Artificial Intelligence Approach for Diabetic Retinopathy Severity Detection

Shalini R*, Sasikala S
Department of Computer Science, IDE, University of Madras, Chennai, India.
Email: rshalini1990@gmail.com, sasikalarams@gmail.com
*Corresponding author

Keywords: Diabetic retinopathy, Deep learning algorithm, Convolutional Neural Network, hyperparameter tuning, cross-validation, medical diagnosis

Received: September 28, 2022

Identifying the diabetic retinopathy (DR) severity in the retina images taken under a variety of imaging conditions is more challenging. Five classes are commonly classified on retinal images based on the severity of DR diseases No DR, mild NPDR (nonproliferative diabetic retinopathy), moderate NPDR, severe NPDR, and PDR (proliferative diabetic retinopathy). Artificial intelligence is an emerging area in the medical diagnosis industry, in specific, deep learning algorithms are used for classifying retina images for accurate diagnosis of disease. The proposed work acquired retina images from the publicly available Kaggle repository and loaded them into an improved grid search Convolutional Neural Network model to accurately diagnose retina images with five different classes. This novel model helps ophthalmologists to classify DR into five stages based on the severity from No DR to PDR. The Experimental study showed that the proposed model performed better than the existing Convolutional Neural Network model with an accuracy of 89%.

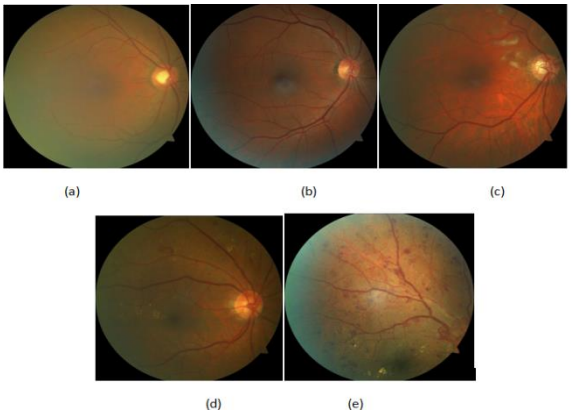
Povzetek: Opisana je metoda globokih nevronskih mrež za ugotavljanje težav pri vidu ljudi s sladkorno boleznijo.

1 Introduction

Diabetic retinopathy (DR) is a microvascular complication [1] caused by high blood sugar levels of type 1 or type 2 diabetes mellitus, which heavily damages retinal blood vessels causing vision loss for diabetes patients. The stages of DR are classified into mild, moderate, and severe Non-Proliferative Diabetic Retinopathy (NPDR) and Proliferative Diabetic Retinopathy (PDR). In the early stage of DR, a tiny red dot visible on the retina called Micro-aneurysms (MAs) is termed mild DR. Moderate DR is characterized by flame-shaped Hemorrhages and Cotton Wool Spots in the eye retina. Severe NPDR is characterized by Hard Exudate, which is a white or yellow deposit, and Intra-Retinal Microvascular abnormalities (IrMAs), which is a growth of a new vessel in retinal tissue. The final stage of DR is Proliferative diabetic retinopathy (PDR), characterized by neovascularization, Retinal Detachment, Vitreous Hemorrhage, and Sub-hyaloid hemorrhage in retinal tissue [2]. Figure 1, shows a visual representation of the different stages of DR in this Figure, it is observed that mild and moderate DR are complicated to classify manually, thus automated detection through deep learning helps ophthalmologists to detect the stages of DR accurately.

Diabetes is a condition caused by insufficient insulin secretion from the pancreas. This affects the blood circulatory system, thus the retina gets affected. The number of adults affected by DR and Vision threatening DR (VTDR) [3] is around 103.12 million, and 28.54 million respectively in 2020. This number is projected to be 160.50 million, and 44.82 million respectively in the year 2045 [4].

Figure 1: Different stages of diabetic retinopathy - (a) No DR (b) Mild NPDR (c) Moderate NPDR (d) Severe NPDR (e) PDR



Diabetes Type 1 is found to be the main cause of diabetic retinopathy. DR is highly asymptomatic in its early stages, during this stage, retinal damage and microvascular changes are observed. The prolonged high blood glucose levels cause blood vessel changes to swell or bleed which affects the eye vision and sometimes causes blindness. The capability of the human body for curing disease causes it to create new blood vessels to

handle the damage. This leads to the formation of new blood vessels in the eye retina, which causes blur in vision and even blindness. Glycemic control and antidiabetic medications should be strictly followed by patients to manage NPDR [5].

It is necessary for diabetic patients for checking vision regularly for very mild NPDR yearly, mild to moderate NPDR every 6 to 12 months, and severe to very severe NPDR must have a close follow-up every 2 to 4 months. Furthermore, prevailing treatments such as laser photocoagulation reduces the threat of causing blindness in DR patient by up to 98% when the patients are treated at an early stage. Thus, early detection of disease must get cured and prevent patients from blindness.

Initial diagnosis of DR is handled through evaluations of retinal function techniques such as electroretinography (ERG) and visual evoked potentials (VEP) are successful for earlier detection of DR. Similarly, a thorough examination of retinal blood flow and retinal blood vessel caliber performed on DR patients.

The other prevailing methods for diagnosis of DR disease through some of the imaging techniques are color fundus images, fluorescein angiography (FA), B-scan ultrasonography, and optical coherence tomography (OCT). From high-resolution images, ophthalmologists can analyze the retina lesion and interpret the severity of the disease.

However, detection of DR from retina images requires expertise to diagnose manually, the availability of experts in highly populated countries is less and the number of DR patients is high and also project to be growing in coming years, thus DR diagnosis requires automated diagnosis tools for fast and accurate detection.

Thus, these factors motivated several researchers on developing an Artificial Intelligence (AI) based system for quick and accurate diagnosis of DR. Recent advancements in the field of artificial intelligence for the healthcare industry are evolving, thus machine learning and deep learning concepts on disease diagnosis are vastly used.

Several artificial intelligence and image processing methods are available for DR detection from retina images. The existing image processing techniques namely Histogram of Oriented Gradients (HOG) [6], SIFT [7], Local binary patterns (LBP) [8], and Gabor filters [9] were used for diabetic retinopathy detection. And these techniques extracted features from the fundus images and performed classification. Similarly, machine learning and deep learning-based models are also implemented for DR detection from fundus images.

The detection of NPDR and PDR is crucial for starting early medications and controlling the disease. The existing techniques failed to detect the DR with higher accuracy. The proposed deep learning algorithm called CNN can extract the disease features and classifies the disease severity automatically from the images, thus preferred for this study.

The main contributions of this proposed work are

- To develop a deep learning model based on a Convolutional Neural Network (CNN), which can classify retina images based on the severity of DR into five classes.
- The proposed system adds novelty by iteratively tuning multiple parameters such as layers, batch size, dropout, epochs, and learning rate to build an effective learning system.
- The cross-validation through the GridSearch mechanism is performed for the best estimator selection.
- To conduct an experimental study and compare the performance of existing CNN models and the proposed improved grid search Convolutional Neural Network (IGSCNN) model for the classification of multi-class DR images.

The objective of this study is to develop a novel CNN algorithm, which can effectively classify retina images as No DR to PDR depending on the severity of the disease. This study focused on the detection and classification of DR images as five classes through the proposed IGSCNN model. The model iteratively tunes multiple hyperparameters in the CNN algorithm which effectively learns the retina images for multi-class classification. Experimental results show that the proposed deep learning model outperforms in DR detection and classification accuracy.

The following chapters discuss related research works in DR classification and detection in Chapter 2. Chapter 3 discusses data pre-processing, tuning the multiple parameters for CNN algorithms, and applying hyper-tuned CNN algorithms for classification. Chapter 4 explains the experimental results and performance of the proposed algorithm over the existing CNN algorithm.

2 Related works

There has been several research represented diabetic retinopathy detection from retina images. Some of the work handled image processing techniques for detection and some of them analyzed machine learning concepts. Some of these works are discussed further in this chapter. The work [10] is a survey of various techniques for diabetic retinopathy detection and classification. The author discussed various deep learning models for DR classification, CNN, UNet, Generative Adversarial Networks (GAN), transfer learning, and an ensemble model. The paper analyzed more pre-processing methods and discussed the importance of deep learning models for DR detection. DR detection using the CNN algorithm was proposed [11] on a small dataset; the authors modeled five-layer architecture of CNN. The work compared CNN, VGG16 pre-trained model, and inceptionV3 model. The dataset considered is a small retina image dataset with 4 classes. The custom CNN model has performed well on 4 classes of DR detection on a small retina dataset.

Another work on DR detection with the CNN algorithm was proposed [12] for five classes of DR classification. For this study, the author used 80,000

retina images from Kaggle and performed 120 epochs of training on the CNN learning model. The result discussed that DR classification with five stages of the disease on the CNN model has achieved 75% accuracy. However, this work has analyzed the DR classification with five stages, the accuracy arrived is less, and this motivated the proposed study to build a CNN model with improved accuracy.

Binary classification of DR is one of the researcher's interests, through deep learning CNN pre-trained model; another work [13] addressed binary classification. This work used the ResNet architecture of the CNN algorithm, which gave the highest accuracy of binary classification.

An ensemble approach to deep learning pre-trained models was addressed in [14] with three algorithms. The work used ResNet50, InceptionV3, Xception, and two dense layers to build an ensemble model. The Multi-class classification was performed with ensemble models and the work achieved 80.8% accuracy for five stages of DR detection through pre-trained ensemble models.

Image processing through feature extraction and classification for three classes namely microaneurysms (MAs), hemorrhages (HMs), and exudates (EX) was addressed in [15]. Candidate lesion detection was performed with the Gabor filter bank technique for feature extraction. M-mediods with Gaussian Mixture Model (GMM) is the classification technique used for classifying it as three classes MAs, HMs, and EX. From this, the results are graded as normal, mild, moderate, and severe as four classes. The grading as four stages of DR is performed based on the count of lesions extracted from the input image. Ternary classification based on the deep learning model, CNN algorithm was proposed in [16], this work classified DR as mild, moderate, and severe based on lesion detection and template matching for MAs detection.

Lesion detection is performed with image processing techniques, Gaussian filtering, and applied sparse PCA (Principal component analysis) for feature reduction. Template matching was proposed with the Gaussian Correlation Coefficients Multi-scale method (MSCF) for detecting MAs. However, this technique can detect only three types of DR stages.

DR detection through optical coherence tomography angiography (OCTA) images was addressed in [17] with a Support Vector Machine (SVM) classifier. This study has addressed binary classification as healthy and affected DR based on multifractal geometry. However, the accuracy was high with SVM and multifractal geometry; this can perform only binary classification. The early stages of DR detection may not be detected accurately with stages. Detection of DR through transfer learning techniques adopted Siamese-like architecture for detection of binary as well as multi-class detection of DR was proposed [18].

Deep learning model InceptionV3 is used for building Siamese-like architecture, whereas InceptionV3 is used for weight sharing and integrated into a fully connected layer. The binary classification is considered on the same network with referable diabetic retinopathy (RDR) as with RDR and without RDR. The proposed work has achieved a specificity of 70.7%.

DR detection based on Neural Networks (NN) for binary classification was addressed in [19]. The feature selection mechanism through Particle Swarm Optimization (PSO) is used to select the important features from the dataset. The attributes include MA detection features, Exudates features, Euclidean distance features, and more. The classes considered are healthy and affected based on these features. Though the binary classification was performed in this work with a feature selection technique, the accuracy is around 76%. The tweets of disaster were efficiently classified with Enhanced Word Embedding Ensemble Convolutional Neural Network Model, which classifies the tweets of various categories [20].

Zhang et al. [21] proposed background segmentation, feature set extraction, feature optimization using Cuckoo search and Convolutional Neural Network (CNN) for DR severity grade classification. For optimization of hyper-parameters of CNN, the Grey wolf optimization algorithm is used [22]. This literature study infers that diabetic retinopathy detection is one of the important research topics among researchers. There are works addressing machine learning algorithms and deep learning algorithms and a combination of image processing with classification. The work studied so far has concentrated on binary classification as healthy or affected with DR. CAD based model used for early DR detection [23].

Some of the work was also addressed with ternary or four-class classification. Though there are few works carried out on the five stages of DR classification, the accuracy of DR detection is less. This problem of DR stage detection with high accuracy is carried out in this proposed work.

2.1 Inferences from the review study

The review study covered various aspects of DR detection and DR severity classification using existing machine learning and deep learning models. The inferences derived from a deep review of various literatures on DR detection, specifically DR severity classification are represented in Table 1. The existing works have performed the DR disease classification with machine learning and deep learning methods. Although these existing methods for detecting diabetic retinopathy have paved the way for more accurate diagnosis and treatment, further improvements are still necessary regarding performance. The prediction of DR with higher accuracy is the major shortcoming of existing models. Hence, this motivated us to propose a model, which can classify multiple stages of DR severity with enhanced accuracy.

3 Research methods

The motive of the proposed work is to perform deep learning classification and detection of diabetic retinopathy using images of the retina. The most existing machine learning algorithms for DR classifies some earlier features and severity stages of DR disease with high computational and error rate. The proposed work builds a hyper-parameter tuned network to innovate an optimized convolutional neural network model with higher accuracy and reduced loss value for DR classification. The work has fixated on multi-class classification on Kaggle retinal dataset [24], to classify 5 different stages normal, mild DR, moderate DR, severe DR, and PDR.

Table 1: Inferences of review study

S.No	Reference	Methods	Advantage	Disadvantage
1.	Tsiknakis et. al.,[10]	This review study on DR detection found models like CNN-UNet, Generative Adversarial Networks, transfer learning, and ensemble.	Several DL systems evolve and become integrated into the clinical practice; they will enable clinicians to treat patients in need more effectively and efficiently.	Still a lack of performance, and interpretability, and trustworthiness from ophthalmologists.
2.	X.Zeng et. al., [18]	A binocular model is also trained for the original five-class DR classification task.	It achieves a kappa score of 0.829 which is higher than existing work on the 10% validation set.	The small size of our data set achieves predictably low performance which can be further improved if more data are collected in the future.
3.	S.Qummar et. al., [14]	An ensemble approach has been addressed for detecting five stages of DR. The work used ResNet50, InceptionV3, Xception, and two dense layers.	In this paper, the classification of the five stages of DR is done with 80.8% accuracy.	There is a need to train specific models for specific stages to improve the accuracy of earlier DR prediction.
4	Samanta et. al., [11]	This five-layer CNN model performs the detections of four DR classes on a small retina dataset. The model compared with other models like CNN, VGG16, and inception V3.	Provides F1 scores of 0.64 and 0.74 on Mild DR and Moderate DR prediction, is found to be low in past works.	Validation done on a small dataset of 419 fundus retinal images to achieve 84.10% accuracy.
5	Pratt Harry et. al.,[12]	The authors have classified five classes of DR using custom CNN.	Achieved 75% accuracy in the validation dataset.	The low sensitivity, mainly from the mild and moderate classes, suggests that the network struggled to learn deep enough features to detect some of the more intricate aspects of DR.

3.1 Dataset

This study used high-resolution retina images of the left and right eye of each patient. These images of the left and right eye are labeled with five classes on a scale of 0 to 4 depending on the severity, whereas 0 represents no DR or normal and 4 represents PDR. There are 9 general DR disease features Microaneurysm, Hemorrhage, Cotton Wool Spot, Hard Exudate, Intra-Retinal

Microvascular Abnormalities, Neovascularisation, Retinal Detachment, Vitreous Hemorrhage, and Sub-hyaloid Hemorrhage. These features are grouped into

five classes based on the DR severity levels as shown in Table 2.

Table 2: DR disease severity scale

Disease Severity Level	Findings upon DR Features
0-No DR	No Diabetic Retinopathy.
1-Mild NPDR	Microaneurysm.
2-Moderate NPDR	Hemorrhage, Cotton Wool Spot.
3-Severe NPDR	Hard Exudate, Intra-Retinal Microvascular Abnormalities.
4 - PDR	Neovascularisation, Retinal Detachment, Vitreous Hemorrhage, Sub-hyaloid Hemorrhage.

3.2 Data preprocessing

The initial preprocessing is resizing the given images to 200X200 dimensions. The image dataset taken here is a color image in jpeg format, which contains three channels (R-Red, B-Blue, G-Green) in each pixel and these color values are stored as numbers in the matrix. These color map values are divided by 255 to convert into 0-1 systems. Image preprocessing helps in improving the quality of retina images and adapting the images to a consistent format to give input for the algorithm. The image taken here is RGB values, thus it is scaled with mix-max normalization by dividing it by 255 and converting it to 0 -1 values, which increases the speed of learning.

3.2.1 Image data augmentation

The next stage of image pre-processing is data augmentation, this helps in generating the synthetic images from the given original images with looped over mini-batches with specified batch sizes. Image data augmentation improves the robustness of learning to attain higher accuracy in classification work. There are many types of techniques available to generate new images including horizontal or vertical flip, rotation, inward or outward scaling, and more.

In this proposed work, the retina images are given for augmentation to increase the number of images with three standard augmentation techniques namely rotation, shift, and zoom. The rotation technique can rotate the images in degrees of 0 to 90 and the extra pixels generated during the augmentation of images are padded. The given retina images are rotated slightly in the degrees of 15 and used for learning. And, the shift technique can be of two types, one is the right shift and the other is the left shift, the shifts of image pixels are done in right or left with the given specific percentage. In this proposed work, the right shift and left shift of 0.1 is used to create new images.

The shift technique does not affect the overall height or width of the image, whereas the generated image takes the adjacent pixel. Zoom augmentation takes input as a float value, the value which is higher than 1 represents zoom out, and a value less than 1 represents zoom in process. The retina images are augmented with a zoom-in process with 0.1 as a parameter for zoom. Figure 2 shows the retina image augmentation. The top left image is a width shift with a range of 0.1, the top right image is a height shift with a range of 0.1, the bottom left image is a rotation with 15 degrees and the bottom right is a zoom-in with a range of 0.1.

The generated retina images are given to the CNN algorithm in a batch size of 8 and accelerated the training and generated image batches are discarded after training. The problem of overfitting can be avoided with the data augmentation process. Keras ImageDataGenerators is the tool used to transform more images from the given original retinal images with shift, rotation, and zoom as the transformation parameters.

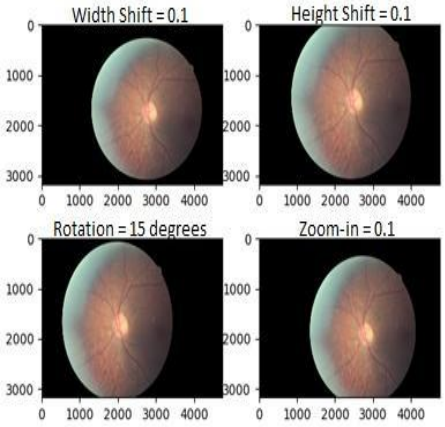


Figure 2: Retina image augmentation

3.2.2 Hyperparameters tuning

Hyperparameter optimization (HPO) can control the performance of the learning model, this is used to search for the best parameters to get the best performance on the taken retina image dataset. The search space is defined with a vector, which presents each point with a specific value used in the optimization process. In the proposed work, Grid search HPO is used for tuning the parameters.

This technique is a method of brute force and takes a finite set of values from the user and evaluates the Cartesian product of these sets. The domain of hyperparameters is given as real-value, integer value, and binary value. The dropout value and learning rate are given as real-valued parameters, whereas the number of layers, batch size, and epochs are integer value parameters. The CNN algorithm can be initialized with N-hyperparameters (i.e layers, batch size, dropout, etc.) and the optimized values of hyperparameters (i.e layers=

128, batch size = 8, dropout = 0.2, etc) can be initialized for presenting the novel CNN learning model.

The proposed work used Grid Search Cross Validation (GSCV) for hyperparameter tuning. GSCV takes the hyperparameter vector as input and cross-validates to get the best model, while cross-validation, the mean and Standard deviation error values are computed, and the hyperparameter, which gets less error is taken for the final learning model. The GSCV trains the CNN model for all combinations of hyperparameter space, this process is guided by score computation for selecting the best estimator.

Thus, the GSCV model has been used for tuning multiple parameters to get the effective learning CNN model. The retina image dataset is split into train and test data. Training retina image data is cross-validated to get the best parameter. Training data takes the various parameters and builds the trained model, in the trained model, test set validation is performed to compute mean and standard deviation (SD) values. This process is repeated till all the parameter's mean and SD values are computed. Figure 3 shows the representation architecture followed for hyperparameter tuning.

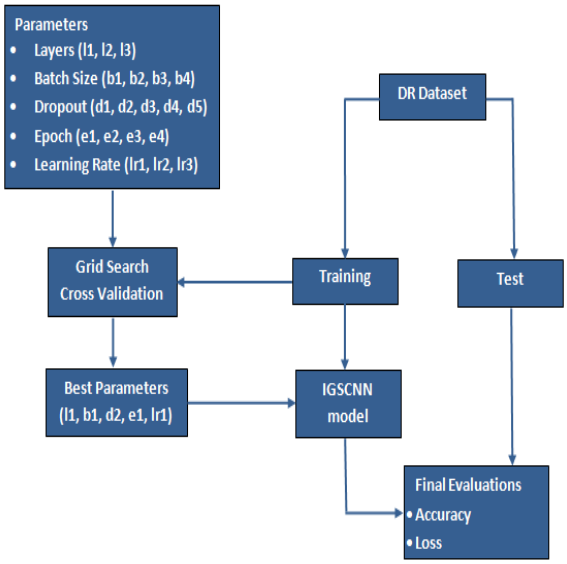


Figure: 3 Cross validation with hyperparameters

The cross-validation technique prediction error of a learning model computes mean and standard deviation error and selects the parameter based on the low error values. The mean is computed as follows in equation (1), where 'b' represents several folds, and 'm' represents the mean value. The standard deviation error is computed based on equation (2) where 'se' represents the standard error.

$$m = \frac{\sum_{i=1}^{b} err_i}{b} \quad \text{------ (1)}$$

$$se = \sqrt{\frac{var(err)}{b}} \quad \text{------- (2)}$$

Minimizing the prediction error is the objective of cross-validation, this technique is used to improve the accuracy of learning models and gives tuning parameters for generic estimators. The cross-validation techniques compute this mean score and SD score for each vector space of hyperparameter, in this proposed work five parameters used are learning rate, batch size, epochs, layers, and drop out. The test means to score and test SD scores are used for choosing the best estimator. The metrics of each hyperparameter combination are evaluated through the test set to get the mean and SD score values.

The new Convolutional Neural Network (CNN) model is defined with all hyperparameter values found by the cross-validation methods. In this proposed work, a total of five hyperparameters are defined and cross-evaluated. Hyperparameters and their vector space for search are given in Table 3. The hyperparameter used for developing an effective learning model for the detection and classification of diabetic retinopathy includes layers, dropout, batch size, epoch, and learning rate. The vector space defined for the layer is 128, 256, and 512 for search through cross-validation to get the best estimator value. The batch size defined for vector space for search is a total of five values, they are 0.1 to 0.5 with an interval of 0.1. The batch size used for the study is 4,8,16,32. The number of epochs defined for hyperparameter search space is 5,10,15 and 20. Similarly, the learning rate used for hyperparameter search space is 0.01, 0.001, and 0.0001. To limit the computation complexity, each hyperparameter search space is limited from 3 to 5 values.

Table 3: Hyperparameter tuning

Hyperparameter	**Hyperparameter Search Space**
Layers	[128, 256, 512]
Dropout	[0.1, 0.2, 0.3, 0.4, 0.5]
Batch Size	[4, 8, 16, 32]
Epochs	[5, 10, 15, 20]
Learning rate	[0.01, 0.001, 0.0001]

3.3 Improved grid search convolutional neural network

Among the deep learning algorithms namely Convolutional Neural Networks (CNN), Deep Neural Networks (DNN) and Recurrent Neural Networks (RNN), CNN is the most suiTable to learn the images. CNN is influenced by human vision and its arithmetic operation named “convolution". CNN uses filters or kernels to calculate convolutions from the input images to generate the features. In the CNN architecture, the convolution part is referred to as feature extraction and the final layer is referred to as the classification of

output. Thus, the proposed DR classification is preferred with the CNN algorithm over any other deep learning models.

Convolutional Neural Network (CNN), a deep learning algorithm used for diabetic retinopathy detection and classification takes an input of retinal image datasets with labels for learning. Necessary pre-process and hyperparameter optimization are used to build this CNN model. The proposed work is implemented with normal CNN model and hyperparameter optimized model. The CNN architecture used is the same for both works. The proposed work has compared both CNN generic algorithm and CNN hyperparameter optimized algorithm with the existing CNN model of Harry Pratt et. al [12] for DR classification. The experimental results proved that the CNN with hyperparameter optimization has achieved the highest classification accuracy. The CNN architecture used for diabetic retinopathy detection is shown in Figure 4.

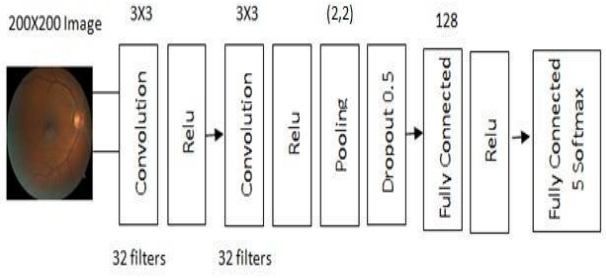


Figure 4: Proposed CNN architecture for DR classification

As the DR detection is performed with generic CNN method and hyperparameter optimized method. The difference between these two models is the input of layer, batch size, epochs, dropout rate, and learning rate. These values are declared with generic CNN, whereas the improved model takes the best parameters computed through Grid search cross-validation analysis.

The input for the generic CNN algorithm is retinal images of dimension 200X200 with training and test data along with labels. The labels are five classes, which represent stages of DR No DR, mild NPDR, moderate NPDR, severe NPDR, and PDR. The first layer is a 3X3 Convolutional layer with 32 filters, this is added with activation Rectified Linear Unit 'relu'. The second layer is a 3X3 Convolutional layer with 32 filters, this layer is added with relu activation and pooling. The pooling layer used pool size (2, 2), and dropout added 0.5. The next layer is a fully connected dense layer with filter size 128 and this layer is added with a relu activation unit. Softmax is used in the last layer for multiclass prediction. Thus here we used the last layer the dense layer with softmax output of class value 5. This generic CNN model with the above architecture is trained and evaluated with a batch size of 32 for the input image, number of epochs 5.

The hyperparameter optimization is performed with Grid Search cross-validation as discussed in the previous section. The parameter values are given through cross-validation and these best estimators are given to build the improved CNN model. The input for the improved CNN algorithm is the same retinal images of dimension 200X200 with training and test data along with labels. The first layer is a 3X3 Convolutional layer with 32 filters, this is added with activation Rectified Linear Unit 'relu'. The second layer is a 3X3 Convolutional layer with 32 filters, this layer is added with relu activation and pooling. The pooling layer uses pool size (2, 2) and dropout is estimated through GSCV techniques, the best estimator given value. The next layer is a fully connected dense layer, for which size is computed by the GSCV technique, and this layer is added with a relu activation unit. The last layer is the dense layer with a softmax output of class value 5 and the learning rate is defined by the GSCV technique. The model is built with epochs computed by GSCV.

Algorithm 1: Proposed algorithm

Input: Retinal image Data (X, Y), where X is the image features and Y is the multi-class (No DR, Mild NPDR, Moderate NPDR, Severe NPDR, PDR); X_train, X_test ϵ X , where X_train is training images and X_test is test images.

Output: Classification of severity of DR from the test set x ϵ X

Step 1: Take input images (X) and resize an image to dimension 200 X 200 with Bi-cubic interpolation area,

$$R(A,B) = \sum_{i=0}^{3}\sum_{j=0}^{3} a_{ij} x^i y^j$$

Step 2: Perform image augmentation on resized images,

- Rotation =15 degrees
- height and width shift = 0.1
- zoom_range=0.1

Step 3: Reshape X_train and X_test using Min-Max normalization,

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}}$$

Step 4: Initialize hyperparameter search space for tuning parameters as layers L = {$l1$, $l2$, $l3$}, batch size B= {$b1$,$b2$,$b3$,$b4$}, dropout D = {$d1$, $d2$, $d3$,$d4$, $d5$}, epochs E={$e1$,$e2$,$e3$,$e4$}, learning rate LR = {$lr1$,$lr2$,$lr3$}.

Step 5: Perform Cross validation through Grid search for the hyperparameter (L, D, B, E, LR) search space defined in step 4

- Compute mean = $\sum_{i=1}^{b} err_{i\cdot}/b$
- Compute std = $\sqrt{var(err)/b}$

Step 6: Repeat Step 5 for all Searchspace

Step 7: Sort mean value, assign high mean score as the best estimator (Bt) to get (L_{Bt}, D_{Bt}, B_{Bt}, $E_{Bt,}$ LR_{Bt})

Step 8: Build CNN model with best_estimator (L_{Bt}, D_{Bt}, B_{Bt}, $E_{Bt,}$ Lr_{Bt}) as parameters

- Let C denote, CNN algorithm with N hyperparameters

- The domain of nth hyperparameter given as Λn= {L, D, B, E, L}, where Λ denotes N-ary operation.
- A vector of hyperparameter is initialized which is acquired from step 7, it is denoted as $\acute{\Lambda} \epsilon \Lambda$.
- CNN model instantiated with hyperparameters is denoted asC_λ.

$$C_\lambda = \Lambda_{Bt}$$

4 Results and analysis

The proposed Diabetic retinopathy classification from five different classes is implemented with two different deep learning models separately they are CNN and hyperparameter optimized CNN and compared with the existing Harry Pratt [12] model. The model is evaluated with a 20% test dataset.

The hyperparameter optimization is implemented with the GridSearch cross-validation technique. The hyperparameter search space for five different learning parameters is considered for this experiment. The GSCV technique computes the mean score and standard deviation score value for every search space and sorts their values to evaluate and select the best estimator.

Table 4: Hyperparameter optimization

Hyper-parameters	Values	Validation	
		Mean_ Score	Std_Score
Layers	**128**	**0.83**	0.01
	256	0.81	0.05
	512	0.80	0.06
Batch Size	**4**	**0.83**	0.01
	8	0.81	0.05
	12	0.80	0.06
	16	0.78	0.08
Dropout	0.1	0.83	0.01
	0.2	**0.83**	0.01
	0.3	0.81	0.02
	0.4	0.80	0.03
	0.5	0.78	0.04
Epochs	**10**	**0.83**	0.01
	15	0.80	0.04
	20	0.78	0.06
Learning Rate	**0.01**	**0.83**	0.01
	0.001	0.81	0.02
	0.0001	0.78	0.04

The Table 4 shows the hyperparameter Optimization for Layer, the parameter search space and the mean score computed are shown. The proposed deep learning model takes the highest mean value as the best estimator, thus the layer value selected by GSCV is 128. Batch size defines the number of input images taken for learning at once. A larger batch size leads to poor generalization, whereas a smaller batch size leads to variance in accuracy, thus optimized value is preferred.

Table 4 shows the metric computed for hyperparameter optimization for Batch size, the parameter search space, and the mean score computed is given. The model takes the highest mean value as the best estimator, thus the batch size selected by GSCV is 4. Hyperparameter dropout optimization helps in avoiding overfitting. Table 4 shows the metric computed for hyperparameter optimization for dropout selection, the parameter search space used was given below along with the mean score computed. The proposed deep learning model takes the highest mean value as the best estimator, thus the dropout selected by GSCV is 0.2.

The number of times the dataset passed to the deep learning model is termed an epoch. The lower epoch leads to an underfitting problem, whereas the higher number of epochs may lead to an overfitting problem. Thus an optimized epoch is required to create the best deep-learning model. Table 4 shows the hyperparameter optimization for epochs selection, the parameter search space used is given in the Table, and the mean score computed is also shown below.

The model takes the highest mean value as the best estimator, thus the epochs selected by GSCV is 10. The learning rate is one of the hyperparameters optimized here, which controls the step size of the learning model and also reduces loss. The Loss function is an evaluation measure for spotting errors between predicted labels and targeted labels. Categorical Cross Entropy is the most commonly used loss function for CNN-based applications. The higher the learning rate, the less loss will not converge easily, the lower the learning rate, the model will take too long to converge.

Table 4 shows the optimization of hyperparameter search space for learning rate selection, the mean score and error computed are listed in the Table. The model takes the highest mean value as the best estimator, thus the learning rate selected by GSCV is 0.01. The random samples of retina images are taken and validated through validation_generator for the classification of DR among five classes. The results represent the true label and predicted label for DR classification. The class value 0, 1, 2, 3, and 4 represents No DR, mild NPDR, moderate NPDR, severe NPDR, and PDR respectively. Moreover, the results show that the early stage of DR is classified prominently through the proposed CNN models as shown in Figure 5.

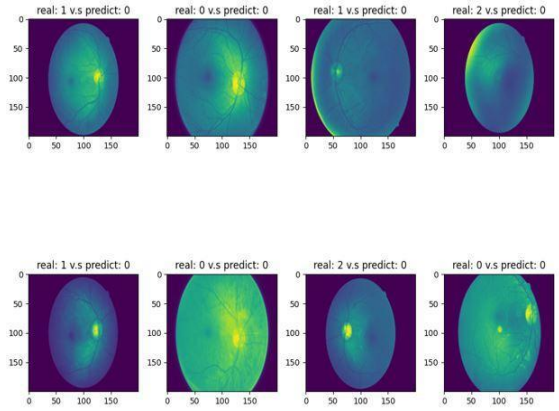
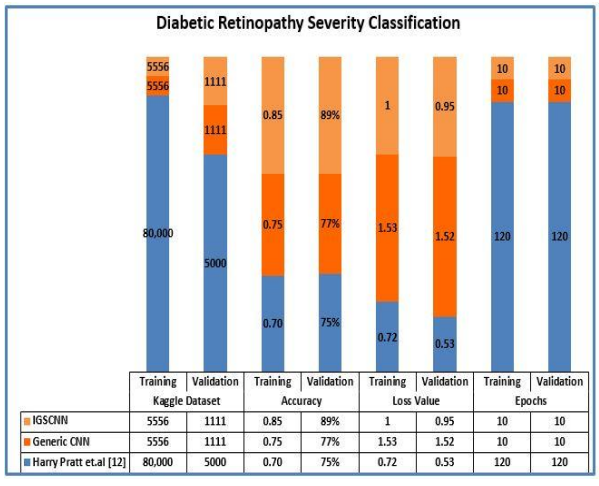


Figure 5: Diabetic retinopathy prediction for random sample from test data

Figure 6 shows the comparative analysis of the accuracy and loss of the existing and proposed CNN learning models for DR severity classification. From the results, it is observed that the proposed IGSCNN model has classified DR images with a higher accuracy of 89%. The error rates of the existing generic CNN and proposed Harry Pratt et.al [12] models are found to be 1.52 and 0.53.

The proposed IGSCNN model gives an error rate of 0.95 which is better than the generic CNN model and close to the Harry Pratt et.al [12] model. The Harry Pratt et.al model has utilized 5000 validation images to achieve 75% accuracy with 120 epochs generating very lesser error rate of 0.53. Hence, the proposed IGSCNN model can be modulated with larger validation dataset with increase in epochs, for achieving lesser error rates.



	Training	Validation	Training	Validation	Training	Validation	Training	Validation
	Kaggle Dataset		Accuracy		Loss Value		Epochs	
IGSCNN	5556	1111	0.85	89%	1	0.95	10	10
Generic CNN	5556	1111	0.75	77%	1.53	1.52	10	10
Harry Pratt et.al [12]	80,000	5000	0.70	75%	0.72	0.53	120	120

Figure 6: Performance of existing and proposed CNN models for DR severity classification

5 Conclusions

Diabetic retinopathy (DR) classification from retina images as five stages is challenging research. Though there is much work representing the DR classification as binary classes and even multi-classes they failed to detect the early stages accurately from the retinal images. The proposed Improved Grid Search Convolutional Neural Network (IGSCNN) based learning model achieved this goal. CNN algorithm is chosen as it extracts features automatically in the convolutional layers. The retinal image dataset is taken from Kaggle for this study, which has five stages of DR classified as No DR, mild NPDR (non-proliferative diabetic retinopathy), moderate NPDR, severe NPDR, and PDR (proliferative diabetic retinopathy). The proposed work analyzed CNN normal model as well as CNN tuned model. The experimental results are conducted on the same dataset for two different models and compared with the existing CNN model. The hyperparameter optimization is done through Grid Search Cross Validation (GSCV) for five different parameters. The results show that the normal CNN model has achieved an accuracy of 77%, whereas the proposed IGSCNN has achieved the highest accuracy of 89% which is better than the existing CNN model with 75% accuracy. From the results, it is also analyzed that the proposed IGSCNN model has detected the early stages of diabetic retinopathy more accurately. In the future, the work can be extended to detecting diabetic retinopathy severity with lesser error rates.

Authors' contributions

All the three authors read and approved the final manuscript.

Declarations

"All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards."

References

[1] https://elsevier.health/en-US/preview/diabetic-retinopathy-co

[2] Wu L, Fernandez-Loaiza P, Sauma J, Hernandez-Bogantes E, Masis M. Classification of diabetic retinopathy and diabetic macular edema. World J Diabetes. 2013;4(6):290-294. doi:10.4239/wjd.v4.i6.290

[3] Zheng, Yingfeng et al. "The worldwide epidemic of diabetic retinopathy." Indian journal of ophthalmology vol. 60,5 (2012): 428-31. doi:10.4103/0301-4738.100542

[4] Teo ZL, Tham YC, Yu M, Chee ML, Rim TH, Cheung N, Bikbov MM, Wang YX, Tang Y, Lu Y, Wong IY, Ting DSW, Tan GSW, Jonas JB, Sabanayagam C, Wong TY, Cheng CY. Global Prevalence of Diabetic Retinopathy and Projection of Burden through 2045: Systematic Review and Meta-analysis. Ophthalmology. 2021 Nov;128(11): 1580-1591.doi: 10.1016/j.ophtha.2021.04.027. Epub 2021 May 1. PMID: 33940045.

[5] Shukla UV, Tripathy K. Diabetic Retinopathy. [Updated 2021 Aug 29]. In: StatPearls [Internet]. Treasure Island (FL): StatPearls Publishing; 2022 Jan-. Available from: https://www.ncbi.nlm.nih.gov/books/NBK560805/

[6] Sarwinda, Devvi & Siswantining, Titin & Bustamam, Alhadi. (2018). Classification of Diabetic Retinopathy Stages using Histogram of Oriented Gradients and Shallow Learning. 83-87. 10.1109/IC3INA.2018.8629502.

[7] Q. Li and X. Wang, "Image Classification Based on SIFT and SVM," 2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS), 2018, pp. 762-765, doi: 10.1109/ICIS.2018.8466432.

[8] Colomer, Adrián et al. “Detection of Early Signs of Diabetic Retinopathy Based on Textural and Morphological Information in Fundus Images.” Sensors (Basel, Switzerland) vol. 20,4 1005. 13 Feb. 2020, doi:10.3390/s20041005

[9] Morales, Y & Nuñez, R & Suarez, J & Torres Moreno, Cesar. (2017). Digital tool for detecting diabetic retinopathy in retinography image using gabor transform. Journal of Physics: Conference Series. 792. 012083. 10.1088/1742-6596/792/1/012083.

[10] Tsiknakis, Nikos & Theodoropoulos, Dimitris & Manikis, Georgios & Ktistakis, Emmanouil & Boutsora, Ourania & Berto, Alexa & Scarpa, Fabio & Scarpa, Alberto & Fotiadis, Dimitrios & Marias, Kostas. (2021). Deep Learning for Diabetic Retinopathy Detection and Classification Based on Fundus Images: A Review. Computers in Biology and Medicine. 135. 104599. 10.1016/j.compbiomed.2021.104599.

[11] Samanta, Abhishek & Saha, Aheli & Satapathy, Suresh & Fernandes, Steven & Zhang, Yu-Dong. (2020). Automated Detection of Diabetic Retinopathy using Convolutional Neural Networks on a Small Dataset. Pattern Recognition Letters. 135. 10.1016/j.patrec.2020.04.026.

[12] Pratt, Harry & Coenen, Frans & Broadbent, Deborah & Harding, Simon & Zheng, Yalin. (2016). Convolutional Neural Networks for Diabetic Retinopathy. Procedia Computer Science. 90. 200-205. 10.1016/j.procs.2016.07.014.

[13] Mehdi Torabian Esfahani, Mahsa Ghaderi, Raheleh Kafiyeh (2018), Classification of diabetic and normal fundus images using new deep learning method, Leonardo Electronic Journal of Practices and Technologies, ISSN 1583-1078

[14] S. Qummar et al., "A Deep Learning Ensemble Approach for Diabetic Retinopathy Detection," in IEEE Access, vol. 7, pp. 150530-150539, 2019, doi: 10.1109/ACCESS.2019.2947484.

[15] Akram, M. & Khalid, Shehzad & Tariq, Anam & Khan, Shoab & Azam, Farooque. (2014). Detection and classification of retinal lesions for grading of diabetic retinopathy. Computers in biology and medicine. 45C. 161-171. 10.1016/j.compbiomed.2013.11.014.

[16] L. Qiao, Y. Zhu and H. Zhou, "Diabetic Retinopathy Detection Using Prognosis of Microaneurysm and Early Diagnosis System for Non-Proliferative Diabetic Retinopathy Based on Deep Learning Algorithms," in IEEE Access, vol. 8, pp. 104292-104302, 2020, doi: 10.1109/ACCESS.2020.2993937.

[17] M. M. Abdelsalam and M. A. Zahran, "A Novel Approach of Diabetic Retinopathy Early Detection Based on Multifractal Geometry Analysis for OCTA Macular Images Using Support Vector Machine," in IEEE Access, vol. 9, pp. 22844-22858, 2021, doi: 10.1109/ACCESS.2021.3054743.

[18] X. Zeng, H. Chen, Y. Luo and W. Ye, "Automated Diabetic Retinopathy Detection Based on Binocular Siamese-Like Convolutional Neural Network," in IEEE Access, vol. 7, pp. 30744-30753, 2019, doi: 10.1109/ACCESS.2019.2903171.

[19] A. Herliana, T. Arifin, S. Susanti and A. B. Hikmah, "Feature Selection of Diabetic Retinopathy Disease Using Particle Swarm Optimization and Neural Network," 2018 6th International Conference on Cyber and IT Service Management (CITSM), 2018, pp. 1-4, doi: 10.1109/CITSM.2018.8674295.

[20] Elangovan, Aarthi, and S. Sasikala. "A Multi-label Classification of Disaster-Related Tweets with Enhanced Word Embedding Ensemble Convolutional Neural Network Model." Informatica 46.7 (2022).

[21] Zhang, Qiu-ming, Jing Luo, and Korhan Cengiz. "An Optimized Deep Learning based Technique for Grading and Extraction of Diabetic Retinopathy Severities." Informatica 45.5 (2021).

[22] Guernine, Akram, and Mohamed Tahar Kimour. "Optimized Training for Convolutional Neural Network Using Enhanced Grey Wolf Optimization Algorithm." Informatica 45.5 (2021).

[23] Jena, Manaswini, et al. "Exploring the Parametric Impact on a Deep Learning Model and proposal of a 2-Branch CNN for Diabetic Retinopathy Classification with Case Study in IoT-Blockchain based Smart Healthcare System." Informatica 46.2 (2022).

[24] https://www.kaggle.com/competitions/diabetic-retinopathy-detection/data