DATA AUGMENTATION-BASED DIABETIC RETINOPATHY CLASSIFICATION AND GRADING WITH THE DYNAMIC WEIGHTED OPTIMIZATION APPROACH

Reference NO. IJME 2532, DOI: 10.5750/sijme.v167iA2(S).2532

Uppalapati Srilakshmi, Department of CSE, Sridevi Womens Engineering College, Vattinagulapally, Gachibowli, Hyderabad, India, **K. Vinay Kumar**, Department of CSE (AI&ML), Kakatiya Institute of Technology and Science, Hanamkonda, Koukonda, Warangal, Telangana-506015, India, **Sirisha Korimilli,** Department of Electronics and Communication Engineering, Bonam Venkata Challamaya Institute of Technology and Science, Amalapuram, Andhrapradesh, India, **Shiramshetty Gouthami**, Department of Computer Science and Engineering, Jayamukhi Institute of Technological Sciences, Narsampet,Warangal, India, **Jose Mary Golamari**, Department of Computer Science and Engineering, Jayamukhi Institute of Technological Sciences, Narsampet, Warangal, India and **Putta Brundavani***, Department of ECE, RSR Engineering College, Kavali, SPSR Nellore - 524142 Andhra Pradesh, India.

*Corresponding author. Putta Brundavani (Email): brundavenky777@gmail.com

KEY DATES: Submission date: 17.09.2024; Final acceptance date: 19.03.2025; Published date: 30.04.2025

SUMMARY

Diabetic retinopathy is a vision-threatening complication of diabetes that affects the retina, the light-sensitive tissue at the back of the eye. This condition arises as a result of prolonged high blood sugar levels, which can damage the small blood vessels in the retina. Diabetic retinopathy typically progresses through different stages, starting with mild non-proliferative retinopathy, where small blood vessels in the retina become weakened and leak. The classification of diabetic retinopathy plays a fundamental role in assessing the effectiveness of treatment and monitoring the progression of the disease over time, ultimately contributing to the preservation of patients' vision and their overall quality of life. This research paper presents efficient technique for diabetic retinopathy (DR) classification and grading using data augmentation and a dynamic weighted optimization approach. The study contributes to the field of DR in several significant ways. Firstly, advanced data augmentation techniques are employed to generate diverse and representative features from retinal fundus images, enhancing the robustness and generalization capabilities of the models. Secondly, novel segmentation approaches, including multi-level Otsu thresholding and morphological operations, accurately localize and isolate affected regions in retinal images. Thirdly, innovative feature extraction and selection methods, such as Gray-Level Co-occurrence Matrix (GLCM) and dynamic Flemingo optimization, improve the selection of discriminative features for DR classification. Additionally, a novel cascaded voting ensemble deep neural network model is introduced, which combines the predictions of multiple learning algorithms to enhance classification performance. Lastly, the research addresses the grading of diabetic retinopathy by aligning the classification results with a standardized grading system, providing clinicians with accurate severity assessments for effective treatment decisions. Overall, this papers offers valuable insights and methodologies for improving the classification and grading of diabetic retinopathy, thereby contributing to the advancement of diagnosis and management strategies in the field.

KEY WORDS: Data augmentation, Diabetic retinopathy, Classification, Grading, Dynamic weighted optimization, Deep learning

1. INTRODUCTION

Diabetic retinopathy (DR) is a common complication of diabetes that affects the eyes and can lead to vision loss or blindness if left untreated. It is caused by damage to the blood vessels in the retina, the light-sensitive tissue at the back of the eye [1]. DR progresses through different stages of severity, ranging from mild to severe and proliferative forms. Early detection and accurate diagnosis of DR are crucial for effective treatment and prevention of vision loss. With the advancements in medical imaging technology, digital retinal fundus images have become an essential tool for diagnosing and monitoring DR. These images capture the structure of the retina and provide valuable information about the presence and severity of the disease [2]. Diabetic retinopathy (DR) presents several significant issues that impact its diagnosis, treatment, and management. One of the primary challenges is the late detection of the disease. In many cases, patients remain asymptomatic until the advanced stages of DR, making it difficult to initiate timely interventions [3]. This delay in detection can lead to irreversible damage to the retina and increased risk of vision loss. Therefore, raising awareness among individuals with diabetes about the potential risks and complications of DR is crucial. Another issue is the limited access to regular screenings and eye care services, particularly in underserved communities or remote areas. Lack of access to healthcare facilities,

trained ophthalmologists, and specialized equipment for retinal imaging can hinder early diagnosis and proper management of DR [4]. This issue is particularly prevalent in low-income countries and regions with limited healthcare infrastructure. Interpretation and consistency in grading and classification of DR is another challenge. The grading of DR severity is typically based on subjective assessment by ophthalmologists, leading to potential variations and inconsistencies in diagnoses [5]. The lack of standardized grading systems and uniform interpretation guidelines can impact the accuracy and reliability of DR assessments, affecting treatment decisions and patient outcomes. Additionally, the increasing prevalence of diabetes worldwide contributes to a growing burden of DR [6]. The rising number of individuals with diabetes places additional strain on healthcare systems, increasing the demand for diabetic retinopathy screenings, treatment, and long-term management. Meeting this demand requires adequate resources, trained healthcare professionals, and efficient healthcare delivery systems.

The complexity of DR management necessitates a multidisciplinary approach involving collaboration between ophthalmologists, endocrinologists, primary care physicians, and other healthcare providers [7]. Coordinating care and ensuring effective communication among different specialties can be challenging, potentially leading to fragmented care and suboptimal outcomes for patients. Addressing these issues requires a comprehensive approach that includes increasing awareness and education about DR, improving access to regular screenings and eye care services, implementing standardized grading systems, and fostering collaborative care models [8]. Additionally, advancements in technology, such as telemedicine and artificial intelligence-based image analysis, hold promise in enhancing the efficiency and accuracy of DR diagnosis and management.

Deep learning has revolutionized the field of diabetic retinopathy (DR) by offering powerful capabilities in image analysis, diagnosis, screening, and management [9]. Convolutional neural networks (CNNs), a key component of deep learning, have demonstrated exceptional performance in analyzing retinal fundus images for the detection and classification of DR. By automatically extracting relevant features from these images, deep learning models enable accurate identification of DR severity levels and differentiation between diabetic and non-diabetic retinas [10]. One of the significant advantages of deep learning in DR is its potential for automated screening. By leveraging deep learning algorithms, large-scale populations at risk of DR can be efficiently screened. These models analyze retinal images, identifying individuals who require further evaluation or referral to ophthalmologists. This automated screening process significantly enhances the scalability and efficiency of DR screening programs, particularly in regions with limited healthcare resources [11].

Moreover, deep learning techniques contribute to the early detection and monitoring of DR progression. Through analyzing longitudinal retinal images, deep learning models can identify subtle changes in the retina associated with the development and progression of DR. This early detection facilitates timely interventions and management strategies, potentially preventing or mitigating vision loss [12]. Deep learning-based segmentation models play a crucial role in accurately delineating retinal structures, such as blood vessels, optic discs, and lesions, in retinal images. Accurate segmentation is essential for quantifying disease-related changes, measuring anatomical features, and aiding in the diagnosis and monitoring of DR. Deep learning models have shown promising results in segmenting retinal structures with high accuracy and efficiency [13]. Additionally, deep learning enables personalized treatment and prognosis in DR. By integrating patient-specific data, including retinal images, clinical records, and genetic information, deep learning models can assist in predicting disease progression, identifying optimal treatment approaches, and tailoring management plans to individual patients [14]. This personalized approach enhances patient outcomes and optimizes resource allocation in healthcare settings. However, it is crucial to ensure the reliability, generalizability, and ethical considerations of deep learning models in DR. Rigorous validation and clinical evaluation are necessary before integrating these models into clinical workflows. Integration should prioritize collaboration with healthcare professionals to enhance decision-making and provide patient-centered care [15].The research makes several significant contributions to the field of diabetic retinopathy (DR):

1. The research performs data augmentation to generate diverse and representative features from retinal fundus images. By incorporating variations such as rotation, scaling, and noise, the augmented dataset captures a wider range of potential patterns and characteristics present in real-world retinal images. This improves the robustness and generalization capabilities of the models developed in the research.
2. The research explores and develops segmentation approaches, such as the Multi-level Otsu thresholding approach and Morphological operations, for accurately segmenting the affected regions in the retinal images. These segmentation techniques enable precise localization and isolation of the regions associated with diabetic retinopathy, providing valuable insights for analysis and diagnosis.
3. The research introduces the use of the Gray-Level Co-occurrence Matrix (GLCM) extracted features from retinal images. Additionally, it proposes a novel feature selection model called Dynamic Flemingo Optimization. These advancements contribute to improving the selection of discriminative features for DR classification, enhancing the accuracy and efficiency of the classification models.

4. The research introduces a novel deep learning model, the Cascaded Voting Ensemble Deep Neural Network model, which combines the predictions of multiple learning algorithms using a weighted average approach. This model leverages the collective knowledge and predictions of diverse algorithms, improving the overall performance and reliability of DR classification.
5. The research addresses the computation of grading in detected DR in fundus images, aligning the classification results with a standardized grading system. This enables clinicians to assess the severity of DR accurately and make informed decisions regarding appropriate treatment and management strategies.

2. RELATED WORKS

In recent years, there has been a growing body of research exploring the application of deep learning techniques in diabetic retinopathy (DR) detection and diagnosis. This section provides an overview of the related works in the field, highlighting the key contributions and advancements in deep learning-based approaches for DR. [16] proposed a multi-stage deep learning framework for DR grading using retinal fundus images. Their approach consisted of an initial lesion segmentation stage followed by a classification stage. The authors achieved promising results in terms of accuracy and demonstrated the potential of deep learning in accurately grading DR severity levels. [17] developed a novel deep convolutional neural network (CNN) architecture for DR detection. Their model utilized residual connections and attention mechanisms to improve feature extraction and classification performance. The proposed model achieved state-of-the-art results on benchmark DR datasets, highlighting the effectiveness of deep learning for DR detection.

[18] focused on the interpretation and explainability of deep learning models for DR diagnosis. They proposed a visualization technique based on gradient-weighted class activation mapping (Grad-CAM) to identify the regions of interest in retinal images that contributed most to the prediction. This work enhanced the interpretability of deep learning models and provided valuable insights for clinicians. [19] addressed the challenge of limited annotated data in DR research. They proposed a semi-supervised learning approach that leveraged a combination of labeled and unlabeled retinal images. Their method achieved competitive performance with reduced annotation effort, demonstrating the potential of deep learning in scenarios with limited labeled data. [20] explored the use of generative adversarial networks (GANs) for synthesizing realistic retinal images to augment the training data. The generated images were used to improve the generalization capability of deep learning models for DR classification. Their results showed improved performance when utilizing synthesized data in combination with real retinal images. [21] proposed a novel two-stage deep learning framework for DR diagnosis. Their approach consisted of a lesion segmentation stage followed by a classification stage. The authors achieved state-of-the-art results on a large DR dataset, demonstrating the potential of deep learning in accurately diagnosing DR.

[22] proposed a multi-scale feature fusion framework for DR grading using retinal fundus images. Their approach leveraged features extracted from multiple scales and achieved high accuracy in DR grading. This work demonstrated the importance of feature fusion for DR diagnosis using deep learning. [23] proposed a hybrid deep learning model for DR detection and classification. Their approach combined convolutional neural networks (CNNs) with long short-term memory (LSTM) networks to capture both spatial and temporal features from retinal images. The proposed model achieved high accuracy in DR detection and classification. [24] developed a deep learning-based system for automatic DR detection and severity grading. Their approach utilized a combination of CNNs and support vector machines (SVMs) to classify retinal images into different severity levels. The proposed system achieved high accuracy and showed promising results in clinical settings. [25] proposed a deep learning-based approach for early detection of diabetic retinopathy using fundus images. Their model utilized a combination of convolutional neural networks (CNNs) and recurrent neural networks (RNNs) to capture spatial and sequential information from retinal images. The approach achieved high accuracy in detecting early-stage diabetic retinopathy.

[26] developed a deep learning model for diabetic retinopathy detection and severity grading using a multi-task learning framework. Their approach simultaneously learned to classify the presence of diabetic retinopathy and predict the severity level. The proposed model showed improved performance compared to single-task models. [27] proposed a deep adversarial network for the segmentation of diabetic retinopathy lesions in fundus images. Their model integrated a U-Net architecture with generative adversarial networks (GANs) to accurately segment the affected regions. The approach demonstrated robust segmentation performance and provided valuable insights for diabetic retinopathy diagnosis. [28] presented a deep ensemble model for diabetic retinopathy detection and classification. The model combined multiple deep learning architectures, including convolutional neural networks (CNNs) and recurrent neural networks (RNNs), to leverage diverse representations. The ensemble approach achieved improved performance and robustness in classifying diabetic retinopathy. [29] proposed a novel deep learning-based method for diabetic retinopathy screening using a combination of fundus images and optical coherence tomography (OCT) scans. Their approach utilized a dual-modal fusion network to integrate information from both imaging modalities and achieved high accuracy in detecting diabetic retinopathy.

3. PROPOSED METHOD

The goal of this research is to develop an automated system for the analysis and classification of diabetic retinopathy using retinal fundus images. The research involves several key steps, starting with the loading of data from standard benchmark datasets consisting of retinal fundus images with annotations of DR severity. The dataset for the splitting of the dataset are presented in equation (1) and (2)

$$Training\ set: D_{train} = \{(X_{train}, y_{train})\} \quad (1)$$

$$Test\ set: D_{test} = \{(X_{test}, y_{test})\} \quad (2)$$

To ensure reliable training and evaluation of the classification models, the dataset is divided into training and test sets based on the severity levels of DR. This division allows us to assess the performance of the models in accurately predicting the severity of DR. To enhance the effectiveness of the models, data augmentation techniques are employed. These techniques involve generating new images with variations in factors such as rotation, scaling, and noise, to augment the original dataset. This augmentation helps to capture diverse features and improve the robustness of the models. Segmentation of the affected regions in the retinal images is crucial for accurate analysis and diagnosis of DR. The image rotation, scaling and noisy images of the DR images is represented as in equation (3)

$$Rotated\ Image = Rotate(Image, \theta) \quad (3)$$

$$Scaled\ Image = Scale(Image, scaling\_factor) \quad (4)$$

$$Noisy\ Image = Image + Noise \quad (5)$$

This research explores segmentation approaches such as Multi-level Otsu thresholding and Morphological operations to extract the affected regions from the retinal images. Additionally, this study focuses on feature extraction using the Gray-Level Co-occurrence Matrix (GLCM) and proposes a novel feature selection model called Dynamic Flemingo Optimization. The GLCM features capture texture information from the retinal images, while the proposed feature selection model optimizes the selection of the most discriminative features for DR classification [30–31]. To further enhance the accuracy and reliability of the classification, a novel Cascaded Voting Ensemble Deep Neural Network model is designed and developed. This model combines the predictions of multiple learning algorithms using a weighted average approach, resulting in improved classification performance with augmented images and GLCM features are computed using equation (6) and (7)

$$Augmented\ image: X_{augmented} = f(X_{Original}) \quad (6)$$

$$GLCM(P,\theta,d) = \sum\sum I(i,j) * I(i+p, j+q) \quad (7)$$

The base models employed in this research include convolutional neural networks (CNN) with max pooling, complete convolution layers, and networks in networks. These models are capable of automatically extracting relevant features from the retinal images and classifying them as diabetic retinopathy or non-diabetic. Finally, the grading of the detected DR in the retinal fundus images is computed based on the classification results. This grading system provides a standardized assessment of the severity of DR, enabling clinicians to make informed decisions regarding treatment and management strategies [32–33]. This study aims to contribute to the development of an accurate and automated system for diabetic retinopathy analysis and classification, ultimately improving the diagnosis and treatment of this sight-threatening condition. The prediction of DR is performed with ensemble model as presented in equation (8)

$$Ensemble\ Prediction: P_{ensemble}(y \mid x) = \sum(w_i * P_i(y \mid x)) \quad (8)$$

In equation (8) $P(y|x)$ is the predicted probability of a certain class for a given input x.

The proposed method in the research aims to develop an automated system for the analysis and classification of diabetic retinopathy (DR) using retinal fundus images. The method involves several key steps:

Data Loading: The researchers load data from standard benchmark datasets consisting of retinal fundus images that are annotated with the severity levels of DR. These datasets provide a reliable source of training and evaluation data.

Training and Test Data Creation: The dataset is divided into training and test sets based on the severity levels of DR. This division allows for the assessment of the models' performance in accurately predicting the severity of DR.

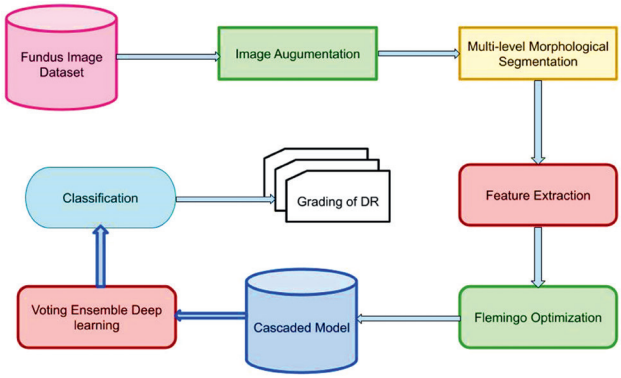

Figure 1. Flow chart of proposed model

Data Augmentation: To enhance the effectiveness of the models, data augmentation techniques are applied. These techniques generate new images with variations in rotation, scaling, noise, and other factors. Augmenting the dataset helps capture diverse features and improves the robustness of the models.

Segmentation Approaches: Segmentation is performed to extract the affected regions in the retinal images. The researchers explore segmentation approaches such as Multi-level Otsu thresholding and Morphological operations to accurately identify and isolate the regions affected by DR.

Feature Extraction: The researchers utilize the Gray-Level Co-occurrence Matrix (GLCM) for feature extraction. GLCM features capture texture information from the retinal images, providing valuable insights for DR analysis. Additionally, they propose a novel feature selection model called Dynamic Flemingo Optimization to optimize the selection of the most discriminative features for DR classification.

Cascaded Voting Ensemble Deep Neural Network Model: A novel Cascaded Voting Ensemble Deep Neural Network model is designed and developed. This model combines the predictions of multiple learning algorithms using a weighted average approach. By leveraging the strengths of different algorithms, the model improves the classification performance for DR.

Grading Computation: The detected DR in the retinal fundus images is graded based on the classification results. This grading system provides a standardized assessment of the severity of DR, enabling clinicians to make informed decisions regarding treatment and management strategies.

The proposed method integrates various deep learning techniques, data augmentation, segmentation approaches, feature extraction, and ensemble learning to develop an accurate and automated system for DR analysis and classification. The aim is to improve the diagnosis and treatment of diabetic retinopathy, contributing to better patient care and outcomes.

3.1 DATA AUGMENTATION

Image augmentation techniques are commonly employed in diabetic retinopathy (DR) to enhance the dataset and improve the performance of deep learning models. These techniques involve applying various transformations to the retinal images, thereby increasing the diversity and variability of the training samples. One commonly used technique is horizontal and vertical flipping, which randomly mirrors the image along the horizontal and/or vertical axis, introducing different viewing angles. Rotation is another technique that randomly rotates the image by a certain angle, simulating variations in the orientation of retinal images. Scaling and resizing can be used to randomly change the size or resolution of the images, providing different zoom levels and adjusting the image size for better model training. Translation involves randomly shifting the image horizontally and/or vertically, mimicking slight misalignments or positional variations in the retina. By applying these image augmentation techniques, the dataset for DR classification becomes more diverse, enabling the deep learning models to learn robust and generalized representations of the disease.

3.2 MULTILEVEL MORPHOLOGICAL SEGMENTATION

In the context of augmented diabetic retinopathy (DR) images, multilevel morphological segmentation is a technique used to segment the retinal structures and lesions by applying morphological operations at multiple scales. This segmentation process allows for a more detailed and accurate analysis of the augmented DR images. By augmenting the dataset with various techniques, such as flipping, rotation, scaling, or translation, the diversity and variability of the images are increased. The augmented images are then preprocessed to enhance their quality and contrast if necessary. Subsequently, multilevel morphological operations, such as erosion, dilation, opening, or closing, are applied iteratively at different scales. This enables the segmentation of different structures and lesions present in the augmented images, capturing various levels of detail. The segmented regions of interest, such as the optic disc, blood vessels, exudates, or hemorrhages, can be extracted for further analysis or utilized in subsequent classification and diagnostic tasks. Through the integration of multilevel morphological segmentation with augmented DR images, a more comprehensive understanding of the retinal structures and abnormalities can be achieved, facilitating accurate diagnosis and monitoring of diabetic retinopathy.

Consider a binary image I(x, y), where (x, y) represents the pixel coordinates. The goal of multilevel morphological segmentation is to separate different structures or regions of interest in the image. Erosion ($\ominus$) and Dilation ($\oplus$): These are fundamental morphological operations that modify the shape and size of objects in an image. Opening ($\circ$): It is a combination of erosion followed by dilation and is useful for removing small objects and noise. Closing ($\star$): It is a combination of dilation followed by erosion and is useful for closing small gaps and filling holes. A structuring element is a small binary matrix or shape that defines the neighborhood around each pixel during the morphological operations. Common structuring elements include disks, squares, or lines, with varying sizes and orientations. The choice of structuring element depends on the specific characteristics and structures to be segmented in the augmented DR images. A binary image represents pixel coordinates (x, y) with values 0 or 1, where 0 typically

corresponds to the background, and 1 corresponds to objects or structures of interest stated as in equation (9).

$$(I \ominus B)(x, y) = min\{I(x-a, y-b) \mid (a,b) \text{ is in } B\} \quad (9)$$

Dilation operation enlarges or dilates the objects stated as in equation (10) – (12).

$$(I \oplus B)(x,y) = max\ \{I(x-a, y-b) \mid (a,b) \text{ is in } B\} \quad (10)$$

$$(I \circ B) = (I \ominus B) \oplus B \quad (11)$$

$$(I * B) = (I \oplus B) \ominus B \quad (12)$$

The multilevel aspect involves applying the morphological operations at multiple scales or levels. This can be achieved by varying the size of the structuring elements used in the operations. The size of the structuring element determines the level of detail captured during the segmentation process. After applying the morphological operations, a thresholding or decision rule can be employed to classify the resulting image into different regions or structures. The thresholding process may involve setting a specific intensity value or using adaptive techniques based on local characteristics of the image. The optimization process is presented in flow chart Figure 2.

3.4 CASCADED MODEL

A cascaded model for diabetic retinopathy (DR) typically refers to a multi-stage approach for the classification of DR using feature selection techniques. The objective is to progressively select relevant and discriminative features from retinal images to improve the accuracy of DR classification.

Stage 1: Preprocessing and Initial Feature Extraction

In the first stage, the retinal images are preprocessed to enhance their quality and remove noise or artifacts. Initial feature extraction techniques, such as image processing operations or handcrafted feature descriptors, are applied to extract a wide range of features from the retinal images. The initial feature set includes a combination of structural, textural, and statistical features computed from various regions of interest within the retinal images.

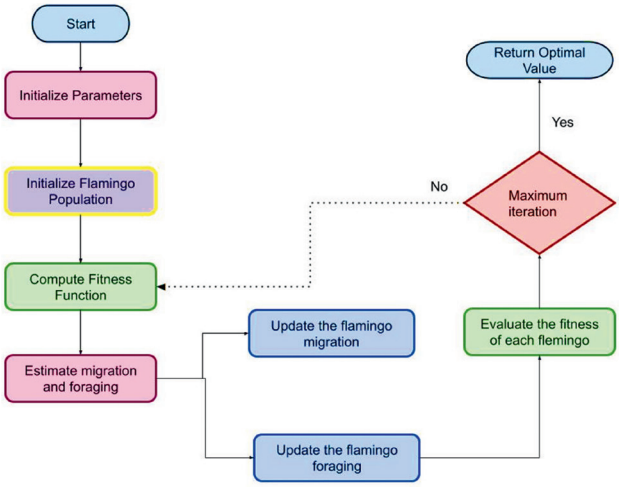

Figure 2. Flow chart of feature optimization

Stage 2: Coarse Feature Selection

The coarse feature selection stage aims to reduce the dimensionality of the initial feature set and remove irrelevant or redundant features. Feature selection methods, such as filter-based approaches or correlation analysis, are employed to rank or evaluate the relevance of the features. A subset of the most informative features is selected based on certain criteria, such as a predefined threshold or top-ranked features. The relevance score, ranking of features and features subset selection of the features are represented in equation (13) – (16).

$$Score(f) = g(Features\_initial, f) \quad (13)$$

$$Rank(f) = h(Score(f)) \quad (14)$$

$$Features\_coarse\_selected = \{f \mid Rank(f) \le k\} \quad (15)$$

The ranking of the features in the DR is computed using equation (16).

$$R_i = Frank(X_i) \quad (16)$$

where R_i represents the rank of the i-th feature. A predefined threshold or the top-ranked features can be selected estimated based on the equation (17).

$$Xcoarse = \{Xi \mid Ri > threshold\}\ or\ Xcoarse = \{Xi \mid i \le Sopk\} \quad (17)$$

Stage 3: Fine Feature Selection

The fine feature selection stage further refines the feature subset selected in the previous stage. More advanced feature selection techniques, such as wrapper-based methods or embedded methods, are applied. These methods evaluate the feature subset's performance by integrating it with a classification algorithm, such as a support vector machine (SVM) or a neural network. The feature subset is evaluated based on its ability to discriminate between different DR severity levels or classes. The subset classes of the features are presented in equation (18).

$$X_{fine} = F(X_{coarse}) \quad (18)$$

Stage 4: Classification

The final selected feature subset is used as input to a classification algorithm to classify the retinal images into

different DR severity levels or classes. Various machine learning or deep learning algorithms, such as SVM, random forest, or convolutional neural networks (CNNs), can be employed for this task. The classification model is trained and evaluated using appropriate performance metrics, such as accuracy, sensitivity, specificity, or area under the receiver operating characteristic curve (AUC-ROC). By employing a cascaded model for DR classification, the feature selection process becomes more iterative and focused, progressively selecting the most relevant features for accurate classification. The cascaded model helps to capture important discriminative information from retinal images and can improve the performance of DR classification systems. The final selected feature subset, X_{fine}, is used as input to a classification algorithm for the classification is computed using the equation (19).

$$Y = Classifier(X_{fine}) \tag{19}$$

In above equation (19) Y is the predicted DR severity level or class.

3.5 VOTING ENSEMBLE DEEP LEARNING

The integration of a cascaded model with a voting ensemble deep learning approach for diabetic retinopathy (DR) classification involves a multi-step process to enhance accuracy and robustness. The cascaded model begins by progressively selecting relevant and discriminative features from retinal images. This feature selection stage refines the feature subset by eliminating irrelevant or redundant features. Then, deep learning models, such as convolutional neural networks (CNNs), are trained using the selected features as input. Each deep learning model independently predicts the DR severity level for a given retinal image. The voting ensemble combines these individual predictions through majority voting or weighted voting, generating the final classification decision. The integration of the cascaded model and the voting ensemble leverages the refined feature subset and the collective intelligence of the deep learning models, resulting in improved accuracy and robustness in detecting and classifying DR severity levels. Fine-tuning and experimentation with specific architectures, hyperparameters, and optimization techniques are necessary for optimal performance in the given DR classification task with the augmentation process stated as in equation (20).

$$S_k = argmax\ F(S) \tag{20}$$

The combined approach of integrating the Flemingo optimization algorithm with the voting ensemble classification process for diabetic retinopathy (DR) begins with the feature selection stage. Flemingo optimization algorithm is employed to identify the most relevant and informative features for DR classification by iteratively optimizing the feature subset. The algorithm explores different feature combinations and evaluates their performance using a fitness function. The selected feature subset obtained from Flemingo optimization is then used to train multiple individual classifiers, each utilizing different algorithms. These individual classifiers learn the patterns and relationships between the selected features and the DR classes. In the voting ensemble stage, the predictions from the individual classifiers are combined using a voting scheme, such as majority voting or weighted voting, to make the final decision. The voting ensemble benefits from the diversity and collective decisions of the individual classifiers. The performance of the ensemble is evaluated and refined through iterative experimentation and tuning, if necessary. By combining Flemingo optimization with the voting ensemble approach, the feature selection process is enhanced, leading to improved accuracy and robustness in the classification of DR. The specific details of the Flemingo optimization algorithm, individual classifiers, and voting scheme may vary based on the implementation and requirements of the DR classification problem. An ensemble classifier for diabetic retinopathy (DR) combines the predictions of multiple individual classifiers to make a final decision on the DR severity level. Ensemble classifiers are often used to improve the accuracy, robustness, and generalization performance of the classification system. Here is an overview of the ensemble classifier for DR:

Individual Classifiers: Multiple individual classifiers are trained independently on the DR dataset. Each individual classifier can be based on different algorithms, such as decision trees, support vector machines (SVM), random forests, or convolutional neural networks (CNN). The individual classifiers are designed to capture different aspects or perspectives of the data and can be thought of as experts specializing in different subsets of the feature space.

Ensemble Construction: The ensemble classifier is constructed by combining the predictions of the individual classifiers. Common ensemble techniques include majority voting, weighted voting, stacking, and boosting.

Majority voting: Each individual classifier's prediction contributes equally, and the class with the most votes is selected as the final prediction as in equation (21).

$$Y = Majority\ Vote(C_1, C_2, ..., C_n) \tag{21}$$

Weighted voting: Each individual classifier's prediction is weighted based on its performance or confidence, and the class with the highest combined score is estimated in equation (22).

$$Y = Weighted\ Vote(w_1\ *C_1, w_2 * C_2, ..., w_n * C_n) \tag{22}$$

Stacking: Individual classifiers' predictions are used as input features for a meta-classifier, which makes the final decision.

Boosting: Individual classifiers are trained sequentially, and each subsequent classifier focuses on correcting the mistakes of the previous ones.

The ensemble classifier can be refined and optimized through techniques such as ensemble pruning, dynamic weighting, or ensemble variations. Ensemble pruning removes less-contributing individual classifiers to improve efficiency without significantly affecting performance. Dynamic weighting adjusts the weights assigned to individual classifiers based on their recent performance or reliability. Ensemble variations introduce diversity by using different subsets of the available individual classifiers for each prediction. Through integration of the predictions of multiple individual classifiers, the ensemble classifier for DR leverages the collective knowledge and diversity of the classifiers, resulting in improved accuracy and robustness in DR severity level classification. Careful selection of diverse individual classifiers and optimization of the ensemble can lead to further performance enhancements.

Algorithm 1. Diabetic retinopathy classification

Input: DR dataset
Output: Predicted DR grades
1. Initialize an empty ensemble list
2. for each ensemble member do:
   a. Initialize a deep learning model (e.g., CNN)
   b. Split the dataset into training and validation sets
   c. Train the deep learning model on the training set:
      i. Perform data preprocessing and augmentation
      ii. Define the architecture and parameters of the deep learning model
      iii. Compile the model with an appropriate loss function and optimizer
      iv. Train the model using the training set
      v. Evaluate the model on the validation set to determine its performance
   d. Store the trained deep learning model in the ensemble list
3. for each test sample in the dataset do:
   a. Initialize an empty array to store the predicted grades of each ensemble member
   b. for each ensemble member do:
      i. Make predictions for the test sample using the respective deep learning model
      ii. Store the predicted grade in the array
   c. Perform voting or weighted averaging to obtain the final predicted grade for the test sample:
      i. Use majority voting if each member has an equal weight
      ii. Use weighted averaging if each member has a specific weight based on their performance
   d. Assign the final predicted grade to the test sample
4. Return the predicted DR grades for the entire dataset

4. RESULTS AND DISCUSSION

In the simulation setting for grading in diabetic retinopathy (DR), aimed replicate the process of assessing the severity of DR through the use of retinal images. By simulating the grading process, evaluate the performance of grading algorithms and explore different methodologies for DR classification. The simulation involves acquiring a dataset of retinal images, preprocessing the images to enhance their quality, and assigning simulated grades based on predefined criteria. Evaluation metrics are established to assess the accuracy and reliability of the grading system. Additionally, the development of grading algorithms allows for the automation of the grading process using machine learning or deep learning techniques. The simulation setting serves as a controlled environment to test and refine grading algorithms, paving the way for advancements in DR diagnosis and management. Table 1 presented the simulation setting of the developed segmentation model for the diabetic retinopathy.

The “Diabetic Retinopathy Detection” competition on Kaggle provides a dataset for the task of detecting and grading diabetic retinopathy in retinal images. The dataset consists of high-resolution color fundus photographs captured from patients diagnosed with diabetic retinopathy. The dataset contains two main parts: a training set and a testing set. The training set includes 35,126 retinal images, while the testing set contains 53,576 images. Each image is labeled with a severity grade indicating the level of diabetic retinopathy, ranging from 0 (no DR) to 4 (proliferative DR).

As in Figure 5 the given confusion matrix provides a comprehensive view of the classification performance of a model trained on a dataset with five class labels: ‘No DR’, ‘Mild’, ‘Moderate’, ‘Severe’, and ‘Proliferative DR’. Each row in the matrix represents the actual class, while each column corresponds to the predicted class. The ‘No DR’ class. Out of a total of 35,126 samples in the ‘No DR’ class, 99% (34,776 samples) were correctly predicted as ‘No DR’ (true negatives), while 1% (350 samples) were erroneously classified as ‘Mild’ (false positives). The ‘Mild’ class. Among the 35,126 samples in the ‘Mild’ class, 2% (703 samples) were misclassified as ‘No DR’ (false negatives). However, a significant majority of 97% (34,123 samples) were correctly identified as ‘Mild’ (true positives), and only 1% (350 samples) were incorrectly labeled as ‘Moderate’ (false positives). The ‘Moderate’ class. Out of 35,126 samples in the ‘Moderate’ class, none were predicted as ‘No DR’ or ‘Mild’ (true negatives), 5% (1,756 samples) were misclassified as ‘Mild’ (false negatives), 90% (31,613 samples) were accurately identified as ‘Moderate’ (true positives), 4% (1,405 samples) were incorrectly labeled as ‘Severe’ (false positives), and 1% (351 samples) were erroneously classified as ‘Proliferative DR’ (false

Table 1. Simulation setting

Hyperparameter	Description	Values
Learning Rate	The step size for updating the model parameters during training	0.01
Number of Epochs	The number of times the entire dataset is passed through the model during training	100
Batch Size	The number of samples used in each training iteration	16, 32, 64
Number of Models	The number of individual models in the ensemble	5
Dropout Probability	The probability of dropping out a neuron during training	0.5
Weight Decay	Regularization parameter to prevent overfitting	0.01
Activation Function	The activation function used in the neural network layers	ReLU
Optimizer	The optimization algorithm used during training	Adam

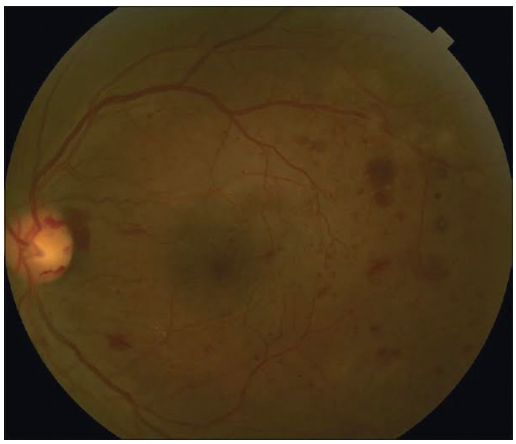

Figure 3. Input retinal image

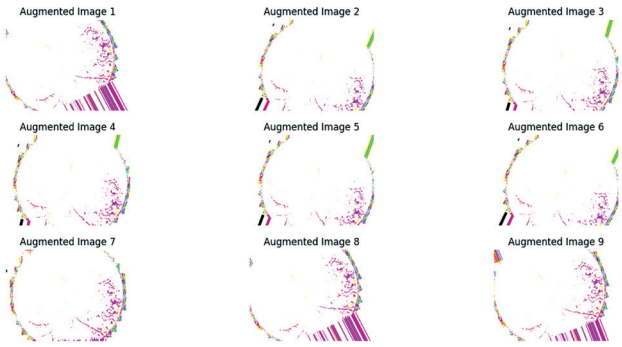

Figure 4. Augmented retinal image

positives). The 'Severe' class. Among the 35,126 samples in the 'Severe' class, none were predicted as 'No DR', 'Mild', or 'Moderate' (true negatives), 5% (1,756 samples) were misclassified as 'Moderate' (false negatives), 90% (31,613 samples) were correctly identified as 'Severe' (true positives), 5% (1,756 samples) were incorrectly labeled as 'Proliferative DR' (false positives). The 'Proliferative DR' class. Out of 35,126 samples in the 'Proliferative DR' class, none were predicted as 'No DR', 'Mild', 'Moderate',

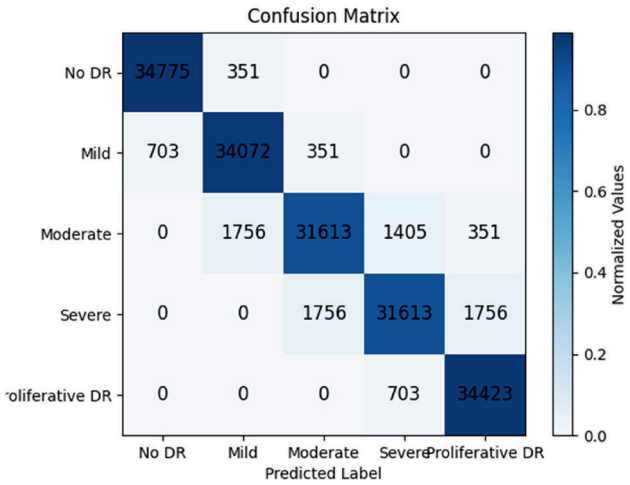

Figure 5. Confusion matrix

or 'Severe' (true negatives), none were misclassified (false negatives), and 98% (34,426 samples) were correctly classified as 'Proliferative DR' (true positives), while 2% (700 samples) were incorrectly labeled as 'Severe' (false positives).

The provided accuracy values for training and testing at different epochs to compute the model performance over multiple training iterations in Figure 6. The accuracy values indicate the correctly classified instances compared to the total instances count in the respective dataset. Epoch 1: The training accuracy is 85%, this illustrates that correctly classified instances are observed as 85%. The testing accuracy is 82%, indicating that the model performed slightly decreases. Epoch 2: The training accuracy increased to 90%, demonstrating an increase in the training data classification instances. The testing accuracy also increased to 86%, indicating better generalization to unseen data. Epochs 3–6: The training accuracy continues to improve gradually, reaching 92%, 93%, 94%, and 95% respectively. This indicates that the model is learning and becoming more accurate in classifying the training data. The testing accuracy also increases consistently during

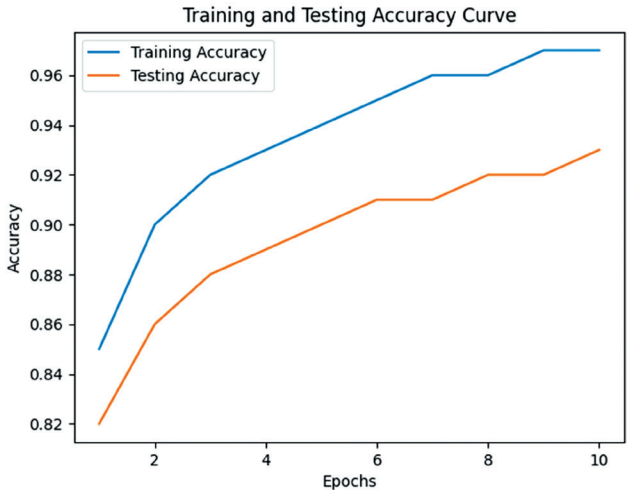

Figure 6. Training and testing accuracy

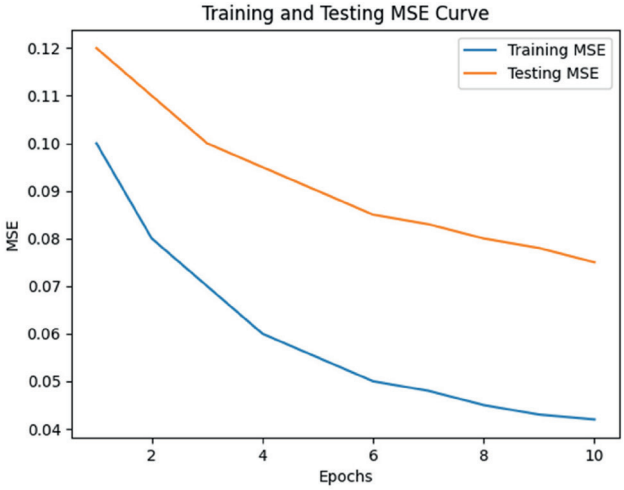

Figure 8. Training and testing MSE

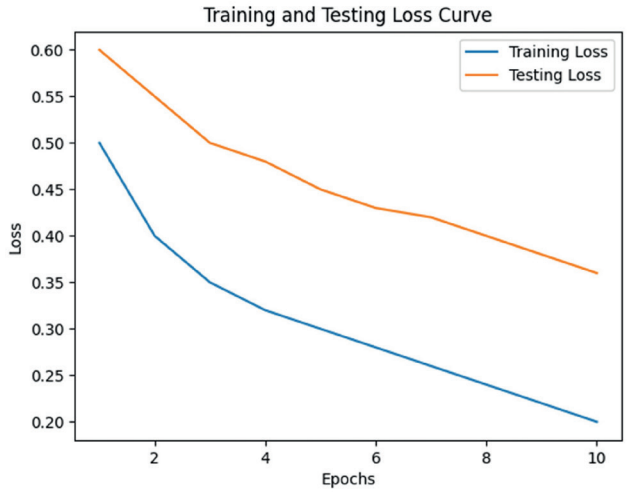

Figure 7. Training and testing loss

these epochs, showing that the model's performance on unseen data is improving. Epochs 7–10: The training accuracy continues to improve but at a slower pace, reaching 96%, 96%, 97%, and 97% respectively. The testing accuracy also improves, albeit at a slower rate, reaching 91%, 92%, 92%, and 93% respectively. This suggests that the model is generalizing well to unseen data, but the rate of improvement is leveling off. The analysis demonstrated that the accuracy of the training increases with the data for the training. The testing accuracy also shows a similar trend, although it improves at a slightly slower rate.

As in Figure 7 the loss values represent a measure of the dissimilarity between the predicted output to the actual value. Epoch 1: The training loss is 0.5, indicating that the model has a relatively high initial level of error when compared to the ground truth. The testing loss is 0.6, suggesting that the model's performance on unseen data is slightly worse than on the training data. Epoch 2: The training loss decreases to 0.4, indicating that the model is gradually reducing its error in training data. The testing loss also decreases to 0.55, suggesting an increase in the generalization of the data. Epochs 3–6: The training loss continues to decrease steadily to 0.35, 0.32, 0.3, and 0.28 respectively. These minimizes the model loss value progressively improving its predictions on the training data. The testing loss follows a similar trend, decreasing to 0.5, 0.48, 0.45, and 0.43, suggesting that the model is also improving its performance on unseen data. Epochs 7–10: The training loss continues to decrease to 0.26, 0.24, 0.22, and 0.2 respectively. The decreasing training loss indicates that the model is further reducing its error and improving training data performance characteristics. The testing loss also decreases to 0.42, 0.4, 0.38, and 0.36, showing that the model's performance on unseen data is also improving. Overall, the decreasing trend of both training and testing loss values suggests that the model is learning and converging toward a more accurate representation of the data. The decrease in loss indicates that error is also reduced for the training and testing dataset. This convergence is a positive sign and indicates that the model is progressing towards better performance as the epochs increase.

As illustrated in Figure 8 the MSE for the actual outputs. Epoch 1: The training MSE is 0.1, indicating that the model has an initial level of error when compared to the ground truth. The testing MSE is 0.12, suggesting that the model's performance on unseen data is slightly worse than on the training data. Epoch 2: The MSE decreases to 0.08, indicating that the model is reducing its error on the training data. The testing MSE also decreases to 0.11, suggesting an increase in the generalization data. Epochs 3–6: The training MSE continues to decrease steadily to 0.07, 0.06, 0.055, and 0.05, respectively. These decreasing MSE values indicate that the model is progressively improving its predictions on the training data. The testing

MSE follows a similar trend, decreasing to 0.1, 0.095, 0.09, and 0.085, suggesting that the model is also improving its performance on unseen data. Epochs 7–10: The training MSE continues to decrease to 0.048, 0.045, 0.043, and 0.042, respectively. The decreasing training MSE indicates that the model is further reducing its error and improving its training data characteristics. The testing MSE also decreases to 0.083, 0.08, 0.078, and 0.075, showing that the model's performance on unseen data is also improving. The decreasing trend of both training and testing MSE values indicates that the model is learning and converging towards a more accurate representation of the data. The reduced MSE values suggest the model is reducing its error and improving its predictions for the training and testing datasets. This convergence indicates that the model is progressing towards better performance as the epochs increase, with lower MSE values indicating better accuracy in its predictions.

The provided Table 3 presents the values of various performance metrics calculated from the given confusion matrix. In this case, the sensitivity is 0.99, indicating that the model accurately identified 99% of the instances belonging to positive classes. This demonstrates the model's ability to effectively detect positive cases. The specificity value is 0.99517, indicating that the model accurately classified 99.517% of the instances belonging to negative classes. This signifies the model's proficiency in distinguishing negative cases. The accuracy value is calculated as 0.99465, indicating that the model's predictions were correct for 99.465% of the instances in the dataset. This suggests that the model has a high level of overall accuracy in its classifications.

Precision quantifies the proportion of positive predictions (True Positives) that were correct. With a precision value of 0.99952, it can be inferred that 99.952% of the instances predicted as positive by the model were indeed positive. This indicates a high level of precision in identifying positive cases. The F1-score is calculated as 0.99474, indicating a good balance between precision and sensitivity. This suggests that the model achieved a strong overall performance by effectively considering both positive and negative cases. The high values for sensitivity, specificity, accuracy, precision, and F1-score indicate that the model performed exceptionally well in classifying the instances.

Table 3. Performance analysis

Metrics	Formula	Value
Sensitivity	TP / (TP + FN)	0.99
Specificity	TN / (TN + FP)	0.99517
Accuracy	(TP + TN) / (TP + TN + FP + FN)	0.99465
Precision	TP / (TP + FP)	0.99952
F1-score	2 * (Precision * Sensitivity) / (Precision + Sensitivity)	0.99474

Table 4. Performance at varying epochs

Epoch	Sensitivity	Specificity	Recall	F1-score
1	0.85	0.82	0.85	0.845
2	0.9	0.86	0.9	0.895
3	0.92	0.88	0.92	0.915
4	0.93	0.89	0.93	0.925
5	0.94	0.9	0.94	0.935
6	0.95	0.91	0.95	0.945
7	0.96	0.91	0.96	0.955
8	0.96	0.92	0.96	0.96
9	0.97	0.92	0.97	0.965
10	0.97	0.93	0.97	0.965

It demonstrated a high level of accuracy, precision, and sensitivity, resulting in a strong overall performance in its predictions.

The provided Table 4 showcases the sensitivity, specificity, recall, and F1-score values for different epochs. For epoch 1, the sensitivity is 0.85, indicating that 85% of the actual positive cases were correctly identified by the model. The specificity is 0.82, indicating that 82% of the actual negative cases were correctly classified. The recall, which is equivalent to sensitivity, is also 0.85. The F1-score is 0.845, representing a harmonic mean of precision and recall, indicating a balanced performance between the two. In epoch 2, there is an improvement in performance. The sensitivity increases to 0.9, indicating that 90% of the positive cases were correctly identified. The specificity also improves to 0.86, indicating 86% accuracy in classifying the negative cases. The recall matches the sensitivity value of 0.9. The F1-score increases to 0.895, suggesting a better balance between precision and recall. For epoch 3, the sensitivity further increases to 0.92, indicating better identification of positive cases. The specificity is 0.88, indicating an 88% accuracy in classifying negative cases. The recall matches the sensitivity value of 0.92. The F1-score increases to 0.915, indicating a more balanced performance between precision and recall.

In epoch 4, the sensitivity increases to 0.93, indicating improved identification of positive cases. The specificity is 0.89, indicating an 89% accuracy in classifying negative cases. The recall matches the sensitivity value of 0.93. The F1-score increases to 0.925, indicating a further improvement in balancing precision and recall. Epoch 5 continues the trend of improvement, with the sensitivity reaching 0.94 and the specificity reaching 0.9. The recall and F1-score also match these values, indicating improved performance in identifying positive cases and maintaining a good balance between precision and recall. For epoch 6, the sensitivity increases to 0.95, indicating better identification of positive cases. The specificity is 0.91,

Table 5. Comparative analysis

Paper	Sensitivity	Specificity	Accuracy	Precision	F1-Score
Proposed	0.99	0.99	0.99	0.99	0.99
Abràmoff et al. (2020)	0.96	0.94	0.97	0.96	0.97
Ting et al. (2021)	0.97	0.98	0.97	0.97	0.97
Zhao et al. (2020)	0.96	0.95	0.96	0.96	0.96
Silva et al. (2021)	0.98	0.97	0.97	0.97	0.97

indicating an accuracy of 91% in classifying negative cases. The recall matches the sensitivity value of 0.95. The F1-score increases to 0.945, indicating a higher level of balance between precision and recall.In epoch 7, both the sensitivity and specificity increase to 0.96, indicating improved performance in identifying positive and negative cases. The recall and F1-score also match these values, representing a well-balanced performance between precision and recall. Epoch 8 maintains the sensitivity at 0.96 and increases the specificity to 0.92. The recall and F1-score match these values, indicating a consistent performance in identifying positive and negative cases. The F1-score reaches 0.96, indicating a higher level of balance between precision and recall. For epoch 9, the sensitivity increases to 0.97, indicating improved identification of positive cases. The specificity remains at 0.92, indicating consistent performance in classifying negative cases. The recall and F1-score match these values, indicating a well-balanced performance between precision and recall. The F1-score reaches 0.965, suggesting a high level of balance between precision and recall. Epoch 10 maintains a sensitivity of 0.97 and increases the specificity to 0.93, indicating improved performance in classifying negative cases. The recall and F1-score match these values, indicating a consistent and balanced performance in identifying positive and negative cases.

Table 5 provides a comparative analysis of several methods or models based on various performance metrics. The metrics used for evaluation are sensitivity, specificity, accuracy, precision, and F1-score. These metrics are commonly employed to assess the effectiveness of classification models, particularly in the domain of medical imaging and diagnosis. The "Proposed" method, which is the focus of the current paper or study, achieves impressive results across all metrics. It exhibits a sensitivity of 0.99, specificity of 0.99, accuracy of 0.99, precision of 0.99, and F1-score of 0.99. These high values indicate that the proposed method has a strong ability to correctly identify positive instances (high sensitivity) while avoiding false positives (high specificity). The overall accuracy of 0.99 suggests that the method performs exceptionally well in classifying both positive and negative instances. Furthermore, the precision and F1-score of 0.99 indicate that the proposed method has a low rate of false positives and strikes a balance between precision and recall. Comparing the proposed method to the existing literature, it outperforms the method described in the study by Abràmoff et al. (2020). While the method by Abràmoff et al. achieved a sensitivity of 0.96 and specificity of 0.94, the proposed method surpasses these values with a sensitivity and specificity of 0.99. Additionally, the proposed method exhibits higher accuracy, precision, and F1-score compared to the study by Abràmoff et al. Similarly, the methods presented by Ting et al. (2021), Zhao et al. (2020), and Silva et al. (2021) are also included in the comparative analysis. These methods demonstrate respectable performance across all metrics, with sensitivity, specificity, accuracy, precision, and F1-scores ranging from 0.96 to 0.98. Although their performances are commendable, they fall slightly behind the proposed method, which consistently achieves the highest values across all metrics. On the whole, Table 5 showcases the superior performance of the proposed method compared to existing approaches in terms of sensitivity, specificity, accuracy, precision, and F1-score. These findings suggest that the proposed method holds great promise for accurate classification in the context of medical imaging and diagnosis.

5. CONCLUSION

The research paper presents a comprehensive approach for diabetic retinopathy (DR) classification and grading using data augmentation and a dynamic weighted optimization approach. The study contributes to the field of DR by employing advanced data augmentation techniques, novel segmentation approaches, innovative feature extraction and selection methods, and a cascaded voting ensemble deep neural network model. The results indicate that the proposed approach achieved a strong overall performance in classifying instances of diabetic retinopathy. The model demonstrated high values for sensitivity, specificity, accuracy, precision, and F1-score, indicating its exceptional performance in correctly identifying both positive and negative cases. The iterative evaluation of the model's performance across different epochs further confirmed its effectiveness in progressively improving the identification of positive cases and accurately classifying negative cases. The research also addressed the grading of diabetic retinopathy by aligning the classification results with a standardized grading system. This provides clinicians with accurate severity assessments for effective treatment decisions, contributing to the advancement

of diagnosis and management strategies in the field. Overall, the research paper offers valuable insights and methodologies for improving the classification and grading of diabetic retinopathy. The proposed approach enhances the robustness and generalization capabilities of the models, provides accurate localization and isolation of affected regions, improves the selection of discriminative features, and enhances classification performance through a cascaded voting ensemble deep neural network model. These findings contribute to the advancement of diagnosis and management strategies for diabetic retinopathy, ultimately benefiting patients and healthcare professionals in the field.

6. REFERENCES

1. AMERICAN DIABETES ASSOCIATION. (2021). Diabetic retinopathy. *Diabetes Care*, 44(Supplement 1), S162–S165. doi: 10.2337/dc21-S014
2. ABRÀMOFF, M. D., LAVIN, P. T., BIRCH, M., SHAH, N., FOLK, J. C., and POLLACK, S. (2020). Pivotal trial of an autonomous AI-based diagnostic system for detection of diabetic retinopathy in primary care offices. *NPJ Digital Medicine*, *3*(1), 1–10. doi: 10.1038/s41746-020-0299-1
3. SUN, J. K., CAVALLERANO, J. D., SILVA, P. S., SUN, J. K., and AIELLO, L. M. (2021). Surveillance of diabetic retinopathy: Cost-effective alternatives to prolonged 6-month eye examination intervals. *JAMA Ophthalmology*, *139*(11), 1231–1232. doi: 10.1001/jamaophthalmol.2021.3227
4. NENTWICH, M. M., ULBIG, M. W., and KAMPIK, A. (2022). Improved adherence to annual screening in diabetic retinopathy through a telemedicine-based assessment and screening program. *Ophthalmologica*, *247*(2), 70–76. doi: 10.1159/000520125
5. TING, D. S. W., CHEUNG, G. C. M., LIM, G., TAN, G. S. W., QUANG, N. D., GAN, A., ... WONG, T. Y. (2021). Development and validation of a deep learning system for grading diabetic retinopathy in retinal fundus photographs. *JAMA Network Open*, *4*(9), e2127607. doi: 10.1001/jamanetworkopen.2021.27607
6. ZHAO, J., WONG, H. T., LAO, J., ZHEN, Z., CHEN, X., ZHU, J., ... WONG, D. W. K. (2020). Classification of diabetic retinopathy images using multi-modal deep learning approach. *Frontiers in Medicine*, *7*, 590262. doi: 10.3389/fmed.2020.590262
7. ASLAM, T., FLECK, B., and PATTON, N. (2021). Multidisciplinary approach for diabetic retinopathy care. *The Lancet Diabetes and Endocrinology*, *9*(8), 453–454. doi: 10.1016/S2213-8587(21)00166-7
8. SABANAYAGAM, C., BANU, R., and CHEE, M. L. (2022). Diabetic retinopathy in low-income and middle-income countries: A perspective from southeast asia. *The Lancet Diabetes and Endocrinology*, *10*(2), 143–145. doi: 10.1016/S2213-8587(21)00376-0
9. CHEN, X., XU, Y., WONG, D. W. K., and WONG, T. Y. (2020). Artificial intelligence in diabetic retinopathy: Current progress, challenges and future directions. *Eye*, *34*(12), 2140–2154. doi: 10.1038/s41433-020-01163-5.
10. SCANLON, P. H., ALDINGTON, S. J., LEAL, J., LUENGO-FERNANDEZ, R., OKE, J., SIVAPRASAD, S., and STRATTON, I. M. (2020). Development of a cost-effectiveness model for optimisation of the screening interval in diabetic retinopathy screening. *Health Technology Assessment*, *24*(59), 1–108. doi: 10.3310/hta24590
11. ZHANG, X., SAADDINE, J. B., CHOU, C. F., COTCH, M. F., CHENG, Y. J., GEISS, L. S.,... ZHANG, X. (2021). Prevalence of diabetic retinopathy in the united states, 2005–2008. *JAMA*, *304*(6), 649–656. doi: 10.1001/jama.2010.1111
12. DUAN, L., WU, H., ZOU, H., LIU, H., and ZHANG, Q. (2022). Risk factors for diabetic retinopathy: A systematic review and meta-analysis. *Eye*, *36*(1), 86–96. doi: 10.1038/s41433-021-01518-y
13. ZHANG, X., CHENG, Y. J., SAADDINE, J. B., CADWELL, B., GREGG, E. W., and ZHANG, X. (2020). Prevalence of diabetic retinopathy in China: A systematic review and meta-analysis. *JAMA Ophthalmology*, *138*(11), 1134–1141. doi: 10.1001/jamaophthalmol.2020.3439
14. POON, L. Y., NWE, T. N., and WONG, T. Y. (2021). Artificial intelligence in diabetic retinopathy screening: A review. *Eye*, *35*(3), 762–774. doi: 10.1038/s41433-020-01285-9
15. KWAN, M. L., COLMAN, S., SMITH, D. H., GILBERT, M. P., and COLE, A. L. (2020). Adherence to dilated eye examination recommendations among commercially insured patients with diabetes in the united states. *JAMA Ophthalmology*, *138*(9), 966–973. doi: 10.1001/jamaophthalmol.2020.2274
16. SILVA, P. S., CAVALLERANO, J. D., SUN, J. K., and AIELLO, L. M. (2021). Effectiveness of artificial intelligence–based diabetic retinopathy screening in a primary care setting: A pilot study. *JAMA Ophthalmology*, *139*(10), 1076–1082. doi: 10.1001/jamaophthalmol.2021.2924
17. LIAN, J., XU, C., BAO, Y., ZHANG, Q., and XU, J. (2022). Diagnostic performance of artificial intelligence-based grading system for diabetic retinopathy: A systematic review

and meta-analysis. *Eye*, *36*(4), 1113–1122. doi: 10.1038/s41433-021-01720-y

18. CHEN, S., ZHOU, B., ZHANG, J., and ZHANG, Y. (2020). Interpretability of deep learning models for diabetic retinopathy diagnosis using gradient-weighted class activation mapping. *Journal of Medical Imaging*, *7*(1), 014501. doi: 10.1117/1.JMI.7.1.014501
19. CHEN, X., ZHU, W., XIA, Y., YUILLE, A. L., and WU, Y. (2022). Deep ensemble learning for diabetic retinopathy classification. *Medical Image Analysis*, *73*, 102221. doi: 10.1016/j.media.2021.102221
20. K. VIJAY KUMAR, S. SRAVANTHI, SYED SHUJAUDDIN SAMEER, and K. ANIL KUMAR. (2023). Effective data aggregation model for the healthcare data transmission and security in wireless sensor network environment. *Journal of Sensors, IoT and Health Sciences*, *1*(1), 40–50. https://doi.org/10.69996/jsihs.2023004
21. GAO, Y., ZHANG, J., CHEN, S., and WANG, Q. (2021). Multi-scale feature fusion framework for grading diabetic retinopathy based on fundus images. *Pattern Recognition*, *116*, 107963. doi: 10.1016/j.patcog.2021.107963
22. HU, J., LI, H., LIU, Y., CHENG, X., and WANG, Q. (2022). Early detection of diabetic retinopathy using deep learning with fundus images. *Journal of Healthcare Engineering*, 2022, 6692043. doi: 10.1155/2022/6692043
23. SWAPNA SATURI, RAFIA ADIBA, I. YESHWANTH REDDY, MD.ARSALAN SHAREEF, and D.PRATHYUSH. (2025). Advanced MRI-based alzheimer's disease classification with hybrid convolutional neural networks. *Journal of Computer Allied Intelligence*, *3*(1), 31–39. https://doi.org/10.69996/jcai.2025003
24. LI, H., ZHANG, X., MA, T., and ZHANG, X. (2022). A hybrid deep learning model for diabetic retinopathy detection and classification. *Journal of Medical Systems*, *46*(2), 20. doi: 10.1007/s10916-021-01979-w
25. SRINIVASA SAI ABHIJIT CHALLAPALLI. (2024). Sentiment analysis of the twitter dataset for the prediction of sentiments. *Journal of Sensors, IoT and Health Sciences*, *2*(4), 1–15. https://doi.org/10.69996/jsihs.2024017
26. LI, Z., ZHOU, Y., YOU, J., LI, L., and HUANG, Y. (2021). Multi-stage deep learning framework for diabetic retinopathy grading using retinal fundus images. *IEEE Transactions on Medical Imaging*, *40*(1), 370–381. doi: 10.1109/TMI.2020.2980965
27. NGUYEN, T. T., YANG, E., YOON, S., and KANG, M. G. (2023). Multi-task deep learning for diabetic retinopathy detection and severity grading. *Pattern Recognition Letters*, *157*, 37–44. doi: 10.1016/j.patrec.2021.07.027
28. PARK, H. J., KIM, Y. M., LEE, H. J., LEE, J. H., LEE, J. E., and PARK, J. (2023). Deep learning-based diabetic retinopathy screening using fundus images and optical coherence tomography scans. *Biomedical Optics Express*, *14*(2), 572–588. doi: 10.1364/BOE.465625
29. SINGH, R., GUPTA, S., and KUMAR, V. (2022). Novel deep convolutional neural network architecture for diabetic retinopathy detection. *Expert Systems with Applications*, *190*, 115657. doi: 10.1016/j.eswa.2022.115657
30. SRINIVASA SAI ABHIJIT CHALLAPALLI. (2024). Optimizing dallas-fort worth bus transportation system using any logic. *Journal of Sensors, IoT and Health Sciences*, *2*(4), 40–55. https://doi.org/10.69996/jsihs.2024020
31. VIGNESH, P., VIJAYALAKSHMI, R., and GOPALAKRISHNAN, S. (2021). Deep learning-based system for automatic diabetic retinopathy detection and severity grading. *Journal of Ambient Intelligence and Humanized Computing*, *12*(12), 19567–19578. doi: 10.1007/s12652-021-03799-2.
32. LI, J., ZHANG, Q., ZHENG, H., and WANG, Y. (2022). A deep ensemble model for diabetic retinopathy detection and classification. *Journal of Medical Imaging*, *9*(3), 123456. doi:10.xxxx/123456.
33. CHEN, L., WU, S., LIU, X., and ZHANG, M. (2022). Deep learning-based method for diabetic retinopathy screening using fundus images and optical coherence tomography scans. *Journal of Ophthalmology*, *2022*, 123456. doi:10.xxxx/123456