

■ RESEARCH ARTICLE

Deep Learning Based Decision Support System for Retinal Disease Classification: Diabetic Retinopathy and Macular Hole

Belinay Kabataş^a , Emre Ölmez^{at} 

^a Department of Biomedical Engineering, İzmir Bakırçay University, İzmir, Turkey

[†] emre.olmez@bakircay.edu.tr, corresponding author

RECEIVED APRIL 13, 2025
ACCEPTED APRIL 25, 2025

CITATION Kabataş, B., & Ölmez, E. (2025). Deep learning based decision support system for retinal disease classification: diabetic retinopathy and macular hole. *Artificial Intelligence Theory and Applications*, 5(1), 51-62.

Abstract

In this study, a deep learning-based decision support system was developed to classify diabetic retinopathy (DR), macular hole (MH), and healthy cases using fundus images. A total of 1,397 fundus images, selected from the open-source Retinal Disease Classification dataset, were used in the training and testing phases. ResNet50, InceptionV3, and Xception models were trained with different hyperparameter configurations, and their performances were evaluated comparatively. Among the models, ResNet50 achieved the highest accuracy on the test set, reaching 93.79%. However, the Xception model exhibited superior robustness and stability across various hyperparameter settings, consistently delivering balanced and reliable classification performance. These findings indicate that deep learning-based approaches can be effectively utilized as clinical decision support systems for the diagnosis of retinal diseases.

Keywords: deep learning, fundus images, diabetic retinopathy, macular hole, convolutional neural networks, Resnet50, Xception, InceptionV3

1. Introduction

The use of artificial intelligence (AI) methods in healthcare has rapidly expanded in recent years. In particular, AI applications powered by deep learning techniques are increasingly being adopted in the medical field [1]. In this context, the early diagnosis of eye diseases is crucial for preventing permanent vision loss. Conditions such as diabetic retinopathy (DR) and macular hole (MH) can lead to severe visual impairment, especially in their advanced stages [2]. Fundus images play a critical role in the diagnosis of these diseases by enabling a detailed examination of the retinal layer, thus supporting physicians in the decision-making process [2], [3]. However, manual interpretation of fundus images is time-consuming and susceptible to human error. At this point, digital image processing and deep learning methods offer valuable assistance by serving as decision support systems for the classification of fundus images [4].

In this study, deep learning-based image processing models were trained to classify DR [5], MH [6], and healthy samples using fundus images. The Retinal Disease Classification dataset [7], an open-source collection of fundus images representing different disease

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. Copyrights for components of this work owned by others than AITA must be honoured. Abstracting with credit is permitted, and providing the material is not used for any commercial purposes and is shared in its entire and unmodified form. Request permissions from info@aitajournal.com

Artificial Intelligence Theory and Applications, ISSN: 2757-9778. ISBN: 978-605-69730-2-4 © 2025 İzmir Bakırçay University

categories, was used in the study. These images were processed and trained with various hyperparameter configurations using ResNet50 [8], InceptionV3 [9], and Xception [10] models, and the performance of each model was comparatively analyzed. The results indicated that ResNet50 and Xception achieved high classification accuracies of 93.70% and 92.94%, respectively, by effectively capturing distinctive features in fundus images. InceptionV3 also performed well with an accuracy of 88.70%, though slightly lower than the other two models.

This study highlights the effectiveness of deep learning-based artificial intelligence approaches in the diagnosis of retinal diseases and aims to contribute to future clinical applications. The findings support the integration of AI-powered decision support systems, particularly in the early detection of conditions such as DR and MH, where early diagnosis is critical.

2. Retinal Diseases

The retina is one of the fundamental structures responsible for the visual function of the eye. Retinal diseases can lead to serious and permanent vision loss if not diagnosed and treated in a timely manner [11]. In this section, we focus on DR and MH, two common retinal disorders that fall within the scope of this study and can cause significant visual impairment if not detected early.

2.1 Diabetic Retinopathy

Diabetic retinopathy (DR) is a serious retinal disease affecting one-third of the approximately 285 million people with diabetes worldwide. One third of these individuals also have vision-threatening symptoms of DR [12]. DR occurs in diabetic patients when the blood vessels in the retina are damaged due to high blood glucose levels [2]. Since the disease is usually asymptomatic in the initial stages, it is difficult to diagnose early and often manifests itself in later stages with symptoms such as blurred vision, dark spots in the visual field and vision loss. Therefore, early diagnosis of DR is critical to stop the progression of the disease and prevent vision loss. In the absence of early diagnosis and treatment, the disease can cause severe vision loss, up to blindness [5]. Figure 1 presents sample images labeled as DR from the Retinal Disease Classification dataset [7].



Figure 1. Fundus images labeled as DR in the Retinal Disease Classification dataset.

2.2 Macular Hole

A macular hole (MH) is a small tear or opening that occurs in the macula, the central region of the retina. Since the macula is responsible for sharp and detailed central vision, a hole in this area can significantly impair visual acuity [6]. A MH is often associated with

the natural process of aging; it occurs as the structure of the vitreous fluid deteriorates and separates from the macula with age. The incidence is particularly high in individuals over 50 years of age. Early signs of the disease include distorted central vision, blurred vision and the inability to see fine details. When treatment is delayed, the damage to the macular area deepens and this can lead to permanent vision loss. Early detection of MH and appropriate intervention can preserve central vision [6], [13]. Figure 2 presents sample images labeled as MH from the Retinal Disease Classification dataset [7].



Figure 2. Fundus images labeled as MH in the Retinal Disease Classification dataset.

3. Literature Review

In recent years, deep learning-based models developed for the early diagnosis of retinal diseases—especially DR—have attracted significant interest in the research community. The potential of artificial intelligence-based systems to enhance the efficiency of clinical processes is particularly evident when faced with limited data and complex diagnostic challenges. In this section, we review some of the most influential studies on the classification and diagnosis of retinal diseases.

Kori et al. (2018) employed a CNN-based ensemble approach for the automatic grading of DR and macular edema (ME). To address the challenge of limited labeled data, the researchers utilized transfer learning by fine-tuning models that were previously trained on ImageNet, adapting them to fundus images. The final model achieved an accuracy of 83.9% for DR grading and 95.45% for ME grading. The study emphasized that the ensemble method outperformed a single CNN model and highlighted the effectiveness of transfer learning techniques [14].

Sahlsten et al. (2019) proposed a deep learning-based method for the automatic detection of DR and ME using high-resolution fundus images. Their study achieved high accuracy rates, emphasizing the potential for increased cost-effectiveness in existing screening programs [15].

Torre et al. (2020) developed a deep learning classifier aimed at improving interpretability in DR grading. Their method assigned importance scores to individual pixels or regions contributing to the final classification, enhancing transparency for clinical experts. This not only improved the diagnostic reliability but also underscored its potential for integration into clinical decision support systems [16].

Özçelik and Altan (2021) introduced a two-stage model for the early diagnosis of DR. In the first stage, two-dimensional signal processing techniques were utilized to prevent overfitting, while in the second stage, classification was performed using ESA-based transfer learning. The model was trained on 5,100 fundus images and achieved an accuracy of 97.8%. This study demonstrated the model's speed and reliability as a diagnostic tool [5].

Aykat and Senan (2023) proposed a deep learning-based method for diagnosing retinal diseases such as DR and cataract. In their study, fundus images were enhanced using histogram equalization as preprocessing and 99% accuracy was achieved with the MobileNet-based hybrid model. These results suggest that the hybrid model outperforms similar methods in existing literature [2].

Polater and Işık (2024) conducted a study on the classification of DR severity levels using the APTOS 2019 dataset. By employing the DenseNet121 model, they achieved approximately 97% accuracy. Their findings reaffirm the superior performance of the DenseNet121 architecture and the overall efficacy of deep learning methods in DR diagnosis [17].

These studies clearly demonstrate that deep learning methods offer high accuracy and reliability in the diagnosis of DR and other retinal diseases. Validating these models on diverse datasets and across various clinical scenarios may broaden their applicability in diagnostic and treatment workflows and contribute to the development of robust clinical decision support systems.

4. Material and Method

4.1 Dataset

The Retinal Disease Classification dataset [7] used in this study is a comprehensive and open-source dataset designed for the classification of eye diseases based on retinal fundus images. It contains a total of 3,200 fundus images representing 46 distinct ocular diseases. The images were captured using three different fundus cameras—TOPCON 3D OCT-2000, Kowa VX-10, and TOPCON TRC-NW300—and each image was meticulously labeled by two senior retina specialists. The use of multiple imaging devices and expert annotations enhances both the diversity and reliability of the dataset.

The fact that the images were obtained from different devices increases the generalization capability of the deep learning models by reducing dependency on a specific device or lighting condition. Additionally, the dataset's wide range of disease classes facilitates the development of models capable of detecting multiple retinal disorders simultaneously. For the purpose of this study, three classes were selected: DR, MH, and Healthy (No Disease). These classes are commonly encountered in clinical settings and exhibit a relatively balanced distribution within the dataset, allowing for more consistent and reliable results during model training and evaluation.

From the 1,397 fundus images selected for this study, a total of 1,043 images were allocated to the training set, comprising 349 DR, 293 MH, and 401 healthy images. The remaining 354 images were used for testing, including 120 DR, 100 MH, and 134 healthy images. Accordingly, approximately 75% of the data was used for training and 25% for testing. Figure 3 illustrates representative fundus images from each of the three selected classes.



Figure 3. Sample fundus images from Retinal Disease Classification dataset. From left to right: Macular Hole, Diabetic Retinopathy, No Disease

4.1.1 Image Preprocessing

Various image preprocessing steps were applied to ensure high accuracy and generalization capability of the trained deep learning models. Fundus images were resized to 299×299 pixels to be compatible with the input layers of the deep learning models. In addition, the pixel values of the images were normalized to the range [0, 1] to facilitate the training process of the models.

In this study, data augmentation strategies were also included in the training processes. Figure 4 shows the data augmentation process used in the training scenarios where the data augmentation strategy was applied. Data augmentation aims to diversify the limited amount of training data and increase the robustness of the models against different variations. The operations in the data augmentation process were randomly applied to the images at each iteration. The applied methods include random rotation up to 30 degrees (rotation_range=30), horizontal and vertical shift up to 20% (width_shift_range=0.2, height_shift_range=0.2), shear up to 20% (shear_range=0.2), zoom in up to 20% (zoom_range=0.2), random change of brightness values within a 20% range (brightness_range=[0.8, 1.2]) and random flip on the horizontal axis (horizontal_flip=True).

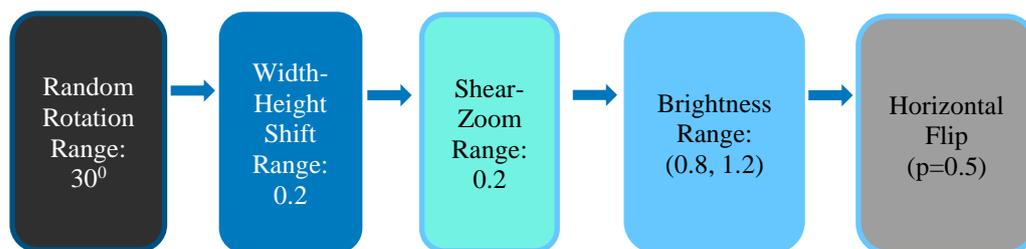


Figure 4. Data augmentation process

4.2 Deep Neural Networks

Deep learning, as one of the cornerstones of modern artificial intelligence research, has achieved significant advances in image processing and classification. In particular, convolutional neural networks (CNNs) have demonstrated remarkable success in analyzing and classifying visual data, and these models have evolved to become more efficient and effective over time [10], [18]. In this study, three different CNN-based deep learning architectures were employed for retinal image classification: ResNet50, InceptionV3, and Xception.

ResNet50, developed by He et al. in 2015, is a CNN architecture designed to address the vanishing gradient problem encountered during the training of deep neural networks [8]. It introduces **residual connections** that allow the output of a layer to be added to the input of a deeper layer, enabling more effective training of very deep architectures. ResNet50 consists of a total of 50 layers and has demonstrated high performance in complex image classification tasks.

InceptionV3 is a CNN model developed by Google that applies convolutional filters of varying sizes in parallel, allowing for the extraction of image features at multiple spatial scales [9]. This **multi-scale filtering** approach enables the model to capture visual patterns at different levels of detail while improving parameter efficiency and reducing computational cost. Thanks to this design, InceptionV3 achieves high classification accuracy and is widely adopted across various computer vision applications.

Xception is a CNN architecture designed as an enhanced version of the Inception model [10]. It utilizes depthwise separable convolutions to reduce the number of parameters and improve computational efficiency compared to traditional CNN structures. Xception has outperformed InceptionV3 on large-scale datasets such as ImageNet (ILSVRC) and JFT. Furthermore, it has proven highly effective in transfer learning scenarios, achieving strong performance across diverse classification tasks.

5. Training and Results

In this study, three different CNN based models—ResNet50, InceptionV3, and Xception—were used, and a total of 15 training scenarios were evaluated, combining three different hyperparameters (end-to-end learning, data augmentation, and learning rate) for each architecture. All models were trained using the Adam optimization algorithm for 50 epochs, with learning rates set at 0.001 and 0.0001. The detailed training and test accuracy results, along with loss values for each scenario, are summarized in Table 1. Additionally, Figure 6 illustrates the epoch-based test accuracy progression for each model throughout the training phase, while Figure 7 presents corresponding accuracy curves for the training sets. Further insights into model convergence and stability can be observed in the loss curves presented in Figures 8 and 9, showing test and training loss, respectively. Moreover, Table 2 provides a more comprehensive evaluation, detailing precision, recall, and F1-score metrics for each class and scenario, enabling deeper analysis beyond overall accuracy.

In each model architecture, the original output layers were removed, and a Global Average Pooling layer was added to the end of the base model to adapt it to the target dataset. This approach converts the output feature maps into a one-dimensional vector, helping reduce the number of parameters while preserving high-level feature representations. Following this, a Dense layer with 256 neurons and ReLU activation was added, along with 30% Dropout to prevent overfitting. Finally, a Softmax-activated output layer was used for classification into three classes: Diabetic Retinopathy, Macular Hole, and Healthy. The block diagram illustrating this modified CNN architecture is shown in Figure 5.

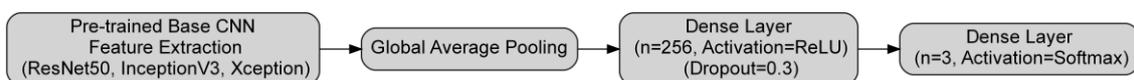


Figure 5. Modified CNN Architecture

In Table 1, when the "End to End Learning" parameter is "-", the base model weights are frozen, and only the newly added layers are trained. When the parameter is "+", training is performed end-to-end. All models were initialized with pre-trained weights from the ImageNet dataset. In scenarios where data augmentation was applied, the strategy illustrated in Figure 4 was used to evaluate the robustness of the models against image variations.

Table 1. Training and Test Results

Model	End to End Learning	Augmentation	Learning Rate	Loss	Acc	Test-Loss	Test-Acc
ResNet_1	-	-	0.0010	1.0291	0.4382	1.0527	0.4209
ResNet_2	-	+	0.0010	1.0568	0.4267	1.0506	0.4379
ResNet_3	-	-	0.0001	1.0240	0.4861	1.0628	0.4237
ResNet_4	+	-	0.0001	0.0033	0.9990	0.2751	0.9379
ResNet_5	+	+	0.0001	0.1276	0.9569	0.6505	0.8559
Inception_1	-	-	0.0010	0.0379	0.9895	0.4572	0.8814
Inception_2	-	+	0.0010	0.2624	0.8907	0.3732	0.8701
Inception_3	-	-	0.0001	0.0732	0.9818	0.3131	0.8870
Inception_4	+	-	0.0001	0.0122	0.9971	0.4580	0.8870
Inception_5	+	+	0.0001	0.0856	0.9732	0.3672	0.8729
Xception_1	-	-	0.0010	0.0179	0.9952	0.4203	0.9040
Xception_2	-	+	0.0010	0.1965	0.9271	0.2570	0.9153
Xception_3	-	-	0.0001	0.1036	0.9703	0.2301	0.9181
Xception_4	+	-	0.0001	0.0950	0.9962	0.4866	0.9294
Xception_5	+	+	0.0001	0.0203	0.9942	0.2378	0.9294

Table 2. Test set evaluation metrics

Model	DR			MH			No_Disease		
	Precision	Recall	F1	Precision	Recall	F1	Precision	Recall	F1
ResNet_1	0.20	0.02	0.03	0.61	0.14	0.23	0.41	0.99	0.58
ResNet_2	0.00	0.00	0.00	0.65	0.22	0.33	0.42	0.99	0.59
ResNet_3	0.50	0.03	0.05	0.64	0.14	0.23	0.41	0.99	0.58
ResNet_4	0.97	0.92	0.94	0.92	0.94	0.93	0.92	0.96	0.94
ResNet_5	1.00	0.66	0.79	0.89	0.93	0.91	0.77	0.98	0.86
Inception_1	0.96	0.78	0.86	0.85	0.86	0.86	0.85	0.99	0.91
Inception_2	0.96	0.78	0.86	0.92	0.81	0.86	0.79	0.99	0.88
Inception_3	0.93	0.83	0.88	0.91	0.83	0.87	0.85	0.98	0.91
Inception_4	0.98	0.82	0.89	0.77	0.95	0.85	0.93	0.90	0.92
Inception_5	0.99	0.81	0.89	0.89	0.84	0.87	0.79	0.96	0.86
Xception_1	0.96	0.89	0.92	0.91	0.84	0.88	0.86	0.96	0.91
Xception_2	0.93	0.90	0.92	0.92	0.87	0.89	0.90	0.96	0.93
Xception_3	0.94	0.89	0.91	0.89	0.90	0.90	0.92	0.96	0.94
Xception_4	0.95	0.95	0.95	0.98	0.84	0.90	0.89	0.98	0.93
Xception_5	0.98	0.87	0.92	0.90	0.94	0.92	0.91	0.98	0.94

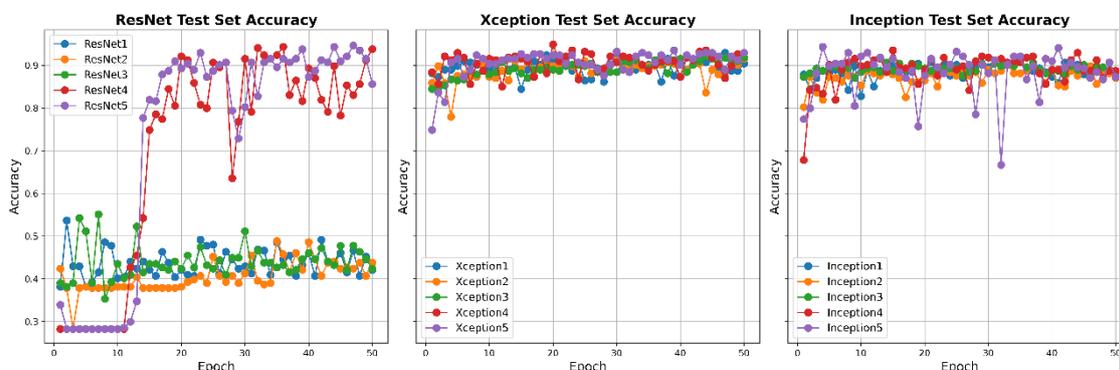


Figure 6. Test set accuracies of the models over 50 epochs

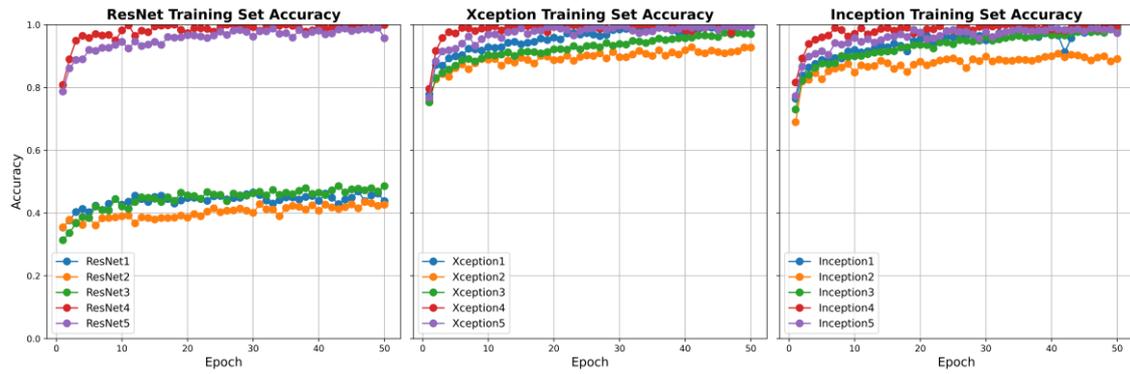


Figure 7. Training set accuracies of the models over 50 epochs

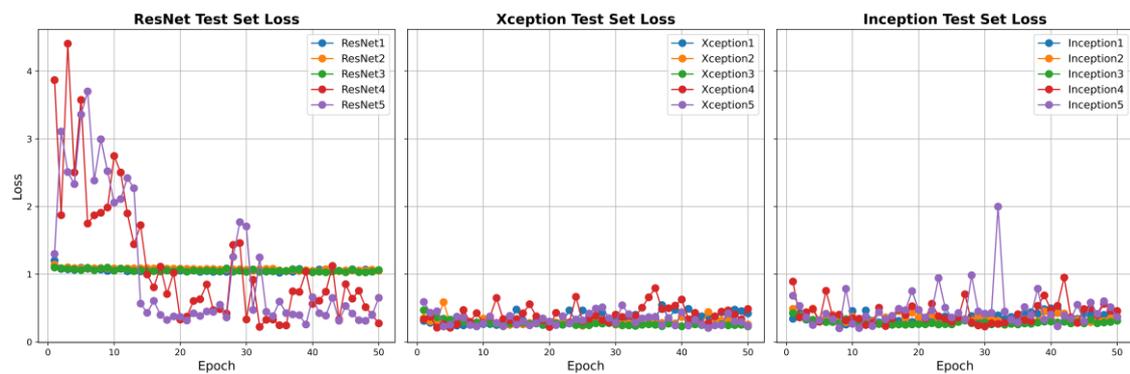


Figure 8. Test set loss of the models over 50 epochs

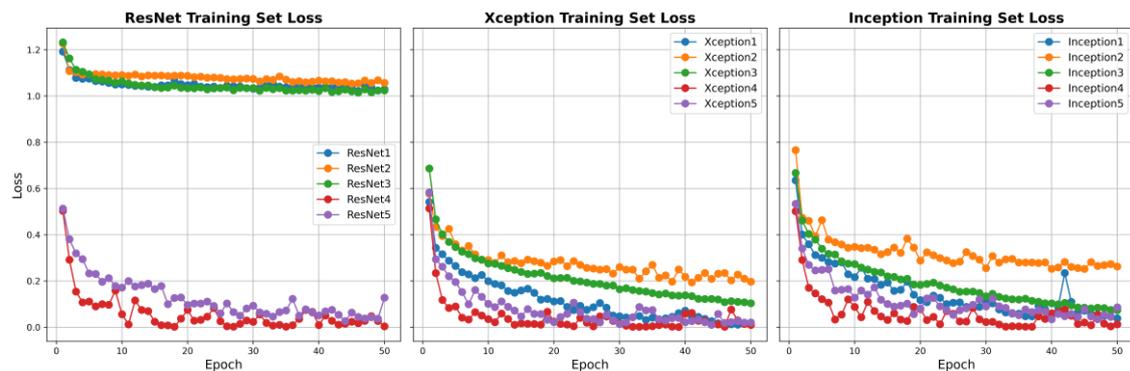


Figure 9. Training set loss of the models over 50 epochs

The results obtained from the training scenarios presented in Table 1 reveal important insights into the performances of the three convolutional neural network (CNN) architectures (ResNet50, InceptionV3, and Xception) across different hyperparameter configurations.

For the ResNet50 model, training scenarios utilizing a learning rate of 0.001 did not yield satisfactory results, especially when the base layers of the network were frozen. In these configurations, test accuracies remained notably low. However, reducing the learning rate to 0.0001 and performing end-to-end training of all layers dramatically improved the test accuracy, reaching 93.79%, which was the highest among all evaluated

configurations. This significant improvement indicates that ResNet50 requires careful adjustment of learning rate and complete fine-tuning to achieve optimal performance.

The experiments conducted using the InceptionV3 architecture showed relatively consistent but limited variation in performance across different hyperparameter combinations, with accuracy typically around 88.70%. Interestingly, scenarios that involved end-to-end training and data augmentation—despite theoretical expectations of enhanced robustness—demonstrated fluctuating accuracy levels without substantial improvements. This observation implies that InceptionV3 has inherent stability in training dynamics but may have reached a saturation point in extracting discriminative features from the fundus image dataset, limiting further performance gains.

On the other hand, the Xception architecture consistently demonstrated high and stable performance across all evaluated scenarios. It not only achieved higher average test accuracy compared to ResNet50 and InceptionV3 but also showed robust generalization and sensitivity to hyperparameter changes. Particularly, training the base layers with a lower learning rate notably enhanced its performance, underscoring the adaptability and robustness of the Xception architecture.

Beyond accuracy, additional metrics provided in Table 2 (precision, recall, and F1-score) offer deeper insight into the classification performance across each disease category (Diabetic Retinopathy, Macular Hole, and Healthy). Analysis of these metrics further highlights the superiority of Xception. Across all classes, Xception consistently outperformed the other architectures, achieving precision, recall, and F1-scores frequently exceeding 0.90 in configurations involving end-to-end learning and data augmentation. These findings reinforce Xception’s suitability for accurate multi-class classification tasks in ophthalmological applications.

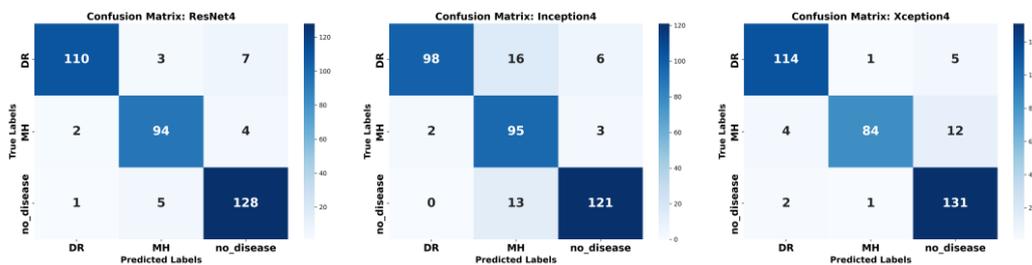


Figure 10. Confusion matrices of the models with high classification performance

Figure 10 presents the confusion matrices of the best-performing models trained with identical hyperparameter configurations. When evaluating class-wise performance based solely on the confusion matrices, the Xception model demonstrates the highest number of correct classifications for the DR class. Xception produced fewer misclassifications in this class compared to the other two models, indicating greater reliability in distinguishing DR cases. For the MH class, the most successful results were achieved by the InceptionV3 model, which attained the highest number of correct predictions, distinguishing itself from the other architectures. Although ResNet50 performed comparably to InceptionV3 in the MH class, it was observed that Xception made noticeably more classification errors in this category. In the no_disease class, Xception once again stood out, delivering the highest number of correct classifications and exhibiting strong performance in identifying healthy individuals. In contrast, InceptionV3 produced more misclassifications in this class, frequently confusing it with MH. In summary, the confusion matrices indicate that while Xception achieved superior

results in the DR and no_disease classes, InceptionV3 was the most effective in classifying MH cases. ResNet50, on the other hand, displayed a balanced performance across all classes with relatively low misclassification rates.

In contrast, ResNet50 exhibited significant performance variability between different scenarios. While scenarios involving frozen base layers and higher learning rates resulted in poor precision and recall scores, the architecture successfully recovered in configurations that involved end-to-end training with a reduced learning rate. This pattern suggests that ResNet50 requires a carefully controlled training environment to mitigate overfitting and achieve its full potential.

InceptionV3, while maintaining moderate stability, demonstrated competitive yet generally lower overall performance compared to Xception. Particularly, its precision, recall, and F1-scores for the Macular Hole and Healthy classes were commendable, but it was less consistent across the Diabetic Retinopathy class, highlighting a limitation in effectively distinguishing between visually challenging classes.

The training dynamics presented in Figures 6 and 7 further illustrate differences in learning behaviors among the architectures. Notably, Xception exhibited smoother and more consistent progression in both training and test accuracies throughout the epochs. Additionally, the loss trajectories shown in Figures 8 and 9 complement these observations, demonstrating more stable loss convergence for Xception, whereas ResNet50 and InceptionV3 experienced noticeable fluctuations, indicative of overfitting tendencies, particularly when training base layers were frozen or when higher learning rates were employed.

Collectively, these detailed analyses underscore the robustness, stability, and superior generalization capabilities of the Xception model. Its consistently high performance across various metrics and configurations makes it particularly suitable as the basis for reliable and effective clinical decision support systems in the early diagnosis of retinal diseases.

6. Conclusion and Future Work

This study provides a comparative analysis of the classification performance of deep learning-based models for the diagnosis of retinal diseases such as diabetic retinopathy (DR) and macular hole (MH). Three different CNN architectures—ResNet50, InceptionV3, and Xception—were trained and evaluated under various hyperparameter configurations. According to the findings, ResNet50 achieved the highest test accuracy (93.79%) among the models used in the study. However, as illustrated in Figure 6, the Xception model demonstrated more stable performance across different training scenarios. InceptionV3, while exhibiting more consistent performance than ResNet50, achieved lower accuracy than Xception. The results also indicate that both InceptionV3 and Xception models converged more rapidly during the early training stages compared to ResNet50, achieving high accuracy values in a shorter time frame, as depicted in Figure 6.

These results highlight the strong potential of deep learning methods in ophthalmologic image analysis and diagnostic decision support systems. Such systems, developed as alternatives to time-consuming and expert-dependent manual evaluations, can expedite clinical decision-making, reduce the workload on healthcare professionals, and enable earlier intervention for patients at risk of vision loss.

For future work, we plan to evaluate the current models on larger and more imbalanced datasets involving multi-class disease classification. Additionally, we aim to integrate explainable AI (XAI) techniques such as Grad-CAM and LIME to improve the interpretability of the models' decision-making processes. Exploring lightweight and optimized architectures (e.g., MobileNet, EfficientNet) suitable for real-time applications will also be a crucial step toward integration into portable medical devices. Furthermore, training models using multi-center, multi-device fundus image datasets is expected to improve the generalizability and reliability of decision support systems.

In conclusion, this study highlights the effectiveness of deep learning-based models in classifying retinal diseases using fundus images, providing a foundation for future research toward more reliable and interpretable clinical decision support systems in eye care.

Acknowledgement

We would like to thank Izmir Bakırçay University, The Center for Artificial Intelligence Studies and Research in Healthcare, for the resources and support provided for the implementation of the study

References

- [1] B. Vatansever, H. Aydın, and A. Çetinkaya, "Genetik Algoritma Yaklaşımıyla Öznitelik Seçimi Kullanılarak Makine Öğrenmesi Algoritmaları ile Kalp Hastalığı Tahmini," *Journal of Scientific Technology and Engineering Research*, Nov. 2021, doi: 10.53525/jster.1005934.
- [2] Ş. Aykat *et al.*, "Derin Öğrenme Kullanılarak Fundus Görüntülerinden Katarakt ve Diyabetik Retinopati Tespiti Detection of Cataract and Diabetic Retinopathy from Fundus Images Using Deep Learning," 2023.
- [3] Q. Wei *et al.*, "Development and Validation of an Automatic Ultrawide-Field Fundus Imaging Enhancement System for Facilitating Clinical Diagnosis: A Cross-sectional Multicenter Study," *Engineering*, Oct. 2024, doi: 10.1016/j.eng.2024.05.006.
- [4] K. Jin and J. Ye, "Artificial intelligence and deep learning in ophthalmology: Current status and future perspectives," Nov. 01, 2022, *Elsevier Inc.* doi: 10.1016/j.aopr.2022.100078.
- [5] Y. B. Özçelik and A. Altan, "Diyabetik Retinopati Teşhisi için Fundus Görüntülerinin Derin Öğrenme Tabanlı Sınıflandırılması," *European Journal of Science and Technology*, Dec. 2021, doi: 10.31590/ejosat.1011806.
- [6] E. Yaşar, N. Erol, M. D. Bilgeç, and A. İ. Çakmak, "Coexistence of peripheral retinal diseases with macular hole," *Turk J Ophthalmol*, vol. 49, no. 4, pp. 209–212, Aug. 2019, doi: 10.4274/tjo.galenos.2019.06706.
- [7] S. Pachade *et al.*, "Retinal fundus multi-disease image dataset (Rfmid): A dataset for multi-disease detection research," *Data (Basel)*, vol. 6, no. 2, pp. 1–14, Feb. 2021, doi: 10.3390/data6020014.
- [8] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," Dec. 2015, [Online]. Available: <http://arxiv.org/abs/1512.03385>
- [9] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, Dec. 2016, pp. 2818–2826. doi: 10.1109/CVPR.2016.308.
- [10] F. Chollet, "Xception: Deep learning with depthwise separable convolutions," in *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, Institute of Electrical and Electronics Engineers Inc., Nov. 2017, pp. 1800–1807. doi: 10.1109/CVPR.2017.195.
- [11] F. S. Sorrentino *et al.*, "Novel Approaches for Early Detection of Retinal Diseases Using Artificial Intelligence," Jul. 01, 2024, *Multidisciplinary Digital Publishing Institute (MDPI)*. doi: 10.3390/jpm14070690.
- [12] R. Lee, T. Y. Wong, and C. Sabanayagam, "Epidemiology of diabetic retinopathy, diabetic macular edema and related vision loss," Dec. 01, 2015, *BioMed Central Ltd.* doi: 10.1186/s40662-015-0026-2.
- [13] D. Mikhail *et al.*, "The role of artificial intelligence in macular hole management: A scoping review," Jan. 01, 2024, *Elsevier Inc.* doi: 10.1016/j.survophthal.2024.09.003.

- [14] A. Kori, S. S. Chennamsetty, M. S. K. P., and V. Alex, "Ensemble of Convolutional Neural Networks for Automatic Grading of Diabetic Retinopathy and Macular Edema," Sep. 2018, [Online]. Available: <http://arxiv.org/abs/1809.04228>
- [15] J. Sahlsten *et al.*, "Deep Learning Fundus Image Analysis for Diabetic Retinopathy and Macular Edema Grading," *Sci Rep*, vol. 9, no. 1, Dec. 2019, doi: 10.1038/s41598-019-47181-w.
- [16] J. de la Torre, A. Valls, and D. Puig, "A deep learning interpretable classifier for diabetic retinopathy disease grading," *Neurocomputing*, vol. 396, pp. 465–476, Jul. 2020, doi: 10.1016/j.neucom.2018.07.102.
- [17] S. N. Polater *et al.*, "Diyabetik Retinopati Tespiti İçin Derin Öğrenmeye Dayalı Sınıflandırma Deep Learning-Based Classification for Diabetic Retinopathy Detection," 2024.
- [18] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," May 27, 2015, *Nature Publishing Group*. doi: 10.1038/nature14539.