

Retrieving clinically relevant diabetic retinopathy images using a multi-class multiple-instance framework

Parag S. Chandakkar¹, Ragav Venkatesan¹, Baoxin Li¹, Helen K. Li^{2,3} MD

¹Computer Science & Engineering, Arizona State University, Tempe, Arizona, USA;

²Weill Cornell Medical College/The Methodist Hospital;

³The University of Texas Health Science Center Houston, and Thomas Jefferson University

ABSTRACT

Diabetic retinopathy (DR) is a vision-threatening complication that arises due to prolonged presence of diabetes. When detected and diagnosed at early stages, the effect of DR on vision can be greatly reduced. Content-based image retrieval can be employed to provide a clinician with instant references to archival and standardized images that are clinically relevant to the image under diagnosis. This is an innovative way of utilizing the vast expert knowledge hidden in archives of previously diagnosed fundus camera images that helps an ophthalmologist in improving the performance of diagnosis. In this paper, with a focus on two significant DR clinical findings, namely, microaneurysm and neovascularization that are representative symptoms of non-proliferate and proliferate diabetic retinopathy, the authors propose a multi-class multiple-instance image retrieval framework that makes use of a modified color correlogram and statistics of steerable Gaussian Filter responses for retrieving clinically relevant images from a database. Experiments are performed using fundus camera images and the results compared with other prior art methods demonstrate the improved performance of the proposed approach.

Keywords: Diabetic retinopathy, image retrieval, multiple-instance learning, Color Correlogram, Steerable Gaussian Filters, Fast Radial Symmetric Transform

1. INTRODUCTION

Diabetic retinopathy (DR) may lead to blindness in people suffering from diabetes [1]. Diabetes is a disease associated with the blood vessels. Glycosylation reaction occurs between sugar and the proteins in the vessel walls, owing to a large amount of glucose coursing through the circulatory system. Organs like the eyes and the kidney have microvasculature that is more susceptible to damage because of glycosylation. Studies have shown that timely DR diagnosis and treatment can significantly reduce the risk of severe vision loss [2]. By improving the quality of DR severity assessment, the quality of DR diagnosis can be significantly improved. Two important clinical findings of DR are microaneurysm (MA) and neovascularization (NV), which are respectively characteristics of the non-proliferate and proliferate DR stages. Hence they are of particular significance, and are the focus of this study.

Figure 1 illustrates a set of typical fundus images diagnosed with varying degrees of non-proliferate diabetic retinopathy (NPDR). While yellow, waxy exudates can be observed in the second image, cotton-wool patches can be observed in the fourth image, very little difference can be observed between the first image and the normal images shown in Figure 2. This elucidates the arduousness of the DR conundrum. Similarly, images in Figure 3 illustrate varying degrees of proliferate diabetic retinopathy (PDR).

In practice, different eye care providers have different abilities in assessing DR, and their accuracy in evaluating the stages of the disease depends upon training and experience. Also, the process was laborious and prone to reviewer fatigue. As a result, many researchers have started exploring automated detection and evaluation of diabetic retinal lesions [3-7]. Unfortunately, to date there is no automated system that can perform DR lesion detection with the accuracy that is comparable to a human expert. Further, a task that is more difficult for ophthalmologists – fast and accurate severity grading of DR – has not been adequately addressed by existing systems, and an ophthalmologist's painstaking visual examination of a digital retinal image and mental or physical comparison with standard images are still the ultimate method for identifying and assessing DR. The proposed research aims at filling this gap in both technical capability and clinical practice by developing a computer-based system with the innovative idea of content-based retrieval and classification of DR images with clinical relevance.

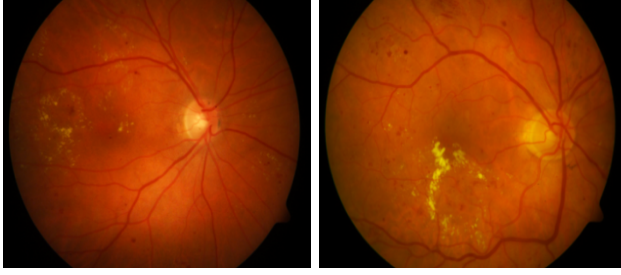


Figure 1. Fundus camera images of an eye with NPDR.

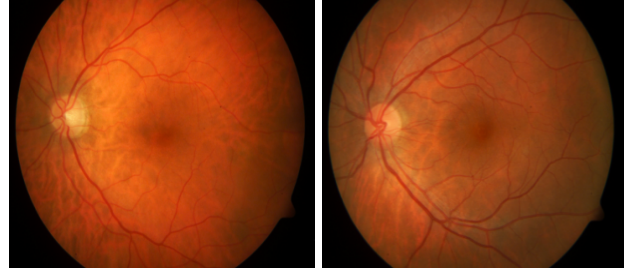


Figure 2. Fundus camera images of a normal eye.

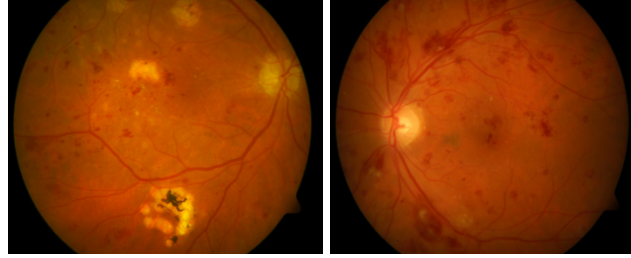


Figure 3. Fundus camera images of an eye with PDR.

The proposed research develops a content-based DR image retrieval system that recalls clinically relevant images to a novel query image to be diagnosed. Such a retrieval system can provide an ophthalmologist with both standard sample images and previously-diagnosed images of similar lesions and severity, and thus effectively enables knowledge sharing/reutilization among experts. A knowledge-sharing scheme like the one discussed above can be useful especially for inexperienced doctors and medical students in training, as they can gain instant access to expert knowledge through the automatically retrieved reference images. It may also be a key technique for mobile DR screening clinics that are not staffed by experienced doctors.

In examples illustrated in Figures 1 to 3, one can see that the distinguishing differences among images of different classes are in general localized to small regions of the images. Meanwhile, in existing practice and public datasets, the diagnosis (or the “class label” assigned to an image) is kept only on a per-image basis, without noting where the symptoms occur in the images. Considering these factors, we propose a multiple instance framework for content-based DR image retrieval. This approach, driven on a feature space of modified color-correlogram (CC) and statistics of SGF response features, provides clinically relevant DR image retrieval. Such a system leads to a novel way of tackling this problem of DR diagnosis and grading.

The remainder of this article is organized as follows: Section 2 cites some related work in the field of CBIR in DR diagnosis. Section 3 describes the proposed algorithm. Section 4 describes the experimental setup. Section 5 provides the results of the new algorithm and Section 6 provides concluding remarks.

2. RELATED WORK

In the research community, several research groups have developed content-based image retrieval (CBIR) systems for use in fields other than ophthalmology (e.g., [8-13]). These CBIR systems typically extract image content and provide query methods for direct image matching using low-level features. For example, Cai et al. [3] developed a prototype retrieving Positron Emission Tomography images based on physiological kinetic features. Chu et al. [4] developed a semantic model for content-based image retrieval for capturing hierarchical, spatial, temporal, and evolutionary semantics of neural image databases. A system by Kelly et al [5] associates each medical image with a signature for capturing textures and histograms of pathologies. This system retrieves images using query-by-example techniques. Fast query using nearest neighbor search is addressed by Korn et al [6] to retrieve medical tumors that are similar to a given pattern.

The system proposed by Nah and Sheu [7] used operational semantics to improve content-based retrieval of brain neuroscience images. In the ASSERT system [8], Shyu et al. designed a suite of computer vision algorithms to extract visual abnormalities and used multidimensional hashing approach to index pathologies of HRCT (high resolution computed tomography) lung images. These pioneering CBIR works are promising in that the papers all report benefits to a clinician or improved performance of a basic system with the respective techniques. However, most of these approaches are domain-specific (and most are for CT, X-ray, or MRI images) and thus not directly adoptable to color fundus image analysis in DR diagnosis.

Previous CBIR-related investigations in ophthalmology include the Structured Analysis of the Retina (STARE) project aimed at automatic diagnosis and comparison of images, including annotating image contents and searching for images similar in content [9]. The similarity metric in this system is based on some basic image features (called primitives) for a particular class of images [10]. While having made significant progresses since the inception of the project, the lack of a unified and systematic solution or system that has been widely accepted in ophthalmology shows that the central problems have not been solved. In particular, when it comes to CBIR for retinal images, the retrieved images given in [10] appear to be largely similar in appearance but not similar in terms of clinical relevance, rendering the results of very limited clinical use.

Another state-of-the-art image representation and large-scale retrieval system works on using scale invariant feature transform (SIFT) features coupled with a bag-of-words (BoW) approach [11]. Similar methods are also developed using the histogram of oriented gradient features [12]. These BoW features when used along with Support vector machines (SVM) are powerful classifiers and are also useful in CBIR. The main constraint of bag of words in image processing is that it cannot capture the spatial location information of feature descriptors that are important in terms of features like SIFT.

Another well-studied feature for CBIR pertaining to general images is the color correlograms [13, 14]. They are considered to be better than histogram features. Their advantage lies in the fact that they have a very low dimensionality and yet provide more assistance to a classifier than the histogram features. Though correlograms are similar to histograms, they incorporate the spatial correlation of image pixels. They also describe the global correlation of local spatial correlation of colors. This is a strong representation of textures with considerably small dimensionality.

Other state-of-the-art image retrieval systems use similarity metrics like Gabor features [15] and Histogram of neighborhood mean moments (HNM) features [16]. Both these features (HNM in particular) are preferred in medical image retrieval but might not be suited for DR images because of the special spectrum that DR images have. Figure 4 shows the histograms of typical DR images. It can be observed that apart from the localized lesions, the red channel is almost always saturated. It can also be observed from Figure 4 that images of varying classes have similar histogram, making it difficult for histogram features. Thereby effective retrieval of DR images intuitively leads to the need for spectrally-tuned color features.

3. PROPOSED APPROACH

3.1 The Features

The color correlogram of an image is a tabular representation of indexed color pairs. The spatial correlation of pixels in a quantized image is described by CC features. When the distribution of only those pixels which lie in the same bin is analyzed, then the features are called color auto-correlogram features (AutoCC). AutoCC either by itself or combined with other features are usually used in many retrieval tasks. Li in [14] proposed one such method which uses AutoCC in combination with other texture features to perform retrieval for natural images, where a quantization scheme modeled after human vision and considering the color spectrum of natural images was proposed. In this quantization scheme, the dimensionality of the AutoCC features is restricted to 44. It can however be seen from Figure 4 that the spectrum of DR images drastically differs from that of natural images considered in [14]. Figure 4 also shows that the spectrum of a typical DR image is saturated in the red channel. It can be observed that DR images have a special spectrum that calls for attention. For CC features to give best results with respect to the DR problem, it is necessary to have uniformly shaded bins over the DR image spectrum. Therefore, we use a modified color correlogram (CC) features proposed in [embc]. These features are combined with statistics of SGF response features [17] along with fast radial symmetric transform (FRST) to form the final feature for our retrieval task. These are elaborated below.

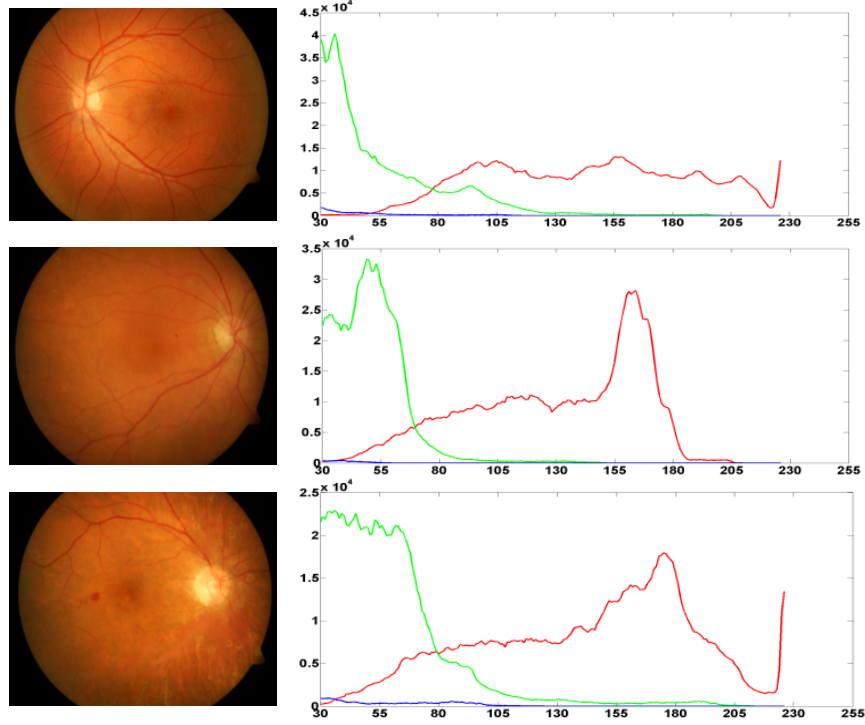


Figure 4. Histograms of typical fundus images. The top image is diagnosed as a NPDR image, the middle image as a normal image and the bottom image PDR.

We utilize SGF features since they can be designed to adaptively filter out the structures in a particular orientation as well as to detect contours. Also, the response of the SGF filter can be used to categorize and identify lines, edges and contours, all being potentially distinguishing features in our application. Though the details behind SGF are beyond the scope of this article, we illustrate in Figure 5 the basis functions of such filters. The 8 basis functions are separated by 45° . Using these bases as kernels the directional derivatives in each direction of the image is calculated. Figure 6 shows the filter response of a typical DR image to basis oriented at 225° . The input image here is an NPDR image. The filter response highlights the nerves and the NPDR spots. We then collect the standard deviation, skewness and kurtosis of the response image as features. This is a three dimensional vector for each angle. For all the 8 bases shown in Figure 5, we have a 24 dimensional feature vector for the given image.

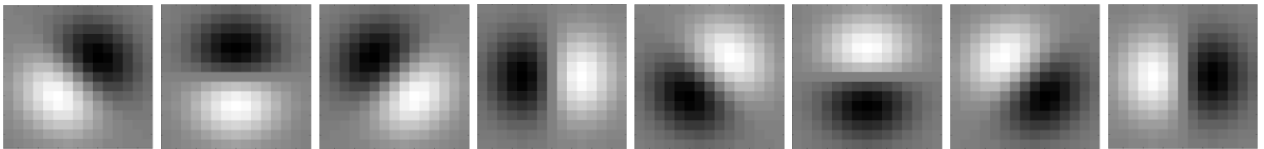


Figure 5. Basis of steerable Gaussians used.

In Figure 6, it is observed that not only are the MA points appreciated but also the nerves. These produce a lot of noise on the feature space as they alter the moments of the filter responses. To avoid these, there is a need to specifically select windows of operation around potential points of interest. To identify such points of interests, we propose to use FRST (Loy & Zelinsky, 2003). FRST makes use of local radial symmetry to highlight the interest points within an image. It calculates the contribution of radial symmetry of the gradients that are at a distance d away from the point under consideration. This implies that the transform estimates the contribution of each pixel to the symmetry around a pixel and not the contribution of the local neighborhood to a central pixel. The transform can be implemented for a set of

radii. In this context to negate the presence of the nerves and other spurious components which might lead to noisy feature space, we use a set of radii – $\{1, 3, 5\}$. Such an implementation is shown in Figure 7.

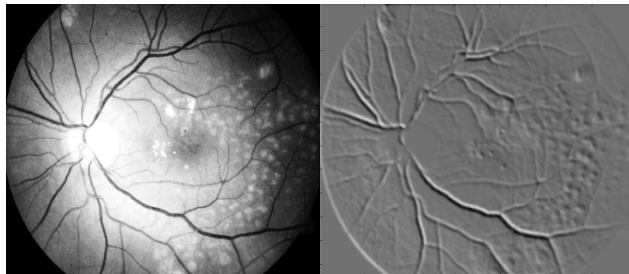


Figure 6. Input MA image (left) and filter response (right) of a typical DR image.

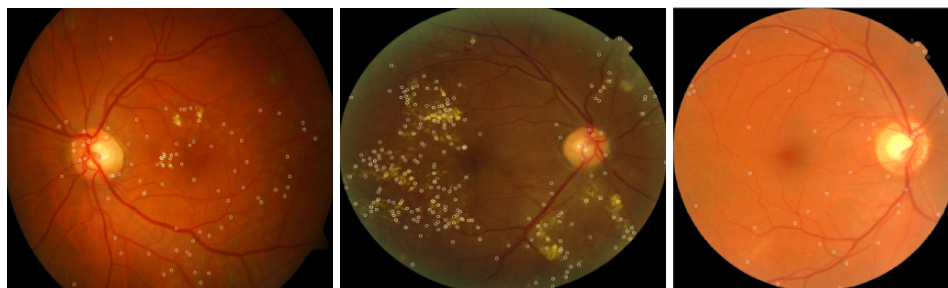


Figure 7. Interest points detection using fast radial transform. The top rows are affected eyes, while the bottom is a normal eye. (Zoom in for better viewing).

Though the majority of detected points are usually lesions of interest, there are false detections too. This is because the radial transform works on local maxima detection. Thus the radial transform feature cannot be used by itself to classify images. Once these points of interests are detected, a neighborhood of 15×15 is chosen on which the prior mentioned statistics of SGF response are extracted. Assuming there are γ number of points on a given image, γ feature vectors will be extracted for the image with each feature vector being 24 dimensional. To provide consistency in the dimensionality of the feature space, the feature space is reduced to a single feature vector of 24 dimensions by taking the first moment over each dimension individually. In the multiple instance framework that is discussed in the following section, this action is justified as each instance is a subset of the image consisting of a very few number of interest points almost always belonging to the same lesion of interest. If in such a situation where no interest points are detected, the feature vector is all zeros making it easier for the proposed McMIL algorithm to not consider them.

3.2 The McMIL Framework for Retrieval

The proposed features are extracted to support a multi-class multiple instance learning retrieval system. As was already discussed, MIL approach is apt for this problem, since only small localized regions of an image affects the class label of the image. In the MIL formulation, a subset of instances in a bag determines the class of the entire bag. This can be seen in Figure 8(a). The instance marked in red is the only instance in the image that characterizes the image as an affected image, while all other instances are normal. Let's call that instance the characterizing instance. The problem now is to identify those characterizing instance(s) in a prospective retrieved image that are clinically relevant to the characterizing instance(s) in the query image and to retrieve the top k images that are closer only in those instances.

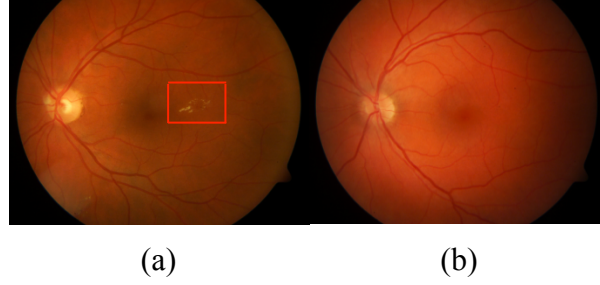


Figure 8. (a) NPDR image. Instance marked in red is the contributing instance (b) Normal image.

In the MIL formulation, a bag has multiple instances and thus corresponds to multiple points in the instance space. In accordance to this formulation the image is divided into 64 blocks on a 8×8 grid. Unwanted instances that contain black background regions are dropped by thresholding. The features discussed in Section 3.1 are extracted for each instance. In case of the statistics of SGF response features, the fast symmetric radial transform is performed on a global level. This identifies regions of interest on a global level, while the statistics of SGF response features are extracted for the detected points on an instance level. This gives us a 280 (256-CC + 24-statistics of SGF features) dimensional feature vector for each instance and the number of instances per image (thereby bag) might vary.

A modified approach originally based on Citation- k NN [19] which falls into the category of lazy learning is proposed to conquer this problem. Citation- k NN uses a modified Hausdorff distance. By definition, Hausdorff distance provides a measure of dissimilarity between two subsets of metric space. Hausdorff distance is very sensitive to outliers and to overcome this problem, the authors of citation- k NN made use of a minimal Hausdorff distance defined by

$$H(A, B) = \min_{b \in B} \min_{a \in A} \|a - b\| \quad (1)$$

where $H(A, B)$ is the Hausdorff distance between any two bags A and B while a and b are instances in bags A and B respectively. $H(A, B)$ gives the least distance between two bags. In the DR problem the distance between a NPDR and a Normal image will be almost equal to the distance between two Normal images, for example, in Figure 11(a) only one instance of an NPDR image (marked by red border) differs from the instances in Normal image shown in Figure 11(b). Thus the use of the minimal Hausdorff distance does not provide satisfactory results. This calls for a new way of measuring the distance between two bags.

Computation of Distance between two bags:

Suppose there are two bags A and B with m and n instances each. The minimum distance between the i^{th} instance of A and all instances of B is given by

$$D_{(i)}(A, B) = \min_{b \in B} d(a_i, b_j) \quad \forall j \in \{1, 2, \dots, n\} \quad (2)$$

Here, $d(\cdot)$ represents correlation distance between two instances. D is a m dimensional vector containing the minimum distances between every instance of A and all instances of B .

Let Q be the bag of instances for a query image for which clinically-relevant images are to be retrieved. The minimum distances between every instance in Q and instances of all the images in the database are calculated as described in Equation 2. This, when sorted individually for every instance, gives a similarity list (SL) for every instance in Q . A bag-level aggregated similarity rank (ASR) is calculated for each bag (image) in the database as the average of ranks in the SL of instances pertaining to that particular bag. This can be termed as the m -Ranking procedure and the rank list thus formed the m -Rank. m -Ranking procedure is explained in the following algorithm.

Algorithm for calculation of m -Rank:

- Step 1. Calculate D between Q (assume Q has m instances) and every image in the database and form the vector $D' = [D_1, D_2, \dots, D_T]$. Here, D' is a $m \times T$ matrix, where $T + 1$ is total number of images in database.
- Step 2. D' is now sorted along columns and the sorted indices gives SL where $SL(u, v) = w$ represents the v^{th} best match for the u^{th} instance of Q and the best match is an instance belonging to image w .
- Step 3. For a new bag N in the database, collect the column indices of all occurrences of N in SL and arrange them in an array col .

Step 4. Calculate ASR for N as $ASR(N) \triangleq avg(col)$.
Step 5. Repeat Steps 3 and 4 for all bags in the database to arrive at the complete ASR .
Step 6. Sort ASR , whose indices will give m -Rank list.

The minimum distance between two bags need not necessarily give optimum retrieval for NPDR images as only few instances contribute towards the label. To avoid this problem, m -Rank is taken one step further by incorporating $citer$ -Rank to it. This is accomplished by obtaining a final $meanRank$ which is the average of m -Rank and $citer$ -Rank of Q .

The top k retrieved images are based on the $meanRank$ list thus created. It can be seen that k doesn't affect the performance of the algorithm as the calculation of $meanRank$ is independent of k .

4. EXPERIMENTAL SETUP

The dataset used to evaluate the proposed system consists of 425 images, put together from well-known databases including DIARETDB0 [20], DIARETDB1 [21], STARE [22] and Messidor (<http://messidor.crihan.fr>). In total, the database contains 160 normal images, 181 MA images and 84 NV images. The labeling in the first three datasets are already well-defined and three fold there by it is easily adapted to fit this combined dataset. The Messidor dataset provides a five-fold label including Normal, mild MA, severe MA, mild NV and severe NV. Out of these we handpicked images from normal, severe MA and severe NV for our normal, NPDR and PDR labeled data respectively.

The proposed system is evaluated against some prior art computer vision retrieval systems using Gabor features [15], semantic of neighborhood color moment histogram features (HNM) [16] along with the original AutoCC features that are proposed in [14]. All of these methods are based on a distance based k -nearest neighbor search systems. The distances used are those that are recommended in the specific articles. These systems are popular and prior art in medical image retrieval literature. All the 425 images in the database are individually queried and the top ($k=$) 5 images are retrieved using the proposed approach. Evaluation metrics are adopted from [23]. If one of the images retrieved has the same label as the queried image then it is a hit.

5. RESULTS

Table 1 presents the results of $\geq k$ hit-rate (HR), which is the percentage of images for which at least k relevant images are retrieved. Table 1 also shows mean accuracy for the retrieval of top 5 images. Another metric used is success at k^{th} rank ($S@k$) which is the probability of finding a relevant image among the top k images. $S@k$ findings are presented in table 2. Table 3 discusses accuracy of images retrieved at each rank.

Table 1. Mean accuracies and $\geq k$ HR in percentages

	Mean Accuracy	≥ 2 HR	≥ 3 HR	≥ 4 HR	≥ 5 HR
AutoCC	60.85	79.06	64.47	42.35	23.29
Gabor	68.61	83.76	75.76	59.52	31.76
HNM	68.04	84.47	74.58	55.76	32.00
Proposed	75.48	88.70	82.58	68.00	43.29

Table 2. Success at k th rank in percentages

	$S@2$	$S@3$	$S@4$	$S@5$
AutoCC	63.88	62.19	61.88	60.85
Gabor	68.47	68.31	68.11	68.61
HNM	71.05	70.27	68.82	68.04
Proposed	77.17	75.60	75.76	75.48

Table 3. Mean Accuracy at kth rank in percentages

	Rank 1	Rank 2	Rank 3	Rank 4	Rank 5
AutoCC	64.24	63.88	62.19	61.88	60.85
Gabor	69.41	67.52	68.00	67.52	70.58
HNM	73.64	68.47	68.70	64.94	64.94
Proposed	77.41	76.94	72.47	76.23	74.35

The reproducibility of results were also tested by repeating the experiments in cross-validation for about 20 times. Figure 10 shows the consistency in hit rates. It can be observed that the algorithm is invariant to cross-validation and shows consistent retrieval performance. This implies that the algorithm can be expected to be consistent with change in dataset so that if in future, more data is to be added or removed on an online basis, the algorithm's performance would remain the unaffected. It can be clearly observed from the mean accuracies reported in table 1 that, the proposed approach outperforms all the competing methods considered. From Table 1, it can also be observed that our $\geq k$ HR is much higher than the other methods. This implies that the proposed method retrieves more relevant images in the top five retrieved images than the other methods. This is of significant assistance to an ophthalmologist in clinical diagnosis, where it is helpful to retrieve maximum clinically relevant images.

HNM fails to retrieve images with clinically similar lesions in particular to that of hemorrhages as was retrieved by the proposed approach. Therefore, this reduces the clinical relevance of the method and thus this method may prove of little assistance to the ophthalmologist. Figure 9 shows 5 retrieved images for a typical PDR image using proposed approach. It can be observed that not only the retrieved images are similar in appearance to the given image; they also all contain PDRs of similar severity, hence achieving the desired clinical relevance.

From Table 2, it can be observed that the successes at each rank are not only higher than the other methods, but also exhibit consistency. This implies that the probability of finding clinical relevance in a retrieved image at any rank is fairly uniform and that the ratio of relevant retrieved images to irrelevant retrieved images is high. This makes the retrieved images more trustworthy for an ophthalmologist as he need not worry about priority of the retrieved images.

Another observation that was drawn from the retrieved images is that, though for NPDR images, PDR images are retrieved occasionally and vice-versa, this is still clinically relevant in the sense that most PDR images contain symptom of NPDR and those NPDR images which are about to turn PDR contain symptoms of the same. This provides more information to an ophthalmologist though the labeling was incorrect. This effectively increases the satisfaction in retrieval. This is a clear advantage from the MIL adaptation.

A mean confusion matrix is created for the proposed approach and shown in Table 4. The following example explains how the confusion matrix is constructed. Suppose 100 normal images are queried and therefore there will be 500 retrieved images. The 1st row of the confusion matrix indicates that 71.5% of the 500 images were normal, 15.87% images were NPDR and 12.62% images were PDR. Similarly the second and third row can be explained. Thus the confusion matrix shows the distribution of top 5 retrieved images for queried images of each category. It can be seen from the confusion matrix that the errors increase while retrieving for normal and PDR images. Some errors in PDR are owing to the fact that certain PDR images in the database are treated for MA symptoms. The retrieval accuracy for NPDR in particular is higher and this is particularly advantageous for a clinician as PDR and normal are usually extreme cases that can be easily diagnosed.

Table 4. Confusion matrix for the proposed approach

	Normal	NPDR	PDR
Normal	71.50	15.87	12.62
NPDR	9.72	83.42	6.85
PDR	17.61	17.14	65.23

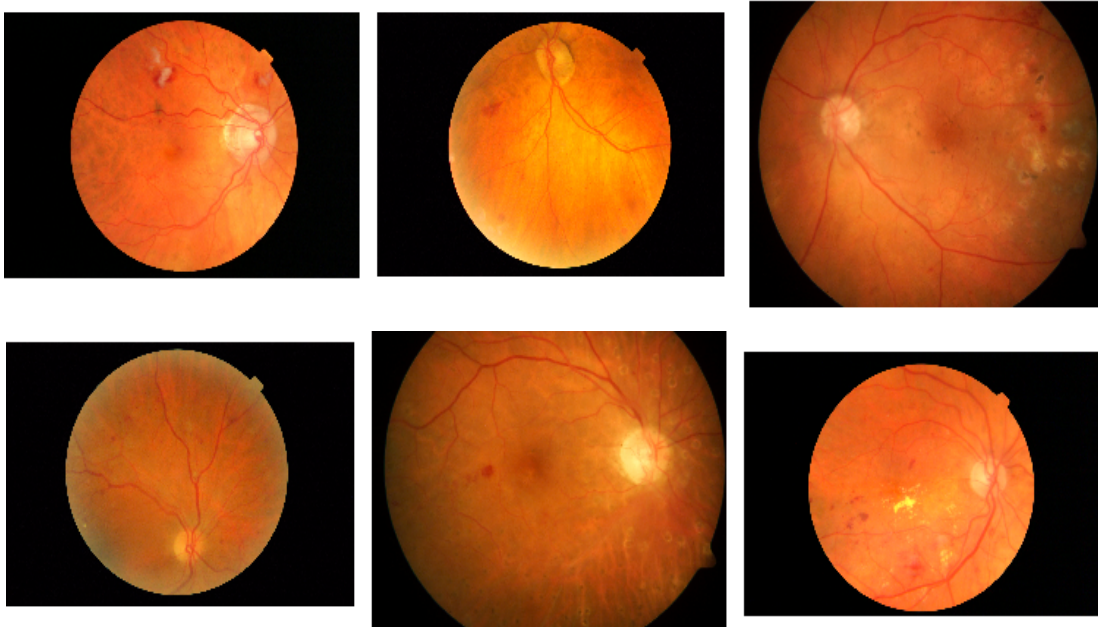


Figure 9. Sample retrieved images for a queried PDR image. The image on top left is the query image and the other images are the retrieved PDR images.

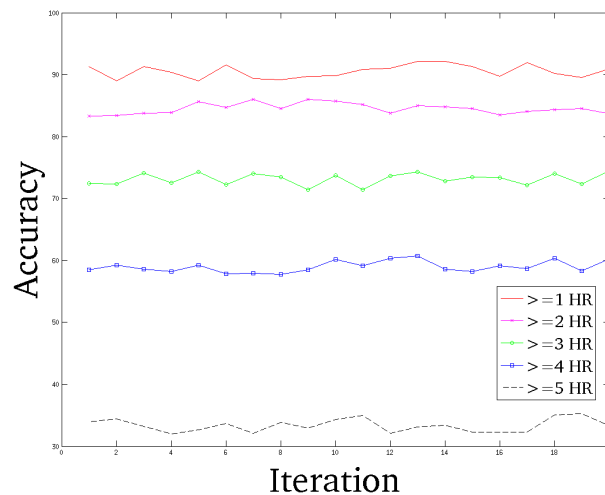


Figure 10. Consistency of hit rates.

6. CONCLUSION

In this paper, we presented a novel approach to retrieval of clinically-relevant DR images. The approach is based on a multi-class multiple-instance learning framework, which is most appropriate for the problem at hand. We also presented new features that are specifically designed for DR images, considering their peculiar color and structural properties. The approach was evaluated on a database of real DR images, and comparison was made with a few leading approaches in the literature. The experimental results demonstrate that the proposed approach can achieve better retrieval performance

than prior-art methods used in content-based image retrieval. In particular, the retrieved images are observed to be of clinical and visual relevance.

REFERENCES

- [1] "Prevalence and incidence of diabetes mellitus in United States, 1980-1997.," Centers for Disease Control and Prevention, Morbidity and Mortality Weekly Report, (MMWR) Rep.39, 1990.
- [2] "Grading diabetic retinopathy from stereoscopic color fundus photographs - An extension of the modified airleie house classification.," Early treatment of diabetic retinopathy study group (ETDRS). Rep. 10 Ophthalmology., 1991.
- [3] Cai, W., Feng, D., and Fulton, R., "Content-based retrieval of dynamic PET functional images.," *IEEE transactions on information technology in biomedicine.*, vol. 4, pp. 152-158, 2000.
- [4] Chu, W., Leong, I. and Taira, R., "A semantic modeling approach for image retrieval by content.," *The VLDB journal-The international journal on very large databases.*, vol. 3, pp. 445-477, 1994.
- [5] Kelly, P., Cannon, T., and Hush, D., "Query by image example: the comparison algorithm for navigating image databases (CANDID) approach.," in *Proceedings of the SPIE*, 1995.
- [6] Korn, P., Sidiropoulos, N., Faloutsos, C., Siegel, E. and Protopapas, Z., "Fast and effective retrieval of medical tumor shapes.," *IEEE transactions on knowledge and data engineering*, vol. 10, pp. 889-904, 1998.
- [7] Nah, Yunmook and Sheu, P. C.-Y., "Image content modeling for neuroscience databases.," in *Proceedings of the 14th international conference on software engineering and knowledge engineering.*, Ischia, Italy, 2002.
- [8] Shyu, C., Brodley, C., Kak, A., Kosaka, A., Aisen, A. and Broderick, L., "ASSERT: a physcian-in-the-loop content-based retrieval system for HRCT image databases.," *Computer vision and image understanding.*, vol. 75, pp. 111-132, 1999.
- [9] Goldbaum, M., Katz, N., Chaudhuri, S. and Nelson, M., "Image understanding for automated retinal diagnosis," in *Proceedings of the annual symposium on computer application in medical care.*, 1989.
- [10] Gupta, A., Moezzi, S., Taylor, A., Chatterjee, S., Jain, R., Goldbaum, L. and Burgess, S., "Content-based retrieval of ophthalmological images.," in *International conference on image processing (ICIP)*, 1996.
- [11] Lowe, D., "Distinctive image features from scale-invariant keypoints.," *International journal of computer vision*, vol. 60, pp. 91-110, 2004.
- [12] Dalal, N., and Triggs, B., "Histograms of oriented gradients for human detection.," in *IEEE computer society conference on computer vision and pattern recognition.*, 2005.
- [13] Huang, J., Kumar, S., Mitra, M., Zhu, W. and Zabih, R., "Image indexing using color correlograms," in *IEEE computer society conference on computer vision and pattern recognition (CVPR)*, 1997.
- [14] Li, M., "Texture moment content-based image retrieval," in *IEEE international conference on multimedia and expo (ICME)*, 2007.
- [15] Manjunath, B., and Ma, W., "Texture features for browsing and retrieval of image data," *IEEE transactions on pattern analysis and machine intelligence (PAMI-special issue on digital libraries)*, vol. 18, pp. 837-842, August 1996.
- [16] Chen, Q., Tai, X., Dong, Y., Pan, S., Wang, X. and Yin, C., "Medical image retrieval based on sementic of neighborhood color moment histogram," in *The 2nd international conference on bioinformatics and biomedical engineering (ICBBE)*, 2008.
- [17] Freeman, W. and Adelson, E., "The design and use of steerable filters," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 13, pp. 891-906, 1991.
- [18] Loy, G. and Zelinsky, A., "Fast radial symmetry for detecting points of interest," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 25, pp. 959-973, 2003.
- [19] Wang, J. and Zucker, J.-D., "Solving the multiple-instance problem: A lazy learning approach," in *17th International conference of Machine Learning*, 2000.
- [20] Kauppi, T., Kalesnykiene, V., Kamarainen, J., Lensu, L., Sorri, I., Uusitaloh, H., Kalviainen, H. and Pietila, J., "Diaretdb0: Evalutaion database and methodology for diabetic retinopathy algorithms," 2006.
- [21] Kauppi, T., Kalesnykiene, V., Kamarainen, J., Lensu, L., Sorri, I., Raninen, A., Voutilainen, R., Uusitalo, H., Kalviainen, H. and Pietila, J., "Diaretdb1: Diabetic retinopathy database and evaluation protocol," *Procdeeings of medical image understanding and analysis (MIUA)*, pp. 61-65, 2007.
- [22] McCormick, B. and Goldbaum, M., "STARE=Structered analysis of the retina: Image processing of TV fundus image.," 1975.
- [23] Sigurbjornsson, B. and Van Zwol, R., "Flickr tag recommendation based on collective knowledge," in *Proceedings of the 17th international conference on world wide web*, 2008.

- [24]Katz, N., Goldbaum, M., Nelson, M. and Chaudhuri, S., "An image processing system for automatic retina diagnosis.," in *SPSE International symposium and exposition on electronic imaging devices and systems*, 1988.
- [25]Chaum, E., Karnowski, T., Govindasamy, V., Abdelrahman, M. and Tobin, K., "Automated diagnosis of retinopathy by content-based image retrieval.," *Retina*, vol. 28, p. 1463, 2008.
- [26]Lalonde, M., Lalibert, F. and Gagnon, L., "RetsoftPlus: A tool for retinal image analysis.," in *IEEE symposium on computer-based medical systems (CBMS)*, 2004.
- [27]Zana, F. and Klein, J., "A multimodeal registration algorithm of eye fundus images using vessels detection and Hough transform.," *IEEE transactions on Medical Imaging.*, vol. 18, pp. 419-428, 1999.
- [28]Ravishankar, S., Jain, A. and Mittal, A., "Automated feature extraction for early detection of diabetic retinopathy in fundus images.," in *IEEE conference on computer Vision and Pattern Recognition (CVPR).*, 2009.