Cureus Journal of Computer Science

Open Access Original Article

Artificial Intelligence-Powered Retinal Fundus Image Evaluation and Synthesis for Ocular Disease Diagnosis and Prevention

Received 02/15/2025
Review began 04/30/2025
Review ended 07/24/2025
Published 08/14/2025

DOI:
https://doi.org/10.7759/s44389-025-03317-w

Wendy Flores-Fuentes [1] , Andre I. Herrera-Chavez [1] , Eder A. Rodríguez-Martínez [1] , Oscar H. Montiel-Ross [2] , Juan C. García-Gallegos [1] , Julio Cesar Rodríguez-Quiñonez [1] , Oleg Yu Sergiyenko [3] , Daniel Hernández-Balbuena [1] , Lenin D. Ochoa de la Paz [4] , Amabile A. Velo-Silvestre [5]

1. Facultad de Ingeniería, Universidad Autónoma de Baja California, Baja California, MEX 2. Centro de Investigación y Desarrollo de Tecnología Digita, Instituto Politécnico Nacional, Baja California, MEX 3. Instituto de Ingeniería, Universidad Autónoma de Baja California, Baja California, MEX 4. Facultad de Medicina, Universidad Nacional Autónoma de México, Ciudad de México, MEX 5. Clínica de Optometría, Universidad Nacional Autónoma de México, Ciudad de México, MEX

Corresponding author: Wendy Flores-Fuentes, flores.wendy@uabc.edu.mx

Abstract

This paper presents a method for generating synthetic fundus images using Neural Style Transfer and Cycle-Consistent Generative Adversarial Network, aimed at enhancing the visualization of retinal diseases, such as diabetic retinopathy. The Cycle-Consistent Generative Adversarial Network approach achieved an overall accuracy of 98.3%, with 100.0% precision, 96.0% recall, and an F1-score of 98.0%. Additionally, a comparison evaluated by health specialists yielded an accuracy of 95.8% for real images. By creating realistic synthetic images, the study seeks to raise awareness about ocular health risks and promote self-care among patients. The effectiveness of the synthetic images is validated against real images through advanced machine learning models, which classify retinopathy stages and identify biomarkers. Evaluations by ocular health specialists further confirm the similarity and utility of the synthetic images for diagnostic standardization and improving AI models in healthcare. The results indicate significant potential for these synthetic approaches to enhance early detection and treatment outcomes in the field of ophthalmology. Limitations and challenges faced were considered and detailed as implications for future work.

Categories: Computer Vision, Image Processing and Analysis, Machine Learning (ML)
Keywords: image processing, synthetic fundus image, retinal disease, stage classification, biomarkers localization and segmentation

Introduction

The importance of sight lies in its ability to provide a visual perception of reality. The eye is the organ that makes this possible and contains many structures, including the retina. The retina has photoreceptor cells that are sensitive to light; these cells capture light and transform it into electrical impulses transmitted via the optic nerve to the brain, enabling vision [1].

However, disorders affecting the retina can diminish quality of life and lead to vision loss. Initially, some sections of the visual field may be affected by sudden onset of disturbances such as floaters and light flickers. This condition is progressive, with symptoms related to changes in vision including altered color perception, blurry vision, and difficulty adapting to different lighting conditions, especially in low illumination or at night [2].

Retinal disease can be caused by physical trauma, aging, and family health history, all of which may lead to retinal disorders. However, its progression can be slowed or even stopped with early diagnosis, appropriate treatment, and a healthy lifestyle. Raising awareness through visualization of retinal damage risk plays a crucial role in prevention.

For this purpose, this work focuses on the synthetic image generation of the main retinal disease, retinopathy, which is primarily caused by complications from diabetes and hypertension. Retinopathy is characterized by fluid leakage resulting from damaged blood vessels in the retina. The goal is to promote prevention by raising awareness of health risks associated with ocular diseases through visual representations that forecast retinal degeneration due to possible proliferation of the disease, taking as base the ocular image with the current stage of the patient.

This approach is motivated by the biomedical sector, where extracting knowledge from images supports decisions in disease diagnosis. Biomedical imaging technology allows collecting big datasets to be investigated to build automatic systems to potentially better support professional decisions [3]. In the

How to cite this article
Flores-Fuentes W, Herrera-Chavez A I, Rodríguez-Martínez E A, et al. (August 14, 2025) Artificial Intelligence-Powered Retinal Fundus Image Evaluation and Synthesis for Ocular Disease Diagnosis and Prevention. Cureus J Comput Sci 2 : es44389-025-03317-w. DOI https://doi.org/10.7759/s44389-025-03317-w

ophthalmic field, the most common imaging modalities are optical coherence tomography, fluorescein angiography, and color fundus photography (CFP) [4].

For early detection, the availability of suitable technology is essential. CFP imaging modality offers advantages such as compatibility with digital smartphone-based photography, as well as widefield and ultra-widefield imaging. These systems can capture detailed retinal images by adding optical accessories like lenses and illumination sources [5].

While artificial intelligence (AI)) revolutionizes image processing for the automatic classification of health conditions [6], it also detects biomarkers for identification and segmentation to lay out their perimeters [7]. This makes possible self-assessments and teleophthalmology instead of complete dependence on specialized health personnel and sophisticated clinical equipment, impacting early treatment decisions to stop proliferation [8].

Besides, these AI developments provide automatic image classification tools for diagnostic standardization. Generative AI, capable of creating images by learning the patterns and structure of training data, can then generate new data with similar characteristics to accomplish this research purpose. Additionally, this development can be utilized as an augmented image tool for further research aimed at improving AI models [9].

The rest of this paper is organized as follows. Section "Related Works" discloses related works and covers a significant brief literature review. Synthetic image generation methods are presented in section "Synthetic Image Generator Methods". Classification, detection, and segmentation methods are described in section "AI Images Evaluation Methods". Section "Retinal Disease Detection Methodology Results" provides the research methodology and results. The evaluation of results by health specialists is summarized in section "Evaluation of Results by Health Specialists". Finally, the conclusion is presented.

Materials And Methods

Related works

Keeping in mind that ocular diseases constitute a major global health issue with increasing prevalence, there is a growing demand not only for early detection and suitable treatment but also for self-care and prevention, goals that can be achieved through accessible automatic technology. This section reviews related works focusing on diabetic retinopathy (DR) stages and biomarkers trademarked in fundus images using computer vision-based approaches, enabling the compilation of relevant materials, methods, and strategies.

Diagnostic techniques can be classified as DR binary detection [10], DR grade detection [11], biomarker detection [12], and biomarker localization and segmentation [13].

Some computational approaches include machine-learning-based methods [14-15], such as neural networks and support vector machines; deep-learning-based methods [16], such as convolutional neural networks (CNNs), multi-layer perceptrons, and transformer-based networks; transfer learning-based methods [17], including InceptionV3, MobileNetV2, DenseNet-121, ImageNet, EfficientNet, VGG, Xception, InceptionResNetV2, NASNetMobile, ResNet, GoogleNet, and AlexNet, to name a few; and synthetic image generation techniques [18], such as Active Shape Model (ASM), patch-based algorithms with model-based texture synthesis, generative adversarial networks (GANs), progressively grown generative adversarial networks (ProGANs), cycle-consistent generative adversarial networks (CycleGAN), style transfer, and diffusion models.

Figure *1* shows a timeline of deep learning-based algorithms in the classification and segmentation of ocular diseases related to retinopathy. The year 2006 is considered the beginning of modern deep learning, when unsupervised layer-wise pretraining was proposed using Deep Belief Networks [19] to train deep neural networks more effectively than traditional methods that suffered from gradient issues. By 2012, deep learning had consolidated with the use of graphics processing units (GPUs) and massive datasets, making backpropagation training feasible. This period saw the breakthrough of the AlexNet architecture in image classification [20]. Between 2012 and 2015, biomedical image analysis began, with 2015 marking the start of retinal vessel and lesion segmentation using architectures like U-Net [21-22], followed by retinal image classification in 2016 [23]. By 2020, advanced classification and segmentation models emerged, incorporating hybrid approaches, transfer learning with pre-trained models, and segmentation using patch-based CNNs and hybrid models [24-25]. New architectures such as transformer-based and multi-modal learning methods were integrated into segmentation and classification pipelines by 2023. The current ongoing trends focus on multitasking, multilabel classification, and generative approaches [26], as proposed in this work.

Image preprocessing techniques related to retinal CFP images include anisotropic diffusion filtering [27], Wiener filtering [28], resizing, green channel extraction, grayscale and gray world conversion [29], and

contrast-limited adaptive histogram equalization [30].

Recognized datasets containing retinal disease CFP images are STARE [31], MESSIDOR [32], EyePACS [33], APTOS [34], IDRiD [35], DIARETDB [36], and ODIR [37].

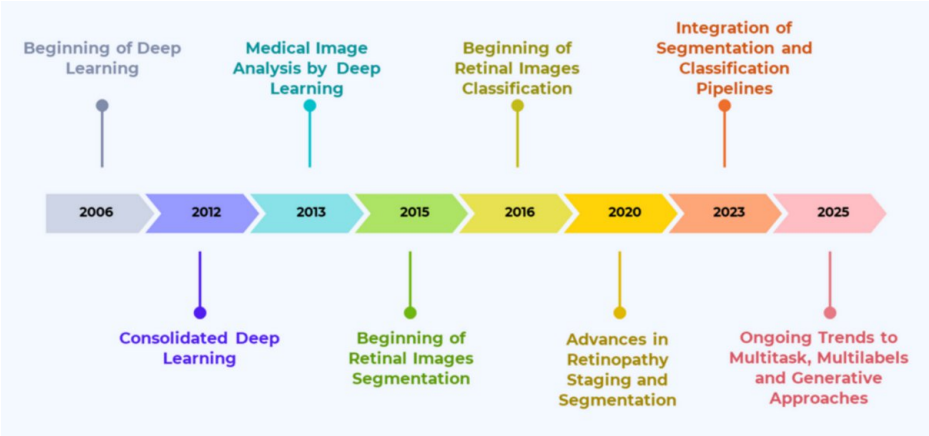

FIGURE 1: Timeline of deep learning-based algorithms for classification and segmentation of diabetic retinopathy

For a further study of previous works, including efforts and achievements with classical AI algorithms and digital image processing techniques, a work by Winder et al. [38] is recommended. This work surveys preprocessing; localization and segmentation of the optic disc; segmentation of retinal vasculature; localization of the macula and fovea; and localization and segmentation of retinopathy using classical methods such as local contrast enhancement by normalization, illumination correction, noise reduction, color normalization, and histogram analysis. Approaches for optic disc localization include vessel position, snakes, principal component analysis, watershed transform, pixel intensity, image thresholding, pyramidal decomposition, edge detection, morphology, Hough transform, point distribution model, hierarchical filter schemes, and maximum contrast. Vessel segmentation techniques involve vessel tracking, matched filtering, morphological analysis, principal component analysis/neural networks, watershed transform, thresholding, filters (steerable/Gaussian), localization via vessel structure, classification/neural networks, edge detection, wavelet transform, and other miscellaneous methods.

Synthetic image generator methods

AI techniques employed to generate synthetic retinal images that visualize stages of ocular diseases are presented.

Neural style transfer (NST) is a technique that synthesizes a new image I_G by optimizing it (Eq. 1) to match the content representation of one image I_C and the style representation of another image I_S#. A pre-trained CNN (typically VGG-19) extracts feature representations from both I_C# and I_S. The content features capture the structural elements of the image, while the style features capture texture and color patterns.

$$I_G \leftarrow I_G - \eta \cdot \alpha \frac{\partial L_{\text{content}}}{\partial I_G} + \eta \cdot \beta \frac{\partial L_{\text{style}}}{\partial I_G} \qquad (1)$$

where η is the learning rate, and α and β are weighting factors that balance the contributions of I_C and I_S to the generated I_G. The optimization minimizes both the content loss L_{content} and the style loss L_{style} [39].

In this study, NST is utilized for visual representations of ocular diseases in normal fundus images. The content image represents the normal fundus, while the style image corresponds to a retina affected by DR [40].

CycleGAN is a variant of GANs developed for unpaired image-based translation, enabling the conversion of images between different domains without requiring matched training samples. In this study, it is used to translate features between stages of DR and normal fundus images. This approach allows for the synthesis of realistic representations of retinal disease progression by transferring pathological features onto real retinal images. It involves two generator networks, $G : X \rightarrow Y$ and $F : Y \rightarrow X$,

which learn mappings from the source domain *X* (e.g., normal fundus images) to the target domain *Y* (e.g., images of fundus with DR), and two discriminators D_X and D_Y, which aim to distinguish between real images and generated images in each domain, through a comprehensive objective function that integrates both adversarial and cycle consistency losses (Eq. 2), where $L_{GAN}(G, D_Y, X, Y)$ and $L_{GAN}(F, D_X, X, Y)$ are the adversarial losses that encourage the generators to synthesize images that are indistinguishable from real images in the target domains. $L_{cyc}(G, F)$ is the cycle consistency loss, which ensures that the mappings *G* and *F* are inverses of each other, thereby preserving the content of the images during translation [41].

$$L(G, F, D_X, D_Y) = L_{GAN}(G, D_Y, X, Y) + L_{GAN}(F, D_X, X, Y) + \lambda L_{cyc}(G, F) \quad (2)$$

Here, λ is a weighting factor that balances the importance of the cycle consistency loss relative to the adversarial losses.

AI images evaluation methods

This section describes the image evaluation methods for both synthetic and real images, based on AI classifiers.

Classification Methods

Deep neural networks consist of input, hidden, and output layers. A network is considered deep when it has several hidden layers connecting the input to the output. The key parameters to consider in these networks encompass the quantity of layers and the choice of activation functions, and how the data flows through the network. When performing classification into distinct classes, the network is designed to estimate the class labels. In contrast, for multi-label classification, the network predicts multiple binary labels for each instance.

There are several approaches to developing models, including building custom models or utilizing transfer learning techniques. Transfer learning leverages pre-trained models that have already learned general features, which can then be fine-tuned for a specific task. In this study, transfer learning was employed using a pre-trained InceptionV3 model (as a result of the previous research [26]), where the top layers were modified to suit both multi-class and multi-label classification tasks.

The performance evaluation metrics used in this study include accuracy (Eq. 3), precision (Eq. 4), true positive rate (also called recall, Eq. 5), F1-score (Eq. 6), and false positive rate (Eq. 7), ensuring comprehensive assessment across both classification tasks.

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (3)$$

$$P = \frac{TP}{TP + FP} \quad (4)$$

$$\mathrm{TPR} = \mathrm{R} = \frac{TP}{TP + FN} \quad (5)$$

$$F1 - Score = 2 \cdot \frac{P \cdot R}{P + R} \quad (6)$$

$$FPR = \frac{FP}{FP + TN} \quad (7)$$

where TP refers to true positive, TN to true negative, FP to false positive, FN to false negative, TPR to True Positive Rate, and FPR to False Positive Rate. All metrics are calculated with ground truths and inferences to evaluate models performances.

The object detection technique was employed to identify and locate specific biomarkers in retinal images using the YOLO (You Only Look Once), a well-known object detection framework and image segmentation model developed by Joseph Redmon and Ali Farhadi at the University of Washington, launched in 2015. Specifically, YOLOv8 model version from Ultralytics [42], through a versatile framework, was designed to cover the entire lifecycle of the machine learning model.

The segmentation method was applied to segment biomarkers of DR in fundus images using a Segment

Anything Model (SAM) [43]. This model consists of three main components: an image encoder, a prompt encoder, and a mask decoder. The image encoder, based on a pre-trained Vision Transformer (ViT) [44-48], processes the input image before any prompts are applied. The prompt encoder supports both sparse prompts (such as points, boxes, and text) and dense prompts (such as masks), using positional encodings for sparse prompts and convolutions for dense ones. The mask decoder merges the image and prompt embeddings to output a mask. Since the generic SAM was not sufficient for accurately segmenting DR biomarkers, fine-tuning was performed to adapt the model to this specific task.

The segmentation process was performed using a GPU on Kaggle. The performance of the fine-tuned model was assessed using the Dice score metric (Eq. 8), the Area Under the Precision-Recall Curve (AUC-PR) (Eq. 9), and the Area Under the Receiver Operating Characteristic Curve (AUC-ROC) (Eq. 10).

$$Dice = \frac{2 \cdot |A \cap B|}{|A| + |B|} \tag{8}$$

Dice score measures the overlap between the predicted segmentation *A* and the ground truth segmentation *B*. The intersection $|A \cap B|$ counts the number of pixels common to both segmentations, while |*A*| and |*B*| are the total number of pixels in the predicted and ground truth masks.

$$AUC - PR = \sum_{i=1}^{n-1} \frac{P_{i+1} + P_i}{2} \cdot (R_{i+1} - R_i) \tag{9}$$

The AUC-PR evaluates the trade-off between precision (Eq. 4) and recall (Eq. 5). The area under a curve is constructed from many discrete points (Precision-Recall pairs), computed by the trapezoidal rule.

$$AUC - ROC = \sum_{i=1}^{n-1} (FPR_{i+1} - FPR_i) \cdot \frac{TPR_{i+1} + TPR_i}{2} \tag{10}$$

The AUC-ROC evaluates the model's capacity to differentiate between classes by plotting TPR (recall) versus FPR.

Results

Retinal disease detection methodology results

From now on, multi-class classification refers to DR stages: C0 (No DR), C1 (mild), C2 (moderate), C3 (severe), and C4 (proliferative), while multi-label classification refers to biomarkers: L0 (no biomarker), L1 (microaneurysm), L2 (hemorrhage), L3 (hard exudate), and L4 (soft exudate), as shown in Figure *2*. The methodology flow is depicted in Figure *3*.

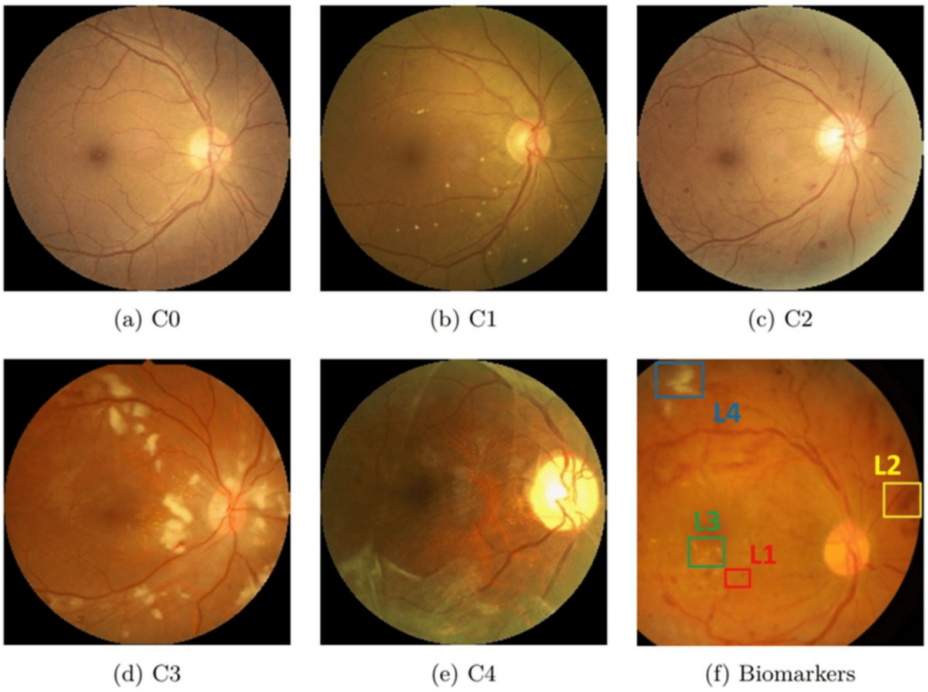

FIGURE 2: Retinal stages and biomakers: (a) no DR, (b) mild, (c) moderate, (d) severe, (e) proliferative, and (f) biomarkers

DR, Diabetic Retinopathy

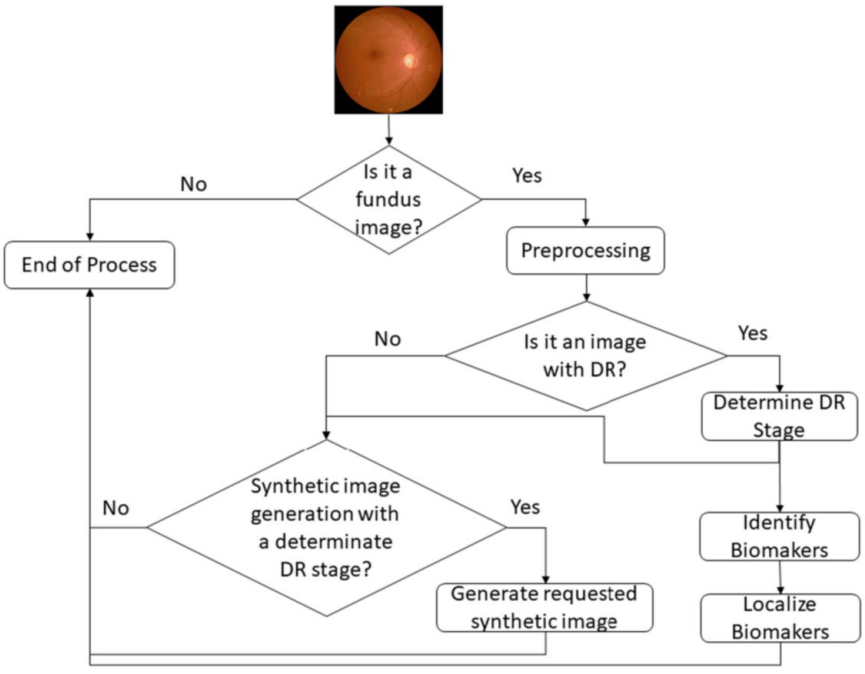

FIGURE 3: Flowchart of the experimentation methodology

DR, Diabetic Retinopathy

Datasets and labeling methods

In this research, the APTOS [34], ODIR-5K [37], EyePacs [33], IDRID [35] and Messidor-2 [32] datasets were employed for different image generation and classification tasks, each tailored to its specific objectives. In

some experiments, the datasets required to be re-labeled according to the model constructed; the labeling methods for each dataset are described below.

The APTOS dataset includes fundus images categorized into five classes, including normal fundus and four stages of DR. For experimentation, 3,031 images were used. This dataset was the primary source for CycleGAN, providing all the images for training.

The ODIR-5K dataset includes 8,000 fundus images covering a variety of retinal diseases. For this study, 1,347 images corresponding to different stages of DR and 1,253 images with normal fundus were selected. This dataset was used for multi-label classification, focusing on the detection of biomarkers. A one-hot vector strategy was employed, where labels were created from scratch. Columns were added for specific biomarkers, including microaneurysms, hemorrhages, hard exudates, and soft exudates. Each image's one-hot vector was assigned a "1" or "0" depending on the presence or absence of these biomarkers. For object detection, 200 images were selected, and bounding boxes for DR biomarkers were manually annotated using the makesense.ai [49] tool, with annotations saved in text format. Finally, these images were used in NST and CycleGAN experiments.

The EyePACS dataset includes a collection of fundus images categorized into five classes: normal fundus and the four stages of DR. The dataset contains a total of 30,262 images, covering both normal fundus and various stages of DR. This dataset was also used for training CycleGAN.

The IDRID dataset provides 81 detailed fundus images containing DR, along with annotations to localize biomarkers. This dataset was used specifically for segmentation tasks.

The Messidor-2 dataset originally contains 1,200 fundus images of DR. However, for this experiment, only 197 images were used, as masks were only available for these images. These masks, provided by Gabriel Lepetit-Aimon, contain detailed annotations of DR biomarkers. The dataset was specifically used for segmentation tasks in this experiment. Table *1* resumes the use of datasets.

Model	Model Function	Dataset Used	Details
Multi-class	Clasify DR stage by CNN model (Inception V3), while categorical cross-entropy with a softmax output layer	APTOS	Images were evenly distributed across the five retinopathy classes using data augmentation. Both single-step and two-step classification approaches were explored. The two-step approach, involving an initial binary classifier followed by a four-class model, yielded the highest performance.
Multi-label	Identify biomarkers present in an image with DR by CNN model (InceptionV3) with binary cross-entropy and a sigmoid output layer	ODIR-5K	A one-hot encoding scheme was implemented to annotate four biomarkers per image. Labels were generated manually, and the best performance was achieved when the model was applied only to previously identified DR cases, reducing false positives from normal images.
Segmentation	Segment specific biomarkers using the SAM model with a ViT-Base backbone and Dice Cross-Entropy loss	IDRID, ODIR-5K, MESSIDOR-2	Each model was trained separately for microaneurysms, hemorrhages, and exudates using manually aligned masks. The SAM model was fine-tuned only at the mask decoder level, preserving pre-trained encoder weights. Performance was evaluated using the Dice coefficient, with favorable results across all biomarkers.
Object Detection	Detect and localize biomarkers using YOLOv8 and bounding box annotations in YOLO format	IDRID, ODIR-5K	Manual annotations were created using makesense.ai, following the YOLO format. The model was trained with a custom configuration file and evaluated using bounding box overlap metrics. The detection of soft exudates achieved the highest F1-score among the biomarkers.
Neural Style Transfer	Transfer the visual style of DR images onto normal fundus content images using VGG19 for feature extraction	ODIR-5K	A single content image and one DR style image were used to synthesize a new fundus image. The model combined content and style representations extracted from intermediate convolutional layers, optimized over multiple iterations with weighted loss functions.
CycleGAN	Generate synthetic DR images from normal fundus images for each DR stage using adversarial training and cycle-consistency loss	APTOS, ODIR-5K, EyePACS	Four independent models were trained to map normal fundus images to each of the four DR stages. The architecture incorporated identity and cycle-consistency constraints to preserve anatomical features while achieving visual fidelity in the generated images.

TABLE 1: Dataset used details

DR, Diabetic Retinopathy; CNN, Convolutional Neural Network; SAM, Segment Anything Model; YOLO, You Only Look Once; NST, Neural Style Transfer; CycleGAN, Cycle Generative Adversarial Networks

Preprocessing for classification and synthetic images generation models

Before training the classification (multi-class and multi-label) and synthetic image generation models, the preprocessing methods (a, b,.., j) in Table *2* were performed as described below to remove noise and improve the effectiveness of subsequent model training. The preprocessed images were stored in a separate folder for better organization and workflow efficiency.

Model	a	b	c	d	e	f	g	h	i	j
Multi-class	•	•	•	-	-	-	•	-	-	•
Multi-label	•	•	•	-	-	-	•	-	-	•
Segmentation	•	•	-	•	-	-	-	•	-	-
Object Detection	•	•	-	-	-	•	-	-	-	-
Neural Style Transfer	•	•	-	-	•	-	-	-	•	-
CycleGAN	•	•	•	-	-	-	-	•	-	-

TABLE 2: Preprocessing steps applied to different models

CycleGAN, Cycle Generative Adversarial Networks

a) Image cropping process to remove irrelevant black borders and highlight the central region of the image.

b) Conversion to RGB color space to standardize the format.

c) Adjusting all images to a standardized size of 224 × 224 pixels.

d) Adjusting all images to a standardized size of 256 × 256 pixels.

e) Adjusting all images to a standardized size of 400 × 400 pixels.

f) Adjusting all images to a standardized size of 640 × 640 pixels.

g) Normalizing pixel values to the range 0, 1.

h) Normalizing pixel values to the range -1, 1.

i) Normalization of pixel values was performed using means of 0.485, 0.456, and 0.406 and standard deviations of 0.229, 0.224, and 0.225.

j) Applying data augmentation techniques such as horizontal flipping and rotation up to 20 degrees.

Architecture for classification models

InceptionV3 [47], a deep CNN pre-trained on the ImageNet dataset, was used in this development for both multi-class and multi-label classification of retinal images.

InceptionV3 was adapted for classifying DR stages and detecting biomarkers in fundus images. The model's convolutional layers were initialized with ImageNet weights and kept frozen to retain general feature extraction capabilities, while the last 32 layers were unfrozen for fine-tuning. The global average pooling and dense layers were modified to suit the specific requirement by batch normalization, for the transfer learning model architecture, considering the change in output layers according to the approaches. The Adam optimizer was applied with an initial learning rate of 0.0001. Early stopping with patience of 10 epochs and the ReduceLROnPlateau callback (reducing the learning rate by 0.2 after 5 epochs without improvement) were used to optimize training. A batch size of 16 was used.

The main difference between the two kinds of classifications is the loss function and final layer output function.

For the multi-class classification, categorical cross-entropy was used, with a softmax output layer corresponding to the approaches shown in Table *3*, to classify the health and four stages of DR.

Model	Accuracy	Precision	Recall	F1-Score
2 Classes	0.9579	0.9568	0.9584	0.9575
4 Classes	0.9297	0.9278	0.9395	0.9332
2 and 4 Classes	0.9444	0.9035	0.9451	0.9215
5 Classes	0.9299	0.9124	0.9320	0.9214

TABLE 3: Multi-class test results

For the multi-label classification, binary cross-entropy was used, with a sigmoid output layer corresponding to the approaches shown in Table 4, for detecting the absence or presence of multiple biomarkers in each image.

Task	Precision	Recall	F1-Score
4 Labels	0.9770	0.9498	0.9551
5 Labels	0.9173	0.9078	0.9049

TABLE 4: Multi-label test results

The architecture for both the multi-class and multi-label classifiers is shown in Figure 4. Both models were trained in the Kaggle environment with GPU, to obtain the best performance, 5-fold cross-validation was employed, and the models were evaluated using the parameters detailed in the "Classification Methods" section.

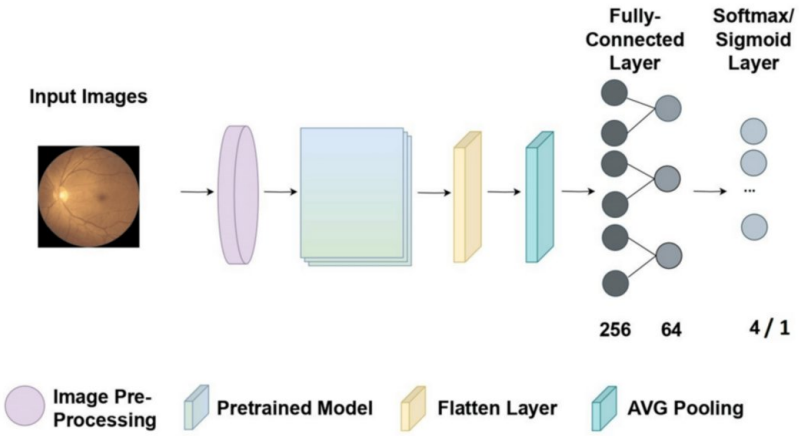

FIGURE 4: CNN architecture of the classifiers multi-classes and multi-labels models

CNN, Convolutional Neural Network

DR stage classification model

This experiment evaluated two approaches for classifying retinal stage images. The first approach used a two-class model to classify images as either normal fundus or DR (Table 3, row 1). If DR was detected, the image was passed to the next model to classify its stage with four classes (Table 3, row 2). If the image was normal, no further processing was required. The results are shown in Table 3, row 3. The second approach employed a single five-class model, which classifies images directly into either normal fundus or one of the four stages of DR (Table 3, row 4).

The goal of comparing these approaches is to determine whether using a two-step process with a simpler initial classification would improve the performance of the stage classification model, concluding that the first approach presents the highest accuracy with a value of 0.9444. Test confusion matrix is shown in Figure *6a*.

Figure *5* shows an example of the prediction of images from the four DR stages for each dataset to visually and analytically validate that the two-step process DR stage classification model.

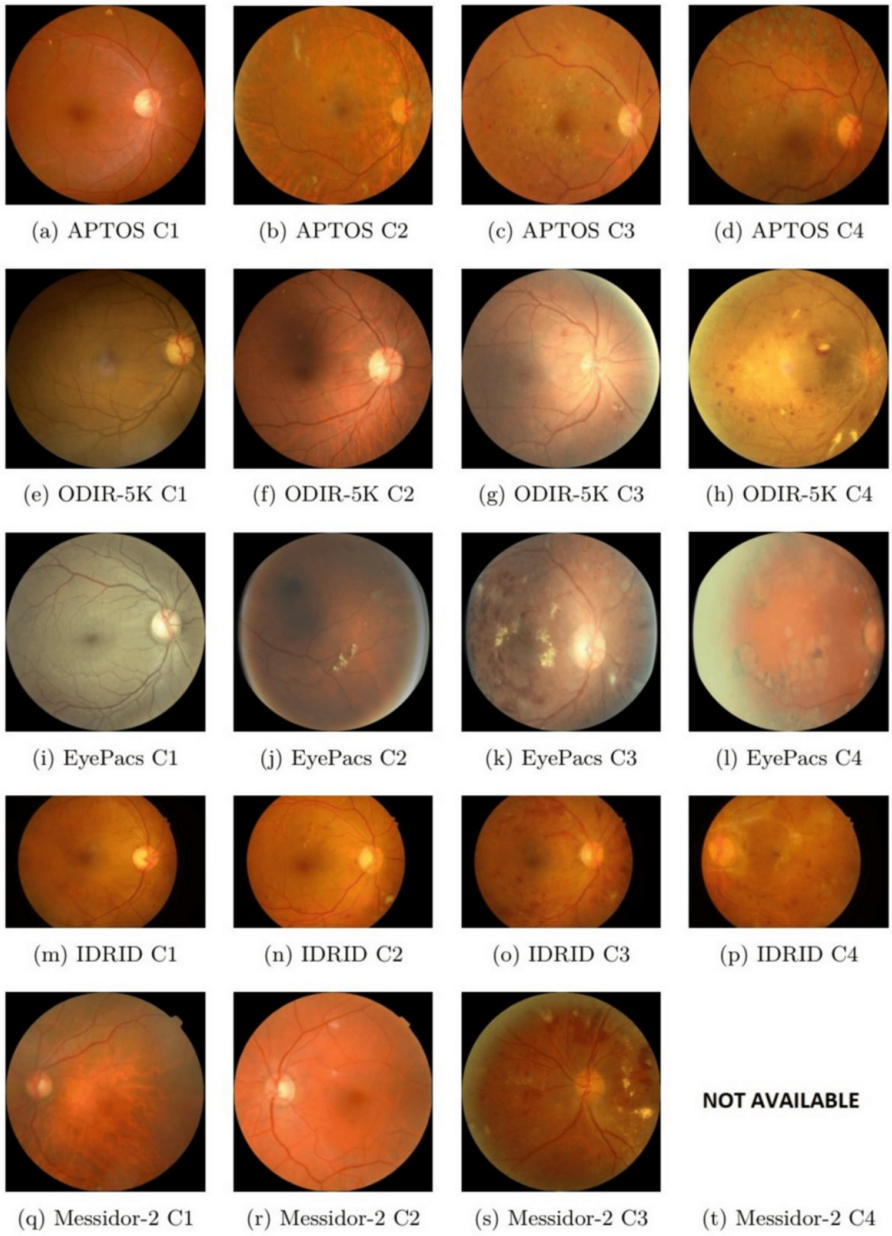

FIGURE 5: Pre-labeled images from diverse datasets, with the DR stage correctly predicted. Row for each datasets, and columns for DR stage

DR, Diabetic Retinopathy

DR biomarkers identification model

This experiment focused on detecting DR biomarkers, such as microaneurysms, hemorrhages, hard exudates, and soft exudates, using the preprocessing described in Table *1* and model architecture detailed in section "Architecture for Classification Models."

In the four-label approach, the biomarker detection model was applied only to images previously

classified as DR (Table 4, row 1). If an image was classified as a normal fundus, the biomarker model was not applied, while a five-class approach was applied directly to all images (Table 4, row 2). Test confusion matrix is shown in Figure 6b. This comparison was conducted to determine whether applying the biomarker model after an initial two-class classification improves performance compared to applying it directly in the five-class classification, concluding the first approach proposed in section "DR stage classification model" is the best methodology.

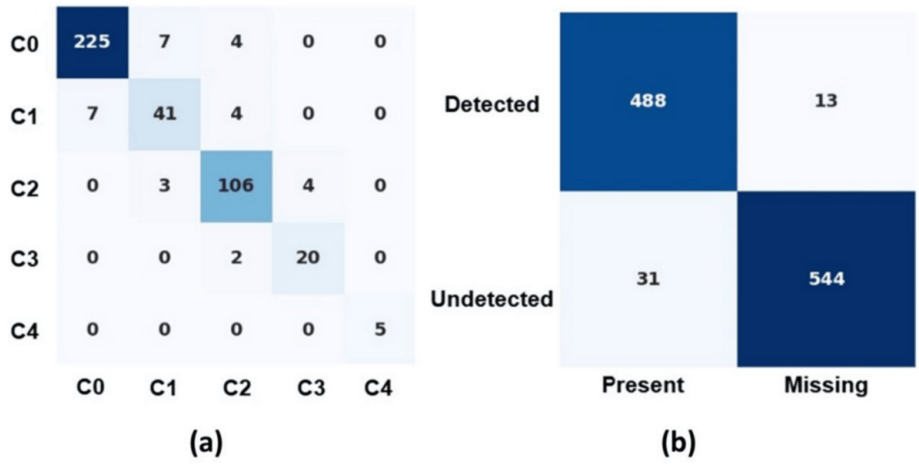

FIGURE 6: Test confusion matrix: (a) DR stage multi-class model. (b) Biomarkers multi-label model

DR, Diabetic Retinopathy

Object detection and segmentation

The object detection process for biomarkers used a pre-trained YOLOv8 model from Ultralytics. New annotations were developed as part of this study, following the YOLO format, which includes information on the class of the object and its localization, with files named to match their corresponding images. After the annotation process, the dataset was split, with 80% of the images (160) used for training and 20% (40) for testing. A .yaml file was created to specify the dataset paths, number, and type of biomarkers. The model was trained for 120 epochs using the YOLOv8 architecture [50] and exported in ONNX format. To evaluate the performance, the predicted bounding boxes were compared with the known bounding boxes in terms of overlap. The evaluation metrics used were precision, recall, and F1-score, with results shown in Table 5.

Biomarkers	Precision	Recall	F1-Score
Microaneurysm (L1)	0.946	0.192	0.321
Hemorrhage (L2)	0.944	0.670	0.784
Hard Exudate (L3)	0.944	0.510	0.662
Soft Exudate (L4)	1.000	0.775	0.873

TABLE 5: Object detection test results

Segmentation was applied to identify specific regions within retinal images corresponding to identified biomarkers. It required training three models, one for each of the biomarkers. No distinction was made between soft and hard exudates because the dataset masks does not distinguish between them.

For each biomarker model, the masks' regions of interest were aligned with retinal images before the cropping process. Bounding boxes were generated from the masks to capture the regions of interest. The dataset was partitioned, with 80% allocated for training and 20% for testing. The bounding boxes and retinal images were fed as inputs for the model training. SAM was used with a ViT base backbone, initialized with pre-trained weights. During fine-tuning, only the parameters of the mask decoder were updated, while the vision and prompt encoders remained frozen to preserve their learned representations.

The model was trained over 30 epochs using the Adam optimizer with a learning rate of 1×10^{-5}. The dice cross-entropy loss function was employed to balance pixel-level accuracy and region-based segmentation performance. Throughout training, the model's predictions were resized to match the ground truth masks, and gradients were computed to update the mask decoder's parameters, thereby enhancing segmentation accuracy. Model performance was assessed using the Dice coefficient, which measures the overlap between predicted masks and ground truth masks. The results, summarized in Table 6, demonstrate the effectiveness of the SAM ViT Base model in accurately segmenting biomarkers within retinal images.

Biomarkers	Dice	AUC-PR	AUC-ROC
Microaneurysm (L1)	0.5454	0.5562	0.8033
Hemorrhage (L2)	0.6481	0.6761	0.8124
Hard Exudate (L3, L4)	0.6635	0.6890	0.8910

TABLE 6: Dice score results for diabetic retinopathy biomarker segmentation

AUC-PR, Area Under the Precision-Recall Curve; AUC-ROC, Area Under the Receiver Operating Characteristic Curve

Synthetic image generation

AI methods were implemented for synthetic image generation. NST was used to blend the content of normal fundus images (Figure 7a) with the style of DR images (Figure 7b). A pre-trained VGG19 model was employed, focusing solely on the feature extraction layers to carry out the style transfer, as described in the article.

The training process was carried out for 250 iterations using the Adam optimizer, with a learning rate of 0.003. The content features and style features were obtained from multiple convolutional layers. The following style features were weighted:

conv1_1: 1.0
conv2_1: 0.75
conv3_1: 0.2
conv4_1: 0.2
conv5_1: 0.2

The content loss weighted 1, and the style loss was weighted at 1e9. At each iteration, the generated NST synthetic image was adjusted based on the combined content and style losses, as seen in Figure 7c.

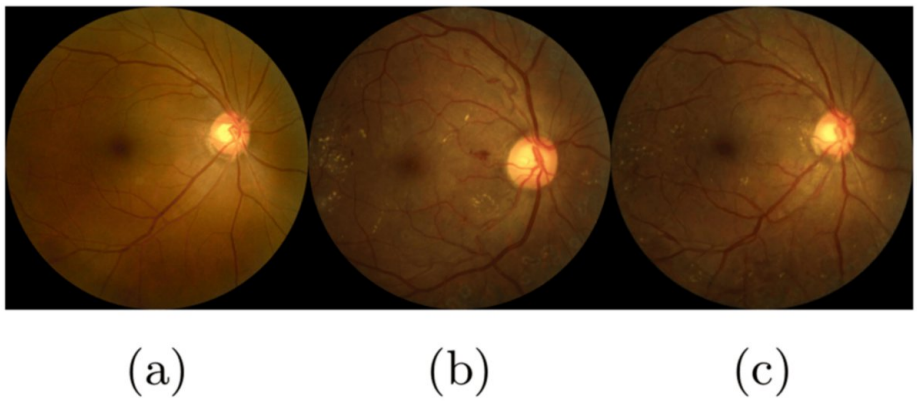

FIGURE 7: Results using NST: (a) Content image (normal fundus). (b) Style image (severe non-proliferative diabetic retinopathy). (c) NST synthetic image

NST, Neural Style Transfer

CycleGAN was employed to convert images between two domains: normal fundus images (Figure 8a) and

synthetic DR images (Figure 8b). Four models were generated, one for each retinopathy stage. For each model, a total of 2,200 images corresponding to the stage of DR and 1,500 normal fundus images were sourced from the APTOS, ODIR-5K, and EyePACS datasets. The architecture consisted of two generator models and two discriminator models, following the structure outlined by Wang et al. [41]. The image resolution was reduced to extract features and then restored to the original size during the generation process.

Cycle consistency loss was applied along with identity loss to ensure that images retained their original characteristics. Additionally, generator loss and discriminator loss were used to evaluate the performance of the model. During training, the Adam optimizer was used with a learning rate of 0.0002 for both the generators and discriminators, running for a total of 110 epochs. The results of this process can be observed in Figure 8c.

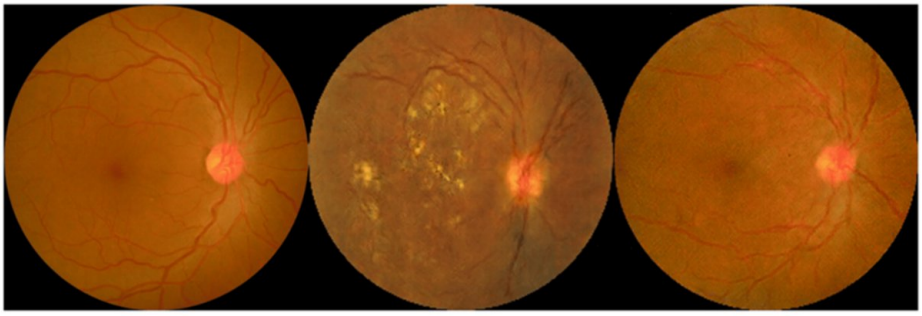

FIGURE 8: Results using CycleGAN: (a) Normal fundus image. (b) Generated image from the severe non-proliferative diabetic retinopathy model. (c) Applying cycle consistency

CycleGAN, Cycle Generative Adversarial Networks

Evaluation of results by health specialists

An experiment was designed to compare the DR stage classes and biomarker labels from the datasets to further evaluate the developed models against health specialists' (HS) criteria, using a double-blind experiment. Researchers from Universidad Autónoma de Baja California and Instituto Politécnico Nacional had no direct contact with the HS and were unaware of their identities. The HS received no information about the project and only provided the requested information. Meanwhile, researchers from Universidad Nacional Autónoma de México linked and coordinated the experimental design with HS.

A batch of 95 images was provided to a group of three HS to a) assign to each of them an DR stages C0, C1, C2, C3, or C4; and (b) assign to each of them all the biomarkers found - L0, L1, L2, L3, L4. The batch was composed of 55 real images and 40 synthetic images, covering all stages and biomarkers. The images were randomly sorted to prevent any pattern that could bias the HS evaluations.

Real Images DR Stages Evaluation

The 55 real images in the batch were made up of 15 C0, 10 C1, 10 C2, 10 C3, and 10 C4 images. These were evaluated by comparing: a) dataset information versus model predictions, obtaining an overall accuracy of 95.8%, precision of 90.8%, recall of 88.0%, and F1-score of 88.4%; and (b) dataset information versus HS detection, with overall accuracy of 82.6%, precision of 51.0%, recall of 52.6%, and F1-score of 50.0%. These results confirm that the models can be efficiently trained to provide consistent results, which can be utilized for diagnostic standardization, while human evaluations present variations, related to natural aspects of humans like perceptions, knowledge, experience, and criteria variables, as well as daily conditions.

Synthetic DR Stages Evaluation

From the 40 synthetic images in the batch, 20 were generated by the NST approach and 20 by the CycleGAN model. For both approaches, no C0 images were generated, and each approach consisted of five

images each for stages C1, C2, C3, and C4. Regarding NST style images DR stage evaluation, results showed 70.8% accuracy, 58.3% precision, 34.3% recall, and 32.8% F1-score for NST generation versus model prediction. Meanwhile, 71.0% accuracy, 49.3% precision, 29.8% recall, and 35.0% F1-score were obtained for NST generation versus HS detection.

Regarding CycleGAN images DR stage evaluation results, an accuracy of 65.0%, precision of 26.3%, recall of 23.8%, and F1-score of 22.3% were obtained for CycleGAN generation versus model prediction. Meanwhile, an accuracy of 70.8%, precision of 29.0%, recall of 25.0%, and F1-score of 27.0% were observed for CycleGAN generation versus HS detection. This indicates that controlling the type of DR stage that is desired to be generated is still a challenge, but they can be generated and classified by the model to be used according to the stage in which they were identified.

Real Images Biomarkers Evaluation

Biomarkers in real images were evaluated using the same batch as in the section "Real Images DR Stages Evaluation." Regardless of the DR stage, the presence (P) and absence (A) of different types of biomarkers in the 55 images were labeled and either detected (D) or undetected (U) by both the biomarker model described in the section "DR biomarkers identification model" and by the HS. Results showed that the model had the highest detection rates for L0, L1, and L2 biomarkers. However, for L3 and L4, which correspond to hard and soft exudates, the model was more conservative than HS in assigning the presence of exudates, regardless of type. Therefore, a new label, L3,4, was evaluated (sum both kinds of exudates). For model detection, an overall accuracy of 92.0%, precision of 94.3%, recall of 83.7%, and F1-score of 86.7% were obtained, compared to HS detection results of 84.3% accuracy, 83.3% precision, 83.0% recall, and 81.6% F1-score.

Of the 15 real images, three were selected to visualize the biomarker detection and segmentation process described in the section "Object Detection and Segmentation," according to the architecture shown in Figure *9*. Figure *10* corresponds to a dataset "Normal Fundus Image" labeled C0, but classified as C1 by the DR stage model, with microaneurysm detected by the biomarker model. Figure *11* corresponds to a dataset labeled "Severe" C3, classified as C3 by the DR stage model, with exudate detected by the biomarker model. Figure *12* also corresponds to a "Severe" C3 dataset, classified as C3 by the DR stage model, with both microaneurysms and exudate detected by the biomarker model.

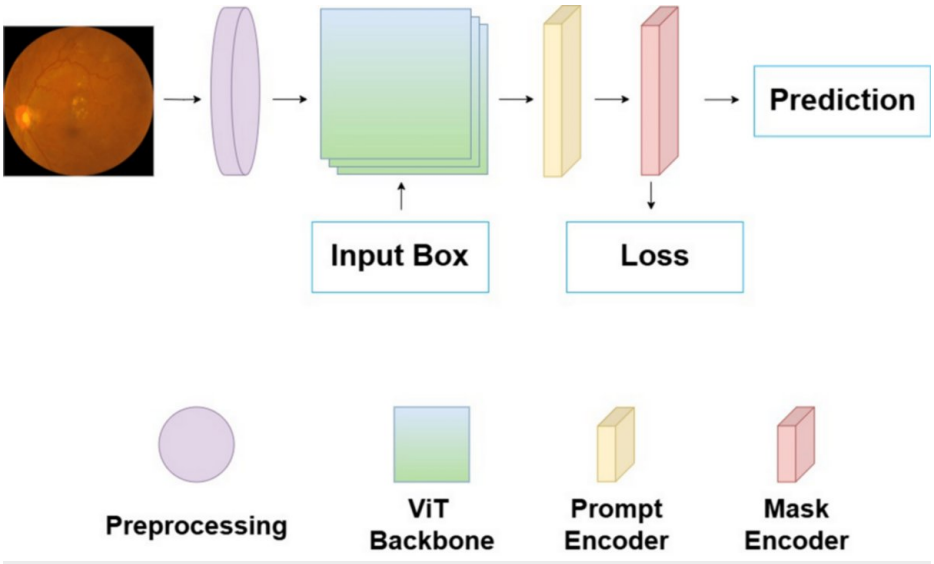

FIGURE 9: Segment Anything Model architecture

ViT, Vision Transformer

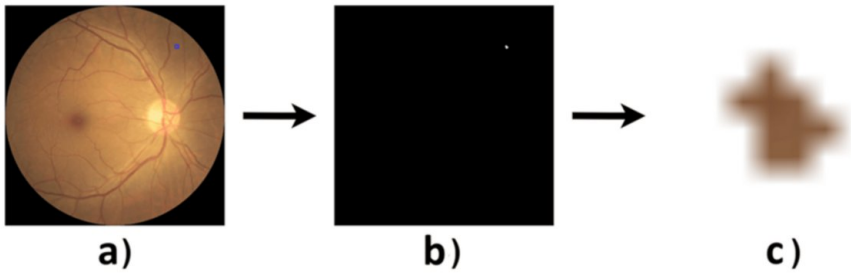

FIGURE 10: Results using SAM: (a) Selected image 1. (b) Mask generated by the segmentation model. (c) Segmented microaneurysm

SAM, Segment Anything Model

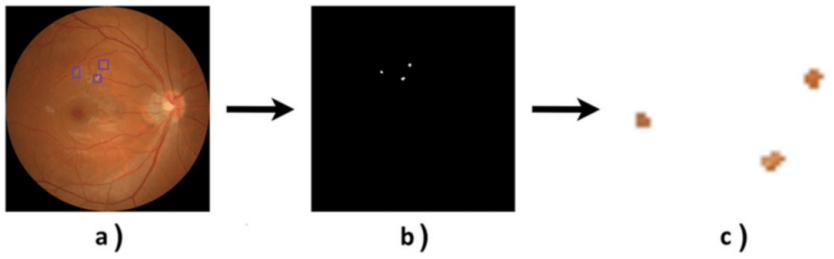

FIGURE 11: Results using SAM: (a) Selected image 2. (b) Mask generated by the segmentation model. (c) Segmented exudates

SAM, Segment Anything Model

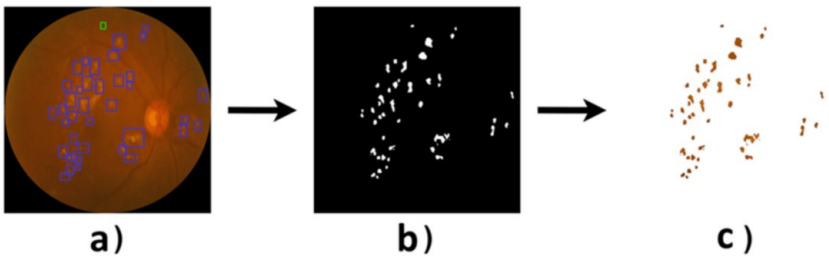

FIGURE 12: Results using SAM: (a) Selected image 3. (b) Mask generated by the segmentation model. (c) Segmented microaneurysm and exudates

SAM, Segment Anything Model

Synthetic Images Biomarkers Evaluation

Biomarkers in synthetic images were evaluated using the same batch as in the section "Synthetic DR Stages Evaluation." Results for the NST approach in model detection showed an overall accuracy of 92.5%, precision of 98.0%, recall of 77.3%, and an F1-score of 94.8%. Meanwhile, for HS detection, accuracy was 69.2%, precision 76.3%, recall 70.3%, and F1-score 67.8%.

Discussion

Experimental setup

The study employed a rigorous experimental methodology to evaluate the effectiveness of synthetic image generation in enhancing the classification, detection, and segmentation of retinal diseases. A

dataset comprising 95 retinal images was utilized, which included 55 real images and 40 synthetic images generated through NST and CycleGAN. The real images were categorized into five stages of DR (C0 to C4), while the synthetic images were designed to simulate these stages and associated biomarkers.

Model evaluation

To assess the performance of the developed models, key metrics such as accuracy, precision, recall, and F1-score were calculated. The properties of the models were evaluated against the expert diagnoses made by HS in a double-blind experiment setting, ensuring unbiased comparisons.

CycleGAN Performance: The CycleGAN achieved a remarkable accuracy of 98.3% with a precision of 100.0% and a recall of 96.0%. This high level of performance indicates the model's ability to accurately identify and classify retinal disease types from synthesized images.

HS Comparison: In contrast, the evaluations made by HS exhibited lesser agreement with the dataset, showing an accuracy of 82.6%. This discrepancy highlights the challenges and variability inherent in human assessments, influenced by personal experience and subjective interpretation of image features.

Discussion of results

The experimental results validate the hypothesis that synthetic images can augment training datasets, leading to improved classification and object detection capabilities in automated diagnostic systems. The high performance of the CycleGAN model illustrates its potential in generating realistic images that maintain the characteristics necessary for effective model training. Additionally, the study revealed that HS could not distinguish the synthetic images from real ones effectively, suggesting that NST- and CycleGAN-generated images had achieved a level of realism that could be beneficial in clinical applications.

Moreover, the comparative analysis underscores the importance of integrating AI with traditional diagnostic methods. While human specialists provide valuable insights based on their expertise, AI models can standardize diagnostic processes, reduce variability in assessments, and deliver rapid results conducive to timely interventions.

Implications for future research

The findings of this study pave the way for several important avenues of future research:

Customization of Models: Future efforts will target the development of customized models tailored to specific healthcare environments, incorporating specific biomarkers and retinal degeneration stages to refine clinical relevance.

Longitudinal Studies: Further longitudinal studies are essential to evaluate the effectiveness of automated diagnostic tools in real-world scenarios and their impact on patient outcomes.

Wider Applications: The methodologies developed in this research could be adapted for other ocular diseases, expanding the application of synthetic image generation beyond DR to conditions such as age-related macular degeneration or glaucoma.

Conclusion of analysis

In conclusion, the experimental analysis confirms that synthetic image generation techniques can significantly contribute to the field of ocular disease diagnostics. By augmenting traditional datasets with high-quality synthetic images, researchers and clinicians can leverage advanced machine learning approaches to improve diagnostic accuracy, optimize patient care, and ultimately mitigate the risks of vision loss.

Conclusions

The findings of this research underscore the critical role of AI in enhancing the diagnosis and management of retinal diseases, particularly DR. By employing NST and CycleGAN for the generation of synthetic retinal images, we have demonstrated a significant potential in augmenting the diversity of datasets used for training diagnostic models. The results indicate that our models achieved an impressive overall accuracy of 98.3% in detecting retinal diseases, with substantial improvements in precision and recall metrics compared to human assessments. The incorporation of synthetic images has proven to be effective, as evidenced by the ability of HS to identify both real and synthetic images without discernible bias, suggesting that our models can be reliably adopted in clinical settings. Moreover, the creation of visually compelling synthetic images serves to raise awareness among patients regarding the risks of ocular diseases, fostering early diagnosis and intervention. Looking ahead, future research will focus on

customizing these models to cater to specific healthcare sectors, incorporating labeled retinal degeneration stages and relevant biomarkers. This tailored approach aims to optimize the applicability of our findings and amplify their social and health impacts. In summary, the integration of advanced image generation techniques represents a transformative step towards improving ocular disease management, enabling timely diagnosis, and ultimately preventing vision loss.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Wendy Flores-Fuentes, Andre I. Herrera-Chavez, Eder A. Rodríguez-Martínez, Oscar H. Montiel-Ross, Juan C. García-Gallegos, Julio Cesar Rodríguez-Quiñonez, Oleg Yu Sergiyenko

Acquisition, analysis, or interpretation of data: Wendy Flores-Fuentes, Andre I. Herrera-Chavez, Eder A. Rodríguez-Martínez, Oscar H. Montiel-Ross, Juan C. García-Gallegos, Daniel Hernández-Balbuena, Lenin D. Ochoa de la Paz, Amabile A. Velo-Silvestre, Julio Cesar Rodríguez-Quiñonez, Oleg Yu Sergiyenko

Drafting of the manuscript: Wendy Flores-Fuentes, Andre I. Herrera-Chavez, Eder A. Rodríguez-Martínez

Critical review of the manuscript for important intellectual content: Wendy Flores-Fuentes, Eder A. Rodríguez-Martínez, Oscar H. Montiel-Ross, Juan C. García-Gallegos, Daniel Hernández-Balbuena, Lenin D. Ochoa de la Paz, Amabile A. Velo-Silvestre, Julio Cesar Rodríguez-Quiñonez, Oleg Yu Sergiyenko

Disclosures

Human subjects: All authors have confirmed that this study did not involve human participants or tissue. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Acknowledgements

Data, analytic methods, and study materials are available upon request.

References

1. Abràmoff MD, Garvin MK, Sonka M: Retinal imaging and image analysis. IEEE Reviews in Biomedical Engineering. 2010, 3:169-208. 10.1109/rbme.2010.2084567
2. American Academy of Ophthalmology. (2025). Accessed: 26 April 2025: https://www.aao.org/.
3. Cesarelli G, Ponsiglione AM, Sansone M, Amato F, Donisi L, Ricciardi C: Machine learning for biomedical applications. Bioengineering. 2024, 11:790. 10.3390/bioengineering11080790
4. Shahriari MH, Sabbaghi H, Asadi F, Hosseini A, Khorrami Z: Artificial intelligence in screening, diagnosis, and classification of diabetic macular edema: A systematic review. Survey of Ophthalmology. 2022, 68:42-53. 10.1016/j.survophthal.2022.08.004
5. Kąpa M, Koryciarz I, Kustosik N, Jurowski P, Pniakowska Z: Modern approach to diabetic retinopathy diagnostics. Diagnostics. 2024, 14:1846. 10.3390/diagnostics14171846
6. Senapati A, Tripathy HK, Sharma V, Gandomi AH: Artificial intelligence for diabetic retinopathy detection: A systematic review. Informatics in Medicine Unlocked. 2024, 45:101445. 10.1016/j.imu.2024.101445
7. Ragab MG, Abdulkader SJ, Muneer A, et al.: A comprehensive systematic review of YOLO for medical object detection (2018 to 2023). IEEE Access. 2024, 12:57815-57836. 10.1109/access.2024.3386826
8. Succar T, Grigg J, Beaver HA, Lee AG: A systematic review of best practices in teaching ophthalmology to medical students. Survey of Ophthalmology. 2015, 61:83-94. 10.1016/j.survophthal.2015.09.001
9. Heinke A, Radgoudarzi N, Huang BB, Baxter SL: A review of ophthalmology education in the era of generative artificial intelligence. Asia-Pacific Journal of Ophthalmology. 2024, 13:100089. 10.1016/j.apjo.2024.100089
10. Padmanayana, Anoop BK: Binary classification of DR-diabetic retinopathy using CNN with fundus colour images. Materials Today Proceedings. 2022, 58:212-216. 10.1016/j.matpr.2022.01.466
11. Chaudhary PK, Pachori RB: Automatic diagnosis of different grades of diabetic retinopathy and diabetic macular edema using 2-D-FBSE-FAWT. IEEE Transactions on Instrumentation and Measurement. 2022, 71:1-9. 10.1109/tim.2022.3140437
12. Maldhure P, Ganorkar SR: Hybrid convolutional neural network for detection of microaneurysms and exudates in retinal images. Journal of Integrated Science and Technology. 2024, 13:1004. 10.62110/sciencein.jist.2025.v13.1004
13. Khojasteh P, Aliahmad B, Kumar DK: Fundus images analysis using deep features for detection of exudates,

hemorrhages and microaneurysms. BMC Ophthalmology. 2018, 18:1-13. 10.1186/s12886-018-0954-4
14. Bhatia K, Arora S, Tomar R: Diagnosis of diabetic retinopathy using machine learning classification algorithm. 2016 2nd International Conference on Next Generation Computing Technologies (NGCT), Dehradun, India. 2016, 347-351. 10.1109/NGCT.2016.7877439
15. Roychowdhury S, Koozekanani DD, Parhi KK: DREAM: diabetic retinopathy analysis using machine learning. IEEE Journal of Biomedical and Health Informatics. 2014, 18:1717-1728. 10.1109/jbhi.2013.2294635
16. Sebastian A, Elharrouss O, Al-Maadeed S, Almaadeed N: A survey on deep-learning-based diabetic retinopathy classification. Diagnostics. 2023, 13:345. 10.3390/diagnostics13030345
17. Kandel I, Castelli M: Transfer learning with convolutional neural networks for diabetic retinopathy image classification. A review. Applied Sciences. 2020, 10:2021. 10.3390/app10062021
18. Bellemo V, Burlina P, Yong L, et al.: Generative adversarial networks (GANs) for retinal fundus image synthesis. Revised Selected Papers. Springer International Publishing, Perth, Australia; 2019. 14:289-302. 10.1007/978-3-030-21074-8_24
19. Hinton GE, Osindero S, Teh YW: A fast learning algorithm for deep belief nets. Neural Computation. 2006, 18:1527-1554. 10.1162/neco.2006.18.7.1527
20. Krizhevsky A, Sutskever I, Hinton GE: Imagenet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems 25 (NIPS 2012). 2012, 25:1097-1105.
21. Ronneberger O, Fischer P, Brox T: U-net: Convolutional networks for biomedical image segmentation. Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015. Navab N, Hornegger J, Wells W, Frangi A (ed): Springer, Cham; 2015. 9351:234-241. 10.1007/978-3-319-24574-4_28
22. Fu H, Xu Y, Lin S, Kee Wong DW, Liu J: DeepVessel: retinal vessel segmentation via deep learning and conditional random field. Medical Image Computing and Computer-Assisted Intervention - MICCAI 2016. Ourselin S, Joskowicz L, Sabuncu M, Unal G, Wells W (ed): Springer, Cham; 2016. 9901:132-139. 10.1007/978-3-319-46723-8_16
23. Pratt H, Coenen F, Broadbent DM, Harding SP, Zheng Y: Convolutional neural networks for diabetic retinopathy. Procedia Computer Science. 2016, 90:200-205. 10.1016/j.procs.2016.07.014
24. Badar M, Haris M, Fatima A: Application of deep learning for retinal image analysis: A review. Computer Science Review. 2020, 35:100203. 10.1016/j.cosrev.2019.100203
25. Vasavi S, Likitha M, Neeraz L, Lin EB: Performance evaluation for classifying type 2 diabetic retinopathy using deep neural network. International Journal of Computer Aided Engineering and Technology. 2022, 17:184-191. 10.1504/ijcaet.2022.125047
26. Herrera-Chavez AI, Rodríguez-Martínez EA, Flores-Fuentes W, et al.: Multi-label image classification for ocular disease diagnosis using k-fold cross-validation on the ODIR-5K dataset. 2024 IEEE 33rd International Symposium on Industrial Electronics (ISIE), Ulsan, Republic of Korea. 2024, 1-6. 10.1109/ISIE54533.2024.10595740
27. Nagpal D, Pattanaik PA, Madaan V, Agrawal P: Enhancement of retinal images through modified anisotropic diffusion. Journal of Medical Artificial Intelligence. 2024, 7:10.21037/jmai-23-80
28. dos Santos JCM, Carrijo GA, dos Santos Cardoso CF, Ferreira JC, Sousa PM, Patrocínio AC: Fundus image quality enhancement for blood vessel detection via a neural network using CLAHE and Wiener filter. Research on Biomedical Engineering. 2020, 36:107-119. 10.1007/s42600-020-00046-y
29. Sahu S, Singh AK, Ghrera SP, Elhoseny M: An approach for de-noising and contrast enhancement of retinal fundus image using CLAHE. Optics & Laser Technology. 2019, 110:87-98. 10.1016/j.optlastec.2018.06.061
30. Mustafa WA, Yazid H: Conversion of the retinal image using gray world technique. Journal of Biomimetics, Biomaterials and Biomedical Engineering. 2018, 36:70-77. 10.4028/www.scientific.net/jbbbe.36.70
31. STARE (Structured Analysis of the Retina). (2025). Accessed: 26 April 2025: https://paperswithcode.com/dataset/stare.
32. Messidor. (2025). Accessed: 26 April 2025: https://www.adcis.net/en/third-party/messidor/.
33. EyePACS. (2025). Accessed: 26 April 2025: https://www.eyepacs.com/data-analysis.
34. APTOS 2019 Blindness Detection. (2025). Accessed: 26 April 2025: https://www.kaggle.com/competitions/aptos2019-blindness-detection.
35. Indian Diabetic Retinopathy Image Dataset. (2025). Accessed: 26 April 2025: https://ieee-dataport.org/open-access/indian-diabetic-retinopathy-image-dataset-idrid.
36. DIARETDB1. (2025). Accessed: 26 April 2025: https://paperswithcode.com/dataset/diaretdb1.
37. Ocular Disease Intelligent Recognition. (2025). Accessed: 26 April 2025: https://odir2019.grand-challenge.org/.
38. Winder RJ, Morrow PJ, McRitchie IN, Bailie JR, Hart PM: Algorithms for digital image processing in diabetic retinopathy. Computerized Medical Imaging and Graphics. 2009, 33:608-622. 10.1016/j.compmedimag.2009.06.003
39. Gupta V, Sadana R, Moudgil S: Image style transfer using convolutional neural networks based on transfer learning. International Journal of Computational Systems Engineering. 2019, 5:53-60. 10.1504/ijcsyse.2019.098418
40. Cai Q, Ma M, Wang C, Li H: Image neural style transfer: A review. Computers & Electrical Engineering. 2023, 108:108723. 10.1016/j.compeleceng.2023.108723
41. Wang G, Shi H, Chen Y, Wu B: Unsupervised image-to-image translation via long-short cycle-consistent adversarial networks. Applied Intelligence. 2023, 53:17243-17259. 10.1007/s10489-022-04389-0
42. Hermens F: Automatic object detection for behavioural research using YOLOv8. Behavior Research Methods. 2024, 56:7307-7330. 10.3758/s13428-024-02420-5
43. Mazurowski MA, Dong H, Gu H, Yang J, Konz N, Zhang Y: Segment anything model for medical image analysis: An experimental study. Medical Image Analysis. 2023, 89:102918. 10.1016/j.media.2023.102918
44. Padmavathi V, Ganesan K: Metaheuristic optimizers integrated with vision transformer model for severity detection and classification via multimodal COVID-19 images. Scientific Reports. 2025, 15:13941. 10.1038/s41598-025-98593-w
45. Yonamine S, Ma CJ, Alabi RO, et al.: Comparison of diagnosis codes to clinical notes in classifying patients with diabetic retinopathy. Ophthalmology Science. 2024, 4:100564. 10.1016/j.xops.2024.100564
46. Lepetit-Aimon G, Playout C, Boucher MC, Duval R, Brent MH, Cheriet F: MAPLES-DR: MESSIDOR anatomical and pathological labels for explainable screening of diabetic retinopathy. Scientific Data. 2024, 11:914.

10.1038/s41597-024-03739-6
47. Agustiono W, Permana KE, Chan C, Dewi DA, Husain MMRA: Rice variety identification based on transfer learning architecture using DENS-INCEP. IEEE Access. 2025, 13:78007-78020. 10.1109/access.2025.3562585
48. Shanmugam M, Sivakumar PD: Improving image inpainting performance with dual-stage adversarial structure and color-aware networks. Journal of Membrane Computing. 2025, 7:165-183. 10.1007/s41965-024-00173-w
49. MakeSense. (2025). Accessed: 26 April 2025: https://www.makesense.ai/.
50. YOLOv8 Architecture. (2025). Accessed: 26 April 2025: https://yolov8.org/yolov8-architecture/.