Applied Soft Computing Journal xxx (xxxx) 109102

Contents lists available at ScienceDirect

Applied Soft Computing Journal

journal homepage: www.elsevier.com

Journal Cover Image

A modified convolutional neural network architecture for diabetic retinopathy screening using SVDD

Ali Karsaz [1,*]

Khorasan Institute of Higher Education, Mashhad, Islamic Republic of Iran

ARTICLE INFO

Article history:
Received 17 November 2018
Received in revised form 9 May 2022
Accepted 26 May 2022

Keywords:
Diabetic retinopathy
Convolutional neural network
Deep learning
Support vector domain description (SVDD)

ABSTRACT

Automatic diabetic retinopathy diagnostic methods are proposed to facilitate the examination process and act as the physician's helper. Most of the traditional convolution neural network (CNN) algorithms use only spatial features for image category recognition. This approach may not be optimal for the screening diabetic retinopathy because the retinal images have generally the same feature maps with minor differences in spatial domain. We propose a new high level image understanding using a modified CNN architecture mixed with modified support vector domain description (SVDD) as a classifier. This new innovative architecture uses two pathways extracting features of the retinal images in both spatial and spectral domains. The standard pre-trained AlexNet is chosen for modification to avoid the time complexity of the training algorithms. In spite the advantages of the modified AlexNet with two pathways configuration and standard SVDD classification, the different SVDD kernel functions affect the performance of the proposed algorithm. By using the appropriate transformed data into two or three dimensional feature spaces, the proposed SVDD can obtain more flexible and more accurate image descriptions. Also, we compared the performance of our approach with that of the commonly used as classification methods such as K-Means, subtractive and FCM clustering. Our proposed architecture achieves more than 98% precision and sensitivity for two class classification.

1. Introduction

Diabetic retinopathy (DR) is an eye disease that damages the retina of patients with long-standing diabetes. By the year 2030, the number of people diagnosed with DR will increase from 126.6 million in 2010 to 191 million, and the number of people with vision-threatening DR (VTDR) will grow from 37.3 million to 56.3 million by the same time [1]. Evidence shows that by diagnosing DR in early stages it can be treated just by diabetes management and can be prevented from further damages to the retina [2,3]. Ophthalmologist diagnoses the presence and severity of DR by carefully investigating fundus images and finding the different symptoms of DR, such as microaneurysms, hemorrhages, neovascularization, and exudates. Finding DR signs is highly subjective, which makes it difficult to diagnose in early stages. The high cost of the physical examination and lack of professional experts are the other obstacles for early DR diagnosis. Therefore, large numbers of early stage DR cases are missing from early diagnosis and treatment [4].

We propose a new modified CNN architecture using an improved support vector machine (SVM) classification for DR screening, making the following contributions:

- We propose a new CNN architecture to increase the performance in DR or NoDR classification by using some modifications in both the type of input data and in the last fully connected layer of the network. This unique approach distinguishes our work from the full training [5,6] and even fine-tuning approaches of a pre-trained CNN methods [7–10].
- Some CNN based approaches use only object features or spatial features for image category recognition [7,11,12]. In this paper we propose a new CNN architecture using a modified AlexNet which uses two pathways extracting retinal image features. These two deep feature extraction (FE) based on a modified standard AlexNet are proposed in two parallel pathways, to extract both the spatial and spectral features of the retinal images.
- There are multiple approaches in the literature using various algorithms to implement diabetic retinopathy classification based on segmentation and detection of different exudates, hemorrhage and blood vessels using some machine learning techniques such as SVM, K-nearest neighbor (KNN), K-Means (KM) algorithm and fuzzy C-Means (FCM) methodologies and etc. To our knowledge, the implementation history

* Associate Professor, Department of Electrical and Electronic Engineering,*Scientific advisor of the Artificial Intelligence and Clinical Laboratory Mathematics (AICLM)*

E-mail address: karsaz@khorasan.ac.ir

[1] IEEE member.

https://doi.org/10.1016/j.asoc.2022.109102
1568-4946/

of a typical CNN architecture in the field of automatic screening of DR is less than a decade and has not been well studied [11].

Few efforts have been made to Modified SVMs to cover the specific requirements of retinal image classification, for example by building nonlinear transformation and suitable kernel functions. In this paper, we proposed a modification to the SVDD's kernel functions that can take into account the difference between the DR and NoDR retinal images by imposing a nonlinear transformation in feature space.

The rest of the paper is organized as follows. The related studies on diabetic retinopathy using state-of-the-art CNN methodologies are described in Section 2. A typical CNN architecture is reviewed in Section 3. Section 4 explains the implementation of the proposed algorithm. The results of the simulations and the comparative analysis of the classifiers are demonstrated in Section 5. Finally, the conclusions and suggestions for future work are mentioned in Section 6.

2. Related works

According to recent studies, DR can be diagnosed accurately and consistently in early stages by applying automatic screening systems. The main purpose of these systems is to classify images into DR and NoDR classes [13]. There are multiple approaches in the literature using various algorithms to implement automatic screening of DR. Typically, these methods use hand-crafted features of retinal images for training their systems and classification. Some of these machine learning algorithms are artificial neural network (ANN), SVM, KNN, KM algorithm and FCM methodologies [14–17]. Supervised classification methods such as SVM [18,19], and KNN [20,21] are trained from different labeled images for segmentation purpose. Because the supervised classification methodologies use the prior knowledge about the image classes, they can improve the classification accuracy. Unsupervised classification algorithms also known as clustering approaches, train themselves from some hidden features in the dataset iteratively so they have this ability to characterize the properties of each class. Unsupervised classification methodologies include KM algorithm [22], expectation–maximization (EM) [23], subtractive clustering [24] and FCM methodologies [25]. The subtractive clustering method was first introduced in the field of extracting fuzzy rules [24]. The subtractive clustering method considers each data point as a potential cluster center and defines a measure of the potential of a data point to serve as a cluster center. The potential of data point is a function of its distance to all other data points. Thus, a high potential data sample would be a data with many neighboring data samples. By selecting *the range of influence*, the data points with the highest potential are determined so that the feature space is covered. Therefore the alignment parameter for subtractive clustering method is *the range of influence* in each of the data dimensions. Medical images are often corrupted by some environment noises, artifacts are caused by operator performance [26]. Since the standard FCM algorithm does not consider local spatial classification, it is very sensitive to noise. Many researchers have incorporated the local spatial information into the standard FCM to remove the effects of noise, such as Wang et al. proposed the FCM distance function as weight sum of distance influenced by local and nonlocal information (LNLFCM) [27]. Using local spatial features increases the computational cost because it needs the computation for each pixel neighborhood. Gong et al. extended fuzzy local information C-Means algorithm by replacing the Euclidean distance in the object function of the FCM by kernel distance-based cost function [28]. As an effort in comparing different available techniques, a comparative analysis of nine common classifier algorithms is implemented in the application of automatic screening of diabetic retinopathy cases [29].

Therefore, recent studies are using the state-of-the-art convolutional neural network (CNN) for various fields, especially in medical image analysis [30]. One of the main reasons for implementing CNN in medical applications is its ability to extract features automatically by using deep multiple layers [7]. Therefore, there has been an increase in using CNN in medical diagnosis applications. For instance, CNN was used for grading brain tumors in magnetic resonance imaging (MRI) scans [5,6,12]. Another CNN-based method was conducted by [31] for feature extraction and ensemble classification for retinal blood vessel segmentation and a study related to severity DR diagnosis using CNN was addressed in [32]. Also, two different comparative studies of two CNN structures for DR screening have been performed in [8] and [33].

The CNN is a class of deep learning models that can learn a complex hierarchy of features by building high-level features from low-level ones. Also, the validation of the performance of the final trained network on clinical data is an important step in performance analysis of the work. Requiring a large amount of medical training images, extensive computational and memory devices, complications about training of a deep CNN such as overfitting and convergence issues made the full training of a CNN (or training from scratch) tedious and in some cases impractical [9,34]. For example, the well-known CNN architecture named AlexNet consists of approximately 60 million parameters within its structure which was trained using 1.2 million images labeled with 1000 separated classes in the large database called ImageNet. An alternative to full training approach is fine-tuning of a pre-trained CNN or modification and customizing CNN based on specific applications [5–11,35,36]. There are different CNN architectures for detection and classification such as CifarNet (Roth et al. 2016; [37], Alexnet [38], GoogLeNet [39], and VGGNet [40]) with different model training parameter values. The GoogLeNet and VGGNet models are significantly more complex than both CifarNet and AlexNet [41]. Some CNN architectures such as VGGNet and GoogLeNet are usually counted as "deeper architectures" [7]. The GoogLeNet model introduces a new module called Inception, which concatenates filters of different sizes and dimensions into a single new filter. In these architecture, Inception layers are repeated many times, leading to a 22-layer deep model in the case of the GoogLeNet (when counting only layers with parameters) where the overall number of layers is about 100 [39].

The main contribution of our work in this paper is the proposal of a deep learning CNN architecture consisting of both spatial and spectral domains image feature extraction mixed with a support vector domain description (SVDD) classification for a retinal image, DR or NoDR recognition.

Although, deeper architecture have recently shown relatively high performance for challenging image processing tasks, but we do not anticipate a significant performance gain through the use of deeper architecture. The objective of our work is to examine the capabilities of two pathways extracting features algorithm (spatial and spectral domains) in comparison with the spatial extracting features approach. Whatever the typical of CNN structure it is necessary to have some modification for better binary classification.

To this end, in this paper, a pre-trained AlexNet architecture was employed and enforced to classify fundus images of a clinical dataset into cases of DR patient or healthy. The first layers of a typical CNN is mostly related to extracting general information from the images such as the edges, while their last layers are specifically trained to extract more detailed features related to the images dataset. Therefore, the type of the input retinal image and the weights of the last layers of a CNN can be modified to adapt the networks for our application and increase the performance accuracy. This new innovative architecture uses two pathways extracting features of the retinal images in both spatial and spectral domains. In addition, the SVDD is a domain description method inspired by SVM algorithm that tries to find the sphere with minimum volume containing almost all objects. By using the appropriate transformed data into two or three dimensional feature space, the proposed SVDD can obtain more flexible and more accurate image descriptions. For any two or three dimensional feature maps obtained by the proposed algorithm, the two-pathway architecture performs much better than individual pathways.

There are a few number of medical images to train a CNN which has a lot of weights needed to be trained. This limitation of the image samples, usually leads to the overfitting problem [12,42]. The objective of this paper is not to use a full training CNN approach but to modify the last layer of a standard CNN and the type of input data to achieve the highest performance for different retinal images.

3. A typical CNN architecture and the standard AlexNet architecture

CNN is a kind of multilayer neural networks which typically consists of convolutional, subsampling, and fully connected (FC) layers [41]. CNN uses some of conventional algorithm for training as other traditional neural networks, i.e. gradient descent approach, backpropagation (BP) algorithm, and it also uses some concepts rather than those used by a traditional neural network such as convolution, dropout, pooling, etc. A complete deep CNN model contains several convolutional layers to form deep architecture. To construct a 3D convolutional layer, a set of small feature extractors will be established, which sweep over their input to extract a stack of higher-level representations. These small feature extractors for convolutional layers usually are known as *convolutional kernels* or in some cases, are inspired by neuroscience, called *local receptive fields* [43] with the predetermined sizes. The convolutional kernel window moves across every feature map of the input layer to create the activations for the next layer. Fig. 1 shows a simplified illustration of some layers which are typically involved in a CNN especially the AlexNet. For simplification we assume that this simplified example of a CNN accepts a $9 \times 9 \times 3$ image patch as input and with the 3×3 kernel.

Typically after each convolutional layer in the standard AlexNet there are some *rectifier linear units* (ReLUs) to improve the network performance. There are different kinds of ReLUs in CNN literatures available to apply. The most common type, is the simple nonlinear unit where accepts the input of a neuron if it is positive, whereas it returns to zero if the input is negative. It has been shown that the use of the ReLUs after the convolutional layers can expedite the training of a CNN [38,44].

Pooling layer as a subsampling layer reduces the dimensionality of each feature map but keeps the most important features. This layer helps to reduce the amounts of learning parameters and is usually placed after the convolutional and ReLU layers (for AlexNet, GoogLeNet and CifarNet). The two most common types of pooling layer are max-pooling which is used in our paper, and mean-pooling. Another layer of processing is called *local response normalization* (Lrn) unit is implemented to enforce competitions between features at the same spatial location across different feature maps. Neurons in a FC layer have full connections to all activations in the previous layer, as seen in a traditional neural network [7,45]. So the FC layer is a common concept between traditional neural network and CNN, where every neuron is connected to each neuron in the next layer. But instead of a simple connection between each input neuron to the next hidden layer in a traditional neural network, a window of neurons in the input layer are connected to one neuron in the next hidden layer in CNN. The output layer of a standard AlexNet is the *softmax* layer, or the regression layer, to generate the final result.

In order to avoid the overfitting, recently introduced a layer is called *dropout layer*, which sets the output of some neurons to zero with probability 0.5. The dropout method prevents complex co-adaptations [46]. By using ReLUs and dropout layers, the outputs of some hidden neurons will be zero, can prevent a large network from overfitting.

The standard AlexNet achieved significantly improved performance over the other non-deep learning methods for ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2012 [38]. The standard AlexNet has approximately 60 million parameters, about 5 million parameters in its convolution layers and approximately 55 million parameters in fully connected layers. The AlexNet computes $11\times11, 5\times5, 3\times3, 3\times3$ and 3×3 convolutions within the same layers of the *Maxpool* and concatenates the output of the whole process to pass it to the *Softmax* layer as the latest layer of the network. Fig. 2 demonstrates schematic diagram of the standard AlexNet, consists of 5 convolutional layers, 3 pooling layers, 3 fully connected layers and 1 *Softmax* layer. This model consists of 12 main layers which includes convolution, pooling, FC and Softmax layers. In Fig. 2, letters S, P and G stand for stride, padding and group, respectively. Stride denotes how many steps we are moving in each steps in convolution. Padding refers to the amount of pixels added to an image when it is being processed by the kernel of a CNN. Group controls the connections between inputs and outputs. We use this CNN architecture with some necessary modifications to obtain the main features of retinal images.

4. Proposed method

AlexNet architecture [38] with 650.000 neurons has been trained on ImageNet as a large database that is basically used for . There are some differences between 1000-class problem and a binary classification such as DR or NoDR screening. A difference is about the softmax

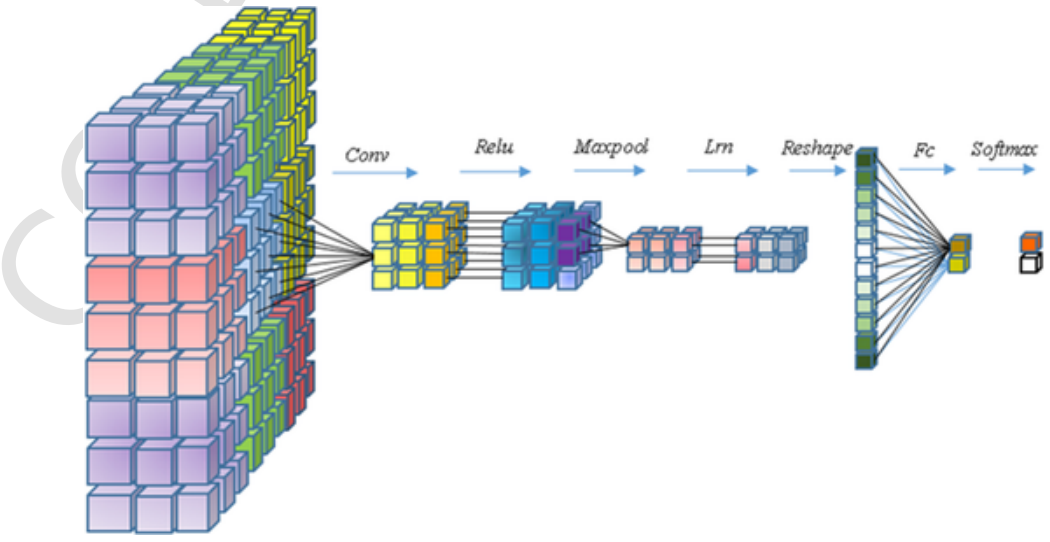

Fig. 1. A simplified illustration of a typical CNN architecture consists of one convolutional layer (conv), one rectifier linear unit (ReLU), one maximum pooling (Maxpool), one local response normalization (Lrn), reshape layer, one fully connected layer (Fc) and finally one softmax layer.

Fig. 2. The main layers of the standard AlexNet [41], consists of 5 convolutional layers (colored red), 3 pooling layers (colored blue), 3 fully connected layers (colored green) and finally one Softmax layer (colored brown).

layer in a typical CNN architecture. This final layer acts as a final classifier of a typical CNN. The Softmax for a 1000-class classification problem is the best solution. Suppose that some vector components could be negative or greater than one, and might not sum to 1, but after applying softmax, each component will be in the interval zero to one. On the other hand we have some stronger classifiers such as SVM. SVM is a professional binary classifier, maps data samples so that the samples of the separate categories are divided by a clear gap that is as wide as possible.

Another difference is about the three FCs in AlexNet architecture. As mentioned before, neurons in a FC layer of a CNN architecture have full connections to all activations in the previous layer, as seen in a traditional neural network. So a FC layer is a common concept between traditional neural network and CNNs, where every neuron is connected to each neuron in the next layer. But there is a minor difference, instead of a simple connection between each input neuron to the next hidden layer in a traditional neural network, a window of neurons in the input layer are connected to one neuron in the next hidden layer in CNN.

Just similar to traditional neural networks there are two different ways of implementation for a specific application. At first, one could use a multiple layers structure with different number of neurons in each layer, and second using a traditional neural network with just one hidden layer with appropriate number of neurons in that layer. In fact, the number of hidden layers in a traditional neural network is not an important point. Our experimental results show that the three fully connected layers (FC6, FC7, FC8) in the AlexNet architecture, act just as multilayers in a traditional neural network. So, we can reduce the number of FCs in our modified CNN.

Three main stages constitute our proposed DR detection algorithm: (1) Image preparation (2) Spatial and spectral features extraction (3) SVDD classification.

Image preparation section itself consists of rescaling, normalizing and finally obtaining a 2D color histogram of a given retinal image. Also, we use the AlexNet in o proposed algorithm as a multi-layer feature extractor with some modification to obtain more spatial and frequency domain complementary information of a retinal image. The last section of our proposed scheme is the improved SVDD classification algorithm which determines the image category for a color fundus image.

The overall structure of the proposed scheme is illustrated in Fig. 3. The input of our system unlike other traditional CNN application, is a 2D histogram of the retinal image, and the output of the system is the DR or NoDR label of the image. After several layers of convolution and pooling, the 2D histogram of the input image can be converted into a 2D or 3D feature vector, which contains the spatial and spectral features within the image. These obtained features are ready to be used in conjunction with classifier such as SVM, FCM or SVDD. For a deep understanding of the new scheme performance, the performances of the 2D and 3D classification will be calculated separately, which are called *2D classification* or *3D classification* case studies.

4.1. Image preprocessing

Medical images usually contain noise and shading artifacts due to interference and other phenomena that affects the process of classification in screening systems [47]. Artifacts due to non-uniform illumination which is a general problem in retinal imaging degrade the efficiency of the image classification as well as the effects of camera variations. Preprocessing is an essential step to reduce the image variation by normalizing and equalization of the irregular illuminations of a color fundus image.

4.1.1. Rescaling the images

To decrease the variation among images due to different camera resolutions and settings, an image preprocessing algorithm is applied to the images. Due to the different retinal picture sizes, the first step of the algorithm is clipping the black borders of some images on the left and right sides. The most important point about these different fundus images are the aspect ratio between the length and width of the retinal circle that should be quite similar. In the next step, a rescaling procedure should be done such that all the input images have the same size. Finally, the color of each pixel is subtracted by the local average, mapping the average to 50% gray. Using this approach, the sharpness of the images will be more unified.

4.1.2. Illumination correction

Several papers have been published in removal of non-uniform of illumination. Nyul [48] compensated the non-uniform illumination by a polynomial surface fitting algorithm based on considering the intensity of the input image as a product of the luminosity and reflectance component. We only use the red and green channels in our algorithm and normalize the R and G values (R stands for red and G for green) in order to reduce the sensitivity to changing light by the following equations [49]:

$$r = R/(R+G+B) \tag{1}$$

$$g = G/(R+G+B) \tag{2}$$

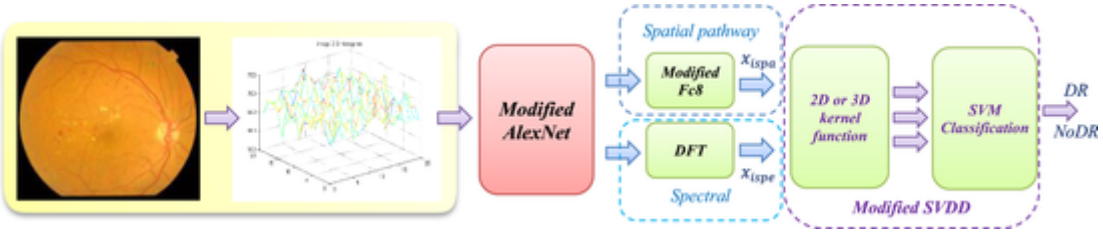

Fig. 3. Schematic diagram of the proposed modified AlexNet.

For example consider two pixels A and B with the same color RGBs but different brightness. Such differences may due to the light bias of camera for the corner and center of the retina. Because a typical retinal has a highly curved surface, the variation may be removed by dividing the three components of a color pixel by its intensity. We can consider $RGB_A = \gamma RGB_B$ for some reasonable γ. By using Eqs. (1) and (2) we would have the same brightness for two pixels as follows:

$$RGB_{Bnorm} = \begin{bmatrix} R_B & G_B & B_B \end{bmatrix} / \left(R_B + G_B + B_B\right) * 255 \tag{3}$$

$$\begin{aligned} RGB_{Anorm} &= \begin{bmatrix} R_A & G_A & B_A \end{bmatrix} / (R_A + G_A + B_A) * 255 \\ &= \begin{bmatrix} \gamma R_B & \gamma G_B & \gamma B_B \end{bmatrix} / (\gamma R_B + \gamma G_B + \gamma B_B) * 255 \\ &= \begin{bmatrix} R_B & G_B & B_B \end{bmatrix} / (R_B + G_B + B_B) * 255 \\ &= RGB_{Bnorm} \end{aligned} \tag{4}$$

This gives the normalized intensity pixels for corner and center ones.

4.1.3. Color histogram

For a given image the color histogram is defined by counting the number of times each color occurs in the image array. Histograms are invariant to rotation about an axis perpendicular to the image plane, and change slowly under change in scale and change of angle of view.

4.2. Spatial and spectral features extraction

4.2.1. Spatial feature extraction

In medical imaging and diagnosis field, it is relatively rare to have an image dataset of sufficient size to completely train a CNN from scratch [7]. In addition, the state-of-the-art CNN included in the GitHub or Keras core library demonstrate a strong ability to be generalized to images outside the ImageNet dataset via transfer learning, such as feature extraction and fine-tuning[2] [40,50]. Therefore, it is very common to fine-tune a CNN that has been trained using a large labeled dataset from a different application to avoid training networks for many general features [51]. Training a CNN from scratch requires a large amount of data as well as extensive computational and memory resources [7]. Training a deep network with small dataset often leads to overfitting and convergence issues. Therefore, in this paper, a pre-trained CNN architecture, AlexNet is modified and is used for DR classification.

The modification of a pre-trained AlexNet is one of the main parts of our algorithm. The modification begins with transferring the weights from a pre-train AlexNet as a part of our final network, the only exception is that the weights of the last fully connected layer (Fc8) is omitted. The completely characteristics of our model as shown in Fig. 3, are listed in Table 1.

Our new scheme begins with producing a 2D histogram 227 × 227, from the original retinal image, and then proceeds with two pairs of convolutional, rectifier linear unit, pooling and local response normalization layers i.e. Cov1, ReLU1, Pool1, Lrn1, Cov2, ReLU2, Pool2, Lrn2, respectively. These latter two pairs of modified AlexNet layers, map the 227 × 227 input histogram to 13 × 13 feature maps. The architecture then proceeds with a sequence of three convolutional and ReLU layers i.e. Cov3, ReLU3, Cov4, ReLU4, Cov5, ReLU5. These layers implement a convolutional layer with 3 × 3 convolutional kernel size. The sequence of the modified AlexNet in our scheme is then followed by three pooling, reshape and ReLU layers i.e. Pool5, Res1 and ReLU6. The modified AlexNet in our algorithm, is completed with only two fully connected layers i.e. Fc6 and Fc7, instead of three FC layers typically used in the standard AlexNet (so the Fc8 is replaced by a modified Fc8 to be considered a typical uniform summation). In our study, we deal with 2-class classification tasks, in two different 2D or 3D feature space. So the new modified FC layer, in up section (spatial domain pathway as shown in Fig. 3) has one neuron, and in down section (spectral domain pathway) has one or two neurons depending on the 2D or 3D scenarios under study. In the spatial domain pathway, only the last layer (Fc8) is substituted with a typical fully connected summation as an activation function.

4.2.2. Spectral feature extraction

As mentioned before, the minor differences between DR and NoDR retinal images are closely related to their frequency domain properties. Our proposed method processes the data in the frequency domain to attain greater accuracy besides to the spatial feature processing. By separating the image feature into different sub-bands, important difference occurs over varying low to high frequencies. When digital images are handled at multiple resolutions, the discrete Fourier transformation (DFT) is viable mathematical tool. So the block diagram DFT in Fig. 3, returns the discrete Fourier transform of the Fc7 output, computed with a fast Fourier transform (FFT):

$$X(k) = \sum_{j=1}^{N} x(j)\, w_n^{(j-1)(k-1)} \tag{5}$$

$$x(j) = \frac{1}{N} X(k)\, w_n^{-(j-1)(k-1)} \tag{6}$$

Where $x(j)$ is a sequence of N complex numbers for $j = 1, \ldots, N$ and $X(k)$ is another sequence of complex numbers and

$$w_n = e^{-\frac{2\pi i}{N}} \tag{7}$$

In this section of our algorithm, we integrate the spectral and spatial features together to construct a powerful framework using 2D or 3D classification.

4.3. SVDD classification

The standard SVM is a supervised learning method that has the aim of determining the location of decision boundaries or hyperplanes that provide the optimal separation of the classes based on statistical theory [52,53]. The SVDD is a domain description method inspired by SVM algorithm that tries to find the sphere with minimum volume containing almost all objects [54]. Let the x_i be a dataset containing N_s sample points as follows:

$$X = \left(x_1, x_2, \ldots, x_i, \ldots, x_{N_s}\right) \tag{8}$$

where X is a data matrix with the size of $N \times N_s$, N represents the dimension of each x_i feature vector and N_s represents the number of feature vectors (the output dimension of networks). Since the output features obtained by the modified CNN architecture are not normally spherically distributed in the input space of the classifier data, the SVDD algorithm uses a nonlinear transformation ($\varphi(.)$) to transform the data from input space to a new high dimensional feature space. The following *Wolfe* dual form, which is a maximization problem respect to α_i ($\alpha_i \geq 0$ is a Lagrange multiplier) will be obtained [55]:

$$W(\alpha_i) = \sum_i \alpha_i \left\langle \varphi(x_i)\, \varphi(x_i) \right\rangle - \sum_{i,j} \alpha_i \alpha_j \left\langle \varphi(x_i)\, \varphi(x_j) \right\rangle \tag{9}$$

Where the $\left\langle \varphi(x_i)\, \varphi(x_j) \right\rangle$ is the inner product can be replaced with an appropriate kernel function $K(x_i, x_j)$ such that satisfies the *Mercer's* theorem [55]. There are different kernel functions; however, the *Gaussian* kernel function is shown to have better performance than the others [54];

[2] *1-https://github.com/fchollet/deep-learning-models.*.

Table 1
The new proposed scheme based on modified AlexNet architecture used in our experiments, consists of preprocessing layers (the yellow rows), 5 convolutional layers (the red rows), 3 pooling layers (the blue rows), 2 fully connected layers (the green rows), 7 rectifier linear units (the pink rows), 2 dropout layers (the gray rows), 2 local response normalization and 1 reshape layers (without color rows) and finally modified FC8 and SVDD layers (the cyan rows).

$$K\left(x_i, x_j\right) = \exp\left(-\frac{\left\|x_i - x_j\right\|^2}{\sigma^2}\right), \sigma \epsilon R^+ \tag{10}$$

The different kernel functions result in different description boundaries in the original input space of the SVM. The generic Gaussian kernels in (10), regard each component of x_i with equal emphasis in their effects into feature space. The problem is to find a suitable nonlinear transformation for each component of x_i to have a larger effect in feature space. To this end, the nonlinear transformation $\varphi\left(x_i\right) = x_i^n$ corresponding to each modified AlexNet outputs (i.e. spatial and spectral features) is used to scale each feature before mapping it into feature space. In our proposed algorithm the Gaussian kernel function with above nonlinear pre-transformation ($\varphi(.)$) is considered to transform the data from input space to a new high nonlinear feature space. So we consider two kernel functions, the first one is a two-dimensional kernel (2D) and the second one is a three-dimensional kernel function (3D) as follows:

$$K_{2D}\left(\varphi\left(x_i\right)\varphi\left(x_j\right)\right) = \exp\left(-\frac{\left\|\varphi_{2D}\left(x_i\right) - \varphi_{2D}\left(x_j\right)\right\|^2}{\sigma^2}\right) \tag{11}$$

$$K_{3D}\left(\varphi\left(x_i\right)\varphi\left(x_j\right)\right) = \exp\left(-\frac{\left\|\varphi_{3D}\left(x_i\right) - \varphi_{3D}\left(x_j\right)\right\|^2}{\sigma^2}\right) \tag{12}$$

In this paper, we present a modified AlexNet model with two pathways retinal image recognition, which extract both spatial and spectral features of images respectively. Experiments show that the two kinds of features contain complementary information for the category recognition of an image as the two-pathway model always achieve better performance than single pathway models [56]. A neural network that consists of two interconnected pathways (a convolutional pathway and a deconvolutional pathway) has been successfully implemented on lesion segmentation in [57].

So each feature space x_i has two main parts, spatial feature space x_{ispa}, obtained by the spatial pathway and spectral feature space x_{ispe}, obtained by the spectral pathway, respectively (as shown in Fig. 3). The two kernel functions (11) and (12) are considered with 2D and 3D nonlinear transformations as follows,

$$\left(x_{ispa}, x_{ispe}\right) \xrightarrow{\varphi_{2D}(x_i)} \left(x_{ispa}^{n_1}, \left|x_{ispe}\right|^{n_2}\right) \tag{13}$$

$$\left(x_{ispa}, x_{ispe}\right) \xrightarrow{\varphi_{3D}(x_i)} \left(x_{ispa}^{n_1}, Re(x_{ispe})^{n_2}, Im(x_{ispe})^{n_2}\right) \tag{14}$$

where the free parameters n_1 and n_2 are the degree of the polynomials, and $Re\left(x_{spe}\right)$ and $Im\left(x_{spe}\right)$ represent the real and imaginary parts of the spectral feature.

5. Experimental results and analysis

5.1. Clinical data and methodology

The main problem in the research of DR screening is the non-availability of a suitable standard datasets for training, testing and evaluation of developed algorithm. Every academic research groups uses their own databases for evaluation and testing with different number of samples, therefore a general comparison with similar studies would not be possible. Many papers have widely reported their results based on retinopathy on-line challenge (ROCh) dataset[3] [58] which is an international competition associated with SPIE medical imaging (MI' 2009) [59,60]. This database provides 50 images for training and 50 images for testing. In this database, only the annotations of the training set (microaneurysms are annotated) are publicly provided. Another well established public dataset is DIARETDB1 which is a standard diabetic retinopathy database[4] has been used by some papers such as [58,59,61–63]. This database consists 89 retinal images, of which 84 images are labeled as DR and remaining 5 images are considered as NoDR. These images were acquired by using a digital fundus camera with 50° field of view (FOV). Both the ROCh and the DIARETDB1 public databases are suitable for feature extraction, microaneurysm, exudate, and macula detection. The publicly available dataset from UCI Diabetic Retinopathy[5] has been used by some papers especially in field of classification [29]. It is important to note that only extracted features are available within the public UCI dataset. The features of this dataset have been extracted from the publicly available Messidor database of 1151 fundus images of patients [64].

Our proposed algorithm is evaluated with the fundus images available in the DIARETDB1 dataset as input to our two-pathway modified AlexNet architecture. Also, our algorithm is applied to diagnose DR cases from real fundus images that are captured from the Navid-Didegan ophthalmology clinic, Iran. The labeling was performed by an experienced independent ophthalmologist. This private dataset contains 94 retinal images, in which 47 of the images are labeled as NoDR and 47 images are labeled as DR class. To compare the results in both the DIARETDB1 and Navid-Didegan datasets, we augmented and randomly missing some NoDR and DR images to the DIARETDB1 dataset, respectively. Therefore, a balanced DIARETDB1 database of 94 retinal images is created in which both DR and NoDR classes were represented equally (47 samples for each class). All images are in different compressed formats such as JPEG, JPG and PNG with two different sizes, 3872 × 2592, 3060 × 2580 pixels for the Navid-Didegan database. Therefore, by applying the labeled images of these two datasets to the modified AlexNet-SVDD in the test step, the performance of the network is examined for the test images, in which all the test images are independent of the train data. We hold out 70% of the dataset for training in classification procedure, while 30% is used to test the performance of the methodology.

5.2. Evaluation criterion for performance analysis

To evaluate the performance of a proposed classification or clustering method for the clinical data, all clustering indices can be used. These different indices can be used in three different studies: internal, relative or external studies [55]. The internal study or sensitive analysis uses some metrics to determine how a particular internal variable affect the performance of the proposed classification method under a given database. The relative study is based on evaluation of the proposed classification results by comparing them with the results of other methods. The external study uses some metrics to evaluate classification performance based on specific reference data. To compare the per-

[3] 1-http://webeye.ophth.uiowa.edu/ROC/.
[4] 2-*http://www.it.lut.fi/project/imageret/diaretdb1.*.
[5] 3-*https://archive.ics.uci.edu/ml/datasets/Diabetic+Retinopathy+Debrecen+Data+Set#*.

formance of the proposed modified AlexNet-SVDD for the clinical data, these three studies have been considered using different indices. Table 2 defines these values as being used in this paper. In this table, *TP*, *FP*, *TN*, and *FN* represent true positive, false positive, true negative, and false negative results of the classification algorithm, respectively. *P* and *N* represent the labeled DR and NoDR samples, respectively, and $P+N$ demonstrates the total number of samples or $P+N=TP+FP+TN+FN$.

In a wide range of medical image classification, the free response operating characteristic (FROC) curve is used [65] as a fundamental index for diagnostic test evaluation [7]. Also, Kappa coefficient of agreement has been used in both internal and external studies [66,67]. Kappa error relations are used to gain insights about who much a clustering method is better than another on a specific dataset [55,68]. Consider N represents the number of normal points (or NoDR) and P represents the number of patient points (or DR) and the contingency table of two classifiers, C_1 and C_2 is as shown in Table 3. Diversity between two classifiers is measured by κ represents the Kappa coefficient as

$$\kappa = (OA - AC) / (1 - AC) \tag{15}$$

Where OA is the observed agreement or accuracy in Table 2 and AC is the agreement by chance. OA shows the probability that the two classifiers will be both either correct or incorrect when classifying a randomly chosen data point. AC is represented for the probability that the two classifiers will agree by chance on a randomly chosen sample point. So the two quantities are as:

Table 2
Performance indices [62].

Accuracy	$(TP+TN)/(P+N)$
Precision or predictive value	$TP/(TP+FP)$
Sensitivity or Recall	$TP/(TP+FN)$
Specificity	$TN/(TN+FP)$

Table 3
The contingency table of two classifiers, C_1 and C_2.

		C_1		
		Correct	Wrong	Total
C_2	Correct	*TP*	*FP*	$TP+FP$
	Wrong	*FN*	*TN*	$FN+TN$
	Total	$TP+FN=P$	$FP+TN=N$	$P+N$

$$OA = (TP+TN)/(P+N) \tag{16}$$

$$AC = (TP+FP)(TP+FN)+(FP+TN)(FN+TN)/(P+N)^2 \tag{17}$$

Substituting (16) and (17) in Eq. (15), the Kappa coefficient obtains:

$$\kappa = \frac{2(TP\times TN - FP\times FN)}{(TP+FP)(FP+TN)+(TP+FN)(FN+TN)} \tag{18}$$

5.3. Preparing dataset

We perform our analysis in retinal images collected from two databases (Navid-Didegan and DIARETDB1 datasets), where the image labeling is provided for the majority of the images. As the train and test datasets include different fundus images taken with different devices to remove the variations in images, preprocessing algorithm described in Section 5.1 is implemented on both train and test points. The result of image preprocessing steps are demonstrated in Fig. 4(a)–(d). Fig. 4 shows two selected examples of DR (the first row) and NoDR (the second row) fundus images and their experimental results of rescaling, illumination normalizing and their obtained 2D histograms from Navid-Didegan dataset which is a private database.

The image preprocessing level as mentioned before, consists of three main parts: *the image rescaling* (Fig. 4(b)), *the RGB equalization* (Fig. 4(c)) and *the 2D histogram extraction* (Fig. 4(d)). The task for *image rescaling* is performed such that all the input images have the same size of $2592\times2592\times3$. At *the normalization* level, the non-uniform brightness of the retinal fundus image will be removed by dividing the three components of a color pixel by its intensity based on Eqs. (1) and (2). The third step of preprocessing level in our proposed algorithm is the choosing red and green components of the retinal image, because these channels contain most information with blood and vessels in a retinal fundus image. In this stage, the color histogram is obtained by counting the number of times each red and green colors occur in the image array. The next step is to apply the preprocessed images to the proposed network that has already been modified and discussed in Sections 5.2 and 5.3.

5.4. Performance analysis of SVDD's parameters on Navid-Didegan dataset

As mentioned in Section 5.3 the different kernel functions result in different description boundaries in the original input space of the SVM. These kernels map the features into the high nonlinear features space

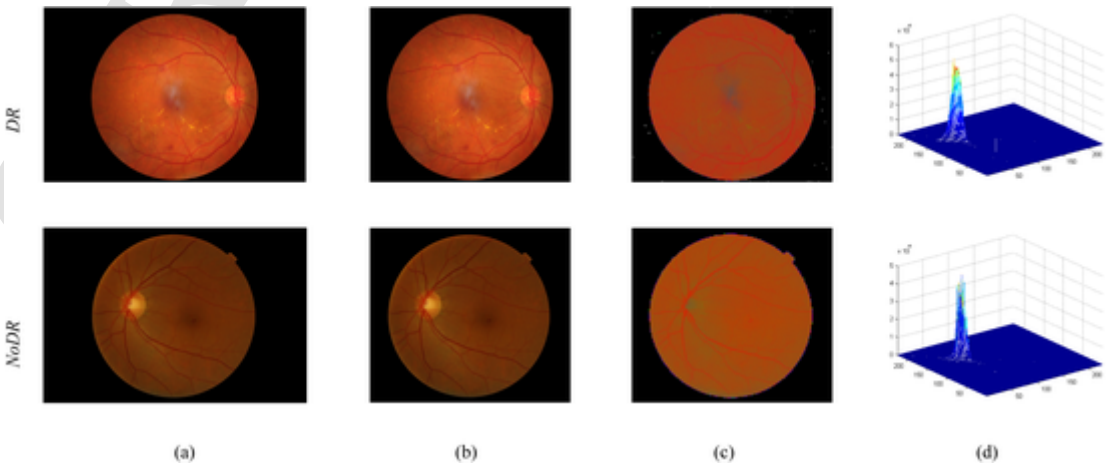

Fig. 4. The image preprocessing level of our proposed algorithm (Section 5.1) (a) The DR (the upper row) and NoDR (the bottom row) retinal fundus images (b) rescaled images (2592×2592 patches) (c) RGB Normalized images and (d) 2D histograms (227×227 patches), (images from Navid-Didegan dataset).

[55]. Due to the computational cost of the meta heuristic algorithms, such as particle swarm optimization (PSO) or genetic algorithm (GA), an exhaustive search method is used in this paper to find the optimum values of the kernel parameters [69]. So the best values of n_1 and n_2 involved in the kernel functions for a number of trial runs are selected. Navid-Didegan dataset is used in this scenario to evaluate the proposed modified AlexNet-SVDD algorithm. The effects of some different kernels with different degrees are shown in Fig. 5 with both 2D and 3D feature maps obtained by the modified AlexNet-SVDD containing 47 × 2 retinal images from Navid-Didegan dataset (47 as DR and 47 as NoDR images). Unlike the traditional SVDD algorithms that the predetermined kernel function acts as a measure to optimize the Lagrange cost function, we use the different kernel functions with some nonlinear transformation directly on input features as mentioned in (13) and (14). Therefore, a simple kernel function can be replaced by a kernel function with a nonlinear pre-transformation into 2D or 3D feature spaces. When a suitable degree of this transformation is chosen, a better and more tighter description can be obtained. In this scenario, two different value of degrees i.e. $n_1 = 1, n_2 = 1$ and $n_1 = 2, n_2 = 4$ are separately studied and the sphere descriptions are computed. Fig. 5 shows the location of the feature maps in input space with the degrees $n_1 = 1$ of the spatial kernel and $n_2 = 1$ for the spectral kernel. The left column of this figure shows the results obtained by using the 2D kernel function, so the inputs of the SVDD algorithm would be the nonlinear transformation over the spatial pathway (i.e. $x_{ispa}^{n_1}$) and spectral pathway (i.e. $\left|x_{ispe}\right|^{n_2}$), respectively. The right column of Fig. 5, shows the results obtained by using 3D kernel function, where the inputs of the SVDD algorithm would be $x_{ispa}^{n_1}$ from spatial pathway and $Re\left(x_{ispe}\right)^{n_2}$ and $Im\left(x_{ispe}\right)^{n_2}$ from the spectral pathway, respectively. Fig. 5(a) shows the spatial pathway outputs (see Fig. 3) with 2D kernel function in the left column and 3D kernel function in the right column (are the same for the both) where the retinal images are 47 as DR (the red stars) and 47 as NoDR (the blue stars), respectively. Also, Fig. 5(b) shows the spectral pathway outputs with the 2D and 3D kernel functions in the left and right columns, respectively. And finally, in Fig. 5(c) the spatial pathway outputs respect to spectral pathway outputs are illustrated. Fig. 5(d) shows the classification results obtained by improved SVDD with 2D and 3D in the left and right columns, respectively. In Fig. 5(d) with only 2D kernel function, the sphere description and support vectors are also plotted. As shown in this figure, these degrees of the kernel functions for n_1 and n_2 (i.e. $n_1 = n_2 = 1$) with both 2D and 3D feature space, do not have the good classification results. Based-on trial and error methodology and over numerous independent trials and observing the performance of the algorithm, the best values for $n_1 = 0.8, n_2 = 4$ are selected. By setting $n_1 = 0.8, n_2 = 4$, as shown in Fig. 6 the location of the feature maps in input space will be changed in order that the SVDD can easily classify them. Also, in Fig. 6(d) with only 2D kernel function, the sphere description and support vectors are plotted.

5.5. Sensitivity analysis of SVDD's parameters on DIARETDB1 dataset

In order to study the impacts of the degree of nonlinear transformation in the kernel functions (i.e., n_1, n_2) on the results of the proposed algorithm, different degrees of kernel function are considered. DIARETDB1 dataset is used in this scenario to evaluate the proposed modified AlexNet-SVDD algorithm. Fig. 7 depicts the accuracy of the algorithm for different degree values of kernel functions. Fig. 7(a)–(b) shows the obtained results for different values of n_1, while the degree of the spectral kernel is assumed to be constant (i.e. $n_2 = 4$). And Fig. 7(c)–(d) shows the obtained results for different values of n_2, while the degree of the spatial kernel is assumed to be constant (i.e. $n_1 = 0.8$). Fig. 7(a) and (c) compares the FROC curves of our approach with different values of n_1 and n_2. To avoid clutter in this figure, only a selected set of representative FROC curves has been shown. Since For small values of n_2, two classes of DR and NoDR images could not efficiently separated in spectral feature space (also shown in Fig. 5(d)), so the SVDD results are not very good (Fig. 7(c–d)). For large values of the n_2 up to 4 (also shown in Fig. 6(d)) the separation between two classes can easily performed by the SVDD. For the n_4 values greater than 4 the Kappa coefficient is started to decrease. Based on this graph, the maximum Kappa coefficient of clustering for DIARETDB1 dataset is about 0.83 with 2D kernel function (Fig. 7(b)) and about 1 for 3D kernel function (Fig. 7(d)).

The proposed modified AlexNet with 3D kernel function has a little better performance than the 2D kernel function in FROC criterion. This result is because the SVDD with 3D kernel function (for $n_1 = 0.1$, $n_1 = 2$ and $n_2 = 0.5$) can efficiently use the information of the input feature space to remove the false-positive candidates. Also, the optimum kernels are obtained as $n_1 = 0.8$ and $n_2 = 4$ for the Navid-Didegan as well as DIARETDB1.

5.6. Comparison to other clustering algorithms (DIARETDB1 dataset)

In this section, the results of the proposed algorithm are compared with the K-Means, subtractive clustering and FCM algorithms as the most frequently used clustering algorithms for the balanced DIARETDB1 dataset as mentioned in Section 5.1. We also use the nonlinear transformation in (11) for 2D feature maps and (12) for 3D feature maps for all classification methods, keeping the results comparable. Therefore, we evaluate and compare the performance of the proposed modified AlexNet with four different classification algorithms in both 2D or 3D proposed nonlinear transformation over the input space (i.e. Modified AlexNet-K-Means, Modified AlexNet-Subtractive, Modified AlexNet-FCM, Modified AlexNet-SVDD). In this scenario, the optimum range of influence for subtractive clustering algorithm has been obtained and implemented. Our experimental results, show that the modified AlexNet using the FCM and Subtractive as two classifiers (Modified AlexNet-FCM, Modified AlexNet-Subtractive) have the comparable performance to the proposed algorithm (Modified AlexNet-SVDD) only in sensitive criterion. Thus, for keeping the results comparable, the optimum parameters for the FCM classification algorithm would be obtained. The Modified AlexNet-FCM has two different parameters which affect the classification results, the number of training samples and the fuzzifier parameter (m), [70]. For our dataset with more similar classes, increasing in the fuzzifier value will cause a decrease in classification result. Our experimental results show, the value $m = 5$ for this parameter seems to be an appropriate value in our proposed methodology. In some paper, the different initial cluster centers of FCM are considered to be study [55], but we let that the initial cluster centers are chosen in a completely random way. In this experimental scenario, all the available samples are split in two training and test sets. Number of training data is set to 70% (i.e., 33 samples per class for each DR and NoDR classes for the balanced DIAREDTB1 dataset) and 30% of the remained data is selected as the test dataset and a comparison based on the Monte-Carlo simulation is made to evaluate the proposed method [71]. In each run, the random testing and training sets are kept the same for all methods, keeping the results comparable. Fig. 8 shows the classification results obtained by the modified AlexNet using K-Means, subtractive, FCM, and SVDD with 2D (the left column) and 3D (the right column) kernel functions. Since some of the clustering algorithms are sensitive to the initial clusters, the best run or initial conditions are illustrated for all methodologies in Fig. 8. In FCM clustering, data points can potentially belong to both clusters. As one can see in Fig. 8(c), the purple stars belong to both clusters DR and NoDR. This means that a given retinal image can be DR to a certain degree as well as NoDR to a certain degree. So, the blue stars in Fig. 8 belong completely to the NoDR, the red pluses belong completely to the DR cluster and the purple color in Fig. 8(c), shows that a sample point belongs to both clusters to a certain degree. Because of the binary classification, a

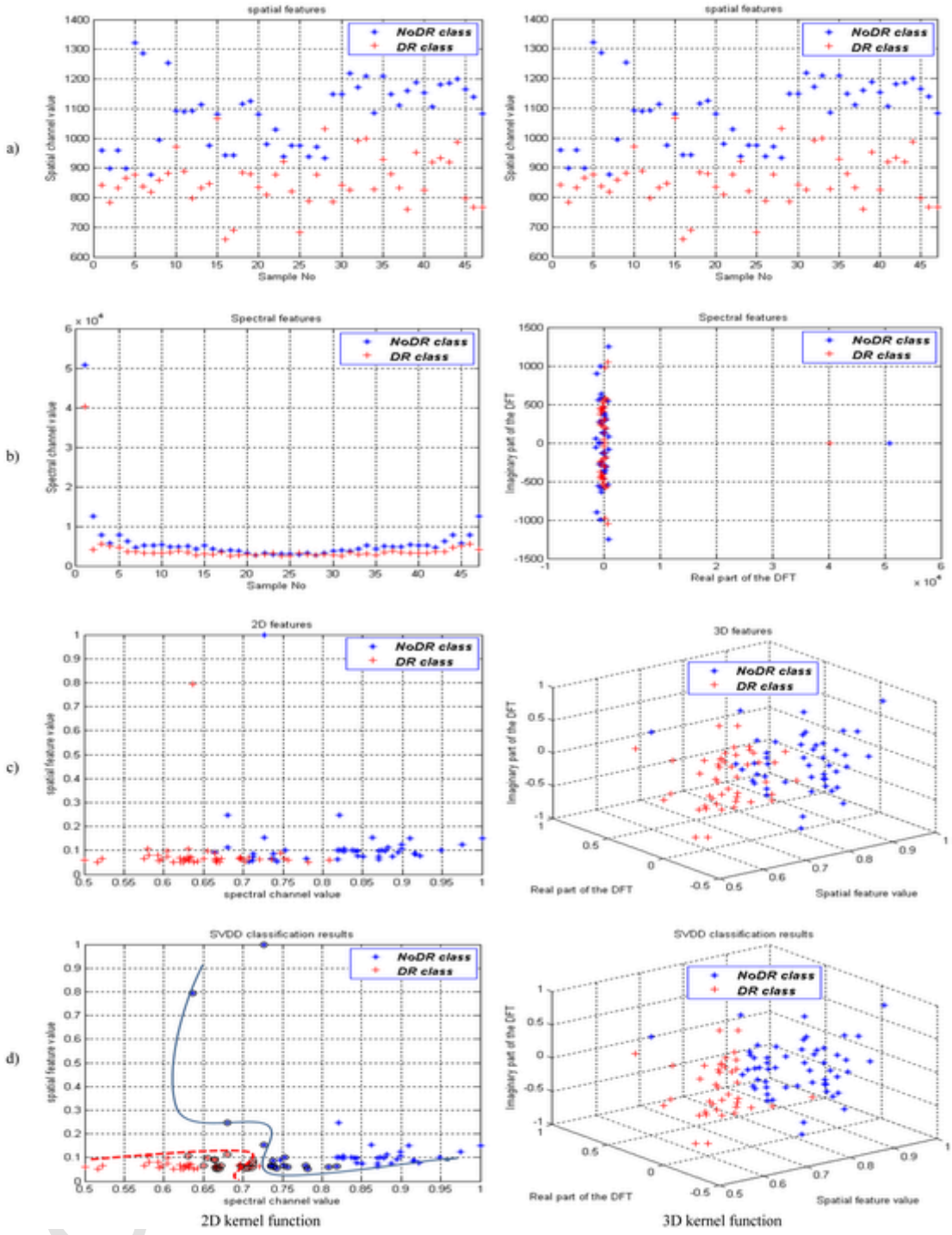

Fig. 5. (a) The spatial pathway output with 2D (the left column) and 3D (the right column), containing 47 sample data as DR the red stars and 47 sample data as NoDR the blue stars, (b) The spectral pathway output, (c) The SVDD input data with degrees $n_1 = n_2 = 1$ of the kernel functions (d) The classification results obtained by the modified AlexNet-SVDD, (Navid-Didegan dataset).

sample data is only considered as DR or NoDR based on which of its membership grade (DR or NoDR) is higher. Therefore, in this figure all the stars belong to NoDR while the pluses belong to DR cluster, whatever its color.

In Fig. 9(a)–(b) we show the results in terms of FROC in function of each classification methods. The Modified AlexNet 2D and 3D with SVDD classification method proves itself to be more efficient in DR and NoDR retinal images detection.

To show the performance of the different classification methods, the performance indices are used based on 250 different runs of Monte-Carlo simulation for balanced DIARETDB1 dataset as mentioned in Section 5.1. Table 4 exhibits the median values of TP, FP, TN, FN (Half of the answers lie below the median and half lie above the median) and 25th-75th percentile is presented in parentheses. The 25th percentile is the value at which 25% of the answers lie below that value, and 75% of the answers lie above that value. Also, the mean values of accuracy, precision, sensitivity, specificity and Kappa coefficient criterions are reported with their standard deviations in parentheses. One feature of the modified AlexNet-SVDD structure is that the Fc7 of the standard AlexNet is forwarded to two separate pathways for better classification.

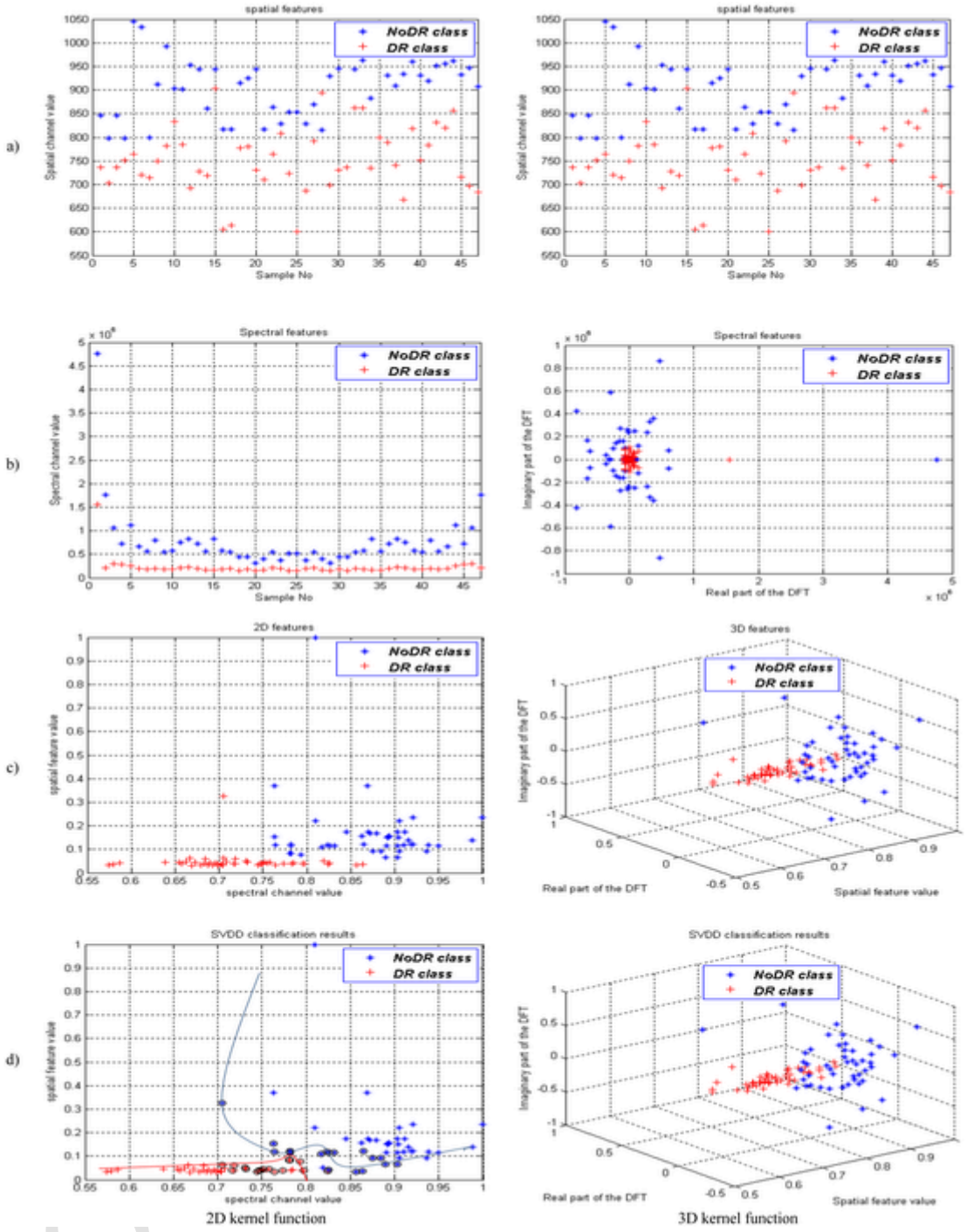

Fig. 6. (a) The spatial pathway output with 2D (the left column) and 3D (the right column), containing 47 sample data as DR (the red stars) and 47 sample data as NoDR (the blue stars) (b) The spectral pathway output (c) The SVDD input data with degrees $n_1 = 0.8, n_2 = 4$ of the kernel functions (d) The classification results obtained by the modified AlexNet-SVDD, (Navid-Didegan dataset).

Similar methodology was used in GoogLeNet where the auxiliary classifiers connected to intermediate layers. To have a deep understanding of the performance of the modified AlexNet-SVDD due to its two-pathway structure, we also measured the performance of the two pathways model separately (i.e. the spatial and spectral pathway models). Table 4 also shows the numerical results of modified AlexNet-SVDD in spatial and spectral pathway structures. Since the two pathways contain complementary information of a retinal fundus image, the two-pathway structure performs much better than the individual pathways. To show that the proposed modified AlexNet-SVDD method substantially improves the automatic screening of DR, the last row of Table 4 shows the numerical results for standard AlexNet structure performance without any classifier algorithm. As is evident, the proposed modified AlexNet-SVDD (2D and 3D feature map) outperforms the standard AlexNet structure in all terms of accuracies. Based on Monte-Carlo simulation on 250 different runs (i.e. 250 different initial training and testing sets), we have obtained a mean accuracy 98.10%, a mean precision 98.14%, a mean specificity 98.05% and Kappa coefficient 0.96 by using the modified AlexNet-SVDD (2D) and a mean sensitivity 98.29% by using the proposed modified AlexNet-SVDD (3D) method. The sensitivity value obtained by the proposed algorithm is comparable with the

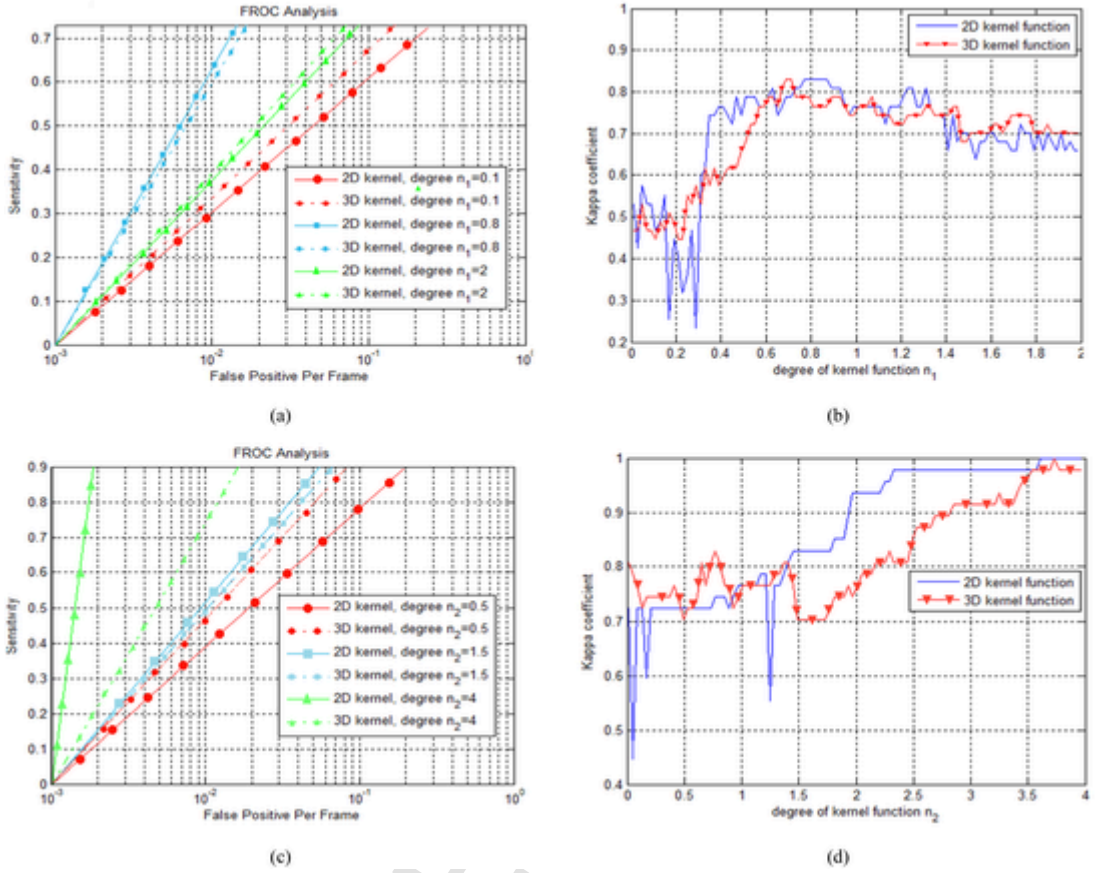

Fig. 7. FROC and Kappa coefficient analysis for the proposed algorithm in DIARETDB1 dataset (a–c) FROC for different values of the degree of SVDD's kernel function (i.e., n_1 and n_2) (b–d) the Kappa coefficient for different values of the degree of SVDD's kernel function.

values obtained by Modified AlexNet-FCM and Modified AlexNet-Subtractive methodologies.

5.7. Comparison with the most related works (based on DIARETDB1 dataset)

Here, we compare the proposed method performance with the most related works in the literatures. There are the works of Franklin and Rajan [62], Adarsh and Jeyakumari [63], Sinthanayothin et al. [72], Wang et al. [73], Walter et al. [74], Osareh et al. [14]. Table 5 summarizes all the results in terms of diagnostic accuracies. Multiple approaches in the literature using various algorithms to implement automatic screening of DR, all of them using only two levels of evaluation including per-lesion and per-image evaluations. Per-lesion evaluation means that the method performance was analyzed in detecting every single lesion such as exudates, microaneurysms or hemorrhages. Typically, these methods use hand-crafted features of retinal images for training their systems. When only a diagnosis was provided for each image, the methodology was evaluated a per-image basis which is more interesting from a screening point of view (discriminating images as DR or NoDR).

In [62] an algorithm to detect the presence of exudates by using an artificial neural network has been presented. The proposed approach is based on feature extraction and clustering technique. The paper is based on pre-lesion exudates detection as a symptoms of DR, while there are some other important symptoms of DR. They have evaluated their works by using 57 color retinal images of DIARETDB1 which contains 5137 objects for training and testing the neural network. As summarized in Table 5, Franklin and Rajan reported a mean accuracy 99.7, mean sensitivity 96.3 and mean specificity 99.8. As during the screening stage, the goal is to save time for the physician while reducing the number of test images and labeling the ones which are suspicious of DR as well as the ones which are close. Having a high recall or sensitivity score means that most of the patients will be screened correctly and their images will be labeled for ophthalmologist consideration. Although, Franklin and Rajan have obtained a better performance indices in terms of accuracy and specificity than the performance of our methodology, but the sensitivity performance index of our algorithm is higher than their results.

Adarsh and Jeyakumari [63] also used feature extraction technique to produce an automated diagnosis for DR through the detection of retinal blood vessels, exudates and microaneurysms. This method achieved the mean scores of 95.3% accuracy, 90.6% sensitivity and 93.65% specificity on the DIARETDB1. Wang et al. reported an image-based accuracy of 100% sensitivity and 70% specificity without any assessment in terms of lesion-based or pixel-resolution accuracies.

An approach based on multi-scale correlation filtering (MSCF) was presented which consists on microaneurysm candidate detection and classification [59]. The experimental results have been evaluated on only 11 images of DIARETDB1 dataset by FROC plots and sensitivity measurement. Only numerical score has been reported is the average false positive per image for DIARETDB1 equal to 0.713.

Another novel method, called Dynamic Shape Feature (DSF) for automatic detection of both microaneurysms and hemorrhages has been evaluated pre-lesion and per-image using six databases including DIARETDB1 [58]. Our proposed method with the modified AlexNet-SVDD classifier trained on the same dataset (DIARETDB1), has achieved better FROC performance than the DFS based method in [58].

Utilizing a modified pre-trained CNN for classification makes Graphics Processing Unit (GPU) and external memories unnecessary. The software package that was used in this paper was Matlab 2017b. All

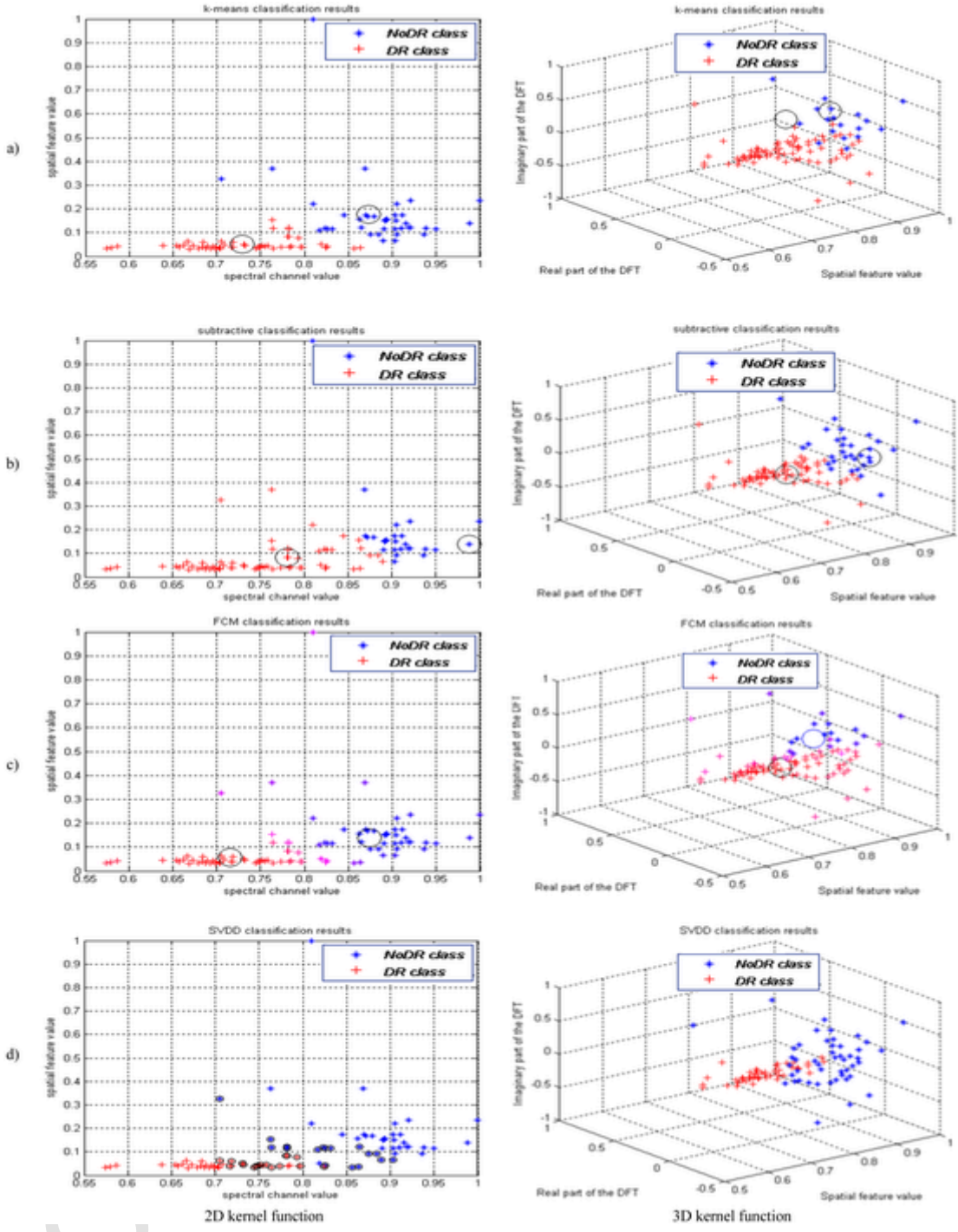

Fig. 8. The classification results obtained by the modified AlexNet using K-Means, Subtractive, FCM, and SVDD with 2D feature maps (the left column) and 3D feature maps (the right column) in Navid-Didegan dataset (a) K-Means clustering method (b) Subtractive clustering algorithm (c) FCM clustering (d) SVDD classification method.

the steps of the proposed method were done by an Intel i7 core CPU, with 8 GB memory, which is considerably advantageous comparing to common CNN training hardware requirements.

6. Conclusions and future work

In this paper, a deep modified CNN learning algorithm was proposed in the diagnosis of DR and NoDR retinal images. The reasoning behind modification of a pre-trained CNN network is to avoid the time complexity of the training process for the convolutional systems. This novel algorithm adapts a modified AlexNet with two pathways for retinal image recognition, which extract spatial and spectral features of images, respectively. These two kind of features contained complementary information of a specific retinal image. The experimental results have shown that, the fusion of the obtained frequency domain features with the spatial features, can introduce an increase in the detection accuracy. For any two or three dimensional feature maps obtained by the proposed algorithm, the two-pathway architecture performed much better than individual pathways. Although our multiple simulation results have shown that the propose methodology is sensitive to the training and testing sizes, but it is not sensitive to the randomly selected training and testing sets.

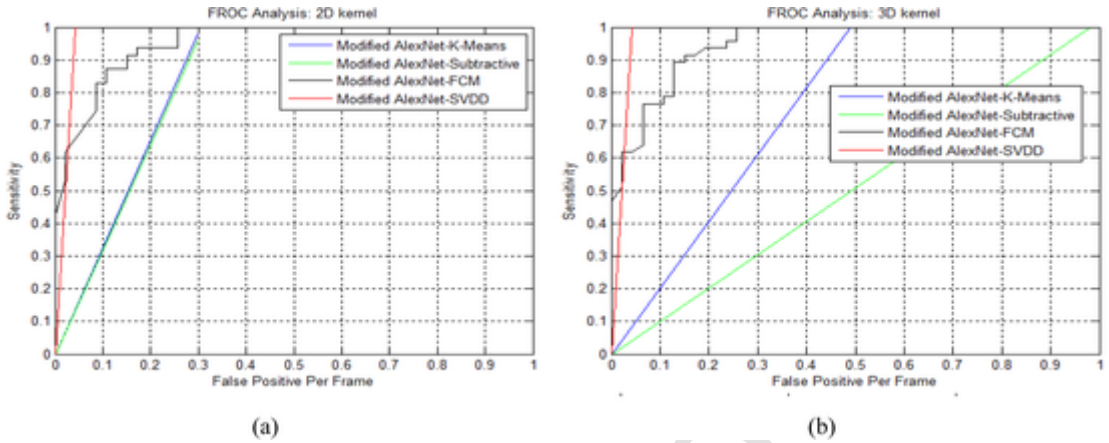

Fig. 9. The results in terms of FROC in function of each classification methods, (a) 2D kernel function analysis (b) 3D kernel function analysis.

Table 4
Performance indices for balanced DIARETDB1 dataset obtained for 250 different Monte-Carlo runs, using different 2D, 3D kernel functions and different classification methods.

Method	Feature space	TP	FP	TN	FN	Accuracy	Precision	Sensitivity	Specificity	κ
Modified AlexNet-K means	2D	35 (32-38)	20 (16-29)	27 (18-31)	12 (9-15)	64.03 (5.95)	63.63 (10.03)	73.44 (9.28)	54.61 (17.21)	28.06 (11.91)
Modified AlexNet-Subtractive	2D	46 (46-46)	16 (15-16)	31 (31-32)	1 (1-1)	81.64 (2.98)	74.31 (2.09)	96.72 (4.77)	66.57 (3.20)	63.29 (5.96)
Modified AlexNet-FCM	2D	45 (45-45)	14 (13-14)	33 (33-34)	2 (2-2)	83.15 (1.74)	76.58 (1.63)	95.74 (1.36)	70.74 (2.45)	66.31 (3.48)
Modified AlexNet—SVDD	**2D**	**46 (45-47)**	**1 (0-1)**	**46 (46-47)**	**0 (0-2)**	**98.10 (1.31)**	**98.14 (2.21)**	**98.12 (1.93)**	**98.05 (2.32)**	**96.21 (2.63)**
Modified AlexNet-K means	3D	30 (28-43)	9 (6-39)	38 (8-41)	17 (4-19)	65.39 (9.2)	69.26 (14.47)	73.51 (16.86)	57.27 (34.02)	30.87 (18.43)
Modified AlexNet-Subtractive	3D	46 (45-46)	16 (16-20	31 (27-31)	1 (1-2)	78.78 (3.84)	71.28 (3.95)	97.27 (1.59)	60.29 (8.59)	57.57 (7.68)
Modified AlexNet-FCM	3D	45 (42-47)	26 (24-29)	21 (18-23)	2 (0-5)	68.91 (2.54)	62.76 (1.97)	93.40 (7.38)	44.42 (6.66)	37.82 (5.08)
Modified AlexNet—SVDD	**3D**	**46 (46-46.75)**	**1 (0.25-1.75)**	**46 (45.25-46.75)**	**1 (0.25-1)**	**97.97 (1.00)**	**96.55 (1.92)**	**98.29 (1.27)**	**97.65 (2.00)**	**94.95 (2.00)**
Modified AlexNet—SVDD	Spatial pathway	35 (35-36)	3 (3-3)	44 (44-44)	12 (11-12)	84.42 (1.24)	91.69 (2.02)	75.82 (3.33)	93.02 (3.02)	68.85 (2.49)
Modified AlexNet—SVDD	Spectral pathway	47 (46-47)	8 (7-8)	39 (39-40)	0 (0-1)	91.37 (0.67)	86.11 (1.17)	98.7 (1.87)	84.04 (1.71)	82.74 (1.34)
Standard AlexNet without classifier	Binary	39 (38-40)	1 (0-2)	46 (45-47)	8 (7-9)	88.38 (5.4)	96.46 (3.72)	80.00 (11.62)	96.76 (3.65)	76.76 (10.81)

Table 5
Comparison of our proposed method against previous techniques based on DIARETDB1 dataset.

Performance indices / *Method*	*Mean accuracy*	*Mean sensitivity*	*Mean specificity*
Franklin and Rajan 2014	**99.7**	96.3	**99.8**
Adarsh et al. 2013	95.3	90.6	93.65
Sinthanayothin 2002	----	88.5	99.7
Wang et al. 2000	----	100	70
Walter et al. 2002	----	92.8	92.4
Osareh et al. 2009	----	93.5	92.1
Proposed method	98.10	**98.12**	98.05

In addition, the algorithm uses the SVDD algorithm with suitable kernel functions to classify the CNN data. A comparative study on different kernel parameters and different classifiers were presented. The comparative study was performed to demonstrate the effect of different degrees of kernel functions on the performance in diagnosing screening diabetic retinopathy cases. Also, to demonstrate the performance of the proposed modified AlexNet-SVDD in clinical applications, Navid-Didegan dataset including 94 fundus images and balanced DIARETDB1 database, were applied and evaluated. The results of the study can be helpful to determine the proposed architecture for screening diabetic retinopathy cases in real clinical cases.

As shown in Table 4, the modified AlexNet-SVDD had the best performance results among the other classification methods, considering all performance indices. The key to success of the modified AlexNet-

SVDD model is the using of the complementary information obtained by the spectral domain, which represent the significant correlations between spatial features such as microaneurysms, hemorrhages, neovascularization, and exudates on retinal images. From the clinical usage perspective of the automatic DR screening approach the sensitivity or recall, which demonstrates the correctness of DR diagnosis, is the most important factors. As during the screening stage, the goal is to save time for the physician while reducing the number of test images and labeling the ones which are suspicious of DR as well as the ones which are close. Having a high recall or sensitivity score means that most of the patients will be screened correctly and their images will be labeled for ophthalmologist consideration.

The proposed algorithm can obtain acceptable results and handle the higher level of dimensionality of the CNN output data by using the proposed kernel functions. Addition of some retinal image characteristics, along with those from CNN feature maps in this work, is expected to further improve the proposed structure and is suggested for further work. The proposed scheme in this paper is not limited to the AlexNet architecture, so the other existing deep learning CNN can be modified and used is also part of further extension of this work. The results reported in this paper can be utilized as a starting point for further research and enhance the accuracy of the DR screening approaches using CNN, while being acceptable from the practical standpoint.

Uncited References

Chen et al. , Ginneken et al. , Liu and Tang , Prasoon et al. , Zheng et al.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The author would like to thank Dr. M. Ansari and Khatam-al-Anbia eye hospital employees for making available the data sets used in this paper. The author would also like to thank Dr. Hamid Khakshur and Navid-Didegan Clinic employees for their participation in preparing and labeling the retinal images to use in this study.

References

[1] N. Congdon, Y. Zheng, M. He, The worldwide epidemic of diabetic retinopathy, Indian J. Ophthalmol. 60 (5) (2012) 428.
[2] B. Antal, A. Hajdu, An ensemble-based system for microaneurysm detection and diabetic retinopathy grading, IEEE Trans. Biomed. Eng. 59 (6) (2012) 1720–1726.
[3] V.V. Kamadi, A.R. Allam, S.M. Thummala, V.N.R. P, A computational intelligence technique for the effective diagnosis of diabetic patients using principal component analysis (PCA) and modified fuzzy SLIQ decision tree approach, Appl. Soft Comput. 49 (2016) 137–145.
[4] A.F.M. Hani, H.A. Nugroho, Gaussian bayes classifier for medical diagnosis and grading: Application to diabetic retinopathy, in: Proc. IEEE Conf. Biomed. Eng. Sci. (EMBS), 2010.
[5] Y. Pan, W. Huang, Z. Lin, W. Zhu, J. Zhou, J. Wong, Z. Ding, Brain tumor grading based on neural networks and convolutional neural networks, in: Proc. IEEE Conf.Eng. Med. Biol. Society (EMBS), 2015, https://doi.org/10.1109/embc.2015.7318458.
[6] B. Menze, M. Reyes, k. V. Leemput, The multimodal brain tumor image segmentation benchmark (brats), IEEE Trans. Med. Imaging 34 (10) (2015) 1993–2024.
[7] N. Tajbakhsh, J.Y. Shin, S.R. Gurudu, R.T. Hurst, C.B. Kendall, M.B. Gotway, J. Liang, Convolutional neural networks for medical image analysis: fine tuning or full training? IEEE Trans. Med. Imaging 35 (5) (2016) 1299–1312, https://doi.org/10.1109/tmi.2016.2535302.
[8] H.H. Vo, A. Verma, New deep neural nets for fine-grained diabetic retinopathy recognition on hybrid color space, in: Proc. IEEE Int. Sym. Multimedia, 2016.
[9] A.S. Razavian, H. Azizpour, J. Sullivan, S. Carlsson, CNN features off-the-shelf: An astounding baseline for recognition, in: Proc. IEEE Conf. Comput. Vis. Pattern Recognit. Workshops, 2014, pp. 512–519.
[10] Margeta. j, A. Criminisi, R.C. Lozoya, D.C. Lee, N. Ayache, Fine-tuned convolutional neural nets for cardiac MRI acquisition plane recognition, Computer Methods Biomechan. Biomed. Eng.: Imag. Visual 5 (5) (2015) 1–11.
[11] D. Doshi, A. Shenoy, D. Sidhpura, P. Gharpure, Diabetic retinopathy detection using deep convolutional neural networks, in: Proc Int. Conf. Comput. Analy. Security Trends (CAST), 2016, pp. 261–266.
[12] D. Nie, L. Wang, E. Adeli, C. Lao, W. Lin, D. Shen, 3-d fully convolutional networks for multimodal isointense infant brain image segmentation, IEEE Trans. Cybern. 99 (2018) 1–14.
[13] M. Niemeijer, B.van. Ginneken, M.J. Cree, A. Mizutani, G. Quellec, C.I. Sanchez, B. Zhang, R. Hornero, M. Lamard, C. Muramatsu, X. Wu, G. Cazuguel, J. You, A. Mayo, Y.Qin.Li. Hatanaka, B. Cochener, C. Roux, F. Karray, M. Garcia, H. Fujita, M.D. Abramoff, Retinopathy online challenge: Automatic detection of microaneurysms in digital color fundus photographs, IEEE Trans. Med. Imaging 29 (1) (2010) 185–195, https://doi.org/10.1109/tmi.2009.2033909.
[14] A. Osareh, B. Shadgar, R. Markham, A computational-intelligence-based approach for detection of exudates in diabetic retinopathy images, IEEE Trans. Inf. Technol. Biomed. 4 (13) (2009) 535–545.
[15] A.P. Bhatkar, G.U. Kharat, Detection of diabetic retinopathy in retinal images using MLP classifier, in: Proc. Int. Symp. Nanoelectronic and Inf. Syst., 2015.
[16] R. Priya, P. Aruna, Diagnosis of diabetic retinopathy using machine learning techniques, ICTACT J. Soft Comput. 3 (4) (2013) 563–575.
[17] K. Saranya, B. Ramasubramanian, S. Kaja, G. Mohideen, A novel approach for the detection of new vessels in the retinal images for screening diabetic retinopathy, in: Proc. Int. Conf. Commun. Signal Process, 2012.
[18] C. Wang, X. Zhang, H. Yang, J. Bu, A pixel-based color image segmentation using support vector machine and fuzzy C–means, Neural Netw. 33 (2012) 148–159.
[19] H. Selvaraj, S.T. Selvi, D. Selvathi, L. Gewali, Brain MRI slices classification using least squares support vector machine, Int. J. Intell. Computer Med. Sci. Image Process 1 (1) (2007) 21–33.
[20] P. Anbeek, K.L. Vincken, G.S. Bochove, M.J. Osch, J. Grond, Probabilistic segmentation of brain tissue in MR imaging, Neuroimage 27 (4) (2005) 795–804.
[21] C.A. Cocosco, A.P. Zijdenbos, A.C. Evans, A fully automatic and robust brain MRI tissue classification method, Med. Image Anal. 7 (4) (2003) 513–527.
[22] M. Khalid, N. Pal, K. Arora, Clustering of image data using K-means and fuzzy K-means, Int. J. Adv. Comp. Sci. Appl. 5 (7) (2014) 160–163.
[23] L. Lalaoui, T. Mohamadi, A. Djaalab, H. Abdelghani, A modified expectation of maximization method and its application to image segmentation, Current Med. Imag. Reviews 11 (2) (2015) 132–137.
[24] S. Chiu, Fuzzy model identification based on cluster estimation, J. Intell. Fuzzy Syst. 2 (3) (1994) 267–278.
[25] J.C. Bezdek, R. Ehrlich, W. Full, FCM: The fuzzy C-Meanss clustering algorithm, Comput. Geosci. 10 (2–3) (1984) 191–203.
[26] C. Singh, S.K. Ranade, K. Singh, Invariant moments and transform-based unbiased nonlocal means for denoising of MR images, Biomed. Signal Process. Control 30 (2016) 13–24.
[27] J. Wang, J. Kong, Y. Lu, M. Qi, B. Zhang, A modified FCM algorithm for MRI brain image segmentation using both local and non-local spatial constraints, Comput. Med. Imag. Graph 32 (8) (2008) 685–698.
[28] M. Gong, Y. Liang, J. Shi, W. Ma, J. Ma, Fuzzy C-means clustering with local information and kernel metric for image segmentation, IEEE Trans. Image Process. 22 (2) (2013) 573–584.
[29] S. Mohammadian, A. Karsaz, Y. M. Roshan, A comparative analysis of classification algorithms in diabetic retinopathy. In Proc. Int. Conf. on Software Engineering and Knowledge Engineering, 2017a.
[30] S. Hussain, S. Anwar, M. Majid, Segmentation of glioma tumors in brain using deep convolutional neural network, Neurocomputing 282 (2018) 248–261.
[31] S. Wang, Y. Yin, G. Cao, B. Wei, Y. Zheng, G. Yang, Hierarchical retinal blood vessel segmentation based on feature and ensemble learning, Neurocomputing 149 (2015) 708–717.
[32] H. Pratt, F. Coenen, D. Broadbent, S. Harding, Y. Zheng, Convolutional neural networks for diabetic retinopathy, in: Proc. Conf. Medical Imag. Understanding and Analysis (MIUA), 2016, https://doi.org/10.1016/j.procs.2016.07.014.
[33] S. Mohammadian, A. Karsaz, Y. M. Roshan, Comparative study of fine-tuning of pre-trained convolutional neural networks for diabetic retinopathy screening, in: 2nd Int. Conf. Biomed. Eng. Iran, 2017b, in preparation.
[34] D. Erhan, P.A. Manzagol, Y. Bengio, S. Bengio, P. Vincent, The difficulty of training deep architectures and the effect of unsupervised pre-training, in: Proc. Int. Conf. Artif. Intell. Stat., 2009, pp. 153–160.
[35] H. Chen, D. Ni, J. Qin, S. Li, X. Yang, T. Wang, P.A. Heng, Standard plane localization in fetal ultrasound via domain transferred deep neural networks, IEEE J. Biomed. Health Inf. 19 (5) (2015) 1627–1636.
[36] G. Carneiro, J. Nascimento, A. Bradley, Unregistered multi view mammogram analysis with pre-trained deep learning models, in: N. Navab, J. Hornegger, W.M. Wells, A.F. Frangi (Eds.), Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, vol. 9351, LNCS, 2015, pp. 652–660.
[37] A. Krizhevsky, Learning multiple layers of features from tiny images (Master's Thesis), Dept. of Comp. Sci. University of Toronto, 2009.
[38] A. Krizhevsky, I. Sutskever, G.E. Hinton, ImageNet classification with deep convolutional neural networks, in: Proc. Conf. Neural Inf. Process. Syst. (NIPS), 2012.
[39] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, A. Rabinovich, Going deeper with convolutions, in: IEEE Conf. on Computer Vision and Pattern Recognition (CVPR), 2015.
[40] K. Simonyan, A. Zisserman,
[41] H. Shin, H. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, R. Summers,

Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning, IEEE Trans. Med. Imaging 35 (5) (2016) 1285–1298.
[42] Y. Bar, I. Diamant, L. Wolf, H. Greenspan, Deep learning with non-medical training used for chest pathology identification, in: Proc. SPIE Med. Imag, 2015, 94140V.
[43] D. Hubel, T. Wiesel, Receptive fields of single neurons in the cat's striate cortex, J. Physiol. 148 (3) (1959) 574–591, https://doi.org/10.1113/jphysiol.1959.sp006308.
[44] V. Nair, G.E. Hinton, Rectified linear units improve restricted boltzmann machines, in: Proc. Int. Conf. Mach. Learn. (ICML-10), 2010, 807–814.
[45] Y. Lecun, Y. Bengio, G.E. Hinton, Deep learning, Nature 521 (7553) (2015) 436–444.
[46] G.E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, R.R. Salakhutdinov, Improving neural networks by preventing coadaptation of feature detectors, Comput. Sci. 3 (4) (2012) 212–223.
[47] Y.M.S. Reddy, R.E. Ravindran, K.H. Kishore, Diabetic retinopathy through retinal image analysis: A review, Int. J. Eng. Technol. 7 (1–5) (2017) 19.
[48] L.G. Nyul, Retinal image analysis for automated glaucoma risk evaluation, in: MIPPR: Med. Imag. Parallel Process. Images, and Optimization Techniques - Invited Paper, vol. 7497, 2009, pp. 891–899.
[49] M. Madhusoodanan, J. Cheriyan, Face recognition using TSF model and DWT based multilevel illumination normalization, Int. J. Sci. Res. (IJSR) 5 (2) (2016) 847–854.
[50] M. Motamedi, P. Gysel, V. Akella, S. Ghiasi, Design space exploration of FPGA-based deep convolutional neural network, in: proc. IEEE conf. Design Automation, 2016.
[51] W. Zhang, R. Li, H. Deng, L. Wang, W. Lin, S. Ji, D. Shen, Deep convolutional neural networks for multi-modality isointense infant brain image segmentation, NeuroImage 108 (2015) 214–224.
[52] C. Cortes, V. Vapnik, Support-vector networks, Mach. Learn. 20 (3) (1995) 273–297.
[53] V.N. Vapnik, The Nature of Statistical Learning Theory, Multimedia Tools and appl (1995) 1–19.
[54] D.M.J. Tax, R.P.W. Duin, Support vector domain description, Pattern Recognit. Lett. 20 (11–13) (1999) 1191–1199.
[55] S. Niazmardi, S. Homayouni, A. Safari, An improved FCM algorithm based on the SVDD for unsupervised hyperspectral data classification, IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens 6 (2) (2013) 831–839.
[56] T. Sun, Y. Wang, J. Yang, X. Hu, Convolution neural networks with two pathways for image style recognition, IEEE Trans. Image Process. 26 (9) (2017) 4102–4113, https://doi.org/10.1109/tip.2017.2710631.
[57] T. Brosch, L.Y.W. Tang, Y. Yoo, D.K.B. Li, A. Traboulsee, R. Tam, Deep 3D convolutional encoder networks with shortcuts for multiscale feature integration applied to multiple sclerosis lesion segmentation, IEEE Trans. Med. Imaging 35 (5) (2016) 1229–1239, https://doi.org/10.1109/TMI.2016.2528821.
[58] L. Seoud, T. Hurtut, J. Chelbi, F. Cheriet, J. Langlois, Red lesion detection using dynamic shape features for diabetic retinopathy screening, IEEE Trans. Med. Imaging 35 (4) (2016) 1116–1126.
[59] B. Zhang, X. Wu, J. You, Q. Li, F. Karray, Detection of microaneurysms using multi-scale correlation coefficients, Pattern Recognit. 43 (6) (2010) 2237–2248.
[60] W. Zhoul, C. Wu, D. Chen, Z. Wang, Y. Yi, W. Du, Automatic microaneurysm detection of diabetic retinopathy in fundus images, in: Proc. IEEE conf. Control and Decision (CCDC), 2017.
[61] T. Kauppi, J. Kämäräinen, L. Lensu, V. Kalesnykiene, I. Sorri, H. Uusitalo, H. Kälviäinen, Constructing benchmark databases and protocols for medical image analysis: Diabetic retinopathy, Comput. Math. Methods Med. 2013 (2013) 1–15.
[62] S. Franklin, S. Rajan, Diagnosis of diabetic retinopathy by employing image processing technique to detect exudates in retinal images, IET Image Process. 8 (10) (2014) 601–609.
[63] P. Adarsh, D. Jeyakumari, Multiclass SVM-based automated diagnosis of diabetic retinopathy, IEEE Int. Conf. Commun. Signal Process (2013) 206–210, https://doi.org/10.1109/iccsp.2013.6577044.
[64] G. Patry, G. Gauthier, B. Lay, J. Roger, D. Elie, ADCIS download third party: Messidor database, in: ADCIS S.a., 2016, 2016 Available: http://messidor.crihan.fr Accessed: Nov. 16, 2016.
[65] D.C. Edwards, M.A. Kupinski, C.E. Metz, R.M. Nishikawa, Maximum likelihood fitting of FROC curves under an initial-detection- and-candidate-analysis model, Med. Phys. 29 (12) (2002) 2861–2870.
[66] L.I. Kuncheva, A bound on Kappa-error diagrams for analysis of classifier ensembles, IEEE Trans. Knowl. Data Eng. 25 (3) (2011) 494–501.
[67] J. Cohen, A coefficient of agreement for nominal scales, Educ. Psychol. Meas. 20 (1) (1960) 37–46.
[68] E. Pasolli, F. Melgani, D. Tuia, F. Pacifici, W. Emery, Svm active learning approach for image classification using spatial information, IEEE Trans. Geosci. Remote Sens. 52 (4) (2014) 2217–2233.
[69] X. Wen, H. Zhang, Z. Jiang, Multiscale unsupervised segmentation of SAR imagery using the genetic algorithm, Sensors 8 (3) (2008) 1704–1711.
[70] J.C. Dunn, A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters, J. Cybernetics 3 (3) (1973) 32–57.
[71] I. Dimov, Monte Carlo Methods for Applied Scientists, World Scientific, Hackensack, N.J., 2008.
[72] C. Sinthanayothin, J.F. Boyce, T.H. Williamson, H.L. Cook, E. Mensah, S. Lal, D. Usher, Automated detection of diabetic retinopathy on digital fundus images, Diabetic Med. 19 (2) (2002) 105–112, https://doi.org/10.1046/j.1464-5491.2002.00613.
[73] H. Wang, W. Hsu, Goh. k, M. Lee, An effective approach to detect lesions in retinal images, in: Proc. IEEE Conf. on Comp. Vis. Pattern Recogn. Hilton Head Island, vol. 2, 2000, pp. 181–187, https://doi.org/10.1109/cvpr.2000.854775.
[74] T. Walter, J. Klein, P. Massin, F. Erginay, A contribution of image processing to the diagnosis of diabetic retinopathy-detection of exudates in color fundus images of the human retina, IEEE Trans. Med. Imaging 21 (10) (2002) 1236–1243, https://doi.org/10.1109/tmi.2002.806290.