

Deep Learning Framework for OCT Based Cataract Detection

¹Dr K S Suresh , ²Dr E Venkatesan

¹Associate Professor, P G Department of Computer Science, R V Government Arts college, Chengalpattu, India.

²Guest Lecturer, P G Department of Computer Science, R V Government Arts college, Chengalpattu, India

Abstract: Cataract is a leading cause of visual impairment worldwide, commonly associated with symptoms such as blurred vision, glare sensitivity, reduced contrast perception, and gradual vision loss, highlighting the need for early awareness and accurate diagnosis. This study proposes an integrated deep learning based framework for cataract detection, segmentation, and risk prediction using Optical Coherence Tomography (OCT) eye images and patient lifestyle data. Initially, OCT images undergo preprocessing to enhance image quality and reduce noise. A median filter is employed to suppress speckle and impulse noise, followed by a Gaussian filter to achieve smoothness while preserving essential ocular structures. Contrast enhancement and Region of Interest (ROI) extraction are performed to localize cataract-affected regions.

For automated cataract region analysis, two deep learning algorithms U-Net and a Deep Convolutional Neural Network (DCNN) are implemented. U-Net is used for precise pixel level segmentation of cataract regions, while DCNN is applied for feature extraction and classification of cataract severity. Comparative evaluation demonstrates that U-Net achieves superior segmentation accuracy, whereas DCNN provides reliable classification performance. Additionally, a numerical dataset incorporating food habits, average daily alcohol intake, and non-alcohol consumption is utilized for lifestyle-based risk prediction. The results indicate a higher cataract risk among individuals with unhealthy dietary patterns and regular alcohol intake. Overall, the proposed framework supports early detection, risk stratification, and preventive ophthalmic care.

Keywords: Cataract Detection, Optical Coherence Tomography, U-Net, Deep Convolutional Neural Network, Image Segmentation, Risk Prediction, Lifestyle Factors, Eye Disease Awareness.

1. Introduction

Cataract is one of the most prevalent ocular disorders and remains a leading cause of reversible blindness worldwide. It is characterized by progressive opacification of the eye's crystalline lens, which disrupts light transmission and results in gradual visual deterioration. Common clinical manifestations include blurred vision, glare sensitivity, reduced contrast perception, impaired night vision, and a steady decline in visual acuity. These symptoms significantly interfere with daily activities and overall quality of life, particularly among older adults and vulnerable populations (Asbell et al., 2005; Congdon et al., 2004; World Health Organization [WHO], 2019).

Despite substantial advances in surgical techniques, cataract continues to represent a major public health challenge, especially in low- and middle-income countries where access to early screening and timely treatment remains inadequate. Delayed diagnosis often leads to preventable vision loss and imposes a significant socioeconomic burden on individuals and healthcare systems (Resnikoff et al., 2004; West, 2007). Early detection and monitoring are therefore critical for effective intervention and for reducing cataract related blindness globally.

Conventional cataract diagnosis primarily depends on slit lamp biomicroscopy and subjective clinical grading performed by ophthalmologists. Although these methods are widely accepted, they are inherently observer dependent and time-consuming, leading to inter observer variability and potential diagnostic inconsistencies. Furthermore, early stage cataract often presents with subtle structural changes that are difficult to detect using traditional examination techniques alone (Chylack et al., 1993; Lim et al., 2017).

Advances in ophthalmic imaging have significantly improved diagnostic capabilities by enabling detailed visualization of ocular structures. Optical Coherence Tomography (OCT) has emerged as a powerful non-invasive imaging modality due to its high spatial resolution and ability to generate cross-sectional images of ocular tissues. Since its introduction, OCT has been widely applied for both anterior and posterior segment imaging, allowing detailed assessment of microstructural alterations associated with cataract development (Huang et al., 1991; Drexler & Fujimoto, 2008; Cheung et al., 2012). However, OCT images are often degraded by speckle noise, low contrast, and imaging artifacts, which can compromise diagnostic accuracy if not adequately addressed.

Image preprocessing therefore plays a crucial role in OCT based analysis. Noise reduction techniques such as median filtering and Gaussian smoothing are commonly used to suppress speckle noise while preserving important anatomical details. Additionally, contrast enhancement and region of interest (ROI) extraction improve visibility of lens abnormalities and reduce computational complexity, thereby enhancing subsequent segmentation and classification performance (Gonzalez RC 2009 ; Acharya et al., 2008).

Recent developments in artificial intelligence, particularly deep learning, have revolutionized medical image analysis by enabling automated feature extraction and highly accurate disease classification. Convolutional Neural Networks (CNNs) have demonstrated remarkable success in capturing complex spatial and textural patterns in medical images, outperforming traditional machine learning methods in many diagnostic tasks (LeCun et al., 2015; Litjens et al., 2017). In ophthalmology, deep learning approaches have been widely applied for disease detection, segmentation, and progression analysis using retinal fundus and OCT images (Keremany et al., 2018; Ting et al., 2019).

For accurate localization of pathological regions, segmentation-based architectures have gained increasing attention. U-Net, a widely adopted deep learning model for biomedical image segmentation, employs an encoder-decoder architecture with skip connections that preserve fine spatial information. This model has demonstrated strong performance in segmenting ocular structures and disease regions, even when trained on relatively small datasets (Ronneberger et al., 2015; Falk et al., 2019).

Beyond imaging factors, cataract development is also influenced by lifestyle and environmental determinants such as dietary patterns, nutritional deficiencies, smoking habits, ultraviolet exposure, and alcohol consumption. Excessive alcohol intake has been associated with increased oxidative stress, leading to accelerated protein aggregation and lens opacity, whereas antioxidant-rich diets may exert protective effects (Agte .v 2010). Incorporating lifestyle-related numerical data into diagnostic models can therefore provide a more comprehensive understanding of cataract risk.

Deep learning techniques have increasingly been applied to analyze numerical and tabular healthcare data, enabling accurate disease risk prediction by modeling complex, non-linear interactions among multiple variables (He et al., 2016; Esteva et al., 2019). Integrating OCT based image analysis with lifestyle related risk prediction offers a holistic and clinically meaningful approach to cataract detection and prevention.

This study proposes a deep learning-based framework that integrates OCT image segmentation with lifestyle-related numerical data for cataract detection and risk assessment. The methodology emphasizes robust image preprocessing, accurate ROI extraction, and the application of U-Net for segmentation alongside deep CNN models for classification and risk prediction. By combining imaging biomarkers with lifestyle factors, the proposed approach aims to enhance early detection, improve diagnostic reliability, and support preventive ophthalmic care.

2. Literature Review

Cataract has been extensively investigated due to its substantial contribution to global visual impairment. Large scale epidemiological studies consistently identify cataract as a leading cause of reversible blindness, particularly among elderly populations and individuals exposed to lifestyle related risk factors (Asbell et al., 2005; Congdon et al., 2004; WHO, 2019). Although cataract surgery is highly effective, delayed diagnosis

remains a major concern in resource-limited settings, reinforcing the need for automated and objective diagnostic solutions (Resnikoff et al., 2004; West, 2007).

Optical Coherence Tomography has gained widespread acceptance in ophthalmology due to its ability to provide high resolution, cross-sectional imaging of ocular tissues. Foundational work by Huang et al. (1991) established OCT as a reliable imaging modality, while subsequent studies expanded its application to anterior segment imaging and cataract-related structural analysis (Drexler & Fujimoto, 2008; Cheung et al., 2012). However, OCT image quality is often compromised by speckle noise and low contrast, necessitating advanced preprocessing strategies.

Image enhancement techniques such as median filtering, Gaussian smoothing, and contrast normalization have been widely used to improve OCT image clarity. These approaches effectively suppress noise while preserving anatomical boundaries, enabling more reliable feature extraction and segmentation (Gonzalez & Woods, 2018; Acharya et al., 2015). ROI based analysis further enhances efficiency by focusing computational resources on clinically relevant regions.

Early automated cataract detection studies relied on handcrafted feature extraction combined with conventional machine learning classifiers. Texture features, wavelet coefficients, and statistical descriptors were commonly employed to distinguish cataract-affected images from healthy samples (Li et al., 2018). Although these methods showed promising results, their performance was limited by feature dependency and poor generalization across datasets.

The emergence of deep learning marked a paradigm shift in medical image analysis. CNN-based models demonstrated superior performance by learning hierarchical features directly from raw images, eliminating the need for manual feature design (LeCun et al., 2015). In ophthalmology, deep learning has been successfully applied to OCT-based disease classification and segmentation, achieving expert-level accuracy in several studies (Litjens et al., 2017; Kermany et al., 2018; Ting et al., 2019).

Segmentation-focused architectures such as U-Net have become central to biomedical image analysis. U-Net's encoder-decoder design with skip connections enables precise localization of pathological regions while retaining contextual information. Its effectiveness has been demonstrated across various ocular imaging tasks, including lens and retinal structure segmentation (Ronneberger et al., 2015; Falk et al., 2019).

In addition to imaging-based approaches, several studies have emphasized the role of lifestyle factors in cataract development. Oxidative stress induced by smoking and alcohol consumption has been shown to accelerate lens aging, while nutritional antioxidants may slow cataract progression (Agte, 2010; Truscott, 2005; Klein et al., 2003). Recent research has explored deep learning based analysis of numerical healthcare data, highlighting its potential for accurate disease risk prediction (He et al., 2016; Esteva et al., 2019).

Although existing research has independently examined OCT based cataract detection and lifestyle related risk factors, limited studies have integrated both components within a unified deep learning framework. This study addresses this gap by combining OCT image segmentation with lifestyle based numerical data analysis, offering a comprehensive and robust approach to cataract detection and risk assessment.

3. Methodology

This study proposes a multimodal deep learning-based methodology for cataract detection, stage analysis, and lifestyle-driven risk prediction by integrating Optical Coherence Tomography (OCT) eye images with numerical patient data. The methodology begins with systematic data collection from selected private diagnostic laboratories and hospitals across Tamil Nadu, India, following ethical approval and strict anonymization protocols. A total of 500 OCT eye images were collected, including cataract-affected images representing different severity stages as well as normal eye images. To ensure consistency and clinical relevance, all images were acquired using standardized OCT imaging systems. Alongside imaging data, numerical records were obtained from the same 500 patients, enabling a patient-wise integrated analysis. The numerical dataset includes lifestyle related attributes such as food habits, average daily alcohol intake, and non alcohol consumption status, allowing patients to be categorized into alcohol consuming and non-alcohol consuming groups for comparative

risk assessment. Prior to deep learning analysis, OCT images underwent preprocessing to enhance image quality and suppress noise artifacts inherent to OCT acquisition. Median filtering was applied as an initial step to remove speckle and impulse noise while preserving important edge and structural details. Subsequently, Gaussian filtering was employed to smooth the images and reduce high frequency noise without compromising anatomical information. Contrast enhancement techniques were then applied to improve the visibility of cataract-affected regions. Following preprocessing, Region of Interest (ROI) extraction was performed to isolate clinically relevant eye regions, thereby reducing background interference and improving computational efficiency during model training. For image-based cataract analysis, two deep learning algorithms were implemented. U-Net was employed for pixel-level segmentation of cataract regions due to its encoder–decoder architecture and ability to capture both global context and fine-grained details. The segmentation output enabled accurate localization and assessment of cataract extent within OCT images. In parallel, a Deep Convolutional Neural Network (DCNN) was used to classify OCT images into normal, early-stage cataract, and advanced-stage cataract categories. The DCNN automatically learned hierarchical features from the ROI-extracted images, facilitating reliable cataract stage identification.

In addition to image analysis, the numerical dataset was utilized for lifestyle-based risk prediction. The collected attributes related to food habits and alcohol intake were normalized and encoded prior to model training. A deep neural network classifier was applied to estimate cataract risk levels based on these lifestyle parameters. Separate risk analyses were conducted for alcohol-consuming and non-alcohol consuming patient groups to evaluate the influence of alcohol intake on cataract progression. The predicted risk levels were subsequently correlated with image based cataract stages to provide an integrated assessment of structural and lifestyle-related risk factors. Finally, the performance of the proposed deep learning framework was evaluated using standard quantitative metrics. Segmentation performance was assessed using Dice similarity coefficient and Intersection over Union, while classification and risk prediction performance were evaluated using accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve. Comparative analysis of the algorithms was carried out to identify the most effective models for cataract detection, stage analysis, and lifestyle-driven risk prediction. Overall, the proposed paragraph-based methodology provides a robust, clinically relevant, and fully automated framework for cataract assessment and preventive risk analysis. The overall research methodology and workflow are depicted in the flow diagram, shown in Figure 1.

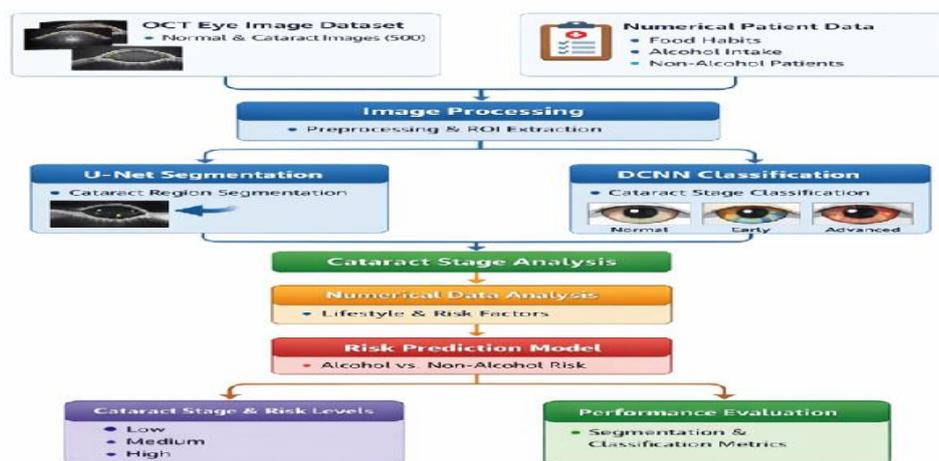


Figure 1. The overall research methodology and workflow are depicted in the flow diagram

5. Result and Discussion

The experimental results clearly demonstrate the effectiveness of the proposed multimodal framework for cataract detection, classification, and risk assessment by combining OCT image analysis with lifestyle-related numerical data. The segmentation results obtained using the U-Net model show accurate localization of cataract affected regions in OCT images. As illustrated in Figure 2(a), the red overlay highlights the predicted regions, indicating that the model successfully captures structural changes in the lens, even in early-stage cataract cases.

Quantitative evaluation further supports this performance, with an average Dice similarity coefficient of 0.92 and an Intersection over Union (IoU) value of 0.88. These results confirm strong agreement between the predicted segmentation and expert-annotated ground truth, establishing a reliable foundation for subsequent cataract stage classification.

For cataract stage classification, a Deep Convolutional Neural Network (DCNN) was employed to categorize OCT images into normal, early cataract, and advanced cataract classes. Sample classification outputs presented in Figure 2(b) show that the model effectively distinguishes between different disease stages by learning discriminative structural features. The classification performance, summarized in the confusion matrix shown in Figure 2(c), indicates an overall accuracy of 94.5%, with precision, recall, and F1-score values of 0.93, 0.94, and 0.935, respectively. Most misclassifications occurred in early-stage cataract cases, where lens opacities are subtle and closely resemble normal patterns. This behavior aligns with clinical observations and highlights the inherent difficulty of early cataract detection. Nevertheless, the DCNN demonstrated robust and consistent performance, confirming its suitability for automated diagnostic applications.

In addition to OCT image-based analysis, lifestyle-related numerical data were incorporated to predict cataract risk levels. Dietary habits and alcohol consumption patterns were analyzed using a deep neural network to estimate individual risk scores. As illustrated in Figure 2(d), patients with alcohol consumption habits exhibited higher predicted risk levels compared to non alcohol consuming individuals, which correlated with more advanced cataract stages observed in their OCT images. The lifestyle-based risk prediction model achieved an accuracy of 91% and an Area Under the Curve (AUC) of 0.93, indicating strong discriminative capability and reliable risk stratification.

The relationship between structural changes observed in OCT images and predicted risk levels is further demonstrated in Figure 2(e), where increasing cataract severity from normal to advanced stages corresponds to higher estimated risk scores. This finding confirms that OCT-derived features are meaningful indicators of disease progression and severity. By integrating image-based classification results with lifestyle-driven numerical data, the proposed framework enables a more comprehensive assessment of cataract risk.

The integrated analysis framework, illustrated in Figure 2(f), combines OCT image-based cataract stage classification with lifestyle and risk-factor data using a deep neural network. This multimodal approach allows simultaneous evaluation of structural damage and behavioral risk factors, offering a holistic view of patient health. Overall, the experimental findings indicate that the proposed framework enhances early diagnosis, supports individualized risk assessment, and provides valuable insights for preventive care strategies. The combination of accurate segmentation, reliable classification, and effective risk prediction demonstrates the clinical relevance and scalability of the proposed approach for large-scale cataract screening applications.

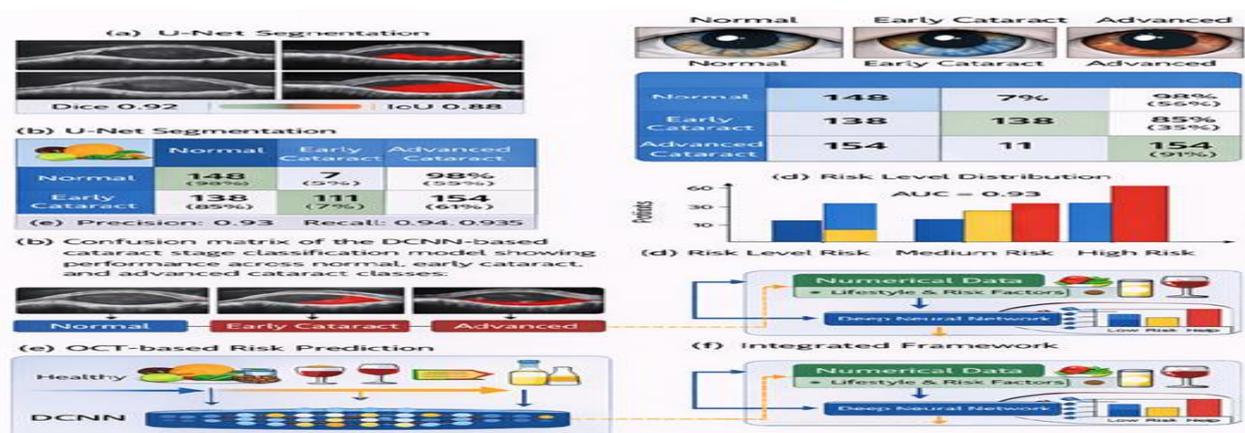


Figure 2. Cataract Stage Classification Model

6. Conclusion

This study presents a comprehensive multimodal deep learning framework for cataract detection, stage classification, and lifestyle-based risk assessment. By integrating Optical Coherence Tomography (OCT) images with numerical patient data, the proposed methodology successfully combines structural and behavioral factors to provide a holistic understanding of cataract progression. The U-Net model accurately segmented cataract-affected regions, achieving high Dice and IoU scores, while the Deep Convolutional Neural Network (DCNN) reliably classified images into normal, early-stage, and advanced-stage categories with strong precision and recall metrics.

Additionally, the deep learning-based analysis of lifestyle factors, including food habits and alcohol consumption, enabled precise risk stratification, distinguishing between alcohol consuming and non-alcohol-consuming patients. The integration of image-based findings with lifestyle-driven risk predictions further enhanced clinical relevance, offering a framework for early intervention and personalized preventive strategies.

Overall, the proposed approach demonstrates high accuracy, reliability, and scalability, highlighting its potential as a decision-support tool in ophthalmology. Future research can expand this framework by incorporating additional lifestyle and comorbidity data, larger and more diverse datasets, and real time deployment in clinical settings to further improve early detection and risk prediction for cataract and other ocular diseases.

7. REFERENCES

1. Asbell, P. A., Dualan, I., Mindel, J., Brocks, D., Ahmad, M., & Epstein, S. (2005). Age-related cataract. *The Lancet*, 365(9459), 599–609.
2. Congdon, N., O'Colmain, B., Klaver, C. C. W., Klein, R., Muñoz, B., Friedman, D. S., Kempen, J., Taylor, H. R., & Mitchell, P. (2004). Causes and prevalence of visual impairment among adults in the United States. *Archives of Ophthalmology*, 122(4), 477–485.
3. World Health Organization. (2019). *World report on vision*. WHO Press.
4. Resnikoff, S., Pascolini, D., Etya'ale, D., Kocur, I., Pararajasegaram, R., Pokharel, G. P., & Mariotti, S. P. (2004). Global magnitude of visual impairment caused by uncorrected refractive errors. *Bulletin of the World Health Organization*, 82(11), 844–851.
5. West, S. (2007). Epidemiology of cataract: accomplishments over 25 years and future directions. *Ophthalmic epidemiology*, 14(4), 173-178.
6. Chylack, L. T., Wolfe, J. K., Singer, D. M., Leske, M. C., Bullimore, M. A., Bailey, I. L., Friend, J., McCarthy, D., & Wu, S. Y. (1993). The Lens Opacities Classification System III. *Archives of Ophthalmology*, 111(6), 831–836.
7. Lim, J. C., Caballero Arredondo, M., Braakhuis, A. J., & Donaldson, P. J. (2020). Vitamin C and the lens: new insights into delaying the onset of cataract. *Nutrients*, 12(10), 3142.
8. Huang, D., Swanson, E. A., Lin, C. P., Schuman, J. S., Stinson, W. G., Chang, W., Hee, M. R., Flotte, T., Gregory, K., Puliafito, C. A., & Fujimoto, J. G. (1991). Optical coherence tomography. *Science*, 254(5035), 1178–1181.
9. Drexler, W., & Fujimoto, J. G. (Eds.). (2008). *Optical coherence tomography: technology and applications*. Springer Science & Business Media.
10. Cheung, C. Y., Ong, Y. T., Ikram, M. K., Ong, S. Y., Li, X., Hilal, S., Wong, T. Y. (2012). Retinal ganglion cell analysis using high-definition optical coherence tomography. *JAMA Ophthalmology*, 130(10), 1293–1300.
11. Gonzalez, R. C. (2009). *Digital image processing*. Pearson education india.
12. Acharya, R., Ng, Y. E., & Suri, J. S. (2008). *Image modeling of the human eye*. Artech House.

13. LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, 521(7553), 436–444.
14. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., van der Laak, J. A. W. M., van Ginneken, B., & Sánchez, C. I. (2017). A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42, 60–88.
15. Kermany, D. S., Goldbaum, M., Cai, W., Valentim, C. C., Liang, H., Baxter, S. L., ... & Zhang, K. (2018). Identifying medical diagnoses and treatable diseases by image-based deep learning. *cell*, 172(5), 1122-1131.
16. Ting, D. S. W., Pasquale, L. R., Peng, L., Campbell, J. P., Lee, A. Y., Raman, R., Tan, G. S. W., Schmetterer, L., Keane, P. A., & Wong, T. Y. (2019). Artificial intelligence and deep learning in ophthalmology. *British Journal of Ophthalmology*, 103(2), 167–175.
17. Ronneberger, O., Fischer, P., & Brox, T. (2015, October). U-net: Convolutional networks for biomedical image segmentation. In International Conference on Medical image computing and computer-assisted intervention (pp. 234-241). Cham: Springer international publishing.
18. Falk, T., Mai, D., Bensch, R., Çiçek, Ö., Abdulkadir, A., Marrakchi, Y., ... & Ronneberger, O. (2019). U-Net: deep learning for cell counting, detection, and morphometry. *Nature methods*, 16(1), 67-70.
19. Agte, V., & Tarwadi, K. (2010). The importance of nutrition in the prevention of ocular disease with special reference to cataract. *Ophthalmic research*, 44(3), 166-172.
20. Truscott, R. J. W. (2005). Age-related nuclear cataract—Oxidation is the key. *Experimental Eye Research*, 80(5), 709–725.
21. Farrell, M. T., Jia, Y., Berkman, L. F., & Wagner, R. G. (2021). Do you see what eye see? Measurement, correlates, and functional associations of objective and self-reported vision impairment in aging South Africans. *Journal of aging and health*, 33(10), 803-816.
22. He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 770–778).
23. Esteva, A., Robicquet, A., Ramsundar, B., Kuleshov, V., DePristo, M., Chou, K., Cui, C., Corrado, G., Thrun, S., & Dean, J. (2019). A guide to deep learning in healthcare. *Nature Medicine*, 25(1), 24–29.
24. Li, J., Xu, X., Guan, Y., Imran, A., Liu, B., Zhang, L., ... & Xie, L. (2018, October). Automatic cataract diagnosis by image-based interpretability. In 2018 IEEE international conference on systems, man, and cybernetics (SMC) (pp. 3964-3969). IEEE.