

Diabetic Retinopathy Classification from Fundus Images

An Minh Tri
BI12-447

Nguyen Cong Quoc
BI12-375

Vu Hung Son
BI12-390

Nguyen Duc Phuong
BI12-357

Truong Tuan Nghia
BI12-314

Abstract—Diabetic Retinopathy (DR) poses a significant threat to vision, especially among individuals with diabetes, as it silently damages the retinal blood vessels, potentially leading to blindness if untreated. Despite its global prevalence, early detection remains challenging due to the absence of obvious symptoms in the initial stages. This project addresses the urgent need for automated DR severity classification systems to aid in timely diagnosis and intervention. Leveraging deep learning techniques, specifically a pre-trained Keras ResNet50 model, this aims to develop a robust classification system capable of accurately assessing DR severity levels from retinal fundus images. By training the model on labeled datasets, the research endeavors to enable accurate differentiation between healthy and advanced stages of DR. The ultimate goal is to contribute to improved patient outcomes through early detection and management, thereby mitigating the risk of vision loss associated with DR.

Index Terms—Diabetic Retinopathy, deep learning, classification, retinal fundus images, Keras ResNet50.

I. INTRODUCTION

A. Background and Problem Statement

Diabetic Retinopathy is a complication of diabetes that silently disrupts the blood vessels in the light-sensitive tissue at the back of the eye, called the retina. High blood sugar levels can damage these delicate blood vessels, causing them to leak fluid or bleed. This damage, if left unchecked, can eventually lead to vision loss and even blindness.

Diabetic Retinopathy is a serious eye disease, and a leading cause of blindness in working-age adults. Often, there are no early warning signs, making early detection crucial. Unfortunately, if left untreated, DR can progress through several stages, each progressively more dangerous.

Diabetic Retinopathy is a global health concern. Alarmingly, the International Diabetes Federation estimated that in 2019, over 25% of people with diabetes have some level of DR. This translates to millions of people worldwide at risk of vision loss. Data specific to Vietnam is limited, but the situation appears similar. A study published in 2022 found a prevalence of DR among type 2 diabetics in Northwest Ethiopia to be 19.48%, with a similar range expected in Vietnam.

B. Objectives

In this project, we aim to develop a system for classifying Diabetic Retinopathy (DR) severity from retinal fundus images. To achieve this, we will leverage the power of deep learning by employing a pre-trained Keras ResNet50 model.

The core tasks will involve training the model on a dataset of labeled fundus images, allowing it to learn the distinctive visual characteristics associated with various DR stages. Our primary objective is to create a model capable of accurately classifying DR severity, potentially ranging from healthy to advanced stages. By achieving this, the project aspires to contribute to the early detection and management of DR. This could lead to improved patient outcomes and potentially prevent vision loss associated with the disease.

C. Project Pipeline

Typical classification pipeline



Fig. 1. Typical Classification Pipeline

Our project pipeline

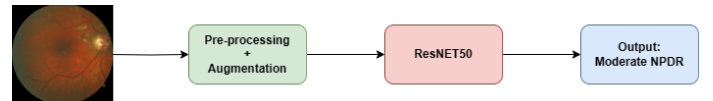


Fig. 2. Our Project Pipeline

II. DATA

A. Dataset Overview and Imbalance Challenge

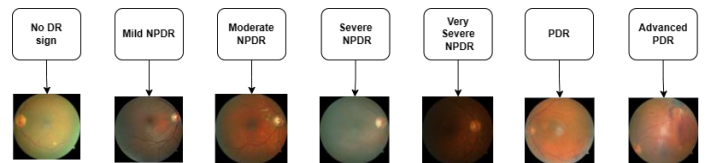


Fig. 3. Data Samples

This dataset consists of 1437 color fundus images obtained from Hospital de Clínicas in Paraguay. Ophthalmologists classified these images using a Zeiss Visucam 500 for Diabetic Retinopathy (DR) research. The dataset is divided into seven classes representing different DR severities:

- **No DR Signs:** 711 images
- **Mild NPDR:** 6 images
- **Moderate NPDR:** 110 images
- **Severe NPDR:** 210 images
- **Very Severe NPDR:** 139 images
- **PDR:** 116 images
- **Advanced PDR:** 145 images

The main challenge associated with this dataset is class imbalance. We can see that the "No DR Signs" class has a significantly higher number of images compared to the other classes, particularly Mild and Moderate NPDR. This imbalance can lead to a model biased towards the majority class and performing poorly on the minority classes, which are crucial for early DR detection.

B. Data Pre-processing

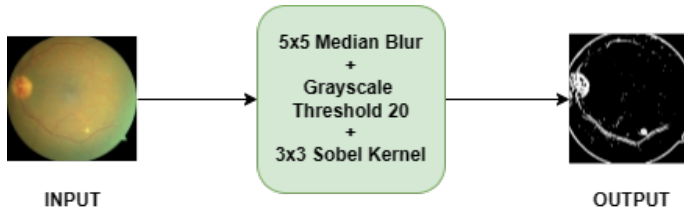


Fig. 4. Pre-processing

Blurring: The code applies a median filter to the image, which can help reduce noise and improve edge detection in the subsequent step.

Edge Detection: The Sobel operator is used to detect edges in the image. This step is likely helpful for the model to focus on important features in the fundus images that differentiate between various Diabetic Retinopathy severities.

Converting to 3-Channel Image: While the original image might be grayscale, the post-processing step replicates the edge-detected image into three channels (likely RGB) to match the expected input format of the ResNet50 model, which is typically trained on RGB images.

Blur Kernel Size: The kernel size in the blur image median function is set to 5. This means that the median blur operation uses a 5x5 pixel neighborhood for the blurring process.

Binary Edge Threshold: The first parameter in the cv2.threshold function is set to 20. This means that pixel

intensities in the edge image greater than 20 are set to the max value (255), and others are set to 0.

Sobel Kernel Size: The ksize parameter in the cv2.Sobel function is set to 3. This means that a 3x3 Sobel operator (or kernel) is used for the edge detection process.

C. Data Augmentation

We implement data augmentation specifically for the under-represented classes ("Mild" to "Advanced PDR") to address class imbalance. Here's a breakdown of the augmentation techniques used:

Flipping: Images are horizontally flipped to introduce additional variations.

Rotation: Images are rotated by various angles (10-360 degrees) to create variations in orientation.

Shifting: Images are shifted slightly horizontally and vertically to simulate minor viewpoint changes.

These techniques artificially create additional images for the minority classes, helping to balance the dataset and improve model performance on these crucial classes.

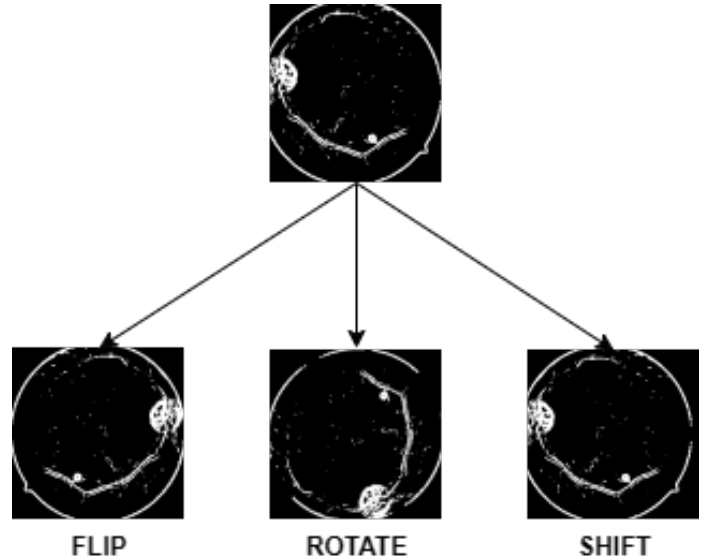


Fig. 5. Augmentation

D. After Data Augmentation

After applying augmentation techniques, the code combines the original image set with all the augmented images for each class. The corresponding labels are also assigned to the augmented images based on their original class. This creates a more balanced dataset for training the classification model.

- **No DR Signs:** 711 images.
- **Mild NPDR:** 714 (+708) images.
- **Moderate NPDR:** 715 (+605) images.
- **Severe NPDR:** 735 (+ 525) images.
- **Very Severe NPDR:** 695 (+579) images.
- **PDR:** 696 (+580) images.
- **Advanced PDR:** 725 (+580) images.

And then we use:

- **Training:** 60% (2994 images).
- **Validation:** 20% (998 images).
- **Testing:** 20% (999 images).

III. THE RESNET50 MODEL

A. Why ResNet50?

We chose the ResNet50 model for this task because it's been trained on a massive collection of images, learning to recognize all sorts of basic visual clues. We can leverage this pre-trained knowledge as a foundation for our DR detection system. It's like giving our system a head start in understanding the building blocks of images.

Instead of training a whole new system from scratch, we can take advantage of ResNet50's expertise. We can then fine-tune the later parts of the system to specifically focus on the unique patterns and features present in retinal images that indicate different DR severities. It's like showing our detective trainee retinal images and highlighting the specific details that signal potential trouble.

This approach offers several benefits. First, it saves us a ton of time. Training a complex system from scratch on a relatively small dataset like ours (think 1437 images) would take ages. Second, it reduces the risk of the system overfitting, which is like the trainee memorizing every detail from the training images and failing to recognize new patterns. Finally, it allows us to achieve good accuracy without needing a massive amount of computing power, which is always a bonus.

B. ResNet50 Architecture Breakdown

ResNet (Residual Network) is a type of CNN architecture designed to address the vanishing gradient problem, a common hurdle in training deep neural networks. It introduces skip connections that allow the gradients to flow directly through the network, facilitating better optimization during training. Here's a simplified breakdown of the key components:

Convolutional Base: The initial layers consist of several convolutional layers with activation functions (e.g., ReLU) and pooling layers (e.g., max pooling). These layers work together to extract features from the input image. Imagine these layers as the initial steps for the "detective" in our analogy, learning to recognize basic shapes, lines, and colors

within the image.

Residual Blocks: The core building block of ResNet is the residual block. A residual block consists of a series of convolutional layers with activation functions, followed by a skip connection that adds the input of the block directly to its output element-wise. This "shortcut" allows the gradients to flow directly through the network, even in very deep architectures, helping the network learn complex features from retinal images. In our analogy, think of these residual blocks as the detective's experience kicking in. They allow the model to build upon the foundational features learned earlier and start recognizing more intricate patterns specific to DR classification.

Global Average Pooling: After the residual blocks, a global average pooling layer is applied. This layer summarizes the features extracted throughout the network into a fixed-size vector. Imagine this as the detective summarizing all the essential clues extracted from the image into a concise report.

Classification Head: Finally, the model has fully-connected layers with activation functions. The last layer has the number of neurons equal to the number of classes (7 in this case for Diabetic Retinopathy classification) and uses a softmax activation function to output the class probabilities for each image. This is like the final step where the detective assigns probabilities (likelihoods) of the image belonging to a specific DR severity class.

In essence, the ResNet50 architecture leverages pre-trained knowledge to extract progressively more specific features from retinal images. The residual blocks help the network learn complex patterns, and the final layers classify the image into one of the seven DR severity classes.

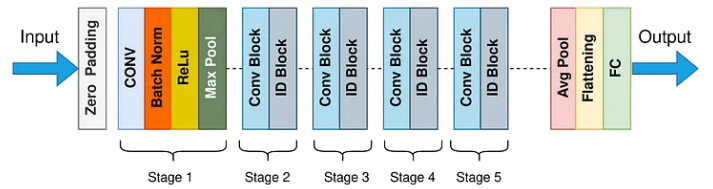


Fig. 6. Model Architecture

IV. EVALUATION

A. Classification Report

The model performs exceptionally well on class 1 compared to other classes. This is likely due to the fact that we used all 6 images of class 1 for augmentation, which may have given an unfair advantage to this class. To address this, we plan to use only 3 images for augmentation and reserve the remaining 3 as new images for validation. This adjustment

TABLE I
CLASSIFICATION REPORT

Class	Precision	Recall	F1-Score	Support
0	0.80	0.86	0.83	143
1	0.98	0.99	0.99	150
2	0.79	0.63	0.70	142
3	0.72	0.78	0.75	149
4	0.73	0.65	0.69	145
5	0.74	0.56	0.64	124
6	0.64	0.88	0.74	146
Accuracy			0.77	999
Macro Avg	0.77	0.76	0.76	999
Weighted Avg	0.77	0.77	0.77	999

should yield better results.

Regarding the classification of Diabetic Retinopathy,, our primary focus is on the recall metric. The current model shows satisfactory performance on classes 0, 3, and 6. However, it falls short on classes 2, 4, and 5. To enhance the model's performance on these classes, we have several strategies at our disposal. We could consider gathering more data for these classes. Alternatively, we could explore using more sophisticated models or employing ensemble methods, which combine the predictions of multiple models for a potentially more accurate result.

B. Confusion Matrix

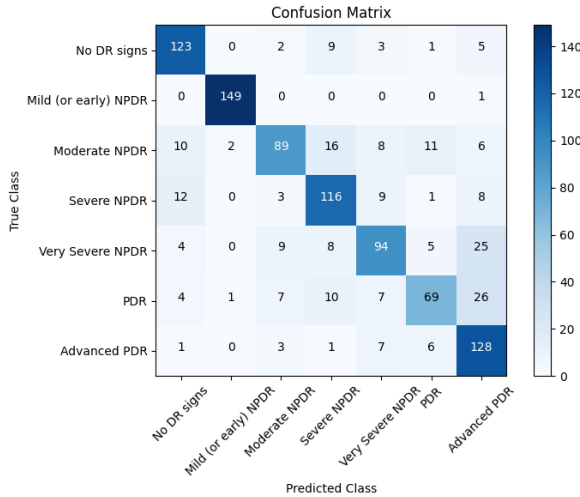


Fig. 7. Confusion Matrix

C. Each Class Accuracy

V. CONCLUSION

This project explored the potential of deep learning for automated Diabetic Retinopathy (DR) classification from retinal fundus images. We utilized a pre-trained ResNet50 model, leveraging its transfer learning capabilities and feature extraction power to address the task on a relatively small

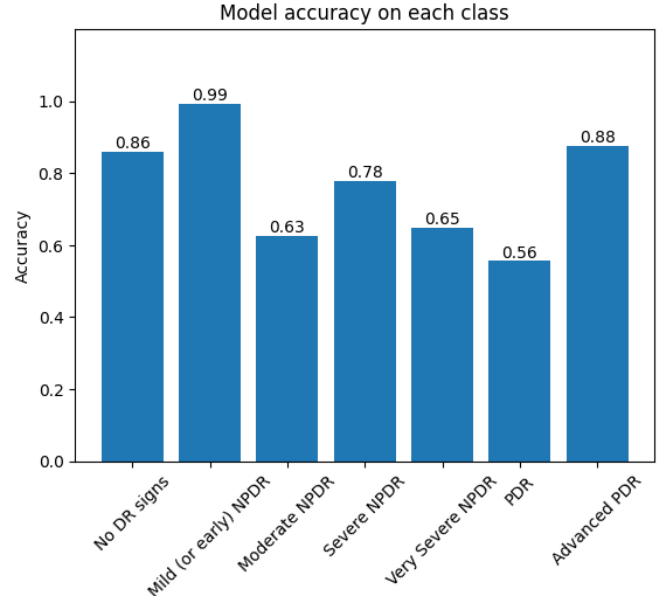


Fig. 8. Accuracy of each class

dataset.

The chosen approach offered several advantages. Firstly, by leveraging pre-trained knowledge, we significantly reduced training time compared to building a model from scratch. Secondly, the initial layers of ResNet50 efficiently extracted generic features from the images, providing a strong foundation for identifying relevant structures and textures in retinal images that differentiate between DR severities. Finally, fine-tuning the later layers of ResNet50 allowed us to specialize the model for the specific task of DR classification.

The evaluation of the model's performance involved analyzing various metrics like accuracy, precision, recall, and F1-score. While the overall accuracy provided a general indication of the model's effectiveness, a more in-depth analysis of class-specific metrics was crucial. This analysis allowed us to identify areas where the model excelled, such as accurately classifying severe DR cases. However, it also revealed potential areas for improvement, particularly in differentiating mild DR cases from healthy retinas.

This project demonstrates the promise of deep learning for automated DR classification. By leveraging pre-trained models and addressing limitations through data augmentation and potential interpretability techniques, this approach has the potential to become a valuable tool for early DR detection. Early diagnosis and intervention are critical for preventing vision loss in diabetic patients. This technology could contribute to improved patient care by enabling faster and more efficient screening, ultimately leading to better clinical outcomes.

REFERENCES

- [1] Castillo Benítez, V. E., Castro Matto, I., Mello Román, J. C., Vázquez Noguera, J. L., García-Torres, M., Ayala, J., Pinto-Roa, D. P., Gardel-Sotomayor, P. E., Facon, J., & Grillo, S. A. Dataset from fundus images for the study of diabetic retinopathy.
- [2] Wang, J., & Perez, L. The Effectiveness of Data Augmentation in Image Classification using Deep Learning.
- [3] He, K., Zhang, X., Ren, S., & Sun, J. Deep Residual Learning for Image Recognition.