ORIGINAL ARTICLE

Transfer learning based robust automatic detection system for diabetic retinopathy grading

Charu Bhardwaj[1] · Shruti Jain[1] · Meenakshi Sood[2]

Abstract

Diabetic retinopathy (DR) can be categorized on the basis of prolonged complication in the retinal blood vessels which may lead to severe blindness. Early stage prediction and diagnosis of DR requires regular eye examination to reduce the complications causing vision loss. Indicative significance of DR forecast and evaluation to help the ophthalmologists in standard screening has prompted the improvement of computerized DR recognition frameworks. This work focuses on automatic DR disease identification and its grading by the means of transfer learning approach using dynamic investigation. Our proposed approach utilizes deep neural network for feature extraction from fundus images and these features are further ensembled with supervised machine learning technique for DR grading. An optimized classification is achieved by applying an ensemble of convolution neural networks (CNNs) with statistical feature selection module and SVM classifier. The learning of classifier is achieved by the feature information transferred from CNN model to the SVM classifier, which results in remarkable performance of the learned models. Statistically optimized feature set utilized for transfer learning technique yields in the classification accuracy of 90.51% with proposed Prominent Feature-based Transfer Learning (PFTL) method employing Inception V3 model. The cost analysis of the proposed model provides a minimum cross-entropy loss of 0.295 consuming the time of 38 min 53 s, thus, maintaining a trade-off. The generalization ability of the proposed model is established by the performance assessment using latest IDRiD dataset that yields accuracy of 90.01% for Inception V3 network providing uniform outcomes for all the evaluation parameters. The diagnosis ability of the proposed transfer learning-based model is justified by comparing the proposed methods with the state-of-the-art methods. The optimized PFTL model (CNN + Statistical Analysis + SVM) outperforms other classification algorithms and provides the maximum accuracy improvement of 16.01% over the state-of-the-art techniques.

✉ Charu Bhardwaj
    176005@juitsolan.in

    Shruti Jain
    shruti.jain@juit.ac.in

    Meenakshi Sood
    meenakshi@nitttrchd.ac.in

1   Department of Electronics and Communication Engineering, JUIT Waknaghat, Solan, H.P, India

2   Department of CDC, NITTTR, Chandigarh, India

1 Introduction

Diabetes is the significant reason for extreme visual blindness among the grown-up population and expanding diabetic population likewise increases the quantity of individuals with vision loss. Over 10% of diabetic patients experience the ill effects of serious vision impairments and are at high risk of developing eye-related problems. As per the WHO overview, 10% patients have genuine visual disabilities and 2% individuals become blind past 15 years of persisting diabetes [1]. Diabetic retinopathy (DR) is the most extreme eye-related disorder which has become a major issue for the diabetic patients around the world. Early finding hugely affects the prognosis of disorders like

DR, assisting the ophthalmologists with taking essential medicinal measures before it causes irreversible visual impairment.

DR condition produces red (microaneurysms (MAs) and hemorrhages (HMs)) and yellow (exudates (EXs) and cotton wools (CWs)) lesions before leading to the most severe DR stage causing neovascularization [2]. In the absence of early stage diagnosis and treatment of DR, its stages progress from mild non-proliferative DR (NPDR) to moderate, severe (NPDR) and proliferative DR (PDR) stages. Normal retina and DR affected retina with several types of lesions are shown in Fig. 1.

Fundus imaging-based DR screening strategies are utilized because of their easier and advantageous acquisition and better lesion perceivability. The gap between profoundly proficient ophthalmologists and diabetic populace is wide, which brings about the requirement for computerized DR recognition frameworks. With the advancement in the technology, much improvement has been made in computerized strategies for mechanized DR determination. Conventional DR recognition approaches utilize different feature attribute extraction techniques to extricate useful handcrafter features from the fundus images. The extracted handcrafted features are only known through professional expertise with the domain knowledge and should be robust to different lesion variations. The DR screening frameworks should guarantee that the lesions are identified in combination as well as isolation with the other lesions in order to make exact and accurate patient referral decision and assist the ophthalmic specialist for further assessment [3]. No clear symptoms indicative of DR are visible by manual detection hence there is the need for automated DR detection to analyze the pattern and characteristics indicative of diabetic retinopathy.

A number of methods are mentioned in the literature for DR lesion identification followed by DR classification. Kumar et al. [4] presented a CLEAR-DR system to enable and aid the clinical decision support system for Diabetic Retinopathy detection. The effectiveness and utility of the proposed system are demonstrated for DR application, thus

improving the clinician's utilities. ElTanboly et al. [5] proposed an enhanced computer-aided system for grading the different stages of NPDR severities. Twelve distinct features were localized using a segmentation technique for extracting shape, intensity and spatial information from the retinal layers. Curvature, reflectivity and thickness features are measured from the segmented layers. A classification network using the deep fusion method was used to classify the normal and the retinopathy grades, which provides the 93% of accuracy value for binary classification and 98% accuracy for the classification of mild/moderate DR grades. A three-stage classification method for the detection of retinopathy lesions was proposed by Tariq et al. [6]. Lesion extraction was done using the filter bank method and the lesion candidates are extracted for each lesion and then the DR lesions are classified. The performance of the system was validated using different performance metrics. A robust segmentation solution was proposed by Poddar et al. [7] for detection and grading of DR and maculopathy. After applying pre-processing steps for removing inter and intra image variability, various parts of retinal image like OD, fovea, MAs, EXs and HMs are segmented. The pattern and extent of abnormalities in abnormal fundus images are classified into different grades of maculopathy and NPDR using concentric circular zones. Fleming et al. [8] proposed an automated grading scheme to reduce human grading. An automated MA detection scheme was proposed which highlights the importance of contrast normalization to improve the capability to distinguish between other noise spots and MAs. The method provides sensitivity of 85.4% and specificity of 83.1% for images containing MAs. The specialized convolutional networks were utilized by Abramoff et al. [9] for the distinction of different retinal anomalies comprising of hemorrhages, microaneurysms, exudates or neovascularization. The pre-trained networks ranging from AlexNet to VggNet have been significantly employed by the authors which outperforms the existing DR screening solutions. Gargeya and Leng [10] developed a customized approach based on deep CNNs for feature extraction as well as classification of non-DR stage from

Fig. 1 Differentiation of normal retina and diabetic retinopathy affected retina
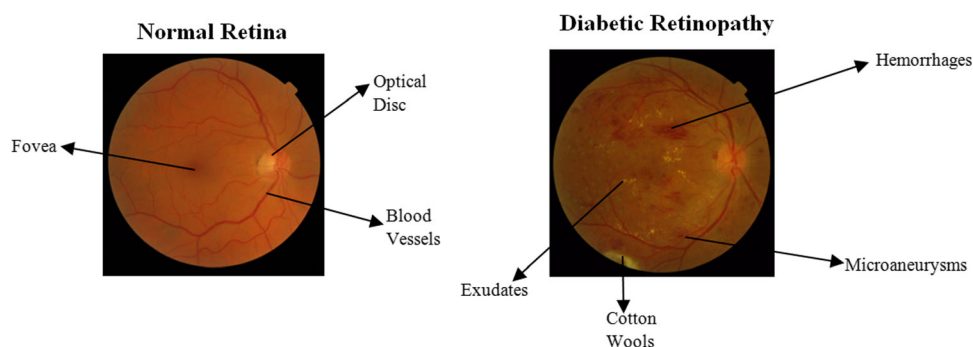
the DR affected stages. The retinal image metadata was combined with the extracted CNN features in order to improve the performance of DR classification method. A heatmap optimization procedure was used by Quellec et al. [11] for CNN training to detect referable DR. It involves image level as well as lesion level classification for the detection of DR-related lesions like hard exudates, soft exudates, small red dots, hemorrhages. An ensembled CNN architecture using Inception-V3 model was proposed by Gulshan et al. [12] for multiple binary decision making to classify between moderate and severe DR, Diabetic macular edema and fully gradable DR. Various segmentation approaches were introduced by Prasad et al. [13] for the detection of retinal blood vessels, hard exudates and microaneurysms. Two-class categorization was accomplished by applying Principal Component Analysis (PCA) for feature selection which is further followed by neural network-based classification employing the back propagation algorithm. Textural analysis technique was proposed by Du and Li [14] for blood vessels and hemorrhage identification. Two-stage deep CNN was employed for automatic DR analysis for red lesion localization and detection in the retinal fundus images followed by severity-level classification. Three-stage classification of fundus images was proposed by Gurudath et al. [15] for DR identification from colored retinal fundus images. Patch-based microaneurysms detection approach was proposed by Cao et al. [16] using SVM, random forest and neural network classifiers for five-stage DR severity grading. Better performance is achieved via dimensionality reduction using principal component analysis (PCA) and random forest feature importance (RFFI). Handcrafter and non-handcrafted features are combined in a computer vision system designed by Nanni et al. [17] to obtain optimal image classification outcomes. PCA, compact binary descriptor (CBD) and CNN are used to extract non-hand-crafted features whereas local phase quantization, completed local binary patterns, rotated local binary patterns algorithms were employed to obtain handcrafted features. A Google Inception v3-based DR assessment approach was introduced by Litjens et al. [18] which yield better results when compared with licensed ophthalmologist's grading. Existing denoising approach was improved by Alban and Gilligan [19] to classify retinal fundus image dataset. Pratt et al. [20] proposed a CNN architecture which can identify complex features for classification of microaneurysms, hemorrhages and exudates. Rahim et al. [21] developed an automatic DR screening system for early detection of microaneurysms. The fuzzy histogram-based technique was used in this method for pre-processing of the fundus images followed by feature extraction and severity grade classification. A computer-aided diagnosis system was developed by Mansour et al. [22] for the detection of

diseased blood vessels and classification of fundus images into different severity levels. AlexNet deep neural network architecture was applied for feature extraction and PCA is used to accomplish dimensionality reduction. Mohammedhasan and Uguz [23] proposes a new strategy for the diagnosis of DR utilizing the residual connections in convolutional neural networks. This method strengthens the CNN model by the using PCA for dimensional reduction improving the diagnostic accuracy of the proposed model. The DR diagnosis accuracy was explored using Kaggle dataset which yields the robust and promising outcomes comparative to the other diagnostic approaches for symptomatic and non-symptomatic DR classification.

High diagnostic relevance of DR classification for better interpretation encouraged various researchers for developing automated DR diagnosis system. Automated DR diagnosis is significant in managing the ophthalmologists' burden by identifying the patients who need critical ophthalmic care [24]. For image categorization and recognition tasks, CNNs have exhibited revolutionary performance and therefore, they are being exploited for DR diagnosis frameworks as CNN ensures proper convergence without over-fitting. Regardless of numerous advances that have been made for utilizing CNN in DR analysis, still these systems present difficulties for clinical applications. Some of the challenges and constraints are being addressed in this article by utilizing a CNN-based approach employing transfer learning.

The DR classification problem is addressed through Convolutional neural networks by using the hierarchical and discriminatory feature learning strategy that automatically classifies the diabetic retinopathy grades. The classical CNN classification approaches for diabetic retinopathy severity grading extracts the complex feature attributes which increases the system complexity. Therefore, an automated DR severity grading model based on transfer learning is proposed in this work which utilizes the CNN's convolutional base for extraction of feature attributes and SVM classifier to aid supervised classification. The concept of transfer learning is fortified from the human capabilities of knowledge transfer in order to perform the tasks more precisely. This method is particularly significant for the task-specific applications dealing with the limited amount of annotated data.

This paper emphasizes on the development of a transfer learning-based CNN system which can detect the DR lesions inevitably and can assess its severity grades into several DR stages. The transfer learning strategy used in this work freezes the convolutional base in its original form and the output of fully connected layers is provided to the machine learning classifier. Our approach aims at proposing two models; image feature based transfer learning (IFTL) model that extracts image feature vector (IFV) from

the fully connected layers of CNN models and another prominent feature-based transfer learning (PFTL) model which reduces the irrelevant features by applying statistical approaches to obtain prominent feature vector (PFV). In the classification stage of our proposed IFTL and PFTL models, the supervised machine learning based classifier is utilized to obtain high classification accuracy maintaining a trade-off with training rate and computational time. The novelty of the proposed mechanism lies in the utilization of the pre-trained architecture as a fixed feature extractor, statistical analysis of extracted features and modification of the network by adding a supervised machine learning based classifier at the end of the network. This innovation of the proposed method is very beneficial in reducing the computational complexity of the network and provides the accurate classification outcomes even for the smaller dataset. The ensemble of convolution neural networks (CNNs) with SVM classifier resulted in remarkable performance for learned model. Feasibility of our proposed approaches is assessed using standard MESSIDOR dataset for retinal fundus images and its effectiveness is evaluated in terms of various performance indices. The performance validation of the proposed method is done over the latest IDRiD dataset justifying the generalization ability of DR grading irrespective of the dataset being used. The robustness of statistically optimized PFTL model is justified using tenfold cross validation. Comparative analysis with other classification algorithms represents that our proposed system outperforms over the state-of-the-art approaches.

The remaining paper is organized as follows: Section 2 discusses the theory and method involved in this article for retinal fundus image classification. Experimental results are discussed in Sect. 3 followed by conclusion in Sect. 4.

2 Materials and methods

The proposed technique aims at building a capable automated DR classification solution for fundus images according to the diseases and their severity levels. The materials and methods utilized in this work to achieve the system for DR severity grade classification are detailed in the following section.

2.1 Materials

The retinal fundus image dataset utilized in this work is described on the basis of number of images, its field of view, size and resolution. The experimentation for this work is done on the benchmark Methods to Evaluate Segmentation and Indexing Techniques in the field of Retinal Ophthalmology (MESSIDOR) dataset for DR

severity grading [25]. There are a total of 1200 fundus images in this dataset which have a 45° Field of View (FOV) and resolution. The 800 images were taken after pupil dilation and other 400 images were taken without pupil dilation. All the images are acquired using standard sizes of 1440 × 960, 2240 × 1488 or 2304 × 1536 which are captured at 8-bit color plane. MESSIDOR dataset provides DR severity categorization guidelines in terms of retinopathy grades and risk for macular edema.

The feasibility of the proposed system is validated using the latest Indian Diabetic Retinopathy Image Dataset (IDRiD) [26] in this research work. The sample images are increased using the retinal fundus images from IDRiD (Porwal et al. 2018) dataset in order to realize the disease progression [27]. The dataset comprises clinical fundus images of thousands of patients examined at Eye Clinic located in Nanded, (M.S.), India, during 2009–2017. Kowa VX-10a digital fundus camera was used to capture these images at 50° FOV while maintaining a distance of 39 mm between the camera lens and the eye. IDRiD dataset contain 454 images with NPDR severities, out of which 168 have no DR sign, 25 images show mild NPDR signs, 168 have moderate and 93 images have severe NPDR symptoms.

In this article, DR severity grading is done based on the retinopathy grades provided by the ophthalmic experts tabulated in Table 1. The dataset comprising a totality of 1200 images consists of 546 normal fundus images having no symptoms of retinopathy, 153 images have symptoms of mild NPDR, 247 images have moderate NPDR symptoms and 254 images have severe NPDR symptoms. Table 1 depicts the DR severity grades along with the number of images present in each grade for the MESSIDOR as well as IDRiD dataset.

2.2 Conventional methodology

Conventional DR classification model used in this work utilized CNN based classification approach. It involves image acquisition, background elimination, DR classification and prediction using pre-trained CNN models followed by model evaluation. The flowchart of the conventional CNN based DR classification and prediction model is shown in Fig. 2.

The basic building blocks of the conventional CNN based DR prediction model are convolution neural network layers which are detailed in the following section.

2.2.1 Convolutional neural networks (CNN)

CNN is significant for image classification tasks providing superior performance. The CNN based DR classification system removes the barrier between experienced clinicians

Table 1 Retinopathy grading for MESSIDOR and IDRiD datasets

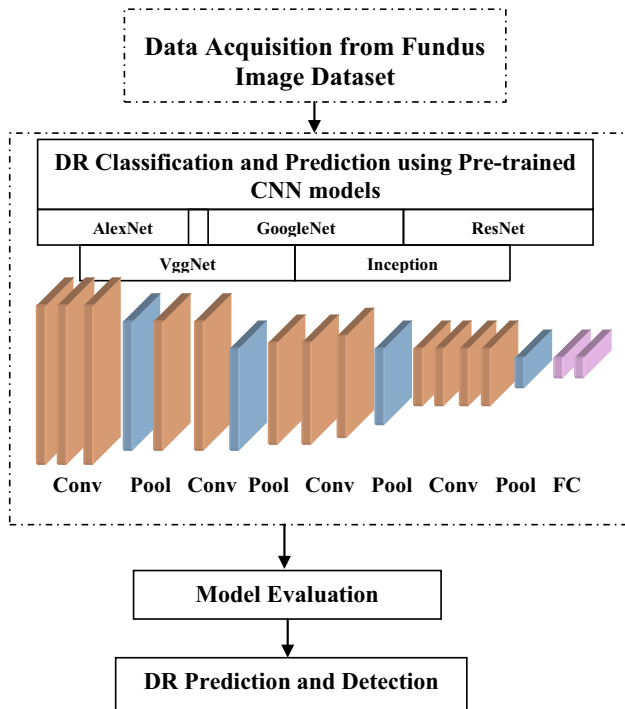| DR severity grades	MESSIDOR dataset images	IDRiD dataset images
Normal fundus with no symptom of DR	546	168
Symptoms of mild NPDR	153	25
Symptoms of moderate NPDR	247	168
Symptoms of severe NPDR	254	93

Fig. 2 Block diagram of convention methodology

and patients. LeCun et al. [28] initiated the CNN based networks which deliver exceptional performance for image-oriented tasks. CNN comprises three basic elements; convolutional layer, pooling layer and fully connected layer.

Convolution layer Convolution layer is the key component of CNN model and it consists of a large set of learnable convolutional filters. For an input image X of size, $a \times b \times c$, $a \times b$ is the image size and c represents the number of channels. The convolutional layer function is expressed by Eq. (1).

$$f(X) = \sum_{k=0}^{c} w_k * X_k + \text{bias} \qquad (1)$$

where w is the weight vector, k is total number of nodes and * denotes the operator of convolution.

Convolution layer is utilized to acquire feature maps by convolving the input image and the weight vector. The weight vector is adjusted during the training phase using the backpropagation algorithm. Rectified linear unit (ReLU) introduces the non-linearity having constant gradient property for positive input in spite of its non-differentiable nature and therefore, is a typical choice for CNNs [29]. The ReLU function is expressed by Eq. (2).

$$\text{ReLU}(X) = \begin{cases} 0 & \text{if } X < 0 \\ X & \text{otherwise} \end{cases} \qquad (2)$$

ReLU function is used to accomplish preferable results over the customary sigmoid or hyperbolic tangent functions by accelerating the training.

Pooling layer Down-sampling is accomplished using max-pooling layer. The input feature map is separated into small patches over the entire feature map using a sliding window considering a specific stride size and maximum of every patch is calculated by the means of max operation. This layer is appended in between adjacent convolution layers for intermediate dimensionality reduction.

Fully connected Layer Fully connected layer act as a conventional neural network and each node is directly associated to every other node of the successive layer. Fully connected layers are appended after the series of subsequent convolution and pooling layers to facilitate the output flattening into a single vector. Fully connected layers consist of large network parameters leading to complex computations. The dropout technique is used to reduce computational complexity by dropping out some nodes and connections of the output layer. Output layers use a particular activation function that permits output interpretation directly as probability likelihood function whose range is in between 0 to 1. The softmax activation function is generally used in conventional CNN model as the output activation function and it is expressed by Eq. (3).

$$\text{Softmax}(X) = \frac{e^{yj}}{\sum_j e^{yj}} \qquad (3)$$

where e^{yj} is exponential function of each element of output vector.

The different pre-trained CNN architectures of different CNN models based on different layers, activation and network parameters are reported in the literature.

2.2.2 Different CNN architectures

Every CNN architecture has its own practical application for different image classification tasks. However, the network parameters can be fine-tuned to exploit its characteristics for another set of data. Initially, the weights are treated as the objective of CNN model which are adjusted as per the desired output and then supervised model training is done on the test data [30]. Every CNN layer extracts meaningful image information depicting a new input image representation.

As per the review of state-of-the-art techniques, there exist various CNN models for different classification tasks including LeNet [28], AlexNet [31], GoogleNet [32], VGGNet [33], ResNet [34] and Inception Networks [35] some of which are depicted in Fig. 3 and are detailed in the following sub-sections.

AlexNet: AlexNet was introduced in 2012 and it was similar to the primary LeNet architecture which was proposed by Lecun et al. [28]. In comparison with the initial one, the AlexNet architecture is much deeper including supplementary filters in each layer. It comprises of variable convolutional filters of size $11 \times 11, 5 \times 5, 3 \times 3$, max pooling and ReLU activations are attached after every convolutional layer [31]. The layer-wise description of AlexNet is given in Table 2.

AlexNet was the first to execute ReLU activation function to advance the training speed while increasing the network performance. The network size was reduced by the pooling layer and dropout layer was used to reduce overfitting. AlexNet reduces the learning rate during the training process when the accuracy value plateaus. To overcome this challenge, GoogleNet was introduced in 2014.

Inception V1 (GoogleNet) GoogleNet [32] was presented in 2014 with a benefit of providing minimized error rate
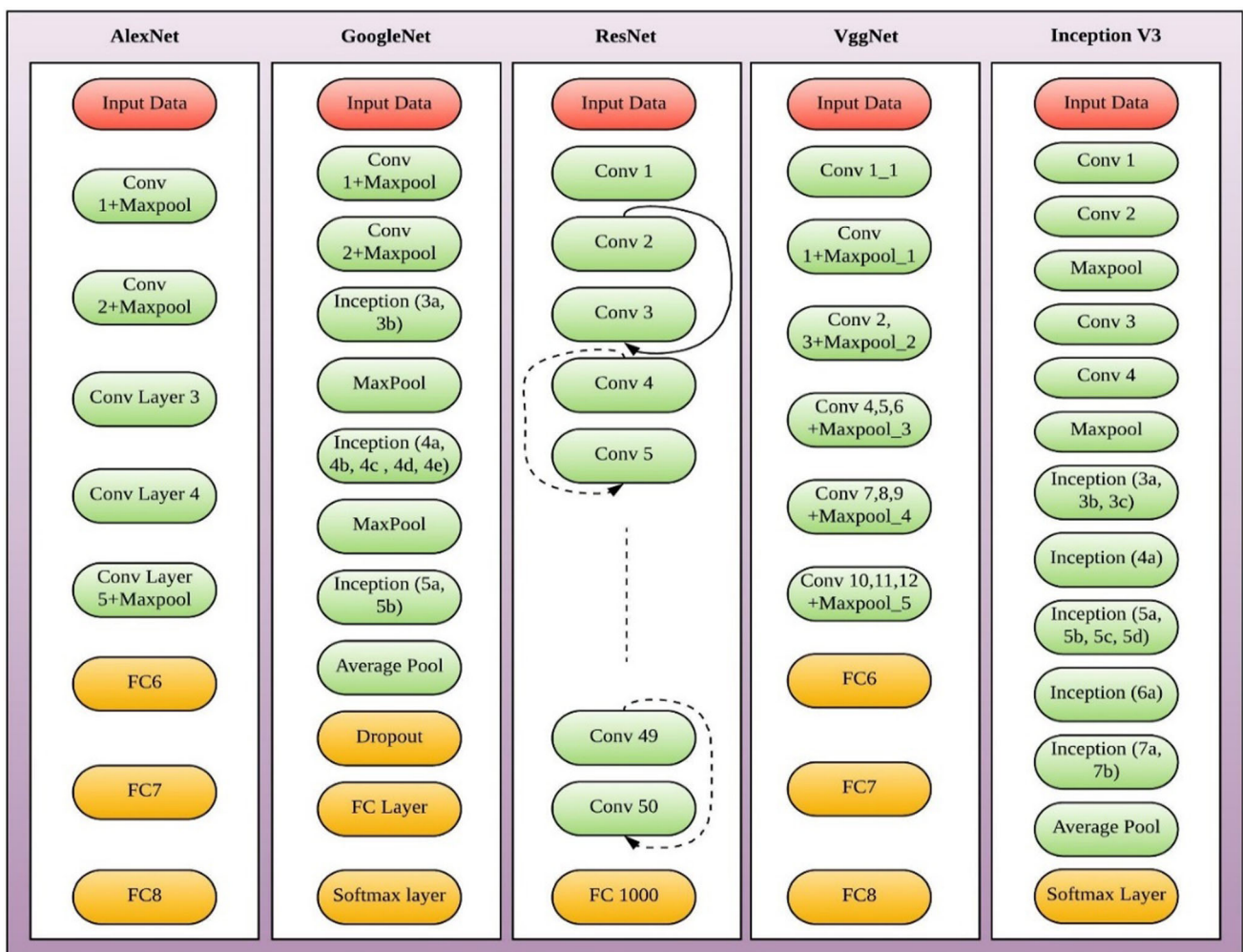

Fig. 3 Architectural representation of different pre-trained CNN models

Table 2 Layer-wise architectural explanation of AlexNet

Layer Index	Type	Activations	Learnables		
			Filter size	Stride	Number of filters/pooling type
0	Input data	$(224 \times 224 \times 3)$	–	–	–
1	Conv. layer 1	$(55 \times 55 \times 96)$	(11×11)	4	96
2	Pooling layer 1	$(27 \times 27 \times 96)$	(3×3)	2	Max pool
3	Conv. layer 2	$(27 \times 27 \times 256)$	(5×5)	1	256
4	Pooling layer 2	$(13 \times 13 \times 256)$	(3×3)	2	Max pool
5	Conv. layer 3	$(13 \times 13 \times 384)$	(3×3)	1	384
6	Conv. layer 4	$(13 \times 13 \times 384)$	(3×3)	1	384
7	Conv. layer 5	$(13 \times 13 \times 256)$	(3×3)	1	256
8	Pooling layer 5	$(6 \times 6 \times 256)$	(3×3)	2	Max pool
9	Fully conn. layer 6	$(1 \times 1 \times 4096)$	–	–	–
10	Fully conn. layer 7	$(1 \times 1 \times 4096)$	–	–	–
11	Fully conn. layer 8	$(1 \times 1 \times 1000)$	–	–	–

relative to another accessible CNN networks. The parameters in the Inception V1 architecture are drastically diminishing to 60 million utilizing fewer number of convolutions. "Network in network" concept is utilized for building a dense network assembling the inception modules rather than assembling individual convolutional layers while expanding the network depth. The disadvantage of this architecture is that it saturates the network accuracy while expanding its depth, hence, ResNet architecture appeared in 2015 that uses the concept of skip connections while preserving the depth of the model.

Residual Neural Network (ResNet) ResNet [33] was presented in 2015, which utilizes the gate skip connections and batch normalization concept for improving the network depth. The powerful representation capability of these networks has boosted the usage of ResNet for various image recognition, object identification and facial recognition tasks. ResNet provides good network performance but its architecture is somewhat intricate and complex. Different CNN variations appeared in 2015, VggNet is one of them which notably has simpler implementation as well as expands the network depth.

Visual Geometry group networks (VggNet) VggNet [34] is a multi-layered deep neural network architecture that is presented in 2015 due to its simple implementation and improved network depth. VggNet comprises of large number of network parameters utilizing higher storage space of 500 MB. VggNet has a constraint of slow training along with higher storage space requirement making its deployment tedious. This impediment is addressed by

another derivative of Inception networks that is Inception-V3 model.

Inception-V3 Incpetion-V3 is a deep neural network intended to categorize 1000 object categories [35]. A wide variety of images are used to train the model and maintaining that training knowledge, the model can be retrained for a smaller dataset. This benefit of Incpetion-V3 CNN model reduces the need for extensive training resulting in higher classification accuracy along with least computational time. The network architecture of Inception-V3 is tabulated in Table 3 which details the type of layers, activation function and the learnables.

The main focus of Inception-V3 network is to remove the bottleneck representation of adjacent network layers that drastically diminishes the input dimensions of the next layer. This is done by utilizing the using factorisation technique for reducing the computational complexity of the network. In the conventional methodology, fundus images are acquired from the image dataset and the classification outcomes for DR prediction and detection are directly obtained from the fully connected predictions classification layer of CNN model. The methodology proposed by the authors in this paper is highlighted in the following section.

2.3 Proposed methodology for transfer learning based automated DR detection system

The proposed technique aims at building a capable automated DR classification solution for segmenting the fundus images according to the diseases and their severity levels. In this paper, the authors proposed a transfer learning based

Table 3 Layer-wise architectural description of Inception-V3 network

Layer Index	Type	Activations	Learnables		
			Filter size	Stride	Number of filters /pooling type
0	Input data	$(299 \times 299 \times 3)$		–	
1	Conv. layer	$(149 \times 149 \times 32)$	(3×3)	2	32
2	Conv. layer	$(147 \times 147 \times 32)$	(3×3)	1	32
3	Conv. layer	$(147 \times 147 \times 64)$	(3×3)	1	64
4	Pooling layer	$(73 \times 73 \times 64)$	(3×3)	2	Max pool
5	Conv. layer	$(73 \times 73 \times 80)$	(1×1)	1	80
6	Conv. layer	$(71 \times 71 \times 192)$	(3×3)	1	192
7	Pooling layer	$(35 \times 35 \times 192)$	(3×3)	2	Max pool
8	Inception module (3A)	$(35 \times 35 \times 256)$		–	
9	Inception module (3B)	$(35 \times 35 \times 288)$		–	
10	Inception module (3C)	$(35 \times 35 \times 288)$		–	
11	Inception module (4A)	$(17 \times 17 \times 768)$		–	
12	Inception module (5A)	$(17 \times 17 \times 768)$		–	
13	Inception module (5B)	$(17 \times 17 \times 768)$		–	
14	Inception module (5C)	$(17 \times 17 \times 768)$		–	
15	Inception module (5D)	$(17 \times 17 \times 768)$		–	
16	Inception module (6A)	$(8 \times 8 \times 2048)$		–	
17	Inception module (7A)	$(8 \times 8 \times 2048)$		–	
18	Inception module (7B)	$(8 \times 8 \times 2048)$		–	
19	Pooling layer	$(1 \times 1 \times 2048)$	(8×8)	8	Avg pool
20	Predictions layer	$(1 \times 1 \times 1000)$		–	

model employing the CNN architectures as feature extractors, a separate feature selection module for dimensionality reduction and supervised SVM classifier is employed for DR identification and prediction. To achieve the system for DR severity grade classification, the proposed approach is divided into several sequential steps for better implementation which are demonstrated in Fig. 4 and detailed in the following section.

Lack of labeled data is a limitation in building good classifier which was overcome by transfer learning approach. Transfer learning aims at extracting the source knowledge and apply that knowledge to the target application. The tradition machine learning method is replaced with the transfer learning module to utilize the knowledge gained from the previous tasks for future learning and categorization. The proposed methodology of transfer learning based automated DR severity grading is detailed in Algorithm 1.

Algorithm 1: Proposed Methodology for Transfer Learning based Automated DR Severity Grading

---

Input: Fundus Images acquired from the dataset.

Output: Diabetic Retinopathy grades distinguishing between different DR severities.

Begin

Step 1: Acquire the retinal fundus images from the image datastore and annotate them into different DR grades.

Step 2: Apply background elimination steps to attain the background segmented image.

   Step 2.1: Pre-processing of fundus images

- Apply the image pre-processing steps like:
  - Color normalization to model the fundus image *(I(x,y))* into corresponding R,G, B channels.
  - Median filtering to distinguish the isolated noise from the normalized image.
  - CLAHE on the denoised image to attain the pre-processed output fundus image ($I_{pre-process}$).

   Step 2.2: Optical Disc (OD) segmentation

- Apply Hue, Saturation and Value (HSV) model conversion on the pre-processed image ($I_{pre-process}$) to obtain the intensity values.
- The largest circular region (I_{OD}) is identified using Morphological closing operation and contour filling.
- The OD segmentation ($I_{ODsegment}$) is obtained after OD removal.

$$I_{ODsegment} = I_{pre-processed} - I_{OD}$$

Step 3: Apply data augmentation including horizontal and vertical flipping, random rotation at 90-180 degrees and random zooming of [0.85, 1.15] on the background eliminated image.

Step 4: Feed the data augmented fundus images to the CNN models: *AlexNet, GoogleNet, ResNet, VggNet and Inception* networks for extraction of features from the fully connected layers.

Step 5: Utilized both Image Feature based Transfer Learning (IFTL) and Prominent Feature based Transfer Learning (PFTL) model for DR classification.

   Step 5.1: The features extracted from the convolutional base are directly fed to the SVM supervised machine learning classifier in case of IFTL model (CNN+SVM).

   Step 5.2: The convolutional base features are initially provided to the feature selection module and then the selected features are fed to the SVM classifier in case of PFTL model (CNN+Statistical Analysis+SVM).

Step 6: Attain the classification results in terms of four classes of Diabetic Retinopathy grades; normal, mild, moderate, severe.

Step 7: Access the performance of the networks by using various evaluation parameters.

Step 8: Compare the proposed statistically optimized PFTL model with conventional CNN, (CNN+kNN) and (CNN+SVM) to justify its superiority.

Step 9: Assess the validation performance of the proposed statistically optimized PFTL model using the latest IDRiD dataset.

Step 10: Obtain a generalized, feasible and robust solution to attain DR prediction and detection.

End

The two models: IFTL and PFTL as depicted in Fig. 4, utilize the transfer learning approach for the detection of DR severity. At the fully connected layer, linear combination of the output from the previous hidden layers acts as the input to the output layer activations. CNN output units
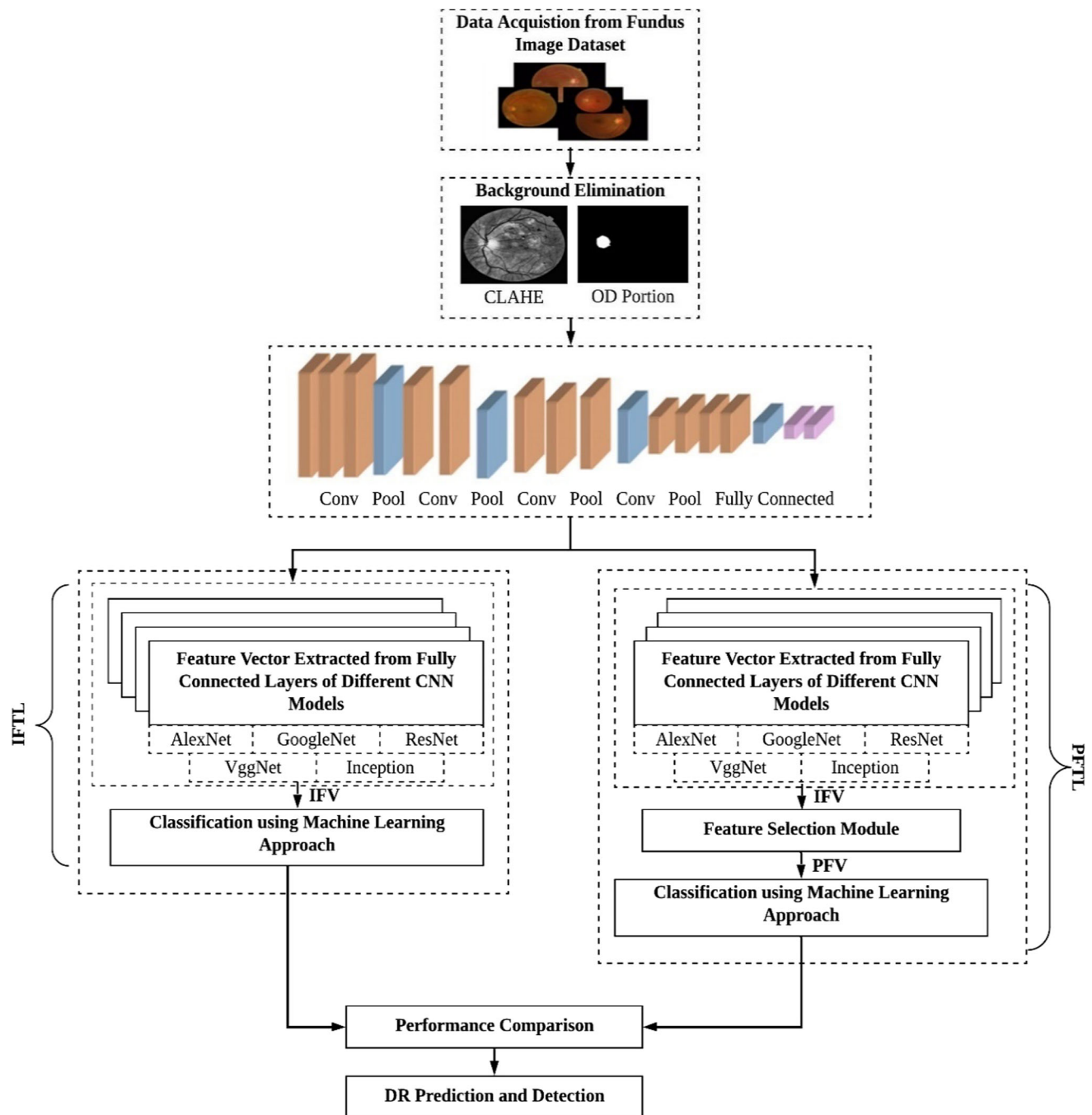
Fig. 4 Automated DR detection system employing transfer learning. *Note: *CLAHE* contrast limited adaptive histogram equalization, *OD* optical disk, *IFV* image feature vector, *PFV* prominent feature vector

calculate the estimated probabilities for the input data depending upon the activation function. The architecture of our proposed models utilizes the outputs of fully connected layers of respective pre-trained CNN models as features, which serves as the input feature vector to the supervised machine learning classifiers for training. Machine learning approach is applied on the image feature vector for detection of DR grades. Once the classifier is trained, it

performs the prediction task for other testing samples exploiting the extracted CNN feature attributes.

2.3.1 image feature based transfer learning (IFTL) model

The proposed IFTL DR prediction model (CNN + SVM) utilizes background elimination technique after data acquisition. Further, background eliminated fundus images

are provided to the CNN network. A novel IFTL approach is proposed in this article, in which Image Feature Vector (IFV) is obtained from the convolutional base of CNN model and they are further subjected to machine learning classification approach using SVM classifier for DR prediction. This proposed approach incorporates the merits of both CNN and SVM classifiers as SVM advances the generalizability of CNN by diminishing the errors and improving the classification accuracy outcomes. Flowchart of the proposed IFTL approach is given in Fig. 5.

Background elimination, extraction of image feature set using CNN and severity grade classification are the three main steps involved in the proposed technique is and their detailed description is given in the following sections.

Background Elimination Raw fundus images consist of various image artefacts due to inadequate image contrast, insufficient background illumination, bad aligning of lens, etc. [36]. Retinal fundus images were taken using different fundus cameras at varying surrounding conditions, resolutions and different field of views. Image pre-processing is needed to lessen these artefacts before optical disk (OD) localization and extraction. Image pre-processing steps used in this work include image variation attenuation, intensity conversion to green channel, medial filtering and contrast limited adaptive histogram equalization (CLAHE) [37, 38]. Pre-processing stage is helpful in differentiating between actual lesion regions and noisy data. Optical disc is also segmented at the initial stage as they are not required for anomaly detection and can be misclassified as pathological symptoms. Morphological closing operation is used for boundary localization where the principal rounded region in the intensity plane is designated as the OD portion. OD removal segmentation is attained by the subtraction of the OD segmented part from the pre-processed fundus [39, 40].

Medical image applications do not provide a large annotated data as it is a laborious and expensive job indulging highly experienced professional ophthalmologists. Different diseased classes have varying number of images so, neural networks also suffer from class imbalance problem. Class imbalance and small dataset problem is mitigated to advance the localization capability of the network using data augmentation step. Various steps are involved in data augmentation stage comprising horizontal and vertical flips, random rotation of 90–180°, random zooming of [0.85 to 1.15] [36]. Training examples and the class size are increased using data augmentation. This process improves the classification performance by avoiding over-fitting. Image pre-processing and data augmentation steps make the CNN model insusceptible from

different orientations, variation attenuation and insufficient illumination.

Image Feature Vector (IFV) Extraction from Fully Connected Layer of CNN In this paper, the authors have implemented AlexNet, GoogleNet, ResNet, VggNet and Inception pre-trained CNN networks to achieve significant improvement in fundus image classification application. In the proposed transfer learning based technique, feature extraction is accomplished using the fully connected layers of various pre-trained CNN architectures. IFVs are extracted from FC8 layer for AlexNet, loss3-classifier layer for GoogleNet, FC1000 for ResNet, FC8 for VggNet and predictions layer for Inception-V3, respectively.

Retinal Fundus Image Classification using IFV Different machine learning based classifier combinations were used in the literature to obtain better accuracy rate for grading features of normal and abnormal fundus images into different grades. Machine learning approaches provide the flexibility of predicting the future events using the past learning from the labeled set of images. The machine learning model can be trained to predict the output by measuring the error between the output obtained and the correct output and thereby modifying the learning algorithm. From the literature review, supervised support vector machine (SVM) Classification method suitable for binary and multi-class discrimination is used for DR prediction problem [41]. SVM is a classification based machine learning algorithm and its goal is to divide the dataset into different classes by finding a maximum marginal hyperplane. In SVM-based machine learning classification approach, testing data are optimally categorized via a hyperplane created using labeled data for training. This technique is appropriate for both binary and multi-class classification [42, 43]. The DR classification problem is dealt using one-against-rest approach in which different classifiers are prepared for each class. For the input vectors denoted by $\vec{i}$, weight vector $\vec{w}$ and bias b, the equation for hyperplane is represented by Eq. (5).

$$\vec{w} \cdot \vec{i} - b = 0 \tag{5}$$

The tuning parameters used to attain improved accurate classification outcomes are regularization parameter I and gamma (γ), which are stated as the error function given in Eq. (6).

$$C \sum_{i=1}^{L} \left(\varepsilon_i^+ + \varepsilon_i^- \right) + \frac{1}{2} w^2 \tag{6}$$

where C is the regularization parameter, ε is the error metric subjected to, $\varepsilon^+ \geq 0$, $\varepsilon^- < 0$ minimization constraint and w = weight vector.
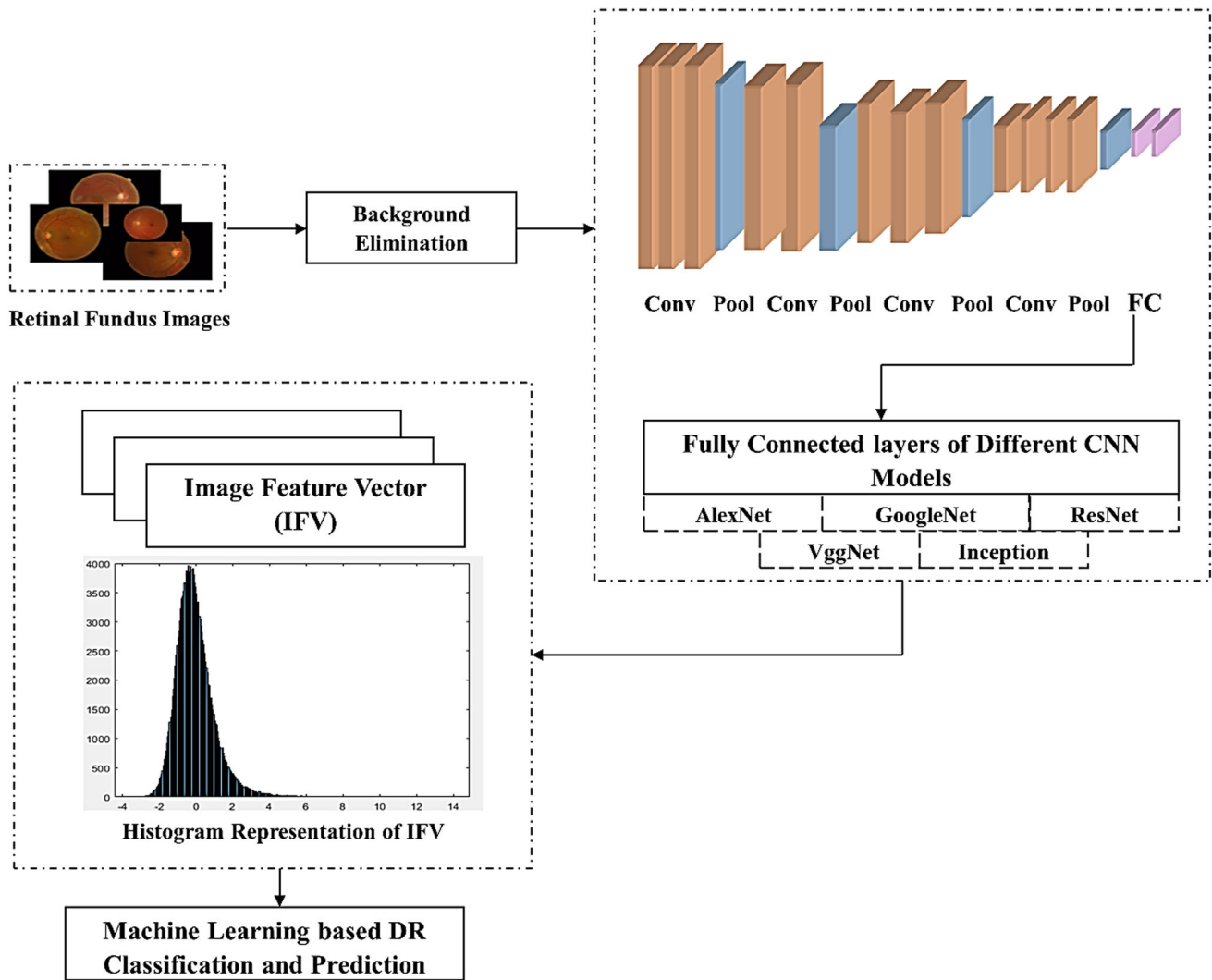
Fig. 5 Flowchart of proposed IFTL (CNN + SVM) model

The selection of CNN as feature extractor and SVM classifier for classification is inspired from the literature Tang [44], who initiated the concept of utilization of features extracted from the convolutional base and improve the network classification performance by training a SVM classifier with these features. The combination of CNN and SVM provides better network performance compared to the other preeminent approaches present in the literature [45, 46]. SVM classifier is an effective supervised learning algorithm that finds the optimal hyperplane for separation of feature space while dealing with smaller dataset. The SVM classifier is selected because of its superior optimization strategy using the stochastic gradient descent and superior regularization effect with less computational complexity comparative to its counterparts like k-nearest neighbor, ensemble classifier, multi-layer perceptron, etc.

2.3.2 Prominent Feature-based Transfer Learning (PFTL) Model

IFV extracted from CNN model are similar and complex increasing the system complexity. However, removal of irrelevant features reduces the complexity resulting in the prominent features for classification. PFTL model (CNN + Statistical Analysis + SVM) proposed in this work utilizes IFV obtained from CNN-based feature extractor and further subjected to feature selection module for dimensionality reduction extracting prominent feature vector (PFV) for classification. The selected PFV is classified using SVM classifier for machine learning-based classification. The flowchart of the proposed PFTL technique is depicted in Fig. 6.

The statistical analysis approaches are used to analyze the spatial distribution of gray level values in an image by computing its local features. The broad classification lies

into the first-order statistics, second-order statistics and higher-order statistics. This work utilizes the first-order statistical methods by exploiting the pixel occurrence probability for finding out the textural feature properties in an image. The major advantage of using this histogram based approach lies in its simplicity as it uses the standard mean and variance descriptors for categorization of image textural data. The histogram distributions for both IFV and PFV derived by optimal statistical analysis are obtained. The distribution of PFV lies in the corresponding range as that of IFV histogram providing better classification results after discarding the irrelevant features. Highly peaked features present in IFV represent the outliers that are removed using feature selection through statistical analysis. Particularly for fundus image classification task, statistical feature reduction methods are significant in reducing the

feature vectors with somewhat identical/ similar elements and transfers the most significant feature vectors.

Optimal feature selection to obtain prominent feature vector (PFV) Optimal feature selection is significant in dimensionality reduction and thereby, improving the performance of the network in terms of computation time and storage. In this work, t test and analysis of variance (ANOVA) statistical tools are used to decrease the dimensional complexity while increasing the level of accuracy. Diagnostic accuracy is improved by selecting only the most prominent features out of the entire feature vector by using the substantial value (p value) analysis indicating the significance of a specific feature and it is expressed in Eq. 4.
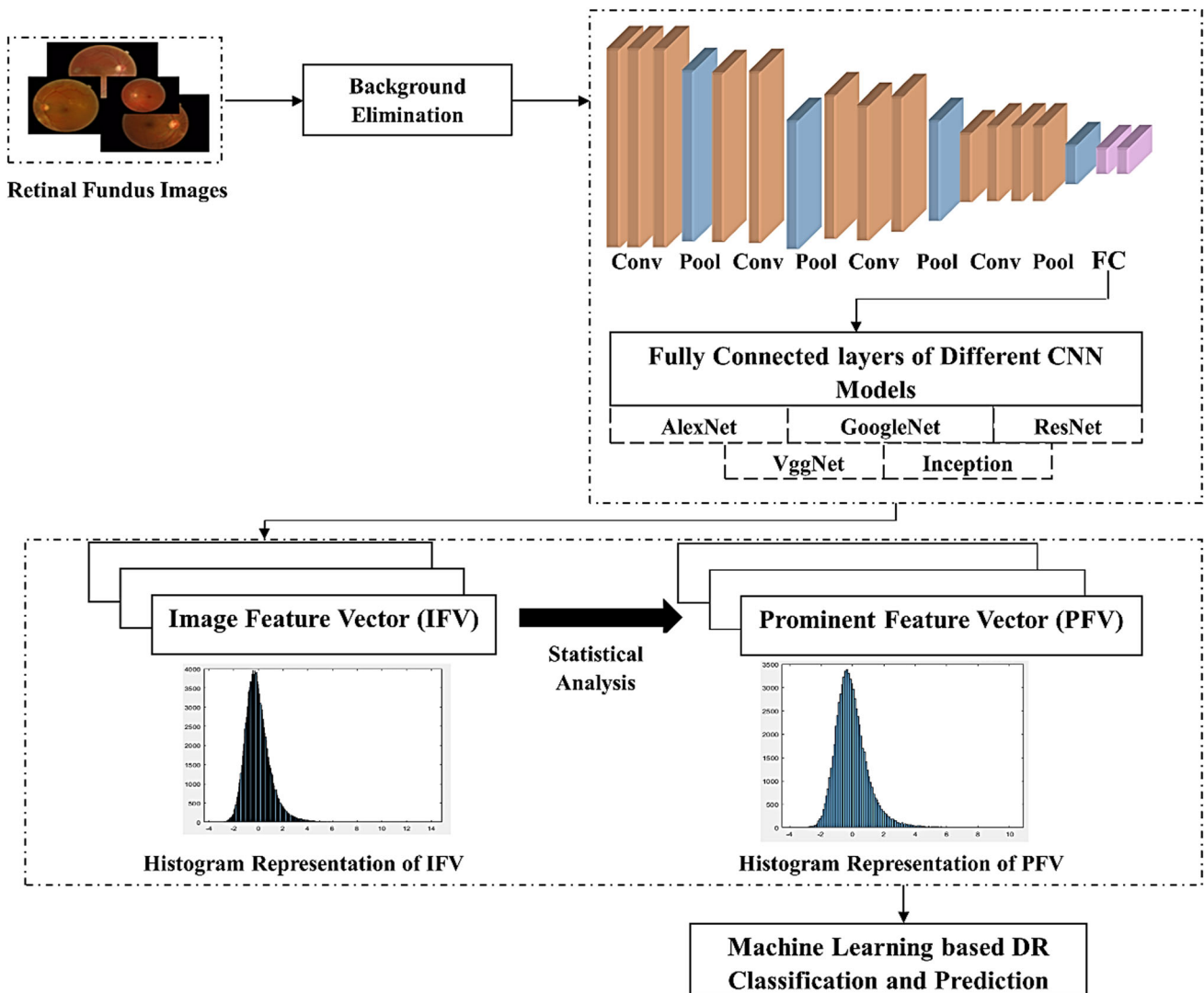

Fig. 6 Flowchart of the proposed PFTL (CNN + Statistical Analysis + SVM) model

$$p = \begin{cases} p \leq 0.05 & \text{indicates Strong Significance} \\ 0.01 < p \leq 0.05 & \text{indicates Moderate Significance} \\ p > 0.05 & \text{indicates Weak Signicance} \end{cases}$$

(4)

where p is the significance value.

If the significance level of a particular feature provides the p value less than 0.05 then that particular feature is considered relevant enough to be used for the classification otherwise for p value > 0.05 moderate or weak significance is shown and that feature is discarded. Features having p value ≤ 0.05 are selected as prominent features in the selection process [47, 48].

The functionality of the statistical approaches shifts the higher dimensional feature space to the lower dimensionality which contains only most important prominent feature vector (PFV). In diabetic retinopathy problem, fundus image mostly contain interrelated features so statistical feature selection approaches yields the PFV containing feature which are different from each other. These selected prominent features are further subjected to supervised SVM classification for precise prediction of diabetic retinopathy. The proficiency of DR severity classification is observed in terms of various performance indices.

2.4 Evaluation parameters

The results of the proposed transfer learning based DR grading system are validated quantitatively using various performance parameters like positive prediction value (PPV), negative prediction value (NPV), sensitivity (Sn.), specificity (Sp.), F-measure, accuracy, area under the receiver operating characteristics curve (AUC) and cross-entropy loss [48]. These performance metrics are defined in terms of true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN) and are tabulated in Table 4.

3 Experimental results and discussion

The experimentation was performed using MATLAB2019b environment on computer system equipped with Intel Core i5 processor, 8 GB RAM and NVIDIA GeForce 4 GB GPU. The setting of hyper-parameters for the experimentation includes the batch size, base learning rate, dropout rate, epoch sizes and iteration count which are tabulated in Table 5.

The MESSIDOR dataset is partitioned into training/testing set based on 70–30% criteria. Out of total 1200 images in the dataset, 840 images are utilized for training and 360 fundus images are considered for testing. tenfold cross validation is ensured to prevent the poor

randomization. The feature vector resulting from the convolutional base of different pre-trained CNN models is utilized to train the SVM classifier. The SVM training process is accomplished using the stochastic gradient descent for network weight adaptation. The complexity analysis of the proposed approach is done in terms of computational time and cross-entropy loss function. The experiments were performed for varying epoch sizes within the range of 1 to 20 and at iteration count ranging between 1 to 200.

To check the validity of the proposed methodology, the conventional CNN based DR classification model is tested and the accuracy values are evaluated for AlexNet, GoogleNet, ResNet, Vgg16, Vgg19 and InceptionV3 pretrained models. Performance assessment of conventional CNN classification model is tabulated in Table 6 in terms of its accuracy value.

To improve the network performance, in terms of various evaluation parameters, authors have proposed two different transfer learning based DR detection models using the feature vector obtained from the CNN network which are further classified using machine learning approach. IFTL (CNN + SVM) transfer learning approach utilizes IFV extracted from fully connected layer of CNN models and use them to train the supervised machine learning SVM classifier. For the proposed PFTL approach (CNN + Statistical Analysis + SVM), IFV is reduced to PFV using statistical analysis methods and these optimal features are further provided to SVM classifier to obtain a robust and feasible performance for severity grade classification in terms of various indices. To justify the superiority of the proposed statistically optimized PFTL (CNN + Statistical Analysis + SVM) model a comparison is drawn with conventional CNN, (CNN + kNN) and IFTL (CNN + SVM) methods. The experimentation performance is also validated over latest IDRiD dataset to make the proposed approach generalized irrespective of dataset being utilized.

The performance comparison of the proposed IFTL and PFTL approaches utilizes various evaluation parameters is tabulated in Table 7.

From the results in Table 7, it is revealed that the classification outcomes after feature selection using optimal Prominent Feature Vector (PFV) and the CNN models achieve phenomenal performance in terms of various parameters for DR classification task. Utilization of PFTL (CNN + Statistical Analysis + SVM) approach provides finest performance for InceptionV3 model and also comparatively better performance is encountered for Vgg16 CNN model for all the parameters. Feature disparities of IFTL are reduced using PFTL approach which utilizes statistical methods to obtain optimal feature vector for proficient DR classification. Maximum 90.51% accuracy is

Table 4 Evaluation Parameter

Performance parameter	Formula
Positive prediction value or precision (PPV)	$Positive\ Prediction\ Value\ (PPV) = \frac{TP}{TP+FP}$
Negative prediction value (NPV)	$Negative Prediction Value\ (NPV) = \frac{TN}{TN+FN}$
Sensitivity or recall rate (Sn.)	$Sensitivity\ (Sen.) = \frac{TP}{TP+FN}$
Specificity (Sp.)	$Specificity\ (Sp.) = \frac{TN}{TN+FP}$
F-measure	$F-Measure = \frac{(2*Recall*Precision)}{(Recall+Precision)}$
Accuracy	$Accuracy\ (Acc.) = \frac{TP+TN}{TP+FN+TN+FP}$
Cross-entropy loss	$\frac{1}{\sum_{t=1}^{M} y_{o,t} \log(p_o,t)}$

where M indicates the classes, true classification is indicated by t for observation o, prediction probability p and y is the binary indicator for precise class label classification.

Table 5 Hyper-parameter setting for proposed Method

Hyper-parameters	Assigned values
Batch size	125
Base learning rate	0.0001
Dropout rate	0.5
Epoch size	20
Iteration count	200

achieved by Inception V3 model and also noteworthy performance is accomplished by utilizing Vgg16, Vgg19 and AlexNet CNN models.

Sensitivity and specificity parameters evaluated in Table 7 are equally important in finding the accurateness of the diagnostic test and both should have larger values for disease classification. It is depicted that the numerical value of Sensitivity is higher for PFTL in comparison with IFTL, signifying the higher probability of DR classification system identifying the actual proportion of patients with DR. Higher sensitivity values of 89.89% and 89.32% are achieved for Vgg16 and InceptionV3 models respectively, utilizing PFTL. Specificity outcomes also provide higher values for PFTL in comparison with IFTL signifying higher probability for the analysis and identification of diabetic retinopathy without giving false positive

outcomes. Maximum of 91.75% specificity value is provided by InceptionV3 CNN model.

The positive prediction value (PPV) indicates the probability of fundus images which truly have any symptom of DR; however, the negative prediction value (NPV) indicates the probability of fundus images which truly does not contain any DR symptom. F-measure provides the capability of the model to detect the false events and the accuracy value represents the proportion of correctly predicted DR cases out of the total population. A higher value of all these parameters is recommendable for a good network performance. A higher PPV, NPV and F-measure values of 92.34%, 93.45% and 90.64%, respectively, are achieved for InceptionV3 model utilizing PFTL approach. Similarly, the accuracy values of 76.61% and 90.51% are achieved for IFTL and PFTL, respectively. The capability of the model to successfully differentiate between the positive and negative detection is given by the area under the curve (AUC) values. AUC achieved for PFTL approach provides a much higher value of 0.92 than the IFTL approach which yields a lower AUC value of 0.75 for the most efficient InceptionV3 model. The outcomes obtained provide the higher values for PFTL in comparison with IFTL indicating higher DR classification system probability for identification of diabetic retinopathy.

Table 6 Performance assessment of conventional CNN based DR classification model

Classifier	Accuracy of conventional CNN based DR classification
AlexNet	73.33%
GoogleNet	65.56%
ResNet	65.83%
Vgg16	82.14%
Vgg19	80.76%
InceptionV3	87.50%

Table 7 Classification outcomes in terms of various evaluation parameters utilizing the proposed IFTL and PFTL models for MESSIDOR dataset

	Evaluation parameters	Different CNN Models					
		AlexNet	GoogleNet	ResNet	Vgg16	Vgg19	InceptionV3
Image Feature Transfer Learning Approach (IFTL)	Sensitivity (Recall)	60.37%	64.00%	69.90%	61.70%	75.00%	75.49%
	Specificity	61.70%	64.00%	71.13%	60.38%	75.00%	69.44%
	PPV (Precision)	64.00%	64.00%	72.00%	58.00%	75.00%	77.00%
	NPV	75.32%	75.83%	83.42%	69.43%	86.43%	87.21%
	F-Measure	62.13%	64.00%	70.93%	59.79%	75.00%	76.24%
	Accuracy	61.01%	64.00%	70.50%	61.10%	75.00%	76.61%
	AUC	0.62	0.65	0.71	0.62	0.73	0.75
Prominent Feature Transfer Learning Approach (PFTL)	Sensitivity (Recall)	88.64%	75.00%	81.82%	89.89%	88.12%	89.32%
	Specificity	90.00%	75.00%	76.92%	89.12%	88.89%	91.75%
	PPV (Precision)	88.64%	75.00%	75.00%	89.00%	89.00%	92.34%
	NPV	90.34%	79.67%	78.98%	91.87%	91.23%	93.45%
	F-Measure	88.64%	75.00%	78.26%	89.44%	88.56%	90.64%
	Accuracy	89.36%	75.00%	79.17%	89.50%	88.50%	90.51%
	AUC	0.90	0.73	0.81	0.90	0.89	0.92

3.1 Comparative analysis of different classification approaches

In this work, the performance of prominent classifiers like k-nearest neighbor (kNN) and multi-layer perceptron neural network (MLPNN) has been compared with the proposed model to justify its prominence for classifying the DR severity grades, comparative to the other combinations. The different classifier combinations compared are: the conventional CNN-based classification method, machine learning based (CNN + kNN) classifier, supervised (CNN + MLPNN) classifier, machine learning-based IFTL approach (CNN + SVM), and optimized PFTL (CNN + Statistical Analysis + SVM) approach. The comparative analysis of different classifier combinations in terms of network accuracy performance is given in Fig. 7.

The proposed PFTL approach provides higher accuracy values comparative to conventional CNN, (CNN + kNN), (CNN + MLPNN) as well as IFTL (CNN + SVM) method and maximum of 90.51% accuracy is achieved by optimal PFTL model utilizing Inception V3 CNN architecture. The Inception V3 CNN model yields 74.19% accuracy for (CNN + kNN) classifier, 75.18% accuracy for (CNN + MLPNN) methods, 76.61% accuracy for IFTL (CNN + SVM) based classification approach, 87.50% accuracy using conventional CNN based classification and highest accuracy performance of 90.51% using the combination of (CNN + Statistical Analysis + SVM)

in PFTL model. The ensemble of CNN with other machine learning algorithms like kNN and MLPNN as they classify the individual class on the basis of similar feature properties. These algorithms when combined with CNN module does not provide promising outcomes as they end up dominating some of the similar features in the feature vector domain and therefore, will classify the patterns of similar type, which is not suitable for DR diagnostic application. However, this evaluation reveals that InceptionV3 CNN model equipped with the combination of (CNN + Statistical Analysis + SVM) in PFTL approach provides the highest network performance for all the performance parameters while maintain a trade-off with computational time.

From the comparative analysis, it is found that Vgg and AlexNet CNN models also provide phenomenal performance for DR classification but the huge requirements of complex computations make the model deployment incompetent both in terms of memory as well as computation time. However, Inception module approximates the redundant CNN activations with a sparse dense construction reducing the computation requirements. Inception model replaces the fully connected layer from the convolutional base with the global average pooling after the last convolutional layer. This modification decreases the network parameters achieving better network performance at much faster rate than Vgg and AlexNet pre-trained CNN models.
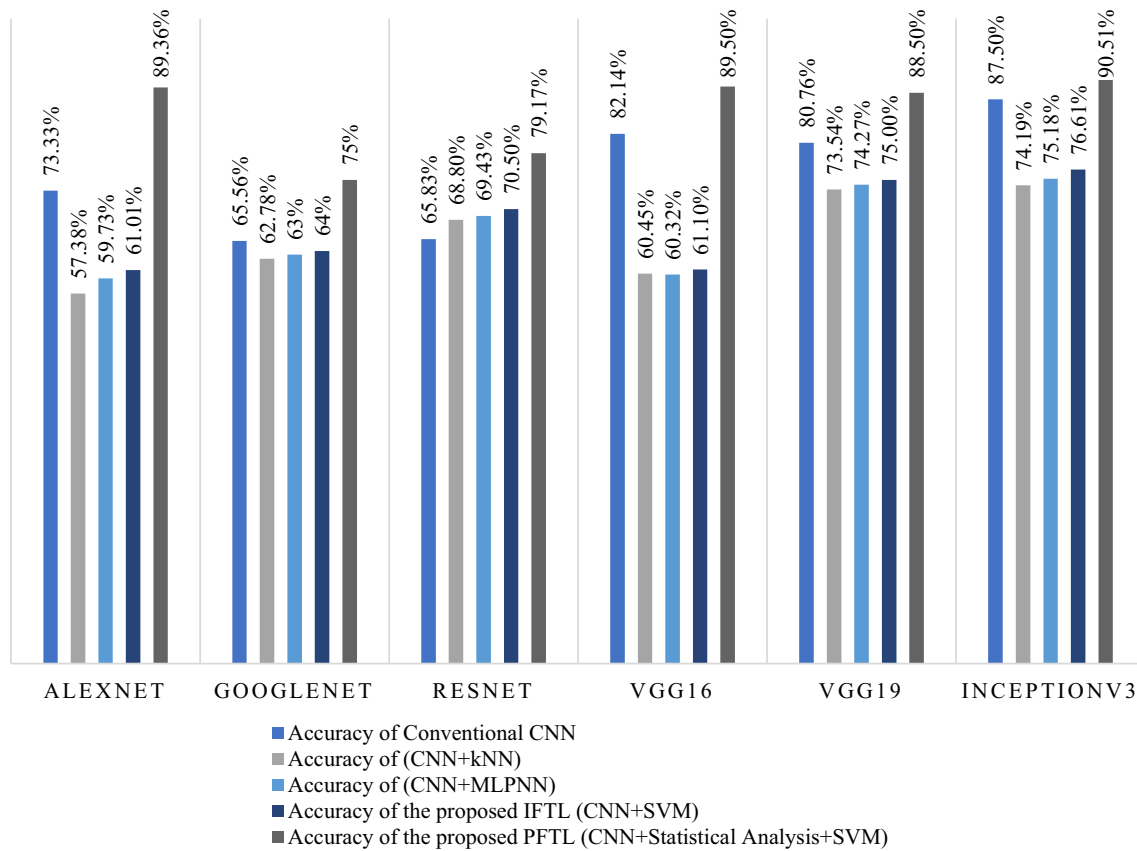
Fig. 7 Comparison of different classification approaches

3.2 Cost analysis for the proposed models

The cost analysis in terms of cross-entropy loss and computational time is formulated in Table 4. It provides the proportion of correctly predicted test data utilizing the trained model and optimizes the model during training. The cost analysis for different CNN model while employing the proposed PFTL model is depicted in Table 8.

The objective of minimizing the cost function is to make the IFTL and PFTL model output nearly close to the desired output values. During the model training, the network weights are iteratively adjusted in order to minimize the cross-entropy loss. Table 8 reveals that the minimized cross-entropy loss of 0.295 is obtained for PFTL model at the final iteration of the network learning which consumes 38 min and 53 s of the computational time. However, for the IFTL model, the same Inception V3 network provides the minimum cross-entropy loss of 0.559 consuming the time of 40 min 51 s, thereby maintaining a trade-off. Among IFTL (CNN + SVM) and PFTL (CNN + Statistical Analysis + SVM) models, PFTL provides the minimized cost function for the Inception V3 CNN model.

3.3 Performance validation with the latest IDRiD dataset

The generalization ability of the proposed IFTL (CNN + SVM) and PFTL (CNN + Statistical Analysis + SVM) models has been validated by the utilization of latest IDRiD dataset. To get an oversight of classification performance in terms of various evaluation parameters, the proposed methodology has been applied on IDRiD dataset and the observations are compared for both proposed IFTL (CNN + SVM) and PFTL (CNN + Statistical Analysis + SVM) models. The classification outcomes in terms of various evaluation parameters utilizing the proposed IFTL and PFTL models for IDRiD dataset are tabulated in Table 9.

The performance metrics evaluated in Table 9 reveal the feasibility of the proposed PFTL model for the IDRiD dataset providing higher values for all the evaluation parameters for Inception V3 CNN network. The proposed system provides 76.36% accuracy and 0.73 AUC for IFTL model (CNN + SVM) while utilizing the Inception V3 CNN architecture which is improved to 90.01% with 0.91 AUC for PFTL (CNN + Statistical Analysis + SVM)

Table 8 Cost analysis for different CNN model while employing the proposed IFTL and PFTL models

Different CNN models	IFTL model		PFTL model	
	Cross-entropy loss	Computational time	CROSS-ENTROPY LOSS	Computational time
AlexNet	0.855	00:18:10	0.675	00:17:05
GoogleNet	0.826	00:24:58	0.686	00:23:53
ResNet	0.768	00:29:45	0.684	00:27:48
Vgg16	0.713	00:17:21	0.638	00:15:19
Vgg19	0.645	00:18:47	0.647	00:16:45
Inception V3	0.559	00:40:51	0.295	00:38:53

model. In this work, the PFTL approach has made the preeminent attempts to address DR severity problem and generalize the grading ability of the proposed method achieving better performance. The performance of IDRiD dataset follows the same trend as obtained for the MESSIDOR dataset which justifies that the proposed models provide uniform results, irrespective of the dataset being used, which in-turn establishes its generalization abilities. The proposed PFTL model is compared with the other automated DR screening methods to obtain a generalized, feasible and robust solution for DR prediction and detection.

## 3.4 Comparative analysis with other state-of-the-art approaches

Traditionally, the main concern for DR diagnosis based research was machine learning based disease classification approaches and not much of the development was encountered in using CNN based classification methods. However, in the recent years, the research has been observed in application of deep learning based approaches for image classification problem.

Figure 8 depicts that comparative results obtained by the proposed PFTL scheme over state-of-the-art methods over a comprehensive MESSIDOR dataset. The maximum percentage improvement of 16.01% is accomplished by the

Table 9 Classification outcomes in terms of various evaluation parameters utilizing the proposed IFTL and PFTL models for IDRiD dataset

	Evaluation parameters	Different CNN models					
		AlexNet	GoogleNet	ResNet	Vgg16	Vgg19	InceptionV3
Image feature transfer learning approach (IFTL)	Sensitivity (Recall)	59.35%	63.00%	70.81%	62.56%	76.09%	78.43%
	Specificity	60.71%	63.00%	70.16%	61.54%	76.09%	70.48%
	PPV (Precision)	63.00%	63.00%	73.00%	59.07%	76.09%	78.08%
	NPV	75.57%	75.63%	83.43%	69.83%	86.93%	87.71%
	F-Measure	62.23%	64.08%	70.95%	59.72%	75.09%	76.67%
	Accuracy	61.78%	64.08%	70.51%	61.47%	75.09%	76.36%
	AUC	0.61	0.63	0.70	0.61	0.70	0.73
Prominent feature transfer learning approach (PFTL)	Sensitivity (Recall)	87.68%	76.02%	82.85%	89.69%	89.13%	89.42%
	Specificity	91.50%	76.02%	76.98%	89.34%	89.86%	90.78%
	PPV (Precision)	87.68%	76.02%	76.05%	90.00%	91.00%	91.95%
	NPV	91.35%	80.67%	78.58%	91.27%	92.24%	91.42%
	F-Measure	87.68%	76.02%	78.57%	89.64%	89.56%	89.81%
	Accuracy	89.35%	76.02%	79.46%	88.56%	88.90%	90.01%
	AUC	0.89	0.72	0.80	0.89	0.89	0.91

Fig. 8 Comparison of existing CNN based techniques with the proposed PFTL model
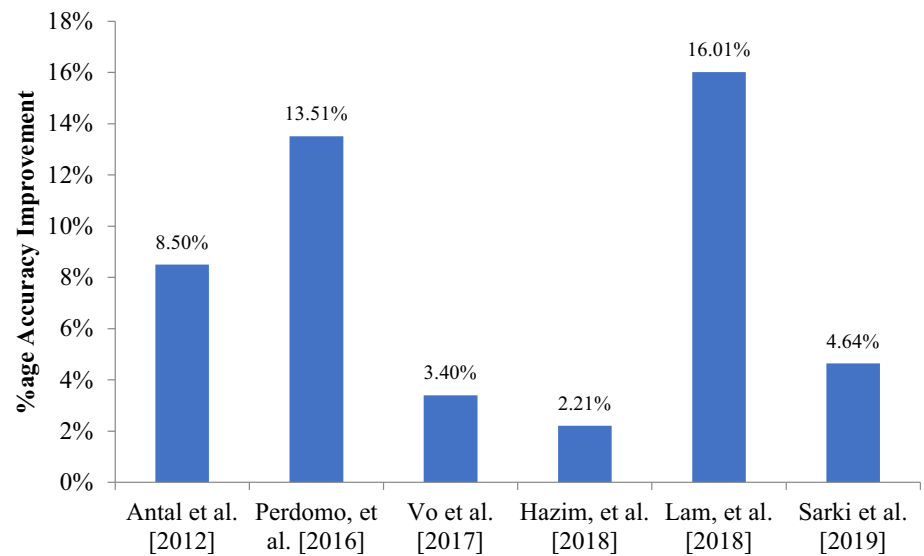
proposed PFTL method over the state-of-the-art technique provided by Lam et al. [49] and minimum 2.21% improvement in accuracy is seen over the technique proposed by Hazim et al. [50]. The comparison with other state-of-the art methods provides accuracy improvement of 8.50%, 13.51%, 3.40% and 4.64% with Antal et al. [51], Perdomo et al. [52], Vo et al. [53] and Sarki et al. [54], respectively. Unlike other state-of-the-art methods, in this article, we have proposed a transfer learning based PFTL approach which uses prominent features reduced after statistical selection for SVM machine learning based DR severity grade classification. Using the proposed PFTL approach for real time DR classification scenario will be the key emphasis of research in the future perspective of this work.

4 Conclusion

The exponential rise in the number of diabetic patients has made DR diagnosis challenging, thus, the practical implementation of DR diagnostic system requires a reliable DR severity grading system to fit into the real-world applications. A transfer learning based DR detection system utilizing CNN architectures for feature extraction and supervised SVM classifier for classification is proposed in this work. The PFTL approach utilizing statistical feature reduction method performs better achieving the maximum accuracy of 90.51% utilizing IncpetionV3 model for MESSIDOR dataset. The cost analysis in terms of cross-entropy loss and computational time reveals that PFTL (CNN + Statistical Analysis + SVM) model provides the minimum cross-entropy loss of 0.295 consuming the time of 38 min 53 s, thereby maintaining a trade-off. A latest

dataset is utilized for establishing the generalization capabilities of the proposed model and it is seen that IDRiD dataset follows the same trend and yields maximum accuracy of 90.01% for Inception V3 network providing uniform outcomes for all the evaluation parameters, irrespective of the dataset being utilized. The experimental results obtained demonstrate the competence of the proposed transfer learning based system which provides quicker and effectual DR analysis. The maximum accuracy improvement of 16.1% is attained for the optimized PFTL model (CNN + Statistical Analysis + SVM) comparative to the state-of-the-art techniques utilizing MESSIDOR dataset. The diagnostic capability of the state-of-the-art method is improved utilizing the proposed PFTL model, as indicated by the comparative analysis. The future perspective of this research work will deal with the model testing and deployment over the more generalized datasets in order to attain improved performance. This research work addresses the grading of non-proliferative DR categories; mild, moderate and severe stages. For future studies, the model generalization will be tested on other derivatives of the diabetic datasets as well as other DR disease's derivatives to improve the diagnostic ability of the learned models. The future initiatives will be focused on abnormal growth of retinal blood vessels like vitreous hemorrhages and retinal detachment.

Declarations

Conflict of interest The authors declare that there is no financial and personal conflict of Interest statement with any person or

organizations that could inappropriately influence (bias) our work. This article does not contain any human or animals participation performed by any of the authors.

References

1. Fong DS, Aiello L, Gardner TW, King GL, Blankenship G, Cavallerano JD, Ferris FL, Klein R (2004) Retinopathy in diabetes. Diabetes Care. 27(1):s84-7
2. Bhardwaj C, Jain S, Sood M (2019) Computer aided hierarchal lesion classification for diabetic retinopathy abnormalities. Int J Recent Technol Eng 8(1):2880–2887
3. Yen GG, Leong WF (2008) A sorting system for hierarchical grading of diabetic fundus images: a preliminary study. IEEE Trans Inf Technol Biomed 12(1):118–130
4. Kumar D, Taylor GW, Wong A (2019) Discovery radiomics with CLEAR-DR: Interpretable computer aided diagnosis of diabetic retinopathy. IEEE Access 7:25891–25896
5. ElTanboly A, Ismail M, Shalaby A, Switala A, El-Baz A, Schaal S et al (2017) A computer-aided diagnostic system for detecting diabetic retinopathy in optical coherence tomography images. Med Phys 44(3):914–923
6. Tariq A, Akram MU, Shaukat A, Khan SA (2013) Automated detection and grading of diabetic maculopathy in digital retinal images. J Digit Imaging 26(4):803–812
7. Poddar S, Jha BK, Chakraborty C (2011) Quantitative clinical marker extraction from colour fundus images for non-proliferative diabetic retinopathy grading. In: 2011 international conference on image information processing. IEEE, pp 1–6.
8. Fleming AD, Philip S, Goatman KA, Williams GJ, Olson JA, Sharp PF (2007) Automated detection of exudates for diabetic retinopathy screening. Phys Med Biol 52(24):7385
9. Abràmoff MD, Reinhardt JM, Russell SR, Folk JC, Mahajan VB, Niemeijer M, Quellec G (2010) Automated early detection of diabetic retinopathy. Ophthalmology 117(6):1147–1154
10. Gargeya R, Leng T (2017) Automated identification of diabetic retinopathy using deep learning. Ophthalmology 124(7):962–969
11. Quellec G, Charrière K, Boudi Y, Cochener B, Lamard M (2017) Deep image mining for diabetic retinopathy screening. Med Image Anal 39:178–193
12. Gulshan V, Peng L, Coram M, Stumpe MC, Wu D, Narayanaswamy A et al (2016) Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA 316(22):2402–2410
13. Prasad DK, Vibha L, Venugopal KR (2015) Early detection of diabetic retinopathy from digital retinal fundus images. In: 2015 IEEE recent advances in intelligent computational systems (RAICS). IEEE, pp. 240–245
14. Du N, Li Y (2013) Automated identification of diabetic retinopathy stages using support vector machine. In: Proceedings of the 32nd Chinese control conference. IEEE, pp. 3882–3886
15. Gurudath N, Celenk M, Riley HB (2014) Machine learning identification of diabetic retinopathy from fundus images. In: 2014 IEEE signal processing in medicine and biology symposium (SPMB). IEEE, pp. 1–7
16. Cao W, Shan J, Czarnek N, Li L (2017) Microaneurysm detection in fundus images using small image patches and machine learning methods. In: 2017 IEEE international conference on bioinformatics and biomedicine (BIBM). IEEE, pp 325–331
17. Nanni L, Ghidoni S, Brahnam S (2017) Handcrafted vs. non-handcrafted features for computer vision classification. Pattern Recogn 71:158–172
18. Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M et al (2017) A survey on deep learning in medical image analysis. Med Image Anal 42:60–88
19. Alban M, Gilligan T (2016) Automated detection of diabetic retinopathy using fluorescein angiography photographs. Report of standford education
20. Pratt H, Coenen F, Broadbent DM, Harding SP, Zheng Y (2016) Convolutional neural networks for diabetic retinopathy. Procedia Comput Sci 90:200–205
21. Rahim SS, Jayne C, Palade V, Shuttleworth J (2016) Automatic detection of microaneurysms in colour fundus images for diabetic retinopathy screening. Neural Comput Appl 27(5):1149–1164
22. Mansour RF (2018) Deep-learning-based automatic computer-aided diagnosis system for diabetic retinopathy. Biomed Eng Lett 8(1):41–57
23. Mohammedhasan M, Uğuz H (2020) A new early stage diabetic retinopathy diagnosis model using deep convolutional neural networks and principal component analysis. Traitement du Signal 37(5):711–722
24. Bhardwaj C, Jain S, Sood M (2020) Two-tier grading system for npdr severities of diabetic retinopathy in retinal fundus images. Recent Patents Eng 14(1):195–206
25. Decencière E, Zhang X, Cazuguel G, Lay B, Cochener B, Trone C, Charton B (2014) Feedback on a publicly distributed image database: the Messidor database. Image Anal Stereol 33(3):231–234
26. Porwal P, Pachade S, Kamble R, Kokare M, Deshmukh G, Sahasrabuddhe V, Meriaudeau F (2018) Indian diabetic retinopathy image dataset (idrid): A database for diabetic retinopathy screening research. Data 3(3):25
27. Bhardwaj C, Jain S, Sood M (2020) Hierarchical severity grade classification of non-proliferative diabetic retinopathy. J Ambient Intell Humanized Comput 12:2649–2670
28. LeCun Y, Boser BE, Denker JS, Henderson D, Howard RE, Hubbard WE, Jackel LD (1990) Handwritten digit recognition with a back-propagation network. In: Advances in neural information processing systems, pp 396–404
29. Xu K, Feng D, Mi H (2017) Deep convolutional neural network-based early automated detection of diabetic retinopathy using fundus image. Molecules 22(12):2054
30. Mohammadian S, Karsaz A, Roshan YM (2017) Comparative study of fine-tuning of pre-trained convolutional neural networks for diabetic retinopathy screening. In: 2017 24th national and 2nd international Iranian conference on biomedical engineering (ICBME). IEEE, pp. 1–6
31. Krizhevsky A, Sutskever I, Hinton GE (2012) Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems, pp 1097–1105
32. Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, et al (2015) Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 1–9
33. Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv409.1556
34. He K, Zhang X, Ren S, Sun J (2016) Deep residual learning for image recognition. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 770–778
35. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z (2016) Rethinking the inception architecture for computer vision. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2818–2826
36. Gao Z, Li J, Guo J, Chen Y, Yi Z, Zhong J (2018) Diagnosis of diabetic retinopathy using deep neural networks. IEEE Access 7:3360–3370

37. Bhardwaj C, Jain S, Sood M (2018) Appraisal of pre-processing techniques for automated detection of diabetic retinopathy. In: 2018 fifth international conference on parallel, distributed and grid computing (PDGC). IEEE, pp. 734–739

38. Bhardwaj C, Jain S, Sood M (2019) Automatic blood vessel extraction of fundus images employing fuzzy approach. Indones J Electr Eng Inf 7(4):757–771

39. Janney B, Meera G, Shankar GU, Divakaran S, Abraham S (2015) Detection and classification of exudates in retinal image using image processing techniques. J Chem Pharm Sci 8:541–546

40. Akyol K, Şen B, Bayır Ş (2016) Automatic detection of optic disc in retinal image by using keypoint detection, texture analysis, and visual dictionary techniques. Comput Math Methods Med

41. Saha R, Chowdhury AR, Banerjee S, Chatterjee T (2018) Detection of retinal abnormalities using machine learning methodologies. Neural Netw World 28(5):457–471

42. Roychowdhury S, Koozekanani DD, Parhi KK (2013) DREAM: diabetic retinopathy analysis using machine learning. IEEE J Biomed Health Inform 18(5):1717–1728

43. Sharma S, Jain S, Bhusri S (2017) Two class classification of breast lesions using statistical and transform domain features. J Global Pharma Technol 26(33):18–24

44. Tang Y (2013) Deep learning using linear support vector machines. arXiv preprint arXiv1306.0239

45. Qomariah DUN, Tjandrasa H, Fatichah C (2019) Classification of diabetic retinopathy and normal retinal images using CNN and SVM. In: 2019 12th international conference on information & communication technology and system (ICTS). IEEE, pp 152–157

46. Basly H, Ouarda W, Sayadi FE, Ouni B, Alimi AM (2020) CNN-SVM learning approach based human activity recognition. In: International conference on image and signal processing. Springer, Cham, pp 271–281

47. Durakovic B (2017) Design of experiments application, concepts, examples: State of the art. Period Eng Nat Sci 5(3):421–439

48. Sood M (2017) Performance analysis of classifiers for seizure diagnosis for single channel EEG data. Biomed Pharmacol J 10(2):795–803

49. Lam C, Yi D, Guo M, Lindsey T (2018) Automated detection of diabetic retinopathy using deep learning. AMIA Summits Transl Sci Proc 2018:147

50. Hazim JM, Hassan HA, Yassin AIM, Tahir NM, Zabidi A, Rizman ZI, Baharom R, Wah NA (2018) Early detection of diabetic retinopathy by using deep learning neural network. Int J Eng Technol 7(441):1997–2004

51. Antal B, Hajdu A (2012) An ensemble-based system for microaneurysm detection and diabetic retinopathy grading. IEEE Trans Biomed Eng 59(6):1720–1726

52. Perdomo O, Otalora S, Rodríguez F, Arevalo J, González FA (2016) A novel machine learning model based on exudate localization to detect diabetic macular edema. In: Proceedings of the ophthalmic medical image analysis international workshop, pp 137–144

53. Vo HH, Verma A (2016) New deep neural nets for fine-grained diabetic retinopathy recognition on hybrid color space. In: 2016 IEEE international symposium on multimedia (ISM). IEEE, pp 209–215

54. Sarki R, Michalska S, Ahmed K, Wang H, Zhang Y (2019) Convolutional neural networks for mild diabetic retinopathy detection: an experimental study. *bioRxiv*: 763136