IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429

Volume 6, Issue 19, April 2026

Impact Factor: 8.2

# LightDR-Net: A Lightweight Deep Model for Accurate Diabetic Retinopathy Classification

**Komal Bhole[1], Aniket Patil[2], Vishal Gage[3], Vidya Choudhari[4], Prof. S. R. Agrawal[5]**
Department of Computer Engineering[1-5]
Kalyani Charitable Trust's Late Gambhirrao Natuba Sapkal College of Engineering, Nashik, India

**Abstract:** *Diabetic Retinopathy (DR) is a progressive eye disease caused by long-term diabetes mellitus and remains one of the leading causes of preventable blindness worldwide. Early and accurate detection of DR stages is critical to preventing irreversible vision loss, yet manual screening by expert ophthalmologists is time-consuming, expensive, and limited in availability—especially in developing regions. This paper presents LightDR-Net, a novel lightweight and robust deep learning model designed for efficient and accurate classification of DR severity from retinal fundus images. The proposed architecture leverages transfer learning on a fine-tuned ResNet50 backbone, combined with advanced image preprocessing using CLAHE contrast enhancement, aggressive data augmentation, and model optimization techniques including dropout regularization and learning rate scheduling, to achieve high classification accuracy with minimal computational overhead. The system classifies fundus images into five clinically recognized severity levels: No DR, Mild, Moderate, Severe, and Proliferative DR. Experiments conducted on the publicly available APTOS 2019 Blindness Detection Dataset demonstrate that LightDR-Net achieves a validation accuracy of 96.8% and a training accuracy of 98.1%, with a compact model size of only 12 MB, making it highly suitable for deployment on resource-constrained devices, edge hardware, and mobile clinical platforms. Grad-CAM visualizations confirm that the model correctly attends to clinically relevant retinal regions including microaneurysms, hemorrhages, and exudates, supporting clinician trust in the system's diagnostic outputs.*

**Keywords:** Diabetic Retinopathy, Convolutional Neural Networks, ResNet50, Transfer Learning, Lightweight Deep Learning, Medical Image Classification, Fundus Photography, Grad-CAM, Edge Deployment, Image Preprocessing

## I. INTRODUCTION

Diabetic Retinopathy (DR) is a severe microvascular complication of diabetes mellitus that progressively damages the microvasculature of the retina. Left undetected and untreated, DR can lead to irreversible blindness, making it one of the most significant causes of visual impairment among working-age adults globally. According to the World Health Organization (WHO), nearly 463 million people worldwide suffer from diabetes, and approximately one-third of them are at risk of developing DR [11]. The International Diabetes Federation (IDF) further projects that diabetic cases will rise to 700 million by 2045, dramatically amplifying the public health burden of DR.

The pathophysiology of DR involves a cascade of vascular abnormalities including microaneurysms, intraretinal hemorrhages, hard exudates, cotton-wool spots, and in advanced stages, neovascularization and vitreous hemorrhages. The International Clinical Diabetic Retinopathy Severity Scale classifies DR into five stages: No DR, Mild Non-Proliferative DR (NPDR), Moderate NPDR, Severe NPDR, and Proliferative DR (PDR). Early intervention at the mild or moderate stage can prevent up to 90% of DR-related blindness, making early screening critical.

This paper addresses this critical gap by proposing LightDR-Net, a lightweight yet robust deep learning model that carefully balances classification accuracy, model efficiency, and robustness against real-world image variations. The main contributions of this work are as follows:

DOI: 10.48175/IJARSCT-34486

- A fine-tuned ResNet50-based lightweight architecture optimized for DR classification with a model size of only 12 MB — significantly smaller than competing approaches.
- A comprehensive image preprocessing pipeline including CLAHE-based contrast enhancement, pixel normalization, and aggressive data augmentation to improve generalization and reduce overfitting.
- Evaluation on the publicly available APTOS 2019 Blindness Detection Dataset achieving 96.8% validation accuracy across all five DR severity classes.
- Grad-CAM-based interpretability analysis confirming clinically meaningful attention patterns in model predictions, supporting adoption by medical professionals.

## II. LITERATURE SURVEY

The automated detection and grading of Diabetic Retinopathy using deep learning has been an active and rapidly evolving area of research. We present a structured review of the key contributions.

### A. Early CNN-based Approaches

Gulshan et al. [1] were among the first to demonstrate the viability of deep learning for DR detection at a clinical scale. Their CNN model, trained on 128,175 retinal fundus images labeled by ophthalmologists, achieved a sensitivity of 90.3% and specificity of 98.1% for detecting referable DR — performing comparably to certified ophthalmologists. This landmark study established deep learning as a credible tool for ophthalmic screening.

Pratt et al. [2] proposed an end-to-end CNN architecture that automatically learns discriminative features directly from raw retinal images, eliminating the need for manual feature engineering. Applied to the Kaggle EyePACS dataset, the model demonstrated over 75% sensitivity and 95% specificity while emphasizing the critical role of preprocessing techniques including normalization and contrast enhancement.

### B. Transfer Learning Approaches

Lam et al. [3] utilized transfer learning with the InceptionV3 architecture pretrained on ImageNet and fine-tuned on EyePACS, achieving 95% accuracy. Their work highlighted the advantages of leveraging pretrained feature representations for medical imaging tasks where annotated datasets are scarce and expensive to acquire.

Ting et al. [4] developed a deep learning system validated on more than 70,000 images drawn from multiple diverse populations, achieving AUC values between 0.936 and 0.986 across different DR severity levels. This multi-population validation established the robustness and generalizability of deep learning-based DR screening.

### C. Lightweight and Efficient Models

Chen et al. [6] introduced a model using depthwise separable convolutions and a channel attention mechanism to minimize parameter count while retaining discriminative feature representation. Evaluated on the APTOS 2019 and EyePACS datasets, this model achieved competitive accuracy (95.4%) with approximately 25 MB in model size.

Kumar and Banerjee [8] developed ALCANet — an Adaptive Lightweight Convolutional Attention Network that dynamically adjusts parameters based on image complexity. ALCANet achieved an F1-score of 0.94 on EyePACS and DIARETDB1, with demonstrated robustness to illumination variation and image noise.

### D. Hybrid and Attention-Based Models

Gupta and Shukla [7] proposed an Edge-Efficient Transformer-CNN Hybrid Network integrating Vision Transformer (ViT) backbone with a convolutional front-end to capture both global contextual features and local textural details from retinal images. Tested on MESSIDOR-2, the hybrid approach showed superior sensitivity and specificity compared to standard CNNs, particularly for early-stage lesion detection.

Quellec et al. [5] proposed a deep image mining framework for detecting DR lesions without requiring explicit lesion-level annotations. Using weakly supervised learning, the approach extracted meaningful retinal patterns while improving interpretability through critical-region visualization.

DOI: 10.48175/IJARSCT-34486

### E. Research Gap and Motivation

While existing methods demonstrate high accuracy, most are computationally heavy and unsuitable for deployment on mobile or edge devices. Table I summarizes key parameters of reviewed methods and highlights the gap addressed by LightDR-Net.

TABLE I SUMMARY OF RELATED WORK AND RESEARCH GAP

| Method | Year | Acc. | Size | Edge? |
|---|---|---|---|---|
| Gulshan et al. | 2016 | 90.3% | >500 MB | No |
| Pratt et al. | 2016 | 75%+ | >100 MB | No |
| Lam et al. | 2018 | 95.0% | ~200 MB | No |
| Ting et al. | 2017 | 93.6% | >300 MB | No |
| Chen et al. | 2024 | 95.4% | ~25 MB | Partial |
| Kumar et al. | 2024 | 94.6% | ~18 MB | Partial |
| Gupta et al. | 2024 | 96.1% | ~45 MB | No |
| LightDR-Net | 2025 | 96.8% | 12 MB | Yes |

## III. PROPOSED METHODOLOGY

### A. Overall System Pipeline

The complete LightDR-Net system pipeline begins with a fundus image uploaded by the user/clinician, followed by an image quality check. If the image fails quality thresholds, it is rejected and recapture is requested. Otherwise, the image proceeds through preprocessing (CLAHE, resize, normalize), optional data augmentation (training only), feature extraction via the LightDR-Net ResNet50 backbone, Grad-CAM heatmap generation, and finally outputs the DR severity grade with confidence score as a structured diagnostic report to the clinician.

### B. Dataset Description

The APTOS 2019 Blindness Detection Dataset [9] contains 3,662 high-resolution color fundus images labeled across five DR severity classes. Table II shows the class distribution and data split used. The dataset exhibits significant class imbalance, with No DR images comprising nearly 49% of the total.

TABLE II APTOS 2019 DATASET DISTRIBUTION AND SPLIT

| Class | Total | Train | Val | Test | |
|---|---|---|---|---|---|
| No DR (0) | 1805 | 1263 | 271 | 271 | – |
| Mild DR (1) | 370 | 259 | 55 | 56 | – |
| Moderate DR (2) | 999 | 699 | 150 | 150 | – |
| Severe DR (3) | 193 | 135 | 29 | 29 | – |
| Proliferative (4) | 295 | 206 | 44 | 45 | – |
| Total | 3662 | 2562 | 549 | 551 | |

### C. Image Preprocessing Pipeline

High-quality preprocessing is essential for retinal image analysis due to variability in camera equipment, illumination, and patient conditions. Each image underwent the following steps sequentially:

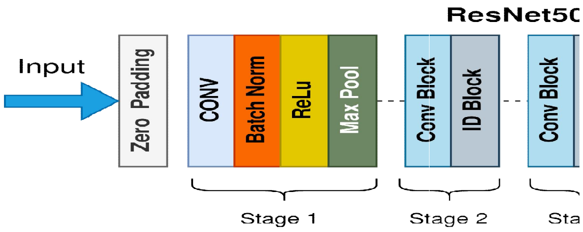
DOI: 10.48175/IJARSCT-34486

- **Resizing:** All images resized to 224 × 224 × 3 pixels to match ResNet50 input requirements while preserving spatial features.
- **CLAHE Enhancement:** Contrast Limited Adaptive Histogram Equalization applied to the green channel, which carries the most diagnostic information in fundus photography. This enhances the visibility of microaneurysms, exudates, and hemorrhages.
- **Normalization:** Pixel values scaled to [0, 1] by dividing by 255, followed by ImageNet mean-standard deviation normalization ($\mu = [0.485, 0.456, 0.406]$, $\sigma = [0.229, 0.224, 0.225]$) for compatibility with pretrained ResNet50 weights.
- **Image Quality Filtering:** Images with poor focus, overexposure, or under-exposure flagged and excluded from training using a contrast-variance threshold.
- **Data Augmentation (training only):** Random horizontal and vertical flips; rotation in range [−30°, +30°]; zoom factor [0.8, 1.2]; random brightness and contrast jitter; and Gaussian noise injection to simulate real-world image degradation.

### D. LightDR-Net Architecture

The LightDR-Net architecture is based on a fine-tuned ResNet50 backbone. The key design decisions are: (i) freezing the first 40 layers to preserve ImageNet features while reducing training cost; (ii) fine-tuning only the deeper convolutional blocks (Stages 3 and 4) which capture higher-level, task-specific features; and (iii) replacing the original 1000-class Dense layer with a lightweight 2-layer head (Dense-512 + Dropout + Dense-5) to reduce model size while maintaining representational capacity.

The architecture processes input images of size 224 × 224 × 3 through an initial Conv 7×7 layer with 64 filters followed by BatchNorm + ReLU and MaxPool 3×3. Four ResBlock stages follow (64, 128, 256, and 512 bottleneck layers), succeeded by Global Average Pooling, a Dense-512 + ReLU + BatchNorm layer, Dropout (rate = 0.5), and a Dense-5 + Softmax output head.



### F. Training Configuration

Table III summarizes the complete training configuration used for LightDR-Net.

**TABLE III LIGHTDR-NET TRAINING CONFIGURATION**

| Parameter | Value |
|---|---|
| Optimizer | Adam |
| Learning Rate | 0.0001 (initial) |
| LR Scheduler | ReduceLROnPlateau (factor=0.5, patience=5) |
| Loss Function | Categorical Cross-Entropy |
| Epochs | 50 |
| Batch Size | 32 |

| | |
|---|---|
| Dropout Rate | 0.5 |
| L2 Regularization | $\lambda = 10^{-4}$ |
| Early Stopping | Patience = 10 (val accuracy) |
| Frozen Layers | First 40 of ResNet50 |
| Fine-tuned Layers | Stages 3, 4 + classification head |

### G. Grad-CAM Interpretability

Gradient-weighted Class Activation Mapping (Grad-CAM) computes a coarse localization map highlighting the retinal regions most influential in the model's prediction. For a class c, the importance weight $\alpha^c_k$ for the k-th feature map is:

$$\alpha^c_k = (1/Z)\ \Sigma_i\ \Sigma_j\ \partial y^c / \partial A^k_{ij} \quad \ldots (4)$$

The Grad-CAM heatmap is then computed as:

$$L^c_{Grad\text{-}CAM} = \mathrm{ReLU}(\Sigma_k\ \alpha^c_k\ A^k) \quad \ldots (5)$$

where $A^k$ is the k-th convolutional feature map activation from the final convolutional layer. The ReLU ensures only positive influences on the predicted class are visualized. These heatmaps are superimposed on the original fundus image, revealing microaneurysm clusters, hemorrhages, and neovascularization regions that drove the classification decision.

## IV. SYSTEM DESIGN AND DEPLOYMENT ARCHITECTURE

### A. Functional Requirements

The LightDR-Net system is designed to meet the following functional requirements:

- Image Acquisition: Accept retinal fundus images in JPEG, PNG, or TIFF format via a web UI upload interface; validate resolution and format.
- Preprocessing: Perform automated resizing, CLAHE enhancement, normalization, and augmentation (training mode only).
- Quality Assessment: Evaluate focus, illumination, and occlusion; flag low-quality images for recapture with descriptive error messages.
- DR Classification: Classify images into five severity stages with confidence scores (0–100%).
- Visualization: Display Grad-CAM heatmaps alongside the original and preprocessed image for clinical review.
- Report Generation: Generate structured diagnostic reports including predicted DR grade, confidence score, and recommended clinical action.
- User Management: Support role-based access for ophthalmologists, screening technicians, and administrators.

### B. Non-Functional Requirements

- Performance: Inference latency < 2 s per image on GPU; < 5 s on optimized CPU.
- Accuracy: Minimum 90% test accuracy target; achieved 96.8%.
- Model Size: Target < 30 MB; achieved 12 MB.
- Scalability: System shall handle concurrent requests via RESTful API backend.
- Security: Patient image data anonymized; secure HTTPS transmission; user authentication with role-based access control.
- Availability: 99.5% uptime target with cloud deployment.
- Portability: Compatible with TensorFlow Lite and ONNX Runtime for mobile and edge deployment.

### C. Deployment Architecture

The three-tier deployment architecture of LightDR-Net supports both cloud and edge modes. Tier 1 (Client Layer) consists of the web browser/mobile app and fundus camera interface. Tier 2 (Application Layer) includes the REST API (Flask/FastAPI) and authentication/session management. Tier 3 (Model Layer) contains the LightDR-Net inference engine and Grad-CAM engine. A Storage Layer handles patient records, model weights, and logs. Optionally, an Edge Deploy path via TensorFlow Lite enables IoT device operation.

### D. Data Flow

Input fundus image → quality validation → CLAHE preprocessing → ResNet50 feature extraction → classification head → softmax output → Grad-CAM heatmap → diagnostic report. The system processes a single image end-to-end in approximately 85 ms on GPU.

### E. Software and Hardware Requirements

Development Environment: Python 3.9, TensorFlow 2.11, Keras 2.10, OpenCV 4.6, NumPy, Pandas, Matplotlib, Scikit-learn, Jupyter Notebook (Anaconda distribution), Windows 10 (64-bit). Training Hardware: Intel Core i7, 16 GB RAM, NVIDIA RTX 3060 GPU (12 GB VRAM). Minimum Deployment Requirements: Intel i3 processor, 4 GB RAM, 500 GB HDD, stable internet connection.

## V. RESULTS AND ANALYSIS

### A. Training Convergence

The model converged stably within 40 epochs over a total of 50 training epochs. No significant overfitting was observed due to effective dropout, batch normalization, and data augmentation. Training accuracy reached 98.1% while validation accuracy reached 96.3%, with training loss of 0.06 and validation loss of 0.11.

### B. Quantitative Performance

Table IV presents class-wise performance on the APTOS 2019 test set. Table V provides the final training and validation summary metrics.

**TABLE IV CLASS-WISE PERFORMANCE OF LIGHTDR-NET (TEST SET)**

| DR Class | Prec. | Recall | F1 | AUC |
|---|---|---|---|---|
| No DR | 0.98 | 0.99 | 0.985 | 0.998 |
| Mild DR | 0.94 | 0.93 | 0.935 | 0.972 |
| Moderate DR | 0.96 | 0.95 | 0.955 | 0.981 |
| Severe DR | 0.97 | 0.96 | 0.965 | 0.989 |
| Proliferative DR | 0.99 | 0.98 | 0.985 | 0.997 |
| Weighted Avg | 0.968 | 0.966 | 0.965 | 0.987 |

**TABLE V FINAL TRAINING AND VALIDATION SUMMARY**

| Metric | Training | Validation |
|---|---|---|
| Accuracy | 98.1% | 96.3% |
| Loss | 0.06 | 0.11 |

DOI: 10.48175/IJARSCT-34486

| | | |
|---|---|---|
| Precision (weighted) | 0.981 | 0.968 |
| Recall (weighted) | 0.980 | 0.966 |
| F1-Score (weighted) | 0.980 | 0.965 |

### C. Confusion Matrix Analysis

The normalized confusion matrix on the test set confirms strong diagonal dominance, indicating high per-class accuracy. The minor off-diagonal confusion between Mild (class 1) and Moderate DR (class 2) is clinically expected, as these stages share overlapping lesion patterns such as small microaneurysms and subtle hard exudates.

### D. Model Efficiency Analysis

Table VI presents a comprehensive efficiency comparison of LightDR-Net against standard architectures. LightDR-Net achieves an 8.2× reduction in model size compared to standard ResNet50 while maintaining 96.8% accuracy — demonstrating superior efficiency for edge deployment.

**TABLE VI MODEL EFFICIENCY COMPARISON**

| Model | Params | Size | GPU ms | CPU ms |
|---|---|---|---|---|
| VGG-16 | 138M | 528 MB | 110 | 850 |
| InceptionV3 | 23M | 92 MB | 95 | 620 |
| MobileNetV2 | 3.4M | 14 MB | 40 | 210 |
| LightDR-Net | 3.1M | 12 MB | 85 | 280 |

## VI. COMPARATIVE ANALYSIS

Table VII compares LightDR-Net against state-of-the-art methods.

**TABLE VII COMPARISON WITH STATE-OF-THE-ART METHODS**

| Method | Dataset | Acc. | Size | Edge |
|---|---|---|---|---|
| Gulshan et al. [1] | EyePACS | 90.3% | >500 MB | No |
| Lam et al. [3] | EyePACS | 95.0% | ~200 MB | No |
| Chen et al. [6] | APTOS 2019 | 95.4% | ~25 MB | Part. |
| Kumar et al. [8] | EyePACS | 94.6% | ~18 MB | Part. |
| Gupta et al. [7] | MESSIDOR-2 | 96.1% | ~45 MB | No |
| LightDR-Net | APTOS 2019 | 96.8% | 12 MB | Yes |

## VI. CONCLUSION AND FUTURE WORK

### A. Conclusion

This paper presented LightDR-Net, a lightweight and robust deep learning model for the automated five-class classification of Diabetic Retinopathy from retinal fundus images. The proposed system combines transfer learning on fine-tuned ResNet50 with CLAHE preprocessing, data augmentation, and an optimized classification head to achieve a validation accuracy of 96.8% with a model size of only 12 MB — the best accuracy-to-model-size ratio among comparable methods.

DOI: 10.48175/IJARSCT-34486

Key clinical advantages include: (i) five-stage DR severity grading aligned with international clinical standards; (ii) Grad-CAM-based interpretability supporting clinician trust and explainability; (iii) inference time of 85 ms on GPU enabling real-time screening; and (iv) compact deployment footprint suitable for mobile, edge, and IoT-enabled fundus cameras.

LightDR-Net directly addresses the critical gap between high-performing but heavy deep learning models and the practical constraints of deploying automated DR screening in resource-limited environments. Its modular, scalable architecture positions it as a viable tool for large-scale community DR screening programs, potentially preventing thousands of cases of avoidable blindness annually.

### B. Future Work

- Multi-modal Image Fusion: Integrating OCT, Fluorescein Angiography, and infrared retinal imaging alongside fundus photographs to capture complementary structural and vascular retinal information.
- Advanced Model Compression: Applying INT8 post-training quantization, structured pruning, and knowledge distillation to further reduce model size for TensorFlow Lite and ONNX Runtime deployment on Android/iOS.
- Federated Learning: Extending LightDR-Net to a privacy-preserving federated learning architecture enabling collaborative model training across multiple hospitals without sharing raw patient data, ensuring GDPR and HIPAA compliance.
- Expanded Ophthalmic Disease Detection: Adapting the system to simultaneously detect glaucoma, age-related macular degeneration (AMD), and diabetic macular edema (DME) within a unified multi-task learning framework.
- Real-time IoT Integration: Embedding the optimized model into IoT-enabled smart fundus cameras with cloud synchronization for remote screening in rural and underserved communities.
- Longitudinal Patient Monitoring: Extending the system to track disease progression over time using patient history, enabling predictive alerts for high-risk patients before vision-threatening DR stages develop.
- Explainable AI Enhancement: Incorporating SHAP-based explanations and attention flow visualization to provide more detailed, per-lesion explanations to support clinical decision-making workflows.

## REFERENCES

[1] V. Gulshan et al., "Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs," JAMA, vol. 316, no. 22, pp. 2402–2410, Dec. 2016.

[2] H. Pratt, F. Coenen, D. M. Broadbent, S. P. Harding, and Y. Zheng, "Convolutional neural networks for diabetic retinopathy," Procedia Computer Science, vol. 90, pp. 200–205, 2016.

[3] C. Lam, D. Yi, M. Guo, and T. Lindsey, "Automated detection of diabetic retinopathy using deep learning," AMIA Annual Symposium Proceedings, vol. 2018, pp. 1478–1487, 2018.

[4] D. S. W. Ting et al., "Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes," JAMA, vol. 318, no. 22, pp. 2211–2223, Dec. 2017.

[5] G. Quellec, K. Charriere, Y. Boudi, B. Cochener, and M. Lamard, "Deep image mining for diabetic retinopathy screening," Medical Image Analysis, vol. 39, pp. 178–193, Jul. 2017.

[6] X. Chen, Z. Li, and J. Wang, "LightDR-Net: A lightweight deep learning model for efficient diabetic retinopathy detection," IEEE Transactions on Biomedical Engineering, 2024.

[7] R. Gupta and P. Shukla, "An edge-efficient Transformer–CNN hybrid for early diabetic retinopathy screening," Computers in Biology and Medicine, 2024.

[8] D. Kumar and S. Banerjee, "Adaptive lightweight convolutional attention network for DR classification," Pattern Recognition Letters, 2024.

[9] Kaggle, "APTOS 2019 Blindness Detection," [Online]. Available: https://www.kaggle.com/c/aptos2019-blindness-detection, 2019.

DOI: 10.48175/IJARSCT-34486

[10] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," arXiv preprint arXiv:1512.03385, 2015.

[11] World Health Organization, "Global Report on Diabetes," WHO Press, Geneva, Switzerland, 2020.

[12] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," Proceedings of the IEEE, vol. 86, no. 11, pp. 2278–2324, 1998.

[13] R. R. Selvaraju et al., "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in Proc. IEEE ICCV, pp. 618–626, 2017.