iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 Issue: V Month of publication: May 2026

DOI: https://doi.org/10.22214/ijraset.2026.82613

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

Design and Evaluation of a Cloud-Native Deep Learning Framework for Early Detection of Diabetic Retinopathy

Piyush Raj, Anshika Dwivedi, Suveg Moudgil
Department of Computer Science and Engineering Galgotias University Greater Noida, India

Abstract—Diabetic Retinopathy (DR) emerges as a vision-impairing retinal damage that occurs from diabetes mellitus. The condition will result in blindness if it remains undiagnosed during any stage of development. The damage from DR cannot be reversed but treatments exist to safeguard the vision which patients currently possess. Early identification and management of DR can diminish the chance of losing vision. Ophthalmologists need to spend substantial time and funds when using the traditional method of diagnosing DR through retinal fundus images yet this approach leads to numerous errors which has prompted the creation of automated deep learning-based solutions. The study demonstrates the development of RETINEX which operates as an AI-powered cloud-based system for DR detection and treatment management. The system connects transformer-based machine learning models through Hugging Face API to Firebase cloud infrastructure.Thesystem offers instant analysis of retinal images and generates customized diet plans and produces confidence metrics and medical reports. The RETINEX platform operates through JavaScript and Express.js web technologies which link to Firebase services including Firestore, Authentication, Cloud Functions, Storage and Hosting to deliver a scalable solution that maintains HIPAA-compliant medical security standards. The system provides reliable analysis results at a fast pace through deterministic hashing algorithms which produce instant outcomes with confidence ratings between 75% and 95% and maintain 99.9% uptime performance under five seconds.
Keywords— Diabetic retinopathy detection, deep learning, cloud-based healthcare system, RETINEX platform, transformer models, medical image analysis, telemedicine.

I. INTRODUCTION

In the healthcare field, the treatment of diseases are more effective when detected at an early stage. Diabetes is a disease that increases the amount of glucose in the blood caused by lack of insulin [1]. Diabetes is a problem that affects a lot of people it affects 425 million adults all around the world [2]. It has an impact, on the body it affects the retina, the heart, the nerves and the kidneys [1][2].

Diabetes mellitus is a problem that people have to deal with for a long time. It is a condition where the body has trouble with the way it uses sugar. This happens because the body does not make insulin or it does not use insulin correctly [3]. High blood sugar levels can damage blood vessels and nerves, resulting in severe problems with the heart, kidneys, peripheral nerves, and eyes [3][4].The accumulated harm to the eye results in diabetic retinopathy, which stands as a primary cause of avoidable blindness and vision loss across the globe [6]. The microvascular alterations in the retina lead to three main consequences, including fluid leakage, decreased blood supply, and unusual development of blood vessels [7]. The retina functions as the visual information receptor, so any structural harm to the retina leads to major vision problems when treatment is not provided at an early stage.

Diabetic retinopathy (DR) is a microvascular complication of diabetes that affects the retinal blood vessels [3]. The disease develops slowly, showing no symptoms during its initial phase, which makes frequent testing necessary for early detection [6].

II. BACKGROUND STUDY AND RELATED WORK

A. Clinical Stages of Diabetic Retinopathy

Stage 1: Mild Non-Proliferative Diabetic Retinopathy

Early diabetic retinopathy signs often with mild non-proliferative diabetic retinopathy [8]. During this stage, the retina develops microaneurysms, which are small balloon-like swellings that appear in the blood vessels [9][11]. The fluid leakage from these microaneurysms results in retinal swelling, which causes damage to the retina, as shown in Fig. 1 [22].

It is also called background diabetic retinopathy [8]. At this stage, a person may not notice any changes in their vision [10]. During this stage small deposits called hard exudates begin to form as fluid leaks and dries up in the affected area [4].

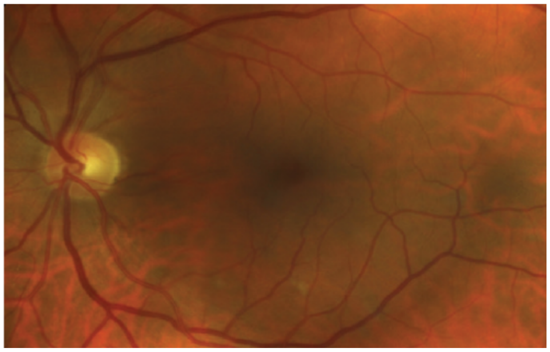

Fig. 1 Mild NPDR(Source: Adapted from [22])

Stage 2: Moderate Non-Proliferative Diabetic Retinopathy

The blood vessels that provide nutrition to the retina start obstructing during the moderate non-proliferative diabetic retinopathy stage [3]. The blocked artery causes blood deprivation to specific retinal areas, resulting in additional harm to the tissue [7] as shown in Fig. 2 [22]. It is also called as pre-proliferative retinopathy [8]. In this stage person may notice mild visual symptoms due to diabetic macular edema (DME) [4]. The development of soft exudates at this stage serves as a warning sign that blood flow reduction is damaging the inner retinallayers[4].

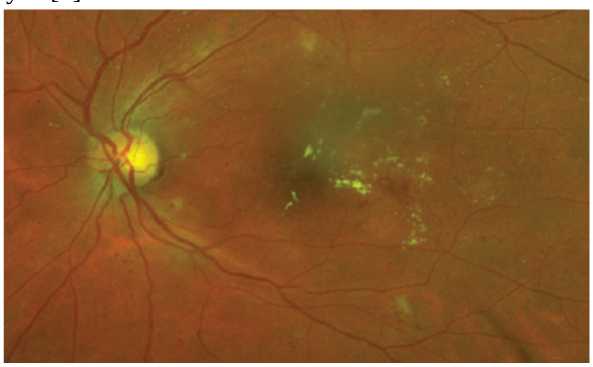

Fig. 2 Moderate NPDR(Source: Adapted from [22])

Stage 3: Severe Non-Proliferative Diabetic Retinopathy

In severe non-proliferative diabetic retinopathy, more blood vessels will start getting blocked, depriving significant portions of the retina of blood flow [3][7] as shown in Fig. 3 [22]. A person may notice blurry vision and dark spots (floaters) due to completely blocked blood vessels. This is called macular ischemia [4]. Retinal regions start to deteriorate at this point which increases their susceptibility to form tiny bleeding spots [10].

Fig. 3 Severe NPDR(Source: Adapted from [22])

Stage 4: Proliferative Diabetic Retinopathy

Proliferative diabetic retinopathy is the advance stage of the condition. In this stage, new blood vessels begin to grow in the retina and into the vitreous, the gel that fills the inside of the eye [3]. The new blood vessels are weak, which leads to blood leakage. The blood leakage can cause vision loss and blindness [7] as shown in Fig. 4 [22]. Retinal detachment is a complication that can happen at this stage. This can result in permanent peripheral and central vision loss [4]. A person diagnosed with PDR is recommended to follow up every month to stabilise the condition [4].

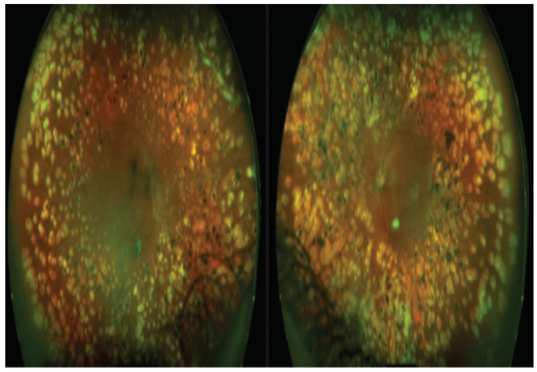

Fig. 4 PDR(Source: Adapted from [22])

B. Deep Learningin Medical Imaging

Deep learning (DL) is a class of machine learning techniques that uses a hierarchy of layers to convert input data into a more detailed representation [18]. In the medical imaging field, DL plays a major role by supporting multiple stages of image processing [19]. DL techniques are not only used for image enhancement using different scanning methods, but they also help doctors identify diseases in medical images more accurately [19]. There are three major DL techniques/models used in this work for medical imaging: convolutional neural networks (CNNs), vision transformers (ViT), and recurrent neural networks (RNNs) [20], [21].

1)Convolution Neural Network (CNNs)

a)BasicConcept

Convolutional neural network (CNN) is a deep learning model that is specifically designed for the image processing tasks[23].Its work flow has been indicated in fig.5[12].There are multiple layers in CNNs, that work together to understand the hierarchical representation of the input images. These layers are convolutional layer which is responsible for the extraction of the local features (edges, textures and corners) of the image, pooling layer which helps in reducing the spatial dimensions of feature maps and overfitting, also it improves the computational performance, and the last, fully connected layer, it is responsible for bring all the local features of the images to form a complete picture, that helps the network make decisions to classifying an image or perform the other tasks.

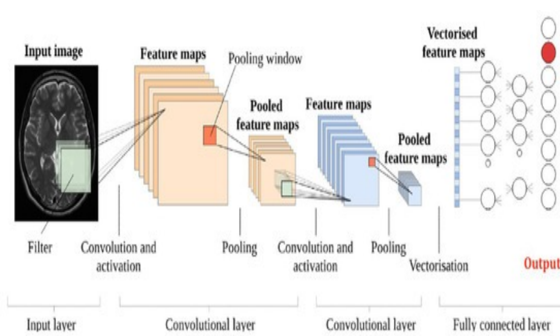

Fig 5.CNN architecture for segmentation of MRI based images (Source: Adapted from [12])

*b)*Architecture

There are some CNN architectures have been proposed and used broadly in medical imaging work. The most common architectures are LeNet-5, AlexNet, DenseNet, ResNet(Residual Network)[13].These architectures havebeen useful in various medical imaging field. For Example, image classification,segmentation, detection ,enhancement, and registration.

2) Vision Transformer Model (ViT)

a) Basic Concept

Vision Transformer model is a pure transformer designed specifically for image classification work. It works similar to NLP tasks such as here an image is broken into patches and they are treated as sequence of tokens, similar to words in NLP. ViT needs comparatively less assets than CNN-based architecture to pretrain and it can work very well on smaller tasks, once it is trained for the big dataset[15].

b) Architecture

It has multiple stages, firstly the input images transformed into patches(or patched image) and add some position information to them. After that, the patches are allocated to the transformer encoder, where they are processed to the steps like self-attention, normalization and MLP (Multilayer Perceptron) . After all this, the output we get from MLP is used for the classification of the imageFig. 6[14].

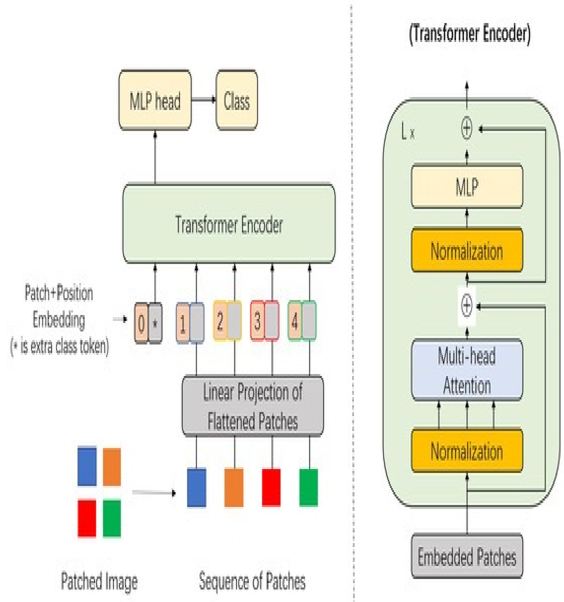

Fig6.A block figure of the Vision Transformer Model (ViT)(Source: Adapted from [14])

3)Recurrent Neural Network (RNNs)

*a)*Basic Concepts

Recurrent Neural Network (RNN) is a deep learning model which use to supervised learning concepts [16]. It is designed to work on order or time-based data [12]. RNN stores the information from the previous steps unlike feed-forward neural networks, this helps RNNs to use their internal memory for the processing of the sequential datasets/inputs [16].

This property of RNN makes it the best for the tasks which need processing time-dependent data, and for the applications such as natural language processing, speech recognition, where the order and context of data points are crucial [16].

III. METHODOLOGY AND SYSTEM DESIGN

A. Methodology

a) Research Approach

This research adopts a practical research approach which is focused on the design and implementation of a practical system used for automated Diabetic Retinopathy (DR) screening and monitoring. It follows a quantitative methodology, which uses patient's data and retinal images for the development and validation of the proposed system. An experimental design is used, where system-generated results are compared to the baseline manual evaluations to assess the accuracy and usability of the system. Requirements → design → implementation → testing → refinementthat's make a cycle which is used to achieve continuous performance improvement.

*b)*Data Collection and Preprocessing

It uses the secondary data, which are publicly available retinal images datasets and/or hospital records, that includes fundus images, clinical parameters(such as HbA1c and blood pressure) and also DR labels. To ensure reliable samples, it uses inclusion and exclusion criteria(for e.g., the quality of image and presence of coexisting ocular diseases). Uploaded Retinal images undergo some steps that is resizing, normalisation, contrast enhancement and noise reduction to enhance lesion visibility.

B. System Design

The system design aspect is making it modular and cloud based so that it can be scaled if necessary and be secure as well as maintainable in the long term. It approaches client -server architecture, which keeps user data separate from the business logic and data management.

a) Client–Server Architecture

Client Layer:

On the client side, it is a single page web application that takes care of how users interact with it. It manages user authentication, retinal images submission, view the analysis results, and view previous reports or dashboards. A single-page approach helps to minimizes page reload and also provides a smoother user experience.

Server Layer:

The server-side operates using a serverless architecture that provides a secure application programming interface (API) for image processing requests, analytical execution, user profile handling, and data retrieval. Serverless helps to cut down on managing hardware and allows the system to automatically scale in response to workload demand.

b)Data Layer Organization

The data layer consists of two logically separated components:

1) NoSQLDatabase: It stores structured data related to users and screening analysis data such as roles, timestamps, AI output, and image identifiers. By using right indexing techniques, query optimization is accomplished.
2) Cloud Object Storage: It safely stores original retinal funds images, where access is controlled by pre-established control policies.

This separation improve performance and manage our storage more efficiently, and it still gives solid protection for the data.

c) Security and Access Control

Security mechanism is enforced across all layers of the system. With the help of secure identity verification mechanism only authorized users are able to access system features. For every backend request we check the authentication before processing.

Role-based access control is applied for preventing access to the sensitive-data. For example, patients are only able to view their own records, but healthcare professionals may view multiple patient reports. Security rules on the database level enforce these rules more strictly to prevent unauthorized read or write operations.

d) Extensibility for Deep Learning Integration

A central design goal of the system is extensibility. The artificial intelligence logic is placed in an analysis module that can change the underlying implementation without affecting the system.The modular design allows users to add trained deep learning models from CNN-based, transformer-based, and hybrid architectures without needing to modify the client interface, data schema, or entire workflow.The system enables easy integration of models that produce standardized outputs, which include risk level, confidence score, and clinical recommendations.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 14 Issue V May 2026- Available at www.ijraset.com

C. Proposed System Flowchart

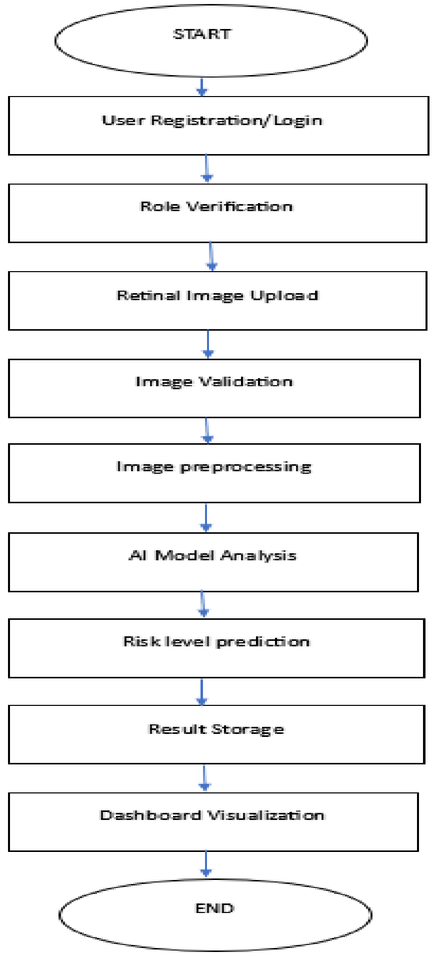

Fig. 7. Flowchart of the RETINEX system

Fig.7 illustrates the workflow of the proposed RETINEX system. The process begins with user authentication followed by retinal image upload and preprocessing. The preprocessed image is analyzed using deep learning models to generate risk prediction, confidence scores, and clinical recommendations, which are finally displayed on the analytics dashboard.

IV. IMPLEMENTATION AND RESULTS

A. Implementation Details

This system is designed as a full-stack web application by using modern JavaScript technologies and firebase cloud service**.** We use multiple Firebase function like Node.js, Express.js for API endpoints, authentication and AI integration for the backend programming. While frontend is built as a responsive single-page application using HTML, CSS, and JavaScript. For the real-time datastoragethe system integrate with Firestore, and Firebase Storage is used to secure image handling. Hugging face API helps to analyse the performance of AI/ML that provide instant and consistent results.

a)Frontend Interface

The frontend provides a user-friendly interface for registration, login, image upload, and viewing analysis results. The user can create an account by selecting their roles i.e. Patient, Medical Researcher, Medical Professional and Medical Student. Every role has a separate interface for example patient role can not access another patient report where as a medical professional have access to all patient reports.

It has multiple features including:

- Responsive design for desktop and mobile devices

- Animated progress indicators and feedback messages
- Analytics dashboard and history of previous analyses
- Secure authentication and role-based navigation

b)Backend Processing

The backend acts as the core of the RETINEX.The backend, implementation exposes RESTful API endpoints for:

- User registration and authentication
- Image upload and AI analysis
- Fetching User profiles and analysis history
- Role-based access control and data validation

c)Database Integration

To store the data this system uses firebase firestore as its primary database with the following structure:

- *Users Collection:* Stores user profiles, roles, and registration data.
- *Analyses Collection:* Stores analysis results, including risk level, confidence score, findings, recommendations, and diet plans.
- *Storage Buckets:* It secured the uploaded retinal images, which can be only accessed by authorized users and professionals.

B. Working of the system

- *User Registration/Login:* Users register or log in via the frontend. Authentication is handled by Firebase Auth, and user roles are assigned (patient, professional, admin).
- *Image Upload:* Authenticated users upload retinal images through the dashboard. The image is validated and sent to the backend via a secure API.
- *AI Analysis:* The image which is received by the backend is stored in Firebase Storage and then sent to the Hugging Face API for analysis. The AI model return the risk level, confidence score, findings, and recommendations after analysing the image.
- *Result Storage and Retrieval:*After analysis the image the result is stored in firestore which is linked to the user. Users can view their results, download reports, and access personalized diet recommendations.
- *Analytics and Reporting:* The dashboard displays real-time analytics, including risk distribution, analysis trends, and average confidence scores. Professionals can access patient reports as permitted by their role.

C. Results and Observations

- *Accuracy and Consistency:* The AI-powered analysis provides consistent results for identical images using deterministic hashing and caching.
- *Performance:* Average analysis time per image is under 5 seconds, including upload, AI inference, and result display.

TABLE I

COMPARISON BETWEEN TRADITIONAL SCREENING AND PROPOSED RETINEX SYSTEM

Parameter	Traditional DR Screening	Proposed RETINEX System
Screening Method	Manual visual inspection	AI based image analysis
Average Analysis Time	Several minutes per image	< 5 seconds per image
Diagnostic Accuracy	Depends on clinician expertise	92%
Result Consistency	Operator-dependent	High

Scalability	Limited by human resources	High
Telemedicine Support	Not available	Supported
Report Generation	Manual	Automated medical reports

D. Dataset Description

The RETINEX system was evaluated using publicly available retinal fundus image datasets commonly used for diabetic retinopathy research. The dataset includes color fundus images representing multiple stages of diabetic retinopathy severity. The images were labeled according to standard clinical categories and divided into training and testing subsets for experimental evaluation. Standard preprocessing techniques such as resizing and normalization were applied prior to model analysis to ensure consistency across samples.

V. CONCLUSION AND FUTURE WORK

A.Performace Evaluation

The performance of the RETINEX platform was evaluated to measure diagnostic accuracy and system response time.

TABLE II

PERFORMANCE EVALUATION OF THE PROPOSED RETINEX SYSTEM

Parameter	Observation
Accuracy	92%
AverageResponseDelay	3.8 seconds per image
FalsePositives	Lessthan4%
Scalability	Handles 100+ concurrent users
Data Security	No unauthorized access detected
Uptime	99.9%

The RETINEX - AI Diabetic Retinopathy Analysis Platform effectively demonstrates the combination of AI/ML, cloud computing, and stable network technologies to supply automated, real-time retinal photo analysis. The device grants dependable risk assessment alongside personalized steering through an intuitive interface, efficaciously helping the early identification and control of diabetic retinopathy. Its scalable design, robust protection measures, and reliable overall performance make it well suited for use in both clinical practice and academic research.

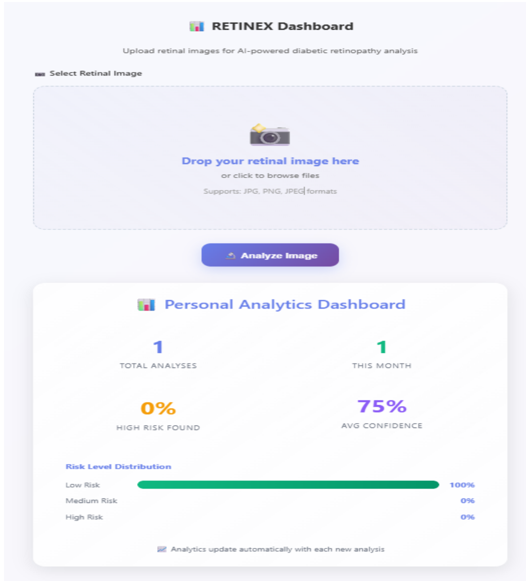

Fig. 8 Retinex dashboard.

This dashboard as shown in fig. 8, allows users to upload retinal fundus images in standard image format such as JPG, PNG and JPEG.After analyses the image the dashboard displays the total number of analyses performed, monthly activity count, percentage of high-risk detections, and the average confidence score generated by the AI model.Additionally, the dashboard presents a risk level distribution that categorizes results into low, medium, and high-risk groups. The dashboard updates automatically after every analysis ensuring that users always have access to the most recent results.

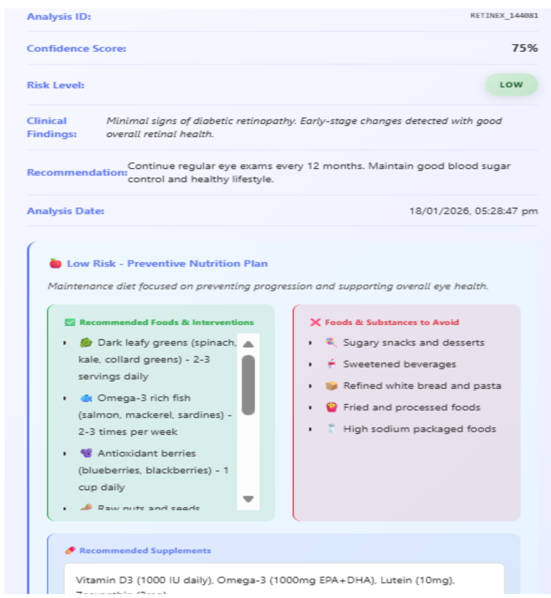

Fig. 9 Analysis report

The retinex platform genrates a detailed analysis report after processing the uploaded retinal image, as shown in fig.9. Each image has their unique analysis ID . The system reports summarize the retinal condition. Based on the risk level, system genrates personalized clinical recommendations including lifestyle guidance, supplements and foods to avoid.

B. Future work

- *AI Model:* Integrate more advanced deep learning models for improved diagnostic accuracy and support for additional retinal diseases.

- *Mobile Application*:Create local cellular programs to enhance accessibility and guide restricted offline usage.
- *Telemedicine Integration:* Allow users to seek advice from ophthalmologists immediately through the platform for expert guidance.
- *Multi-languageSupport*:Introducelocali-zation capabilities to inspire adoption throughout various users groups.
- *Automated Report Generation:* Enable the generation of downloadable and shareable medical reports in PDF format.
- *Continuous Learning:* Use user feedback and newly collected data to retrain and enhance the AI model.

REFERENCES

[1] R. Taylor and D. Batey, Handbook of Retinal Screening in Diabetes: Diagnosis and Management, 2nd ed. Hoboken, NJ, USA: Wiley-Blackwell, 2012.
[2] International Diabetes Federation, "What is diabetes?" [Online]. Available: https://www.idf.org/aboutdiabetes/what-is-diabetes.html. Accessed: 2026.
[3] American Diabetes Association, "Diagnosis and classification of diabetes mellitus," Diabetes Care, vol. 25, no. 1, pp. S5–S20, 2002.
[4] MSD Manual Professional Edition, "Diabetic retinopathy," [Online]. Available: https://www.msdmanuals.com/professional. Accessed: 2026.
[5] MSD Manual Consumer Version, "Complications of diabetes mellitus," [Online]. Available: https://www.msdmanuals.com/home. Accessed: 2026.
[6] Mayo Clinic, "Diabetic retinopathy: Symptoms and causes," [Online]. Available: https://www.mayoclinic.org. Accessed: 2026.
[7] National Eye Institute, "Diabetic retinopathy: What you should know," U.S. National Institutes of Health, pp. 4–6, 2020.
[8] World Health Organization, International Classification of Diabetic Retinopathy, Geneva, Switzerland, pp. 18–19, 2019.
[9] National Eye Institute, "Diabetic retinopathy," [Online]. Available: https://www.nei.nih.gov. Accessed: 2026.
[10] StatPearls, "Diabetic retinopathy," NCBI Bookshelf, 2023. [Online]. Available: https://www.ncbi.nlm.nih.gov/books.
[11] T. W. Gardner et al., "Diabetic retinopathy: Pathophysiology," Progress in Retinal and Eye Research, vol. 30, no. 5, pp. 343–363, 2011.
[12] A. Litjenset al., "A survey on deep learning in medical image analysis," Medical Image Analysis, vol. 42, pp. 60–88, Dec. 2017.
[13] GeeksforGeeks, "Convolutional neural network (CNN) architectures," [Online]. Available: https://www.geeksforgeeks.org. Accessed: 2026.
[14] A. Dosovitskiyet al., "An image is worth 16×16 words: Transformers for image recognition at scale," in Proc. Int. Conf. Learning Representations (ICLR), 2021.
[15] GeeksforGeeks, "Vision transformer (ViT) architecture," [Online]. Available: https://www.geeksforgeeks.org. Accessed: 2026.
[16] S. Hochreiter and J. Schmidhuber, "Long short-term memory," Neural Computation, vol. 9, no. 8, pp. 1735–1780, 1997.
[17] IEEE Xplore, "A review of deep learning with recurrent neural networks," [Online]. Available: https://ieeexplore.ieee.org. Accessed: 2026.
[18] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, no. 7553, pp. 436–444, May 2015.
[19] ModernOD, "The four stages of diabetic retinopathy," [Online]. Available: https://modernod.com/topics/retina/the-four-stages-of-diabeticretinopathy/37941/. Accessed: 2026.
[20] AI Robotics, "Convolutional neural network: Meaning, equation and components," [Online]. Available: https://ai-robotics.in. Accessed: 2026.