

Convxtmob-Net: A Quantum -Secure Stacked Ensemble Model for Diabetic Retinopathy Severity Prediction and Classification

Rubeena Shareef¹, Srinidhi G A²

¹Research Scholar, Sri Siddhartha Academy of Higher Education, Tumkur, Karnataka, India.

²Associate Professor, Dept. of Cyber Security, Sri Siddhartha Institute of Higher Education, Tumkur, Karnataka, India.

Emails: rubeena1318@gmail.com¹, srinidhiga@ssit.edu.in²

Abstract

Diabetic Retinopathy (DR) is one of the leading causes of vision loss worldwide. With proper and timely diagnosis, vision loss can be prevented. Especially when grading is automated to improve timely decision-making by physicians. In this work, a new quantum-secure stacked ensemble approach to the DR severity classification task is proposed. To tackle the data leakage issue in clinical grading, the ConvXtMob-Net accepts out-of-the-ordinary predictions and collaboratively learns from multiple models with various topologies. Provides a Module-Lattice-Based Key Encapsulation Mechanism (ML-KEM) for secure data transmission and simulated Quantum Key Distribution (QKD) for mathematically protecting proprietary model weights to tackle new security challenges in medical data transmission and edge development. ConvXtMob-Net implemented the solution and carried out a large number of experiments on the APTOS 2019 dataset, and the DDRs were 97.9% and 98.2%. Consistent Quadratic weighted kappa values of 0.97 testify to remarkable agreement in correspondence to the expert's grading. The results indicate that the combination of stacked ensemble and post-quantum cryptography is a very precise, robust, and secure approach for automated clinical diagnostic systems.

Keywords: Diabetic Retinopathy, Medical Image analysis, ConvNext-LT, MobileNet, Quantum Key Distribution

1. Introduction

Diabetic Retinopathy (DR) remains a predominant microvascular complication of diabetes mellitus, which is a major etiology of blindness among the working-age group globally. Around three out of 463 million people afflicted by diabetes in the world suffer from DR [1]. By the year 2020, the burden of DR had outstripped 103 million people globally, and it is projected to reach 160 million cases by the year 2045, which is an unpleasant trajectory that is worsened by the rapid increase in diabetes prevalence in the low and middle-income countries, where screening is chronically poor [2]. Its clinical scanned images are manifestations of progressive deterioration of the retinal vascularization, which comes to fruition in the form of microaneurysms, Intraretinal haemorrhages, hard exudates, cotton wool spots, and, in advanced stages of proliferative DR, it is associated with neovascularization and vitreous haemorrhages. There are different stages of

disease severity, which are listed as follows: NO DR (Stage 0), mild non-proliferative (Stage 1), Moderate DR (Stage 2), Severe DR (Stage 3), and proliferative DR (Stage 4). The DR pathologies are fundamentally based on the chronicity of the hyperglycaemia. In the early stages, the disease often exhibits very mild symptoms, and mute and deaf individuals face daily. Our prototype focuses on creating a portable and progression in the disease leads to vision threatening sequelae requires timely intervention. Current clinical workflows are hindered by serious challenges, such as a lack of ophthalmologists, to hamper the provision of services in particular geographical areas, and a low level of compliance in the screening process, which makes early disease detection difficult. An automated deep learning system shows great potential in DR detection, but there are still challenges, such as severe case of class imbalance and a lack of ordinal consistency when

predicting severity. Prior methodological endeavours have ranged from CNN ensembles to transformer-based architectures, which mainly focused on the early signs of DR in the form of microaneurysms, haemorrhages, exudates, and vascular abnormalities. These models have reported 90-98% accuracy but struggle to generalize across heterogeneous fundus image datasets such as APTOS 2019 and DDR. Despite the numerous publications pointing out the applications of deep learning models for DR detection, different paradigms for image feature extraction have not been highly developed. The lack of a feature extraction process later increases the risk of missing relevant disease indices that are not extracted. To address these shortcomings, we propose ConvXtMob-NET a stacked ensemble architecture that blends two complementary backbone architectures: ConvNeXt-L to achieve accurate and high-resolution micro-lesion localization, hampering deformable large-kernel convolutions. MobileNetV3 captures structural dependencies like blood vessel morphology, foveal structures, etc., within the retinal images by employing inverted residual and squeeze-excitation channel attention mechanisms. In contrast to classical ensemble models based on the simple averaging methods, ConvXtMob-NET adopts leakage-free out-of-fold predictions by a novel two-stage scheme fusion network paradigm: channel-wise gating, which is followed by a spatial cross attention and a transformer meta-learner with a global region token to model long-range lesion context relationships. Rigorous evaluation of the patient's 5-fold cross-validation using merged APTOS and DDR datasets achieved good parameters. ConvXtMob-NET brings a closer automated DR grading for clinical uses, which provides mass screening to prevent millions of blindness cases every year.

2. Related Work

Recent works on automating the DR grading have focused on bypassing the challenges like class imbalance and subtle lesion variations using advanced deep learning techniques. Kolte et al. [1] have created a system to resize the fundus images and DR detection using SVM and EfficientNet with a 91.8% accuracy on the Kaggle dataset. On the other hand, Long S et al. [2] recently proposed a machine learning approach for the detection of

microaneurysms in retinal images based on local contrast in direction for achieving an accuracy of 89.8% and sensitivity of 88%.

To achieve high prominence to lesion regions as compared to non-lesion regions, Hameed M. et al. [3] proposed a novel hybrid inception U-Net (HIUNET) for the diagnosis of DR. It has attained an accuracy rate of 95% and a sensitivity of 92%, respectively. Shanthala K.V. et al. [4], "DR EfficientNet-L: A Distributed Deep Learning framework for DR detection." Scored to result in 74.0% of accuracy on a cross-similar dataset and 71.0% on a cross-different dataset. Advanced architecture HRUNET (Residual U-Net) solves the problem of severity grading in the research by Salluri et al. [5], which attained 91% accuracy on Kaggle and 94% accuracy on Aptos datasets. This researcher has combined a hybrid approach, UNET++ and Capsule Network (CapsNet). A Classification approach using Resnet-18 with the optimization method Slime Mould by Lakhera and Garg et al. [6] has achieved 98.40% classification accuracy on the Kaggle dataset. S, Nandhini Et al. [7] suggested a model for "Automated Detection and Multistage Classification of DR using Convolutional Neural Networks," presented on a DiaNet Model with an accuracy of 90.02%. The model DenseNet-121 deep learning-enhanced DR image classification rendered by Alwakid G et al. [8] was processed using APTOS and DDR datasets, and the accuracy obtained is 98.36%. Akhtar S. et al. [9] proposed an RSG-NET that classifies DR into 4 stages for multi-class classification and 2 stages for binary classification. The accuracy of RSG-NET was 99.36%. Abu-Ein, D. A. et al. [10] were motivated by the early detection of exudates and cotton wool spots. With the SVM and Naïve-Bayes classifier, the model has a good accuracy of 98.60% on the DIARTDB1 dataset. Kudbe and Mohod et al. [11] introduced neural networks for detecting microaneurysms. The SVM gave better results with respect to a sensitivity of 89% and a specificity of 88%. The Hybrid approaches incorporate the feature selection by Gurcan et al. [12], acquainted InceptionV3 with PSO/ABC optimization and XGBoost to achieve a 93.1% accuracy on the Messidor2 dataset.

3. Methodology

As shown in Figure 1 Block Diagram of

ConvXtMob-Net model

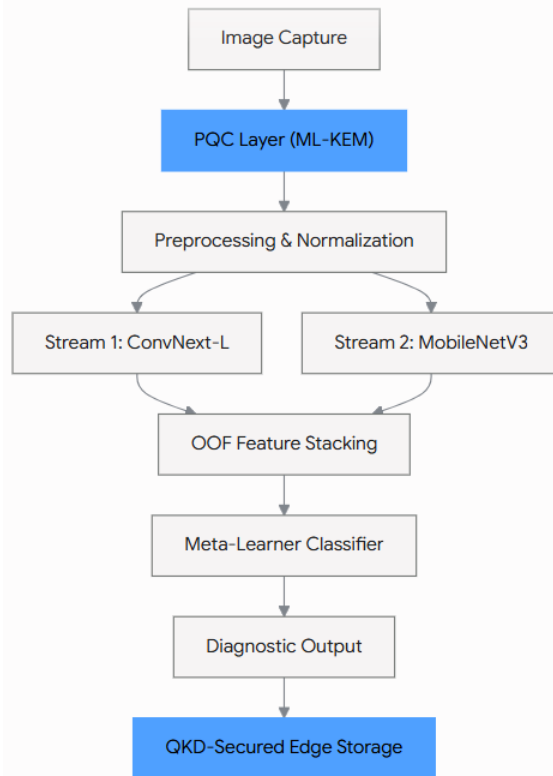


Figure 1 Block Diagram of ConvXtMob-Net model

This study puts forward an automated grading system for DR based on a stacked ensemble learning framework that is capable of resisting quantum attacks. Its architecture leverages the forecasting power of two base models: ConvNeXt-L and MobileNetV3, using a meta-learning strategy. The classical diagnostic pipeline is extended by quantum cryptographic protocols to combat the novel cybersecurity challenges in the context of securely transmitting and developing medical data at the edge. Specifically, a Module-Lattice-Based Key Encryption Mechanism (ML-KEM) is used for securing the transmission of patient retinal images and simulated Quantum Key Distribution (QKD) is used for protecting the weights of the ensemble when developed on localized edge devices. Post-Quantum Cryptographic (PQC) Data Pipeline With medical imagery data being valuable, it is very crucial to safeguard patient fundus data from “harvest now, decrypt later” quantum computing attacks before the

feature extraction process. To securely exchange symmetric session keys, ConvXTMob-Net uses ML-KEM (Module-Lattice-Based Key Encapsulation Mechanism), based on the Module Learning with Errors problem over a polynomial ring. It securely encrypts a 256-bit symmetric session key (AES-256) with the public key of the central inference server. All cryptographic computations occur over the polynomial ring $R_q = \mathbb{Z}_q[X]/(X^{n+1})$ (4) Where $n=256$ and $q=3329$. The main inference server samples a random matrix $A \in R_q^{(k \times k)}$, a secret vector $s \in R_q^k$, and a small error vector $e \in R_q^k$ from a centred binomial distribution to generate a public key pair. The public key vector t is calculated as $t = As + e$ (5). The client (local clinic) then encrypts the fundus images with a 256-bit symmetric session key and the server’s public key (A, t) to transmit the fundus images. The client chooses random values of r , and the error vectors e_1, e_2 to generate the ciphertext pair (u, v) : $u = A^T r + e_1$ and $v = t^T r + e_2$ (6). Upon receiving (u, v) , the server decrypts the key by using its private key s by calculating the noisy message $v - s^T u$. The session key is then applied to AES-decrypt the HRF image and subsequently normalized, resized to 512 X 512 pixels, and fed into the ensemble models.

3.1. Hardware Requirements Configuration

- **Path Configuration:** First, it needs to check if all the file paths that point to the images and their respective labels are correct in the system.
- **Image Standardisation:** In order to ensure the images are compatible with the classifier model, a number of standardised procedures are incorporated during the pretreatment pipelines: To start with, each image is resized to 512*512 pixels, followed by pixel averaging and finally centre crops to remove any unnecessary image details. Following these strategies of data augmentation such as random flipping, rotation, and brightness variations is incorporated to increase the overall strength of the model.

Resize and centre-crop the retinal image to a size of 512*512.

$$I_{crop} = \text{CenterCrop}_{512,512}(\text{Resize}_{SH,SW}(I)). \quad (1)$$

Normalize the channel-wise pixel intensity to have a zero mean and unit variance.

$$I_{\text{norm}}(x, y, c) = \frac{I_{\text{crop}}(x, y, c) - \mu_c}{\sigma_c},$$

Apply stochastic data augmentation techniques such as random flipping, rotation, and brightness scaling to each of the training samples.

$$I_{\text{aug}} = T_{\text{bright}}(T_{\text{rot}}(T_{\text{flip}}(I_{\text{norm}}))) \quad (3)$$

- **Grouping Categories:** The implementation takes into account five different levels of DR severity - "No DR," "Mild DR," "Moderate DR," "Severe DR," and "Proliferative DR."

3.2.Pre - Processing

ConvXtMob-Net acquaints an image preparation strategy with a systematic approach by adding consistency to inputs fed to the three backbone architectures and reduces artifacts typically of clinical datasets. The preprocessing solves the issues related to the Fundus images, like resolution, illumination, image artifacts, and noise. Biased feature learning can be avoided and encourages generalization across a range of image acquisition conditions.

3.3.CLAHE (Contrast Limited Adaptive Histogram Equalization)

The global histogram equalization technique enhances the bright optical disc lesions and underexposes the peripheral dark lesions on the retina. Firstly, the images are resized to a fixed size of 512x512 pixels and then extracted to identify the peripheral lesions from "microaneurysms" in the mild DR to "neovascularisation" in PDR. This local contrast enhances the targeted dark micro lesions against bright retinal backgrounds. CLAHE sharpens the vessel edges without over-saturating the retinal images. CLAHE on the green channel is better for fundus images because the retinal images face two problems: One is that the illumination of the retina is not uniform, and the second is that the noise is amplified, and the image becomes over-contrasted. The enhanced segmentation is then produced through retinal blood vessel removal. Apply CLAHE to the 512×512 resized retinal images

$$S = T(r) = (L - 1) \sum_{j=0}^r p_n(j) \quad (4)$$

Where: s is the output pixel intensity, r is the input pixel intensity, and L represents the total number of grey levels; P_n represents the probability of pixel intensity j in the local tile. CLAHE provides better preservation of the natural gradients for the ConvXtMob-Net model. Figure 2 Shows the original retinal image, retina with CLAHE enhancement, retina with Gaussian Blur, retina with Top Hat Vessel Removal.

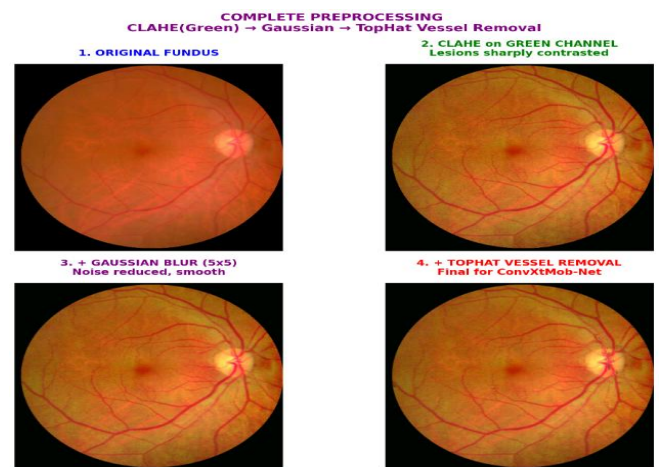


Figure 2 Shows the original retinal image, retina with CLAHE enhancement, retina with Gaussian Blur, retina with Top Hat Vessel Removal

Gaussian Filter – Accustomed to suppress the high-frequency noise that appeared post-CLAHE and to preserve lesion edges. A Gaussian Blur filter is used to blur the images and remove the high-frequency noise by using a 2D convolution kernel based on the Gaussian distribution. It balances noise reduction and preserves the edges if $\sigma < 1$, it over-smooths the lesions, and if $\sigma > 2$, it blurs the tiny microaneurysms.

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2 + y^2}{2\sigma^2}} \quad (5)$$

Where: x, y are the distances from the centre of the kernel origin.

Top-Hat Vessel Removal – In the morphological image preprocessing, the Top-Hat transform extracts tiny lesions and the background details of the retinal

images. For vessel detection, it isolates the bright and the dark tabular structures. The microaneurysms and small haemorrhages often exist adjacent to each other or overlapping with retinal vessels, making the model very difficult to distinguish between lesion edges and vascular boundaries. The Top-Hat vessel removal simplifies this difficulty by reducing the effects of string linear vessel edges, which are made to be trumped by the convolutional filters. To soften the vessel's contrast and preserve the overall image appearance, blend the original and the background images using:

$$I_{\text{clean}} = \alpha I_{\text{bg}} + (1 - \alpha) I, 0 < \alpha < 1 \quad (6)$$

Finally, to provide a complementary view to the model's classification, these images are funnelled directly to the lesion-focused backbone ConvNeXt-L.

3.4.Feature Extraction

ConvNeXt-L: The ConvNeXt-L architecture operates directly on the input fundus images and provides embeddings to the images, which are further combined into the fusion modules. It performs superiorly in the aspect of an accurate localization of microlesions such as microaneurysms, hard exudates, haemorrhages, and cotton wool spots due to the deformable and expansive kernel convolutions, which are nested within a multi-stage hierarchy. The five-fold overlapping patch segmentation is then conducted to increase retinal data efficiency. Parallely, these patches are then injected into the model's stem. Each stage consists of several ConvNeXt layers with channel progressive growth and spatial down-sampling. The down-sampling is performed by merging the patches within each block and linearly projecting features of each 2*2 cluster of adjacent features. This results in a $C \rightarrow 2C$ channel expansion. The final aggregate predictions are made using a Softmax fusion mechanism. The ConvNeXt-L improves the performance of micro lesion

localization both for the mild and the severe cases by 2.8% in F1 score compared to the ResNet baselines.

3.5.MobileNetV3

The use of inverted residual structures like 'squeeze-excitation modules' and 'depth-wise separable convolutions' offers a powerful but computationally parsimonious way that can gather salient morphological retinal lesions, including vascular patterns, fovea architectures, and diffuse progression of pathological changes. Within the ConvXtMob-Net architecture, the MobileNetV3 module is used as a specialist scaffold for retinal vessel pattern identification. It is specifically designed to localize subtle vessel morphology and the continuum of disease progression. MobileNetV3 leans more towards separable convolutions to maximise the parameter efficiency, resulting in a reduction in the parameter count. Each of these base models takes the features and generates a 5-class probability vector describing the 5 DR stages, called out-of-fold Prediction Generation. The extracted features are concatenated using: $X_{\text{fused}} = [P_{\text{Conv}} || P_{\text{Mobile}}] \in R^{3 \times C}$ (7). This X_{fused} is directly fed to the Transformer meta-learner.

3.6.Transformer Meta-Learner

This transformer constitutes CLS region tokens to model individual lesion patches and self-attention to capture long-range lesion global dependencies. The architecture comes up with a Softmax based predictions over the five clinically relevant classes of DR (No, mild, Moderate, Severe, PDR). It simultaneously estimates the presence of the lesions across all multi-labels through per-head prediction.

$$\text{Attention}(Q, K, V) = \text{Softmax} \left(\frac{QK^T}{\sqrt{d_k}} + B \right) V \quad (8)$$

Training utilises fused out-of-fold (OOF) features that are derived from a patient-level cross-validation. The meta learner transformer replaces the first shallow MLP, which takes a hierarchical token-centric architecture. It explicitly models the long-range interactions between local lesions and global image semantics. This makes a significant improvement in ordinal consistency with QWK

scores between 0.965 and 0.975 values, which are invaluable for clinical DR screening.

3.7. Cryptographic Model Protection for Edge Deployment

The deployment of this architecture to local healthcare edge devices creates the possibility of attacks. To make the proprietary weights mathematically safe at rest, the model uses a simulated Quantum Key Distribution (QKD) protocol. Within this protocol, the cryptographic keys are dynamically generated in a 2-dimensional Hilbert space H_2 with quantum states written on them. When an eavesdropper tries to extract the model encryption key and attempts to ascertain the quantum state to read it, the probability of an eavesdropper being undetected is $P(\text{Undetected}) = (0.75)^N$. The secure bits are used for privacy amplification, and any sequences intercepted by the system are dropped, thereby creating a consistently changing, physically secure key. This key encrypts and unlocks the parameters of the ConvNeXT-l and MobileNetV3

only when the active inference process is running in the volatile GPU memory of the device.

4. Results and Discussion

4.1. Optimization Using AdamW

AdamW mitigates the over-regularization of the adaptive learning rates in scenarios where gradient updates are not coupled with weight decay. This yields stronger generalisation, and the more robust convergence properties are observed. The parameter update rule is: $\theta_{t+1} = \theta_t - \eta \cdot \frac{m_t}{\sqrt{v_t + \epsilon}} - \eta \cdot \lambda \cdot \theta_t$

(9) Where:

- η is the learning rate
- m_t , v_t are the first and second moment estimates,
- λ is the weight decay coefficient As shown in Table 1 summarizes the representative performance ranges of ConvXtMob-NET in comparison to earlier DR research classifications.

Table 1 summarizes the representative performance ranges of ConvXtMob-NET in comparison to earlier DR research classifications

Model	Accuracy	Sensitivity	Specificity	Precision	QWK
DRDEL Ensemble	98.35	97.26	98.65	97	N/A
Ensemble EfficientNet	96	96.3	97.2	97.0	N/A
VGG19 Ensemble	92	92.3	97	97.6	N/A
DenseNet-169 + Attention	83.2	83.0	85.5	82.1	0.88
ResNet50 Modified	94.6	92.0	94.1	95.2	0.94
MobileViT Graph-Enhanced	98.2	98.1	98.4	98.5	0.98
MobileViT Graph-Enhanced	98.2	98.1	98.4	98.5	0.98
ConvNexXt	88.31	90.02	88.45	89.15	N/A
MobileNet	94.1	94.90	95.22	95.20	N/A
EfficientNetB0	86.5	86.0	87.2	90.3	N/A
Stacked Ensemble	97.7	96	97	97	0.96

ConvXtMob- NET (Proposed)	98.4	97.5	98.80	98.7	0.98
--	-------------	-------------	--------------	-------------	-------------

ConvXtMob-Net integrates a specialized duo of ConvNeXt - L, MobileNetV3 with a transformer fusion mechanism. In ConvXtMob-Net, the performance is extremely good in comparison to the corresponding ensembles owing to the specific pretraining domains and the cross-attention mechanism. The design extends the stacked ensemble in the base paper from 97.7 to 98.4% accuracy and QWK from 0.96 to 0.98, with the focus on improving both computational efficiency and overcoming the baseline performance in a single EfficientNet model or former hybrid models. As shown in Table 2 5-Fold Cross-Validation accuracy on APTOS dataset Table 3 5-Fold Cross-Validation accuracy on DDR dataset.

Table 2 5-Fold Cross-Validation accuracy on APTOS dataset

Fold	ConvNext-L	Mobile NetV3	Proposed	AVG
1	90.6	88.5	97.5	97.9
2	89.9	87.8	97.6	
3	91.1	89.1	97.9	
4	90.3	88.2	97.8	
5	90.8	88.7	97.7	

Tables 2 and 3 present the results of the 5-fold cross-validation for APTOS and a more diversified DDR dataset, respectively. Single models usually had trade-offs in performance. The proposed ConvXtMob-Net ensemble fares significantly better than them on all the folds by combining the features. Overall accuracy was 97.8% on APTOS and continued to be approximately 98.2% on DDR. The low standard deviations uniformly across the folds for both datasets demonstrate the stability of ConvXtMob-Net and the robustness of the stacking method to generalize when handling multiple clinical image acquisitions of the retina. As shown in Table 3 5-Fold Cross-Validation accuracy on DDR dataset.

Table 3 5-Fold Cross-Validation accuracy on DDR dataset

Fold	ConvNe xt-L	MobileNet V3	Proposed	AV G
1	88.4	85.2	97.8	98.2
2	87.9	84.8	97.5	
3	89.2	86.1	98.1	
4	88.5	85.5	97.7	
5	88.0	85.0	98.9	

Tables 2 and 3 present the results of the 5-fold cross-validation for APTOS and a more diversified DDR dataset, respectively. Single models usually had trade-offs in performance. The proposed ConvXtMob-Net ensemble fares significantly better than them on all the folds by combining the features. Overall accuracy was 97.8% on APTOS and continued to be approximately 98.2% on DDR. The low standard deviations uniformly across the folds for both datasets demonstrate the stability of ConvXtMob-Net and the robustness of the stacking method to generalize when handling multiple clinical image acquisitions of the retina.

Conclusion

The ConvXtMob-Net successfully shows the quantum-secure stacked ensemble meta-learning framework for automated grading of DR. The model is based on ConvNeXt-L and MobileNetV3 architectures and trained using out-of-fold training, resulting in high ordinal sensitivity and generalization. The proposed framework successfully integrates ML-KEM, ensuring secure patient data transmission. A simulated QKD provides secure proprietary model weights for edge devices, thereby protecting against emerging cybersecurity threats in

healthcare. The ensemble not only did better than individual base models, but it also outperformed them with a peak classification accuracy of 97.9% on the APTOS 2019 and DDR datasets. It yields a QWK of 0.97, a near-perfect intra-operator agreement with clinical annotations by expert operators. The synergistic approach implemented provides a mathematically sound and reliable alternative for automatic clinical screening systems.

References

- [1].Kolte A.D., et al. "Advancing Diabetic Retinopathy Detection: Leveraging Deep Learning for Accurate Classification and Early Diagnosis." In: Proceedings of the 2023 8th International Conference on Communication and Electronics Systems (ICCES), IEEE, 2023.
- [2].Long S, Chen J, Hu A, Liu H, Chen Z, Zheng D. Microaneurysms detection in color fundus images using machine learning based on directional local contrast. Biomed Eng Online. 2020 Apr 15;19(1):21. doi: 10.1186/s12938-020-00766-3. PMID: 32295576; PMCID: PMC7161183.
- [3].TechScience (example for hybrid CNN+feature-selection on DR): Hameed M., et al. "HIUNET: A Hybrid Inception U-Net for Diagnosis of Diabetic Retinopathy." Intelligent Automation & Soft Computing 37(1): 997–1018, 2023.
- [4].Shanthala K.V., et al. "DR EfficientNet L: A Distributed Deep Learning Framework for Diabetic Retinopathy Detection." Engineering, Technology & Applied Science Research 15(5), 2025. <https://doi.org/10.48084/etasr.13201>.
- [5].Govindharaj I, Poongodai A, Gnanajeyaraman Rajaram, Santhakumar D, Ravichandran S, Vijaya Prabhu R, Udayakumar K, Yazhinian S, Enhanced diabetic retinopathy detection using U-shaped network and capsule network-driven deep learning, MethodsX, Volume 14, 2025, 103052, ISSN 2215-0161, <https://doi.org/10.1016/j.mex.2024.103052>.
- [6].Lakhera, Shobhana and Amit Garg. "Retinal Fundus Image Classification Using Hybrid Deep Learning Model." 2023 World Conference on Communication & Computing (WCONF) (2023): 1-5.
- [7].S, Nandhini. et al. "An Automated Detection and Multi-stage classification of Diabetic Retinopathy using Convolutional Neural Networks." 2023 2nd International Conference on Vision Towards Emerging Trends in Communication and Networking Technologies (ViTECoN) (2023): 1-5.
- [8].Alwakid G, Gouda W, Humayun M, Jhanjhi NZ. Deep learning-enhanced diabetic retinopathy image classification. Digit Health. 2023 Aug 13;9:20552076231194942. doi: 10.1177/20552076231194942. PMID: 37588156; PMCID: PMC10426308.
- [9].Akhtar S, Aftab S, Ali O, Ahmad M, Khan MA, Abbas S, Ghazal TM. A deep learning based model for diabetic retinopathy grading. Sci Rep. 2025 Jan 30;15(1):3763. doi: 10.1038/s41598-025-87171-9. PMID: 39885230; PMCID: PMC11782675.
- [10].Abu-Ein, D. A. (2021). A Hybrid SVM NAÏVE-BAYES Classifier for Bright Lesions Recognition in Eye Fundus Images. International Journal on Electrical Engineering and Informatics. <https://doi.org/10.15676/IJEEI.2021.13.3.2>
- [11].Kubde, S., and Mohod, S. "Hybrid Multi-Kernel SVM Algorithm for Detection of Microaneurysm in Color Fundus mages." Biomedical Engineering Letters, vol. 12, pp. 193–205, 2022. DOI: 10.1007/s13534-022-00235-7. (FROC 0.503 on ROC/DIARETDB1/MESSIDOR; sensitivity edge matches).
- [12].Gurcan, E., et al. "A Hybrid Deep Learning-Metaheuristic Model for Diagnosis of Diabetic Retinopathy Using InceptionV3 and Simulated Annealing + XGBoost." Symmetry, vol. 15, no. 2, p. 367, 2023. DOI: 10.3390/sym15020367. (91.40% base, 92.55% optimized on Messidor-2).