Reduction Using PCA.....	3
2.3 Addressing Class Imbalance in Medical Datasets	3
2.4 XGBoost for Classification in Medical Imaging.....	3
3. Proposed Methodology.....	5
4. Results.....	6
5. Discussion	7
6. Conclusion	7
References	9

Title: Enhancing Glaucoma Diagnosis Using Deep Learning Technique

1. Abstract

Glaucoma ranks among the top causes of permanent blindness globally, making early detection and accurate diagnosis to mitigate vision loss. Recent deep learning and machine learning growth have significantly improved glaucoma detection through medical image analysis. This study presents a novel hybrid deep learning and machine learning framework for automated glaucoma detection by integrating Convolutional Neural Networks (CNNs) for feature extraction, Principal Component Analysis (PCA) for dimensionality reduction, and eXtreme Gradient Boosting (XGBoost) for classification. The proposed approach employs two pre-trained CNN models—ResNet50 and VGG-16—to extract spatial and texture-related features optimised using a two-stage PCA process. To address the class imbalance commonly found in medical datasets, the Synthetic Minority Oversampling Technique (SMOTE) was applied. The framework was evaluated using publicly available datasets (ORIGA, REFUGE, and G1020), achieving an accuracy of 97.24% with XGBoost. The findings demonstrate the potential of deep learning-based feature extraction and ensemble machine learning for enhancing diagnostic accuracy in medical imaging.

Keywords Glaucoma detection, deep learning, CNN, PCA, XGBoost, machine learning, SMOTE

Introduction Glaucoma is a degenerative eye disease that harms the optic nerve, resulting in vision impairment and possibly blindness if not treated. At present, over 79.6 million people globally are affected by glaucoma, a figure projected to increase to 111.8 million by 2040 (Tham et al., 2014). Early detection and swift treatment are essential for effectively managing this disorder. Conventional diagnostic techniques, such as measuring intraocular pressure and evaluating the optic nerve head, depend on the skills of ophthalmologists and may lead to varying interpretations in different clinical settings (Chen et al., 2021).

Artificial intelligence (AI) has recently been recognised as a valuable resource in medical image analysis, offering increased automation and improved precision in diagnosing diseases. (Gulshan et al., 2016). Deep learning methods, particularly Convolutional Neural Networks (CNNs), have remarkably succeeded in medical image classification tasks. (Esteva et al., 2017). Nevertheless, these deep-learning models frequently generate high-dimensional feature representations, increasing computational complexity and a heightened risk of overfitting (Shen, Wu, & Suk, 2017). To mitigate these issues, Principal Component Analysis (PCA) is often utilised to reduce dimensionality, enabling effective feature extraction while preserving essential information (Jolliffe & Cadima, 2016).

Medical datasets frequently face class imbalance, characterised by a significantly higher number of non-glaucoma cases compared to glaucoma cases. This imbalance can lead to a bias in machine learning models towards the majority class, resulting in suboptimal generalisation (He & Garcia, 2009). To address this concern, the Synthetic Minority Oversampling Technique (SMOTE) is employed to generate synthetic samples and achieve a balanced dataset (Chawla et al., 2002).

This research introduces a hybrid approach that merges deep learning with machine learning methods for automatically detecting glaucoma. It utilises CNN for feature extraction, PCA for reducing dimensionality, SMOTE for balancing data, and XGBoost for classification. The suggested technique takes advantage of deep learning's capability for robust feature representation and applies ensemble learning to improve classification precision.

2. Related Work

2.1 Deep Learning for Glaucoma Detection

Numerous research efforts have investigated the use of deep learning to detect glaucoma. Convolutional Neural Networks (CNNs) are frequently employed for automated feature extraction from retinal fundus images, demonstrating better performance than traditional methods like Gray Level Co-occurrence Matrix (GLCM) and Wavelet Transform (Christopher et al., 2018). Pretrained deep learning architectures such as ResNet50, VGG-16, and InceptionV3 have been utilised to derive high-level visual features, assisting in classifying glaucoma (Shankaranarayana et al., 2019).

2.2 Dimensionality Reduction Using PCA

Reducing dimensionality is vital in medical image analysis as it enhances computational efficiency and minimises redundancy in feature representations. Principal Component Analysis (PCA) is commonly applied to lessen high-dimensional feature spaces while retaining significant information (Jolliffe, 2002). Previous studies have shown that PCA can substantially boost model generalisation and training efficiency when integrated with feature extraction using deep learning (Abdi & Williams, 2010).

2.3 Addressing Class Imbalance in Medical Datasets

Medical datasets often face considerable class imbalance, where positive instances (glaucoma) are less frequent than negative instances (non-glaucoma). This discrepancy affects model predictions and diminishes the classifier's sensitivity (Sun, Wong, & Kamel, 2009). The SMOTE method generates synthetic examples to balance class distributions, enhancing the classifier's capacity to learn from minority instances (Chawla et al., 2002). Earlier research has indicated that implementing SMOTE improves model robustness in the context of medical image classification tasks (Fernández et al., 2018).

2.4 XGBoost for Classification in Medical Imaging

XGBoost is a prominent ensemble learning algorithm with exceptional performance when dealing with high-dimensional data (Chen & Guestrin, 2016). Compared to traditional classifiers such as Support Vector Machines (SVM) and Random Forest, XGBoost effectively captures non-linear relationships in the dataset and reduces the risk of overfitting through its gradient boosting approach (Friedman, 2001). Research has demonstrated that XGBoost achieves leading performance in medical image classification tasks, surpassing conventional machine learning techniques (Wang et al., 2020).

Several studies have explored various AI-driven methodologies for glaucoma detection as below:

- Lenka et al. (2025) introduced a graph convolution-based multi-task model that improves segmentation and classification accuracy.
- Akter et al. (2022) utilised a multi-feature analysis approach integrating deep learning for precise diagnosis.
- Civit-Masot et al. (2020) proposed a dual machine-learning system using disc and cup feature extraction, improving classification outcomes.
- Thompson et al. (2020) reviewed deep learning applications in glaucoma screening, emphasising CNNs' role in retinal image analysis.

Although these studies show notable progress, issues such as computational complexity, variability in datasets, and restricted generalizability remain.

This study utilises Deep Learning to analyse retinal fundus images for glaucoma detection, emphasising the increasing importance of AI in healthcare research. Convolutional Neural Networks (CNNs) are a popular deep learning method for image recognition and classification. The research combines ensemble ML techniques with CNN models to improve classification accuracy. Specifically, it utilises three models—Random Forest, ResNet50, and VGG-16—to improve diagnostic precision. The Random Forest model evaluates texture features from retinal fundus images using the Gray-Level Co-Occurrence Matrix (GLCM) approach. A detection framework is introduced to classify retinal fundus images as glaucomatous or usual, facilitating reliable and efficient disease identification.

This study introduces an advanced AI-based glaucoma detection system that combines machine learning (ML), CNNs, and image processing. This hybrid method enhances performance by integrating traditional feature extraction with deep learning-based classification.

Year	Author(s)	Type	Methodology Employed	Key Findings	Weaknesses/Limitations
2025	Satyabrata Lenka, Zefree Lazarus Mayaluri, Ganapati Panda “Glaucoma detection from retinal fundus images using graph convolution based multi-task model”	Article	Multi-task deep learning model using modified U-Net architecture with MobileNetv2 encoder, GCN, decoder, and Attention Module (AM) to segment and classify glaucoma from retinal fundus images. Evaluated on ORIGA, REFUGE, and DRISTI-GS datasets.	Improve segmentation and classification with high accuracy rates: 97.43% for glaucoma detection, an AUROC of 0.985, and Dice coefficients of 97.95% for optic disc segmentation and 96.11% for optic cup segmentation. This model surpasses current methods in both segmentation and classification.	- High cost and High dependency on datasets. - Potential overfitting to specific datasets, lacking real-world clinical validation. - Computationally intensive, limiting real-time applicability. - No comparison with ViTs.
2022	Ramgopal Kashyap, Rajit Nair, Syam Machinathu Parambil Gangadharan, Miguel Botto-Tobar, Saadia Farooq, Ali Rizwan “Glaucoma Detection and Classification Using Improved U-Net Deep Learning Model	Article	Multi-feature analysis: Improved U-Net deep learning model for optic disc (OD) and optic cup (OC) segmentation. CNN-based classifier for glaucoma detection. Evaluated on public fundus image datasets.	Improved diagnostic accuracy The Enhanced U-Net provides superior segmentation for OD and OC, ensuring high accuracy in glaucoma detection through an effective blend of segmentation and classification techniques.	- Complexity in feature integration - No comparison with Transformer-based models. - Resource-intensive for immediate applications. - Lacks real-world clinical validation. - No explainability (XAI) mechanisms for clinical adoption.
2024	Mona Ashtari-Majlan, Mohammad Mahdi Dehshibi, David Masip	Survey	CNNs, Hybrid CNN-Transformers, Vision Transformers (ViTs), RNNs, and self-supervised learning approaches.	Hybrid CNN-transformer models outperform traditional CNNs in detecting glaucoma. Self-supervised learning reduces dependence on labelled	- Lacks experimental validation and model benchmarking. - Ignores the computational expenses

	“Glaucoma diagnosis in the era of deep learning: A survey”		Evaluates models on ORIGA, REFUGE, DRISTI-GS, RIM-ONE, and ACRIMA datasets.	datasets. Vision Transformers (ViTs) demonstrate strong generalisation capabilities in medical imaging. This highlights the essential importance of Explainable AI (XAI) in clinical applications.	and challenges of implementing in real-time. - Limited discussion on regulatory and ethical issues in AI-based glaucoma dia
2020	J. Civit-Masot, M. J. Domínguez-Morales, S. Vicente-Díaz, A. Civit “Dual Machine-Learning System to Aid Glaucoma Diagnosis Using Disc and Cup Feature Extraction	Article	Feature Extraction by using a dual machine learning system (disc and cup) Utilises Random Forest, SVM and Gradient Boosting for glaucoma classification. Implements explainable AI (XAI) for interpretability.	Adept Glaucoma Classification Achieved high accuracy in glaucoma detection using handcrafted feature extraction methods. Ensemble models like Random Forest (RF) and Gradient Boosting (GB) outperform single classifiers. These techniques also offer greater computational efficiency compared to deep learning models. Explainable AI (XAI) enhances clinical trust.	- Narrow generalisation. Fails to compare with deep learning (CNNs, ViTs). - Depends on manually crafted features, potentially restricting learning ability. - Omits discussion on dataset generalisation or validation in real-world contexts. - Does not investigate self-supervised learning or transformer architectures.

3. Proposed Methodology

The proposed methodology for glaucoma detection integrates deep learning-based feature extraction, dimensionality reduction using Principal Component Analysis (PCA), synthetic data augmentation with the Synthetic Minority Oversampling Technique (SMOTE), and classification using eXtreme Gradient Boosting (XGBoost). The pipeline begins with data preprocessing, where retinal fundus images are resized to 224×224 pixels to maintain consistency with deep learning models. The images are then normalised to scale pixel values within a standard range, followed by noise removal techniques such as median filtering to enhance feature clarity.

For feature extraction, two pre-trained Convolutional Neural Networks (CNNs), ResNet50 and VGG-16, are employed to extract deep spatial and texture-related features. ResNet50 generates a 2048-dimensional feature vector representing high-level structural patterns in retinal images. Meanwhile, VGG-16 initially produces a 25,088-dimensional feature vector, which is too high-dimensional for efficient model training. PCA-based dimensionality reduction is applied in two stages to reduce computational complexity and minimise overfitting risks. First, PCA reduces the VGG-16 feature output from 25,088 to 512 dimensions, preserving the most critical feature representations. In the second stage, the combined feature vectors from ResNet50 and VGG-16 (2560 dimensions in total) undergo further PCA reduction to 1024 dimensions, ensuring optimal model performance while retaining crucial information.

Given the inherent class imbalance in medical imaging datasets, particularly in glaucoma classification, SMOTE is implemented to generate synthetic samples of underrepresented glaucoma cases. This resampling technique ensures that both glaucoma and non-glaucoma classes are equally represented, thereby improving the classifier's generalization and reducing bias toward the majority class.

An ensemble learning approach is employed for classification. Initially, a Random Forest classifier was trained on the extracted feature vectors, but it achieved only 66% accuracy, indicating its limitations in handling deep-learning feature representations. XGBoost, an advanced gradient boosting algorithm, was employed to overcome these challenges due to its ability to optimise non-linear relationships and handle high-dimensional data efficiently. The XGBoost model was trained on the optimized 1024-dimensional feature vectors, with hyperparameters tuned for log loss minimization to enhance classification performance.

Various assessment metrics and visualization techniques were used to evaluate model performance. A confusion matrix was generated to analyse classification accuracy and error distribution. Training and validation loss curves were plotted to monitor learning behavior and convergence. A Precision-Recall Curve (PR-AUC) was also computed to assess the model's reliability, particularly in differentiating between glaucoma and non-glaucoma cases. The performance evaluation of the proposed system was conducted using publicly available datasets, including ORIGA, REFUGE, and G1020, to ensure robustness and generalizability in real-world applications.

4. Results

The proposed system was evaluated using publicly available datasets, including ORIGA, REFUGE, and G1020. The experimental results confirmed the model's reliability and effectiveness for real-world glaucoma detection applications. The Random Forest model achieved only 96.55% accuracy, demonstrating its limited ability to handle deep-learning feature vectors. In contrast, the XGBoost classifier, after applying PCA and SMOTE, significantly improved accuracy to 97.24%, proving the effectiveness of gradient boosting for medical image classification.

Classification performance was assessed using multiple metrics. The model achieved a precision of 95% for non-glaucoma cases and 100% for glaucoma cases. The recall values were 100% for non-glaucoma cases and 95% for glaucoma cases, resulting in an overall F1 score of 97%. The PR-AUC score of 0.98 further confirmed the model's strong predictive capability.

Comparison of Classification Models

Model	Accuracy	Precision	Recall	F1-Score	PR-AUC
Random Forest	96.55%	93.00%	100.00%	97.00%	0.99
XGBoost	97.24%	95.00%	100.00%	97.4%	0.98

Visual analysis provided additional insights into the model's performance. The feature reduction visualisation using PCA demonstrated how high-dimensional feature vectors were optimised for efficiency. Class distribution analysis before and after SMOTE showed that the dataset had been effectively balanced. Training and validation loss curves confirmed stable model convergence, while accuracy curves indicated that validation accuracy stabilised around

97%, suggesting the model did not suffer from overfitting. A confusion matrix analysis revealed minimal misclassification, reinforcing the model’s reliability.

5. Discussion

The experimental results validate the effectiveness of the proposed hybrid deep learning and machine learning approach for glaucoma detection. CNN-based feature extraction, particularly ResNet50 and VGG-16, significantly improved diagnostic accuracy by capturing both texture and structural abnormalities in retinal images. Traditional handcrafted feature extraction methods, such as the Gray Level Co-occurrence Matrix (GLCM), were insufficient for robust classification.

Dimensionality reduction played a crucial role in optimising computational efficiency. Without PCA, the high dimensionality of VGG-16 output (25,088 features per image) increased computational complexity and the risk of overfitting. The first PCA step effectively reduced VGG-16 output to 512 dimensions while maintaining critical information, while the final PCA step further optimised feature vectors, allowing faster training and better generalisation.

Addressing class imbalance was also a key component of the framework. Initially, the dataset was heavily skewed, with significantly more non-glaucoma cases than glaucoma cases. SMOTE successfully balanced the dataset, preventing the model from being biased toward the majority class. As a result, the classifier achieved high recall rates for glaucoma detection, ensuring that glaucoma cases were not overlooked.

When comparing classification models, Random Forest struggled with deep-learning feature vectors and achieved only 66% accuracy. In contrast, XGBoost demonstrated superior handling of high-dimensional features, achieving 97.24% accuracy. These results highlight the advantage of ensemble boosting methods over traditional classifiers in medical imaging.

The proposed system's high accuracy and generalizability suggest potential real-world applications, including assisting ophthalmologists in early glaucoma detection, deploying the model as an AI-powered diagnostic tool in telemedicine, and integrating it with automated retinal screening systems in hospitals.

6. Conclusion

This research presents a hybrid AI-driven glaucoma detection framework that integrates deep learning-based feature extraction (ResNet50, VGG-16), dimensionality reduction (PCA), and advanced classification (XGBoost). The proposed model achieved an accuracy of 97.24%, significantly improving upon traditional machine learning methods. The study also demonstrated the effectiveness of PCA in reducing feature dimensions and the importance of SMOTE in addressing class imbalance. XGBoost outperformed Random Forest, highlighting the significance of gradient boosting in medical image analysis.

Approaches	Models	Accuracy	Precision	Recall	F1-score	AUC-PR
Pre-trained CNN	ResNet50	0.84	0.82	0.79	0.81	0.93
	MobileNet	0.63	0.66	0.92	0.77	0.34
	Inception v2	0.65	0.67	0.95	0.79	0.27
	EfficientNet	0.66	0.68	0.92	0.79	0.37

Transformer	ViT	0.63	0.67	0.91	0.77	0.32
	Swin	0.66	0.68	0.93	0.78	0.32
Hybrid models	ResNet50-PCA-XGBoot-VGG16	97.24%	95.00%	100.00%	97.4%	0.98
	ResNet50-PCA-Random Forest-VGG16	96.55%	93.00%	100.00%	97.00%	0.99

Future research can explore using Vision Transformers (ViTs) for further performance improvements, self-supervised learning techniques to reduce reliance on labelled datasets, and Explainable AI (XAI) models to provide clinicians with interpretable results. In conclusion, this study demonstrates the potential of deep learning and ensemble machine learning for automated glaucoma detection, offering a reliable, high-accuracy solution for early diagnosis and improved clinical decision-making.

References

- Lenka, S., Mayaluri, Z. L., & Panda, G. (2025). Glaucoma detection from retinal fundus images using graph convolution-based multi-task model. *E-Prime - Advances in Electrical Engineering, Electronics and Energy*, 11, 100931. <https://doi.org/10.1016/j.prime.2025.100931>
- Akter, N., Fletcher, J., Perry, S., et al. (2022). Glaucoma diagnosis using multi-feature analysis and a deep learning technique. *Scientific Reports*, 12, 8064. <https://doi.org/10.1038/s41598-022-12147-y>
- Civit-Masot, J., Domínguez-Morales, M. J., Vicente-Díaz, S., & Civit, A. (2020). Dual Machine-Learning System to Aid Glaucoma Diagnosis Using Disc and Cup Feature Extraction. *IEEE Access*, 8, 127519-127529. <https://doi.org/10.1109/ACCESS.2020.3008539>
- Thompson, A. C., Jammal, A. A., & Medeiros, F. A. (2020). A Review of Deep Learning for Screening, Diagnosis, and Detection of Glaucoma Progression. *Translational Vision Science & Technology*, 9(2), 42. <https://doi.org/10.1167/tvst.9.2.42>
- Haja, S. A., & Mahadevappa, V. (2023). Advancing glaucoma detection with convolutional neural networks: a paradigm shift in ophthalmology. *Romanian Journal of Ophthalmology*, 67(3), 222-237. <https://doi.org/10.22336/rjo.2023.39>