Journal of Artificial Intelligence and Engineering Applications

Website: https://ioinformatic.org/

June 2026. Vol. 5. No. 3; e-ISSN: 2808-4519

Eye Disease Classification System Based on Fundus Images Using the InceptionV3 Architecture

Annisa Aulia[1]*, Hermawan Syahputra[2], Yulita Molliq Rangkuti[3], Insan Taufik[4], Kana Saputra S[5]

[1,2,3,4,5] Department of Computer Science, Faculty of Mathematics and Natural Sciences, Universitas Negeri Medan, Indonesia
annisaauliaaa@mhs.unimed.ac.id[1], hsyahputra@unimed.ac.id[2]*
yulitamolliq@unimed.ac.id[3], insantaufik@unimed.ac.id[4], kanasaputras@unimed.ac.id[5]

Abstract

This study aims to develop an automated eye disease classification system based on retinal fundus images using the InceptionV3 deep learning architecture. The dataset consists of four classes: cataract, diabetic retinopathy, glaucoma, and normal, collected from public sources and clinical data. The proposed method applies several preprocessing techniques, including background segmentation, data augmentation, data normalization, and an 80:20 data split to improve model performance and generalization. Transfer learning is implemented by utilizing pretrained ImageNet weights and modifying the final layers to suit the classification task. The model is trained using the Adam optimizer with a learning rate of 0.001 and categorical cross-entropy loss function. Evaluation results show that the model achieves an accuracy of 96%, with average precision, recall, and F1-score values of 0.97, 0.96, and 0.97, respectively. The confusion matrix analysis indicates that most predictions are correctly classified, demonstrating strong performance across all classes. Furthermore, the model is successfully integrated into a web-based system that enables users to upload fundus images and obtain classification results automatically. These findings indicate that the proposed system can effectively assist in early detection of eye diseases and support clinical decision-making.

***Keywords**: Cataract; Deep Learning; Diabetic Retinopathy; Fundus Image; Glaucoma*

1. Introduction

All humans possess five senses, known as the five senses; one of these is the sense of sight, or the eyes, which function to receive visual images and transmit them to the brain for processing. The sense of sight is the sense most frequently used in daily life, such as at work, in school, and so on [1] [2]. The eye is one of the most complex organs in the human body, relied upon to see and understand the surrounding world. However, certain eye diseases can develop and lead to the loss of vision, making it crucial to identify and treat these conditions [3].

Indonesia ranks third globally in the prevalence of blindness, with a rate of 1.47% [4]. According to the Rapid Assessment of Avoidable Blindness survey, approximately 3% of adults over the age of 50 in Indonesia are blind, amounting to roughly 1.6 million people. Based on data from the Ministry of Health of the Republic of Indonesia, priority vision disorders include cataracts, glaucoma, refractive errors, and diabetic retinopathy. However, this study focuses on three diseases: cataracts, glaucoma, and diabetic retinopathy [5].

Cataracts, which are the leading cause of blindness, can be treated through cataract surgery. For cataract patients who undergo surgery and experience no other complications, the chances of regaining vision are very high [6]. Similar to cataracts, glaucoma also requires immediate treatment to prevent the risk of blindness. Glaucoma is an eye disease that attacks the optic nerve and can cause permanent blindness [7]. Additionally, there is another eye condition known as diabetic retinopathy, which is often a manifestation of systemic diseases such as diabetes. Diabetic retinopathy is estimated to account for 5% of blindness cases worldwide [8].

To diagnose these conditions, an examination of the retina using ophthalmoscopy is required; this test allows a medical professional to see the back of the eye more clearly and then diagnose the specific condition the patient is suffering from. Ophthalmoscopy is also known as funduscopy, and the result of this test is a fundus image [9].

Retinal fundus images often vary in quality due to noise, uneven lighting, and low contrast, resulting in unclear visual information that can hinder the image interpretation process during diagnosis. This issue poses a significant challenge in the diagnosis of eye diseases[10]. Furthermore, medical examinations to detect eye diseases are relatively time-consuming because they are performed manually by doctors who examine fundus images of the patient's retina; however, these images often fail to provide clear information [11].

In this context, one effective approach is the use of deep learning architectures, such as InceptionV3. This architecture has proven capable of analyzing images with high accuracy, making it an effective tool for supporting medical diagnosis. The use of deep learning algorithms like InceptionV3 has been successfully applied by many researchers for image classification [12].

2. Literature Review

2.1. Eyes and Eye Diseases

As one of the five primary senses, the eyes play a crucial role in supporting human activities and survival. The eyes are the organs of vision that detect light; they can only distinguish between light and dark. The retina, the part of the eye most sensitive to light, contains receptors that transmit the images received from light. This is what enables people to see. If the eyes are affected by disorders or diseases, daily activities can be disrupted. There are many vision disorders, ranging from the mildest to the most severe [13].

2.1.1. Cataract

A condition in which the normally clear and transparent lens of the eye becomes cloudy is known as a cataract. As a result, vision becomes blurry or vision is completely lost [14]. Blurred vision caused by aging is a sign of cataracts. Cloudiness in the eye's lens reduces visual function. Clumps of protein in the eye cause cloudiness [15]

2.1.2. Diabetic Retinopathy

One of the complications of diabetes is diabetic retinopathy, which can lead to vision loss or even blindness [16].

2.1.3. Glaucoma

Increased intraocular pressure (IOP), optic nerve head atrophy, and narrowing of the visual field are signs of an eye condition known as glaucoma. A person with glaucoma may notice that their pupils appear bluish-green. This eye condition is caused by an increase in the amount of aqueous humor produced by the ciliary body and a decrease in the amount of fluid drained from the area at the angle of the eye or the pupil [17].

2.2. Machine Learning

In machine learning, rather than being explicitly programmed to perform specific tasks, systems are trained using data so that they can recognize patterns and make decisions based on new data they receive. Machine learning is a process in which computers use data to improve their performance on specific tasks. This process involves the use of algorithms that can recognize patterns in the data and then make predictions or decisions based on those patterns [18].

2.3. Deep Learning

Deep learning involves the use of artificial neural network algorithms—or artificial neural networks—that are capable of extracting and learning representations from large amounts of unstructured data [19].

2.4. Convolutional Neural Network

CNNs have the ability to learn abstract representations of object features, particularly in spatial data, and to perform identification processes more efficiently. This capability is supported by the characteristics of CNNs in modeling data through nonlinear transformations and the use of nonlinear activation functions. CNNs are capable of learning abstract representations of object features, particularly in spatial data, and identifying them more efficiently. This is made possible by their ability to represent data through nonlinear transformations and the application of nonlinear activation functions [20].

2.5. InceptionV3

InceptionV3 consists of five initial convolutional layers (stem) with valid padding, namely from conv2d_0 to conv2d_4, where each convolution operation is always followed by a ReLU activation and BatchNormalization. After that, the architecture continues with eleven Inception modules, ranging from mixed0 to mixed10, which use same padding for each convolution. These eleven modules are designed using various forms of convolution factorization, including 1×1, 3×3, 1×3, 3×1, 5×5, 1×7, and 7×7 [21].

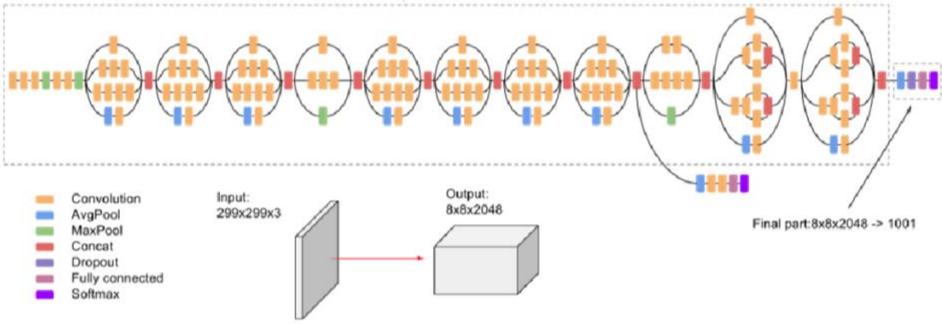

Fig. 1: Architecture Diagram of InceptionV3

2.6. Confusion Matrix

A confusion matrix is a process used to evaluate the accuracy of a classification model designed to identify data across different classes. Using a confusion matrix, we can calculate metrics such as accuracy, precision, recall, and F1-score, which provide a more detailed picture of the model's performance [22]. The formulas for accuracy, precision, recall, and F1-score are shown in the equations below:

$$Accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (1)$$

$$Precision = \frac{TP}{(TP+FP)} \quad (2)$$

$$Recall = \frac{TP}{(TP+FN)} \quad (3)$$

$$F1 - Score = 2 \times \frac{Recall \times Precission}{Recall+Precission} \quad (4)$$

3. Research Methodology

The research workflow can be observed in the flowchart presented below.

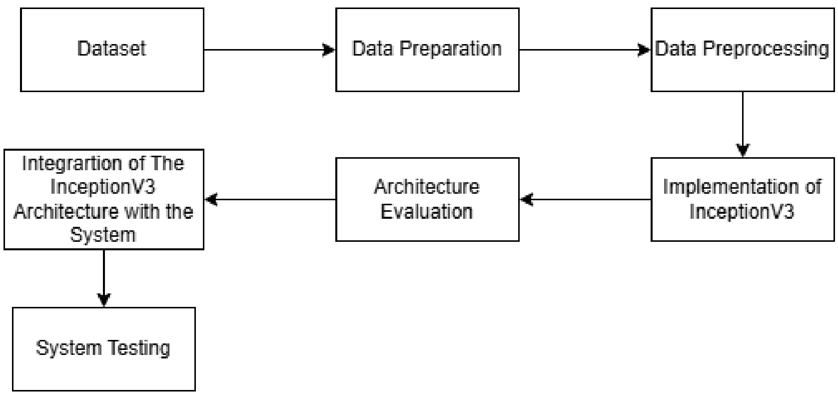

Fig. 2: Reasearch Workflow

3.1. Dataset

The dataset used in this study consists of retinal fundus images categorized into four classes: cataract, diabetic retinopathy, glaucoma, and normal. The dataset was obtained from public sources and served as the primary data for training and testing the eye disease classification model. Additional data was also obtained from the SPEC Eye Clinic in Medan, which was used for testing on the website.

3.2. Data Preparation

The data preparation phase is conducted to ensure that the dataset used can be processed effectively in the subsequent stages. This process involves inspecting the dataset, organizing the data into folders by class, and checking the quality of the images to be used in the study.

3.3. Data Preprocessing

The preprocessing steps performed in this study include background segmentation, data splitting, data augmentation, and data normalization. By performing these preprocessing steps, it is hoped that the model will be able to learn retinal image patterns more effectively, thereby improving the performance of eye disease classification.

3.4. Implementation of the InceptionV3

The retinal images, after undergoing preprocessing, are fed into the model and processed through several convolutional layers to generate a more complex feature representation. At the end of the model, several dense layers with a softmax activation function are added to generate classification probabilities for each eye disease class.

3.5. Architecture Evaluation

The evaluation was conducted using test data to measure the model's ability to correctly predict image classes. The evaluation metrics used in this study include accuracy, precision, recall, and F1-score, as well as an analysis using a confusion matrix to examine the distribution of the model's predictions across each class.

3.6. Integration of the InceptionV3 Architecture with the System

After the model demonstrated good performance, it was integrated into a web-based system. The integration was carried out using the Flask framework so that users could use the model directly to classify retinal images.

3.7. System Testing

In this phase, system testing is conducted on the system that has been designed and implemented using a method known as black-box testing.

4. Results and Discussion

4.1. Dataset

Dataset yang digunakan pada penelitian ini berasal dari dua sumber, yakni dataset pertama dikumpulkan dari dataset publik Kaggle (www.kaggle.com) yang berisi citra fundus dengan empat kelas penyakit mata, yaitu Cataract, Diabetic Retinopathy, Glaucoma, Normal yang digunakan sebagai data training dan data testing. Dataset kedua berasal dari Klinik Mata SPEC dengan jumlah 8 data fundus mata, yang akan digunakan sebagai data uji untuk pengujian pada sistem. Jumlah data citra fundus pada setiap kelas terdiri dari 1038 citra untuk kelas cataract, 1098 citra untuk kelas diabetic retinopathy, 1007 citra untuk kelas glaucoma, dan 1074 citra untuk kelas normal.

4.2. Data Preprocessing

Data preprocessing is performed to prepare the retinal images before they are used for model training. The steps applied in this study include background segmentation, data splitting, data augmentation, and data normalization.

First, background segmentation—or background removal—was performed to eliminate areas outside the retina, ensuring that the processed images focus more on the eye's primary structures. This step helps reduce visual distractions that could affect the model's learning process, such as background variations or uneven image edges. Next, the data in this study was divided into two main parts: training data and validation data. This study uses an 80:20 split. Third, data augmentation is performed; the data augmentation techniques in this study are applied to enrich the variation in the training dataset to improve the model's generalization ability and minimize the risk of overfitting. This data augmentation is implemented dynamically or on-the-fly. As explained by Al-Fahrezi [23], The on-the-fly method means that the augmented images are not physically saved but are generated dynamically during each training epoch. The dataset consists of 3,374 original images used for training, with a batch size of 32, resulting in 106 batches per epoch. The augmentation process is applied to increase data variation, including horizontal flipping, small rotations (±10%), and image scaling. Additionally, adjustments to contrast and brightness, position shifts, and the addition of Gaussian noise are performed to enhance the model's robustness against variations in image conditions. Finally, data normalization: in this study, image pixel values were normalized from the range of 0–255 to -1 to 1 to ensure a more balanced data distribution centered around zero. This method is an extension of an empirical approach involving dividing the pixel values and transforming them according to the equation used [24]This normalization helps deep learning models understand color differences more consistently and proportionally.

4.3. Implementation of InceptionV3

In this study, a retinal image classification model was built using the InceptionV3 architecture with a transfer learning approach. The model utilized initial weights from the ImageNet dataset and was modified at the output layer by adding Global Average Pooling, a Dense layer, and Dropout to adapt it to the four-class classification task.

The training parameters used include a learning rate of 0.001, the Adam optimizer, and the categorical cross-entropy loss function. The training process was conducted for up to 100 epochs with the application of early stopping, which automatically halted training at the 50th epoch when validation performance no longer improved. The use of these parameters aims to maintain a balance between convergence speed and model stability during the training process.

The model training results are visualized through accuracy and loss graphs shown in Figure 2. These graphs demonstrate that training and validation accuracy consistently increased from the beginning to the middle of the epochs before eventually reaching a stable state. On the other hand, the loss values showed a significant downward trend, particularly during the early training phase, and then leveled off as the model began to reach convergence.

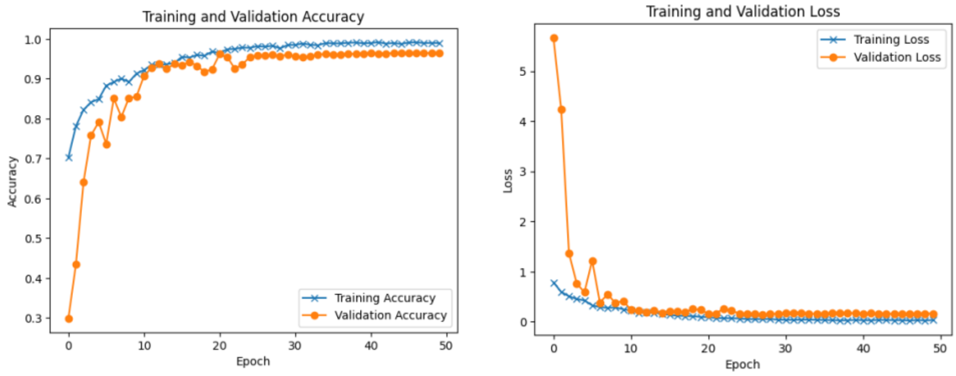

Fig. 3: Accuracy and Loss Curves

Several key findings can be drawn from Fig 3. First, the validation accuracy curve, which follows the pattern of the training curve, indicates that the model has good generalization capabilities on unseen data. Second, the absence of a significant difference between training and validation accuracy suggests that the model is not significantly overfitted. Third, the use of early stopping proved effective in halting training when the model's performance began to plateau, thereby preventing a decline in performance due to overtraining.

4.4. Model Evaluation

An architectural evaluation was conducted to assess the performance of the InceptionV3 model in classifying fundus images into four categories: cataract, diabetic retinopathy, glaucoma, and normal. Performance was measured using the metrics accuracy, precision, recall, and F1-score, along with a visual analysis using a confusion matrix to examine the distribution of the model's predictions across each category.

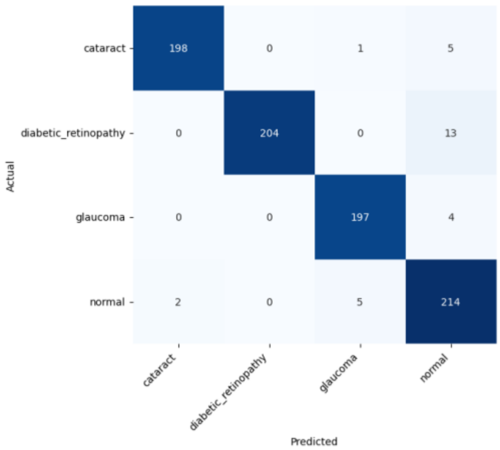

Fig. 4: Confusion Matrix

Based on the confusion matrix in Fig 4, it can be seen that most of the data falls along the main diagonal, indicating that the model is able to correctly classify the data into their respective classes. This suggests that the InceptionV3 architecture is effective at recognizing the distinctive patterns of each fundus image category. The normal and glaucoma classes show very high prediction rates with minimal errors, while the cataract and diabetic retinopathy classes still exhibit some classification errors, though the number is not significant.

Classification errors typically occur due to similarities in visual characteristics between classes, such as the presence of certain patterns in the images that resemble those of other eye conditions. Nevertheless, the relatively low error rate indicates that the model has a fairly good ability to distinguish key features in retinal images. The accuracy, precision, recall, and F1-score values for all classes are presented in Table 1 below:

Table 1: InceptionV3 Architecture Evaluation Results

Class	Precision	Recall	F1-Score
Cataract	0,99	0,97	0,98
Diabetic_retinopathy	1,00	0,94	0,97
Glaucoma	0,97	0,98	0,98
Normal	0,91	0,97	0,94
Rata-Rata	0,97	0,96	0,97
Akurasi		0,96 (96%)	

Thus, it can be concluded that the InceptionV3 architecture used in this study is capable of delivering optimal performance and demonstrates good generalization capabilities on new data.

4.5. Integration of the InceptionV3 Architecture with the System

The trained InceptionV3 model was then integrated into a web-based system using the Flask framework. The best model was saved in .keras format and loaded into the system to perform automatic fundus image classification.

The system accepts user-uploaded images as input, which are then preprocessed by resizing them to 256×256 pixels and normalizing them as required by the model. Next, the images are processed by the model to generate class predictions along with their probability values.

1. Home

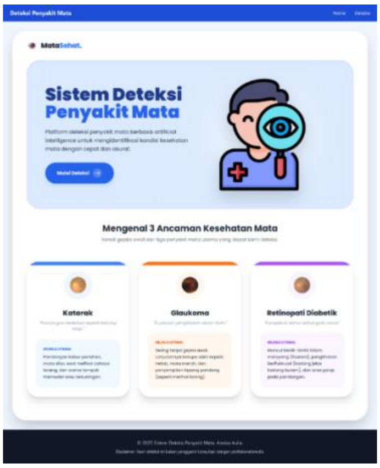

Fig. 5: Home Page

The home page serves as the initial interface that allows users to easily access the fundus image classification feature.

2. Classification Page

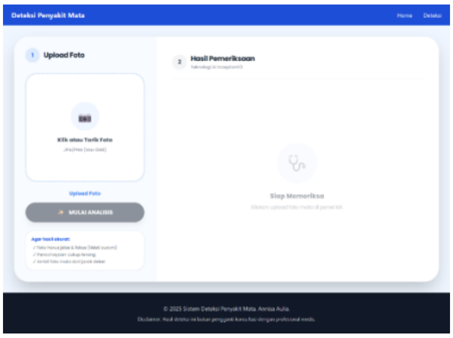

Fig. 6: Classification Page

On the classification page, users can upload images and view the prediction results, which are displayed as disease classes along with their confidence levels.

3. Classification Results Page

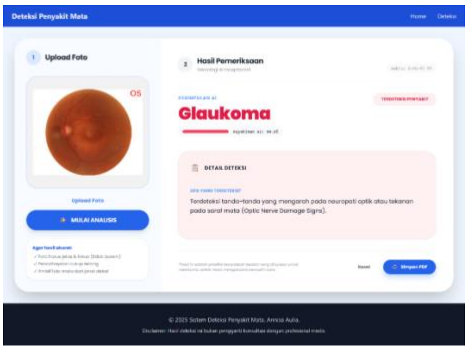

Fig. 7: Classification Results Page

The detection results page displays the classification output in the form of the disease categories predicted by the model—namely, cataract, diabetic retinopathy, glaucoma, or normal—along with the probability values (confidence scores).

4.6. System Testing

At this stage, the system will undergo testing using the black-box testing method. Black-box testing is a software testing method aimed at testing the system's functionality without requiring knowledge of the system's code structure. There are several types of black-box testing, including functional testing, non-functional testing, and regression testing. In this study, the author uses functional testing, in which tests are conducted to verify every function present in the system. Table 4.6 shows the results of the functional testing on this system.

Table 2: System Testing Results

Tested Feature	Test Scenario	Input	Expected Output	Result
Image Upload	User uploads an eye image in JPG/PNG format	Fundus image (JPG/PNG < 5MB)	Image is displayed in the preview panel	Success
Invalid Image Upload	User uploads a non-image file	PDF/ZIP/DOCX	System rejects the file and displays an error message	Success
Oversized Image Upload	User uploads a file larger than 5MB	JPG file (7MB)	System rejects the file and displays a maximum size notification	Success
Start Analysis	User clicks the analysis button after uploading an image	Button click	System sends a request to the backend and processes the image	Success
Analysis Without Image Upload	User clicks "Start Analysis" without uploading an image	Button click without input	System displays a warning to upload an image	Succes
Display Detection Results	System performs classification using an AI model	Valid image	Page displays the detected disease type and confidence score	Success
Display Detailed Explanation	System shows disease description	-	Brief explanation is displayed correctly	Success
Reset Examination	User clicks the reset button	Button click	Page returns to initial state (no image & no results)	Success
Save Results to PDF	User clicks "Save PDF"	Button click	System downloads a PDF file containing the examination results	Success
Menu Navigation	User selects Home / Detection menu	Menu click	System navigates to the selected page	Success

5. Conclusion

The conclusions drawn from this study are as follows:
1. The training results using the InceptionV3 architecture with the following parameters: Adam optimizer, a learning rate of 0.001, and a batch size of 32 over 100 epochs. Using these parameters, the model achieved an accuracy of 96% and average precision, recall, and F1-score values of 0.96. These results demonstrate that the model and system can assist ophthalmologists in making decisions.
2. The trained architecture was successfully integrated into a web-based eye disease detection system, capable of processing uploaded images, performing automatic predictions, and clearly displaying the results.

References

[1] D. Marcella And S. Devella, "Klasifikasi Penyakit Mata Menggunakan Convolutional Neural Network Dengan Arsitektur Vgg-19," Vol. 3, No. 1, Pp. 60–70, 2022.
[2] R. Nurlizah, A. E. Minarno, And G. W. Wicaksono, "Klasifikasi Penyakit Katarak Pada Mata Manusia Menggunakan Metode Convolutional Neural Network," *Repositor*, Vol. 4, No. 4, Pp. 491–496, 2022.
[3] Medline, "Bethesda (Md): National Library Of Medicine."
[4] H. Imaduddin And A. R. Sakina, "Eye Disease Detection Using Transfer Learning Based On Retinal Fundus Image Data," *Indonesian Journal Of Electrical Engineering And Computer Science*, Vol. 36, No. 1, Pp. 509–516, Oct. 2024, Doi: 10.11591/Ijeecs.V36.I1.Pp509-516.
[5] Kemenkes Ri, "Begini Strategi Pengentasan Gangguan Penglihatan," Https://Sehatnegeriku.Kemkes.Go.Id/Baca/Rilis-Media/20221005/1241204/Begini-Strategi-Pengentasan-Gangguan-Penglihatan/.
[6] Kemenkes Ri, "Katarak Penyebab Terbanyak Gangguan Penglihatan Di Indonesia," Https://Sehatnegeriku.Kemkes.Go.Id/Baca/Umum/20211012/5738714/Katarak-Penyebab-Terbanyak-Gangguan-Penglihatan-Di-Indonesia/.
[7] Infosehat Fkui, "Akademisi Ui: Glaukoma, Penyebab Dan Cara Mengobatinya," Https://Fk.Ui.Ac.Id/Infosehat/Akademisi-Ui-Glaukoma-Penyebab-Dan-Cara-Mengobatinya/.
[8] Honestdocs, "Retinopati - Tanda, Penyebab, Gejala, Cara Mengobati," Https://Www.Honestdocs.Id/Retinopati.
[9] N. Bhandary And A. Adnani, "Eye Disease Detection Using Resnet," *International Research Journal Of Engineering And Technology*, 2020, [Online]. Available: Www.Irjet.Net
[10] S. Sengupta, A. Singh, H. A. Leopold, T. Gulati, And V. Lakshminarayanan, "Ophthalmic Diagnosis Using Deep Learning With Fundus Images – A Critical Review," *Artif. Intell. Med.*, Vol. 102, P. 101758, Jan. 2020, Doi: 10.1016/J.Artmed.2019.101758.
[11] S. H. Abdullah, R. Magdalena, And R. Y. N. Fu'adah, "Klasifikasi Diabetic Retinopathy Berbasis Pengolahan Citra Fundus Dan Deep Learning," *Journal Of Electrical And System Control Engineering*, Vol. 5, No. 2, Pp. 84–90, Feb. 2022, Doi: 10.31289/Jesce.V5i2.5659.
[12] N. Khasanah, D. U. E. Saputri, F. Aziz, And T. Hidayat, "Enhancing Skin Cancer Classification Using Optimized Inceptionv3 Model," *Journal Medical Informatics Technology*, Sep. 2023, Doi: 10.37034/Medinftech.V1i3.14s.

[13] F. N. Cahya, N. Hardi, D. Riana, And S. Hadianti, "Klasifikasi Penyakit Mata Menggunakan Convolutional Neural Network (Cnn)," *Sistemasi: Jurnal Sistem Informasi*, Vol. 10, Pp. 618–626, 2021, [Online]. Available: Http://Sistemasi.Ftik.Unisi.Ac.Id

[14] N. W. Cahyana, *Katarak & Penanganannya Digital Repository Universitas Jember*. 2020.

[15] A. E. Suwanda And D. Juniati, "Klasifikasi Penyakit Mata Berdasarkan Citra Fundus Retina Menggunakan Dimensi Fraktal Box Counting Dan Fuzzy K-Means," 2022.

[16] Who, "Diabetic Retinopathy Screening: A Short Guide," 2020.

[17] S. Ilyas And S. R. Yulianti, *Ilmu Penyakit Mata (H. Utama (Ed.)*, 5th Edition. Badan Penerbit Fk Ui, 2019.

[18] L. Alzubaidi *Et Al.*, "Review Of Deep Learning: Concepts, Cnn Architectures, Challenges, Applications, Future Directions," *J. Big Data*, Vol. 8, No. 1, Dec. 2021, Doi: 10.1186/S40537-021-00444-8.

[19] R. H. Pramestya, "Deteksi Dan Klasifikasi Kerusakan Jalan Aspal Menggunakan Metode Yolo Berbasis Citra Digital.," Institut Teknologi Sepuluh Nopember, 2018.

[20] A. Ghosh, A. Sufian, F. Sultana, A. Chakrabarti, And D. De, "Fundamental Concepts Of Convolutional Neural Network," In *Intelligent Systems Reference Library*, Vol. 172, Springer, 2019, Pp. 519–567. Doi: 10.1007/978-3-030-32644-9_36.

[21] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, And Z. Wojna, "Rethinking The Inception Architecture For Computer Vision," In *2016 Ieee Conference On Computer Vision And Pattern Recognition (Cvpr)*, 2016, Pp. 2818–2826. Doi: 10.1109/Cvpr.2016.308.

[22] C. Goutte And E. Gaussier, "A Probabilistic Interpretation Of Precision, Recall And F-Score, With Implication For Evaluation," In *Advances In Information Retrieval*, D. E. Losada And J. M. Fernández-Luna, Eds., Berlin, Heidelberg: Springer Berlin Heidelberg, 2005, Pp. 345–359.

[23] M. A. Al-Fahrezi, "Pengaruh Augmentasi Data Terhadap Akurasi Pelatihan Model Cnn Untuk Klasifikasi Jenis Ikan," *Jitsi : Jurnal Ilmiah Teknologi Sistem Informasi*, Vol. 6, No. 2, Pp. 177–185, Jun. 2025, Doi: 10.62527/Jitsi.6.2.471.

[24] X. Pei *Et Al.*, "Robustness Of Machine Learning To Color, Size Change, Normalization, And Image Enhancement On Micrograph Datasets With Large Sample Differences," *Mater. Des.*, Vol. 232, Aug. 2023, Doi: 10.1016/J.Matdes.2023.112086.