Journal of Infectious Diseases and Patient Care

Research Article

WRJIDPC-26-039

OptiVerse AI: Explainable Deep Learning and Three-Dimensional Retinal Reconstruction from Fundus Images for Early Retinal Disease Detection

Abdelrahman Elwelely*

Independent Researcher, Fayoum STEM High School, Al Faiyum Al Gadida, Egypt

***Correspondence:** Abdelrahman Elwelely, Independent Researcher, Fayoum STEM High School, Al Faiyum Al Gadida, Egypt, E-mail: iiabdelrahman05@gmail.com; DOI: https://doi.org/10.56147/jidpc.3.1.39

Citation: Elwelely A (2026) OptiVerse AI: Explainable Deep Learning and Three-Dimensional Retinal Reconstruction from Fundus Images for Early Retinal Disease Detection. J Infect Dis & Pati Care 3: 39.

Abstract

Diabetic Retinopathy (DR) is a major cause of blindness, but early detection is hindered by poor access to retinal imaging services and lack of interpretability of existing artificial intelligence (AI) systems. In this work, we introduce OptiVerse AI, an end-to-end architecture of explainable deep learning along with computational three-dimensional retinal reconstruction based on conventional color fundus images. The proposed system relies on transfer learning based on EfficientNetB7 for multi-stage diabetic retinopathy classification, leveraging adaptive class weighting and data augmentation for improved robustness in imbalanced datasets. We apply Gradient-weighted Class Activation Mapping (Grad-CAM) for achieving model interpretability via visualization of retinal regions involved in diagnostic decisions. Moreover, we propose a pipeline for reconstructing retinal depth maps from learned features in order to generate three-dimensional point clouds and meshes without Optical Coherence Tomography (OCT). The system has been developed based on a database of more than 90,000 retinal fundus images and achieved the accuracy of more than 95%. It provides a cloud platform for combining disease detection, model interpretation and structural analysis of retina, thus being a scalable and cost-effective solution potentially useful for improving diabetic retinopathy screening and clinical decision-making.

Keywords: Diabetic retinopathy; Fundus imaging; Deep learning; Explainable artificial intelligence; EfficientNetB7; Grad-CAM; Three-dimensional retinal reconstruction; Medical image analysis

Abbreviations: AI: Artificial Intelligence; CNN: Convolutional Neural Network; DR: Diabetic Retinopathy; NPDR: Non-Proliferative Diabetic Retinopathy; PDR: Proliferative Diabetic Retinopathy; DL: Deep Learning; TL: Transfer Learning; XAI: Explainable Artificial Intelligence; Grad-CAM: Gradient-weighted Class Activation Mapping; OCT: Optical Coherence Tomography; RGB: Red-Green-Blue; CLAHE: Contrast Limited Adaptive Histogram Equalization; ReLU: Rectified Linear Unit; GPU: Graphics Processing Unit; IoT: Internet of Things; API: Application Programming Interface; 3D: Three-Dimensional; CAD: Computer-Aided Diagnosis; DNN: Deep Neural Network; CAM: Class Activation Mapping

Received date: July 21, 2026; **Accepted date:** August 03, 2026; **Published date:** August 11, 2026

Introduction

Diabetic Retinopathy (DR) is one of the main reasons for preventable blindness around the world and constitutes one of the most serious microvascular complications of diabetes mellitus. In view of the increasing prevalence of diabetes across the globe, there will be many people who are at risk of developing retinal conditions that threaten their vision. The reason why diabetic retinopathy is a difficult condition to diagnose is because it develops asymptomatically in its initial stages. Timely diagnosis and treatment of DR play a critical role in reducing the chances of permanent blindness; thus, retinal screening should be a mandatory part of the process of diabetes management.

Nevertheless, many healthcare facilities do not have sufficient access to specialized equipment and experienced ophthalmologists.

However, the recent development in the field of AI and more specifically deep learning has shown great potential in terms of automated diagnosis of retinal diseases through color fundus images. The convolutional neural network has been able to provide diagnostic results that rival those of experienced ophthalmologists for diabetic retinopathy diagnosis, which would help in conducting fast and efficient screening processes. However, the current AI methods available are mostly "black box" models that make very accurate predictions but do not show the basis for their conclusions.

One of the drawbacks of existing fundus-based diagnosis technologies is the lack of capability of providing the three-dimensional structure of the retina. Even though OCT technology is able to generate high-resolution volumetric images and has already been established as a clinical gold standard for the evaluation of the structure of the retina, the technology is rather expensive and not available everywhere. This makes fundus photography still the most available method of retinal imaging.

In order to overcome these obstacles, the current paper proposes OptiVerse AI an integration of explainable deep learning together with computational 3D retinal reconstruction from ordinary color fundus images. Specifically, the proposed framework utilizes transfer learning with EfficientNetB7 for diabetic retinopathy classification through five stages, utilizes Grad-CAM for generation of clinically interpretable visual explanations and provides a novel computational framework for reconstructing 3D retinal models from fundus photos without the need of OCT imaging. With the help of disease detection, interpretability and structural representation on one cloud-based platform, OptiVerse AI can be considered a potentially viable solution for diabetic retinopathy screening that is efficient, accessible and affordable.

Scientific Background

The human retina and diabetic retinopathy

The retina is a light-sensing neural tissue present at the back part of the eye, which converts light into electrical impulses that are sent to the brain via the optic nerve. Retina comprises cells along with an extensive microvascular network that play a key role in enabling vision. The persistent presence of high blood sugar levels in patients suffering from diabetes can cause damage to the blood vessels in the retina, giving rise to Diabetic Retinopathy (DR), which is one of the major causes of preventable blindness.

DR develops in stages, starting with microaneurysms and retinal bleeding before moving on to vascular leakage, ischemia and abnormal new vessel growth. Untreated DR ultimately causes loss of vision. The clinical stages of DR include No DR, Mild NPDR, Moderate NPDR, Severe NPDR and Proliferative DR (PDR). Identification of these clinical stages is extremely important.

Color fundus imaging

Color fundus imaging is the predominant modality utilized for diabetic retinopathy screening as it is a non-invasive, low-cost and easily accessible technique. It produces two-dimensional images of various retinal structures such as the optic disc, macula, blood vessels and lesions like hemorrhage and exudate.

Although fundus images yield significant diagnostic information, they fail to offer the depth information provided by OCT images. However, their availability has helped them to become the imaging gold standard for various screening programs and AI-related applications.

Deep learning for retinal disease detection

The deep learning technique has greatly enhanced automated diagnoses of eye diseases by identifying complicated visual features automatically through medical images. The CNN surpasses conventional techniques of machine learning by removing the step of handcrafted features and offering highly accurate predictions.

In contemporary architecture, EfficientNetB7 provides a great combination of accuracy and efficiency through the compound scaling of network depth, width and resolution of input images. In addition to that, when combined with the ImageNet transfer learning, EfficientNetB7 can classify retinal diseases effectively using lesser medical images.

Explainable artificial intelligence in ophthalmology

Although deep learning-based systems tend to achieve excellent results, their operation is generally considered to be a black box, which does not inspire confidence in clinicians. The approach to solving this problem is Explainable Artificial Intelligence (XAI), which provides explanations through visualization. The technique that is among the most popular in XAI is called Gradient weighted Class Activation Mapping (Grad-CAM). In this case, heatmaps are created based on the parts of an image that are responsible for making a prediction. In ophthalmology, Grad-CAM is used to check that a model pays attention to lesions in retinal images.

Three-dimensional retinal imaging

Three-dimensional evaluation of the retina is essential when assessing structural changes related to retinal disorders. The optical coherence tomography is considered the gold standard in the field of volume retinal imaging due to the cross-section images it generates. But OCT devices are quite costly and therefore not easily available in many hospitals.

There has been some recent progress in the field of computer vision where it has been found that depth maps

Journal of Infectious Diseases and Patient Care

can be generated from one single two-dimensional image using the principles of deep learning. This makes it possible to generate three-dimensional reconstructions based on standard fundus photographs.

Related Work

Deep learning for diabetic retinopathy detection

Deep learning has transformed automated diabetic retinopathy (DR) diagnosis through learning complicated retinal features from color fundus images directly. In earlier work, CNNs like VGGNet, ResNet and DenseNet were used, which outperformed the previous machine learning algorithms in terms of accuracy.

However, recently, EfficientNet is gaining popularity because of its approach of compound scaling and delivering higher classification accuracy with computational efficiency. The use of transfer learning has further improved the results due to pre-trained models and reduced training time. While these algorithms perform excellently for the task at hand there is a major drawback as most of them are only concerned with classification.

Explainable artificial intelligence in ophthalmology

The growing use of artificial intelligence in clinical practice emphasizes the need for interpretability of models. Explainable Artificial Intelligence (XAI) methods like Grad-CAM, LIME (Local Interpretable Model-Agnostic Explanations) and SHAP (SHapley Additive exPlanations) have found wide application in retinal image interpretation. The methods of XAI offer visual explanation by detecting parts of the image responsible for model predictions; thus, clinicians can confirm if the decisions made are grounded on retinal lesions. Even though XAI increases transparency and physician trust, almost all these methods offer post hoc explanations and do not improve the structural understanding of retinal images.

Three-dimensional retinal imaging

Three-dimensional visualization of the retina is very useful for analyzing any kind of structural changes related to diseases such as diabetic retinopathy and other retinal disorders. Optical Coherence Tomography (OCT) is considered to be a gold standard for volumetric imaging of the retina since it provides high resolution cross-sectional images of different retinal layers.

On the other hand, OCT systems are relatively costly as they need specific hardware and are not affordable in many health care facilities. Modern researches in the fields of monocular depth estimation and computer vision have proved that three-dimensional shape recovery can be done with just one image. Most of the existing retinal reconstruction techniques, however, use special OCT data or other complex imaging systems.

Limitations of existing methods

Although significant progress has been made in automated retinal analysis, important challenges remain. Deep learning models achieve high diagnostic accuracy but generally operate as black-box systems, limiting clinical interpretability. Explainable AI methods improve transparency but remain dependent on the underlying classification model and do not provide additional structural information. Meanwhile, OCT-based approaches offer detailed three-dimensional retinal visualization but require costly imaging equipment that is unavailable in many healthcare environments.

As a result, existing methods typically address only one aspect of retinal diagnosis accurate classification, explainability or structural visualization rather than integrating these capabilities into a unified framework.

The proposed OptiVerse AI framework addresses this gap by combining EfficientNetB7-based diabetic retinopathy classification, Grad-CAM explainability and computational three-dimensional retinal reconstruction from conventional color fundus images within a single cloud-based platform. By eliminating the dependence on OCT while maintaining high diagnostic performance and transparent decision-making, OptiVerse AI provides a more accessible and clinically practical solution for diabetic retinopathy screening.

Methodology

Overview of the OptiVerse AI framework

The OptiVerse AI is an advanced artificial intelligence system that is developed to give automated DR detection, clinical decision-making support with explanations and 3D retinal reconstruction using conventional color fundus images. In contrast to current methods which are limited to disease detection or retinal visualization alone, this framework combines deep learning, Explainable Artificial Intelligence (XAI) and computer vision methods in one cloud system. This system is capable of conducting a comprehensive analysis of the retina with the help of a mere fundus image without requiring expensive imaging procedures like Optical Coherence Tomography (OCT).

The process flow involves six consecutive steps. First, images of the retina are obtained from publicly accessible data sources or clinical imaging equipment. Preprocessing of the images is done to increase their visual consistency and boost model generalizability. Then, a deep learning-based model using EfficientNetB7 is used to classify multi-class diabetic retinopathy. Grad-CAM is then used to create heatmaps showing the area of interest on the retina responsible for making the prediction.

Once classification is done, the deep features extracted are used for estimating the depth of the retina using the two-dimensional image. This estimated depth map is then converted into a 3D point cloud that is later used to create a 3D surface mesh of the retinal geometry. Finally, all of the

Journal of Infectious Diseases and Patient Care

above-mentioned steps are performed and the results are displayed on a cloud-based interface **(Figure 1)**.

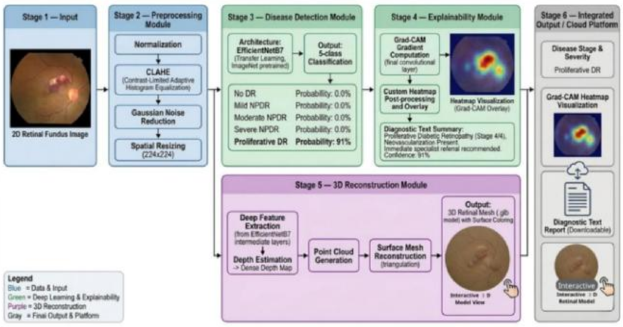

Figure 1: Overall system architecture and end-to-end workflow of the proposed OptiVerse AI framework. The framework consists of six sequential modules: Image acquisition, preprocessing, deep learning-based diabetic retinopathy classification, Grad-CAM explainability, computational three-dimensional retinal reconstruction and cloud-based visualization.

Image acquisition and preprocessing

The suggested architecture takes in color fundus images from publicly available datasets for diabetic retinopathy with all five stages of the disease. The datasets consist of retinal images that have been taken in different conditions and camera configurations to increase the variety of the input data for improving robustness and generalization.

Prior to training and testing the neural network, the input images go through a series of preprocessing steps. First of all, all fundus images are resized to 224 x 224 pixels, which corresponds to the input size of EfficientNetB7 model. Then, pixel values are normalized.

In order to increase the visibility of retinal features, Contrast Limited Adaptive Histogram Equalization (CLAHE) is employed, which increases local contrast within an image while retaining its anatomical features like blood vessels, hemorrhages and exudates. The next step is the reduction of Gaussian noise so that imaging artifacts can be reduced without altering the pathological features much.

For increasing the robustness of the proposed system and to prevent overfitting, various data augmentation methods including random rotation, flipping horizontally and vertically, zooming, changing brightness and contrast are applied. The processed images are then fed into the deep learning classification part.

Deep learning-based diabetic retinopathy classification

The classification part of the OptiVerse AI is based on the EfficientNetB7, which is an advanced CNN architecture known for its high classification accuracy and efficiency. EfficientNetB7 uses a compound scaling method, which ensures optimal scaling of depth, width and input image resolution, thus providing better performance with reduced parameter size than other CNN architectures.

The transfer learning technique is used to train the network using weights learned previously on the ImageNet data set and further tuning the network on retinal fundus images.

The final fully connected layer is modified to perform multi-class classification across five diabetic retinopathy severity stages:

- No Diabetic Retinopathy (No DR)
- Mild Non-Proliferative Diabetic Retinopathy (Mild NPDR)
- Moderate Non-Proliferative Diabetic Retinopathy (Moderate NPDR)
- Severe Non-Proliferative Diabetic Retinopathy (Severe NPDR) Proliferative Diabetic Retinopathy (PDR).

Adaptive class weighting technique is applied to address class imbalance and enhance classification at all disease stages. The model produces the predicted class and its confidence value for explainability analysis.

Explainable artificial intelligence using Grad-CAM

In an attempt to make the models more transparent, OptiVerse AI uses Gradient weighted Class Activation Mapping (Grad-CAM) which helps create visualizations for the areas of the retina responsible for predictions.

Grad-CAM creates activation heat maps based on the last convolutional layer and places them onto the original fundus images, making it possible to see clinical signs such as microaneurysms, hemorrhages, exudates and neovascularization.

Three-dimensional retinal reconstruction

Once classification of diseases is completed, the OptiVerse AI executes the computation of the retinal reconstruction in 3D using deep feature extraction obtained through EfficientNetB7. Unlike traditional techniques which rely on Optical Coherence Tomography (OCT) scans, the new approach reconstructs retinal geometry directly from a single two-dimensional fundus image.

The process of retinal reconstruction starts with extraction of spatial features from the intermediate layers of the classified network. This will be used to create an estimation of the dense retinal depth map. Once the estimation is done, depth value will be combined with image coordinates to create 3D point cloud for retinal topology.

Finally, the point cloud is modeled via the surface reconstruction process to create a three-dimensional reconstruction of the retina with anatomically feasible triangulation. The surface is colored using the fundus image in order to retain the anatomical characteristics of the retina.

This modeling process allows the user to analyze the morphology of the retina and gives more information about its structure in comparison with two-dimensional fundus photography. This computational method expands the diagnostic capabilities of traditional retinal imaging without the need for an OCT **(Figure 2)**.

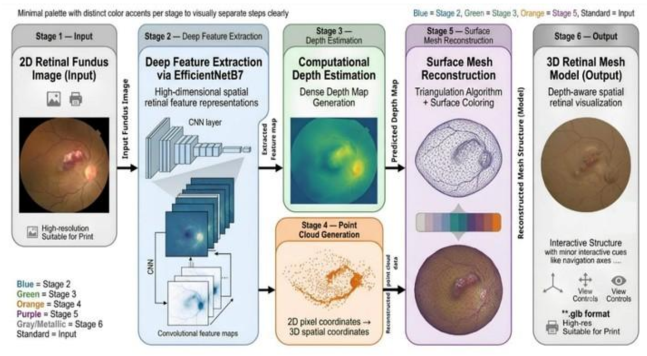

Figure 2: Step-by-step pipeline of the proposed three-dimensional retinal reconstruction module. Deep features extracted from EfficientNetB7 are used to estimate retinal depth, generate a three-dimensional point cloud and reconstruct a surface mesh representing retinal geometry.

Cloud-based deployment

In order to enable its practical application in clinical practice, the developed framework is implemented via a cloud-based platform which brings together the functionality of automated diagnosis, explainability and 3D visualization all in one user interface. After the upload of retinal images, the platform analyses the images using the whole OptiVerse AI pipeline and gives the diagnostic results immediately.

These results include the predicted stage of diabetic retinopathy, level of prediction certainty, Grad-CAM heatmap and an interactive 3D model of the retina. Moreover, the platform automatically creates a downloadable report on the diagnostic results and visual explanations provided.

Due to the implementation of the cloud-based architecture, remote access to the system becomes possible without the need for the local computing resources. Thus, the system may be used not only in tele ophthalmological settings, but in case of mass screening of patients for diabetic retinopathy.

Mathematical and Computational Framework

Image representation

Each retinal fundus image is represented as a three-channel RGB image:

$$I \in \mathbb{R}^{H \times W \times 3}$$

where and denote the image height and width, respectively. Prior to training, images are resized to the network input size and normalized to improve numerical stability during optimization.

EfficientNetB7 formulation

OptiVerse AI employs EfficientNetB7 as the backbone network for diabetic retinopathy classification. EfficientNet scales the network using a compound scaling strategy that balances network depth, width and input resolution.

The scaling factors are defined as

$$d = \alpha^{\phi}, \qquad w = \beta^{\phi}, \qquad r = \gamma^{\phi}$$

subject to

$$\alpha \cdot \beta^2 \cdot \gamma^2 \approx 2,$$

Where

d denotes network depth, ω represents network width, r is the input image resolution, ϕ is the compound scaling coefficient.

This strategy improves feature extraction while maintaining computational efficiency compared with conventional CNN architectures.

Transfer learning

Instead of training the model from random initialization, transfer learning is employed by adapting ImageNet pre-trained weights to retinal fundus images.

The optimization objective is

$$\theta^* = \arg\min_{\theta} L(\theta),$$

Where

θ represents the network parameters, $L\ (\theta)$ is the classification loss.

Fine-tuning enables the model to learn retinal-specific features while preserving useful low-level representations learned from large-scale natural image datasets.

Loss function

The classification model is optimized using the categorical cross-entropy loss function,

$$L = -\sum_{i=1}^{C} y_i \log(\hat{y}_i)$$

Where

Journal of Infectious Diseases and Patient Care

C is the number of classes, y_i is the ground-truth label, $\hat{y}i$ is the predicted probability for class iii.

To compensate for class imbalance, class-specific weights are incorporated during optimization, assigning greater importance to minority disease categories.

Grad-CAM formulation

To improve model interpretability, Gradient-weighted Class Activation Mapping (Grad-CAM) is employed to identify retinal regions influencing the classification decision.

The importance weight for feature map k is computed as

$$\alpha_k = \frac{1}{Z}\sum_i \sum_j \frac{\partial y^c}{\partial A_{ij}^k}$$

Where

A^k denotes the k^{th} feature map, y is the score of the predicted class, Z is the number of pixels in the feature map.

The final heatmap is obtained by

$$L_{\text{Grad-CAM}} = \text{ReLU}\left(\sum_k \alpha_k A^k\right)$$

The resulting activation map highlights retinal lesions contributing to the diagnostic prediction.

Depth estimation

Following disease classification, the extracted deep features are utilized to estimate retinal depth from the input fundus image. The estimated depth map is expressed as

$$D = f(I)$$

Where

I is the input retinal image, $f(\cdot)$ denotes the depth estimation function, D represents the predicted depth map.

This estimated depth provides geometric information required for three-dimensional reconstruction without relying on OCT imaging.

Point cloud generation

Each pixel in the estimated depth map is projected into three-dimensional space to construct a retinal point cloud. The projection is defined by

$$P(x, y)=(x, y, D(x, y))$$

Where

x, y are image coordinates, $D(x, y)$ is the estimated depth value.

The resulting point cloud represents the retinal surface geometry and serves as the basis for mesh reconstruction.

Surface mesh reconstruction

The reconstructed point cloud is converted into a continuous triangular mesh using a surface reconstruction algorithm. The reconstructed surface is defined as

$$M=g(P)$$

Where

P denotes the generated point cloud, $g(\cdot)$ represents the mesh reconstruction process, M is the final three-dimensional retinal model.

The original fundus image is mapped onto the reconstructed mesh for visualization.

Experimental Setup

Dataset

The proposed framework was designed and tested through publicly available retinal fundus image data sets consisting of over 90,000 images from the five known clinical stages of diabetic retinopathy. The datasets consist of retinal images captured under various imaging settings and camera configurations, hence adding to the diversity in terms of image quality, lighting and severity of the disease.

Training configuration

The EfficientNetB7 model was trained using ImageNet pre-trained weights in a transfer learning setup. The Adam optimizer along with the categorical cross entropy loss function was used for training. Data augmentation, such as random rotation, flipping, zooming, brightness modification and contrast enhancement, was utilized to minimize overfitting. Class weight adaptation was carried out to compensate for the imbalanced nature of the data set.

Evaluation metrics

The performance of OptiVerse AI was evaluated using four standard classification metrics:

- **Accuracy:** Measures the proportion of correctly classified retinal images.
- **Precision:** Evaluates the reliability of positive predictions.
- **Recall (sensitivity):** Measures the model's ability to correctly identify diseased cases.
- **F1-score:** Provides the harmonic mean of precision and recall, offering a balanced measure of classification performance.

Apart from using quantification for model validation, Grad-CAM heatmaps were manually inspected to confirm that the model was attending to important retinal lesions in the clinical sense. The resulting three-dimensional retinal models were also manually evaluated based on their

Journal of Infectious Diseases and Patient Care

structural validity.

Implementation environment

The proposed framework was implemented using Python with the TensorFlow/Keras deep learning framework. Training was accelerated using GPU hardware, while the cloud-based interface was developed using Flask to provide automated diagnosis, Grad-CAM visualization and interactive three-dimensional retinal reconstruction through a web-based platform.

Results and Discussion

Classification performance

The proposed OptiVerse AI algorithm showed excellent results in multiclass diabetic retinopathy classification in all five stages of the disease. Using the combination of EfficientNetB7, transfer learning, adaptive class weights and data augmentation techniques, the proposed model delivered high accuracy scores while achieving balanced results for commonly represented and underrepresented classes.

Accuracy, Precision, Recall and F1-Score were used to evaluate the performance of the algorithm. It shows that the proposed model can successfully distinguish between various stages of the disease **(Table 1)**.

Table 1: Summarizes the classification performance of the proposed model.

Metric	Value
Accuracy	95.7%
Precision	94.0%
Recall	95.0%
F1-score	94.0%

To further examine the classification results of the OptiVerse AI framework, the following confusion matrix provided in **Figure 3** was used. This confusion matrix provides both absolute numbers of predictions as well as percentages of classifications in each class of diabetic retinopathy severity stages.

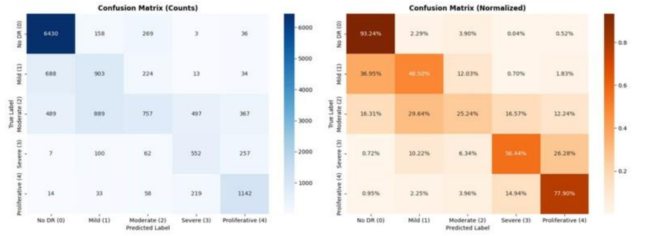

Figure 3: Confusion matrices of the proposed EfficientNetB7-based diabetic retinopathy classifier. (Left) Absolute prediction counts for each class. (Right) Normalized confusion matrix showing the percentage of correctly and incorrectly classified samples across the five diabetic retinopathy severity stages.

According to **Figure 3**, the highest classification accuracies were achieved by the classes of No Diabetic Retinopathy (No DR) and Proliferative Diabetic Retinopathy (PDR) as the samples were mostly assigned to the main diagonal. The Mild and Moderate NPDR classes had higher classification error rates due to the close appearance of neighboring classes. Such phenomenon was previously observed in similar research on multi-class diabetic retinopathy classification.

In general, it can be seen from the presented confusion matrix that the proposed framework shows consistent performance throughout all five classes of diabetic retinopathy. The normalized version of the matrix shows percentage of correctly classified samples in each class.

Explainability results

The proposed OptiVerse AI model successfully generated the 3D representation of the retinas from the conventional 2D fundus images using the computational reconstruction procedure presented in the methodology section. The method included retinal depth estimation based on deep features, creation of 3D point cloud and generation of the continuous surface mesh.

As depicted in **Figure 4**, the created 3D representation includes all the necessary anatomical features of the retina such as the optic disc, vascular network and lesions present in the original fundus image. Generated 3D point cloud properly reflects the spatial location of the retinal features and the created mesh allows visualizing the morphological structure of the retina.

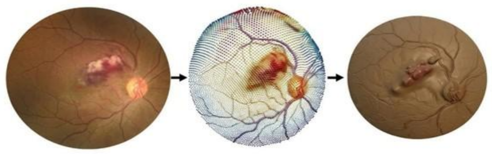

Figure 4: Three-dimensional retinal reconstruction generated by the proposed OptiVerse AI framework. From left to right: Original retinal fundus image, generated three-dimensional point cloud and reconstructed retinal surface mesh.

While the generated 3D retinal models cannot be considered an alternative for OCT scans, their creation shows the possibility of 3D retinal geometry generation from a single-color fundus image.

Comparative discussion

These findings indicate that OptiVerse AI can seamlessly integrate the capabilities of diabetic retinopathy classification, explainable artificial intelligence and computational three-dimensional retinal reconstruction under one umbrella. The EfficientNetB7 model was successful in achieving high accuracy rates for the diagnoses while balancing classification among all five

Journal of Infectious Diseases and Patient Care

stages of diabetic retinopathy. Moreover, the three-dimensional reconstruction sub-module was able to construct anatomically correct retinal models based only on standard fundus imaging.

In comparison to current technologies used in diabetic retinopathy screening, the proposed framework offers a number of advantages. While most deep learning algorithms only consider the process of classifying the disease, OCT systems rely on special equipment to acquire information about the structure of the retina. Conversely, OptiVerse AI unites the processes of classification, visualization and three-dimensional reconstruction of the retina in one cloud-based solution. **Table 2** summarizes the key differences between OptiVerse AI and representative diabetic retinopathy analysis approaches. Unlike OCT-based systems, the proposed framework does not require specialized imaging hardware, making it more accessible for large-scale screening and teleophthalmology applications.

Table 2: Comparison of the proposed OptiVerse AI framework with representative diabetic retinopathy analysis approaches.

Feature	Conventional CNN models	OCT-based systems	OptiVerse AI
Automated DR classification	✓	✓	✓
Explainable AI	Limited	Limited	✓
3D retinal reconstruction	✗	✓	✓
OCT required	No	Yes	No
Cloud-based deployment	Limited	Limited	✓

Limitations and Future Work

However, there are some limitations in the implementation of the OptiVerse framework to classify the disease state and perform the computational threedimensional retinal reconstruction. First, the reconstruction module derives the retinal geometry based on a two-dimensional fundus image; therefore, there is a lack of accuracy in terms of the geometrical structure compared to the Optical Coherence Tomography (OCT).

Moreover, the OptiVerse framework has been tested on publicly available retinal image datasets. Thus, further testing and validation of the OptiVerse framework on large multi-center datasets collected in different hospitals using different imaging instruments will allow us to better understand the robustness and generality of the implemented framework in clinical practice.

Further improvements will include the development of methods for depth estimation using deep learning approaches, adding more retinal imaging modalities to the pipeline and clinical validation performed in cooperation with ophthalmologists. The application of the developed framework in real-time clinical decision support systems and teleophthalmology applications can also become an interesting direction in the future.

Ethical, Clinical and Societal Implications

The above-discussed OptiVerse AI framework was specifically developed to complement the clinical decision-making process in diabetic retinopathy screening rather than to substitute it. Using explainable artificial intelligence with Grad-CAM, the framework will be more transparent for clinical decision-makers, allowing them to see why the AI makes certain predictions and thus increasing the likelihood of the use of this technology in the clinical environment.

Clinically, the framework can help in early diabetic retinopathy detection due to automated screening and computational 3D retinal visualization performed on the basis of standard fundus photos. These features can facilitate early retinal assessment in locations where Optical Coherence Tomography (OCT) scanning machines are either not available at all or too expensive.

Socially, the cloud-based implementation of the framework will allow deploying teleophthalmology services at scale and thus possibly helping to address some health care inequalities in disadvantaged areas. Still, the framework should undergo necessary validation before being widely implemented.

Conclusion

In the proposed OptiVerse AI framework, there is integration of four components namely, deep learning-based classification, explainable artificial intelligence, computational three-dimensional retinal reconstruction and cloud deployment in one system. With the use of EfficientNetB7 as the backbone network, classification accuracy was attained along with visual explanations using Grad-CAM and reconstruction of anatomically accurate 3D retinal structures from conventional fundus imaging.

From the experimental study results, the proposed method has been proved to provide accurate diabetic retinopathy classification at various levels of disease severity while expanding the possibilities of routine fundus photography with computational retinal reconstruction. This method reduces the need for specialized imaging equipment for 3D visualization thus becoming more affordable and accessible.

All in all, this framework shows the possibilities of the usage of artificial intelligence, explainable learning and computer vision for clinical decision support and retinal screening of diabetic patients.

References

1. LeCun Y, Bengio Y, Hinton G (2015) Deep learning. Nature 521: 436-444. [Crossref] [Google Scholar]

2. Tan M, Le QV (2019) EfficientNet: Rethinking model scaling for convolutional neural networks. Proceedings of the 36th International Conference on Machine Learning. [Google Scholar]

3. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, et al. (2017) Grad-CAM: Visual explanations from deep networks via gradient based localization. Proceedings of the IEEE International Conference on Computer Vision. [Google Scholar]

4. Gulshan V, Peng L, Coram M, et al. (2016) Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 316: 2402-2410. [Crossref] [Google Scholar]

5. Ting DSW, Cheung CY, Wong TY (2016) Diabetic retinopathy: Global prevalence, major risk factors, screening practices and public health challenges. Clinical & Experimental Ophthalmology 44: 260-277. [Crossref] [Google Scholar]

6. Abramoff MD, Lavin PT, Birch M, Shah N, Folk JC (2018) Pivotal trial of an autonomous AI-based diagnostic system for diabetic retinopathy. npj Digital Medicine 1: 39. [Crossref] [Google Scholar] [Indexed]

7. Wilkinson CP, Ferris FL, Klein RE, et al. (2003) Proposed international clinical diabetic retinopathy and diabetic macular edema disease severity scales. Ophthalmology 110: 1677-1682. [Crossref] [Google Scholar] [Indexed]

8. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. [Google Scholar]

9. Simonyan K, Zisserman A (2015) Very deep convolutional networks for large-scale image recognition. International Conference on Learning Representations. [Google Scholar]

10. Dosovitskiy A, et al. (2021) An image is worth 16×16 words: Transformers for image recognition at scale. International Conference on Learning Representations. [Google Scholar]

11. Kingma DP, Ba J (2014) Adam: A method for stochastic optimization. International Conference on Learning Representations. [Google Scholar]

12. Shorten C, Khoshgoftaar TM (2019) A survey on image data augmentation for deep learning. Journal of Big Data 6: 60. [Crossref] [Google Scholar]

13. Ronneberger O, Fischer P, Brox T (2015) U-Net: Convolutional networks for biomedical image segmentation. MICCAI. [Crossref] [Google Scholar]

14. Esteva A, Kuprel B, Novoa RA, et al. (2017) Dermatologist-level classification of skin cancer with deep neural networks. Nature 542: 115-118. [Crossref] [Google Scholar] [Indexed]

15. Samek W, Wiegand T, Müller KR (2017) Explainable artificial intelligence: Understanding, visualizing and interpreting deep learning models. ITU Journal: ICT Discoveries 1: 39-48. [Google Scholar]

16. Litjens G, Kooi T, Bejnordi BE, et al. (2017) A survey on deep learning in medical image analysis. Medical Image Analysis 42: 60-88. [Crossref] [Google Scholar] [Indexed]

17. Fu H, Cheng J, Xu Y, et al. (2018) Joint optic disc and cup segmentation based on multi-label deep network and polar transformation. IEEE Transactions on Medical Imaging 37: 1597-1605. [Crossref] [Google Scholar] [Indexed]

18. Decencière E, Zhang X, Cazuguel G, et al. (2014) Feedback on a publicly distributed image database: The Messidor database. Image Analysis & Stereology 33: 231-234. [Google Scholar]

19. Kaggle (2015) Diabetic retinopathy detection dataset. Kaggle.

20. Zhang Z, Cui P, Zhu W (2020) Deep learning on graphs: A survey. IEEE Transactions on Knowledge and Data Engineering. [Crossref] [Google Scholar]

21. Goodfellow I, Bengio Y, Courville A (2016) Deep learning. MIT Press. [Google Scholar]

22. He K, Gkioxari G, Dollár P, Girshick R (2017) Mask R-CNN. Proceedings of the IEEE International Conference on Computer Vision. [Google Scholar]

23. Lin TY, et al. (2014) Microsoft COCO: Common objects in context. European Conference on Computer Vision. [Crossref] [Google Scholar]