IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received March 7, 2020, accepted March 30, 2020, date of publication April 3, 2020, date of current version April 17, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2985543

Intelligent Automated Detection of Microaneurysms in Fundus Images Using Feature-Set Tuning

IMRAN USMAN, (Senior Member, IEEE), AND KHALED A. ALMEJALLI
College of Computing and Informatics, Saudi Electronic University, Riyadh 11673, Saudi Arabia

Corresponding author: Imran Usman (i.usman@seu.edu.sa)

This work was supported by the Deanship of Scientific Research (DSR) at Saudi Electronic University, Saudi Arabia, for its In-House Project through the Automated Detection of Micro-Aneurysms (MAs) Using Machine Learning for Early Prevention of Diabetic Retinopathy (DR), May 27, 2018, under Grant 15.

ABSTRACT Due to the widespread of diabetes mellitus and its associated complications, a need for early detection of the leading symptoms in the masses is felt like never before. One of the earliest signs is the presence of microaneurysms (MAs) in the fundus images. This work presents a new technique for automatic detection of MAs in color fundus images. The proposed technique utilizes Genetic Programming (GP) and a set of 28 selected features from the preprocessed fundus images in order to evolve a mathematical expression. Through the binearization of the fitness scores, the optimal expression is evolved generation by generation through a stepwise enhancement process. The best expression is then used as a classifier for real world applications. Experimental results using three publically available datasets validate the usefulness of the proposed technique and its ability to outperform the state of the art contemporary approaches.

INDEX TERMS Automatic microaneurysms detection, diabetic retinopathy, fundus images, genetic programming.

I. INTRODUCTION

As the world's human population continues to multiply, the need for finding efficient automated health solutions for the masses is felt like never before. Among some of the most prevalent ailments in the world is diabetes mellitus, and its associated complications. In a 2014 estimate by the World Health Organization (WHO) [1], approximately 422 million people in the world were suffering from diabetes, which is predicted to reach 592 million by 2035 [2]. In 2016, an estimated 1.6 million people died worldwide due to diabetes [1].

One of the most persistent complications attributable to diabetes is Diabetic Retinopathy (DR). It is the leading cause of blindness among adults between the ages of 20 to 60 years [3]. Vision impairment due to DR is progressive with almost no early complaints by the patients. However, if DR is detected and treated at the earlier stages, vision loss is preventable. One of the earliest signs of DR is the presence of microaneurysms (MAs), which are tiny red circular spots on the surface of the retina in the human eye [4]. They are very minute with a diameter range of 10μm to 100μm [5]. MAs are out-pouches of capillaries and sometimes result from the leakage of tiny blood vessels in the retina. They are the only lesions which are present at the earliest stage of the disease and remain until the development of the DR [6]. Therefore, early detection of MAs for the detection and treatment of DR is of prime importance.

An ophthalmologist uses fundoscopy or fundus images, captured with a specialized camera, in order to detect MAs. However, there are several difficulties associated with the manual screening of the masses. Firstly, the availability of ophthalmologist is not guaranteed everywhere, especially in remote locations. Secondly, manual screening of a large population is a cumbersome and time-consuming work. Thirdly, it is not cost-efficient and is sometimes prone to errors. Hence, an automated computer-aided diagnosis tool is required for accurate detection and classification of MAs. There are certain other elements in fundus images that resemble MAs which makes it difficult to differentiate them from MAs. Examples of these include small circular

The associate editor coordinating the review of this manuscript and approving it for publication was Weiping Ding.

spots at the crossing of thin blood vessels, noise, and other external distortions introduced during transmission of the images. In addition, MAs that are irregular in shape, clustered together, close to the periphery of the image or in the macula, are also difficult to detect.

One of the earliest works on the automated detection of MAs was by Baudoin *et al.* [7] who proposed a mathematical morphology based technique using fluorescein angiograms (FAs). This method was costly, time-consuming, invasive and not suitable for everyone due to FAs. Based on FAs, Spencer *et al.* [8] proposed a technique that utilized adaptive filters to calculate several features of MAs such as shape, size, and intensity in order to detect MAs. In another work [9], Spencer *et al.* used a bilinear top-hat transformation and Gaussian matched filters to segment the digitized FA image initially, and then applied a region growing algorithm for delineating and analyzing candidate MAs for detection. However, this technique also had the drawbacks of using FAs. Fleming *et al.* [10] used various contrast normalization methods for differentiating between MAs and other dots in the fundus image. Their algorithm had low sensitivity and specificity values, which makes it unsuitable for efficient MA detection.

There are many recent studies on the detection of MAs with improved results compared to earlier techniques. Antal *et al.* [11] proposed an ensemble-based detection technique by placing all available preprocessing steps and the classifiers in an ensemble pool and then selecting an acceptable subset of combinations from the ensemble pool based on results' evaluation. Rocha *et al.* [12] proposed a novel technique based on visual words dictionary representing all possible points of interests associated with DR and already marked by human experts. Fundus images are then classified based on the presence or absence of these points of interests. An interactive multi-agent based method for MA segmentation is presented in [5]. Lazar and Hajdu's [13] work is based on analysis of cross-section profiles and peak detection method to find a set of attributes related to shape, size and height of the peaks. Feature sets are constructed and used in a naive Bayes classifier. Wu *et al.* [3] extended the work presented in [13] by introducing a total of 27 characteristic features and using Kth Nearest Neighbor (KNN) classifier to classify true MAs from false ones. Srivastava *et al.* [14] used Frangi filters to extract features from preprocessed sub-images of the green channel. A Support Vector Machine (SVM) classifier then uses these features to predict the presence of MAs. Rosas-Romero *et al.* [15] used bottom-hat transform and hit-or-miss transform to first find the candidate MAs in a preprocessed image, then used Radon transform and Principal Component Analysis (PCA) to detect the real MAs. Nevertheless, the accuracy of these techniques is not adequate to be used for a DR diagnostic tool in real-world scenario.

Recently many deep learning based algorithms are proposed in the literature [16]. Haloi [17] proposed a deep neural network based technique using maxout activation function and the dropout training process. Their technique provides good detection accuracy but is computationally expensive. Orlando *et al.* [18] proposed a technique based on a combination of deep learning and domain knowledge. Handcrafted features are supported by features learned through a convolutional neural network (CNN), which are later used by a Random Forest Classifier to find true lesions. Recently, Eftekhari *et al.* [19] proposed a technique based on a two-stage process and CNN. First, normal samples are selected from a probability map generated from the first CNN. Next, the second CNN classifies each pixel in the test images as MA or non-MA. Test results on online datasets show that their technique solves the imbalanced data problem with shorter training time in comparison to the previous techniques.

The recent advancements in the accuracy of MAs detection show the usefulness of intelligent techniques, such as those inspired by nature, in better diagnosis of DR. Based on these, in this work we propose a Genetic Programming based novel method of MA detection. In addition to enhanced image pre-classification compared to [3], the proposed method is capable of exploring hidden dependencies among the features, which are otherwise overlooked in the conventional classifiers. It generates a generalized mathematical expression for the classification of candidate MAs in a given image. Through substantial experimentation, we enhance the work presented in [3] and [13] and include a number of features in our feature set which highly contribute to the classification accuracy of the proposed technique. In addition, classification accuracy is also improved by steering the solution space towards the optimum using a finely crafted fitness function and the binarization module. Experimental results on publically available datasets demonstrate that the proposed technique outperforms some state of the art contemporary techniques in terms of different performance measures. In addition, it provides a generalized mathematical formula based on feature dynamics through the GP based training mechanism. We call our proposed technique as Intelligent Feature Set Tuning (IFST) technique.

The rest of this paper is organized as follows. Section 2 presents our proposed technique with the presentation of the general architecture and explanation of the constituting modules. Experimental results and related discussion are presented in section 3. Finally, in section 4, we present the conclusion of our work.

II. PROPOSED IFST TECHNIQUE

This section presents the proposed IFST technique in detail. We first present the general architecture of the proposed technique, which includes description of both the training and the testing phases of the GP based evolution process. Next, we present the details of the individual modules in the subsequent sub sections.

A. THE IFST TECHNIQUE: GENERAL ARCHITECTURE

Figure 1 shows the general architecture of the IFST technique. It constitutes of the traning phase and the testing phase. In the training phase, GP based simulation is carried out to

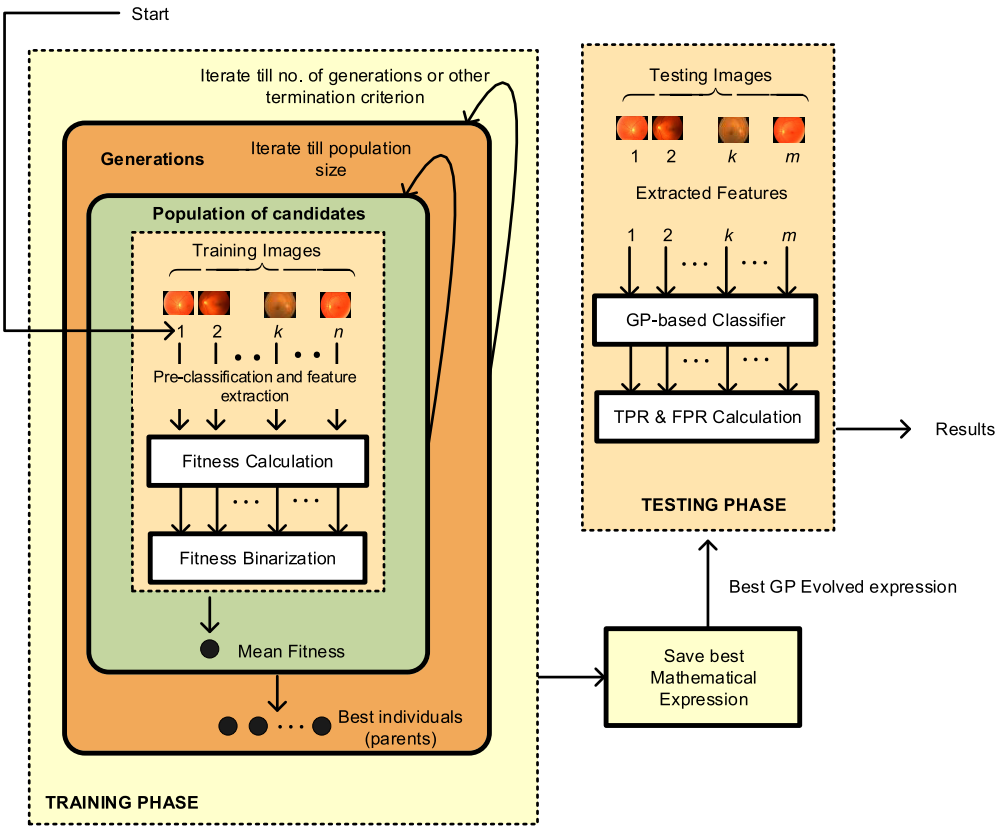

FIGURE 1. General architecture of the IFST technique.

evolve an optimized mathematical expression for accurate detection of MAs using a number of training images from publicaly available datasets. Details of these data sets are presented in section III.A. The detection of MAs is a two step process. First we detect a set of potential candidate MAs by applying a selection of image preprocessing techniques and generate the feature set. We call this stage as the pre-classification stage. This feature set is then passed on to the GP based classifier in order to refine MAs detection through classification. For a population of candidates (GP based mathematical expressions), each of the candidates is scored and ranked using a predefined fitness function, initially, and then converted to a binary number through the fitness biniarization process. Based on the fitness performance, the best individuals are selected as the parents for the next generation. Genetic operators of reproduction, crossover and mutation are applied on the parents to produce the new generation. The process continues to evolve the generalized expression, generation by generation, until a predefined termination criterion is met. This best evolved expression is then saved and used in the testing phase to validate the accuracy of the IFST technique on the test images.

B. THE TRAINING PHASE

The training phase comprises of the fundus image preprocessing, candidate extraction, feature extraction and the classification modules. It is to be noted that the preprocessing, candidate extraction and feature extraction phases can be a part of the training phase or isolated operations. Since there is no impact of these on the training and the overall temporal cost of the proposed technique, we incorporate them into the training phase. Figure 2 shows the detailed architecture of the training phase with the description of modules in the following subsections.

1) FUNDUS IMAGE PREPROCESSING

Candidate MAs appear in high contrast in the green color band of the RGB fundus images [20], [21]. Therefore, we use this color channel for the detection of MAs. In order to remove the salt and pepper noise, a 3×3 median filter is applied initially. With the focus on balancing contrast with the surrounding background and to make the interesting areas of the image more visible, contrast limited adaptive histogram equalization (CLAHE) technique is applied next [22].

To further reduce the effect of noise, we apply a Gaussian filter with the kernel size of 5×5 in order to generate the enhanced image I_{enh}. Next, we apply the shade correction technique adopted from [20] using the following equation:

$$I_{sc} = I_{enh}/I_{bk} \tag{1}$$

where I_{sc} is the image after shade correction, and I_{bk} is the background estimated image calculated by applying a median

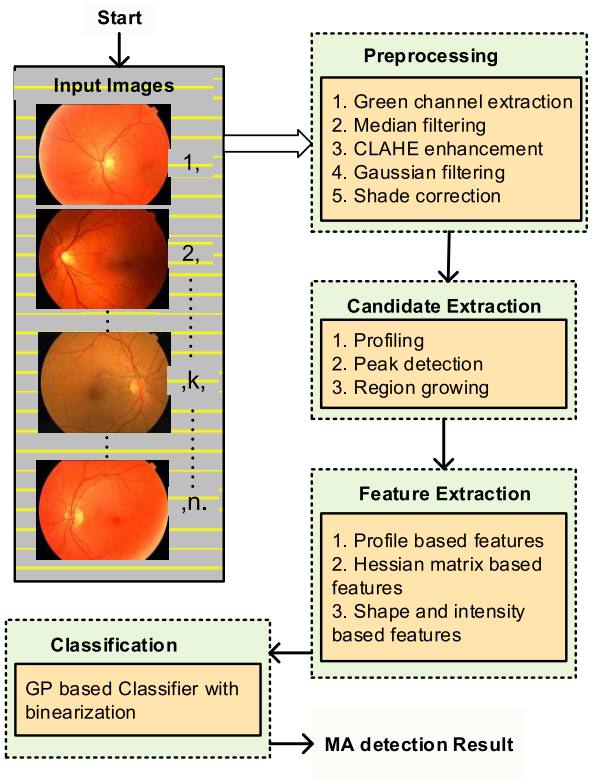

FIGURE 2. Detailed architecture of the training phase.

filter to I_{enh} with the kernel size of 68 × 68. Lastly, I_{sc} is divided by its standard deviation to produce global-contrast normalized image.

$$I_{preproc} = \frac{I_{sc}}{SD(I_{sc})} \tag{2}$$

where, $SD(I_{sc})$ represents the standard deviation of I_{sc} and $I_{preproc}$ is the final preprocessed image of this module which is used in the subsequent modules.

2) CANDIDATE MICROANEURYSMS EXTRACTION

Before the detection of MAs, extraction of the candidate MAs is a very important step. Any true MAs missed in this stage cannot be recovered later on. The primary objective of this module is to minimize this loss and to discard a number of objects that do not resemble MAs. In the work, we utilize the peak detection to extract the initial candidates as described by Wu *et al.* [3] and Lazar and Hajdu [13]. This technique is based on the fact that the MA region consists of at least one local maximum pixel. These regional maximum points can be considered as good MA candidates. However, this can also lead to the extraction of a large amount of noise which can make the detection task more complex and time consuming. In order to cope with this issue, we adopt the idea of Lazar and Hadju [13] who proposed a method for individual candidate profiles. For each of the local maximum pixel, its circular neighborhood pixels are examined using a set of line detectors of different orientations. This leads to a set of cross-sectional

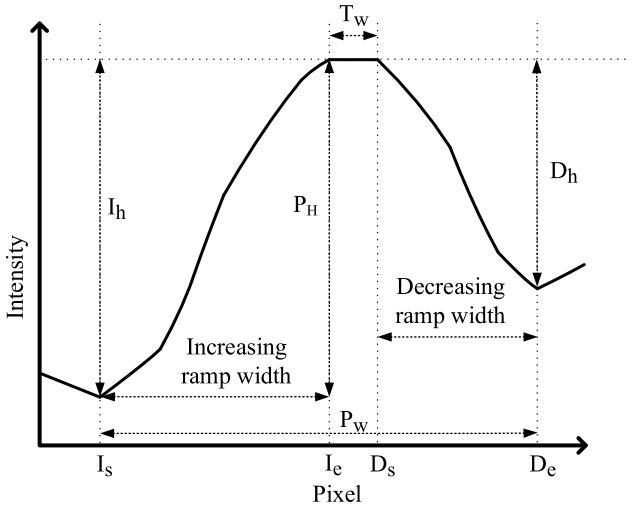

FIGURE 3. Intensity distribution of a profile and the definitions of various measures.

intensity profiles. An example of such intensity distribution is shown in figure 3.

Consequently, peak detection is applied to each of the profile to assess whether the peak is present at the center of the profile. For each of the candidates, we generate 30 profiles using 30 line detectors of length 31 pixels with 6^0 angular resolution following the approach adopted in [13]. A candidate is eliminated if in any of the profiles the absolute difference between all of the consecutive pixels is less than 4. The value of 4 is chosen empirically through experimentation. With the aim to characterize the pathologies found in the original retinal image, region-growing techniques are applied to grow the candidate MA back to the original pathology shape at the candidate extraction step. In order to achieve this, we apply the dynamic transformation based technique as used in [3].

3) FEATURE-SET EXTRACTION

To achieve better classification accuracy, appropriate set of features are selected and further exploited exhaustively through the GP simulation to find hidden dependencies in them. We use 17 profile features, 7 of which were proposed by Lazar and Hadju [13], and 3 proposed by Wu *et al.* [3], along with other important features pertaining to local curvature, shape and intensity of the candidate MA. For each of the candidate regions where a peak is detected, the 17 profile based features are:

1. μ_{HR}, the mean value of HR, and HR is a set containing the increasing and decreasing ramp height values denoted by I_H and D_H respectively. Here, $I_H = P[I_e] - P[I_s]$ and $D_H = P[D_s] - P[D_e]$ as shown in figure 3.
2. σ_{HR}, the standard deviation of HR.
3. $c\upsilon_{HR}$, the coefficient of variation for HR and is defined as: $c\upsilon_{HR} = \sigma_{HR}/\mu_{HR}$.
4. μ_{SR}, the mean value of SR, and SR is a set containing the increasing and decreasing ramp slope values denoted by I_{rS} and D_{rS} respectively. Here, $I_{rS} = I_H/(I_e - I_s)$ and $D_{rS} = D_H/(D_e - D_s)$.

5. σ_{SR}, the standard deviation of SR.
6. cv_{SR}, the coefficient of variation for SR and is defined as: $cv_{SR} = \sigma_{SR}/\mu_{SR}$.
7. μ_{WT}, the mean value of WT, and WT is a set containing the values of top width, T_W, which is the gap between the increasing and decreasing ramp and is defined as: $T_W = D_s - I_e$.
8. σ_{WT}, the standard deviation of WT.
9. cv_{WT}, the coefficient of variation for WT and is defined as: $cv_{WT} = \sigma_{WT}/\mu_{WT}$.
10. μ_{WP}, where μ represent the mean value of WP, and WP is a set containing the values of peak width, P_W, which is the gap between the start and end indices of the peak and is defined as: $P_W = D_e - I_s$.
11. σ_{WP}, the standard deviation of WP.
12. cv_{WP}, the coefficient of variation for WP and is defined as: $cv_{WP} = \sigma_{WP}/\mu_{WP}$.
13. μ_{HP}, the mean value of HP, and HP is a set containing the values of peak height, P_H, which is the difference between the intensity of the central pixel and a baseline that connects the start and the end of the profile. It is defined as: $P_H = P[center] - (P[D_e] - P[I_s])/P_W.(center - I_s) + P[I_s]$.
14. σ_{HP}, the standard deviation of HP.
15. cv_{HP}, the coefficient of variation for HP and is defined as: $cv_{HP} = \sigma_{HP}/\mu_{HP}$.
16. σ_{SEID}, the standard deviation of $SEID$, and $SEID$ is a set containing the intensities of start of the increasing ramp and the end of the decreasing ramp values defined by $P[I_s]$ and $P[D_e]$ respectively.
17. H_r, where $H_r = (H_{\max} - H_{\min})/H_{\min}$. Here, $H_{\max}$ and $H_{\min}$ represent the maximum and the minimum values in the set HR respectively.

These profile based features provides information regarding various inherent characteristics such as the intensity transition, sharpness of the contours, variance in the surroundings, presence of blood vessels, symmetry of the object, and vessel bifurcations, to name a few. Apart from these profile features, local features also play a vital role in classification accuracy. In order to describe the local curvature of a function with many parameters, we use Hessian matrix, $H(x, y)$, as described in [3]. Let κ_1 and κ_2 denote the two eigenvalues of the Hessian matrix, such that $|\kappa_1| \leq |\kappa_2|$. The probability map for an MA-like structure can be defined as:

$$P_{map}(x, y; \sigma) = \begin{cases} 0 & \kappa_1 > 0 || \kappa_2 > 0 \\ \frac{2}{\pi} \arctan\left(\frac{|\kappa_2| + |\kappa_1|}{|\kappa_2| - |\kappa_1|}\right) & \kappa_1 \neq \kappa_2 \\ 1 & \kappa_1 = \kappa_2 < 0 \end{cases} \tag{3}$$

and consequently,

$$P_{map}(x, y) = \max_{\sigma_{\min} < \sigma < \sigma_{\max}} P_{map}(x, y; \sigma) \tag{4}$$

where higher value of P_{map} corresponds to the higher likelihood of a pixel (x, y) belonging to the MA. For each of the candidate regions, the following Hessian matrix based 6 features are extracted:

18. $\mu_{P_{map}}$, the mean of P_{map}.
19. $\mu_{|H|}$, the mean of $|H|$, where $|H| = \kappa_{1\max} \times \kappa_{2\max}$. Here, $\kappa_{1\max}$ and $\kappa_{2\max}$ represent the eigenvalues for which P_{map} is the maximum.
20. $MAX(P_{map})$, where MAX represent the maximum value.
21. $MAX(|H|)$
22. $\sigma_{P_{map}}$, the standard deviation of P_{map}.
23. $\sigma_{|H|}$, the standard deviation of $|H|$.

Apart from Hessian matrix based features that define the curvature of the MA region, other local features are also considered that describe the shape and intensity of the MA. These 5 features are described as:

24. The area $A = \sum_{i \in n} p_i$, where n is the set of pixels in the candidate region.
25. The mean contrast of edge pixels, μ_E.
26. The mean intensity of MA candidate region, μ_{cand}.
27. The difference between the average intensity of the candidate region and the average intensity of its background. $\delta = \mu_{cand} - \mu_{bg}$
28. The aspect ratio $\lambda = l/s$, where l and s are the largest and the second largest eigenvalues of the covariance matrix of the candidate region respectively.

The above 28 features make the feature set corresponding to a candidate MA in this module. These features have different ranges and values. Therefore, we normalize the feature set values to zero mean and unit variance before classifying using the following:

$$\tilde{\Im}_j = \frac{\Im_j - \mu_j}{\sigma_j} \tag{5}$$

where, $\Im_j$ is the j th feature for $j = \{1, 2, \ldots, 28\}$. μ_j is the mean of the jth feature and σ_j is the standard deviation. After the feature set is extracted and normalized, it is presented to the GP-based classifier as described in the following subsection.

4) MICROANEURYSM DETECTION

We use GP to automatically evolve the classifier with the primary task to classify a candidate MA belonging to the set of labelled MAs. In order to generate the potential mathematical expressions population through GP, we first need to define suitable GP *function set* (*function_set*) and GP *terminal set*. In our proposed IFST, the GP function set comprises of mathematical functions such as plus, minus, multiplication, protected division, sine, cosine, log, exponent, maximum and minimum. The GP terminal set can be further categorized as constant terminal set and the variable terminal set. For the constant terminal set, we use random constants in the range [-1 1], whereas the variable terminal set comprises of the 28 features in the feature set after normalizing through equation 5. These functions and terminals make the non-leaf and leaf nodes in the GP expression tree respectively, which generates the candidate mathematical formula when

traversed. For each candidate expression constructed by the GP simulation, the functional dependency F of the features pertaining to the classification is defined as:

$$fitness = Eval(F(\bar{\Im}_j, [-11], function\_set)) \tag{6}$$

where *Eval* represent the evaluation of the function and *fitness* represents any numerical value that will help evolving a suitable expression generation by generation. In order to drag the fitness value towards a binary classification, we introduce the concept of *binarization*. In order to achieve this, we let $Flag = 1$ if $fitness > 1$ and $Flag = 0$ otherwise. In addition, we assume that $Flag = 1$ represents the presence of MA and $Flag = 0$ represents its absence. We also denote the actual class labels in similar manner, whereby the value of 1 represent the presence of MA and its absence is denoted by 0. Next we take the *XNOR* binary operation between the *Flag* and the actual class label represented in binary. This generates a value of 1 for the true positive (TP) and true negative (TN) cases and a value of zero for false positive (FP) and false negative (FN) cases. For the entire set of candidate MAs in an image, we define the final fitness score, $Score_{fitness}$ based on the contribution from all the candidates as follows:

$$Score_{fitness} = sensitivity = \frac{TP}{TP + FN} \tag{7}$$

The value of $Score_{fitness}$ ranges from 0 to 1. The larger this value is, the better is the performance of a candidate solution. Based on this, the complete population is scored and ranked in a particular generation. The best individuals from a given population are chosen as parents and genetic operators of crossover, mutation, and reproduction are applied to form the new generation. The solution is enhanced generation by generation until we reach the termination criterion, which is based on fitness value, or time limit or the number of generations. At the end of the GP simulation, the best-evolved mathematical expression is saved and used in the testing phase as a classifier function. Figure 4 shows a sample of candidate solution in the form of GP tree and the corresponding mathematical expression.

C. THE TESTING PHASE

After the best expression is evolved in the training phase, we use it in the testing phase to classify a number of test images. The results of this phase also validate the generalization capability of the proposed IFST technique. Based on this, several performance measures are computed that verify the usefulness of the proposed technique for the real world applications, as discussed in the next section.

III. EXPERIMENTAL RESULTS AND DISCUSSION

The proposed technique is implemented in MATLAB programming environment using the GPLAB toolbox [23] for the implementation of GP. We used the Intel Core i7 8th generation processor with SSD memory of 1TB and 32GB of RAM. In addition, the machine used for simulation was also equipped with 512GB graphics card. Table 1 presents different GP parameter settings used in our GP simulations using the GPLAB tool. For the purpose of evaluation and comparison with existing techniques, we use three standard datasets that are described in the following subsection. For each of the data set used, several GP simulations are carried out in order to get the best evolved expression with different training and testing images combinations. The core parameter settings, as presented in table 1, are kept the same.

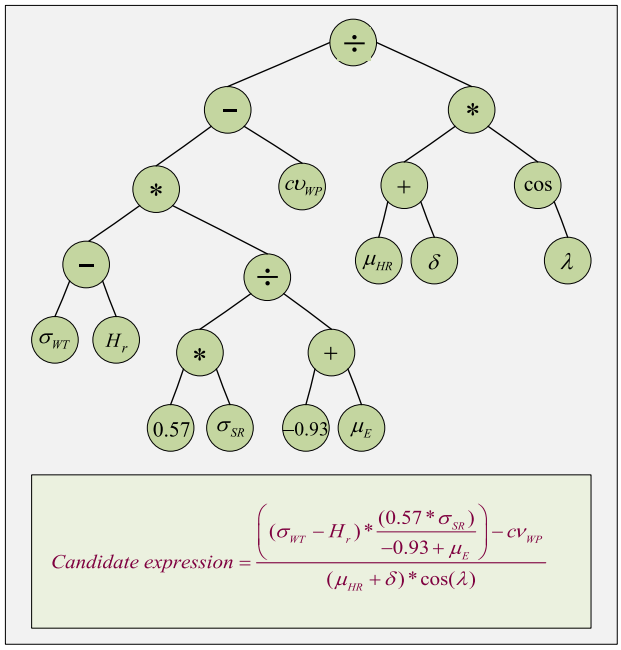

FIGURE 4. A sample GP candidate solution represented as a GP Tree, and the corresponding mathematical expression.

A. DIABETIC RETINOPATHY DATASETS

1) MESSIDOR DATASET

The Messidor dataset [24] consists of 1200 losslessly compressed images in TIFF format at three different resolutions and is publically available. These images were captured using 8 bits per color plane at 1440 × 960, 2240 × 1488 or 2304 × 1536 pixels. Corresponding to each image, a grading score is provided ranging from 0 to 3 that represents the severity of the disease. A patient with grade 0 has no DR. Patients with grade 1 and 2 have mild and severe levels of non-proliferative DR respectively. Whereas, grade 3 represents the most serious condition of DR. In the entire dataset, 540 images (46%) are of grade 0, 153 images (12.75%) are of grade 1, 247 (20.58%) correspond to grade 2 and 260 (21.67%) represent grade 3. The images are captured with a 45° field of view (FOV). Some sample images from the Messidor dataset with their grades are shown in figure 5.

TABLE 1. GP parameter settings for the IFST technique.

Function Set	+, -, *, protected division, EXP, SIN, COS, LOG, MAX and MIN
Terminal Set	Constant terminals: random constants in range of [-1, 1]
	Variable Terminals: Normalized values of μ_{HR}, σ_{HR}, cv_{HR}, μ_{SR}, σ_{SR}, cv_{SR}, μ_{WT}, σ_{WT}, cv_{WT}, μ_{WP}, σ_{WP}, cv_{WP}, μ_{HP}, σ_{HP}, cv_{HP}, σ_{SEID}, H_r, $\mu_{P_{map}}$, $\mu_{\|H\|}$, $MAX(P_{map})$, $MAX(\|H\|)$, $\sigma_{P_{map}}$, $\sigma_{\|H\|}$, A, μ_E, μ_{cand}, δ and λ.
Fitness Function	$fitness = Eval(F(\overline{\Im}_j, [-1 \quad 1], function\_set))$
Selection	Generational
Population size	100
Initial max. tree depth	8
Initial population creation type	Ramped Half-and-Half
Operator Probability type	Variable
Sampling	Tournament
Expected no. of offspring method	Rank 89
Survival mechanism	Keep best
Real max. level	30
Termination criterion	Generation 50

2) DIARETDB1 DATASET

The DIARETDB1 dataset is another freely available and widely used dataset that is abbreviated from DIAbetic RETinopathy DataBase - Calibration Level 1 (DIARETDB1) [25]. This database consists of 89 color fundus images in PNG format with a resolution of 1500 × 1152 pixels. Out of these, 84 images contain non-proliferative signs of the DR in the form of MAs and 5 are considered as normal images. It includes labels for both the MAs and exudates. Images in this database are captured using 50^o FOV of the fundus camera with various other settings that depict scenarios which are comparable to real life situations of image acquisition when screening the masses for DR. In addition, it is possible to quantitatively assess the accuracy of the method to detect each type of lesion.

3) E-OPHTHA DATASET

The e-ophtha database [26], [27] is made up of two sub-datasets called e-ophtha-MA and e-ophtha-EX which correspond to the presence of microaneurysms and exudates respectively. This work utilized e-optha-MA which comprises of 148 color fundus images with MAs or small hemorrhages and 233 images with no lesion in JPEG format. The image sizes are 2544 x 1696 and 1440 x 960. These images are manually annotated by expert ophthalmologists. In this work, we used half of the images as training and images while the remaining half is utilized for testing.

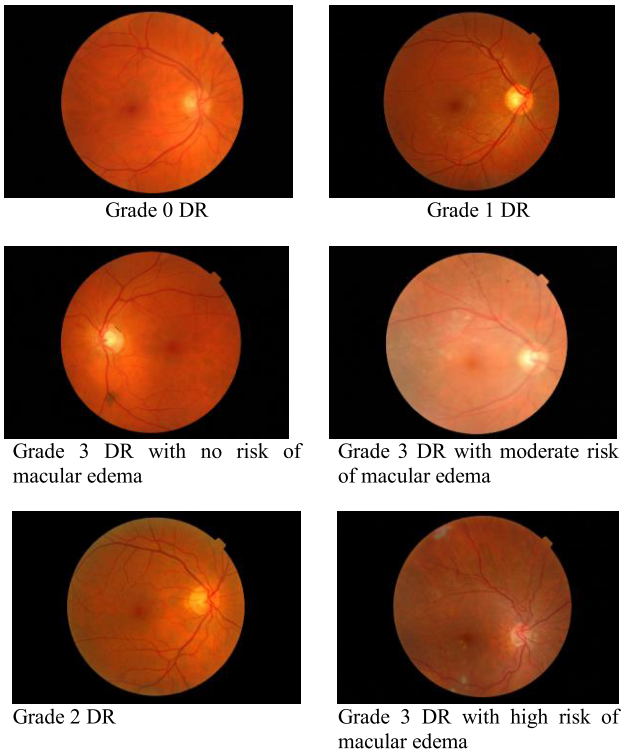

FIGURE 5. Sample images from the Messidor data set corresponding to different clinical conditions.

B. CLASSIFICATION PERFORMANCE ASSESSMENT

In order to assess the classification performance of the proposed technique on the available datasets, we use different performance metrics such as sensitivity, specificity, accuracy, area under the curve (AUC), receiver operating characteristic curve (ROC) and free-response receiver operating characteristic (FROC). Sensitivity is defined in equation (7) and relate to the true positive rate or the probability of detection. Specificity measures the proportion of actual negatives that are correctly identified as negatives. It is defined as:

$$specificity = \frac{TN}{FP + TN}. \tag{8}$$

Accuracy, on the other hand, refers to the degree of closeness of measurements of a quantity to that quantity's true value and is given by:

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}. \tag{9}$$

ROC is the receiver operating characteristic curve which is plotted between sensitivity and 1-specificity (false positive rate). AUC is the area under this curve. The closer its value is to 1, the better is the performance of the classifier in terms of positive predictions. FROC is a relatively new measure that determines the free-response receiver operating characteristic curve. It is a plot of operating points displaying the possible

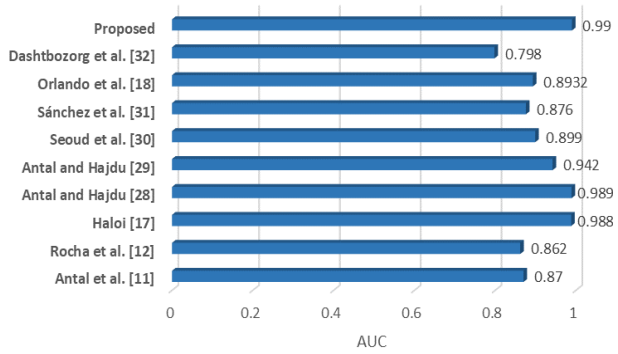

FIGURE 6. Performance comparison in terms of AUC for the Messidor dataset.

TABLE 2. Performance comparison in terms of sensitivity, specificity and accuracy for the Messidor data set.

Technique	Sensitivity (%)	Specificity (%)	Accuracy (%)
Antal et al. [11]	76	88	82
Haloi [17]	97	**96**	96
Antal and Hajdu [28]	90	91	90
Antal and Hajdu [29]	94	90	90
Seoud et al. [30]	93.9	50	-
Orlando et al. [18]	91.09	50	-
Proposed	**98**	**96**	**97**

tradeoff between the sensitivity versus the average number of false positive detection per image.

Figure 6 demonstrates the performance of the proposed technique in terms of AUC using the Messidor dataset. Some state of the art techniques from the literature are used for the comparison. It can be observed that the proposed technique out performs the existing works in terms of the AUC with a value of 0.99. For techniques such as [31] and [32], there is a significant difference. Table 1 presents the comparison in terms of sensitivity, specificity and the accuracy using the Messidor dataset. The fields marked with '-' represent the unavailability of the results for those measures. Risk stratification is not taken into account by any of the techniques. For a particular measure, the bold values represent the best results. It can be observed that for the sensitivity and accuracy, the proposed IFST technique gives the best results with values of 98% and 97% respectively. For the specificity, the proposed technique and the technique proposed by Haloi [17] yield the same result of 96%.

Figure 7 shows the comparison of the FROC curves of different techniques using the e-ophtha-MA dataset. It can be observed that the sensitivity of the proposed technique increases with the increase in the number of false positives per image. The proposed technique performs slightly better than the one proposed by Eftekhari *et al.* [19], and much higher performance is demonstrated with respect to the other techniques such as [33] and [3]. This improvement with a wide margin in performance is attributable to the fact that the proposed technique exploits hidden dependencies in the search space by exhaustively searching the attributes and

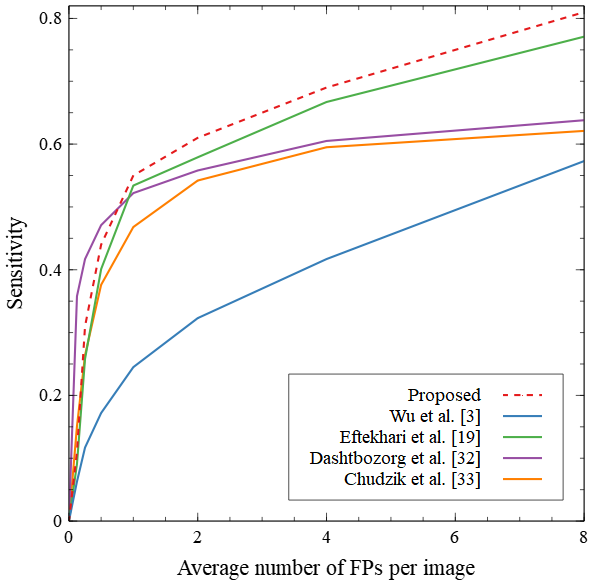

FIGURE 7. Performance comparison in terms of FROC for the e-ophtha-MA dataset.

TABLE 3. Performance comparison in terms of sensitivity and specificity using DIARETDB1 data set.

Technique	Sensitivity (%)	Specificity (%)
Romero et al. [15]	92.32	93.87
Manjaramkar and Kokare [34]	84.15	93.5
Adal et al. [35]	64.62	-
Proposed	**93.62**	**94.6**

their effects using various mathematical dependencies and expressions, which are otherwise overlooked in the traditional methods.

For the DIARETDB1 database, we used 50 images for the training of the proposed technique and the remaining 39 images are used for the testing purpose. Table 3 presents the results achieved by using the DIARETDB1 database. In order to compare classification performance, sensitivity and specificity values are computed using different techniques proposed in literature. It can be seen that the proposed IFST technique yields enhanced values of sensitivity and specificity as compared to the contemporary techniques. Comparatively, the technique proposed by Rosas-Romero *et al.* [15] yields close enough sensitivity with a value of 92.32% compared to the proposed technique's sensitivity of 93.62%. Also, their specificity value of 93.87% is very close to the proposed technique's value of 94.6%. However, there is a notable over all enhancement in results using the proposed technique.

Figure 8 demonstrates the ROC of different techniques using the DIARETDB1 dataset. It can be observed that there is a significant improvement in the ROC curve generated by the proposed technique when compared to the techniques proposed by Manjaramkar and Kokare [34] and Adal *et al.* [35]. On the other hand, there is a marginal improvement by

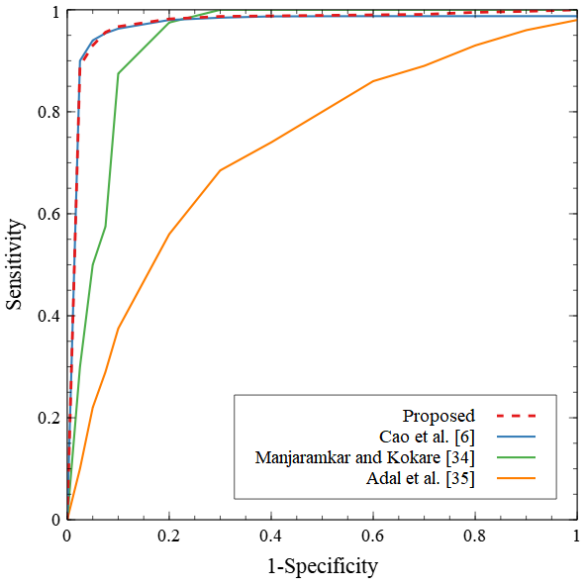

FIGURE 8. Performance comparison in terms of ROC for the DIARETDB1 data set.

the proposed technique when compared to the work by Cao *et al.* [6]. From table 3 and figure 8, it is noticeable that the improvement in classification is less in the case of DIARETDB1 as compared to other data sets. This could be primarily due to the initial random population generated by GP or the size and quality of the dataset itself. Nevertheless, similar behavior in results is observed by the other techniques likewise.

IV. CONCLUSION

This work presents a new intelligent technique for the automatic detection of MAs. At the preprocessing stage, green channel is extracted initially and median filtering is applied to remove random noise. In order the make certain areas more visible, CLAHE based enhancement is utilized which is followed by Gaussian filtering and shade correction in order to remove residual noise and enhance the image contrast, respectively. The proposed technique then utilizes GP and a set of 28 features to exploit hidden dependencies among the features which are otherwise overlooked by the conventional techniques. The best mathematical expression is then evolved following the binearization of the fitness score and GP based stepwise enhancement over the generations. Experimental results and comparison with state of the art recent techniques validate that the evolved expression is generalized and can be used as a classifier for real world applications. The implication of the current study on clinical applications is yet to be established and is a subject of future research.

REFERENCES

[1] *Diabetes*. Accessed: Aug. 22, 2019. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/diabetes

[2] L. Guariguata, D. R. Whiting, I. Hambleton, J. Beagley, U. Linnenkamp, and J. E. Shaw, "Global estimates of diabetes prevalence for 2013 and projections for 2035," *Diabetes Res. Clin. Pract.*, vol. 103, no. 2, pp. 137–149, Feb. 2014.

[3] B. Wu, W. Zhu, F. Shi, S. Zhu, and X. Chen, "Automatic detection of microaneurysms in retinal fundus images," *Computerized Med. Imag. Graph.*, vol. 55, pp. 106–112, Jan. 2017.

[4] N. Salamat, M. M. S. Missen, and A. Rashid, "Diabetic retinopathy techniques in retinal images: A review," *Artif. Intell. Med.*, vol. 97, pp. 168–188, Jun. 2019.

[5] C. Pereira, D. Veiga, J. Mahdjoub, Z. Guessoum, L. Gonçalves, M. Ferreira, and J. Monteiro, "Using a multi-agent system approach for microaneurysm detection in fundus images," *Artif. Intell. Med.*, vol. 60, no. 3, pp. 179–188, Mar. 2014.

[6] W. Cao, N. Czarnek, J. Shan, and L. Li, "Microaneurysm detection using principal component analysis and machine learning methods," *IEEE Trans. Nanobiosci.*, vol. 17, no. 3, pp. 191–198, Jul. 2018.

[7] C. E. Baudoin, B. J. Lay, and J. C. Klein, "Automatic detection of microaneurysms in diabetic fluorescein angiography," *Rev. Epidemiol. Sante Publique*, vol. 32, nos. 3–4, pp. 254–261, 1984.

[8] T. Spencer, R. P. Phillips, P. F. Sharp, and J. V. Forrester, "Automated detection and quantification of microaneurysms in fluorescein angiograms," *Graefe's Arch. Clin. Experim. Ophthalmology*, vol. 230, no. 1, pp. 36–41, Jan. 1992.

[9] T. Spencer, J. A. Olson, K. C. McHardy, P. F. Sharp, and J. V. Forrester, "An image-processing strategy for the segmentation and quantification of microaneurysms in fluorescein angiograms of the ocular fundus," *Comput. Biomed. Res.*, vol. 29, no. 4, pp. 284–302, Aug. 1996.

[10] A. D. Fleming, S. Philip, K. A. Goatman, J. A. Olson, and P. F. Sharp, "Automated microaneurysm detection using local contrast normalization and local vessel detection," *IEEE Trans. Med. Imag.*, vol. 25, no. 9, pp. 1223–1232, Sep. 2006.

[11] B. Antal, I. Lazar, A. Hajdu, Z. Torok, A. Csutak, and T. Peto, "Evaluation of the grading performance of an ensemble-based microaneurysm detector," in *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, Aug. 2011, pp. 5943–5946.

[12] A. Rocha, T. Carvalho, H. F. Jelinek, S. Goldenstein, and J. Wainer, "Points of interest and visual dictionaries for automatic retinal lesion detection," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 8, pp. 2244–2253, Aug. 2012.

[13] I. Lazar and A. Hajdu, "Retinal microaneurysm detection through local rotating cross-section profile analysis," *IEEE Trans. Med. Imag.*, vol. 32, no. 2, pp. 400–407, Feb. 2013.

[14] R. Srivastava, D. W. K. Wong, L. Duan, J. Liu, and T. Y. Wong, "Red lesion detection in retinal fundus images using frangi-based filters," in *Proc. 37th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc. (EMBC)*, Aug. 2015, pp. 5663–5666.

[15] R. Rosas-Romero, J. Martínez-Carballido, J. Hernández-Capistrán, and L. J. Uribe-Valencia, "A method to assist in the diagnosis of early diabetic retinopathy: Image processing applied to detection of microaneurysms in fundus images," *Computerized Med. Imag. Graph.*, vol. 44, pp. 41–53, Sep. 2015.

[16] N. Asiri, M. Hussain, F. Al Adel, and N. Alzaidi, "Deep learning based computer-aided diagnosis systems for diabetic retinopathy: A survey," *Artif. Intell. Med.*, vol. 99, Aug. 2019, Art. no. 101701.

[17] M. Haloi, "Improved microaneurysm detection using deep neural networks," 2015, *arXiv:1505.04424*. [Online]. Available: http://arxiv.org/abs/1505.04424

[18] J. I. Orlando, E. Prokofyeva, M. del Fresno, and M. B. Blaschko, "An ensemble deep learning based approach for red lesion detection in fundus images," *Comput. Methods Programs Biomed.*, vol. 153, pp. 115–127, Jan. 2018.

[19] N. Eftekhari, H.-R. Pourreza, M. Masoudi, K. Ghiasi-Shirazi, and E. Saeedi, "Microaneurysm detection in fundus images using a two-step convolutional neural network," *Biomed. Eng. OnLine*, vol. 18, no. 1, Dec. 2019.

[20] M. M. Habib, R. A. Welikala, A. Hoppe, C. G. Owen, A. R. Rudnicka, and S. A. Barman, "Detection of microaneurysms in retinal images using an ensemble classifier," *Informat. Med. Unlocked*, vol. 9, pp. 44–57, Jan. 2017.

[21] U. Budak, A. Sengür, Y. Guo, and Y. Akbulut, "A novel microaneurysms detection approach based on convolutional neural networks with reinforcement sample learning algorithm," *Health Inf. Sci. Syst.*, vol. 5, no. 1, p. 14, Dec. 2017.

[22] A. M. Reza, "Realization of the contrast limited adaptive histogram equalization (CLAHE) for real-time image enhancement," *J. VLSI Signal Process.-Syst. Signal, Image, Video Technol.*, vol. 38, no. 1, pp. 35–44, Aug. 2004.

[23] *GPLAB—A Genetic Programming Toolbox for MATLAB*. Accessed: Nov. 22, 2019. [Online]. Available: http://gplab.sourceforge.net/
[24] E. Decencière, X. Zhang, G. Cazuguel, B. Lay, B. Cochener, C. Trone, P. Gain, R. Ordonez, P. Massin, A. Erginay, B. Charton, and J.-C. Klein, "Feedback on a publicly distributed image database: The Messidor database," *Image Anal. Stereology*, vol. 33, no. 3, p. 231, 2014.
[25] *Diaretdb1—Standard Diabetic Retinopathy Database*. Accessed: Dec. 12, 2019. [Online]. Available: http://www2.it.lut.fi/project/imageret/diaretdb1/
[26] *TeleOphta-Homepage*. Accessed: Dec. 12, 2019. [Online]. Available: http://www.cmm.mines-paristech.fr/Projects/TeleOphta/index.html
[27] *TeleOphta: Machine Learning and Image Processing Methods for Teleophthalmology—ScienceDirect*. Accessed: Jun. 11, 2018. [Online]. Available: https://www-sciencedirect-com.sdl.idm.oclc.org/science/article/pii/S1959031813000237?via%3Dihub
[28] B. Antal and A. Hajdu, "An ensemble-based system for automatic screening of diabetic retinopathy," *Knowl.-Based Syst.*, vol. 60, pp. 20–27, Apr. 2014.
[29] B. Antal and A. Hajdu, "An ensemble-based system for microaneurysm detection and diabetic retinopathy grading," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 6, pp. 1720–1726, Jun. 2012.
[30] L. Seoud, T. Hurtut, J. Chelbi, F. Cheriet, and J. M. P. Langlois, "Red lesion detection using dynamic shape features for diabetic retinopathy screening," *IEEE Trans. Med. Imag.*, vol. 35, no. 4, pp. 1116–1126, Apr. 2016.
[31] C. I. Sánchez, M. Niemeijer, A. V. Dumitrescu, M. S. A. Suttorp-Schulten, M. D. Abràmoff, and B. van Ginneken, "Evaluation of a computer-aided diagnosis system for diabetic retinopathy screening on public data," *Investigative Opthalmology Vis. Sci.*, vol. 52, no. 7, p. 4866, Jun. 2011.
[32] B. Dashtbozorg, J. Zhang, F. Huang, and B. M. ter Haar Romeny, "Retinal microaneurysms detection using local convergence index features," *IEEE Trans. Image Process.*, vol. 27, no. 7, pp. 3300–3315, Jul. 2018.
[33] P. Chudzik, S. Majumdar, F. Calivá, B. Al-Diri, and A. Hunter, "Microaneurysm detection using deep learning and interleaved freezing," in *Proc. Med. Imag. , Image Process.*, Mar. 2018, Art. no. 105741.
[34] A. Manjaramkar and M. Kokare, "Statistical geometrical features for microaneurysm detection," *J. Digit. Imag.*, vol. 31, no. 2, pp. 224–234, Apr. 2018.
[35] K. M. Adal, D. Sidibé, S. Ali, E. Chaum, T. P. Karnowski, and F. Mériaudeau, "Automated detection of microaneurysms using scale-adapted blob analysis and semi-supervised learning," *Comput. Methods Programs Biomed.*, vol. 114, no. 1, pp. 1–10, Apr. 2014.

IMRAN USMAN (Senior Member, IEEE) received the B.E. degree in software engineering from Foundation University, Pakistan, in 2003, the M.S. degree in computer system engineering from the Ghulam Ishaq Khan Institute of Engineering Sciences and Technology, Pakistan, in 2006, and the Ph.D. degree from the Pakistan Institute of Engineering and Applied Sciences, in 2010. He joined the Pakistan Institute of Engineering and Applied Sciences as a Research Scholar. From 2009 to 2010, he was an Assistant Professor with the Department of Computing and Technology, Iqra University, Islamabad, Pakistan. From 2010 to 2012, he was an Assistant Professor and a Senior In-Charge Graduate Program with the Department of Electrical Engineering, COMSATS Institute of Information Technology Islamabad, Pakistan. He is currently an Assistant Professor with the College of Computing and Informatics, Saudi Electronic University, Saudi Arabia. His current research interests include machine learning, digital image processing, evolutionary computation, and digital watermarking. He has a number of research articles to his credit and has supervised many B.S., M.S., and Ph.D. Students.

KHALED A. ALMEJALLI received the master's degree from the Department of Computing, University of Bradford, U.K., in 2011, and the Ph.D. degree from the AI Research Centre, University of Bradford, in 2011. Prior to his master's and Ph.D. study, he worked at the Road Traffic Control Department, Riyadh, Saudi Arabia. He is currently an Assistant Professor with the Faculty of Computing and Informatics, Saudi Electronic University (SEU), Saudi Arabia, where he is also the Dean of IT. His research interests include the areas of optimization and decision support technologies, AI applications, evolutionary computation, and road-traffic control applications. He has published some interesting articles in peer-reviewed journals and conference proceedings.