# Identify Diabetic Retinopathy by Color Textures

**Holly Vo [1], Abhishek Verma [2]**
1) California State University, Fullerton, CA, USA, hhvo@csu.fullerton.edu
2) California State University, Fullerton, CA, USA, averma@csu.fullerton.edu, ecs.fullerton.edu/~averma

***Abstract:*** *Diabetic retinopathy (DR) is a common eye disease that could lead to irreversible vision loss but hard to be noticed by carriers in early stages. Instead of isolating DR signs for DR recognition, this paper examines discriminant texture features obtained by color multi-scale uniform local binary pattern (LBPs) descriptors on five common color spaces and two proposed hybrid color spaces. The extracted features are evaluated by the enhanced Fisher linear discriminant, EFM. Experiments are done on a large dataset of 35,126 training images and 53,576 testing images that have been taken by different devices with high variance in dimensions, quality and luminance.*

***Keywords*:** diabetic retinopathy recognition, texture, local binary pattern, EFM, color channels, color spaces

## 1. SIGNIFICANCE OF DIABETIC RETINOPATHY RECOGNITION

Diabetic retinopathy (DR) is a common eye disease which affects one in three diabetes carriers in America. DR is caused due to high glucose level which damages small blood vessels in retina. In consequence, blood, extra fluid, cholesterol, and other fats leak in the retina and swollen the macula. To replace damaged capillaries, the retina grows new abnormal fragile blood vessels; however, the new vessels are usually accompanied with scar tissue which may wrinkle or detach the retina and distort vision. Although timely treatment can reduce the risk of severe vision loss by over 90%, DR carriers do not notice vision changes until the late stages and the manual process of grading a retina image consumes time and labor.

Fig. 1 shows a fundus image with labeled signs of diabetic retinopathy. Microaneurysms (MAs) are tiny bulges in blood vessels and appear as deep-red dots. Haemorrahages are small spots of blood discharge. Hard exudates are leakage of lipid and protein in the retina. Hard exudates typically emerge as bright, reflective lesions. Hard exudates and microaneuryms around macula might block vision, damage the macula, and leads to permanent vision loss [7, 19]. Depending on the presence of DR signs and their complexity, a fundus image can be marked by an ophthalmologist as normal or one of four DR stages that are grouped in two types of retinopathy: nonproliferative diabetic retinopathy (NPDR) and proliferative diabetic retinopathy (PDR). NPDR is further classified into three stages: mild, moderate, and severe.

This paper studies DR recognition on a real-world retinopathy image dataset provided by EyePACS, a free platform for retinopathy screening, through Kaggle website. The dataset originally consists of 35,126 training images and 53,576 testing images. These images are taken by different models and types of camera under different conditions and stored in various, high resolutions. Each image has been examined on the presence of DR by a clinician to be labeled with a DR stage from 0 to 4, corresponding to no DR, mild, moderate, severe, and proliferative DR as being described in the introduction section. Within the train dataset, there are 74% images of stage 0 (no DR), 7% of stage 1 (mild), 15% of stage 2 (moderate), 2% of stage 3 (severe) and 2% of stage 4 (proliferate DR) approximately. Images in the test datasets are split to 5 stages in the similar ratios as in the train dataset.

Beside the large scale, the challenges of this dataset are its large variance in resolution, intensity, and quality. By examining the train dataset, image heights vary from 289 to 3456 pixels, while their widths fluctuate from 400 to 5184 pixels with the range of ratios between height and width is between 0.66 and 1.00. The average of image intensity spreads from 1 to 192 around the mean of 63. Low intensity images are stored in 8KB while other images can allocate up to 2MB files.

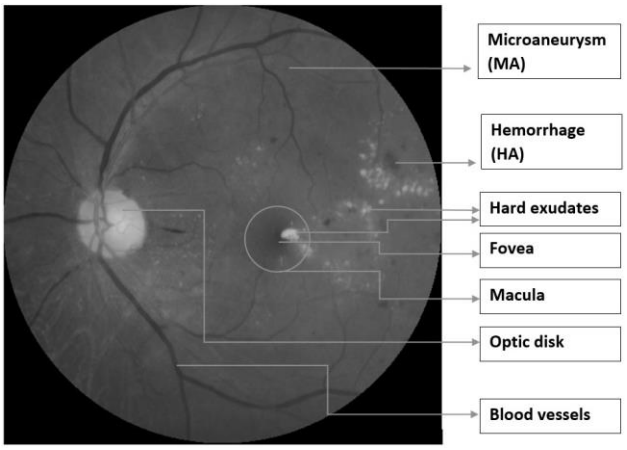


**Fig. 1 - DR signs and main structure in a retina image**

The rest of the paper is organized as follows. Section 2 reviews background of current feature extraction and classification techniques on DR recognition. Section 3 proposes the methodology to extract color texture features and classification techniques to identify DR stages. Experimental results and discussion present in section 4. Section 5 concludes the research with future direction.

## 2. RELATED WORK

With the clinical fact that MAs are the earliest signs of diabetic retinopathy [3], most DR papers focus on extracting clinical features by localizing and segmenting lesions, blood vessels, optic disks, and macula one by one. Basic point operators are applied to balance and enhance local contrast, and linear filters and neighborhood operators such as morphological operators, median filters, and Gaussian filters are convoluted on images in pre-processing as indicated in surveys [7, 12].

Watershed transformation is applied in [22] to overcome over-segmentation caused by thresholds. Other techniques such as active contour models and recursive region-growing technique (RRGT) are used in the domain researches to isolate blood vessels and other interested regions [12].

In addition to segmentation, contrast texture is extracted together with isolated areas of MAs and HAs in [13] to classify DR. In recent years, local binary pattern (LBP) texture is started being used for DR detection on small retina datasets with less than 100 images in [6, 4]. In other domains such as face and scene recognition, texture descriptor such as local binary pattern texture has been proven to contribute significant performance [9].

When sampling on a large set of images which are taken by different devices under various conditions of light and intensity, it is crucial for a robust vision system to adapt a discriminant color space. HSI is applied for Messidor and DB-rect DR datasets in [8] to extract MAs and exudates, and selected by [20] to locate fovea. Green component in RGB is focused to extract blood vessel structure in [1, 13]. All channels of RGB are separately examined in [11] with morphology operations to extract the total area and perimeters of blood vessels, HAs, and MAs. Ram and Jayanthi [18] consider multiple color spaces such as RGB, L*u*v*, HSV and HSI to extract lesion pixel values.

In classification stage, support vector machine (SVM) and artificial neural network (ANN) are two popular techniques in DR recognition problems [8, 1, 13, 11, 12, 6, 4]. Acharya et al. [1] applies SVM on a dataset of 331 retina images to identify five DR stages with an overall accuracy of 86%. Back propagation neural network is applied in [11] to identify four DR stages with an accuracy of 84% on a dataset of 124 images.

## 3. METHODOLOGY

In this work, we examine DR classification performance by fusion of features that are extracted by uniform LBP descriptors at multiple scales. Discriminating power of color spaces and their individual color channels are investigated by applying the method on commonly used color spaces such as RGB, HSI, L*a*b*, rgb, and $I_1I_2I_3$, and two hybrid color spaces a*SI and bSI. The proposed method consists of three major parts: image preprocessing, feature extraction on a color space, and classification via an enhanced Fisher model (EFM) built upon principal component analysis (PCA) and Fisher's linear discriminant method (FLD) (Fig. 2).

As being mentioned in the dataset description section that input images are in a large range of dimension, the main task in image preprocessing is converting any input image of any size to a fixed square image, whose dimension is also the diameter of the circular eye shape. For a given image, a circumscribing rectangle of the eye is determined by scanning pixels along the horizontal and vertical mid-lines of the image. The image is then cropped around the center of the determined circumscribing rectangle to extract a square image whose dimension is the shorter side of the boundary rectangle. Next, the square image is scaled by bicubic interpolation method to a 512x512 image before being clipped around its center by a radius r=256 to guarantee that retina content is captured in a full inscribed circle of the final output square image. In the final step of preprocessing part, training images in DR stages 1-4 are also increased by flipping and rotation to balance training image quantities for all DR stages and avoid over-train on large classes.

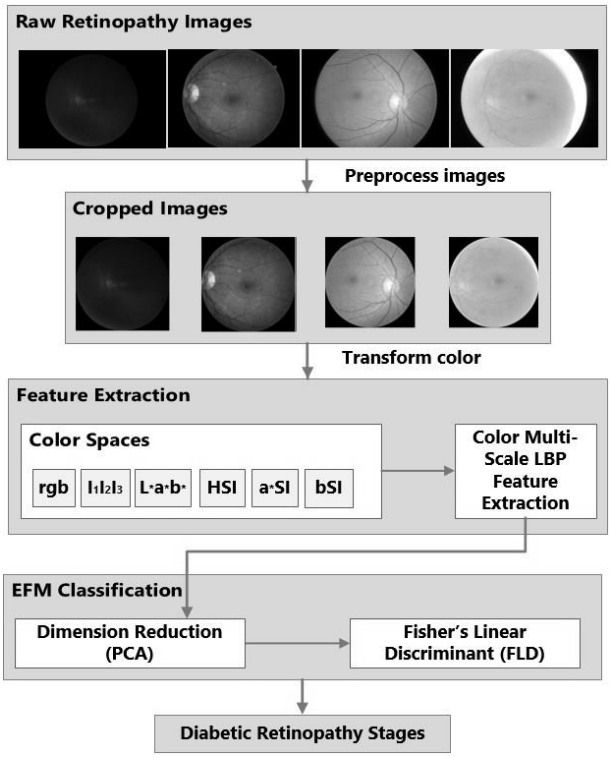


**Fig. 2 - Methodology Overview**

Feature extraction is based on local binary pattern (LBP) descriptors, a texture technique that was proposed and improved by Ojala et al. [16, 17]. LBP has been proven as a robust feature descriptor in biometrics, face detection, and scene recognition [2, 5]. The basic LBP focuses on each 3x3 neighborhood to form an ordered 8-bit LBP code by comparing the surrounding pixels' gray value with its center. The local binary pattern around a pixel P by eight neighbor $P_x$ is encoded as follows:

$$LBP = \sum_{x=1}^{8} s(x)2^{x-1} \qquad (1)$$

$$s(x) = \begin{cases} 1, P_x > P \\ 0, otherwise \end{cases}$$

As the operator focuses on the signed differences of gray values and disregards the value difference, it is invariant to changes in mean luminance. For scale invariance improvement, LBP operator is extended to consider a circularly symmetric neighbor set of P pixels on a circle of radius R surrounding the center pixel, denoted as $LBP_{P,R}$. The top middle neighbor is the most significant bit in LBP code, and other neighbors are ordered clockwise. For each neighbor point whose coordinators are not exactly in the center of pixels, its gray value is estimated by interpolation rather than the nearest pixel's value.

The extension in [16] defines a so-called *uniform pattern*, denoted as $LPB_{P,R}^{riu2}$, which contains at most two spatial transitions in its circular chained binary pattern, "1-0" and "0-1". $LPB_{P,R}^{riu2}$ does not only improve rotation

invariance but also significantly reduces LBP dimension by preserving a single bin for all nonuniform patterns.

In the real diabetic retinopathy recognition problem, images are taken by different devices under different conditions of light and quality. In addition, besides the main structure of a retina in the images, DR signs are fine-grain and their granularity diversifies. Thus, uniform local binary pattern detection should be applied at different scales to capture discriminant features that are invariant to rotation, global intensity, and scales.

In our experiments, each 512x512 retina image is divided into four regions. Texture features are extracted at 4 scales by $LPB_{8,2}^{riu2}$, $LPB_{16,4}^{riu2}$, $LPB_{24,6}^{riu2}$, and $LPB_{32,10}^{riu2}$ descriptors on each 256x256 region at each color channel of a color space. The extracted LBP features are standardized at each scale. The final 1056-dimension feature vector for a color retina image is formed by standardizing the concatenation of features at all scales.

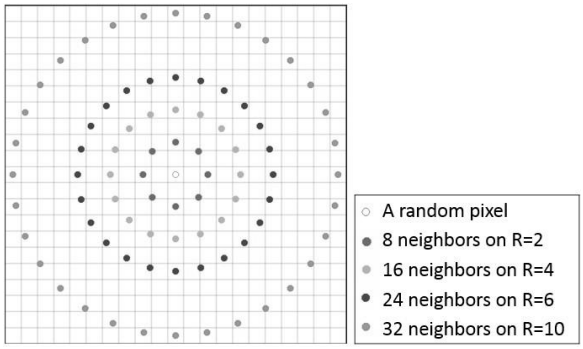


**Fig. 3 - (8,2), (16,4), (24,6), and (32,10) LBP neighborhoods.**

The aforementioned feature extraction process will be applied on different color spaces. Colors are wavelengths of light that are reflected by object surfaces and perceived by human eyes. A color space is a mathematical model to organize colors in a way that relates to the perceived colors. Each color space possesses specific characteristics with different discriminating power and suitable for selective visual tasks. The fundamental RGB color space is built on three primary color components close to red, green, and blue wavelengths and used in reproduction systems [14].

The $I_1I_2I_3$ color space is obtained by the decorrelation of RGB color components through Karhumen Lo´eve [15]. The color space was found in 1980 in the experiments of region segmentation. The color components are ordered by their segmentation effectiveness and formed by the following transformation:

$$\begin{aligned} I_1 &= (R+G+B)/3 \\ I_2 &= (R-B)/2 \\ I_3 &= -(G+B)/2 \end{aligned} \tag{2}$$

The rgb color space is a chromaticity space defined by normalizing RGB components to reduce the sensitivity of luminance [5]. The color space is represented by the proportion of red, green, and blue in the original RGB.

$$\begin{aligned} r &= R/(R+G+B) \\ g &= G/(R+G+B) \\ b &= B/(R+G+B) \end{aligned} \tag{3}$$

HSI is a perceptual color space whose components intuitively approximate the perceived hue, saturation, and intensity in order. HSI components are obtained by the following equations

$$\begin{aligned} H &= \begin{cases} \alpha, if\ (b < g) \\ 2\pi - \alpha, otherwise \end{cases} \\ S &= 1 - \min(r, g, b) \\ I &= (R+G+B)/3, \end{aligned} \tag{4}$$

where and r, g, and b are normalized RGB components obtained by Eq. 3 and

$$\alpha = \cos^{-1}\left( \frac{0.5*[(r-g)+(r-b)]}{[(r-g)^2+(r-b)(g-b)]^{1/2}} \right). \tag{5}$$

L*a*b* color space is directly derived from CIE XYZ, a color system that was built upon imaginary primary colors [X Y Z] to form a device-independent color space with better descriptive properties [14]. L*a*b* mimic the logarithmic response of the human vision system. The L* channel represents luminance in the range from 0 to 100, while a* and b* channels represent chrominance opponents. L*a*b* components can be derived as follows

$$\begin{aligned} L^* &= 116 f(Y/Y_n) - 16 \\ a^* &= 500[(X/X_n) - (Y/Y_n)] \\ b^* &= 200[(Y/Y_n) - (Z/Z_n)], \end{aligned} \tag{6}$$

where $X_n$, $Y_n$, and $Y_n$ are the CIE XYZ component values of the reference white point and

$$f(t) = \begin{cases} t^{1/3} & t > \left(\frac{6}{29}\right)^3 \\ \frac{1}{3}\left(\frac{6}{29}\right)^2 t + \frac{4}{29} & otherwise \end{cases}. \tag{7}$$

In addition to the existing color spaces, experiments are similarly conducted on two hybrid colors that are derived from HSI: a*SI and bSI. The only difference between the hybrid colors and the original color space HSI is the first channel. The first channel of a*SI is the a* channel of L*a*b*, and the first channel of bSI is the normalized blue channel of rgb.

Classification of the extracted features on each color space is based on the Enhanced Fisher model, EFM, a classification model that is built upon PCA and FLD. In short, PCA is a common technique to linearly transform data to a lower dimensionality space and reduce data noise while trying to preserve the most variance of data [23]. Although PCA is a popular technique in pattern recognition, it is not optimized for class separability. Instead, the alternative technique, FLD, has been proposed to model the difference between classes of data [10, 21]. FLD is a popular discriminant criterion that defines a projection to reduce within-class scatter and enlarge the between-class scatter for an L-class problem.

The FLD method encounters overfitting drawback when there are insufficient sample data for generalization. EFM overcomes this issue by combining PCA and FLD in the proper balance of the selected eigen features for an adequate representation of raw data and the requirement that the eigenvalues of the within-class scatter matrix in the reduced PCA are sufficient large for generalization

[9]. In EFM classifier, discriminant features, *Z*, are obtained by projecting the PCA reduced feature on the optimal FLD projection matrix. Each discriminant feature is assigned to the nearest class by measuring its cosine distances to all class centers in our experiments.

Besides the proven proven effectiveness of EFM in the domains of face and scene recognition, the classifier is selected for DR recognition experiments on multiple colors and color channels because of its simplicity in terms of computation and parameters. Since the number of classes in DR recognition is small, the maximum FLD feature dimension, $m = L-1 = 4$, is chosen for the best generalization of within and between class relationship. The balanced PCA criterion for the selected FLD feature dimension is determined by 5-fold cross-validation on the train dataset.

## 4. EXPERIMENT RESULTS

Experiments are initially conducted on gray scale and five different color spaces. The performance is measured by the correct classification rate on the test dataset. Experiment results obviously show that all color spaces surpass gray scale performance. Among examined color spaces, intuitive color space L*a*b* and HSI outperforms RGB, rgb, and $I_1I_2I_3$. The multiscale LBP descriptor on HSI, denoted as HIS-LBPs, achieves the best performance, 71.45%, among the five descriptors, and it exceeds gray scale by 5.47%. The results support the idea of selecting HSI as the color space to extract MAs and exudates in [8]. It is apparent to understand that the rgb color space performs worst (66.30%) due to the complete absence of luminance. Performance on the proposed hybrid color spaces, a*SI and bSI, is 71.49%, that is slightly better than HSI performance as summarized in Fig. 4. Additionally, Fig. 5 shows sample RGB images that are not correctly classified by RGB-LBPs and Lab-LBPs descriptor but correctly classified by HSI-LBPs, aSI-LBPs, and bSI-LBPs descriptors. One can visually realize that these example images are very different in terms of light and color.

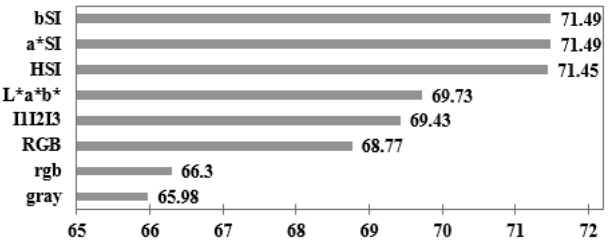


**Fig. 4 - EFM classification performance on color spaces.**

**Table 1. LBP-EFM performance on color channels and color spaces**

| | Color Space | Channel 1 | Channel 2 | Channel 3 |
|---|---|---|---|---|
| HSI | 71.45 | 65.92 | 68.79 | 65.93 |
| L*a*b* | 69.73 | 66.03 | 63.69 | 61.18 |
| I123 | 69.43 | 65.93 | 62.21 | 61.99 |
| RGB | 68.77 | 63.98 | 65.64 | 61.73 |
| rgb | 66.30 | 61.59 | 61.05 | 60.50 |
| gray | 65.98 | 65.98 | | |

For more insights on color LBPs descriptors, experiments are further carried on individual channels of each aforementioned color (Table 1). In rgb color space, b channel performs worst. The best performance, 61.59%, on the r channel is only 1% above the worst channel but 5% below the performance of the color space. It means that different discriminant features could be arranged on every channel of rgb, thus, there is no significantly strong channel and rgb significantly outperforms its individual channels.

In the other four color spaces, improvement of each color performance over its best channel is not as significant as in rgb. The performance improvement on these color spaces is from 2.7% to 3.8%. There is a dominant channel that outperforms the worst channel from 2.9% to 3.9% in each color space. In RGB, green channel achieves the best performance and this result could explain why it is the selected channel to extract blood vessel structure in [1, 13]. The first channel of $I_1I_2I_3$ performs better than other two channels because it the channel holding most chrominance and luminance information. In L*a*b* color space, the luminance channel L* outperforms chrominance channels.

In HSI and proposed HSI-like hybrid color spaces, the most discriminant features are identified on the saturation channel. Although intensity and hue perform at the same accuracy rate, a* and b channels performance is at least 2% less than hue. It could be implied that there are more discriminant features that are not found in saturation channel could be identified in channel I than in channel H. It should be noticed that channel I of HSI is also the strongest channel of $I_1I_2I_3$.

## 5. CONCLUSION

We have examined in this paper different color multi-scale local binary pattern texture descriptors, LBPs, for diabetic retinopathy recognition on a large dataset. The enhanced Fisher linear discriminant is applied to identify the promising color spaces and color channel candidates to obtain the most discriminant LBPs features for DR recognition. Empirical results show that HSI-LBPs descriptor and its variants, a*SI-LBPs and bSI-LBPs descriptors outperform other color LBPs and gray LBPs descriptors.

For the future plan, the candidate color LBPs descriptors can be combined with features from other region or gradient detectors to improve DR performance. Other classification and ensemble techniques will be explored to achieve better accuracy.

## 6. REFERENCES

[1] U. Acharya, C. Lim, E. Ng, C. Chee, and T. Tamura. Computer-based detection of diabetes retinopathy stages using digital fundus images. *Proceedings of the Institution of Mechanical Engineers, Part H: Journal of Engineering in Medicine*, 223(5):545–553, 2009.

[2] T. Ahonen, A. Hadid, and M. Pietikainen. Face description with local binary patterns: Application to face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(12):2037–2041, 2006.

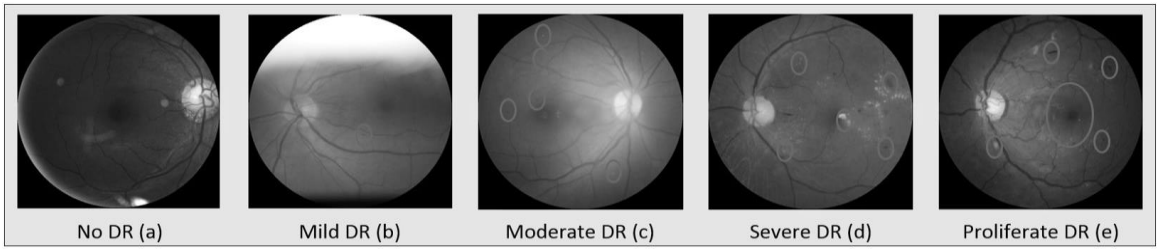


**Fig. 5 - Example DR images correctly classified by using HSI-LBPs, aSI-LBPs, bSI-LBPs descriptors but not by RGB-LBPs or Lab-LBP descriptors.**


[3] B. Antal and A. Hajdu. An ensemble-based system for microaneurysm detection and diabetic retinopathy grading. *IEEE Transactions on Biomedical Engineering*, 59(6):1720–1726, 2012.

[4] M. N. Ashraf, Z. Habib, and M. Hussain. Texture feature analysis of digital fundus images for early detection of diabetic retinopathy. In *Computer Graphics, Imaging and Visualization (CGIV), 2014 11th International Conference on*, pages 57–62. IEEE, 2014.

[5] S. Banerji, A. Verma, and C. Liu. Novel color lbp descriptors for scene and image texture classification. In *15th International Conference on Image Processing, Computer Vision, and Pattern Recognition, Las Vegas, Nevada*, pages 537–543. Citeseer, 2011.

[6] J. de la Calleja, L. Tecuapetla, M. A. Medina, E. Ba´rcenas, and A. B. U. Na´jera. Lbp and machine learning for diabetic retinopathy detection. In *Intelligent Data Engineering and Automated Learning–IDEAL 2014*, pages 110–117. Springer, 2014.

[7] O. Faust, R. Acharya, E. Y. K. Ng, K. H. Ng, and J. S. Suri. Algorithms for the automated detection of diabetic retinopathy using digital fundus images: a review. *Journal of medical systems*, 36(1):145–157, 2012.

[8] J. Lachure, A. Deorankar, S. Lachure, S. Gupta, and R. Jadhav. Diabetic retinopathy using morphological operations and machine learning. In *Advance Computing Conference (IACC), 2015 IEEE International*, pages 617–622. IEEE, 2015.

[9] C. Liu and H. Wechsler. Robust coding schemes for indexing and retrieval from large face databases. *IEEE Transactions on Image Processing*, 9(1):132–137, 2000.

[10] C. Liu and H. Wechsler. Gabor feature based classification using the enhanced fisher linear discriminant model for face recognition. *IEEE Transactions on Image Processing*, 11(4):467–476, 2002.

[11] W. L.Yun, U. R. Acharya, Y. V. Venkatesh, C. Chee, L. C. Min, and E. Y. K. Ng. Identification of different stages of diabetic retinopathy using retinal optical images. *Information Sciences*, 178(1):106–121, 2008.

[12] M. R. K. Mookiah, U. R. Acharya, C. K. Chua, C. M. Lim, E. Y. K. Ng, and A. Laude. Computer-aided diagnosis of diabetic retinopathy: A review. *Computers in Biology and Medicine*, 43(12):2136–2155, 2013.

[13] J. Nayak, P. S. Bhat, R. Acharya, C. M. Lim, and M. Kagathi. Automated identification of diabetic retinopathy stages using digital fundus images. *Journal of medical systems*, 32(2):107–115, 2008.

[14] M. S. Nixon and A. S. Aguado. *Feature extraction & image processing for computer vision*. Academic Press, ISBN: 9780123965493, 2012.

[15] Y. Ohta, K. Takeo, and S. Toshiyuki. Color information for region segmentation. *Computer Graphics and Image Processing*, 13(3):222–241, 1980.

[16] T. Ojala, M. Pietikainen, and T. Maenpaa. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(7):971–987, 2002.

[17] T. Ojala, M. PietikAd'inen, and D. Harwood. A comparative study of texture measures with classification based on featured distributions. *Pattern Recognition*, 29(1):51–59, 1996.

[18] K. Ram and J. Sivaswamy. Multi-space clustering for segmentation of exudates in retinal color photographs. In *Proc. Annu. Int. Conf. IEEE Eng. Med. Biol. Soc*, pages 1437–1440, 2009.

[19] S. Ravishankar, A. Jain, and A. Mittal. Automated feature extraction for early detection of diabetic retinopathy in fundus images. In *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, pages 210–217. IEEE, 2009.

[20] C. Sinthanayothin, J. F. Boyce, H. L. Cook, and T. H. Williamson. Automated localisation of the optic disc, fovea, and retinal blood vessels from digital colour fundus images. *British Journal of Ophthalmology*, 83(8):902–910, 1999.

[21] D. L. Swets and J. J. Weng. Using discriminant eigenfeatures for image retrieval. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(8):831–836, 1996.

[22] T. Walter and J.-C. Klein. Segmentation of color fundus images of the human retina: Detection of the optic disc and the vascular tree using morphological techniques. In *Medical Data Analysis*, pages 282–287. Springer, 2001.

[23] I. H. Witten and E. Frank. *Data Mining: Practical machine learning tools and techniques*. Morgan Kaufmann, ISBN: 9788131200506, 2005.