International Journal of Scientific Research and Engineering Development-– Volume 9 Issue 2, Mar-Apr 2026
Available at www.ijsred.com

RESEARCH ARTICLE OPEN ACCESS

A Systematic Review of Deep Learning Algorithms for Automated Diabetic Retinopathy Detection and Classification

Shikha Methil[1], Dr. Mukul Shrivastava[2]
[1]Research Scholar, [2]Assistant Professor
[1,2]Department of Electronic and Communication Engineering
[1,2]Bansal Institute of Science and Technology Bhopal

Abstract
Diabetic Retinopathy (DR) is one of the leading causes of preventable blindness worldwide, making early detection and accurate classification critical for effective clinical intervention. In recent years, deep learning (DL) algorithms have demonstrated remarkable potential in automating DR diagnosis using retinal fundus images. This systematic review focuses exclusively on the development, comparison, and performance evaluation of various deep learning algorithms applied to DR detection and classification. The study categorizes existing approaches into Convolutional Neural Networks (CNNs), transfer learning models, ensemble methods, attention-based architectures, and hybrid frameworks. Each category is analyzed in terms of architectural design, learning capability, computational complexity, and diagnostic accuracy. CNN-based models form the foundation of most systems, while transfer learning improves performance under limited data conditions. Ensemble and hybrid models enhance robustness and predictive accuracy, whereas attention mechanisms improve interpretability by focusing on clinically relevant regions. The review also examines commonly used evaluation metrics such as accuracy, sensitivity, and AUC. Finally, future research directions are discussed, emphasizing lightweight architectures, explainable AI, and real-time deployment for practical clinical applications.

Keywords: Diabetic Retinopathy, Deep Learning, CNN, Transfer Learning, Ensemble Learning, Attention Mechanism, Medical Image Analysis.

1. Introduction
Diabetic Retinopathy (DR) is a progressive microvascular complication of diabetes that affects the retinal blood vessels and can ultimately lead to irreversible vision loss if not detected at an early stage. With the global rise in diabetes prevalence, the burden of DR has increased significantly, placing immense pressure on healthcare systems. Early screening and timely diagnosis are therefore essential to prevent disease progression and reduce the risk of blindness. However, traditional screening methods rely heavily on manual examination of retinal fundus images by ophthalmologists, which is not only time-consuming but also subject to inter-observer variability and diagnostic inconsistency. In recent years, deep learning (DL) algorithms have emerged as a transformative approach in medical image analysis, particularly in ophthalmology. These algorithms, especially convolutional architectures, have demonstrated the ability to automatically learn complex and hierarchical feature representations directly from raw image data. Unlike conventional machine learning methods that depend on handcrafted features, DL models eliminate the need for manual feature extraction, thereby improving both efficiency and accuracy. This capability makes them highly suitable for detecting subtle retinal abnormalities such as microaneurysms, hemorrhages, and exudates associated with DR. This paper presents a systematic review that focuses exclusively on the algorithmic advancements in deep learning for automated DR detection and classification. It critically examines various model architectures and learning strategies while deliberately excluding dataset-specific and hardware-oriented discussions, thereby providing a clear and focused perspective on algorithmic innovations in this domain.

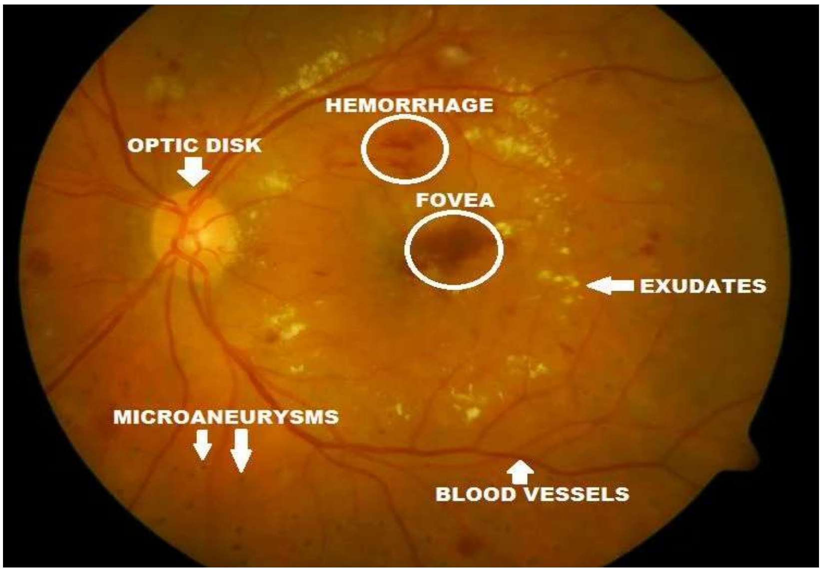

Figure 1: Automated Diabetic Retinopathy Detection Using Deep Learning

2. Methodology of Review

This systematic review adopts a structured and rigorous methodology to analyze recent advancements in deep learning algorithms for automated diabetic retinopathy (DR) detection and classification. The primary objective is to provide a comprehensive algorithm-centric evaluation while ensuring the inclusion of high-quality and relevant research contributions. To achieve this, a well-defined selection process was followed, incorporating clear inclusion and exclusion criteria along with a systematic categorization of algorithms. The methodology emphasizes analytical consistency and comparability across studies, enabling a critical understanding of how different deep learning architectures contribute to performance improvements in DR diagnosis. Furthermore, the review prioritizes recent literature to reflect current trends and innovations in the field, particularly focusing on models developed between 2020 and 2026. By narrowing the scope to algorithmic advancements, the study avoids confounding factors such as dataset variability and hardware dependencies, thereby maintaining a focused and academically robust evaluation framework.

2.1 Inclusion Criteria

The inclusion criteria for this review were carefully defined to ensure that only relevant and high-quality studies were considered. First, only those research works that explicitly utilize deep learning techniques for diabetic retinopathy detection and classification were included. This encompasses models such as Convolutional Neural Networks (CNNs), transfer learning architectures, attention-based networks, and hybrid deep learning frameworks. Studies focusing on both binary classification (DR vs. non-DR) and multi-class classification (grading severity levels) were considered, provided they demonstrated clear algorithmic contributions.
Second, priority was given to studies that emphasize algorithmic design, architectural innovation, or performance optimization. This includes works proposing new model architectures, improving existing deep learning models, or combining multiple approaches to enhance predictive accuracy.
Third, only peer-reviewed journal articles, conference proceedings, and reputed preprints (such as those from recognized repositories) were included to maintain academic credibility. Additionally, recent publications were preferred to capture the latest advancements in the domain. Studies reporting quantitative evaluation metrics such as accuracy, sensitivity, specificity, and AUC were particularly favored, as they enable objective comparison of algorithmic effectiveness. This selective approach ensures that the review remains focused, relevant, and scientifically rigorous.

2.2 Exclusion Criteria

To maintain the specificity and clarity of the review, several exclusion criteria were applied during the study selection process. Firstly, studies that do not employ deep learning techniques were excluded. This includes traditional machine learning approaches such as Support Vector Machines (SVM), Random Forests, and handcrafted feature-based methods, as the focus of this review is strictly on deep learning algorithms. Secondly, research papers that primarily emphasize datasets, data augmentation techniques, or preprocessing methods without contributing significantly to algorithmic development were omitted. While such studies are valuable, they fall outside the scope of this algorithm-focused review.

Additionally, studies involving non-medical imaging applications or unrelated domains were excluded to maintain domain relevance. For example, works applying deep learning to general object detection or non-retinal imaging tasks were not considered. Papers lacking sufficient experimental validation or those without clear performance metrics were also excluded, as they do not provide a reliable basis for comparison. Furthermore, duplicate studies or extended versions of previously published work were carefully filtered to avoid redundancy. By applying these exclusion criteria, the review ensures a high level of precision and relevance, focusing solely on meaningful contributions to deep learning-based DR detection.

2.3 Algorithm Categorization

To facilitate a systematic and comparative analysis, the selected studies were categorized based on the type of deep learning algorithms employed. This classification enables a structured understanding of how different algorithmic approaches contribute to DR detection and classification performance. The first category includes Convolutional Neural Networks (CNNs), which form the foundational architecture for most image-based medical diagnosis systems due to their strong feature extraction capabilities. The second category comprises transfer learning models, where pre-trained networks such as ResNet, VGG, and EfficientNet are fine-tuned for DR tasks, significantly reducing training time and improving performance on limited datasets.

The third category involves ensemble learning methods, which combine multiple deep learning models to enhance prediction accuracy and robustness. These approaches leverage the strengths of individual models while minimizing their weaknesses. The fourth category focuses on attention-based architectures, including models that integrate spatial, channel, or self-attention mechanisms to improve interpretability and focus on clinically relevant regions of retinal images. Finally, the fifth category includes hybrid and emerging models, which integrate multiple techniques such as CNNs with transformers, recurrent networks, or classical classifiers. This categorization provides a clear framework for analyzing algorithmic trends, strengths, and limitations across different approaches, thereby enabling a comprehensive and insightful review.

3. Convolutional Neural Networks (CNNs)

3.1 Basic CNN Architectures

Convolutional Neural Networks (CNNs) form the fundamental backbone of most automated diabetic retinopathy (DR) detection systems due to their exceptional capability in processing image data. A typical CNN architecture consists of convolutional layers for feature extraction, pooling layers for dimensionality reduction, and fully connected layers for classification. These models automatically learn hierarchical feature representations, capturing low-level features such as edges and textures, as well as high-level patterns like lesions and abnormalities in retinal images. Popular early architectures such as LeNet, AlexNet, and VGGNet have been widely adopted and adapted for DR classification tasks. Their primary advantages include high classification accuracy and reduced reliance on manual feature engineering. However, these models require large annotated datasets for effective training and involve substantial computational cost, limiting their applicability in resource-constrained environments.

3.2 Deep CNN Variants

Advanced deep CNN architectures have significantly enhanced the performance of DR detection systems by addressing key limitations of basic CNN models. Architectures such as ResNet, DenseNet, and Inception Network introduce innovative design strategies that enable deeper and more efficient networks. For instance, residual connections in ResNet mitigate the vanishing gradient problem, allowing models to train effectively even with hundreds of layers. DenseNet improves feature reuse by establishing direct connections between all layers, enhancing information flow and reducing redundancy. Inception networks utilize parallel convolutional operations to capture multi-scale features. These innovations collectively contribute to improved accuracy, robustness, and generalization. Empirical studies demonstrate that deep CNN variants frequently achieve accuracy levels exceeding 90% in

DR classification tasks, making them highly effective for clinical decision support systems.

4. Transfer Learning Algorithms

4.1 Concept of Transfer Learning

Transfer learning is a widely adopted approach in deep learning that leverages knowledge gained from models pre-trained on large-scale datasets and adapts it to specific tasks such as diabetic retinopathy (DR) detection. Instead of training a model from scratch, transfer learning involves fine-tuning an existing network—typically trained on datasets like ImageNet—to classify retinal fundus images. This approach is particularly beneficial in medical imaging, where labeled data is often limited and expensive to obtain. By reusing learned feature representations such as edges, textures, and patterns, transfer learning accelerates convergence and enhances model performance. It also reduces computational cost and training time while maintaining high accuracy, making it a practical and efficient solution for DR classification tasks in real-world clinical applications.

4.2 Common Pre-trained Models

Several pre-trained deep learning architectures have been successfully applied to DR detection through transfer learning. Notable models include VGG16 and VGG19, which are known for their simplicity and uniform architecture; ResNet50, which introduces residual connections to enable deeper learning; InceptionV3, which captures multi-scale features efficiently; and EfficientNet, which optimizes model scaling for better accuracy and efficiency. These models significantly reduce training time and perform well even with limited datasets, while also improving generalization. However, they may not fully capture domain-specific retinal features and can lead to overfitting if not properly fine-tuned. Despite these limitations, transfer learning models consistently outperform basic CNNs, achieving high sensitivity and specificity in DR classification.

5. Ensemble Learning Algorithms

5.1 Concept

Ensemble learning is an advanced machine learning paradigm that enhances predictive performance by combining multiple models rather than relying on a single classifier. In the context of diabetic retinopathy (DR) detection, ensemble methods integrate outputs from several deep learning models—typically convolutional neural networks—to produce a more accurate and robust final prediction. The underlying principle is that different models capture different aspects of the data, and their combination reduces individual model biases and errors. This approach is particularly beneficial in medical image analysis, where variability in retinal images can affect model performance. By aggregating predictions through voting or averaging mechanisms, ensemble learning improves generalization and minimizes the risk of misclassification, making it a highly effective strategy for reliable DR diagnosis.

5.2 Types of Ensembles

Ensemble learning techniques can be broadly categorized into bagging, boosting, and stacking, each offering unique advantages in DR classification tasks. Bagging (Bootstrap Aggregating) involves training multiple models on different subsets of the dataset, such as random CNN variations, and combining their predictions to reduce variance. Boosting methods, such as AdaBoost, sequentially train models where each new model focuses on correcting the errors of the previous ones, thereby improving overall accuracy. Stacking, on the other hand, combines multiple base models and uses a meta-learner to determine the optimal way to integrate their outputs. Ensemble methods offer improved robustness, reduced variance, and significantly higher accuracy compared to individual models. However, they come with increased computational cost and implementation complexity. Despite these challenges, ensemble models frequently achieve state-of-the-art performance, often exceeding 95% accuracy in DR classification tasks.

9. Comparative Analysis of Algorithms

The comparative analysis of deep learning algorithms for diabetic retinopathy (DR) detection provides a structured understanding of their relative strengths and limitations across key performance dimensions. Different algorithmic approaches exhibit varying levels of accuracy, computational complexity, data dependency, and interpretability, making their suitability context-dependent. Basic CNN models offer simplicity but are limited in performance, whereas advanced architectures such as ResNet and DenseNet achieve higher accuracy with moderate complexity. Transfer learning reduces data requirements, while ensemble and hybrid models maximize predictive performance at the cost of increased computational overhead. Attention-based models uniquely enhance interpretability, which is critical in medical diagnosis.

Table 1: Comparative analysis of deep learning algorithms for diabetic retinopathy detection based on accuracy, computational complexity, data requirements, and interpretability.

Algorithm Type	Accuracy	Complexity	Data Requirement	Interpretability
Basic CNN	Medium	Low	High	Low
Deep CNN (ResNet, DenseNet)	High	Medium	High	Medium
Transfer Learning	High	Medium	Low	Medium
Ensemble Models	Very High	High	High	Low
Attention-Based Models	High	High	Medium	High
Hybrid Models	Very High	Very High	High	Medium

10. Conclusion

This systematic review demonstrates that deep learning algorithms have substantially transformed the field of automated diabetic retinopathy (DR) detection and classification, offering significant improvements over traditional diagnostic approaches. Among the various algorithmic categories analyzed, Convolutional Neural Networks (CNNs) serve as the foundational framework, enabling effective feature extraction from retinal fundus images. Building upon this foundation, transfer learning techniques have proven particularly valuable in scenarios with limited annotated data, as they leverage pre-trained models to achieve high performance with reduced training effort. Similarly, ensemble learning methods have consistently delivered superior accuracy and robustness by combining the strengths of multiple models, making them highly suitable for clinical-grade applications.

In addition, attention-based architectures have emerged as a promising direction, enhancing model interpretability by focusing on clinically relevant regions such as microaneurysms and hemorrhages. This is especially important in medical settings, where explainability plays a critical role in gaining clinician trust. Hybrid models that integrate multiple algorithmic strategies further push the boundaries of performance, although they often come with increased complexity and computational demands.

Despite these advancements, no single algorithm can be considered universally optimal for all scenarios. The effectiveness of a model depends on several factors, including dataset size, image quality, class distribution, and available computational resources. Moreover, challenges such as overfitting, class imbalance, and lack of interpretability still persist. Therefore, future research should focus on developing lightweight, efficient, and explainable models that can be deployed in real-world clinical environments. Emphasis should also be placed on scalability and generalization to ensure reliable performance across diverse populations and imaging conditions.

References

1. S. Zhu, C. Xiong, and Q. Zhong, “Diabetic Retinopathy Classification with Deep Learning via Fundus Images: A Short Survey,” IEEE Access, 2024. (ResearchGate)
2. M. Bappi et al., “Deep learning-based diabetic retinopathy recognition with context gating,” Elsevier / ICIP, 2025. (ScienceDirect)
3. S. Akhtar et al., “A deep learning-based model for diabetic retinopathy grading,” Scientific Reports, 2025. (Nature)
4. M. Akram et al., “Uncertainty-aware diabetic retinopathy detection using DenseNet-based transfer learning,” Scientific Reports, 2025. (Nature)
5. S. Tyagi et al., “Enhanced diabetic retinopathy detection using ViT, CLIP and hybrid deep learning models,” 2026. (ScienceDirect)
6. S. Prathibha et al., “Advancing diabetic retinopathy diagnosis with fundus imaging: A survey,” ScienceDirect, 2024. (ScienceDirect)
7. “Deep Learning-Based Classification of Diabetic Retinopathy,” ACM Conference, 2024. (ACM Digital Library)
8. N. Sikder et al., “Severity classification of diabetic retinopathy using ensemble learning,” Symmetry, 2023. (PMC)
9. A. M. Moustari et al., “Two-stage deep learning classification for diabetic retinopathy using Grad-CAM,” 2024. (PMC)

10. H. Shakibania et al., "Dual-branch deep learning network for DR detection and grading," Biomedical Signal Processing, 2024. (PMC)
11. L. Dai et al., "Deep learning system for predicting DR progression," Nature Medicine, 2024. (PMC)
12. Y. Jin et al., "Deep learning-based automated quality assessment of OCTA images for DR," 2024. (PMC)
13. B. Ebrahimi et al., "Optimizing OCTA layer fusion for DR classification," Biomedical Optics Express, 2023. (PMC)
14. A. Butt et al., "Hybrid deep learning features for DR detection," Diagnostics, 2023. (PMC)
15. S. Mondal et al., "Ensemble deep learning technique for DR detection and classification," Diagnostics, 2023. (PMC)
16. A. Abbood et al., "Hybrid retinal image enhancement and deep learning model for DR diagnosis," IEEE Access, 2023. (PMC)
17. A. Bilal et al., "CNN-based feature selection and classification for diabetic retinopathy," 2023. (PMC)
18. N. Nasir et al., "Deep DR: CNN-based diabetic retinopathy detection," IEEE Conference, 2023. (PMC)
19. H. Kaushik et al., "Stacked generalization of deep models for DR diagnosis," IEEE Access, 2023. (PMC)
20. I. Al-Kamachy et al., "Classification of diabetic retinopathy using pre-trained deep learning models," arXiv, 2024. (arXiv)
21. P. Zhang et al., "Deep learning-based detection of referable DR using ultra-widefield imaging," MICCAI, 2024. (arXiv)
22. S. Malik et al., "Deep learning approaches for retinal image grading and anomaly detection: A systematic review," 2024. (arXiv)
23. "Deep Learning in Automatic Diabetic Retinopathy Detection and Grading Systems: A Survey," 2024. (ResearchGate)
24. "Deep Learning for Diabetic Retinopathy Detection: Multimodal Data Fusion Approaches," 2025. (ResearchGate)
25. "Detecting Diabetic Retinopathy using Deep Learning," International Journal of Intelligent Systems, 2024. (IJISE)