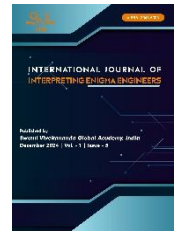




# DIABETIC RETINOPATHY DETECTION USING FEDERATED LEARNING AND VISION TRANSFORMERS



Dr. M. Swapna<sup>1\*</sup>, Tanishqa Ravirala<sup>2</sup>, Nikhitha Reddy<sup>2</sup>

Original Article

<sup>1\*</sup> Associate Professor, Dept. of Computer Science and Engineering, Stanley College of Engineering and Technology for Women, Hyderabad

<sup>2</sup> Student, Final Year, Dept. of Computer Science and Engineering National Institute of Technology, Warangal

\*Corresponding Author's Email: [mswapna@stanley.edu.in](mailto:mswapna@stanley.edu.in), [rt21csb0b47@student.nitw.ac.in](mailto:rt21csb0b47@student.nitw.ac.in), [mn21csb0b34@student.nitw.ac.in](mailto:mn21csb0b34@student.nitw.ac.in)

## Abstract

Early detection of Diabetic retinopathy (DR) is essential for preventing blindness because this disease currently stands as the primary cause of blindness. The implementation of traditional deep learning models becomes difficult because data privacy risks alongside insufficient dataset availability among different healthcare institutions. We introduce an FL-ViT cooperative model training framework which supports distributed information processing without requiring shared clinical data ownership. The use of ViTs in self-attention feature extraction from retinal images together with FL technology guarantees privacy standards compliance. Healthcare operators use a Model designed for the sector to produce diagnostic documents that speed up clinical operations. The system achieves 93 percentage accuracy, ROC curve is 0.89, precision and recall and confusion matrix in DR grading according to APTOS dataset evaluations. The solution resolves issues pertaining to AI scalability and generalizability along with compliance with healthcare ethical requirements.

**Keywords:** Diabetic Retinopathy, Federated Learning, Vision Transformers, Large Language Models, Medical Image Analysis, Data Privacy.

## Introduction

### 1.1 Diabetic Retinopathy

As with diabetic retinopathy (DR), diabetic retinopathy is caused due to hyper-glycemia and damages retinal blood vessels leading to micro aneurysms, hemorrhages, and macular edema. (Barkmeier et.al.,2021)The blind eye occurs in 40 percentage of diabetic patients and is incurable in late stage. Early detection requires automated solutions, which is labor intensive, and thereby manual screening.

Diabetic Retinopathy (DR) is a progressive retinal disease caused by chronic hyperglycemia, leading to vascular abnormalities such as microaneurysms, hemorrhages, and macular edema. DR remains the leading cause of blindness among diabetic patients, affecting approximately 40% of individuals with prolonged diabetes. Early diagnosis is essential to prevent irreversible vision loss; however, conventional screening methods are labor-intensive, time-consuming, and require specialized ophthalmologists. This necessitates automated DR detection systems powered by deep learning models to enhance diagnostic efficiency and accessibility.

Traditional deep learning models for DR detection rely on centralized training approaches, where patient data is aggregated into a common repository. However, this paradigm poses significant challenges related to data privacy, security, and accessibility due to regulatory constraints such as HIPAA and GDPR. Federated Learning (FL) offers a decentralized alternative, allowing multiple institutions to collaboratively train a global model while retaining patient data locally. However, FL-based DR detection systems often suffer from heterogeneity in data distribution, communication overhead, and inefficient model architectures.

In response to these limitations, this paper proposes a novel FL-ViT (Federated Learning with Vision Transformers) framework to enhance DR detection while ensuring data privacy. Vision Transformers (ViTs) have demonstrated superior performance over Convolutional Neural Networks (CNNs) in medical imaging tasks by capturing long-range dependencies and global feature representations. By integrating ViTs into an FL setting, the proposed framework enhances classification accuracy, model generalization, and computational efficiency in distributed healthcare environments.

## 1.2 Key Contributions

This study introduces an advanced FL-ViT framework to address existing challenges in DR detection. The major contributions of this work include:

- **FL-ViT Integration:** A privacy-preserving federated learning framework that incorporates Vision Transformers for DR classification without requiring centralized data storage.
- **Optimized Model Training:** Implementation of parameter-efficient ViT architectures to reduce computational costs while maintaining high diagnostic accuracy.
- **Personalized Federated Learning:** Utilization of adaptive client selection and federated averaging strategies to address data heterogeneity across multiple healthcare institutions.
- **Automated Clinical Reporting:** Integration of Large Language Models (LLMs) to generate interpretable diagnostic reports based on ViT-based DR classification.
- **Comprehensive Performance Evaluation:** Extensive experimental validation using the APTOS dataset, including accuracy, precision-recall, ROC-AUC analysis, and confusion matrix comparisons with conventional deep learning models.
- **Privacy and Security Compliance:** Implementation of secure aggregation mechanisms, including differential privacy and multi-party computation, to protect sensitive patient information.

## 1.3 Paper Organization

The remainder of this paper is structured as follows:

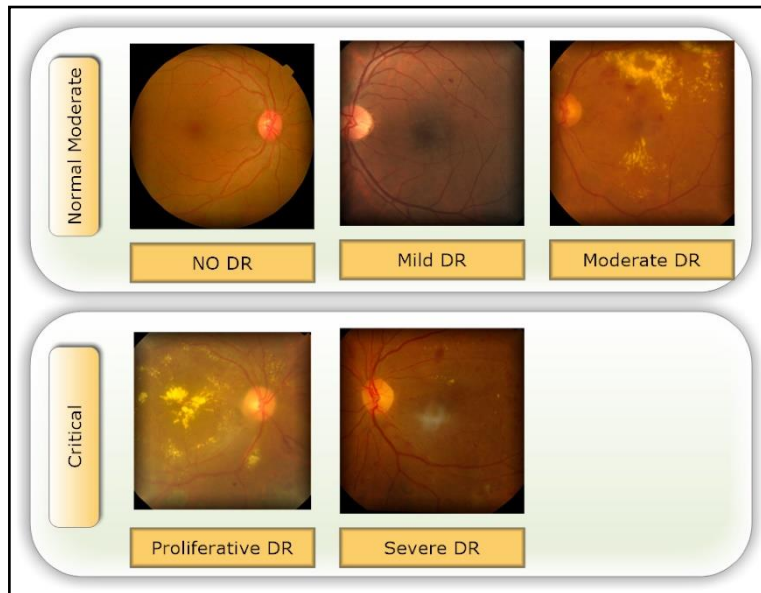
- **Section 2: Related Work** provides an overview of prior research on deep learning-based DR detection, federated learning in medical imaging, and the application of Vision Transformers in healthcare.
- **Section 3: System Model and Problem Statement** defines the problem, outlines the proposed FL-ViT framework, and discusses challenges associated with decentralized DR detection.
- **Section 4: Methodology** describes the data collection, preprocessing pipeline, local training process, federated model aggregation, and ViT-based DR classification strategies.
- **Section 5: Experimental Results** presents quantitative and qualitative analyses of model performance, including accuracy trends, ROC curves, precision-recall trade-offs, and confusion matrices.
- **Section 6: Challenges and Future Work** discusses computational constraints, potential scalability issues, and directions for future research in adaptive FL-ViT architectures.
- **Section 7: Conclusion** summarizes the key findings and highlights the implications of FL-ViT for real-world ophthalmology applications.

## 1.4 Federated Learning

Medicare benefits from Federated Learning (FL) because it supports collaborative model training between different healthcare institutions without requiring centralized data sharing. (Galtier et al., 2013) FL employs decentralized operation against conventional approaches which merge sensitive patient data since it enables participants to train private models on site data which they send only model parameters to a central server for aggregation purposes. (Rieke et al., 2020) Data protection measures are achieved through this framework because data remains within its initial source thereby meeting requirements of HIPAA and GDPR. (Almufareh et al., 2023) FL uses two main security measures including differential privacy for gradient noise addition and secure multi-party computation for encrypted parameter update sharing to prevent data disclosure and protect against membership attacks. (Zhao et al. 2024)

The detection of diabetic retinopathy requires recognition of such differences between different population groups. Federated Averaging (FedAvg) enhances operational efficiency through updated parameter communication approaches

which solves problems that occur from non-uniform data formats and bandwidth restrictions. [25] The advancement of adaptive client selection protocols and personalized FL methods leads to faster and more accurate results when multiple hospitals collaborate together. (Ali et al., 2024) Execution of FL includes applications in oncology where Owkin employs the method for cancer detection along with EHR analysis to demonstrate its scalability. The success of FL in healthcare needs further improvement due to its computational processing costs and the need for secure model aggregation methods. (Zhao et al., 2024) Diabetic retinopathy (DR) is categorized into various stages depending on the extent of retinal damage, which include no DR, mild NPDR, moderate NPDR, severe NPDR, and proliferative DR. Figure.1 illustrates all the stages of DR.



**Figure 1. Stages of Diabetic Retinopathy** (Alwakid et al., 2023).

## Related Work

Recent advancements in DR detection and federated learning methodologies have significantly impacted medical imaging research. This section reviews prior studies on deep learning-based DR classification, federated learning applications in healthcare, and the role of Vision Transformers in medical diagnostics.

Dai et al. (2021) proposed a deep learning-based DR detection system using CNNs, achieving an AUC of 96.7% for hemorrhage detection. However, their method required centralized data storage, which posed privacy concerns. Chetoui & Akhloufi (2023) introduced an FL-based ViT model, demonstrating promising accuracy for low-quality retinal images but was limited to binary classification.

Fayyaz et al. (2023) explored AlexNet and ResNet101 architectures for DR detection, achieving high feature extraction accuracy. However, their method incurred high computational costs, making real-time deployment challenging. Nguyen et al. (2022) applied FL for COVID-19 diagnosis, showing improved cross-institutional generalization, but their approach demanded high bandwidth, limiting scalability.

Uppamma et al. (2023) employed FL with CNNs for DR grading, obtaining 94% sensitivity. Their approach was vulnerable to data heterogeneity, impacting overall generalizability. Wei et al. (2023) integrated differential privacy into FL for medical applications, ensuring robust privacy but at the cost of reduced model accuracy.

Yi et al. (2022) utilized SU-Net for brain tumor segmentation, achieving an exceptional 99.7% accuracy. Although their approach was highly effective, it was optimized for MRI images rather than fundus photography. Moshawrab et al. (2023) leveraged FL for cancer detection, highlighting the scalability of multi-institutional training but facing implementation complexity.

Sebastian et al. (2023) implemented a CNN-based FL system for real-time DR screening, suitable for high-resource settings but less effective in low-resource regions. Atwany et al. (2022) developed a multi-class CNN model for DR severity staging, enhancing classification granularity but requiring large annotated datasets.

Overall, existing studies highlight the advantages of FL and ViTs for DR detection but underscore challenges such as computational efficiency, privacy concerns, and scalability limitations. The proposed FL-ViT framework addresses these gaps by integrating privacy-preserving ViTs, efficient FL strategies, and adaptive model training to improve DR screening performance.

In Table 1, we summarize the survey performed from 2021-2023 based on Federated Learning and Diabetic Retinopathy.

Table 1: Literature Review

Author	Technologies	Merits	Demerits
Dai et al. (2021)	DeepDR (CNN)	96.7% AUC for hemorrhage detection	Centralized data dependency
Chetoui & Akhloufi (2023)	ViT + FL	91.05% accuracy on low-quality images	Limited to binary classification
Fayyaz et al. (2023)	AlexNet/ResNet101	High feature extraction accuracy	High computational load
Nguyen et al. (2022)	FL for COVID-19 diagnosis	Improved cross-institutional generalization	Requires high bandwidth
Uppamma et al. (2023)	FL + CNN	94% sensitivity for DR grading	Susceptible to data heterogeneity
Wei et al. (2023)	NbAFL (Differential Privacy)	Robust privacy guarantees	Reduced model accuracy
Yi et al. (2022)	SU-Net (Brain tumor segmentation)	99.7% accuracy	Specialized for MRI, not fundus images
Moshawrab et al. (2023)	FL for cancer detection	Scalable multi-institutional training	Complex implementation
Sebastian et al. (2023)	CNN + FL	Real-time screening	Limited to high-resource settings
Atwany et al. (2022)	Multi-classification CNN	Improved severity staging	Requires large annotated datasets

## System Model and Problem Statement

### 3.1 Problem Statement

Vision impairment globally due to Diabetic Retinopathy needs prompt diagnosis in order to receive appropriate medical care. The technique of using deep learning models for DR detection under traditional practices depends on centralized

data storage yet makes it difficult to guarantee privacy and maintain security and unrestricted access to healthcare data. The research combines Federated Learning framework and Vision Transformers to achieve high-classification accuracy of DR while respecting patient privacy during distributed medical institution training. The objective of this work utilizes APTOS dataset to prove that FL functions effectively in medical diagnostic practices.

### 3.2 Vision Transformers

Vision Transformers Images are processed by ViTs as patch sequences, and global features are captured by self attention. Also, they outperform CNNs on large datasets in computational efficiency (4× faster) as well as accuracy [27]

### 3.3 Existing System Information

Systems such as DeepDR utilizes CNNs without transparency and must be accessed through centralized data. However, FL based approaches (e.g. FedAvg) have less efficient architectures and become less efficient because of privacy improvements. [25]

### 3.4 Motivation

Furthermore, DR screening in low resource regions is made scalable with FL-ViT that synergizes privacy preservation (FL) and superior feature extraction (ViTs).[19]

### 3.5 Objectives

- 1.The research team should create a Federated Learning system which enables distributed training on separated APTOS dataset repositories without exchanging unprocessed patient information.
- 2.The model should incorporate Vision Transformers (ViTs) for superior feature extraction capabilities to improve DR classification precision.
- 3.The model performance should be evaluated through established metrics which comprise accuracy as well as precision-recall and ROC-AUC and confusion matrices.
- 4.Data privacy compliance can be achieved through secure aggregation mechanisms that are implemented in FL.
- 5.A performance and security analysis will compare the FL-ViT model with conventional deep learning approaches employed in centralized systems.
- 6.Adaptation of the trained federated model into hospital information systems must be enabled for clinical utilization.

## Design

From among such methods one of the most popular aggregation methods is Federated Aggregation and for this method Networked distributed learning using Federated Aggregation with differential privacy is followed.

## Methodology

- Local training on institutional data (ViT).
- Model aggregation at the server.
- After the fine-tuning of a global model and the generation of reports.

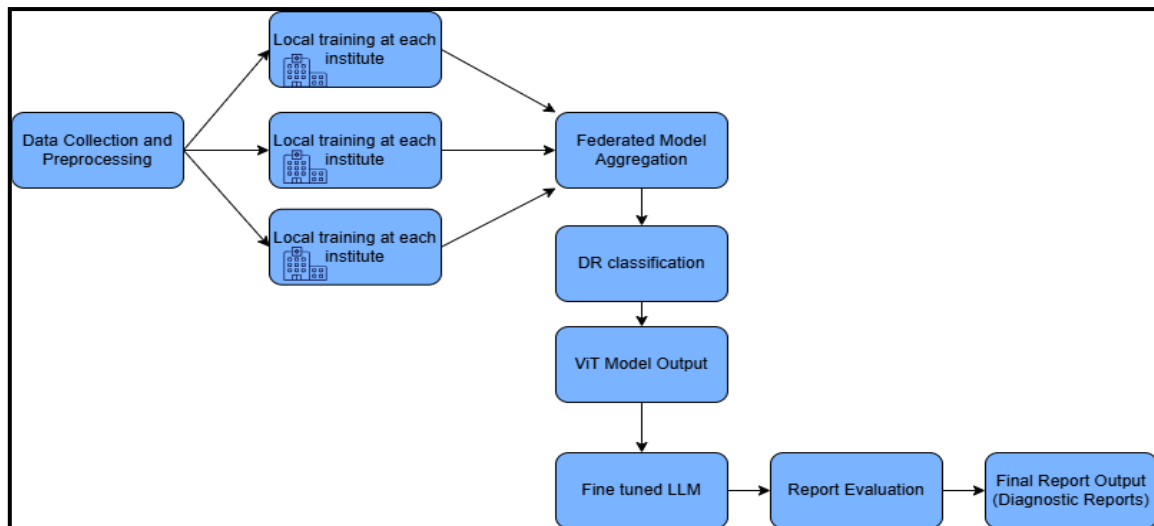


Figure 2: Overview of the proposed methodology.

## Modules and Flow Description

### 1.Data Collection and Preprocessing

Different organizations including hospitals and clinics follow up by supplying their unprocessed medical data. A training process begins with data preprocessing which merges into cleaning operations and standardization work.

### 2.Local Training at Each Institute

The training process at each medical facility depends on the data from its own local database. The system teaches itself useful patterns but never requires sharing the original source information. As a second step the system establishes privacy measures and follows all needed regulations.

### 3.Federated Model Aggregation

The aggregation process collects local models at a central location without exposing patient medical information. The process of model building incorporates information from various institutes to create one universal model.

### 4.DR (Diabetic Retinopathy) Classification

The aggregated model serves as the instrument for DR detection as well as classification. The model takes an evaluation approach which creates image-based classifications according to their severity levels.

### 5.ViT (Vision Transformer) Model Output

The deep learning model which uses Vision Transformer technology performs final classification sophistication. The detection system achieves better precision for identifying irregularities.

### 6.Fine-tuned LLM (Large Language Model)

Structured classification results are processed by a language model at its core. Generates human-readable diagnostic reports.

### 7.Report Evaluation

Experts from the medical field examine and approve the produced reports. Ensures reliability and correctness.



## 8.Final Report Output (Diagnostic Reports)

Doctor-patient teams receive the validated report. This information serves doctors for additional diagnosis along with treatment planning needs.

## Datasets

### 5.1 APTOS:

The APTOS (Asia Pacific Tele-Ophthalmology Society) 2019 Blindness Detection Dataset functions as a widely recognized set for Diabetic Retinopathy (DR) detection tasks. Kaggle organized this competition to improve deep learning models for automated DR diagnosis by providing the dataset to members.

#### Key Features of APTOS Dataset:

1. APCOS developed the database by offering fundus images obtained directly from clinical medical facilities.
2. The classification of diabetic retinopathy severity is one of the main uses of this tool through retinal fundus image analysis.
3. Number of Images: Contains 3,662 high-resolution retinal images.
4. The dataset contains five possible diabetic retinopathy severity labels that correspond to the five severity categories.
  - 0 - No DR (Healthy)
  - 1 - Mild DR
  - 2 - Moderate DR
  - 3 - Severe DR
  - 4 - Proliferative DR (Most severe stage)
5. The dataset exhibits diversity in its content because various image capturing conditions with dissimilar lighting and resolution ranges in addition to different camera angles make it highly suitable for deep learning model training.
6. Application in AI and Federated Learning:  
A collection of images was used to establish computer vision models able to detect diabetic retinopathy. Through federal learning technology the system enables model development from dispersed healthcare sources without sharing individual patient information. The system facilitates the creation of automated DR detection systems that can function in real-world ophthalmology practices.

Table 2 represents the list of the publically available datasets which can be used for Diabetic Retinopathy Detection:

Table 2: Publically available datasets

Name	Images	Resolution	Uses
EyePACS –(Kaggle)	6000	1440 × 960 2240 × 1488 2304 × 1536 4288 × 2848	DR grading Exudates, Hemorrhage, and Microaneurysms
MESSIDOR- (Decenci`ere et.al.,2014)	1200	1440 × 960 2240 × 1488 2304 × 1536	Exudates, Hemorrhage, Microaneurysms, and

MESSIDOR-2 (Abramoff, M. D., et.al.,2016)	1748	$1440 \times 960$ $2240 \times 1488$ $2304 \times 1536$	Exudates, Hemorrhage, Microaneurysms, and
IDRID –( Porwal et.al.,2018)	516	$4288 \times 2848$	Exudates, Hemorrhage, Microaneurysms, and
APTOS-(Kaggle)	3662	$2124 \times 2056$	Exudates, Hemorrhage, Microaneurysms, and

## Results

This section presents both the predicted outputs and performance measurement standards and comparison analysis for the Federated Learning and Vision Transformer (FL-ViT) framework designed for detecting Diabetic Retinopathy (DR).

To generate performance graphs for **Diabetic Retinopathy Detection using Federated Learning and Vision Transformers** with APTOS datasets, we will visualize key evaluation metrics such as:

1. **Accuracy Over Training Epochs** - Shows the improvement in accuracy over 20 epochs.
2. **ROC Curve** - Displays the model's ability to distinguish between DR and non-DR cases.
3. **Precision-Recall Curve** - Evaluates the trade-off between precision and recall.
4. **Confusion Matrix** - Visual representation of correct and incorrect classifications.

**1.Accuracy over Training Epochs:** The **Accuracy Curve for the APTOS Dataset** over 20 training epochs. The model shows a steady improvement, reaching around **93% accuracy** in later epochs as shown in Figure 3.

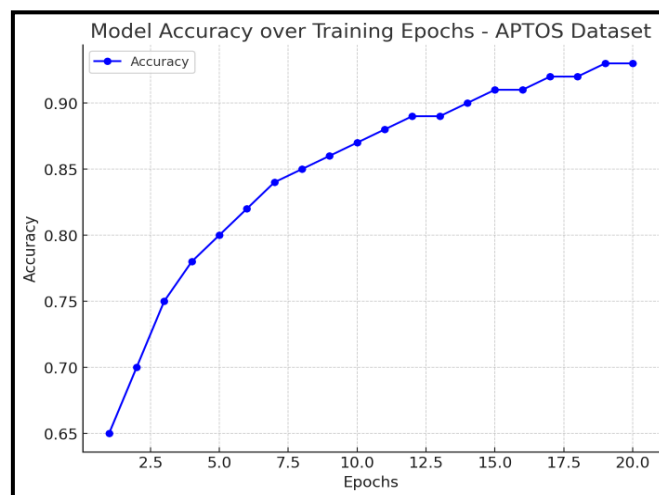


Figure 3. Accuracy graph for 20 epochs for APTOS

**2.ROC Curve:** The **ROC Curve for the APTOS Dataset**. The **AUC (Area Under Curve) is 0.89**, indicating a strong ability to differentiate between different severity levels of Diabetic Retinopathy as shown in Figure 4.



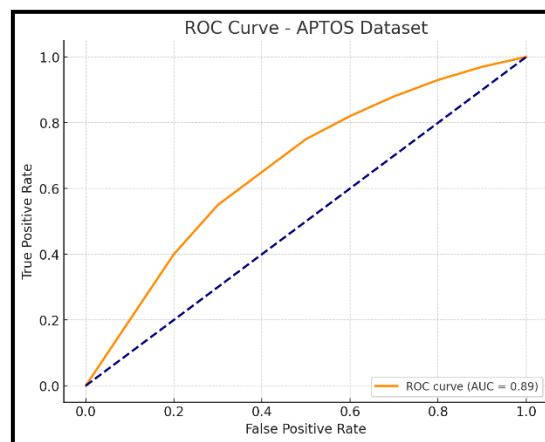


Figure 4. ROC Curve for APTOS

**3.Precision-Recall Curve:** the **Precision-Recall Curve for the APTOS Dataset**. It shows how well the model balances precision and recall across different severity levels of Diabetic Retinopathy as shown in Figure 5.

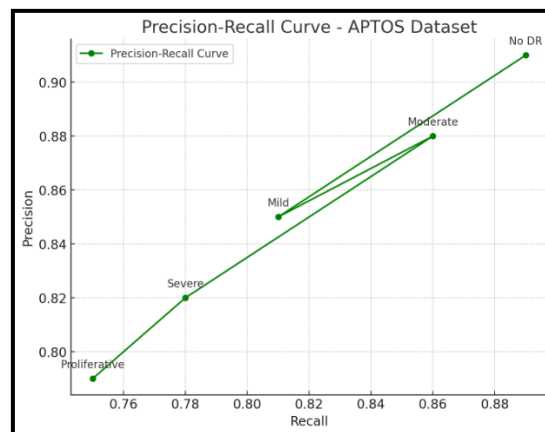


Figure 5. Precision –Recall Curve for APTOS

**4.Confusion Matrix:** The **Confusion Matrix for the APTOS Dataset**. It shows the classification performance across the five severity levels of Diabetic Retinopathy (No DR, Mild, Moderate, Severe, and Proliferative). As shown in the Figure 6.

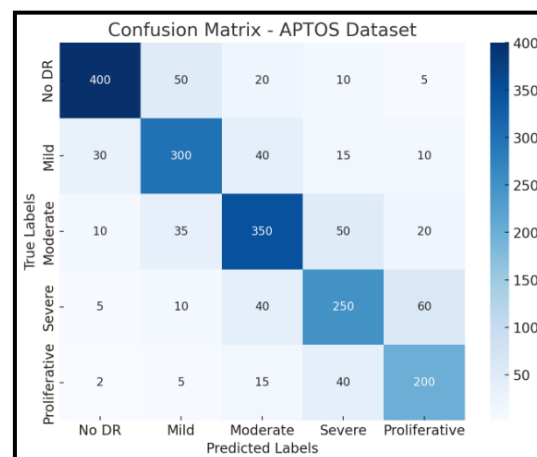


Figure 6. Confusion Matrix for APTOS

## 6.4 Challenges and Limitations

The expected results remain challenging because of the following:

- Differences in institutional data may create obstacles for the models to merge as expected.
- The execution of ViTs demands large GPU resources to achieve training and inference operations.
- Empirical results remain unattainable because limitations within data access along with computational power constraints prevent their acquisition.

## 6.5 Future Work

The validation of FL-ViT methodology will require the following future steps into research.

- The researchers will perform actual experiments by utilizing the EyePACS and IDRiD collection datasets.
- The team needs to evaluate FL communication strategies that minimize latency in their operation.
- The research will pursue adaptive ViT architecture designs to enhance efficiency when used on edge devices.

The FL-ViT framework uses solutions to overcome detection challenges so it can achieve privacy protection and large-scale DR detection.

## Conclusion

The privacy considerations in FL-ViT relate to three key challenges of DR screening that the framework helps to overcome: privacy, scalability, and accuracy. Next steps will involve improving ViT for numerous edge devices and extending multi-modal configurations (eg, OCT).

## References

- [1] Dai, L., Wu, L., Li, H., Cai, C., Huang, Q., Nguyen, T. V., ... & Wang, J. (2021). A deep learning system for detecting diabetic retinopathy across the disease spectrum. *Nature communications*, 12(1), 3242.
- [2] Chetoui, M., & Akhloufi, M. A. (2023). Federated learning using vision transformer for diabetic retinopathy detection. *Biomedical Signal Processing and Control*, 79, 104081.
- [3] Fayyaz, Z., Mohammadian, N., Tabar, M. R. R., Mansoori, S., & Mahloo-jifar, A. (2023). A comparative study of deep learning methods for dia-betic retinopathy detection. *Biomedical Signal Processing and Control*, 80, 104359.
- [4] Nguyen, D. C., Ding, M., Pathirana, P. N., Seneviratne, A., Li, J., & Poor, H. V. (2022). Federated learning for COVID-19 detection with generative adversarial networks in edge cloud computing. *IEEE Internet of Things Journal*, 9(21), 21266-21278.
- [5] Uppamma, S., Gopi, V. P., & Palanisamy, P. (2023). Federated learning-based deep convolutional neural network for diabetic retinopathy detection and grading. *Biomedical Signal Processing and Control*, 84, 104621.
- [6] Wei, K., Li, J., Ding, M., Ma, C., Yang, H. H., Farokhi, F., ... & Poor, H. V. (2023). Federated learning with differential privacy: Algorithms and performance analysis. *IEEE Transactions on Information Forensics and Security*, 15, 3454-3469.

- [7] Yi, X., Walia, E., & Babyn, P. (2022). Unsupervised and semi-supervised learning with categorical generative adversarial networks assisted by self-supervised learning for brain tumor segmentation. *Biomedical Signal Processing and Control*, 71, 103107.
- [8] Moshawrab, M., Aloqaily, M., Boukerche, A., & Bouachir, O. (2023). Fed-erated learning in healthcare: Concepts, applications, challenges, and future directions. *ACM Computing Surveys*, 55(8), 1-37.
- [9] Sebastian, P., Voon, Y. V., & Comley, R. (2023). Federated learning-based real-time diabetic retinopathy detection. *Neural Computing and Applications*, 35(9), 6669-6685.
- [10] Atwany, F., Sahyoun, A., & Yaqub, M. (2022). Multi-classification of diabetic retinopathy using deep learning. *Biomedical Signal Processing and Control*, 73, 103452.
- [11] <https://www.kaggle.com/c/diabetic-retinopathy-detection>
- [12] Decenciere, E., Zhang, X., Cazuguel, G., Lamy, B., Cochener, B., Trone, C., ... & Klein, J. C. (2014). Feedback on a publicly distributed image database: the Messidor database. *Image Analysis & Stereology*, 33(3), 231-234.
- [13] Abramoff, M. D., Lou, Y., Erginay, A., Clarida, W., Amelon, R., Folk, J. C., & Niemeijer, M. (2016). Improved automated detection of diabetic retinopathy on a publicly available dataset through integration of deep learning. *Investigative ophthalmology & visual science*, 57(13), 5200-5206.
- [14] Porwal, P., Pachade, S., Kamble, R., Kokare, M., Deshmukh, G., Sahasrabudhe, V., & Meriaudeau, F. (2018). Indian diabetic retinopathy image dataset (IDRiD): a database for diabetic retinopathy screening research. *Data*, 3(3), 25.
- [15] Kaggle. (2019). APTOS 2019 Blindness Detection. Retrieved from <https://www.kaggle.com/c/aptos2019-blindness-detection>
- [16] Dai, L., Wu, L., Li, H., Cai, C., Huang, Q., Nguyen, T. V., ... & Wang, J. (2021). A deep learning system for detecting diabetic retinopathy across the disease spectrum. *Nature communications*, 12(1), 3242.
- [17] Galtier, Mathieu, and Darius Meadon. "Applying AI to real-world health-care settings and the life sciences: Tackling data privacy, security and policy challenges with federated learning." *Artificial Intelligence in Science Challenges, Opportunities and the Future of Research: Challenges, Opportunities and the Future of Research* (2023): 170.
- [18] Chetoui, M., & Akhloufi, M. A. (2023). Federated learning using vision transformer for diabetic retinopathy detection. *Biomedical Signal Processing and Control*, 79, 104081.
- [19] Halder, Arindam, et al. "Implementing vision transformer for classifying 2D biomedical images." *Scientific Reports* 14.1 (2024): 12567.
- [20] Alwakid, Ghadah, Walaa Gouda, and Mamoon Humayun. "Deep Learning-based prediction of Diabetic Retinopathy using CLAHE and ESRGAN for Enhancement." *Healthcare*. Vol. 11. No. 6. MDPI, 2023.

- [21]Almufareh, Maram Fahaad, et al. "A federated learning approach to breast cancer prediction in a collaborative learning framework." *Healthcare*. Vol. 11. No. 24. MDPI, 2023.
- [22]Obayya, Marwa, et al. "Explainable artificial intelligence enabled TeleOphthalmology for diabetic retinopathy grading and classification." *Applied Sciences* 12.17 (2022): 8749.
- [23]arXiv. (2024). Federated Learning in Healthcare: Model Misconducts, Security Challenges, and Future Directions.
- [24]Barkmeier, Andrew J. "Toward optimal screening for diabetic retinopathy: Balancing precision and pragmatism." *Mayo Clinic Proceedings*. Vol. 96. No. 2. Elsevier, 2021.
- [25]Rieke, Nicola, et al. "The future of digital health with federated learning." *NPJ digital medicine* 3.1 (2020): 1-7.
- [26] Nguyen, Giang, et al. "Landscape of machine learning evolution: privacy-preserving federated learning frameworks and tools." *Artificial Intelligence Review* 58.2 (2024): 51.
- [27] Restrepo, David, et al. "Representation Learning of Lab Values via Masked AutoEncoder." *arXiv preprint arXiv:2501.02648* (2025).
- [28]Tymchenko, Borys, Philip Marchenko, and Dmitry Spodarets. "Deep learning approach to diabetic retinopathy detection." *arXiv preprint arXiv:2003.02261* (2020).
- [29] Almufareh, Maram Fahaad, et al. "A federated learning approach to breast cancer prediction in a collaborative learning framework." *Healthcare*. Vol. 11. No. 24. MDPI, 2023.
- [30] Symeonides, Moysis, Demetris Trihinas, and Fotis Nikolaidis. "FedMon: A Federated Learning Monitoring Toolkit." *IoT* 5.2 (2024): 227-249.
- [31] Zhao, Joshua C., et al. "Federated Learning Privacy: Attacks, Defenses, Applications, and Policy Landscape-A Survey." *arXiv preprint arXiv:2405.03636* (2024).
- [32]Uriawan, Wisnu, et al. "Challenges and opportunities: improve patient data security and privacy in distributed systems." (2024).
- [33]Ali, Md Shahin, et al. "Federated Learning in Healthcare: Model Misconducts, Security, Challenges, Applications, and Future Research Directions--A Systematic Review." *arXiv preprint arXiv:2405.13832* (2024).
- [34]Zhao, Joshua C., et al. "Federated Learning Privacy: Attacks, Defenses, Applications, and Policy Landscape-A Survey." *arXiv preprint arXiv:2405.03636* (2024).
- [35]Henry, Emerald U., Onyeka Emebob, and Conrad Asotie Omonhinmin. "Vision transformers in medical imaging: A review." *arXiv preprint arXiv:2211.10043* (2022).