Nanotechnology Perceptions
ISSN 1660-6795
www.nano-ntp.com

Machine Learning To Deep Learning For Detection Of Diabetic Retinopathy

Harpreet Kaur[1] [0009-0006-7527-8754], Maninder Singh[2] [0009-0008-7752-9160] and Sunil Kumar[3]

[1] *Department of Computer Science & IT, Apeejay College of Fine Arts, Jalandhar.*
[2] *Department of Computer Science, Punjabi University, Patiala.*
[3] *Directorate of Livestock Farms, Guru Angad Dev Veterinary and Animal Sciences University, Ludhiana.*

Diabetic retinopathy (DR) is the term for the formation of malformed blood channels in the retina; complications can result in significant loss of vision. Early detection of DR and lowering the incidence of disease succession can prevent future visual loss. Consequently, in order to detect the illness without such limitations, a computerized method is required which is more accurate, impartial, and dependable than professional diagnosis. Convolution Neural Networks (CNN), in particular, are the fifth generation of deep neural networks that have been harnessed to solve computer vision issues across multiple domains and have produced the highest accuracy in the stratification of diabetic retinopathy. Medical image analysis employing deep learning and machine learning techniques for diagnosing DR are providing accurate solutions. The intended study will illustrate the benefits of using deep learning analysis to the DR diagnosis problem. This paper reviews the several stages of diabetic retinopathy and emphasizes the need of early detection. The methods, difficulties, and shortcomings of classical Machine Learning (ML) are then emphasized. Next, an review of the terms and methods utilized while developing convolutional neural networks (CNNs) and the current applications of deep neural networks, particularly CNNs, is given. Finally, results about the deep learning technique's categorization of diabetic retinopathy are explored. The suggested work aims to investigate several approaches for the grading of images according to diabetic retinopathy.

Keywords: Diabetic Retinopathy, Machine Learning, Deep Learning, Classification, CNN, Filters.

1. Introduction

Diabetic-Retinopathy (DR) is the most constantly occurring complication of diabetes mellitus and remains a leading cause of vision loss. Approximately 3 million people ages 40 years or older have Vision-Threatening Diabetic Retinopathy(VTDR) in India [1] and major cause of new blindness person aged 25-74 years in the United States [2]. A recent survey by United Health Ministry has divulge that diabetic retinopathy was present in almost 17% of the Indian population with 3.6% of it being sight-threatening [3]. Specific to the rural population, a Shankar Nethralaya DR study identified that 10.4% had diabetes and around

10.3% had diabetic retinopathy [4]. By 2030, According to WHO predictions, the number of cases with diabetic retinopathy will rise by 50% [3]. The rising population of diabetics will cause retinopathy in those who already have it.

National Eye Institute categorizes DR in four stages:

1. **Mild Diabetic Retinopathy**: When condition first manifests, little balloon-like bulge in the retina known as microaneurysms—tiny blood vessels—occur. The retina may see fluid leakage from these microaneurysms. We call this capillary leakage.

2. **Moderate Diabetic Retinopathy:** The arteries of the retina may enlarge and warp as the condition worsens. They might also stop being able to move blood. Both disorders result in distinct alterations to the retina's appearance, may exacerbate macula enlargement, and also induce blood to leak from farther away capillaries.

3. **Severe Diabetic Retinopathy**: Several additional blood arteries are obstructed, preventing blood flow to certain retinal regions known as “capillary non-perfusion”. Growth factors re-leased from these areas instruct the neural layer of the eye to produce new blood- vessels.

4. **Proliferative Diabetic Retinopathy**: New blood vessels begin to proliferate as a consequenceof the capillary non-perfusion, growing into the vitreous gel that fills the optic, additionally along the inside membrane. Because the new blood capillaries are substational, capillary leak-age is more likely to occur. Retinal

Detachment probably caused from the accompanying scar tissue contracting. Permanent blindness might be attributed from retinal detachment [5]. Fig 1.1 shows the different stages of diabetic retinopathy.

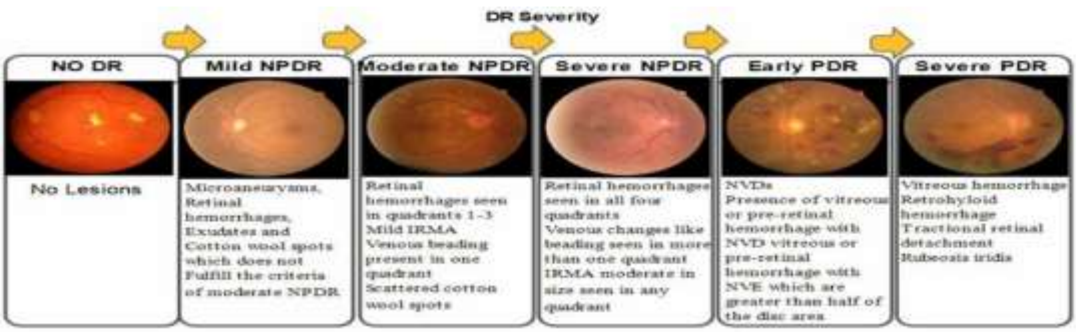
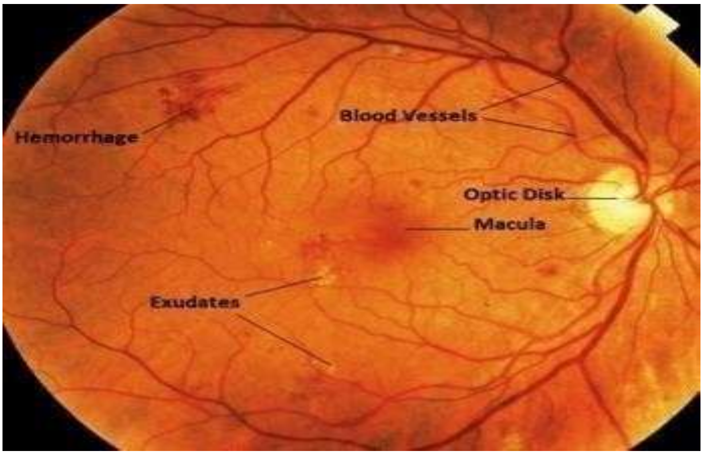


Fig 1.1: Stages of Diabetic Retinopathy

The technology that may stimulate the hasty screening process for this disease and lessen the need for expert personnel would likely be advantageous for both the patients and ophthalmologists. In computer vision research, artificial intelligence (AI) has emerged as a key frontier. AI refers to the ability of computers to complete tasks mostly without the need for human involvement. Similar to how people learn, artificial intelligence (AI) systems must be exposed to fundus images in order to grade diabetic retinopathy. Fundus images are visual records which document the current ophthalmoscopic appearance of a patients retina.

According to NCBI (National Center for Biotechnology Information) Retinal fundus image can be taken of the eye with fundus camera [6]. In addition to that, fundus photography is important for diagnosing and treating various posterior segments and other ocular disease. According to Caroline R., the retina can be seen in color or without red in fundus photography [7] and is mostly utilized for clinical studies. Deep learning convolutional neural network architecture-based computerized diagnostic tools are suggested in order to



detect the severity of the disease and learn DR patterns by image classification from fundus images. Fig 1.2 shows the retina fundus image.

Fig1.2 Retina Fundus Image

Traditionally, DR is best diagnosed with comprehensive dilated eye examination, in which doctors diagnose it manually. For any automated diagnosis method to be adopted in a widespread manner its operation must be equivalent or better to any manual process. Otherwise, even though the method may be justified from the economic, time efficiency perspective and accuracy of diagnosis. Many studies employing ML and DL, as well as ongoing research is to classify different forms of diabetic retinopathy. According to certain research, Deep convolutional neu-ral networks perform better at classifying images than people do. In order to achieve this, fundus and optical coherence tomography pictures are employed to capture the retina and the features can be mined from the images using ML and DL techniques, there are numerous examples of DL being implemented correctly and efficiently in the context of image classification at the right time. Through DL, a computer model has the ability to execute classification tasks directly from images, sometimes out performing human performance with state-of-the-art accuracy. DL models are typically referred to as deep neural networks because most DL models are skilled by utilizing labeled data and neural network architectures that learn features directly from the data without the requirement for manual feature extraction. In general, the phrase "deep" describes how many deep layers a neural network has. While deep networks can have upto150 layers, traditional neural networks consist of merely two or three hidden layers. Convolutional Neural Networks (CNN) has consequent swiftly surfaced as the preferred technique for interpretation of medical images, primarily for medical image categorization. Typically, medical datasets consist of hundreds or thousands of images instead of millions, making them smaller than ordinary computer vision datasets. CNNs discover highly abstract characteristics in training images by using labeled data, and these features are frequently utilized to create a classification model.

Table1, summarizes the different techniques of deep learning and machine learning, training and testing dataset description that contains number of images used, performance evaluation and their future directions explains about their future scope.

Table 1. Various techniques used by different author for the classification and detection of diabetic retinopathy

Reference	Network	Training Dataset	Testing Dataset	Performance Evaluation	Future Directions
Nayak et al.[11]	Proposed computer-based sys- tem to identify normal,NP D R and PDR	140fundus im ages		Accuracy:93% Senstivity:90% Specificity:100%.	Accuracy can be improved using more input features, diverse im- ages and good environmental illu- mination condition.
Gardner et al. [12]	Artificial Neural Network	147 dia- Betic and 32 normal images	200dia- Betic and 101normal images	Detection rates Exudates: 93.1% Haemorrhages73.8 %.	

Adarsh et al. [18]	Featureex-tractionand Sup-port VectorMa-chine (SVM)cla-ssifier	Twodata-bases DI-ARETDB 1 :28colorfu ndu s im-agesDI-ARETB0: 42colorfun du s im-ages	Two data-bases DI-ARETDB1 :61 color fundusim-ages DIARETB 0 : 88 color fundus im-age	DIARETDB0 Accuracy: 96% Specificity:94.3% Senstivity:90% DIARETDB1: Accuracy: 94.6% Specificity: 93% Senstivity:91.2. AUC:0.96	
Acharya et al. [19]	Feature extraction us-inghigh orderspec-tramethod SVMclas-sifier	200 fundus images	100 fundus images	Accuracy:82% Senstivity:82% Specificity:88%	
Gulshan et al. [22]	DeepCon-volutional Neural Network	EyePACS dataset:99 63 im-ages Messidor-2dataset:		EyePACS dataset Senstivity:90.3% Specificity:98.1% Messidor-2dataset	To determine the feasibility of ap- plying this algorithm in the clinical setting and to determine weather use of the algorithm could lead to improved care and outcomes com- pared with current ophthalmo-logic.
		1748		Sentivity:87.0% Specificity:98.5%.	

GargeyaR. et al. [23]	DeepCon - volutional NeuralNe t work	75137fun - dusimage s	MESSI- DOR 2 andE- OPH- THAdata- bases.	AUC0.97 Sentivity:94% Specificity: 98%.Testingdatas et Messidor2AUC0.9 4 E-OphthaAUC 0.95	
Sridhar et al. [24]	CNN- basedmo del	Kaggleda - taset 100 images	100 images	Training dataset Accu- racy:86% Testing DatasetAccurac y:85%	Using superior calculations for im- age processing so that images canbe arranged better than the present classifier. Utilization of other available algo- rithms for pattern extraction and classification to strengthen the accuracy.
Luo et al.[25]	Proposed Novel DCNN Framewo rk using multi- view Fundusi m- ages	15,468im - ages		MVDRNet Accuracy60.23% MVDR Net_SKAccu- racy77.75% MVDRNet_SEAc cu- racy76.60%	We intend to train our network intended with lesion annotations in order to upgrade the performance of multi-view DR detection.
Qureshi et al. [26]	Multi- layer archi- tecture of Activedee p learn- ing convo- lutional neural net- work(AD L -CNN)	35000 im- Agesfrom EyePACS dataset	54000im- agesfrom EyeP ACS datasetco mpar edwith thestate- of- artmethod s	Accuracy: 98.0%Specificity: 95.10 %Sentivity:92.2% F-measure:93.0% ACC98%	TrytoextendourADL- CNNmulti- layer architecture to various multi media applications such as data mining, data exchange, video tracking and image dehaze.

Khan et al.[31]	Various deeplearn- ing net- works like VGG-net, Incep- tion V3 and ResNetwit h trans- fer learn- ing	3700 im- ages		InceptionV3hadbest results the Accuracy: 81.2% on training phase. Accuracy:79.4% intesting phase.	Performance can be improved by training the model with more data and advanced algorithms.
Andriman et al.	Local bi-	5592 im-		Accuracy	Willconcentrateonassessingal-
[34]	Nary pat- terns(LBP) to extracttex- tual fea- tures DL techniques ResNet,Det Net,VGG16 andDense- Net werescrutinize detection andclassi- ficationof DR.	Ages from APTOS2019 Blindness detection dataset		byResNet34:96.25 % DetNet59: 93.99%VGG16:76.21% DenseNet121: 84.05% respectively.	ternative LBP iterations, including circular-LB Porother enhanced methodologies, and plan to metic- ulously assess each network's per- formance using arrange of hy- perparameter settings and adjust- ments.

Kele Xu et al. [35]	Deepcon- volutional neuralnet- work methodol- ogy	Kaggle Commu -nity		Accuracy by different methods GBM (Gradi- ent Boosting Trees- based) Hard exudates+GBM:89.4 %. Redlesions+GBM:88 .7% Micro-aneu- rysms+GBM:86.2% Bloodvesselsdetec- tion + GBM:79.1% CNN withoutdata augmentation:91.5% CNN with data augmentation: 94.5%	
Lam et al.[38]	Usescon- volutional neural networks. Transfer Learning on pre- rained GoogleNet and AlexNet models from ImageNet			Senstivity:95% Accuracies :74%to 2-ary 68.8%to3-ary 57.2% to 4-ary classi- fication models re- spectively.	
Ting et al.[39]	DeepLear- ing System(D L S)	71896 im- ages	14880 im- ages	AUCofDLSwasref- erable DR was 0.936 Senstivity:90.5% Specificity 91.6%. For glaucoma AUC was 0.942	To evaluate the applicability of the DL in health care settings and the utility of the DLS to improve vi- sion outcomes.

				Senstivity:100% Specificity:91.1%. For AMD AUC was 0.942 Senstivity: 95% Specificity:88.7%	
ManojSHet al. [53]	Transfer Learning withCNN architec- ture. Demon- strated in the APTOS20 1 9 Blindness Detection Competi- tion	Eyeclinic in Mah a- rashtra, India:516 images. Kaggle(EyeP ACs,2015) : 35,126im- ages.	KaggleAP TO S 2019 BlindnessD ete ction 13000	Achievedhighquad- ratic appascoreof0.92546. Accuracyfordifferent segmentations: Blood Vessel:0.9922 HardExudate:0.6663 Soft Exudate:0.9981 Haemorrhage:0.9958 Microaneurysms:0.9 967 OpticalDisc:0.9989	The research scope extends to rev- olutionizing DR detection method- ologies, integrating cutting-edge technologiesandcontributingto the wider field of medical imaging and automated diagnostics.

2. Traditional approaches for diabetic retinopathy classification

Within AI field, ML aims to educate machines to learn from data and improve with experience. Aiming to create the best judgments and predictions, it trains algorithms to find patterns and correlations on datasets. Using conventional/classical methods, machine learning can be used to recognize and categorize diabetic retinopathy from clinical images by completing the following steps:

2.1 Data acquisition

Preparing and cleaning data is the process of data gathering. One way to collect data is by the use of a computer vision system to take images of individuals having DR. Moreover, the data can be gathered from the range of additional sources, such as online scraping, fundus cameras, sensors, questionnaires, and many more. The data may be available in raw, CSV, or spreadsheet format. DR are initially categorized in this manner. After being trained upon the given data, ML algorithms can utilize a range of image processing techniques to either identify or categories the DR.

2.2 Image preprocessing

Image pre-processing is the process of improving the image and uses a range of methods and applications. For instance, Quantization and sampling are two pre-processing techniques. In this step, several types of noise, such as distortion, clothes, and hair, are eliminated from the images. Different strategies for removing noise are outlined as follows:

2.2.1 Noise removal

Duringtheimageacquisitionprocessusingfunduscameras,noisypixelsappearinthefundusimages. Inorderto remove noise from the images, there exist two distinct kinds of image noise suppression filters: time domain and frequency domain. The time domain includes the mean, median, Gaussian, lateral, and weiner transformswhereas the frequency domain includes the fourier and wavelet transforms. Moreover, the primary cause ofnoise in fundus images is similarly uneven illumination. Salt and pepper noise as well as gaussian noise are the main sources of artifacts in fundus imaging. Here's an explanation:

2.2.1.1 Mean Filter

This technique, which is also termed to as box filtering and averaging, is the straightest forward, understandable, and simple to use for smoothing images. This filtering technique is an example of linear filter. Its primary application is image noise reduction. It substitutes the aver-age "mean" pixel values of its neighbors for thecenter value between the windows and it lessens the degree of intensity variance that exists between consecutive pixels. The three main categories: of mean filters are harmonic, geometric, and arithmetic mean filters. Local fluctuation is smoothed and noise is decreased by blurring with arithmetic mean filter. While harmonic mean filters perform poorly against pepper noise but well against salt noise, geometric mean filters preserve moredetail in images [54].

2.2.1.2 Median Filter

In 1977, Tukey proposed the median filter. This filtering method is executed on signals and images to remove noise. This technique of filtering is an illustration of a nonlinear filter. Due to its well-known ability to preserve edges during noise reduction, this filter is extremely important when processing images. Since it reduces noise significantly with a less blurring than a linear smooth-ing filter of comparable size, this filter is the predominantly used one. The histogram technique is the foundation for median filters, which estimate the median value. When impulse noise, commonly known as salt and pepper noise, is present, median filters work especially well [54].The fundamental problem with the median filter is that it cannot handle various types of noise, such as Gaussian noise, well.

2.2.1.3 Gaussian Filter

Carl Friedrich Gauss was the one who first proposed the Gaussian filter. A Gaussian filter is a linear filter and is commonly used effecting graphics of software, usually for noise reduction. Noise produced by a normal distribution can be effectively eliminated with this filter. The linear filter blurs and eliminates noise using the Gaussian function. There are two varieties of Gaussian filters: high pass and low pass. High pass filters eliminate frequency noise and are used to sharpen images while low pass filters eliminate low frequency noise and smooth the image for eliminating salt and pepper noise.

2.2.1.4 Weiner Filter

Norbert Weiner proposed the Weiner filter in1940, and it was published in 1949. It can at times be recognized as the least square error or minimum mean square error filter. A Weiner filter is the MSE-optimal stationary linear filter. The Weiner filter reduces the predicted value of the squared error signal, which is its primary benefit when used for noise reduction. The restoration technique considers both statistical characteristic noise along with degradation function. More-over, filter processing an image has a high-quality output that lessens noise effects while pre-serving signal properties. Additionally, it lessens noise impacts and generates images of excel-lent quality. Simultaneously, it reverses the blurring and eliminates the additional noise. It performs superior to the inverse filter. Due of their need for frequency domain work, Weiner filters are relatively low to implement.

2.2.2 Segmentation

Oliver Brock and DOV Katz proposed the concept of image segmentation. After data collecting and pre- processing, segmentation comes next. Segmentation is key to success of future image processing tasks. Global and local segmentation are the two categories of image segmentation. While only a portion of the image is incorporated into consideration in local segmentation, the entire image is incorporated in global segmentation. The ultimate success or failure of automated analytic techniques is determined by the accuracy of segmentation. Clinical images are segmented so that each pixel is mapped to an object and the images are divided in to distinct segments. Accordingly, segmentation gives each pixel in an image a label; pixels with the same label share similar properties.

The majority of segmentation algorithms rely on two fundamental characteristics of intensity values that is discontinuity and similarity. The first method divides an image according to sharp contrasts, like edges. These could method divides an image in to sections based on predetermined criteria that are comparable to each other. The following is an explanation of these approaches:

2.2.2.1 Edge Detection

The edge detection operates by detecting discontinuities in luminance. The most wide spread use case for it is image segmentation based on localized intensity variations. Two varieties of edge detection operators exist, gradient based operator and Gaussian based operators. Gradient operator computes first-order derivative and Gaussian based operator computes second-order derivative.

First-order or second-order derivatives can be harnessed to recognize intensity variations in order to locate edges.

Intensity profiles are applied to categorize edge models. Common edge detection algorithms include Prewitt, Canny, Sobel, Roberts, and fuzzy logic methods. During edge identification, the following three essential processes are carried out: i) Image smoothing intended for noise reduction ii) Locating edge points iii) Location of edges.

2.2.2.2 Boundary Detection and Edge Linking

The aspiration of a fore mentioned is to identify regional borders using discontinuities in intensity levels. In practice, edge connection is approached from two basic perspectives: local and global-processing through transform. Local processing analyses the traits of pixels in a small neighborhood of 3*3 or 5*5 about every point that has endured edge detection and global processing can be done using HOG transform

2.2.2.3 Region based segmentation

Segmenting an image in to distinct sections is the aim of the process. The segmentation methods based on directly identifying the region are explained as follows:

2.2.2.4 Region-Based Growing

The workflow of "region growing" involves organizing pixels or subsections according to predetermined growth criteria into larger regions. The kind of picture data that is available influences the assortment of resemblance criteria besides to the problem that is being studied. To create a zone by assembling pixels of the same intensity.

2.2.2.5 Region splitting and merging

In order to satisfy the condition of segmentation, an image is subdivided into a set of arbitrary disjoint regions and then merges the regions. The splitting technique has a convenient representation in the form of so-calledquad trees, the trees in which each node has exactly four descend-ant’s also known as quad regions. If only partitioning is used, the ultimate partition contains adjacent regions with identical properties, so this drawback can be solved by merging as well as split-ting. Splitting is simple and merging is more complex [54].

2.2.3 Thresholding

Thresholding is a procedure that divides an image based on intensity levels into two or more sections so that desiredfeaturescanbeeasilydiscoveredandextracted.Usingthistechnique,grayscaleimagesareconvertedinto binary images in order to extract specific objects. An image that is binary has just two values for each of its pixels: 0 and 1. Essentially, thresholding operates under the premises that objects of interest encompass a rangeof intensity levels that differ from the back drop. Using thresholds based on, the distribution of pixel attributes, such as color or intensity levels, segmentation was achieved [54].The explanation of thresholding types is as follows:

2.2.3.1 Local thresholding

It is also referred as, regional thresholding or adaptive thresholding. If the value T at any point (x,y) within an image relies on the attributes of the neighborhood of (x,y) means neighboring properties is also taken into account. It creates binary image by segmenting a grayscale image into a particle region and a background region.

2.2.3.2 Global thresholding

When the amplitude distribution of objects along with background are adequately distinct then it is better to harness global thresholding. The foundational global thresholding algorithm iteratively detects the best threshold value for segmenting. Its main parameter that, it is derived entirely on computations implemented on the histogram of an image is Optimum Global Thresholding using OTSU method.

2.2.3.3 Multiple Thresholds

This approach is employed when it fragments a reason to believe that the problem can be solved effectively with two or more thresholds. It comprises of two categories for this approach: multi-level thresholds and bi-level thresholds. A single threshold is used in the bi-level technique to fragment an image into two groups, while two or more thresholds are used in the multi-level technique to divide an image in to several groups. This method aims to segment the image within three sections: the portions in shadow, the areas with light, and the dark background.

3 Feature extraction

A key component of ML aimed at image classification is feature extraction. The primary goal of feature extraction techniques is to extract features from big image data sets while preserving as much information as possible. The variety of techniques for feature extraction includes texture features, color features and shape features. These methods divide the features into high and low level categories. While low level features may automatically extract features from an image without form knowledge whereas top tier features are employed to detect objects and huge shapes in the image. In essence, it's a kind of data reduction method [42]. The following is an explanation of various techniques:

3.1 Color Feature Extraction

The most usual strategy for determining an image's color characteristic is to use the color histogram. The two categories of color histograms: local and global. The local color histogram considers the spatial distribution, whereas the global color histogram analysis copious statistical colour frequency distributions. Problems with translation, rotation, and angle view are also resolved by the global colour histogram. Other methods for collecting color information include histogram interaction, colour histogram for K-means, correlogram, color co- occurance matrix, chromaticity, dominant colour descriptor, and many more [43].

3.2 Texture Feature Extraction

Texture features can be identified to clean up the data by using the Grey Level Co-occurance

Matrix (GLCM), Grey Level Run-Length (GLRM), and Histogram of Oriented Gradients (HOG). These processes help to extract information from the data that is useful. On the other hand, GLCM is the most often used technique for figuring outanimage'suniquegreyleveldependency.Theentirenumberofgreylevelsinanimageisexactlyeq ualtothe count of rows and columns in GLCM. Many statistical metrics, such as contrast, mean, variance, stand-ard deviation, and many more, are evaluated in order to compute the GLCM features [44].

3.3 Shape Feature Extraction

The statistical characteristics of an image, such as its size, perimeter, and circularity, are designated to as shape features. It is essential for feature extraction in many different contexts, including shape alignment, registration, and approximation, as well as shape retrieval, identification, and classification. Many methods, such as contour- based and regions-based approaches, space do-main and transform domain methods, are adopted for shape description [44].

Nayak et al. [11] apply artificial neural networks to detect features such as area of the blood vessels, area of hard exudates and the contrast in order to perform a three-class classification of DR. The proposed model was validated on an extremely small datasets of50 images. This engaged extracting facts from fundus images and feeding those data into neural networks to classify the images into three categories: proliferative retinopathy, normal, and non-proliferative retinopathy. The results were confirmed by comparing the categorization results with ex-pert ophthalmologists' grades.

Artificial Neural Networks (ANN) were employed by Gardner et al. [12] to predict the existence of vessels, exudates, and haemorrhages within fundus images. The neural network was trained using patches of fundus images,allofwhichindividuallyrequired a manual grading to determine whether or not it contained a certain DR feature. Then, the patches on the test images were compared to see if a fundus image had disease-related characteristics. Nevertheless, 23 images were excluded from the analysis because an ophthalmologist had disagreed with the first evaluation. Additionally, it might be argued that the tests' small sample size of 100images is not indicative of the normal screening data set.

4 Classification

ClassificationisthelaststageoftheMLmodel,wheretheretrievedfeaturesareutilizedtocategorizet heimages. Diverse productivity criteria are used to evaluate the accuracy of different approaches, including confusion metrics, sensitivity, specificity, and accuracy. Support Vector Machines (SVM), K-Nearest Neighbour (K-NN), and Decision Trees are the three most popular classification techniques. They are explained as follows:

4.1 K-Nearest Neighbor (K-NN)

Cover and Hart introduced the concept of the K-nearest neighbor in 1968. K-NN has been applied to statistical problems. It adheres to the non-parametric methodology [45]. The implementation of this approach is quite simple. For example, if"C" has a K-nearest case where features space and the majority of them have the identical label "B," then "C" belongs to "B." The closest neighbor determines the class of the new case. The effectivenessoftheKNNmethodforlarge-scaleandhigh-dimensionaldatasetsisoneofitskeydrawbacks, despite its great strength and ease of implementation. Its slow learning is the chief cause of its disadvantage, and the outputs, have a significant computational cost [46]. Fig 1.3 clarifies the K-NN algorithm.

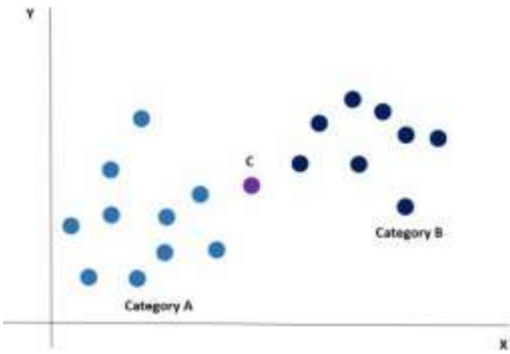


Fig 1.3 K-Nearest Neighbor (K-NN)

Numerous more methods of detecting microaneurysms and grading DR involving K-NN, SVM, and ensemble based methods have yielded sensitivities and specificities within the 90% range using feature extraction techniques and preprocessing procedures [15, 16, 17].

4.2 Support Vector Machine (SVM)

V. Vapnik and his colleagues invented SVM in 1985 at AT&T Bell Labs [48]. It has a strong foundation in mathematics. A different name for it is Maximum Margin classifiers. SVMs are a class of related labeled data learning methods used to address challenging problems in outlier identification, regression, and classification. The SVM classification yields more valid results when contrasted with other techniques. Furthermore, it functions well with high-dimensional data, including images [47]. SVM generates decision boundaries, or hyperplanes, for classification. Hyper planes are employed to split multiple classes, and their sizes change according on the number of features. Moreover, SVM is exceptional in that it may simultaneously, lessen the empirical classification error and maximise the geometric margin. Based on labeled data training, Kernal comes inadiversityof forms.SVMusesanumberofKernelfunctions,suchasthepolynomial,GaussianRBF,standard deviation, and sigmoid. The mitigation of structural risk is its corner stone [49]. One of the problems with this technique is how long it takes. Support Vector Machine is explained in the Fig 1.4.

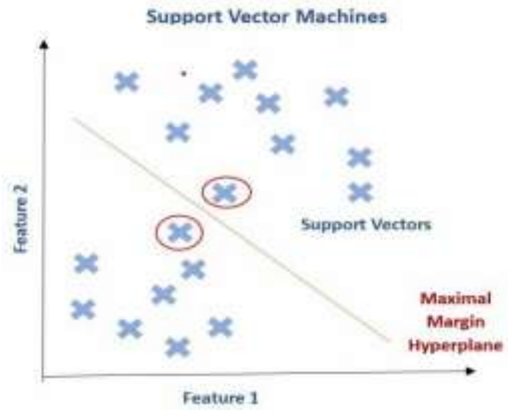


Fig 1.4 Support Vector Machine

Support vector machines (SVM) classifiers and feature extraction were employed by Adarsh et al. [18] to diagnose DR in five groups. The identification of textural characteristics, exudates, microaneurysms, and retinal blood vessels was conducted using "image processing techniques." The multi-class SVM feature vector was constructedusingthelesions'areaandtexturedata.With89and130images,respectively,inthepublic databases DIARETB0 and DIARETDB1, this obtained stated accuracies of 96% and 94.6% for disease classification. Adarsh verified that fundus imageries may be used to extract characteristics for the five class diagnosis problem. However, even a multi-class SVM was used and the datasets used were small so the robustness of the feature extraction algorithms for detection of disease are suspicious.

Acharya and colleagues [19] devised an automated technique to discuss the five categories of delayed release. Using the high spectrum approach, features were retrieved from the raw data, quantified, and then input

into an SVM classifier. There reports have been published that 82% accuracy over five classes, 82% sensitivity, and 88% specificity have been attained. 100 fundus images were used to assess the procedure. Additionally, the paper showed that DR char-acteristics could be echoed using computer vision techniques and proposed that a CNN model should be capable of doing so.

4.3 Naïve Bayes

The Bayes Theorem, put forward by Thomas Bayes in 1700, is the foundation associated withnaïve Bayes classifier. The following equation defines the Bayes Theorem:

Where, x is input variables/feature vector and y output variable. The early 1960s, text retrieval was one of its uses.This method, which finds the best-fitted classification for a given piece of data inside the issue domain, is based on a mathematical classification technique and uses a series of probabilistic computations. Applications for naive Bayes classifiers are many and include weather forecasting, classifying health conditions, and evaluating customer credit, among many others. These classifiers are utilized in ML models, which may be quickly and readily programmed to generate predictions in real time. Finding the best mapping between a new piece of data and a group of classifications is the key objective. The main flaw of Naïve Bayes is that it cannot handle issues with zero frequency because it assigns zero probability to categorical variables in training datasets. If categorical variables of a category that was missing from the training and data set reveal in the test data set, the Naive Bayes model accredit zero probability and will not be applicable [50]. Fig 1.5 clarify about Naive Bayes theorem.

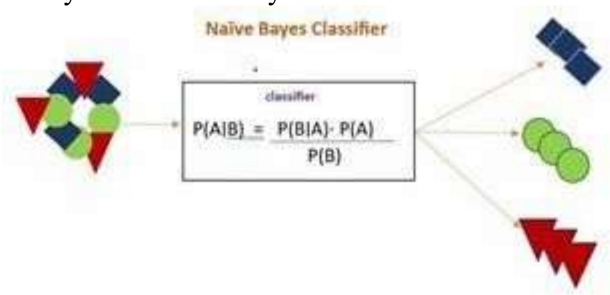


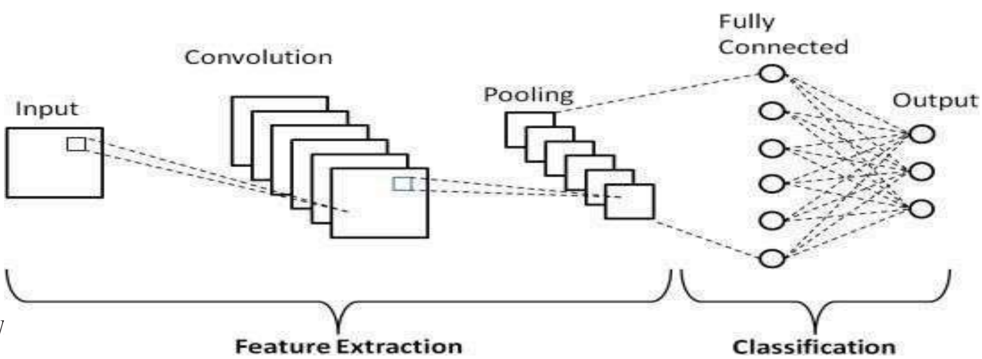
Fig: 1.5 Naïve Bayes Classifier

5. Deep Learning approaches for DR classification

Over recent years, DL techniques are a subclass of ANN that have numerous hidden layers and have been used to computer vision problems. Applications of DL are Deep neural networks include, convolutional neural networks (CNNs), which just need raw input images. CNN developed a single unit that encompassed all stages of customary ML. When utilised for object classification and recognition tasks that need little or no preprocessing, little processing and segmentation, automatically retrieved features, strong CPUs, and GPUs, deep neural networks are very useful for feature learning. In essence, CNN were developed to lessen the proportion of computing resources needed by a model to process the input.

5.1 Convolution Neural Network(CNN)

Around 1980, Yannle CUn, a post-doctoral computer science researcher at the University of Toronto, created and utilized CNN for the first time. Convolution, pooling, and fully connected layers are some of the many building elements that CNN uses to automatically learns patial hierarchy through backpropagation, from low level patterns of data to high level patterns [51].



N

Fig 1.6 Convolution Neural Network

While fully connected layers translate the retrieved features into final outputs like classification, convolutional and pooling layers carry out feature extraction. In CNN, a convolution layer is crucial which is based on mathematical processes and a particular kind of linear operation. CNN is incredibly effective at processing images in characteristics can appear any where in the image.

5.1.1. CNN Layers

The CNN building piece that is primarily in charge of computation is the convolution layer. The architecture of CNN is made up of several layers. The details of each layer in CNN architecture are provided below:

5.1.1.1 Convolution Layer

The essential component of the CNN architecture that manages feature extraction is the convolution layer. This system consists of convolution and activation functions, which combine linear and non-linear processes [51, 52]. In a convolution, the input is an array of components referred as a tensor, to which a small array of numbers known as a kernel is applied. First, the kernel traverses the whole image in both vertical and horizontal directions. The subsequent step consist computing the dot product between the kernel and the input image, multiplying and summing the results to calculate a single scalar value.

When two matrices are merged to create a third matrix, convolution takes place in CNNs. Convo-lutions like this are used to find information and filter input data. Every pixel in the convolution's receptive are transformed into a single value. Until there is no more sliding available, the whole process is repeated to build an arbitrary count of feature maps. The convolution, process reduces the height and width of the output feature map by stopping each kernel's centre from overlappingtheoutermostelementoftheinputtensor.Zeropaddingisappliedtotheinputtensortoad drowsandcolumns of zeros to either side in order to solve this problem.

Modern CNN architecture usually uses zero padding to preserve dimensions while applying more layers. In CNN models, the convolution layer is utilised to identify which kernels work best for a particular task based on a particular training technique. A kernel learns new information automatically. Before the training process starts, the following hyper parameters must be known: the number-of kernels, kernel size, padding, and stride. Following equation illustrates how the convolution layer's output is calculated:

Where P and S are the padding and stride respectively. N*N are size of image and K*K are size of kernel respectively.The count of feature map obtained depends upon the number of kernel. More benefits of convolution layer are sparse connectivity and weight sharing [51]. The ultimate result of the convolution layer is a vector.

Activation Function

An essential component of a neural network's architecture is its activation functions. To transform the feature map into an activation function, the result of the linear convolution procedure is passed through a nonlinear function, such as a sigmoid or hyperbolic tangent. The rectified linear unit (ReLU) is currently the most commonly utilized nonlinear function. Every work must be taken in to consideration when selecting an activation function. There are numerous varietiesofactivationfunctions,includingsigmoid,linearfunction,softmax,andmanymore.Thesoftmaxfunctionisan activationfunctionthatismadetoidentifytheactivationfunctionthatisutilizedforregressiontocontinuousvalues.Itis used in multiclass classification tasks to normalize the output real values from the last fully connected layer to the probability of the target class, where all values add up to one and the value ranges from 0 to 1 [51]. The binary and multiclass classification processes use the sigmoid activation function. The pooling layer, which comes after the activation function, can be utilised to further reduce the feature map's dimension while preserving its crucial data which is explained as follows:

5.1.1.2 Pooling Layer

Reducing the size of the feature map by means of a pooling layer speeds up computation by lowering the number of training parameters. In the CNN model, adding a pooling layer results in less overfitting, more efficiency, and quicker training periods. Three hyperparameters are present in the pooling layer because it lacks learnable parameters: filter size, stride,andpadding.Thepoolinglayerprovidesadownsamplingprocedurethatreducesthein-planedimensionalityofthe feature map. Max-pooling and average pooling were the most widely used pooling strategies.

The average sampling simply takes the average of all the components in each feature map while max-pooling with filter size 2*2 and stride of 2 down samples the feature map's in-plane dimension by a factor of 2. The main advantage of using average pooling is that it allows the CNN to accept inputs of different sizes and reduces the amount of learnable parameters..Convolution and pooling layer output are three-dimensional tensor feature maps.

5.1.1.3 Fully Connected Layer

When performing classification tasks, a subset of fully connected layers maps the features that were retrieved by convolution layers, downscaled by pooling layers, and ultimately mapped to the net-work's final outputs. It is a sort of feed-forwardartificialneuralnetworkthatessentiallyusesthesamemethodologyasatraditionalmultiplelayerperceptron neural networks. At the very end of every CNN architecture comes this Layer. Every completely connected layer's output is routed through an on-linear activation function.

Training a network

For the training a network back propagation method is used and involves the optimisation of gradient descent and loss function. After the network design is developed, the network needs to be trained on the label dataset. The objective function, sometimes known to as the cost function or loss function which generates a loss value by comparing the network prediction with the true labels. Cross entropy is the name of the loss function used in multiclass classification. In contrast, regressionusesmeansquarederrortopreservevalues.Themodel'sperformanceisascurtainedunder particularkerneland weight combinations using a loss function. Gradient descent is a popular optimization technique that iterativelyrecondition the learnable parameters and kernels and weights, to minimize network loss. The optimisation model and generalisation are thought to be the best. While optimisation refers to changing learnable parameters on a training dataset, generalization relates to the performanceof a learned model on a new, unknown dataset.

Deep learning models have two main issues: underfitting and overfitting. If at all feasible, increase the degree of training data to assist the model better understand the underlying patterns in the data and address underfitting. In addition, underfitting happens when the model is too straight forward to identify the fundamental patterns in the data. Moreover, other methods such as lengthening train-ing sessions are employed to lessen underfitting. Nonetheless, the issue of overfitting and under-fitting can be resolved by using CNN regularisation, which is explained as follows:

Regularization of CNN

Underfitting develops when a model is unable to learn complex patterns. This is can be resolved by selecting a more capable model. The over fitting model performs well on the validation data set in contrast to the training-set. The regularisation process is assisted by a number of initiative con-cepts to prevent both over-and under fitting.

a) **Drop weights**: Instead of deleting neurons, the connections between are dropped throughout every training epoch.

b) **Data augmentation**: This method applies random geometric transformations, such as scaling, rotation, translation, flipping, and so on, to enhance the amount of training data that is now available. The simplest method to prevent over- fitting of the model is to train it on a sizable amount of data.

c) **Batchnormalization**:Its distribution is unit-wise Gaussian. Moreover, itconsiderablycurtailsthe time imperativefor network convergence and aims to reduce training dependence across hyperparameters. Since it affects regularization more greatly, the major advantage is a lower chance of over-fitting [52]. Numerous improvements have been made to CNN to make it more effective at solving difficult problems. Even if CNNs operate well and serve the purpose, they are lessfrequentsincedeepneuralnetworksneedenormousamountsofdatatotrain,andobtainingbigdatasetsofimagesin the medical field is quite challenging. Currently, image classification uses very efficient CNN architectures.

Convolutional neural networks produce the most effective and efficient findings; prior CNN re-search on DR fundus images produced binary classifications or sensitivities and specificities in the 80–90% range. Historically, handwriting recognition was the initial use of CNNs using image data. Currently, one significant aspect of CNN is its ability to achieve "spatial l invariance". Moreover, it has the ability to identify and extract image features from data and photos, as well as finish extraction directly from images [20]. CNN is therefore an effective deep-learning tool for getting precise results [21].

Sridhar et al. [24] proposed the use of CNN architecture in a DR detection system. Retinal fundus images were categorized by a CNN-based model based on their features, indicating whether or not they had DR and how severe it was. An image dataset that is accessible to the general public on Kaggle was used to instruct the model. The model improved the accuracy of DR detection and produced remarkable results.

Luoetal.[25] suggested the integration of multi view fundus mages into a unique convolutional network for theautomatic identification of diabetic retinopathy (DR). The predicted technology makes use of retinal lesion features in contrast to single view DCNN-based DR detection methods that are presently in use. Moreover, for efficient feature extraction, give significant weights to the network's key channels. It included a multi-view DR dataset with 15,468 images in it. Compared to other bench marking techniques, this one is better.

6. Evaluation Metrices

A variety of measures are employed to evaluate the efficacy of varied algorithms, including those for segmentation, classification, and object detection. The four main ideas, which form the basis for the definition of the assessment metrics are true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). The true positives are the samples that are correctly identified as authentic, or samples that are true in reality. The real negatives are the cases that are classified as negative and are predicted to be negative. False positives represent cases that are categorized as negative, even if the network forecast was positive. False negatives represent cases that the network labels as positivebut projects as negative.

For classification tasks, accuracy, sensitivity, specificity, precision, and F1 score are the most often utilized metrics. Based on the accurate classification of each image's class, the classification accuracy is determined.

The following is a brief explanation and definition of the evaluation metrics equation:

Accuracy

It gives the overall accuracy of the model and is designed by dividing the total number of correctly categorized imageclasses by the total number of images.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Sensitivity

It gauges the model's ability to accurately forecast the real positive cases.

$$Senstivity = \frac{TP}{TP + FN}$$

Precision

Precision calculates the number of true positives divided by the digit of all positives.

$$Precision = \frac{TP}{TP + FN}$$

F1 Score

This matrix is ideal for imbalanced datasets as it integrates recall and precision in to a single matrix. It is the precision and recall harmonic mean.

$$F1\ score = 2 \times \left(\frac{Precision \times recall}{Precision + recall}\right)$$

7. Conclusion

An automated screening method can diagnose patients faster, saving ophthalmologist's time and money moreover, also improving patient outcomes. ML and DL are both utilized in the identification of diabetic retinopathy by different approaches. DL makes use of neural networks, specifically convolutional neural networks, which automatically learn features from data, in contrast to ML, which uses classic methods like support vector machines, decision trees, and many more which require human feature extraction and selection. Deep learning techniques, particularly CNN, can directly learn complicated patterns from the data, recent research indicates that they perform better on this job than classic ML techniques. We can let the machine choose the optimal features on its own by using deep learning to remove the need to extract features that were designed by humans. In future all these patterns for binary classification and multiclass classification will be employed by adopting basic DL models and developing hybrid model.

8. References

[1] https://www.thelancet.com/journals/langlo/article/PIIS2214-109X(22)00411-9/

[2] https://emedicine.medscape.com/article/1225122-overview?form=fpf

[3] https://pharmeasy.in/conditions/diabetes/diabetic-retinopathy-treatment-india/

[4] https://www.sankaranethralaya.org/

[5] https://www.nei.nih.gov/sites/default/files/health-pdfs/diabeticretino.pdf

[6] https://www.ncbi.nlm.nih.gov/books/NBK585111/

[7] https://www.sciencedirect.com/topics/medicine-and-dentistry/fundus-photography\

[8] https://towardsdatascience.com/using-convolutional-neural-network-for-image-classification.

[9] https://www.ibm.com/topics/convolutional-neural-networks.

[10] Philip, S., Fleming, A.D., Goatman, K.A., Fonseca, S., Mcnamee, P., Scotland, G.S., Prescott, G.J., Sharp, P.F., Olson, J.A.: The efficacy of automated 'disease/no-disease'

grading for diabetic retinopathy in a systematic screening programme. British Journal of Ophthalmology91(11), 1512–1517 (2007). https://doi.org/10.1136/bjo.2007.119453
[11] Nayak, J., Bhat, P.S., Acharya, U.R., Lim, C.M., Kagathi, M.: Automated identification of diabetic retinopathy stages using digital fundus images. Journal of Medical Systems32(2), 107–115 (2008). https://doi.org/10.1007/s10916-007-9113-9
[12] Gardner, G.G., Keating, D., Williamson, T.H., Elliott, A.T.: Automatic detection of diabetic retinopathy using an artificial neural network: A screening tool. British Journal of Ophthalmology80(11), 940–944 (1996). https://doi.org/10.1136/bjo.80.11.940
[13] Mookiah, M.R., Acharya, U.R., Chua, C.K., Lim, C.M., Ng, E.Y.K., Laude, A.: Computer-aided diagnosis of diabetic retinopathy: A review. Computers in Biology and Medicine43(12), 2136–2155 (2013). https://doi.org/10.1016/j.compbiomed.2013.10.007
[14] Wang, S., Tang, H., Hu, Y., Turk, A.I., Sanei, S., Saleh, G.M., Peto, T.: Localizing microaneurysms in fundus images through singular spectrum analysis. IEEE Transactions on Biomedical Engineering64(5), 990–1002 (2017). https://doi.org/10.1109/TBME.2016.2585344
[15] Abramoff, M.D., Reinhardt, J.M., Russell, S.R., Folk, J.C., Mahajan, V.B., Niemeijer, M., Quellec, G.: Automated early detection of diabetic retinopathy. Ophthalmology117(6), 1147–1154 (2010). https://doi.org/10.1016/j.ophtha.2010.03.046
[16] Antal, B., Hajdu, A.: An ensemble-based system for microaneurysm detection and diabetic retinopathy grading. IEEE Transactions on Biomedical Engineering59(6), 1720–1726 (2012). https://doi.org/10.1109/TBME.2012.2193126
[17] Noronha, K., Acharya, U.R., Nayak, K.P., Bhandary, S.V.: Decision support system for diabetic retinopathy using discrete wavelet transform. Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine227(3), 251–261 (2013). https://doi.org/10.1177/0954411912470240
[18] Adarsh, P., Jeyakumari, D.: Multiclass SVM-based automated diagnosis of diabetic retinopathy. In: Proceedings of the International Conference on Communication and Signal Processing (ICCSP), Melmaruvathur, India, pp. 206–210 (2013). https://doi.org/10.1109/ICCSP.2013.6577044
[19] Acharya, U.R., Chua, C.K., Ng, E.Y.K., Yu, W., Chee, C.: Application of higher order spectra for the identification of diabetic retinopathy stages. Journal of Medical Systems32(6), 481–488 (2008). https://doi.org/10.1007/s10916-008-9154-8
[20] Saleh, M.D., Eswaran, C., Mueen, A.: An automated blood vessel segmentation algorithm in histogram equalization and automatic threshold selection. Journal of Digital Imaging24(4), 564–572 (2010). https://doi.org/10.1007/s10278-010-9302-9
[21] ThinkPalm Technologies: Convolutional Neural Network: Are they the future of Artificial Intelligence? https://www.thinkpalm.com/blogs/convolutional-neural-network-future-artificial-intelligence (Accessed: 23 Jul 2025)
[22] Gulshan, V., Peng, L., Coram, M., Stumpe, M.C., Wu, D., Narayanaswamy, A., Venugopalan, S., Wider, K., Madams, T., Cuadros, J., Kim, R., Raman, R., Nelson, P.C., Mega, J.L., Webster, D.R.: Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. JAMA316(22), 2402 (2016). https://doi.org/10.1001/jama.2016.17216

[23] Gargeya, R., Leng, T.: Automated identification of diabetic retinopathy using deep learning. Ophthalmology124(7), 962–969 (2017). https://doi.org/10.1016/j.ophtha.2017.02.008

[24] Sridhar, S., Kandhasamy, J.P., Sinthuja, M., Minish, T.N.S.: Diabetic retinopathy detection using convolutional neural networks algorithm. Materials Today: Proceedings (2021). https://doi.org/10.1016/j.matpr.2021.01.059

[25] Luo, X., Pu, Z., Xu, Y., Yeung, W., Su, J., Dou, X.Y., Hu, J., Mou, L.: MVDRNet: Multi-view diabetic retinopathy detection by combining DCNNs and attention mechanisms. Pattern Recognition120, 108104 (2021). https://doi.org/10.1016/j.patcog.2021.108104

[26] Qureshi, I., Ma, J., Abbas, Q.: Diabetic retinopathy detection and stage classification in eye fundus images using active deep learning. Multimedia Tools and Applications80(8), 11691–11721 (2021). https://doi.org/10.1007/s11042-020-10238-4

[27] Butt, M.M., Iskandar, D., Abdelhamid, S., Latif, G., Alghazo, R.: Diabetic retinopathy detection from fundus images of the eye using hybrid deep learning features. Diagnostics12(7), 1607 (2022). https://doi.org/10.3390/diagnostics12071607

[28] Menaouer, B., Zoulikha, D., Kebir, N.E.H., Matta, N.: Diabetic retinopathy classification using hybrid deep learning approach. SN Computer Science3(5) (2022). https://doi.org/10.1007/s42979-022-01240-8

[29] Gunasekaran, K., Pitchai, P., Chjaitanya, G.K., Selvaraj, D., Sheryl, S.A., Almoallim, H.S., Alharbi, S.A., Raghavan, S.S., Tesemma, B.G.: A deep learning framework for earlier prediction of diabetic retinopathy from fundus photographs. BioMed Research International2022, 1–15 (2022). https://doi.org/10.1155/2022/3163496

[30] Fang, L., Qiao, H.: Diabetic retinopathy classification using a novel DAG network based on multi-feature of fundus images. Biomedical Signal Processing and Control77, 103810 (2022). https://doi.org/10.1016/j.bspc.2022.103810

[31] Khan, A., Kulkarni, N., Kumar, A., Kamat, A.: D-CNN and image processing based approach for diabetic retinopathy classification. In: Springer eBooks, pp. 283–291 (2021). https://doi.org/10.1007/978-981-16-2008-9_27

[32] Kobat, S.G., Baygin, N., Yusufoglu, E., Baygin, M., Barua, P.D., Dogan, S., Yaman, O., Celiker, U., Yildirin, H., Acharya, U.R.: Automated diabetic retinopathy detection using horizontal and vertical patch division – based pre-trained DenseNet with digital fundus images. Diagnostics12(8), 1975 (2022). https://doi.org/10.3390/diagnostics12081975

[33] Huang, G., Liu, Z., Van Der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. arXiv preprint arXiv:1608.06993 (2016). https://doi.org/10.48550/arXiv.1608.06993

[34] Adriman, R., Muchtar, K., Maulina, N.: Performance evaluation of binary classification of diabetic retinopathy through deep learning techniques using texture feature. Procedia Computer Science179, 88–94 (2021). https://doi.org/10.1016/j.procs.2020.12.012

[35] Xu, K., Feng, D., Mi, H.: Deep convolutional neural network – based early automated detection of diabetic retinopathy using fundus image. Molecules22(12), 2054 (2017). https://doi.org/10.3390/molecules22122054

[36] Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. Communications of the ACM60(6), 84–90 (2017). https://doi.org/10.1145/3065386
[37] Chandore, V., Asati, S.: Automatic detection of diabetic retinopathy using deep convolutional neural network. International Journal of Advance Research3(4), 3 (2017). https://www.ijariit.com/manuscripts/v3i4/V3I4-1321.pdf
[38] Lam, C., Yi, D., Guo, M., Lindsey, T.: Automated Detection of Diabetic Retinopathy using Deep Learning. AMIA Jt Summits Transl. Sci. Proc. 2018 May 18; 2017:147–155 (2018). PMID: 29888061; PMCID: PMC5961805.
[39] Ting, D.S.W., Cheung, C.Y.L., Lim, G., Tan, G.S.W., Quang, N.D., Gan, A., et al.: Development and Validation of a Deep Learning System for Diabetic Retinopathy and Related Eye Diseases Using Retinal Images From Multiethnic Populations With Diabetes. JAMA **318**(22), 2211 (2017). https://doi.org/10.1001/jama.2017.18152
[40] Pratt, H., Williams, B., Coenen, F., Zheng, Y.: FCNN: Fourier Convolutional Neural Networks. In: LNCS, pp. 786–798 (2017). https://doi.org/10.1007/978-3-319-71249-9_47
[41] Ravishankar, S., Jain, A., Mittal, A.: Automated Feature Extraction for Early Detection of Diabetic Retinopathy in Fundus Images. In: IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2009). https://doi.org/10.1109/cvpr.2009.5206763
[42] Goel, R., Kumar, V., Srivastava, S., Sinha, A.: A Review of Feature Extraction Techniques for Image Analysis. Int. J. Adv. Res. Comput. Commun. Eng. **6**(2), 152–155 (2017)
[43] Srivastava, D., Wadhavani, R., Gyanchandani, M.: A Review: Color Feature Extraction Methods for Content-Based Image Retrieval. Int. J. Comput. Eng. Manage. **18**(3) (2015). ISSN: 2230-7893
[44] Sk, P.S.: Extraction of Texture Feature Using GLCM and Shape Feature Using Connected Regions. Int. J. Eng. Technol. **8**(6), 2926–2930 (2019)
[45] ThangNoi, P., Kappas, M.: Comparison of Random Forest, K-Nearest Neighbor and Support Vector Machine Classifiers for Land Cover Classification Using Sentinel-2 Imagery. Sensors **18**, 18 (2018). https://doi.org/10.3390/s18010018
[46] Kataria, A., Singh, M.D.: A Review of Data Classification Using K-Nearest Neighbor Algorithm. Int. J. Emerg. Technol. Adv. Eng. **3**(6) (2013). ISSN 2250-2459
[47] Kumar, M., Chakravarty, S.: Diabetic Retinopathy Image Classification Using Support Vector Machine. University of Western Ontario. In: IEEE Xplore (2020)
[48] Wang, Y., Hu, B.: Hierarchical Image Classification Using Support Vector Machine. In: ACCV 2002: 5th Asian Conference on Computer Vision, Melbourne, Australia, Jan 23–25 (2002)
[49] Srivastava, D.K., Bhambhu, L.: Data Classification Using Support Vector Machine. J. Theory. Appl. Inf. Technol. (2009)
[50] Yang, F.: An Implementation of Naïve Bayes Classifier. In: Int. Conf. Comput. Sci. Comput. Intell. (CSCI), IEEE (2018). https://doi.org/10.1109/CSCI46756.2018.00065
[51] Yamashita, R., Nishio, M., Gian Do, R.K., Togashi, K.: Convolutional Neural Networks: An Overview and Application in Radiology. Insights Imaging **9**, 611–629 (2018). https://doi.org/10.1007/s13244-018-0639-9

[52] Alzubaidi, L., Zhang, J., Humaidi, A.J.: Review of Deep Learning: Concepts, CNN Architectures, Challenges, Applications, Future Directions. J. Big Data **8**, 53 (2021). https://doi.org/10.1186/s40537-021-00444-8
[53] Manoj, S.H., Bhosle, A.: Detection and Classification of Diabetic Retinopathy Using Deep Learning Algorithms for Segmentation to Facilitate Referral Recommendation for Test and Treatment Prediction. The National Institute of Engineering, Mysore
[54] Gonzalez, R.C., Woods, R.E.: Digital Image Processing. 4th edn. Pearson Education (2018).