A Machine-learning approach to Retrieving Diabetic Retinopathy Images

[1]Parag S. Chandakkar, [1]Ragav Venkatesan, [1]Baoxin Li and [2]Helen K. Li
{pchandak, ragav.venkatesan, baoxin.li} @asu.edu; hli@communityretina.com
[1]Computer Science and Engineering, Arizona State University
[2]Weill Cornel Medical College/The Methodist Hospital, University of Texas, and Thomas Jefferson University

INTRODUCTION

- DR is a leading cause of blindness among diabetics.
- Vision loss is a devastating complication despite the advancement in diabetic care.
- Timely diagnosis and treatment of DR can significantly reduce the risk of vision loss.
- Unfortunately, there is a dearth of computer-based systems that can match the level of performance achieved by experienced ophthalmologists.
- Vast knowledge hidden in archived images can be exploited by retrieving clinically relevant images.

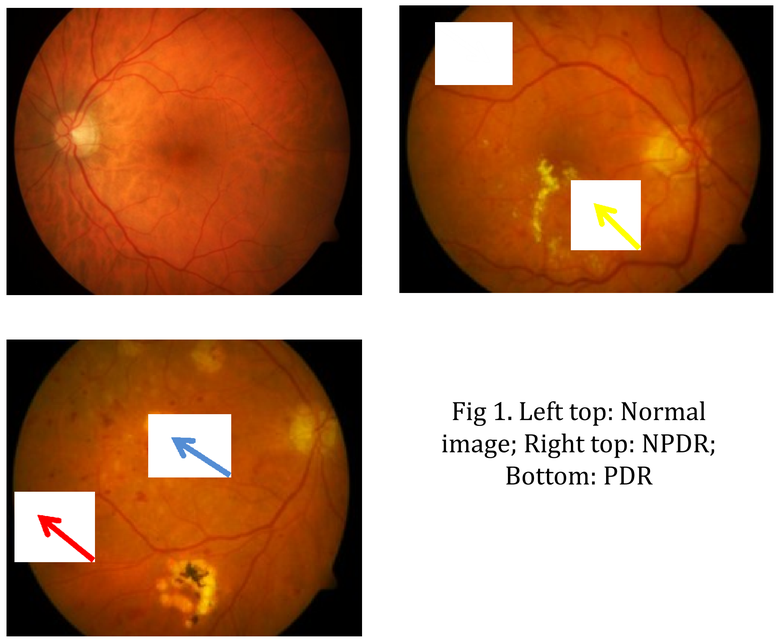

Fig 1. Left top: Normal image; Right top: NPDR; Bottom: PDR

- Red and blue arrow indicate hemorrhages and cotton wool spots respectively. White and yellow arrow indicate microaneurysms and yellow, waxy exudates.
- We propose a multi-class multiple-instance DR image retrieval framework that makes use of a modified color correlogram (CC) and statistics of steerable Gaussian filter (SGF) responses.

RELATED METHODS

- Several attempts have been made to develop content-based medical image retrieval systems.
- Structured Analysis of Retina (STARE) project aims searching for images similar in content [1].
- These images are largely similar in appearance but they are not clinically relevant.
- This method may not convey clinically useful information. This renders the clinical use of this method.

PROPOSED APPROACH

Features:-

- Color correlogram (CC) is a well-studied feature for image retrieval.
- CC features are unsuitable for DR images due to their unique color spectrum.
- A modified CC feature, which adapts well to DR images is proposed in [2].

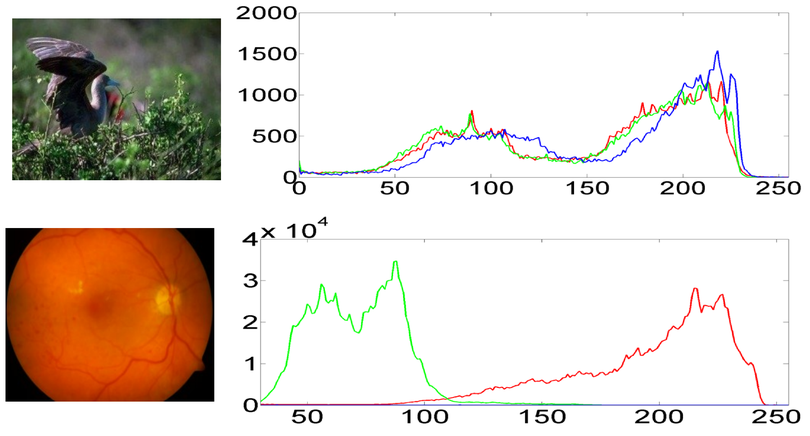

Fig 2. Color spectrum of a natural and a DR image.

- Along with the said CC features, fast radial symmetric transform and steerable Gaussian filter responses are also used to pick out regions of interest and retinal landmarks and provide good features.

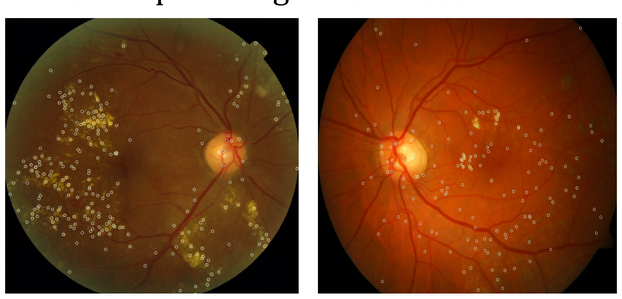

Fig 3. Fast radial symmetric transform.

Retrieval:-

- The learning approach used is the Multi-Class Multiple Instance Learning (McMIL).
- McMIL is already well studied for the purposes of DR image classification [2].

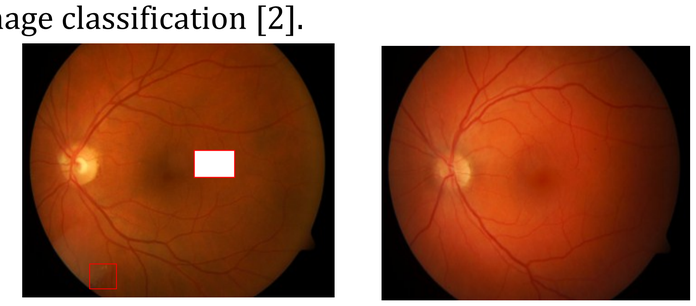

Fig 3. Necessity for a MIL Framework.

- A very small localized region in an otherwise normal image is enough to label an image as affected.
- An image is divided into instances and we work on those instances using Haussdorff distances as in [3].
- Minimum distance between the i^{th} instance of bag A and every instance of bag B is given by:

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

- A retrieval system is developed based on learning using instance level distances while retrieving bag level images.
- A similarity list (SL) is thus produced giving a similarity measure between every instance of a query image and all the images in the database.
- The mean of all such ranks in the SL for any image is the *m-rank.*
- Final *meanRank* is the average of *m-Rank* and *citer-Rank* of the query image.

EXPERIMENTAL RESULTS

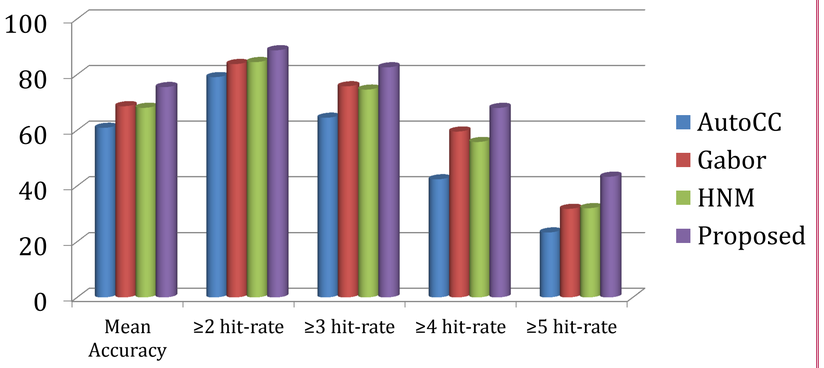

Fig 5. Mean Accuracy and Hit-rates in percentages

- Experiments were performed on a manually assembled database containing 425 images including 160 Normal, 181 NPDR and 84 PDR images.

CONCLUSION

- An image retrieval algorithm was proposed based on a modified CC feature space and a multi-class multiple instance distance framework.
- The proposed approach is found to retrieve images of better clinical relevance than the prior art algorithms.
- Further research is under progress to improve the algorithm and make it a system of clinical use.

REFERENCES

[1] A. Gupta , S. Moezzi, A. Taylor, S. Chatterjee, R. Jain, L, Goldbaum and S. Burgess, "Content-based retrieval of opthalmological images," *in the Proc. of ICIP,* 1996.

[2] R. Venkatesan, P. S. Chandakkar, B. Li and H. Li, "Classification of Diabetic Retinopathy Images Using Multi-Class Multiple-Instance Learning Based on Color Correlogram Features," *in Proc. of IEEE EMBS, San Diego,* 2012.

[3] J. Wang and J.-D. Zucker, "Solving the multiple-instance problem: A lazy learning approach," in *Proc. of 17th International conference of Machine Learning,* 2000.