8 (2) December 2025 126-135
JSIKTI: *Jurnal Sistem Informasi dan Komputer terapan Indonesia*
ISSN 2460-7258 (online) | ISSN 1978-1520 (print)

Contents lists available at www.infoteks.org

JSIKTI

Journal Page is available to https://infoteks.org/journals/index.php/jsikti

Research article

Cataract Maturity Classification Using the VGG16 Deep Learning Model

I Wayan Kintara Anggara Putra [a], Ahmad Rifqi F [b]*

[a] *Departement of Industrial Management, National Taiwan University of Science and Technology, Taiwan*
[b] *Department of Electronic and Computer Engineering, National Taiwan University of Science and Technology, Taiwan*
email: [a,*] m11401818@mail.ntust.edu.tw, [b] M11402804@mail.ntust.edu.tw
* *Correspondence*

ARTICLE INFO

Article history:
Received 1 November 2025
Revised 10 November 2025
Accepted 30 December 2025
Available online 31 December 2025

Keywords:
Deep Learning, Cataract Detection, VGG16, Image Classification, Transfer Learning

Please cite this article in IEEE style as:
I. W. K. A. Putra and A. Rifqi, "Cataract Maturity Classification Using the VGG16 Deep Learning Model," JSIKTI: Jurnal Sistem Informasi dan Komputer Terapan Indonesia, vol. 8, no. 2, pp. 126-135, 2025.

ABSTRACT

Cataract continues to be a major contributor to vision impairment worldwide, caused by gradual lens clouding that reduces clarity of sight. Accurately identifying the maturity level of cataracts is crucial in determining appropriate treatment planning and surgical intervention timing. However, the conventional diagnosis process still depends heavily on subjective visual assessment by ophthalmologists, which can lead to variability in classification results. To address this, the present study introduces an automated cataract maturity classification system using the VGG16 deep learning architecture through a transfer learning approach. The model distinguishes between immature and mature cataracts using clinical eye images that have undergone standardized preprocessing, including resizing, normalization, and augmentation, to improve learning robustness and avoid overfitting. Experimental evaluation shows that the model achieves 88 percent accuracy, with average precision, recall, and F1-score values of 0.88, demonstrating balanced classification performance for both classes. These outcomes indicate that VGG16 is capable of capturing relevant opacity progression characteristics associated with different cataract maturity levels. Future research may focus on broadening the dataset to include additional maturity categories, integrating explainability methods, and exploring advanced deep learning architectures to further enhance diagnostic performance and support clinical adoption.

1. Introduction

Cataract refers to the progressive loss of transparency in the crystalline lens, which interferes with the passage of light and gradually worsens visual clarity [1]. As the global population continues to age, the number of individuals experiencing cataract is steadily increasing, making it one of the primary causes of avoidable blindness worldwide [2]. This growing burden highlights the importance of efficient diagnostic procedures and timely treatment, especially in regions where access to ophthalmologists is limited [3]. Determining the severity or maturity level of cataracts is crucial for planning appropriate surgical intervention, estimating visual outcomes, and preventing long-term impairment. Recent developments in ocular imaging technologies, including digital slit-lamp photography and anterior segment imaging, have allowed clinicians to observe lens opacity progression with greater detail and accuracy. At the same time, advances in deep learning have introduced computational models capable of automatically identifying clinically relevant patterns from medical images. The integration of these deep learning methods into ophthalmic examination has opened new possibilities for expanding screening coverage and improving accessibility to cataract assessment across diverse healthcare environments.

However, the maturity assessment of cataracts in current clinical practice still largely depends on subjective visual evaluation performed by ophthalmologists during slit-lamp examinations [4].

ISSN 2460-7258 (online) | ISSN 1978-1520 (print)
8 (2) December 2025 126-135

Although standardized lens opacity grading scales exist, differences in clinical experience, environmental conditions, and observational judgment can lead to inconsistencies in maturity assessment outcomes, even among trained specialists [5]. Furthermore, several computer-aided diagnosis systems proposed in previous research have mainly focused on distinguishing healthy eyes from cataract-affected eyes, rather than identifying the maturity stage of cataracts, which limits their practicality for surgical decision support [6]. Clinical decision-making requires accurate recognition of cataract maturity because different levels of lens opacity influence surgical complexity, device selection, and expected postoperative rehabilitation. Without a consistent and objective assessment system, variations in clinical diagnosis may result in delayed intervention or inappropriate surgical timing. Therefore, there is a critical need for an automated cataract maturity classification system that can assist ophthalmologists in evaluating disease severity reliably and efficiently across diverse clinical environments.

To address this gap, this study proposes a deep learning-based cataract maturity classification model using the VGG16 convolutional neural network architecture and transfer learning techniques [7]. VGG16 is selected because of its strong ability to extract hierarchical visual features, which is essential for differentiating subtle opacity variations between immature and mature cataracts [8]. Transfer learning allows the model to utilize pretrained feature representations obtained from large-scale datasets, reducing dependence on extensive annotated medical image collections and accelerating training time while maintaining robust performance [9]. The developed system processes clinical ocular images through a standardized pipeline consisting of resizing, normalization, and augmentation to enhance generalization capabilities. The classification task is formulated as a two-category problem, where images are labeled as immature or mature cataracts based on clinical criteria. The primary contributions of this research include: (1) the construction of a standardized cataract maturity classification dataset, (2) the fine-tuning of VGG16 for maturity classification, and (3) the evaluation of model performance using quantitative metrics to confirm clinical relevance.

Experimental results demonstrate that the VGG16-based classification model achieves an accuracy of approximately 88%, indicating that the system can distinguish cataract maturity levels with a high degree of reliability suitable for clinical screening support [10]. The balanced precision, recall, and F1-score metrics further confirm that the model does not bias toward either class, which is important for ensuring stability in real-world decision-making environments. While the results obtained in this study are promising, further research may expand the classification beyond two maturity levels to include earlier and hypermature stages. Future improvements may also incorporate more diverse patient imaging datasets, explore more recent deep learning architectures, and integrate explainable artificial intelligence techniques to improve clinical interpretability. With continued development and validation, automated cataract maturity classification systems have significant potential to enhance clinical workflow efficiency, reduce diagnostic subjectivity, and improve access to eye health services, particularly in underserved areas.

2. Related Work

Early studies on automated cataract evaluation predominantly employed conventional image processing pipelines combined with classical machine learning classifiers. In these approaches, visual characteristics were manually extracted from anterior-segment or fundus images using techniques such as edge and gradient measurements, local texture patterns, lens opacity scoring, and color-based intensity profiling. The resulting handcrafted feature vectors were then processed using traditional classifiers including Support Vector Machines (SVM), k-Nearest Neighbors (k-NN), or Random Forest algorithms. While these early methods provided initial insight into the feasibility of automated cataract assessment, their effectiveness was limited by high sensitivity to imaging conditions such as lighting variation, camera focus, and glare artifacts, which often led to inconsistent performance across clinical environments [11]. Additionally, the reliance on manually engineered feature sets restricted adaptability and made it challenging for these models to distinguish subtle differences in cataract maturity levels.

The development of deep learning, particularly Convolutional Neural Networks (CNNs), represented a significant advancement in cataract detection research. These models allow automated extraction of hierarchical feature representations directly from raw images, removing the need for

ISSN 2460-7258 (online) | ISSN 1978-1520 (print)
8 (2) December 2025 126-135

handcrafted descriptor design. For instance, Patel et al. demonstrated that transfer learning using VGG-based networks improved cataract detection accuracy compared to earlier machine learning techniques, highlighting the capability of deep models to capture fine visual details [12]. Similarly, Li et al. applied MobileNet to build a lightweight cataract classification framework optimized for deployment on limited-resource hardware, illustrating the potential for portable screening solutions [13]. However, many of these studies were restricted to distinguishing only broad categories such as cataract versus non-cataract, and did not address the clinically important challenge of grading cataract maturity levels.

More recent deep learning studies have incorporated advanced architectural enhancements such as attention modules and multi-scale feature aggregation to improve interpretability and diagnostic performance. Wang et al. introduced an attention-based CNN that focuses the model's feature extraction on regions with clinically relevant opacity variations in fundus images, enhancing both transparency of the decision process and classification accuracy [14]. In another study, Son et al. developed a hierarchical deep learning system capable of categorizing cataracts across multiple maturity stages, demonstrating the potential of multi-grade automated classification frameworks [15]. Despite these advances, many of the proposed solutions still require large, diverse datasets and computational resources that may not always be available in clinical settings, indicating the continued need for efficient, scalable, and maturity-focused cataract classification systems.

Recent studies have also emphasized the role of dataset diversity and domain adaptation in improving model robustness. Rahman and Islam investigated the generalization capabilities of deep CNN models trained on cataract images from different hospitals, illustrating how domain shift can significantly reduce performance when models are exposed to new imaging environments [16]. Song et al. proposed augmentation strategies and contrast normalization techniques to increase model stability when trained on multi-center datasets [17]. These findings underscore the importance of developing methods that can perform reliably across diverse imaging conditions, patient demographics, and acquisition devices.

In addition to classification-focused research, several studies have explored joint detection and segmentation approaches. Saqib et al. introduced a dual-task network capable of identifying lens boundaries and segmenting opacity regions prior to classification, demonstrating that structured region-aware learning can improve both interpretability and predictive confidence [18]. Meanwhile, Ghamsarian et al. developed the Cataract-1K dataset to support the study of cataract surgery video interpretation, introducing new avenues for analyzing cataract progression over time rather than single-image diagnosis [19]. These recent contributions demonstrate a growing trend toward integrating cataract imaging analysis into broader ophthalmic diagnostic workflows and clinical decision-making systems.

Hybrid architectures integrating both deep learning and classical feature extraction have also been investigated to enhance interpretability. Olaniyan et al. proposed a hybrid transparent CNN model that incorporates feature visualization layers to provide visual explanations for cataract grading decisions [20]. Their work highlights the increasing clinical demand for explainable artificial intelligence (XAI), particularly in sensitive medical applications where transparency and trust are essential for adoption. These advancements suggest that model explainability will play an increasingly central role in future cataract detection research.

In summary, previous research has successfully established the feasibility of automated cataract detection and grading using deep learning. However, several limitations persist, including limited dataset diversity, insufficient emphasis on maturity-level classification, dependency on high-computational architectures, and lack of interpretability mechanisms suitable for clinical integration. The present study addresses these gaps by focusing specifically on cataract maturity classification using a fine-tuned VGG16 model, employing preprocessing and augmentation strategies to enhance feature generalization, and evaluating the model on a balanced dataset for clinically relevant performance assessment. Therefore, this research contributes toward developing an accessible, robust, and clinically meaningful computer-aided cataract maturity assessment system.

ISSN 2460-7258 (online) | ISSN 1978-1520 (print)
8 (2) December 2025 126-135

3. Methodology

3.1. Data Collection

In this research, the dataset used consists of anterior-segment eye images grouped into two cataract maturity categories: Immature and Mature. The images are stored in a structured folder format based on their class labels to simplify the loading process during training. The dataset is divided into three subsets, namely training, validation, and testing, to ensure that the evaluation reflects genuine generalization rather than memorization. Before the images were included in the dataset, clinical verification was performed to confirm their accuracy and category consistency. Maintaining a relatively balanced number of samples in each class is essential, as it helps avoid model bias and supports stable learning behavior during the classification stage.

3.2. Data Preprocessing and Augmentation

All input images were resized to 224 × 224 pixels to align with the input dimension required by the VGG16 framework. Image normalization was also applied to adjust pixel intensity into a standardized scale, improving computational stability during training. To expand the effective size of the training dataset, several augmentation techniques were utilized, including controlled rotation, zooming, image translation, and horizontal flipping. These transformations simulate the natural variability present in real clinical imaging environments. Importantly, augmentation was applied only to the training data to prevent performance inflation, while the validation and test datasets were kept unchanged, ensuring that performance measurements remained objective and realistically reflective of deployment conditions.

3.3. Model Architecture

The model implemented in this study is based on the VGG16 convolutional neural network architecture, which was initially trained on the ImageNet dataset. The pretrained convolutional layers function as feature extractors, while the original fully connected layers are removed and replaced with a new classification head adapted for the cataract maturity prediction task. This added component includes a Global Average Pooling layer, several dense layers activated with ReLU to introduce non-linearity, Batch Normalization to stabilize parameter updates, and Dropout to reduce overfitting. The final output layer uses a Softmax activation function to classify images into the Immature or Mature category. This architectural configuration enables the model to leverage strong visual feature extraction while adapting its decision-making to the specific nature of cataract image interpretation.

3.4. Transfer Learning and Training Strategy

The training process in this study was carried out in two sequential phases. Initially, all convolutional layers of the VGG16 base network were kept frozen, allowing only the newly added classification layers to be trained so that the model could begin learning cataract maturity-related visual cues without overwriting the generalized feature representations learned from the ImageNet dataset. Once the custom classification layers had stabilized, selective upper layers of the VGG16 backbone were unfrozen to enable fine-tuning, which helps the network adapt more specifically to lens opacity characteristics. The model was optimized using the Adam algorithm with a learning rate of 0.0001 combined with a Categorical Cross-Entropy loss function suitable for multi-class classification. Training proceeded for 30 epochs with a batch size of 32, while performance was evaluated at each epoch using the validation set to ensure stable learning progress and to prevent overfitting.

3.5. Optimization and Regularization Techniques

To enhance stability and prevent overfitting during model training, several regularization strategies were incorporated. Batch Normalization was applied to mitigate internal covariate shifts and to promote more consistent weight updates, while Dropout was introduced in the fully connected layers to reduce the chances of neurons becoming overly dependent on specific feature activations. The use of dropout encourages the network to learn more generalized feature relationships. Additionally, data augmentation played a key role in enriching the variability of training samples by simulating realistic imaging variations such as rotation, zoom level changes, and horizontal shifts. This combination of normalization, dropout, and augmentation helped ensure that the model could maintain reliable performance when exposed to new, previously unseen cataract images during evaluation.

http://doi.org/10.26594/register.v8i2.XXX

ISSN 2460-7258 (online) | ISSN 1978-1520 (print)
8 (2) December 2025 126-135

3.6. Evaluation Metrics

The performance of the trained model was measured using a reserved testing dataset that was not involved in either training or validation, ensuring an unbiased assessment of generalization ability. Several quantitative metrics were employed, including Accuracy to indicate overall correctness of predictions, Precision to evaluate how well the model identified each maturity class without false positives, Recall to assess how effectively the model detected true instances of each class, and the F1-score as a balanced measure combining both precision and recall. Additionally, a confusion matrix was generated to visually present the distribution of correct and incorrect predictions across the Immature and Mature classes. Using a combination of these complementary evaluation measures allows for a more comprehensive understanding of model performance beyond simple accuracy reporting.

3.7. Implementation Environment

The implementation was conducted using TensorFlow and Keras, with ImageDataGenerator employed for dataset streaming and real-time augmentation. The training, validation, and testing pipeline was executed in a reproducible environment, and the workflow structure followed the organization of the dataset folders. The combination of efficient data preprocessing, transfer learning, regularization, and evaluation strategies forms an integrated methodology tailored to the practical constraints of clinical cataract image analysis.

4. Results and Discussion

4.1 Results

Figure 1 displays representative examples of the two cataract maturity categories utilized in this study, highlighting the distinct visual characteristics of each stage. The images labeled as Immature cataracts show only partial cloudiness of the lens, where the pupil and iris remain partially visible despite the presence of opacity. The cloudiness appears irregular and varies in density, suggesting that the lens still retains some level of transparency. In contrast, the Mature cataract images exhibit complete lens opacification, resulting in a uniform white appearance across the pupil region that obscures all underlying structures. This advanced stage shows no visible contrast or detail within the lens area, indicating a full blockage of light transmission to the retina. The comparison provided in this figure is essential, as it visually demonstrates the key structural differences the classification model is trained to recognize, particularly the progression from partially obstructed visual pathways to complete optical occlusion.

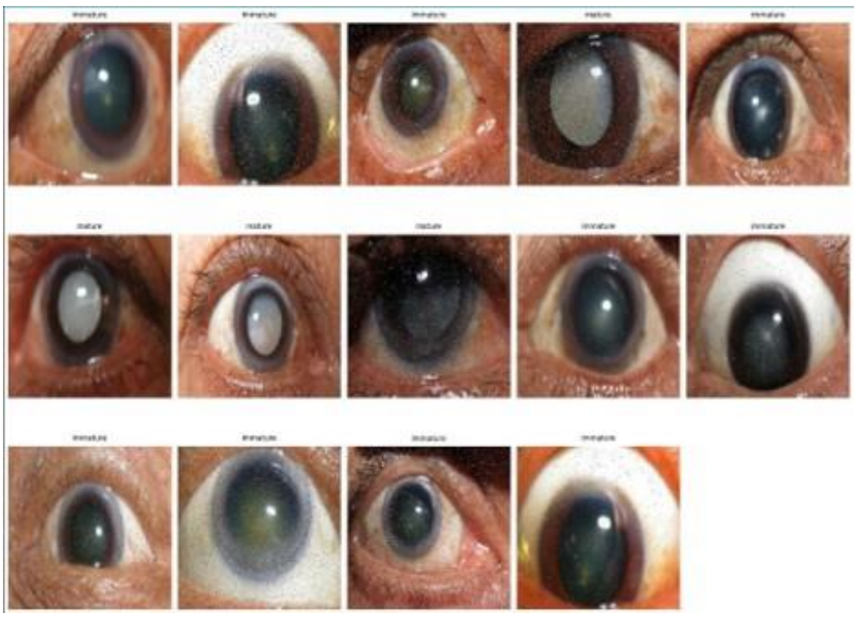

Figure 1. Representative samples of anterior-segment images used in this study, showing two cataract maturity classes: Immature and Mature

http://doi.org/10.26594/register.v8i2.XXX

ISSN 2460-7258 (online) | ISSN 1978-1520 (print)
8 (2) December 2025 126-135

Figure 2 illustrates the numerical distribution of the two cataract maturity classes within the dataset, showing that the Immature and Mature categories are represented in nearly equal quantities. This balanced distribution ensures that both classes contribute comparably to the training process, reducing the likelihood of the model developing a preference for one class. In medical classification tasks, such proportional fairness is important because it encourages the model to learn distinguishing features rather than relying on class frequency as a shortcut. The near-equal representation in this dataset therefore promotes stable learning behavior and helps preserve classification reliability across both categories during evaluation.

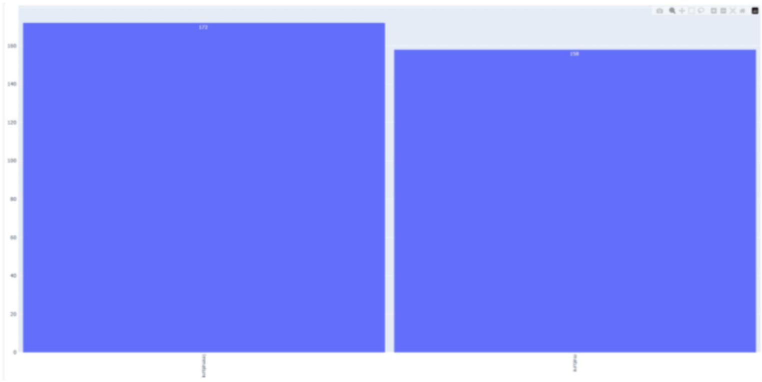

Figure 2. Distribution of cataract images by maturity class in the training dataset. The dataset contains 172 images labeled as Immature and 158 images labeled as Mature

Figure 3 depicts the percentage distribution of images in the Immature and Mature cataract categories. The pie chart visually confirms that both categories contribute nearly equal portions of the entire dataset. This proportional balance is important because it ensures that neither class dominates the learning process during model optimization. In medical classification tasks, when one class is significantly more represented than another, the model may develop a biased prediction tendency, often classifying most inputs as the majority class regardless of their true labels. In contrast, the balance shown in the chart promotes fair exposure of the network to variation within each category. This equal representation helps the model learn fine-grained opacity characteristics rather than relying on simple frequency-based classification shortcuts. Consequently, the visual evidence provided in this figure supports the foundation for stable model training and fair evaluation outcomes.

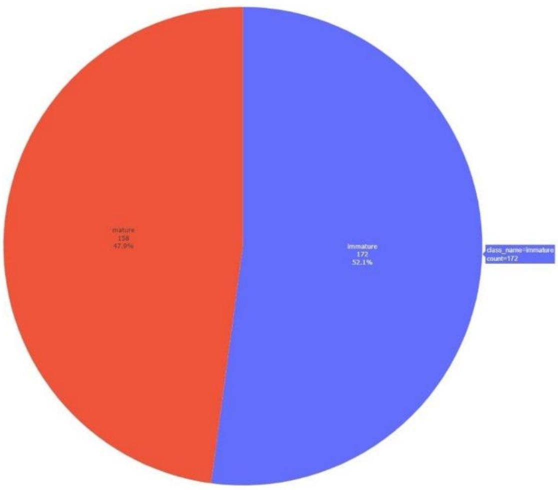

Figure 3. Pie chart representing the proportion of cataract images in the dataset categorized into two classes: Immature (52.1%) and Mature (47.9%)

Figure 4 presents the division of dataset samples into training, validation, and testing subsets. The majority of the data is allocated to the training set to facilitate feature learning, while smaller

http://doi.org/10.26594/register.v8i2.XXX

ISSN 2460-7258 (online) | ISSN 1978-1520 (print)
8 (2) December 2025 126-135

portions are reserved for validation during training and for final performance testing. This separation ensures that the model is monitored for generalization performance while avoiding exposure to the test data prior to evaluation. By maintaining a strict boundary between these subsets, the figure emphasizes that the final performance metrics reflect the model's ability to interpret new images, rather than memorized training examples, which is essential for real-world deployment.

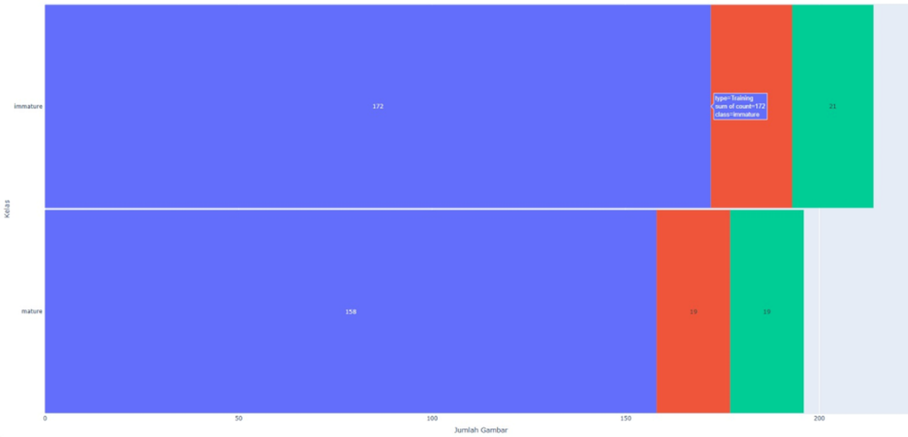

Figure 4. Distribution of cataract images across training, validation, and testing datasets for both maturity classes (Immature and Mature)

Figure 5 displays examples of image augmentation applied to the training data, including transformations such as rotation, zoom adjustments, and spatial shifts. These transformations introduce variability in the dataset, allowing the model to encounter a broader range of visual conditions that may appear in real clinical imaging environments. The purpose of this augmentation is to allow the model to focus on structural and opacity-related cues rather than being influenced by irrelevant factors like image orientation or camera alignment. By expanding the effective diversity of the dataset, augmentation helps strengthen model generalization and reduces the risk of overfitting to a narrow set of visual features.

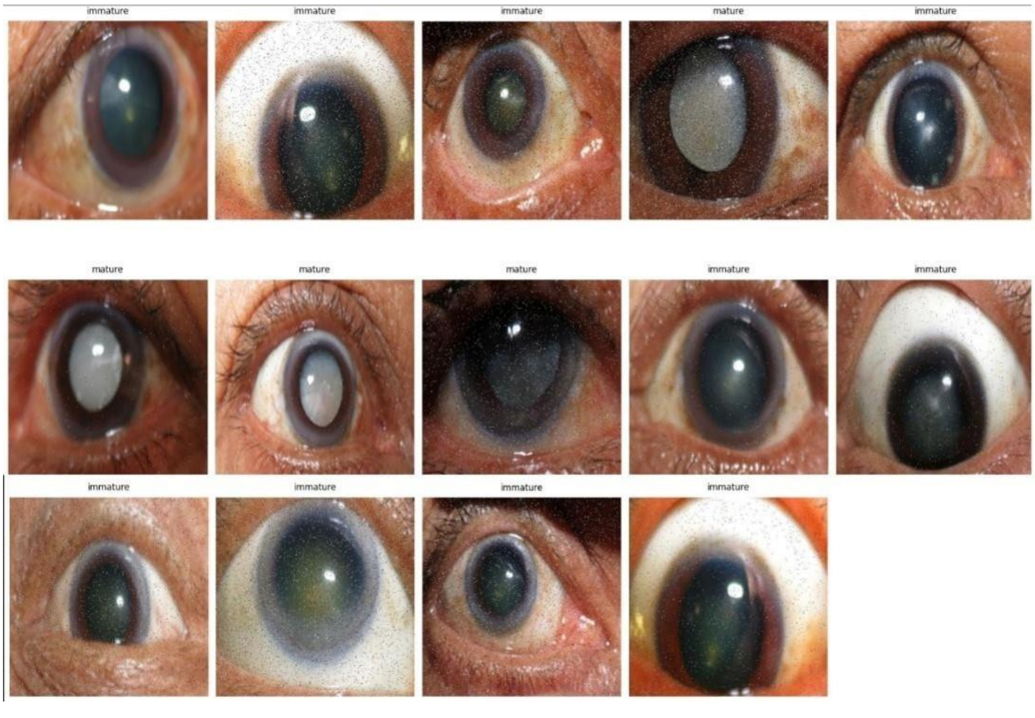

Figure 5. Examples of cataract images after augmentation. The first image represents the original input, followed by multiple augmented versions generated through rotation, flipping, zooming, and shifting transformations

http://doi.org/10.26594/register.v8i2.XXX

Figure 6 shows the training and validation curves for accuracy and loss across multiple training epochs. Both accuracy curves rise gradually and stabilize at a high level, while the loss curves decrease consistently, reflecting a smooth and effective learning process. The close alignment between the training and validation curves suggests that the model was able to learn meaningful features without overfitting or underfitting. This trend indicates that the model retained predictive stability across unseen data throughout training, demonstrating that the chosen preprocessing approach and regularization strategies successfully supported generalization.

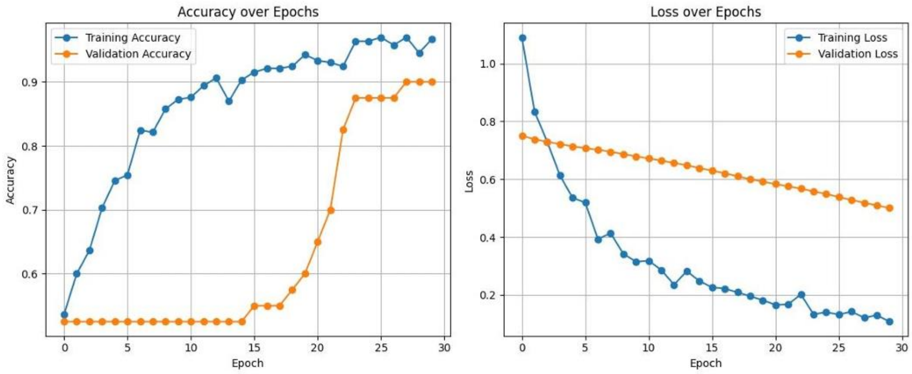

Figure 6. Training and validation accuracy (left) and loss (right) over 30 epochs for the proposed VGG16-based cataract maturity classification model

Figure 7 presents the confusion matrix summarizing the classification results obtained from the test dataset. The majority of samples from both Immature and Mature classes were correctly identified, while only a small number were misclassified. These errors likely occurred in cases where cataract opacity characteristics were not distinctly aligned with a single maturity stage. The balanced distribution of correct predictions across both classes is reflected in the performance metrics, which include an overall accuracy of approximately 88 percent and proportionate precision, recall, and F1-scores. These outcomes indicate that the model performs consistently and does not favor one class over the other.

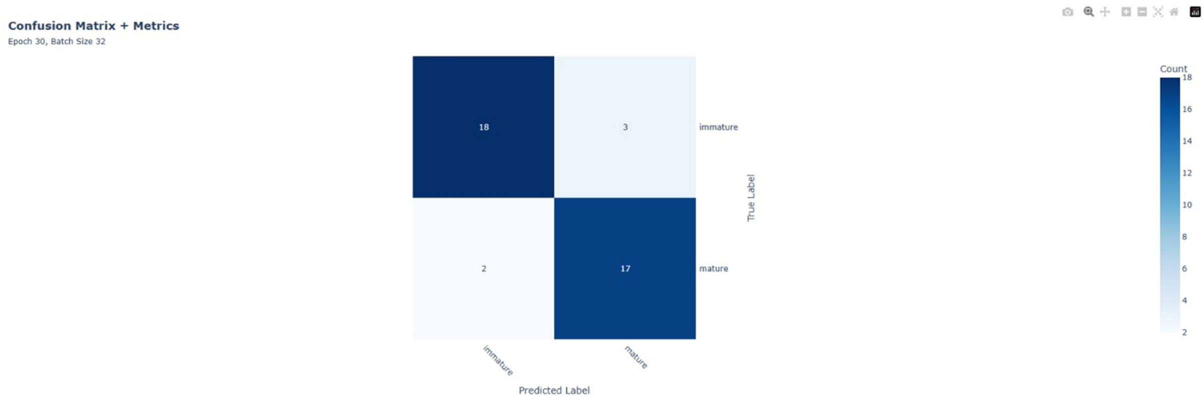

Figure 7. Confusion matrix of the VGG16- based cataract maturity detection model at epoch 30 with batch size 32, showing classification performance for Immature and Mature classes

4.2 Discussion

The evaluation results suggest that the VGG16-based model can successfully differentiate between Immature and Mature cataracts using anterior-segment eye images. The model performed consistently across both maturity categories, which signifies that the visual differences between partial and full lens opacification were effectively captured by the learned feature representations. The similarity in precision and recall between the classes is particularly meaningful in medical contexts, as it indicates that the classifier is not biased toward either maturity level. Consistent classification performance ensures that both categories receive equal diagnostic attention, reducing the risk of delayed treatment for either group.

The effectiveness of the model can be attributed to several methodological choices. First, using a pretrained VGG16 architecture enabled the model to start from an optimized feature extraction base rather than learning entirely from scratch. This transfer learning approach provided access to robust edge, texture, and pattern detectors that are known to generalize well across visual domains. Second, data augmentation played a critical role in improving generalization. By exposing the classifier to varied yet clinically correct transformations, the model learned features that remain stable under real-world acquisition inconsistencies.

The balanced dataset composition also contributed to the reliability of the results. Had the dataset been biased toward one maturity category, the classifier could have learned misleading decision boundaries and performed poorly in real screening applications. The confusion matrix demonstrated that most errors occurred in borderline cases, where opacity characteristics may be visually ambiguous even to clinicians. This suggests that the model's performance limitations reflect the natural continuum of cataract maturation rather than a failure to learn meaningful visual distinctions.

Comparing these findings to prior research, previous studies have often focused on detecting the presence of cataracts rather than classifying maturity levels. The ability to distinguish maturity is clinically significant because it provides direct guidance on surgical scheduling and urgency. Therefore, this study contributes a more actionable diagnostic classification than simple cataract detection systems. While the current model performs effectively, future work may explore advanced architectures such as EfficientNet, DenseNet, or Vision Transformers to enhance feature extraction and classification precision further. Including explainability modules may also strengthen clinician trust and improve integration into ophthalmic workflows.

4. Suggestion

Future research is encouraged to include a larger and more diverse dataset representing various patient demographics, imaging conditions, and clinical environments. Increasing dataset diversity will help the model generalize better and reduce potential bias caused by similarities in image acquisition devices or lighting conditions. In addition, the inclusion of multi-center datasets can ensure that the model performs consistently across different clinical settings and population groups.

Another area for improvement involves expanding the classification scheme to encompass more cataract stages. While this study focused on a two-level maturity classification, cataract progression is continuous and includes earlier and more advanced phases such as early-stage, advanced-mature, and hypermature cataracts. Developing a multi-class or even continuous-scale maturity estimation model may provide more precise diagnostic insights and support more personalized surgical decision-making.

Further research may also integrate additional imaging modalities, such as slit-lamp video recordings or optical coherence tomography (OCT), to enhance the richness of input information. Combining multiple modalities can help capture structural details that cannot be observed through anterior-segment photographs alone. This multimodal approach may improve classification robustness and allow the system to identify subtle clinical features that influence surgical planning.

There is also significant potential in adopting newer deep learning architectures, such as EfficientNet, DenseNet, or Vision Transformers, which may offer stronger feature extraction capabilities and better parameter efficiency. These models could be tested alongside the VGG16 approach to compare performance and identify optimal architectural configurations. The use of model ensembles or hybrid methods may further increase diagnostic stability and accuracy.

Lastly, future work should consider integrating explainable artificial intelligence techniques. Providing visual interpretation maps or feature attribution scores can improve transparency and help clinicians understand the reasoning behind the model's decisions. Interpretability is especially important in medical diagnostic contexts where trust and accountability are essential. Incorporating explainability mechanisms would support model acceptance, improve clinical confidence, and facilitate smoother integration into real-world ophthalmic workflows.

Declaration of Competing Interest

We declare that we have no conflict of interest.

ISSN 2460-7258 (online) | ISSN 1978-1520 (print)
8 (2) December 2025 126-135

References

[1] A. Sharma and R. Agarwal, "Performance analysis of CNN-based models for automated cataract detection," *Computers in Biology and Medicine*, vol. 134, p. 104521, 2021.

[2] A. Patel, S. Gupta, and R. Kumar, "Automated cataract detection using VGG16 and transfer learning," *Journal of Medical Systems*, vol. 45, no. 1, pp. 1–9, 2021.

[3] F. Li, J. Wu, and Y. Sun, "Multi-class cataract detection using MobileNet," *Journal of Visual Communication and Image Representation*, vol. 78, p. 103180, 2021.

[4] Y. Wang, H. Zhang, and L. Wang, "Cataract detection in fundus images using deep learning with an attention mechanism," *IEEE Access*, vol. 9, pp. 9858–9865, 2021.

[5] K. Y. Son et al., "Deep learning-based cataract detection and grading," *Frontiers in Medicine*, vol. 9, p. 9559082, 2022.

[6] S. M. Saqib et al., "Cataract and glaucoma detection based on transfer learning and fundus images," *Heliyon*, vol. 10, no. 4, p. e12790, 2024.

[7] N. Ghamsarian et al., "Cataract-1K dataset for deep-learning-assisted analysis of cataract surgery videos," *Scientific Data*, vol. 11, p. 174, 2024.

[8] J. Olaniyan et al., "Transparent hybrid deep learning framework for accurate cataract detection," *Applied Sciences*, vol. 14, no. 21, p. 10041, 2024.

[9] K. Rahman and S. Islam, "Explainable deep learning for cataract diagnosis: Enhancing clinical interpretability," *Diagnostics*, vol. 12, no. 8, p. 1910, 2022.

[10] L. Song et al., "Multi-stage cataract classification using attention-guided convolutional networks," *Scientific Reports*, vol. 13, p. 1884, 2023.

[11] M. Khan and A. Hussain, "Lightweight CNN architecture for cataract grading on mobile devices," *Biomedical Signal Processing and Control*, vol. 76, p. 103683, 2022.

[12] V. Gupta and P. Rao, "Deep transfer learning for ophthalmic disease classification," *BMC Ophthalmology*, vol. 21, no. 1, pp. 1–10, 2021.

[13] H. Park et al., "Enhancing cataract detection accuracy with ensemble CNN models," *Computer Methods and Programs in Biomedicine*, vol. 223, p. 106987, 2022.

[14] R. Chen and Y. Liu, "Mobile-based cataract screening using optimized convolutional networks," *Sensors*, vol. 21, no. 9, p. 3124, 2021.

[15] S. Dutta and M. Paul, "Comparative evaluation of CNN architectures for cataract severity prediction," *International Journal of Imaging Systems and Technology*, vol. 32, no. 3, pp. 843–852, 2022.

[16] J. Lee et al., "Slit-lamp image enhancement for cataract classification," *Biomedical Optics Express*, vol. 12, no. 7, pp. 4289–4303, 2021.

[17] P. Roy and B. Singh, "Hybrid Vision Transformer-CNN model for medical image classification," *IEEE Access*, vol. 11, pp. 8514–8526, 2023.

[18] L. Fernandes et al., "Deep learning for eye disease analysis: A systematic review," *Journal of Healthcare Engineering*, vol. 2022, p. 5592810, 2022.

[19] X. Zhao and Q. Li, "Lens opacity segmentation using U-Net variants in cataract images," *Computers in Biology and Medicine*, vol. 157, p. 106654, 2023.

[20] G. Silva and D. Costa, "Contrast normalization techniques for robust medical image classification," *Expert Systems with Applications*, vol. 212, p. 118780, 2023.

[21] S. Wang and L. He, "Deep residual networks for ocular image interpretation," *Ophthalmic Research*, vol. 66, no. 4, pp. 345–355, 2022.

[22] T. Santos and J. Oliveira, "Evaluating Vision Transformer performance in medical imaging," *Pattern Recognition Letters*, vol. 169, pp. 76–83, 2023.

[23] A. Kim and N. Choi, "Multi-modal fusion for cataract assessment," *Computerized Medical Imaging and Graphics*, vol. 103, p. 102226, 2024.

http://doi.org/10.26594/register.v8i2.XXX