

Diabetic retinopathy detection through novel tetragonal local octa patterns and extreme learning machines

Tahira Nazir^a, Aun Irtaza^a, Zain Shabbir^b, Ali Javed^c, Usman Akram^d,
Muhammad Tariq Mahmood^{e,*}

^a Department of Computer Science, University of Engineering and Technology Taxila, Pakistan

^b Department of Electrical, Electronics and Communication Engineering, University of Engineering and Technology, Lahore (Faisalabad Campus), Pakistan

^c Department of Software Engineering, University of Engineering and Technology Taxila, Pakistan

^d Department of Computer and Software Engineering, National University of Science and Technology, Islamabad, Pakistan

^e School of Computer Science and Engineering, Korea University of Technology and Education, 1600 Chungjeolno, Byeogchunmyun, 31253 Cheonan, South Korea

ARTICLE INFO

Keywords:

Diabetic retinopathy
Tetragonal local octa patterns
Extreme learning machines
Content based image retrieval

ABSTRACT

Diabetic retinopathy (DR) is an eye disease that victimize the people suffering from diabetes from many years. The severe form of DR results in form of the blindness that can initially be controlled by the DR-screening oriented treatment. The effective screening programs require the trained human resource that manually grade the fundus images to understand the severity of the disease. But due to the complexity of this process, and the insufficient number of the trained workers, the precise manual grading is an expensive process. The CAD-based solutions try to address these limitations but most of the existing DR detection systems are as evaluated over small sets and become ineffective when applied in real scenarios. Therefore, in this paper we proposed a novel technique to precisely detect the various stages of the DR by extending the research of the content-based image retrieval domain. To achieve the human-level performance over the large-scale DR-datasets (i.e. Kaggle-DR), the fundus images are represented by the novel tetragonal local octa pattern (T-LOP) features, that are then classified through the extreme learning machine (ELM). To justify the significance of the method, the proposed scheme is compared against several state-of-the-art methods including the deep learning-based methods over four DR-datasets of variational lengths (i.e. Kaggle-DR, DRIVE, Review-DB, STARE). The experimental results confirm the significance of the DR-detection scheme to serve as a stand-alone solution for providing the precise information of the severity of the DR in an efficient manner.

1. Introduction

Diabetic retinopathy (DR) is a disease that affects the blood vessels of the eye-retina of the diabetic patients. According to the world health organization (WHO), 347 million diabetic patients around the world are at a risk of developing the DR [1]. Only in USA, over 40% of 29.1 million diabetic patients suffers from the different stages of the DR. are suffering from diabetes and among them 40 to 45% are also the victim of DR over its different stages. The worst effect the DR causing is the blindness in its victims, which can be prevented if proper treatment is provided to the patients at its earlier stages, but it is very difficult to detect it at its initial stage as it exhibits the few signs [1]. The main reason of DR is the variations in blood vessels of the eye which causes the retinal disorder and that eventually leads toward the blindness. There are the main four levels of DR [2]:

1. Mild non-proliferative retinopathy: In the initial stage of DR, the small blood vessels inside the eye produce the tiny areas of balloon-like swelling which are known as micro-aneurysms. This micro-aneurysm may cause the leakage of fluid in retina.
2. Moderate non-proliferative retinopathy: In this level the micro-aneurysms [3] may cause the leakage of both fluid and blood from retina due to the hemorrhages.
3. Severe non-proliferative retinopathy: At this stage, the blood supply from numerous vessels is blocked which also affect the supply of blood to various areas of retina. These areas exude growth factors that causes the retina to produce the new blood vessels [4].
4. Proliferative diabetic retinopathy (PDR): Growth factors secreted by the retina trigger the proliferation of new blood vessels that grow along the inside surface of the retina and into the vitreous gel, the fluid that fills the eye. The new blood vessels are fragile that makes

* Corresponding author.

E-mail address: tariq@koreatech.ac.kr (M.T. Mahmood).

<https://doi.org/10.1016/j.artmed.2019.07.003>

Received 2 November 2018; Received in revised form 17 January 2019; Accepted 15 July 2019

0933-3657/ © 2019 Elsevier B.V. All rights reserved.

them more likely to leak and bleed. Accompanying scar tissue can contract and cause retinal detachment that leads to permanent vision loss [5].

To deal with the DR, digital fundal photography-based screening programs are used throughout the world. But, the cost-factor is still highlighted in these screening practices to enable them to be used for the larger populations [6]. Meanwhile, the rapid increase in the victims of DR is also making the capabilities of current DR screening programs questionable; as these programs are highly dependent over the manual grading that take the significant amount of time for every case. Therefore, to support the grading process in large populations, the research has focused on identifying the ways to detect the DR in an automatic way.

In literature, a lot of attention has been given to develop the automatic grading systems for DR. In [7], an automatic grading system for DR was proposed through the CNN and pattern matching detector. For learning the patterns of DR, the training images were annotated, and suspicious regions were highlighted in the training repository. The annotated images were then passed to the CNN which in turns predicts the respective DR severity. In [8], various features i.e. blood vessel density, blood vessel caliber, and size of the foveal avascular zone were extracted over optical coherence tomography angiography (OCTA) images to detect DR through the SVM classifier. In [9], CNN based algorithm was proposed for DR detection. The main contribution of the paper was Zoom-in-Net process that mimics the zoom-in process to highlight the suspicious regions and predicts the severity of disease. In [10], a hybrid statistical framework for DR detection was proposed that combines the probabilistic SVM-based kernel with scaled Dirichlet distributions. In [11], sliding band kernel was utilized for the segmentation of optic disk. In [12], infinite perimeter contour model was used for the segmentation of blood vessels for feature extraction and moderate retinopathy detection. In [13], segmentation of blood vessels was done through the morphological bit plane slicing and centerline detection by obtaining the first order derivative of the Gaussian kernel. Whereas in [14], the similar goal was achieved through the morphological operations. In [15] Quéllec et al. proposed a CNN-based method to create heatmaps to reflect pixels in images that play a role in the image-level predictions. The back-propagation method was used to produce the high quality heatmaps, and the method was assessed on the Kaggle-DR dataset. In [16] a basic framework was proposed for solving multiple-instance problems through CNN and the significance of the framework was demonstrated in the circumstance of diabetic retinopathy screening process through Kaggle-DR. In [17], the network with CNN structure and data augmentation was presented to identify the elaborated features involved in the DR-classification task by validating the method over 5000 Kaggle-DR images. In [18] Lei Zhou et al. improved the model of dense conditional random field (CRF) by discriminative unary features and combined it with CNN to represent long-range pixel interactions. The proposed method in [18] consisted of four steps, first was to eliminate the strong edges around the FOV and normalization was performed inside the FOV. In second step, CNN was trained by linear model. In Third step, the filters were applied to enhance the vessels. Finally, adopted the dense CRF model for segmentation of retinal vessels by the discriminative features of the enhanced image. In [19] field programmable gate arrays (FPGA) were proposed for the implementation of blood vessels segmentation, and the performance of the method was evaluated over the DRIVE and STARE datasets. In [20] the automated detection of red lesions or hemorrhages was presented through the gamma correction, and global thresholding techniques. In [21], the Probabilistic Graphical Model (PGM) was proposed for retinal vascular tree extraction. The proposed algorithm comprised of a two steps process i.e. tracking and classification. In [22], the author presented an improved method for vessel centerline tracking that was the combination of geometrical topology information and intensity distribution information. The method in [31] resolve the issues

of inaccuracy, inflexion and discontinuity of extracted centerline.

Most of the above approaches either detect specific conditions of DR or need noise-free images with clear anatomical structures for accurate detection in multi-class environment. However, in real-world scenarios the ultra-high-resolution images captured in perfect conditions are not always available; therefore, the actual application of these methods become limited. Moreover, usually these systems are as evaluated over the smaller repositories, consequently, in real application scenarios accuracy of these approaches further decreases.

In this paper we present a reliable method for DR-detection, which is independent of the anatomical extractions for classifying various stages of the DR. In the present work, we extend the feature extraction and multi-category recognition concepts of content-based image retrieval (CBIR) research for DR detection. For this, as a first step, we obtain the anatomy independent feature representation of the fundus images through the novel tetragonal local-octa patterns. The reason for the anatomy independent feature extraction is that the powerful image representations are generated in an efficient way, as we suppress the known requirements of the precise anatomical structure detection (i.e. blood vessel extraction and hemorrhage detection) that ultimately reduces the feature generation time. Once the feature extraction is done, we train the extreme learning machine (ELM) classifier over the training repository to learn the patterns of all known classes of the DR. ELMs are also unexplored classification methods in the DR detection research; whereas, the research is competing for more powerful classification approaches i.e. CNNs to achieve the reliable performance in the medical solutions. Afterwards, the feature vector of the query fundus image is classified as one of the DR class members. Additionally, we also return the associated case histories for making the informed decisions by computing the distance measure. The proposed method is then evaluated on several DR-repositories including the Kaggle-DR having 88,704 images representing more than 44,000 DR-patients. The DR-datasets used in existing approaches of DR detection [23] also enjoy the benefits of noise free instances with clear vascular structures over which DR detection is far simpler than the Kaggle-DR that has also noisy and low-resolution images. The similar phenomena described in this work can also be used to make relevant predictions regarding presence, manifestations, and severity of several retinal diseases in a deterministic manner. Here is the summary of our contributions.

- A novel technique is proposed to precisely detect the various stages of the DR by extending the research of the content-based image retrieval domain.
- For feature extraction, the concept of tetragonal local octal patterns (T-LOP) is introduced. T-LOP is a novel image representation scheme that particularly is designed for analyzing the fundus images that is a composition of blood vessels and textured lesions.
- To achieve the human-level performance over the large-scale DR-datasets (i.e. Kaggle-DR), the fundus images are represented by the proposed tetragonal local octa pattern (T-LOP) features, that are then classified through the extreme learning machine (ELM).

Rest of the paper is organized as follows: in Section 2 motivation behind using CBIR for DR detection and the proposed methodology of DR detection are discussed in detail. Experimental results are provided in Section 3. Finally, in Section 4, we conclude our work.

2. Proposed methodology

2.1. Motivation

The block diagram of the proposed scheme is shown in Fig. 1. In the proposed method, the DR detection is considered as a two-step process. In the first step we represent images in a robust way and train extreme learning machine (ELM) classifier over the feature repository. For feature extraction, we introduce the concept of tetragonal local octal

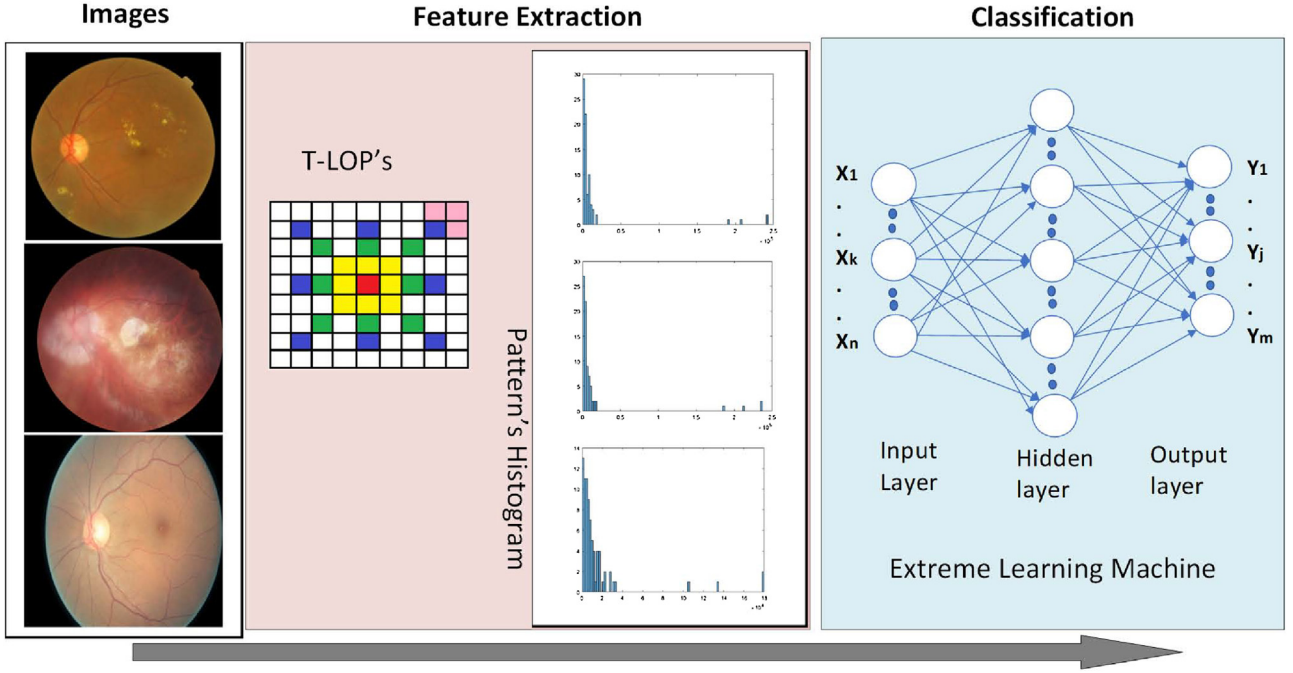


Fig. 1. The block diagram for the proposed methodology.

patterns (T-LOP). T-LOP is a novel image representation scheme that particularly is designed for analyzing the fundus images that is a composition of blood vessels and textured lesions. Once the training is done, we present a query image to the trained model that classifies the input fundus image as one of DR class members. For DR positive images we obtain the pathological information in the image by involving the query neighborhood and return the system output.

In the second step pathology detection is performed by incorporating the query neighborhood. In content-based image retrieval (CBIR) the query image contents are analyzed and semantically similar images are returned from large image collections [24]. Image semantics are determined based on the intrinsic features by estimating the objects comprising the query image or generating the global image representations [25]. The obtained feature representations are then compared against the feature representations of the repository images and top-matched images are returned as the semantic response of the system [25]. But as the image semantics are concerned with the analysis of images with dense objects, complex backgrounds, and overlapping semantics, thus, state-of-the-art in CBIR also utilize the supervised learning capabilities. Thus, in modern CBIR systems, image retrieval is considered as an image classification problem where semantic classifiers are defined to determine the underlying semantics of the images. Moreover, a guided approach in the form of feedback rounds by the system user can also be incorporated to ensure the precise image retrieval output [25].

The application of the concept of semantic classification in CBIR for DR detection is the main motivation behind the proposed scheme. In the proposed method, the DR detection occurs in the similar way as semantic response occurs in CBIR. In both domains, image representation occurs in the form of low level features as a first step that are then considered for the association with a particular class [23]. But as CBIR is intended to deal with far larger repositories and semantic categories compared to the DR detection, therefore, the efficiency in feature extraction and semantic association of CBIR can particularly be adopted to fulfill the need to deal with larger number of patients to facilitate the human graders. In CBIR relevant images are returned as semantic response of the CBIR system, whereas in DR-detection we can return the relevant case histories to support the informed decisions for determining the prescriptions.

2.2. Input data

Given an image-base having fundus images belonging to the categories of DR. The DR detection system generates the low-level feature representations of all the images and divides the image-base into l_i subsets (where $i = 1, 2, \dots, n$) known as bag of images (BOI). Each BOI contains $D_i = (D/n) * q$ image representations considered as positive examples and all other examples present in other BOIs as negative examples. The parameter q ranges from 0 to 1 determines the size of the positive BOI in respect to the i th category. The positive and negative examples are defined as follows:

$$D_i^+ = \{x_1, x_2, \dots, x_{D_i}\} \quad (1)$$

and

$$D_i^- = \{x_1, x_2, \dots, x_{D_i-D_i^+} \mid D_i^+ \cap D_i^- = \emptyset\} \quad (2)$$

where D_i^+ represents the positive examples, and D_i^- represents the negative examples in the i th BOI and x_i represents the feature vectors present in the corresponding BOI. For effective screening of DR, we need to ensure that the classifier learning occurs on the robust representation of fundus images. Hence, a more principled scheme of fundus image representation in form of feature vectors is required to ensure precise DR detection.

2.3. Tetragonal local octa patterns based feature extraction

In order to generate more discriminative features for fundus image representation, we introduce tetragonal local octa patterns (T-LOP) by extending the concepts provided in LTrPs [26]. A brief review of local binary patterns is provide in Appendix A. T-LOP descriptor considers the horizontal, vertical, and diagonal derivative directions and contains the uniform patterns of various diagonals to represent a region. Uniform patterns at one end reduce the feature vector length and allow us to generate image representations that are rotation invariant. Whereas, the diagonals of various sizes make the representation scale invariant. The main motivation behind the utilization of the diagonal derivative direction along with LTrPs is that: the LTrPs as consider only horizontal and vertical directions, therefore, when applied over the fundus images depict insensitivity against diagonally shaped blood vessels; hence do

not encodes them effectively. Therefore, to overcome this shortcoming we have extended the LTrPs by computing the four derivative directions instead of two directions. The n_{th} order T-LOP descriptor pre-computes the $(n - 1)$ th order derivatives along 0° , 45° , 90° , and 135° as suggested in [26] and then eight distinct values at g_c are computed as:

$$I_D^{n-1}(g_c) = \begin{cases} 1, I_{0^\circ}^{n-1}(g_c) \geq 0 \wedge I_{90^\circ}^{n-1}(g_c) \geq 0 \wedge (I_{45^\circ}^{n-1}(g_c) \vee I_{135^\circ}^{n-1}(g_c)) \geq 0 \\ 2, I_{0^\circ}^{n-1}(g_c) < 0 \wedge I_{90^\circ}^{n-1}(g_c) \geq 0 \wedge (I_{45^\circ}^{n-1}(g_c) \vee I_{135^\circ}^{n-1}(g_c)) \geq 0 \\ 3, I_{0^\circ}^{n-1}(g_c) < 0 \wedge I_{90^\circ}^{n-1}(g_c) < 0 \wedge (I_{45^\circ}^{n-1}(g_c) \vee I_{135^\circ}^{n-1}(g_c)) \geq 0 \\ 4, I_{0^\circ}^{n-1}(g_c) \geq 0 \wedge I_{90^\circ}^{n-1}(g_c) < 0 \wedge (I_{45^\circ}^{n-1}(g_c) \vee I_{135^\circ}^{n-1}(g_c)) \geq 0 \\ 5, I_{0^\circ}^{n-1}(g_c) \geq 0 \wedge I_{90^\circ}^{n-1}(g_c) \geq 0 \wedge (I_{45^\circ}^{n-1}(g_c) \vee I_{135^\circ}^{n-1}(g_c)) < 0 \\ 6, I_{0^\circ}^{n-1}(g_c) < 0 \wedge I_{90^\circ}^{n-1}(g_c) \geq 0 \wedge (I_{45^\circ}^{n-1}(g_c) \vee I_{135^\circ}^{n-1}(g_c)) < 0 \\ 7, I_{0^\circ}^{n-1}(g_c) < 0 \wedge I_{90^\circ}^{n-1}(g_c) < 0 \wedge (I_{45^\circ}^{n-1}(g_c) \vee I_{135^\circ}^{n-1}(g_c)) < 0 \\ 8, I_{0^\circ}^{n-1}(g_c) \geq 0 \wedge I_{90^\circ}^{n-1}(g_c) < 0 \wedge (I_{45^\circ}^{n-1}(g_c) \vee I_{135^\circ}^{n-1}(g_c)) < 0 \end{cases} \quad (3)$$

Once the direction of the referenced pixel is computed, the n th order T-LOPs can be defined as:

$$TLOP_P^n(g_c) = \{f_6(I_D^{n-1}(g_p), I_D^{n-1}(g_c))\} | p = 1, 2, \dots, P \quad (4)$$

where

$$f_6(I_D^{n-1}(g_p), I_D^{n-1}(g_c)) = \begin{cases} I_D^{n-1}(g_p), & \text{if } I_D^{n-1}(g_p) \neq I_D^{n-1}(g_c) \\ 0, & \text{else} \end{cases} \quad (5)$$

From Eq. (5), we get T-LOP code that is further divided into 7 binary patterns based on the direction of center pixel.

$$TLOP_P^n|_{\{\bar{D} | \forall D, \exists \neg I_D^{n-1}(g_c)\}} = f_7(TLOP_P^n(g_c))|_{\{\bar{D} | \forall D, \exists \neg I_D^{n-1}(g_c)\}} \quad (6)$$

$$f_7(TLOP_P^n(g_c))|_{\bar{D} \in \bar{D}} = \begin{cases} 1, & \text{if } TLOP_P^n(g_c) = \bar{D} \\ 0, & \text{else} \end{cases} \quad (7)$$

where $\bar{D}$ is the set of all quadrants(except the quadrant of the referenced pixel) and $\bar{D}$ is the one of the quadrants of $\bar{D}$. Afterwards, the T-LOP code can be generated as:

$$TLOP_P^n|_{\{\bar{D} | \forall D, \exists \neg I_D^{n-1}(g_c)\}} = \sum_{p=1}^P 2^{(p-1)*f_7}(TLOP_P^n(g_c))|_{\{\bar{D} | \forall D, \exists \neg I_D^{n-1}(g_c)\}} \quad (8)$$

For each of remaining 7 directions an 8-bit pattern is achieved, that is merged with magnitude pattern and it is calculated as:

$$MP = \sum_{p=1}^P 2^{n-1*f_8}(M_I(g_p) - M_I(g_c)) \quad (9)$$

where

$$M_I(g_p) = \sqrt{\sum \left((I_{0^\circ}^{n-1}(g_p))^2 + (I_{45^\circ}^{n-1}(g_p))^2 + (I_{90^\circ}^{n-1}(g_p))^2 + (I_{135^\circ}^{n-1}(g_p))^2 \right)} \quad (10)$$

Fig. 2 illustrates the possible Local Octa Patterns for a center pixel (shown in red color) in a segment of an image with neighbors (shown in green color). D(c), D(1), D(2) are the directions of center, first neighbor, and second neighbor pixels; whereas, M(c), M(1), M(2) are magnitudes of center, first neighbor, and second neighbor pixels and so on. If direction of neighbor pixel is that of center pixel then TLOP bit is taken as 0 otherwise it is same as the direction of neighbor pixel. So, the TLOP pattern comes out to be 41383183. This pattern is then subdivided into seven binary patterns; first pattern is achieved by replacing 2 with 1 and all other values with 0 in TLOP, second pattern is achieved by replacing 3 with 1 and other values with 0 and so on. Magnitude pattern is achieved by comparing magnitude values of neighbor pixels with center pixel and pattern achieved in this case is 11100001. These eight patterns are used to describe the texture of an image.

Higher order T-LOP patterns are expected to extract more

information, however it has been found in simulations that the second order patterns give the best results. Moreover, in the pictorial explanation, nearest neighbors are considered for pattern calculation. Neighbors at the diagonal 5 and 7 from the center pixel can also be considered for pattern calculation thus forming tetragonal octa patterns. Fig. 3 shows neighbors at different diagonal locations; red represents center pixel; yellow pixels are nearest neighbors at diagonal 3; green pixels are neighbor pixels at diagonal 5; blue pixels are neighbors at diagonal 7; whereas pink pixels represent horizontal, vertical and diagonal pixels of the blue cornered pixel.

By considering these observations, the feature vectors for robust DR classification are extracted and normalized as:

$$x = \frac{x_i - \mu_i}{\delta_i} \quad (11)$$

where x_i the corresponding feature in the feature vector is x , μ_i represents the mean, and δ_i represents the standard deviation of the feature vector.

2.4. Extreme learning machines based classifier

Once feature set is extracted, a classifier is learned by using extreme learning machines (ELM) [27]. It was originally proposed for single hidden layer feed-forward networks (SLFNs) where the hidden layer is not needed like a neuron; hence hidden layer is not tuned. For binary classification problems the output function of ELM for single output unit is:

$$f_L(x) = \sum_{i=1}^L \beta_i h_i(x) = h(x)\beta \quad (12)$$

where $\beta = \{\beta_1, \dots, \beta_L\}^T$ is the vector having output weights between the hidden layer of L nodes and the output node and $h(x) = [h_1(x), \dots, h_L(x)]$ is the output vector. For binary classification problems the decision function of ELM is:

$$f_L(x) = \text{sign}(h(x)\beta) \quad (13)$$

In order to have the better generalization performance of the network, the ELM targets to reach the smallest training error and smallest norm of the output weights by minimizing the following objective function:

$$\text{Minimize: } \|H\beta - T\|^2 \text{ and } \|\beta\| \quad (14)$$

where H represents the hidden layer output matrix

$$H = \begin{bmatrix} h(x_1) \\ \vdots \\ h(x_N) \end{bmatrix} = \begin{bmatrix} h_1(x_1) \cdots h_L(x_1) \\ \vdots \\ h_1(x_N) \cdots h_L(x_N) \end{bmatrix} \quad (15)$$

To minimize the norm of the output weights, $\|\beta\|$ maximizes the distance of the separating margins of both classes in ELM feature space: $2/\|\beta\|$ by defining the minimal least square method as:

$$\beta = H^\dagger T \quad (16)$$

where $H^\dagger$ is the moore-penrose generalized inverse of matrix that can be computed through the orthogonalization method, orthogonal projection method, and singular value decomposition. The multi class scenario consists of minimization of the following objective function:

$$\text{Minimize: } L_{PELM} = \frac{1}{2} \|\beta\|^2 + C \frac{1}{2} \sum_{i=1}^N \|\xi_i\|^2 \quad (17)$$

Subject to: $h(x_i)\beta = t_i^T - \xi_i^T, i = 1, \dots, N$

where $t_i = [t_{i,1}, \dots, t_{i,m}]^T$ is the target output vector and $\xi_i = [\xi_{i,1}, \dots, \xi_{i,m}]^T$ is the training error vector of the m output nodes. Training of the ELM consists of solving the following dual optimization problem based on KKT theorem:

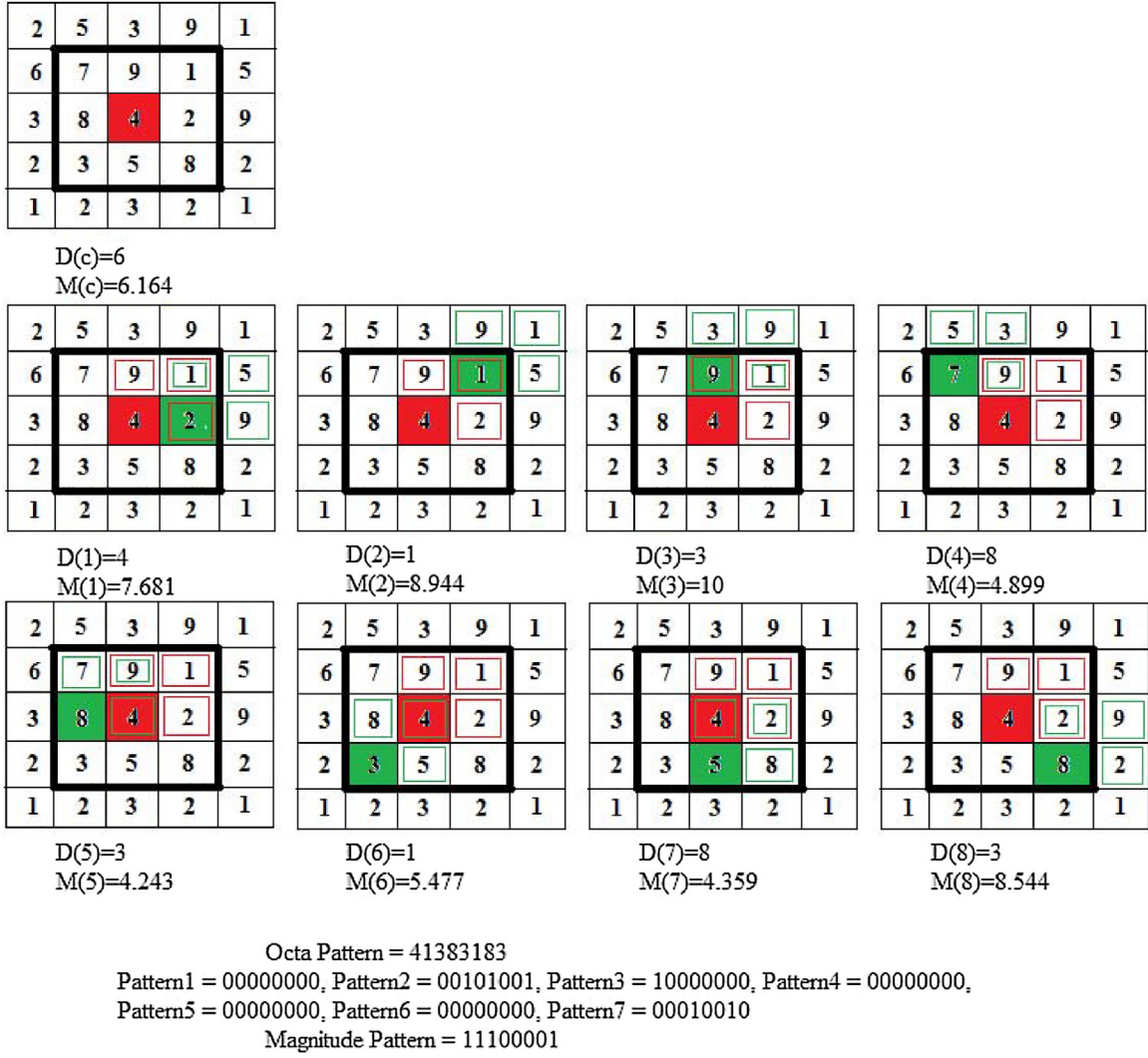


Fig. 2. Illustration of the tetragonal local octa patterns.

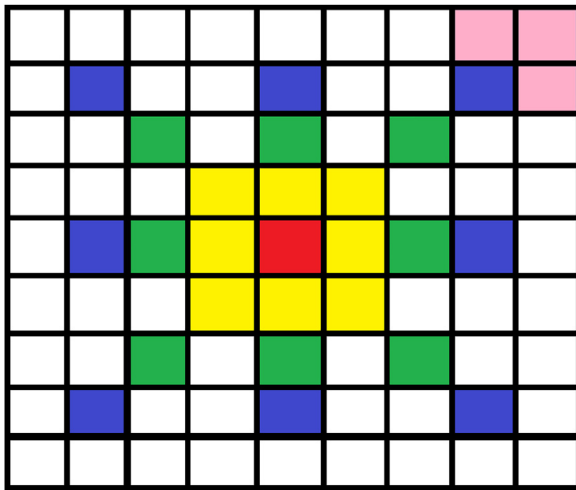


Fig. 3. Neighbors at different diagonal locations of a center pixel; yellow pixels are nearest neighbors at diagonal 3; green pixels are neighbor pixels at diagonal 5; blue pixels are neighbors at diagonal 7; whereas pink pixels represent horizontal, vertical and diagonal pixels of the blue cornered pixel. (For interpretation of the references to color in text, the reader is referred to the web version of the article.)

$$L_{DELM} = \frac{1}{2} \|\beta\|^2 + C \frac{1}{2} \sum_{i=1}^N \|\xi_i\|^2 - \sum_{i=1}^N \sum_{j=1}^m \alpha_{ij} (h(x_i) \beta_j - t_{ij} + \xi_{ij}) \quad (18)$$

where β_j is the vector of weights linking hidden layer to the j th output unit and $\beta = [\beta_1, \dots, \beta_m]$. The corresponding optimality conditions based on KKT are as follows:

$$\begin{cases} \frac{\partial L_{DELM}}{\partial \beta_j} = 0 \rightarrow \beta_j = \sum_{i=1}^N \alpha_{ij} h(x_i)^T \rightarrow \beta = H^T \alpha \\ \frac{\partial L_{DELM}}{\partial \xi_i} = 0 \rightarrow \alpha_i = C \xi_i, i = 1, \dots, N \\ \frac{\partial L_{DELM}}{\partial \alpha_i} = 0 \rightarrow h(x_i) \beta - t_i^T + \xi_i^T = 0, i = 1, \dots, N \end{cases} \quad (19)$$

where $\alpha_i = [\alpha_{i1}, \dots, \alpha_{im}]^T$ and $\alpha = [\alpha_1, \dots, \alpha_N]^T$. From equation (31) we have:

$$\begin{cases} \beta = CH^T \xi \\ \xi = \frac{1}{C} (H^T)^{\dagger} \beta \end{cases} \quad (20)$$

$$\begin{cases} H\beta - T + \frac{1}{C}(H^T)^*\beta = 0 \\ H^T\left(H + \frac{1}{C}(H^T)^*\right)\beta = H^TT \\ \beta = \left(\frac{1}{C} + H^TH\right)^{-1}H^TT \end{cases} \quad (21)$$

The output function of the ELM classifier is:

$$f(x) = h(x)\beta = h(x)\left(\frac{1}{C} + H^TH\right)^{-1}H^TT \quad (22)$$

For a given testing sample the class-label is the index of output node that has highest output.

2.5. CBIR based diabetic retinopathy detection

Once the query image is classified through ELM classifier, we obtain top-query neighborhood within the fundus images of the same class-label by computing Euclidean-distance based similarity. The purpose of returning the top-query neighborhood is to support the informed decisions about the associated cases regarding the prescriptions and affects. In the present work, although we do not have the pathology history of the fundus images, but association of this information is our target for the future work. The proposed DR-framework also help-out users to explore various case examples based on feature similarity. For this we generate a second set of fundus images based on feature similarity between query and repository images and enable relevance feedback over it [25]. In relevance feedback user label few output images as relevant through mouse clicks and rest of the output is considered as irrelevant. The classifier training occurs again on the labeled output that now serves as the input for classifier training. The process continues until the reliable output is achieved for the query image.

3. Experiments and results

3.1. Data sets

The system evaluation experiments were performed on following datasets: Kaggle-DR, DRIVE, STARE, and Review-DB. The Kaggle-DR dataset was taken from a 2015 DR detection competition sponsored by the California Healthcare Foundation [1]. The dataset comprises of 88,704 fundus images of either 5184×3456 or 3888×2592 resolution size and graded into five stages of DR by ophthalmologists [28]. Moreover, the images in the Kaggle-DR dataset were captured through different fundus cameras under different illumination conditions that makes the DR detection even more challenging. Other datasets we used for performance evaluation purposes are of varying lengths as shown in Table 1. The DRIVE dataset comprises of the images obtained from a diabetic retinopathy screening program occurred in the Netherlands. The dataset contains 40 colored retinal images randomly selected from the screening set of 400 DR patients. The images were grouped into training and test sets with 20 images in each set. Among these 40 images, 7 images belong to the mild non-proliferative DR class, whereas, 33 images have no sign of the DR (i.e. No-DR class). The images are of size 768×584 pixels, 8 bits per color channel, have a field of view (FOV) of approximately 540 pixels in diameter, and compressed in JPEG-format [29]. As the JPEG-compression is the lossy

compression method, therefore images lose significant amount of information, and feature extraction occurs imprecisely; hence, DR detection becomes a challenging task even in the smaller image repositories. Although the main purpose of the DRIVE dataset was to enable the segmentation of the retinal blood vessels but ultimately the segmentation results in the form of association of images with either non-proliferative DR class or the no-DR class, as the images in the dataset belongs to only these two classes.

The STARE dataset consists of 400 PPM images that are digitized slides captured by a TopCon TRV-50 fundus camera with 35° FOV that is approximately 650×550 pixels [30]. Each slide was digitized to produce 605×700 pixel image with 24 bits per pixel. The STARE dataset covers all classes of DR. The images have representation of large regions of exudates, multiple hemorrhages, and vessel occlusions. Annotations at image level are available. The Review-DB (Retinal Vessel Image set for Estimation of Widths database) was used to facilitate the development and comparison of vessel measurement algorithms and comprises of 16 images with 193 vessel segments.

The foremost reason for also incorporating the smaller sets for evaluation of the proposed method is that: the good performance over smaller training-sets confirms the robustness of a classifier and the faster convergence; whereas, the weak learners need large training sets and fails in case of inadequate training. Few healthy and affected fundus images from different datasets are shown in Fig. 4.

3.2. Evaluation of tetragonal local octa patterns

The performance of the proposed local pattern T-LOP is compared with other local patterns e.g. LBP [31], LDP [32] and LTrP [26]. In this experiment, classifiers are trained using various local patterns and applied to detect DR from the STARE dataset. The results in terms of accuracy, precision and recall are shown in Table 2. From the table, it can be observed that the T-LOP outperforms the LBP, LDP and LTrP based methods. Our algorithm overcomes the limitations that other patterns have e.g. LBP is not invariant to rotations and provides limited structural information; LDP is rotation invariant but it is also unable to effectively capture the structural information i.e. blood vessels; and LTrP can be further improved by considering the diagonal pixels for derivative calculations in addition to horizontal and vertical directions to capture the finer image regions (i.e. blood vessels). Further, the histograms of patterns (features) obtained from LBP, LDP, LTrP and L-TOP for an image from STARE dataset are shown in Fig. 5. We can observe from that the proposed T-LOP features effectively captures the structural information including the (thin and thick) blood vessels and color lesions and generates more discriminative features; whereas, LTrP captures the structural information better than LBP and LDP that ignores the directional information altogether; hence only captures the color lesions and discounts the blood vessels. Therefore, the LTrP performs better than the LBP and LDP but performs lower than the proposed method.

3.3. Stage-wise performance

The capability of the proposed system in detecting the stage of DR is also evaluated through the experiments. The trained single classifier through the ELM is applied on the proposed T-LOP features for the images from Kaggle dataset. The stage-wise DR-detection performance of the proposed method, in terms of accuracy, precision, recall, and F1-score, is presented in Table 3. It clearly shows that the proposed approach has achieved the remarkable accuracy, precision, recall, and F1-score values, and exhibited the lower error rates. The reason for the good classification performance signifies the effectiveness of proposed feature extraction approach that screens each class in a viable manner. In addition, features for various stages are shown in Fig. 6. It can be observed that the No-DR and mild-DR are closely related, but mild-DR exhibits higher peaks at some points. Whereas, the severe DR grows

Table 1
Data Sets used for comparative study.

Dataset	Instances	Resolution
Kaggle DR	88,704	5184×3456 , 3888×2592
DRIVE	40	565×584
Review-DB	16	3584×2438 , 1360×1024
STARE	400	605×700

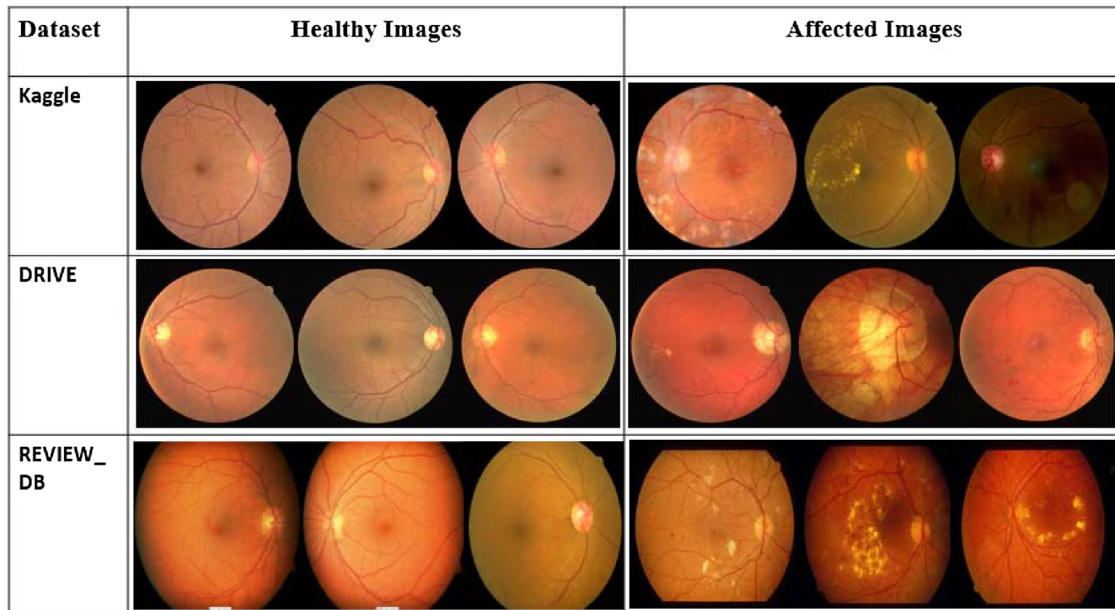


Fig. 4. Sample images from different datasets.

Table 2

Performance comparison of the proposed tetragonal local octa patterns with other local patterns.

Methods	Accuracy	Precision	Recall
LBP [31]	90.13	0.892	0.902
LDP [32]	90.33	0.903	0.92
LTrP [26]	94.0	0.941	0.93
Proposed T-LOP	99.5	0.991	0.9932

from the moderate DR, but in severe-DR due to the bleeding blood vessels dense texture appears and feature discrimination becomes easier comparatively. On the other hand, as the proliferative-DR occurs after severe DR, therefore some association is found between both classes, but still both classes are recognizable. Hence, due to the effective feature extraction our classification model gives the robust stage-wise DR performance that highlights the effectiveness of the proposed approach.

3.4. Comparative analysis

In the present work, we have evaluated the proposed method by a 10-fold cross validation approach, we randomly partitioned the data into 10 equal size sub samples. The average results in terms of performance evaluation measures were then considered. For comparison, we considered the published results reported in the comparative studies. The brief description of the comparative analysis based on different data sets is provided in the following subsections.

3.4.1. Performance comparison over Kaggle-DR

For performance evaluation over Kaggle-DR, we compared our method against the work of Quéllec et al. [15], Gargaya et al. [16], and Prett et al. [17]. The reason to select these methods for comparison is that these approaches are based on deep learning (DL); and in medical imaging DL had shown highest accuracy over tasks that previously required medical experts.

The results presented in Table 4 elaborates that the proposed approach has achieved the highest precision, recall, accuracy, AUC rates (i.e. 0.991, 0.9932, 99.96, 0.995 respectively) that signifies the reliability of the proposed approach for DR detection in comparison against other solutions. Additionally, the comparative approaches have reported the results without quantifying the uncertainty factor in a

decision. In contrast, a physician consults more experienced colleagues and refers case histories with similar background when needed. Therefore, in our work we addressed this fundamental requirement by returning the symmetrical case histories in form of visually-similar images to effectively mimic the human behavior. Moreover, our method only takes 20 min to train on Kaggle-DR with Intel Core I-5 machine with 16-GB ram; whereas, the comparative DL-based approaches on average take 13 h for training over NVIDIA GeForce GTX 970 and a GeForce 412 GTX 1080 with cuda versions 7.5/8 and cuDNN 4/5 machines that clearly shows the efficiency of our method. Moreover, as our solution is independent of the expensive hardware requirements, therefore, can easily be adopted in environments with budget-constraints.

3.4.2. Performance comparison over DRIVE and STARE datasets

Table 5 presents the comparative results of proposed method over DRIVE and STARE datasets. The results show that the proposed method outperforms the comparative methods in terms of all performance measures. The fundamental performance difference can be observed in terms of Recall rates in both DRIVE, and STARE datasets; where in case of the DRIVE dataset, our method achieved the average recall rate of 0.9945; whereas the average recall rate of the comparative methods is 0.726. Similarly, in case of the STARE dataset our method achieved the recall rate of 0.994; whereas, the comparative methods achieved the average recall rate of 0.7433. Therefore, we can clearly observe that our method provided approximately 27% performance gain for DRIVE dataset, and 25% for STARE datasets. Hence, we can say that our method is robust in terms of model association for various stages of DR. Moreover, as described earlier that the images in the DRIVE dataset are compressed through the JPEG-compression, therefore, the high evaluation rates also signify the robustness of the proposed method against image compression. Hence, the method can easily be applied for the screening repositories where compression is a mandatory step.

3.4.3. Performance comparison over Review-DB

To further evaluate the performance of proposed system, the retinal images from REVIEW database were also used. Table 6 presents the comparison of proposed method against two methods i.e. [21,22]. The obtained results show the higher performance gain of proposed method in contrast with comparative methods in terms of average recall rates, accuracy and time. By inspecting Table 4, it can be observed that our

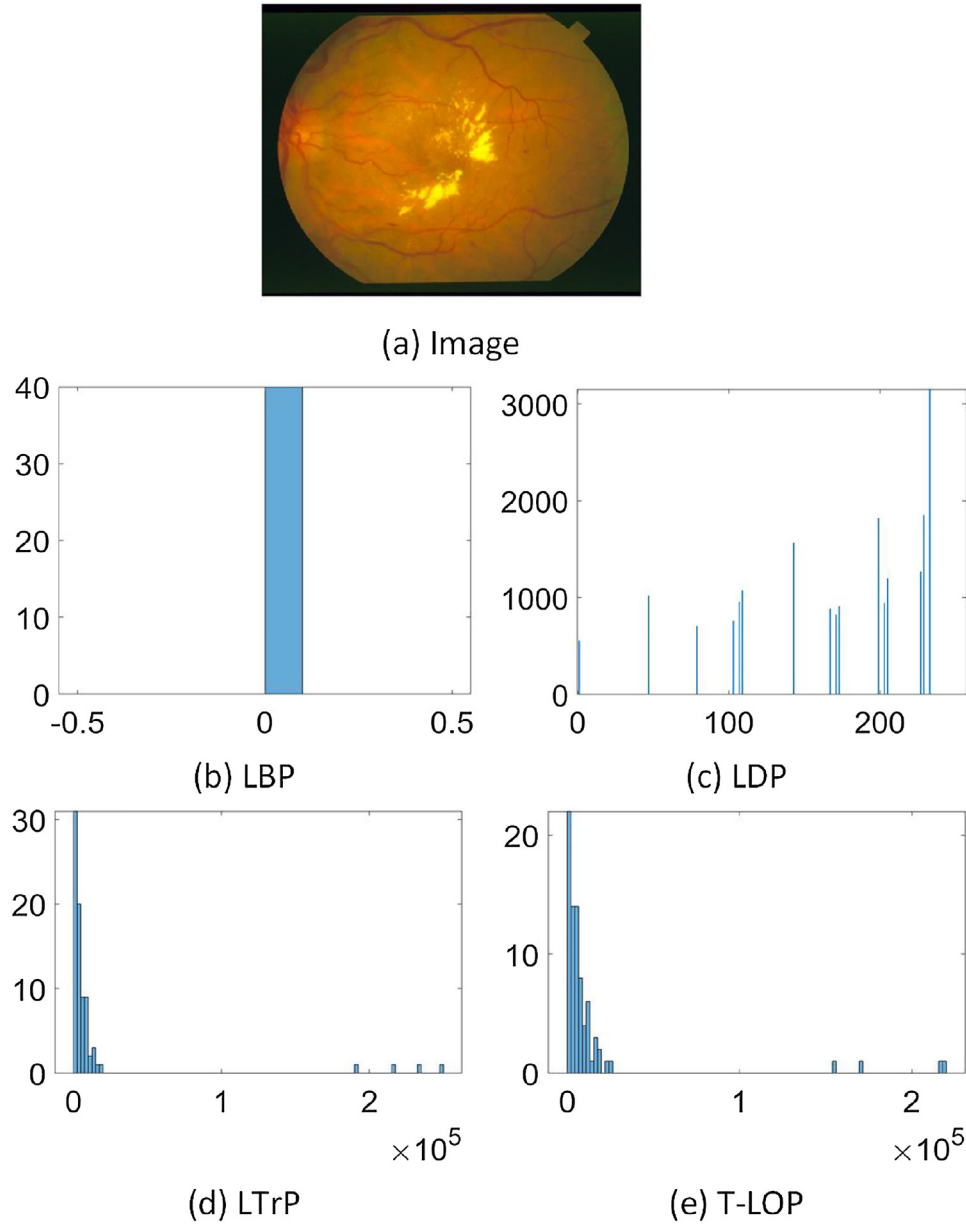


Fig. 5. Comparison of the feature patterns in terms of histograms: (a) a sample image taken from STARE dataset, (b) LBP, (c) LDP, (d) LTrP, and (e) T-LOP.

Table 3

Stage-wise performance of proposed method.

Stages	Accuracy	Precision	Recall	F1-score	Error-rate
No DR	99.75	0.987	0.989	0.99	0.01
Mild DR	98.83	0.979	0.995	0.993	0.007
Moderate DR	99.79	0.996	0.987	1	0
Severe DR	99.85	0.999	0.998	0.998	0.002
Proliferative DR	99.76	0.997	0.997	0.997	0.003

proposed method achieved 0.9883 as average recall rate that is approximately 10% higher than [21] and 32% than [22]. Whereas, average accuracy of our proposed method is 98.94% that is also higher than the comparative methods. These statistics clearly demonstrates the supremacy of our method and its capability to produce acceptable and competitive accuracies at minimum computational time. The average processing time for proposed method on an image of REVIEW dataset is 20 s with Intel Core I-5 machine with 16-GB ram; whereas, the average processing time for Kalaie et al.'s method was 28.1 s and for Xu et al.'s

method was 27.7 s on one image.

3.5. Cross-dataset validation

We conclude the experimentation section by presenting the cross-dataset validation of the proposed method. Through cross dataset validation, we can measure the robustness of the proposed scheme in terms of overcoming the training and testing challenges and prove it as a suitable candidate for the real-world scenarios. For the cross-dataset validation, we considered the following scenarios: (a) training over the Kaggle and test over the DRIVE datasets (Fig. 7(a)); (b) training over the DRIVE and testing over the STARE (Fig. 7(b)); (c) training over the Kaggle and test over STARE (Fig. 8); (d) training over DRIVE and test over Review-DB (Fig. 9). In cross dataset validation experiment, we achieved average accuracy of 99.9% for training and 99% accuracy on test sets (Fig. 7). To further elaborate the performance gain on standard dataset, we considered box and whisker plots to understand the distribution of accuracy on validation sets, as shown in Figs. 8 and 9. The box and whisker plot portrays the spread of accuracy across the number

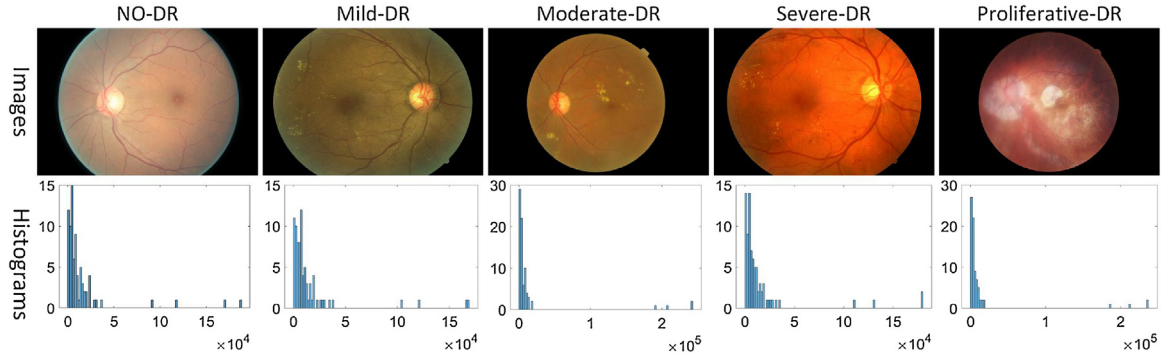


Fig. 6. Feature patterns of DR stages in Kaggle-DR.

Table 4

Performance comparison over Kaggle-DR.

Methods	Accuracy	Precision	Recall	AUC
Quelleg et al. [15]	95.4	0.952	0.945	0.955
Gargeya et al. [16]	97.0	0.98	0.94	0.97
Pratt et al. [17]	75	0.3	0.95	–
Proposed T-LOP	99.6	0.991	0.9932	0.995

Table 5

Performance comparison over DRIVE and STARE.

Methods	DRIVE			STARE		
	Accuracy	Precision	Recall	Accuracy	Precision	Recall
Zhou et al. [18]	94.6	0.967	0.827	95.98	0.9761	0.8065
Cavinato et al. [19]	92.93	–	0.645	90.3	–	0.7291
Mumtaz et al. [20]	94.5	0.9801	0.7067	95.26	0.9819	0.6944
Proposed T-LOP	99.96	0.997	0.9945	99.95	0.993	0.994

Table 6

Performance comparison over REVIEW-DB.

Methods	Accuracy	Precision	Recall
Kalaie et al. [21]	98.4	0.886	0.886
Xu et al.'s [22]	97	0.642	0.666
Proposed T-LOP	98.94	0.9871	0.9883

line into four quartiles and median. The achieved results represent the higher prediction rate in term of accuracy on cross dataset validation. The median accuracy of our technique is at 0.996 on Kaggle training set

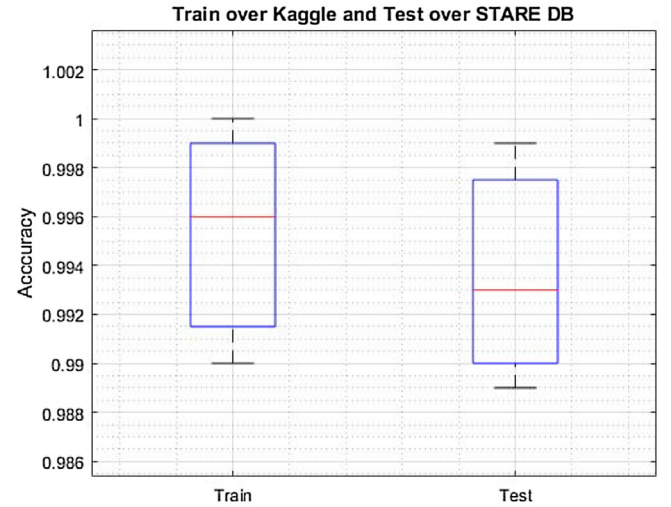


Fig. 8. Training over Kaggle and test over STARE DB.

and 0.993 at STARE testing set. Whereas, median accuracy of our technique is at 0.994 on DRIVE training set and 0.993 at Review-DB testing set. Hence, on the basis of the cross-dataset validation, we can say that our method can reliably be applied in real-world scenarios to address any condition of the DR to support the ophthalmologists.

4. Conclusion

In this paper, we presented a novel technique to precisely detect various stages of the DR by extending the research of the content-based image retrieval domain. For feature extraction, the concept of

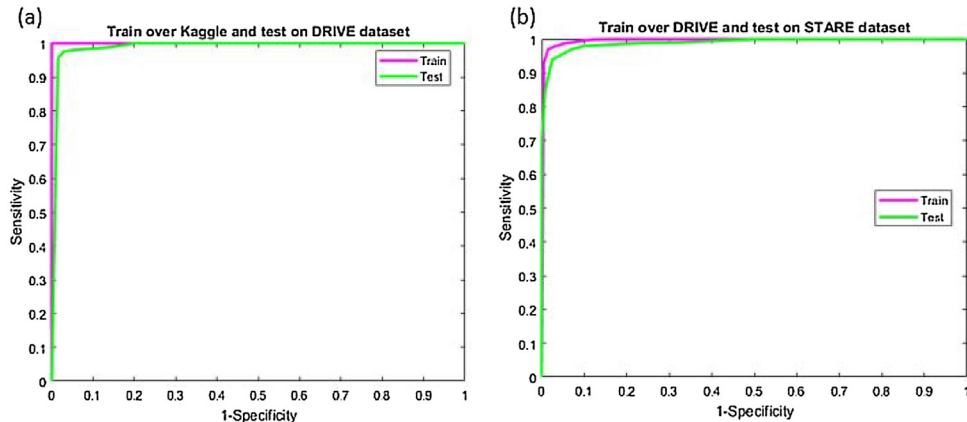


Fig. 7. (a) Training over Kaggle and test over DRIVE DB, (b) training over DRIVE and test over STARE.

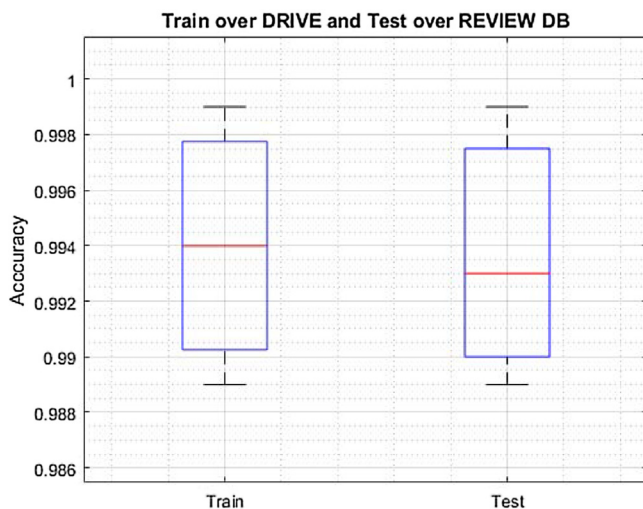


Fig. 9. Training over DRIVE and test over REVIEW DB.

tetragonal local octal patterns (T-LOP) is introduced. T-LOP is a novel image representation scheme that particularly is designed for analyzing the fundus images that is a composition of blood vessels and textured lesions. To achieve the human-level performance over the large-scale DR-datasets (i.e. Kaggle-DR), the fundus images are represented by the proposed tetragonal local octa pattern (T-LOP) features, that are then classified through the extreme learning machine (ELM). The proposed approach has achieved highest accuracy, precision and recall rates (i.e. 99.6%, 0.991, 0.993 respectively). The proposed technique supports in diagnosing DR from fundus images and precisely classifies the severity of the disease by also presenting the similar case histories for effective prescription. In contrast to the state-of-the-art DR detection methods, the proposed framework is capable to compute discriminative features from low-resolution and noisy images; and hence improved the detection and recognition of Diabetic Retinopathy.

Conflict of interest

The authors have no conflict of interest.

Acknowledgments

This work was supported by Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education (2016R1D1A1B03933860).

References

- [1] Diabetic-retinopathy. Available from: <https://www.kaggle.com/c/diabetic-retino>. [Accessed 14 September 2018].
- [2] Carrera EV, González A, Carrera R. Automated detection of diabetic retinopathy using SVM. 2017 IEEE XXIV International Conference on Electronics, Electrical Engineering and Computing (INTERCON). 2017. p. 1–4.
- [3] Wu B, Zhu W, Shi F, Zhu S, Chen X. Automatic detection of microaneurysms in retinal fundus images. *Comput Med Imaging Graph* 2017;55:106–12.
- [4] Akbar S, Sharif M, Akram MU, Saba T, Mahmood T, Kolivand M. Automated techniques for blood vessels segmentation through fundus retinal images: a review. *Microsc Res Tech* 2019;82(2):153–70.
- [5] García M, Sánchez CI, López MI, Abásolo D, Hornero R. Neural network based detection of hard exudates in retinal images. *Comput Methods Programs Biomed* 2009;93(1):9–19.
- [6] Faust O, Acharya R, Ng EY-K, Ng K-H, Suri JS. Algorithms for the automated detection of diabetic retinopathy using digital fundus images: a review. *J Med Syst* 2012;36(1):145–57.
- [7] Desbiens J, Gupta S, Stevenson J, Alderman A, Trivedi A, Buehler P. Deep annotated learning, harmonic descriptors and automated diabetic retinopathy detection. 2018.
- [8] Sandhu HS, Eladawi N, Elmoghy M, Keynton R, Helmy O, Schaal S, El-Baz A. Automated diabetic retinopathy detection using optical coherence tomography angiography: a pilot study. *Br J Ophthalmol* 2018. [bjophthalmol-2017-025177](https://doi.org/10.1136/bjophthalmol-2017-025177).
- [9] Wang Z, Yin Y, Shi J, Fang W, Li H, Wang X. Zoom-in-net: deep mining lesions for diabetic retinopathy detection. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. 2017. p. 267–75.
- [10] Bourouis S, Zaguia A, Bouguila N. Hybrid statistical framework for diabetic retinopathy detection. *International Conference Image Analysis and Recognition*. 2018. p. 687–94.
- [11] Dashtbozorg B, Mendonça AM, Campilho A. Optic disc segmentation using the sliding band filter. *Comput Biol Med* 2015;56:1–12.
- [12] Xu J, Chutatape O, Sung E, Zheng C, Kuan PCT. Optic disk feature extraction via modified deformable model technique for glaucoma analysis. *Pattern Recogn* 2007;40(7):2063–76.
- [13] Fraz MM, Barman SA, Remagnino P, Hoppe A, Basit A, Uyyanonvara B, Rudnicka AR, Owen CG. An approach to localize the retinal blood vessels using bit planes and centerline detection. *Comput Methods Programs Biomed* 2012;108(2):600–16.
- [14] Fraz MM, Remagnino P, Hoppe A, Uyyanonvara B, Owen CG, Rudnicka AR, Barman S. Retinal vessel extraction using first-order derivative of Gaussian and morphological processing. *International Symposium on Visual Computing*. 2011. p. 410–20.
- [15] Quéllec G, Charrière K, Boudi Y, Cochener B, Lamard M. Deep image mining for diabetic retinopathy screening. *Med Image Anal* 2017;39:178–93.
- [16] Gargeya R, Leng T. Automated identification of diabetic retinopathy using deep learning. *Ophthalmology* 2017;124(7):962–9.
- [17] Pratt H, Coenen F, Broadbent DM, Harding SP, Zheng Y. Convolutional neural networks for diabetic retinopathy. *Procedia Comput Sci* 2016;90:200–5.
- [18] Zhou W, Wu C, Chen D, Wang Z, Yi Y, Du W. Automatic microaneurysm detection of diabetic retinopathy in fundus images. 2017 29th Chinese Control and Decision Conference (CCDC). 2017. p. 7453–8.
- [19] Cavinato L, Fidone I, Bacis M, Del Sozzo E, Durelli GC, Santambrogio MD. Software implementation and hardware acceleration of retinal vessel segmentation for diabetic retinopathy screening tests. 2017 39th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC). 2017. p. 1226–9.
- [20] Mumtaz R, Hussain M, Sarwar S, Khan K, Mumtaz S, Mumtaz M. Automatic detection of retinal hemorrhages by exploiting image processing techniques for screening retinal diseases in diabetic patients. *Int J Diabetes Dev Ctries* 2018;38(1):80–7.
- [21] Kalaie S, Gooya A. Vascular tree tracking and bifurcation points detection in retinal images using a hierarchical probabilistic model. *Comput Methods Programs Biomed* 2017;151:139–49.
- [22] Xu Y, Zhang H, Li H, Hu G. An improved algorithm for vessel centerline tracking in coronary angiograms. *Comput Methods Programs Biomed* 2007;88(2):131–43.
- [23] Esnaashari M, Monadjemi SA, Naderian G. A content-based retinal image retrieval method for diabetes-related eye diseases diagnosis. *Int J Res Rev Comput Sci* 2011;2(6):1222.
- [24] Chaum E, Karnowski TP, Govindasamy VP, Abdelrahman M, Tobin KW. Automated diagnosis of retinopathy by content-based image retrieval. *Retina* 2008;28(10):1463–77.
- [25] Irtaza A, Adnan SM, Ahmed KT, Jaffar A, Khan A, Javed A, Mahmood MT. An ensemble based evolutionary approach to the class imbalance problem with applications in CBIR. *Appl Sci* 2018;8(4):495.
- [26] Murala S, Maheshwari R, Balasubramanian R. Local tetra patterns: a new feature descriptor for content-based image retrieval. *IEEE Trans Image Process* 2012;21(5):2874–86.
- [27] Ding S, Zhao H, Zhang Y, Xu X, Nie R. Extreme learning machine: algorithm, theory and applications. *Artif Intell Rev* 2015;44(1):103–15.
- [28] Leibig C, Ailken V, Ayhan MS, Berens P, Wahl S. Leveraging uncertainty information from deep neural networks for disease detection. *Sci Rep* 2017;7(1):17816.
- [29] Niemeijer M, Staal JJ, van Ginneken B, Loog M, Abramoff MD. Drive: digital retinal images for vessel extraction. *Methods for evaluating segmentation and indexing techniques dedicated to retinal ophthalmology*. 2004.
- [30] Hoover A. Structured analysis of the retina stare. 2015.
- [31] Ojala T, Pietikainen M, Maenpää T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans Pattern Anal Mach Intell* 2002;24(7):971–87.
- [32] Schölkopf B, Platt JC, Shawe-Taylor J, Smola AJ, Williamson RC. Estimating the support of a high-dimensional distribution. *Neural Comput* 2001;13(7):1443–71.