Localising Microaneurysms in Fundus Images Through Singular Spectrum Analysis

Su Wang, Hongying Lilian Tang*, Lutfiah Ismail Al turk, Yin Hu, Saeid Sanei, Senior Member, IEEE, George Michael Saleh and Tunde Peto

Abstract—Goal: Reliable recognition of microaneurysms is an essential task when developing an automated analysis system for diabetic retinopathy detection. In this work, we propose an integrated approach for automated microaneurysm detection with high accuracy. Methods: Candidate objects are first located by applying a dark object filtering process. Their cross-section profiles along multiple directions are processed through singular spectrum analysis. The correlation coefficient between each processed profile and a typical microaneurysm profile is measured and used as a scale factor to adjust the shape of the candidate profile. This is to increase the difference in their profiles between true microaneurysms and other non-microaneurysm candidates. A set of statistical features of those profiles is then extracted for a K-Nearest Neighbour classifier. Results: Experiments show that by applying this process, microaneurysms can be separated well from the retinal background, the most common interfering objects and artefacts. Conclusion: The results have demonstrated the robustness of the approach when testing on large scale datasets with clinically acceptable sensitivity and specificity. Significance: The approach proposed in the evaluated system has great potential when used in an automated diabetic retinopathy screening tool or for large scale eye epidemiology studies.

Index Terms—Computer-aided diagnosis, image classification, microaneurysm detection, retinal image, singular spectrum analysis, diabetic retinopathy.

I. INTRODUCTION

Diabetic retinopathy (DR) is the most common microvascular complication of diabetes and remains the leading cause of vision loss in the working-age population [1]. Early diagnosis through regular screening helps prevent vision loss. In a DR screening programme, a very large number of digital retinal images need to be examined for the presence of DR by human experts. Pathological signs of DR in the digital colour fundus images are dark lesions including microaneurysms (MAs) and haemorrhages, as well as bright lesions such as exudates and cotton wool spots (see Fig. 1). An automated system for separating healthy and diseased regions in the image can efficiently reduce the workload associated with large scale screening. Over the last two decades, research in DR image analysis has been constantly increasing. Many studies of automated DR screening systems have appeared in the literature, e.g. [2]–[6].

One critical stage in these automated image processing systems is microaneurysm detection. MAs are the first visible signs of DR and they appear as small circular reddish dots in the retina. The quantity of MAs indicates the progression of DR [7]. The complexity of MA recognition lies in the fact that MAs can be located anywhere in the retina: in isolation, in clusters, close to vasculature, around macula or among exudates. Meanwhile, their local contrast is very low compared to their surrounding background and their edges are not well defined in the image. In addition, MAs have very similar intensity and morphological characteristics to other DR signs and anatomical features such as haemorrhages, thin vessel junctions, visually disconnected vessel fragments, local darkenings on the vessels or retinal background noise [5] (See Fig. 2 for some examples). Retinal images of patients from different ethnic groups also pose challenges for MA detection by varying background colour, introducing new disease patterns but often new non-DR diseases that are unknown to the automated system. Fig. 3 shows the retinal images from different populations including Kenya [8], Botswana [9], Mongolia [1], China [1], Saudi Arabia [1], Italy [1], Lithuania [1] and Norway [1]. In this work, our aim is to develop an automated system to recognise MAs in large scale fundus images with clinically acceptable accuracy regardless of their quality, ethnic origins and the types of cameras used for capturing the images.

In the past, a number of different methods for automated...
detection of MAs have been proposed. Baudoin et al. [10] used a mathematical morphology approach to eliminate vessels in fluorescein angiography images so as to extract MA candidates. They used a top-hat transformation to distinguish non-connected and circular dark lesions from the elongated vasculature. Variations of this method have been explored to improve its performance by several authors [11]–[14].

The contrast between MAs and background in fluorescein angiography images is higher than that in digital colour photographs. As there is a mortality of 1 : 222,000 associated with the intravenous use of fluorescein [15], the application of fluorescein angiography images for larger scale screening purposes is impractical [16]. Hipwell et al. [17] extended the top-hat based method to high-resolution red-free fundus photographs. From now on, all methods discussed in this paper are for digital colour fundus images only.

Walter et al. [18] proposed to remove dark objects based on their diameter criteria. This was followed by a top-hat based method to generate initial object segments. MA candidates were then extracted through an automatic threshold scheme. This approach was also to reduce false candidates located on tortuous vessels. At the classification stage, a kernel density estimation with variable bandwidth was used to classify candidates into true MAs and other objects.

Niemeijer et al. [19] proposed a hybrid detection system by combining a supervised pixel classification and the top-hat based algorithm to detect dark lesions in colour fundus photographs. Intensity and shape descriptors and a k-nearest neighbour (kNN) classifier were used to ameliorate MA detection. This method allowed detection of MAs and larger dark lesions (i.e., haemorrhages) using the same system.

Other approaches than top-hat based methods have also been reported. Quellec et al. [20] applied a template matching method to detect MAs in the wavelet domain. The authors considered that the wavelets could be effectively used to distinguish between lesions and non-lesion areas. The MAs were modelled with a 2D rotation-symmetric generalised Gaussian function, and the optimal adapted wavelet transform for MA detection was obtained by applying the lifting scheme framework [21]. MAs were validated by applying a threshold on the matching result. Other popular MA detection methods include region grow [6], [22], Gaussian masks and multiscale correlation coefficients [22], and dictionary learning with sparse representation [23].

An intensity profile is a sequence of pixel intensities when scanning an image along a certain direction. As the intensity profiles across MA objects have local minima, they can be modelled as an inverted 2D Gaussian shape. An understanding of the intensity profiles of an MA in an image plays an important role for an effective separation between the MA and other similar objects. A few methods following this approach for MA detection have been proposed by several authors [24], [25]. In [24], the cross-section profiles were extracted by detecting the local maximum pixels in an inverted image. A naïve Bayes (NB) classification was then used to remove false MA candidates based on a set of statistical measures of cross-section profiles, including the size, height and shape of the directional cross-sections of MAs.

Although these reported MA extraction approaches have some advantages, they still have difficulty in extracting MAs that are located close to blood vessels and discriminating MAs from the most common interfering structures such as vessel crossings and elongated haemorrhages. In this paper, we propose an effective method to detect MAs which addresses these problems. We tested our approach on large scale data from eight populations where diverse pathological cases present to the system in addition to those inherent variations during the image acquisition process. There are three main contributions in this proposed work. First, a candidate extraction scheme is proposed to extract more MA candidates including those close to vessels. Second, for every candidate, its cross-section profiles along 12 directions are obtained. Singular spectrum analysis (SSA) [26] is used to decompose each profile and reconstruct a new one that is of a slow varying trend. Third, each filtered profile is scaled using the correlation coefficient between itself and an ideal Gaussian shape assuming this candidate is a true MA. This will enable an enhancement of the profiles in all directions for true MA candidates while decrease the similarity among profiles in all directions for non-MA candidates. Features are then extracted from the scaled profiles of each candidate for MA/non-MA classification.

The rest of this paper is organised as follows. Section II describes our proposed approach on MA detection. Key parameter setups are justified in Section III, together with the overall performance of the proposed method. In Section IV, we carry out discussion and conclude this paper.

II. PROPOSED METHOD

The proposed method is performed on the green channel of retinal images as MAs, haemorrhages and vessels normally present the highest contrast against the surrounding background in this colour channel [27]. MA detection is divided into preprocessing, candidate extraction through multilayered dark object filtering, candidate cross-section profile analysis based on SSA, feature extraction and classification. The details of these stages are provided in the following sections.

A. Image Preprocessing

Preprocessing attenuates the effects of noise and preserves the true information of MAs. A Gaussian filter is applied to the green channel $I_g$ to enhance the small and dark structures, resulting in a filtered image $I_{g,g}$ as shown in Fig. 4 (b). The Gaussian filter ($width = 3, variance = 1$) removes many
tiny structures occurring as a results of noise while preserving those corresponding to MAs or vessels.

In addition, when bright lesions or regions are close together, the small gaps between them can be wrongly recognised as MAs in the later stages of the processing [19]. In order to prevent these false positives (FPs), a shade correction algorithm [11], [13] is extended to remove any bright region from image $I_{gg}$.

1) First, estimate a background image $I_{bg}$ (Fig. 4 (c)) by applying a median filter ($35 \times 35$) to $I_{gg}$.

2) Subtract $I_{bg}$ from $I_{gg}$ to obtain image $I_{sc}$ (Fig. 4 (d)). Any pixel in $I_{sc}$ that has a positive value means, in $I_{gg}$, the corresponding pixel has higher intensity value than its neighbouring retinal background intensity. These pixels are used to locate bright regions in $I_{gg}$.

3) All bright pixels in $I_{gg}$ indicated by $I_{sc}$ are replaced by the values of their corresponding pixels in $I_{bg}$ resulting in an image $I_{pp}$ (Fig. 4 (e)).

Since this process removes bright regions (including bright lesions) from $I_{gg}$, it also removes the ambiguity when bright lesions or regions cluster together causing the gaps among them to be considered as MAs. Those pixels in $I_{sc}$ with a negative or zero value are processed for candidate extraction in the next stage.

### B. Candidate Extraction Using Multilayered Dark Object Filtering

Candidate extraction is a process which aims to identify any object in the image showing MA-like characteristics. These candidates will then be further analysed or classified into MAs and non-MAs. Previous reported approaches utilised different characteristics of MAs to extract MA candidates. In [17]–[19], a morphological transformation such as top-hat was employed to eliminate the vasculature from a retinal image aiming to leave possible MA candidates untouched. These methods were able to extract those isolated MAs away from other dark objects including vessels. However, when an MA is next to other dark objects, it was often not detected but considered as part of the neighbouring objects. In [24], a pixel was regarded to be a local maximum (in an inverted image) if its eight-neighbouring pixels have lower or the same intensity. The use of these local maxima made it easier to obtain more MAs. However, it also brings in much more common interfering structures as MA candidates such as vessel crossings as well as many small background regions due to high local intensity variation.

We have largely addressed these limitations by using a multilayered dark object filtering method. After preprocessing, all pixels $(x, y)$ with negative values in $I_{sc}$ are regarded as initial positions to examine dark objects in image $I_{pp}$, which include vessels, dark lesions and noise. Inspired by [28], which was initially proposed for identifying vessel candidates, we estimate the position of a given dark pixel $O(x, y)$ within a dark object by examining all its neighbours’ darkness through calculating a connected neighbourhood strength, defined by Eq. (1):

$$
\eta_i = sgn(\cos \theta_i), \quad \eta_2 = sgn(\sin \theta_i), \quad i = 1, \ldots, 8
$$

$$
p_i(x, y, \theta_i) = I_{pp}(x + \eta_1 - \eta_2 w, y + \eta_2 + \eta_1 w) + I_{pp}(x + \eta_1 + \eta_2 w, y + \eta_2 - \eta_1 w) - 2I_{pp}(x + \eta_1, y + \eta_2),
$$

Fig. 3. The first row shows the retinal images from different populations. (a) Kenya, (b) Botswana, (c) Mongolia, (d) China, (e) Saudi Arabia, (f) Italy, (g) Lithuania and (h) Norway. The second row shows the corresponding detailed subimages in the white boxes in the first row. Note that the images shown here were preprocessed by removing the black borders around the field of view. All images have been scaled to equal height for display.

Fig. 4. (a) The green channel of the fundus image. (b) The Gaussian filtered image of (a). (c) The estimated background image $I_{bg}$. (d) Shade correction image $I_{sc}$ was accomplished by subtracting the median filtered image $I_{bg}$ from $I_{gg}$. The white parts indicate bright regions (including bright lesions). (e) The preprocessed image $I_{pp}$ (the bright pixels in $I_{gg}$ indicated in $I_{sc}$ are replaced by the values of corresponding pixels in $I_{bg}$).
where $\theta_i$ is the clockwise angle from x-axis to the vector $OC_i$ as illustrated in Fig. 5. $C_i$ is one of the eight-connected neighbours of pixel $O(x, y)$. In Fig. 5, $L$ and $R$ are equidistant pixels from $C_i$ along the line perpendicular to $OC_i$. Their distance from $C_i$ is defined by $w$ as illustrated.

All pixels $(x, y)$ with negative values in $I_{sc}$ are processed sequentially according to Eq. (1) to produce a confidence map as described in Algorithm 1. Fig. 6 shows the process of generating a confidence map.

In order to ensure only the pixels sufficiently darker than their background are chosen, the connected neighbourhood strength is thresholded (line 3, 6 in Algorithm 1). This threshold is denoted as $\delta = \mu + k\sigma$, where $\mu$ and $\sigma$ are the mean and the standard deviation of the connected neighbourhood strengths of all negative valued pixels indicated by $I_{sc}$. In this study, $k$ was fixed to 1.

As shown in line 5-14 in Algorithm 1, if a pixel $O(x, y)$ has $N_{x,y}$ neighbouring pixels that have strength greater than $\delta$, it will appear in candidate layers 1 through $N_{x,y}$. If all of its neighbouring pixels’ strengths have higher values than $\delta$, the current pixel will appear in all layers. On the contrary, if none its neighbours’ strengths are higher than $\delta$, this pixel cannot appear in any layer.

- A higher $N_{x,y}$ value means pixel $O(x, y)$ has more darker neighbours, i.e., the dark pixel $(x, y)$ itself is more likely in the middle of a dark object.
- A lower $N_{x,y}$ means pixel $O(x, y)$ is more likely to be on the edge of a dark object.
- As the confidence layers $M_j$ are generated based on the number of connected dark neighbours, lower layers (lower $j$) such as that in Fig. 6 (c) ($M_1$) contains more pixels at the edge of a large dark object. These include those true MAs right next to the vessel network. These MA pixels, however, may have been merged into the neighbouring dark objects together with those retinal background pixels between MAs and neighbouring objects. This is due to less constraints on the connected neighbourhood strength in this layer. Those isolated MAs further away from other dark objects will easily be obtained at this layer. In addition, some low contrast MAs can only appear in lower layers.
- Higher layers (higher $j$, such as $M_5$ in Fig. 6 (g)) contain more small but well separated candidates even if they are very close to another object. This layer tends to miss those low contrast MAs, but may include small vessel fragments.
- Fig. 6 (c), (e) and (g) show the instances of such confidence layers ($j = 1, 3$ and 5). The dark pixel information in different layers complements each other and is combined later in the final confidence map.

Any candidate object in each layer that is too large to be an MA is removed (lines 16-21 in Algorithm 1). To determine a threshold $\Delta(\Delta = 100)$, an observation was made based on the average MA lesion size in the training set. The remaining objects include a number of MAs and small

### Algorithm 1 Multilayered dark object filtering

1: Process all candidate pixels in $I_{pp}$ which are negative value in $I_{sc}$
2: For each candidate pixel $O(x, y)$, calculate each of its eight connected neighbourhood strength $p_i$ in $I_{pp}$ according to Eq. 1
3: Calculate the global threshold $\delta = \mu + k\sigma$ for all $p_i$
4: Create 8 empty binary images as initial candidate layers $M_j (j = 1 \ldots 8)$
5: for each pixel $O(x, y)$ do
6: Count the number ($N_{x,y} = 0 \ldots 8$) of its neighbours whose strength $p_i$ is larger than $\delta$
7: if $N_{x,y} > 0$ then
8: for $j = N_{x,y}; j >= 1; j --$ do
9: Include the pixel $O(x, y)$ in the $j$th layer $M_j$
10: end for
11: else
12: The pixel $O(x, y)$ cannot appear at any layer
13: end if
14: end for
15: Create another empty binary image as the final confidence map $I_{bin}$
16: for each individual layer $M_j (j = 1 \ldots 8)$ do
17: Find connected components as candidate objects
18: if the size of a candidate object as smaller than $\Delta$ then
19: Add this candidate to the final confidence map $I_{bin}$
20: end if
21: end for
22: Remove any object (connected component) in $I_{bin}$ that is greater than $\Delta$
vessel fragments (Fig. 6 (d), (f) and (h)). To obtain a higher number of true positive candidates, all remaining pixels in each layer are combined into a final map in the binary image $I_{bin}$. Then, all candidate objects in the form of connected components in $I_{bin}$ are validated again. Those larger than $\Delta$ are removed. Integrating all layers can sufficiently reduce false MA candidates and extract more dark lesions located close to the vasculature. Fig. 6 (i) visualises all candidates when superimposing $I_{bin}$ on $I_{pp}$. The connected pixels on the map indicate positions of MA candidate objects.

C. Extraction of SSA-Based Cross-Section Profiles

The candidate objects on the confidence map may include MAs and the most common interfering objects, such as vessel crossings, retinal haemorrhages and small vessel fragments. Since these extracted candidates indicate the positions of the actual objects, region growing techniques [11], [13], [19] could be used to obtain the actual shape of the objects. However, this procedure can encroach into neighbouring MAs and other objects and it highly depends on the local threshold setup. To avoid this situation, we use the candidates’ profiles to extract a set of statistical features that can discriminate true MAs from non-MAs.

Work based on cross-sectional scanning of pixel intensities was been reported in [24], [25]. A limitation in the published methods is that many small vessels or vessel crossings can lead to spurious MAs. Our proposed approach based on singular spectrum analysis has largely overcome this limitation by considering candidate objects’ major axis when constructing their profiles. According to our observations, the major axis has significant difference in their profiles between circular structures (such as true MAs) and elongated structures (such as small vessels or vessel crossings). The different directions of cross-section lines can be denoted as:

$$\alpha[i] = \beta + \frac{180^\circ}{n} \times i, \ i = 1, \ldots, n$$  \hspace{1cm} (2)

where $\alpha[i]$ is the $i$th direction of the cross-section scanning lines, $\beta$ is the angle between the major axis of the candidate object and $x$-axis and $n$ is the total number of profiles to be generated. In our implementation, $n$ was set empirically to 12. We do not need to consider extra directions of profiles, since this proposed approach can ensure that the cross-section lines contain the major axis and the minor axis of this object. Fig. 7 illustrates how the cross-sections are established.

The choice of cross-section length ($r$ as indicated in Fig. 7) also affects the detection outcome. Through experimental observation, we choose $r = 31$. This is in line with what was suggested by Lazer and Hajdu [24]. Based on the above setup, we can obtain 12 sets of profiles of 31 pixels along 12 directions for each candidate object.

SSA [26] is able to decompose these profile series and then reconstruct them in order to remove the noise and enhance meaningful signals. The main advantage of SSA over other subspace-based methods such as Principal Component Analysis (PCA) is that, the size of the generated Hankel matrix as shown below can vary according to the expected number of underlying signal components, while this is fixed for PCA. Thus, a significantly better data decomposition into subspaces is achieved through SSA. After applying SSA, the key characteristics and differences between the profiles of MAs and non-MAs are more prominent. A comparison of the different filtering methods is given in Section III. A brief description of the SSA technique follows. Two complementary steps of the basic SSA algorithm include decomposition and reconstruction.

1) Decomposition: This step consists of embedding operation and singular value decomposition (SVD). The cross-sectional intensity profile $f$ with length $r$ is mapped into an $l \times k$ matrix by applying the embedding operation:

$$X = [x_1, x_2, \ldots, x_k] = \begin{bmatrix} f_0 & f_1 & f_2 & \ldots & f_{k-1} \\ f_1 & f_2 & f_3 & \ldots & f_k \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ f_{l-1} & f_l & f_{l+1} & \ldots & f_{r-1} \end{bmatrix}$$  \hspace{1cm} (3)

where $k = r - l + 1$, $l$ denotes the window length ($1 \leq l \leq r$). The trajectory matrix $X$ is a Hankel matrix which has equal elements for all the diagonals $i + j = constant$, where $i$ and $j$ are indices of rows and columns. In our implementation, $l$
was empirically set to half of \( r \) \((l = 15)\).

Then, the SVD of the trajectory matrix is created and represented as the sum of rank-one biorthogonal elementary matrices. Let \( C_x = X X^T \) and assume \( \lambda_1, \lambda_2, \ldots, \lambda_l \) are the eigenvalues of the covariance matrix \( C_x \) in decreasing order \((\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_l \geq 0)\) and denote the corresponding eigenvectors by \( u_i \) to \( u_l \). If we denote \( v_i = X^T u_i / \sqrt{\lambda_i} \), where the length of \( v_i \) is \( 1 \times l \), then, the SVD of the \( C_x = X X^T \) is calculated. The collection \((\sqrt{\lambda_i}, u_i, v_i)\) is considered as the \( i \)th eigentriple of SVD. Sanei et al. [29] mentioned that the first few eigenvalues often carry the maximum energy of the data.

2) Reconstruction: In this step, the elementary matrices are firstly divided into a number of groups and summed within each group. The indices corresponding to the \( p \) eigenvalues of the desired component are defined as \( I = \{i, \ldots, i + p - 1\} \). Then, let \( X_I = \sum_{j=i}^{i+p-1} X_j \) be the matrix \( X_I \) corresponding to group \( I \). Therefore, the overall sequence \( X \) can be written as

\[
X = \sum_{j=I_i}^{I_{m}} X_I
\]

where index \( m \) refers to the \( m \)th subgroup of eigentriples. The procedure for selection of the sets \( I_1, \ldots, I_m \) is called eigentriple grouping. For a given group \( I \), \( \sum_{j=I} \lambda_i / \sum_{i=1}^{p} \lambda_i \) is regarded as the contribution of the component \( X_I \) in expansion (4). If we are to select all components to form one group, then the original series will be reconstructed. Using the obtained Hankel matrix, the cross-section profile can be reconstructed by diagonal averaging [26].

By application of SSA we can decompose the original data into a set of principal components (PCs). Our objective is to denoise and obtain the key profiles of MA candidates. Since the eigenvalues in the decomposition step are equal to the variance of the signal in the direction of their corresponding PCs, the smallest eigenvalues are regarded as noise and the largest eigenvalues belong to the signal subspace. After testing on many images, we observed that \( \lambda_1 \gg \lambda_2 \). Hence, we chose the first eigenvector \( u_1 \) for reconstruction.

Due to the local variation, some retinal background regions, as well as the junctions or crossings of very fine vessels are extracted as MA candidates in the previous dark object filtering stage. Fig. 8 shows the SSA-based scanning profiles of an MA, a retinal background region, an elongated haemorrhage and a blood vessel crossing. As shown in Fig. 8 (a), unlike those non-MA objects, the MA profiles show an obvious inverse Gaussian-like extremum in all directions. In contrast, the sets of profiles in Fig. 8 (b), (c) and (d) are shallower and less regular. Comparing these non-MA profiles, the MAs can be discriminated from the most common interfering structures through analysing the respective sets of profiles of the candidate objects. For example, Fig. 8 (c) and (d) show that the extracted cross-sections contain the major and minor axes of the object. These filtered SSA profiles will provide the basis to extract the necessary features for classifying the candidate objects into MAs and non-MAs.

D. Feature Extraction and Classification System

As compared to those works that focused on pure pixel intensity profiles, SSA generated profiles further highlight the candidate object information and minimise the impact from the noise. After observing SSA-based profiles, we found the shapes of MA profiles are more similar in all directions than those of non-MAs. In order to increase the difference between MAs and non-MAs, a dissimilarity score was assigned to each cross-section profile of a candidate object.

A correlation coefficient formula is used to calculate the dissimilarity score between an estimated MA profile and each of the 12 scanning profiles \( f_r \). A Gaussian function \( G \) is used to generate the estimated cross-section of MA which exhibits a Gaussian shape:

\[
G = (f_{rc} - f_{rb}) \exp \left(-\frac{d_E^2}{2\sigma_f^2}\right) + f_{rb}
\]

where \( f_{rc} \) and \( f_{rb} \) are the mean values of the pixel intensities of the region inside and outside the MA candidate in \( I_{pp} \). \( d_E \) is the distance between the centroid of the candidate region and a point \( K \) within the region, and \( \sigma_f \) defines the spread of the estimated profile and is set to the value equal to half of the local candidate’s radius. An estimated intensity profile of a given MA candidate is shown in Fig. 9.

The correlation coefficient (cc) is then used to measure the similarity between an MA candidate and the Gaussian function. This value varies within \([-1, 1]\). The cc value will be
high if the two profiles match well and vice versa. This means a true MA candidate will have high scores for its profiles along all cross-section directions, while non-MA candidates will have lower scores for at least some of its profiles in respective directions. The correlation coefficient is denoted as:

$$cc = \frac{\sum_{i}(f_r[i] - \overline{f_r})(G[i] - \overline{G})}{\sqrt{\sum_{i}(f_r[i] - \overline{f_r})^2(\sum_{i}(G[i] - \overline{G})^2)}}$$

(6)

where $f_r$ and $G$ are mean values of $f_r$ and $G$, and $i$ is the pixel point along the profile scanning line, $i = 1, 2, \ldots, 31$. Each cross-section profile $f_r$ is scaled by its own dissimilarity score (cc) and then a set of features is extracted from the new profile for classification. The scaled profile $f_d$ is determined by:

$$f_d[i] = \frac{f_r[i]}{cc}, \quad i = 1, \ldots, 31.$$  

(7)

If the value of cc is close to one, the scaled profile $f_d$ should be similar to the original profile $f_r$; otherwise, $f_d$ will become more different from $f_r$ if cc is close to zero or negative. These changes will enable the features to be more discriminating for classification of MAs and non-MAs. Fig. 10 shows the comparison between the cross-section profiles with and without being adjusted using their dissimilarity scores. Comparing with the original profiles (Fig. 10 (b), (c) and (d)), the variations between adjusted profiles for the same object (Fig. 10 (f), (g) and (h)) are larger for non-MA candidates. The changes of the profile directly impact the features for classification (see Section III for the classification performance with and without such dissimilarity scores).

Once the profiles have been scaled, the slopes are defined as either increasing or decreasing slopes based on whether the sign of the difference between sequential values of a profile is positive or negative. Fig. 11 illustrates a graphical interpretation of the slopes on a scaled profile. The peak width measures the spread of MA candidates in the considered direction, denoted by $w_{peak} = slope\_inc - slope\_dec$. The values of $slope\_dec$ denote the start point of the decreasing slope, and $slope\_inc$ corresponds to the end point of the increasing slope. In total, 11 features are extracted from the scaled profiles for classification, they are:

1. The mean and standard deviation ($\mu_{w_{peak}}$ and $\sigma_{w_{peak}}$) of the peak widths of all cross-section profiles of the object.
2. The mean and standard deviation ($\mu_{h_{dec}}$ and $\sigma_{h_{dec}}$) of the heights of decreasing slopes of all cross-section profiles of the object.
3. The mean and standard deviation ($\mu_{h_{inc}}$ and $\sigma_{h_{inc}}$) of the heights of increasing slopes of all cross-section profiles of the object.
4. The compactness of an object is denoted as: $v = \sqrt{\sum_{j=1}^{n}(d_j - \overline{d})/n}$, where $d_j$ is the distance from its $j$th edge point of slope ($slope\_inc$, $slope\_dec$) to the centroid of the profile and $\overline{d}$ is the mean of the distance from each edge point to the centroid. Here $n$ is the total number of edge points ($n = 24$ in our implementation, as each candidate has 12 profiles, each profile has two edge pixels).
5. The mean and standard deviation ($\mu_{\lambda_1}$ and $\sigma_{\lambda_1}$) of the largest eigenvalues of all profiles.
6. The mean and standard deviation ($\mu_{\lambda_2}$ and $\sigma_{\lambda_2}$) of the aspect ratio, $r = \lambda_1/\lambda_2$. $\lambda_1$ and $\lambda_2$ are the respective values of the first and second largest eigenvalues of a profile.

Features (1)-(3) are typical properties of Gaussian shaped profiles. They give information about the sharpness and spread of the profiles and their difference from the surrounding regions. Compared to some published cross-sectional approaches [24], [25], we added several new features. The compactness of an object (Feature (4)) is regarded as an extra shape feature to help remove elongated structures. Furthermore, some of the false positives come from the region of the optic disc [24]. These MA-like structures have very high contrast in this region. Features (5) and (6) are used to remove these MA-like objects. The largest eigenvalue (Feature (5)) provides the object intensity level. Since the largest eigenvalue $\lambda_1$ captures most of the profile information and generally $\lambda_1 \gg \lambda_2$, the
aspect ratio (Feature (6)) indicates the variation in a given profile. The aspect ratio has a higher value for true MA profiles while a lower value means much less variation in a profile, which often belongs to retinal background or a large dark object, for example a blood vessel profile but running through the vessel direction. Detailed value ranges of these features are given in Section III.

Using these features, we compared kNN, support vector machines (SVMs) and naïve Bayes (NB) classifiers for classifying those candidates into true MAs and non-MAs. Our experiments showed that the kNN classifier gave the best performance. For the rest of the experiment, kNN is used as the main classifier. Details of the classifier comparison are provided in Section III. The final outputs of our proposed approach include a set of coordinates of detected MAs and their corresponding probabilities for being an MA.

III. Experiments and Results

A. Setups

In this work, we experimented on three sets of data, i.e. Retinopathy Online Challenge (ROC) [16], DiaretDB1 2.1 [30] and Moorfields Eye Hospital datasets. Images from different datasets were automatically resized and cropped by bilinear interpolation, to the same resolutions of those from the ROC dataset, either 768 × 576 (images of fundus with top and bottom parts missing, such as Fig. 3 (c), (e), (f), (g) and (h)) or 1389 × 1383 (full fundus images, such as Fig. 3 (a), (b) and (d)). The diameter of the circular fundus region in former resolution is approximately 800 pixels and in later resolution is approximately 1300 pixels. Despite their size and intensity variation, MAs are of similar properties so the algorithms developed in this work are suitable for both types of images.

We used 50 training images from the ROC to choose the parameters and classifier of our proposed method. While this public dataset has provided ground truth for MAs with their pixel coordinates, we manually prepared non-MA training samples. The non-MA samples contain the previously presented false positives, such as vessel bifurcations and crossings, small disconnected vessels fragments and retinal haemorrhages. The final training sets contain both true and false MA examples, including 336 MAs and 403 non-MA objects. We made observations on this training set and noted the pixel properties of key dark objects as shown in Table I. These observations helped us to establish object size references for all processing. The following gives details of various experimental setups as well as additional experiments to justify some of our design choices.

In the preprocessing stage, in order to enhance the small and dark objects that correspond to MAs, the size of a Gaussian filter should be less than the MAs diameter (5 – 10 pixels). In this study, the Gaussian filter (width = 3, variance = 1) is used. In the next step when estimating the background image $I_{bg}$, based on the observations in Table I, the size of the median filter should be wider than the widest vasculature in retinal images. It is set to $35 \times 35$.

After preprocessing, a confidence map is generated by using Algorithm 1. To find the optimal values for parameter $w$ in Eq. (1) and $k$ in Algorithm 1, an optimisation process is applied on the ROC training set. The optimisation objective is the detection accuracy on the 50 training images. Table II listed the first, second and third rank of the parameters based on the numbers of extracted candidate objects and missed true MAs. $w = 3$ and $k = 1$ give the least number of missed MAs. These values have been applied to all other experiments described in this paper. According to Table I, a size threshold $\Delta$ of 100 pixels was found to include all true MAs. Most of the vasculature will be removed in Algorithm 1 as their sizes are larger than 300 pixels, leaving the rest of candidates to be mostly small vessel fragments, small haemorrhages and true MAs.

During SSA, we need to decide the cross-section length $r$ and window length $l$ in Eq. (3). If $r$ is too small, it is hard to distinguish MAs from haemorrhages and parts of the vasculature. Furthermore, large $r$ results in increasing the computational time. We varied the value of $r$ on the training data with all the odd numbers between 1 to 100 and found that 31 pixels gives consistent results. Selecting a proper window length $l$ relies on a priori knowledge about the size characteristics of dark objects, such as MA, haemorrhages, retinal blood vessels and noise. With regards to $l$, a sufficiently large window length can effectively remove noise and preserve the original patterns of different objects. Theoretically, $l$ should not exceed $r/2$. In our implementation, $l$ was empirically set to half of the length of cross-sections ($l = 15$). Based on $l$ and the SVD of the trajectory matrix ($15 \times 15$), we obtained 15 ordered eigentriples in the decomposition.

The benefit of applying SSA is to obtain more smooth profiles and preserve the original patterns of different types of candidates. To demonstrate this, we compared SSA with other approaches of a similar nature, such as median filters (with an optimal size $M = 7$) and PCA. Fig. 12 visualises these approaches against the original intensity profile for smoothing two types of candidate profiles. Table III shows the sensitivities at 1 FP/image of various classifiers using different profiling approaches including using just original profiles. Original profiles include much local intensity variation. In order to characterise the main signal and extract the features for classification, in the experiment when dealing with the

<table>
<thead>
<tr>
<th>TABLE I</th>
<th>THE DETAILS OF THE DETECTED CANDIDATE OBJECTS OF THE TRAINING SET</th>
</tr>
</thead>
<tbody>
<tr>
<td>MA</td>
<td>Haemorrhage</td>
</tr>
<tr>
<td>Diameter</td>
<td>5-10 pixels</td>
</tr>
<tr>
<td>Area</td>
<td>&lt;35 pixels</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE II</th>
<th>COMPARISON OF THE RESULTS OF CANDIDATE OBJECT EXTRACTION USING DIFFERENT VALUES OF $w$ IN EQ. (1) AND $k$ IN ALGORITHM 1</th>
</tr>
</thead>
<tbody>
<tr>
<td>$w$, $k$</td>
<td>Total number of extracted candidate objects</td>
</tr>
<tr>
<td>3, 1</td>
<td>3571</td>
</tr>
<tr>
<td>3, 1.2</td>
<td>3123</td>
</tr>
<tr>
<td>2, 1.5</td>
<td>8456</td>
</tr>
</tbody>
</table>
original profiles we ignored very small changes in them. We also duplicated another original profile based approach proposed by Lazer and Hajdu [24], where the original profiles were thresholded to eliminate noise or artefacts as well as small changes between consecutive values of the profile and the slope height. The performance results based on the original profiles are shown in the first two columns in Table III. The filter size of the median filters is very sensitive to the intensity changes in candidate objects. Using a small size cannot remove the noise and larger ones will lose MA pattern. In our experiment, the median filter size 7 gives the best performance when using the filtered profiles for classification, compared with other sizes between values 5-15. In PCA, the first PC is used to reconstruct the candidate’s profiles. However, it retains more noise (Fig. 12 (a)) and cannot fully represent object patterns (Fig. 12 (b)). As seen in Table III, the classification performance based on SSA outperformed the rest of the profiling approaches.

In this paper we aim to investigate the benefit of applying the whole process from preprocessing to feature extraction for classification. It is not critical in terms of choosing any particular classification. In order to choose one for this work, we compared the performances of kNN, SVM and NB classifiers based on the features listed in Section II D. In this experiment, we used the training data in the ROC dataset as it contains 50 images with pixel based ground truth. Half of these images were used for training and the rest were used for testing. As seen from Table III, kNN outperformed the rest of the classifiers. In the following experiment, we use kNN for MA and non-MA classification. kNN has also higher computation speed on larger datasets. The benefit of using the proposed feature set is justified in the following evaluation on various datasets when comparing with other systems. An initial examination of the feature vectors for all ROC training data are given in Table IV.

<table>
<thead>
<tr>
<th>TABLE III</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>------------------</td>
</tr>
<tr>
<td>KNN (k = 15)</td>
</tr>
<tr>
<td>SVM</td>
</tr>
<tr>
<td>NB</td>
</tr>
</tbody>
</table>

The full training data of the ROC and the parameter setups discussed in this section have been used in all evaluations as presented in the following section (no further training sets were added for any experiment).

B. Evaluation on Various Retinal Image Datasets

1) Retinopathy Online Challenge: The Retinopathy Online Challenge is to compare the capabilities of MA detectors under the same conditions. This dataset contains 50 training and 50 testing images, and all MAs were annotated by human experts. These images were acquired by three different cameras and had different resolutions and sizes, ranging from 768 x 576 to 1389 x 1383. The results of the proposed method are used to generate a free-response operating characteristic (FROC) curve which plots the sensitivity against average number of false positives (FP) per image. Fig. 13 demonstrates the performance of our proposed MA detection with and without the use of correlation coefficient scores. The final score is measured by the average sensitivity with seven predefined FPS per image (1/8, 1/4, 1/2, 1, 2, 4, and 8 FP rate). The scores were measured independently by the Retinopathy Online Challenge team after we submitted our results on the testing data.

Table V shows the ranked results of the various methods submitted to the ROC. Our proposed method obtained an average score of 0.464, which is slightly higher than other published methods. Table VI shows the sensitivities at seven specific FP rates by our proposed method, an ensemble-based method [31] and a cross-section based method [24]. Our proposed method achieved higher sensitivity at low FP rates (1/8, 1/4, 1/2, 1 and 2 FPS/image), and our overall FROC score is slightly higher. Although the ensemble-based method achieved higher sensitivity at 4 and 8 false positives per image, according to Niemeijer et al. [16], the false positive rate of 1.08 FP/image is regarded as an indication of ‘clinically acceptable’ FP rate. Higher FP/image rates are not desirable in clinical practice.

2) DiaretdB1 2.1 Dataset: Many factors can influence the performance of MA detection, such as their local contrast, colour and location. Our proposed method was further evaluated on different lesion categories in DiaretdB1 2.1 dataset. This dataset contains 89 uncompressed retinal images with 1500 x 1152 resolution. Each MA was assigned into a different
The performance of our proposed method on the DiaretDB1 dataset and tested all images. Table VII demonstrates the sensitivities of our proposed method (without and with a correlation coefficient) at 1 FP rate for various categories of MAs. It shows our proposed method can effectively recognise MAs. However, we still have lower performance on some categories (i.e., subtle and periphery). Fig. 14 shows the performance of our proposed method on the DiaretDB1.2 dataset, comparing with the ensemble-based method [31].

3) Dataset from Moorfields Eye Hospital: The proposed method was further tested on the dataset we collected at Moorfields Eye Hospital, London. The images originally came from different population based studies (such as Kenya (Fig. 3 (a), size: 3,008 x 2,000, number: 9,878 images), Botswana (Fig. 3 (b), size: 3,456 x 2,304, number: 500 images), Mongolia (Fig. 3 (c), size: 3,888 x 2,592, number: 1,636 images), China (Fig. 3 (d), number: 579 images size: 1,380 x 1,180), Saudi Arabia (Fig. 3 (f), size: 4,288 x 2,848, number: 67 images), Italy (Fig. 3 (g), size: 1,024 x 1,024, number: 3,365 images), Lithuania (Fig. 3 (h), size: 3,072 x 2,048, number: 4,962 images) and Norway (Fig. 3 (g), size: 2,196 x 1,958, number: 840 images)) and were anonymised to the point that nobody would be able to identify the individual patients. This anonymised dataset consists of 21,536 retinal images. The images were then independently graded by human graders as normal or abnormal based on the presence or absence of MAs. In this work on the Moorfields Eye Hospital datasets, we used the same training data from the ROC and the same setups for testing on the ROC and DiaretDB1.2 datasets.

The test focused mainly on whether an image contains MAs or not. A receiver operating characteristic (ROC) curve analysis is able to evaluate this type of performance. We measured the sensitivity and specificity of the proposed MA detection: Sensitivity = TP/(TP + FN), Specificity = TN/(TN + FP), where TP is true positive, FN is false negative, TN is true negative and FP is false positive. Fig. 15 shows the ROC curve of our presented approach. The average sensitivity and specificity can achieve 96.3% and 88.4% respectively. Table VIII shows the respective Kappa values (K) of different populations. According to Altman [37], K = 1 means complete agreement and K = 0 is no agreement. The strength of agreement is considered to be ‘Good’ if K is between 0.61 – 0.80.

### C. Analysis of Misclassification

We now review the FN images that contain MAs but were labelled as normal. Most of these missed MAs are subtle, of
D. Discussion

The proposed candidate filtering process is able to significantly reduce the number of non-MA candidates and sufficiently extract more candidates located close to the vasculature. We take the advantage of a basic SSA method to filter MA candidates’ profiles. In places where a single channel signal is available, SSA separates the data into its constituent components of different subspaces efficiently. Moreover, unlike other signal decomposition methods, here, changing the embedding dimension resulting a lower decomposition error (better noise restoration). In our application, given the expected number of underlying components, SSA automatically decomposes and reconstructs each individual profile. These filtered profiles are then automatically scaled based on the correlation coefficient value to achieve more discriminative features for MAs and non-MA candidates. Based on the proposed set of features we achieved a robust detection of MAs on the data from various sources with different resolutions, quality and of different ethnic origins.

We examined our proposed algorithm using the ROC dataset. It obtained an overall slightly higher score than all other published methods, and achieved higher sensitivity at low FP rates (1/8, 1/4, 1/2, 1 and 2 FPs per image). With DiaretDB1 2.1 dataset, we evaluated our method on different lesion categories. At the rate of 1 FP/Image which is clinically acceptable, our method can effectively recognise MAs from the most common categories. We have further evaluated our system on 21,536 images provided by the Moorfields Eye Hospital. The results prove the robustness of our proposed approach, demonstrating a promising sensitivity and specificity when applied to larger datasets.

Although some vessel bifurcations and crossings are treated as MA candidates in the preprocessing stage, our proposed feature set is found to discriminate true MAs from those most common interfering candidate objects. Additionally, these features can be used to remove MA-like structures in the optic disc region without the detection of the optic disc first.

The processing time of our proposed method is approximately one minute, on a computer equipped with Intel Core i5 processor, running at 2.2 GHz. The method is currently implemented using MATLAB. This is much faster than our previous system [5], because around 50% of false MA candidates are removed after the preprocessing stage.

IV. Conclusion

In this work, we have demonstrated the scalability of our approach for localising MA in digital fundus images. Only a small number of image samples from the public domain were used for training (50 images) and the system was then tested on 21,536 previously unseen images. In our ongoing research project, we have been evaluating an automated system for retinal image analysis on very large datasets collected through diabetic retinopathy screening programmes and eye epidemiology studies from Africa, Asia (including Far East and Middle East countries) and Europe. The data used in this paper are from part of this project.

Our proposed MA detection achieved a good sensitivity and specificity on a per image basis. This is especially meaningful when this method is integrated into a reliable automated system for detecting abnormality in digital fundus images.

ACKNOWLEDGMENT

This work has been carried out in collaboration with the Reading Centre, Department of Research and Development,
NIHR Biomedical Research Centre at Moorfields Eye Hospital NHS Foundation Trust and UCL Institute of Ophthalmology, United Kingdom. The authors would like to thank the Mongolia, Kenya Norway, China, Saudi Arabia, Botswana, Italy and Lithuania medical study teams and those in the UK who contributed either by providing images or carrying out manual grading for this study. This project was supported by the NSTIP strategic technologies program in the Kingdom of Saudi Arabia (Project No.: 10-INF1262-03). The authors also acknowledge with thanks the Science and Technology Unit, King Abdulaziz University for technical support. The authors would like to thank the Engineering and Physical Sciences Research Council (EPSRC) in the UK for supporting the foundation of this work. The authors thank the Retinopathy Online Challenge team and DiaretDB1 team for making their data publicly available.

REFERENCES


