

Research Article

Development of a New Computational System Supported by Artificial Intelligence for Detection of Real-Time Retinal Diseases

Hasan Memis and Emrullah Acar

Abstract— Retinal diseases such as choroidal neovascularization (CNV), diabetic macular edema (DME), and drusen are among the leading causes of vision loss worldwide, requiring early and accurate diagnosis to prevent irreversible damage. Optical Coherence Tomography (OCT) provides high-resolution imaging of retinal structures, making it a valuable tool in ophthalmological diagnosis. This study presents a novel artificial intelligence (AI)-supported computer-aided diagnostic system for the real-time classification of retinal diseases using OCT images. The proposed system integrates a DenseNet-201 deep learning model with a hash-based data integrity mechanism and a user-friendly interface for clinical deployment. The DenseNet-201 model achieved superior performance with an accuracy of 94.42%, an F1-score of 0.9442, and an AUC of 1.00, outperforming other widely used models such as GoogleNet, ResNet50, and EfficientNetB0. Unlike existing systems, our approach includes automatic image validation, eliminates data redundancy through hashing, and is optimized for practical use via the Gradio interface. These features address major limitations in prior studies, such as a lack of real-time capability, data inconsistency, and insufficient clinical integration. The system not only improves diagnostic accuracy but also reduces clinician workload, ensuring faster and more reliable decision-making in the detection of retinal diseases. This work demonstrates the feasibility of deploying AI-powered diagnostic tools in real-world ophthalmic settings and lays the groundwork for future development of integrated, scalable healthcare solutions.

Index Terms—Retinal Diseases, Optical Coherence Tomography (OCT), Artificial Intelligence (AI), Deep Learning, Computer-Aided Diagnosis, Clinical Decision Support System, Real-Time Classification, Medical Image Analysis, DenseNet-201, GoogleNet, ResNet50, EfficientNetB0, Gradio Interface

I. INTRODUCTION

IN MANY parts of the world, especially in Turkey, changes in eating habits, a decline in physical and social activities, and excessive engagement with technology have not only reduced overall quality of life but also contributed to serious health problems [1]. One of the most significant issues is the partial or

total loss of vision. Additionally, with the rapid advancement of technology, screens have become deeply integrated into our daily lives, leading to prolonged screen exposure throughout the day. This prolonged exposure results in excessive eye fatigue and may contribute to future vision-related complications [2]. The primary objective of this study is to explore how the benefits of modern technology—particularly artificial intelligence—can be leveraged to support healthcare, rather than merely highlighting health warnings or precautions. Since the introduction of AI into everyday life, it has found applications across numerous fields, with a particularly significant impact in healthcare. Therefore, this study aims to provide a comprehensive and precise evaluation while also optimizing diagnostic efficiency and reducing time consumption.

Among all human functions, vision is undoubtedly one of the most critical, as it provides meaning to our interaction with the environment. With vision, the human environment gains meaning. In addition, thanks to the ability to see, muscle functions are fulfilled without restriction. We can explain this by the fact that a person who has lost their eyesight moves very slowly. Therefore, problems that may occur in eyesight will affect our quality of life significantly negatively. It will limit our activities considerably and will cause us to barely maintain our vital functions [3].

The human eye is one of the most complex and essential organs, responsible for facilitating communication with the external world through vision. It operates via a series of intricate biological processes and is composed of various protective and sensory structures. Because vision is highly sensitive, even minor damage or dysfunction in any of its anatomical components can lead to severe visual impairment. To prevent such ocular health issues, it is essential to understand the eye's anatomy and its associated functions. For this purpose, several key terms related to the diseases in question, which are the subject of this study, need to be examined simply [4].

The eye is the most complex sensory organ responsible for enabling vision and facilitating interaction with the external world. Its primary anatomical components include the cornea,

iris, lens, sclera (outer fibrous layer), choroid (vascular layer), retina, vitreous gel, and optic nerves. Briefly, the process of vision begins when light rays from an object enter the eye through the cornea, pass through the iris—which adjusts in size according to light intensity—then pass through the lens, and finally converge onto the retina at the back of the eye as an inverted image. These rays falling on the yellow spot on the

Hasan Memis, is with Department of Electrical Engineering of Batman University, Batman, Turkey, (e-mail: hasanmms21@gmail.com).

 <https://orcid.org/0009-0006-4366-4722>

Emrullah Acar, is with Department of Electrical Engineering of Batman University, Batman, Turkey, (e-mail: emrullah.acar@batman.edu.tr).

 <https://orcid.org/0000-0002-1897-9830>

Manuscript received June 5, 2025; accepted Jul 19, 2025.

DOI: [10.17694/bajece.1715185](https://doi.org/10.17694/bajece.1715185)

retina stimulate the photoreceptors in this layer and send optical signals to the brain to be evaluated, and as a result, a meaningful image is formed from these signals [5][6].

The retina is a highly differentiated neuroectodermal tissue located at the innermost and posterior part of the eyeball, where light is focused. It consists of two distinct regions: the central retina (macula) and the peripheral retina. Photoreceptor cell bodies are situated in the outer nuclear layer, while bipolar, horizontal, and amacrine cell bodies are located in the inner nuclear layer [7]. Signals generated by the photoreceptors are transmitted to the brain via the optic nerves. There are two main types of photoreceptors in the retina: rods and cones, which absorb light and enable sharp, color-rich vision. The most common retinal diseases include choroidal neovascularization (CNV), diabetic macular edema (DME), and drusen. CNV refers to the formation of abnormal new blood vessels beneath the retina. These vessels often leak fluid, leading to dark spots in the visual field, distorted lines, and sudden vision loss. It is most frequently observed in the advanced stages of age-related macular degeneration [8]. DME is a complication of diabetic retinopathy characterized by fluid accumulation in the macula, resulting in blurred vision. Since the macula is responsible for detailed central vision, swelling in this area can cause significant visual impairment [9][10]. Drusen are yellowish deposits that accumulate under the retinal pigment epithelium and are considered an early indicator of age-related macular degeneration. Their presence may cause symptoms such as central blurring, faded color perception, and scotomas in the visual field [11].

Optical coherence tomography (OCT) is an advanced imaging technique that enables high-resolution visualization of biological tissue layers. It generates cross-sectional images by calculating the intensity and time delay of infrared light waves emitted from the device and reflected from various tissue structures. The light beam from the source is split by a semi-permeable mirror called a beamsplitter; one part is directed to a reference mirror, while the other part is directed into the patient's eye. The light reflected from the retinal layers and the reference mirror is then combined in an interferometer. The resulting time delay is calculated and used to construct a detailed tomographic image [12].

Some of the studies using OCT images with artificial intelligence are as follows: Li et al. (2019) applied deep transfer learning with VGG-16 on 109,312 OCT images and achieved 98.6% accuracy [13]. Serener et al. (2019) classified dry AMD using ResNet with 99.5% accuracy [14]. Motozawa et al. (2019) used CNN and transfer learning to distinguish normal and AMD subtypes with 99% accuracy [15]. Sun et al. (2020) employed ResNet-50 with an attention mechanism and reached 98.17% accuracy on the Duke dataset [16]. Acar et al. (2021) developed an automatic cataract detection system using VGGNet and DenseNet on fundus images, achieving up to 97.94% accuracy and highlighting the potential of deep learning in teleophthalmology [17]. Shi et al. (2021) developed the Med-XAI-Net model for geographic atrophy detection and achieved 91.5% accuracy [18]. Rajagopalan et al. (2021) trained a CNN model on 12,000 OCT images and achieved 97.01% accuracy [19]. Sahin (2022) compared a proposed CNN model with MobileNet50 and reported 94% accuracy [20]. Elsharkawy et al. (2022) used ResNet-50 on the Kermany dataset and reached 96.21% accuracy [21]. He et al. (2022) achieved 99.78% accuracy on the UCSD dataset using ResNet-

50 and the LOF algorithm [22]. Aykat et al. (2023) developed a novel hybrid model combining EfficientNetV2S and Xception (EffXceptNet) for classifying CNV, DME, and drusen in OCT images, achieving 99.90% accuracy and outperforming conventional CNN architectures [23]. Baharlouei et al. (2023) applied a CNN model based on wavelet scattering transform and achieved 97.1% accuracy on the Heidelberg and Duke datasets [24]. Kulyabin et al. (2024) used VGG16 on the OCTDL dataset and reported 85.9% accuracy [25]. Gencer et al. (2025) developed a hybrid model integrating EfficientNetB0 and Xception with SE blocks, achieving 99.58% and 99.18% accuracy on the UCSD and Duke datasets, respectively [26].

Existing AI-based methods, although successful in controlled settings, often suffer from limitations such as a lack of real-time functionality, redundancy in data processing, high computational cost, and an insufficient user interface design for clinical deployment. Most prior models also do not incorporate effective mechanisms to prevent data duplication or ensure image type validation, which are essential for consistent clinical performance. The proposed method overcomes these limitations by combining a high-performing DenseNet-201 deep learning model with a hash-based data integrity mechanism and a streamlined user interface via Gradio, enabling clinicians to diagnose retinal diseases such as CNV, DME, and drusen with high precision (accuracy: **94.42%**, F1-score: **0.9442**, AUC: **1.00**) in real time. Furthermore, the model's ability to focus solely on OCT images, validate them automatically, and avoid saving duplicates enhances both accuracy and efficiency, making the system a robust and scalable solution for clinical environments.

A key distinction from previous studies in visual data processing is that the system assigns a unique hash value to each tomographic image. This prevents data repetition and redundancy, reduces unnecessary system load, and enables faster access to images. Moreover, restricting the system to accept only OCT (Optical Coherence Tomography) images improves the efficiency of the analysis process by ensuring the model focuses solely on relevant medical data. Storing patient information alongside the images enhances the integrity of medical records. The proposed model will utilize an optimized convolutional neural network (CNN) architecture that is capable of performing both diagnosis and segmentation tasks with a reduced number of deep layers. The model's performance will be evaluated comparatively using pre-trained architectures (GoogleNet, ResNet, EfficientNet, DenseNet) and expert assessments. With these components combined, the study aims to develop an AI-assisted retinal disease diagnosis system with high accuracy based on OCT images.

II. MATERIAL AND METHOD

a. Dataset Collection

The OCT-2017 dataset consists of retinal images obtained using a medical imaging technique called Optical Coherence Tomography OCT. The dataset is published on the Kaggle platform. The dataset contains images from four different retinal disease categories: DME (Diabetic Macular Edema), Drusen, CNV (Choroidal Neovascular Membrane), and Normal. In this framework:

- CNV: Choroidal Neovascular Membrane category includes 37,205 training images and 242 test images.
- DME: Diabetic Macular Edema category includes 11,350

training images and 242 test images.

- DRUSEN: There are 8,616 training images and 242 test images in the Drusen category.

- NORMAL: There are 26,318 training images and 242 test images in the Normal category [27].

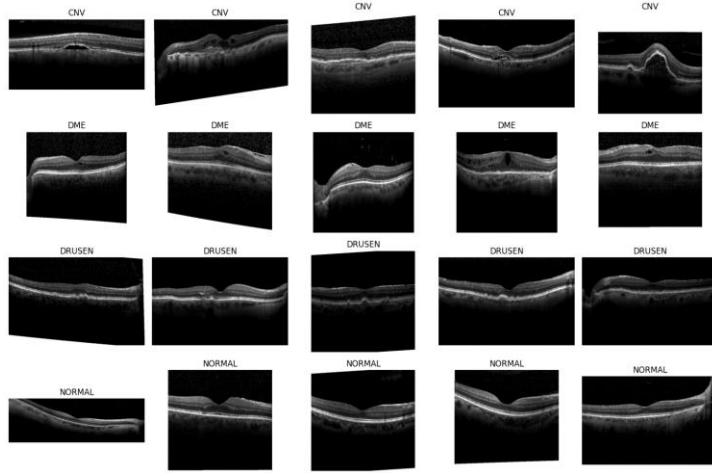


Fig 2.1. Examples from the OCT dataset [27]

b. Basic Working Principle

The basic working principle of the system describes the process of a client (doctor, specialized personnel or other service) submitting a decision support request via Gradio, as illustrated by a sequence diagram in Figure 2.1. The process starts with the client sending a request to Gradio. If the uploaded data is in image format, the request is forwarded to the Image Validator layer. If the image format is not supported, the uploaded image is not a colour or OCT image, and it provides feedback to the user with an unsupported image warning. If it is a supported image, the validator checks the accuracy of the OCT retinal image. Then, the AI model makes a prediction based on the input image. During the prediction process, the AI model makes a prediction using the data. Finally, the prediction result and the image are saved in the database and sent to the client via Gradio. However, if the image sent is not valid, i.e., has an unsupported format or structure, the validator responds to the client with an Unsupported Media Type (415) error. Similarly, if the incoming token is invalid, in this way, a doctor or other expert or service can send a request to the decision support system and receive an estimate. It is also worth mentioning that if the uploaded image has already been evaluated, it is re-evaluated, and the result is retrieved again. However, the image is not saved in the database again to avoid data redundancy, as it was previously saved with the given hash.

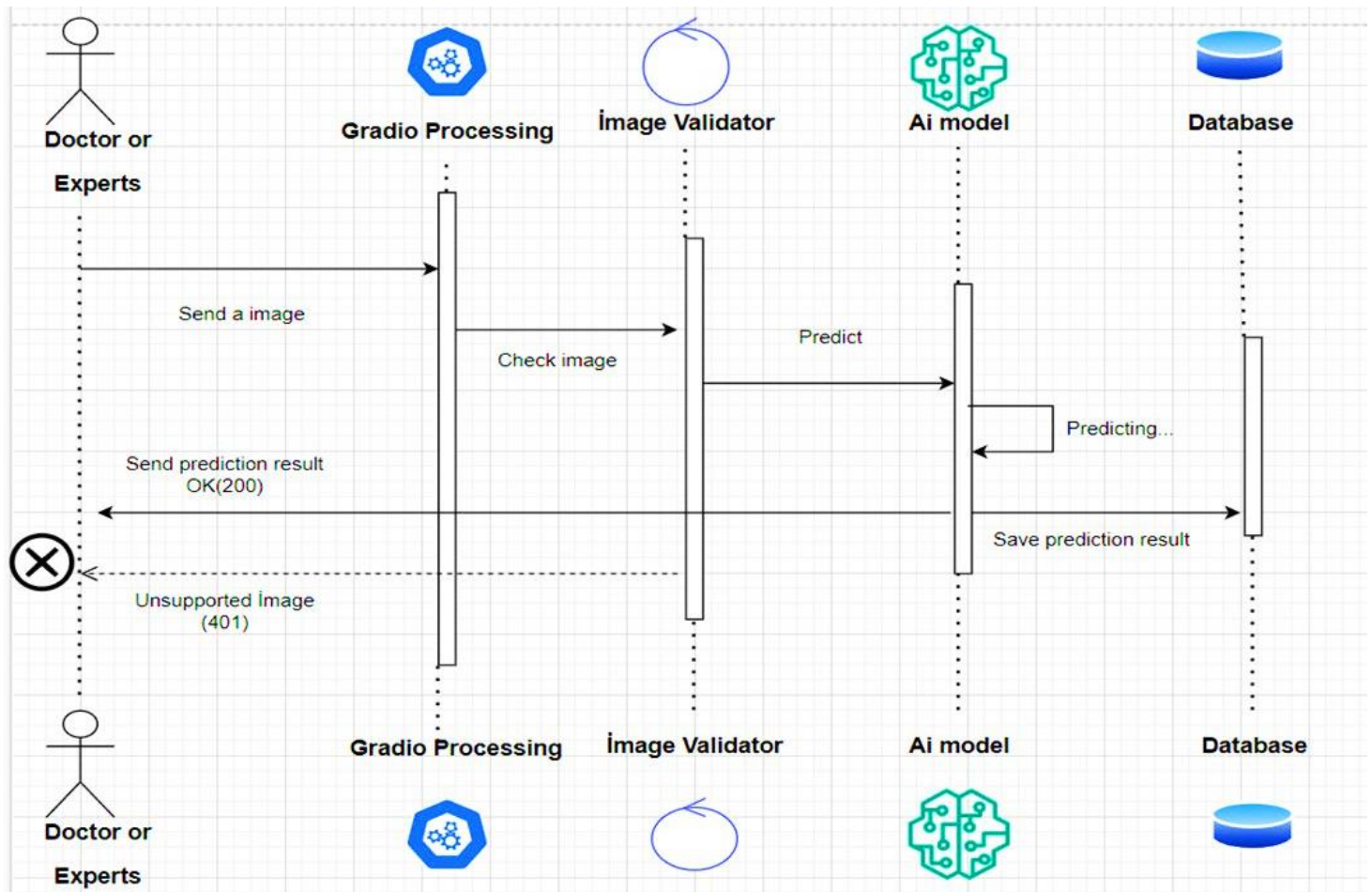


Fig. 2.2. System Basic Working Principle

c. Deep Learning

Artificial intelligence involves automating processes that require human reasoning and logic, such as mathematical computations, and human-specific capabilities, including event interpretation,

experience-based benefits, learning, and inference. It can be said that artificial intelligence, which is a modelling of the human brain, is an effort to bring human and superhuman abilities to machines. Machine learning is the self-learning of operations

based on the codes written by users, without waiting for commands from the user or without the need to write new code [28]. Deep learning is a subset of machine learning that consists of a large number of hidden layers, and the output of one hidden layer is the input of the next hidden layer. It is a subfield of machine learning based on artificial neural networks that tries to mimic the way the human brain works. It aims to create systems that can automatically learn complex tasks by training with large data sets [29].

1. Convolutional Neural Networks (CNN)

Convolutional Neural Networks (CNN) are a multilayer forward artificial neural network, which is one of the deep learning architectures that needs fewer parameters and less training in large network models. It is used in image processing analysis, object recognition, and natural language processing. At first, it was used for object recognition because it could not meet the requirements due to hardware deficiencies, but later, with the development of graphical processing units, it has been frequently preferred in areas such as image processing and recognition, sound processing, computer vision, face recognition, text processing and classification, and its popularity has increased. The crowding and complexity of data, along with the development of technology, have underscored the importance of convolutional neural networks and increased their use [30]. Convolutional neural networks perform better on complex data than fully connected layers, as they progress hierarchically by adding and superimposing what they have learnt in the previous layer. CNN architectures consist of input layer (Input Layer), convolution layer (Convolution Layer), flattened linear unit layer (RELU), pooling layer (Pooling Layer), fully connected layer (Fully Connected Layer), dropout layer (Dropout Layer) and classification layer (Classification Layer) [31]. The number of layers depends on the work of the person designing the model and the study. In convolutional neural networks, much less preprocessing is needed to create the filters to be trained compared to traditional techniques due to the automatic self-training of different features [32].

In this study, deep learning architectures such as GoogleNet, ResNet, EfficientNet, and DenseNet are used, and they are compared with each other, and the comparison results are briefly mentioned. As can be seen from the results, the most efficient result is obtained with DenseNet, so the graphs and evaluations are based on this model.

reducing computational and memory costs. In the 22-layer architecture, 1×1 convolution layers are used to manage the number of parameters and computational complexity [34].

b) ResNet

ResNet was developed to solve the vanishing gradient problem seen in deep networks and is differentiated by its residual block structure. This structure provides a deeper and more efficient learning by learning the input and output differences instead of direct input-output mapping. There are versions with different depths (ResNet18, ResNet34, ResNet50, ResNet101, ResNet152) [35].

c) EfficientNet

EfficientNet is an architecture that scales the depth, width and input resolution of the network in a balanced way. Convolutional layers, bottleneck layers and inverse residual blocks are used to increase efficiency. It aims to achieve high accuracy with fewer parameters than traditional CNN architectures [36].

d) DenseNet

The images were trained with the DenseNet201 artificial intelligence model for image classification and retinal disease detection. DenseNet-201 has an input size of $224 \times 224 \times 3$, where 224 is the width, 224 is the height, and 3 is the depth (number of color channels). DenseNet201 is a deep learning model where each layer receives features from all previous layers as input, thus improving information flow and gradient propagation [37]. This dense connection structure makes the model more efficient and provides high performance with fewer parameters. In the tests performed with the DenseNet201 model in this study, successful results were obtained. The main reason why the model is preferred is that it has more layers than other DenseNet models. Since it has 201 layers, the model can be trained in many more layers, and much more accurate results can be obtained. In classical convolutional neural network models, the layers are connected in such a way that the output of one layer depends on the input of the other layer, but in the DenseNet architecture, the outputs of all previous layers can be given as input for any layer, offering much more effective results in terms of performance and accuracy [38].

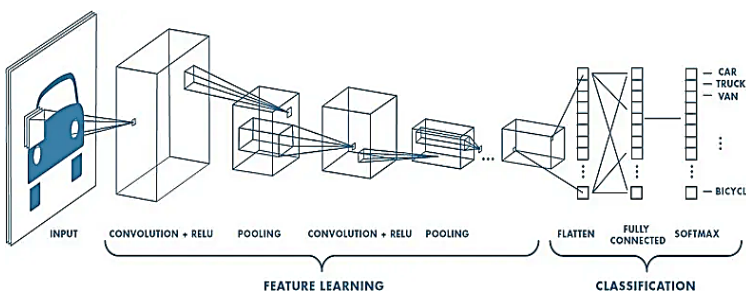


Fig. 2.3. DenseNet Architecture [33]

a) GoogleNet

Instead of stacking convolution and pooling layers directly on top of each other, GoogleNet is structured with Inception modules connected in parallel. This approach aims to increase learning capacity with non-linear activation functions while

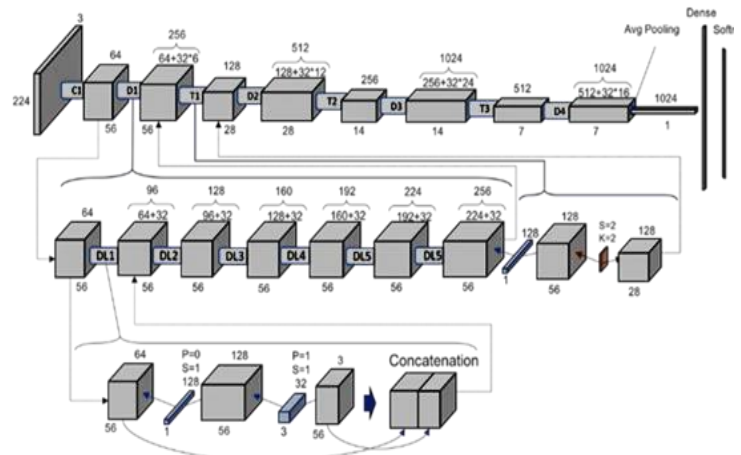


Fig. 2.4. DenseNet Architecture [39]

III. RESULTS AND DISCUSSION

During the model development studies, various deep learning models, including GoogleNet, ResNet (ResNet50), EfficientNet (EfficientNetB0), and DenseNet (Densenet201), were applied. As can be seen from the detailed comparison and analyses below, the most efficient results were obtained with the DenseNet-201 model. Therefore, the graphs and analyses in this study are mostly related to the DenseNet-201 model.

Train accuracy shows how accurately the model predicts the training data. In other words, it is the success of the model on the data it sees during learning. Validation accuracy indicates the model's success on unseen data (validation). This evaluates the generalization ability of the model. If train accuracy is high but validation accuracy is low, there may be overfitting. If both are low, there may be underfitting. If both are high and close to each other, the model has learnt well and can generalise. Based on this information, it is understood that the DenseNet-201 model gives by far the most successful result among the other models for the graph in Figure 3.1. The lowest accuracy is observed in EfficientNetB0 and GoogleNet models. In addition, since Densenet201, the model with the highest F1 score in the F1 score comparison graph in Figure 3.2, gives very balanced and successful results in terms of both precision and recall, this model stands out as the model that can discriminate between classes best. Since EfficientNetB0 has the lowest F1 score with a ratio of 0.83, this model seems to be less successful compared to the others. In this parameter and many similar parameters, it is clearly seen that the Densenet201 model shows the best result among the models tested in this study.

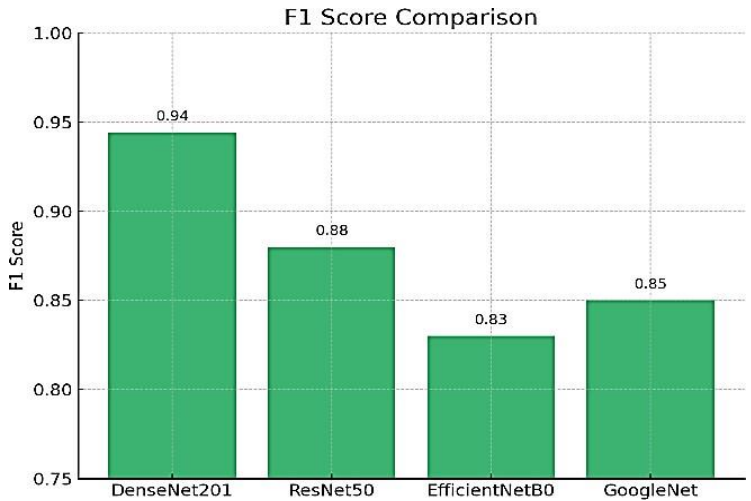


Fig. 3.2. F1 Score Comparison of Models

The two most important factors that determine the complexity of deep learning models are the number of layers and the number of parameters. A high number of layers allows the model to learn more complex patterns, which is especially advantageous in tasks such as image processing, where edges at the lower level, shapes at the middle level and objects at the upper level can be recognized. A higher number of parameters allows the model to learn more information; however, this requires more data, more memory and more computational power. In this context, the DenseNet201 model stands out among other models with 201 layers and approximately 20 million parameters. In comparison, GoogleNet has 22 layers and 7 million parameters, ResNet50 has 50 layers and 25.6 million parameters, and EfficientNetB0 has 82 layers and 5.3 million parameters. This technical evaluation shows that DenseNet201 is the best-performing model in this study. However, it is not correct to say that DenseNet201 is the best model in all cases; many factors, such as the size and complexity of the dataset, data imbalance, and hardware, affect the model's performance. Nevertheless, in this study, it was observed that the best results were obtained with DenseNet201. Therefore, the technical analyses, results and graphical presentations of the study were based on this model.

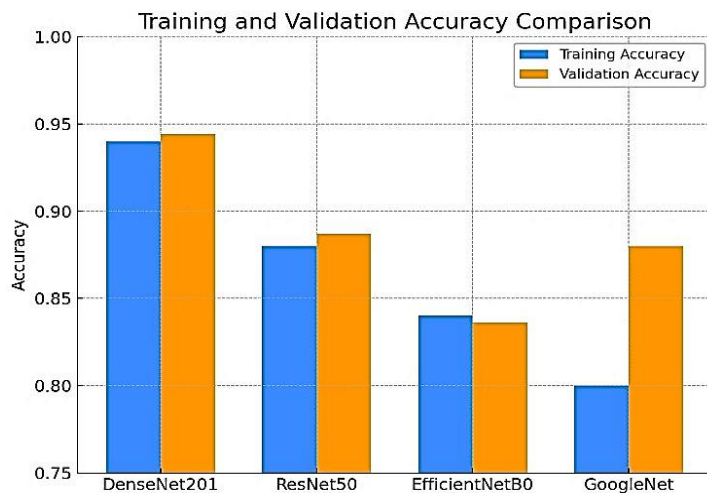


Fig. 3.1. Comparison of Train accuracy and Validation accuracy of the models

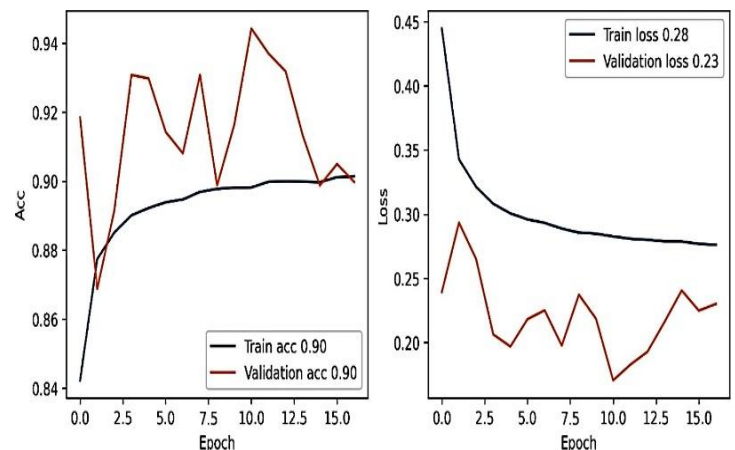


Fig. 3.3. Model accuracy and loss curves of DenseNet

Figure 3.3 shows the accuracy and loss curves obtained during the validation process of the model training for the DenseNet-201 model.

In this study, the training and validation loss curves of the DenseNet201 model, as shown in Figure 3.3, were analyzed to evaluate the model's learning behavior over time. The training loss curve exhibited a consistent downward trend, indicating that the model effectively learned from the training data and progressively improved its performance. In contrast, the validation loss curve showed more fluctuations, suggesting potential overfitting risk and a relatively unstable generalization performance on unseen data. To address this issue, an early stopping mechanism was employed. Although the training cycle was initially set for 128 epochs, the training process was automatically halted around the 15th epoch due to a lack of significant improvement in the validation loss. This decision was based on a predefined "patience" parameter, which monitors the model's performance on the validation set and stops the training if no further improvement is observed for a specific number of epochs. This technique effectively prevented overfitting, preserved the model's generalization capability, and ensured optimal training efficiency.

The observed trends in Figure 3.3 confirm this mechanism: while training accuracy continued to improve, validation accuracy plateaued and then stabilized. Similarly, validation loss ceased to decrease meaningfully, indicating that the model had reached a critical threshold in learning. By terminating training at this point, unnecessary training time and computational resources were avoided. Furthermore, early stopping contributed to maintaining the model's robustness, reliability, and clinical applicability—key considerations in a medical imaging context where unseen data performance is paramount.

Overall, the use of early stopping played a crucial role in balancing training performance with generalization ability, optimizing resource usage, and enhancing the clinical readiness of the proposed AI system.

In order to evaluate the classification performance, in other words, the success of the study, some metric values are used as shown in Table 3.1. The mathematical ratio of true positive (TP), true negative (TN), false positive (FP) and false negative (FN) to each other provides some metric results to show the success of the study. Knowing some metrics such as accuracy, sensitivity, recall, specificity, precision, false positive rate, and F1 score will contribute significantly to the judgment of the study's performance.

TABLE I
PERFORMANCE RESULTS OF DENSENET

Metrics	Values
Accuracy	0.9442
Sensitivity	0.9442
Specificity	0.9814
Precision	0.9475
False positive rate	0.0185
F1 score	0.9442

Table 3.1 shows the performance results and success measures of the classification model. Accuracy, or the rate at which the model makes correct predictions, is 94.42%. This indicates that the majority of instances are correctly classified. Sensitivity, or the ability of the model to correctly recognize true positives, is also 94.42%. It shows the percentage of actual positives that the model correctly predicts as positive. Specificity, which refers to the model's ability to correctly recognize true negatives, is 98.14%. It indicates the percentage of actual negatives that the model correctly predicts as negative. Precision is the ratio of the number of instances that the model predicts as positive to the number of positive cases, and it is 94.75%. It shows the proportion of samples classified as positive that are indeed positive. The false positive rate, which is the rate at which the model incorrectly classifies true negatives as positives, is 1.85%. It shows the percentage of negative samples that the model incorrectly identifies as positive. The F1 score, which balances sensitivity and precision, is 94.42%. It indicates how well the classification model balances recall and precision. Overall, these results demonstrate that the model performs well. High accuracy, sensitivity, specificity, and precision suggest that the model can accurately classify both positive and negative cases.

In this study, the classification performance of the proposed DenseNet201-based model was evaluated using Receiver Operating Characteristic (ROC) curves. ROC curves are widely used to assess the performance of classification models by illustrating the trade-off between the true positive rate (TPR) and the false positive rate (FPR) across various threshold settings. In Figure 3.4, ROC curves for each class (CNV, DME, DRUSEN, NORMAL) are presented on a single graph using the one-vs-rest approach, allowing a comprehensive and comparative analysis of the model's discriminative power across all classes.

The results indicate that all classes achieved an AUC (Area Under the Curve) value of 1.00, which suggests that the model distinguishes each class from the others with perfect accuracy.

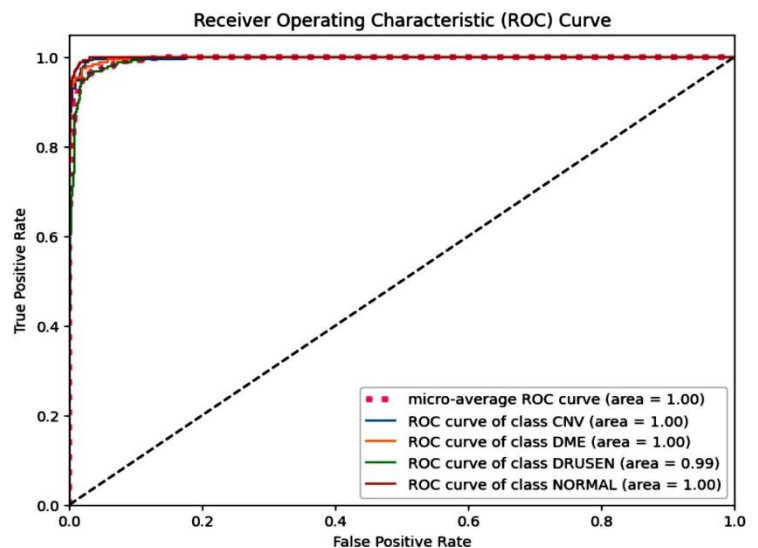


Fig. 3.4. ROC curve of the model

The ROC curves cluster closely around the upper-left corner of the graph, signifying both high sensitivity and a low false positive rate—hallmarks of exceptional classification performance.

In multi-class classification problems, visualizing ROC curves for all classes in a single graph is a common and effective practice in the academic literature. For instance, Rajagopalan et al. (2021) utilized this approach to evaluate the overall diagnostic strength of their deep learning models by comparing the ROC curves of each class collectively [19]. This strategy enables direct visual comparison of class-level performance, particularly beneficial for assessing critical clinical classes such as CNV and DME. Although presenting separate ROC curves for each class may offer visual clarity, it limits holistic comparison and may lead to misinterpretation due to inconsistent axis scales. Therefore, displaying all ROC curves within a single unified graph ensures scale consistency and facilitates direct, comparative interpretation of model performance across classes. This makes it a valuable method for evaluating the overall diagnostic reliability of the model, especially in clinical contexts.

IV. CONCLUSION

In this study, an artificial intelligence-based computer-aided diagnostic system was developed for the automated detection of retinal diseases using Optical Coherence Tomography (OCT) images. Various state-of-the-art deep learning models, including GoogleNet, ResNet, EfficientNet, and DenseNet, were implemented and comparatively evaluated. Among these models, DenseNet-201 achieved the highest classification performance with an accuracy of 94.42%, an F1 score of 94.42%, and an AUC of 1.00. These results indicate that DenseNet-201 is highly effective in distinguishing between different retinal disease categories and normal cases.

The system was designed not only to achieve high diagnostic accuracy but also to optimize data management through hashing mechanisms and to ensure usability through a user-friendly interface. Moreover, the system was developed with real-world clinical applicability in mind, aiming to support healthcare professionals in early and accurate diagnosis of retinal diseases. By reducing the workload of clinicians and accelerating the diagnostic process, the system has the potential to improve patient outcomes and contribute to more efficient clinical workflows.

The proposed system aims to address the shortcomings observed in previous studies in the literature. Although some studies report high accuracy rates, they often fall short in terms of ease of use, real-time processing, and clinical applicability. Key distinguishing features of this study include a hash-based data recording mechanism that enhances data security, acceptance of only OCT (Optical Coherence Tomography) images to ensure analysis efficiency, intuitive and training-free user interfaces, and integration potential with national health software such as e- Nabız. Another notable advantage is the ability of the system to be embedded into OCT devices via embedded software, enabling direct output generation from the device. Patient-specific OCT images are securely stored in the MongoDB database with a unique hash value assigned at the time of upload. If the same image is uploaded again, only the prediction result is displayed without saving a duplicate, thereby

reducing data redundancy and improving system performance. The Gradio-based interface is designed to reduce the workload of healthcare professionals and save time. Furthermore, the developed patient interface allows individuals to communicate directly with their doctors via email.

Despite these strengths, the system has several practical limitations. If the system is to be installed individually on the computers located in relevant departments of the Ministry of Health, the process may lead to excessive time consumption and face limitations due to insufficient hardware capabilities, as not every institutional computer may support the operational requirements of the model. Alternatively, deploying it via centralized servers could increase operational costs. Additionally, OCT devices may not be available in every healthcare unit, limiting their widespread use. Most importantly, since patients' medical images are classified as sensitive personal health data, using such images for analysis requires special authorization from the Ministry of Health. This administrative requirement is often a lengthy and complex process, posing a significant barrier to real-world implementation. Therefore, while the current version of the system cannot yet fully achieve large-scale clinical integration, it lays a strong conceptual foundation for future development and adoption.

Looking ahead, several directions for future improvement are evident. Expanding the dataset with more diverse and real-world clinical images will be crucial to enhancing the model's generalization capability. Integration of hybrid approaches, such as combining deep learning with traditional image processing techniques or clinical metadata, may further boost diagnostic performance. Efforts should also be made to optimize the system for real-time performance and to ensure seamless integration with existing healthcare platforms and electronic medical record systems. Collaborations with hospitals and healthcare providers will be essential to test and refine the system in actual clinical practice. Ultimately, embedding the model directly into OCT devices or national healthcare platforms could significantly increase its accessibility and impact. Furthermore, more effective use and widespread access may be possible by integrating the model directly into OCT devices or health system software and the e-Nabız platform. However, adequate technical infrastructure, financial support and official permissions are required to realize these goals.

In conclusion, this study demonstrates the feasibility and potential of an AI-supported OCT image analysis system for retinal disease detection. While further work is needed to fully realize its clinical utility, the results obtained provide a strong foundation for future advancements and broader adoption in ophthalmological practice.

V. REFERENCES

- [1] World Health Organization. (2019). World report on vision. Geneva: WHO.
- [2] Resnikoff, S., Lansingh, V. C., Washburn, L., Felch, W. C., & Gauthier, T. M. (2020). Vision loss and its impact on quality of life. *Ophthalmic Epidemiology*, 27(2), 85–90.
- [3] Lamoureux, E. L., & Fenwick, E. K. (2016). Health-related quality of life and visual impairment. *Current Opinion in Ophthalmology*, 27(3), 238–243.
- [4] Forrester, J. V., Dick, A. D., McMenamin, P. G., Roberts, F., & Pearlman, E. (2015). *The Eye: Basic Sciences in Practice*. Elsevier

- Health Sciences.
- [5] Berger, John, Ways of Seeing, Penguin Books, UK 2008, p.7-33.
 - [6] Kolb, H. (2005). Simple Anatomy of the Retina. Webvision.
 - [7] Curcio, C.A., Sloan, K.R., Kalina, R.E., Hendrickson, A.E. (1990) Human photoreceptor topography. The Journal of Comparative Neurology, 292 (4), 497-523.
 - [8] Dandekar SS, Jenkins SA, Peto T, et al.: Autofluorescence imaging of choroidal neovascularization due to age-related macular degeneration. Arch Ophthalmol. 2005;123:1507-1513.
 - [9] Bhende, M, Shetty, S, Parthasarathy, M. K., & Ramya, S. (2018). Optical coherence tomography: A guide to interpretation of common macular diseases. Indian journal of ophthalmology, 66(1), 20.
 - [10] Erdoğan, Alper, Epiretinal Membran Cerrahisinde Prognozu Etkileyen Faktörler, Uzmanlık tezi, s.8.
 - [11] Deutman AF, Hansen LMAA: Dominantly inherited drusen of Bruch's membrane. Br J Ophthalmol 1970; 34:373-382.
 - [12] Aydın, Ali, Bilge, A. Hamdi, Optik Koherens Tomografinin Glukomda Yeri, s.78.
 - [13] Li, F., Chen, H., Liu, Z., Zhang, X.-D., Jiang, M.-S., Wu, Z.-Z., and Zhou, K.-Q. (2019). Deep learning-based automated detection of retinal diseases using optical coherence tomography images. Biomedical Optics Express, 10(12), 6204–6226. <https://doi.org/10.1364/BOE.10.006204>
 - [14] Serener, A., and Serte, S. (2019). Dry and wet age-related macular degeneration classification using OCT images and deep learning. 2019 Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science (EBBT), 1–5. <https://ieeexplore.ieee.org/document/8741768>
 - [15] Motozawa, N., An, G., Takagi, S., Kitahata, S., Mandai, M., Hiram, Y., Yokota, H., Akiba, M., Tsujikawa, A., Takahashi, M., & Kurimoto, Y. (2019). Optical coherence tomography- based deep-learning models for classifying normal and age- related macular degeneration and exudative and non-exudative age-related macular degeneration changes. Ophthalmology and Therapy, 8(4), 527–539. <https://doi.org/10.1007/s40123-019- 00207-y>
 - [16] Sun, W., Zhang, H., & Yao, Z. (2020). Automatic diagnosis of macular diseases from OCT volume based on its two- dimensional feature map and convolutional neural network with attention mechanism. Journal of Biomedical Optics, 25(9), 096004. <https://doi.org/10.1117/1.JBO.25.9.096004>
 - [17] Acar, E., Türk, Ö., Ertuğrul, Ö. F., and Aldemir, E. (2021). Employing deep learning architectures for image-based automatic cataract diagnosis. Turkish Journal of Electrical Engineering & Computer Sciences, 29(5), 2649–2662. <https://doi.org/10.3906/elk-2103-77>
 - [18] Shi, X., Keenan, T. D. L., Chen, Q., De Silva, T., Thavikulwat, A. T., Broadhead, G., Bhandari, S., Cukras, C., Chew, E. Y., and Lu, Z. (2021). Improving interpretability in machine diagnosis: Detection of geographic atrophy in OCT scans. Ophthalmology Science, 1(3), 100038. <https://doi.org/10.1016/j.xops.2021.100038>
 - [19] Rajagopalan, N., Venkateswaran, N., Josephraj, A. N., and Srithaladevi, E. (2021). Diagnosis of retinal disorders from Optical Coherence Tomography images using CNN. PLoS ONE, 16(7), e0254180. <https://doi.org/10.1371/journal.pone.0254180>
 - [20] Şahin, M. E. (2022). A deep learning-based technique for diagnosing retinal disease by using optical coherence tomography (OCT) images. Turkish Journal of Science & Technology, 17(2), 417–426. <https://doi.org/10.55525/tjst.1128395>
 - [21] Elsharkawy, M., Sharafelddeen, A., Soliman, A., Khalifa, F., Ghazal, M., El-Daydamony, E., Atwan, A., Sandhu, H. S., & El- Baz, A. (2022). A novel computer-aided diagnostic system for early detection of diabetic retinopathy using 3D-OCT higher- order spatial appearance model. Diagnostics, 12(2), 532. <https://doi.org/10.3390/diagnostics12020532>
 - [22] He, T., Zhou, Q., and Zou, Y. (2022). Automatic detection of age-related macular degeneration based on deep learning and local outlier factor algorithm. Diagnostics, 12(2), 532. <https://doi.org/10.3390/diagnostics12020532>
 - [23] Aykat, Ş., & Şenar, S. (2023). Advanced detection of retinal diseases via novel hybrid deep learning approach. Traitement du Signal, 40(6), 2367–2382. <https://doi.org/10.18280/ts.400604>
 - [24] Baharlouei, Z., Rabbani, H., & Plonka, G. (2023). Wavelet scattering transform application in classification of retinal abnormalities using OCT images. Scientific Reports, 13, 19013. <https://doi.org/10.1038/s41598-023-46200-1>
 - [25] Kulyabin, M., Zhdanov, A., Nikiforova, A., Stepichev, A., Kuznetsova, A., Ronkin, M., Borisov, V., Bogachev, A., Korotkich, S., Constable, P. A., & Maier, A. (2024). OCTDL: Optical coherence tomography dataset for image-based deep learning methods. Scientific Data, 11(365). <https://doi.org/10.1038/s41597-024-03182-7>
 - [26] Gencer, G., & Gencer, K. (2025). Advanced retinal disease detection from OCT images using a hybrid squeeze and excitation enhanced model. PLOS ONE, 20(2), e0318657. <https://doi.org/10.1371/journal.pone.0318657>
 - [27] Mooney, P. T. (2018). Kermany 2018 OCT Dataset. <https://www.kaggle.com/datasets/paultimothymooney/kermany2018>
 - [28] C. Cortes, X. Gonzalvo, V. Kuznetsov, M. Mohri, and S. Yang. Adanet: Adaptive structural learning of artificial neural networks. arXiv preprint arXiv:1607.01097, 2016.
 - [29] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In CVPR, 2016
 - [30] Urmamen H E. “Optik Koherans Tomografi Görüntüleri İle Retinal Hastalıkların Evrişimsel Sinir Ağı Kullanılarak Teşhis Edilmesi“. Yüksek Lisans Tez Çalışması, 2023, s.22
 - [31] Long, E. Shelhamer, and T. Darrell. Fully convolutional networks for semantic segmentation. In CVPR, 2015
 - [32] Özdemir E, Türkoğlu İ. Yazılım Güvenlik Açıklarının Evrişimsel Sinir Ağları (CNN) ile Sınıflandırılması, 2022
 - [33] He, X. ve Dong, F. (2023). A deep learning-based mathematical modeling strategy for classifying musical genres in musical industry. Nonlinear Engineering, 12, 20220302. <https://doi.org/10.1515/nleng-2022-0302>
 - [34] Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V., & Rabinovich, A. (2015). Going deeper with convolutions. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 1–9.
 - [35] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 770–778.
 - [36] Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking model scaling for convolutional neural networks. In Proceedings of the 36th International Conference on Machine Learning (ICML), PMLR, 6105–6114.
 - [37] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2017). *Densely Connected Convolutional Networks*. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 4700-4708.
 - [38] Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., & Chen, T. (2018). Recent Advances in Convolutional Neural Networks. Pattern Recognition, 77, 354-377
 - [39] Gajera, H. ve Trada, P. (2020). An AI Approach for Disease Detection in Plants. International Research Journal of Engineering and Technology (IRJET), 7(7), 5287–5290.

VI. BIOGRAPHIES



HASAN MEMİŞ, born in Diyarbakır in 1988, completed his high school education in Diyarbakır. In 2011, he was admitted to the Electrical and Electronics Engineering Department at Gaziantep University and graduated in 2017. In 2023, he started to study his master's degree programme in Electrical

and Electronics Engineering at Batman University. He researches topics such as artificial intelligence and image processing. He has also been teaching as an electrical and electronic engineering teacher at the Ministry of National Education in Şırnak since 2019.



EMRULLAH ACAR received the B.S. degree in electrical and electronics engineering from Çukurova University, Adana, Turkey, in 2009, the M.S. degree in biomedical engineering from Istanbul Technical University, Istanbul, Turkey, in 2010, and the M.S. and Ph.D. degrees in electrical–electronics engineering

from Dicle University, Diyarbakır, Turkey, in 2012 and 2017 respectively. He is currently an Associate Professor of electrical–electronics engineering with Batman University and also the head of the electronics division. His research interests include digital image processing, machine learning, and remote sensing applications. Dr. Acar received awards and honors include The Scientific and Technological Research Council of Turkey (TUBITAK) Grant, Erasmus Mobility Grant, Sweden, Internship Grant, Germany and Visiting Researcher, Czech Republic.