

Early Detection of Diabetic Retinopathy Using a Mobile-Based Deep Learning Framework

Ahmed Ben Ali^{1*}, Samira Gharbi¹, Nabil Jebali²

¹ Department of Ophthalmology and Telemedicine, Faculty of Medicine, University of Tunis, Tunis, Tunisia.

² Department of Diabetic Retinopathy Screening and AI, Faculty of Medicine, University of Sfax, Sfax, Tunisia.

*E-mail ✉ ahmed.benali@fmed.u-tunis.tn

Abstract

Diabetic retinopathy (DR) is a major contributor to blindness worldwide, particularly in areas with limited access to ophthalmic services. Timely identification is critical to preventing permanent vision impairment and improving clinical outcomes. This research introduces a lightweight, real-time Assisted Mobile Diagnostic (AMD) platform for DR screening that embeds a refined deep learning model directly onto a mobile device. Unlike traditional AI solutions that rely on powerful servers and fixed fundus imaging equipment, the proposed setup pairs a non-mydratic retinal camera with a smartphone, enabling on-site evaluations. Retinal photographs undergo initial processing steps, including smoothing and contrast adjustment, before classification by a fine-tuned DenseNet-121 network. Model training leveraged a combination of an internal hospital dataset and two prominent public repositories: APTOS (Asia Pacific Tele-Ophthalmology Society) and EyePACS (Eye Picture Archive Communication System). The method performed strongly, achieving 97.38% accuracy on APTOS, 90.90% on EyePACS, and 98.61% on the private dataset. On mobile hardware, the system processes images in an average of 162.5 ms, enabling efficient point-of-care screening. This deep learning-powered mobile tool forms a core component of a broader tele-ophthalmology ecosystem encompassing tablet and desktop versions, promoting widespread and remote access to DR diagnosis.

Keywords: Deep learning-based mobile application, Earlier screening of diabetic retinopathy, Real-time eye diagnosis, Handheld fundus camera

Introduction

Prompt identification of diabetic retinopathy (DR) is vital to prevent blindness, yet limited access to specialized imaging tools often hinders it, especially in resource-limited regions. As diabetes continues to spread rapidly — impacting 463 million individuals in 2019 and projected to affect 700 million by 2025 [1] — DR has emerged as a primary driver of vision loss. Particularly elevated rates are observed in Africa (35.9%), the Middle East and North Africa (32.9%), and North America (33.3%) [2]. Fundus photography is a key diagnostic tool

for DR, enabling detection of early disease indicators before permanent harm [3, 4].

Compact, handheld retinal cameras represent a flexible modern option compared to traditional fixed systems, supporting convenient and scalable screening in mobile environments [5]. These devices can acquire sharp images instantly and integrate seamlessly with smartphones or tablets for swift capture and evaluation. This capability is especially valuable in remote and rural communities where specialist care is scarce [6]. Coupling such devices with automated analysis tools would enhance their utility by quickly flagging urgent cases and streamlining referrals to experts. Overall, these innovations expand screening coverage and play a significant role in safeguarding vision among vulnerable populations [7].

The current study forms part of a broader initiative to create an intelligent automated framework for identifying and diagnosing eye conditions through fundus images

Access this article online

<https://smerpub.com/>

Received: 06 April 2025; Accepted: 22 June 2025

Copyright CC BY-NC-SA 4.0

How to cite this article: Ben Ali A, Gharbi S, Jebali N. Early Detection of Diabetic Retinopathy Using a Mobile-Based Deep Learning Framework. J Med Sci Interdiscip Res. 2025;5(1):204-25. <https://doi.org/10.51847/81LWc9jJmf>

obtained via portable retinal cameras. The detectable lesions in these images primarily stem from the initial phases of DR and manifest as microaneurysms, hemorrhages, soft exudates, and hard exudates. Microaneurysms represent the earliest indicator arising from weakened retinal capillaries. Hemorrhages occur following microaneurysm rupture, soft exudates result from lipid leakage due to vessel damage, and hard exudates form from accumulated lipids, reflecting more prolonged retinal alterations in later DR stages [8, 9]. Detecting these changes early allows for timely intervention to halt disease advancement and minimize irreversible vision loss. The envisioned platform is intended for deployment within Tele-Ophthalmology networks, with an initial focus on the Fez region in Morocco, where approximately 1.5 million people currently have diabetes—a figure anticipated to rise to 2.6 million by 2030 [9].

Progress in artificial intelligence has shown that deep learning-based automated systems can deliver outstanding precision in interpreting medical images, including those for DR screening [10], as well as in other demanding domains such as neurodevelopmental disorder detection and brain scan evaluation [11, 12]. To meet the demand for widely available screening, a Mobile-Aided Diagnosis (MAD) solution is required—one that is portable, affordable, and straightforward for eye care professionals, orthoptists, or nurses to use. This system should enable immediate acquisition of fundus images using non-mydratic portable cameras linked to smartphones or tablets, delivering on-the-spot screening outcomes to support early DR recognition, particularly at the mild non-proliferative stage. After the initial assessment during a visit, ophthalmologists can then conduct a detailed evaluation at a distant specialized center [13].

For the core intelligence component, DenseNet-121 was chosen as the primary deep learning architecture because of its strong track record in medical image categorization, particularly for retinal conditions. Its design promotes efficient feature reuse while maintaining a relatively low parameter count, balancing high performance with computational efficiency. Compared with conventional CNNs, DenseNet's dense layer connections improve gradient flow and feature transmission—advantages that are especially beneficial with the limited or imbalanced medical datasets typical of DR studies. Initial testing across several architectures, including MobileNet-V2, ResNet50, VGG16, and InceptionV3, confirmed that

DenseNet-121 delivered superior accuracy and robustness, justifying its adoption.

The planned Tele-Ophthalmology mobile application will serve as an effective means of delivering mobile eye care services in underserved rural zones around Fez, Morocco. As an initial version, it will connect to a pilot implementation at the Tertiary Ophthalmology Center and Vision Center at Omar Drissi Hospital, CHU HASSAN II, Fez. This effort seeks to address the critical shortage of ophthalmologists—with only 330 practicing in the public sector—while noting that over 80% of private-sector specialists are concentrated along the Casablanca-Kenitra corridor [14, 15].

Because images from portable retinal cameras are often affected by blur, noise, or other acquisition artifacts that can impair both human interpretation and the reliability of real-time deep learning systems for DR detection, a central focus of this work is to maximize input image quality for mobile use. Accordingly, the main contributions of this study are as follows:

- A detailed comparative evaluation of noise suppression and contrast improvement techniques tailored for retinal images captured by portable fundus cameras. Each approach is measured using standard quality indicators such as PSNR (Peak Signal-to-Noise Ratio), SSIM (Structural Similarity Index Measure), and RCEF (Relative Contrast Enhancement Factor) to identify the most effective options under practical capture conditions.
- Assessment of the best-performing preprocessing pipeline on the public APTOS and EyePACS datasets, followed by validation on a private dataset gathered at CHU Hassan II Hospital in Fez, confirming its suitability for real-world early DR detection.
- Creation of a complete mobile-aided diagnostic (MAD) platform that combines optimized deep learning models (DenseNet-121, Inception-v3, and MobileNetV2) within a tele-ophthalmology application for instant screening and referral. The solution is implemented as a mobile app to support timely diabetic retinopathy identification in remote and underserved locations.

The remainder of the paper is structured as follows: Section 2 offers a thorough overview of prior studies on early DR screening with non-mydratic handheld fundus cameras and relevant public retinal image datasets. Section 3 details our comparative analysis of image

enhancement techniques for fundus photographs, including the datasets employed. Section 4 presents the proposed methodology, encompassing deep transfer learning models for early DR identification, performance benchmarking against existing solutions, and the development of the mobile application, highlighting its value for fast and equitable DR screening. Section 5 provides concluding remarks.

Related work

Diabetic retinopathy is classified into varying degrees of severity, broadly divided into non-proliferative DR (NPDR) and proliferative DR (the more advanced form). NPDR itself is subdivided into mild, moderate, and severe stages [8-16]. Over time, researchers have proposed a wide range of computational methods for DR detection. In recent years, deep learning (DL) approaches, especially those based on transfer learning (TL), have shown strong potential for image analysis tasks [17].

For instance, one study [18] investigated how image enhancement and hyperparameter optimization influence DR classification performance with pre-trained CNNs. Using CLAHE (contrast-limited adaptive histogram equalization) combined with Bayesian optimization on AlexNet, GoogLeNet, and SqueezeNet, they reported classification accuracies reaching 99.81% in multi-class DR grading, underscoring the impact of preprocessing and tuning. Numerous computer-aided systems have also been developed for preliminary binary DR screening. In Ref. [19], a CNN built with standard trainable and non-linear layers was applied to DR detection. Wang *et al.* [20] evaluated deep transfer learning models, including AlexNet, VGG-16, and InceptionNetV3, on retinal images from the Kaggle dataset, achieving maximum accuracies of 37.43% (AlexNet), 50.03% (VGG-16), and 63.23% (InceptionNetV3). Pamadi *et al.* [21] utilized MobileNetV2 for binary and multi-class DR categorization, achieving 78% accuracy on the Kaggle APTOS dataset for binary detection. Zhou *et al.* [22] introduced a multi-cell network architecture to boost training efficiency and accuracy, reaching 63.22% for DR diagnosis. Deepa *et al.* [23] developed an ensemble of multiple deep learning models, attaining 96.2% accuracy via 5-fold cross-validation. Imran Qureshi *et al.*

[24] introduced an active deep learning (ADL) technique for distinguishing “DR” from “No DR,” reporting 92.20% sensitivity, 95.10% specificity, 93.1% F1-score, and 98.1% accuracy.

Transfer learning strategies are particularly advantageous as they deliver high performance with reduced training duration, lower memory demands, and simpler model design. Several reviews [25, 26] have summarized the application of DL algorithms and transfer learning-based DCNN models for fundus image-based DR classification and grading.

Portable retinal cameras have proven indispensable for advancing Tele-Ophthalmology and enhancing DR detection rates [27]. Adaptive fine-tuning with transfer learning has emerged as a valuable strategy in this domain [28]. Miden *et al.* [29] found that the Optomed Aurora portable camera surpassed conventional tabletop models in sensitivity and specificity for DR screening. Another comparative study [30] assessed three portable devices—iNview, Peek Retina, and Pictor Plus— and demonstrated solid performance in identifying DR and other ocular conditions during Tele-Ophthalmology sessions. In Ref. [31], researchers employed the Pictor Plus handheld camera alongside the Pegasus deep learning system (Visulytix Ltd., UK) for DR grading, obtaining AUROC values of 89.4% for referable DR (RDR) and 94.3% for proliferative DR (PDR) within the Mexican Advanced Imaging Laboratory for Ocular Research (MAILOR) cohort.

The present work seeks to establish a telemedicine infrastructure for timely DR detection. The overarching aim is to build a mobile system featuring a tele-ophthalmology application that leverages deep learning models and portable retinal cameras for screening and early diagnosis. The initial focus is a mobile app that supports rapid binary classification (“DR” or “No DR”) to enable swift clinical decisions and timely treatment. This project involves collaboration with the Ophthalmology Center and Vision Center at Omar Drissi Hospital, CHU HASSAN II, Fez [32, 33]. Beyond DR, the framework can be readily extended to other ocular conditions, as demonstrated in prior research on glaucoma [34] and age-related macular degeneration (AMD) [35].

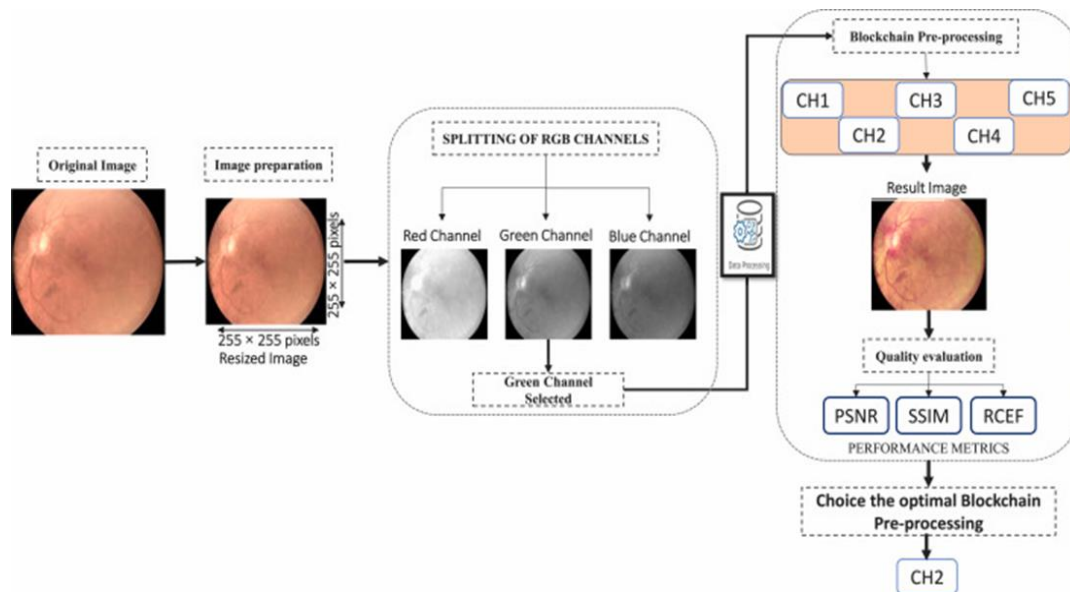


Figure 1. Diagram illustrating the workflow for improving retinal image quality via the preprocessing module.

Comparative analysis of fundus image enhancement techniques

Fundus photographs taken with portable handheld cameras often exhibit quality issues such as blur, insufficient contrast, noise, and inconsistent lighting, especially under field conditions or in low-resource settings. These degradations make it difficult to clearly observe critical diagnostic elements such as microaneurysms, hemorrhages, and exudates, which are crucial for identifying DR in its initial phases [36, 37]. Recent investigations have underscored the value of preprocessing steps to improve image clarity before automated analysis. In particular, contrast-improvement methods have been shown to enhance lesion visibility and, in turn, improve the effectiveness of deep learning models used for DR screening [38, 39]. Moreover, adaptive pipelines that integrate denoising with CLAHE have yielded notable improvements in both sensitivity and specificity. Consequently, developing a reliable yet computationally efficient image enhancement workflow is crucial to maintain the diagnostic value of field-acquired fundus images.

In the present work, we conduct an extensive comparison of multiple preprocessing pipelines aimed at refining retinal images obtained during actual mobile screening sessions. The study evaluates combinations of noise reduction, contrast adjustment, and adaptive gamma correction techniques to determine the optimal setup for improving diagnostic image quality. The selected pipeline prioritizes the green channel to retain fine

vascular structures, and its effectiveness is measured using established quantitative indicators, including PSNR, SSIM, and RCEF. **Figure 1** provides a visual overview of the complete image enhancement process.

The suggested preprocessing sequence for retinal fundus images consists of multiple stages. Initially, the captured retinal image is resized to a suitable dimension to streamline computational demands. The image is then separated into its red, green, and blue color channels. Special attention is given to the green channel because it highlights detailed structural information within the retina [28]. Further refinement is performed by applying a series of smoothing and sharpening filters. Ultimately, the quality of the processed images is quantified using three key metrics: PSNR, SSIM, and RCEF. These measures provide objective evaluations of image fidelity and help gauge the contribution of each processing stage.

Private dataset description

Dataset collection process

In recent DR research, several initiatives have concentrated on building proprietary databases with portable retinal cameras to gain deeper insights into the disease. Acquiring data locally using such devices allows detection algorithms to be better tailored to specific regional traits and patient demographics [28, 30]. Hence, developing private datasets is highly valuable, as they provide contextually relevant information suited to particular environments. Our partnership with the Department of Ophthalmology at Omar Drissi Hospital

(CHU Hassan II) in Fez, Morocco, adopts this strategy by using a portable retinal camera to collect targeted data, supporting more precise and individualized DR detection.

The private dataset consists of retinal fundus images acquired with the oDocs Nun IR camera—a compact, non-mydratic device optimized for DR screening—as shown in **Figure 2**. To ensure clinical representativeness, images were collected from individuals with different stages of DR, as verified by comprehensive eye examinations. We applied strict inclusion criteria: participants were adults aged 18 years or older with a documented diagnosis of diabetes. We excluded individuals with additional retinal disorders or unrelated ocular conditions to keep the dataset focused on DR.

Unlike widely available datasets such as APTOS and EyePACS, which are typically collected in controlled clinical settings with professional-grade equipment, our private collection captures greater diversity and authenticity. We recorded images at Omar Drissi Hospital (CHU Hassan II) in Fez, Morocco, using a handheld camera during outreach screening programs for residents of rural and semi-urban communities. Consequently, the dataset reflects a broad spectrum of illumination variations, image qualities, and acquisition angles that mirror practical difficulties often absent from standard benchmarks. Furthermore, the study population encompasses a wide demographic mix across age ranges, socio-economic levels, and ethnic backgrounds, with no selective bias, offering a sample that accurately represents the Moroccan population. This diversity strengthens the generalizability and resilience of the proposed detection system in under-resourced and underrepresented scenarios.

The selection of the oDocs Nun IR camera was driven by its notable strengths. It delivers a high resolution of 2880×2160 pixels and a broad field of view (FOV) spanning 45° to 55° , facilitating detailed observation of retinal structures. The device supports manual focus adjustment from $-20D$ to $+20D$ for achieving sharp focus. Its lightweight, portable design and seamless integration with Android smartphones make it particularly well-suited for deployment in remote locations where access to advanced ophthalmic infrastructure is limited.



Figure 2. oDocs Nun IR retinal camera [40].

The oDocs Nun IR device effectively records fine pathological features, including microaneurysms and hemorrhages, which are indispensable for precise identification and grading of diabetic retinopathy (DR). This equipment balances strong imaging capabilities with practical accessibility, making it especially useful in geographically isolated or underserved regions. In addition, the developed preprocessing pipeline demonstrates compatibility with a variety of other portable retinal cameras, such as Vistaview, Volk iNview, and Peek Retina. This versatility accommodates diverse clinical scenarios and supports reliable DR detection under varying operational conditions. The overall system was designed using modular elements for image capture, preprocessing, and model inference. Such modularity enables straightforward adaptation to additional hardware with only minor modifications. This design also promotes expandability, allowing easy incorporation of different mobile phones or fixed fundus cameras that may produce images with differing resolutions and characteristics.

Data quality and preprocessing

The private dataset used in this study contains 1319 retinal fundus images with a range of quality levels. Representative examples from this collection are shown in **Figure 3**. Superior-quality images (**Figure 3a**) exhibit sharp structural details, low noise levels, and good contrast, making them well-suited for detailed examination. In contrast, the dataset also includes several lower-quality samples affected by blurring (**Figure 3b**), noise (**Figure 3c**), and poor contrast (**Figure 3d**). To mitigate these limitations and bring all images to an acceptable level for clinical review and deep learning

processing, we implemented supplementary enhancement methods—notably deblurring and contrast adjustment techniques.

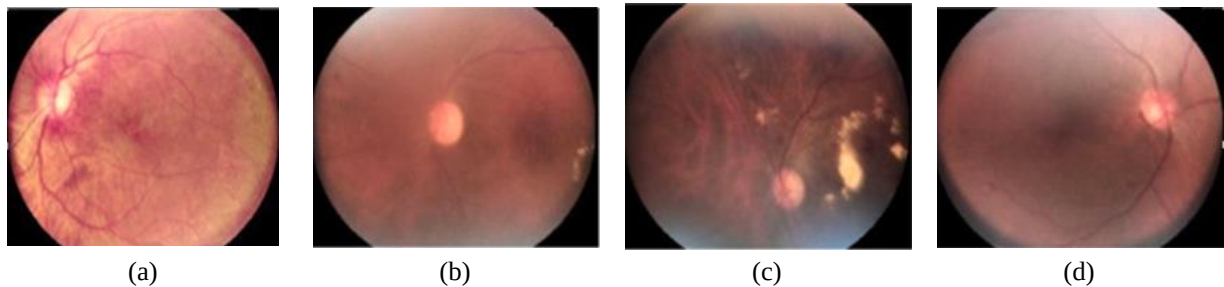


Figure 3. Samples illustrating image clarity (a), blurriness (b), contrast issues (c), and brightness inconsistencies (d) from the private dataset.

Dataset annotation and labeling

The data acquisition process fully complied with ethical requirements to protect patient privacy and meet global data protection regulations. An independent ethics review board approved the study, and every participant provided written informed consent before inclusion. To protect sensitive information, we applied strong safeguards, such as anonymizing images and associated metadata, storing data on encrypted servers, and limiting access to authorized team members. These precautions ensured the dataset met rigorous privacy and security standards in line with established medical research protocols. Nonetheless, qualified healthcare professionals must operate the devices under direct supervision and responsibility, including in remote locations, to uphold ethical standards and safeguard patient information. A dedicated clinical protocol will be established during consultations to confirm its role as a supportive tool for early diagnosis.

Three seasoned ophthalmologists annotated the dataset, independently evaluating and assigning each image to one of five DR severity categories: no DR, mild non-proliferative diabetic retinopathy (NPDR), moderate NPDR, severe NPDR, and proliferative diabetic retinopathy (PDR). To maintain consistency across raters, we resolved discrepancies through discussion and consulted a fourth specialist when needed. The resulting labeled collection comprises: 256 images without DR, 381 images with mild NPDR, 335 images with moderate NPDR, 253 images with severe NPDR, and 94 images with PDR. Labeling followed established international guidelines for DR classification to guarantee uniformity, precision, and reproducibility. Additionally, the “No

DR” category was kept separate, while the four stages featuring retinal abnormalities were grouped under a single “DR” label. This binary grouping simplifies early detection while still accommodating more advanced cases.

Public datasets description

Alongside the private dataset, this study incorporated two well-established public repositories: APTOS [41] and EyePACS [42]. Growing interest in training and validating deep learning models has encouraged the creation and public release of specialized retinal image collections [43]. Both APTOS and EyePACS are highly regarded within the research community for their large volumes of high-quality fundus photographs.

APTOS

The APTOS [41] Blindness Detection dataset serves as an important asset for investigators and clinicians working on DR. It includes a broad set of fundus images representing the five recognized stages of the disease. Each image carries a severity score ranging from 0 to 4, where label 0 indicates no DR, label 1 corresponds to mild DR, label 2 to moderate DR, label 3 to severe DR, and label 4 to proliferative DR. The complete dataset contains 3662 retinal images, broken down as follows: 1805 images with no DR, 370 with mild DR, 999 with moderate DR, 193 with severe DR, and 295 with proliferative DR. All images maintain a high resolution of 3216×2136 pixels, providing excellent visual quality suitable for thorough analysis and model assessment.

EyePACS

Similar to APTOS, the EyePACS [42] dataset offers a vast array of fundus images covering all five stages of DR progression. It includes 35,126 retinal photographs (capturing both left and right eyes) at 3888×2951 pixels, providing rich detail for research applications. The images are distributed across the DR stages as follows: 25,810 labeled as no DR, 2,443 as mild DR, 5,292 as moderate DR, 873 as severe DR, and 708 as proliferative DR. For model training, a random subset was selected consisting of 1,534 images from the no DR class and 2,764 images from the DR class.

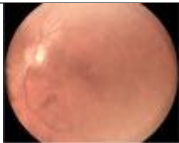
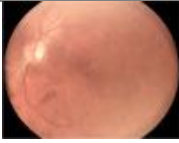
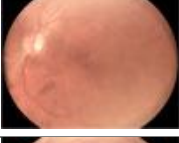
Proposed flowchart and analysis of enhancement techniques

Image quality improvement methods seek to minimize unwanted artifacts, highlight clinically relevant structures, and elevate overall visual clarity. A common strategy is to increase the distinction between darker and brighter areas. This is typically accomplished using various filters, which are standard tools for restoring low-quality images. Filters can soften appearance, sharpen details, or accentuate boundaries. Earlier studies have assessed the performance of 11 basic filters for refining retinal images [39]. These filters proved effective at clarifying blood vessel structures. Before applying filters,

we prepare images by resizing them and extracting the green channel, which carries the richest structural information. In the current study, we selected five particularly effective filters based on strong results: the median filter (MF), non-local means filter (NLMF), Gaussian filter (GF), CLAHE, and bilateral filter (BF). **Table 1** presents comprehensive descriptions of these filters, including their mathematical formulations, alongside example outputs.

The median filter suppresses noise while preserving edges by replacing each pixel with the median value from its surrounding neighborhood. This property makes it well-suited for retinal images, as it efficiently eliminates outlier noise values. CLAHE (Contrast Limited Adaptive Histogram Equalization), on the other hand, locally adjusts contrast based on regional pixel distributions, helping emphasize important features such as vessels. However, it requires careful tuning to avoid over-amplifying noise. Both techniques significantly enhance the interpretability of retinal photographs. In addition, we tested an enhanced median filter (EMF) and a combination of median and average filters (MMF) [44]. However, the MMF combination performed worse than the five primary filters across PSNR, SSIM, and RCEF metrics [39] and is therefore not detailed further here.

Table 1. Filters with their mathematical expressions and output images.

Filters	Mathematical expressions	Output images
MF	$g(x,y)=\text{median}\{f(n,m)\}$ Where $f(n,m)$ returns the pixel value at position (n,m) in the image.	
GF	$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-(x^2+y^2)/2\sigma^2}$ Where $G(x,y)$ is the value of the Gaussian filter at a given position (x,y). σ: the standard deviation of the Gaussian filter.	
BF	$I = \frac{1}{wP} \sum_{xi} I(xi) fi(\ I(xi) - I(x)\)$ Where wP is the normalization factor. $I(x)$ is the intensity of the pixel being processed. xi are the neighboring pixels around the pixel being processed.	
NLMF	$\text{NLMF}(x) = \left(\frac{1}{Z(x)}\right) * \sum (w(x,y) * f(y))$ Where $Z(x)$ is a normalization factor. $\sum (w(x,y) * f(y))$ represents the weighted sum of similar pixel values $f(y)$ based on their weights $w(x,y)$.	
CLAHE	$g = [g_{\max} - g_{\min}] * p(f) + g_{\min}$ Where $g_{\max}, g_{\min}$ are the maximum and minimum grey levels. $p(f)$ is the cumulative probability of grey level f in a local region of the input image.	

Enhancement chains

To develop the preprocessing sequences, we selected the top-performing filter configurations. These sequences are composed of successive operations encompassing noise suppression, contrast improvement, and adaptive gamma correction. The filters that performed well, as summarized in **Table 1**, formed the basis of these chains. MF, GF, BF, and NLMF primarily function as smoothing filters, whereas CLAHE serves as a sharpening filter. Consequently, the chains integrate a strategic mix of both smoothing and sharpening operations. Adaptive gamma correction (AGC) modifies pixel intensity levels in an image according to a gamma correction curve (GCC) [45]. For an image with dimensions $m \times n$ pixels, this correction applies the following transformation to the intensity value of each pixel at location (x, y) :

$$I_{gam}(x, y) = I(x, y)^\gamma \quad (1)$$

Where $I_{gam}(x, y)$ denotes the gamma-adjusted intensity at position (x, y) in the original image $I(x, y)$, and γ represents the variable gamma parameter that determines the mapping behavior. When $\gamma > 1$, darker areas in the source image appear brighter after correction. Conversely, when $\gamma < 1$, brighter regions become darker in the corrected image [38]. Importantly, the gamma value is determined individually for each image according to its specific intensity distribution.

Performance metrics

The effectiveness of image enhancement algorithms and filters is assessed using several quantitative indicators. Metrics including PSNR, RCEF, and SSIM help identify which techniques deliver the best outcomes. PSNR is widely regarded as the standard measure for enhancement quality and is defined by:

$$PSNR = 10 \log_{10} \frac{255^2}{\frac{1}{MN} \sum_{i,j} [g_{i,j} - f_{i,j}]^2} dB \quad (2)$$

Where f and g represent the original and enhanced images, respectively, while M and N denote the number of rows and columns. The indices i and j refer to specific row and column positions. Higher PSNR values indicate better performance.

Structural similarity between the original and processed images is quantified by SSIM, calculated as:

$$SSIM(A, B) = \frac{(2 \mu_A \mu_B + c_1)(2 \sigma_{AB} + c_2)}{(\mu_A^2 + \mu_B^2 + c_1)(\sigma_A^2 + \sigma_B^2 + c_2)} \quad (3)$$

Where μA and μB are the mean intensity values of images x and y computed over a local window, σA^2 , σB^2 , and σAB represent the variances and covariance, and c_1 and c_2 are small stabilization constants. SSIM ranges from -1 to 1 , with 1 indicating perfect structural correspondence and lower values signifying reduced similarity.

The expansion of the histogram's dynamic range is evaluated using RCEF, given by:

$$RCEF = \frac{\frac{\sigma_B^2}{\mu_B}}{\frac{\sigma_A^2}{\mu_A}} \quad (4)$$

Where σA and σB are the standard deviations of the original and enhanced images, and μA and μB are their respective mean intensity values. An RCEF greater than 1 indicates that the enhanced image has improved contrast and a broader distribution of pixel intensities than the original.

While enhancement methods generally improve visual quality, overly aggressive sharpening or filtering can introduce noise or compromise subtle pathological details. Therefore, objective metrics (PSNR, SSIM, and RCEF) were used to select the preprocessing sequence that best preserves critical diagnostic information, especially early indicators of diabetic retinopathy.

Image quality evaluation

The five proposed enhancement chains were individually assessed, with outcomes summarized in **Table 2**. Key observations from the different combinations are noted. Among them, the chain combining GF + CLAHE + AGC demonstrated the strongest overall results. Chain 2 achieved an SSIM of 0.9808 and a PSNR of 33.2950 , reflecting excellent structural fidelity to the original image. Owing to its superior metrics, Chain 2 was adopted for all subsequent detection and classification steps described in Section 4. This preprocessing workflow was incorporated into the classification pipeline, improving model performance across the evaluated datasets. **Figure 4** presents the results of the optimal preprocessing chain, showing input images (a) and (b) alongside their corresponding enhanced outputs (c) and (d). The sequence applies a Gaussian filter to reduce noise and smooth the image, followed by channel

merging to produce the final improved image. This process is designed to elevate retinal image quality, supporting more accurate analysis and interpretation. Early DR screening distinguishes between two main phases: non-proliferative DR (NPDR) and proliferative DR (PDR). NPDR constitutes the earlier phase and is marked by lesions such as microaneurysms, hemorrhages, and exudates. PDR represents the more advanced phase and features the growth of abnormal new blood vessels on the retina [46]. Captured images allow

clear observation of typical DR lesions and help determine disease severity [47]. In this research, the three NPDR stages (mild, moderate, and severe) along with PDR were consolidated into one “DR” category, while the “No DR” category remained distinct. **Figure 5** displays original images from the private dataset (a), APTOS (b), and EyePACS (c), together with the corresponding outputs after applying the selected optimal preprocessing pipeline (d), (e), and (f).

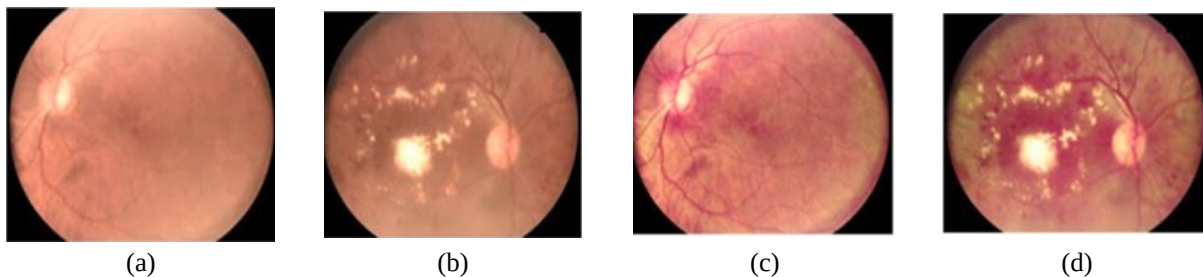


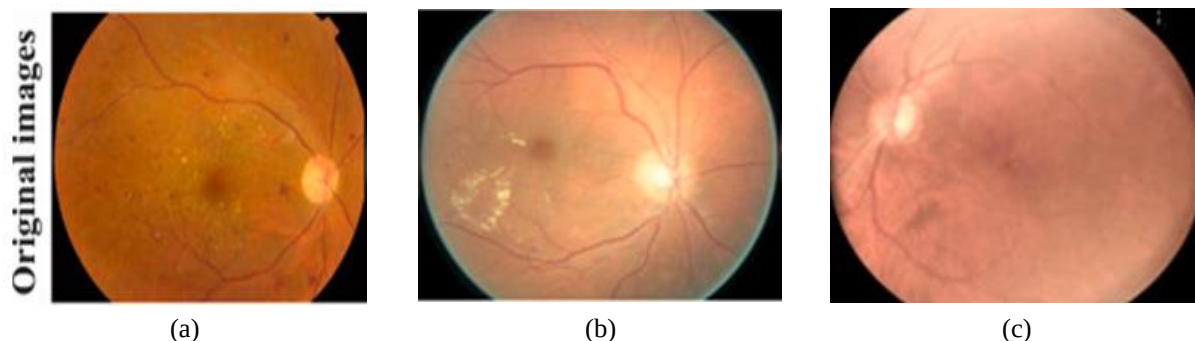
Figure 4. The best enhancement chain involves Gaussian Filter, CLAHE, and Gamma correction to improve retinal image quality: (a) and (b) are the original images, (c) and (d) are the output results corresponding to (a) and (b), respectively.

Table 2. Performance evaluation of image quality using PSNR, SSIM, and RCEF metrics.

Empty Cell	Blockchain	SSIM	PSNR	RCEF
Chain 1	MF+CLAHE+AGC	0,9804	33,2531	1,0177
Chain 2	GF+CLAHE+AGC	0,9808	33,2950	1,0476
Chain 3	BF+CLAHE+AGC	0,9803	33,2317	1,0378
Chain 4	NLM+CLAHE+AGC	0,9804	33,2096	1,0381
Chain 5	MMF+CLAHE+AGC	0,9802	33,2641	1,0276

Given that enhancing image quality was a core objective, we further examined which model architecture would integrate most effectively with the image enhancement pipeline. The primary aim was to investigate how different enhancement methods affect classification accuracy across models and to select a strong, reliable

pairing optimized for real-world use. This strategy tackles the two central issues in DR diagnosis: suboptimal image quality and the need for consistent classification performance in variable real-world conditions.



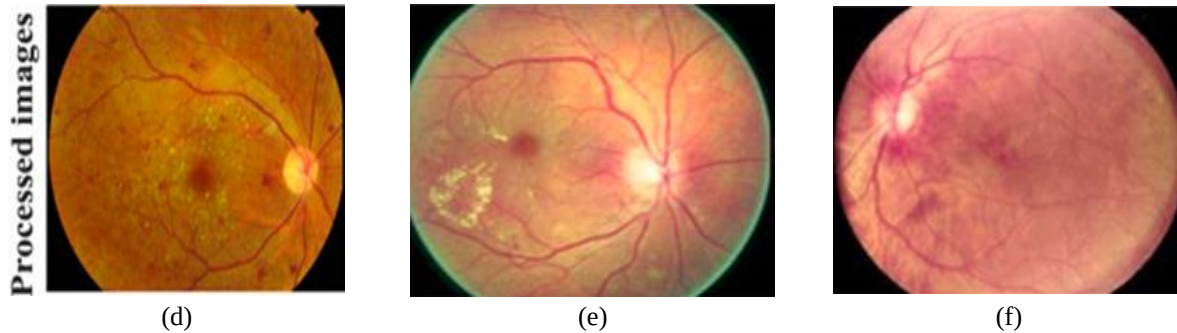


Figure 5. Preprocessed retinal fundus images: (a) and (d) APTOS dataset, (b) and (e) EyePACS dataset, (c) and (f) private dataset.

Proposed Method

Image classification involves assigning a label or category to a previously unseen image. This task typically relies on pre-trained models that predict new samples using knowledge gained from labeled training data. Deep neural networks (DNNs) are among the most effective architectures for this purpose, as they automatically learn multi-level features—including edges, textures, shapes, and their spatial arrangements. These networks excel at identifying complex patterns in visual data, which makes them highly suitable for accurate image recognition and classification tasks.

The approach introduced in this work consists of three main stages. The first stage focuses on data preparation to ready the model for the classification task. In the second stage, the optimal preprocessing pipeline identified in the study referenced in Section 3 is applied, combined with data augmentation techniques to enhance dataset variety and boost model generalization. Additionally, 5-fold cross-validation is employed to ensure reliable performance assessment. Three deep learning models—DenseNet-121, Inception-v3, and MobileNet-v2—are utilized through transfer learning and fine-tuning. These architectures are adapted to capture a broad spectrum of lesion characteristics from retinal fundus images, thereby improving the detection of

subtle indicators associated with diabetic retinopathy (DR).

The early detection of diabetic retinopathy has been integrated into a mobile application capable of identifying the presence of DR. A detailed flowchart illustrating the overall proposed methodology is presented in **Figure 6**.

The selection of these pre-trained models was guided by their strengths in recognizing the two DR stages (“NO DR” or “DR”), with each architecture contributing complementary feature extraction capabilities tailored to the structural properties of DR lesions. DenseNet-121 benefits from its dense connectivity, which promotes effective information flow across layers and supports the extraction of intricate, fine-grained features such as microaneurysms, exudates, and hemorrhages in retinal images. Similarly, Inception-v3 leverages its Inception modules to perform multi-scale feature extraction, enabling it to detect complex lesion patterns including blood vessel abnormalities, regions of ischemia, and other retinal alterations typical of the disease. In contrast, MobileNet-v2 offers an efficient, lightweight design optimized for resource-constrained environments like mobile devices. This allows it to identify key retinal lesion features while maintaining strong performance and supporting real-time DR stage detection on mobile platforms.

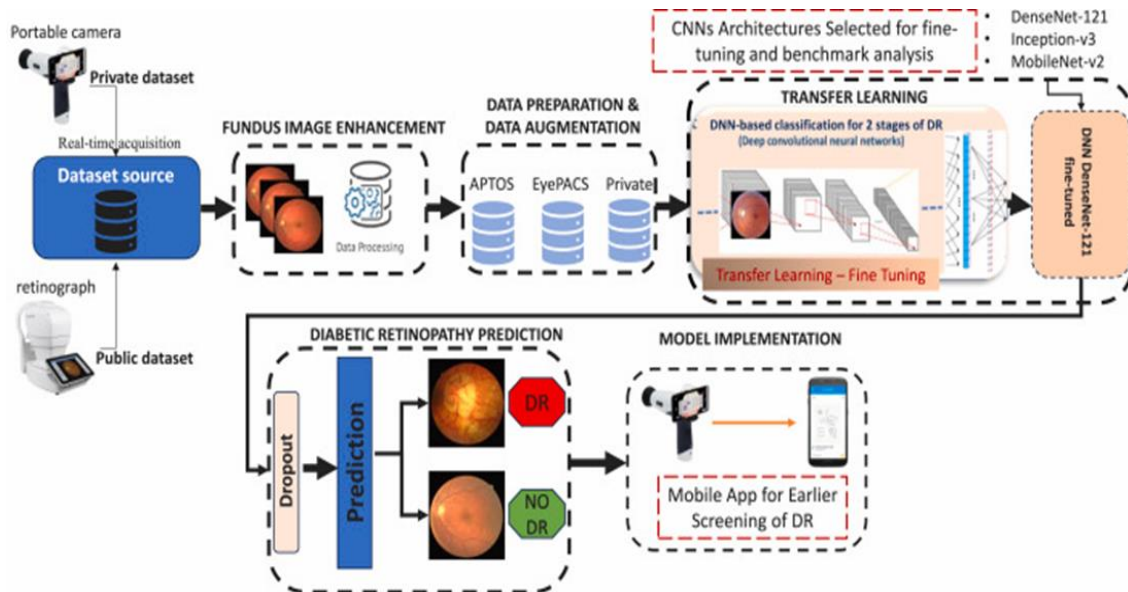


Figure 6. Flowchart of a deep learning-based mobile app for early screening of diabetic retinopathy using real-time handheld fundus imaging and validation with public datasets.

Data preparation

In the diabetic retinopathy (DR) detection study, the three utilized datasets were each split into binary classes labeled “NO DR” and “DR”. Images free of any retinopathy indicators were assigned to the “NO DR” class. In contrast, the “DR” class gathered all samples across the four graded severity levels (1 through 4). This organization clearly separates unaffected cases from those with varying disease intensity, supporting thorough examination and precise model-based classification.

Furthermore, to overcome limitations caused by inconsistent quality and irregularities commonly found in publicly available data, problematic images with visible artifacts were deliberately removed from both the APTOS and EyePACS datasets. We performed these removals primarily because saturation issues could compromise the reliability and effectiveness of the trained system. Overall, this cleaning phase aimed to preserve only clear, high-fidelity retinal images for the modeling stages ahead.

Table 3. Dataset distribution before and after data augmentation.

Datasets	Before data augmentation		After data augmentation
Empty Cell	Initial dataset	Selected images	Used images
Private	1321	1321	7200
APTOS	3662	2880	9566
EyePACS	35 126	4290	20 000

Data augmentation and dataset splitting

During preprocessing, the images were resized from 255×255 pixels to 224×224 pixels to meet the DNN model's input requirements. We then applied a range of data augmentation methods to expand the dataset, including vertical and horizontal flips, rotations at multiple angles, and brightness and color adjustments. We carefully balanced these augmentation strategies at every stage to address the imbalance between the “NO DR” and “DR” classes. By applying these techniques while maintaining equilibrium between the “DR” and

“NO DR” categories, the goal was to increase diversity and variability in the training data. This process aimed to strengthen the model's generalization and help it capture varied real-world scenarios, thereby boosting robustness and accuracy across cases. **Table 3** presents the image counts for each dataset split, including the numbers before and after augmentation.

Various research efforts have used different ratios for partitioning datasets into training, validation, and test sets. Properly dividing the data into these subsets is crucial for effective model training and reliable

evaluation of performance and generalization. The chosen allocation prioritizes a larger share for training to allow the model to learn from a wide variety of examples, while still reserving adequate portions for validation and testing. A widely adopted split is 80% for training, 10% for validation, and 10% for testing [48, 49]. Other works have used 75% for training, with the remainder split between validation and testing [50]. In this study, we applied a 75% training, 15% validation, and 10% testing distribution consistently across all three datasets, based on their respective image totals. For the private dataset of 7200 images, this resulted in 5400 training, 1080 validation, and 720 testing images. The APTOS dataset (9566 images) was divided into 7174 training, 1433 validation, and 959 testing images. Similarly, the EyePACS dataset (20,000 images) was split into 15,000 training, 3000 validation, and 2000 testing images.

To reduce class imbalance and improve the model's responsiveness to all DR severity levels, we performed targeted data augmentation specifically on the underrepresented classes in the private dataset. This produced a balanced training set across the five DR grades, minimizing bias toward the dominant NoDR class that prevails in the original data. Importantly, augmentation was applied only to the training portion after the initial split into training, validation, and test sets. This ensured no augmented samples entered the validation or test sets, avoiding data leakage and maintaining an unbiased performance assessment.

designing the DNN model for early DR screening

A Deep Neural Network (DNN) for binary DR detection was developed using transfer learning with three distinct architectures integrated as separate blocks in the proposed framework (**Figure 6**): DenseNet-121 [51], Inception-v3 [52], and MobileNetV2 [53]. The main aim was to reliably classify images as DR-present or DR-absent. Transfer learning from these architectures allowed the model to benefit from knowledge acquired from large image datasets, thereby improving its capacity to identify early DR signs for initial screening.

The models were trained for 200 epochs with the RMSProp optimizer and Binary Cross-Entropy (BCE) loss function. These choices were selected because they are well suited to training the architectures over the full 200 epochs. A learning rate of 0.001 was used, along with a batch size of 64 during training and 32 for validation. The final layer employed a 'sigmoid' activation function, which is standard for binary classification as it outputs

probabilities between 0 and 1. To combat overfitting, early stopping was implemented (halting training after 50 epochs without validation accuracy improvement), and a dropout layer was added. Validation accuracy was monitored each epoch, and the model achieving the best validation accuracy was saved via the Keras ModelCheckpoint callback.

Performance results and discussion

Implementation framework

All experiments were executed on a system with an NVIDIA Corporation GP102GL [Quadro P6000] GPU, an Intel (R) Xeon(R) Gold 6240R CPU (20 cores at 2.40 GHz), and 377 GB of RAM. The models were implemented in Python 3.10.9 using deep learning libraries including Keras 3.10.9 with TensorFlow backend, OpenCV, and Scikit-Learn.

Evaluation metrics

Model performance was assessed using five standard metrics: accuracy, sensitivity, specificity, precision, and F1-score, as defined in Eqs. (5)-(9), respectively. These measures offer a comprehensive view of the classification model's effectiveness.

$$\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN} \quad (5)$$

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (6)$$

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

$$\text{Precision} = \frac{TP}{TP + FP} \quad (8)$$

$$\text{F1 - Score} = 2 \times \frac{\text{Precision} \times \text{sensitivity}}{\text{Precision} + \text{sensitivity}} \quad (9)$$

Results and Discussion

Fine-tuning involves adjusting the weights of a pre-trained model on a more focused, smaller dataset to detect the initial phases of diabetic retinopathy (DR). This strategy builds on a model originally trained for broader purposes and tailors it to identify the distinctive characteristics linked to DR. By doing so, fine-tuning enhances the model's generalization capabilities while minimizing reliance on extensive labeled data, which is especially valuable in medical imaging, where high-quality annotated datasets are often scarce.

The proposed system's classification performance was assessed using the datasets outlined in Section 3.1.3. As

shown in **Table 4**, the evaluation metrics highlight DenseNet-121's clear advantage over other models across all datasets. Specifically, DenseNet-121 delivered an accuracy of 98.61%, sensitivity of 98.33%, specificity of 100.00%, precision of 98.88%, and an F1-score of 70.03%. Particularly strong outcomes were recorded on the APTOS and EyePACS datasets, with DenseNet-121 attaining accuracy levels of 97.38% and 90.90%, respectively. These outcomes underscore DenseNet-121's strong effectiveness in DR detection and its robust ability to generalize across varied datasets despite differences in image quality and patient characteristics.

Table 5 provides additional insights into the training and validation accuracy of various deep learning models across the three datasets. DenseNet-121 consistently surpassed alternatives such as Inception-v3 and MobileNetV2, illustrating its stability during both training and validation. The model's outstanding results stem from its dense connectivity pattern, which promotes efficient feature reuse and propagation, enabling it to capture intricate patterns and deliver highly accurate classifications. This steady performance gain throughout the process further validates DenseNet-121 as a dependable choice for demanding classification challenges.

Table 4. Comparative analysis of performance metrics for three architectures on Private, APTOS, and EyePACS Datasets.

Dataset	Test prediction	Architectures		
Empty Cell	Eval. metrics	DenseNet-12	MobileNetv2	Inception-v3
Private	Acc	98.61	96.11	96.52
	Sen	98.33	94.72	94.72
	Spe	100.00	99.72	99.72
	Pre	98.88	98.27	98.27
	F1-score	70.03	66.97	66.97
APTOS	Acc	98.03	97.49	97.38
	Sens	97.76	96.42	97.09
	Spe	99.78	100.00	100.00
	Pre	98.20	98.40	97.53
	F1-score	67.98	70.21	65.78
EyesPACS	Acc	90.90	87.85	89.14
	Sens	93.90	85.30	88.10
	Spe	99.40	98.90	98.90
	Pre	88.58	89.88	89.89
	F1-score	69.37	70.00	68.37

Our results show that DenseNet-121's strong capabilities can be applied not only in sophisticated medical diagnostic settings but also in resource-constrained mobile scenarios that require efficient computation and real-time processing. This is especially relevant for teleophthalmology applications, where integrating such models into portable devices can support rapid and widely accessible screening for diabetic retinopathy (DR). Although architectures like Inception-v3 and MobileNetV2 are commonly adopted, DenseNet-121 offers a superior trade-off between high accuracy and computational requirements. Its dense connectivity promotes effective feature reuse, lowering the parameter count, reducing overfitting risk, and ultimately improving accuracy and generalization.

Preprocessing techniques also contributed substantially to the overall classification success. For DenseNet-121, with its deep and densely interconnected layers, these steps enhance the signal-to-noise ratio and highlight key retinal structures. This combination supports better extraction of relevant features at multiple levels, allowing the model to detect subtle disease indicators more reliably than networks with shallower or less interconnected designs. Furthermore, when compared to lighter models such as MobileNetV2, DenseNet-121 delivers markedly better classification results while still providing practical inference speeds suitable for contemporary mobile devices. Its deeper structure also enables the identification of intricate retinal patterns that simpler convolutional approaches might overlook.

In relation to previous studies, Wang *et al.* [20] applied transfer learning using AlexNet and VGG-16, achieving reasonable accuracy but facing constraints due to their shallower, less optimized architectures. Deepa *et al.* [23] investigated ensemble approaches with InceptionV3 and Xception, which yielded better performance but required considerably more computational resources. By contrast, our adoption of DenseNet-121 strikes a more effective balance among accuracy, efficiency, and generalization ability, especially when tested on diverse and heterogeneous datasets. **Figure 7** presents the training and validation accuracy and loss curves for the Private, APTOS, and EyePACS datasets, where sustained high

accuracy and consistent loss reduction demonstrate the model's robustness and effective transferability across varied data sources.

The confusion matrix is an important tool for understanding model behavior, offering detailed insights into classification accuracy for both DR and No DR classes. We evaluated these results on the three datasets using confusion matrices. As depicted in **Table 6**, the matrices for each dataset reveal the outstanding and consistent performance of the DenseNet-121 model. Based on these observations, DenseNet-121 was chosen as the best-performing architecture thanks to its exceptional predictive power.

Table 5. Training and validation results comparison of the three architectures.

Dataset	DenseNet-121		MobileNet2		Inception-v3	
Empty Cell	Training	Validation	Training	Validation	Training	Validation
Private	99.81	99.37	99.07	98.12	98.88	98.42
APTOS	98.91	98.81	98.68	98.51	98.61	98.60
EyePACS	94.06	93.92	94.54	92.16	91.30	91.43

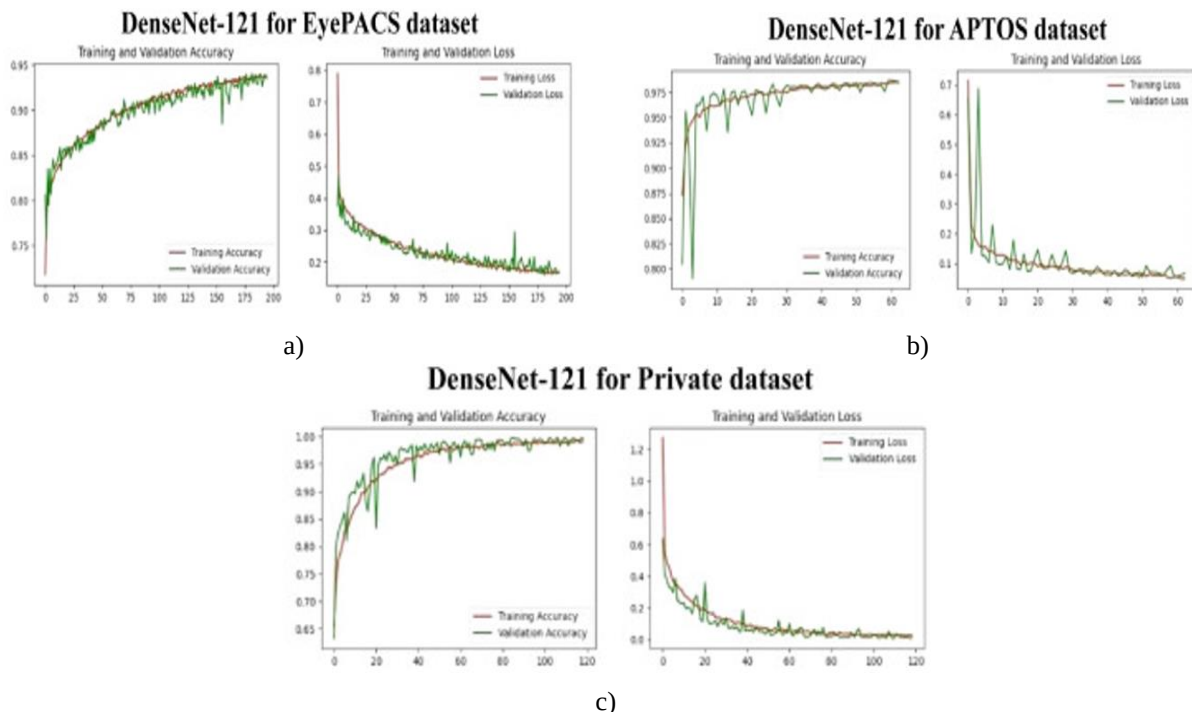


Figure 7. Training and validation accuracy and loss curves for the DenseNet-121 model on (a) the private dataset, (b) the APTOS dataset, and (c) the EyePACS dataset.

Table 7 presents additional performance metrics for the transfer learning approaches, providing a detailed comparison of DenseNet-121, MobileNet-v2, and Inception-v3. Across all three datasets (Private, APTOS,

and EyePACS), DenseNet-121 consistently delivered outstanding results, achieving notably high AUROC scores. On the Private dataset, for instance, DenseNet-121 reached an AUROC of 0.9991, outperforming

MobileNet-v2 (0.9974) and Inception-v3 (0.9945), which showed competitive yet lower values. These outcomes underline the models’ overall effectiveness in reliable dataset classification. Moreover, the recorded

inference times—such as 161.62 ms for DenseNet-121 on the Private dataset—provide valuable insight into each model’s computational demands and practical efficiency.

Table 6. Confusion matrix of DenseNet-121 architecture on the Private, APTOS, and EyePACS datasets.

Empty Cell		DenseNet-121 prediction					
Empty Cell		Predicted values					
Empty Cell		Private dataset		APTOS dataset		EyePACS dataset	
Empty Cell		NO DR	DR	NO DR	DR	NO DR	DR
True values	NO DR	356	4	459	11	879	121
	DR	6	354	13	435	61	939

Figure 8 illustrates the AUROC scores achieved by DenseNet-121, MobileNet-v2, and Inception-v3 on the various datasets. DenseNet-121 attained the highest AUROC values across the board, confirming its superior classification capabilities relative to the competing models. While MobileNet-v2 and Inception-v3 recorded marginally lower AUROC scores, they still delivered strong, competitive results. These observations offer a clear view of the performance-efficiency balance among the tested architectures. Overall, DenseNet-121 emerged as the leading model in terms of AUROC, reflecting its strong classification performance. In contrast, MobileNet-v2 and Inception-v3 offer computational speed benefits due to their lighter, simpler designs,

resulting in faster inference times. This information is vital when choosing the most suitable transfer learning model for the target application. Previous studies have investigated various transfer learning approaches for diabetic retinopathy (DR) detection [18-24]. Ultimately, **Table 8** offers an in-depth comparison between our method and the current state-of-the-art (SOTA) techniques. The comparison emphasizes our method’s success in enhancing the quality of retinal images acquired through mobile devices or fundus photography. This improvement in image quality is critical for enabling precise DR detection and has significant potential to prevent blindness, especially among elderly patients.

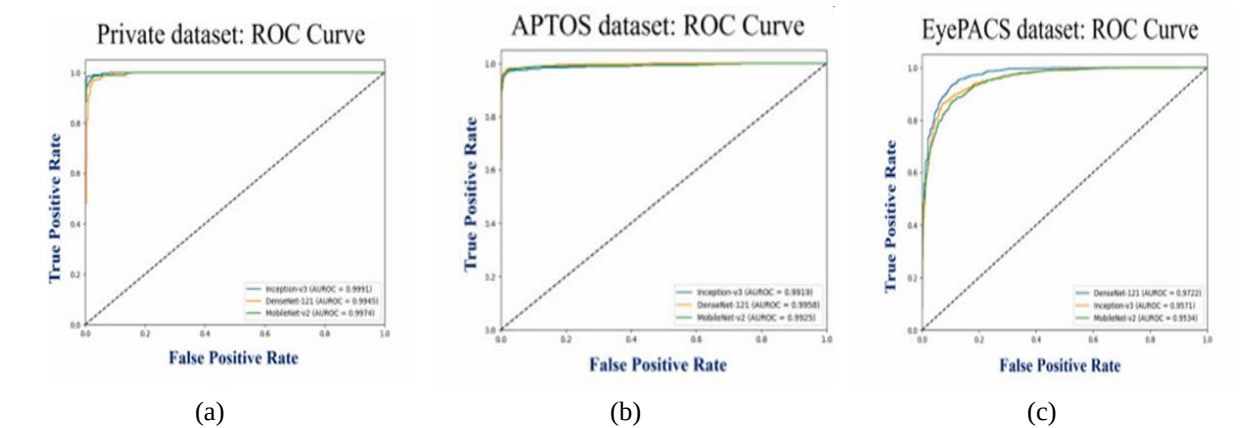


Figure 8. ROC curves obtained from models fitted to the Private (a), APTOS (b), and EyePACS (c) datasets.

Table 7. Summary of performance indices for benchmark analysis: Evaluating model complexity and inference time.

Dataset	Model	Model depth	Model complexity	Memory size (MB)	AUROC	Inference time (ms)
Private	DenseNet-121	431	32.72	223.99	0.9991	161.62
	MobileNet-v2	158	34.37	254.09	0.9974	118.65

	Inception-v3	315	48.01	284.02	0.9945	133.50
	DenseNet-121	431	32.72	223.99	0.9958	210.03
APTOS	MobileNet-v2	158	34.37	254.09	0.9925	157.93
	Inception-v3	315	48.01	284.02	0.9919	170.84
	DenseNet-121	431	32.72	223.99	0.9722	137.52
EyePACS	MobileNet-v2	158	34.37	254.09	0.9534	104.72
	Inception-v3	315	48.01	284.02	0.9571	105.54

Table 8. A comparison with state-of-the-art methods.

Works	Technique	Dataset	Performance metrics
Gulshan <i>et al.</i> [18]	CNNs	EyePACS	Sen = 90.3% , Spe = 98.1%
Wang <i>et al.</i> [20]	TL AlexNet, VGG-16 and InceptionNetV3	EyePACS	Acc = 37.43%, 50.03%, 63.23%
Pamadi <i>et al.</i> [21]	TL MobileNetV2	APTOS	Acc = 78%
Zhou <i>et al.</i> [22]	Multi-Task learning	EyePACS	Acc = 63.22%
Deepa <i>et al.</i> [23]	Ensemble method (InceptionV3, Xception)	EyePACS	Acc = 96.2% for Ensemble method, Acc = 85.4% for InceptionV3
Qureshi <i>et al.</i> [24]	Active Deep Learning (ADL)	EyePACS	Acc = 95.10%, Sen = 92.20%, Sep = 95.10%
Our approach	Fine-tuned DenseNet-121	Private, EyePACS, and APTOS	Acc = 98.61%, 98.03%, 90.90%; Sen = 98.33%, 97.76%, 93.90%; Sep = 100.00%, 99.78%, 99.40%

Real-time mobile app-based teleophthalmology

Creating and implementing a real-time diagnostic tool for diabetic retinopathy (DR) involves multiple key steps. Initially, retinal images are acquired live using an oDocs nun IR handheld fundus camera. These images then undergo enhancement through the designated preprocessing algorithms. The selected deep learning model, DenseNet-121, was optimized and transformed into TensorFlow Lite format to enable efficient operation on mobile devices. Once deployed on smartphones, the full classification pipeline—encompassing preprocessing and inference—completes in an average of 1.2 seconds per image. This speed is well-suited for real-time screening during patient examinations, where prompt feedback is essential. Inference timing was evaluated using Python's time module on 100 test images. These results confirm that the model supports real-time operation on smartphones, fulfilling the practical demands of mobile-based DR screening systems [29].

The DenseNet-121 model, chosen after evaluating several alternatives, was adapted and fine-tuned to reliably detect the presence or absence of DR. This model was subsequently integrated into a mobile smartphone application. Such accessibility allows ophthalmologists to utilize the app efficiently for DR classification and

diagnosis. Diagnostic outcomes are transmitted instantly to a pilot site connected with the Faculty of Medicine at HASSAN II Hospital in Fez, facilitating centralized oversight and ongoing evaluation. This comprehensive setup provides valuable clinical support, improves decision-making, and promotes early identification. Better management of DR. **Figure 9** depicts the workflow for early DR screening in remote healthcare facilities or regions with limited access to ophthalmology specialists.

The mobile platform connects seamlessly with digital tablets and desktop applications via a cloud-based server. This infrastructure supports instantaneous data synchronization across devices, making diagnostic results from the mobile app immediately accessible on both tablet and desktop systems. Cloud-enabled connectivity ensures the solution can be used effectively in diverse medical environments, from remote locations with minimal infrastructure to well-equipped hospital settings. Additionally, the tablet and desktop versions include advanced functionalities for healthcare providers, such as improved data visualization and integration with patient management systems, thereby increasing the system's versatility for both portable and fixed deployments.

The compact, portable nature of the proposed system greatly expands its usefulness in resource-limited settings where traditional fundus imaging equipment is scarce. Utilizing a lightweight handheld retinal camera paired with an efficient mobile application allows deployment in rural or semi-urban regions where patients often lack regular access to eye care specialists. Even under difficult conditions such as inconsistent lighting

and patient movement, the system sustains strong diagnostic accuracy thanks to its robust image enhancement pipeline. Moreover, on-device real-time processing delivers immediate results, enabling quicker clinical decisions and timely patient referrals. This portability helps close the gap in early DR detection for underserved populations, promoting fairer access to healthcare services.

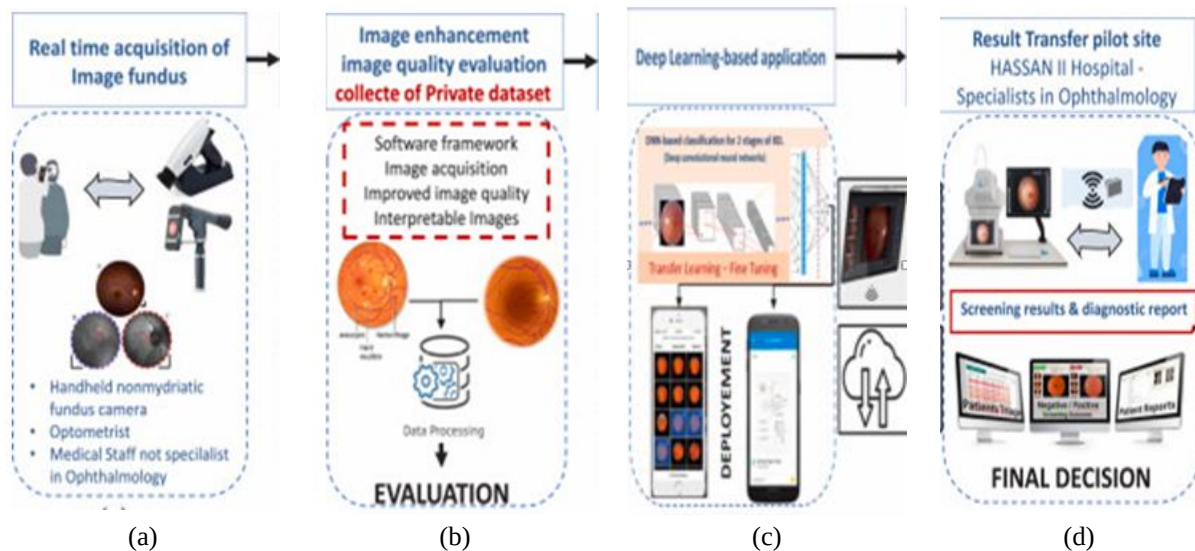


Figure 9. Early DR screening workflow in remote medical centers or resource-limited settings: A complete process covering real-time image acquisition (a), image enhancement (b), deep learning model inference (c), and screening with result transmission (d).

The DenseNet-121 model, valued for its strong performance, has been incorporated into a mobile application to provide a convenient, portable diagnostic tool. A thorough planning phase defined the app's goals, assessed end-user needs—particularly clinicians working in field conditions such as medical caravans—and identified essential functional requirements. The development uses contemporary technologies, with JavaScript as the primary programming language, to enable seamless integration of the DenseNet-121 model. The application supports efficient image processing and reliable automated diagnosis while featuring an intuitive interface that reduces user workload, proving especially useful in fast-paced and resource-scarce environments. Important functionalities include real-time image capture, advanced image quality improvement, and clear presentation of results. At present, the system has been evaluated only by the research team and a small group of medical personnel. In subsequent phases, following official validation by relevant authorities, the application

will be released to a broader group of users who may not be familiar with it. We will then collect feedback and further assess performance. Our team will soon carry out this evaluation process in collaboration with other users, allowing a proper assessment of the user experience. These aspects will form the basis of future work.

Figure 10 outlines the deployment process for the optimized and fine-tuned DenseNet-121 model within a mobile application prototype. The model is first converted to .tflite format using the TensorFlow Lite interpreter. It is then implemented on smartphones using Java in the Android Studio Integrated Development Environment (IDE). Libraries such as TensorFlow, Keras, and OpenCV support efficient image handling and model execution. This successful integration on mobile devices provides a practical and powerful approach to early DR detection.

Deploying the system on smartphones offers several benefits, particularly its portability and ease of access for screening without specialized hardware. These attributes

are especially valuable in environments with limited medical resources, including rural communities and medical outreach programs. Furthermore, extending the platform to desktops and tablets increases its flexibility and potential impact. **Figure 11** illustrates the model running successfully on an Android smartphone, highlighting the system's ability to generate real-time, accurate results. Looking ahead, our objective goes

beyond DR detection alone. After successful validation for diabetic retinopathy, the plan is to expand and generalize the model to identify additional eye diseases, including glaucoma, age-related macular degeneration (AMD), and others. This will involve training on pathology-specific datasets, ultimately enabling clinicians to use the application as a comprehensive tool for diagnosing a broad spectrum of ocular conditions.

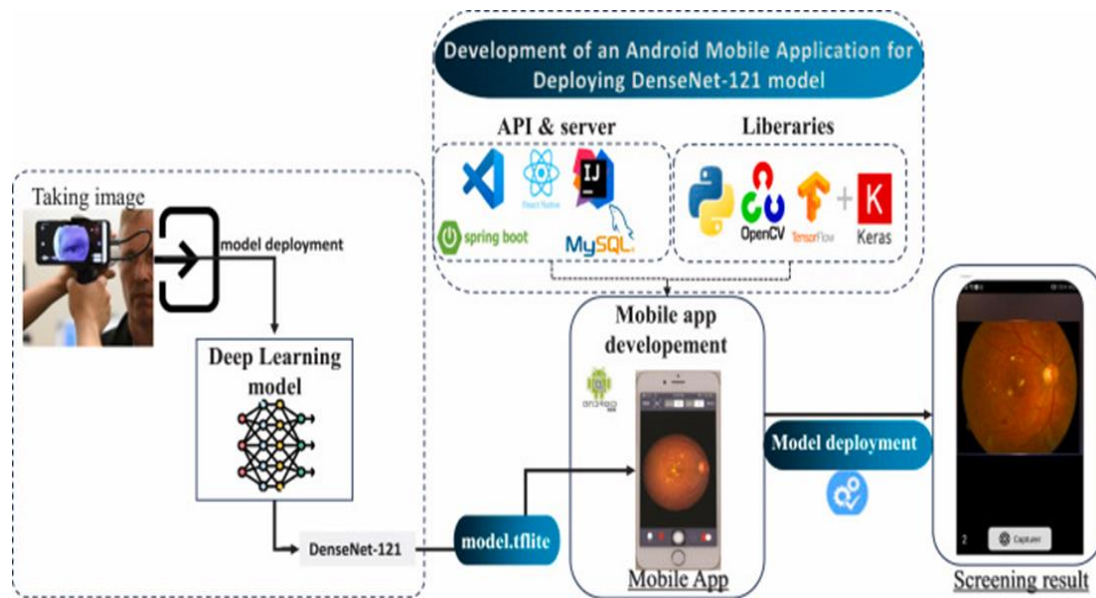


Figure 10. The workflow for implementation on smartphone, desktop, and tablet platforms.

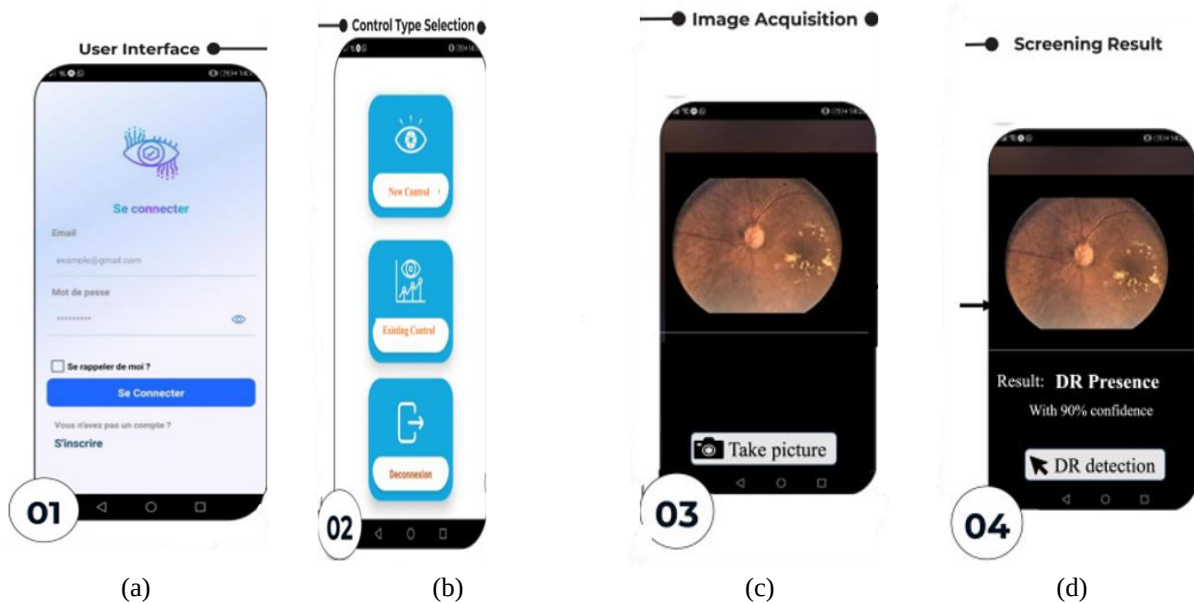


Figure 11. Example of the smartphone interface for diabetic retinopathy detection, showing the connection page (a), selection of control mode (b), image capture step (c), and final screening outcome display.

Fusing the publicly accessible APTOS and EyePACS datasets with our internal private collection has played a key role in confirming the reliability of the findings. Using several heterogeneous datasets helps mitigate bias that could arise from relying on a single data source. APTOS and EyePACS bring valuable diversity through differences in patient populations, acquisition settings, and image qualities, which boosts the model's capacity to generalize effectively. Incorporating our private dataset further adds authentic clinical cases from the CHU Hassan II population, tailoring the model more closely to environments with comparable patient profiles. This multi-dataset strategy significantly improves the system's robustness and reliability, enabling consistent performance across a wide range of situations and reinforcing its suitability for real-time early screening via mobile tools.

To increase the generality of the proposed solution, several improvements remain necessary. Chief among them is expanding the image database with additional samples acquired from different cameras and introducing greater variety in patient age, gender, and geographical backgrounds. User input will also be carefully reviewed to refine the application's ergonomics, simplicity, and level of interactivity. Upcoming research will emphasize uncertainty quantification, along with broader validation using varied fundus cameras, diverse populations, and different clinical contexts, to enhance overall resilience and real-world usefulness.

Conclusion

In this study, we presented a Deep Convolutional Neural Network (DCNN) that uses transfer learning combined with fine-tuning for early identification of diabetic retinopathy (DR). The system centers on an optimized DenseNet-121 architecture and delivers excellent accuracy across public datasets (APTOS and EyePACS) and our private collection. Specifically, the model attained 98.61% accuracy on the private dataset, 98.03% on APTOS, and 90.90% on EyePACS. These outcomes validate the proposed method's robustness and strong generalization potential. Key contributions of this work encompass: (i) creating and refining a compact, high-accuracy DCNN for DR categorization; (ii) embedding the model into a mobile app for on-device real-time processing paired with a portable fundus camera; and (iii) conducting thorough testing on multiple datasets to

verify adaptability and stable performance regardless of image origin.

This research advances medical image analysis by highlighting the advantages of transfer learning across varied data sources and showing that sophisticated deep learning solutions can be implemented in resource-constrained settings. On the applied side, the mobile screening platform is lightweight, economical, and capable of rapid inference (161.62 ms per image), positioning it as an ideal tool for areas lacking sufficient specialist care. Consequently, it supports timely DR detection where expert access is limited, which can enhance patient outcomes and help avert unnecessary blindness. Looking ahead, we will seek to boost model efficiency by shortening inference times and extending compatibility to tablets and desktop computers. We also intend to launch a real-world validation initiative in the Fez region of Morocco to test the system in everyday clinical practice. Lastly, we plan to expand the dataset and conduct in-depth user experience studies to develop further an application that is more user-friendly, ergonomic, and accessible.

Acknowledgments: None

Conflict of Interest: None

Financial Support: None

Ethics Statement: None

References

1. World Health Organization. Blindness and visual impairment [Internet]. 2023 [cited 2025 May 7]. Available from: <https://www.who.int/news-room/fact-sheets/detail/blindness-and-visual-impairment>
2. Teo ZL, Tham YC, Yu M, Chee ML, Rim TH, Cheung N, et al. Global prevalence of diabetic retinopathy and projection of burden through 2045: systematic review and meta-analysis. *Ophthalmology*. 2021;128(11):1580-91.
3. Bilal A, Liu X, Shafiq M, Ahmed Z, Long H. NIMEQ-SACNet: A novel self-attention precision medicine model for vision-threatening diabetic retinopathy using image data. *Comput Biol Med*. 2024;171:108099.

4. Bilal A, Imran A, Baig TI, Liu X, Long H, Alzahrani A, et al. DeepSVDNet: A deep learning-based approach for detecting and classifying vision-threatening diabetic retinopathy in retinal fundus images. *Comput Syst Sci Eng*. 2024;48(2).
5. Burton MJ, Ramke J, Marques AP, Bourne RRA, Congdon N, Jones I, et al. The Lancet Global Health Commission on global eye health: vision beyond 2020. *Lancet Glob Health*. 2021;9(4):e489-551.
6. Begum T, Rahman A, Nomani D, Mamun A, Adams A, Islam S, et al. Diagnostic accuracy of detecting diabetic retinopathy by using digital fundus photographs in the peripheral health facilities of Bangladesh: validation study. *JMIR Public Health Surveill*. 2021;7(3):e23538.
7. Elam AR, Tseng VL, Rodriguez TM, Mike EV, Warren AK, Coleman AL, et al. Disparities in vision health and eye care. *Ophthalmology*. 2022;129(10):e89-113.
8. An D, Tan B, Yu DY, Balaratnasingam C. Differentiating microaneurysm pathophysiology in diabetic retinopathy through objective analysis of capillary nonperfusion, inflammation, and pericytes. *Diabetes*. 2022;71(4):733-46.
9. Walsh L, Hong SC, Chalakkal RJ, Ogbuehi KC. A systematic review of current teleophthalmology services in New Zealand compared to the four comparable countries of the United Kingdom, Australia, United States of America (USA) and Canada. *Clin Ophthalmol*. 2021;15:4015-27.
10. Bangyal WH, Shakir R, Ashraf A, Qayyum ZU, Rehman NU. Automatic detection of diabetic retinopathy from fundus images using machine learning based approaches. In: 2023 25th International Multitopic Conference (INMIC). IEEE; 2023. p. 1-6.
11. Ashraf A, Zhao Q, Bangyal WH, Raza M, Iqbal M. Female autism categorization using CNN based NeuroNet57 and ant colony optimization. *Comput Biol Med*. 2025;189:109926.
12. Ashraf A, Qingjie Z, Bangyal WHK, Iqbal M. Analysis of brain imaging data for the detection of early age autism spectrum disorder using transfer learning approaches for Internet of Things. *IEEE Trans Consum Electron*. 2023;70(1):4478-89.
13. Journal de la société marocaine d'ophtalmologie [Internet]. 2025 [cited 2025 May 7]. Available from: <https://revues.imist.ma/index.php/jsmo>
14. El Alami H, Haddou I, Benaadi G, Lkhider M, El Habchi D, Wakrim L, et al. Prevalence and risk factors of chronic complications among patients with type 2 diabetes mellitus in Morocco: a cross-sectional study. *Pan Afr Med J*. 2022;41(1).
15. Labiris G, Panagiotopoulou EK, Kozobolis VP. A systematic review of teleophthalmological studies in Europe. *Int J Ophthalmol*. 2018;11(2):314.
16. Modaresnia Y, Torghabeh FA, Hosseini SA. Enhancing multi-class diabetic retinopathy detection using tuned hyper-parameters and modified deep transfer learning. *Multimed Tools Appl*. 2024;83(34):81455-76.
17. Quelled G, Lamard M, Erginay A, Chabouis A, Massin P, Cochener B, et al. Automatic detection of referral patients due to retinal pathologies through data mining. *Med Image Anal*. 2016;29:47-64.
18. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A, et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA*. 2016;316(22):2402-10.
19. Wang X, Lu Y, Wang Y, Chen WB. Diabetic retinopathy stage classification using convolutional neural networks. In: 2018 IEEE International Conference on Information Reuse and Integration (IRI). IEEE; 2018. p. 465-71.
20. Wang D, Chen X, Jiang M, Du S, Xu B, Wang J. ADS-net: An attention-based deeply supervised network for remote sensing image change detection. *Int J Appl Earth Obs Geoinf*. 2021;101:102348.
21. Pamadi AM, Ravishankar A, Nithya PA, Jahnvi G, Kathavate S. Diabetic retinopathy detection using mobilenetv2 architecture. In: 2022 International Conference on Smart Technologies and Systems for Next Generation Computing (ICSTSN). IEEE; 2022. p. 1-5.
22. Zhou K, Gu Z, Liu W, Luo W, Cheng J, Gao S, et al. Multi-cell multi-task convolutional neural networks for diabetic retinopathy grading. In: 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). IEEE; 2018. p. 2724-7.
23. Deepa V, Sathish Kumar C, Cherian T. Ensemble of multi-stage deep convolutional neural networks for automated grading of diabetic retinopathy using image patches. *J King Saud Univ Comput Inf Sci*. 2022;34(8):6255-65.

24. Qureshi I, Ma J, Abbas Q. Diabetic retinopathy detection and stage classification in eye fundus images using active deep learning. *Multimed Tools Appl.* 2021;80(8):11691-721.
25. Abou Baker N, Zengeler N, Handmann U. A transfer learning evaluation of deep neural networks for image classification. *Mach Learn Knowl Extr.* 2022;4(1):22-41.
26. Tsiknakis N, Theodoropoulos D, Manikis G, Ktistakis E, Boutsora O, Berto A, et al. Deep learning for diabetic retinopathy detection and classification based on fundus images: A review. *Comput Biol Med.* 2021;135:104599.
27. Chin EK, Ventura BV, See KY, Seibles J, Park SS. Nonmydriatic fundus photography for teleophthalmology diabetic retinopathy screening in rural and urban clinics. *Telemed J E Health.* 2014;20(2):102-8.
28. Iqbal S, Khan TM, Naveed K, Naqvi SS, Nawaz SJ. Recent trends and advances in fundus image analysis: A review. *Comput Biol Med.* 2022;151:106277.
29. Midena E, Zennaro L, Lapo C, Torresin T, Midena G, Pilotto E, et al. Handheld fundus camera for diabetic retinopathy screening: a comparison study with tabletop fundus camera in real-life setting. *J Clin Med.* 2022;11(9):2352.
30. Lu L, Ausayakhun S, Ausayakuhn S, Khunsongkiet P, Apivatthakakul A, Sun CQ, et al. Diagnostic accuracy of handheld fundus photography: A comparative study of three commercially available cameras. *PLOS Digit Health.* 2022;1(11):e0000131.
31. Rogers TW, Gonzalez-Bueno J, Garcia Franco R, Lopez Star E, Méndez Marín D, Vassallo J, et al. Evaluation of an AI system for the detection of diabetic retinopathy from images captured with a handheld portable fundus camera: the MAILOR AI study. *Eye.* 2021;35(2):632-8.
32. Khodris C, Ahmed B, Fouad C, Meriem A, Idriss BA, Tairi H. Artificial intelligence in ophthalmology: the ophthalmologist's opinion. In: 2020 International Conference on Intelligent Systems and Computer Vision (ISCV). IEEE; 2020. p. 1-5.
33. Hassan M, Ahmed B, Fouad C, Meriem A, Benatiya A, Faouzi BM. Ganglion cell layer analysis in nonarteritic anterior ischemic optic neuropathy in diabetic patients: From the acute to resolving phases. *J Res Med Dent Sci.* 2021;9(11).
34. Mvoulana A, Kachouri R, Akil M. Fully automated method for glaucoma screening using robust optic nerve head detection and unsupervised segmentation based cup-to-disc ratio computation in retinal fundus images. *Comput Med Imaging Graph.* 2019;77:101643.
35. Ben Sayadia S, Elloumi Y, Kachouri R, Akil M, Ben Abdallah A, Bedoui MH. Automated method for real-time AMD screening of fundus images dedicated for mobile devices. *Med Biol Eng Comput.* 2022;60(5):1449-79.
36. Takahashi H, Tampo H, Arai Y, Inoue Y, Kawashima H. Applying artificial intelligence to disease staging: Deep learning for improved staging of diabetic retinopathy. *PLoS One.* 2017;12(6):e0179790.
37. Bilal A, Imran A, Baig TI, Liu X, Long H, Alzahrani A, et al. Improved support vector machine based on CNN-SVD for vision-threatening diabetic retinopathy detection and classification. *PLoS One.* 2024;19(1):e0295951.
38. Dash S, Senapati MR. Enhancing detection of retinal blood vessels by combined approach of DWT, tyler coye and Gamma correction. *Biomed Signal Process Control.* 2020;57:101740.
39. Henry AGP, Jude A. Convolutional neural-network-based classification of retinal images with different combinations of filtering techniques. *Open Comput Sci.* 2021;11(1):480-90.
40. oDocs Eye Care. oDocs nun IR [Internet]. 2025 [cited 2025 May 7]. Available from: <https://odocseyecare.shop/products/odocs-nun-ir>
41. Kaggle. APTOS 2019 blindness detection [Internet]. 2019 [cited 2025 May 7]. Available from: <https://kaggle.com/competitions/aptos2019-blindness-detection>
42. Papers with Code. Kaggle EyePACS dataset [Internet]. 2025 [cited 2025 May 7]. Available from: <https://paperswithcode.com/dataset/kaggle-eyepacs>
43. Acharya UK, Kumar S. Swarm intelligence based adaptive gamma corrected (SIAGC) retinal image enhancement technique for early detection of diabetic retinopathy. *Optik.* 2021;247:167904.
44. University of Southampton. Median filtering - computer vision website header [Internet]. 2025 [cited 2025 May 7]. Available

- from: https://www.southampton.ac.uk/~msn/book/new_demo/median/
45. Ningsih DR, Sari YP, Harahap M, Siregar B, Simanjuntak J. Improving retinal image quality using the contrast stretching, histogram equalization, and CLAHE methods with median filters. *Int J Image Graph Signal Process.* 2020;14(2):30.
 46. Mohanty C, Mahapatra S, Acharya B, Kokkoras F, Gerogiannis VC, Karamitsos I, et al. Using deep learning architectures for detection and classification of diabetic retinopathy. *Sensors.* 2023;23(12):5726.
 47. Qureshi I, Sharif M, Yasmin M, Raza M, Javed MY. Computer aided systems for diabetic retinopathy detection using digital fundus images: A survey. *Curr Med Imaging Rev.* 2016;12(4):234-41.
 48. Mahmoud HAH. RETRACTED: Diabetic retinopathy progression prediction using a deep learning model. *Axioms.* 2022;11(11):614.
 49. Mookiah MRK, Acharya UR, Chua CK, Lim CM, Ng EYK, Laude A. Computer-aided diagnosis of diabetic retinopathy: A review. *Comput Biol Med.* 2013;43(12):2136-55.
 50. Gargeya R, Leng T. Automated identification of diabetic retinopathy using deep learning. *Ophthalmology.* 2017;124(7):962-9.
 51. Huang G, Liu Z, Maaten LVD, Weinberger KQ. Densely connected convolutional networks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.* Vol. 1. 2017. p. 3-12.
 52. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition;* 2016. p. 2818-26.
 53. Sandler M, Howard A, Zhu M, Zhmoginov A, Chen LC. MobileNetV2: Inverted residuals and linear bottlenecks. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition;* 2018. p. 4510-20.