Hierarchical Local Binary Pattern for Branch Retinal Vein Occlusion Recognition

Zenghai Chen[1], Hui Zhang[2], Zheru Chi[1], and Hong Fu[1,3]

[1]Department of Electronic and Information Engineering, The Hong Kong Polytechnic University, Hong Kong
[2]Yancheng Institute of Health Sciences, Yancheng, Jiangsu, P.R.China
[3]Department of Computer Science, Chu Hai College of Higher Education, Hong Kong
zenghai.chen@connect.polyu.hk, zhanghui@ycmc.edu.cn,
chi.zheru@polyu.edu.hk, hongfu@chuhai.edu.hk

Abstract. Branch retinal vein occlusion (BRVO) is one of the most common retinal vascular diseases of the elderly that would dramatically impair one's vision if it is not diagnosed and treated timely. Automatic recognition of BRVO could significantly reduce an ophthalmologist's workload, make the diagnosis more efficient, and save the patients' time and costs. In this paper, we propose for the first time, to the best of our knowledge, automatic recognition of BRVO using fundus images. In particular, we propose Hierarchical Local Binary Pattern (HLBP) to represent the visual content of an fundus image for classification. HLBP is comprised of Local Binary Pattern (LBP) in a hierarchical fashion with max-pooling. In order to evaluate the performance of HLBP, we establish a BRVO dataset for experiments. HLBP is compared with several state-of-the-art feature presentation methods on the BRVO dataset. Experimental results demonstrate the superior performance of our proposed method for BRVO recognition.

1 Introduction

Branch retinal vein occlusion (BRVO) is the second most common retinal vascular disease after diabetic retinopathy, with a prevalence range from 0.6% to 1.1% in the population [1]. BRVO is a blockage of the small veins in the retina. It usually occurs in the elderly, and can be caused by hypertension, cardiovascular disease, obesity, etc. Without timely treatment, BRVO would lead to macular edema, intraretinal hemorrhage, surface wrinkling retinopathy, and vitreous hemorrhage, which can then cause vision impairment or even blindness to the patients. These severe complications can be prevented or alleviated if BRVO is diagnosed early and treated timely. The diagnosis of BRVO is mainly made by analyzing a patient's fundus images or fluorescein angiography images. Although fluorescein angiography images provide more details about one's retinal conditions, the acquisition process is invasive and costly. By contrast, the acquisition of fundus images is non-invasive and inexpensive [2]. Moreover, an ophthalmologist can diagnose BRVO effectively by analyzing fundus images only. The recognition of BRVO is thus based on fundus images in this paper.

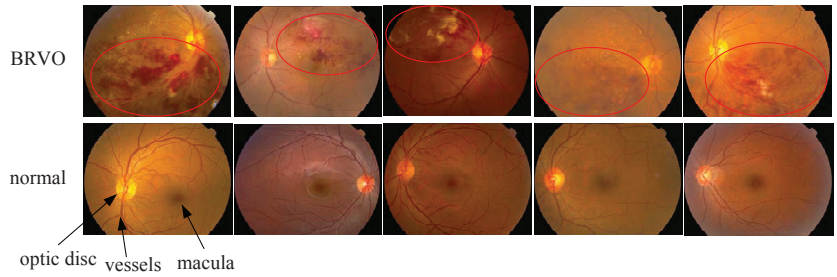

Fig. 1. BRVO fundus images and normal fundus images.

Analyzing every fundus image by an ophthalmologist can get high diagnosis accuracy. But this would increase the ophthalmologist's workload, in particular when the ophthalmologist has a large number of images to examine every day. Furthermore, the patients need to pay more money and spend more time to waiting for the results if all the images are analyzed by the ophthalmologist. If the fundus images can be automatically recognized and diagnosed with an acceptable accuracy, the workload of an ophthalmologist and the cost and time of the patients can be significantly reduced. Automatic recognition of fundus images can be a powerful auxiliary diagnosis tool to an ophthalmologist. For instance, the automatic recognition system could make preliminary results for all the patients' cases first. The ophthalmologists could then concentrate on cases whose results look suspicious or cases specially required by the patients. It has been shown that automatic retinopathy recognition could significantly reduce the workload of manual image graders (usually ophthalmologists) by 50% [3]. In addition to that, automatic recognition systems would help improve the diagnosis conditions of rural areas that are short of ophthalmologists. Other advantages of an automatic recognition system for fundus images include release from fatigue and improved repeatability [4]. It is meaningful, therefore, for us to propose an automatic BRVO recognition method based on fundus images in this paper.

Figure 1 shows some color fundus images of BRVO in the first row, and some normal eyes' images in the second row for a comparison. The main components of a normal retina include blood vessels, optic disc and macula. For BRVO, the blockage of small veins in the retina causes retinal hemorrhages, retinal edema and intraretinal microvascular abnormalities etc. There is why we can see from Figure 1 that BRVO fundus images have abnormal regions, as pointed out by red circles. The purpose of our work is to automatically and precisely distinguish the BRVO images from the normal images. The fundus images are captured in different positions and illuminations, which would increase the recognition difficulty. There are some research papers working on the detection of microaneurysms and diabetic retinopathy [2, 4–6]. To the best of our knowledge, however, no paper working on the automatic recognition of BRVO has been published. One main method to process fundus images, such as the detection of diabetic retinopathy [4–6], performs the detection or segmentation of vessels and/or optic disc first,

and then conducts further recognition based on the detection or segmentation result. One drawback of this method is that the the performance of the detection or segmentation of vessels and/or optic disc remarkably affects the final recognition result. In light of that, we would like to process a fundus image by extracting its global visual features, without any detection or segmentation of vessels and/or optic disc.

Computer vision achieved rapid developments in the past two decades. Numerous successful feature representation methods have been proposed, e.g., Local Binary Pattern (LBP) [7], Histograms of Oriented Gradients (HOG) [8], SIFT [9], Spatial Pyramid Matching (SPM) [10], Gist [11], and CENTRIST [12] etc., among which LBP is a simple and efficient texture descriptor. LBP has been successfully applied to various computer vision tasks, e.g., pedestrian detection [13], segmentation [14], face analysis and recognition [15–17], biomedical image analysis [18], etc. Different variations of LBP have been proposed as well [19–22]. As shown in Figure 1, the BRVO images have strong texture features. It is well known that LBP is a powerful texture descriptor. Therefore, we choose LBP for solving the BRVO recogniton problem. In this paper, we propose Hierarchical Local Binary Pattern (HLBP) for BRVO recognition. The architecture of HLBP is motivated by deep learning, a type of promising machine learning algorithm boomed in recent years.

Deep learning started to attract increasingly significant attention in both academic and industrial communities since 2006 when Hinton and Salakhutdinov proposed a novel algorithm to train deep neural networks effectively [23]. It has been demonstrated that deep architectures can extract high-level and more abstract features than shallow architectures [24], making deep architectures a good solver for image recognition problems. Convolutional neural networks (CNNs) are one of the most successful deep architectures for image recognition. They have shown state-of-the-art performance for different problems, e.g., image classification [25], scene labeling [26], facial point detection [27], video classification [28], etc. A CNN consists of multiple (usually two or three) stages. Each stage has a convolution layer and a subsampling layer. In general, a CNN is a hierarchical combination of coding (convolution) and pooling (subsampling) operators. By following this idea, we constructs HLBP, where LBP acts as the coding operator, for BRVO recognition. The architecture of HLBP will be elaborated in Section 2.2. In order to evaluate HLBP, we establish a BRVO dataset for experiments, and compare the performance of HLBP with several state-of-the-art feature representation methods. The rest of this paper is organized as follows. Section 2 describes LBP and the proposed HLBP. Section 3 introduces a BRVO dataset, and reports experimental results. Section 4 concludes this paper, and points out future research work. Before ending this introductory section, it is worth mentioning the contributions of this paper as follows.

1. We propose for the first time, to the best of our knowledge, automatic recognition of BRVO using fundus images. The automatic recognition of BRVO could significantly reduce an ophthalmologist's workload, improve the diagnosis conditions of rural areas, and save the patients' time and costs.

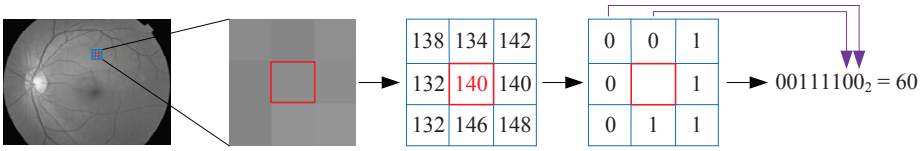

Fig. 2. Calculation of local binary pattern.

2. We propose a new feature representation method, termed hierarchical local binary pattern, to effectively characterize the visual content of a fundus image for recognition.
3. We establish a BRVO dataset to evaluate the performance of the proposed method and other state-of-the-art methods. The dataset is made public for free use, such that people can conduct experiments and comparisons conveniently.

2 Methodology

This section presents the methodology used in our work. In particular, LBP is introduced first, and then the proposed HLBP is elaborated.

2.1 Local Binary Pattern (LBP)

LBP is a local texture descriptor. LBP is calculated by comparing a central pixel's intensity with its neighboring pixels. Figure 2 illustrates the calculation procedure of LBP. For each 3×3 gray patch, compare the intensity value of the central pixel and its eight neighboring pixels. If a neighbor is bigger than or equal to the central pixel, a bit 1 is set in the corresponding neighbor's position. Otherwise, a bit 0 is set. After that, all the eight neighbors are set to binary bits. The eight binary bits are then concatenated in a clockwise or counterclockwise order, to form a 8-bit binary number. The binary number can be finally converted to a decimal number in [0 255]. The decimal number is the LBP value of the central pixel. Mathematically, the LBP of a central pixel c can be formulated as follows [7]:

$$LBP_c = \sum_{p=0}^{7} 2^p s(g_p - g_c), \quad s(x) = \begin{cases} 1, if\ x \geq 0; \\ 0, otherwise. \end{cases} \tag{1}$$

where g_c is the intensity of the central pixel c, and g_p is the intensity of the neighboring pixel p. Using Eq. (1), we can have LBP values for all the pixels (the boundary pixels could be removed), and calculate a histogram for the LBP values as the representation of an image.

Figure 2 and Eq. (1) show the calculation of LBP with eight neighboring pixels. There are also schemes that consider circles with larger radiuses as neighborhood and use more than eight neighbors for the calculation of LBP [7]. In

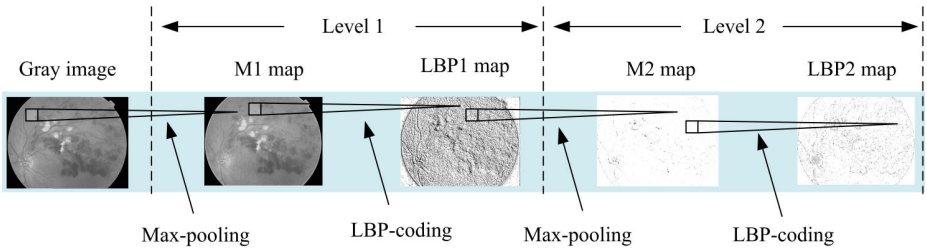

Fig. 3. The proposal hierarchical local binary pattern method.

this paper, however, we only take into account eight-pixel neighborhood, as this is the most simple and popular scheme for the calculation of LBP.

LBP is a local texture coding method. It can be roughly regarded as a convolution or filtering operator as well. The convolution of CNN for the central pixel of a 3×3 patch can be defined as follows:

$$y_c = f(\sum_{p=1}^{9} k_p x_p + b), \tag{2}$$

where b is a bias, and $f(\cdot)$ is a transfer function. By observing Eq. (1) and Eq. (2), we can find that when the bias $b = 0$ and $f(\cdot)$ is a linear function, LBP is similar to a convolution operator. The main differences include: (1) LBP performs "convolution" on the differences between the central pixel and its neighbors, rather than on the original pixels; (2) LBP uses fixed weights 2^p, while the weights k_p in CNN are trainable. Intuitively, we can construct hierarchical local binary pattern by following the architecture of CNN: hierarchical combination of coding and pooling.

2.2 Hierarchical Local Binary Pattern (HLBP)

The proposed HLBP is shown in Figure 3. HLBP consists of two levels, each of which has a max-pooling layer and a LBP-coding layer. HLBP process an image as follows. Firstly, Level 1 receives an gray image as input. Max-pooling is first performed on the input image to generate a feature map M1. LBP is then conducted on M1 map in the LBP-coding layer to generate a LBP1 feature map. Secondly, LBP1 map is fed to Level 2 to perform max-pooling and LBP-coding, as a result of which M2 map and LBP2 map are produced. Thirdly, the histogram of LBP1 map and the histogram of LBP2 map are calculated, respectively. The two histograms are concatenated to form a feature vector as the representation of the input image.

The formulation of LBP-coding has specified in Eq. (1); while the max-pooling with a $m \times m$ window is defined as follows:

$$y = \max(x_i), \quad i \in \{1, 2, \cdots, m \times m\}, \tag{3}$$

where x_i is the gray intensity of pixel i. The reason for us to choose max-pooling rather than other pooling schemes is that max-pooling is invariant to small translations [26]. The max-pooling window slips over the whole image or LBP1 map with a step size of one pixel to generate M1 or M2 map. Thus, given an $n \times n$ image (or LBP1 map) and $m \times m$ pooling window, the resulting M1 (or M2) map would be in size $(n-m+1) \times (n-m+1)$. The pooling window size is an essential component that would affect the final performance. An intuitive consideration is that the pooling size should not be too large. Otherwise, the pooling result would lose an image's important local properties. The effect of pooling size on the recognition accuracy will be studied in Section 3.

It is worth pointing out that CNN uses a coding-pooling order, while HLBP uses a pooling-coding order. There are two reasons for us to use a pooling-coding order. Firstly, it was demonstrated that patch-based LBP achieves better performance than LBP [29]. Performing LBP on max-pooling results is like performing LBP on image patches, as each max-pooling value is calculated in a small window (patch). Secondly, according to our observation, HLBP using pooling-coding (max-LBP) order performs much better than that using coding-pooling (LBP-max) order. One possible explanation could be that max-pooling-based LBP or patch-based LBP is able to incorporate spatial information of neighboring small regions/patches, since LBP is run on small regions/patches. The extracted features are hence more discriminative. The operation of max-pooling increases the input values. So the feature maps in Level 2 are much whiter (higher gray values) than that of Level 1, as shown in Figure 3. After two levels, all of the feature maps' values tend to be 255 (the highest value of an 8-bit gray image). In other words, it is hardly to extract useful features after Level 2. There is why we only consider two levels in HLBP. The working mechanism of HLBP could be as follow. LBP2 can be regarded as a LBP of LBP, or a feature of feature. It is more discriminative but sensitive. In the recognition phase, LBP1 is to classify most of the relatively easy samples; while for samples that seem difficult to LBP1, LBP2 is expected to provide better results.

3 Experiments

This section presents the experiments. A BRVO datasest is first described, and then the experimental results and corresponding analysis are presented.

3.1 Experimental Dataset

With the help of a hospital, we establish a BRVO dataset for our experiments.[1] The dataset has in total 200 fundus images acquired from 200 persons, 100 BRVO fundus images and 100 normal fundus images. All the images are of size 768×576. Some of the images are shown in Figure 1. All the original fundus images are in color. But only gray information is used in this paper. We use

[1] The dataset can be downloaded on: http://pan.baidu.com/s/1ntohK5V

Table 1. Accuracy rates (%) of different methods using SVM with linear kernel.

Fold no.	Gist [11]	HOG [8]	SPM [10]	CENTRIST [12]	LBP [7]	HLBP
1	85	90	95	95	95	95
2	90	90	95	85	85	100
3	95	100	100	95	100	100
4	95	100	95	90	90	95
5	100	100	95	90	95	95
6	100	85	100	85	100	100
7	95	95	85	90	80	95
8	95	95	95	85	85	95
9	100	95	100	90	85	100
10	100	90	90	85	85	95
Std	4.97	5.16	4.71	3.94	7.07	**2.58**
Mean	95.5	94.0	95.0	89.0	90.0	**97.0**

Matlab built-in function rgb2gray to convert the original RGB images into gray images.

We adopt ten-fold evaluation scheme. That is, the dataset is randomly partitioned into ten folds. Each fold has ten BRVO images and ten normal images. Each time one fold is used for testing, and the remaining nine folds for training. As a result, ten results of ten testing folds are obtained. The average accuracy of the ten results is used as the final performance for a comparison. SVM classifier [30] is employed for classification. The parameters of SVM is determined by two-fold cross-validation on the training data. Both linear kernel and RBF kernel will be investigated.

3.2 Experimental Results

HLBP will be compared with LBP, CENTRIST, Gist, HOG, and SPM. The two histograms of LBP1 and LBP2 of HLBP are normalized to sum to one, respectively. Thus, each HLBP feature vector sums to two. Both LBP and CENTRIST are normalized to sum to one as well. Note that CENTRIST highly relates to LBP, with a difference in bit ordering. The features of Gist, HOG, and SPM are normalized to [-1 1], respectively.

The recognition accuracy rates of different methods for ten testing folds and the overall performances are tabulated in Table 1 and Table 2. Table 1 uses linear kernel, while Table 2 uses RBF kernel. The HLBP compared in Table 1 and Table 2 utilizes a 4×4 max-pooling size. As can be seen, HLBP outperforms the other methods in terms of mean accuracy rate and standard deviation, for both linear kernel and RBF kernel. In particular, the improvement of HLBP over LBP is significant, indicating that the hierarchical architecture could enhance the discriminative power of LBP. As CENTRIST is similar to LBP, they achieve similar overall accuracy rates. The standard deviation of LBP with linear kernel is relatively high. The main reason is that it performs poor on Fold 7. Gist, HOG, and SPM achieve better performance than LBP and CENTRIST. One

Table 2. Accuracy rates (%) of different methods using SVM with RBF kernel.

Fold no.	Gist [11]	HOG [8]	SPM [10]	CENTRIST [12]	LBP [7]	HLBP
1	95	90	95	95	95	95
2	90	90	95	85	85	100
3	90	100	95	90	100	100
4	95	100	100	95	95	95
5	100	100	95	90	95	95
6	100	90	100	90	95	100
7	95	95	85	95	80	90
8	95	95	95	85	90	95
9	95	95	100	95	90	100
10	100	95	90	80	90	100
Std	3.69	4.08	4.71	5.27	5.80	**3.50**
Mean	95.5	95.0	95.0	90.0	91.5	**97.0**

reason could be that their feature coding algorithms encode spatial information between regions. For example, SPM exploits spatial pyramid to encode spatial relationships of regions in different scales. HOG organizes oriented gradient using cells and blocks. Gist is originally designed to represent the dominant spatial structure of an image. Thanks to the max-pooling and the hierarchical architecture, HLBP is able to incorporate the spatial information of an image, as a result of which high performance can be achieved. By comparing Table 1 and Table 2, we can see that RBF kernel cannot always guarantee better results than linear kernel.

In order to investigate the effect of max-pooling size on the recognition accuracy, we test several pooling sizes, and the results of using linear kernel are depicted in Figure 4. Figure 4 shows the curves of recognition accuracy rates for three cases: (1) using LBP1 only; (2) using LBP2 only; and (3) using both LBP1 and LBP2 (i.e., HLBP). As shown in Figure 4, LBP1 always outperforms LBP2. Their combination LBP1+LBP2 can obtain better results for a moderate pooling size. This indicates that LBP2 is complementary to LBP1. LBP2 could recognize samples that seem difficult to LBP1. By comparing Figure 4 and Table 1, we can see that for pooling sizes from 3×3 to 6×6, LBP1 achieves higher accuracy than LBP, meaning that performing LBP on the max-pooling result is a better choice than performing LBP on the original image pixels. HLBP (LBP1 + LBP2) achieves the best performance on the BRVO dataset when 4×4 max-pooling size is adopted. The performance gradually reduces when the pooling size decrease or increase from 4×4 size. This is reasonable, because if the size is too small (e.g., 2×2), the spatial information incorporated is not significant enough; while if the size is too large, local spatial properties would lose. When the pooling size is too large (e.g., 7×7 or 8×8 in Figure 4), the accuracy rate of LBP2 drops remarkably, and the combination of LBP1+LBP2 does not make a better result than LBP1. Figure 4 demonstrates that it is crucial to adopt a moderate (e.g., from 3×3 to 5×5) max-pooling size, if HLBP is expected to achieve a good performance.

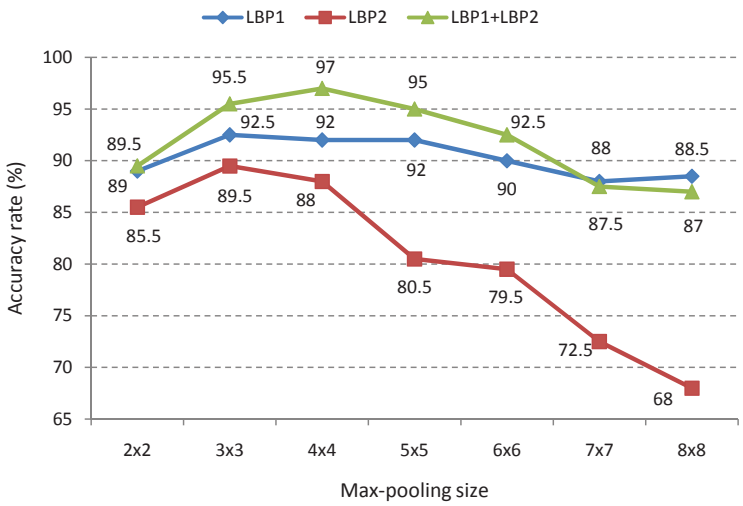

Fig. 4. Accuracy rates (%) of HLBP using different max-pooling sizes and linear kernel.

Figure 5 shows the six images that are misclassified by HLBP using a 4×4 pooling size and linear kernel. As can be seen, the two misclassified BRVO images have only a few irregular dark ares. They look similar to normal images. For the four normal images, different illuminations make them difficult to recognize.

4 Conclusion and Future Work

We propose a new feature representation method HLBP, based on LBP and max-pooling, for automatic recognition of BRVO in this paper. The proposed method is simple but efficient. We establish a BRVO dataset to evaluate the proposed HLBP. The performance of HLBP is compared with widely-used feature presentation methods. Experimental results demonstrate that HLBP with a 4×4 max-pooling size outperforms other methods. The improvement of HLBP over LBP is significant. The performance of HLBP reduces as the max-pooling size decreases or increases from 4×4. When the max-pooling size is too large, combining LBP1 and LBP2 would not produce a better result. It is, therefore, important to choose a proper pooling size. We recommend that the max-pooling size should be from 3×3 to 5×5.

Although HLBP performs well on the BRVO dataset, more experiments on various types of images need to be conducted to examine the performance of HLBP. In the future work, we would like to apply HLBP to other image recognition tasks, e.g., face recognition, large-scale scene classification etc. Moreover, the BRVO dataset used in this paper is not very large. We will continue to collect fundus images to enlarge the dataset. We would like to establish a large retinopathy dataset that contains images of not only BRVO, but also other retinopathies

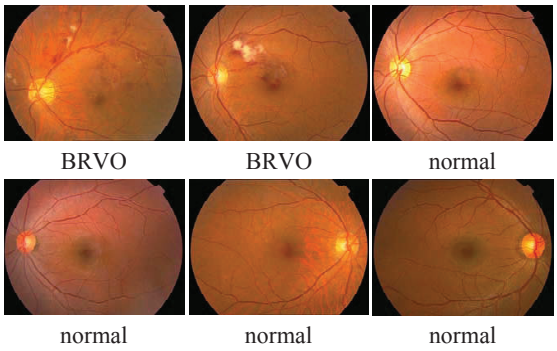

Fig. 5. Images misclassified by HLBP using a 4×4 pooling size and linear kernel. The caption underneath each image is the ground truth.

such as diabetic retinopathy etc., to help the research on the automatic recognition and diagnosis of retinopathy using computer vision methods.

Acknowledgement. This work was supported by a research grant from The Hong Kong Polytechnic University (Project Code: G-YL77). The authors thank Yancheng Third People's Hospital for providing the BRVO and normal color fundus images.

References

1. Ehlers, J.P., Decroos, F.C., Fekrat S.: Intravitreal bevacizumab for macular edema secondary to branch retinal vein occlusion. Retina−The Journal of Retinal and Vitreous Diseases 31, 1856−1862 (2011)
2. Walter, T., Massin, P., Erginay, A., Ordonez, R., Jeulin, C., Klein, J.C.: Automatic detection of microaneurysms in color fundus images. Medical Image Analysis 11, 555−566 (2007)
3. Fleming, A.D., Philip, S., Goatman, K.A., Olson, J.A., Sharp, P.F.: Automated assessment of diabetic retinal image quality based on clarity and field definition. Investigative Ophthalmology and Visual Science 47, 1120−1125 (2006)
4. Tavakoli, M., Shahri, R.P., Pourreza, H., Mehdizadeh, A., Banaee, T., Bahreini Toosi, M.H.: A complementary method for automated detection of microaneurysms in fluorescein angiography fundus images to assess diabetic retinopathy. Pattern Recognition 46, 2740−2753 (2013)
5. Akram, M.U., Khalid, S., Khan, S.A.: Identification and classification of microaneurysms for early detection of diabetic retinopathy. Pattern Recognition 46, 107−116 (2013)
6. Zhang, B., Karray, F., Li, Q., Zhang, L.: Sparse representation classifier for microaneurysm detection and retinal blood vessel extraction. Information Sciences 200, 78−90 (2012)
7. Ojala, T., Pietikainen, M., Maenpaa, T.: Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. TPAMI 24, 971−87 (2002)

8. Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: CVPR, pp. 886−893 (2005)
9. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. IJCV 60, 91−110 (2004)
10. Lazebnik, S., Schmid, C., Ponce, J.: Beyond bags of features: spatial pyramid matching for recognizing natural scene categories. In: CVPR, pp. 2169−2178 (2006)
11. Oliva A., Torralba, A.: Modeling the shape of the scene: A holistic representation of the spatial envelope. IJCV 42, 145−175 (2001)
12. Wu, J., Rehg, J.M.: CENTRIST: A Visual Descriptor for Scene Categorization. TPAMI 33, 1489−1501 (2011)
13. Wang, X., Han, T.X., Yan, S.: An HOG-LBP human detector with partial occlusion handling. In: ICCV, pp. 32−39 (2009)
14. Li, M., Staunton, R.C.: Optimum Gabor filter design and local binary patterns for texture segmentation. Pattern Recognition Letters 29, 664−672 (2008)
15. Heusch, G., Rodriguez, Y., Marcel, S.: Local binary patterns as an image preprocessing for face authentication. In: FG, pp. 9−14 (2006)
16. Moore, S., Bowden, R.: Local binary patterns for multi-view facial expression recognition. Computer Vision and Image Understanding 115, 541−558 (2011)
17. Zhang, B.C., Gao, Y.S., Zhao, S.Q., Liu, J.Z.: Local derivative pattern versus local binary pattern: face recognition with high-order local pattern descriptor. TIP 19, 533−544 (2010)
18. Sorensen, L., Shaker, S., de Bruijne, M.: Quantitative analysis of pulmonary emphysema using local binary patterns. IEEE Transactions on Medical Imaging 29, 559−69 (2010)
19. He, Y., Sang, N., Gao, C.: Pyramid-based multi-structure local binary pattern for texture classification. In: ACCV, pp. 133−144 (2010)
20. Zhu, C., Bichot, C.E., Chen, L.: Multi-scale color local binary patterns for visual object classes recognition. In: ICPR, pp. 3065−3068 (2010)
21. Chen, J., Kellokumpu, V., Zhao, G., Pietikainen, M.: RLBP: Robust local binary pattern. In: BMVC, (2013)
22. Guo, Y., Zhao, G., Zhou, Z., Pietikainen, M.: Video texture synthesis with multiframe LBP-TOP and diffeomorphic growth model. TIP 22, 3879−3891 (2013)
23. Hinton, G.E., Salakhutdinov, R.R.: Reducing the dimensionality of data with neural networks. Science 313, 504−507 (2006)
24. Bengio, Y.: Learning deep architectures for AI. Foundations and Trends in Machine Learning 2, 1−127 (2009)
25. Krizhevsky, A., Sutskever, I., Hinton, G.E.: ImageNet classification with deep convolutional neural networks. In: NIPS, (2012)
26. Farabet, C., Couprie, C., Najman, L., LeCun, Y.: Learning hierarchical features for scene labeling. TPAMI 35, 1915−1929 (2013)
27. Sun, Y., Wang, X., Tang, X.: Deep convolutional network cascade for facial point detection. In: CVPR, pp. 3476−3483 (2013)
28. Karpathy, A., Toderici, G., Shetty, S., Leung, T., Sukthankar, R., Fei-Fei, L.: Large-scale video classification with convolutional neural networks. In: CVPR, pp. 1725−1732 (2014)
29. Wolf, L., Hassner, T., Taigman, Y.: Descriptor based methods in the wild. In: ECCV workshop on Faces in Real-Life Images, (2008)
30. Chang, C.C., Lin, C.J.: LIBSVM: a library for support vector machines. ACM Transactions on Intelligent Systems and Technology 2, 27:1−27:27 (2011)