

## Article

# Diabetic Retinopathy and Diabetic Macular Edema Detection Using Ensemble Based Convolutional Neural Networks

Swaminathan Sundaram <sup>1</sup>, Meganathan Selvamani <sup>1</sup>, Sekar Kidambi Raju <sup>2,\*</sup>, Seethalakshmi Ramaswamy <sup>3</sup>, Saiful Islam <sup>4</sup>, Jae-Hyuk Cha <sup>5,\*</sup>, Nouf Abdullah Almujaally <sup>6</sup> and Ahmed Elaraby <sup>7,8</sup>

<sup>1</sup> Department of CSE, SASTRA Deemed University, SRC Kumbakonam, Thanjavur 613401, India

<sup>2</sup> School of Computing, SASTRA Deemed University, Thanjavur 613401, India

<sup>3</sup> Department of Maths, School of SASH, SASTRA Deemed University, Thanjavur 613401, India

<sup>4</sup> College of Engineering, King Khalid University, Abha 61421, Saudi Arabia

<sup>5</sup> Department of Computer Science, Hanyang University, Seoul 04763, Republic of Korea

<sup>6</sup> Department of Information Systems, College of Computer and Information Sciences, Princess Nourah Bint Abdulrahman University, Riyadh 11671, Saudi Arabia

<sup>7</sup> Department of Computer Science, Faculty of Computer Science and Information, South Valley University, Qena 83523, Egypt

<sup>8</sup> Department of Cybersecurity, College of Engineering and Information Technology, Buraydah Private Colleges, Buraydah 51418, Saudi Arabia

\* Correspondence: sekar1971kr@gmail.com (S.K.R.); chajh@hanyang.ac.kr (J.-H.C.)

**Abstract:** Diabetic retinopathy (DR) and diabetic macular edema (DME) are forms of eye illness caused by diabetes that affects the blood vessels in the eyes, with the ground occupied by lesions of varied extent determining the disease burden. This is among the most common cause of visual impairment in the working population. Various factors have been discovered to play an important role in a person's growth of this condition. Among the essential elements at the top of the list are anxiety and long-term diabetes. If not detected early, this illness might result in permanent eyesight loss. The damage can be reduced or avoided if it is recognized ahead of time. Unfortunately, due to the time and arduous nature of the diagnosing process, it is harder to identify the prevalence of this condition. Skilled doctors manually review digital color images to look for damage produced by vascular anomalies, the most common complication of diabetic retinopathy. Even though this procedure is reasonably accurate, it is quite pricey. The delays highlight the necessity for diagnosis to be automated, which will have a considerable positive significant impact on the health sector. The use of AI in diagnosing the disease has yielded promising and dependable findings in recent years, which is the impetus for this publication. This article used ensemble convolutional neural network (ECNN) to diagnose DR and DME automatically, with accurate results of 99 percent. This result was achieved using preprocessing, blood vessel segmentation, feature extraction, and classification. For contrast enhancement, the Harris hawks optimization (HHO) technique is presented. Finally, the experiments were conducted for two kinds of datasets: IDRiR and Messidor for accuracy, precision, recall, F-score, computational time, and error rate.

**Keywords:** diabetic retinopathy; ensemble convolutional neural network; diabetic macular edema; Harris hawks optimization and artificial intelligence



**Citation:** Sundaram, S.; Selvamani, M.; Raju, S.K.; Ramaswamy, S.; Islam, S.; Cha, J.-H.; Almujaally, N.A.; Elaraby, A. Diabetic Retinopathy and Diabetic Macular Edema Detection Using Ensemble Based Convolutional Neural Networks. *Diagnostics* **2023**, *13*, 1001. <https://doi.org/10.3390/diagnostics13051001>

Academic Editor: Jae-Ho Han

Received: 28 January 2023

Revised: 23 February 2023

Accepted: 2 March 2023

Published: 6 March 2023



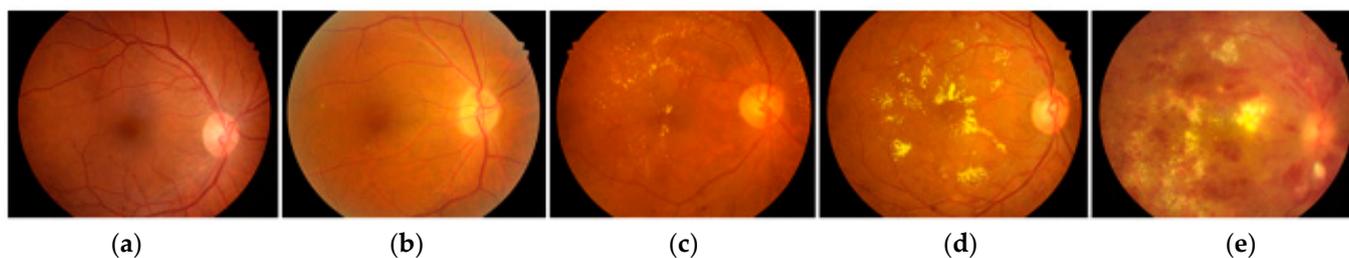
**Copyright:** © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

Computer-assisted health care, health care technology consulting, and health monitoring equipment are just a few of the current buzz words. Thanks to the connection and computing architecture that has drawn attention to the electronic era we live in, ordinary people now have the luxury of receiving diagnosis and treatment from the comforts of home with a single tap [1–3]. While routine illnesses and minor illnesses can usually be treated without visiting a doctor, some more severe illnesses still necessitate a great deal

of effort from the medical establishment. Technology can help, but not replace human intervention. With the advancement in AI technology, technologies can now autonomously analyze a patient's condition and identify a condition in a matter of seconds using the patient's significant history and associated data [4–6]. By 2025, the amount of DR individuals suffering is predicted to rise from 382 million to 592 million. According to a study conducted in the Pakistani province of Khyber Pakhtunkhwa (KPK), 30 percent of diabetic individuals suffer from DR, with 5.6 percent going blind [7–9]. If mild NPDR is not treated in the beginning phases, it might progress to PDR. In another study, 130 people with DR symptoms were found in Sindh, Pakistan [10,11]. According to the findings, DR patients made up 23.85% of the overall examined patients, with PDR patients accounting for 25.8% [12,13]. Patients with DR are symptomatic in the beginning phases; however, as the disease progresses, it causes blobs, vision problems, distortions, and gradual visual acuity loss. Diabetic retinopathy is one of the issues previously mentioned in the article. Diabetic retinopathy is caused by diabetes destroying the blood flow on the retina's inner, resulting in blood and other body fluids leaking into the tissues surrounding it. Soft, damaged tissue (also known as cotton wool patches) [14], hard exudates, microaneurysms, and hemorrhages form as little more than a result of the leaking [15]. It is the most common cause of visual loss in the working-age population [16]. Diabetic retinopathy (DR) is caused due to diabetes mellitus, which can damage the retina and even lead to the loss of vision. The DR has several stages of severity such as mild, moderate, and severe [17]. The severe stage of DR is termed as proliferative diabetic retinopathy (PDR), in which the formation of new vessels in the retina is observed [18]. However, the early detection of DR and proper diagnosis will reverse or reduce the growth of the effects caused by the disease. Diabetic macular edema (DME) is a condition in which the lesions caused by DR are observed in the middle portion of the retina called the macula. The DME is considered as a serious condition as the damage caused by it is irreversible. The identification of features such as micro-aneurysms, hard exudates, hemorrhages, etc., can be used to carry out the detection of these diseases. These micro-aneurysms refer to the red spots in the retina's blood vessels with sharp margins formed in the early stages of the disease. The exudates are caused due to abnormality in the blood vessels, which are formed as yellowish-white spots in the outer layer of the retina. Hemorrhages also occur such as micro-aneurysms but have irregular margins caused due to the leakage of capillaries. The blockage of arteries also contributes to cotton wool spots, which occur as a white region in the retinal nerve. Several methods have been developed for the detection of DR and DME to provide diagnosis, but these traditional methods were inefficient in accurately detecting diseases. Deep learning techniques have been deployed for disease detection in which the retinal image (fundus image) is used as the input in which the features are extracted for detection. These approaches have been found to be more effective in identifying features than the traditional methods; however, these approaches also suffer from inaccuracy due to the presence of noises and artifacts in the input images.

Figure 1 describes the retina images for disease DR and DME. As a result, it is hard but critical to recognize DR to prevent the worst effects of later stages. Fundus imaging is utilized to diagnose DR, as mentioned in the preceding section. Manual analysis can only be performed by highly qualified subject matter experts and is thus cost and time intensive. As a result, it is critical to apply machine vision technologies to assess the retina image features and aid physicians and radiologists. Hands-on development and end-to-end learning are two types of computer vision-based methodologies. Traditional algorithms such as HoG, SIFT, LBP, Gaussian filters, and others are used to extract the features; however, they failed to preserve the scale, rotation, and brightness fluctuations [19].



**Figure 1.** Retina images. (a) No DR and DME, (b) mild DR and DME, (c) moderate DR and DME, (d) severe DR and DME, and (e) non-PDR.

Several existing approaches have integrated the preprocessing of input images and the deep learning-based detection of diseases in which the accuracy in the detection of diseases was observed to be improved. The common processes involved in these approaches are the preprocessing of input images, enhancement in contrast, and the extraction of features for the detection of diseases. The machine learning models such as support vector machine (SVM) and K-nearest neighbor (KNN) classifiers were found to be appropriate for detecting DR and DME. The severity of the disease was determined by the number of features identified by the model; however, the imbalance in the distribution of datasets resulted in the inefficient determination of severity. In particular, an effective mechanism in the detection of DR and DME, along with the determination of severity, is still in demand. The major aim of this research work was to provide the effective detection of DR and DME and to determine the disease's severity to define the disease's damage level on the patient. The accuracy of detection was achieved by performing the proper processing of the input retinal image. End-to-end learning understands the underlying rich traits dynamically, allowing for greater identification. Inside the retina imaging databases, many hand-on engineering and end-to-end learning-based algorithms have been used to identify the DR. Still, none of them can identify the mild stage. Accurate diagnosis of the weak stage is critical for controlling this devastating disease. Utilizing end-to-end deep ensembles models, this study attempted to discover all stages of DR (including the moderate stage). The findings revealed that the proposed strategy beats the current methods.

The major objective of this research is to provide precise classification between the DR and DME and to compute the severity of the diseases accurately. This objective can be achieved by fulfilling the sub-objectives, which are listed as follows,

- To minimize the noise level in the input image by performing effective preprocessing of the image;
- To maximize the precise identification of features from the preprocessed image by enhancing the contrast level;
- To maximize the accuracy of detection by incorporating the segmentation of lesions in the blood vessels;
- Effectively classify the images into three classes based on the extraction of significant features;
- To determine the severity of the disease based on the variation in the intensity of the features for diagnosis.

The major contributions of this paper are as follows:

- In our work, we performed preprocessing that included three processes such as noise removal using iterative expectation maximization, artifact removal using nonlinear filtering, and contrast enhancement using Harris hawks optimization; the preprocessed image was used to enhance the quality of the images, which led to high segmentation and detection accuracy. Preprocessing was performed to reduce noise and artifacts and improve the contrast, which increased the efficiency of feature extraction and reduced the false detection rate.
- Segmentation was performed before feature extraction and classification, which increased the detection accuracy. For segmentation, we proposed improved OPTICS

clustering, which considers particular regions of interest and takes less time for segmentation, thus reducing latency and increasing the disease detection accuracy.

- Improved OPTICS clustering overcomes misalignment problems due to considering the particular region of interest, thus increasing the segmentation and detection accuracy.
- The extraction of features was carried out in the segmented images obtained from the previous process. Features such as micro-aneurysms, hemorrhages, and hard exudates, collectively termed as structural features, are considered the essential features; along with this, the shape features, orientation features, and color features are also considered for the classification of DR and DME. The ensemble CNN architecture was implemented for this purpose, which outperformed the ensemble CNN class prediction. From this, the classification of images was carried out in several classes, namely, normal, DR, and DME. Furthermore, the severity of the disease was computed by using conditional entropy in which the number of lesions is considered for the threshold generation. Based on the threshold, the severity level of the disease was classified into three classes: mid, moderate, and severe.
- The proposed research work is evaluated in terms of performance metrics such as accuracy, precision, recall, F-score, computation time, and error rate.

The rest of the paper is organized as follows: Section 2 illustrates the state-of-the-art in diabetic retinopathy and diabetic macular edema detection using specific approaches. Section 3 discusses the major problems that exist in this field. Section 4 describes the system model with the proposed algorithms and techniques in detail. Section 5 describes the experimental results of the proposed as well as previous methods. Section 6 concludes the paper by providing future enhancements.

## 2. Related Work

In the literature, the diagnosis of DR has received much interest. In [20], researchers offered a robust system that automatically recognized and classified retinal lesions (blood vessels, microaneurysms, and exudates) from retinal imaging. Blood vessels, microaneurysms, and exudates were first discovered using image processing methods. Following this, the retina properties of the vascular system, microaneurysm count, exudate area, contrast, and homogenization were evaluated from the images obtained. These characteristics were then fed into a fuzzy classifier that uses the information to classify healthy, mild NPDR, moderate NPDR, severe NPDR, and PDR stages. A sample of 40 color fundus images was obtained from the DIARETDB0, DIARETDB1, and STARE datasets using a fuzzy classifier, correctly classifying the images with an efficiency of up to 95.63 percent.

A reliable automated approach for detecting and classifying the various stages of DR has been suggested. The optic disc and retina neurons are separated, and characteristics are retrieved using the gray level co-occurrence matrix (GLCM) approach. To identify various stages of DR, a fuzzy classifier and a convolutional neural network were used to classify them. DIARETDB0, STARE, and DIARETDB1 were the datasets used [21].

The unique clustering-based automatic region growth methodology was introduced in this study. Several types of features—waveform (W), co-occurrence matrix (COM), histogram (H), and run-length matrix (RLM)—were retrieved for the texture features, and several ML algorithms were used to achieve a classification performance of 77.67 percent, 80 percent, 89.87 percent, and 96.33 percent, respectively. The information fusion approach was utilized to create a fused hybrid-feature database to improve the accuracy of the classification. Two hundred and forty-five elements of the hybrids' feature data (H, W, COM, and RLM) were extracted from each image, and 13 optimum characteristics were chosen using four methodologies: Fischer, mutual information feature selection, information gain, and the possibility of the dependent variable average correlation [22]. The number of DR patients outnumbered the number of practitioners by a large margin. As a result, manual clinical diagnosis or screening takes a long time. To avoid this problem, follow-up scanning is performed regularly, and automated DR identification and intensity

classification are required. Several strategies for detecting retinopathy and classifying its severity and likelihood are presented here [23].

Exudates are the diagnostic indications of diabetic retinopathy, a retina condition caused by long-term diabetes that can lead to eyesight problems if not detected early. The procedure of recognizing and categorizing exudates from a retinal image has been made easier thanks to a medical screening program. The exudates are first segregated using the FCM technique and then transformed into discrete mother wavelets. The classifier is fed the texture textural properties retrieved by the grey-level co-occurrence matrix. The suggested program's efficiency was evaluated by comparing it to the data from the publicly available dataset IDRID. MATLAB was used to formulate and construct a GUI [24].

This research has the proposed texture feature extraction characteristics of the GLDM method (contrast, angular second moments, density, median, and inverse difference moment) feature and feed-forward neural net classifier as a machine learning-based approach for DR detection and evaluation. According to the results of the trials and performance assessment, the suggested methodology had a detection performance of 95% [25].

Diabetes is responsible for 50 deaths per 1000 live births amongst individuals over the age of 70. The identification of diabetes at a preliminary phase and the implementation of a suitable therapy may minimize the visual loss among the sufferers. Once symptoms of DR have been identified, the severity of the disease must be defined to recommend the appropriate treatment. Mild nonproliferative diabetic retinopathy (NPDR), moderate NPDR, severe NPDR, proliferative diabetic retinopathy (PDR), and no DR are the five phases of diabetic retinopathy severity. The techniques and issues associated with DR identification are summarized in this publication [26]. In [27], the authors proposed diabetic retinopathy classification using retinal images through an ensemble learning algorithm. The proposed work includes the following processes: retinal image collection, preprocessing, feature extraction, and feature selection and classification. In preprocessing, the noisy images, duplicate images, and black borders are removed from the images. Tone mapping is used to increase the contrast and luminance in the images. Two sets of features are extracted from images such as the histogram-based feature and GLCM feature extraction. Then, the features are concatenated to select the relevant features. Here, the GA algorithm is used for feature selection. Finally, classification was undertaken by the XGBoost algorithm using the selected features. Here, genetic algorithm (GA) was used for feature selection; it takes a lot of time to select the features, thus increasing feature selection and classification latency.

Early detection of diabetic retinopathy using retinal images for diabetes is presented in [28]. The proposed method includes four processes: preprocessing, segmentation, feature extraction, and classification. The preprocessing includes noise removal and contrast enhancement using histogram equalization (HE). The segmentation is performed by Gaussian derivative and Coye filter, which segments the EX, MA, and HM. The features are extracted from the segmented image and extract features such as EX, MA, and HM values. Finally, SVM is used to classify the images using the extracted features. Here, SVM was used for classification, which takes a lot of time for training when considering larger datasets, thus leading to classification latency.

A histogram equalization method for the early detection of diabetic retinopathy was presented in [29]. The proposed algorithm included three methods: histogram clipping, RIHE-RVE, and RIHE-RRVE, which addressed the issues of the illumination of the retinal images. To avoid enhancement, the histogram clipping algorithm was proposed. The simulation result showed that the proposed method achieved a high performance compared to the other state-of-the-art methods. Here, the histogram equalization method was proposed, however, it is an unselective process that may increase the background noise contrast while decreasing the functional input image.

The authors in [30] proposed CANet to detect diabetic retinopathy and macular edema for diabetes. The proposed work used ResNet50 to produce a feature map with various resolutions including a cross-disease attention network, disease-specific attention module, and disease-dependent attention module. The disease-specific attention module was used

to learn the features of the two diseases. In this stage, the inter special relationship was evaluated to detect the diseases. A disease dependent attention module was used to evaluate the internal relationship between the DR and DME diseases. Here, raw images were considered for training and testing, thus increasing the high false positive rate due to the presence of noise and low contrast, also reducing the detection accuracy.

The authors in [31] proposed a deep learning algorithm to detect diabetic retinopathy disease in diabetic patients. The proposed method included two processes: diagnosing DR severity and the feature extraction of DR. The proposed system hierarchical multitask learning architecture aims to detect both the DR severity and DR feature extraction. Finally, the fully connected layer provides the output, and it considers the hybrid loss, cross-entropy loss, and kappa loss for reducing the errors in the levels of DR severity. The simulation results showed that the proposed model achieved a higher performance using traditional deep learning methods. Here, the traditional deep learning method was used to detect the DR severity levels and feature extraction of DR; however, it generated multiple convolutional layers, thus increasing the complexity and latency.

In [32], the authors proposed a modified contrast enhancement approach from the effective identification of features in detecting diabetic retinopathy and diabetic macular edema. The limitations of conventional contrast limited the adaptive histogram equalization (CLAHE) technique such as the fixed clip limit and region of context, resulting in the inefficient identification of minute features, but can be overcome by implementing modified particle swarm optimization (MPSO) to determine the optimal clip limit and region of context, thereby resulting in the precise identification of features that further help in the accurate detection of diseases. The global best solution of all the operating particles was computed by comparing the output provided by all the particles in the iteration, which resulted in enhanced image contrast. The optimization of the clip limit and region of context was performed by the MPSO algorithm for the purpose of enhancing the contrast of the input image, but the proposed algorithm possessed slow convergence and is stuck in the optimal local solution.

In [33], the authors proposed an approach for the detection of diabetic macular edema in an automatic manner. The macular edema was identified, and the severity of the disease was determined by implementing mathematical morphology. The retinal image was used as the input from the detection process that was carried out. Initially, the preprocessing of the input image was performed from the removal of noise and enhancement of the contrast. Furthermore, the localization of the macula was executed by removing the optic disc and locating the center of the fovea. Then, the exudates in the region of the macula were identified in order to determine the severity of the disease. The removal of artifacts such as reflection due to lighting was removed as a post-processing step to achieve an accurate determination of severity. The detection of the macula in the input retinal image was carried out by using mathematical morphology, but this approach resulted in less accuracy in the detection of the macula region.

In [34], the authors proposed a probability-based construction of the future retinal image in detecting diabetic retinopathy. The difficulty in identifying the future instances of lesions in the retinal image was addressed. Initially, the segmentation of lesions and vessels was carried out to identify the severity of the disease from the input retinal image. Then, the probability of future lesion location was computed by the construction of a probability map. Furthermore, the generated probability map, along with the structure of vessels, was considered for the systemization of future lesions in the retina. This method was found to be effective in predicting future lesions based on the progression of the severity of diseases. The future severity of diabetic retinopathy was determined by using the probability map and the features of the current vessels, but the lack of noise removal in the input image reduced the efficiency of this approach.

### 3. Problem Statement

An input fundus image is used to perform the identification of diabetic retinopathy and diabetic macular edema; however, the accuracy of the system is decreased by the increased false detection rate of the existing techniques. In addition, the following issues are encountered in the best detection of DR and DME, which are listed as:

- **Difficulty in feature differentiation:** The detection of DR and DME is based on various features such as hard exudates, hemorrhages, and micro-aneurysms, but the differentiation of these minute features from each other is a hard task, which degrades the computation of the accurate severity of diseases.
- **Class Overlapping:** Current techniques also consider illness severity; however, the sparse training data for each severity leads to class imbalance issues that degrade the classification accuracy.
- **Inadequate preprocessing:** Using the current methods for effective contrast enhancement with traditional preprocessing leads to difficulties distinguishing features from the background.

In [35], the authors proposed diabetic retinopathy detection using a deep convolution neural network (DCNN) for nonproliferative diabetic retinopathy. The proposed work includes three phases: preprocessing, candidate lesion detection, and candidate extraction. In preprocessing, the image contrast is enhanced using curve transformation. Then, the images are smoothed by a bandpass filter. In the lesion, the detection process includes four stages: optical disc removal, candidate lesion detection, vessel extraction, and preprocessing. In candidate extraction, the micro-aneurysms are detected to measure the coefficient between every pixel using Gaussian kernels. For this, a PCA algorithm was proposed to reduce the dimensionality. Finally, classification was undertaken by DCNN. In this way, the proposed work achieved high accuracy of nonproliferated diabetic retinopathy. The major issues determined in this paper are as follows:

- Here, preprocessing was performed to enhance the quality of the retinal images; however, the retina image still has noise due to the implementation of traditional contrast enhancement techniques, thus reducing the image quality, which leads to a high false detection and reduced detection accuracy.
- DCNN is used for feature extraction and the detection of nonproliferated diabetic retinopathy. Still, DCNN focuses on the whole image for the extraction of features without any particular region of interest, thus increasing the high latency for feature detection.
- The PCA algorithm was used to reduce the dimensionality, but the number of principal components must be selected otherwise it may cause information loss, thus reducing the detection accuracy.

The authors in [36] proposed a data augmentation method to improve the detection rate of proliferative diabetic retinopathy. The NVs were inserted onto pixels located on vessels. Vessel segmentation was performed by Otsu thresholding and the U-Net deep learning algorithm, and then optic disc segmentation was performed. The count of NVs was determined by selecting random values using a threshold. The next process is semi-random blood vessel generation, which is based on the tree structure. This process considers the shape and orientation of the NVs. For the vessel color assignment color, a matrix was proposed that calculates the weighted average of the RGB values of the images. Finally, DR grading and data augmentation was proposed to improve the NVs. Some of the significant problems in this research are as follows:

- Here, the Otsu thresholding method was used for vessel segmentation, which performed well; however, it did not provide an optimal result for noisy images. First, the noise is removed from the images, and then the thresholding is applied; otherwise, this method will fail, thus reducing the performance of vessel segmentation.
- The detection of diabetic retinopathy was carried out by performing the segmentation of neovessels in the retina. However, performing detection based on a single feature results in a high false detection rate.

- Here, the U-Net algorithm was also used for vessel segmentation, which takes a lot of time to learn the vessels from the retinal images at the middle layers, thus leading to high latency.

The authors in [37] proposed the analysis of retinal images to detect eye diseases for diabetes using the deep learning method. The proposed method considered two processes: detection and localization, and the segmentation of localized regions. For localization, the author proposed the FRCNN method, which extracts the features from the images that evaluate the affected portions. For the segmentation process, the author proposed the FKM clustering algorithm. The ground truth was generated for detecting the affected regions during training. Finally, the DME is classified into two classes such as DME and background. The serious issues in this paper are as follows:

- Here, raw images were considered for the localization and segmentation process, thus reducing segmentation and detection accuracy due to low contrast and the presence of noise in the retinal images.
- Faster RCNN was implemented for the extraction of features but the lack of pixel-to-pixel alignment in the region of interest caused misalignment, resulting in the degradation of the detection accuracy.
- The proposed approach was used for diabetic-based disease detection in the eye, but the detection of various diseases from the limited number of trained images resulted in class imbalances.

The authors in [38] presented an efficient framework for the detection of macular edema disease for diabetes. The proposed work used the combination of a deep convolution neural network (DCNN) and a meta-heuristic algorithm for feature extraction and feature selection, respectively. At the stage of feature extraction, the proposed work reduced the feature extraction complexity by reducing the prior knowledge. The SMOTE algorithm was used to perform class imbalance. The generic algorithm and binary particle swarm optimization algorithm were used to select the relevant features. The drawbacks in this paper are as follows:

- Here, the features were extracted from the noisy images, thus reducing the quality of the images and leading to poor feature extraction, thus increasing the macular edema's false detection rate.
- The integration of the genetic algorithm and binary particle swarm optimization was used to determine the subset size. However, implementing these two algorithms increases the complexity and time consumption, thereby increasing the latency.
- DCNN was used for feature extraction and the detection of nonproliferated diabetic retinopathy, however, DCNN focuses on the whole image for the extraction of features without any particular region of interest, thus increasing the high latency for feature detection.

#### 4. Proposed Model

In this research work, we concentrated on accurately detecting the DR and DME from the input fundus images. The severity of the disease is also determined based on the features extracted from the images. Figure 2 shows the architectural view of the proposed work. The description of the dataset is provided below:

The properties of the blood vessels in the retinal image enable the ophthalmologist to assess retinal disease. The presence of lesions on the fundus image is the first sign of diabetic retinopathy. The preprocessing technique is mainly used to remove unwanted noise and enhance some image features.

The fundamental idea underlying OPTICS is to find the points associated by density to extract the cluster structure of a dataset. The approach generates a density-based representation of the data by constructing a reachability graph, an ordered collection of points. Each location in the list has a reachability distance associated with it, which

measures how simple it is to get to that site from other points in the collection. Points with comparable accessibility distances are most likely in the same category.

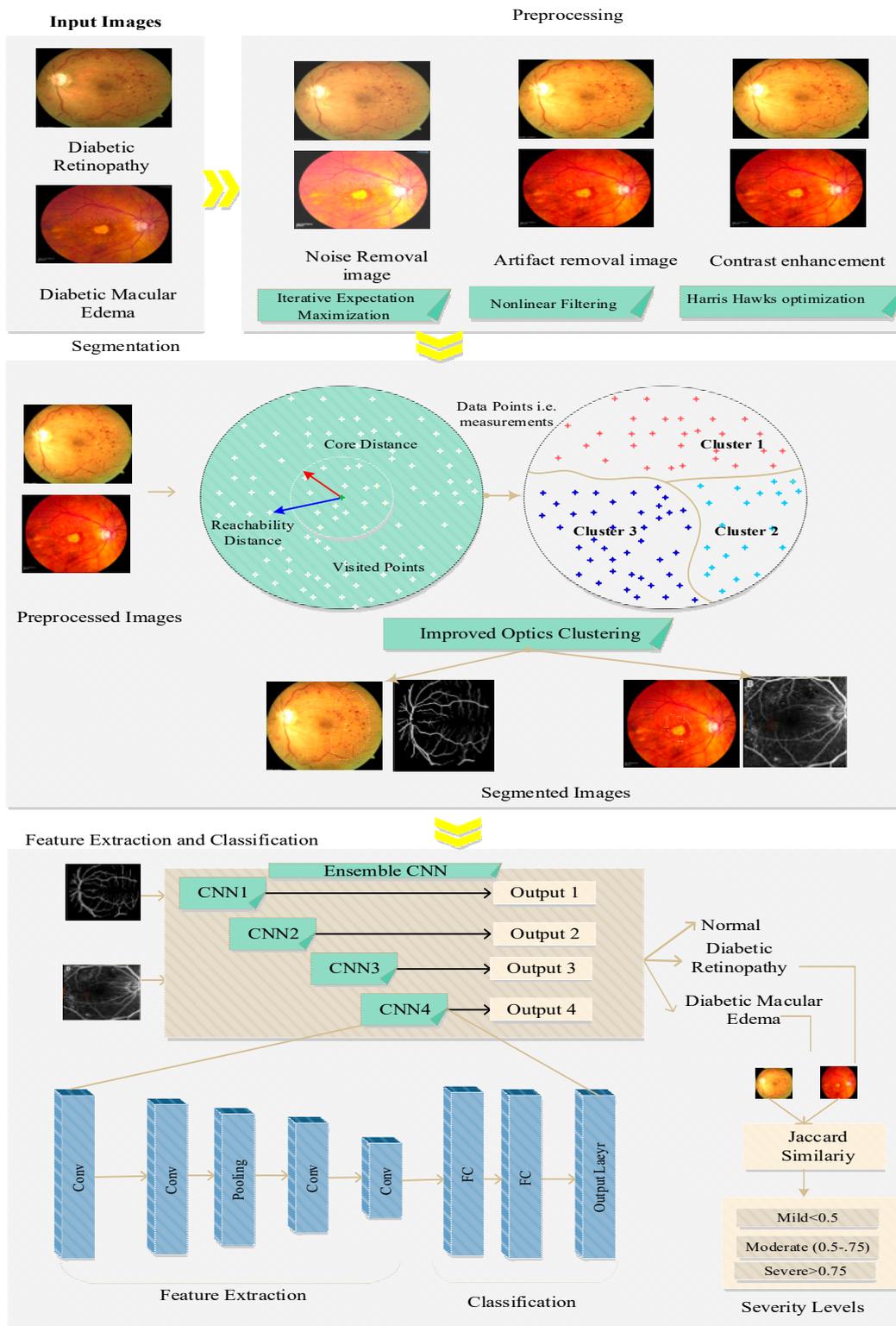


Figure 2. System model.

Before sharing our preprocessed image with CNN, we converted the image to an array and mapped that array's values in the range of 0 to 1 as the epoch was set at 235 to reach a deep network. The initial learning rate was kept at  $1 \times 10^3$ , which is the default value for the

Adam’s optimizer, and the die stack size was 32. We trained our model with more pictures, obtained only a few hundred of images for training, and generated more images from the existing dataset by passing parameters such as the rotating range, width changing range, height changing range, scissors range, zoom range, and pan on image data generator.

The classification of diabetic retinopathy is classified into two types: nonproliferative and proliferative. The term “proliferative” refers to whether the retina has neovascularization (abnormal blood vessel growth). Nonproliferative diabetic retinopathy refers to early illness without neovascularization (NPDR).

Dataset Collection: For accurate prediction of diabetic retinopathy and diabetic macular edema, we applied two kinds of retina fundus images: IDRiD and MESSIDOR. The description of these two datasets is as follows:

- IDRiD: Based on the presence of DR and DME disease, 516 images were loaded in the dataset. In addition, images were acquired through the field of view and stored in JPG format, and the size of each image was 800 KB. This dataset contained 81 color fundus images with the sign of DR. With this dataset, hard exudates (EX), microaneurysms (MA), soft exudates (SE), and hemorrhage (HE) based images are stored.
- MESSIDOR: This dataset was used, whose scope is to develop the DR and DME detection of images. In total, 1200 eye fundus images were used with the multiple pixel rates of  $1440 \times 960$ ,  $2240 \times 1488$ , and  $2304 \times 1536$ .

The following steps implement a prediction of DR and DME.

#### 4.1. Preprocessing

This is an initial step for DR and DME detection. To enhance the information for the disease diagnosis system, it is necessary to use some of the preprocessing steps as follows:

(a) Noise Filtering: Fundus images are cropped by salt and pepper noises, which are removed from the input images using the iterative expectation maximization (IEM) approach. In this approach, uncertainty is overwhelmed by using IEM variables. Noise is removed in the zig-zag trajectory and edge, and the corner position of the image is denoised using IEM variables. A dynamic threshold was computed and adjusted accordingly for noise removal since the acquisition of each image was different with their resolution.

The proposed inverse dual tree initial ranging (IDTIR) procedure uses the iterative expectation maximization (IEM) algorithm. The IEM algorithm is an iterative method that effectively estimates the parameters of the statistical model. In the IEM algorithm, two major steps are executed to estimate the parameters accurately. These steps can be explained as follows:

- E-step—This step determines the current estimate of parameters by creating a function for the expectation of log-likelihood. The expectation step is the base of the proposed IEM algorithm.
- M-step—This step is the final step that computes the parameters in such a way that the expected log-likelihood function can be maximized (i.e., the likelihood function determined in the E-step is maximized to calculate the parameters).

The above two steps were iteratively executed to determine the final parameters. Let  $\omega_{\mathcal{Y}, \mathcal{L}}$  be the parameter vector, and it can be represented as  $\omega_{\mathcal{Y}, \mathcal{L}} = [h_{\mathcal{Y}, \mathcal{L}}, T_{\mathcal{Y}, \mathcal{L}}] \in \mu$  for the  $\mathcal{L}^{\text{th}}$  active channel path of the given ranging code. The set of parameters is represented as  $\mu$ . The latest estimated parameter set is denoted as  $\hat{\mu}$  and can be formulated as follows,

$$\hat{\omega}_{\mathcal{Y}, \mathcal{L}} = [\hat{h}_{\mathcal{Y}, \mathcal{L}}, \hat{T}_{\mathcal{Y}, \mathcal{L}}] \in \hat{\mu} \tag{1}$$

#### E-Step Computation

In this step, the expected value is calculated as

$$G(\omega_{\mathcal{Y}, \mathcal{L}} | \hat{\mu}) \triangleq \ln \mathfrak{P}(Y | \omega_{\mathcal{Y}, \mathcal{L}}, \hat{\mu}) \alpha - \|(Y - \hat{K}_{\mathcal{Y}, \mathcal{L}}) - h_{\mathcal{Y}, \mathcal{L}} \Gamma(T_{\mathcal{Y}, \mathcal{L}}) \mathcal{C}_{\mathcal{Y}}\|^2 \tag{2}$$

Here,

$$\hat{K}_{\mathcal{D}}, \triangleq \sum_{N=1}^N \sum_{s=1}^{N_{\mathcal{D}}} \hat{h}_{N,s} \Gamma(\hat{T}_{N,s}) \mathcal{C}_N - \hat{h}_{\mathcal{D},\mathcal{L}} \Gamma(\hat{T}_{\mathcal{D},\mathcal{L}}) \mathcal{C}_{\mathcal{D}} \tag{3}$$

M-Step Computation

In this step, the expected value is maximized as follows:

$$\hat{\omega}_{\mathcal{D},\mathcal{L}} = \operatorname{argmax}\{G(\omega_{\mathcal{D},\mathcal{L}}|\hat{\mu})\} \tag{4}$$

After parameter estimation, the estimated parameters are updated in the parameter vector. These two steps are executed until the terminating condition is met. The channel coefficient is derived from the parameter vector by letting the derivative equal zero with the fixed timing offset.

(b) Artifact Removal: Blurriness, poor edges, and illumination are called artifacts, which are removed using the nonlinear diffusion filtering algorithm, which eliminates all kinds of artifacts and ensures the image quality in terms of illumination correction and edge preservation.

(c) Contrast Enhancement: Low contrast is one of the important issues of image classification. In this work, we considered contrast enhancement as an optimization problem with the intention of optimizing the pixel values based on the contrast level of the input image. To enhance the contrast level of the input image, we proposed the Harris hawks optimization algorithm, which improves the performance of the image brightness.

H2O is a recently developed meta-heuristic algorithm that performs better in solving optimization problems. The H2O algorithm mimics the cooperative strategy and chasing style of the Harris hawks in nature. Since it has an intelligent searching strategy and fast convergence rate, it works better than the conventional genetic algorithm, particle swarm optimization algorithm, etc. Due to the benefits of the H2O algorithm, it was adapted for contrast enhancement using the pixel intensity rate in the proposed system. The proposed H2ORSS algorithm detects the optimum threshold value for replacing the pixel intensity values with normal ones. The proposed H2ORSS algorithm involves three major processes: initialization, fitness value estimation, and update of hawks.

Initially, the image matrix is initialized as hawks with the population size of PS. For each hawk ( $\mathbb{X}_i$ ) in the population, the fitness function is estimated. The fitness function is determined in terms of the pixel intensity, neighbor intensity, and resolution. The fitness function of the  $i$ th hawk is expressed as follows, Once the fitness is computed for all hawks, then three sequential phases are executed to select the optimal solution.

Phase 1: Exploration Phase

This phase relies on waiting, searching, and detecting prey. In every step, each Harris hawk is considered as the alternative solution. Based on the fittest solution, the position for each Harris hawk is updated as follows:

$$\mathbb{X}(\operatorname{iter} + 1) = \begin{cases} \mathbb{X}_{\text{rand}}(\operatorname{iter}) - r_1 |\mathbb{X}_{\text{rand}}(\operatorname{iter}) - 2r_2 \mathbb{X}(\operatorname{iter})| & \text{if } \sigma \geq 0.5 \\ \mathbb{X}_{\text{p}}(\operatorname{iter}) - \mathbb{X}_{\text{a}}(\operatorname{iter}) - r_3(l_b + r_4(u_b - l_b)) & \text{if } \sigma < 0.5 \end{cases} \tag{5}$$

The location of the hawks in the next iteration is denoted as  $\mathbb{X}(\operatorname{iter} + 1)$  and  $r_1, r_2, r_3, r_4$  are the present location vectors of the hawks. Furthermore,  $\sigma$  is the random number selected in the range of 0 and 1, and  $u_b, l_b$  are the upper bound and lower bound, respectively. The average location of hawks ( $\mathbb{X}_{\text{a}}(\operatorname{iter})$ ) can be estimated from the following expression:

$$\mathbb{X}_{\text{a}}(\operatorname{iter}) = \frac{1}{\text{PS}} \sum_{i=1}^{\text{PS}} \mathbb{X}_i(\operatorname{iter}) \tag{6}$$

Phase 2: Transformation from Exploration to Exploitation

Next, the algorithm transforms the state from exploration to exploitation. In this transformation, the energy of the prey is dissipated due to evading behavior. The energy level of the prey is ( $E_p$ ), which is expressed as follows:

$$E_p = 2E_0 \left( 1 - \frac{\text{iter}}{T_m} \right) \tag{7}$$

Here,  $E_0$  is the initial state energy of the prey and  $t_m$  is the maximum iteration. By varying the tendency of  $E_0$ , the state of the prey can be judged.

Phase 3: Exploitation

After judging the state of the prey, the Harris hawks attack the selected prey. In practice, the prey changes the evading behavior, frequently changing the attacking behavior. Four strategies are constructed in the H2ORSS algorithm for attacking prey based on evading behavior. Here, soft besiege and hard besiege are the basic strategies to attack the prey, which is decided as follows: If  $|E_p| \geq 0.5$ , then a soft besiege occurs, and if  $|E_p| < 0.5$ , then a hard besiege occurs.

Soft Besiege

This attacking strategy is selected when  $E_p \geq 0.5$  and  $r \geq 0.5$  by Harris hawks. This soft besiege attacking strategy is modeled as follows:

$$X(\text{iter} + 1) = \Delta X(\text{iter}) - E_p |\mathfrak{J} X_p(\text{iter} - X(\text{iter}))| \tag{8}$$

Here,  $\mathfrak{J}$  is the jump intensity of the prey during the evading process, and it is given as  $\mathfrak{J} = 2(1 - r_5)$  and  $\Delta X(\text{iter})$  represents the difference in the location vector of prey in each iteration. This difference is estimated by using the following expression:

$$\Delta X(\text{iter}) = X_p(\text{iter}) - X(\text{iter}) \tag{9}$$

Hard Besiege

If  $|E_p| < 0.5$  and  $r \geq 0.5$ , the hard besiege strategy is selected to attack the prey. In general, these probability values show that the prey's energy is dissipated and has low evading energy. In this case, the position of Harris hawks is updated by the following equation:

$$X(\text{iter} + 1) = X_p(\text{iter}) - E_p |\Delta X(\text{iter})| \tag{10}$$

Soft Besiege with Progressive Rapid Dives

This strategy is selected when the prey has sufficient energy to evade form the attack. This situation is explained as  $|E_p| \geq 0.5$  and  $r < 0.5$ . Based on this behavior, the next position of the hawks is updated as follows:

$$\mathfrak{J} = X_p(\text{iter}) - E_p |\mathfrak{J} X_p(\text{iter} - X(\text{iter}))| \tag{11}$$

As this strategy involves progressive dives, the hawk's dive is formulated as follows:

$$Z = \mathfrak{J} + \mathcal{B} * \text{lf}(d) \tag{12}$$

where  $\mathcal{B}$  represents the random vector;  $\text{lf}(d)$  represents the levy flight with the dimension  $d$ . Thus, the next position is updated as follows:

$$(\text{iter} + 1) = \begin{cases} \mathfrak{J} & \text{if } f(\mathfrak{J}) < f(X(\text{iter})) \\ Z & \text{if } f(Z) < f(X(\text{iter})) \end{cases} \tag{13}$$

Hard Besiege with Progressive Rapid Dives

This situation is defined as the prey has not sufficient energy to escape. This situation is formulated as  $|E_p| < 0.5$  and  $r < 0.5$ . The rule for this situation is formulated as follows:

$$\mathbb{X}(\text{iter} + 1) = \begin{cases} \mathcal{Y} & \text{if } f(\mathcal{Y}) < f(\mathbb{X}(\text{iter})) \\ \mathcal{Z} & \text{if } f(\mathcal{Z}) < f(\mathbb{X}(\text{iter})) \end{cases} \quad (14)$$

Here,  $\mathcal{Y}$  is estimated using the upcoming Equation,

$$\mathcal{Y} = \mathbb{X}_p(\text{iter}) - E_p |\mathfrak{F}| \mathbb{X}_p(\text{iter} - \mathbb{X}_a(\text{iter})) \quad (15)$$

Based on the above rules, the position of each hawk is updated, and the optimal solution is derived over iteration. Finally, the optimum threshold value was computed for the prediction of contrast values throughout the images.

Algorithm 1 deals with Generalized Linear Model (GLM), which is used for regression and classification tasks, is one of the key algorithms in H2O. GLM is a versatile and effective modeling approach that can deal with different data kinds and distributions.

---

#### Algorithm 1 Pseudocode for H2O

---

Input : PS,  $\text{Max}_{\text{ite}}$

Output: Optimal Threshold

---

Begin

Initialize  $\rightarrow$  hawks population  $\mathbb{X}_i$  (C.U.<sub>i</sub>);

While (Stopping Condition Not Met) do

  Compute  $\rightarrow$  fitness function

  For ( $\mathbb{X}_i \in \mathbb{X}_{PS}$ ) do

    Update  $\rightarrow E_o$  and  $\mathfrak{F}$ ;

    Update  $\rightarrow E_p$  using Equation (8);

  End For

  If ( $E_p \geq 1$ ) Then

    Update position using Equation (9);

  End If

  If ( $E_p < 1$ ) Then

    If ( $r \geq 0.5 \ \&\& \ |E_p| \geq 0.5$ )

      Update  $\rightarrow$  position using Equation (10);

    Else If ( $r \geq 0.5 \ \&\& \ |E_p| < 0.5$ ) Then

      Update  $\rightarrow$  position using Equation (11);

    Else If ( $r < 0.5 \ \&\& \ |E_p| \geq 0.5$ ) Then

      Update  $\rightarrow$  position using Equation (12);

    Else If ( $r < 0.5 \ \&\& \ |E_p| < 0.5$ ) Then

      Update  $\rightarrow$  position using Equation (13);

  End If

  End If

  End While

End

---

#### 4.2. Blood Vessel Segmentation

Blood vessels are important in computing the image intensity, edges, texture, and other analyses of image features. Analyzing the diagnosis over the segmented area increases the accuracy and precision rate of any disease. Hence, the optic disk is removed from the contrast-enhanced image, and then the blood vessels are extracted using improved mask RCNN, in which ROI alignment is the first step that predicts the region of interest from the input image. In this work, pixel-wise softmax was applied for accurate segmentation of blood vessels, which was better than the CNN, RNN, RCNN, and DCNN algorithms [39,40].

OPTICS clustering stands for ordering points to identify clustering structure. It is more similar to DBSCAN clustering. OPTICS algorithm includes two measurements, which are defined as follows,

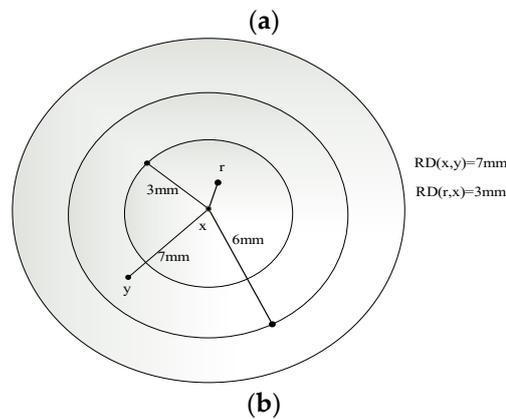
- Core distance: This represents the minimum values of the radius essential to classify the given point as a core point. If the considered point is not a core point, then its core distance is indeterminate.
- Reachability distance: This is represented with respect to another cluster data point. The reachability distance between two points (x,y) is the highest of the core distance and then the Euclidean distance between the two points (x, y). The reachability distance is not defined if the y point is not a core point. Figure 3 represents the calculation of the reachability distance. The general procedure of M-OPTICS is defined as follows:

---

```

OPTICS (SetofObject, ε, Minpts, ordered file)
OderedFile.open()
For i from 1 to SetofObjects. Size do
object:=SetofObjects.get (i)
If nor object.processed then
Expand cluster order (SetofObjects, object, ε, Minpts, ordered file)
Ordered file.close()
end
    
```

---



**Figure 3.** (a) OPTICS clustering algorithm. (b) Reachability distance.

Next, the proposed M-OPTICS explanation is defined as follows: M-OPTICS considers three important conditions: maximum radius, distance, and number of cluster points including the core distance, core points, and reachability distance. In the M-OPTICS algorithm, the point P is known as the core point when the point is on MinPts. The reachability distance and core distance calculations are given as follows:

$$CD(o) = \begin{cases} \infty & , |o, \epsilon| < MinPts \\ MinPts - D(o), & otherwise \end{cases} \tag{16}$$

$$R.D.(p, o) = \max(CD(o), D(p, o)) \tag{17}$$

where p represents the object and o represents the center point. The core distance represents the lowest value, which is the radius. From the radius, the core point is different. RD represents the reachability distance, which is estimated as the highest core distance, and ε represents the radius of the data. The reachability distance data are clustered separately. The data similarities were measured by Jaccard similarity, which calculates the similarity between a finite set of samples. The calculation of the Jaccard similarity is defined as follows:

$$JD = 1 - J_{\mu}(A, B) \tag{18}$$

$$J_{\mu}(A, B) = \frac{\mu(A \cap B)}{\mu(A \cup B)} \tag{19}$$

where A and B represent the two points obtained from the blood vessels.

The CNN-based ensemble learning model was incorporated due to two major unique features: shared weights and local connections. The extraction of features from the input data using convolutional layers and determining the relationship between the obtained features using the pooling layers was implemented, which can be formulated as:

$$a_q^l = \sum_{p=1}^V a_p^{l-1} * J_{pq}^l + y_q^l \tag{20}$$

where  $J_{pq}^l, y_q^l$  denote the trainable parameters, and V denotes the input features. The output provided by the nonlinear layer is computed as:

$$x^d = f(v^d) \tag{21}$$

where function  $f(v^d)$  denotes the output of the rectified linear unit. The performance of the model can be further improved by executing batch normalization. The dataset comprising of fused images of R dimensions comprised of a T number of training samples can be denoted as  $H = \{(h_i, c_i) | 1 \leq i \leq T\}$ , where the classes are  $c_i \in Cl = \{1, 2, \dots, M\}$  and the maximum count of classes is denoted as M. In the ensemble model, each model's training is performed randomly. The input of each CNN will be  $\tilde{H} = \{(\tilde{h}_i, c_i) | 1 \leq i \leq T\}$ , which comprises r "R feature subspaces that are randomly selected.

For instance, i and j are two identical features with dimension d, and for that similarity function  $sim_{p_d}^{(i,j)}$ , which is computed by:

$$sim_{p_d}^{(i,j)} = \frac{vector_i \times vector_j}{\|vector_i\| \times \|vector_j\|} \tag{22}$$

For a different number of CNN layers and the operations involved in this study, computational complexity was evaluated, which is described as follows:

$$P_{v_{s(i)}} = (1 - \mu) \times O(N) + \mu \times O(N) = O(N) \tag{23}$$

where  $O(N)$  represents the sum of iterations for performing the feature extraction and classification  $\mu \in [0, 1]$  and then S.S.<sub>upd</sub> with respect to the  $f_x$  as follows:

$$S.S._{upd} = \arg(\max_{i_p \in id_{x_i}} (f_x(P_{V_{sn}}))) = O(N) \tag{24}$$

where  $O(N)$  represents the sum of iterations for S.S.<sub>upd</sub>, which provides the near optimum feature matches from the trained set. Once the features are extracted, they are then updated by the presented method.

The output obtained from each CNN is denoted as  $x = CNN(\tilde{H})$ ; the collective outputs obtained from the individual CNN are denoted as  $X = \{x_1, x_2, \dots, x_L\}$ , where L denotes the ensemble's size, and the global output of the ensemble model is obtained by using weighted averaging of the output of the individual CNN. The weighted average of the output of the individual CNN is formulated as:

$$G = \frac{\sum_{j=1}^T w_j s_j}{T} \text{ with } w_j \geq 0$$

$$\sum_{j=1}^T w_j = 1 \tag{25}$$

where  $s_j$  denotes the score and  $w_j$  denotes the weight of the  $j - th(j = 1, 2, 3)$  model.

The classifier diversity between any two CNN models is computed as:

$$C_D(i,j) = \frac{T_w}{N_T} \tag{26}$$

where  $N_T$  and  $T_w$  denote the total number of test samples and the difference of results caused by the samples. The diversity of the ensemble model is computed as the average of the classifier diversities, which can be formulated as:

$$E_D = \frac{\sum_{i=1}^M \sum_{j=1}^M C_D(i,j)}{L}, i \neq j \tag{27}$$

where  $E_D$  denotes the diversity of the ensemble model, and  $C_D$  denotes the classifier diversity. The classification output achieved from the weighted averaging of the individual CNN models possessed increased accuracy than the individual CNN models. Figure 4 presented the SMDTR-CNN-based land cover classification for identifying normal, DR and DME.

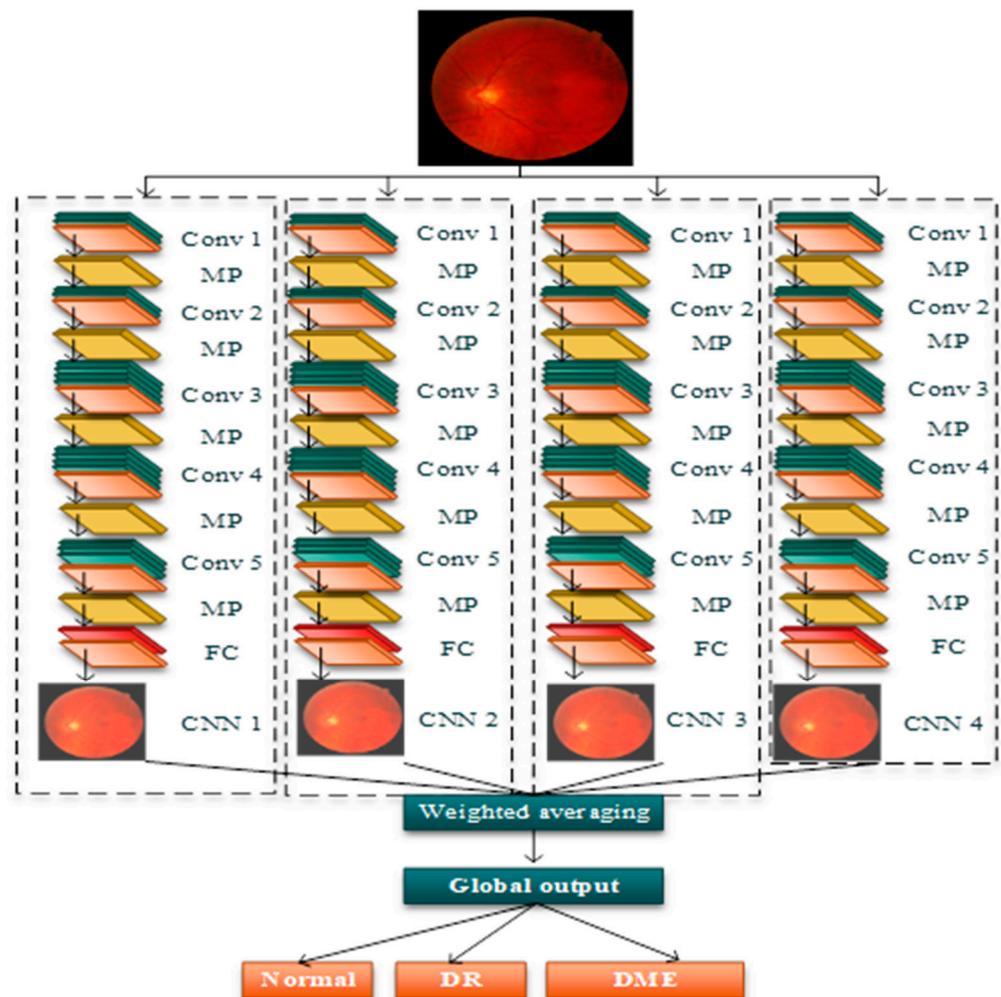


Figure 4. SMDTR-CNN-based land cover classification.

Table 1 addresses the ensemble deep learning model below with their filters, filter size, stride, padding, and output image size. A CNN’s fundamental building block is a convolutional layer and includes a series of filters, the parameters of which must be learned throughout the training process. The filters are often smaller in size than the real image. The pooling layer’s function is to lower the spatial size of the representation in order to reduce the number of parameters and calculations in the network; it operates independently on each feature map (channels). Maximum pooling and average pooling are the two types

of pooling layers. Max pooling is a procedure commonly used for the individual CNN convolution layers listed below when they are added to a model. Maxpooling minimizes the picture dimensionality by lowering the number of pixels in the preceding convolution layer's output. The rectified linear activation unit (ReLU) is one of the few milestones in the deep learning revolution. It is basic, but it is superior to the activation features of its predecessors such as sigmoid or tanh.

**Table 1.** Layers of the convolutional neural network.

Operational Layer		Filters	Filter Size	Stride	Padding	Output Image Size
Preprocessed image		-	-	-	-	$224 \times 224 \times 3$
Convolutional layer (2 times)	Convolutional	64	$3 \times 3 \times 3$	$1 \times 1$	$1 \times 1$	$224 \times 224 \times 64$
	ReLU	-	-	-	-	$224 \times 224 \times 64$
Pooling layer	Max pooling	1	$2 \times 2$	$2 \times 2$	0	$112 \times 112 \times 64$
Convolutional layer (2 times)	Convolutional	128	$3 \times 3 \times 64$	$1 \times 1$	$1 \times 1$	$112 \times 112 \times 128$
	ReLU	-	-	-	-	$112 \times 112 \times 128$
Pooling layer	Max pooling	1	$2 \times 2$	$2 \times 2$	0	$56 \times 56 \times 128$
Convolutional layer (4 times)	Convolutional	256	$3 \times 3 \times 128$	$1 \times 1$	$1 \times 1$	$56 \times 56 \times 256$
	ReLU	-	-	-	-	$56 \times 56 \times 256$
Pooling layer	Max pooling	1	$2 \times 2$	$2 \times 2$	0	$28 \times 28 \times 256$
Convolutional layer (4 times)	Convolutional	512	$3 \times 3 \times 256$	$1 \times 1$	$1 \times 1$	$28 \times 28 \times 512$
	ReLU	-	-	-	-	$28 \times 28 \times 512$
Pooling layer	Max pooling	1	$2 \times 2$	$2 \times 2$	0	$14 \times 14 \times 512$
Convolutional layer (4 times)	Convolutional	512	$3 \times 3 \times 512$	$1 \times 1$	$1 \times 1$	$14 \times 14 \times 512$
	ReLU	-	-	-	-	$14 \times 14 \times 512$
Pooling layer	Max pooling	1	$2 \times 2$	$2 \times 2$	0	$7 \times 7 \times 512$
Inner product layer	Fully connected	-	-	-	-	4096
	ReLU	-	-	-	-	4096

## 5. Results and Discussion

The E-CNN performance was estimated with the accuracy, precision, recall, F-score, error rate, and computational time.

### 5.1. Accuracy

Accuracy is defined as the ratio of the received input image inventive classification scheme by the assessed classification scheme, which can be formulated as:

$$A = \frac{\mathcal{T}_1 + \mathcal{T}_2}{\mathcal{T}_1 + \mathcal{T}_2 + \mathcal{F}_1 + \mathcal{F}_2} \quad (28)$$

From the above equation,  $\mathcal{F}_1, \mathcal{F}_2$  denote the false positive and false negative values, respectively; and  $\mathcal{T}_1, \mathcal{T}_2$  denote the true positive and true negative values, respectively. Accuracy is the significant metric for calculating the performance of the system.

### 5.2. Precision

Precision is computed by the ratio of excluding the significant classification result from the overall classification outcome. The meticulousness of the system can be measured using precision, which can be formulated as:

$$\mathcal{P} = \frac{\mathcal{T}_2}{\mathcal{T}_1 + \mathcal{F}_1} \quad (29)$$

### 5.3. Recall

The recall is defined as the ratio of excluding the same classification result to the recovered results. The recall is used for measuring the comprehensiveness of the system, which can be formulated as:

$$\mathcal{R} = \frac{\mathcal{T}_1}{\mathcal{T}_1 + \mathcal{F}_2} \quad (30)$$

### 5.4. F-Score

The F-score is computed by using the parameters of recall and precision by jointly assessing them. The results accuracy can be computed using F-score, which can be formulated as:

$$F_S = \frac{2 * \mathcal{P} * \mathcal{R}}{\mathcal{P} + \mathcal{R}} \quad (31)$$

### 5.5. Computation Time

Computation time is the amount of time needed to complete a computational operation. Computation time is calculated by calculating the time elapsed between the classification completion and computation. The system's efficacy is assessed in terms of computation time. It is appreciated whether the study obtained a greater accuracy with better precision of outcome in a shorter computing period.

### 5.6. Error Rate

In Table 2, the results analysis of all models is furnished in the numerical form for better understanding. The error rate is defined as the ratio of errors in the sample to the overall samples. The error rate is used to determine the system's performance. A good system has a much lower error rate, which can be formulated as:

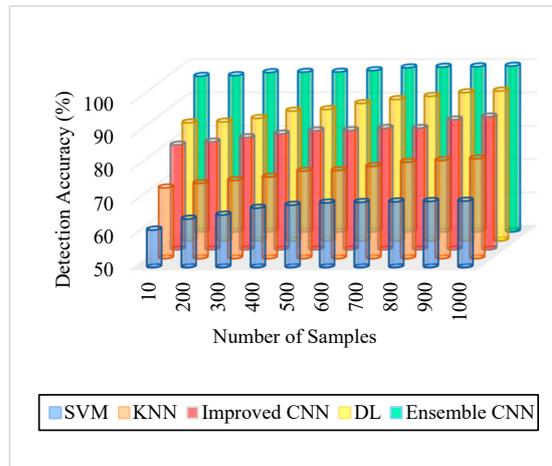
$$\text{Error Rate} = \frac{\text{No of Errors}}{\text{No of Samples}} \quad (32)$$

**Table 2.** Results analysis of Figures 5–10.

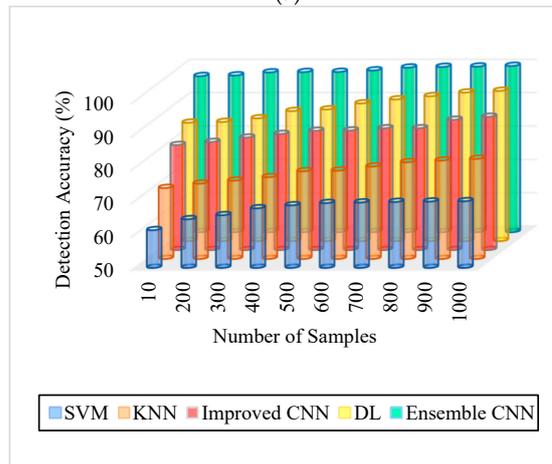
Figures	SVM	KNN	Improved CNN	DL	E-CNN
Figure 5. Detection Accuracy. (a) IDRiR	61.3–69.96%	65–71%	79–83%	83–91%	94–98%
Figure 5. Detection Accuracy. (b) Messidor	54–59%	65–71%	79–84%	84–92%	96–98%
Figure 6. Precision. (a) IDRiR	60.5–69%	70–77.5%	80.12–87.5%	84–91.25%	92.5–97%
Figure 6. Precision. (b) Messidor.	60.5–69%	70–77.5%	80.12–87.5%	84–91.25%	92.5–97%
Figure 7. Recall. (a) IDRiR	60.5–67.5%	70.5–76.8%	80.25–88%	82.5–88.5%	91.2–97.5%
Figure 7. Recall. (b) Messidor.	60.5–67.5%	70.5–76.8%	80.25–88%	82.5–88.5%	91.2–97.6%
Figure 8. F-Score. (a) IDRiR	61.25–70%	70.69–79.5%	81–87.5%	85–91.2%	93–98.5%
Figure 8. F-Score. (b) Messidor	61.25–70%	70.69–79.5%	81–87.5%	85–91.2%	93–98.5%
Figure 9. Computation Time. (a) IDRiR (seconds)	11	14	10.5	8.8	2.6
Figure 9. Computation Time. (b) Messidor (seconds)	14	11	9.5	8.2	2.4
Figure 10. Error Rate. (a) IDRiR	0.985	0.774	0.865	0.923	0.15
Figure 10. Error Rate. (b) Messidor	0.99	0.792	0.865	0.923	0.19

As can be seen in Figures 5–10, we evaluated the proposed E-CNN to various state-of-the-art approaches such as SVM, KNN, enhanced CNN, and deep learning (DL). When analyzing performance, the optic disk (OD) is eliminated because it is a non-lesion area. The numerical findings suggest that our proposed E-CNN was superior. E-CNN had a mean

accuracy of 99.84 percent, which was 4.38 percent greater than the benchmark. Although its effectiveness was equivalent to that of the Messidor database, it performed poorly in blood vessel segments.

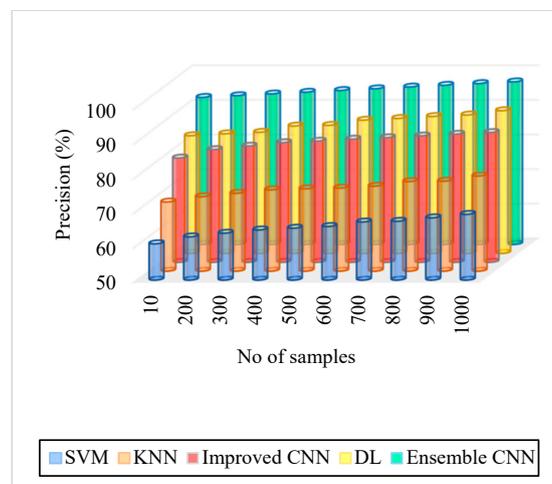


(a)



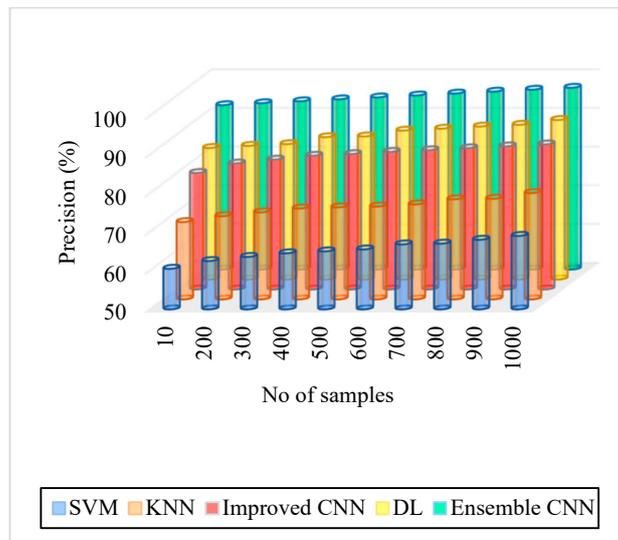
(b)

Figure 5. Detection accuracy (a). IDRiR and (b). Messidor.



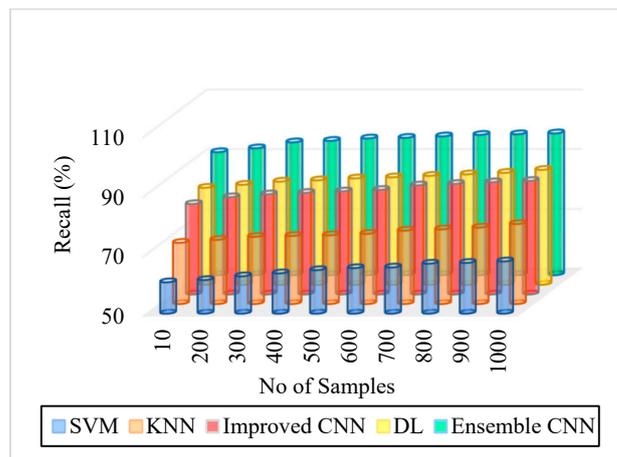
(a)

Figure 6. Cont.

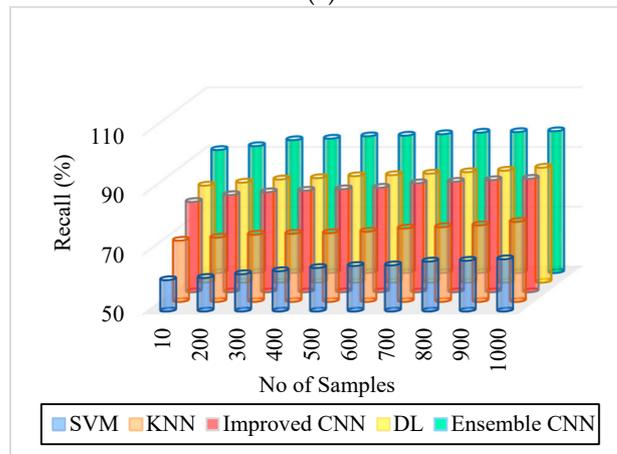


(b)

Figure 6. Precision. (a) IDRiR and (b) Messidor.

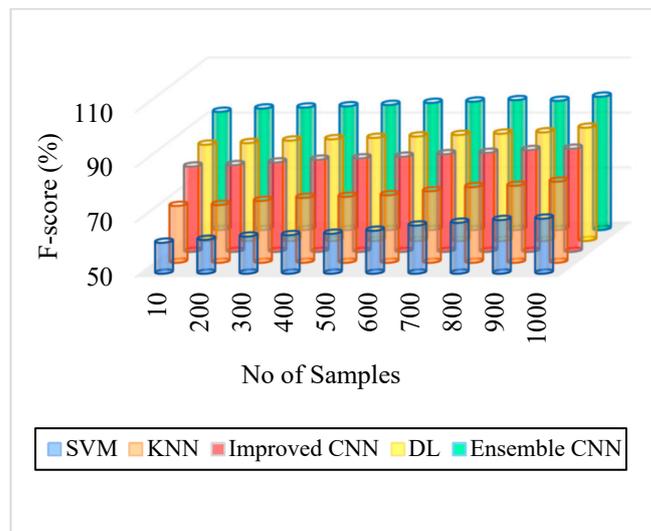


(a)

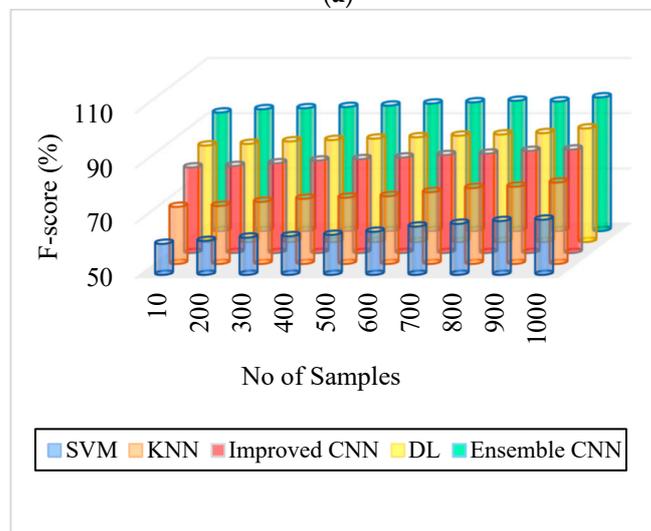


(b)

Figure 7. Recall. (a) IDRiR and (b) Messidor.

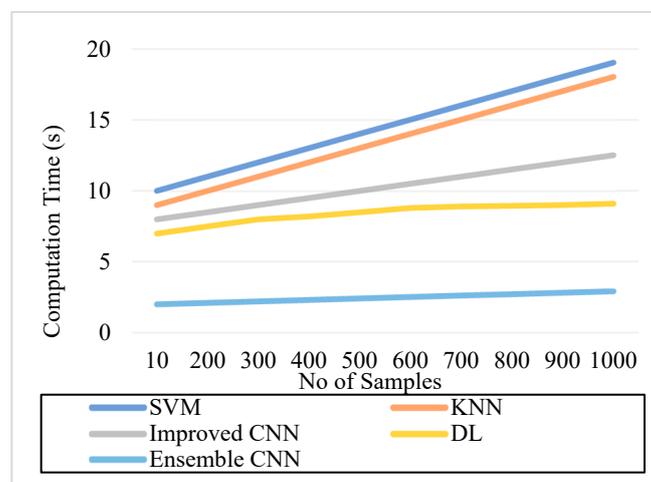


(a)



(b)

Figure 8. F-Score. (a) IDIRiR and (b) Messidor.



(a)

Figure 9. Cont.

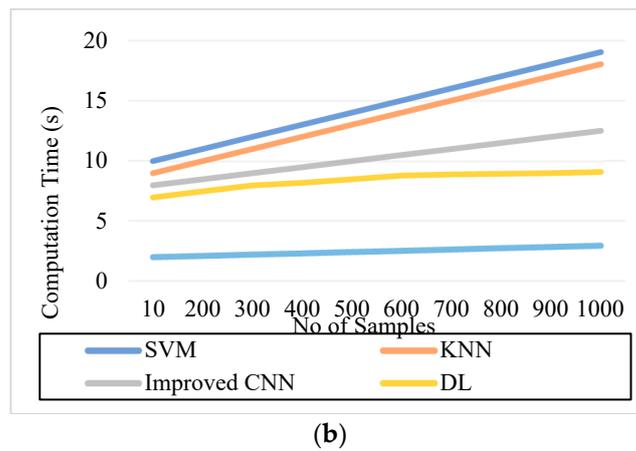


Figure 9. Computation time. (a). IDRiR and (b) Messidor.

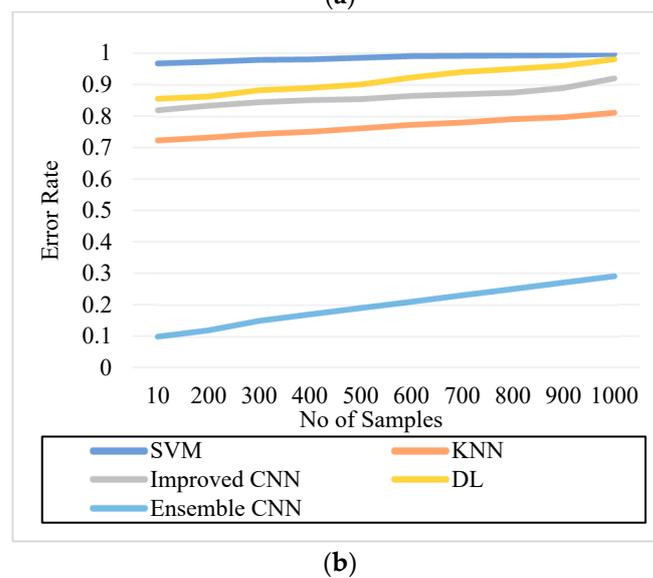
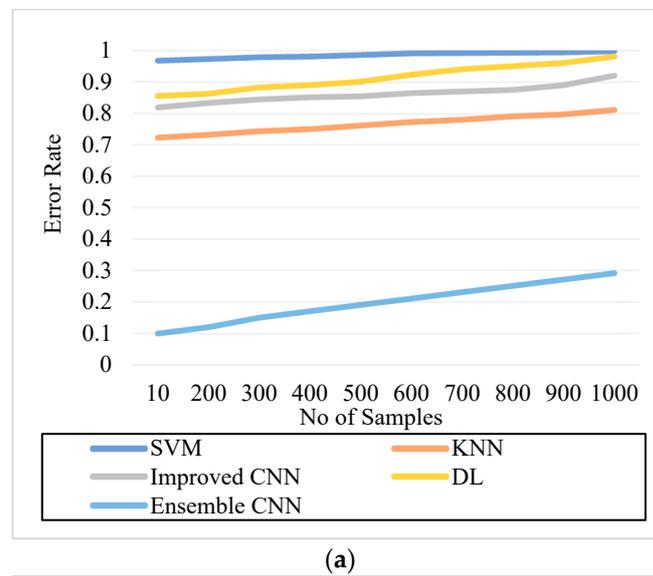


Figure 10. Error Rate. (a) IDRiR and (b) Messidor.

Furthermore, as shown in Figures 5–10, the accuracy, precision, recall, f-score, computational time, and error rate produced by E-CNN were much better than those acquired by

other approaches. The difficulty of misclassification was exacerbated in the lesions DR and DME due to fewer samples; however, our proposed technique could still meet this obstacle. The quantitative class labels are also shown in Figure 5 to further illustrate the suggested strategy's efficiency. In the DR lesion segmentation challenge, one can see that the E-CNN was much more accurate and robust. We also performed an ablation experiment to prove the accuracy of the proposed E-CNN. The SVM is referred to as the baseline approach for convenience. The suggested strategy has been demonstrated to generate significant improvements over the baseline regarding four targets, as shown in Figures 5–10. The addition of preprocessing also improved the performance. The mean precision improved by 3.83 percent in comparison to the benchmark. Our proposed technology, in particular, can be simply integrated into other encoder–decoder networks, which we wish to conduct soon. Additionally, the proposed E-CNN achieved the greatest accuracy values in DR and DME diagnosis, demonstrating the efficacy of our proposed method.

In this work, ensemble convolutional neural networks (ECNNs) were used to classify images of diabetic retinopathy. A recently developed meta-heuristic method, the Harris hawks optimization (HHO) algorithm, was used to optimize the ECNN hyperparameters. Then, the Harris hawks optimization technique was used to improve the feature extraction and classification processes to obtain the most significant features. Compared to previous systems, the deep learning model provides extremely satisfactory results regarding the specificity, precision, accuracy, and recall.

## 6. Conclusions

All of the studies on the DR classification issue can be divided into two groups. The first is a binary DR diagnosis in which the individual possesses or does not. The problem with this technique is that after we realize a person has DR, we cannot tell how serious the disease is. Multi-class identification is the answer to this challenge. As previously mentioned, we classified DR into five classes or phases using multi-class classification. However, almost all of the associated studies, particularly in the early stages of DR, have been unable to appropriately define every one of the stages of DR. It is critical to identify the DR at a very early stage to treat the disease, as treating the disease at a much later date is challenging and can result in death. To our understanding, no other study has employed the IDRiR and Messidor databases to identify the milder phases of DR that we used in our study. Our approach outperformed the present advancements in detecting the mild stage. Furthermore, no one else has demonstrated the impact of a balanced dataset in previous research. The unbalanced dataset may have caused the classification accuracy to be skewed. The network can be trained on features correctly when samples in the classes are evenly distributed such as in a balanced dataset; however, in the case of asymmetrical distributions, the network performs for heavily tested classes. Furthermore, the present CNN architectures for DR identification do not consider the impact of varied hyperparameter tweaking (meta-learning) as well as its consequences. In the future, we plan to use some other deep-learning techniques for DR and DME disease classification.

Recently, CNN-based methodology has been considered to learn features for classification. However, tuning non-trainable hyperparameters for such networks is manual, intuitive, and non-trivial. In the future, a technique based on DR and DME will be proposed to adjust the CNN architecture parameters. The convolution and pooling layer number, the kernel number, and the kernel size of the convolution layer are determined by the upcoming proposed technique. Therefore, the number of untrainable hyperparameters can be reduced. There are some challenges in adapting DR and DME to a CNN. Based on the dimension of the input image, the maximum and minimum sizes of the kernel must be specified for clear classification.

**Author Contributions:** Software, S.S.; Validation, M.S.; Formal analysis, S.R.; Investigation, S.I.; Writing—review & editing, J.-H.C.; Data curation, N.A.A.; Resources, A.E.; Writing—original draft, S.K.R. All authors have read and agreed to the published version of the manuscript.

**Funding:** This work was supported by “Human Resources Program in Energy Technology” of the Korea Institute of Energy Technology Evaluation and Planning (KETEP), granted financial resources from the Ministry of Trade, Industry & Energy, Republic of Korea. (No. 20204010600090). The funding of this work was provided by Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R410), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** For this research work, datasets were taken from the kaggle repository site, available online at [kaggle.com/datasets/mariaherrerot/idrid-dataset](https://www.kaggle.com/datasets/mariaherrerot/idrid-dataset) and <https://www.kaggle.com/datasets/google-brain/messidor2-dr-grades> (19 October 2022).

**Acknowledgments:** The authors are thankful to Princess Nourah bint Abdulrahman University Researchers Supporting Project number (PNURSP2023R410), Princess Nourah bint Abdulrahman University, Riyadh, Saudi Arabia.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Thiagarajan, A.S.; Adikesavan, J.; Balachandran, S.; Ramamoorthy, B.G. Diabetic Retinopathy Detection using Deep Learning Techniques. *J. Comput. Sci.* **2020**, *16*, 305–313. [\[CrossRef\]](#)
2. Skariah, S.M.; Arun, K.S. A Deep Learning Based Approach for Automated Diabetic Retinopathy Detection and Grading. In Proceedings of the 2021 4th Biennial International Conference on Nascent Technologies in Engineering (ICNTE), Navi Mumbai, India, 15–16 January 2021; pp. 1–6.
3. Chowdhury, P.; Islam, M.R.; Based, M.A.; Chowdhury, P. Transfer Learning Approach for Diabetic Retinopathy Detection using Efficient Network with 2 Phase Training. In Proceedings of the 2021 6th International Conference for Convergence in Technology (I2CT), Maharashtra, India, 2–4 April 2021; pp. 1–6.
4. Feng, L.; Wang, Y.; Xu, T.; Dong, L.; Yan, L.; Jiang, M.; Zhang, X.; Jiang, H.; Wu, Z.; Zou, H. Deep learning-based automated detection for diabetic retinopathy and diabetic macular oedema in retinal fundus photographs. *Eye* **2022**, *36*, 1433–1441.
5. Bhardwaj, C.; Jain, S.; Sood, M. Deep Learning–Based Diabetic Retinopathy Severity Grading System Employing Quadrant Ensemble Model. *J. Digit. Imaging* **2021**, *34*, 440–457. [\[CrossRef\]](#) [\[PubMed\]](#)
6. Snehil, K. Diabetic Retinopathy Diagnosis with Ensemble Deep-Learning. In Proceedings of the 3rd International Conference on Vision, Image and Signal Processing, Vancouver, BC, Canada, 26–28 August 2019; pp. 1–5.
7. Orlando, J.I.; Prokofyeva, E.; Fresno, M.D.; Blaschko, M.B. Learning to Detect Red Lesions in Fundus Photographs: An Ensemble Approach based on Deep Learning. *arXiv* **2017**, arXiv:abs/1706.03008.
8. Yan, Y.; Gong, J.; Liu, Y. A Novel Deep Learning Method for Red Lesions Detection Using Hybrid Feature. In Proceedings of the 2019 Chinese Control And Decision Conference (CCDC), Nanchang, China, 3–5 June 2019; pp. 2287–2292.
9. Safwan, M.; Chennamsetty, S.S.; Kori, A.; Alex, V.; Krishnamurthi, G. Classification of Breast Cancer and Grading of Diabetic Retinopathy & Macular Edema using Ensemble of Pre-Trained Convolutional Neural Networks. In *Lecture Notes in Computer Science Book Series*; Springer: Berlin/Heidelberg, Germany, 2018; Volume 10882.
10. Harangi, B.; Tóth, J.; Hajdu, A. Fusion of Deep Convolutional Neural Networks for Microaneurysm Detection in Color Fundus Images. In Proceedings of the 2018 40th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), Honolulu, HI, USA, 18–21 July 2018; pp. 3705–3708.
11. Ramchandre, S.; Patil, B.; Pharande, S.; Javali, K.; Pande, H. A Deep Learning Approach for Diabetic Retinopathy detection using Transfer Learning. In Proceedings of the 2020 IEEE International Conference for Innovation in Technology (INOCON), Bangalore, India, 6–8 November 2020; pp. 1–5.
12. Shorfuzzaman, M.; Hossain, M.S.; El Saddik, A. An Explainable Deep Learning Ensemble Model for Robust Diagnosis of Diabetic Retinopathy Grading. *ACM Trans. Multimed. Comput. Commun. Appl.* **2021**, *17*, 1–24. [\[CrossRef\]](#)
13. Rajkumar, R.S.; Jagathishkumar, T.; Ragul, D.; Selvarani, A.G. Transfer Learning Approach for Diabetic Retinopathy Detection using Residual Network. In Proceedings of the 2021 6th International Conference on Inventive Computation Technologies (ICICT), Coimbatore, India, 20–22 January 2021; pp. 1189–1193.
14. Al-Smadi, M.; Hammad, M.M.; Baker, Q.B.; Al-Zboon, S.A. A transfer learning with deep neural network approach for diabetic retinopathy classification. *Int. J. Electr. Comput. Eng.* **2021**, *11*, 3492–3501. [\[CrossRef\]](#)
15. Dela Pava, M.; R’ios, H.; Rodr’iguez, F.J.; Perdomo, O.J.; Gonz’alez, F.A. A deep learning model for classification of diabetic Retinopathy in eye fundus images based on retinal lesion detection. In Proceedings of the Symposium on Medical Information Processing and Analysis, Mexico City, Mexico, 17–19 November 2021.
16. Chetoui, M.; Akhloufi, M.A. Explainable end-to-end deep Learning for diabetic retinopathy detection across multiple datasets. *J. Med. Imaging* **2020**, *7*, 044503. [\[CrossRef\]](#)

17. Mohamed, E.; Elmohsen, M.N.; Basha, T.A. Improved Automatic Grading of Diabetic Retinopathy Using Deep Learning and Principal Component Analysis. In Proceedings of the 2021 43rd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC), Scotland, UK, 1–5 November 2021; pp. 3898–3901.
18. Khan, A.A.; Kulkarni, N.; Kumar, A.; Kamat, A. D-CNN and Image Processing Based Approach for Diabetic Retinopathy Classification. In *Advances in Intelligent Systems and Computing*; Springer: Berlin/Heidelberg, Germany, 2021.
19. Han, Y.; Tao, M.; Zheng, X. Ensembling Learning for Automated Detection of Diabetic Retinopathy. In *Proceedings of 2021 International Conference on Medical Imaging and Computer-Aided Diagnosis (MICAD 2021) Medical Imaging and Computer-Aided Diagnosis*; Springer: Singapore, 2022.
20. Afrin, R.; Shill, P.C. Automatic Lesions Detection and Classification of Diabetic Retinopathy Using Fuzzy Logic. In Proceedings of the 2019 International Conference on Robotics, Electrical and Input Image Processing Techniques (ICREST), Dhaka, Bangladesh, 10–12 January 2019; pp. 527–532.
21. Chaudhary, S.; Ramya, H.R. Detection of Diabetic Retinopathy using Machine Learning Algorithm. In Proceedings of the 2020 IEEE International Conference for Innovation in Technology (INOCON), Bangaluru, India, 6–8 November 2020; pp. 1–5.
22. Ali, A.; Qadri, S.; Mashwani, W.K.; Kumam, W.; Kumam, P.; Naeem, S.; Göktas, A.; Jamal, F.; Chesneau, C.; Anam, S.; et al. Machine Learning Based Automated Segmentation and Hybrid Feature Analysis for Diabetic Retinopathy Classification Using Fundus Image. *Entropy* **2020**, *22*, 567. [[CrossRef](#)]
23. Kumar, K.; Megha, P.; Meenakshy, K. Diabetic Retinopathy Detection & Classification Techniques: A Review. *Int. J. Sci. Technol. Res.* **2020**, *9*, 1621–1628.
24. Ammal, M.; Gladis, D. Perception of hard exudates using Fuzzy Optimization and Discrete Wavelet Transformation in Fundus Image. In Proceedings of the 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), Tirunelveli, India, 4–6 February 2021; pp. 1034–1039.
25. Tala, E.B.; Thabet, B. Diabetic Retinopathy Recognition System based on GLDM Features and Feed Forward Neural Network Classifier. *Al-Qadisiyah J. Pure Sci.* **2022**, *27*, comp1-16. [[CrossRef](#)]
26. Patil, S.; Kulkarni, P. Diabetic Retinopathy Detection: Methods and Challenges. In Proceedings of the 2019 IEEE Pune Section International Conference (PuneCon), Pune, India, 18–20 December 2019; Volume 2022, pp. 1–3.
27. Qiao, L.; Zhu, Y.; Zhou, H. Diabetic Retinopathy Detection Using Prognosis of Microaneurysm and Early Diagnosis System for Nonproliferative Diabetic Retinopathy Based on Deep Learning Algorithms. *IEEE Access* **2020**, *8*, 104292–104302. [[CrossRef](#)]
28. Araújo, T.; Aresta, G.; Mendonça, L.; Penas, S.; Maia, C.; Carneiro, A.; Mendonça, A.M.; Campilho, A. Data Augmentation for Improving Proliferative Diabetic Retinopathy Detection in Eye Fundus Images. *IEEE Access* **2020**, *8*, 182462–182474. [[CrossRef](#)]
29. Nazir, T.; Irtaza, A.; Javed, A.; Malik, H.; Hussain, D.; Naqvi, R.A. Retinal Image Analysis for Diabetes-Based Eye Disease Detection Using Deep Learning. *Appl. Sci.* **2020**, *10*, 6185. [[CrossRef](#)]
30. Chalakkal, R.J.; Hafiz, F.M.; Abdulla, W.; Swain, A. An Efficient Framework for Automated Screening of Clinically Significant Macular Edema. *Comput. Biol. Med.* **2021**, *130*, 104128. [[CrossRef](#)] [[PubMed](#)]
31. Sikder, N.; Masud, M.; Bairagi, A.; Arif, A.S.; Nahid, A.; Alhumyani, H. Severity Classification of Diabetic Retinopathy Using an Ensemble Learning Algorithm through Analyzing Retinal Images. *Symmetry* **2021**, *13*, 670. [[CrossRef](#)]
32. Abdelmaksoud, E.; Barakat, S.; Elmogy, M.M. A comprehensive diagnosis system for early signs and different diabetic retinopathy grades using fundus retinal images based on pathological changes detection. *Comput. Biol. Med.* **2020**, *126*, 104039. [[CrossRef](#)]
33. Singh, N.; Kaur, L.; Singh, K. Histogram equalization techniques for enhancement of low radiance retinal images for early detection of diabetic Retinopathy. *Eng. Sci. Technol. Int. J.* **2019**, *22*, 736–745. [[CrossRef](#)]
34. Li, X.; Hu, X.; Yu, L.; Zhu, L.; Fu, C.; Heng, P. CANet: Cross-Disease Attention Network for Joint Diabetic Retinopathy and Diabetic Macular Edema Grading. *IEEE Trans. Med. Imaging* **2020**, *39*, 1483–1493. [[CrossRef](#)]
35. Wang, J.; Bai, Y.; Xia, B. Simultaneous Diagnosis of Severity and Features of Diabetic Retinopathy in Fundus Photography Using Deep Learning. *IEEE J. Biomed. Health Inform.* **2020**, *24*, 3397–3407. [[CrossRef](#)]
36. Aurangzeb, K.; Aslam, S.; Alhusein, M.A.; Naqvi, R.A.; Arsalan, M.; Haider, S.I. Contrast Enhancement of Fundus Images by Employing Modified PSO for Improving the Performance of Deep Learning Models. *IEEE Access* **2021**, *9*, 47930–47945. [[CrossRef](#)]
37. Rajput, G.G.; Reshmi, B.; Rajesh, I. Automatic detection and grading of diabetic maculopathy using fundus images. *Procedia Comput. Sci.* **2020**, *167*, 57–66. [[CrossRef](#)]
38. Ahn, S.; Pham, Q.T.; Shin, J.; Song, S.J. Future Image Synthesis for Diabetic Retinopathy Based on the Lesion Occurrence Probability. *Electronics* **2021**, *10*, 726. [[CrossRef](#)]
39. Qummar, S.; Khan, F.G.; Shah, S.; Khan, A.; Shamshirband, S.; Rehman, Z.U.; Ahmed Khan, I.; Jadoon, W. A Deep Learning Ensemble Approach for Diabetic Retinopathy Detection. *IEEE Access* **2019**, *7*, 150530–150539. [[CrossRef](#)]
40. Nazir, T.; Nawaz, M.; Rashid, J.; Mahum, R.; Masood, M.F.; Mehmood, A.; Ali, F.; Kim, J.; Kwon, H.; Hussain, A. Detection of Diabetic Eye Disease from Retinal Images Using a Deep Learning based CenterNet Model. *Sensors* **2021**, *21*, 5283. [[CrossRef](#)] [[PubMed](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.