Automatic Detection of Laser Marks in Retinal Digital Fundus Images

João G. R. Almeida e Sousa, Carlos Manta Oliveira, Luis A. da Silva Cruz

Department of Physics
University of Coimbra, Portugal
Email: uc2010129800@student.uc.pt

Retmarker S.A.
Coimbra, Portugal
Email: caoliveira@retmarker.com

Department of Electrical and Computer Engineering
University of Coimbra, Portugal
Email: lcruz@deec.uc.pt

Abstract—Diabetic retinopathy (DR) is the most frequent complication of diabetes mellitus that affects vision to the point of causing blindness. In advanced stages its progress can be delayed with laser photocoagulation which leaves behind marks on the retina. Modern screening programs rely on automatic diagnostic algorithms to detect signs of DR in patients. These systems performance may be impaired when patient retina presents marks from previous laser photocoagulation treatments. Since these patients are already being treated, it is desirable to detect and remove them from the screening program. An algorithm that automatically detects the presence of laser marks in retinal images using tree-based classifiers is proposed and the results on its performance are obtained and described. Two new public accessible datasets containing retinal images with laser marks are provided in this paper.

Keywords—Diabetes, Biomedical image processing, Feature extraction, Classification algorithms.

I. INTRODUCTION

Diabetic retinopathy is characterized by a set of lesions on the retina caused by complications that accompany diabetes mellitus and it is the leading cause of preventable blindness amongst working age population. It is recommended that each diabetic patient participates in a DR screening program annually by taking non-invasive digital retinal fundus photographs. Typically an optometrist performs a first grading, and if the images contain captured significant presence of microaneurysms (MA) or exudates (EX), then the patient is referred to an ophthalmologist for further follow-up. Strategies and recommendations to implement cost-effective DR screening programs have been increasing in the last years. The main objective of these programs is to reduce the workload and manual burden of the specialists and ensure a high coverage of the target population in a short period of time [1] with the help of automated diagnostic algorithms. Some of these automatic DR detection algorithms are shortly described and reviewed by Dawn Sim et al. [2].

In DR screening initiatives, such as the one in the centre region of Portugal, every person with diabetes mellitus is called for screening and occasionally patients that already underwent laser surgery show up. These patients either do not remember having been treated or are aware of the fact that they were treated, but still believe they should participate. Such situations pose two problems: if lesions are detected, the patient will unnecessarily be scheduled for a medical appointment, when he is already being treated; if no lesions are detected, the patient will be rescheduled for annual re-screening. In either case this causes an unproductive burden to the health care system. It is therefore valuable to automatically detect laser marks that result from photocoagulation treatments on retinal fundus images. Doing so, not only treated patients are removed from the screening programs and directed to appropriate follow-up, but also a filtering step is performed before unnecessary processing of the images to detect DR lesions.

II. RELATED WORKS

Despite the interest in being able to detect the presence of laser marks left by photocoagulation treatment, very few publications have been produced describing methods for their automatic detection. In [3] Dias et al. a method is described which is an adaptation of a previously developed retinal image quality evaluation algorithm [4] to detect the presence of laser marks in digital fundus images. In [5] Syed et al. proposed an algorithm which is based on classification using support vector machines (SVM) with inputs consisting of 3 color-domain features, 2 texture-domain features and 4 shape features. In [6] Tahir et al. proposed a classifier based on minimum distance clustering to decide if an image represented by 10 scalar features contains laser marks. The features used include a measure of spatial compactness of the presumed laser marks as well as nine other values quantifying color and intensity such as maximum value of hue and saturation, mean and maximum of intensity (luminance) as well as mean and maximum values of the red and green color channels. The classifiers used were trained and tested using a non-publicly available dataset containing 380 retinal images.

III. AUTOMATIC LASER MARKS DETECTION ALGORITHM

The laser marks detection algorithm proposed classifies an input retinal image as either "Laser" or "No Laser" (as illustrated in Fig. 1). The approach followed is typical for this type of image classification problem and involves a pre-processing stage that prepares the images for several segmentation steps that identify candidate laser marks, for which several
features are computed. The classification is done using tree-based algorithms that operate on those features. The datasets used to develop the classification algorithm and assess its performance are now described.

A. Materials

Eight public datasets and three proprietary datasets were used in this work. All the public datasets contained exclusively images without laser marks and thus were labeled as "No Laser" and the proprietary datasets contain images with and without laser marks.

Table I briefly describes the datasets containing retinal images available to the scientific community and Table II shows the proprietary datasets used in this work.

<table>
<thead>
<tr>
<th>TABLE I. SUMMARY OF THE PUBLIC DATASETS USED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dataset</td>
</tr>
<tr>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>Messidor (M) [7]</td>
</tr>
<tr>
<td>e-ophtha MA (EOMA) [9]</td>
</tr>
<tr>
<td>e-ophtha No MA (EONMA) [9]</td>
</tr>
<tr>
<td>e-ophtha EX (EOEX) [9, 10]</td>
</tr>
<tr>
<td>e-ophtha No EX (EONEX) [9, 10]</td>
</tr>
<tr>
<td>Vessel-Based Registration (VBR) [11]</td>
</tr>
<tr>
<td>50 Healthy People (HP) [12]</td>
</tr>
<tr>
<td>Foveal Avascular Zone Detection (FAZD) [13]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TABLE II. SUMMARY OF THE PROPRIETARY DATASETS USED</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dataset</td>
</tr>
<tr>
<td>---------------------------------------------</td>
</tr>
<tr>
<td>Laser Marks Dataset - DR Screening (LMD-DRS)</td>
</tr>
<tr>
<td>Laser Marks Dataset - Before and After Photocoagulation Treatment (LMD-BAPT)</td>
</tr>
<tr>
<td>Proprietary Dataset - Juie Disc (PIJD) [14]</td>
</tr>
</tbody>
</table>

The datasets LMD-DRS and LMD-BAPT were made publicly available by the authors and they can be accessed and downloaded at http://beam.to/lmd.

B. Pre-processing

Before classification each image is pre-processed to ease the following segmentation stage. The image is first circularly cropped leaving only a central region of interest (ROI). This operation uses a mask similar to the one shown in Fig. 2(a). Fig. 2(b) shows an example of a resulting cropped image. During the photocoagulation treatment medical doctors avoid damaging the Vascular Network and the Optic Disc (OD) and so there will be no laser marks on the OD region or over the vascular network and its vicinity. This fact is used to define a binary mask containing the pixels of the OD region and of the vascular tree (see Fig. 2(c)). These pixels will not be considered during the identification of the laser mark candidates. OD detection is based on the method from [15] and the blood vessels extraction is based on the Contourlet Transform as defined in [16].

Poor illumination of the retina and other problems originated during image capture can result in fundus images with large spatial variations in the locally averaged luminances. These problems are minimized by first converting the image to the L*a*b* colorspace followed by the application to each pixel of a local averaging operation with a square kernel with constant value. This filtering operation corrects most of the uneven illumination problems as it can be seen in the example of Fig. 2(d) and Fig. 2(e). Then, an adaptive histogram equalization is applied to the corrected L* channel to improve the contrast of the image after which the image is converted back to the RGB colorspace. Additionally three information channels, Red, Hue and Saturation, are computed for use in the calculation of several intensity-based features. The last pre-processing step is the application of a 5x5 median filter to the green channel to reduce its noise. The resulting luminance image is shown in Fig. 2(f). For more detailed information about the pre-processing steps please refer to [8, subsection 4.1].

Fig. 2. Pre-processing of retinal image from Fig. 1 (a) ROI mask (b) Cropped (c) Vessels and OD mask (d) Luminance channel before uneven illumination correction (e) Luminance channel after uneven illumination correction (f) Fully pre-processed image.

C. Candidate Laser Mark Regions Identification

An observation of retinal images of patients that underwent photocoagulation treatment revealed that laser marks are not randomly distributed over the entire retinal area but tend to occur in clusters in the periphery regions, away from the optical centre and usually have a circular shape. Therefore one way to decide whether an image has laser marks is to identify image blobs and determine if their spatial distribution (expressed through different features) is consistent with that of laser marks. Thus the first step to be carried out is the identification of small patches of pixels that are likely to be laser marks. Three different segmentation algorithms were used in this step as described in the following paragraph.

The Circular Hough Transform (CHT) [17] is very frequently used to detect circles in images from diverse fields.
of study. In this case, the CHT detects the marks that have a bright circular shape. Compared to the classical CHT, instead of using a gaussian filter, we used a bilateral filter because the gaussian filter smoothes the image and eliminates important information while the bilateral filter enhances the edges [18]. The CHT radius range used goes from 6 to 30 pixels for images with a resolution of 584x768 pixels. As a second segmentation algorithm we used the Frangi Vesselness Filter (FVF) described in [19] which was able to isolate further blobs missed by the CHT. The FVF is based on computation of the eigenvalues of the image’s Hessian matrix, and was originally proposed to identify tubular structures in the vascular tree of angiographies, but it can also be used to detect the vascular network and other dark blob-like pixel groups, such as dark laser marks, in our target retinal images. The remaining segmentation algorithm used to identify possible laser marks is the method which we called Laser Mark Segmentation (LMS) and described in [20, pages 133–138]. This method was modified to use only the roundness ($f_1$) and solidity ($f_2$) parameters instead of the original four parameters. As conditions of detection we used $1 < f_1 < 2.5$ and $f_2 > 0.8$.

For every candidate region, the center and radius were computed, together with a likelihood parameter which indicates how similar to a circle is the region. Potential laser marks located on the OD or superimposed to the vascular network were removed. Fig. 3 shows the original retinal image with the candidates that resulted from the application of the three segmentation algorithms superimposed and color coded according to the originating detection method.

![Retinal image with detected candidate regions superimposed. Key: CHT - yellow, FVF - blue, LMS - green.](image)

**D. Feature Computation**

After identifying the regions candidate to be laser marks, a total of 65 features are computed and used as input to a classifier that determines if the retinal image belongs to the "Laser" or to the "No Laser" categories.

The features used are divided in four categories:

- **Geometrical Descriptors** - 12 features that include the number of candidate regions detected by each segmentation algorithm (3 features), the total area of blobs detected by each method (3 features), average radius (1 feature), radius variance (1 feature), and averages of the likelihood values described in the previous subsection, as well as two measures that are the weighted areas of CHT and FVF detected regions where the weighting factors are the likelihoods computed before.

- **Texture Descriptors** - 27 features that represent texture statistics. This is the only category where the features are computed directly from the processing image and not from the candidate regions. There are six texture descriptors described in Gonzalez and Woods in [21, pages 464–467] computed on the green channel level histogram and the remaining 21 are Gray Level Co-occurrence Matrix (GLCM) based features: 11 Haralick features [22] plus 10 features based on [23] and [24].

- **Spatial Distribution Descriptors** - 10 features that describe the distribution of the candidate regions in the retinal image. These features represent the distances between the laser marks and measure their dispersion, clustered nature or distribution randomness. These features are computed based on spatial measures of the convex hull [25] enclosing the candidate regions, the covariance matrix of the centroids coordinates and the Moran I spatial autocorrelation [26].

- **Intensity-based Descriptors** - 16 features based on the intensity values of the candidate regions for each channel computed during the pre-processing step (red, hue and saturation) plus on the median filtered green channel of the pre-processed image. Four intensity-related features are computed for each of those channels following the approach delineated by Tahir’s work [6].

It is important to state that some of these features are normalized according to the image resolution and to the OD size so that the proposed system is robust to variations of the image resolution. A full description of these 65 features can be found in subsection in [8, subsection 4.4].

**E. Feature Selection and Classification Procedures**

Feature selection [27] is an important process in any classification problem because if properly done it allows the reduction of the dimensionality of the feature space by the removal of redundant, irrelevant and noisy data. Even though the 65 features described intuitively seem relevant, it is possible that only a small subset of them are in fact relevant for the classification. As described later a relevant subset of features was identified based on their information gain and gain-ratios. The features selected are then used as input to each of the four tree-based classifiers chosen in this work: pruned C4.5 Decision Tree (DT) [28], Random Forest with 5 trees (RF5), Random Forest with 50 trees (RF50), Random Forest with 500 trees (RF500) [29]. The training dataset used was a combination of the retinal images in the LMD-DRS and EONMA datasets. These two datasets were chosen because their images have different resolution and were captured by different cameras, and so the training data are diverse thus improving the training process resulting in a robust classification system.

As mentioned before the features were selected by their usefulness for the classification. This selection was done by first ranking all 65 features according to their information gains (IG) and gain ratios (GR). A threshold $t = 0.1$ was used with both the IG and GR values and the features were declared relevant only if their GR or IG were higher than $t$. Therefore two different subsets of features were obtained after this filtering step. To remove the duplicates, these two subsets
are used as input to the wrapper method, which returns a predictive performance of the subset of features that provides the best accuracy for each of the four tree-based classifiers by making use of the classifier itself and performing an inner stratified 5-fold cross-validation on the training dataset (LMD-DRS + EONMA). The search method applied was the sequential forward selection with a stop criterion of 15 nodes.

Table III shows the best subset of features and the respective accuracy obtained for each classifier used during the application of the wrapper method. The headings indicate which information estimator achieved the best accuracy for each classifier. It is noteworthy that no texture descriptor was selected.

TABLE III. BEST SUBSET OF FEATURES FOR EACH TREE-BASED CLASSIFIER AND RESPECTIVE ACCURACY

<table>
<thead>
<tr>
<th>Feature</th>
<th>Laser (45.0/5.0)</th>
<th>Laser (6.0/1.0)</th>
<th>Laser (109.0/1.0)</th>
<th>Laser (15.0)</th>
</tr>
</thead>
<tbody>
<tr>
<td>weighted_area_CHT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean_laser_trace</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>max_laser_trace</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean_laser_green</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean_laser_red</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>var_laser_green</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>var_laser_red</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>number_LMS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean_laser_green</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean_laser_red</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>weighted_area_FVF</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean_laser_trace</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>max_laser_trace</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean_laser_green</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean_laser_red</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>var_laser_green</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>var_laser_red</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>number_LMS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean_laser_green</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean_laser_red</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>weighted_area_CHT</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean_laser_trace</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>max_laser_trace</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean_laser_green</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean_laser_red</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>var_laser_green</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>var_laser_red</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>number_LMS</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean_laser_green</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>mean_laser_red</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Accuracy: 92.8% 93.0% 93.3% 93.1%

IV. RESULTS

The quality of the proposed algorithm was evaluated by the Sensitivity (SENS) and Specificity (SPEC) performance metrics. These two parameters measure its ability to correctly classify instances and are computed according to the expressions in Equation 2. The positive case is a retinal image classified as having laser marks.

\[
SENS = \frac{TP}{TP + FN} \quad \text{and} \quad SPEC = \frac{TN}{TN + FP} \quad (1)
\]

where
- TP are true positives, meaning the number of retinal images with laser marks classified as "Laser".
- TN are true negatives, meaning the number of retinal images without laser marks classified as "No Laser".
- FP are false positives, meaning the number of retinal images without laser marks classified as "Laser".
- FN are false negatives, meaning the number of retinal images with laser marks classified as "No Laser".

A stratified 5-fold cross validation was performed on the merged dataset (LMD-DRS + EONMA) using each of the four tree-based classifiers with the corresponding selected features. Moreover, each classifier was trained using the dataset (LMD-DRS + EONMA) and tested on the remaining datasets. The performance measures were computed and their values are listed in Table IV.

The retinal images of the 9 datasets used for testing were merged and resulted in a single dataset containing a total of 1749 images: 135 images "Laser" and 1614 images "No Laser". This merged dataset was used to test the tree-based classifier that showed the best performance, which was the decision tree, and the results are shown in Table V where PPV is the Positive Predictive Value and NPV is the Negative Predictive Value.

\[
PPV = \frac{TP}{TP + FP} \quad \text{and} \quad NPV = \frac{TN}{TN + FN} \quad (2)
\]

Analysis of the results for LMD-BAPT, one realises that retinal image 07_B2.jpg was the only false positive. As for the 13 false negatives, it can be concluded at least one "Laser" image per patient was correctly classified by the algorithm, except for patient 09 where none of the 3 "Laser" images was correctly classified. Therefore, 7 out of 9 patients were accurately detected using the trained DT classifier.

Training the pruned C4.5 classifier using the dataset (LMD-DRS + EONMA) resulted in the DT model shown in Fig. 4 which uses only 8 of the pre-selected 13 features.
V. CONCLUSION

Photocoagulation treatments of diabetic retinopathy leave scars in the retinal tissue. These marks need to be detected both to adapt further image processing operations done in the context of automatic diagnostic as well as to avoid repeated screening of patients already treated. This work proposes a simple and effective algorithm to detect these laser marks. Results show stable performance across heterogeneous datasets and robustness to changes of resolution. Furthermore it is shown that the best performance is attained with a computationally simple tree-based classifier using only 8 input features.

Comparison with the only published work who uses publicly available datasets [3] shows that the method proposed here has a higher sensitivity of 97% vs. 63%.

A second important contribution of this work to the retinal image research area is the offer for public use of a fully classified set of test images [14], containing both images with and without laser marks, at different resolutions.

ACKNOWLEDGMENT

The authors would like to thank AIBILI and CCC, in the persons of their directors, Dr. José Cunha-Vaz and Dr. António Travassos for providing and authorizing the use and publication of the LMD-DRS and LMD-BAPT image sets. The authors also thank Retmarker S.A. for the support provided as well as FCT suport of R&D Unit UID/EEA/50008/2013 and FCT grant SFRH/BSAB/113682/2015.

REFERENCES


