VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480 /jitk.v11i4.7502

JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

COMPARISON OF CONVOLUTIONAL NEURAL NETWORK ARCHITECTURES EFFICIENTNET-B4 AND MOBILENETV2 IN CATARACT DISEASE DETECTION

Yuanita Puspitasari[1]; Jong Jek Siang[1*]; Andhika Galuh Prabawati[1]

Information System[1]
Universitas Kristen Duta Wacana, Yogyakarta, Indonesia[1]
www.ukdw.ac.id[1]
yuanita.puspitasari@si.ukdw.ac.id, jjsiang@staff.ukdw.ac.id*, andhika.galuh@staff.ukdw.ac.id

(*) Corresponding Author
(Responsible for the Quality of Paper Content)

Abstract— *Cataracts are the leading cause of blindness worldwide, with 94 million cases reported in 2023. Conventional cataract identification relies on visual examination methods that are prone to error due to their subjective nature. This study compares the performance of two Convolutional Neural Network (CNN) architectures, MobileNetV2 and EfficientNet-B4, in detecting cataract images. The dataset used was sourced from Kaggle and consisted of 1,074 normal images and 1,038 cataract images. The stages included preprocessing, augmentation, and the application of transfer learning with weights from ImageNet. The models were evaluated using accuracy, loss, precision, recall, F1-score, error rate, and visual interpretation using Grad-CAM metrics. The results showed that MobileNetV2 achieved 96% accuracy with an error rate of 4.05%, balanced precision, recall, and F1-score of 0.96, and a loss of 0.60. Meanwhile, EfficientNet-B4 achieved an accuracy of 96.5% with an error rate of 3.47%, balanced precision, recall, and F1-score of 0.97, and a lower loss of 0.12. Further evaluation indicates that EfficientNet-B4 has the same error rate on both training and test data (3.47%) with a loss difference of 0.02, suggesting that the model performs well and does not experience overfitting. In MobileNetV2, the difference in error rate between training (3.28%) and test (4.05%) is relatively small (0.77%), indicating that this model also does not exhibit overfitting. Grad-CAM visualization reveals that EfficientNet-B4 focuses more on clinically relevant areas, whereas MobileNetV2 tends to capture global patterns. Thus, EfficientNet-B4 is considered superior in terms of accuracy and generalization, while MobileNetV2 is more computationally efficient.*

Keywords: *Cataract, Efficientnet-B4, Grad-CAM, Image Classification, Mobilenetv2*

Intisari— *Katarak merupakan penyebab utama kebutaan didunia dengan 94 juta kasus pada tahun 2023. Identifikasi katarak secara konvensional menggunakan metode pemeriksaan visual yang rentan terhadap kesalahan karena bersifat subjektif. Penelitian ini membandingkan performa dua arsitektur Convolutional Neural Network (CNN), yaitu MobileNetV2 dan EfficientNet-B4, dalam mendeteksi citra katarak. Dataset yang digunakan berasal dari Kaggle terdiri dari 1.074 citra normal dan 1.038 citra katarak. Tahapan meliputi tahap preprocessing, augmentasi, serta penerapan transfer learning dengan bobot dari ImageNet. Model dievaluasi dengan metrik akurasi, loss, presisi, recall, F1-score, error rate, serta implementasi visual dengan Grad-CAM. Hasil penelitian menunjukkan bahwa MobileNetV2 mencapai akurasi 96% dengan error rate 4,05%, presisi, recall dan F1-score seimbang di angka 0,96, serta loss 0,60. Sementara EfficientNet-B4 mencapai akurasi 96,5% dengan error rate 3,47%, presisi, recall dan F1-score seimbang di angka 0,97, serta loss yang lebih rendah, yaitu 0,12. Evaluasi lebih lanjut menunjukkan bahwa EfficientNet-B4 memiliki error rate yang sama pada data train dan test (3,47%) dengan selisih loss sebesar 0,02, menandakan model bekerja dengan baik dan tidak mengalami overfitting. Pada MobileNetV2, perbedaan error rate antara train (3,28%) dan test (4,05%) relative kecil (0,77%), sehingga model ini juga dapat dikatakan tidak overfitting. Visualisasi Grad-CAM menunjukkan EfficientNet-B4 lebih fokus pada area klinis relevan, sedangkan MobileNetV2*

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7502

cenderung menangkap pola global. Dengan demikian, EfficientNet-B4 dinilai lebih unggul dari segi akurasi dan generalisasi, sementara MobileNetV2 lebih efisien secara komputasi.

***Kata Kunci**: Katarak, Efficientnet-B4, Grad-CAM, Klasifikasi Citra, Mobilenetv2*

INTRODUCTION

Cataracts are an eye disease that can cause blindness if not treated properly and promptly [1]. According to WHO data from 2023, 2.2 billion people worldwide suffer from long-distance or near vision impairment, with 1 billion of these cases still being preventable or treatable. Among these, cataracts are the leading cause of long-distance vision impairment and blindness, accounting for 94 million cases [2]. In general, ophthalmologists identify cataracts by assessing the brightness of the fundus image through visual examination, a method that is prone to physician subjectivity. One widely used method is the Lens Opacities Classification System (LOCS) III, which utilizes retro-illumination techniques where light from a slit lamp is reflected back to assess lens opacity [3]. Despite providing various optical techniques for visualizing ocular anatomy and pathology, slit-lamp biomicroscopy relies heavily on manual control, visual interpretation, and the clinician's ability to mentally reconstruct three-dimensional structures, making the examination inherently subjective and highly dependent on the examiner's expertise [4].

To overcome these limitations, one application of AI in the medical field is deep learning, particularly Convolutional Neural Networks (CNNs), which are capable of recognizing patterns in image data and automating classification [1]. CNN architecture has been widely used in medical image processing for tasks such as segmentation (e.g., U-Net for brain tumors) and classification (e.g., predictions in MRI images) [5]. In this context, selecting the appropriate CNN architecture significantly affects the model's accuracy and efficiency. For instance, MobileNetV2 is specifically designed to reduce computational and memory requirements while maintaining accuracy [6], whereas EfficientNet has been proven to offer a combination of high efficiency and superior accuracy [7]. However, MobileNetV2 can be prone to overfitting on small datasets [8], and EfficientNet's higher network complexity can also trigger overfitting, particularly on imbalanced data [9].

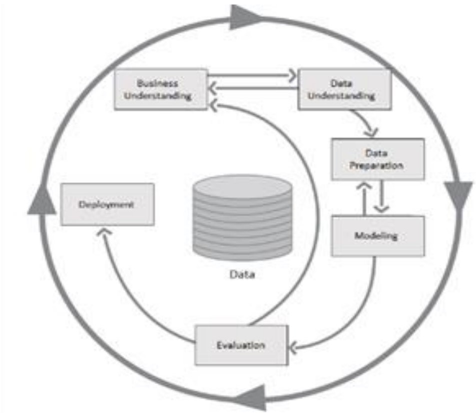
Several previous studies have explored the use of CNNs and their architectures in the medical field, particularly for disease image classification. Firdaus et al. [10] developed a web-based system for identifying cataracts and normal eyes using a CNN with epoch variations, achieving a best accuracy of 99.74% at 25 epochs. Similarly, Weni et al. [1] also investigated cataract identification using a CNN with epoch variations, achieving a highest accuracy of 95% at 50 epochs. In another domain, Wibowo et al. [11] classified brain tumors using MRI images with EfficientNet-B0 and EfficientNet-B7 through hyperparameter tuning, achieving 91-98% accuracy across four different scenarios. Wahyuningsih et al. [8] explored dental caries classification by comparing EfficientNet-B0, ResNet-50, MobileNetV2, and Inception, where ResNet-50 achieved the highest accuracy (99%), while EfficientNet-B0 and MobileNetV2 both recorded 98%.

Furthermore, Eryana et al. [12] researched malaria detection using EfficientNet-B0 and MobileNetV2 and found that MobileNetV2 provided better results with an accuracy of 97.27% and an F1-score of 97.29%. CNN architectures and their derivatives have proven effective in medical image classification. For instance, EfficientNet has achieved up to 98% accuracy in brain tumor classification, while MobileNetV2 demonstrates computational efficiency with high accuracy (97%) for other diseases. Although most existing research focuses on single architectures or disease domains other than cataracts, these results confirm the efficacy of both MobileNetV2 and EfficientNet, suggesting that similar methodologies hold potential for application in cataract detection, an area that remains underexplored to date.

Selecting a CNN architecture that balances accuracy and efficiency is crucial for overcoming the challenges of medical image classification. To enhance model transparency, Grad-CAM is applied to visualize the image regions on which the model focuses, thereby providing a visual interpretation of prediction results and facilitating evaluation of whether the model concentrates on clinically relevant areas [13]. In this context, MobileNetV2 employs depthwise separable convolutions with linear bottlenecks and shortcut connections to reduce computational load while maintaining accuracy, whereas EfficientNet-B4 optimizes model scaling by balancing depth, width, and resolution [12], [14], [15], [16]. The objective of this study is to compare MobileNetV2 and EfficientNet-B4 for cataract detection, an area that has not been extensively explored, while utilizing data augmentation and transfer learning to improve model performance and address dataset limitations [11], [17].

MATERIALS AND METHODS

CNN (Convolutional Neural Network) is an artificial neural network that processes spatial data such as images, functioning similarly to the brain's visual system by using convolution for greater efficiency and robustness in pattern recognition [18]. This study utilizes two CNN architectures: MobileNetV2 and EfficientNet-B4. MobileNetV2 was selected for its ability to reduce computational complexity and memory requirements without significant loss of accuracy [6]. EfficientNet-B4 was chosen because it implements a compound scaling strategy that balances the network's depth, width, and resolution, enabling high accuracy with good efficiency [19]. The primary parameters used in this study include an image input size of 224x224 pixels, a batch size of 32, a learning rate of 0.0001, and the Adam optimizer.



Source: (Shimaoka et al., 2024)[20]

Figure 1. Framework

This research employs the Cross-Industry Standard Process for Data Mining (CRISP-DM) framework. This framework was selected for its flexibility, applicability across various industries, and iterative cycle that facilitates sequential and adaptive project management and development [20]. The Deployment stage was excluded from this study, as the focus was confined to the development and evaluation of eye image classification models rather than their direct implementation in clinical systems. The implemented stages are illustrated in Figure 1, with the following detailed steps:

Business Understanding

Cataracts are a leading cause of blindness worldwide. A diagnostic aid system based on Convolutional Neural Networks (CNNs) can improve the accuracy and efficiency of cataract detection. This study compares the performance of MobileNetV2 and EfficientNet-B4 in classifying normal and cataractous eye images. The comparison results are expected to support the development of more reliable and interpretable models for medical applications.

Data Understanding

The dataset used in this study was obtained from the publicly available Kaggle repository [21] under the name "eye_diseases_classification". The original dataset comprises 4,217 eye images across four categories. This study utilized only two classes from the dataset, comprising 1,074 normal eye images and 1,038 cataract images, resulting in a total of 2,112 images used for analysis. During this stage, a filtering process was applied to remove a limited number of images that could potentially affect Grad-CAM interpretation results.

Data Preparation

All data underwent preprocessing, which included resizing the images to 224x224 pixels using the `target_size` parameter and normalizing pixel values to a range of 0-1 using `rescale=1./255`. The dataset was then split into training, validation, and test sets with a ratio of 70:15:15, employing `random_state=42` for reproducibility. To increase data variety, augmentation was performed on the dataset using techniques including rotation, translation, channel shifting, zooming, brightness adjustment, cropping, and flipping.

Modeling

A transfer learning approach was implemented by freezing all layers in MobileNetV2 and 75% of the layers in EfficientNet-B4, while utilizing pre-trained weights from ImageNet. To mitigate overfitting risk, regularization techniques including dropout and batch normalization were applied. The training process was conducted using Python 3.11.12 with the TensorFlow 2.18.0 framework in a Google Colab environment with GPU acceleration. For the EfficientNet-B4 architecture, the default input size (380x380) was adjusted to 224x224 to accommodate device capacity limitations, matching the input dimensions used for MobileNetV2. Supporting libraries such as NumPy, scikit-learn, and Matplotlib were utilized for data processing, evaluation, and visualization. Additionally, callbacks including EarlyStopping, ModelCheckpoint, and ReduceLROnPlateau were employed to enable early training termination, preserve the best model weights, and dynamically adjust the learning rate to accelerate model convergence.

JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)
VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7502

Evaluation

Model performance was evaluated using accuracy, loss, precision, recall, and F1-score metrics. The error rate is additionally calculated using the following formula:

$$Error\ Rate = \frac{Number\ of\ Incorrect\ Predictions}{Total\ Number\ of\ Predictions} x\ 100\% \quad (1)$$

These metrics provide a more comprehensive assessment of the model's performance, evaluating not only accuracy but also the balance between positive and negative predictions. To improve model interpretability, this study also utilizes the Gradient-weighted Class Activation Mapping (Grad-CAM) method. Grad-CAM visualizes the areas of the eye image that receive the primary focus in the model's classification decisions.

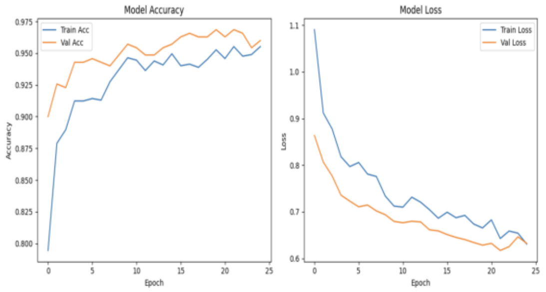
This visualization makes the classification results more transparent and helps medical personnel understand the basis for the model's decisions. By highlighting clinically relevant features, reducing the "black-box" nature of deep learning models, facilitating collaborative decision-making, and enabling model validation and error analysis, this approach increases model reliability and user confidence in artificial intelligence-based systems [22].

RESULTS AND DISCUSSION

Training Performance:

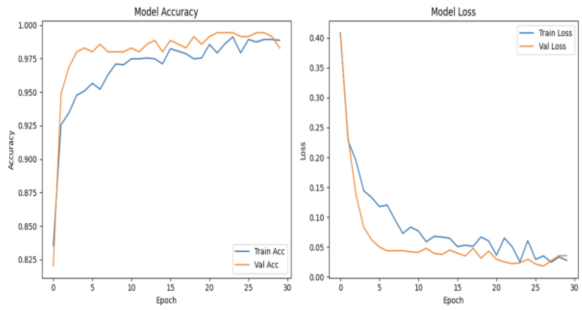
The initial training results were demonstrated by MobileNetV2. At the beginning of the epochs, the training accuracy remained low (0.79-0.90), while the validation accuracy had already reached 0.9 (Figure 2, left). As the number of epochs increased, both accuracies improved steadily until the end of training, reaching values of 0.95 (training) and 0.96 (validation), with no significant indication of overfitting (Figure 2, left).

The loss curve decreased consistently throughout training, with the validation loss generally lower than the training loss. In the final epoch, the validation loss was slightly higher than the training loss (0.646 vs. 0.639), although this minor discrepancy does not indicate substantial overfitting (Figure 2, right). This indicates that the model learned effectively and remained relatively stable, with a training duration of approximately 2 hours and 25 minutes. The accuracy and loss curves for MobileNetV2 are presented in Figure 2.



Source: (Research Results, 2025)
Figure 2. Accuracy and Loss of MobileNetV2

The subsequent training results are demonstrated by EfficientNet-B4, which achieved high initial accuracy (0.82) and showed a sharp increase to >0.95 within just a few epochs (Figure 3, left). After the 10th epoch, the accuracy stabilized between 0.97-0.99, with validation accuracy slightly exceeding training accuracy, indicating good model generalization without signs of overfitting (Figure 3, left). The initial loss of approximately 0.35-0.4 decreased dramatically to <0.05, with validation loss generally remaining lower than training loss (Figure 3, right). Although the final epoch showed a slightly higher validation loss than training loss (0.035 vs. 0.024), this minor difference does not indicate overfitting. This confirms optimal learning supported by regularization techniques, including Batch Normalization and Dropout, alongside augmentation applied exclusively to training data. The training process required approximately 4 hours and 30 minutes. The accuracy and loss graphs for EfficientNet-B4 are shown in Figure 3.



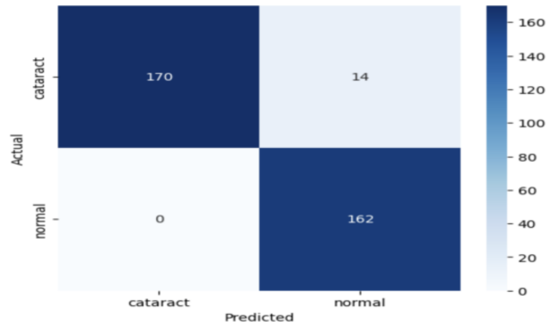
Source: (Research results, 2025)
Figure 3. Accuracy and Loss of EfficientNet-B4

Confusion Matrix Analysis:

The confusion matrix demonstrates the model's classification performance. The MobileNetV2 confusion matrix in Figure 4 indicates that the model achieves very high sensitivity (recall) for the normal class (100%) and performs well for the cataract class (92.4%). Classification errors occur only in a small proportion of cataract images misclassified as normal. This demonstrates that the

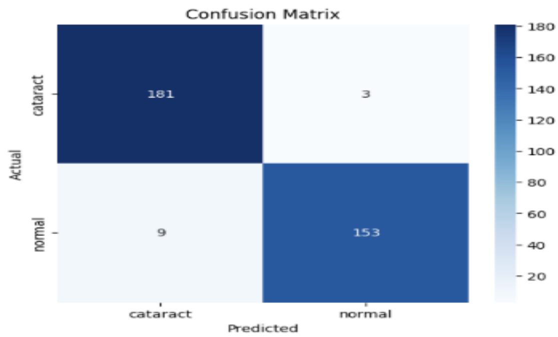
VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480 /jitk.v11i4.7502

model can classify cataract and normal images with excellent accuracy, where most image data are predicted correctly according to their true classes.



Source: (Research results, 2025)
Figure 4. MobileNetV2

The confusion matrix for EfficientNet-B4 demonstrates balanced performance with excellent capability in recognizing cataract images (98.9%) and good performance in identifying normal eye images (97.5%). The classification error is relatively minimal, with only 2 cataract images misclassified as normal (False Negative) and 4 normal eye images misclassified as cataracts (False Positive). The corresponding graph is presented in Figure 5.



Source: (Research results, 2025)
Figure 5. EfficientNet-B4

Evaluation on Test Data:

The testing results for MobileNetV2 and EfficientNet-B4 are presented in Table 1.

Table 1. Testing Results

Model	MobileNetV2	EfficientNet-B4
Akurasi	96%	96,5%
Loss	0,64	0,12
Error Rate	4,05%	3,47%
Presisi	0,96	0,97
Recall	0,96	0,97
F1-score	0,96	0,97
Waktu	2 h 25 min	4 h 30 min

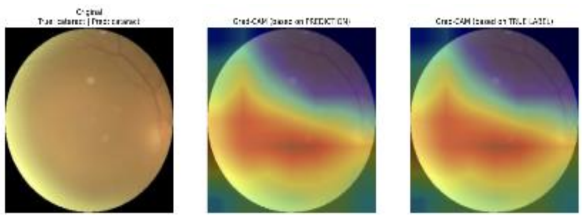
Source: (Research results, 2025)

The test results indicate that MobileNetV2 achieved 96% accuracy with an error rate of 4.05%, along with balanced precision, recall, and F1-score of 0.96, although the loss value remained relatively high (0.64), which could lead to prediction errors in some samples. The high loss value is due to the fact that most samples have high confidence, but a small number of samples (3-4%) have very low probabilities in the correct class, resulting in high cross-entropy and increasing the average loss. This probability pattern is consistent with the confusion matrix results, where the model produced 14 false negatives and 0 false positives, indicating a tendency for MobileNetV2 to miss cataract cases. Misclassified samples generally show low-contrast cataracts that are difficult for lighter architectures to detect.

On the other hand, EfficientNet-B4 recorded an accuracy of 96.5% with an error rate of 3.47% and a lower loss (0.12), accompanied by balanced precision, recall, and F1-score at 0.97, demonstrating more consistent performance with minimal prediction errors. Although the accuracy during training reached 98% (Figure 3, left), the evaluation results on both the training and test data showed a consistent accuracy of 96.5%, with a small difference in loss values (0.10 for training data and 0.12 for test data). This consistency indicates that the model remains relatively stable without exhibiting overfitting, despite the evaluation accuracy being lower than the training accuracy. Based on the confusion matrix, EfficientNet-B4 produced fewer false negatives (3), although there are 9 false positives, indicating higher sensitivity of the model to the cataract class. The two samples misclassified by both models show very subtle cataract characteristics with low contrast levels, making them difficult to detect by both architectures. Regarding efficiency, MobileNetV2 required a shorter training time (2 h 25 min), while EfficientNet-B4 required nearly double that duration (4 h 30 min).

Grad-CAM Visualization:

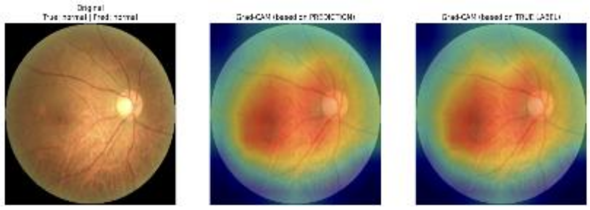
The interpretation of Grad-CAM visualizations will be explained according to the class and CNN architecture model used.


Source: (Research results, 2025)
Figure 6. Grad-CAM of Cataract (MobileNetV2)

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

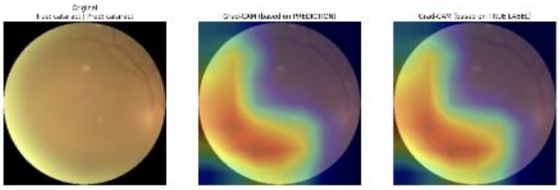
JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7502

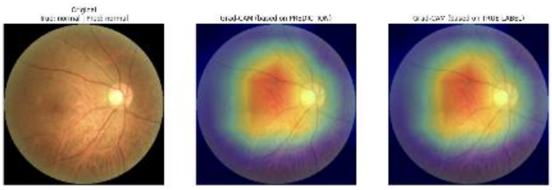


Source: (Research results, 2025)
Figure 7. Grad-CAM of Normal (MobileNetV2)

Figures 6 and 7 present the Grad-CAM visualization results for MobileNetV2. The model generally demonstrates a tendency to focus globally on cloudy areas in cataract images (Figure 6), while in normal images, the primary focus is on the retinal center and optic disc (Figure 7). The consistency between Grad-CAM predictions and true labels confirms that the model does not make random guesses but recognizes relevant visual patterns, although its approach remains generalized and lacks the detail required for more refined cataract detection, which may represent a limitation of this model.



Source: (Research results, 2025)
Figure 8. Grad-CAM of Cataract (EfficientNet-B4)



Source: (Research results, 2025)
Figure 9. Grad-CAM of Normal (EfficientNet-B4)

Grad-CAM visualization for EfficientNet-B4 demonstrates that in cataract images, the model concentrates on blurred areas at the bottom and left side of the fundus, aligning with characteristics of lens opacity (Figure 8). Conversely, in normal images, the focus shifts to the optic disc and surrounding blood vessels as key anatomical structures (Figure 9). The consistent heatmap patterns between predictions and ground truth labels confirm that the model's decisions are not arbitrary but based on clinically relevant visual features, establishing EfficientNet-B4 as a reliable architecture for distinguishing between cataract and normal images.

Discussion and Comparison:

Overall, the results indicate that EfficientNet-B4 demonstrates superior performance in accuracy, generalization, and Grad-CAM consistency, while MobileNetV2 offers advantages in computational efficiency and training speed. The consistency is evidenced by stable and reproducible heatmap patterns in clinically relevant regions, particularly the lens center and anterior capsule, observed across nearly all samples and consistently aligning with both predictions and true labels. In contrast, MobileNetV2 generates more diffuse and variable activation patterns, with greater heterogeneity in hotspot localization between images. Based on visual interpretation results, Grad-CAM reveals that EfficientNet-B4 produces more localized and clinically relevant activations. Therefore, model selection should be determined by specific requirements: MobileNetV2 for efficiency-oriented applications, and EfficientNet-B4 for high accuracy clinical implementations.

CONCLUSION

The CNN architectures demonstrate capability in detecting cataract images with high accuracy (≥96%) and balanced precision, recall, and F1-score values, indicating reliable classification performance and good generalization. MobileNetV2 achieves 96% accuracy with precision, recall, and F1-score values of 0.96 each, excelling in time efficiency and computational requirements, though it relies more on global patterns and consequently lacks detail in distinguishing cloudy areas. In comparison, EfficientNet-B4 achieves 96.5% accuracy with an error rate of 3.47%, along with precision, recall, and F1-score values of 0.97 each, demonstrating superior generalization capability. Grad-CAM visualization reveals that EfficientNet-B4 focuses more precisely on clinically relevant areas, whereas MobileNetV2 tends to capture global patterns. Although the accuracy difference is relatively small, EfficientNet-B4 can be considered the superior model despite requiring greater computational resources and time. The limitations of this study include the relatively small dataset size, inconsistent labeling quality, as well as hardware and time constraints that limited thorough parameter exploration. Future research should consider expanding the dataset with consistent labeling and conducting studies in environments with adequate resources to enable optimal parameter tuning and development of superior models.

ACKNOWLEDGEMENTS

This publication was fully supported by the Faculty of Information Technology, Universitas Kristen Duta Wacana

REFERENCE

[1] I. Weni, P. E. P. Utomo, B. F. Hutabarat, and M. Alfalah, "Detection of Cataract Based on Image Features Using Convolutional Neural Networks," *IJCCS (Indonesian Journal of Computing and Cybernetics Systems)*, vol. 15, no. 1, pp. 75–86, Jan. 2021.

[2] World Health Organization, "Blindness and vision impairment," World Health Organization. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/blindness-and-vision-impairment. [Accessed: Sep. 15, 2025].

[3] M. Mirzaie, E. Bahremani, N. Taheri, Z. Khamnian, and B. Kharrazi Ghadim, "Cataract grading in pure senile cataracts: Pentacam versus LOCS III," *J. Ophthalmic Vis. Res.,* vol. 17, no. 3, pp. 337–343, Aug. 2022, doi: 10.18502/jovr.v17i3.11570.

[4] P. A. R. Meyer, "Living histopathology - interrogation of ocular tissues by light: a celebration of the slit-lamp and a repertoire of clinical techniques," Mar. 01, 2025, *Springer Nature*. doi: 10.1038/s41433-024-03139-5.

[5] A. Crimi and S. Bakas, Eds., BrainLes, "Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries", Part I, vol. 12962, *Lecture Notes in Computer Science. Cham, Switzerland*: Springer, 2022. doi: 10.1007/978-3-031-08999-2.

[6] S. Velu, "An efficient, lightweight MobileNetV2-based fine-tuned model for COVID-19 detection using chest X-ray images," *Mathematical Biosciences and Engineering*, vol. 20, no. 5, pp. 8400–8427, 2023, doi: 10.3934/mbe.2023368.

[7] A. G. Pradana, D. R. I. M. Setiadi, and A. R. Muslikh, "Fine tuning model convolutional neural network EfficientNet-B4 dengan augmentasi data untuk klasifikasi penyakit kakao," *J. Inf. Syst. Appl. Dev*., vol. 2, no. 1, pp. 1–11, Mar. 2024, doi: 10.26905/jisad.v2i1.11899.

[8] Wahyuningsih, G. S. Nugraha, and R. Dwiyansaputra, "Classification of dental caries disease in tooth images using a comparison of EfficientNet-B0, MobileNetV2, ResNet-50, and InceptionV3 architectures," *J. Tek. Inform.,* vol. 5, no. 4, pp. 177–185, Aug. 2024, doi: 10.52436/1.jutif.2024.5.4.2187.

[9] M. W. Purbandanu, R. Yanuarta, and A. Kurniawan, "Optimization of Skin Cancer Detection to Improve Accuracy with the Application of Efficient Convolutional Neural Network and EfficientNetB2 Models," *Journal of Intelligent Computing & Health Informatics*, vol. 5, no. 2, pp. 43–50, Apr. 2024, doi: 10.26714/jichi.v5i2.14338.

[10] D. H. Firdaus, B. Imran, L. D. Bakti, and E. Suryadi, "Klasifikasi Penyakit Katarak Pada Mata Menggunakan Metode Convolutional Neural Network (CNN) Berbasis Web," *Jurnal Kecerdasan Buatan dan Teknologi Informasi (JKBTI)*, vol. 1, no. 3, pp. 18–26, Dec. 2022, doi: 10.69916/jkbti.v1i3.6.

[11] M. A. P. Wibowo, M. B. Al Fayyadl, Y. Azhar, and Z. Sari, "Classification of brain tumors on MRI images using convolutional neural network model EfficientNet," *J. RESTI (Rekayasa Sist. Teknol. Inf.)*, vol. 6, no. 4, pp. 538–547, Aug. 2022, doi: 10.29207/resti.v6i4.4110.

[12] A. Eryana, Sriyanto, and Firmansyah, "Prediksi malaria menggunakan metode pre-trained model algoritma EfficientNet-B0 dan MobileNet-V2," *J. Ilm. Komputasi,* vol. 22, no. 1, pp. 71–80, Mar. 2023, doi: 10.32409/jikstik.22.1.2967.

[13] M. Denny Richardo and H. Satria, "Comparative Analysis of Explainable AI Models for Pneumonia Detection in Chest X-rays Using Grad-CAM," *Sistem Informasi dan Komputer*, vol. 14, no. 04, pp. 475–483, Oct. 2025, doi: 10.32736/sisfokom.v14i4.2450.

[14] R. Preetha, M. J. P. Priyadarsini, and J. S. Nisha, "Automated Brain Tumor Detection From Magnetic Resonance Images Using Fine-Tuned EfficientNet-B4 Convolutional Neural Network," *IEEE Access*, vol. 12, pp. 112181–112195, Jul. 2024, doi: 10.1109/ACCESS.2024.3442979.

[15] Y. Miftahudin and F. Zaelani, "Perbandingan Metode Efficientnet-B3 dan Mobilenet-V2 Untuk Identifikasi Jenis Buah-buahan Menggunakan Fitur Daun," *Fiqry Zaelani S Jurnal Ilmiah Teknologi Informasi Terapan*, vol. 9, no. 1, pp. 1–11, Dec. 2022, doi: 10.33197/jitter.vol9.iss1.2022.911.

[16] A. Paisal, J. Jasril, S. Sanjaya, L. Handayani, and F. Syafria, "Klasifikasi Daging Sapi dan Daging Babi Menggunakan CNN dengan Arsitektur EfficientNet-B4 dan Augmentasi Data," *Jurnal Informatika Universitas*

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.7502

Pamulang, vol. 8, no. 2, pp. 165–176, Jun. 2023, doi: 10.32493/informatika.v8i2.31884.

[17] A. N. Nafisa, E. N. D. B. Purba, F. A. A. Harahap, and N. A. Putri, "Implementasi Algoritma Convolutional Neural Network Arsitektur Model MobileNetV2 dalam Klasifikasi Penyakit Tumor Otak Glioma, Pituitary dan Meningioma," *Jurnal Teknologi Informasi, Komputer dan Aplikasinya*, vol. 5, no. 1, pp. 53–61, Mar. 2023, doi: 10.29303/jtika.v5i1.234.

[18] O. Colliot, Ed., *Machine Learning for Brain Disorders*, vol. 197. New York, NY: Springer US, 2023.

[19] M. W. Purbandanu, R. Yanuarta, A. Kurniawan, M. Munsarif, and A. Ilham, "EfficientNet for Medical Image Classification: Performance vs. Efficiency in Skin Cancer Detection," *Journal of Intelligent Computing & Health Informatics*, vol. 5, no. 2, Apr. 2024, doi: 10.26714/jichi.v5i2.14338.

[20] A. M. Shimaoka, R. C. Ferreira, and A. Goldman, "The evolution of CRISP-DM for Data Science: Methods, Processes and Frameworks," *SBC Reviews on Computer Science*, vol. 4, no. 1, pp. 28–43, Oct. 2024, doi: 10.5753/reviews.2024.3757.

[21] G. V. Doddi, "Eye diseases classification," Kaggle, 2022. [Online]. Available: https://www.kaggle.com/datasets/gunavenkatdoddi/eye-diseases-classification. [Accessed: Apr. 23, 2025].

[22] F. M. Talaat, S. A. Gamel, R. M. El-Balka, M. Shehata, and H. ZainEldin, "Grad-CAM Enabled Breast Cancer Classification with a 3D Inception-ResNet V2: Empowering Radiologists with Explainable Insights," *Cancers (Basel)*, vol. 16, no. 21, Nov. 2024, doi: 10.3390/cancers16213668.