# Vision Transformer-Based Diabetic Retinopathy Detection with Web Deployment for Clinical Decision Support

Adrian Richardy Kurniawan
*Information Systems Study Program*
*Universitas Multimedia Nusantara*
Banten, Indonesia
adrian.richardy@student.umn.ac.id

Irmawati
*Information Systems Study Program*
*Universitas Multimedia Nusantara*
Banten, Indonesia
irmawati@umn.ac.id

***Abstract*—Diabetic retinopathy (DR) is a severe complication of diabetes mellitus that can lead to vision loss if not diagnosed early. Traditional DR detection methods rely on expert interpretation of retinal fundus images, which can be time-consuming and prone to human error. This study proposes a web-based automated DR detection system utilizing a Data-Efficient Image Transformer (DeiT) model. The system classifies retinal fundus images into five DR severity levels: No DR, Mild, Moderate, Severe, and Proliferative DR. The model was trained using the APTOS2019 dataset, enhanced with data augmentation techniques to improve classification performance. The proposed approach achieved an accuracy of 89.60% on the test dataset, demonstrating its effectiveness in DR classification. The system is integrated into a web-based platform, enabling real-time image processing and user-friendly access for both patients and healthcare professionals. This research contributes to the development of automated and accessible DR detection systems to facilitate early diagnosis and treatment.**

***Keywords—diabetic retinopathy, deep learning, data-efficient image transformer, web-based system, medical image classification***

## I. Introduction

Diabetes is a chronic metabolic disease that requires serious attention and is characterized by elevated blood sugar levels due to the body's inability to transport glucose effectively, leading to its accumulation in the bloodstream blood stream[1][2]. Diabetes is classified into two types based on its cause: type 1, which results from an autoimmune response, and type 2, which occurs due to insufficient insulin production or insulin resistance in body cells[3]. Indonesia is among the top five countries with the highest number of diabetes cases, with approximately 19.5 million people affected[4]. This number is projected to increase to 40.7 million by 2045 if immediate measures are not taken[5]. Although the overall prevalence of diabetes remains relatively low, it has risen significantly from 6.2% in 2009 to 10.8% in 2021[6]. If left untreated, diabetes can lead to severe complications, including retinal damage, a condition known as diabetic retinopathy (DR), which may result in blindness or, in extreme cases, death [7]. DR can be prevented through regular monitoring, such as the A1C test, which provides an average blood sugar level over the past three months [7]. This shows that early detection of diabetes retinopathy can prevent potential blindness and even death[8]. However, current DR detection methods still rely on manual image interpretation by specialists, which is time-consuming and prone to human error, potentially delaying treatment and worsening the patient's condition [9][10].

DR detection is crucial in preventing blindness caused by diabetes. A web-based system supported by deep learning models provides a fast and accurate solution for DR classification. The application of deep learning is expected to enhance the speed and accuracy of DR detection, enabling earlier and more effective medical intervention.

This study focuses on developing a web-based system for automated DR detection using retinal fundus images. The system allows users to upload retinal images, which are then processed by a deep learning model to classify the severity of DR and provide classification accuracy. This system is designed to offer a reliable and efficient solution for the early detection of DR, improving diagnostic accuracy and facilitating timely medical treatment.

## II. Literature Review

Recent studies have explored various algorithms and models for DR detection, yet their accuracy remains below 90%. Another study by Faraq et al. employed the DenseNet169+CBAM model, resulting in an accuracy of Recent studies have explored various algorithms and models for DR detection, yet their accuracy remains below 90%. A survey conducted by Wejdan L. Ayoubi et al. 2021 utilized the EfficientNetB0 model, achieving an accuracy of 84% [11]. Another study by Mohamed Faraq et al. employed the DenseNet169+CBAM model, resulting in an accuracy of 82% [12]. Additionally, Sharmin Majumder investigated the SEDenseNet model and achieved an accuracy of 87.43% [13].

Our previous work examined using the ConvNext model combined with a Gabor filter, which improved classification performance. The model achieved a validation accuracy of 85.87% and a testing accuracy of 86.74%, outperforming the same model without a Gabor filter, which only attained a validation accuracy of 82.83% and a testing accuracy of 77.45%[14]. This research aims to surpass previous achievements by implementing Data Efficient Image Transformers (DeiT).

DeiT is a type of Vision Transformer (ViT) designed primarily for image classification tasks to achieve high accuracy while requiring less training data and computational resources than other ViT models [15][16]. DeiT offers several advantages, including data efficiency, as it requires fewer training samples to reach comparable accuracy levels. Additionally, it incorporates an efficient training strategy based on a teacher-student approach that leverages attention mechanisms. This method enables DeiT to learn effectively with fewer resources while maintaining competitive performance. Furthermore, DeiT has demonstrated state-of-the-art accuracy on various image classification benchmarks, including CIFAR-10 and ImageNet, making it a powerful and efficient alternative to traditional Vision Transformers [17][18].

In this study, the DeiT model is used to develop a classification system trained on the APTOS2019 dataset, which has been augmented to balance the number of images to 2,000 per class. After model development, it will be integrated into a web-based system that allows users to upload retinal images for classification. The web application will be built using Hypertext Markup Language (HTML), Hypertext Processor (PHP), JavaScript, Cascading Style Sheets (CSS), Standard Query Language (SQL), and Flask, with XAMPP as the local server. The findings from this research are expected to contribute to developing more reliable and efficient automated systems for DR detection.

## III. METHODS AND MATERIALS

The flowchart of this study is shown in Fig. 1. The dataset is divided into three folders: train, test, and validation.

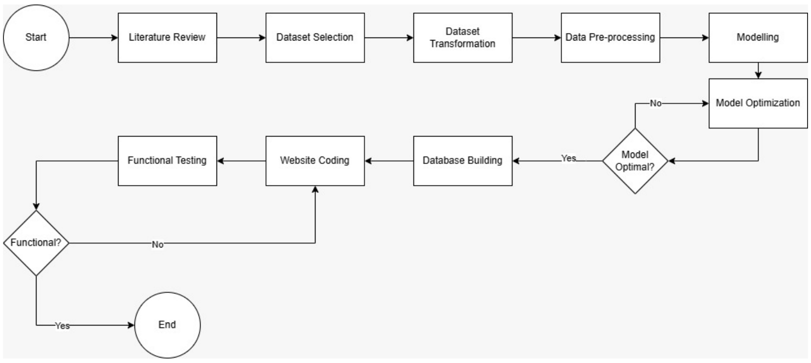


Fig. 1. System development workflow.

### A. Literature Review

A literature review was conducted to analyze existing studies on diabetic retinopathy detection. The review examined various approaches, methodologies, and models to identify existing limitations and potential improvements. Relevant studies were gathered from open-access journal databases such as Google Scholar, Semantic Scholar, and DOAJ. The findings from this review establish a foundation for selecting an appropriate deep learning approach for this study.

### B. Dataset Selection

The dataset used in this study consists of 10,000 labeled high-resolution fundus images obtained from the APTOS2019 dataset, which has undergone augmentation[19]. APTOS2019, provided by the Asia Pacific Tele-Ophthalmology Society (APTOS) as part of a Kaggle competition, is widely used for diabetic retinopathy detection research. The dataset comprises retinal fundus images collected from real-world clinical settings, captured under varying conditions such as different lighting, contrast levels, and image resolutions. These variations introduce challenges in image preprocessing and model training, making it an ideal benchmark for deep learning-based classification tasks. The dataset is categorized into five classes based on the severity of diabetic retinopathy: No DR (normal retina with no signs of diabetic retinopathy), Mild (presence of microaneurysms, indicating early-stage DR), Moderate (more significant microaneurysms and hemorrhages, suggesting progressive DR), Severe (increased hemorrhages and abnormal blood vessel growth, requiring immediate medical attention), and Proliferative DR (advanced-stage DR characterized by neovascularization, which may lead to blindness if untreated). Due to the inherent class imbalance in the dataset, augmentation techniques such as image rotation, contrast adjustments, and noise reduction were applied to improve model generalization and performance. These preprocessing steps ensure that the model is robust in handling variations in real-world fundus images. Examples of images for each class are shown in Fig. 2.


Fig. 2. Sample fundus images of DR severity levels.

### C. Dataset Transformation

The dataset was initially partitioned, with 10% of the total data reserved as an independent test set. The remaining data was divided into training (70%) and validation (30%) subsets. The dataset-splitting process used the Python split-folders library to ensure balanced class distribution across all subsets. The test set was allocated before splitting the training and validation data to maintain an unbiased evaluation and ensure the model was assessed on completely unseen data. A stratified sampling approach was also applied to preserve the proportional distribution of diabetic retinopathy severity levels, mitigating class imbalance and enhancing model reliability during training.

### D. Data Pre-processing

Before model development, the dataset underwent a pre-processing stage to ensure consistency in input size, enhance generalization, and improve model performance. The pre-processing steps included image resizing, data augmentation, and normalization, which were applied separately to the training, validation, and test datasets.

All images were resized to a fixed $224 \, x \, 224$ pixels to standardize input dimensions. Subsequently, pixel values were normalized using the mean [0.485,0.456,0.406][0.485, 0.456,0.406][0.485,0.456,0.406] and standard deviation [0.229,0.224,0.225][0.229,0.224, 0.225][0.229,0.224,0.225], which are precomputed values from the ImageNet dataset, ensuring stable model training. Four data augmentation techniques were applied to the training dataset to enhance model robustness and mitigate overfitting, including random horizontal flip, random rotation, color jitter, and random resized crop. In contrast, the validation and test datasets underwent only resizing and normalization, without augmentation, to maintain evaluation consistency.

### *E. Modelling*

This study employed the DeiT model as the backbone for image classification. Specifically, the DeiT-Base Distilled Patch16-224 architecture was selected due to its ability to efficiently learn from limited training data while maintaining high classification performance. The model was initialized with pre-trained weights from the ImageNet dataset to leverage transfer learning and accelerate convergence.

Modifications were made to the output layer of the pre-trained DeiT model to adapt the model for the diabetic retinopathy classification task. By default, the DeiT-Base Distilled architecture contains 1,000 output neurons, corresponding to the 1,000 classes in ImageNet. Since this study involves a five-class classification problem, the model's output layer was restructured by replacing the default classification head (model.head) with a fully connected layer containing five output neurons to align with the target categories. Additionally, since the DeiT Distilled variant includes a distillation token, the secondary classification head (model.head_dist) was modified to match the five-class output structure. These adaptations allow the model to extract task-specific features while retaining generalizable representations learned from pre-training, enhancing classification accuracy and robustness for diabetic retinopathy detection.

### *F. Database Architecture*

The database was designed and implemented using SQL Server with XAMPP, utilizing phpMyAdmin for database management. The database structure was developed to efficiently store and manage system-related data while ensuring scalability, integrity, and optimized performance. It consists of nine tables. The admin table stores system administrator credentials, while the doctor and patient tables manage user-related data. The appointment and appointment_history tables handle scheduling and historical records of medical consultations. The classification_history table maintains diabetic retinopathy classification results, ensuring traceability of model predictions. Additionally, the feedback table allows users to submit system performance evaluations, and the model table stores metadata for the deep learning models integrated into the system.

The database schema was structured to facilitate efficient data retrieval and storage, ensuring that patient records, appointment details, and classification results remain accessible while upholding data integrity and security. Patients can submit fundus images for classification using the deep learning model added by the administrator, with the results stored in the classification_history table. Patients can also schedule appointments with a selected doctor.

The administrator oversees system management, including user accounts, model updates, and database maintenance. Doctors primarily handle patient appointments and review patient feedback to improve healthcare services.

### *G. Website Development*

Following the database implementation, the study proceeded with the website development phase. The system architecture and functionalities were designed to ensure a seamless user experience for patients and healthcare providers.

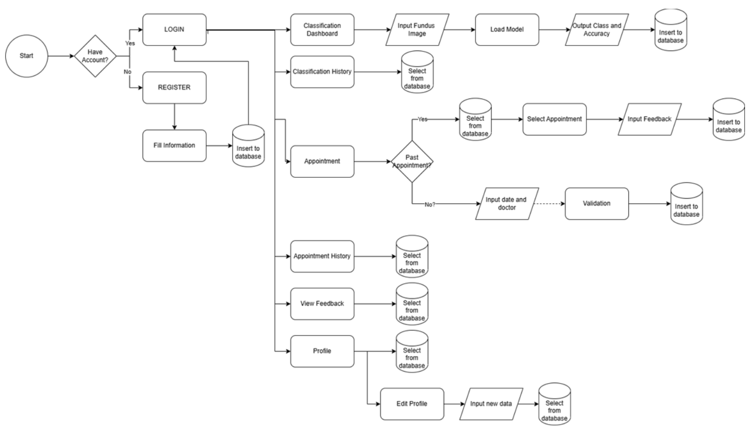


Fig. 3. Flowchart of the DR prediction system.

The patient flowchart, illustrated in Fig. 3, outlines the steps involved in the patient interaction process. Patients begin by logging into the system using their email and password. If they do not have an account, they can register a new one. Upon successful login, patients are directed to a dashboard where they can upload fundus images for classification by the model. The classification history menu allows them to view previous classification results. Additionally, patients can book appointments, check their appointment history, provide feedback on completed appointments, and manage their profiles by viewing and editing personal information.

Administrators begin by logging into the system, after which they can access five main menus: patient management, model management, doctor management, appointment management, and feedback review. Administrators can add new models and remove unused ones in the model management section. The doctor management section allows administrators to add, edit, and delete doctor records. Additionally, administrators can manage patient appointments and review patient feedback to monitor system performance and service quality.

Doctors must log in to access their dashboard to manage patient records, handle appointments, provide medical notes, and review patient feedback. The system allows doctors to assess patient feedback, improving service quality and patient care.

### *H. Black-Box Testing*

The website's functionality was evaluated using the black-box testing method, which assesses the system's outputs without examining the internal code structure. This approach ensures the system meets the expected functional requirements based on user interactions and predefined test cases. The test results determine the next steps in the development process. If the system satisfies all functional requirements, the study is considered complete. However, if any issues are identified, further development and refinement will be conducted to address the detected shortcomings before re-evaluating the system's functionality.

## IV. Result and Discussion

After training, the model's performance was evaluated using key classification metrics, including accuracy, precision, recall, and F1-score. Accuracy measures the overall correctness of the model by calculating the proportion of correctly classified instances relative to the total instances. Precision represents the proportion of correctly predicted positive cases out of all predicted positive cases, indicating the model's ability to minimize false positives. Recall measures the proportion of correctly identified positive cases among all actual positive cases, assessing the model's capability to detect true positives. The F1-score, which is the harmonic mean of precision and recall, provides a balanced evaluation of the model's performance, particularly for imbalanced datasets [20]. Once the model was evaluated and optimized, it was integrated into a web-based system using Flask, enabling real-time predictions through a user-friendly interface. This integration allows users to upload retinal fundus images and receive automated classification results, enhancing accessibility and efficiency in diabetic retinopathy detection.

### *A. DeiT Model*

During DeiT model training, early stopping was employed to prevent overfitting by terminating the training process when validation performance ceased to improve. Additionally, the training, validation, and test datasets underwent preprocessing, including various transformation techniques. These preprocessing steps ensured uniform image dimensions, applied augmentation techniques to enhance generalization, and normalized pixel values to improve learning efficiency. As a result, these optimizations contributed to the overall robustness and performance of the model.

The model was trained using the following hyperparameters: num_classes = 5, learning rate = 1e-4, batch size = 8, epochs = 100, and patience = 10. A learning rate of 1e-4 was chosen due to its stability and effectiveness in training Vision Transformers. A batch size of 8 was used to accommodate limited memory resources, while the training was conducted for a maximum of 100 epochs to allow the model sufficient time to converge. To prevent overfitting, early stopping with a patience value of 10 was applied, halting the training process if the validation accuracy did not improve over 10 consecutive epochs. These settings were found to balance training efficiency, stability, and model generalization. The training lasted 10 hours and was terminated at epoch 54 due to early stopping.

The model's performance was evaluated on the training, validation, and testing datasets. A detailed classification report for each diabetic retinopathy class is provided in Table I, summarizing precision, recall, and F1-score for each category. The overall training, validation, and testing performance is presented in Table I.

TABLE I. Classification Performance For Different Severity Levels Of DR

| Class | Precision | Recall | F1-score |
|---|---|---|---|
| 0 – No DR | 0.93 | 0.99 | 0.96 |
| 1 – Mild | 0.92 | 0.85 | 0.89 |
| 2 – Moderate | 0.83 | 0.81 | 0.82 |
| 3 – Severe | 0.89 | 0.90 | 0.90 |
| 4 – Proliferate DR | 0.90 | 0.93 | 0.91 |

The model achieved the highest performance for Class 0 (No DR) with an F1-score of 0.96, precision of 0.93, and recall of 0.99, indicating a strong ability to identify non-DR cases with minimal misclassification. Class 4 (Proliferative DR) also demonstrated high performance with an F1-score of 0.91, suggesting that the model effectively recognizes the most severe cases, which is critical for early medical intervention.

The lower performance for Moderate DR suggests potential challenges in feature differentiation, which could be addressed through additional data augmentation, improved feature extraction, or model fine-tuning. The results indicate that the model effectively classifies non-DR and severe DR cases, while moderate cases may require further refinement for improved accuracy.

TABLE II. Overall Performance Of The DR Classification Model

| Phase | Accuracy | Loss |
|---|---|---|
| Training | 93.50% | 0.1709 |
| Validation | 87.16% | 0.4007 |
| Testing | 89.60% | 0.3433 |

The overall performance of the DR classification model can be seen in Table II. During the training phase, the model achieved a high accuracy of 93.50% with a low loss of 0.1709, indicating effective learning from the training data. However, in the validation phase, the accuracy dropped to 87.16%, while the loss increased to 0.4007, suggesting potential overfitting. The higher validation loss indicates that the model's performance on unseen validation data was lower than the training set, possibly due to memorization of training samples rather than generalization. In the testing phase, the model attained an accuracy of 89.60% with a loss of 0.3433, demonstrating a good balance between performance and generalization. The improvement in accuracy compared to the validation set suggests that the model successfully adapted to new data, though some overfitting may still be present.

The confusion matrix for the validation and test sets, illustrated in Fig. 4, provides insights into the model's performance in classifying diabetic retinopathy severity levels.

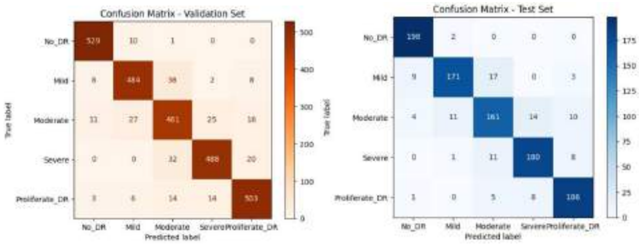


Fig. 4. Confusion matrix for the diabetic retinopathy classification model.

The validation set results indicate high classification accuracy, particularly for the No DR and Proliferative DR classes, with minimal misclassification. However, moderate misclassification is observed between Mild and Moderate categories and between Severe and Moderate, suggesting some degree of feature similarity between these severity levels. A similar pattern is observed in the test set, where the model maintains high accuracy for No DR and Proliferative DR. However, occasional misclassification occurs between adjacent severity levels, such as Mild to Moderate and Severe to Moderate, indicating challenges in distinguishing between these intermediate stages.

To improve clinical relevance and trust, future research should incorporate interpretability methods such as Grad-CAM or Score-CAM to provide visual explanations of the model's decisions. This can help highlight clinically significant regions and support medical professionals in decision-making. Additionally, to address the observed overfitting—evident from the gap between training and validation accuracy—future work may explore regularization techniques like dropout, data augmentation, or weight decay. Expanding the dataset, applying domain-specific transfer learning, and implementing cross-validation could further enhance the model's generalization and robustness.

## B. *Web Application*

The web-based system includes three user roles: admin, patient, and doctor, each with distinct privileges and functionalities. The admin has the highest level of control and is responsible for managing patients, doctors, deep learning models, patient appointments, and user feedback. The interface includes email and password fields for authentication and a role selection dropdown to specify whether the user is an admin, doctor, or patient. The "Login" button initiates the authentication process, while additional options allow new users to register and existing users to reset their password if needed. The system ensures secure access and role-based functionalities to maintain data integrity and privacy.

The Admin dashboard includes several key features, such as Manage Doctors, allowing administrators to add, edit, or remove doctor profiles. The Manage Models option enables the management of predictive models for diagnosing diabetic retinopathy. The View Feedback feature also grants access to user feedback, ensuring continuous system improvement. Administrators can also manage patient records through the Manage Patients option and monitor scheduled appointments using the View Appointments feature. Finally, the Logout button allows administrators to exit the system securely. This structured interface ensures efficient system administration and smooth operational management.

The Doctor Dashboard is designed to enhance patient management efficiency by enabling doctors to monitor appointments, review patient records, and respond to feedback, thereby improving the overall effectiveness of diabetic retinopathy diagnosis and patient care.

Fig. 5 displays the patient dashboard in the diabetic retinopathy prediction system. The left panel provides various navigation options, including classification history, appointment history, list of doctors, view feedback, profile settings and logout options. The main section of the screen features an image upload function, allowing patients to upload their retinal images for analysis by the system. Once uploaded, the system processes the image and provides a classification result indicating the severity of diabetic retinopathy.

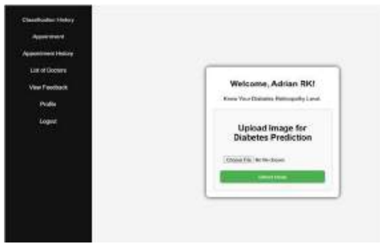


Fig. 5. Patient dashboard of the DR prediction system.

Fig. 6. displays the result of the classification process with the class and accuracy shown on the bottom of the 'Upload Image' button.


Fig. 6. Result of the DR prediction system.

The website undergoes the final phase of this study, which involves testing using the black-box method with the result shown in Table III. The results for all core functionalities for the roles are working as expected. In addition to functional validation, the system also incorporates basic security measures. Passwords are securely stored using hashing techniques, and role-based access control (RBAC) is implemented to limit access according to user privileges. To further enhance database security, all SQL queries are executed using prepared statements, mitigating the risk of SQL injection attacks. While the current deployment uses a local server environment (XAMPP), future deployment to public hosting would require additional security enhancements such as enabling HTTPS, implementing session timeout after 15–30 minutes of inactivity, and introducing login attempt limits to protect against brute-force attacks.

TABLE III. BLACK-BOX TESTING RESULT

| Role | Functionality | Output | Result |
|---|---|---|---|
| Patient,Doctor,Admin | Login | Redirect to dashboard | Pass |
| Patient | Register | Redirect to login | Pass |
| Patient | Image Classification | Result shown | Pass |
| Patient | Book/Cancel Appointment | Appointment booked | Pass |
| Patient | View Appointment History | History shown | Pass |
| Patient | View Doctor List | List shown | Pass |

## V. CONCLUSION

This study successfully developed and evaluated a Diabetic Retinopathy Prediction System using a DeiT model to enhance classification accuracy. The system was designed with three primary user roles: Admin, Doctor, and Patient, each with distinct functionalities to ensure efficient operation. The APTOS2019 dataset was used for model training, and the dataset underwent augmentation to improve performance. The DeiT model achieved a training accuracy of 93.50%, a validation accuracy of 87.16%, and a testing accuracy of 89.60%, demonstrating its ability to generalize well to unseen data. The F1 scores across different diabetic retinopathy severity levels ranged from 0.82 to 0.96, indicating strong classification performance.

## ACKNOWLEDGMENT

The authors would like to thank the support provided by Universitas Multimedia Nusantara Jakarta, Indonesia during this study.

## REFERENCES

[1] T. E. Tan and T. Y. Wong, "Diabetic retinopathy: Looking forward to 2030," *Front Endocrinol (Lausanne)*, vol. 13, no. January, pp. 1–8, 2023, doi: 10.3389/fendo.2022.1077669.

[2] Z. Yang, T. E. Tan, Y. Shao, T. Y. Wong, and X. Li, "Classification of diabetic retinopathy: Past, present and future," *Front Endocrinol (Lausanne)*, vol. 13, no. December, pp. 1–18, 2022, doi: 10.3389/fendo.2022.1079217.

[3] Corinne O'Keefe Osborn, "Type 1 and Type 2 Diabetes: What's the Difference?" [Online]. Available: https://www.healthline.com/health/difference-between-type-1-and-type-2-diabetes

[4] International Diabetes Federation, "Indonesia - International Diabetes Federation," 2021. [Online]. Available: https://idf.org/our-network/regions-and-members/western-pacific/members/indonesia/

[5] M. Wahidin *et al.*, "Projection of diabetes morbidity and mortality till 2045 in Indonesia based on risk factors and NCD prevention and control programs," *Sci Rep*, vol. 14, no. 1, pp. 1–17, 2024, doi: 10.1038/s41598-024-54563-2.

[6] D. W. Soeatmadji, R. Rosandi, M. R. Saraswati, R. P. Sibarani, and W. O. Tarigan, "Clinicodemographic Profile and Outcomes of Type 2 Diabetes Mellitus in the Indonesian Cohort of DISCOVER: A 3-Year Prospective Cohort Study," *J ASEAN Fed Endocr Soc*, vol. 38, no. 1, pp. 68–74, 2023, doi: 10.15605/jafes.038.01.10.

[7] NHS, "Diabetic retinopathy," 2021. [Online]. Available: https://www.nhs.uk/conditions/diabetic-retinopathy/#:~:text=Diabetic retinopathy is a complication,it could threaten your sight.

[8] Z. Khan *et al.*, "Diabetic Retinopathy Detection Using VGG-NIN a Deep Learning Architecture," *IEEE Access*, vol. 9, pp. 61408–61416, 2021, doi: 10.1109/ACCESS.2021.3074422.

[9] N. H. Shabrina *et al.*, "A novel dataset of potato leaf disease in uncontrolled environment," *Data Brief*, vol. 52, p. 109955, 2024, doi: 10.1016/j.dib.2023.109955.

[10] A. Sebastian, O. Elharrouss, S. Al-Maadeed, and N. Almaadeed, "A Survey on Deep-Learning-Based Diabetic Retinopathy Classification," *Diagnostics*, vol. 13, no. 3, pp. 1–22, 2023, doi: 10.3390/diagnostics13030345.

[11] W. L. Alyoubi, M. F. Abulkhair, and W. M. Shalash, "Diabetic retinopathy fundus image classification and lesions localization system using deep learning," *Sensors*, vol. 21, no. 11, pp. 1–22, 2021, doi: 10.3390/s21113704.

[12] M. M. Farag, M. Fouad, and A. T. Abdel-Hamid, "Automatic Severity Classification of Diabetic Retinopathy Based on DenseNet and Convolutional Block Attention Module," *IEEE Access*, vol. 10, no. Ml, pp. 38299–38308, 2022, doi: 10.1109/ACCESS.2022.3165193.

[13] S. Majumder and N. Kehtarnavaz, "Multitasking Deep Learning Model for Detection of Five Stages of Diabetic Retinopathy," *IEEE Access*, vol. 9, pp. 123220–123230, 2021, doi: 10.1109/ACCESS.2021.3109240.

[14] A. R. Kurniawan, "Detection of Diabetic Retinopathy in Retinal Fundus Images Using Gabor Filters and ConvNeXt," pp. 47–52, doi: 10.1145/3705927.3705936.

[15] Hugo Touvron, Matthijs Douze, Francisco Massa, Alex Sablayrolles, and Hervė Jegou, "Data-efficient image Transformers: A promising new technique for image classification," 2020. [Online]. Available: https://ai.facebook.com/blog/data-efficient-image-transformers-a-promising-new-technique-for-image-classification/

[16] Huggingface, "DeiT." [Online]. Available: https://huggingface.co/docs/transformers/en/model_doc/deit

[17] D. Schumacher, "Data-efficient Image Transformer."

[18] X. Chen, Q. Cao, Y. Zhong, J. Zhang, S. Gao, and D. Tao, "DearKD: Data-Efficient Early Knowledge Distillation for Vision Transformers," *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2022-June, pp. 12042–12052, 2022, doi: 10.1109/CVPR52688.2022.01174.

[19] S. Kumar, "APTOS2019 - 10K Augmented Images."

[20] N. S. Chauhan, "Model Evaluation metrics in Machine learning," 2020. [Online]. Available: https://www.kdnuggets.com/2020/05/model-evaluation-metrics-machine-learning.html