

Genetic Algorithm Based Feature Selection Combined with Dual Classification for the Automated Detection of Proliferative Diabetic Retinopathy

R.A. Welikala^{a,*}, M.M. Fraz^a, J. Dehmeshki^a, A. Hoppe^a, V. Tah^b, S. Mann^c, T.H. Williamson^c, S.A. Barman^a

^a Digital Imaging Research Centre, Faculty of Science, Engineering and Computing, Kingston University, London, United Kingdom.

^b Medical Retina, Oxford Eye Hospital, Oxford, United Kingdom.

^c Ophthalmology Department, St Thomas' Hospital, London, United Kingdom.

* Corresponding author. Tel.: +44 7944990065. E-mail address: Roshan.Welikala@kingston.ac.uk (R.A. Welikala).

Abstract:

Proliferative diabetic retinopathy (PDR) is a condition that carries a high risk of severe visual impairment. The hallmark of PDR is the growth of abnormal new vessels. In this paper, an automated method for the detection of new vessels from retinal images is presented. This method is based on a dual classification approach. Two vessel segmentation approaches are applied to create two separate binary vessel maps which each hold vital information. Local morphology features are measured from each binary vessel map to produce two separate 4-D feature vectors. Independent classification is performed for each feature vector using a support vector machine (SVM) classifier. The system then combines these individual outcomes to produce a final decision. This is followed by the creation of additional features to generate 21-D feature vectors, which feed into a genetic algorithm based feature selection approach with the objective of finding feature subsets that improve the performance of the classification. Sensitivity and specificity results using a dataset of 60 images are 0.9138 and 0.9600 respectively on a per patch basis and 1.000 and 0.975 respectively on a per image basis.

Keywords:

Retinal images, proliferative diabetic retinopathy, new vessels, dual classification, feature selection, genetic algorithm.

1 INTRODUCTION:

Diabetes mellitus is a disorder of sugar metabolism caused by an absolute lack of insulin or an insufficient action of insulin and hence is characterized by raised levels of glucose in the blood. High blood glucose levels (Hyperglycemia) can damage the vessels that supply blood to vital organs. Diabetic retinopathy (DR) is the resultant disorder affecting the retinal vasculature, leading to progressive retinal damage that can end in loss to vision and blindness [1]. DR is the most frequent cause of new cases of blindness among adults aged 20-74 years [2]. The problem is increasing in its scale, with diabetes identified as a significant growing global public health problem [3]. In the United Kingdom alone, three million people are estimated to have diabetes and this figure is expected to double in the next 15-30 years [4].

Diabetic patients are required to attend regular eye screening appointments in which DR can be assessed, with the intention of early detection of the disease to allow for timely intervention [5,6]. During these appointments retinal images are captured and figure 1 shows examples of such images. These images then undergo various stages of manual assessment by trained individuals [7]. This assessment can be a very time consuming and costly task due to the large diabetic population. Therefore this is a field that would greatly benefit from the introduction of automated detection systems [8].

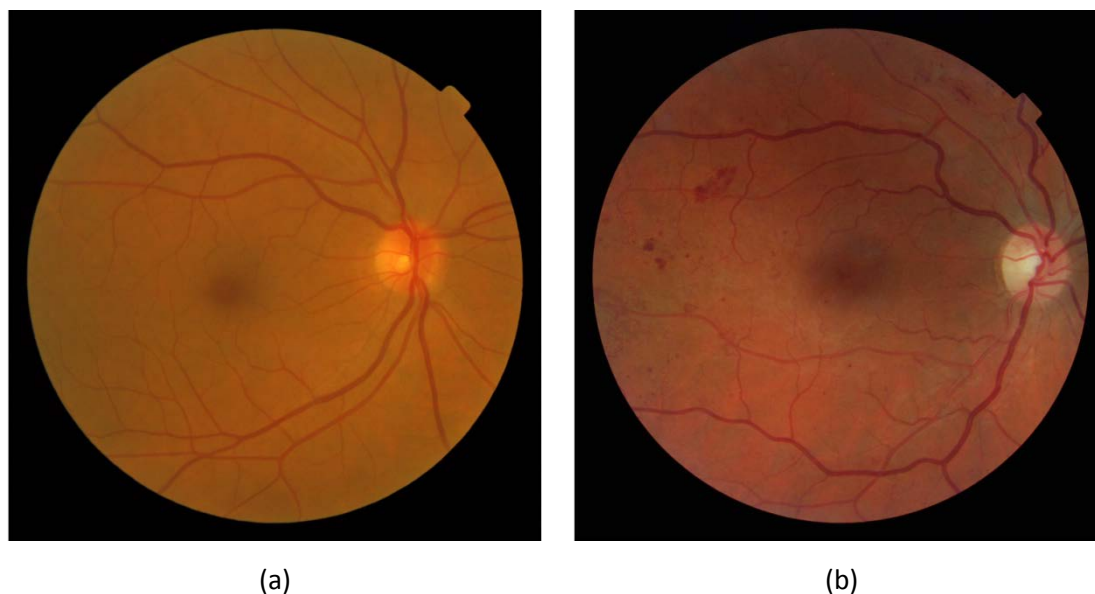


Figure 1: (a) Healthy retinal image. (b) Retinal image with DR.

The damage to the retinal blood vessels will cause blood and fluid to leak on the retina and form features such as microaneurysms, haemorrhages, exudates, cotton wool spots and venous loops [9]. With progression, the blockages and damage to blood vessels deprive areas of the retina with their blood supply. These areas of the retina send signals to the body to grow new blood vessels for nourishment. New vessels are the hallmark of proliferative diabetic retinopathy (PDR), which is the most advanced stage of DR. PDR poses a high risk of severe vision loss due to the fragile nature of the new vessels making them prone to bleed and cause pre-retinal and vitreous haemorrhages [10]. Patients presenting PDR require an urgent referral to an ophthalmologist.

New vessels are divided into two categories, new vessels at the optic disc (NVD) and new vessels elsewhere (NVE). They tend to be fine in calibre and are more tortuous and convoluted than normal vessels. Initially they appear as loops or small networks that are located on the optic disc or near a vein. As they grow they form dense lacy networks which usually pass across the underlying veins and arteries [1]. New vessels tend to grow away from the retinal surface and hence can appear out of the focal plane of the photograph, which can result in a blurry and obscure appearance. Examples of new vessels are shown in figure 2.

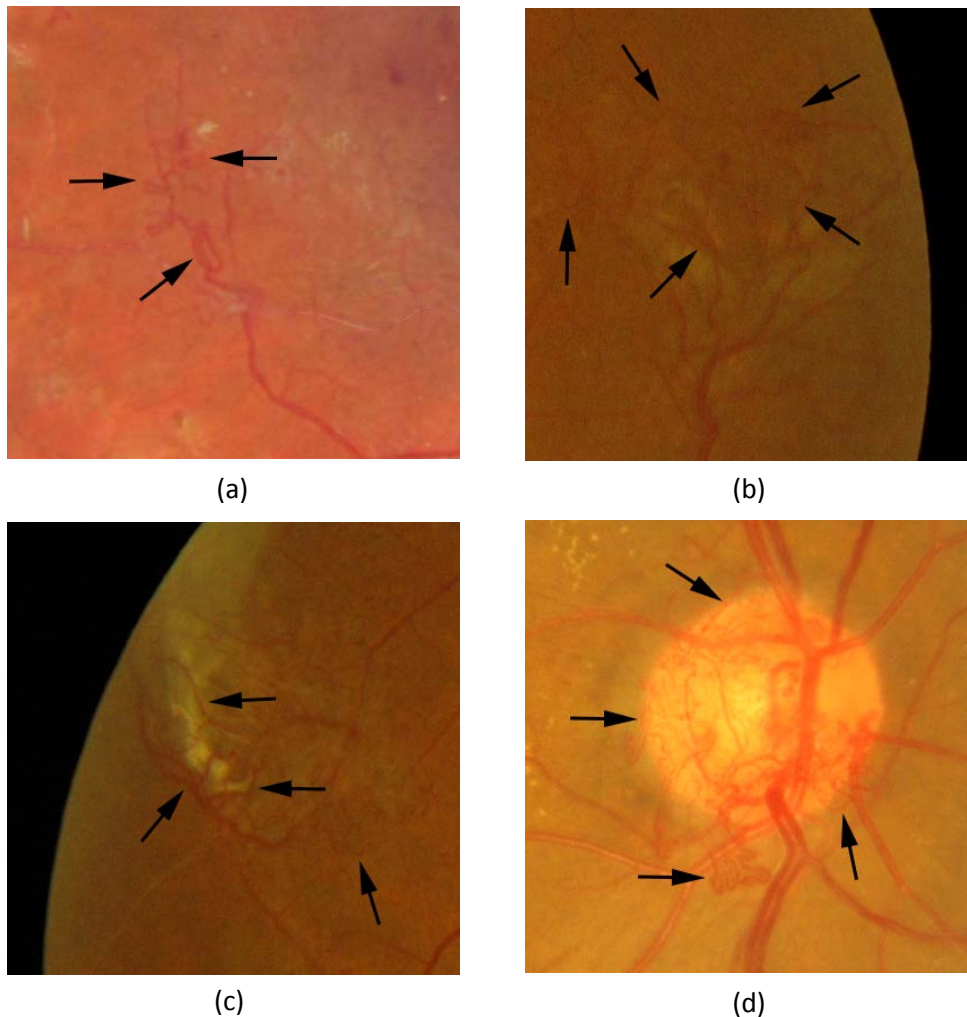


Figure 2: Zoom-in regions of retinal images. (a)-(c) New vessels elsewhere, NVE. (d) New vessels at the optic disc, NVD. Arrows indicate examples of new vessels within the images.

Large scale audits of disease/no disease automated grading systems have shown the benefits they provide [11]. An additional aim is the detection of different stages of DR, which should include the capability of detecting and prioritising PDR images to ensure immediate referral to a specialist. There are many studies investigating the automatic detection of DR focused on microaneurysm and haemorrhage detection [12-16], and exudate detection [17-20]. Conversely, research on the detection of PDR is relatively rare.

New vessel detection methods can be split into two categories, based on whether vessel segmentation is performed or not. Those methods based on vessel segmentation are developed with the purpose of analysing the morphology of the binary vessel map in search of abnormality. The other category is methods based on extracting textural information from the images and therefore avoiding the difficulties that arise from vessel segmentation.

Vessel segmentation has received the largest share of attention in the field of retinal image analysis, studies include [21-28]. A comprehensive review of this mature field of vessel segmentation is provided by Fraz [29]. Studies have shown that vessel calibre relates to hypertension and cardiovascular disease [30]. The main driving force for accurate segmentation has been for the quantification of vessel calibre [31] for cardiovascular studies. Vessel segmentation also forms the backbone for many automated systems aimed at diagnosing ophthalmic disease. However, vessel segmentation techniques struggle to extract new vessels due to their fine calibre and irregular appearance. Also most techniques do not put enough emphasis on removing false responses due to artefacts and other lesions.

The following vessel segmentation techniques were designed with PDR taken into consideration. Ramlugun [32] described a small vessel extraction technique, the main contribution was the varying of the clip limit for contrast limited adaptive histogram equalization (CLAHE) to allow more contrast for small vessels. B.Zhang [33] applied the matched filter with the first-order derivative of the Gaussian. The main emphasis was not on the increased segmentation of new vessels, but instead the reduction of the false response to exudates which can cause large local densities on the segmented map and therefore can be mistaken for new vessels. L.Zhang [34] proposed a modified matched filter that used double sided thresholding to reduce the false response to exudates. Figure 3 shows an example of exudates, also known as bright lesions. Akram [35] proposed the use of the Gabor wavelet for vessel enhancement followed by a multilayered thresholding technique.

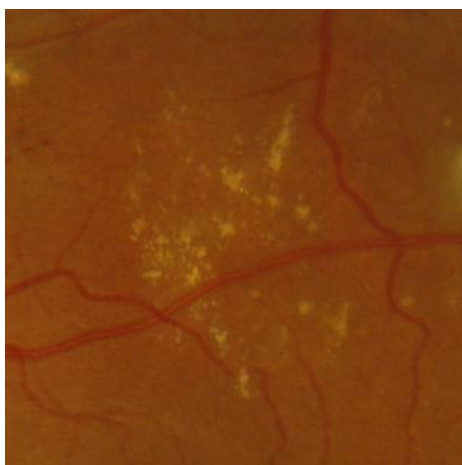


Figure 3: Zoom-in region of a retinal image presenting exudates.

The following new vessel detection methods are categorised as those performing vessel segmentation prior to the described analysis methods. Hassan [36] proposed a region based technique where the number of vessels and the area of vessels within a small scanning sub-window were used to indicate new vessels. Welikala [37] also produced a region based technique using five

local morphology features. Also included was the prior step of straight vessel removal in order to remove the majority of normal vasculature and therefore simplifying new vessel detection. A comprehensive set of 15 features was developed by Goatman [38] including the number of vessel segments, the mean vessel wall gradient and various tortuosity measures in order to detect NVD. Jelinek [39] used data obtained from the application of the derivatives of Gaussian wavelets to the vessel skeleton to extract morphological based features. Daxer [40] described the retinal vasculature as a fractal and used the fractal dimension to quantify its complexity to indicate the presence of new vessel growth. Karperien [41] furthered this with the analysis of local dimensions using the local connected fractal dimension. Akram [42] proposed a multivariate m-Medoids based classifier with a ten dimensional feature set based on morphological, intensity and gradient based values. Saranya [43] created a feature vector that involved the use of Hu moments for the detection of new vessels. Oloumi [44] proposed the use of modelling the major temporal arcade for the diagnosis of proliferative diabetic retinopathy.

The methods described next are categorised as not performing vessel segmentation and therefore avoid their associated difficulties. Frame [45] applied statistical texture measures, calculated using the grey level co-occurrence matrix (GLCM), to identify irregular distributions of pixel intensities associated with neovascularisation. Acharya [46] calculated texture features from the run length matrix, as well as the GLCM to identify the stage of DR. Multi-scale amplitude modulation frequency modulation (AM-FM) methods were utilised by Agurto [47] for spectral texture analysis to characterise different retinal structures, including new vessels. However, Agurto [48] extended their work to involve AM-FM along with vessel segmentation and granulometry to detect NVD. Vatanparast [49] compared the performance of six different texture based methods for the detection of neovascularisation. These include GLCM, AM-FM, Gabor filters, Contourlet transform and local binary patterns. Lee [50] proposed a new vessel detection method which combined statistical texture analysis, high order spectrum analysis and fractal analysis. However vessel segmentation was required for multi-fractal analysis (not for mono-fractal analysis).

There exist numerous techniques [51,52] that measure general aspects of the morphology of the retinal vasculature that were not directly designed to detect new vessels, however certain aspects are relevant. There also exist techniques developed from other research topics that are relevant to PDR detection. An automated method for the quantification of micro-vessel density within the inner surface of egg shells in order to study the angiogenesis in developing chick embryos was created by Doukas [53]. Measures included branching points, vessel length and GLCM textural information. A system using edge contour analysis is presented by Zutis [54] for detecting abnormal retinal capillary regions, with the focus on telangiectasia.

The contribution of [55] was the novel application of a dual classification approach to independently process the binary maps from two different vessel segmentation methods with the aim to detect new vessels and reduce false responses caused by other retinal features. This included a novel modified line operator, based on double sided thresholding, designed to segment vessels whilst reducing false responses to non-vessel edges.

The main contribution of the proposed method is the exploration of features suitable for the classification of proliferative diabetic retinopathy combined with the dual classification approach [55]. This is achieved by expanding the feature vector which now includes morphology, gradient and

intensity based features. This is followed by feature selection using a genetic algorithm based approach to find the most suitable feature subsets. The organization of this paper is as follows: section 2 provides details of the methodology, including a concise overview of the dual classification approach; section 3 presents the experimental evaluation; finally a discussion and conclusion is given in section 4.

2 METHODOLOGY

2.1 Modified Line Operator and Dual Classification

The architecture of dual classification approach is shown in figure 4. Following spatial normalization and pre-processing the system splits into two pathways as two different vessel segmentation methods were applied to create two binary vessel maps. Each pathway had its own feature set produced, picked from the same pool of local features. Independent classification was performed for each pathway using a support vector machine classifier. The system produced a final decision by combining the two individual classification outcomes in which regions of the retina were labelled as new vessels or non-new vessels. This section describes a very concise overview, for a full explanation with all parameters listed see [55].

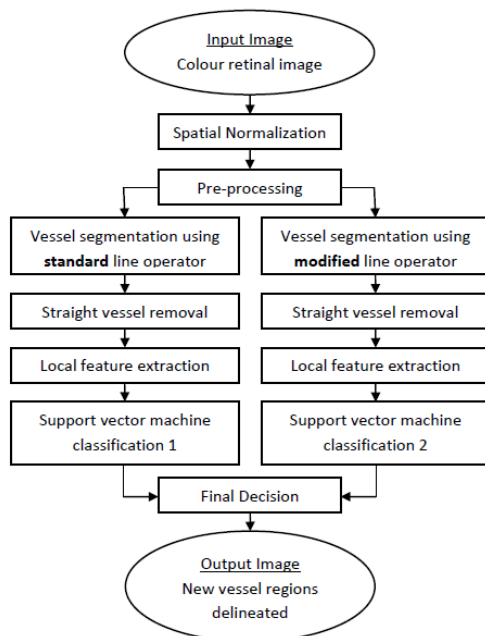


Figure 4: System architecture of the dual classification approach.

2.1.1 Line Operator/Modified Line Operator:

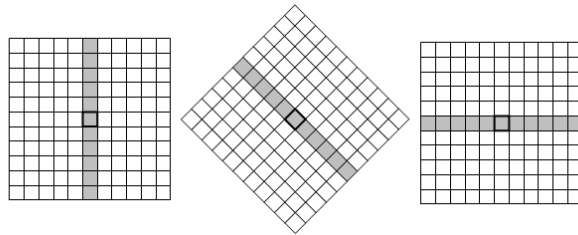


Figure 5: Standard line operator at 3 different orientations. The coefficients on the line were each valued 1 and the remaining coefficients of the mask were each valued 0.

The line operator was used in the proposed method to perform vessel segmentation. The standard line operator is illustrated in figure 5. The average grey-level of the pixels along a line passing through the target pixel was calculated for multiple orientations. The line strength of the pixel was obtained from the orientation with the largest value subtracted by the average grey-level of the similar orientated neighbourhood. The line strength was large if the winning line was aligned with a vessel. An empirically derived threshold was applied to the line strength image to produce the segmentation of the vessels. The modified line operator differed by being based on three line strength measures as opposed to one. This included measures that separately took into account the left and right side of the neighbourhood. For full details and equations see [55].

Figures 6(d)-(f) show that the standard line operator segments normal and new vessels with a high level of accuracy, however it also responds to non-vessel edges (mainly bright lesions and reflection artefacts). These false vessel detections often cause large local densities and large curvatures which are indistinguishable to new vessels. Figures 6(g)-(i) show the segmentation generated by the modified line operator has significantly reduced the false vessel detections to the extent that non vessel edges were distinguishable to new vessels. Unfortunately the segmentation of new vessels has worsened. This meant that new vessels were no longer so distinguishable to the normal vasculature. Both of the segmentation methods showed disadvantages and therefore neither method alone was suitable for the detection of new vessels. However extraction of information from both maps could be used effectively and this lead to the development of the dual classification approach.

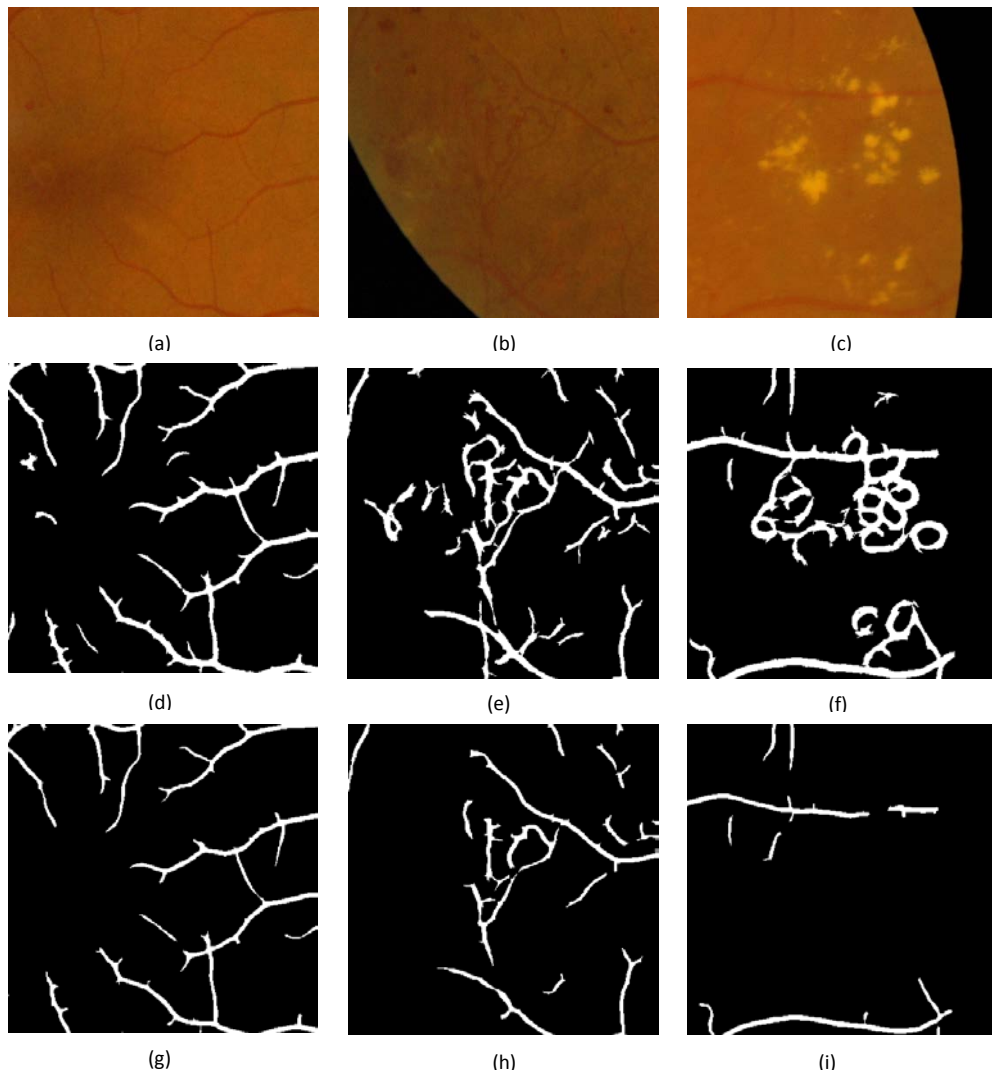


Figure 6: Zoom-in regions of retinal images: (a) Normal vessels. (b) New vessels. (c) Bright lesions. (d)-(f) Segmentation of (a)-(c) using the standard line operator. (g)-(i) Segmentation of (a)-(c) using the modified line operator.

2.1.2 Feature Extraction

The design of this method is aimed at the classification of image regions that contain new vessels. These image regions can be described as containing many vessel segments, which are closely spaced and appear in multiple orientations.

Prior to measurements of features, the binary maps were simplified. This included straight vessel removal and the creation of vessel segments (single pixel in thickness). A sub window of size 151 x 151 pixels was created in order to calculate local features associated with the morphology of the vasculature. This sub window was scanned through the image and at each pixel position the following four features were calculated: number of vessel pixels, number of vessel segments, number of vessel orientations and vessel density. This same set of features was measured from the

binary vessel map from each the standard and modified approach to produce two separate feature sets.

2.1.3 Dual Classification

All features were normalised so that each feature had zero mean and unit standard deviation. Independent classification was performed for each of the two feature sets using a support vector machine (SVM) classifier [56,57]. Each classifier independently labelled the candidate pixel as new vessels or non-new vessels. The system produced a final decision by combining the outcomes. The candidate pixel achieved a new vessel label only when both classifications agreed on its identity being new vessels, otherwise it achieved a non-new vessel label. When complete, all pixels labelled as new vessels were dilated with a structuring element the size of the sub window to illustrate the new vessel regions.

Whilst only two classes were used, new vessels and non new vessels, both feature sets and their independent classification were not intended to distinguish the same two cases. Classifier 1, associated with the feature set measured from the standard line operator approach, was intended to distinguish new vessels from normal vessels. Classifier 2, associated with the feature set measured from the modified line operator approach, was intended to distinguish new vessels from exudates. Combining the outcomes then removed the false new vessel responses that each classifier made.

It should be noted that the optimal SVM kernel and parameters were determined by a cross-validation grid search, with the kernel types of linear, Gaussian radial basis function and polynomial being tested. The grid search was not an aspect covered in the prior work reported in [55], therefore results described will not correspond exactly to those reported from [55].

From the above it is evident that two individual decisions had to be made, hence keeping the feature sets separate was important to ensure an improved performance was achieved. Alternatively, a conventional single classification could have been applied by combining all features into a single feature set and the classifier's decision boundary could attempt to deal with making both decisions. However, such a boundary would compromise each of the decisions.

Consider that there were three cases, new vessels, normal vessels and exudates. Hence, a one-versus-one multiclass SVM has similarities to our approach as it is based on independent classifications between each of the classes. However it would be disadvantaged as each classification would be made within a combined single feature space, where as the dual approach used only the features relevant to each classification.

To validate our dual classification system its performance was compared to the performance from a single SVM classification and a one-versus-one multiclass SVM classification. The results for this comparison were not documented in [55].

2.2 Feature Vector Expansion

There are other characteristics that can help to distinguish new vessels aside from those mentioned in section 2.1.2. New vessels tend to appear finer in calibre, shorter in length and possess a tortuous nature. New vessels also appear less homogeneous than normal vessels. Therefore, additional features were created based on these characteristics. This formed a 21-D feature vector which contained morphology, intensity and gradient based features.

All features were measured from information extracted from the sub window. This sub window scanned through the image and at each pixel position a 21-D feature vector was calculated. This same set of features was measured for each of the pathways to produce two separate feature vectors. Some features were designed specifically for a particular classification, however for simplicity both feature vectors remained identical.

The description of the features is listed below. The term segment refers to the skeletonised vessel segments.

1) Number of vessel pixels

The sum of all segment pixels within the sub window.

2) Number of vessel segments

The number of whole and partially included segments within the sub window.

3) Number of vessel orientations

The end points of a segment were connected by a straight line. The angle the line made with the x-axis that fell within the range -90° to 90° of the unit circle was calculated. The calculated angle was accordingly dropped into one of eight bins, each representing a range of angles. This was done for each segment within the sub window and the number of non-empty bins represented the number of orientations.

4) Vessel density

A segment was dilated with a disk structuring element with a radius of 20 pixels. The number of vessel pixels within the dilated area was divided by the number of pixels within the segment to give its vessel density. This was done for each segment within the sub window and the mean vessel density was calculated.

5) Tortuosity 1

The tortuosity of each segment was calculated using the true length (measured with the chain code) divided by the Euclidean length. The mean tortuosity was calculated from all the segments within the sub window.

6) Tortuosity 2

The maximum tortuosity amongst all segments in the sub window.

7) Vessel length

The mean true length (in pixels) of all segments within the sub window.

8) Number of bifurcation points

The number of bifurcation points removed within the sub window when creating segments.

9) Grey level mean

The mean grey level of all segment pixels within the sub window.

10) Grey level coefficient of variation

The ratio of the standard deviation to the mean of the grey level of all segment pixels within the sub window.

11) Gradient mean

The mean gradient magnitude along all segment pixels within the sub window calculated using the Sobel gradient operator applied on the pre-processed image.

12) Gradient coefficient of variation

The ratio of the standard deviation to the mean of the gradient of all segment pixels within the sub window.

13) Line strength

The mean line strength of all segment pixels within the sub window.

14) Vessel width

Skeletonization correlates to vessel centre lines. The distance from the segment pixel to the closest boundary point of the vessel using the vessel map prior to skeletonization. This gives the half-width at that point which is then multiplied by 2 to achieve the full vessel width. The mean is calculated for all segment pixels within the sub window.

15) Mean vessel wall gradient

As for the vessel width above, the closest boundary point was assumed to be the vessel wall. The mean gradient magnitude along all vessel wall pixels within the sub window.

16) Vessel wall gradient coefficient of variation

The ratio of the standard deviation to the mean of the vessel wall gradient along all vessel wall pixels within the sub window.

17) Compactness

The full vessel map prior to skeletonization and straight vessel removal was used. Area and perimeter within the sub window were measured and used to calculate circularity ($4\pi \cdot \text{area} / \text{perimeter}^2$). Circularity is also a measure of compactness.

18) Connectivity

The full vessel map prior to skeletonization and straight vessel removal was used. The vessel area was divided by the number of objects within the sub window.

19) Local grey level mean

The mean grey level of all pixels within the sub window using the pre-processed image.

20) Local grey level max

The maximum grey level amongst all pixels within the sub window using the pre-processed image.

21) Local grey level variation

The standard deviation of all pixels within the sub window using the pre-processed image.

2.3 Genetic Algorithm Based Feature Selection and Parameter Optimization

Feature selection is the process of selecting the smallest subset of features that is necessary for accurate prediction (classification or regression). It achieves this by the removal of redundant and irrelevant features. Redundant features are those which provide no further information than the currently selected features, and irrelevant features provide no useful information in any context. The benefits of feature selection are improving the classifier's performance and providing a better understanding of the underlying process that generated the data [58]. Using a small number of features will also save significant computation time and builds models that generalise better for unseen data.

An exhaustive search of all possible subsets of features to find the one which maximises the classifier's performance is the simplest approach. However, this approach is far too computationally expensive and is only suitable for the smallest of feature sets. Filter methods are also simple, as well as computationally fast. A statistical test is performed for individually ranking the features according to their relevance and low ranked features are then removed (filtered). However they suffer from ignoring feature dependencies and also ignore interaction with the classifier. Wrapper methods do interact with the classifier, utilizing them to score feature subsets and also model feature dependencies. Two of the most popular wrapper methods are greedy search strategies known as sequential forward selection and sequential backward selection. Genetic algorithms can also be applied as a wrapper method and shall be discussed further later. [59] provides a comprehensive review of feature selection techniques.

Before proceeding, it's worth touching on the main principle of the SVM classification. SVMs seek a linear decision surface (hyperplane) that can separate classes of objects and has the largest distance (largest gap or margin) between border-line objects (that are also called support vectors). If the classes are not linear separable the SVM maps the data in a higher dimensional space known as the feature space, where the separating linear decision surface exists and can be determined. The feature space results from a mathematical construction known as the kernel trick. There are numerous different kernel functions. Besides the standard linear kernel, the most popular kernel functions are Gaussian radial basis function kernel (RBF) and the polynomial kernel. The majority of kernels possess parameters which need to be selected. Another parameter associated with SVMs is the soft margin parameter C which is tuned to deal with noisy measurements and outliers. Therefore, the effectiveness of SVMs depends on the selection of the kernel function, the kernel parameters and the soft margin parameter C . A grid search is the conventional approach for selecting the optimal SVM parameters.

The feature subset selected influences the appropriate SVM parameters and vice versa [60]. To elaborate, each time a new feature subset is selected the input space is altered. Therefore the SVM parameters have to be re-explored to find an optimal selection. However, performing a grid search of SVM parameters each time a new feature subset is selected is a very time consuming task. To solve this problem we followed a methodology by Huang [61] which used genetic algorithms to allow for the exploration of the optimal feature subset and SVM parameters to occur simultaneously. The only requirement was the pre-selection of the kernel function type.

Genetic algorithms (GA) is a general adaptive optimization search methodology first presented by Bledsoe [62] and mathematically formalized by Holland [63], which takes inspiration from Darwin's theory of evolution. Each gene represents a variable and a sequence of genes is referred to as a chromosome. Each chromosome may be a potential solution to the optimization problem. A population of chromosomes is randomly initialized. Chromosomes are evaluated for their quality according to a predefined fitness function. New chromosomes are produced by selecting high performing chromosomes to produce offspring which retain many aspects of their parents. These offspring are formed by using two main genetic operators, crossover and mutation. Crossover is a mechanism for exchanging genes between two selected chromosomes to create new offspring. Mutation operates by modifying one or more components of a selected chromosome. Therefore, chromosomes are competing with each other and only the fittest survive. The GA obtains the chromosome providing the optimal solution after a series of iterative computations. GA can deal with large search spaces efficiently, therefore convergence is more likely to avoid local optimal solutions.

2.3.1 Chromosome Design

In the context of our optimization problem, each chromosome comprised of three parts. This was the feature subsets, the soft margin parameter C and the kernel parameters. These are each described in further detail below.

Feature subsets:

Two feature subsets were required, one for each classification in the dual process. The chromosome contained 42 genes used to represent the two sets of 21 features. These genes were denoted with value '1' to indicate the feature was selected or '0' to indicate the feature was not selected.

Soft margin parameter C :

This was represented with a gene with an integer value ranging from 1-5 which correlated to the C values of 0.01, 0.1, 1, 10, 100.

Kernel parameters:

The type of kernel parameters depended on the kernel function used, with kernel types of linear, Gaussian radial basis function (RBF) and polynomial being assessed. The type of kernel function used was pre-selected. No kernel parameters were required for the linear kernel. For the RBF kernel the scaling factor (γ) was required and was represented by a gene with an integer value ranging from 1-5 which correlated to γ of 0.0001, 0.001, 0.01, 0.1, 1. For the polynomial kernel the order of the polynomial was required and this was represented by a gene with an integer value ranging from 1-4 which correlated to the order.

It should be noted that C and γ are continuous variables. However, they were represented by discrete values in order to simplify the optimization problem by the reduction of the search space.

The SVM kernels that were chosen to be assessed, as well as the value range for the SVM parameters were based on a strategy for SVM model selection provided by Statnikov [64].

2.3.2 Fitness Function

The optimal solution that the GA produced was a chromosome with the optimal feature subsets and SVM parameters that maximised the fitness function. The criterion used to design the fitness function is shown in equation 1, with p representing the classification performance and n representing the combined number of selected features from both feature sets. Thus, a chromosome with a high classification performance and a small number of features produced a high fitness value. An empirically derived scaling factor of 1/100 was applied to n in order to ensure the main priority was on a high classification performance. The classification performance was assessed from the final decision achieved from combining the two classification outcomes. The measure used to quantify the classification performance shall be discussed in a later section.

$$Fitness\ function = p - \frac{n}{100} \quad (1)$$

2.3.3 System Architecture for the Genetic Algorithm Based Approach

A flow chart presenting the system's architecture is provided in figure 7. All features were normalised so that each feature had zero mean and unit standard deviation. The kernel function

type was pre-selected, therefore this process was separately performed for the linear, RBF and polynomial kernel functions. An explanation of the main steps follows:

- (1) Consider the box, from figure 7, labelled population as the starting point. A population of chromosomes was randomly initialized. Each chromosome comprised of a feature subset and SVM parameters.
- (2) All chromosomes in the population were each evaluated individually with the following procedure. The selected feature subset was used to accordingly alter the training and testing sets. The training set along with the selected SVM parameters were used to train the SVM classifier, while the testing set was used to calculate the classification performance. The chromosome was evaluated by the fitness function which was derived using the classification performance along with number of selected features.
- (3) The fitness functions of all chromosomes in the population were checked to see if the stopping criteria was reached. The stopping criteria was set at the maximum possible fitness function. If reached the process was ended, otherwise it proceeded to the genetic operations.
- (4) The genetic operations produced the next generation (new population). This was done by first selecting the elite chromosomes as parents, and then crossover and mutation were used to produce offspring (new population).

Steps (2)-(4) were performed iteratively until the stopping criteria was met or the preset maximum number of generations was reached.

The GA settings were chosen at a population size of 40 (initial population of 80), elite parent count of 2 and the number of generations of 10. The operations of single point crossover (crossover fraction 0.75), uniform mutation (probability rate 0.01) and roulette wheel selection were performed.

The efficiency of GA is greatly dependent on the parameters selected. There are several publications that state recommended standard parameter settings. However, no recommended settings are truly universal, instead parameters have to be generally tuned to specific problems. Therefore, the GA parameter values stated above were achieved by starting with the standard values provided by De Jong [65] and Grefenstette [66]. This was then followed by parameter tuning. A limitation was the number of generations had to remain low due to time constraints.

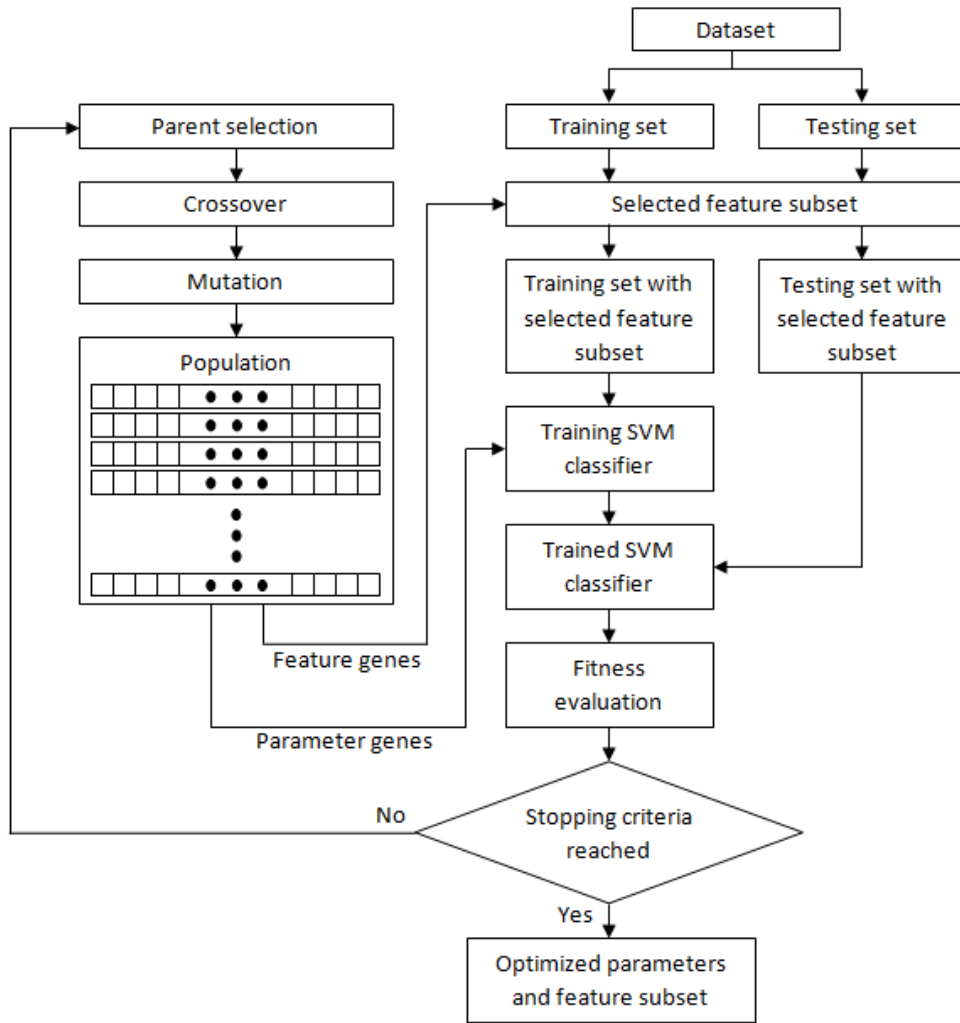


Figure 7: Architecture of the GA based feature selection and SVM parameter selection system, based on [61].

3 EXPERIMENTAL EVALUATION:

3.1 Materials:

A dataset of 60 images was used for this study, 20 images with confirmed new vessels and a further 40 images without new vessels. The images were gathered from two sources:

- 1) Publicly available MESSIDOR retinal image database, provided by the Messidor program partners [67]: 5 new vessel images, 20 normal images and 20 images with the large majority showing other DR pathology (mainly bright lesions) and the remainder showing strong reflection artefacts. These images were acquired from a colour video 3CCD camera on a

Topcon TRC NW6 fundus camera with a 45 degree field of view (FOV) and an image resolution of 2240 x 1488 pixels.

- 2) St Thomas' Hospital ophthalmology department: 15 new vessel images acquired with a Nikon D80 digital SLR camera on a Topcon TRC NW6 fundus camera with a 45 degree FOV and an image resolution of 2896 x 1944 pixels. Ethical approval was obtained for the use of these images.

Images were spatially normalized using a technique proposed by [68] along with bicubic interpolation and anti-aliasing. This was based on normalizing the FOV width, with the requirement that all images were captured with the same FOV angle. All images were normalized to have a FOV width of 1379 pixels. Cropping was used to remove some of the surrounding black border to produce images of size 1479 x 1479 pixels.

For training data, a specific selection of pixels was chosen from the dataset. It can be convenient to refer to a pixel as an image patch, considering features were extracted using information from the local neighbourhood contained within the sub window centred over the target pixel. These image patches were labelled as either new vessels or non-new vessels by an ophthalmologist. Separate training data was used for each classifier. Classifier 1 was trained with 50 new vessel patches and 50 normal vessel patches. Classifier 2 was trained with 50 new vessel patches and 50 patches made up of a variety of bright lesions, dark lesions and reflection artefacts. Note that each new vessel patch correlated to a pixel selected in the centre of a new vessel region.

Testing was performed across the whole of each retinal image, in terms of the classification process being performed at every pixel location. Because of the limited size of the dataset, leave-one-out cross validation was applied. This meant the classifiers were trained using all the patches from all the images except those from the single test image, and this process was repeated for each image. The feature value normalization was also recalculated each time, leaving out the test image.

As mentioned above, the system made decisions on a pixel basis using information extracted from the local neighbourhood. Therefore, to represent these local neighbourhoods all positive pixels were then dilated by the size of the sub window. This resulted in the delineation the new vessels regions. However, the performance from a per image basis is more useful from a clinical point of view. An image simply achieved a new vessel label if it contained any delineated regions. Prior to this, all images had been labelled by an ophthalmologist using the same labels as before but on a per image basis.

To get a more detailed insight into the system's performance the obvious choice would be to also evaluate the performance of delineation. This involves assessing the outcome for every pixel across all images and therefore is evaluation on a per pixel basis. However the performance of delineation was not evaluated and the reason why shall be discussed in section 4. Instead testing was performed on the selection of pixels used for training using the leave-one-out cross validation method. A few additional pixels from outside the training data were also tested. As stated above a pixel can be referred to with the term patch. Therefore, we refer to this assessment as evaluation on a per patch basis.

3.2 Performance Measures:

As mentioned above, the performance was separately assessed on a per image and per patch basis. Any image or patch was classified as either new vessels or non-new vessels. Consequently there are four outcomes, two classifications and two misclassifications which are defined in table 1. The algorithm was evaluated in terms of sensitivity (SN), specificity (SP) and accuracy (Acc). These are often used in machine learning and are measures of the quality of binary classification. These metrics are defined in table 2.

	New vessels present	New vessels absent
New vessels detected	True positive (TP)	False positive (FP)
New vessels not detected	False negative (FN)	True negative (TN)

Table 1: New vessel classification.

Measure	Description
SN	$TP/(TP+FN)$
SP	$TN/(TN+FP)$
Acc	$(TP+TN)/(TP+FP+TN+FN)$

Table 2: Performance measures for new vessels detection.

The use of the receiver operating characteristic (ROC) curve allows for the visualization of the performance of a binary classifier system, expressing the trade off between increased detection and false alarms. This was created by plotting the true positive rate (SN) versus the false positive rate (1-SP) at various threshold levels of the probability score of the classifier. The SVM calculated a new vessel probability score using the distance to the decision boundary.

With a dual classifier approach and therefore two probability scores, the creation of ROC curves was not a straight-forward task. This was resolved with the addition of a third axis to the conventional 2D ROC plot to accommodate for varying the threshold of the probability score of the additional classifier that arises in the dual classification approach. The resultant was a 3D ROC surface that explored all combinations of thresholds for the dual classification. Information from this 3D ROC surface was extracted to create a conventional 2D ROC curve. Full details of the procedure are described in [55]. From this 2D plot the area under the curve (AUC) was extracted and used as a performance measure.

Maximum Acc was the criteria used to select the optimal operating point from the ROC curve. However, this was only the case for per patch assessment. For per image assessment it was evident the algorithm could reach a SN of 100% at a high SP. This was important as from a clinical point of view a SN of 100% was considered an essential requirement from a per image basis. Max Acc may not always equate to an operating point with a SN of 100%. Therefore an application specific performance measure was created for the per image basis, in which the operating point with the highest SP at a SN of 100% was selected as the optimal operating point.

The measures used to quantify classification performance for involvement in the fitness function used in feature selection were the same as those stated above. Max Acc for per patch assessment and the highest SP at a SN of 100% for per image assessment. Each time a chromosome was assessed the ROC curve was assessed and the optimal operating point was found.

3.3 Results:

3.3.1 Modified Line Operator and Dual Classification

The performance on a per image and per patch basis for the original dual classification approach is presented in table 3. The new addition of performing a grid search for the optimal kernel type and SVM parameters means that these results do not match those reported in [55].

Method	Level	AUC	Acc	SN	SP	Kernel	Parameters
Welikala [55]	Image	0.9693	0.9500	1.0000	0.9250	Linear	$C=0.01$
Welikala [55]	Patch	0.9616	0.9235	0.8793	0.9440	Linear	$C=1$

Table 3: Performance measures and SVM parameters for [55].

The performance of the dual classification system against single classification and multiclass classification is presented in table 4. The table also lists the SVM kernel type and SVM parameters selected to reach optimal performance for each classification method, determined using a grid search. The comparison of classification methods is assessed on a per image basis.

Classification Method	SN	SP	Kernel	Parameters
Single	1.0000	0.8750	Polynomial	$C=0.01$, Order=2
Multiclass	1.0000	0.9000	Linear	$C=0.01$
Dual	1.0000	0.9250	Linear	$C=0.01$

Table 4: Comparison between single, multiclass and dual classification on a per image basis.

3.3.2 Feature expansion and selection

The Wilcoxon rank sum test, a non-parametric statistical test, was used to infer the discrimination ability of the features. This is done by assessing whether the median feature values for the two classes differ significantly. Results for both classifications are shown in table 5, with a smaller p-value indicating better discrimination ability. These calculations were not used for feature selection in this paper, however they were performed to provide a useful insight.

Wilcoxon rank sum test			
Classification 1		Classification 2	
Feature	p-value	Feature	p-value
8	2.0889×10^{-17}	2	4.3315×10^{-16}
2	2.1501×10^{-17}	1	1.8680×10^{-14}
4	6.7195×10^{-17}	3	7.1091×10^{-14}
1	1.4700×10^{-16}	4	6.4934×10^{-11}
3	7.1501×10^{-16}	17	1.3213×10^{-11}
17	7.2511×10^{-14}	8	1.1658×10^{-10}
7	9.4306×10^{-13}	19	6.8748×10^{-8}
14	1.8945×10^{-12}	6	1.0735×10^{-5}
6	5.3935×10^{-11}	16	3.3293×10^{-4}
16	2.6465×10^{-10}	20	5.0906×10^{-4}
10	1.0324×10^{-8}	7	0.0011
11	8.1172×10^{-7}	18	0.0018

21	5.7432×10^{-4}	10	0.0019
13	0.0017	14	0.0060
18	0.0030	13	0.0084
12	0.0193	21	0.0122
5	0.1030	9	0.0189
9	0.3293	11	0.0204
19	0.4340	5	0.1411
15	0.7174	15	0.1938
20	0.8066	12	0.3683

Table 5: Performance of the 21 features for each classification assessed using the Wilcoxon rank sum test. Listed in order of p-value, with the lowest value at the top of the table.

The features and SVM parameters selected for each classification that maximise the fitness function of the GA based system are shown in table 6. The resulting ROC curves of the proposed system under these settings on a per image and per patch basis are depicted in figures 8 and 9. The AUC value for the per image basis is 0.9914. The optimal operating point according to the application specific performance measure is a SN of 1.0000 and a SP of 0.9750. For a per patch basis the AUC value is 0.9600. The operating point with maximum Acc of 0.9454 gives a SN of 0.9138 and a SP of 0.9600. Table 7 shows these results along with the reported results from other new vessel detection methods. Examples of classified images are given in figure 11. Classified new vessel regions have also been indicated with a white boundary. Images containing any delineated regions are classified as new vessel images. Examples of classified patches are shown in figure 12. Note that these patches were individually classified and have been assembled together to create a single image only for visualization purposes. As stated previously, the term patch can be used to refer to a pixel. Therefore figure 12 represents the classification of the central pixel of each image patch.

Level	Classification	Kernel	SVM parameters	Features
Image	1	RBF	$C=1, \gamma=1$	4,8,10,13,14,16,19
	2	RBF	$C=1, \gamma=1$	1,2,4,6,8,15,17
Patch	1	RBF	$C=10, \gamma=1$	2,3,4,7,9,10,14,16,20
	2	RBF	$C=10, \gamma=1$	1,6,7,14,16,20

Table 6: Features and SVM parameters selected for each classification by the GA based system.

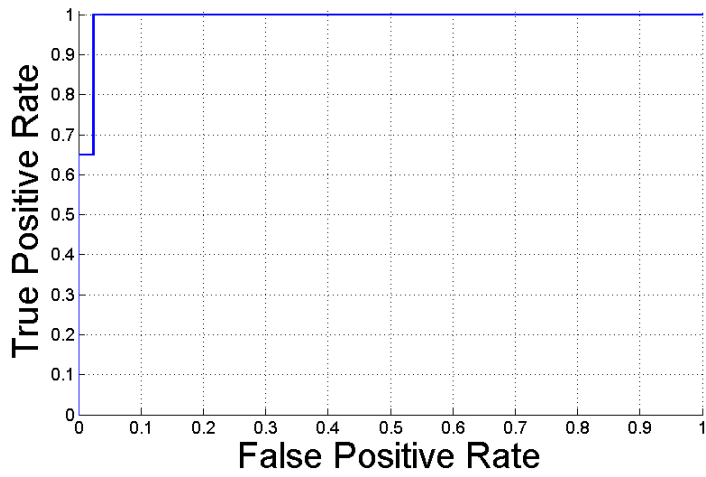


Figure 8: ROC curve for the performance on a per image basis, proposed method.

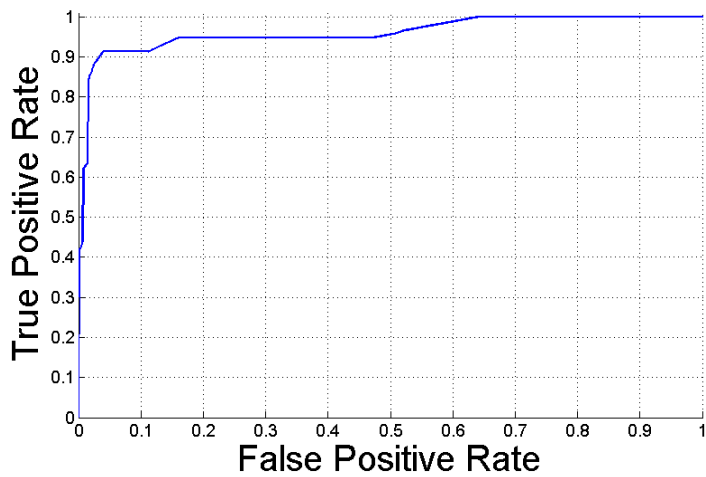


Figure 9: ROC curve for the performance on a per patch basis, proposed method.

Method	Year	Database	Level	AUC	Acc	SN	SP
Hassan [36]	2011	Combination	Pixel	0.7045	-	0.6390	0.8940
Welikala [37]	2013	Public	Image	-	-	1.0000	0.7000
Goatman [38]	2011	Local	Image	0.9110	-	0.8420	0.8590
Jelinek [39]	2007	Local	Image	0.9000	-	0.9400	0.8200
Akram [42]	2013	Public	Segment	0.9800	0.9800	0.9800	0.9700
Agurto [48]	2012	Local	Image	0.9400	-	0.9600	0.8300
Vatanparast [49]	2012	Local	Patch	-	-	0.9962	0.9661
Lee [50]	2013	Public	Image	0.9930	0.9850	0.9630	0.9910
Welikala [55]	2014	Combination	Image	0.9693	0.9500	1.0000	0.9250
Welikala [55]	2014	Combination	Patch	0.9616	0.9235	0.8793	0.9440
Proposed	-	Combination	Image	0.9914	0.9833	1.0000	0.9750
Proposed	-	Combination	Patch	0.9600	0.9454	0.9138	0.9600

Table 7: Reported results for new vessel detection methods.



Figure 10: Location and approximate delineation of new vessel regions marked by an ophthalmologist.

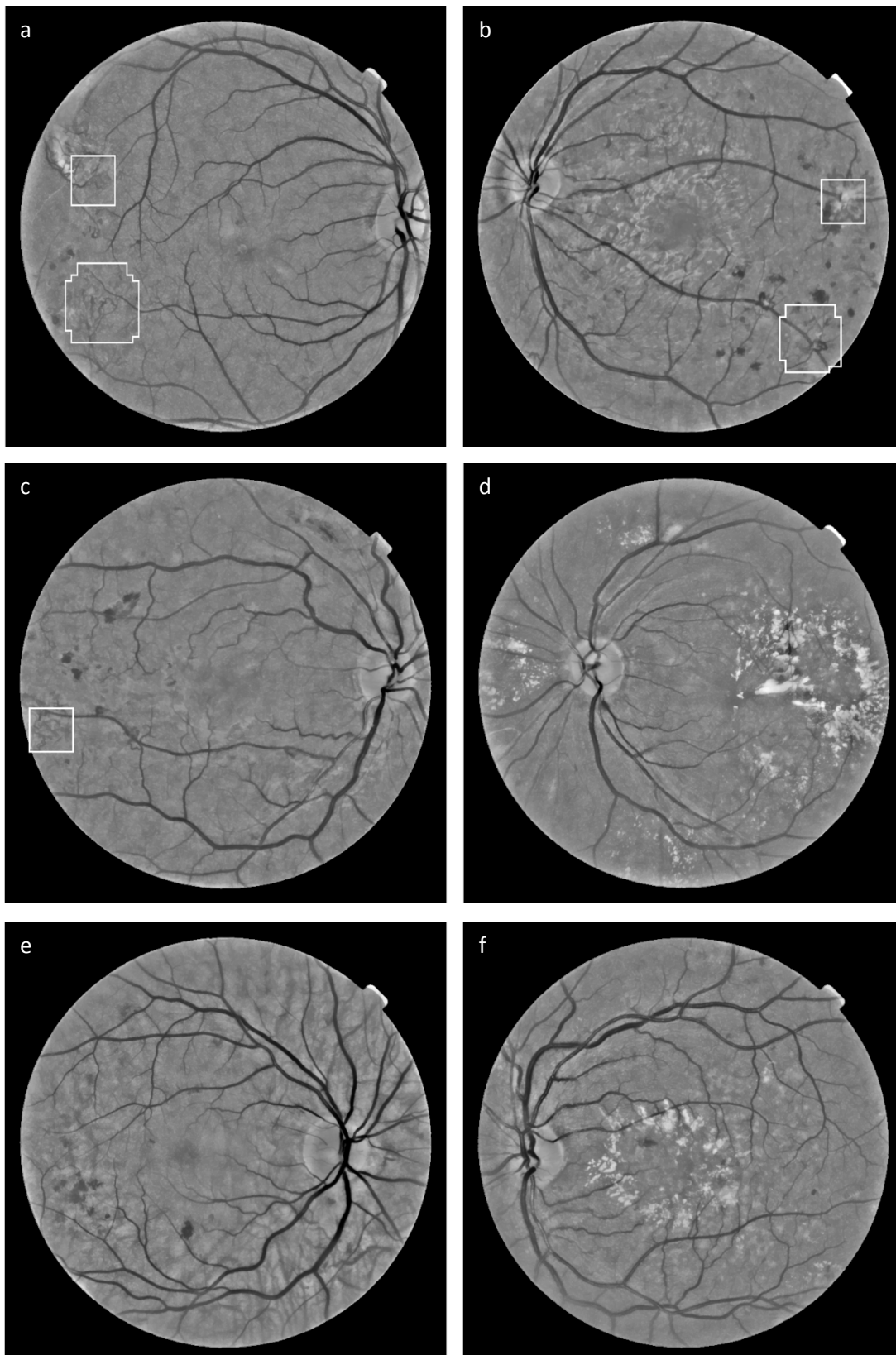


Figure 11: Results of the proposed system on a per image basis. (a)-(c) True positive images, (d)-(f) true negative images.

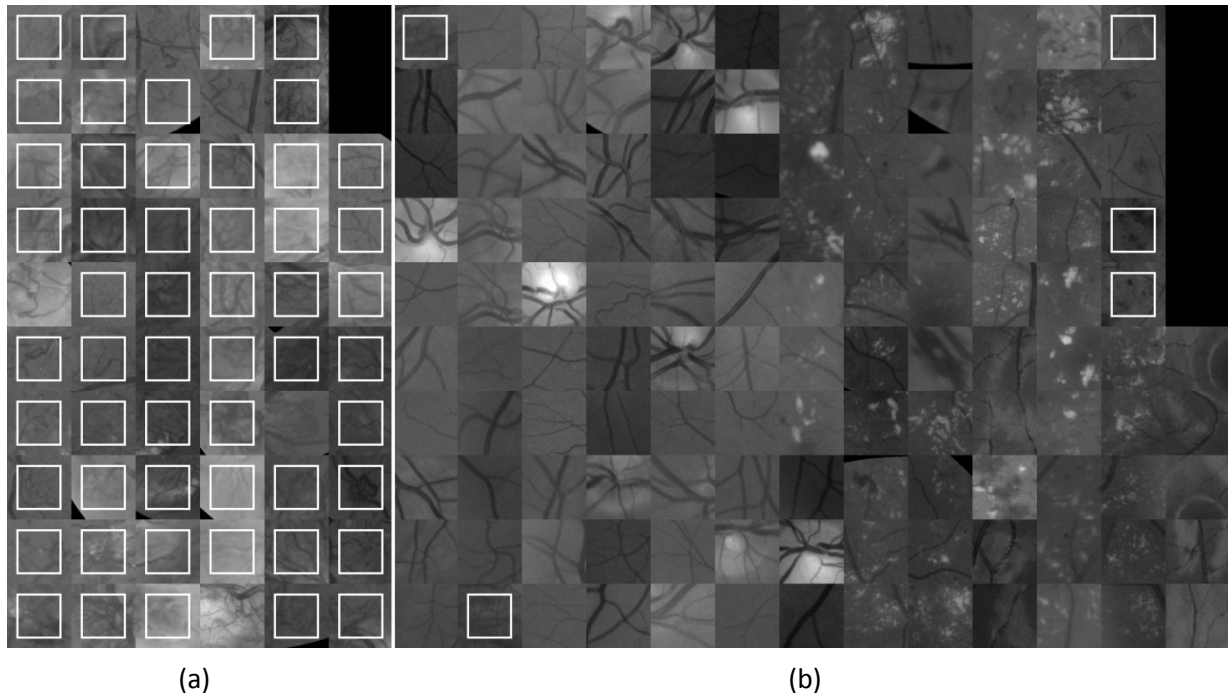


Figure 12: Results of the proposed system on a per patch basis. (a) New vessel patches, (b) non new vessel patches. White boxes indicate patches classified with a new vessel label.

4 DISCUSSION AND CONCLUSION:

In this paper, we have presented an effective new vessel detection method based on dual classification combined with feature vector expansion and feature selection. This includes the exploration of a 21-D feature vector which contained morphology, intensity and gradient based features.

Prior to discussing feature vector expansion and selection consider first the original methodology of [55]. The addition of performing a grid search for the optimal kernel type and SVM parameters does show an improvement in performance to that previously reported. Also table 4 compares the dual classification system's performance against the performance from a single SVM classification and a one-versus-one multiclass SVM classification. These results validate the dual classification approach, showing this process to be superior to others.

The first main contribution of this paper is the creation of a large feature set in combination with the dual classification method. The 21-D feature set looks at many different aspects in order to find suitable features for discrimination, these include morphology, intensity and gradient based features. The Wilcoxon test results in table 5 shows the original 4 features from [55], as well as the number of bifurcation points, compactness and tortuosity 2 all possess good discrimination ability for both classifications. Other features show particularly good discrimination ability for just a single

classification, for example features 19 and 20 which are based on the local grey level have a significantly better discriminating ability for classification 2 than for classification 1.

The second main contribution of this paper is the selection of features in combination with the dual classification method. A small feature subset not only has the potential to improve the classifier's performance, but is also better for generalization. However with feature selection, overfitting can still occur and some techniques are more prone than others. The results from the Wilcoxon test were not used to apply a filter approach for feature selection as this approach ranks features individually and therefore ignores feature dependencies. For example, after filtering only high rank features would remain and some of these features may be highly correlated. Thus resulting in the possible inclusion of many redundant features and therefore failing to detect the most useful features. GA is a much more thorough feature selection approach that interacts with the classifier and can model feature dependencies. Therefore, it has a better chance of finding the most useful features than the filter approach and greedy search strategies, however this puts it at high risk of overfitting. The GA approach is also highly beneficial as it allows for SVM parameter selection to be incorporated.

Table 6 provides the selection achieved from this GA based approach. The selected feature subsets are relatively small in size. On a per image basis the feature subsets contained 7 features for each classification. Feature subsets of 6 features and 9 features are selected on a per patch basis. The RBF kernel with a γ of 1 was selected on both levels (image and patch).

From table 6, it is apparent that many of the features that were deemed to have a high discriminating ability by the Wilcoxon test have been excluded from the feature subset due to their redundancy. Also it's apparent that almost all features possessing a low discriminating ability by the Wilcoxon test have been excluded due to them being irrelevant. However, each feature subset does still retain a couple of lower ranked features. This was expected because features which are completely irrelevant individually can provide significant performance improvements when considered with other features.

The results in table 7 show the performance of the proposed system is superior than [55], achieving a SN of 1.0000 and a SP of 0.9750 compared to a SN of 1.0000 and a SP of 0.9250 on a per image basis. Also achieving a SN of 0.9138 and a SP of 0.9600 compared to a SN of 0.8793 and a SP of 0.9440 on a per patch basis. From the examples of classified images shown in figure 11 it is clear that the algorithm responds well to a range of new vessel formations. Also evident is the algorithm's ability to avoid false responses despite the presence of other pathology (bright lesions, dark lesions etc.).

The results that achieve maximum performance on each level (image and patch) do not correspond to the same operating point. The reported per image performance of 1.0000 and 0.9750 for SN and SP respectively corresponds to a per patch performance of 0.3103 and 0.9920 for SN and SP respectively. This illustrates, for the per image performance, that the system puts no emphasis on correctly detecting all new vessels. Instead identifying any part of any new vessel region in the image is sufficient for the image to achieve a new vessel label. Figure 11(b) illustrates how there is no requirement to identify all new vessels in the image, with only two out of the three new vessel networks being identified. Figure 10(b) shows the same image with the location of all three new vessel networks marked by an ophthalmologist. Such an approach assists in ensuring a higher

specificity is obtained. Whilst this approach may hold certain risks, it is still a viable option for clinical application. Niemeijer [13] follows the same approach but in respect to dark lesion detection.

From figure 12 it is clear that the emphasis on a per patch basis is to detect as many new vessel patches as possible. This further highlights how the process at the two levels (image and patch) differs. The actual classification process in terms of SVM boundaries is designed to label pixels, this equates to a patch level (recall a pixel can be referred to with the term patch). Therefore, unlike on a per image basis, enhancing the results on a per patch basis is based on a direct enhancement of the SVM boundaries. The fact that the features selected for the per image and per patch basis (table 6) vary significantly, further suggests that the requirements for classification on each level differ significantly. From a computer science point of view the development and evaluation from a per patch basis should guide future work.

Whilst our algorithm delineates new vessel regions to a certain degree, it was decided not to assess evaluation of its performance in these terms. Our clinical team have approximately manually delineated all new vessel regions in all images (for examples see figure 10). Accurate manual delineation would have been a very time intensive task and high accuracy would have been difficult to achieve due to the obscurity of new vessel regions. Consider accurate manual delineation had have been achieved, then all pixels from all images could have been used not only to evaluate (test), but also to train the classifiers as opposed to the relatively small selection of pixels (patches) used. However, doing so would make the algorithm more susceptible to false positives. Therefore, our main concern is the identification of new vessel regions as opposed to their accurate delineation. Hence all new vessel patches used for training correlated only to sub windows directly centred over the middle of new vessel regions. This is a more targeted approach which suits the clinical requirements for the identification of new vessels. Hence, evaluation on a per patch basis (testing on training patches using leave-one-out cross validation) was more relevant and therefore preferred to the evaluation of delineation to get a more detailed insight into the system's performance than that provided by performance on a per image basis.

The proposed method does achieve better performance metrics than most of the other published methods, as shown in table 7. True comparisons are difficult to make as there exists variability in terms of their application. Goatman [38] and Agurto [48] seek to only detect NVD. Jelinek [39] applied their methods on fluorescein images as opposed to conventional retinal images. More importantly there exist no standard datasets that have been used for testing. Our particular dataset was not solely created from publicly available data sets due to their limited inclusion of images containing new vessels. Akram [42] created a dataset containing 52 images with new vessels from the four main publicly available retinal image databases. However, the image file names used are not available and we were unable to find the number of new vessel images reported.

Future developments of this method will include further exploration into the selection of SVM parameters. Currently a single set of SVM parameters were selected by the GA based system for both classifications 1 and 2. These are two separate classifications with their own distinct feature sets, therefore the optimal SVM parameters should be searched individually for each. This should lead to further enhancement of the classifier's performance. This holds the risk of causing overfitting, whilst our current approach is better in terms of generalization. Another aspect that

needs attention is the number of generations. This value needs to be set significantly higher to explore whether the fitness function could be further increased. Current limitations restricting the number of generations are the computational expense of performing assessment by leave-one-out cross validation and the large amount of operating points to assess that arise from dual classification. These limitations were the reason why the optimization problem was simplified with the use of discrete values to represent continuous variables.

In conclusion, this paper has demonstrated an automated system that is capable of detecting the presence of new vessels whilst reducing false responses to bright lesions, dark lesions and reflection artefacts. This involved the use of a genetic algorithm based system to perform feature selection and SVM parameter selection.

5 REFERENCES:

- [1] P.M. Dodson, Oxford Diabetes Library: Diabetic Retinopathy, Oxford University Press (2008).
- [2] D.S. Fong, L. Aiello, T.W. Gardner, G.L. King, G. Blankenship, J.D. Cavallerano, et al., Retinopathy in diabetes, *Diabetes care* 27 Suppl 1 (2004) S84.
- [3] H.R. Taylor and J.E. Keeffe, World blindness: a 21st century perspective, *British Journal of Ophthalmology* 85 (2001) 261-266.
- [4] S. Wild, G. Roglic, A. Green, R. Sicree, H. King, Global prevalence of diabetes: estimates for the year 2000 and projections for 2030. *Diabetes Care* 27 (2004) 1047-1053.
- [5] J.P. De La Cruz, J.A. Gonzalez-Correa, A. Guerrero, Pharmacological approach to diabetic retinopathy, *Diabetes/metabolism research and reviews* 20 (2004) 91-113.
- [6] T. Bek and M. Erlandsen, Visual prognosis after panretinal photocoagulation for proliferative diabetic retinopathy. *Acta Ophthalmologica Scandinavica* 84 (2006) 16.
- [7] U.K National Screening Committee, Essentials Elements in Developing a Diabetic Screening Programme. Workbook Version 4.3, (2009).
- [8] D.C. Klonoff and D.M. Schwartz, An economic analysis of interventions for diabetes, *Diabetes Care* 23 (2000) 390.
- [9] R.N. Frank, Diabetic retinopathy, *Progress in Retinal and Eye Research* 14 (1995) 361-392.
- [10] J.J. Kanski and B. Bowling, *Clinical Ophthalmology: A Systematic Approach*. Seventh edition. Elsevier Limited (2011).
- [11] A.D. Fleming, K.A. Goatman, S. Philip, G.J. Prescott, P.F. Sharp, J.A. Olson, Automated grading for diabetic retinopathy: a large-scale audit using arbitration by clinical experts, *British Journal of Ophthalmology* 94 (2010) 1606-1610.

- [12] T. Spencer, J.A. Olson, K.C. McHardy, P.F. Sharp, J.V. Forrester, An image-processing strategy for the segmentation and quantification of microaneurysms in fluorescein angiograms of the ocular fundus, *Computers and Biomedical Research* 29 (1996) 284-302.
- [13] M. Niemeijer, B. van Ginneken, J. Staal, M.S.A. Suttorp-Schulten, M.D. Abramoff, Automatic detection of red lesions in digital color fundus photographs, *IEEE Transactions on Medical Imaging* 24 (2005) 584-592.
- [14] A. Sopharak, B. Uyyanonvara, S.A. Barman, Simple hybrid method for fine microaneurysm detection from non-dilated diabetic retinopathy retinal images, *Computerized Medical Imaging and Graphics* (2013).
- [15] L. Tang, M. Niemeijer, J.M. Reinhardt, M.K. Garvin, M.D. Abramoff, Splat feature classification with application to retinal hemorrhage detection in fundus images, *IEEE Transactions on Medical Imaging* 32 (2013) 364-375.
- [16] B. Zhang, X. Wu, J. You, Q. Li, F. Karray, Detection of microaneurysms using multi-scale correlation coefficients, *Pattern Recognition* 43 (2010) 2237-2248.
- [17] T. Walter, J. Klein, P. Massin, A. Erginay, A contribution of image processing to the diagnosis of diabetic retinopathy-detection of exudates in color fundus images of the human retina, *IEEE Transactions on Medical Imaging* 21 (2002) 1236-1243.
- [18] A. Sopharak, B. Uyyanonvara, S.A. Barman, Automatic exudate detection from non-dilated diabetic retinopathy retinal images using fuzzy C-means clustering, *Sensors* 9 (2009) 2148-2161.
- [19] A. Sopharak, M.N. Dailey, B. Uyyanonvara, S.A. Barman, T. Williamson, K.T. Nwe, et al., Machine learning approach to automatic exudate detection in retinal images from diabetic patients, *Journal of Modern Optics* 57 (2010) 124-135.
- [20] A. Osareh, M. Mirmehdi, B. Thomas, R. Markham, Automated identification of diabetic retinal exudates in digital colour images, *British Journal of Ophthalmology* 87 (2003) 1220-1223.
- [21] S. Chaudhuri, S. Chatterjee, N. Katz, M. Nelson, M. Goldbaum, Detection of blood vessels in retinal images using two-dimensional matched filters, *IEEE Transactions on Medical Imaging* 8 (1989) 263-269.
- [22] A.M. Mendonca and A. Campilho, Segmentation of retinal blood vessels by combining the detection of centerlines and morphological reconstruction, *IEEE Transactions on Medical Imaging* 25 (2006) 1200-1213.
- [23] G.B. Kande, P.V. Subbaiah, T.S. Savithri, Unsupervised fuzzy based vessel segmentation in pathological digital fundus images, *Journal of Medical Systems* 34 (2010) 849-858.
- [24] J. Staal, M.D. Abramoff, M. Niemeijer, M.A. Viergever, B. van Ginneken, Ridge-based vessel segmentation in color images of the retina, *IEEE Transactions on Medical Imaging* 23 (2004) 501-509.
- [25] J.V.B. Soares, J.J.G. Leandro, R.M. Cesar Jr, H.F. Jelinek, M.J. Cree, Retinal vessel segmentation using the 2-D Gabor wavelet and supervised classification, *IEEE Transactions on Medical Imaging* 25 (2006) 1214-1222.

- [26] E. Ricci and R. Perfetti, Retinal blood vessel segmentation using line operators and support vector classification, *IEEE Transactions on Medical Imaging* 26 (2007) 1357-1365.
- [27] M.M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A.R. Rudnicka, C.G. Owen, et al., An ensemble classification based approach applied to retinal blood vessel segmentation, *IEEE Transactions on Biomedical Engineering* 59 (2012) 2538-2548.
- [28] D. Marín, A. Aquino, M.E. Gegúndez-Arias, J.M. Bravo, A new supervised method for blood vessel segmentation in retinal images by using gray-level and moment invariants-based features, *IEEE Transactions on Medical Imaging* 30 (2011) 146-158.
- [29] M.M. Fraz, P. Remagnino, A. Hoppe, B. Uyyanonvara, A.R. Rudnicka, C.G. Owen, et al., Blood vessel segmentation methodologies in retinal images - A survey, *Computer methods and programs in biomedicine* 108 (2012) 407-433.
- [30] M.K. Ikram, F.J. de Jong, J.R. Vingerling, J.C.M. Witteman, A. Hofman, M.M.B. Breteler, et al., Are Retinal Arteriolar or Venular Diameters Associated with Markers for Cardiovascular Disorders? The Rotterdam Study, *Investigative ophthalmology & visual science* 45 (2004) 2129-2134.
- [31] M. Saez, S. Gonzalez-Vazquez, M. Gonzalez-Penedo, M.A. Barcelo, M. Pena-Seijo, G.C. de Tuero, et al., Development of an automated system to classify retinal vessels into arteries and veins, *Computer methods and programs in biomedicine* 108 (2012) 367-376.
- [32] G.S. Ramlugun, V.K. Nagarajan, C. Chakraborty, Small retinal vessels extraction towards proliferative diabetic retinopathy screening, *Expert Systems with Applications* 39 (2012) 1141-1146.
- [33] B. Zhang, L. Zhang, L. Zhang, F. Karray, Retinal vessel extraction by matched filter with first-order derivative of Gaussian, *Computers in Biology and Medicine* 40 (2010) 438-445.
- [34] L. Zhang, Q. Li, J. You, D. Zhang, A modified matched filter with double-sided thresholding for screening proliferative diabetic retinopathy, *IEEE Transactions on Information Technology in Biomedicine* 13 (2009) 528-534.
- [35] M.U. Akram, A. Tariq, S.A. Khan, Detection of neovascularization for screening of proliferative diabetic retinopathy, In *Proceedings of the 9th international conference on Image Analysis and Recognition - Volume Part II* (2012) 372-379.
- [36] S.S.A. Hassan, D.B.L. Bong, M. Premseenthil, Detection of neovascularization in diabetic retinopathy, *Journal of digital imaging : the official journal of the Society for Computer Applications in Radiology* (2011).
- [37] R.A. Welikala, V. Tah, T.H. Williamson, A. Hoppe, J. Dehmeshki, S.A. Barman, Differing matched filter responsivity for the detection of proliferative diabetic retinopathy, *Proceedings of the IASTED International Conference Signal Processing, Pattern Recognition and Applications* (2013) 356-364.
- [38] K.A. Goatman, A.D. Fleming, S. Philip, G.J. Williams, J.A. Olson, P.F. Sharp, Detection of new vessels on the optic disc using retinal photographs, *IEEE Transactions on Medical Imaging* 30 (2011) 972.

- [39] H.F. Jelinek, M.J. Cree, J.J. Leandro, J.V. Soares, R.M. Cesar Jr, A. Luckie, Automated segmentation of retinal blood vessels and identification of proliferative diabetic retinopathy, *Journal of the Optical Society of America.A, Optics, image science, and vision* 24 (2007) 1448-1456.
- [40] A. Daxer, Characterisation of the neovascularisation process in diabetic retinopathy by means of fractal geometry: diagnostic implications, *Graefes archive for clinical and experimental ophthalmology* 231 (1993) 681-686.
- [41] A. Karperien, H.F. Jelinek, J.J.G. Leandro, J.V.B. Soares, R.M. Cesar Jr, A. Luckie, Automated detection of proliferative retinopathy in clinical practice, *Clinical Ophthalmology* 2008 (2008) 109.
- [42] M.U. Akram, S. Khalid, A. Tariq, M.Y. Javed, Detection of neovascularization in retinal images using multivariate m-Medoids based classifier, *Computerized Medical Imaging and Graphics* 37 (2013) 346.
- [43] K. Saranya, B. Ramasubramanian, S. Kaja Mohideen, A novel approach for the detection of new vessels in the retinal images for screening Diabetic Retinopathy, *International Conference on Communications and Signal Processing* (2012) 57-61.
- [44] F. Oloumi, R.M. Rangayyan, A.L. Ells, Computer-aided diagnosis of proliferative diabetic retinopathy via modeling of the major temporal arcade in retinal fundus images, *Journal of Digital Imaging* 26 (2013) 1124-1130.
- [45] A.J. Frame, P.E. Undrill, J.A. Olson, K.C. McHardy, P.F. Sharp, J.V. Forrester, Texture analysis of retinal neovascularisation, *IEE Colloquium on Pattern Recognition (Digest No. 1997/018)* (1997) 5/1-5/6.
- [46] U.R. Acharya, E.Y.K. Ng, J. Tan, S.S. Vinitha, K. Ng, An integrated index for the identification of diabetic retinopathy stages using texture parameters, *Journal of Medical Systems* 36 (2012) 2011-2020.
- [47] C. Agurto, V. Murray, E. Barriga, S. Murillo, M. Pattichis, H. Davis, et al., Multiscale AM-FM methods for diabetic retinopathy lesion detection, *IEEE Transactions on Medical Imaging* 29 (2010) 502-512.
- [48] C. Agurto, Y. Honggang, V. Murray, M.S. Pattichis, S. Barriga, W. Bauman, et al., Detection of neovascularization in the optic disc using an AM-FM representation, granulometry, and vessel segmentation, *Annual International Conference of the IEEE, Engineering in Medicine and Biology Society (EMBC)* (2012) 4946-4949.
- [49] M. Vatanparast and A. Harati, A feasibility study on detection of Neovascularization in retinal color images using texture, *2nd International eConference on Computer and Knowledge Engineering (ICCKE)* (2012) 221-226.
- [50] J. Lee, B.C.Y. Zee, Q. Li, Detection of neovascularization based on fractal and texture analysis with interaction effects in diabetic retinopathy, *PLoS One* 8 (2013) e75699.
- [51] A. Perez-Rovira, T. MacGillivray, E. Trucco, K.S. Chin, K. Zutis, C. Lupascu, et al., VAMPIRE: Vessel assessment and measurement platform for images of the REtina, *Conf Proc IEEE Eng Med Biol Soc* 2011 (2011) 3391-3394.

[52] M.E. Martinez-Perez, A.D. Highes, A.V. Stanton, S.A. Thorn, N. Chapman, A.A. Bharath, et al., Retinal vascular tree morphology: a semi-automatic quantification, *IEEE Transactions on Biomedical Engineering* 49 (2002) 912-917.

[53] C.N. Doukas, I. Maglogiannis, A.A. Chatziioannou, Computer-supported angiogenesis quantification using image analysis and statistical averaging, *IEEE Transactions on Information Technology in Biomedicine* 12 (2008) 650-657.

[54] K. Zutis, E. Trucco, J.P. Hubschman, D. Reed, S. Shah, J. van Hemert, Towards automatic detection of abnormal retinal capillaries in ultra-widefield-of-view retinal angiographic exams, *International Conference on Engineering in Medicine and Biology* (2013).

[55] R.A. Welikala, J. Dehmeshki, A. Hoppe, V. Tah, S. Mann, T.H. Williamson, et al., Automated detection of proliferative diabetic retinopathy using a modified line operator and dual classification, *Computer methods and programs in biomedicine* 114 (2014) 247-261.

[56] B.E. Boser, I.M. Guyon, V.N. Vapnik, A training algorithm for optimal margin classifiers, *Proceedings of the fifth annual workshop on Computational learning theory* (1992) 144-152.

[57] N. Cristianini and J. Shawe-Taylor, *An introduction to support vector machines and other kernel-based learning methods*, Cambridge university press (2000).

[58] I. Guyon and A. Elisseeff, An Introduction to Variable and Feature Selection, *Journal of Machine Learning Research* 3 (2003) 1157-1182.

[59] Y. Saeys, I. Inza, P. Larrañaga, A review of feature selection techniques in bioinformatics, *Bioinformatics* 23 (2007) 2507-2517.

[60] H. Frohlich, O. Chapelle, B. Scholkopf, Feature selection for support vector machines by means of genetic algorithm, *Proceedings. 15th IEEE International Conference on Tools with Artificial Intelligence* (2003) 142-148.

[61] C.L. Huang and C.J. Wang, A GA-based feature selection and parameters optimization for support vector machines, *Expert Systems with Applications* 31 (2006) 231-240.

[62] W.W. Bledsoe, The use of biological concepts in the analytical study of systems, Paper presented at ORSA-TIMS National Meeting, San Francisco (1961).

[63] J.H. Holland, *Adaptation in natural and artificial systems*, MIT Press (1992).

[64] A. Statnikov, C.F. Aliferis, D.P. Hardin, I. Guyon, *A gentle introduction to support vector machines in biomedicine*, World Scientific Publishing (2011).

[65] K.A. De Jong and W.M. Spears, An analysis of the interacting roles of population size and crossover in genetic algorithms, in *Parallel problem solving from nature*, Anonymous, pp. 38-47 (Springer, 1991).

[66] J.J. Grefenstette, Optimization of control parameters for genetic algorithms, *Systems, Man and Cybernetics*, *IEEE Transactions on* 16 (1986) 122-128.

[67] MESSIDOR [Online], <http://messidor.crihan.fr>.

[68] X. Zhang, G. Thibault, E. Decencière, G. Quellec, G. Cazuguel, A. Erginay, et al., Spatial normalization of eye fundus images, IEEE International Symposium on Biomedical Imaging (2012).