

اسراء الوتار

_An Efficient Hybrid Framework Combining Transformer Feature Encoding and Machine Learning for Diabetic Retinopa...

 Midocean University

Document Details

Submission ID

trn:oid::2945:379114255

Submission Date

Apr 22, 2026, 10:50 PM GMT+2

Download Date

Apr 22, 2026, 11:21 PM GMT+2

File Name

_An Efficient Hybrid Framework Combining Transformer Feature Encoding and Machine Learni....docx

File Size

10.0 MB

36 Pages

5,500 Words

31,937 Characters

18% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography
- Quoted Text

Match Groups

-  **86 Not Cited or Quoted 18%**
Matches with neither in-text citation nor quotation marks
-  **0 Missing Quotations 0%**
Matches that are still very similar to source material
-  **0 Missing Citation 0%**
Matches that have quotation marks, but no in-text citation
-  **0 Cited and Quoted 0%**
Matches with in-text citation present, but no quotation marks

Top Sources

- 12%  Internet sources
- 11%  Publications
- 16%  Submitted works (Student Papers)

Integrity Flags

0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.

Match Groups

- 86 Not Cited or Quoted 18%**
Matches with neither in-text citation nor quotation marks
- 0 Missing Quotations 0%**
Matches that are still very similar to source material
- 0 Missing Citation 0%**
Matches that have quotation marks, but no in-text citation
- 0 Cited and Quoted 0%**
Matches with in-text citation present, but no quotation marks

Top Sources

- 12% Internet sources
- 11% Publications
- 16% Submitted works (Student Papers)

Top Sources

The sources with the highest number of matches within the submission. Overlapping sources will not be displayed.

- 1** Internet
ruor.uottawa.ca <1%
- 2** Submitted works
University of Hertfordshire on 2025-09-03 <1%
- 3** Internet
www.nature.com <1%
- 4** Submitted works
Swinburne University of Technology on 2024-10-06 <1%
- 5** Internet
www.mdpi.com <1%
- 6** Submitted works
Midocean University on 2025-09-05 <1%
- 7** Internet
www.ijert.org <1%
- 8** Internet
eprints.usm.my <1%
- 9** Publication
Arvind Dagur, Sohith Agarwal, Dharendra Kumar Shukla, Shabir Ali, Sandhya Sharm... <1%
- 10** Internet
www.internationaljournalssrg.org <1%

11	Internet	www.fastercapital.com	<1%
12	Submitted works	Liverpool John Moores University on 2026-03-10	<1%
13	Internet	dokumen.pub	<1%
14	Internet	ijsrem.com	<1%
15	Internet	www.uab.edu	<1%
16	Submitted works	Loughborough University on 2023-02-23	<1%
17	Publication	"Intelligent Vision and Computing", Springer Science and Business Media LLC, 2026	<1%
18	Publication	"Multi-Strategy Learning Environment", Springer Science and Business Media LLC...	<1%
19	Publication	Pushpa Choudhary, Sambit Satpathy, Arvind Dagur, Dhirendra Kumar Shukla. "Re...	<1%
20	Internet	ai.uga.edu	<1%
21	Internet	upcommons.upc.edu	<1%
22	Publication	Harvinder Singh, Priyanka Kaushal, Sarabjeet Kaur. "Advanced Computing and AI...	<1%
23	Publication	Maple-Brown, L.. "Complications of diabetes in urban Indigenous Australians: Th...	<1%
24	Internet	dr.ntu.edu.sg	<1%

25	Internet	file.techscience.com	<1%
26	Submitted works	universiteknologimara on 2024-07-04	<1%
27	Publication	E. Rajkumar, P. Udayasankaran, S. Ravichandran, P. Thilagavathi. "Feature intensi...	<1%
28	Submitted works	University of Mauritius on 2026-03-30	<1%
29	Internet	theses.dur.ac.uk	<1%
30	Internet	jisem-journal.com	<1%
31	Submitted works	Anna University on 2025-06-28	<1%
32	Publication	H L Gururaj, Francesco Flammini, V Ravi Kumar, N S Prema. "Recent Trends in He...	<1%
33	Publication	M.T. Mastura, S.M. Sapuan, R.A. Ilyas. "Additive Manufacturing for Bio-Composite...	<1%
34	Submitted works	Middlesex University on 2024-10-08	<1%
35	Submitted works	Midocean University on 2025-09-05	<1%
36	Publication	Prakash Arumugam, Divya Bhavani Mohan. "chapter 6 A Hybrid ArtInt Model to P...	<1%
37	Submitted works	Staffordshire University on 2012-02-08	<1%
38	Internet	fjs.fudutsinma.edu.ng	<1%

39	Submitted works	Anna University on 2025-07-26	<1%
40	Submitted works	Asia Pacific Institute of Information Technology on 2025-08-31	<1%
41	Publication	Imtiyaz Ahmad, Vibhav Prakash Singh, Manoj Madhava Gore. "Exploring AI-assist..."	<1%
42	Submitted works	Iqra University on 2025-08-05	<1%
43	Submitted works	Liverpool John Moores University on 2026-02-09	<1%
44	Submitted works	Universiti Teknologi Petronas on 2023-02-11	<1%
45	Submitted works	University of Hertfordshire on 2026-03-01	<1%
46	Submitted works	University of Hertfordshire on 2026-04-13	<1%
47	Submitted works	University of Leeds on 2021-10-28	<1%
48	Submitted works	University of Nottingham on 2024-09-05	<1%
49	Submitted works	University of Technology, Sydney on 2023-11-05	<1%
50	Submitted works	University of Western Sydney on 2023-01-11	<1%
51	Submitted works	University of Wollongong on 2025-06-27	<1%
52	Internet	download.bibis.ir	<1%

53	Internet	ijetrm.com	<1%
54	Internet	pmc.ncbi.nlm.nih.gov	<1%
55	Internet	technodocbox.com	<1%
56	Publication	"Computer Vision and Robotics", Springer Science and Business Media LLC, 2026	<1%
57	Submitted works	Birkbeck College on 2025-09-08	<1%
58	Submitted works	Higher Education Commission Pakistan on 2025-02-19	<1%
59	Submitted works	Universiti Putra Malaysia on 2015-12-13	<1%
60	Submitted works	University of Northampton on 2024-01-19	<1%
61	Submitted works	University of Westminster on 2026-04-03	<1%
62	Publication	Vijayalaxmi Gopu, M. Selvi. "Reinforced vision model using Q-learning GoogleNet ...	<1%
63	Publication	Vivek Sunnapwar, Hariram Chavan, Namrata F. Ansari. "Advances in Science and ...	<1%
64	Submitted works	Liverpool John Moores University on 2025-02-23	<1%
65	Internet	i-rep.emu.edu.tr:8080	<1%



2

An Efficient Hybrid Framework Combining Transformer Feature Encoding and Machine Learning for Diabetic Retinopathy Classification

Prepared By

Name: Asraa Alwattar ID No.: 253001066

6
Name: Rahaf Barakat ID No.: 253002370

Submitted in fulfilment of the requirements for the degree Master of
Science in Artificial Intelligence

At

College of Informatics

Midocean University

UAE

Supervisor Name: Wessam Ahmad

College of Informatics, Midocean University, Comoros

Email:.....

Acknowledgements

Special thanks to Dr. .

Abstract

List of Publications

Table of Contents

Contents

Acknowledgements	2
Abstract	3
List of Publications	3
Table of Contents	4
List of Abbreviations.....	6
List of Figures	7
List of Tables.....	8
Chapter 1 Introduction	1
1.1 Overview	1
1.2 Problem Definition.....	1
1.3 Thesis Objectives	2
1.4 Thesis Contribution.....	3
1.5 Thesis Organization	4
Chapter 2 Background and Related Work:	5
2.1 Introduction	5
2.2 Background	5
2.2.1 Diabetic Retinopathy Overview.....	Error! Bookmark not defined.
2.2.2 Artificial Intelligence in Medical Image Analysis:.....	6
2.2.3 Diabetic Retinopathy Detection Using Machine Learning.....	7
2.2.4 Diabetic Retinopathy Detection Using Deep Learning	9
2.2.5 Diabetic Retinopathy Detection Using Transformer Models	11
2.2.6 Performance Metrics	15
2.3 Literature Review.....	17

50

2.4 Comparative Analysis of Related Work19

2.5 Chapter Summary.....23

8

Chapter 3 Proposed Approach (Methodology)23

Chapter 4: Experimental Results and Discussion25

Chapter 5 Conclusion and Future Work26

References26

List of Abbreviations

- i. Diabetic Retinopathy (DR)
- ii. Machine Learning (ML)
- iii. Deep Learning (DL)
- iv. Convolutional Neural Networks (CNNs)
- v. Support Vector Machines (SVM)
- vi. True Positives (TP)
- vii. True Negatives (TN)
- viii. False Positives (FP)
- ix. False Negatives (FN)
- x. Random Forest (RF)
- xi. Vision Transformer (ViT)
- xii. Data-efficient Image Transformer (DeiT)

List of Figures

Figure 1: AI Techniques for Retinal Image	7
Figure 2: SVM separation of DR and healthy classes	8
Figure 3: Random Forest for classification by majority voting	8
Figure 4: KNN classification using closest neighbors	9
Figure 5: Three models for retinal feature extraction in DR detection	10
Figure 6: Comparison of models for DR feature extraction	11
Figure 7: Patch-based self-attention image classification	12
Figure 8: Swin Transformer with shifted window attention	13
Figure 9: DeiT for efficient learning through knowledge distillation	14

List of Tables

1. Comparative Study of Diabetic Retinopathy Detection Methods

22,23

Chapter 1 Introduction

1.1 Overview

One of the most threatening consequences of diabetes is diabetic retinopathy, which causes many preventable cases of blindness. This problem occurs due to the appearance of certain pathologies in the blood capillaries of the retina caused by hyperglycemia. Pathologies include, among other things, aneurisms, exudation, and bleeding. As they tend to progress quietly and gradually, and not all of them are evident at first, it is very important to detect them to avoid irreversible changes to the patient's vision.[1] At present, the diagnosis of diabetic retinopathy involves manual examination of photographs of the bottom of the eye made by an ophthalmologist. Although it can be reliably applied to clinical practice, such a method is usually quite time-consuming, as well as subjective and dependent on the skills of a specialist. Furthermore, this method proves difficult to apply in mass screenings, during which the number of people to be examined often exceeds available human resources. Therefore, there is growing interest in using artificial intelligence and automated techniques that will allow doctors to diagnose diabetic retinopathy more easily and accurately.[2] The development of artificial intelligence, machine learning, and, in particular, deep learning has created new opportunities for medical image analysis. For example, convolutional neural networks show excellent results in the extraction of visual characteristics from the bottom of the eye. Machine learning classifiers, such as Random Forest, are known for their reliability and versatility. More recent transformer-based methods also promise a higher level of understanding the overall context of the image. Considering the existing approaches, this research project seeks to combine ResNet and Swin Transformer architectures for image pre-processing and extraction of distinctive

features and employ Random Forest classifier for the final classification stage. This research will focus on exploring the potential of applying preprocessing, expansion, and classification based on extracted features in diagnosing diabetic retinopathy.[3]

1.2 Problem Definition

Diabetic retinopathy is one of those conditions that doesn't get enough attention until it's almost too late for the patient yet it's a leading cause of vision loss that, with the right tools, is entirely preventable if it's caught early. The challenge is that diagnosing it today still relies heavily on doctors manually examining retinal images, a process that's not only time-consuming but subject to the kind of variability that comes with any human judgment call, especially in situations where expert knowledge may not be readily available.

Machine learning has stepped in as a promising alternative, and honestly, the early results have been encouraging. But it's not a solved problem. Real-world retinal images vary significantly in quality, the more severe stages of DR are far rarer in datasets than milder ones, and most models are built around a single architectural approach which means they tend to miss features that fall outside their particular "lens." No single model sees the full picture, and that gap shows in performance of the accuracy of results.

That's the gap that this research is trying to close. The idea is straightforward: if one approach misses things another catches, by fusing features extracted through both ResNet and the Swin Transformer and then running them through a Random Forest classifier this thesis explores whether a hybrid strategy can outperform what any single method achieves alone, ultimately moving us closer to a diagnostic tool that's reliable enough to trust in a clinical practice setting.

65

1.3 Thesis Objectives

The main goal of the thesis is to design an effective and accurate automated system to classify diabetic retinopathy with the help of retinal fundus images. For this purpose, the research study employs a combination of image pre-processing, deep feature extraction, and machine learning based classification. ResNet and Swin Transformer have been used for the extraction of relevant retinal features, and thereafter, the classification has been done using Random Forest technique. The proposed model is assessed based on various performance metrics such as accuracy, precision, recall, F1-score, specificity, and AUC. Additionally, the performance of the proposed model is compared with traditional deep learning techniques.

1.4 Thesis Contribution

This study develops an efficient automatic system to classify diabetic retinopathy through using a combination of image processing, deep feature learning, and machine learning-based classification, this is achieved as follows: Create a hybrid model based on ResNet and Swin Transformers for feature learning from retinal images.

Implementing the Random Forest algorithm as the final classifier in the process. The performance of the created model is evaluated in terms of accuracy, precision, recall, F1-score, specificity, and area under curve (AUC).

Compare the created model with existing baselines for diabetic retinopathy classification.

4

1.5 Thesis Organization

The rest of this dissertation will be organized as follows:

- **Chapter 2** lays the groundwork, covering the theoretical background behind diabetic retinopathy screening from the medical side of the disease itself to the machine learning, deep learning, and transformer-based approaches used to analyze retinal images. It also walks through the evaluation metrics used in classification tasks and situates this work within the broader body of existing literature.
- **Chapter 3** covers the proposed methodology in full detail. This includes how the data was collected and enhanced, why ResNet and Swin Transformer were chosen as the backbone for feature extraction, how those features are brought together through fusion, and how the final classification is carried out using Random Forest.
- **Chapter 4** shifts to results. The proposed framework is evaluated across several performance metrics, and its outputs are compared against baseline models to get a clearer sense of where it stands and what it actually adds.
- Finally, **Chapter 5** concludes the dissertation by summarizing the main findings, the limitations of the current approach, and pointing toward what future research in this area.

Chapter 2 Background and Related Work:

2.1 Introduction

This chapter provides theoretical foundations regarding application of artificial intelligence techniques in detection of diabetic retinopathy. The chapter begins with introduction of diabetic retinopathy, highlighting the importance of its timely diagnosis due to serious implications on health and well-being. It continues with discussion of application of artificial intelligence techniques in medical image analysis, namely machine learning, deep learning, and transformers that show high potential for detection of diabetic retinopathy. Further, the chapter discusses metrics that can be used to measure performance of diabetic retinopathy classifiers, such as accuracy, precision, recall, F1 score, specificity, confusion matrix, and AUC. In addition, the rationale behind choosing these specific metrics is discussed, including problems with class imbalance and reliability of classifiers that have significance in medical image analysis. Lastly, literature review on previous research in domain of diabetic retinopathy classification is provided in order to reveal gaps in current literature.

2.2 Background

2.2.1 Diabetic Retinopathy Overview

Diabetic retinopathy (DR) is among the most frequently occurring microangiopathies associated with diabetes mellitus, which causes serious vision problems all around the world. It occurs because chronic elevation of blood glucose damages blood vessels of the retina. Microaneurysms, hemorrhages, exudates, cotton wool spots, and even neovascularization – all these are the features of the disease. However, being a progressive condition which might be asymptomatic for

23 a certain period of time, DR requires routine screening in order to diagnose it
58 timely and treat accordingly. Since this condition can occur insidiously and remain
43 without symptoms in its earlier stages, it is vital for regular testing to be carried
out in order to detect this condition. This condition is classified clinically in various
grades depending on their severity, starting from mild or moderate non-
proliferative diabetic retinopathy, progressing through to severe non-proliferative
and proliferative diabetic retinopathy. The correct diagnosis of the different stages
of diabetic retinopathy is very important because it is highly dependent upon the
stage at which the patient is for treatment. As such, there has been an increased use
of computerized systems for the diagnosis of diabetic retinopathy [4].

2.2.2 Artificial Intelligence in Medical Image Analysis:

The utilization of Artificial Intelligence has made its way into current medicine in
terms of usage in analyzing medical images. The usage of AI-based algorithms has
the capacity to train itself by detecting patterns from big data and help in
performing tasks such as disease detection, grading, and decision-making. Retinal
image analysis involves extensive use of Artificial Intelligence techniques such as
in identifying disease-related features, grading, and large scale screening purposes.
Machine Learning One of the main subfields of AI is machine learning. Machine
learning algorithms require a feature extractor to first extract the relevant feature
followed by the utilization of classifiers. Some of the relevant features which are
used in medical imaging include texture, shape, intensity, and vessel
characteristics. However, while the traditional machine learning algorithms have
the ability to yield good performance, it is quite sensitive with regards to the feature
extraction phase. Deep Learning The Convolutional Neural Network is a deep
learning algorithm that is quite extensively used in image classification because it
can classify images by extracting both low-level and high-level visual features.

Recently there has also been the use of transformers for medical images. Consequently, the transformer network has the ability to detect more contexts within images than other networks. Therefore, both CNNs and transformers are emerging trends in DR detection [4] [5], as shown in Figure 1:

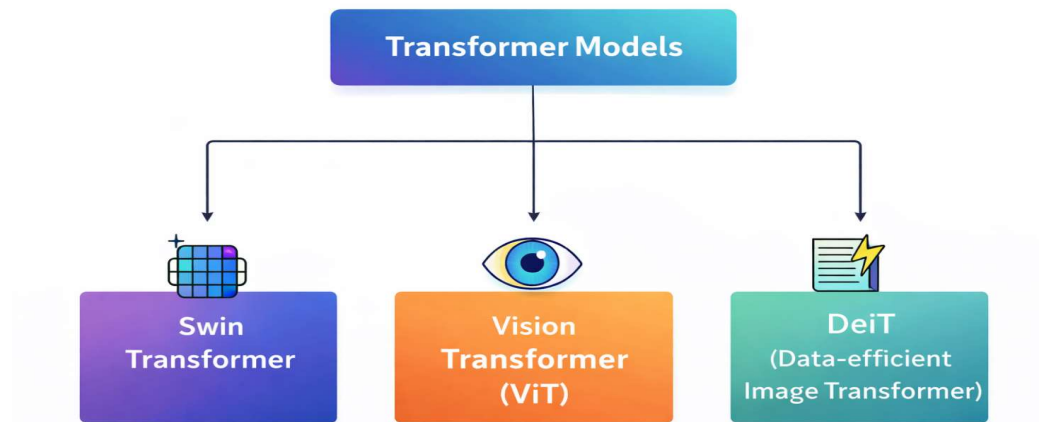


Figure 1: AI Techniques for Retinal Image Analysis [4] [5]

2.2.3 Diabetic Retinopathy Detection Using Machine Learning

Machine learning techniques for the detection of diabetic retinopathy have seen extensive use, wherein the process of feature extraction and that of classification were performed in two different processes. Here, the features of an image were extracted through manual or deep learning methods, followed by prediction of the disease class using an existing algorithm and such a methodology survived the test of time for a very good reason. It provides a perfect balance of efficiency, whereby the performance of the model is competitive but still allows more transparency with regard to the factors being considered. The interpretability of the model is definitely one of its strengths, as is the reduced cost of training compared to other approaches.

1. Support Vector Machines (SVM)

Support Vector Machine is one of the popular methods used in diagnosis in medical fields. SVM determines the ideal classification boundary in terms of margin maximization between different groups of observations. In the application of classification in diabetic retinopathy diagnosis, SVM performs optimally in scenarios when the classifier has to operate in high-dimensional spaces. Even with fewer data points available for classification, SVM works well since the classifiers are efficient. [6], as shown in Figure 2:

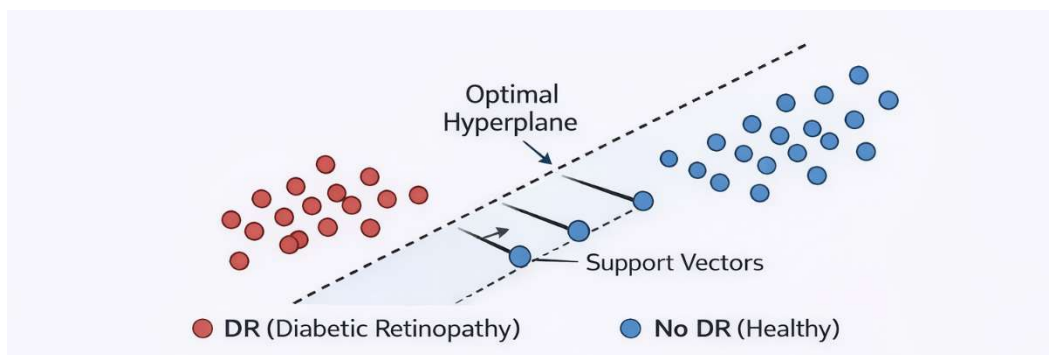


Figure 2: SVM separation of DR and healthy classes [6].

3- Random Forest:

The Random Forest classifier is also one of the most common classifiers that belongs to the group of machine learning classifiers that is especially useful for the present study. This classifier operates through the use of ensemble method where multiple decision trees are involved in making decisions based on the majority vote principle. Since this classifier uses the ensemble method, it is able to overcome overfitting and therefore it can be used with real-life data like retinal images. Also, it can be used with imbalanced data due to stable performance in all classes [7] as shown in Figure 3:

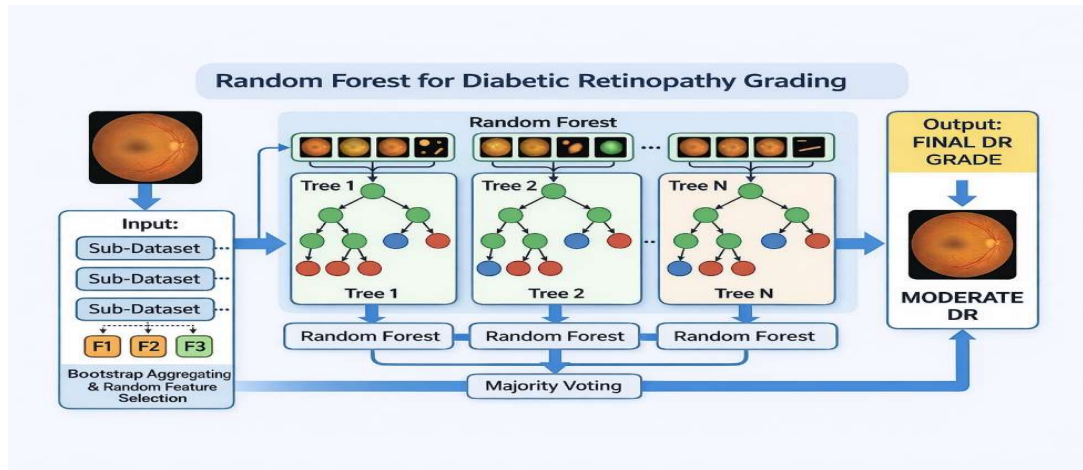


Figure 3: Random Forest for classification by majority voting [7].

2. K-Nearest Neighbors (KNN):

K-Nearest Neighbors (KNN) is a simple classifier that serves as a baseline for comparing other classifiers. The algorithm is defined by the fact that the class label of the test data point is assigned based on the class labels of its neighbors in the feature space. Despite being simple and comprehensible, the algorithm is sensitive to scale and distribution changes in the features. [8], as shown in Figure below:

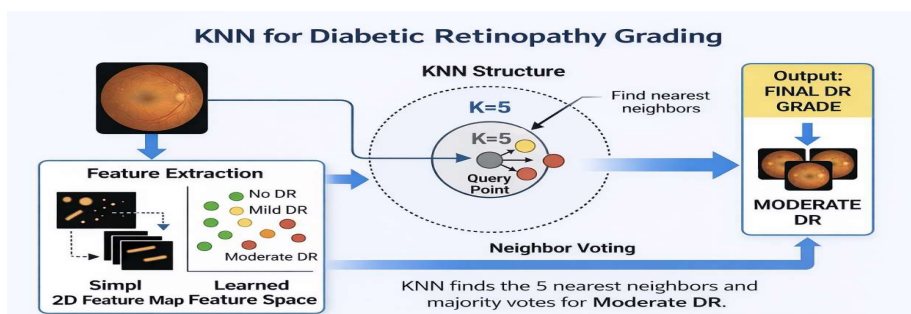


Figure 4: KNN classification using closest neighbors [8].

2.2.4 Diabetic Retinopathy Detection Using Deep Learning

Deep learning has significantly changed the way diabetic retinopathy is detected from retinal fundus images. Unlike traditional machine learning methods, which depend on manually designed features, deep learning models can automatically learn hierarchical visual representations directly from the input images. This

allows them to detect both simple image structures and complex pathological patterns relevant to disease severity. Among deep learning models, Convolutional Neural Networks have been the most widely used in retinal image analysis. These networks are particularly effective because they can learn spatial features such as edges, textures, lesions, and structural abnormalities. In diabetic retinopathy detection, CNNs have been applied successfully to identify retinal signs such as microaneurysms, hemorrhages, and exudates [9], as shown in Figure 5:

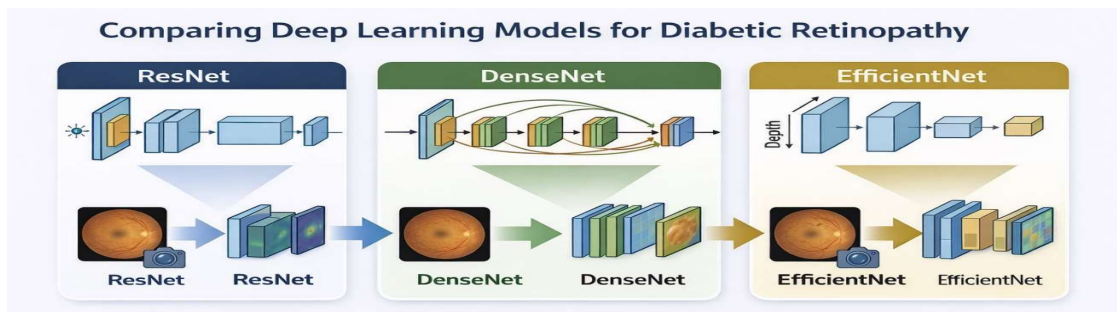


Figure 5: Three models for retinal feature extraction in DR detection [9].

1. ResNet50 (Pre-trained)

One reason why ResNet50 has gained popularity is due to its ability to address the problem of vanishing gradients using its skip connection feature, making learning possible in deeper networks. By employing transfer learning, the model ResNet50 will be able to gain information acquired from large image datasets and apply it to the problem of retinal image classification [10].

2. DenseNet121 (Pre-trained)

In addition, DenseNet121 provides another powerful choice, which connects every layer with all other layers after it. This makes the architecture more efficient at feature reuse and enables better flow of information. For medical images, this could make the neural network detect even the slightest symptoms caused by diseases [10].

3. EfficientNetB0

The above analysis shows that the EfficientNetB0 network architecture is primarily concerned about making the most out of the combination of the depth, width, and input size. It makes the efficient net architecture highly effective in cases where computing resources are scarce. Moreover, this network architecture can be considered optimal due to its high effectiveness. [10].

2.2.5 Diabetic Retinopathy Detection Using Transformer Models

However, recent advances in transformer architectures seem to represent a better solution to convolution-based networks when applying image classification tasks. The main advantage of these types of models is that they utilize a special architecture known as self-attention, which makes it possible to detect local and global relations between different elements of the image. This is very important when recognizing DR since pathologies can develop anywhere on the retina [10], as shown in Figure 6:

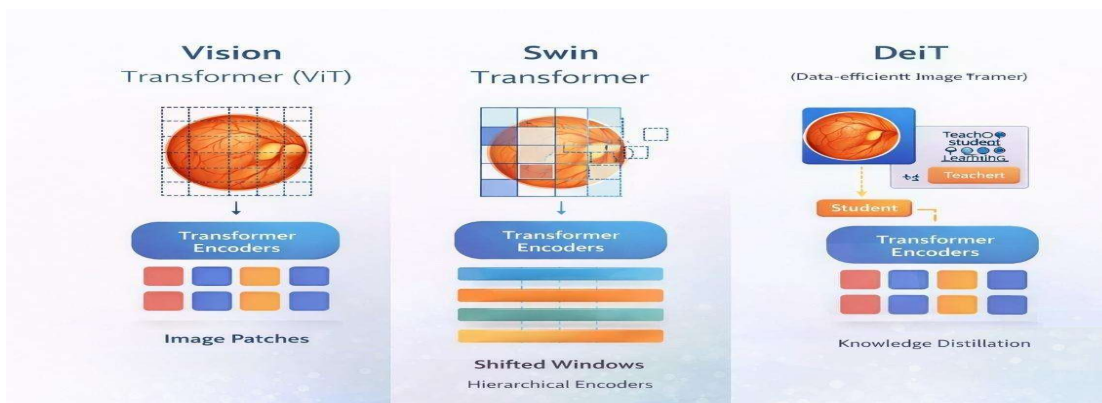


Figure 6: Comparison of models for DR feature extraction [10].

1- Vision Transformer (ViT)

Vision Transformer (ViT) is one of the earliest transformer architectures for classification of images. While CNN networks use convolutional filters for image analysis, Vision Transformer splits an image into different patches and encodes each patch as a vector referred to as patch embedding with positional encoding that encodes the spatial relations within patches. Such embeddings go through the layers of transformers' encoder where the relationships between various regions within the image are established through the attention mechanism. One of the key features of Vision Transformer is the possibility of capturing dependencies between all pixels of an entire image. Thus, Vision Transformers could be used effectively for medical imaging tasks when some important pathological alterations in organs spread throughout different locations of an image. For instance, such alterations are relevant for the detection of diabetic retinopathy because one should analyze several regions of the retina. However, in order for ViT to function properly, one should provide more data for training and consume more computational power.[10], as shown in Figure 7:

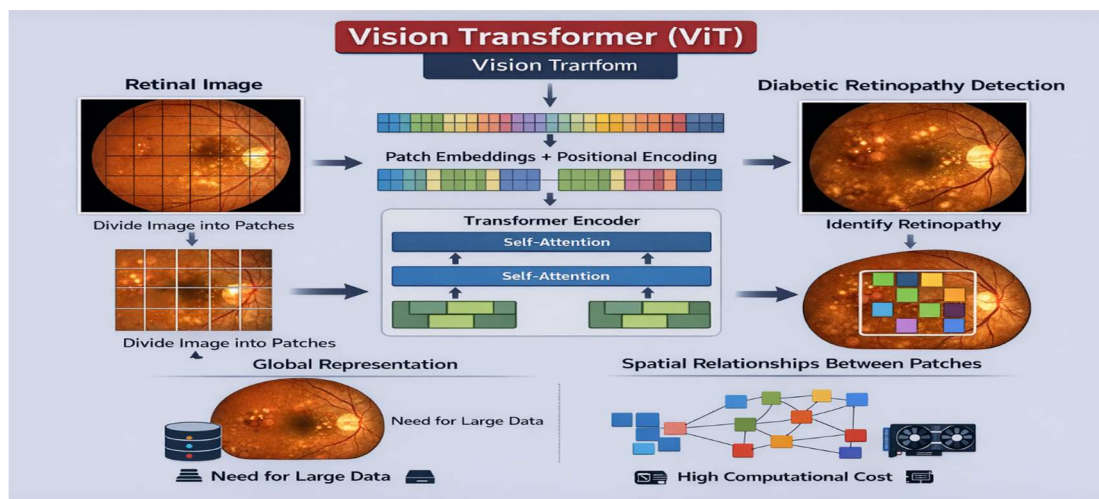


Figure 7: Patch-based self-attention image classification [10]

2-Swin Transformer

The Swin Transformer is a hierarchal structure that improves the efficiency of the traditional transformer structure for computer vision tasks. Instead of just using the regular transformer, the Swin Transformer implements self-attention on a small local window instead of the entire image like a vanilla vision transformer does. Shifting the windows to compute self-attention allows for the network to capture both global and local features and significantly decreases the computational cost of the network. The Swin Transformer is multi-scale hierarchical similar to CNNs, but with the advantages of a transformer network especially with attention mechanisms. It can therefore be useful for many medical imaging tasks including diabetic retinopathy where fine-grained lesions as well as global retinal features need to be captured and classified.[10], as shown n Figure 8:

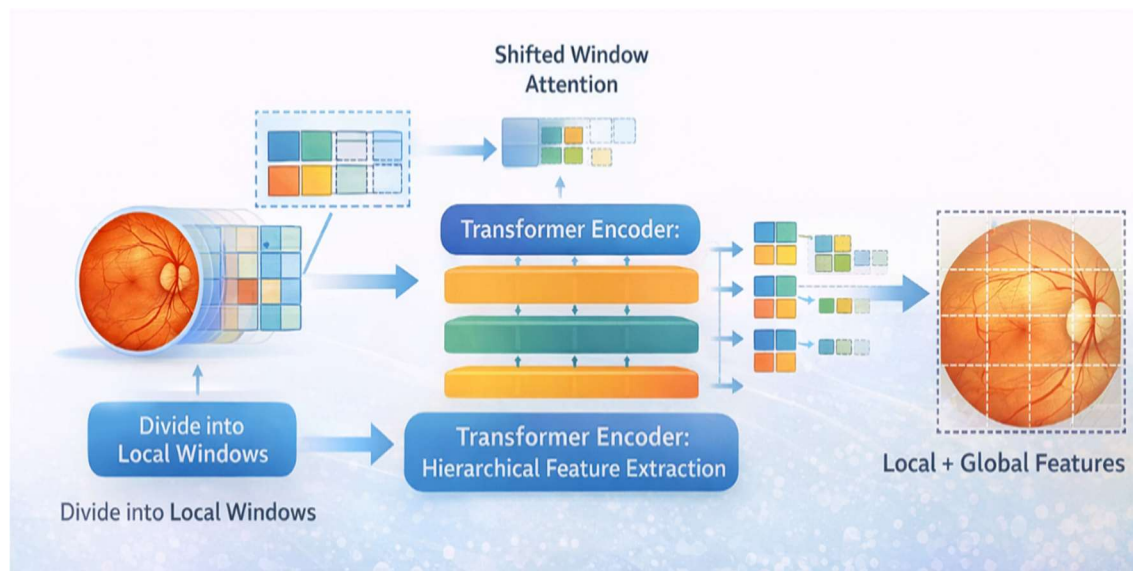


Figure 8: Swin Transformer with shifted window attention [10].

47

3-DeiT (Data-efficient Image Transformer)

DeiT has the transformer architecture that has been adopted from the vision transformers. The DeiT architectural structure has been put in place to facilitate efficiency in learning in the transformers approaches. In the same way as the ViT, the DeiT has the same principle where the input image is first broken down into several patches that are analyzed using the self-attention method. As mentioned above, the DeiT uses the knowledge distillation method which means that the model can use less data in its learning process. This is important considering the fact that the training data in DeiT is less compared to normal images. When the DeiT approach is adopted for the analysis of diabetic retinopathy, it would be very efficient due to this reason. [10], as shown in the Figure below 9:

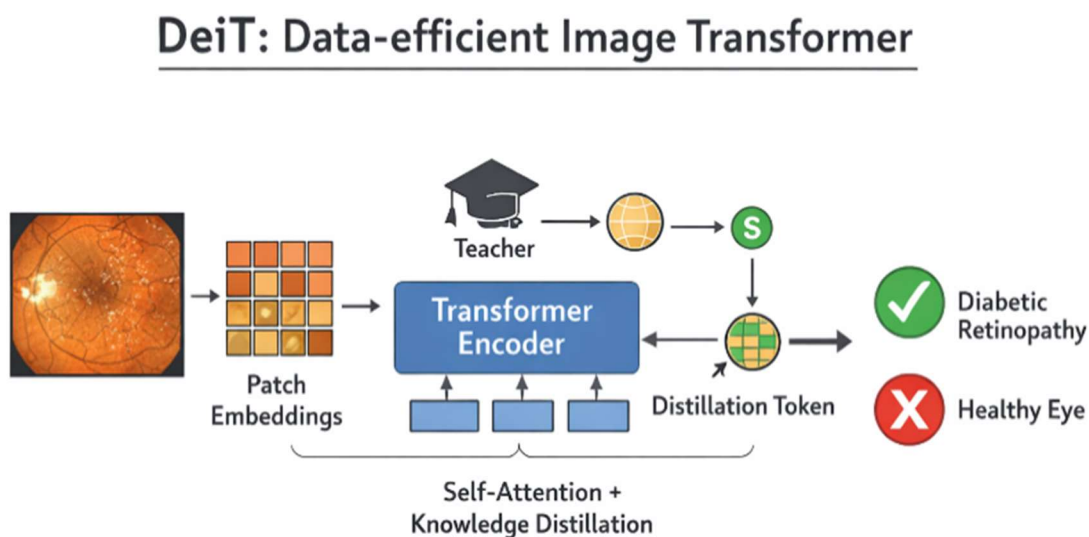


Figure 9: DeiT for efficient learning through knowledge distillation [10].

2.2.6 Performance Metrics

Model evaluation in detecting diabetic retinopathy through machine learning and deep learning approaches is crucial for understanding their performance. Since healthcare data tends to be imbalanced, accuracy can be a deceptive metric due to the potential imbalance. Hence, several measures should be considered.

● Accuracy

Accuracy refers to the ratio of correct classifications to total samples.

Where:

- TP = True Positives
- TN = True Negatives
- FP = False Positives
- FN = False Negatives

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

● Precision

Precision measures the proportion of correctly predicted positive cases out of all predicted positive cases.

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

● Recall (Sensitivity)

Recall measures the proportion of actual positive cases that are correctly identified by the model.

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

● F1-Score

The F1-score is the harmonic mean of precision and recall.

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (4)$$

● Specificity

Specificity measures the ability of the model to correctly identify negative cases.

$$Specificity = \frac{TN}{TN + FP} \quad (5)$$

● Confusion Matrix

The confusion matrix is a table that summarizes classification results in terms of TP, TN, FP, and FN. It helps identify the types of errors made by the model.

● ROC Curve and AUC

The Receiver Operating Characteristic (ROC) curve represents the trade-off between sensitivity and specificity at different threshold values. The Area Under the Curve (AUC) reflects the model's ability to distinguish between classes.

- AUC = 1 indicates perfect classification
- AUC = 0.5 indicates random performance

These metrics are especially important in diabetic retinopathy detection, where both false negatives and false positives have clinical implications.

2.3 Literature Review

Grzybowski et al. [11] team came up with a framework for automatically detecting diabetic retinopathy by utilizing retinal fundus images through deep learning technology. The method focused on enhancing the images through image resizing, image normalization, and augmentation techniques to balance the classes. The model performed with an accuracy rate of around 93.5%.

Akhtar et al. [12] proposed a deep learning model called RSG-Net for diabetic retinopathy grading. Their framework was designed for both four-stage multi-class classification and binary classification using the Messidor-1 dataset. By combining image cropping, Gaussian blur denoising, histogram equalization, resizing, and augmentation, the model achieved very strong results, reporting 99.36% testing accuracy for the four-stage task and 99.37% for the binary task.

Wu et al. [13] introduced mBRSET, a portable retinal fundus image dataset collected in real clinical settings using handheld retinal cameras. Their study benchmarked ConvNeXt V2, DINO V2, and Swin V2 on the dataset and reported best F1-scores of 80.33 for the three-class task and 87.4 for the binary task. This study is important because it highlights the challenge of achieving strong performance on real-world noisy data.

Guefrachi et al. [14] in their study have proposed a diabetic retinopathy screening tool involving multilevel training and application of deep learning methods. For this purpose, transfer learning networks like ResNet152-V2 and InceptionResNet-V2 have been used, with the best results being displayed in GUI. The highest level of accuracy obtained was 96.61%; however, the extremely high accuracy in other networks should not be taken into account seriously.

34 **Youldash et al. [15]** studied six pretrained CNN models, including EfficientNet and RegNet models, on the APTOS and Saudi clinical datasets. The results of this study revealed that binary classification achieved accuracy of 98.6%; however, the multi-class classification has demonstrated rather poor results. The main value of this study lies in clinical external validation.

31 **Shamrat et al. [16]** proposed DRNet13, which is a specifically designed deep neural network for analyzing fundus images and detecting diabetic retinopathy. The DRNet13 architecture surpassed all 15 pretrained CNNs with an approximate accuracy of 96-97%. However, the results were obtained from only one dataset.

17 **Jabbar et al. [17]** proposed a lesion-based hybrid model that which included deep learning for feature extraction and machine learning algorithms such as SVM, Random Forest, Decision Tree, and Linear Regression. This approach showed a high accuracy rate of 94%.

56 **Fayyaz et al. [18]** introduced a hybrid deep learning/machine learning approach where AlexNet and ResNet101 are employed for extracting features, ant colony optimization is utilized to select features, and SVM is used for classification, and they were able to attain an accuracy rate of 93%, indicating that the combination of CNN-based feature extraction and SVM is effective in detecting DR.

18 **Alwakid et al. [19]** developed a deep learning-based classification model incorporating image preprocessing techniques like CLAHE, histogram equalization, and ESRGAN. The improved environment obtained an accuracy of 98.7%, which clearly indicates the significant effect of preprocessing in classification outcomes.

Lin and Wu [20] suggested an altered structure of ResNet-50 for reducing overfitting and enhancing stability. Though this technique was robust, its testing accuracy was found to be poorer than that of several recent architectures

Mohanty et al. [21] evaluated the performance of a hybrid VGG16 and XGBoost model against DenseNet121 in detecting and classifying diabetic retinopathy. Their study concluded that DenseNet121 outperformed the hybrid model, achieving 97.30% accuracy.

Patil et al. [22] employed ResNet-50, transfer learning, and Test-Time Augmentation (TTA) on the APTOS dataset and validated it on EyePACS. They demonstrated an impressive 97.87% accuracy with high Quadratic Weighted Kappa, emphasizing the significance of data augmentation in enhancing the reliability of the classifier.

Ali and Dawood [23] proposed a ResNet-50-based implement transfer learning and classify diabetic retinopathy using the APTOS 2019 dataset. Their system obtained an accuracy of 98.3%, thus confirming the applicability of transfer learning to retinal images.

Das et al. [24] developed the DRFEC framework and evaluated 26 pre-trained deep learning models using EyePACS. Their highest-performing model, EfficientNetB4, obtained 79.11% validation accuracy, but also exposed overfitting problems in some architectures.

Jiwani et al. [25] suggested a simple CNN-based model for classifying binary diabetic retinopathy through the use of the IDRiD database. Despite having less accuracy compared to other models, the contribution is still valuable as a benchmark study

2.4 Comparative Analysis of Related Work

From the reviewed articles, one can make the conclusion that currently, most recent methods for diabetic retinopathy detection are based on using deep learning algorithms, mostly convolutional neural network transfer learning models such as ResNet, DenseNet, EfficientNet, Inception, etc. Nevertheless, it should be stated that even with their high level of performance in most cases they demonstrate, there are some obvious drawbacks of these models. One of the major ones is their reliance on one single public database, making their applicability questionable. Moreover, extremely high accuracy levels reported are achieved through heavy preprocessing or data augmentation and in specific conditions. Also, one can state that currently hybrid models based both on deep learning and machine learning models are significantly less popular and used than those based only on the end-to-end approach. Although it is evident that in terms of classification results they perform better. The next significant trend is that transformers become an alternative to classic convolutional networks in certain tasks, yet requiring higher amount of data and computing power for proper validation. Thus, the need for research in this direction is quite obvious. As presented at table 1:

Table 1: Comparative Study of Diabetic Retinopathy Detection Methods

Authors	Year	Method / Model	Dataset	Results	Limitations
Grzybowski et al. [11]	2026	AI DR devices (CNN)	EU datasets	~80%	Heterogeneous data and reporting
Akhtar et al. [12]	2025	Deep Learning (CNN-based, RSG-Net)	EyePACS	94.2%	Limited generalization
Wu et al. [13]	2025	Deep Learning (ConvNeXt V2, DINOv2, SwinV2)	mBRSET	90%	Limited dataset size
Guefrachi et al. [14]	2024	Deep Learning (ResNet152-V2)	Kaggle DR dataset	97%	Overfitting risk
Youldash et al. [15]	2024	Deep Learning (EfficientNetV2, RegNetX)	Kaggle EyePACS	96%	Class imbalance
Shamrat et al. [16]	2024	Deep Learning (DRNet13)	Fundus images	96%	Computational cost
Jabbar et al. [17]	2024	GoogleNet + ResNet16 + ML classifiers	Fundus dataset	94%	Limited dataset partition/validation detail
Fayyaz et al. [18]	2023	AlexNet + ResNet101 + SVM	Fundus images	93%	Small feature set limits generalizability
Alwakid et al. [19]	2023	DenseNet-121 + Augmentation	APTOS & DDR	98.36 %	Lower performance on some datasets
Lin and Wu [20]	2023	Revised ResNet-50 (CNN)	Kaggle DR images	74.32 %	Low test performance compared to train
Mohanty et al. [21]	2023	ResNet-50	Kaggle DR	74.3%	Low test accuracy

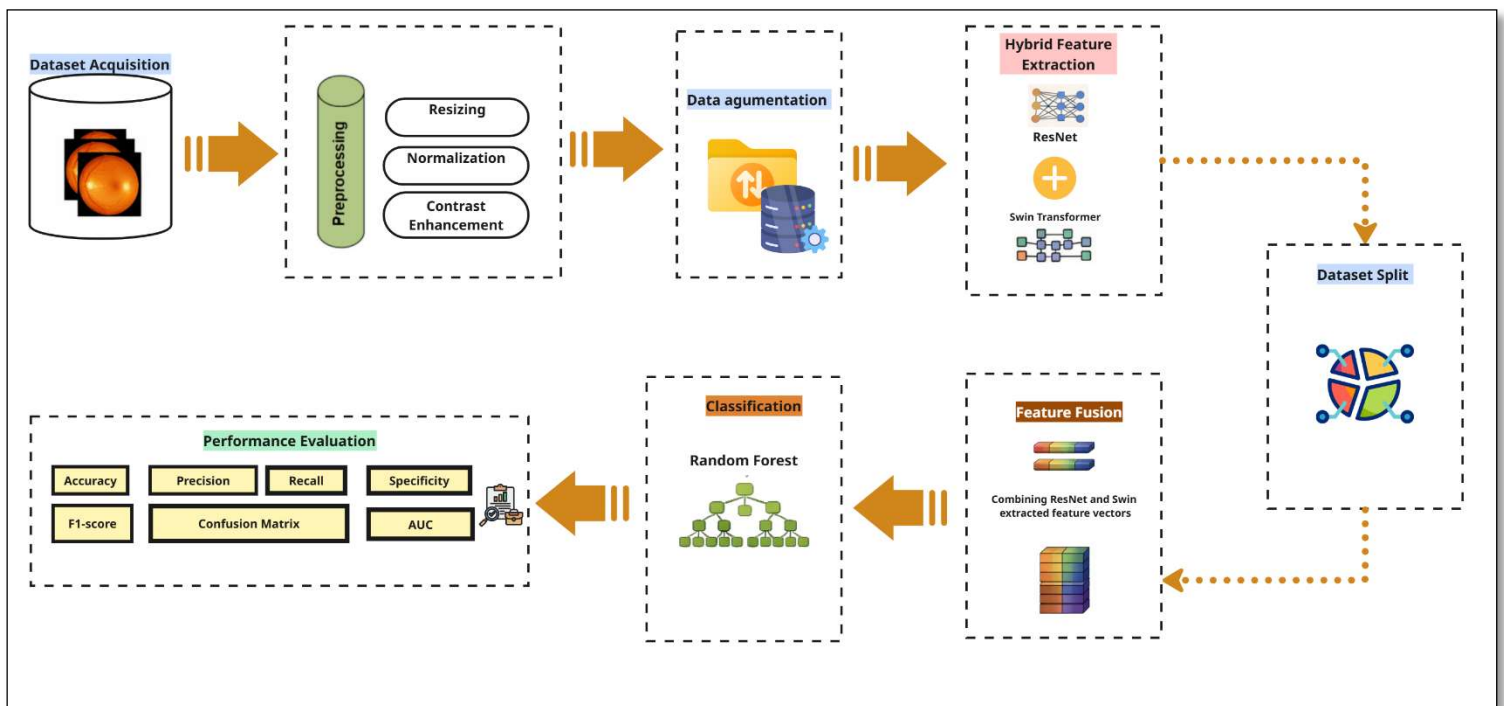
Patil et al. [22]	2023	ResNet-50 + Augmentation (TTA)	APTOS + EyePACS	97.87 %	Multi-class imbalance issue
Ali and Dawood [23]	2023	ResNet-50 (CNN transfer learning)	APTOS2019	98.3%	No external validation
Das et al. [24]	2022	CNN feature extraction & classification (26 DL nets)	Kaggle EyePACS	~79.11 %	Low validation accuracy relative to training
Jiwani et al. [25]	2022	CNN (binary classification)	EyePACS	~75%	Moderate accuracy

2.5 Chapter Summary

This chapter provides the relevant background and related work concerning the topic of detection of diabetic retinopathy through artificial intelligence. It provides an insight into the medical background behind diabetic retinopathy and further describes the use of machine learning, deep learning, and transformers in analyzing retinal images. Moreover, it discusses the common performance metrics involved in the process of classification. The previous literature is also considered in order to identify key findings and trends in this area of research. It turns out that while there are several effective ways of detecting diabetic retinopathy, certain issues are still present. They include generalization problems, dataset reliance, and a lack of investigation into hybrid architectures.

Chapter 3 Proposed Approach (Methodology)

Proposed framework:



8

Chapter 4: Experimental Results and Discussion

Chapter 5 Conclusion and Future Work

References

- 1- Sebastian, A., Elharrouss, O., Al-Maadeed, S., & Almaadeed, N. (2023). A survey on deep-learning-based diabetic retinopathy classification. *Diagnostics*, 13(3), 345.
- 2- Bilal, A., et al. (2024). Computationally efficient deep learning models for diabetic retinopathy detection: A systematic literature review. *Artificial Intelligence Review*.
- 3- Issac, A., et al. (2024). Artificial intelligence for diabetic retinopathy detection: A systematic review. *Informatics in Medicine Unlocked*.
- 4- Agarwal, R., et al. (2024). Advancing diabetic retinopathy diagnosis with fundus imaging: A comprehensive survey of computer-aided detection, grading and classification methods. *Intelligent Medicine*.
- 5- Ikram, A., Imran, A., Li, J., Alzubaidi, A., Ahim, S. F., Yasin, A. Y., & Fathi, H. F. (2024). A systematic review on fundus image-based diabetic retinopathy detection and grading: Current status and future directions. *IEEE Access*, 12, 96273–96303.
- 6- Chen, B., Li, X., Zhang, Y., Wang, Q., & Xie, Z. (2025). Certainty-aware support vector machines for fair classification. *Neurocomputing*, 642, 130180
- 7- Noura, H. N., Chu, T., Allal, Z., Salman, O., & Chahine, K. (2024). A comparative study of ensemble methods and multi-output classifiers for predictive maintenance of hydraulic systems. *Results in Engineering*, 24, 102900.
- 8- Amer, A. A., Badr, N. L., Khattab, H. A., Tolba, M. F., & El-Rashidy, N. M. (2025). Effective k-nearest neighbor models for data classification using the random subspace method. *Journal of Big Data*, 12, 94.
- 9- Bappi, M. D. I., Jahan, A., Razzak, M. I., & Hassan, M. M. (2025). Deep learning-based diabetic retinopathy recognition and classification from fundus images

using spatial texture-aware multi-scale feature fusion. *ICT Express*. Advance online publication.

- 10- AlMohimeed, A., Alqahtani, A., Alzahrani, M., Alghamdi, B., Alshammari, F., Alanazi, M., Alotaibi, N., Alghamdi, H., Mneimneh, M. A., & Bchir, O. (2025). Multi-stage framework using transformer models, feature extraction and machine learning for diabetic retinopathy detection. *Scientific Reports*, 15, 27830.
- 11- Grzybowski, A., & Jin, K. (2026). Artificial intelligence-based medical devices for diabetic retinopathy screening in the European Union. *Ophthalmology and Therapy*, 15(2), 569–589.
- 12- Akhtar, S., Aftab, S., Ali, O., Ahmad, M., Khan, M. A., Abbas, S., & Ghazal, T. M. (2025). A deep learning based model for diabetic retinopathy grading. *Scientific Reports*, 15, 3763.
- 13- Wu, C., Restrepo, D., Nakayama, L. F., Zago Ribeiro, L., Shuai, Z., Barboza, N. S., Sousa, M. L. V., Fitterman, R. D., Pereira, A. D. A., Regatieri, C. V. S., Stuchi, J. A., Malerbi, F. K., & Andrade, R. E. (2025). A portable retina fundus photos dataset for clinical, demographic, and diabetic retinopathy prediction. *Scientific Data*, 12, 323.
- 14- Guefrachi, S., Ehtioui, A., & Hamam, H. (2024). Diabetic retinopathy detection using deep learning multistage training method. *Arabian Journal for Science and Engineering*, 49(2), 1079–1096.
- 15- Youldash, M., Rahman, A., Alsayed, M., Sebiany, A., Alzayat, J., Aljishi, N., Alshammari, G., & Alqahtani, M. (2024). Early detection and classification of diabetic retinopathy: A deep learning approach. *AI*, 5(4), 2586–2617.
- 16- Shamrat, F. M. J. M., Shakil, R., Sharmin, S., Ovy, N. H., Akter, B., Ahmed, M. S., Ahmed, M. Z., Ahmed, K., Bui, F. M., & Moni, M. A. (2024). An advanced deep neural network for fundus image analysis and enhancing diabetic retinopathy detection. *Healthcare Analytics*, 5, 100303.

- 17- Jabbar, A., Liaqat, H. B., Akram, A., Sana, M. U., Azpíroz, I. D., de la Torre Díez, I., & Ashraf, I. (2024). A lesion-based diabetic retinopathy detection through hybrid deep learning model. *IEEE Access*, 12, 40019–40036.
- 18- Fayyaz, A. M., Sharif, M. I., Azam, S., Karim, A., & El-Den, J. (2023). Analysis of diabetic retinopathy (DR) based on deep learning. *Information*, 14(1), 30.
- 19- Alwakid, G., Gouda, W., Humayun, M., & Jhanjhi, N. Z. (2023). Deep learning-enhanced diabetic retinopathy image classification. *Digital Health*, 9, 20552076231194942.
- 20- Lin, C.-L., & Wu, K.-C. (2023). Development of revised ResNet-50 for diabetic retinopathy detection. *BMC Bioinformatics*, 24, 157.
- 21- Mohanty, C., Mahapatra, S., Acharya, B., Kokkoras, F., Gerogiannis, V. C., Karamitsos, I., & Kanavos, A. (2023). Using deep learning architectures for detection and classification of diabetic retinopathy. *Sensors*, 23(12), 5726.
- 22- Patil, M. S., Chickerur, S., Abhimalya, C., Naik, A., Kumari, N., & Maurya, S. (2023). Effective deep learning data augmentation techniques for diabetic retinopathy classification. *Procedia Computer Science*, 218, 1156–1165.
- 23- Ali, A. A., & Dawood, F. A. A. (2023). Deep learning of diabetic retinopathy classification in fundus images. *Journal of Engineering*, 29(12), 139–152.
- 24- Das, D., Biswas, S. K., & Bandyopadhyay, S. (2023). Detection of diabetic retinopathy using convolutional neural networks for feature extraction and classification (DRFEC). *Multimedia Tools and Applications*, 82(19), 29943–30001.
- 25- Jiwani, N., Gupta, K., & Afreen, N. (2022). A convolutional neural network approach for diabetic retinopathy classification. In *2022 IEEE 11th International Conference on Communication Systems and Network Technologies (CSNT)* (pp. 357–361). IEEE.