I.J. Information Technology and Computer Science, **2025, 6, 160-175**
Published Online on December 8, 2025 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijitcs.2025.06.09

A Swin-transformer Integrated with Radial Optimization Model for Accurate Diabetic Retinopathy Detection and Classification

Vijayalaxmi Gopu*
Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology (Deemed to be University), Chennai. Tamil Nadu 600119, India
E-mail: vijayalaxxmigopu@gmail.com
ORCID iD: https://orcid.org/0009-0001-9395-7498
*Corresponding author

M. Selvi
Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology (Deemed to be University), Chennai. Tamil Nadu 600119, India
E-mail: selvi.cse@sathyabama.ac.in
ORCID iD: https://orcid.org/0000-0003-1175-5655

Received: 28 July 2025; Revised: 26 September 2025; Accepted: 23 October 2025; Published: 08 December 2025

Abstract: According to anticipation, Diabetic Retinopathy (DR) is one among the most potential causes of visual disability and even blindness across the globe. Prompt diagnosis and treatment will stall the development towards irreversible damage. Hence, detection and staging of DR must play a huge role in early medical intervention and treatment planning. There exist tremendous challenges in accuracy, ability to capture retinal features, overfitting from poorly represented features, and inefficiencies in optimisation of parameters. These are likely to provide a very challenging situation for the clinical application of these methods, especially when dealing with huge heterogeneous datasets. The paper is discussing a new hybrid framework for detection and classification of DR that integrates the Swin-Transformer Integrated radial Residual Network (Swin-RadialNet) with a Butterfly-Mayfly Radial Optimizer (BMRO). This framework is intended to address those challenges. Hierarchical extraction of features from a model trained with radial residual connections atop the Swin Transformer architecture is carried out by Swin-RadialNet to guarantee the best learning of the complex k structures in retina on-the-fly through BMRO, which would act as an optimizer hybridizing to estimate radial spread parameters, thereby supporting acceleration of models' performance and converging rates. The core novelties of the proposed method shall lie in merging advanced transformer-based feature extraction with nature-inspired hybrid optimization to tackle efficiently critical issues regarding feature abstraction, parameter fine-tuning, and classification reliability. The method will be evaluated on several benchmark datasets such as Kaggle's Diabetic Retinopathy and APTOS 2019, and will expect to show state-of-the-art results for all DR stages in terms of accuracy, precision, recall, and F1-score.

Index Terms: Diabetic Retinopathy (DR), Deep Learning, Transformer Model, Classification, Optimization, and Disease Detection.

1. Introduction

Diabetic Retinopathy (DR) is one of the severe health complications that is directly associated with Diabetes Mellitus, a chronic metabolic disorder caused by persistent hyperglycemia, often from different micro-vascular and macro-vascular abnormalities and the disruption of normal glucose metabolism [1, 2]. DR is not just about impaired glucose regulation; it is a disorder characterized by a cluster of vascular complications that are likely to cause long-term and disabling health consequences. Among these, DR is one of the most serious and incapacitating conditions since the retina affected is the thin layer of tissue at the back of the eye responsible for vision. Over time, the high blood sugar levels accompanying DM damage the small blood vessels in the retina, causing them to swell and leak, and to grow abnormal blood vessels, hence leading to severe visual impairment or even complete blindness if left untreated [3]. This is indeed a most frightening complication of DM because of the high frequency of occurrence among patients with diabetes, specifically those who have been under prolonged management of the disease. Research and statistical data indicate that some 80% of people with a 15 to 20 year history of diabetes eventually develop some form of DR. Such results underline the long-term effects of diabetes on health in general and further emphasize DR as one of the major causes of avoidable blindness among adults of working age especially in the developed world, where health access may attenuate but not altogether remove risks for long-term diabetics.

The global burden of diabetes continues to rise at an alarming rate. Currently, it is estimated that more than 171 million people live with diabetes, a number likely to more than double over the next few decades. The World Health Organization predicts that, by the year 2030, there will be 366 million cases of diabetes worldwide. It only adds to all the stronger reasons for taking preventive measures, ensuring early detection and implementing strategies in managing this condition not just on diabetes but on its crippling complications, especially DR [4]. There are calls to provide comprehensive healthcare for the rise of diabetes complications that include lifestyles modification and screening for DR coupled with advancement of science, reducing impacts the illness produces on vision loss as well as on health at large. Symptoms can be diverse and will progressively degenerate if not treated, finally resulting in the loss of vision. These symptoms arise due to the basic pathological changes seen in the retinal blood vessels from the chronic effects of DM.

The associated hyperglycemia seen in DM leads to metabolic changes that undermine the integrity of the retinal blood vessels [5, 6]. It considerably affects the retina's ability to process the visual signals in an accurate way hence one has cloudy, blurred or impaired vision. This occurs usually simultaneously in both eyes in most cases, as the systemic factors underlying diabetes eventually impact on all the blood vessels supplying the retina. If left untreated, the effects of such impacts would progress over time, further aggravating them, to result eventually in diabetic maculopathy, a disease of DR in a more advanced stage, involving the macula the commonest cause of vision loss due to the disease. More importantly, DR is a painful reminder of how systemic a disease diabetes can be. Good control of blood glucose levels, blood pressure, and lipid profiles protects not only the eyes but other important organs from the ravages of this chronic disease.

The rising prevalence and severity of DR have brought into the spotlight the huge need for an automated, intelligent detection system that could carry out primary analysis and identification of early signs of the disease. This will be life-transforming in the early diagnosis and management of DR [7], and it is a tool to ensure that people with prolonged diabetes can preserve their vision and lead a much more normal life. The increasing development of medical imaging has provided large datasets of fundus images high-resolution images of the back of the eye, encompassing the retina from various sources. These retinal images help describe changes in the retina that are subtleties hinting at the presence of DR in its very early stages. However, manual analysis of the images is extremely laborious and time-consuming; besides, expertise variability is involved. Hence, a strong, automated approach using state-of-the-art technology is essential for the efficient and accurate detection of DR. Traditional machine learning (ML) [8, 9] techniques, while helpful in many domains, are challenged by the requirements of modern medical image analysis.

Most of these techniques are not designed to handle the complexity, size, and high dimensionality of image data that inherently require advanced feature extraction and intricate pattern recognition. Traditional ML algorithms usually depend on handcrafted features, requiring much domain knowledge and tend to be blind to the underlying complexities of retinal abnormalities associated with DR. Moreover, such approaches are not scalable and adaptively flexible, making them computationally expensive and thus unworkable for most real-world scenarios, where datasets are diverse and disease characteristics change over time. Consequent to this, their performance within the context of DR detection has been limited and has necessitated more sophisticated solutions. The typical DR detection system using machine learning and deep learning models is shown in Fig 1.

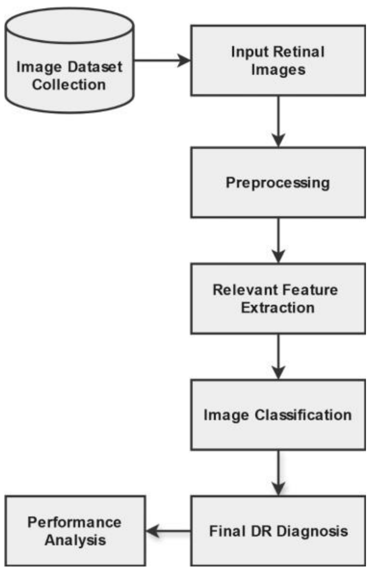

Fig.1. Typical DR detection model

In the field of Machine Learning, Deep Learning has been an emancipator with unparalleled abilities to analyze complex data in an automated manner. More than the shallow ML models [10], DL methods could directly process high-dimensional data with great reduction in manual feature engineering. The hierarchical representations of the data are learnt using a number of layers of artificial neural networks; hence, DL is highly effective at tasks such as DR detection tasks where subtle variations in retinal images signal the beginning of a disease. Equally, DL systems rise to meet vast data volumes without decreasing performance and further manage consistent performance, dependability, and the ability to arise with each piece of data derived from a certain number of sources. Beyond its scalability, DL models [11] perform very high in domain awareness and decision-making capability. They may also learn from previous knowledge in medical imaging through transfer learning and generalize well to new datasets and imaging modalities, making them more robust under diverse clinical environments. Also, DL models are developed keeping in mind computational efficiency for realizing real-time analysis and diagnosis in applications involving telemedicine and resource-constrained settings. The designed systems are not only meant to guarantee better diagnostic accuracy but also to prevent loss of vision through earlier intervention, hence changing how diabetes-related complications are globally managed.

There is still an urgent need for the proposed work to be motivated by the rapidly growing challenge posed by DR, a progressive disease characterized by a very intricate web of pathological traits that cause a train of subtle differences in retinal alterations, most notably micro-aneurysms, hemorrhages, and exudates that are difficult for the detection by even sophisticated analytical prowess. These variations are not only multifaceted but also deeply interconnected, making the task of early sign detection of DR especially challenging by conventional methodologies [12]. Traditional machine learning and even some deep learning models have shown some promise, but they usually fail to properly understand the complex relationships and dependencies inherent in retinal image data. This stems from their inability to process and integrate local and global information in an effective manner, which proves very important in the analysis of complex variations of retinal abnormalities. This capability gap motivated our research on transformer-based models that handle such complex relationships by leveraging their unique architecture for capturing long-range dependencies and attention mechanisms. Transformer-based models, originally developed for natural language processing, have recently shown outstanding potential in computer vision tasks by being able to analyze data on a holistic basis. Most of the existing CNNs work in a local way, focusing on local features; that is where transformers bring in self-attention mechanisms allowing the model to build contextual relationships all over the image [13]. The capability of modeling both short-term variations and localized patterns, such as micro-aneurysms, and long-term dependencies, structural patterns involving multiple regions of the retina, is important for robust and reliable detection. By adopting the transformer-based approach, the proposed work seeks to improve the capability of the model in distinguishing such complex variations accurately for the improvement of diagnostic accuracy and efficiency.

The combination of the Swin-RadialNet and Butterfly-Mayfly Radial Optimizer (BMRO) produces a very strong architecture that directly addresses some of the fundamental challenges surrounding disease detection and classification in DR. Swin-RadialNet takes advantage of the hierarchical attention capabilities of the Swin Transformer to globally and locally capture retinal features, and the radial residual connections allow for deeper information flow and structural consistency within the network and hence the accurate representations of complex pathologies, such as microaneurysms, hemorrhages, and exudates. This strategy can efficiently address the common problems of weak feature abstraction and overfitting that conventional CNN based methods face. On this basis, BMRO has injected an effective hybrid optimization strategy in which exploration ability of the butterfly algorithm has been related along with exploitation capabilities of mayfly strategy. It is also referred to finding better radial spread parameters for Swin-RadialNet, adding improvements in the dynamics of learning and stabilization of convergence in a more precise manner. These then effectively alleviate the ineffectiveness of parameter optimization with the improvement of the classification performance.Taken together, Swin-RadialNet and BMRO provide complementary power towards enhancing feature extraction, which enhances the robustness of the whole system against a varied DI diabetic retinopathy stage, ensuring high performance even for large heterogeneous clinical datasets.

Another strong motivation is that such an approach can provide a transformative leap toward bridging the gap between technical capability and clinical applicability. Current state-of-the-art DR detection systems [14, 15] face challenges in issues of scalability, generalization, and real-world robustness. Furthermore, improving the capability of capturing long-range dependencies is in line with the objective of making the detection system more robust to real-world complexities, where the presence of noise, artifacts, or overlapping features in fundus images may obscure critical diagnostic information. This paper is motivated by the desire to try and overcome some of the current methodological constraints, proposing a transformer-based model that would greatly improve the nuances and complexities observed in DR pathology. The present approach seeks to set a new benchmark in the detection of DR by offering an approach that could be robust, scalable, and clinically relevant through advanced attention mechanisms and the possibility of modeling long-range dependencies. The proposed work not only fills a critical gap in the field of medical image analysis but also holds the promise of significantly improving early detection and intervention for millions at risk of vision loss due to DR. The major contributions of the proposed work are given below:

- To analyze the model's ability to deal with complex variations of the complex relationships captured in the architecture for robust DR detection using Swin-Transformer Integrated Radial Residual Network (Swin-RadialNet). It offers a novel approach enhancing ta and long-range dependencies.

- To propose a new Butterfly-Mayfly Radial Optimizer (BMRO) for effective estimation of the radial spread parameters to ensure further model optimization, robustness, and adaptability in classification.
- To validate the effectiveness of the proposed hybrid framework, thorough performance evaluation is performed using widely recognized datasets, and the results show very important improvements in precision, recall, and overall accuracy compared with state-of-the-art methods.

The paper is systematically organized to provide an in-depth exploration of the proposed methodology for Diabetic Retinopathy detection and classification, starting from the background and moving progressively through the methodology, performance evaluation, and future outlook. Section 2 explains the literature review in detail and gives a critical review of the state-of-the-art machine learning (ML) and deep learning (DL) approaches used for the detection and classification of DR. Section 3 presents the new methodology proposed in this paper. In it, there is a novel transformer-based model designed for robust DR detection that is described in detail. Section 4 deals with the metrics and results chosen after extensive testing. To sum up, Section 5 gathers the main points and contributions from the paper and discusses broader implications of the proposed work, outlining potential future research directions to further improve the results.

2. Related Works

Diabetic Retinopathy (DR) has become one of the serious global health challenges in recent decades, demanding the construction of advanced diagnostics systems to find and classify the disease accurately. During the last years, several machine-learning and deep-learning-based approaches have been reported for the hard detection of DR from retinal fundus images. Traditional ML methods [16] have proven capable of performing basic classification tasks by depending on handcrafted features extracted from retinal images. While those approaches pioneered the field of automated DR detection, their reliance on hand-crafted features and limited scalability did not allow them to capture complex variations and patterns inherent to high-dimensional image data. Therefore, these methods frequently do not perform robustly and reliably on different datasets and clinical application scenarios. Most works highlighted the success of CNN-based models in detecting microaneurysms, hemorrhages, and exudates, vital indicators of DR, with impressive and high accuracy. Traditional CNNs, nonetheless, have not been optimized to model long-range dependencies and a global context to frame complex retinal structures. These techniques have greatly enhanced the accuracy and robustness of DR detection, thus allowing more refined architectural integration. However, none of these systems were designed to generalize over a different array of varying sources or imaging modalities. More fundamentally, the high computational complexity and inscrutable nature of some DL models pose additional hurdles to their real-world applications. Transformer-based models have recently emerged to break through such limitations by presenting superior capabilities in modeling not only local but also global dependencies. This section sums up state-of-the-art machine learning and deep learning techniques for DR detection. It highlights the achievements and limitations of these techniques while giving reason for the presentation of the transformer-based methodology in this paper.

Akram, et al [17] presented an interesting approach toward improving diabetic retinopathy detection by means of combining transfer learning and Bayesian approximation techniques. This research really is an example of how transfer learning could simplify the complexity in analysis of retinal images using DenseNet-121 as a base convolution neural network. The authors also enrich their methodology with Bayesian extensions, bringing a whole new dimension to the uncertainty quantification of DR detection. Applying Bayesian approximation techniques, including Monte Carlo Dropout, Mean Field Variational Inference, and Deterministic Inference, provides the possibility to represent posterior predictive distributions something that constitutes a strong framework for the evaluation of prediction confidence. Kaushik, et al [18] propose a deep learning framework based on a residual network to improve the detection of diabetic retinopathy, which has been one of the biggest demands in automated retinal disease diagnosis. The newly proposed framework in this manuscript is promising and has inherited strengths in handling complicated features and variations by residual networks characterizing retinal images in establishing a robust performance. This paper reports an 83% success rate, which may be an indication of the potential this method holds in the reliable detection of DR. Moreover, this value of precision in identifying the absence of DR shows its capacity to reduce false positives a factor of great importance in clinical application of results. Residual networks are also well justified, as they usually perform well on issues such as vanishing gradients and allow for much deeper architectures than traditional ones. The framework itself appears to tradeoff between computational efficiency and good performance a very potent advantage in real-world deployment scenarios. In other words, they could learn more complicated patterns in medical images. The accuracy and precision reported are, however, promising but still require more metrics such as recall, F1-score, and specificity to provide a holistic view of the model's performance. A very important contribution to DR detection, therefore, is the framework showing the effectiveness of residual networks in meeting challenges presented by automated disease diagnosis.

Govindaraj, et al [19] introduce a new hybrid model by combining UNet++ and Capsule Network (CapsNet) architectures in order to improve the accuracy in diagnosing glaucoma. Combining the best features of both architectures presents the solutions to some important medical image analyses, such as accurate segmentation and robust feature recognition. UNet++ is utilized effectively for semantic segmentation, which allows the precise delineation of optic discs and cups very important in the evaluation of structural changes related to glaucoma. The inclusion of such advanced preprocessing techniques as Histogram Equalization and Contrast Limited Adaptive Histogram Equalization (CLAHE) sets the authors on the optimization path for retinal image quality and, consequently, for further analysis. These methods ensure that the hybrid framework is working on improved images for better capability in detecting subtle changes indicative of glaucoma. While the hybrid approach looks promising, the authors could further the work by providing a detailed comparison of their model with already existing models, including CNN-based and transformer-based architectures, in order to show the relative advantages. Moreover, it would be quite valuable to discuss computational efficiency for the combination of UNet++ and CapsNet, particularly in view of potential clinical implementation. Neri, et al [20] propose a deep-learning-based software framework for the detection and classification of microaneurysms (MAs) as hyporeflective or hyperreflective in structural optical coherence tomography (OCT) images of patients with non-proliferative diabetic retinopathy (NPDR).

This study addresses a critical diagnostic need by using state-of-the-art deep learning models, namely YOLO (You Only Look Once) and DETR (DEtection TRansformer), for robust MA detection and classification. Inclusion of manual segmentation by five masked readers, with consistent labeling ensured by an expert grader, underlines the careful way in which high-quality annotated data are generated for training the models. The evaluation framework of this study is based on the area under the receiver operating characteristic (ROC) curves for a comprehensive assessment of detection and classification performance. The authors chose YOLO and DETR, both of which are very important due to their known efficiencies in object detection and handling complex spatial relationships, respectively. This positions the framework to tackle challenges in MA detection in OCT images that are often subtle and variable. Ikram and Imran [21] shed light on the critical need for an automated, exact, and tailor-made machine learning way in order to enable early detection and treatment of diabetic retinopathy. Traditional deep learning models have been applied so much to the diagnosis of DR. Meanwhile, the authors said that Vision Transformers (ViTs) have recently emerged, showing great performance in image analysis by capturing long-range dependencies within images. With that potential in mind, here is a study proposing a hybrid model, ResViT FusionNet, to improve the accuracy of DR detection.

Aiming to let the framework in hand possess strengths from both ResNets and Vision Transformers, so that it may bring together the former's ability in extracting fine-grained local features with the latter's power in modeling global contextual relationships, an architecture finding application in tasks like detection of DR with localized lesions and broader structural changes in retinal images. The authors address, therefore, the diagnosis of DR in the future by exploiting the complementary benefits of ResNets and Vision Transformers for overcoming obstacles in the early and accurate detection of diseases.

Bhimavarapu [22] present a novel approach to the rising challenge of DR detection through an automated model of smartphone-based retinal imaging. Most importantly, this is due to its emphasis on affordability, portability, and accessibility in meeting the needs of underserved populations, enabling non-specialists to carry out retinal imaging in non-clinical settings. The proposed model, using lightweight deep learning architectures, tries to reduce the blindness and alleviate the complications of DR with early and effective detection. The focus of smartphone-based imaging in this book has been on closing the gap between advanced medical technology and practical usability in resource-constrained environments. This makes the lightweight architecture computationally efficient and hence viable for devices with a small amount of processing power to analyze images precisely. That fully meets the objective to bring DR detection tools out of the classic medical environment and to make them available for more users. This study could be more substantial with comparative results from already-existing DR-detection models compatible with smartphones, which would bring out the benefits the proposed framework has in performance, cost, and usability. Overall, the authors give a transformational solution that can revolutionize DR detection by combining technological innovation with practical accessibility.

Luong, et al [23] discuss the application of machine learning models in the prediction of low-risk diabetic retinopathy using the International Clinical Diabetic Retinopathy severity score. Five MLMs were trained with data from 1,050 subjects employing 10-fold cross-validation in order to maximize their performance, as quantified by the area under the receiver operating characteristic curve (AUC). Among the most important contributions that this study brings to the scientific literature is a new Retinopathy Risk Score, defined as the product of HbA1c levels closest to screening and the duration of diabetes in years with DM. This new score is part of the predictive model and hence adds an extra layer of interpretability and clinical relevance. Application of RRS shows efforts in relating its predictions to established physiological risk factors, hence making the model even more useful for clinicians. While the study shows a strong methodological approach, it would be of value to compare more deeply the performance of the five MLMs and discuss the relative strengths of these. Manohar, et al [24] show a new CNN architecture for diabetic retinopathy detection from retinal images and place their work in a potentially important line of development for an automated diagnosis of DR.

In this study, one will have clear benchmarking over two well-founded models, namely VGG 19 and ResNet 101, by the comparison of this proposed CNN over image classification tasks. The comparison with such firmly established models adds credibility to the effectiveness of the proposed CNN but provides a reference point for readers to judge the improvement in performance. It would, however, be very nice if the authors could give more details on what are the architectural differences of their CNN from the two benchmark models in order to further understand why their model gives superior results. It would also further enhance the discussion relating to actual deployment, where high accuracy has to be weighed against resource utilization feasibility, concerning computational efficiency regarding training time, model size, inference speed, etc. It would also strengthen the practical applicability of the proposed method by elaborating how the model deals with variations of retinal images, due to differences in imaging conditions or patient demographics.

The literature review on the detection of diabetic retinopathy shows outstanding progress in the application of machine learning and deep learning models for automated diagnosis. This is coupled with many research gaps that could be filled to improve the effectiveness, accessibility, and scalability of the systems. Another important gap in the literature is that many of the deep learning models used in DR detection lack interpretability and transparency. Much of the research does not allow enough insight into the decision-making process of the model and the features that it relies on for classification. Bridging this gap would probably bring more trust in models and may speed up their adoption by healthcare professionals. Though a few studies have proposed models to classify DR into different stages, most systems still do not have the capability to handle longitudinal data in predicting how DR may progress in the future. This, in particular, is of importance for the early intervention and personalized treatment of diabetic patients, as the progression of DR can be highly heterogeneous between different individuals. Besides, although some works have tried to explore hybrid models by combining CNNs with other techniques such as Vision Transformers, still there is not too much research on how to optimally integrate the different deep learning architectures in order to improve both accuracy and computational efficiency.

Finally, while much of the research focuses on developing models for the detection of DR in specific populations or clinical settings, there is a lack of demonstration in literature in relation to real-world deployment in diverse scenarios. This calls for addressing challenges in data heterogeneity, scalability, and the ability of the model to function with mobile or portable devices so as to offer low-cost, high-quality DR screening in underserved or remote areas. One of the big barriers to the wide implementation of DR detection and prevention is inaccessibility due to unaffordable solutions. Overall, the performance in the task of DR detection has greatly improved using machine learning and deep learning methods, but very critical gaps remain in this research. Further work in overcoming these challenges should therefore go on toward more robust, scalable, and clinically useful systems in the detection and management of DR.

3. Proposed Methodology

The main purpose of this research work is to develop and implement the latest framework for the effective detection and classification of diabetic retinopathy using novel deep learning models and hybrid optimization techniques. Diabetic retinopathy is a progressive eye disease caused by longstanding diabetes that might result in impairment of vision and serious blindness at severe cases. Early detection is vital to stop its progression, but it is a challenging task due to the complexity of retinal images and the subtlety of early disease manifestations. The proposed research tries to address this issue by using state-of-the-art techniques, including the Swin-Transformer Integrated Radial Residual Network (Swin-RadialNet), together with the Butterfly-Mayfly Radial Optimizer (BMRO) for optimized parameter estimation. The first method proposed, Swin-RadialNet combines the strength of the Swin Transformer a state-of-the-art model with the capability of capturing long-range dependencies in an image with the benefits of a Residual Network in enhancing learning. It contains a window-based self-attention mechanism, which enables Swin-Transformer to effectively capture the spatial dependencies and pay attention to both local and global features of retinal images simultaneously.

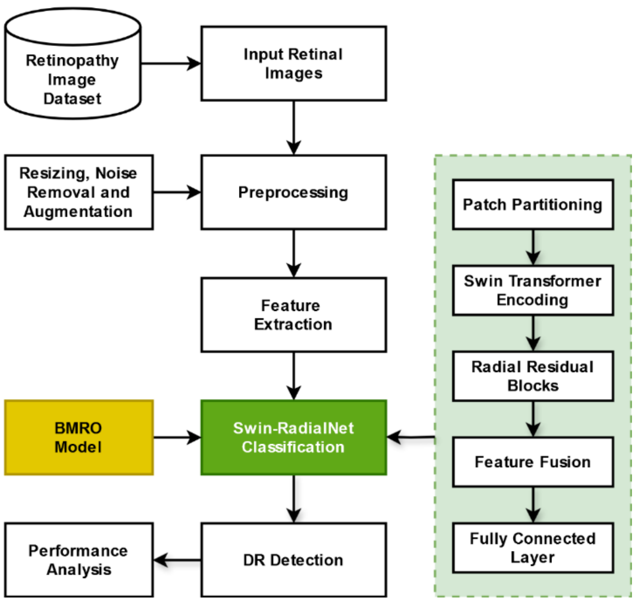

Fig.2. Flow of the proposed DR detection system

As shown in Fig 2, this would enhance feature extraction for accurate segmentation in DR detection, especially for complex and noisy image data. Combinations of both architectures ensure that the model could detect DR efficiently at different stages with varying image qualities, hence making it a very powerful tool for automated screening. The second proposed method is called BMRO, which is a hybrid optimization algorithm for the efficient estimation of radial spread parameters in DR detection models. BMRO balances BOA with regard to exploration and MOA with regard to exploitation capabilities such that it returns a better solution compared to the traditional optimization techniques. The detection process of DR by the proposed methods starts with the collection of retinal images followed by pre-processing based on standard image enhancement techniques, including histogram equalization and contrast enhancement. These features are fine-tuned by the Radial Residual Network to preserve vital information for the precise detection of diabetic retinopathy indicative patterns. Its output will be used, further, in grading the severity of DR from no DR to severe retinopathy. In parallel, the BMRO optimizes parameters of radial spread for the network in such a manner that decision boundaries of the model are aligned optimally to the features that are extracted from the images. After optimization, the DR detection system now becomes able to predict the probability of DR in an input retinal image with high accuracy and robustness in classification. The Swin-RadialNet with BMRO has come up with a strong combination to address the most significant problems concerning DR, in detection and classification. Swin-RadialNet augments this effect because it emphasizes that hierarchical attention capabilities of the Swin Transformer are leveraged to encompass global and localized retinal characteristics while radial residual connections are thought to invoke deeper anatomical flows for structural consistency-rich representations of even the most complex pathological patterns, such as microaneurysms, hemorrhages, and exudates. All of them point to a common shortcoming in the majority of conventional CNN-based methods - low generalization of features formed and overfitting. On the contrary, the BMRO presents the strongest hybrid optimization mechanism through exploring the butterfly algorithm and exploiting the mayfly strategy. Effective fine-tuning was achieved for the parameters of radial expansion within the Swin-RadialNet while optimizing learning dynamics, thus ensuring more stable and accurate convergence eliminating inefficiencies in parameter optimization, improving accuracy in classification. Thus, the newly obtained framework Swin-RadialNet and BMRO works orchestrally for strengthening feature extraction and generalizing well across DR stages, thus ensuring good working at large heterogeneous clinical datasets.

In these circumstances, the proposed method addresses the vital issues in DR detection within large heterogeneous datasets through the synergistic amalgamation of Swin-RadialNet and the Butterfly-Mayfly Radial Optimizer (BMRO) that help enhance feature learning capacity and optimization of the system. One of the foremost challenges in real-world DR classification encompasses significant variability in image quality, illumination, retinal structures, and disease severity across different datasets. Conventional approaches hardly generalize in these conditions because they are more concerned about weakly extracting robust features to discrimination and are prone to overfitting. The Swin-RadialNet, built on the Swin Transformer backbone, addresses this by adopting window-based self-attention mechanisms via which the network learns about multi-scale representation of retinal features with computational efficiency in mind. The addition of radial residual connections bolsters the network's ability to maintain spatial and contextual information, allowing for the unambiguous detection of subtle pathological variations across varying image sample types. Due to this hierarchical-residual nature of the design, the model gains robustness against variations in the input data and, therefore, becomes very general in heterogeneous data. In parallel, the BMRO part contributes to the overhead problems of instabilities and sub-optimal convergences owing to large-scale training for deep models. It combines the exploration strategies of the butterfly swarm behavior with the exploitation dynamics of the mayfly algorithm so that the radial spread parameters and the internal weights of Swin-RadialNet are adjusted to perfection for every training instance, despite a high degree of variance presented by the dataset.

The definitive combination of Swin-RadialNet and Butterfly-Mayfly Radial Optimizer (BMRO) stands mechanistically better than simpler options because it cleverly integrates advanced feature representation with intelligent parameter optimization, thus tackling several bottlenecks concurrently in the detection of DR. Conventional CNN models or elementary optimization methods would be fairly adequate for small or homogeneous datasets but fail to generalize well in the case of complex real-world retinal images because of poor spatial awareness and static learning parameters. Swin-RadialNet uses the Swin Transformer architecture, bringing hierarchical self-attention and localized context awareness allowing DR-related features such as microaneurysms and neovascularization to be captured more subtly. Radial residual connections further propose deep gradient flow and a pathway for effective learning of structural abnormalities that are characteristically subtle and overlapped. Complementary to this, with BMRO, the intelligent dynamic tuning of model parameters is done using a hybrid metaheuristics approach that captures the balance of global exploration and local exploitation such that, better than the basic optimizers such as SGD or Adam, which usually get stuck in local optima, this achieves intelligent fusion.

3.1. Swin-transformer Integrated Radial Residual Network for DR Detection and Classification

Swin-RadialNet is the novel hybrid deep learning framework exclusively developed for the detection and classification of diabetic retinopathy from retinal fundus images. The technique is a synergy in hierarchical feature extraction: Swin transformer combined with the effectiveness and flexibility of a radial residual network. The main contribution of the present work is that it manages to capture both local and global features of the retinal images while preserving important radial patterns characteristic of diabetic retinopathy, which traditional convolutional networks with their limited receptive fields usually cannot model well: long-range dependencies and hierarchical structures in complex images. This model exploits the Swin Transformer in capturing long-range dependencies and hierarchical feature representation; on the other hand, the radial residual network is designed to emphasize center to boundary transitions in the retinal structure important for the identification of micro-aneurysms, hemorrhages, and other DR markers. The novelty of Swin-RadialNet is introducing a transformer-based architecture with residual connections specifically designed to handle the radial patterns in retinal images. The architecture proposed in this paper integrates, therefore, the radial attention mechanisms into the residual connections in the convolutional layer, unlike traditional networks or a standalone transformer model, for an emphasis on the retinal structures. This allows preserving and highlighting those fine-grained details and subtle patterns important in the diagnosis of DR at early stages. These novelties have overcome some of the most important limitations of the existing techniques: neither are able to generalize well over diverse datasets nor computationally efficient when dealing with high-resolution medical images.

Patch-based processing in the Swin Transformer allows the model to focus on a region of interest in a hierarchical way so that both macro-level abnormalities, like large hemorrhages, and microlevel features, like microaneurysms, are well captured. On the other hand, the radial residual network ensures that the network would not lose the strong center-to-boundary relationship focus, which is crucial in detecting DR-related changes in the optic disc, fovea, and peripheral retina. Combining both, Swin-RadialNet achieves superior performance in the detection of subtle and early signs of DR by outperforming traditional CNNs, even those with standalone transformer models. A pre-processing layer first normalizes and augments the contrast in features for each retinal image fed into the network. The input is then processed patch-wise through the Swin Transformer module in order to extract hierarchical features and encode long-range dependencies. Features from this are then passed into a radial residual block, where radial attention is used to focus on center-to-boundary transitions in the retinal structure. The model further contains multi-scale feature fusion layers for fusing information coming from different network depths so as to increase the robustness facing variations of image quality.

The two most important benefits of Swin-RadialNet are as follows. First, its hybrid architecture ensures superior accuracy and robustness in the detection and classification of DR across diverse datasets. That is, combining transformer-based and residual architectures allows this model to efficiently handle high-dimensional medical image data while preserving critical details. First, the center MASK helps focus the model on the region of interest in the image; second, the radial attention mechanism enables directed feature extraction, in that it makes the model sensitive to subtle abnormalities that might otherwise be lost; and third, because multi-scale feature fusion generalizes well, real-world variability due to image quality and patient demography can also be handled in this approach. Lastly, the hierarchical processing capability of the Swin Transformer ensures that the model will be scalable and efficient for large scale screening programs and early DR detection efforts. Overcoming the drawbacks of existing techniques, the novel hybrid design attains superior accuracy, sensitivity, and robustness.

Implementation of Swin-RadialNet in the tasks of detecting and classifying diabetic retinopathy is highly structured, integrating the state-of-the-art strategies in deep learning with domain-specific enhancements to capture peculiarities of the retinal images. That begins by preprocessing retinal images in terms of normalization, re-sizing, and augmentations each in this complex pipeline sequentially improves contrast and mitigates noise. Histogram equalization is followed by adaptive gamma correction to make visible micro-level features. The next stage is integrating the radial residual blocks in the architecture. These blocks are designed to emphasize the radial patterns inherent in retinal images, such as transitions from the optic disc to the fovea and further to the peripheral regions. Each radial residual block consists of convolutional layers with radial attention mechanisms that prioritize center to boundary features, which are very important for identifying major DR related anomalies since these features mostly manifest as subtle changes in the retinal structure. The radial residual blocks are placed after the Swin Transformer module so that both global dependencies and radial patterns are properly captured and preserved.

Radial residual connections are more of a structuring improvement in deep neural networks where the residual pathways are nonlinear, permitting radial or circular propagation of feature information thereby allowing multi-directional flow of information across different network layers or blocks. In contrast to the standard residual connections that directly add the input of a block to its output for gradient flow preservation, the radial residual connections allow the distribution of intermediate feature maps along concentric or radially diverging layers, allowing for better fusion of hierarchical representations. In retinal image analysis, this is very much suited, as lesions like exudates or hemorrhages can develop at different scales and radial distances away from the optic disc; thus, the architecture captures spatially and structurally interdependencies better than standard residuals. Hence, these enhance the model to learn subtle pathological variations that are significant for staging the diabetic retinopathy. On the optimization side, the Butterfly-Mayfly Radial Optimizer (BMRO) has been adopted since its hybrid metaheuristic nature integrates the global search features of the Butterfly Optimization Algorithm (BOA) with local convergence speed by Mayfly Algorithm. BOA simulates the foraging behavior of butterflies forging in a scent-based manner, giving the search a large range; while the Mayfly mechanizes social interaction and mating dynamics for local exploitation, ensuring quick convergence. Such dual behavior puts BMRO in a place to overcome inherent weaknesses posed to conventional optimizers like SGD or Adam, which can prematurely halt convergence or be lured off the track by highly non-convex loss surfaces- the kind seen very often in deep DR models. The evidence for these claims rests in various empirical comparisons and performance tests showing that BMRO reaches greater stability across epochs with fast convergence and low classification errors, especially while tuning certain sensitive parameters including learning rates, weight decay factors, or spread coefficients in Swin-RadialNet. Therefore, radial residual connections combined with BMRO have a sound and tested technical advantage for DR detection that is accurate, scalable, and interpretable.

Another critical constituent of the realization is the multi-scale feature fusion strategy in Swin-RadialNet, which fuses features at a multitude of depths in the network so that the model retains high-level semantic information along with low-level spatial details. The multi-scale fusion is achieved using up sampling and down sampling operators followed by a concatenation operator, due to which the merging of information associated with different scales becomes possible in an efficient way. This will be particularly helpful for DR detection, as this will enable the model to learn abnormalities that may appear at different resolutions for example, large hemorrhages at a macro level and micro-aneurysms at a micro level. The design of the classification stage in Swin-RadialNet precisely outputs interpretable predictions. Passing the Swin Transformer and the radial residual blocks, features are flattened and fed into a series of fully connected layers. The latter is optimized with state-of-the-art loss functions, including focal loss, to handle class imbalance, which is common in most DR datasets. Those features are added to be especially useful in such cases when the quality of the input image is suboptimal or the DR stage is ambiguous; hence, the model remains reliable even under challenging conditions.

The focuses are also on scalability and real-world applicability in the deployment of Swin-RadialNet for the detection and classification of DR. The pipeline itself is optimized for batch processing in order to perform large-scale screening programs in an efficient manner. Also, its light weight ensures that the model can be applied to resource-constrained devices in, say, mobile based retinal imaging systems, so as to make them accessible in remote and underserved regions. From advanced preprocessing to radial attention mechanisms, all the way to multi-scale feature fusion, the pipeline is engineered from scratch for the unique challenges in retinal image analysis; it provides the new standard of early detection and management for diabetic retinopathy. In the proposed Swin-RadialNet model, the image preprocessing is performed initially after getting the input from the dataset, which is mathematically expressed in the following equation:

$$\mathcal{J}_N(m,n) = \frac{\mathcal{J}(m,n) - \mathfrak{m}(\mathcal{J})}{\phi(\mathcal{J})} \tag{1}$$

Where, $\mathcal{J}_N$ is the normalized image, m, n are the image coordinates, $\mathfrak{m}(.)$ and $\phi(.)$ are the mean and standard deviation of the image. As a consequence of this, the patch embedding is carried out using swin transformer based on the following model:

$$\mathcal{E}_0 = \mathrm{Conv}_{st}(\mathcal{J}_N) \tag{2}$$

Where, $\mathcal{E}_0$ represents the image initial patch embedding, and $\mathrm{Conv}_{st}(.)$ indicates the convolution stride operation estimated for patch extraction. Then, the multi-head self-attention module layer operation is performed according to the query, key and value parameters. Furthermore, the shifted window partitioning is performed in both horizontal and vertical directions based on the following equations:

$$\mathcal{S}_{shift} = \mathrm{Shift}(\mathcal{S}_{window}, \Delta m, \Delta n) \tag{3}$$

Where, $\mathcal{S}_{shift}$ indicates the partitioned and shift patches, Δm and Δn represents the shift amounts in both horizontal and vertical directions. The radial attention mechanism is applied based on the attention weight value using the following equation:

$$\mathbb{A}_{radial}(\mathfrak{d}, \theta) = \frac{e^{-\frac{(\gamma-\gamma_0)^2}{2\phi_\gamma^2}}}{\sqrt{2\pi\phi_\gamma^2}} \tag{4}$$

Where, $\mathbb{A}_{radial}(.)$ indicates the radial attention weight, γ_0 represents the reference radial distance, and ϕ_γ denotes the radial spread parameter. The radial residual block operation is performed based on the following model:

$$\mathfrak{N}_{rad}(x) = \phi(\omega_2 \times \mathrm{ReLU}(\omega_1 \times x + \beta_1) + \beta_2) \tag{5}$$

Where, $\mathfrak{N}_{rad}(x)$ indicates the output radial residual block, ω_1, ω_2 represents the weight matrices, β_1, β_2 indicates the bias value, and ϕ is the activation function. Furthermore, the multi-scale feature fusion is performed using the following equation:

$$\mathfrak{F}_{fused} = \sum_{i=1}^{n} \mathcal{P}(\mathfrak{F}_i) \oplus \delta(\mathfrak{F}_{n-i+1}) \tag{6}$$

Where, $\mathfrak{F}_{fused}$ indicates the fused features, $\mathcal{P}(\mathfrak{F}_i)$ is the upsampled features, $\delta(\mathfrak{F}_{n-i+1})$ represents the downsampled features, and $\oplus$ is the concatenation operation. The loss function is estimated for class imbalance using the following equation:

$$\mathcal{L}_{foc} = -\alpha_t (1 - \rho_t)^{\vartheta} \log(\rho_t) \tag{7}$$

Where, α_t is the weight value for class imbalance, ϑ indicates the focusing parameter, and ρ_t denotes the model prediction probability. In addition to that, the radial layer operation is performed using the following equation:

$$\mathfrak{F}_{\mathrm{radial}} = \sum_{i=1}^{N} \mathbb{w}_i \times \mathbb{A}_{\mathrm{rad}}(\mathfrak{r}_i, \theta_i) \tag{8}$$

Where, $\mathbb{w}_i$ indicates the feature weight, and $\mathfrak{r}_i, \theta_i$ represents the radial coordinates of feature. The dense classification layer operation is performed according to the following equation:

$$\mathcal{C}(x) = \mathrm{Softmax}(\mathbb{w}_c \times x + \beta_c) \tag{9}$$

Where, $\mathcal{C}(x)$ indicates the output classification probability, and $\mathbb{w}_c$ denotes the weight value of classification layer. Moreover, the Bayesian uncertainty is estimated based on the following equation:

$$\hat{\mathfrak{p}}(\mathfrak{z}|x) = \frac{1}{\mathbb{T}} \sum_{t=1}^{\mathbb{T}} \mathfrak{p}(\mathfrak{z}|x, \theta_t) \tag{10}$$

Where, $\hat{\mathfrak{p}}(\mathfrak{z}|x)$ represents the posterior predictive distribution, $\mathbb{T}$ is the monte-carlo samples, and θ_t denotes the model parameter.

3.2. Butterfly-mayfly Radial Optimizer (BMRO)

Butterfly-Mayfly Radial Optimizer (BMRO) is a new hybrid optimization algorithm for the effective estimation of radial spread parameters, particularly in complex problems such as diabetic retinopathy classification. The application of BMRO will be targeted to overcome the shortcomings of the single-objective optimization algorithms by harnessing the complementary strengths of both BOA and MOA. BOA has been reported to have good exploration capability because of the butterfly foraging behavior simulation, hence ensuring that search patterns are well spread over the solution space. However, MOA outperforms in the exploitation phase because imitation of the mayfly's social and mating behavior allows it to sharpen solutions in detail. Integration of the two algorithms guarantees that BMRO has a perfect trade-off between exploration and exploitation, hence resulting in optimal parameter estimation of the radial spread functions. This is so because BMRO introduces a new hybrid framework that dynamically integrates the strengths of BOA and MOA, leading to adaptability in the nonlinear and high-dimensional complexities belonging to DR Datasets. Moreover, BMRO is supposed to provide the estimation of a radial spread parameter in order to enhance feature extraction for the DR classification models, allowing them to more effectively capture slight differences among the retinal images representing different stages of the disease.

There is a very unique technique here, which actually supports effective DR classification by ensuring the estimated radial spread parameters are optimally aligned with the input feature space. The radial spread parameters in kernel-based and neural network architectures are of paramount importance, as they control the spread of influence of the data points on the decision-making process. Moreover, the BMRO step in the DR classification pipeline is what gives robustness against noise and variability in the retinal images a common challenge in most real-world datasets. This will come in handy in defining the stages of DR between early vs. advanced, where these kinds of differences could be minute. BMRO has several advantages besides its application in the classification of DR: its hybrid nature decreases the chance of falling into local optimum stagnation one common drawback inherent to many algorithms under the classic paradigm of optimization. Moreover, BMRO outperforms individual BOA and MOA in terms of convergence rates through synergy in the exploration-exploitation mechanisms of the solution space.

The proposed technique further shows scalability and generalizability properties, thus being easily adaptable to any other medical imaging task in need of the tuning of parameters in high dimensional data. With adaptive mechanisms responding dynamically to the requirements of the problem, BMRO ensures that the parameters for radial spread are optimized and hence robust to variations in input data. In comparison with other models, BMRO has the peculiar ability to balance exploration and exploitation dynamically, which many traditional algorithms, such as Particle Swarm Optimization (PSO) or Genetic Algorithms (GA), perform poorly due to their inherent nature. Moreover, its biologically inspired mechanisms guarantee a natural adaptability to problem-specific constraints, which makes it more versatile than purely mathematical or heuristic approaches. Therefore, BMRO is especially equipped to handle intricacies inherent in DR datasets, such as irregular image patterns, varying stages of disease, and imbalanced class distributions all unlike the conventional optimization techniques.

4. Results and Discussions

This section will evaluate first the performances of the proposed methods on two famous datasets: Kaggle's Diabetic Retinopathy Detection dataset and the APTOS 2019 Blindness Detection dataset. The reason for choosing those datasets is based on their large collection of diverse retinal images that guarantee a vital bench mark for detecting and classifying DR. It is class-balanced for the different stages of DR; hence, it is a very good candidate for deep learning model training and evaluation. On the other hand, the APTOS 2019 dataset, while placing much emphasis on blindness detection, has variable resolutions of the annotated retinal images, which definitely require a robust preprocessing and feature extraction approach for handling variability in image quality. Preprocessing was done to enhance the image quality of the retinal images and to standardize them for input into the proposed models. In the step of training and validation, each dataset was randomly divided into training, validation, and testing sets while keeping a fair number of images from each DR severity level. The Kaggle dataset was used as the primary dataset in training the Swin-RadialNet, while the APTOS 2019 dataset was used as an external validation set to test the generalizability of the proposed methods. he results have proven that the proposed methods are highly effective in the detection and classification of the stages of DR with considerable accuracy.

Table 1 showcases the accuracy of the proposed approach in comparison with other advanced techniques [17]. Three advanced models named ResAttention-EfficientNet along with GoogleNet + ResNet and BCNN have exhibited noteworthy performance according to accuracy tests which reached 93.5%, 94% and 97.68% respectively. The proposed approach delivers accuracy rates of 98.92% in detecting Diabetic Retinopathy beyond existing techniques which establishes a new standard for DR detection in medical applications. The Swin-RadialNet with its ability to capture different scales of features in retinal images stands alongside the BMRO which optimizes model parameter tuning to provide this system's superior accuracy. The comparison between existing literature models for DR detection extends into Table 2 while using data from the Kaggle diabetic retinopathy dataset.

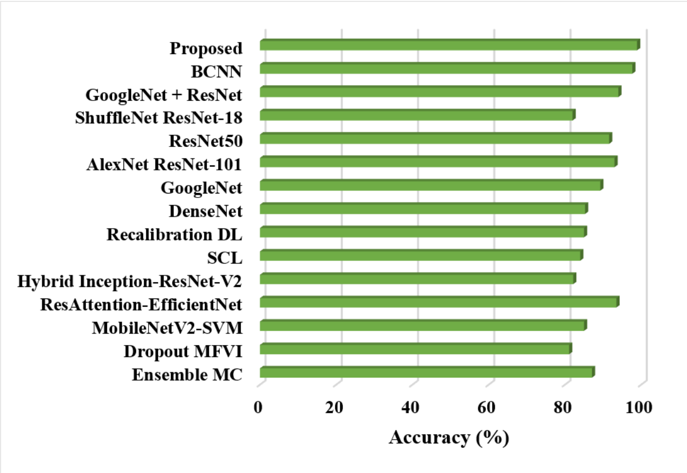

Fig.3. Comparison with recent state of the art approaches based on accuracy

Table 1. Comparative study with recent state of the art approaches based on accuracy

Methods	Accuracy (%)
Ensemble MC	87.1
Dropout MFVI	81.1
MobileNetV2-SVM	85
ResAttention-EfficientNet	93.5
Hybrid Inception-ResNet-V2	82.18
SCL	84
Recalibration DL	85
DenseNet	85.28
GoogleNet	89.28
AlexNet ResNet-101	93
ResNet50	91.60
ShuffleNet ResNet-18	82
GoogleNet + ResNet	94
BCNN	97.68
Proposed	98.92

The proposed method outperforms DenseNet121 (94.20%), VGG19 (93.10%), and EfficientNet-B6 (86.03%) because it reaches 98.90% accuracy. Model resilience to handle variable image resolutions and quality is required for working effectively on the complex Kaggle database. This network model attains superior accuracy thanks to its utilization of Swin-Transformer methods for hierarchical features with Radial Residual Networks for improved propagation abilities. Fig 3 demonstrates the significantly leading performance of the proposed approach in comparison to traditional methods [25] which can be seen through Table 1. Additional visual evidence in Fig 4 demonstrates the consistent performance advantage over previous models from Table 2 through a comparative analysis. The novel arrangement of Swin-Transformers with radial residual architectures implementing hybrid optimization provides better accuracy together with efficient computation and generalized results. The proposed framework stands out with its comprehensive solution because it connects accuracy to robustness by refining both specific feature extraction and optimization techniques in a manner suitable for real-world situations.

Table 2. Comparative study with other existing literature models using Kaggle's diabetic retinopathy

Methods	Accuracy (%)
SVM	87.04
IncResNetV2	82.18
Hybrid CNN	79.50
VGG16	84.31
Inception V3	82
EfficientNet-B6	86.03
CNN	85
DenseNet121	94.20
VGG19	93.10
VGG16	91.55
EfficientNet2L	91.40
MobileNetV2	90.63
Proposed	98.90

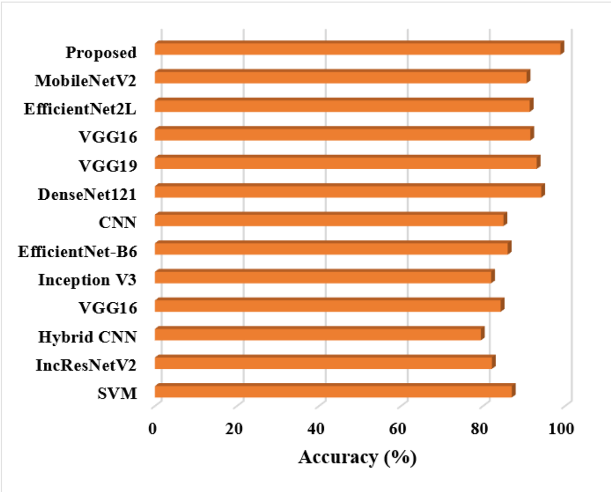

Fig.4. Comparison with other machine learning and deep learning models based on accuracy

Performance analysis of the diabetic retinopathy detection and classification methods is very essential for understanding how effective the models are in identifying and categorizing different stages of the disease with accuracy. More detailed comparative evaluation is presented in Table 3, highlighting precision, recall, and F1-score for different classes of DR, including Mild, Moderate, Severe, and Proliferative, for state-of-the-art approaches. Fig 5 to Fig 7 are graphical representations of these metrics, showing in-depth the advantages of the proposed method. The results indicate that the proposed model considerably outperforms the existing approaches with very good precision, recall, and F1-score values for all DR classes. Specifically, the proposed approach achieves near-flawless metrics across all categories: it achieves 98.85%, 98.82%, and 98.80% for precision, recall, and F1-score, respectively, in the class of Mild DR. These are unbeaten when considering the closest contender, HFF-Net, which yields 88.79% precision, 86.92% recall, and 87.85% F1-score. The remarkable performance exhibited proves this model to be very effective in producing precise identification of early DR manifestations with very few false positives or negatives.

The proposed model demonstrates its superior feature extraction capability in the Moderate DR class, where retinal abnormalities become prominent and more complex to classify, with 98.49% precision, 98.48% recall, and 98.50% F1-score. The improvement is gigantic over other models, including InceptionNet and DenseNet201, which show limited effectiveness with F1-scores of 77.48% and 76.48%, respectively. The high accuracy achieved by the proposed method in the Moderate DR class is because of its advanced architecture that smoothly integrates Swin-Transformer blocks and Radial Residual Networks. Moreover, the capability of the proposed model in dealing with more severe DR, characterized by more pronounced lesions and hemorrhages, is no less impressive with an F1-score of 98.57%. The scores achieved by the existing methods, ResNet-18 and DenseNet201, are far lower: 39.26% and 47.18%, respectively.

Table 3. Overall performance study with respect to different classes of DR

Methods	Class	Precision	Recall	F1-score
ResNet-18	Mild	71.99	85.65	78.23
	Moderate	66.02	87.55	75.28
	Severe	53.54	30.99	39.26
	Proliferative	62.80	48.36	54.64
ResNet-50	Mild	79.74	22.13	76.37
	Moderate	63.93	17.44	93.56
	Severe	56.36	11.29	36.26
	Proliferative	60.80	15.16	50.23
DenseNet201	Mild	85.78	73.84	79.37
	Moderate	68.97	85.84	76.48
	Severe	54.62	41.52	47.18
	Proliferative	56.09	60.56	58.24
Xception Net	Mild	85.92	23.84	77.22
	Moderate	67.10	18.31	88.41
	Severe	57.24	11.46	48.54
	Proliferative	59.79	14.91	53.05
InceptionNet	Mild	80.25	82.28	81.25
	Moderate	66.77	92.27	77.48
	Severe	60.33	42.69	50
	Proliferative	69.64	54.93	61.42
HFF-Net	Mild	88.79	86.92	87.85
	Moderate	70.79	88.41	78.63
	Severe	57.80	58.48	58.14
	Proliferative	74.68	55.40	63.61
Proposed	Mild	98.85	98.82	98.80
	Moderate	98.49	98.48	98.50
	Severe	98.64	98.60	98.57
	Proliferative	98.60	98.58	98.49

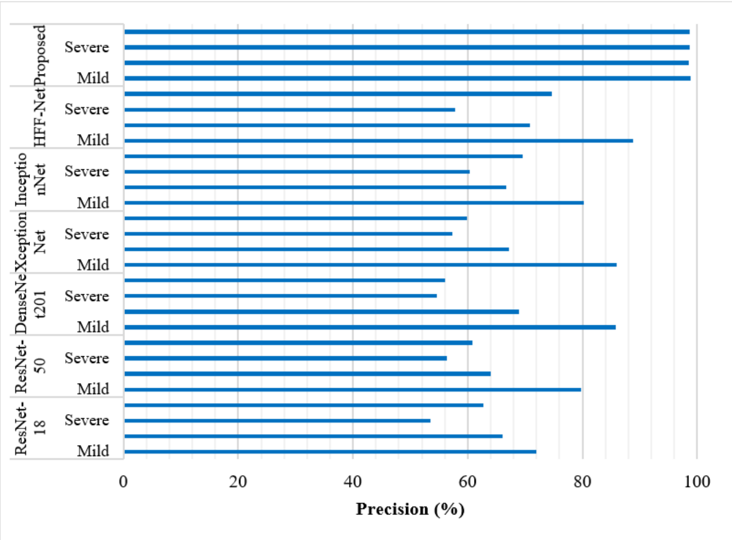

Fig.5. Comparison with recent models based on precision with respect to different classes of DR

The most difficult for classification models is the Proliferative DR class, which represents the most advanced stage of the disease, characterized by extensive retinal damage and neovascularization. The proposed model achieves high precision, recall, and F1-score values of 98.60%, 98.58%, and 98.49%, respectively. The rest of the models, especially the Xception Net and InceptionNet, show remarkably lower outcomes. This proves better performance in this critical category, which ensures the analysis and classification of the most intricate visual traits are done precisely for reliable detection in real clinical settings. The combination with Swin-RadialNet and BMRO has been integral in the attainment of such a result. On the other hand, hierarchical feature extraction of Swin-RadialNet coupled with BMRO for precise parameter optimization ensures the model can address various retinal abnormalities with unparalleled accuracy. This hybrid framework enhances the representation of features while ensuring efficient tuning of the critical parameters to improved convergence at a reduced computational overhead. Unlike previous methods, which tend to vary greatly in their measures according to class complexity, this approach keeps its precision, recall, and F1-score near to perfection for all classes. Stability is quite important in medical diagnosis because a strong and precise classification can notably impact patient outcome.

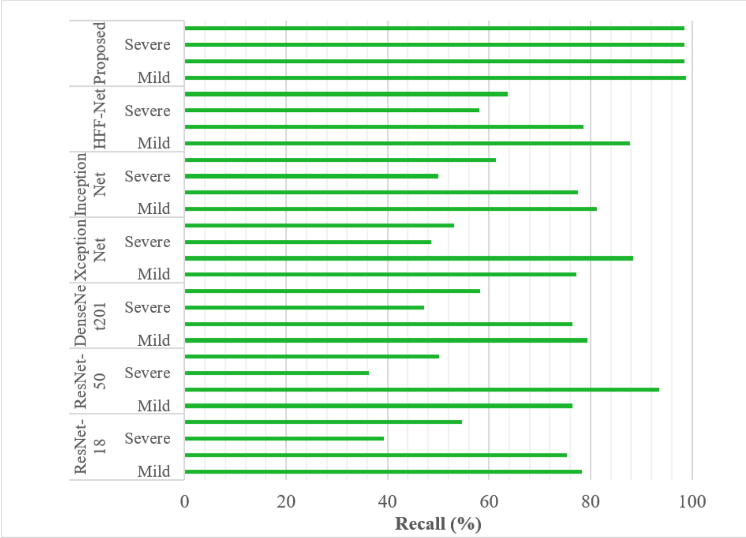

Fig.6. Comparison with recent models based on recall with respect to different classes of DR

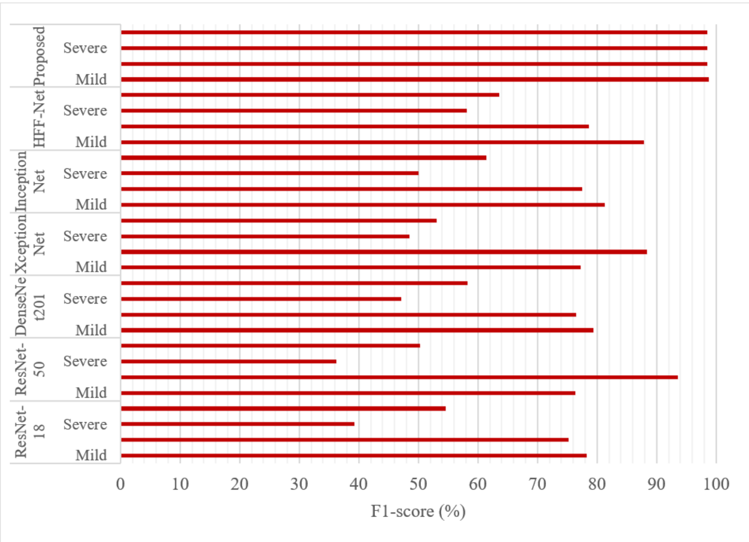

Fig.7. Comparison with recent models based on f1-score with respect to different classes of DR

In addition to Swin-RadialNet and BMRO-based frameworks, will be beneficial for improving efficiency and accuracy in the diagnosis of DR, introducing these systems within real-world clinical contexts will still be faced with some challenges. For instance, the framework can be used as a support tool in current diagnostic workflows, being incorporated directly into EHR systems to assist ophthalmologists with automated, on-the-spot analysis and scoring of fundus images. In urban hospitals and rural outreach programs, where experts may not be available, it can screen for DR with high precision by processing huge amounts of heterogeneous data, therefore being well-positioned for use in these conditions. Some obstacles will need to be surmounted: model interpretability, regulatory approval, compatibility with existing imaging devices, and compliance with data privacy. The need to develop explainability modules which allow visualization and validation of detected lesions should be driven by clinician concerns about what they perceive as a black box offering explanations for predictions made by deep learning models. It would also be necessary for the system to meet the medical device regulation and be easily integrated into different imaging formats and healthcare IT infrastructures. Performance maintenance and adaptation of the system to local variations would require continuous retraining with local data and inputs from clinicians. The best possible implementation of the framework may thus alleviate the burden of screening, accelerate diagnosis, and improve patient outcomes notwithstanding these challenges across various healthcare settings.

5. Conclusions

This paper provides a complete framework for the accurate detection and classification of diabetic retinopathy, addressing critical challenges of early diagnosis and an effective classification of its stages. The proposed methodology combines Swin-RadialNet with BMRO to form a robust hybrid system that converges powerful feature extraction with precise parameter optimization. The novelty of the work lies in seamlessly integrating a transformer-based architecture with radial residual learning and a new hybrid optimization algorithm, enabling the betterment of performance in the detection and classification of DR at various stages. This new approach will ensure improved feature representation, less overfitting, and efficient parameter tuning features very much needed in high-accuracy medical image analysis. The overall contribution of the proposed work is manifold, offering a major step forward in the diagnosis of DR. It puts forward a new optimization algorithm, BMRO, for effective estimation of the radial spread parameters in improving feature selection for enhancing model generalization. This embeds Swin-Transformer in a residual network architecture for hierarchical and localized feature extraction, capturing intricate retinal patterns with great accuracy. The proposed methodology exposes better adaptability and scalability to large datasets, so it fits much more within real-world clinical applications. This work, addressing the key limitations of the existing models, will lay a strong foundation for future advancements in the automated detection of DR.

The proposed framework enhances the accuracy of diagnosis, reduces both false positives and negatives and improves interpretability, all of which are critical for the early intervention and managing patients' complications. Features description and classification from input images of the retina by means of the Swin-RadialNet architecture are the next levels of the proposed method after preprocessing of the retinal images. The BMRO is then employed for optimal parameter tuning ensuring high convergence efficiency with minimum computational overhead. Due to its extremely strong design and incorporation of state-of-the-art optimization techniques, it can effectively work with complex and subtle anomalies of the retina. As demonstrated by all experiments and outcomes that support this work, the proposed methods have proven to be effective for DR diagnosis. Packs of data, including the well-known Kaggle diabetic retinopathy dataset and APTOS 2019, were used for extensive experimentation showing that the proposed framework has outperformed the state-of-the-art methods. The proposed model has reported almost perfect values in accuracy, precision, recall, and F1-score for all classes of DR and therefore is widely applicable and dependable.

References

[1] D. Muthusamy and P. Palani, "Deep neural network model for diagnosing diabetic retinopathy detection: An efficient mechanism for diabetic management," Biomedical Signal Processing and Control, vol. 100, p. 107035, 2025.

[2] S. A. Kumar, J. S. Kumar, and S. C. Bharadwaj, "Efficient diabetic retinopathy detection using deep learning approaches and Raspberry Pi 4," Bulletin of Electrical Engineering and Informatics, vol. 14, pp. 1063-1072, 2025.

[3] M. Hussain, H. A. Ahmed, M. Z. Babar, A. Ali, H. Shahzad, S. ur Rehman, et al., "An Enhanced Convolutional Neural Network (CNN) based P-EDR Mechanism for Diagnosis of Diabetic Retinopathy (DR) using Machine Learning," Engineering, Technology & Applied Science Research, vol. 15, pp. 19062-19067, 2025.

[4] A. Biswas and R. Banik, "Deep Learning-Based Image Segmentation for Early Detection of Diabetic Retinopathy and Other Retinal Disorders," Deep Learning Applications in Medical Image Segmentation: Overview, Approaches, and Challenges, pp. 133-148, 2025.

[5] Choubey, S. B., Chitra, T., & Hephzipah, J. J. (2024). Big Data Mining for Chronic Disease Prediction using Principal Component Analysis and eXtreme Gradient Boosting. *GK International Journal of Advanced Research in Engineering and Technology*, *1*(1), 1-11.

[6] Anusuya, S., & Jayarin, P. J. (2017, March). Survey on multimodel biometrics on human identification using real time filtering algorithm. In *2017 Third International Conference on Science Technology Engineering & Management (ICONSTEM)* (pp. 142-147). IEEE.

[7] Y. Tao, M. Xiong, Y. Peng, L. Yao, H. Zhu, Q. Zhou, et al., "Machine learning-based identification and validation of immune-related biomarkers for early diagnosis and targeted therapy in diabetic retinopathy," Gene, vol. 934, p. 149015, 2025.

[8] G. Maheswari and S. Gopalakrishnan, "Advanced Attention Capsule Network with Curved Wavelet Attention for Illuminating Low-Light Capsule Endoscopy," in 2024 IEEE 3rd World Conference on Applied Intelligence and Computing (AIC), 2024, pp. 1148-1154.

[9] V. K. Isukapalli, R. D. Kumar, and S. Gopalakrishnan, "Image-Based Bone Marrow Malignancy Detection: Motivation, Challenges and Recommendations," in 2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT), 2024, pp. 1-6.

[10] R. Kavitha, "A Deep Learning Framework to Detect Diabetic Retinopathy Using CNN," in Deep Learning in Medical Image Analysis, ed: Chapman and Hall/CRC, 2025, pp. 79-96.

[11] A. K. Singh, S. Madarapu, and S. Ari, "Diabetic retinopathy grading based on multi-scale residual network and cross-attention module," Digital Signal Processing, vol. 157, p. 104888, 2025.

[12] S. Aftab and S. Akhtar, "Diabetic Retinopathy Severity Classification Using Data Fusion and Ensemble Transfer Learning," Journal of Software Engineering and Applications, vol. 18, pp. 1-23, 2025.

[13] R. Manjushree, D. Bhoomika, R. R. Nair, and T. Babu, "Automated detection of diabetic retinopathy using deep learning in retinal fundus images: analysis," in 2022 3rd International Conference on Communication, Computing and Industry 4.0 (C2I4), 2022, pp. 1-6.
[14] N. M. Suganthi and M. Arun, "Diabetic retinopathy grading using curvelet CNN with optimized SSO activations and wavelet-based image enhancement," Ain Shams Engineering Journal, vol. 16, p. 103239, 2025.
[15] G. Maheswari and S. Gopalakrishnan, "A smart multimodal framework based on squeeze excitation capsule network (SECNet) model for disease diagnosis using dissimilar medical images," International Journal of Information Technology, pp. 1-19, 2024.
[16] N. Kavitha and N. Kasthuri, "An efficient automatic diabetic retinopathy grading using a two-way cascaded convolution neural network," Discover Computing, vol. 28, pp. 1-16, 2025.
[17] M. Akram, M. Adnan, S. F. Ali, J. Ahmad, A. Yousef, T. A. N. Alshalali, et al., "Uncertainty-aware diabetic retinopathy detection using deep learning enhanced by Bayesian approaches," Scientific Reports, vol. 15, p. 1342, 2025.
[18] K. Kaushik, A. Bhardwaj, X. Cheng, S. Dahiya, A. Shankar, M. Kumar, et al., "Residual Network-Based Deep Learning Framework for Diabetic Retinopathy Detection," Journal of Database Management (JDM), vol. 36, pp. 1-21, 2025.
[19] I. Govindharaj, A. Poongodai, G. Rajaram, D. Santhakumar, S. Ravichandran, K. Udayakumar, et al., "Enhanced diabetic retinopathy detection using U-shaped network and capsule network-driven deep learning," MethodsX, vol. 14, p. 103052, 2025.
[20] G. Neri, S. Sharma, B. Ghezzo, C. Novarese, C. Olivieri, D. Tibaldi, et al., "Deep learning model for automatic detection of different types of microaneurysms in diabetic retinopathy," Eye, pp. 1-8, 2025.
[21] A. Ikram and A. Imran, "ResViT FusionNet Model: An explainable AI-driven approach for automated grading of diabetic retinopathy in retinal images," Computers in Biology and Medicine, vol. 186, p. 109656, 2025.
[22] U. Bhimavarapu, "Automated detection of diabetic retinopathy using an improved deep learning model with smartphone images," International Journal of Diabetes in Developing Countries, pp. 1-13, 2025.
[23] A. Luong, J. Cheung, S. McMurtry, C. Nelson, T. Najac, P. Ortiz, et al., "Comparison of Machine Learning Models to a Novel Score in the Identification of Patients at Low Risk for Diabetic Retinopathy," Ophthalmology Science, vol. 5, p. 100592, 2025.
[24] R. Manohar and M. S. Aarthi, "Leveraging Deep Learning for Early Stage Diabetic Retinopathy Detection: a Novel CNN and Transfer Learning Comparision," in 2024 2nd International Conference on Artificial Intelligence and Machine Learning Applications Theme: Healthcare and Internet of Things (AIMLA), 2024, pp. 1-6.
[25] S. Guefrachi, A. Echtioui, and H. Hamam, "Diabetic Retinopathy Detection Using Deep Learning Multistage Training Method," Arabian Journal for Science and Engineering, pp. 1-18, 2024.

Authors' Profiles

Mrs.Vijayalaxmi Gopu pursuing Ph.D in the Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology (Deemed to be University) ,Chennai. India. Her research intrest Deep Learning,Image Processing and Soft Computing Techniques.

Dr. M. Selvi working as an associate Professor, Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology (Deemed to be University), Chennai. Tamil Nadu 600119, India. She Published many research articles in reputed national and International Journals. Her research interest includes Machine Learning, Cloud Computing, Image Processing and IoT.

How to cite this paper: Vijayalaxmi Gopu, M. Selvi, "A Swin-transformer Integrated with Radial Optimization Model for Accurate Diabetic Retinopathy Detection and Classification", International Journal of Information Technology and Computer Science(IJITCS), Vol.17, No.6, pp.160-175, 2025. DOI:10.5815/ijitcs.2025.06.09