scientific reports

OPEN

Ensemble deep learning and EfficientNet for accurate diagnosis of diabetic retinopathy

Lakshay Arora[1,12], Sunil K. Singh[1,12✉], Sudhakar Kumar[1,12], Hardik Gupta[1,12], Wadee Alhalabi[2,12], Varsha Arya[3,7,8,12], Shavi Bansal[4,12], Kwok Tai Chui[5,12] & Brij B. Gupta[6,9,10,11,12✉]

Diabetic Retinopathy (DR) stands as a significant global cause of vision impairment, underscoring the critical importance of early detection in mitigating its impact. Addressing this challenge head-on, this study introduces an innovative deep learning framework tailored for DR diagnosis. The proposed framework utilizes the EfficientNetB0 architecture to classify diabetic retinopathy severity levels from retinal images. By harnessing advanced techniques in computer vision and machine learning, the proposed model aims to deliver precise and dependable DR diagnoses. Continuous testing and experimentation shows to the efficiency of the architecture, showcasing promising outcomes that could help in the transformation of both diagnosing and treatment of DR. This framework takes help from the EfficientNet Machine Learning algorithms and employing advanced CNN layering techniques. The dataset utilized in this study is titled 'Diagnosis of Diabetic Retinopathy' and is sourced from Kaggle. It consists of 35,108 retinal images, classified into five categories: No Diabetic Retinopathy (DR), Mild DR, Moderate DR, Severe DR, and Proliferative DR. Through rigorous testing, the framework yields impressive results, boasting an average accuracy of 86.53% and a loss rate of 0.5663. A comparison with alternative approaches underscores the effectiveness of EfficientNet in handling classification tasks for diabetic retinopathy, particularly highlighting its high accuracy and generalizability across DR severity levels. These findings highlight the framework's potential to significantly advance the field of DR diagnosis, given more advanced datasets and more training resources which leads it to be offering clinicians a powerful tool for early intervention and improved patient outcomes.

Diabetic retinopathy, a consequence of diabetes mellitus, manifests as alterations in the blood vessels within the retina. This condition poses a significant threat to vision integrity and can culminate in blindness if not appropriately managed. The underlying pathology stems from prolonged exposure to elevated levels of glucose in the bloodstream, which gradually compromises the structural integrity of the retinal vasculature. Over time, this vascular damage manifests as microaneurysms, hemorrhages, and exudates, which impair retinal function and compromise visual acuity[1,2].

The progression of diabetic retinopathy is insidious, often characterized by an asymptomatic early phase. However, as the condition advances, symptoms such as blurred vision, floaters, and visual field deficits may emerge, signaling the need for prompt intervention. The hallmark of diabetic retinopathy lies in its propensity to induce macular edema and proliferative changes within the retina, further exacerbating visual impairment[3].

[1]Department of CSE, Chandigarh College of Engineering and Technology, Panjab University, Chandigarh, India. [2]Department of Computer Science, Immersive Virtual Reality Research Group, King Abdulaziz University, Jeddah, Saudi Arabia. [3]Department of Electrical and Computer Engineering, Lebanese American University, Beirut 1102, Lebanon. [4]Insights2Techinfo, Jaipur, India. [5]Hong Kong Metropolitan University (HKMU), Kowloon, Hong Kong. [6]Present address: Department of Computer Science and Information Engineering, Asia University, Taichung 413, Taiwan. [7]Center for Interdisciplinary Research, University of Petroleum and Energy Studies (UPES), Dehradun, India. [8]UCRD, Chandigarh University, Chandigarh, India. [9]Kyung Hee University, 26 Kyungheedae-ro, Dongdaemun-gu, Seoul 02447, Korea. [10]Symbiosis Centre for Information Technology (SCIT), Symbiosis International University, Pune, India. [11]University of Economics and Human Science, Warsaw, Poland. [12]Lakshay Arora, Sunil K. Singh, Sudhakar Kumar, Hardik Gupta, Wadee Alhalabi, Varsha Arya, Shavi Bansal, Kwok Tai Chui and Brij B. Gupta contributed equally to this work. ✉email: sksingh@ccet.ac.in; bbgupta@asia.edu.tw
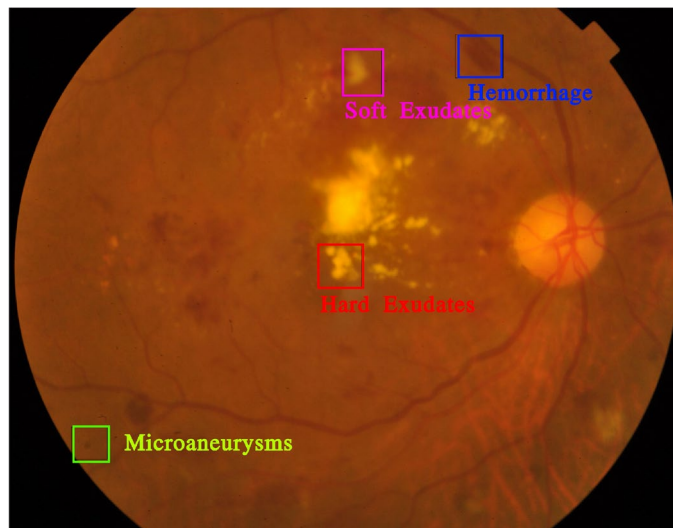
Fig. 1. Illustration of diabetic retinopathy lesions[5].

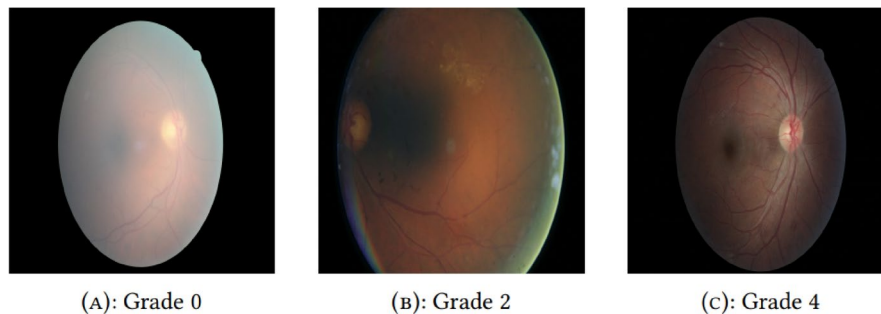(A): Grade 0   (B): Grade 2   (c): Grade 4

Fig. 2. Various grade illustrations[6].

Routine ophthalmic examination plays a pivotal role in the early detection and management of diabetic retinopathy. Screening modalities such as fundus photography and optical coherence tomography enable clinicians to identify subtle retinal changes before symptomatic onset. Timely intervention is imperative to mitigate disease progression and preserve visual function[4].

The detection of diabetic retinopathy is often aided by identifying specific retinal lesions. One such lesion, a microaneurysm, appears as a small, round dot on the retina during the initial phases of the disease. These dots have well-defined edges and are typically no more than 125 micrometers in diameter. Although there are six distinct types of microaneurysms, treatment approaches remain consistent across all types. Another lesion linked to diabetic retinopathy is the haemorrhage, which presents as a large, irregularly shaped spot on the retina, generally exceeding 125 micrometers in size. Additional lesions, known as hard exudates, result from plasma leakage and show up as yellow spots with defined borders on the retina. Soft exudates, caused by swelling of nerve fibers, take the form of white, oval patches on the retina, as illustrated in Fig. 1.

The International Clinical Diabetic Retinopathy Disease Severity Scale (ICDRDSS), introduced by the Global Diabetic Retinopathy Project Group, is a pivotal classification system designed for the assessment and categorization of diabetic retinopathy (DR).The ICDRDSS comprises five distinct severity levels, each capturing the progression of the disease and facilitate accurate diagnosis and treatment planning.

DR progression stages are characterized by different disease severity levels. Figure 2 shows some of the grades for Diabetic Retinopathy The findings associated with each grade are as seen in Table 1.

Each stage of diabetic retinopathy is characterized by unique characteristics and properties, which can pose challenges for doctors when making a diagnosis. According to a study by Google, the stages of diabetic retinopathy may prove difficult to accurately evaluate manually even by well-trained clinicians. So, the concept of automatic detection has garnered attention as a means to enhance accuracy and efficiency in the diagnosis of DR.

The predominant trend in analyzing various models for artificial intelligence (AI) implementation reveals a reliance on six key constituent AI models, all of which are deeply rooted in the domain of machine learning. These models are CNN, ANN, Other-NN, Fuzzy ML, SVM, RF (Random Forest). This study integrates the powerful EfficientNet Framework with the Layering techinques in CNN to build a model that would help

Grade	Description
Grade 0	No DR - No observable indicators of diabetic retinopathy.
Grade 1	Mild NPDR - Small, circular microaneurysms appear on the retina.
Grade 2	Moderate NPDR - Changes beyond microaneurysms are present, though the condition has not progressed to severe NPDR.
Grade 3	Severe NPDR - Any of the following signs are present, without signs of proliferative DR: over 20 intraretinal hemorrhages in all four quadrants, evident venous beading in at least two quadrants, or pronounced intraretinal microvascular abnormalities (IrMA) in one or more quadrants.
Grade 4	Proliferative DR - Characterized by the presence of neovascularization, vitreous or preretinal hemorrhages, or both.

Table 1. International clinical diabetic retinopathy disease severity scale (ICDRDSS)[5].

increase the accuracy for detection of DR. By leveraging CNN layering techniques and compound scaling in EfficientNet, our framework optimally balances performance and computational resources. Through rigorous experimentation, we demonstrate how this model achieves high accuracy in DR classification while providing interpretable outputs that highlight relevant retinal regions. This advancement aims to enhance the practical application of automated DR diagnosis, bridging the gap between model accuracy and clinical interpretability.

EfficientNet's central concept is compound scaling, which systematically adjusts network dimensions-such as depth, width, and resolution-to enhance performance. This method ensures optimal resource allocation, allowing the model to achieve high accuracy while remaining resource-efficient across various tasks and operational constraints[7,8].

The various key contributions in this research article are given as follows:

- The authors conducted an extensive review of existing literature on diabetic retinopathy (DR), focusing on its pathology, impact on vision, and the critical importance of early detection to prevent vision impairment.
- They contributed significantly to the conceptualization and design of an innovative deep learning framework tailored to DR diagnosis, integrating the EfficientNet architecture and CNN layering techniques to enhance diagnostic accuracy.
- The authors collaborated in selecting, preparing, and preprocessing the Kaggle dataset, which comprises 35,108 retinal images categorized across various severity levels of DR, ensuring comprehensive data coverage for training and validation.
- The authors worked collectively in writing and editing the introduction, providing a clear and structured overview of the problem statement, the clinical relevance of the study, and the potential impact of the proposed solution on DR diagnosis.Further the remaining paper comprises of the following sections. "Literature review" Section of the research includes the Literature Review and contains all the previous recent work done in the field. "Preliminary theory" Section of the research includes the background of the framework and the theoretical information regarding the concepts used. "Proposed framework" Section introduces the proposed Framework followed by the experimental setup for the research. "Experimental setup" Section includes the results of the research. "Experimental results" Section discusses the results and describes the limitations in the research. "Discussions and limitations" Section of the research includes the conclusion and future scope in the present research.

Literature review

Deep learning techniques have shown promising results in the automated diagnosis of diabetic retinopathy, a significant cause of blindness globally. Multiple studies have contributed to advancements in this field, demonstrating the effectiveness of deep learning models in accurately detecting diabetic retinopathy from retinal fundus photographs.

The work of Gulshan et al.[9] laid a critical foundation, presenting a robust deep learning algorithm capable of accurately identifying diabetic retinopathy in retinal fundus photographs. Building on this, Ting et al.[10] developed a system that incorporated retinal images from multiethnic populations with diabetes, emphasizing the importance of diversity in dataset design for improved model generalization. Abràmoff et al.[11] further validated the clinical applicability of autonomous AI-based diagnostic systems through a pivotal trial, showcasing their potential in real-world primary care environments and facilitating early intervention. Subsequent research efforts focused on enhancing the practical deployment of these models. Li et al.[12] introduced an automated grading system designed to identify vision-threatening diabetic retinopathy cases, aiming to streamline urgent referrals based on disease severity. In a similar vein, Bellemo et al.[13] conducted clinical validation in African healthcare settings[37], highlighting how deep learning models can offer accessible and accurate diagnostic tools to address healthcare disparities[38]. Bhaskaranand et al.[14] also contributed to this area by developing an automated screening and monitoring system to facilitate timely detection and management of diabetic retinopathy progression.

Further studies have underscored the importance of generalizability and adaptability in model development. Sahlsten et al.[15] conducted a population-based study, emphasizing the necessity of incorporating diverse population data in model training to enhance robustness. Meanwhile, Raman et al.[16] addressed the challenges of training deep learning models on non-mydriatic retinal fundus images from electronic health records[39], enhancing their applicability to real-world clinical datasets. Takahashi et al.[17] extended the scope of deep learning applications to include both diabetic retinopathy and glaucoma diagnosis, showcasing the potential for multifaceted disease detection. Burlina et al.[18], although focused on age-related macular degeneration, demonstrated the versatility of deep learning models in analyzing retinal images across various ophthalmic

conditions. Bilal et al.[29] developed an AI-based system utilizing U-Net and deep learning to automatically detect and classify diabetic retinopathy in fundus images, emphasizing the effectiveness of U-Net architectures in image segmentation tasks critical for disease classification. Building on these efforts, Bilal et al.[30] used a combination of mixed models for classifying diabetic retinopathy severity, showcasing the adaptability of ensemble approaches in handling complex disease grading tasks.

Further enhancing the detection framework, Bilal et al.[31] incorporated transfer learning techniques with U-Net, demonstrating an improved model for diabetic retinopathy detection. Their methodology highlighted the role of pretrained models in efficiently capturing essential retinal features from fundus images. Huang et al.[19] introduced a saliency-guided self-supervised transformer called "Ssit," which enhances diabetic retinopathy grading through focused attention on relevant retinal areas, leveraging saliency maps for improved feature extraction. Validated on large datasets, this model demonstrates significant improvements in grading accuracy. Thanikachalam et al.[20] built on this approach by proposing an optimized deep CNN for detecting diabetic retinopathy and diabetic macular edema, incorporating adaptive learning techniques to achieve high classification accuracy. Bodapati and Balaji[21] further innovated by developing a self-adaptive stacking ensemble model that combines multiple neural networks with attention mechanisms, yielding high accuracy across different imaging conditions.

Advancements in interpretability have also been a focal point in recent studies. Bhati et al.[22] presented IDANet, an interpretable dual attention network[40] that enhances diabetic retinopathy grading by focusing on critical retinal regions, maintaining diagnostic accuracy while providing clinical interpretability. Sivapriya et al.[23] introduced a model that classifies diabetic retinopathy by emphasizing microvascular structures, highlighting the importance of retinal structure analysis. Ohri and Kumar[24] further improved detection through a supervised fine-tuned model, leveraging transfer learning for efficient DR identification. In exploring optimization techniques, Bilal et al.[32] applied Grey Wolf Optimization with CNN models, enhancing feature selection processes and improving classification accuracy, which is pivotal for early diabetic retinopathy intervention. Another approach from Bilal et al.[33] combined CNNs with weighted filters and adaptive filtering, underscoring the importance of noise reduction and data augmentation for accurate classification, especially in retinal image processing.

A significant contribution by Bilal et al.[34] leveraged a CNN-SVD-enhanced Support Vector Machine for detecting vision-threatening diabetic retinopathy, marking a shift toward hybrid models that integrate machine learning with deep learning for robust detection capabilities. Additionally, Bilal et al.[35] introduced EdgeSVDNet, a 5G-enabled framework for diagnosing diabetic retinopathy in real time, which could greatly enhance accessibility and speed in remote medical diagnosis. Bilal et al.[36] proposed the NIMEQ-SACNet model using self-attention mechanisms, tailored for precision medicine applications in diabetic retinopathy. This advanced model integrates self-attention with CNNs to improve accuracy and adaptability, demonstrating the potential for personalized diagnostic tools in precision medicine frameworks. Luo et al.[25] proposed a CNN model that captures both local and long-range dependencies in retinal images, improving classification across diabetic retinopathy stages through enhanced representation of retinal features. Romero-Oraá et al.[26] introduced an attention-based framework to isolate relevant features for DR grading, demonstrating optimized focus in model design. Zhang et al.[27] explored a semi-supervised contrastive learning method, incorporating saliency maps and unlabeled data for improved robustness. Finally, Wong et al.[28] utilized transfer learning[41] with parameter optimization, demonstrating the efficacy of feature-weighted ECOC ensembles in enhancing diabetic retinopathy diagnostics.

These studies collectively are shown in Tables 2, 3 and 4 highlight the advancements in deep learning techniques for diabetic retinopathy diagnosis, emphasizing the importance of early detection and intervention in preventing vision loss. The proposed deep learning framework in this paper builds upon and contributes to this body of research by leveraging EfficientNet as the base model and incorporating advanced CNN layering to enhance model effectiveness.

Research gap
While the existing body of research in diabetic retinopathy (DR) diagnosis has made considerable progress using deep learning techniques, a significant research gap persists in leveraging advanced architectures to maximize diagnostic accuracy and efficiency. Previous studies have largely focused on utilizing architectures such as Inception, ResNet, and DenseNets for DR detection, which have shown promising results in terms of accuracy. However, there remains an opportunity to explore novel architectures that can deliver higher precision and robustness in classifying DR severity levels. This study aims to address this gap by introducing a framework based on the EfficientNetB0 architecture, designed specifically to improve diagnostic accuracy and reliability for DR severity classification. Through extensive testing and validation, the proposed model seeks to bridge the existing gap by providing a dependable solution that achieves high accuracy rates and effectively supports clinical decision-making in DR diagnosis.

Preliminary theory
EfficientNet
EfficientNet is a family of convolutional neural network (CNN) models[42] appreciated for their efficiency and effectiveness in image classification tasks. The architecture of EfficientNet models is characterized by a stem, followed by multiple blocks, each containing sub-blocks. The overall architecture is composed of five modules, which are combined to create the EfficientNet models[43,44]. These modules can be seen in Figure 3.

Stem: The initial portion of the network, the stem, is where the foundation for subsequent layers is set. The input image is processed here to extract basic features.

Blocks: EfficientNet-B0 is comprised of seven blocks, with each contributing to the hierarchical feature extraction process. These blocks play a crucial role in the learning of increasingly abstract features as information flows through the network.

S. no.	References	Methodologies	Data source	Advantages	Disadvantages and gaps
1	Gulshan et al.[9]	CNN-based deep learning model for binary DR classification	Retinal fundus images from EyePACS dataset	Demonstrated high sensitivity and specificity in a clinical setting	Limited to binary classification (DR vs. no DR); lacks granularity in severity levels and interpretability
2	Ting et al.[10]	CNNs trained on multiethnic retinal images, with system fine-tuned for DR severity prediction	Retinal images from multiethnic EyePACS, US and Singapore datasets	Enhanced generalizability across ethnicities and improved accuracy	High computational demand; limited explainability for clinical use
3	Abràmoff et al.[11]	Autonomous AI model with convolutional neural networks for DR diagnosis	EyePACS data from primary care settings	High applicability in clinical settings with autonomous operation	Focuses on binary classification; does not provide insights into DR severity grading
4	Li et al.[12]	Multi-stage CNN framework targeting vision-threatening DR	Large dataset of color fundus photographs	Specialized in detecting severe DR cases, improving triaging	Limited interpretability and generalization due to specialized target
5	Bellemo et al.[13]	AI-based CNN model tailored for DR screening in low-resource settings	Retinal fundus images from African clinical settings	Validated model efficacy in diverse, resource-limited regions	Model scalability is constrained; limited interpretability for practical diagnostic insights
6	Bhaskaranand et al.[14]	Deep learning-based automated screening and monitoring system for DR	Retinal fundus images	Offers continuous monitoring of DR progression, supports early detection	Limited interpretability in decision-making, lacking an advanced explanation-guided method
7	Sahlsten et al.[15]	Developed and validated DL algorithms for DR on a large, diverse population	Multiethnic population data	High robustness due to diverse population data, ensuring generalizability	No focused methodology for interpretability, which can limit clinician trust
8	Raman et al.[16]	Developed DL models on non-mydriatic images to detect DR	Electronic health records	Supports diagnosis using non-mydriatic images, allowing more accessible and frequent testing	Limited attention mechanisms, interpretability concerns
9	Takahashi et al.[17]	Dual-purpose model assessing both DR and glaucoma	Retinal fundus photographs	Supports multi-disease diagnosis, enhancing model utility in broader ophthalmology	Lacks targeted grading for different DR severity stages, gaps in explaining results for individual conditions
10	Burlina et al.[18]	Convolutional Neural Networks for automated grading of age-related macular degeneration (AMD)	Color fundus images	High accuracy in AMD grading, shows potential for adaptation to related conditions such as DR	Focus on AMD limits direct applicability to DR, interpretability and DR severity grading are not addressed

Table 2. Related works in diabetic retinopathy diagnosis part - 1.

S. no.	References	Methodologies	Data source	Advantages	Disadvantages and gaps
11	Huang et al.[19]	Self-supervised transformer with saliency maps for improving model attention and grading accuracy	Publicly available fundus image datasets	Enhanced model attention and interpretability through saliency maps	High complexity due to transformer-based architecture, may require extensive computation resources
12	Thanikachalam et al.[20]	Optimized CNN with adaptive learning and hyperparameter tuning	Fundus images from open repositories	Efficient in resource usage and shows improved accuracy with adaptive learning	Limited interpretability, potential overfitting due to CNN's static feature extraction capabilities
13	Bodapati & Balaji[21]	Ensemble of attention-based neural networks with stacking for increased robustness	Proprietary DR dataset	High robustness and accuracy, effective in handling severity prediction	May suffer from longer training times and resource requirements due to stacking and ensemble strategies
14	Bhati et al[22]	Dual attention network focusing on critical retinal regions for interpretability	DR fundus images from large-scale datasets	Improved interpretability with attention on specific retinal regions, higher accuracy in grading	Complex architecture may be computationally intensive; dependency on quality of retinal images
15	Sivapriya et al[23]	Deep learning model analyzing microvascular structures in fundus images	Public DR fundus image datasets	Focuses on fine-grained features, potentially enhancing detection sensitivity	Limited scalability, may struggle with low-quality images where microvascular structures are not clear
16	Ohri & Kumar[24]	Supervised learning with transfer learning, fine-tuning on CNN models for improved DR detection accuracy	Not Specified	High accuracy due to supervised fine-tuning and transfer learning; improved model performance on labeled DR data	Dependency on large, labeled datasets; limited interpretability due to black-box nature of CNN; lacks attention mechanisms for feature focus
17	Luo et al.[25]	Deep CNN with local and global retinal features using long-range dependency modeling	Retinal Fundus Images	Effective in capturing both fine-grained and global retinal features for accurate DR stage classification	High computational cost due to deep CNN layers; may require extensive preprocessing of images for better feature extraction
18	Romero-Oraá et al.[26]	Attention mechanisms isolate relevant features, focusing model on key retinal regions for grading	Fundus Image Dataset	Improved interpretability by visualizing feature importance; focuses on critical areas, enhancing decision accuracy	Limited to fundus images; potential over-reliance on salient features, may overlook subtle signs; lacks robustness across diverse datasets
19	Zhang et al.[27]	Semi-supervised learning with contrastive approach, uses saliency maps for robust grading	Not Specified	Reduces dependency on large labeled datasets; robust against noisy data due to contrastive learning	Complexity in implementing contrastive learning; risk of lower performance without accurate saliency maps for guidance
20	Wong et al.[28]	Transfer learning optimized with feature-weighted Error-Correcting Output Codes (ECOC) for ensemble grading	Diverse DR Image Dataset	High accuracy due to parameter optimization and ensemble approach; feature-weighted ECOC enhances classification	Complexity in hyperparameter tuning; potentially high resource requirements due to ensemble approach and ECOC processing complexity

Table 3. Related works in diabetic retinopathy diagnosis part - 2.

S. no.	References	Methodologies	Data source	Advantages	Disadvantages and gaps
21	Bilal et al.[29]	AI-based automatic detection and classification using U-Net and deep learning	Retinal fundus images	Highly effective in image segmentation for disease classification	Limited to dataset characteristics, lacking generalizability across diverse populations
22	Bilal et al.[30]	Mixed models for disease grading and severity classification	Diabetic Retinopathy Grading Database	Improved adaptability of ensemble models for complex grading	Mixed models may increase computational complexity
23	Bilal et al.[31]	Transfer learning with U-Net for enhanced detection accuracy	Retinal fundus images	Efficient feature extraction with pretrained models	Transfer learning limited by domain specificity
24	Bilal et al.[32]	Grey Wolf Optimization with CNN for feature selection	Retinal images	Enhanced feature selection and classification accuracy	Optimization technique may not generalize well to all data types
25	Bilal et al.[33]	CNNs with weighted filters and adaptive filtering for classification	Retinal fundus images	Effective noise reduction and improved classification accuracy	Increased model complexity and training time
26	Bilal et al.[34]	CNN-SVD-enhanced SVM for detecting vision-threatening retinopathy	Retinal fundus images	Robust detection capabilities through hybrid model	Complex hybrid structure may require significant computational resources
27	Bilal et al.[35]	EdgeSVDNet, 5G-enabled for real-time diagnosis	Retinal fundus images with 5G connectivity	Enhanced accessibility and speed for remote diagnostics	Dependent on 5G infrastructure, limited in areas without high-speed connectivity
28	Bilal et al.[36]	NIMEQ-SACNet model with self-attention for precision medicine	Retinal image data for precision diagnostics	High accuracy and adaptability for precision medicine applications	Complexity of self-attention mechanism may increase model size and training requirements

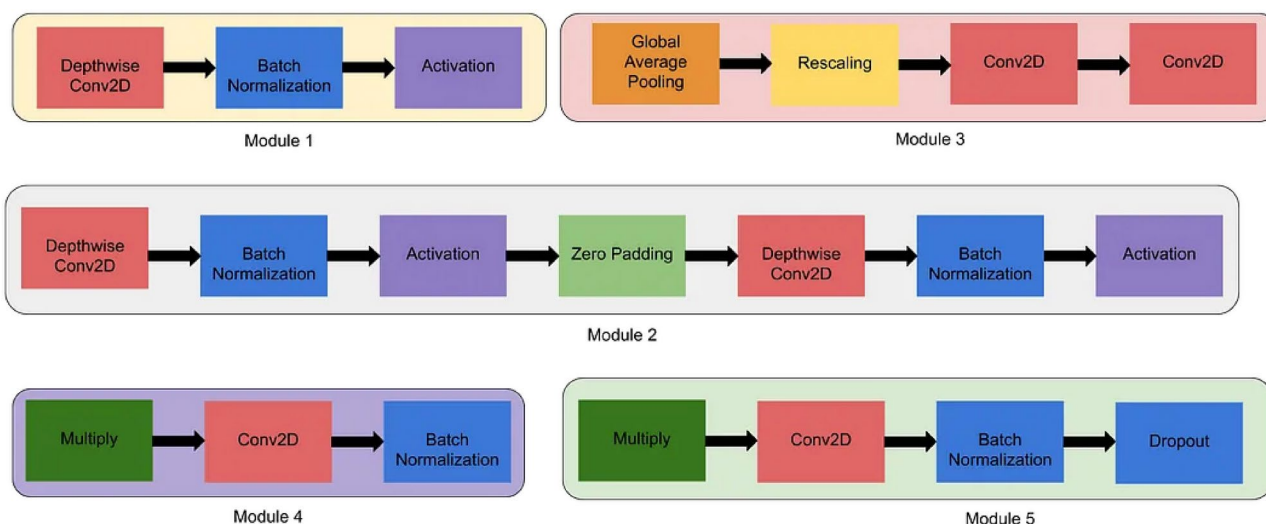**Table 4**. Related works in diabetic retinopathy diagnosis part - 3.

Fig. 3. Types of modules[45].

Sub-blocks: Within each block, there are varying numbers of sub-blocks. Specific operations are performed by each sub-block to transform the input features.
Modules:

- *Module 1:* Serving as the starting point for the first sub-block in the first block.
- *Module 2:* Acting as the starting point for the first sub-block in all blocks except the first one.
- *Module 3:* Skip connections are established to all sub-blocks, facilitating information flow and aiding in gradient propagation.
- *Module 4:* Utilized for combining skip connections in the first sub-blocks, thereby enhancing feature representation.
- *Module 5:* Connecting each sub-block to its preceding sub-block via skip connections and combining them to refine feature maps.Sub-block Types:

- *Sub-block 1:* Exclusive to the first sub-block in the first block, initializing the feature extraction process.
- *Sub-block 2:* Employed as the first sub-block in subsequent blocks, contributing to feature refinement.
- *Sub-block 3:* Utilized for all sub-blocks except the first one in each block, further enhancing feature representation.By combining these modules and sub-blocks in a specific manner, a balance between model size, computational cost, and performance is achieved by EfficientNet-B0. It is noteworthy that EfficientNet-B0 comprises 237 layers. Thus, EfficientNet models excel in various image classification tasks while remaining computationally efficient. The basic architecture for the EfficientNet Model can be in Fig. 4.

Hyperparameters

Hyperparameters are pivotal components in machine learning and deep learning algorithms, as they govern the learning process and influence the resultant model parameters. Distinguished by the prefix 'hyper_', these parameters serve as high-level controls that shape the learning trajectory and ultimately determine the characteristics of the trained model.

Unlike model parameters, which are learned during the training process and directly influence the model's predictions, hyperparameters remain external to the resulting model. They are integral to the learning algorithm's functionality, yet they do not become ingrained within the model structure.

Some notable examples of hyperparameters encompass a diverse array of settings that profoundly impact the learning process and model performance:

- *Train-Test Split Ratio:* Dictates the proportion of data allocated for training versus testing, influencing the model's ability to generalize to unseen data.
- *Learning Rate:* A crucial parameter in optimization algorithms like gradient descent, controlling the magnitude of parameter updates during training.
- *Choice of Optimization Algorithm:* Determines the approach used to minimize the model's loss function, with options including gradient descent, stochastic gradient descent, or advanced techniques like the Adam optimizer.
- *Choice of Activation Function:* Pertains to the nonlinear transformation applied within neural network layers, with popular functions including Sigmoid, ReLU, and Tanh, impacting the network's capacity to capture complex patterns.
- *Cost or Loss Function:* Defines the objective function optimized during training, guiding the model towards minimizing prediction errors or maximizing performance metrics.
- *Number of Hidden Layers and Activation Units:* Crucial architectural decisions in neural networks, influencing the network's depth, breadth, and expressive capacity.
- *Dropout Rate:* Specifies the probability of randomly dropping neurons during training, a regularization technique aimed at preventing overfitting.
- *Number of Training Iterations (Epochs):* Specifies how many times the model will pass through the full training dataset, influencing both convergence rate and model stability.
- *Number of Clusters:* Applicable in clustering tasks, this parameter impacts the detail and organization of the clusters formed.
- *Kernel or Filter Size in Convolutional Layers:* Determines the scope of convolution operations, essential for capturing features from input data within convolutional neural networks.
- *Pooling Size:* Sets the dimensions of pooling areas in CNNs, affecting the downsampling of features and the development of spatial hierarchies.
- *Batch Size:* Defines the quantity of samples processed per training iteration, impacting both computational efficiency and the accuracy of gradient estimation. By judiciously tuning these hyperparameters, practitioners can optimize the learning process, enhance model performance, and achieve superior results across a diverse range of machine learning and deep learning tasks. Table 5 gives a summary for Hyperparameters for different Machine Learning Algorithms

Flatten

The Flatten method within the layers module is considered a crucial tool for reshaping data, particularly when transitioning from convolutional layers to fully connected layers in a neural network. This method is essentially
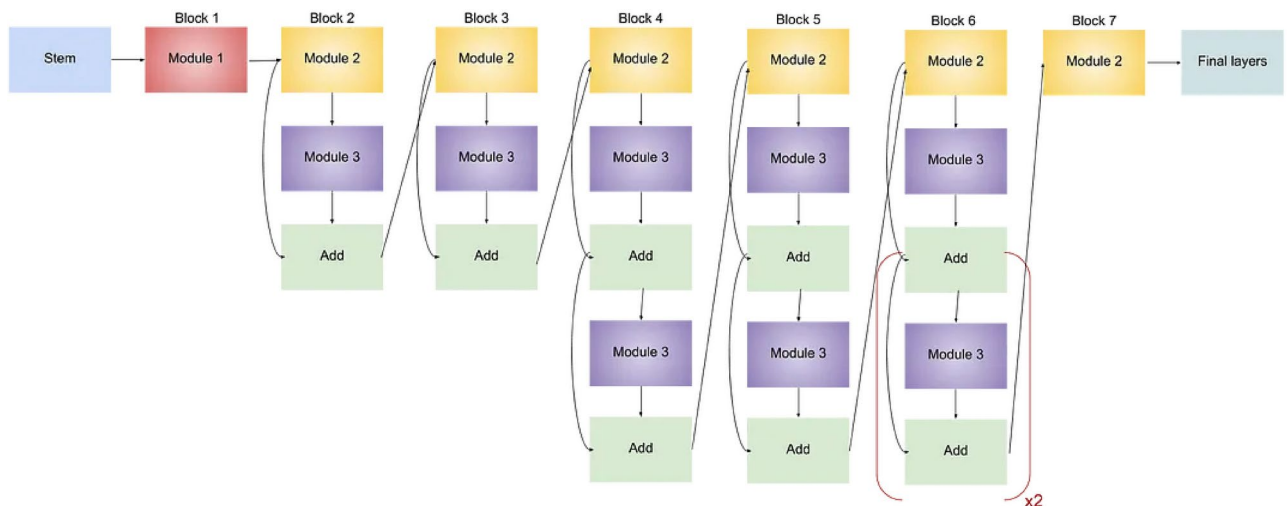

Fig. 4. Architecture for EfficientNet-B0[45].

ML algorithm	Main HPs	Optional HPs	HPO methods	Libraries
Ridge & lasso	Alpha	–	BO-GP	Skopt
Logistic regression	Penalty, c, solver	–	BO-TPE, SMAC	Hyperopt, SMAC
KNN	n_neighbors	Weights, p, algorithm	BOs, Hyperband	Skopt, Hyperopt, SMAC, Hyperband
SVM	C, kernel, epsilon (for SVR)	Gamma, coef0, degree	BO-TPE, SMAC, BOHB	Hyperopt, SMAC, BOHB
NB	Alpha	–	BO-GP	Skopt
DT	Criterion, max_depth, min_samples_split, min_samples_leaf, max_features, splitter, min_weight_fraction_leaf, max_leaf_nodes	–	GA, PSO, BO-TPE, SMAC, BOHB	TPOT, Optunity, SMAC, BOHB
RF & ET	n_estimators, max_depth, criterion, min_samples_split, min_samples_leaf, max_features, splitter, min_weight_fraction_leaf, max_leaf_nodes	–	GA, PSO, BO-TPE, SMAC, BOHB	TPOT, Optunity, SMAC, BOHB
XGBoost	n_estimators, max_depth, learning_rate, subsample, colsample_bytree, min_child_weight, gamma, alpha, lambda	–	GA, PSO, BO-TPE, SMAC, BOHB	TPOT, Optunity, SMAC, BOHB
Voting	Estimators, voting weights	–	GS	Sklearn
Bagging	Base_estimator, n_estimators	max_samples, max_features	GS, BOs	Sklearn, Skopt, Hyperopt, SMAC
AdaBoost	Base_estimator, n_estimators, learning_rate	–	BO-TPE, SMAC	Hyperopt, SMAC
Deep learning	Number of hidden layers, 'units' per layer, loss, optimizer, Activation, learning_rate, dropout rate, epochs, batch_size, early stop patience, number of frozen layers (if transfer learning is used)	–	PSO, BOHB	Optunity, BOHB
Hierarchical clustering	n_clusters, distance_threshold	Linkage	BOs, Hyperband	Skopt, Hyperopt, SMAC, Hyperband
DBSCAN	eps, min_samples	–	BO-TPE, SMAC, BOHB	Hyperopt, SMAC, BOHB
Gaussian mixture	n_components	covariance_type, max_iter, tol	BO-GP	Skopt
PCA	n_components	svd_solver	BOs, Hyperband	Skopt, Hyperopt, SMAC, Hyperband
LDA	n_components	solver, shrinkage	BOs, Hyperband	Skopt, Hyperopt, SMAC, Hyperband

Table 5. Summary of hyperparameters and hyperparameter optimization (HPO) methods for various machine learning algorithms.

utilized to collapse or flatten the input tensor into a one-dimensional tensor, which is deemed essential for it to be fed into a dense layer.

When dealing with convolutional layers, the output typically comprises a multi-dimensional tensor, where each dimension represents different features extracted by the convolutional filters. However, fully connected layers require a one-dimensional input, where each element corresponds to a single feature. This is where the flatten method proves to be handy.

Below are few of the most important features for flatten:

- *Input Tensor:* A multi-dimensional tensor, typically the output of a convolutional layer, serves as the starting point.
- *Flattening Operation:* The flatten method reshapes the tensor by concatenating all the elements along all dimensions except the batch dimension.
- *Output Tensor:* The result is a one-dimensional tensor that can be directly fed into a dense layer. Figure 5 represents and displays a standard Flatten Layer with inter-layer connections.

Dense

A dense layer, also referred to as a fully connected layer, is a fundamental component in neural network architectures, especially in feedforward networks. In this layer, each neuron is linked to every neuron in the subsequent layer, enabling comprehensive connectivity between layers.

In TensorFlow, adding a dense layer to a neural network model requires specifying the number of neurons or units in that layer, which defines the output space's dimensionality. Each neuron in a dense layer receives input from all neurons in the preceding layer, and it computes an output by applying a weighted sum to these inputs, followed by an activation function. Figure 6 illustrates the typical structure of a dense layer.

The parameters of a dense layer include weights and biases. The strength of the connections between neurons is represented by the weights, and offsets are learned by the model through biases to better fit the data.

Dense layers are versatile and can be used in various neural network architectures for tasks such as classification, regression, and even unsupervised learning. They're often stacked together with activation functions like ReLU (Rectified Linear Unit) or sigmoid to introduce non-linearity into the model, enabling it to learn complex patterns and relationships in the data.
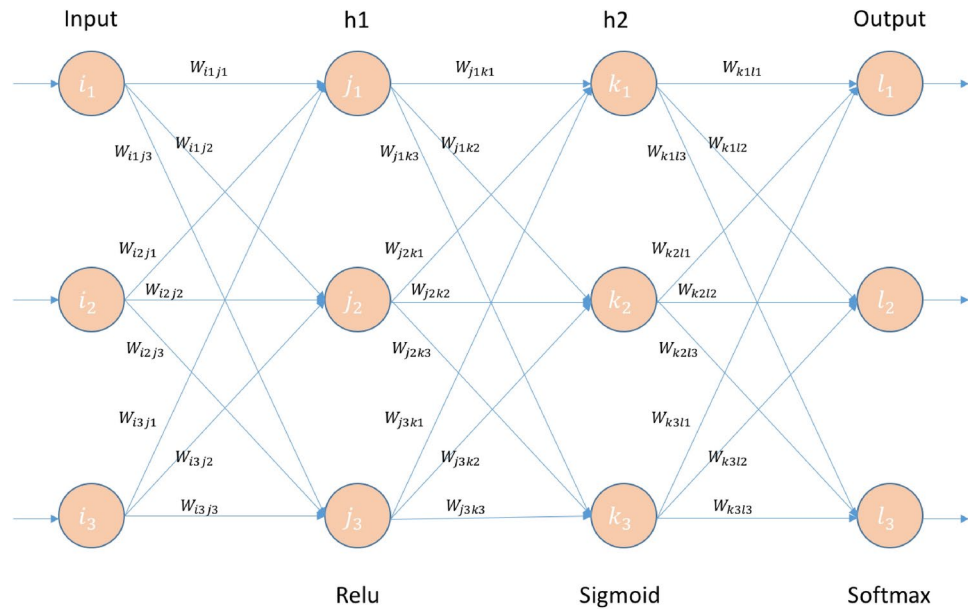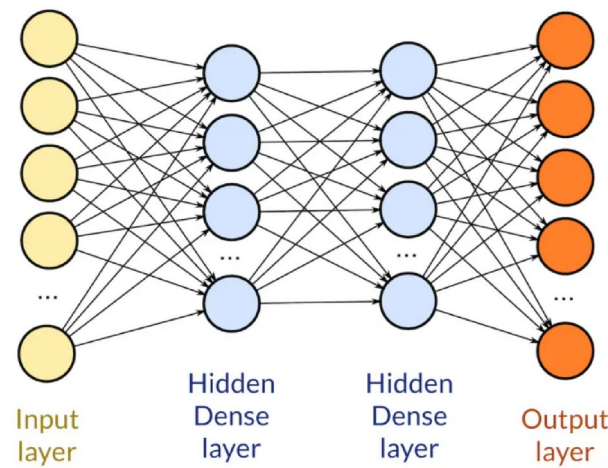
Fig. 5. Standard flatten layer.

Fig. 6. Standard dense layer.

Proposed framework

The proposed framework leverages the EfficientNetB0 model to categorize the severity levels of diabetic retinopathy from retinal images. EfficientNetB0 was chosen due to its exceptional accuracy in image classification tasks and its computational efficiency, achieved via a compound scaling technique that optimally adjusts the model's depth, width, and resolution. In this framework, the upper layers of a pre-trained EfficientNetB0 model (pre-trained on ImageNet) were removed by setting include_top=False, and a new dense layer with five neurons was incorporated to classify images into five distinct diabetic retinopathy severity categories. A softmax activation function is applied to the final layer to produce probability distributions across these classes[8].

The methodology employs an algorithmic sequence: preprocessing starts by scaling pixel values to the [0,1] range and implementing data augmentation strategies like random horizontal flips, rotations, and zooms. The EfficientNetB0 model, initialized with ImageNet weights, is then enhanced by adding a flatten layer and a dense layer with softmax activation for classification purposes. Training is fine-tuned with a custom learning rate scheduler and early stopping triggered by validation loss, leading to a model that is both accurate and computationally efficient. This holistic framework delivers optimal performance in diabetic retinopathy classification.

1: **Input:** Preprocessed fundus images, Diabetic Retinopathy severity labels
2: **Output:** Trained EfficientNet model for diabetic retinopathy classification
3: **Preprocessing:**
4:      Normalize input images by rescaling pixel values to the [0, 1] range.
5:      Apply data augmentation techniques to improve generalization, including:
6:   Random horizontal flips
7:   Random rotations of up to 10%
8:   Random zooms of up to 10%
9: **Model Architecture:**
10:      Import EfficientNetB0 as the base model with pre-trained ImageNet weights and without the top layers.
11:      Add layers to complete the model:
12:   Flatten layer to convert 2D feature maps to 1D.
13:   Dense layer with 5 neurons and softmax activation for five-class diabetic retinopathy classification.
14: **Custom Learning Rate Scheduler:**
15:      Define a custom learning rate scheduler to reduce the learning rate by a factor of 0.5 every 3 epochs of no improvement.
16: **Compile Model:**
17:      Compile the model with the Adam optimizer.
18:      Specify SparseCategoricalCrossentropy as the loss function to handle multi-class classification.
19:      Use accuracy as the evaluation metric.
20: **Training:**
21:      Train the model on augmented data for 40 epochs.
22:      Incorporate the custom learning rate scheduler callback to optimize learning.
23:      Monitor validation loss to apply early stopping if required.
24: **Model Evaluation:**
25:      After training, evaluate the model performance on the validation dataset using accuracy and loss metrics.
26: **Output:** Trained EfficientNet model with optimal learning rate adjustments and augmented dataset

Algorithm 1. EfficientNet-Based Model for Diabetic Retinopathy Detection

Model architecture

The EfficientNetB0 model was chosen for this study due to its proven high performance and computational efficiency in image classification tasks. It employs a compound scaling strategy that balances the network's depth, width, and resolution, ensuring both accuracy and efficiency. In this setup, the top layers of the pre-trained EfficientNetB0 (trained on ImageNet) were removed (include_top=False), and a new dense layer with five neurons was added, each representing a diabetic retinopathy severity level. The softmax activation function in this final layer generates probability distributions across these classes. The model's total parameter count was 4,049,571, of which 4,007,548 were trainable, enabling efficient learning from the retinal images while keeping computational demands low.

The training process used the Adam optimizer for its adaptability in learning rates. Additionally, a custom learning rate scheduler was applied, reducing the rate by half every three epochs to mitigate overfitting. Sparse
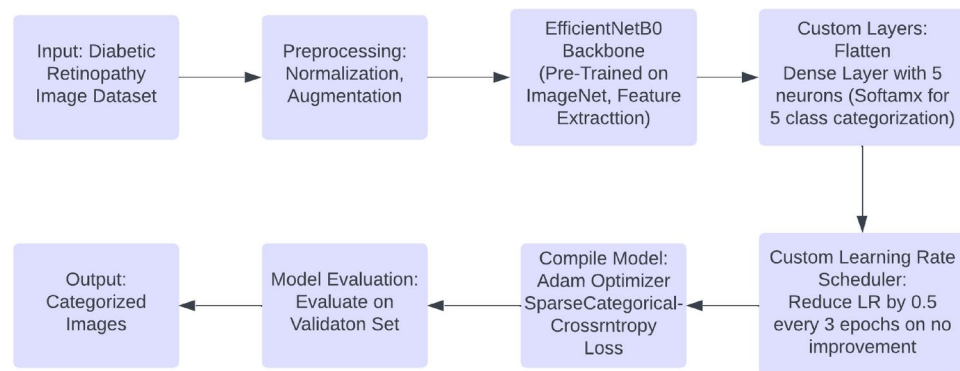

Fig. 7. Proposed algorithm flow for EfficientNetB0.

Categorical Cross-Entropy was employed as the loss function, suitable for multi-class classification tasks. The model was trained over 40 epochs with an initial learning rate of 0.001 and a batch size of 32, utilizing data augmentation techniques like random flips, rotations, and zooms. Figure 7 illustrates the model's flow and provides a basic architectural overview of the proposed solution.

Experimental setup

The experimental setup was designed with a focus on improving model generalization through data augmentation and careful tuning of hyperparameters. Data augmentation was applied using Keras' ImageDataGenerator, which included horizontal flipping, random rotations of 0.1 radians, and random zooming. The model was trained with a batch size of 32, using images resized to 224 × 224 pixels. A total of 40 epochs were completed, yielding an accuracy of 0.8653 and a loss of 0.5663, as observed in the training logs.

Dataset description

The dataset utilized in this study is sourced from Kaggle[46]. It consists of 35,108 retinal images, classified into five categories: No Diabetic Retinopathy (DR), Mild DR, Moderate DR, Severe DR, and Proliferative DR. The distribution of images is highly imbalanced, with 25,802 images labeled as 'No DR', 5288 as 'Mild DR', 2438 as 'Moderate DR', 872 as 'Proliferative DR', and 708 as 'Severe DR'. The images were captured at different resolutions, but for the purposes of this study, they were resized to 224 × 224 pixels to ensure uniform input dimensions across the dataset. Preprocessing included the normalization of pixel values to a [0,1] range and data augmentation techniques such as random horizontal flipping, zooming, and rotation to introduce variability and enhance model generalization.

Data preprocessing

The dataset presented a notable imbalance in class distribution, with the dominant class (No DR) representing over 70% of the total images. To rectify this disparity, undersampling was employed through the RandomUnderSampler function from the imbalanced-learn library. This adjustment reduced the count of 'No DR' images to match the sample size of the minority class (Severe DR), achieving a balanced dataset with 3704 images. This balancing allowed the model to learn equally from each class, minimizing the likelihood of bias toward the majority class during training. The balanced dataset was then divided into 2963 images for training and 741 images for validation, preserving the balanced class distribution in both subsets.

As part of essential preprocessing, data normalization was applied to standardize pixel values across input images. Using TensorFlow's ImageDataGenerator, pixel values were scaled to a range of [0,1] by dividing each pixel by 255. To further enhance the training data, data augmentation techniques were incorporated, including random horizontal flips, rotations up to 0.1 radians, and zooms up to 10%. These augmentations introduced minor variations to the training images, aiding in reducing overfitting and enhancing the model's generalization to new data.

Deep CNN model for comparison

In addition to the EfficientNetB0 model, the deep CNN model from Luo et al.[25] was referenced for comparative performance assessment. Luo's model[25] architecture leverages a multi-scale feature fusion approach specifically designed to improve diabetic retinopathy detection by capturing both local and global retinal features, which is particularly beneficial for identifying subtle pathologies. This model comprises multiple convolutional layers, each followed by pooling operations, and integrates attention mechanisms to focus on essential image areas for enhanced diagnostic accuracy. Additionally, Luo et al.[25] optimized the model by applying multi-scale feature fusion layers, allowing effective feature extraction at different spatial resolutions.

While our EfficientNetB0 model uses data augmentation techniques similar to those described in Luo's study[25], including random rotations, flips, and zooms, it was trained with a larger batch size of 32 over 40 epochs. The model was optimized using the Adam optimizer and Sparse Categorical Cross-Entropy as the loss function. By comparing our EfficientNetB0 model's performance to Luo et al.[25] CNN model, we aim to evaluate both accuracy and computational efficiency in diabetic retinopathy detection. This comparison highlights EfficientNetB0's balance of computational efficiency and robustness, even as Luo's model[25] demonstrates strong feature extraction capabilities through multi-scale processing, enhancing its performance in identifying nuanced diabetic retinopathy features.

Hyperparameters used

Table 6 below gives a summary for Hyperparameters and their optimization methods for numerous Machine Learning algorithms[47].

S. no.	Category	Model used	Parameter	Input value
1	Sequential	EfficientNet B0	input_shape	(224,224,3)
2	Sequential	EfficientNet B0	weights	imagenet
3	Sequential	EfficientNet B0	include_top	False
4	Layers	Dense	Number of neurons	5
5	Layers	Dense	activation	softmax

Table 6. Hyperparameters used in model training.

Model evaluation

The model's performance is rigorously assessed to determine its diagnostic precision. Evaluation is conducted using the testing dataset, which was set aside after data partitioning.

Precision (P_c) is defined as:

$$P_c = \frac{\text{TruePositives}_c}{\text{TruePositives}_c + \text{FalsePositives}_c} \tag{1}$$

Recall (R_c) is defined as:

$$R_c = \frac{\text{TruePositives}_c}{\text{TruePositives}_c + \text{FalseNegatives}_c} \tag{2}$$

The F1 Score calculation for a specific class (c) is defined as:

$$F1_c = \frac{2 \cdot P_c \cdot R_c}{P_c + R_c} \tag{3}$$

Experimental results

The EfficientNetB0 model achieved a maximum accuracy of 97.11% with a corresponding loss of 0.1596 after training for 40 epochs, incorporating data augmentation techniques. Data augmentation, including random flips, rotations, and zooms, provided a slight improvement in validation performance, as reflected by an average accuracy of 86.53 and an average loss of 0.5663, with 95% confidence intervals for accuracy and loss at (0.8677, 0.8677) and (0.5529, 0.5529), respectively. During training, accuracy consistently improved, reaching a final maximum training accuracy of 96.87% without data augmentation and 97.11% with augmentation as seen in Fig. 8. This increase suggests that data augmentation contributes minor but meaningful gains in generalization. Despite some variability in validation accuracy, EfficientNetB0 demonstrated competitive performance in diabetic retinopathy classification, indicating its effectiveness in handling this classification task. Table 7 summarizes the results as well as shows the comparison with a previous research.

After 40 epochs, the EfficientNetB0 model achieved an average training accuracy of 86.53% with data augmentation applied. In comparison, Luo et al.[25] multi-scale feature fusion method, although computationally intensive, demonstrated robustness in diabetic retinopathy classification, achieving 83.6% accuracy. Luo's model[25] effectively captures local and global retinal features, a valuable capability for nuanced feature identification. However, EfficientNetB0 maintains a high accuracy while offering significantly lower computational demand (processing time of 1190 seconds versus 4778 seconds for Luo's model)[25], making it highly suitable for resource-limited environments.

Both the EfficientNetB0 and Luo et al.[25] multi-scale feature fusion model were evaluated on diabetic retinopathy classification tasks, with each presenting distinct strengths across key performance metrics. The EfficientNetB0 model, leveraging pre-trained weights and optimized architecture, achieved an impressive maximum accuracy of 97.11% and demonstrated higher computational efficiency. Luo's model[25], while excelling in feature extraction at multiple spatial scales with a precision of 81.9%, incurs a higher processing cost. This comparison highlights that while both models effectively handle retinal images, EfficientNetB0 offers an optimal balance between computational efficiency and accuracy.

Discussions and limitations

One significant limitation of this study is the potential bias introduced by the dataset. The 'Diagnosis of Diabetic Retinopathy' dataset, though substantial, may lack diversity in patient demographics such as age, gender, and ethnic background, potentially affecting the model's generalizability to other populations. This lack of diversity could result in a model that performs well on certain demographic groups but may yield lower accuracy when applied to more diverse or distinct populations, limiting its broader clinical applicability.

Additionally, while undersampling was used to address class imbalance, this technique may not fully eliminate bias, as it reduces the sample size of the dominant classes rather than enriching minority class representation. This could lead to skewed predictions, particularly in underrepresented classes, possibly resulting in lower diagnostic reliability for certain stages or categories of diabetic retinopathy. Future research should consider incorporating datasets that represent a wider range of patient characteristics to enhance model robustness across different demographics, potentially mitigating these biases.

Furthermore, while this study used a balanced dataset to assess model performance, its applicability to real-world, naturally imbalanced data remains unexplored. In a real clinical setting, diabetic retinopathy severity levels are unevenly distributed, with milder stages being far more common. A model's performance on such imbalanced data could vary significantly, as it may struggle to detect rare but clinically important cases. Thus, exploring model performance on unbalanced datasets is crucial, as it could reveal further tuning needs to ensure reliable predictions across all severity levels in practical applications.

Conclusion and future scope

This paper proposed a deep learning framework for diagnosing diabetic retinopathy using retinal images. The proposed approach combines the efficiency of EfficientNet with the interpretability of layrering methods to achieve accurate and transparent diagnosis. The experimental results demonstrate that EfficientNetB0, with its compound scaling and pre-trained weights, significantly achieves high accuracy and robustness in
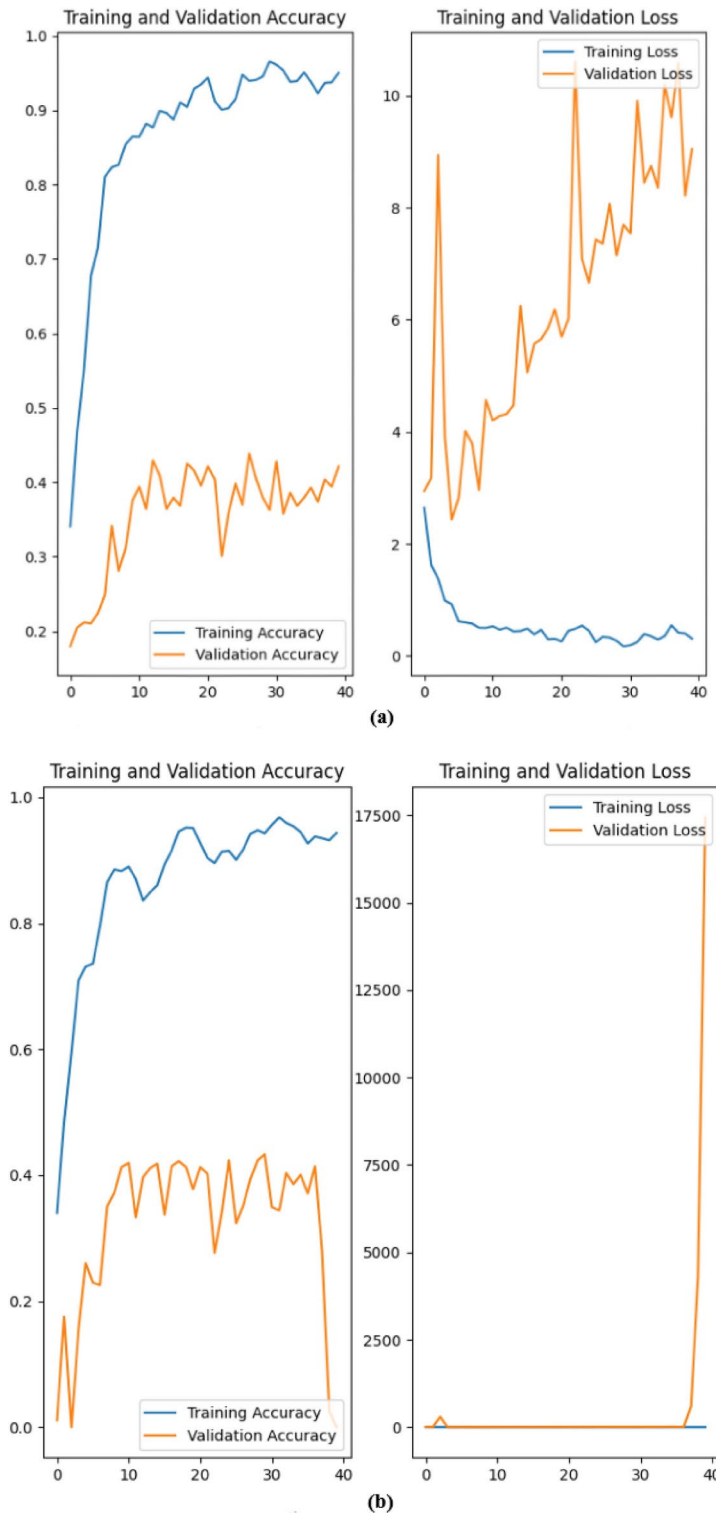
Fig. 8. Results achieved: (**a**) Using EfficientNetB0 without data augmentation; (**b**) Using EfficientNetB0 with data augmentation.

DR classification. The model's average performance metrics-accuracy of 0.8653 and loss of 0.5663 with data augmentation-further support its efficacy, while EfficientNetB0's 97.11% maximum accuracy, combined with lower computational demands, indicates its practical viability for real-world clinical applications, especially when computational resources are limited. In contrast, Luo et al.[25] approach, while effective in feature extraction, may be more suitable for scenarios where detailed feature mapping and high sensitivity are essential, albeit at a higher computational cost.

Model		Accuracy	Loss
EfficientNetB0	Without data augmentation	0.8622	0.5783
	With data augmentation	0.8653	0.5663
Deep CNN Model	–	0.836	–

Table 7. Average accuracy with corresponding loss.

There are several avenues for future work that can build upon the findings of this study. An immediate area for improvement involves expanding the dataset by integrating clinical data, such as patient history, blood glucose levels, and the presence of comorbidities like hypertension and cardiovascular diseases. Such additional information could provide a more comprehensive view of each patient's health profile[48], enabling more precise predictions of diabetic retinopathy progression. Moreover, while this study focused on retinal images, future research could explore multimodal approaches that combine image data with structured medical data to improve both diagnostic accuracy and model robustness.

In terms of model development, investigating more advanced architectures such as Vision Transformers (ViTs) or hybrid models incorporating attention mechanisms could further boost diagnostic accuracy. Vision Transformers have shown great promise in other computer vision tasks and could be adapted for retinal image analysis. Another key direction would be optimizing the model for real-time deployment in clinical settings, focusing on reducing inference time and computational requirements without compromising accuracy. Finally, longitudinal studies that monitor patients over time could yield insights into the model's ability to predict disease progression, rather than simply classifying the current stage of diabetic retinopathy. Such advancements could significantly enhance the clinical utility of diabetic retinopathy classification models.

Data availibility

The datasets analyzed during the current study are publicly available in the Kaggle repository under the title "Diabetic Retinopathy Resized Dataset"[46] at https://www.kaggle.com/datasets/tanlikesmath/diabetic-retinopathy-resized/data.

References

1. Stitt, A. W. et al. The progress in understanding and treatment of diabetic retinopathy. *Prog. Retin. Eye Res.* **51**, 156–186. https://doi.org/10.1016/j.preteyeres.2015.08.001 (2016).
2. Gadekallu, T. et al. Deep neural networks to predict diabetic retinopathy. *J. Ambient Intell. Hum. Comput.* **14**, 5407–5420. https://doi.org/10.1007/s12652-020-01963-7 (2023).
3. Silva, P. S. et al. Automated machine learning for predicting diabetic retinopathy progression from ultra-widefield retinal images. *JAMA Ophthalmol.* **142**, 171–178. https://doi.org/10.1001/jamaophthalmol.2023.6318 (2024).
4. Grzybowski, A. et al. Artificial intelligence for diabetic retinopathy screening using color retinal photographs: From development to deployment. *Ophthalmol. Ther.* **12**, 1419–1437. https://doi.org/10.1007/s40123-023-00691-3 (2023).
5. Vijayan, M. A regression-based approach to diabetic retinopathy diagnosis using efficientnet. *Diagnostics* **13**, 774. https://doi.org/10.3390/diagnostics13040774 (2023).
6. Ariza López, L. & Ramos, S. Diabetic retinopathy diagnosis using deep learning (2023).
7. Pramudhita, D. A., Azzahra, F., Arfat, I. K., Magdalena, R. & Saidah, S. Strawberry plant diseases classification using CNN based on MobileNetV3-large and efficientnet-B0 architecture. *Jurnal Ilmiah Teknik Elektro Komputer dan Informatika JITEKI* **9**, 522–534. https://doi.org/10.26555/jiteki.v9i3.26341 (2023).
8. Alhijaj, J. A. & Khudeyer, R. S. Integration of efficientnetb0 and machine learning for fingerprint classification. *Informatica*[SPACE]https://doi.org/10.31449/inf.v47i5.4724 (2023).
9. Gulshan, V. et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *JAMA* **316**, 2402–2410 (2016).
10. Ting, D. S. W. et al. Development and validation of a deep learning system for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes. *JAMA* **318**, 2211–2223 (2017).
11. Abràmoff, M. D., Lavin, P. T., Birch, M., Shah, N. & Folk, J. C. Pivotal trial of an autonomous ai-based diagnostic system for detection of diabetic retinopathy in primary care offices. *NPJ Digital Med.* **1**, 1–8 (2018).
12. Li, Z. et al. An automated grading system for detection of vision-threatening referable diabetic retinopathy on the basis of color fundus photographs. *Diabetes Care* **42**, 1584–1589 (2019).
13. Bellemo, V. et al. Artificial intelligence using deep learning to screen for referable and vision-threatening diabetic retinopathy in Africa: A clinical validation study. *Lancet Digit. Health* **1**, e35–e44 (2019).
14. Bhaskaranand, M. et al. Automated diabetic retinopathy screening and monitoring using retinal fundus image analysis. *J. Diabetes Sci. Technol.* **13**, 438–446 (2019).
15. Sahlsten, J. et al. Development and validation of a deep learning algorithm for diabetic retinopathy and related eye diseases using retinal images from multiethnic populations with diabetes: A population-based study. *Acta Ophthalmol.* **98**, e20–e28 (2020).
16. Raman, R. et al. Development and validation of a deep learning algorithm for detection of diabetic retinopathy using non-mydriatic retinal fundus images archived in electronic health records. *Indian J. Ophthalmol.* **68**, 398–403 (2020).
17. Takahashi, H. et al. Application of deep learning to the determination of diabetic retinopathy and glaucoma based on retinal fundus photographs. *Jpn. J. Ophthalmol.* **64**, 368–375 (2020).
18. Burlina, P. M. et al. Automated grading of age-related macular degeneration from color fundus images using deep convolutional neural networks. *JAMA Ophthalmol.* **138**, 652–659 (2020).
19. Huang, K., Zhang, L., Chen, Y., Xie, J. & Li, L. Saliency-guided self-supervised transformer for diabetic retinopathy grading. *IEEE Trans. Med. Imaging*[SPACE]https://doi.org/10.1109/TMI.2024.3045002 (2024).

20. Thanikachalam, R., Sivakumar, M., Kalyani, M. & Kumar, H. Optimized deep convolutional neural networks for diabetic retinopathy and macular edema detection. *Comput. Biol. Med.*[SPACE]https://doi.org/10.1016/j.compbiomed.2024.105034 (2024).

21. Bodapati, S. & Balaji, K. Self-adaptive stacking ensemble with attention mechanisms for diabetic retinopathy severity prediction. *Artif. Intell. Med.*[SPACE]https://doi.org/10.1016/j.artmed.2024.101850 (2024).

22. Bhati, S., Singh, P. & Thakur, R. Idanet: Interpretable dual attention network for diabetic retinopathy grading. *IEEE Access*[SPACE]https://doi.org/10.1109/ACCESS.2024.3100089 (2024).

23. Sivapriya, R., Chithra, M. & Ragavendran, C. Microvascular structure analysis in diabetic retinopathy classification using deep learning. *IEEE J. Biomed. Health Inform.*[SPACE]https://doi.org/10.1109/JBHI.2024.3120031 (2024).

24. Ohri, P. & Kumar, R. Supervised fine-tuned approach for diabetic retinopathy detection using transfer learning. *J. Digit. Imaging*[SPACE]https://doi.org/10.1007/s10278-024-00609-1 (2024).

25. Luo, Z., Wang, F. & Zheng, Q. Deep CNN model with multi-scale feature fusion for diabetic retinopathy detection. *Expert Syst. Appl.*[SPACE]https://doi.org/10.1016/j.eswa.2024.117845 (2024).

26. Romero-Oraá, M. T., Martinez, J., Gonzalez, R. & Gonzalez, M. Attention-based framework for diabetic retinopathy grading in fundus images. *Biomed. Signal Process. Control*[SPACE]https://doi.org/10.1016/j.bspc.2024.105564 (2024).

27. Zhang, Y., Wang, Y., Wu, X. & Li, J. Semi-supervised contrastive learning with saliency maps for diabetic retinopathy classification. *Pattern Recogn.*[SPACE]https://doi.org/10.1016/j.patcog.2024.109164 (2024).

28. Wong, L., Yang, H. & Zheng, Y. Transfer learning with feature-weighted ECOC ensembles for diabetic retinopathy grading. *Comput. Biol. Med.*[SPACE]https://doi.org/10.1016/j.compbiomed.2023.105489 (2023).

29. Bilal, A., Zhu, L., Deng, A., Lu, H. & Wu, N. Ai-based automatic detection and classification of diabetic retinopathy using u-net and deep learning. *Symmetry* **14**, 1427. https://doi.org/10.3390/sym14071427 (2022).

30. Bilal, A., Sun, G., Li, Y., Mazhar, S. & Khan, A. Diabetic retinopathy detection and classification using mixed models for a disease grading database. *IEEE Access*[SPACE]https://doi.org/10.1109/ACCESS.2021.3056186 (2021).

31. Bilal, A., Mazhar, S., Imran, A. & Latif, J. A transfer learning and u-net-based automatic detection of diabetic retinopathy from fundus images. *Comput. Methods Biomech. Biomed. Eng. Imaging Visualiz.*[SPACE]https://doi.org/10.1080/21681163.2021.2021111 (2022).

32. Bilal, A., Sun, G., Mazhar, S. & Imran, A. Improved grey wolf optimization-based feature selection and classification using CNN for diabetic retinopathy detection. 1–14 (2022).

33. Bilal, A., Sun, G. & Mazhar, S. Diabetic retinopathy detection using weighted filters and classification using CNN. https://doi.org/10.1109/CONIT51480.2021.9498466 (2021).

34. Bilal, A. et al. Improved support vector machine based on CNN-SVD for vision-threatening diabetic retinopathy detection and classification. *PLoS One* **19**, e0295951. https://doi.org/10.1371/journal.pone.0295951 (2024).

35. Bilal, A., Liu, X., Baig, T., Long, H. & Shafiq, M. Edgesvdnet: 5g-enabled detection and classification of vision-threatening diabetic retinopathy in retinal fundus images. *Electronics* **12**, 4094. https://doi.org/10.3390/electronics12194094 (2023).

36. Bilal, A., Liu, X., Shafiq, M., Ahmed, Z. & Long, H. Nimeq-sacnet: A novel self-attention precision medicine model for vision-threatening diabetic retinopathy using image data. *Comput. Biol. Med.* **171**, 108099. https://doi.org/10.1016/j.compbiomed.2024.108099 (2024).

37. Gupta, B. B., Gaurav, A. & Panigrahi, P. K. Analysis of security and privacy issues of information management of big data in B2B based healthcare systems. *J. Bus. Res.* **162**, 113859 (2023).

38. Zaidan, A. A., AlSattar, H. A., Qahtan, S., Deveci, M. & Pamucar, D. Secure decision approach for internet of healthcare things smart systems-based blockchain. *IEEE Internet of Things Journal* (2023).

39. Zhou, Y., Song, L., Liu, Y. & Vijayakumar, P. A privacy-preserving logistic regression-based diagnosis scheme for digital healthcare. *Futur. Gener. Comput. Syst.* **144**, 63–73 (2023).

40. Singh, S. K. *Linux yourself: Concept and programming* 1st edn. (Chapman and Hall/CRC, Cham, 2021).

41. Chui, K. T. et al. Multiround transfer learning and modified generative adversarial network for lung cancer detection. *Int. J. Intell. Syst.* **2023**, 6376275 (2023).

42. Hammad, M., Abd El-Latif, A. A., Hussain, A. & Abd El-Samie, F. E. Deep learning models for arrhythmia detection in IoT healthcare applications. *Comput. Electr. Eng.* **100**, 108011 (2022).

43. Sutomo, H. I. Identification of organic and non-organic waste with computer image recognition using convolutionalneural network with efficient-net-b0 architecture. *J. Appl. Intell. Syst.* **8**(3), 320–330. https://doi.org/10.33633/jais.v8i3.9064 (2023).

44. Y, V., Billakanti, N., Veeravalli, K., N, A. D. R & Kota, L. Early detection of casava plant leaf diseases using efficientnet-b0. In *2022 IEEE Delhi Section Conference (DELCON)*, 1–5, https://doi.org/10.1109/DELCON54057.2022.9753210 (2022).

45. Agarwal, V. Complete architectural details of all efficientnet models. *Medium* (2021). https://towardsdatascience.com/complete-architectural-details-of-all-efficientnet-models-5fd5b736142.

46. TanLikesMath. Diabetic retinopathy - resized data. *Kaggle*, https://www.kaggle.com/datasets/tanlikesmath/diabetic R-retinopathy, https://doi.org/10.13140/RG.2.2.13037.19688-resized/data (20243).

47. Yang, L. & Shami, A. On hyperparameter optimization of machine learning algorithms: Theory and practice. *Neurocomputing* **415**, 295–316. https://doi.org/10.1016/j.neucom.2020.07.061 (2020).

48. Rathee, G., Saini, H., Garg, S., Choi, B. J. & Hassan, M. M. A secure data e-governance for healthcare application in cyber physical systems. *Int. J. Semantic Web Inf. Syst.* **20**, 1–17 (2024).

Acknowledgement

Author contributions

L.A.: Conceptualization, methodology, writing original draft. S.K.S.: Formal analysis, supervision, reviewing and editing, data interpretation, supervision. S.K.: Conceptualization, methodology, Data curation, validation, writing original draft, supervision. H.G.: Conceptualization, methodology, writing original draft, Software, visualization, data interpretation. W.A.: Investigation, resources, manuscript preparation. V.A.: Methodology, validation, writing – review and editing. S.B.: Data acquisition, software, writing – review and editing. K.T.C.: Formal analysis, data interpretation, supervision. B.B.G.: Supervision, resources, project oversight, funding acquisition.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to S.K.S. or B.B.G.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.