

## PAPER

# Comparative Evaluation of Transfer Learning Architectures for Glaucoma Classification

Helmi Imaduddin  (✉),  
Widi Widayat , Ihsan  
Cahyo Utomo, Nur Rizqi  
Febriandika , Amri Zadi  
Hudaya, Alleyda Aurora  
Yafentri 

Universitas Muhammadiyah  
Surakarta, Surakarta,  
Indonesia

[helmi.imaduddin@ums.ac.id](mailto:helmi.imaduddin@ums.ac.id)

## ABSTRACT

Glaucoma is a leading cause of irreversible blindness worldwide, primarily resulting from elevated intraocular pressure (IOP) that leads to progressive optic nerve damage. Early detection is essential to prevent permanent vision loss; however, it remains challenging due to the disease's asymptomatic nature in its initial stages. This study explores the application of deep learning in glaucoma detection, utilizing the Xception architecture to classify glaucoma from retinal fundus images. A total of 6,540 retinal images were obtained from the publicly available EyePACS-AIROGS-light-V2 dataset variant designed for glaucoma detection and systematically divided into training, validation, and testing subsets. Data augmentation techniques such as image rotation, flipping, and brightness adjustments were applied to enhance model robustness and reduce the risk of overfitting. The performance of the Xception model was benchmarked against other popular convolutional neural network architectures, including EfficientNet B7 and ResNet-50, using key evaluation metrics: accuracy, precision, sensitivity, and specificity. The Xception model outperformed the others, achieving an accuracy of 93%, a precision of 95%, a sensitivity of 91%, and a specificity of 94%. These results underscore the model's effectiveness in identifying glaucomatous patterns within retinal images. The comparative analysis reveals that Xception achieves the best balance between accuracy and computational efficiency, owing to its depthwise separable convolutional design. Beyond quantitative results, this study provides insight into how architectural choices affect model performance in ophthalmic imaging, offering guidance for developing scalable AI-assisted screening systems.

## KEYWORDS

fundus imaging, EfficientNet B7, glaucoma, ResNet50, Xception

## 1 INTRODUCTION

Eye health affects many aspects of life. The Global Eye Health Commission on eye health defines that maximizing eye health contributes to well-being and health to achieve a good quality of life. Eye health also plays a vital role in the process of

Imaduddin, H., Widayat, W., Utomo, I. C., Febriandika, N. R., Hudaya, A. Z., Yafentri, A. A. (2026). Comparative Evaluation of Transfer Learning Architectures for Glaucoma Classification. *International Journal of Online and Biomedical Engineering (iJOE)*, 22(2), pp. 76–89. <https://doi.org/10.3991/ijoe.v22i02.58101>

Article submitted 2025-08-05. Revision uploaded 2025-11-05. Final acceptance 2025-11-06.

© 2026 by the authors of this article. Published under CC-BY.

realizing the Sustainable Development Goals (SDGs). Eye health problems that arise based on data in 2020: around 596 million individuals in the world suffer from distance vision problems, around 510 million individuals suffer from near vision problems, and around 43 million individuals experience blindness [1]. Problems that arise as a threat to eye health include myopia, diabetic retinopathy, retinopathy of prematurity, cataracts, macular degeneration, and glaucoma.

Glaucoma is a significant global health issue characterized by increased intraocular pressure (IOP) that leads to progressive optic nerve damage and potentially irreversible vision loss [2], [3]. It is the leading cause of irreversible blindness worldwide, affecting millions of individuals and posing a substantial burden on healthcare systems [4], [5]. The disease encompasses various subtypes, including primary open-angle glaucoma (POAG), angle-closure glaucoma, normal-tension glaucoma, and secondary glaucoma, each with distinct pathophysiological mechanisms and risk factors such as age, family history, ethnicity, and certain medical conditions such as type 2 diabetes mellitus and cardiovascular diseases [6]. Early detection and timely treatment are crucial to prevent the progression of glaucoma and subsequent vision loss.

Nonetheless, diagnosing and recognizing glaucoma, especially at early stages, will always be tricky chiefly because of the absence of distinctive symptoms and the time and resource-consuming nature of monitoring [7]. In order to assist the healthcare practitioner in identifying the disease, conventional diagnostic modalities include OCT, visual perimetry, and IOPs [8]. New methods of treatment and other methods of diagnosis, including automated detection and classification systems based on the analysis of retinal fundus images, have begun to be developed to optimize and facilitate the diagnosis of glaucoma. In particular, a model based on a Stacked Attention U-Net has been proposed for optic disc segmentation, which has also demonstrated effective results with respect to glaucoma detection. This automated method attempts to overcome the problems of manual detection, such as the lack of human resources and the tediousness of the detection procedure, thereby enabling timely management, which lowers the chances of incurring vision problems [9]. In another research by Mukhlif et al., it was noted that such medical image classification models as Xception had advantages over VGG16 and ResNet50-based models, assuming the images contained skin cancer data from ISIC2020 and breast cancer data from ICIAR2018.

The work presented in this study emphasized the model's better than average accuracy and specificity, which makes it a good candidate for classification of glaucoma [10]. The efforts of Tang and Sheykhahmad are also positive in regard to the Xception model, where it was reported that the model yielded 94.206% in average accuracy of skin cancer diagnosis using the ISIC dataset, exceeding other current methods [11]. Similarly, Sharma's work was also able to show that ResNet50 and DenseNet-201 are useful in the classification of glaucoma with 85.43% across the HVD Dataset [12]. These models are characterized by deep architectures and the ability to learn complex features from medical images. The work done on the application of transfer learning is not restricted to medical imaging only; other domains, including crop classification and code-related works, are also being performed on this. Antonijević et al. illustrated the usefulness of this method for crop classification accuracy by retraining models on specific datasets [13].

In the same manner, Mastropaolo et al. drew attention to the usefulness of transfer learning for the performance of models building on code tasks [14]. Lots of things can be achieved also in different domains, which are cross-transferring learning. Also, the application of transfer learning algorithms in the analyses of the fundus

photographs has made it possible to recognize minute variations in the retinal structure indicating glaucoma damage.

Transfer learning has yielded promising results in the use of retinal fundus images for enhancing the performance of glaucoma screening. It has been shown that high levels of accuracy are achieved with the use of deep learning architectures such as DenseNet-121 in the detection of eye conditions such as glaucoma, cataracts, and diabetic retinopathy. Some previous works explain the usage of neoconstructs, in particular pre-trained models such as ResNet18 and ResNet50, with advanced metrics of precision, recall, and F1 score in retinal disease classification. They show great potential in the screening of these conditions [15]–[17]. In spite of these technological improvements, further efforts, such as the creation of more diverse and large datasets and definitions of appropriate uniform diagnostic standards, are still required to push these automated systems further and make them effective in populations of different ethnicities. It is also necessary to highlight the role of primary care providers, who identify and manage large populations of patients at high risk for glaucoma, referring them for evaluation and treatment by specialists when necessary. New treatment modalities such as topical drugs, laser interventions, and surgical procedures are being developed aimed to decreasing IOP and avoiding further damage to the optic nerve [18]. Given the background and the aims and objectives of this study, this study will use Xception, the most recent version of Xception architecture, to perform glaucoma disease detection.

Due to the optimization of both efficiency and performance, Xception is particularly effective in deployment scenarios with resource constraints. Afterward, the efficiency is tested, and it will actually be assessed how this Xception architecture compares with other transfer learning architectures such as EfficientNet and ResNet50 on the classification of glaucoma data. The results will present rationale in detail through the statistical tests of accuracy, precision, sensitivity, and specificity that particularly focus on the classification of the model.

This study presents a structured comparative analysis of three transfer learning architectures, namely 1) Xception, 2) EfficientNet B7, and 3) ResNet 50, for automated glaucoma detection using a large retinal fundus image dataset. It provides the first systematic evaluation of these architectures within this specific context. Furthermore, the study introduces an extensive data augmentation framework incorporating image rotations, flips, shifts, zooming, and brightness adjustments, which significantly enhances the model's robustness to variations in image quality and acquisition conditions. The work also details finely tuned hyperparameters and regularization strategies, including adaptive learning-rate scheduling and dropout, to secure stable training convergence and strong generalization. By demonstrating effective transfer-learning workflows tailored for clinical scenarios with limited resources, this study paves the way for scalable, AI-driven early-screening solutions in ophthalmology.

## 2 MATERIALS AND METHODS

This study includes several important steps as follows: The first step is the collecting of retinal images, followed by preprocessing of the data. The dataset was divided into training, validation, and testing subsets for purposes of creating and testing the model. This phase of the research entails the creation of the machine learning model that is capable of classifying an image in terms of whether it indicates glaucoma or not. The functional capabilities of the model, including accuracy, precision, recall,

and other measures, are then measured and evaluated using the confusion matrix. Figure 1 presents the general research process of the study, outlining the complete process that has been undertaken in order to meet the aims of the study.

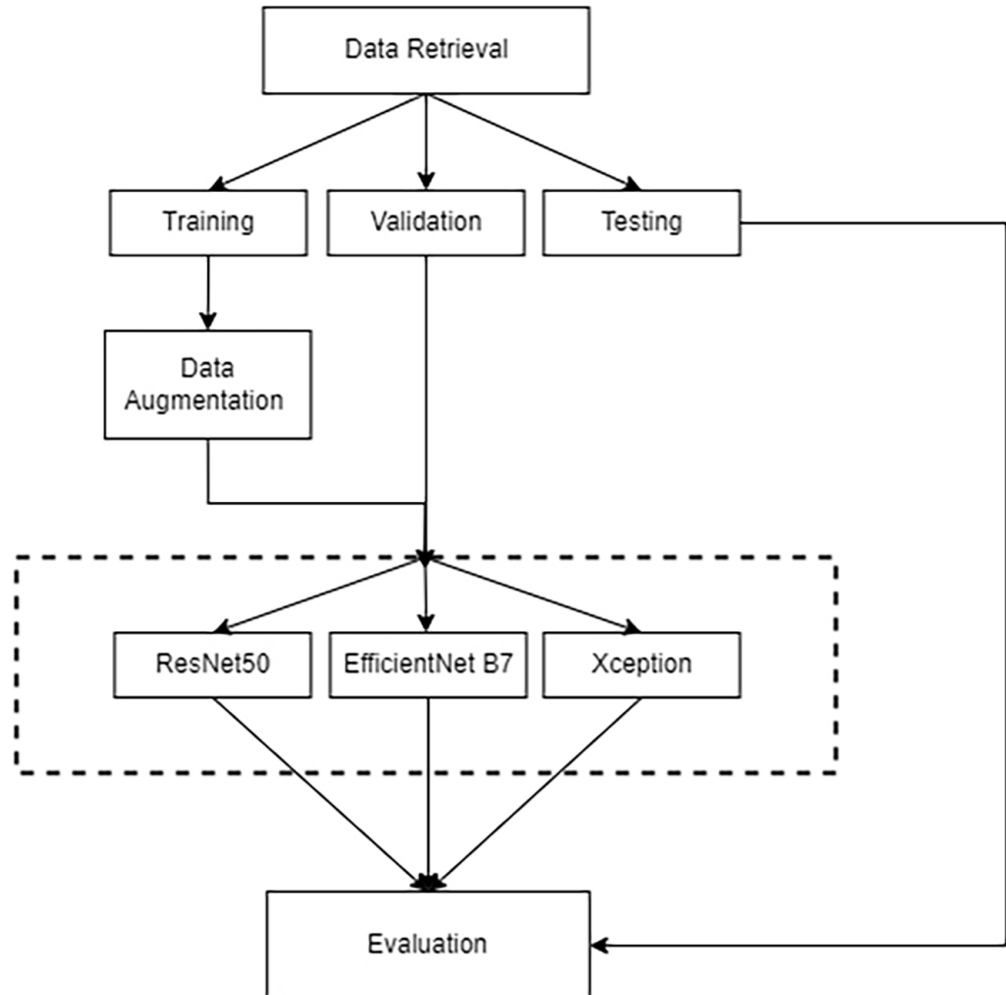


Fig. 1. Research process

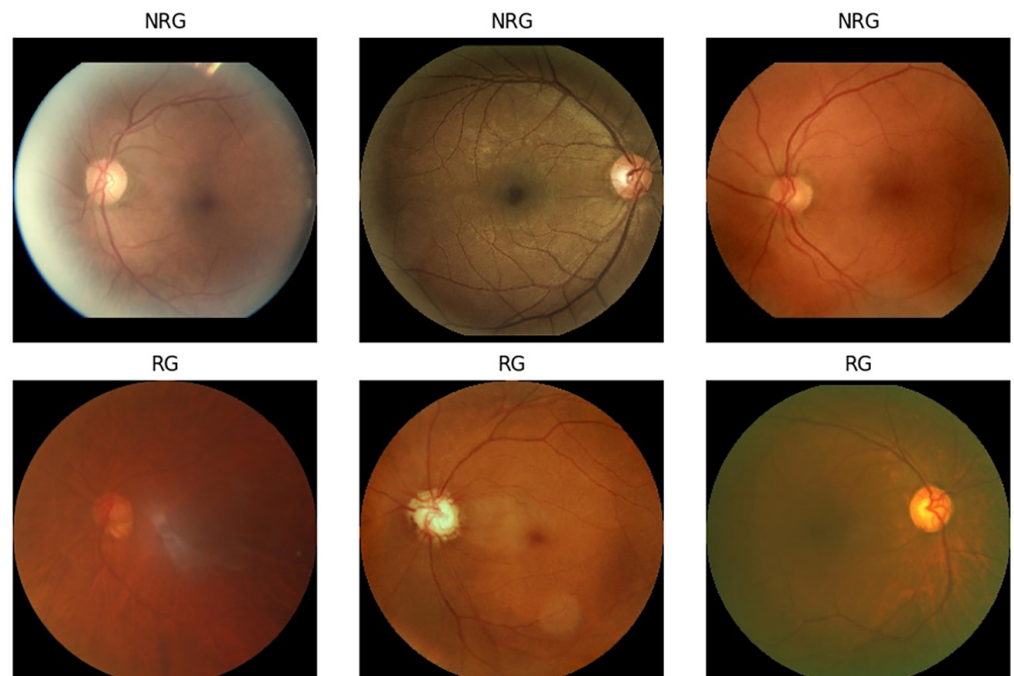
## 2.1 Data retrieval

The images used in this study were retinal fundus images available in the Kaggle dataset. Although the original EyePACS dataset is for diabetic retinopathy, this study specifically uses the EyePACS-AIROGS-light-V2 variant designed for glaucoma detection [19], [20]. A total of 6,540 retinal fundus images were included in the dataset, comprising 3,270 glaucomatous and 3,270 non-glaucomatous samples, ensuring a balanced dataset. In order to achieve adequate training, validation, and testing of the machine learning frameworks developed in this study, the dataset was systematically divided into these subsets. A total of 6,540 images of the retinal fundus were contained in the dataset, out of which three main subsets were formed. The training set comprised 5,000 images, the testing set had 1,000 images, and the validation set contained 540 images. The proportion of the dataset is shown in Table 1.

**Table 1.** Dataset proportion

Subset	Images
Training	5,000
Testing	1,000
Validation	540
Total	<b>6,540</b>

The dataset division is comprehensively described to improve the performance and generalization ability of the machine learning model. The training set occupies the largest portion so that the model is exposed to a variety of retinal images, both positive and negative for glaucoma, so that it can learn the various characteristics of each class. The validation set acts as a bridge between the training and testing phases, providing an objective evaluation for model performance and facilitating hyperparameter tuning during the training process. Meanwhile, the testing set is used to assess the overall prediction accuracy and the model's ability to generalize to completely new data. This procedure is crucial to ensure that the model not only recognizes glaucoma patterns in the training data but also remains reliable when applied to retinal images in real clinical scenarios [21]. Figure 2 shows a visual representation of retinal images with referable glaucoma (RG) and non-referable glaucoma (NRG) conditions, illustrating the morphological differences that form the basis of the model's classification. This visual recognition not only facilitates manual verification by ophthalmologists but also serves as a benchmark for the algorithm's success in extracting and distinguishing pathological features that are crucial for early glaucoma diagnosis.

**Fig. 2.** Retinal images dataset

In addition, the careful splitting of the data is also emphasized for models that are expected to yield high reliability when performance metrics are reported. Data partitioning in terms of training, validation, and test data partitions prevents the

challenge of overfitting, in which a model would simply try to memorize every piece of data presented at training and not learn how to generalize. Therefore, refraining from training the execution of the task while keeping the test data separate from the training data allows for a reliable evaluation of the model's performance in glaucoma detection within retinal fundus images that were not included in the training set.

## 2.2 Data augmentation

The augmentation propels the available information into the learning style explanatory transformation processes as part of the steps. This process ensures that the given dataset can be more effective and more focused so as to enhance the efficiency of the classification model that is about to be constructed. In this work, we used several data augmentation techniques to enhance the number of training images so that overfitting of the model is minimized and the generalization ability of the model is enhanced. This implies that any augmentation considered has to be preceded by the process of resizing the images to  $224 \times 224$  for ResNet-50,  $299 \times 299$  for Xception, and  $600 \times 600$  for EfficientNet-B7. Adjusting the image size ensures uniform input dimensions and compatibility across different deep learning model architectures.

This section will discuss how data augmentation is carried out and the parameters that were utilized and deliberately set. First, we previously used a `width_shift_range` and `height_shift_range` of 0.5 each. This technique enables the surface of the image to be shifted up to fifty percent of the image's dimension in both horizontal and vertical orientations so that the position of the objects can be changed in the image, and thus the model trained can be able to tolerate changes in the location of the objects. Vertical and horizontal alterations are also performed, namely `horizontal_flip` and `vertical_flip`, which allow the image to be changed both horizontally and vertically, thereby enhancing variations with respect to the position of the objects. This is useful in cases when the objects do not have a consistent direction, and the direction of the objects has to be changed to assist the model in learning better.

In addition, we implement a `zoom_range` of 0.15. This means that the image can be enlarged or reduced to 15% of the original size, and therefore the model will learn to regard objects at various perspectives. To inject into the empty pixels that remain due to augmentation operations such as shift and zoom, `fill_mode` with the value of "constant" and `cval` of 0.1 is applied in this case. It makes sure that blank spaces surrounding the main image (due to image augmentations) are filled by a certain value so as to protect the data integrity. Finally, all images were resized considering the following rescaling factor =  $1/255$ . All pixel values in this range [0,1] normalization of the temperature decreases the training time of the model as well as enhances convergence of gradient descent optimization by eliminating the initial pixel values range of [0, 255], which might hinder the model's convergence. Overall, this exhaustive preprocessing not only helps build an effective and accurate classification model but also extends to making great impacts, especially in the area of medical imaging and disease diagnosis, such as glaucoma.

## 2.3 Transfer learning

Transfer learning has grown into a prominent machine learning paradigm, especially in image classification tasks [22]. Such models ease the application of existing

general knowledge from big data sources to specific problem areas where only a few data resources are available in order to minimize unnecessary errors. This method is very useful in medical imaging, where the major hindrance is lack of ample data. The following sections study how different CNNs classify glaucoma and the significance of transfer learning in other areas related to medicine.

Especially in medical imaging, large annotated datasets are seldom accessible, thus disfavoring traditional machine learning methods; transfer learning comes in handy. Reis et al. demonstrated that using pretrained models such as InceptionV3 and DenseNet121 improved the classification of medical datasets, including chest X-ray and diabetic retinopathy images [23]. Medical imaging has one area that transfer learning enhances—glaucoma imaging. Researchers can substantially advance the diagnostic power in this area by employing already trained CNNs and engaging in novel approaches like classifier fusion and teacher-student systems. Although this section centers around medical contexts, it should be emphasized that the concepts of transfer learning are not limited to this sphere, and such features of learning can be observed in many other fields as well. Building upon these concepts, the next stage involves the training and fine-tuning of selected architectures.

## 2.4 Training model

This study incorporates three well-known deep architectures that have proven to have the best results in the image classification tasks; these include Xception, EfficientNet B7, and ResNet50. These three models are utilized in a binary classification problem, where the data set used comprises a total of 6,540 images that have undergone an augmentation procedure. Data augmentation improves and adapts the images with the use of enhancements such as rotation, flipping, zooming in, or altering the lighting conditions. This technique is useful to improve the generalization ability of the model, especially in instances where one has a small dataset.

The data distribution used comprises 5000 images for training, 1000 images for testing, and 540 images for validation. Thus, the percentage of validation data was determined in a way that it would, at all costs, facilitate proper and fair validation of the model during the training phase, so much so that the underlying cause of reducing training data would not be compelling. To ensure fairness across architectures, each model was retrained using its optimal native resolution:  $224 \times 224$  for ResNet-50,  $299 \times 299$  for Xception, and  $600 \times 600$  for EfficientNet-B7 while keeping all other hyperparameters consistent. This design mitigates potential resolution-induced bias and yields more reliable inter-model comparisons.

The training processes adopted a two-class bowl with a sigmoid activation function in the output layer in order to help the model produce probabilities on the two classes depending on the nature of the task, which is binary. The prediction error of the model is computed using a binary cross-entropy loss function, which is a function of only two classes and measures the difference between the emission and the target variable of the model. To combat the problem of overfitting, a ratio of 0.5 as a dropout level was used on the last fully connected layer during the training process. Dropout is one of the useful techniques whereby during each iteration, a certain number of units are turned off, resulting in low correlation between neurons and hence enhancing the performance of the model on new data. Although models were trained

for 15 epochs in this study, validation accuracy trends suggest possible under-fitting. Future work will extend training with adaptive early stopping to ensure full convergence.

For optimization purposes, the model was initialized with a static learning rate of 0.001. In practical terms, since it is necessary for the training process to reach the tolerable convergence criteria, there exists a learning rate reduction policy that must be triggered when the metric does not improve for three successive epochs. A reduction factor of 0.25 is employed, which is aimed at lowering the learning rate in increments lest the model need to readjust in the final stages of training. The total number of epochs used to train the model was 15, wherein each epoch the model weights were changed with respect to the training data that was made available. The parameters involved in the training of the model are presented in Table 2.

**Table 2.** Parameter training

Parameter	Value
Learning rate	0.001
Loss	Cross entropy
Epochs	15
Dropout	0.5
Activation function	Sigmoid

## 2.5 Evaluation

The evaluation is one of the most critical steps in any machine learning study meant to quantify the classification achieved by the model. The focus here is on the ability of the model to convert the information obtained from the training data to the new data through the application of the predictions from the testing data. As far as the present research is concerned, some of the metrics used for evaluation include accuracy, sensitivity, specificity, and precision.

Accuracy is the metric predominantly used, which computes the ratio of all the correct predictions occupiers [24]. Class imbalance is common in most health-related datasets; therefore, additional metrics beyond accuracy are necessary for a meaningful evaluation. Sensitivity represents the ratio of true positives to all actual positive samples, while specificity measures the proportion of true negatives among all actual negative cases. Precision quantifies the ratio of true positives to all predicted positives. Using these complementary metrics, this study provides a comprehensive assessment of each model's capability to classify retinal images for glaucoma detection.

## 3 RESULT AND DISCUSSION

In this study work, the researchers employed three transfer learning architectures, namely, Xception, EfficientNet B7, and ResNet50, to identify the disease of glaucoma using retinal images. From the classification results, it was observed

that all three models were observed to have performed well in detecting various forms of glaucoma, which suggests that the models have some level of effectiveness that could see use in a clinical setting. As per the tests done, the Xception model was noted to be the best in glaucoma image classification among the other models. In particular, the model yielded 93% accuracy, 95% precision, 91%, and 94% for sensitivity and specificity, respectively. The resultant metrics confirmed the model’s extreme distinction between glaucoma and no-glaucoma retinal image potentials, which were rated higher than other models, including the Xception model.

The EfficientNet B7 model, on the other hand, was also impressive, as its performance analysis showed an accuracy of 91%, with precision, sensitivity, and specificity values of 92%, 91%, and 92%. Though slightly lower than the Xception model, these results still show that, although at a reasonable percentage level, the level of glaucoma detection was still quite high. The same approach was applied to the ResNet50 model that performed the worst among the tested models at 89%, 88%, 90%, and 88% on accuracy, precision, sensitivity, and specificity, respectively. However, in the other two models, these values still exhibit acceptable performance but are not the highest. The performance evaluation of all models is presented in detail in Table 3. It is also the key result of the study that deep learning models, and especially the Xception model, have a notable potential in increasing the accuracy of glaucoma detection by analyzing retinal images.

**Table 3.** The performance of models

Architecture	Accuracy	Precision	Sensitivity	Specificity	F1-Score	AUC-ROC
Xception	0.93	0.95	0.91	0.94	0.93	0.96
EfficientNet B7	0.91	0.92	0.91	0.92	0.91	0.95
Resnet50	0.89	0.88	0.90	0.88	0.89	0.93

Based on Table 3, we can see that the Xception model does not only have good accuracy but also has good precision, sensitivity, and specificity too. High sensitivity means that the model is good at identifying patients suffering from glaucoma, and high specificity means that the model has mastered the ability to separate normal patients from those with the disease. To relax even more these restrictions on the Xception model’s authors and provide more coverage to this section of the paper. Figures 3, 4, and 5 present the training accuracy graphs of the EfficientNet-B7, ResNet-50, and Xception models, respectively, over 14 epochs. In Figure 3, the EfficientNet-B7 model demonstrates a smooth and consistent improvement in training accuracy, with validation accuracy surpassing 91% early in the training phase, indicating strong generalization performance. Figure 4 shows the ResNet-50 model achieving a steady rise in training accuracy from approximately 58% to 89%, while the validation accuracy remains flat in the early epochs before increasing significantly after epoch 6, suggesting a delayed convergence due to deeper network adaptation. Meanwhile, in Figure 5, the Xception model achieves rapid improvements in both training and validation accuracy within the first few epochs, maintaining 93% performance throughout the training process. This reflects the effectiveness of depth-wise separable convolutions in accelerating convergence while preserving model accuracy and stability.

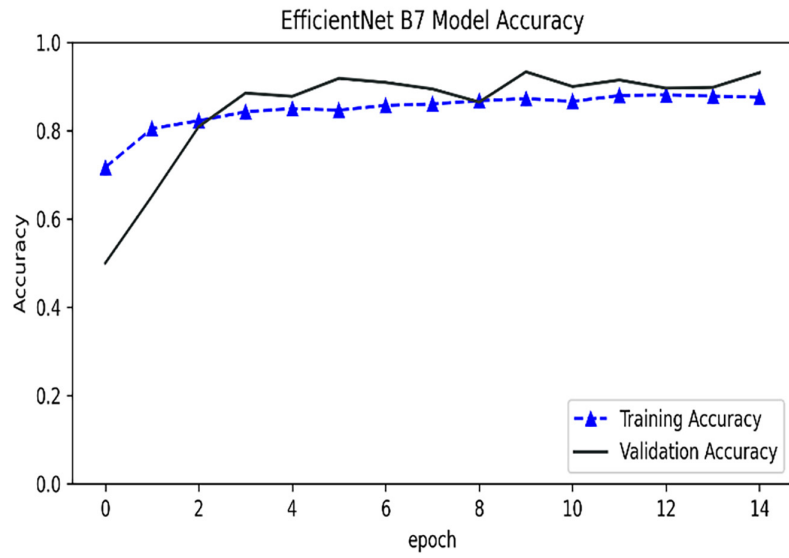


Fig. 3. Training process of EfficientNet B7

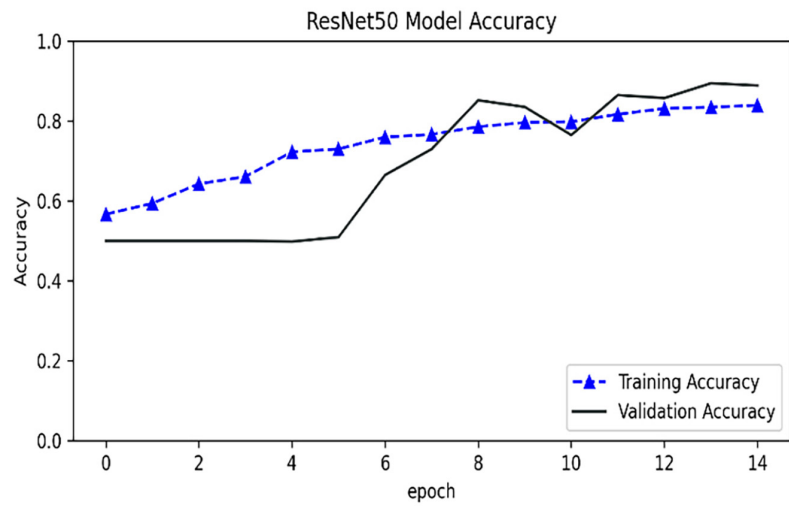


Fig. 4. Training process of ResNet-50

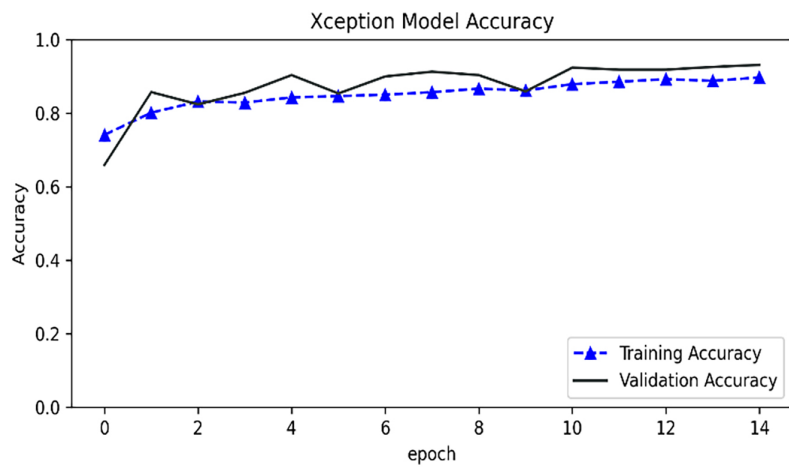


Fig. 5. Training process of Xception

The results indicate that the Xception architecture demonstrates superior performance in glaucoma classification compared to EfficientNet-B7 and ResNet-50. This superiority can be attributed to its use of depth-wise separable convolutions, which decompose standard convolutions into spatial and channel-wise operations. This design enables the network to capture fine-grained, localized features of the optic disc and cup regions, which are critical biomarkers for glaucomatous damage, while maintaining computational efficiency.

In contrast, the ResNet-50 architecture, which relies on residual connections to preserve gradient flow in deep networks, primarily focuses on global contextual information. While this aids in feature stability, it may reduce sensitivity to subtle local variations essential for early glaucoma detection. Similarly, EfficientNet-B7, which employs compound scaling to optimize depth, width, and resolution simultaneously, generally performs best at higher image resolutions. Since our dataset resolution was standardized for computational efficiency, EfficientNet-B7's potential was likely underutilized in this context.

From a clinical perspective, these findings underscore the value of architectures optimized for localized feature extraction in ophthalmic imaging tasks. The Xception model's balance between diagnostic accuracy and computational cost suggests its strong suitability for real-world deployment in resource-limited screening environments where rapid inference and model efficiency are crucial.

Overall, this comparative evaluation provides not only empirical evidence of Xception's effectiveness but also architectural insight into how CNN design choices influence feature learning behavior in medical image analysis. These observations may guide future research toward more targeted model selection and optimization for disease-specific diagnostic applications. Clinically, the results indicate that Xception provides an effective trade-off between accuracy and computational demand, making it suitable for integration into screening tools deployed in low-resource environments where real-time processing is essential.

## 4 CONCLUSION

The results provide insight into the architectural trade-offs among leading transfer learning models, guiding researchers in choosing efficient frameworks for real-world ophthalmic screening deployments. We have created and evaluated a classification model for glaucomatous disease based on the Xception approach, which is highly efficient for image classification tasks. The results validate our initial hypothesis that the Xception architecture, with its depth-wise separable convolutions, yields superior performance in glaucoma detection tasks. The algorithm created is capable of realizing a significant level of accuracy in the detection of glaucoma from retinal images, with values reaching 93% accuracy, 95% precision, 91% sensitivity, and 94% specificity. This is consistent with previous studies, which report that there can be improved performance in the diagnosis of eye disease when transfer learning approaches are utilized.

To conclude, the findings provide evidence that in order to further improve the ability to classify disease using advanced deep learning architectures, including the Xception, it will be revolutionary to change the clinical practice of managing ocular disease, especially glaucoma. It is acknowledged that more sophisticated neural networks have the ability to improve the processes of diagnosis and thereby enhance the efficiency and the effectiveness of disease control measures, which include the timing of the diagnosis and, by extension, the outcomes of the patients, which in some cases can be life-threatening.

Furthermore, effectively extending the generalization performance of the developed glaucoma detection models will depend on the incorporation of the general datasets since a more general dataset includes a wider diversity of retinal images from several populations and conditions. Though large datasets mainly serve to improve the learning ability of the models, they also add more power to the testing procedures, indicating how well the model will perform. This is of special importance because it is likely that these models will be deployed in a clinical setting where it is ideal for broader generalization to novel data outside those witnessed during training. In addition, use of additional types of information, such as patient age, wiping gender information, or medical records, including treatments conducted and diagnoses made, could be added potentially in the future and can prove important. This would facilitate the construction of more complete models that would recognize multiple more aspects of glaucoma risk and allow more efficient diagnosis of the disease.

## 5 ACKNOWLEDGMENTS

This study was funded entirely by Riset Kompetitif from Universitas Muhammadiyah Surakarta, reference number 303037/A/3-III/LRI/X/2023. The funders aided in the conduct of the study, as funds were available for the entire study, including the collection and analysis of data and dissemination of the research work. Many thanks for this generous funding: this study would not have been possible.

## 6 REFERENCES

- [1] M. Bansal, M. Kumar, M. Sachdeva, and A. Mittal, "Transfer learning for image classification using VGG19: Caltech-101 image data set," *Journal of Ambient Intelligence and Humanized Computing*, vol. 14, no. 4, pp. 3609–3620, 2021. <https://doi.org/10.1007/s12652-021-03488-z>
- [2] M. J. Burton *et al.*, "The lancet global health commission on global eye health: Vision beyond 2020," *The Lancet Global Health*, vol. 9, pp. e489–e551, 2021. [https://doi.org/10.1016/S2214-109X\(20\)30488-5](https://doi.org/10.1016/S2214-109X(20)30488-5)
- [3] L. J. Coan *et al.*, "Automatic detection of glaucoma via fundus imaging and artificial intelligence: A review," *Survey of Ophthalmology*, vol. 68, no. 1, pp. 17–41, 2023. <https://doi.org/10.1016/j.survophthal.2022.08.005>
- [4] H. Imaduddin, F. Y. A'la, A. Fatmawati, and B. A. Hermansyah, "Comparison of transfer learning method for COVID-19 detection using convolution neural network," *Bulletin of Electrical Engineering and Informatics*, vol. 11, no. 2, pp. 1091–1099, 2022. <https://doi.org/10.11591/eei.v11i2.3525>
- [5] R. Mahum *et al.*, "A novel hybrid approach based on deep CNN to detect glaucoma using fundus imaging," *Electronics*, vol. 11, no. 1, p. 26, 2022. <https://doi.org/10.3390/electronics11010026>
- [6] A. Haider, M. Arsalan, C. Park, H. Sultan, and K. R. Park, "Exploring deep feature-blending capabilities to assist glaucoma screening," *Applied Soft Computing*, vol. 133, p. 109918, 2023. <https://doi.org/10.1016/j.asoc.2022.109918>
- [7] S. Pathan, P. Kumar, R. M. Pai, and S. V. Bhandary, "Automated segmentation and classification of retinal features for glaucoma diagnosis," *Biomedical Signal Processing and Control*, vol. 63, p. 102244, 2021. <https://doi.org/10.1016/j.bspc.2020.102244>

- [8] A. A. Reshi *et al.*, “An efficient CNN model for COVID-19 disease detection based on X-Ray image classification,” *Complexity*, vol. 2021, p. 6621607, 2021. <https://doi.org/10.1155/2021/6621607>
- [9] S. Wang *et al.*, “Stacked dilated convolutions and asymmetric architecture for U-Net-based medical image segmentation,” *Computers in Biology and Medicine*, vol. 148, p. 105891, 2022. <https://doi.org/10.1016/j.compbiomed.2022.105891>
- [10] A. A. Abdulrahman, M. Belal, M. Abed, and M. Mohammed, “Incorporating a novel dual transfer learning approach for medical images,” *Sensors*, vol. 23, no. 2, p. 570, 2023. <https://doi.org/10.3390/s23020570>
- [11] X. Tang, F. Rashid, and S. Sheykahmad, “Boosted dipper throated optimization algorithm-based Xception neural network for skin cancer diagnosis: An optimal approach,” *Heliyon*, p. e26415, 2024. <https://doi.org/10.1016/j.heliyon.2024.e26415>
- [12] D. Lakhan and S. Sharma, “Multi-stage glaucoma classification using pre-trained convolutional neural networks and voting-based classifier fusion,” *Frontiers in Physiology*, vol. 14, p. 1175881, 2023. <https://doi.org/10.3389/fphys.2023.1175881>
- [13] O. Antonijević, B. Bajat, and M. Kilibarda, “Transfer learning approach based on satellite image time series for the crop classification problem,” *Journal of Big Data*, vol. 10, no. 1, p. 19, 2023. <https://doi.org/10.1186/s40537-023-00735-2>
- [14] A. Mastropaolo *et al.*, “Using transfer learning for code-related tasks,” *IEEE Transactions on Software Engineering*, vol. 49, pp. 1580–1598, 2022. <https://doi.org/10.1109/TSE.2022.3183297>
- [15] M. Agarwal, K. S. Gill, D. Upadhyay, and S. Dangi, “Charting new frontiers using transfer learning in OCT image analysis for retinal health,” in *International Conference on Sustainable Expert Systems (ICSSES)*, Tumakuru, India, 2024, pp. 1–5. <https://doi.org/10.1109/ICSSES62373.2024.10561346>
- [16] R. M. Bhavadharini *et al.*, “Eye disease detection using deep learning models with transfer learning techniques,” *ICST Transactions on Scalable Information Systems*, vol. 11, 2024. <https://doi.org/10.4108/eetsis.5971>
- [17] K. R. Ummah *et al.*, “Effect of image pre-processing method on convolutional neural network classification of COVID-19 CT scan images,” *International Journal of Innovative Computing, Information and Control*, vol. 18, no. 6, pp. 1895–1912, 2022. <https://doi.org/10.24507/ijicic.18.06.1895>
- [18] S. H. Wang *et al.*, “COVID-19 classification by CCSHNet with deep fusion using transfer learning and discriminant correlation analysis,” *Information Fusion*, vol. 68, pp. 131–148, 2021. <https://doi.org/10.1016/j.inffus.2020.11.005>
- [19] R. Kiefer *et al.*, “A catalog of public glaucoma datasets for machine learning applications: A detailed description and analysis of public glaucoma datasets available to machine learning engineers tackling glaucoma-related problems using retinal fundus images and OCT images,” in *Proceedings of the 2023 7th International Conference on Information System and Data Mining*, 2023. <https://doi.org/10.1145/3603765.3603779>
- [20] C. de Vente, K. A. Vermeer, N. Jaccard, B. van Ginneken, H. G. Lemij, and C. I. Sánchez, “Rotterdam EyePACS AIROGS train set,” Zenodo, 2021. <https://doi.org/10.5281/zenodo.5793241>
- [21] N. N. Fatihah, G. Ariyanto, A. J. Latipah, and D. M. Pangestuty, “Face recognition using local binary pattern and nearest neighbour classification,” in *2018 International Symposium on Advanced Intelligent Informatics (SAIN)*, IEEE, 2018, pp. 142–147. <https://doi.org/10.1109/SAIN.2018.8673375>
- [22] J. Pardede and A. S. Purohita, “The advantage of transfer learning with pre-trained model in CNN towards CT-scan classification,” *Khazanah Informatika: Jurnal Ilmu Komputer Dan Informatika*, vol. 9, no. 2, pp. 155–161, 2023. <https://doi.org/10.23917/khif.v9i2.19872>

- [23] H. C. Reis *et al.*, “MediNet: Transfer learning approach with MediNet medical visual database,” *Multimedia Tools and Applications*, vol. 1, pp. 1–44, 2023. <https://doi.org/10.1007/s11042-023-14831-1>
- [24] S. S. Mawarni, M. Murinto, and S. Sunardi, “Medical external wound image classification using support vector machine technique,” *Khazanah Informatika: Jurnal Ilmu Komputer Dan Informatika*, vol. 9, no. 2, pp. 98–103, 2023. <https://doi.org/10.23917/khif.v9i2.22541>

## 7 AUTHORS

**Helmi Imaduddin** is currently serving as a Lecturer in Informatics Engineering at the Universitas Muhammadiyah Surakarta, where he actively engages in both teaching and research. His research interests are medical image processing, natural language processing (NLP), and machine learning. His academic journey began in 2015 when he earned a Bachelor of Science degree in Informatics Engineering from Universitas Muhammadiyah Surakarta, where he developed a solid foundation in computer science and engineering. He pursued and completed his Master of Science degree in Information Technology at Gadjah Mada University, Yogyakarta, in 2020 (E-mail: [helmi.imaduddin@ums.ac.id](mailto:helmi.imaduddin@ums.ac.id)).

**Widi Widayat** is an Assistant Professor at the Informatics Department, Universitas Muhammadiyah Surakarta, Indonesia. He received his Bachelor degree in Informatics Engineering from Universitas Islam Indonesia, Indonesia, in 2014 and then his Master degree in Information technology from Universitas Gadjah Mada, Indonesia, in 2019. His research areas are NLP (Natural Language Processing), sentiment analysis, word embedding, and machine learning (E-mail: [widi.widayat@ums.ac.id](mailto:widi.widayat@ums.ac.id)).

**Ihsan Cahyo Utomo** is currently a Lecturer at the Faculty of Communication and Informatics, University of Muhammadiyah Surakarta. His primary research interests lie in the areas of Computer Networks, the Internet of Things, and Information Systems. He earned his Bachelor of Science degree in Informatics Engineering from Muhammadiyah University of Surakarta in 2011. Subsequently, in 2016, he completed his Master’s degree in Informatics Engineering from Amikom Yogyakarta University. For communication purposes (E-mail: [icu886@ums.ac.id](mailto:icu886@ums.ac.id)).

**Nur Rizqi Febriandika** is currently a Lecturer/researcher at Islamic Economic Law, Muhammadiyah University of Surakarta. He hold two master’s degrees, namely a Master of Islamic Economics (M.SEI) and a Master of Business Administration (MBA). After graduation, his research has predominantly focused on the marketing area, with some contributions to digital marketing and e-commerce (E-mail: [nrf679@ums.ac.id](mailto:nrf679@ums.ac.id)).

**Amri Zadi Hudaya** is a student pursuing Bachelor’s Computer Science at the Faculty of Communication and Informatics, Universitas Muhammadiyah Surakarta, Indonesia. His research interests include computer vision, object detection, and image processing. He is particularly interested in developing intelligent vision-based systems that enhance automation and visual data analysis in real-world applications.

**Alleyda Aurora Yafentri** is a student pursuing Bachelor’s in Computer Science at the Faculty of Communication and Informatics, Universitas Muhammadiyah Surakarta, Indonesia. Her research interests include software engineering and data mining. She is particularly interested in exploring intelligent data-driven systems to improve software development processes and decision-making in practical applications.