World Academy of Science, Engineering and Technology
International Journal of Computer and Information Engineering
Vol:8, No:3, 2014

Similarity Based Retrieval in Case Based Reasoning for Analysis of Medical Images

M. Das Gupta, S. Banerjee

***Abstract*—**Content Based Image Retrieval (CBIR) coupled with Case Based Reasoning (CBR) is a paradigm that is becoming increasingly popular in the diagnosis and therapy planning of medical ailments utilizing the digital content of medical images. This paper presents a survey of some of the promising approaches used in the detection of abnormalities in retina images as well in mammographic screening and detection of regions of interest in MRI scans of the brain. We also describe our proposed algorithm to detect hard exudates in fundus images of the retina of Diabetic Retinopathy patients.

***Keywords*—**Case based reasoning, Exudates, Retina image, Similarity based retrieval.

I. Introduction

CASE based reasoning (CBR) is a methodology for solving problems by utilizing previous experiences. The underlying assumption is that similar problems have similar solutions, an idea which essentially embodies a physician's approach in the diagnosis and therapy planning of medical ailments. The knowledge of medical experts is a mixture of textbook knowledge and experience acquired through real life clinical cases. Thus, there is a growing interest in using CBR [1] in developing medical decision support systems. The basic idea involved is to retrieve relevant cases from a case data base and establish the relevance between candidate cases and prototype cases of the data base through a similarity measure. When the case history involves the analysis and classification of sets of longitudinal series of multimedia image sets, automatic indexing using digital content, referred to as Content Based Image Retrieval (CBIR) [2] is a possible solution for defining similarity measures. Similarity-based image retrieval is part of the case-based reasoning scenario [3] which allows for the retrieval of images from a database that are similar in some way to a given query image.

II. Decision Trees

In [4], authors present a Case Based Reasoning (CBR) system for the retrieval of medical cases made up of a series of images with contextual information (such as the patient age, sex and medical history). When designing a CBR system to retrieve such cases, several problems arise: we have to aggregate heterogeneous variables (images, nominal and continuous variables), and moreover, we sometimes have to deal with missing information. Decision trees (generally used for classification) are well suited to solve both these problems. So a retrieval framework from decision trees is derived [4], which are well suited to process heterogeneous and incomplete information. In [4] it emerges that a retrieval system based on several trees is also more accurate than system based on single tree.

M. Das Gupta and S. Banerjee are with the West Bengal University of Technology, Kolkata, India (e-mail: madhuparnadgupta@gmail.com, sreeparna.banerjee@wbut.ac.in).

A. Images in Decision Trees

The integration of images in a DT was inspired by CBIR. CBIR involves

1) building a signature for images by extracting image features, and
2) defining a distance measure associated with the signature.

Thus, measuring the distance between two images comes down to measuring the distance between two signatures. Similarly, we could segment cases according to an "image attribute" by clustering the corresponding signatures, and assign each cluster to a child node.

In [4] authors proposed to compute a signature for images (i.e. a feature vector summarizing image content) from their wavelet transform (WT) [5]. These signatures model the distribution of the WT coefficients in each sub-band of the decomposition. The associated distance measures D [5] compute the divergence between these distributions. These signatures and distance measures were used to build the DTs.

Any clustering algorithm can be used, provided that the distance measure between feature vectors can be specified. FCM (Fuzzy C-Means) [6] is used, one of the most common algorithms, and replaced the Euclidian distance by D.

B. Decision Tree Theory

Decision trees (DTs) [7], [8] are used to divide a population of cases into homogeneous groups, according to a set of discriminant features; these features are automatically searched for (by a learning process) amongst all the available features, as explained below. The case population is segmented in a hierarchical way; hence a tree with such structure is built:

- each non-terminal node corresponds to a test on a single feature (e.g. what is the patient sex ?)
- each edge corresponds to a test outcome (e.g. male/female)
- each leaf corresponds to a cluster of cases that provide a similar answer to each test

At the beginning of the learning process, the tree is made up of a single node containing the whole case population. Then, for each leaf L of the developing tree, the most discriminant feature is searched for and the population in L is split into new child nodes, one for each outcome of the test. The

discriminant power of a test can be measured by the Shannon entropy gain G [7] (see (1)) obtained when dividing the current node into child nodes,

$$\begin{cases} G = \left(\sum_{n=1}^{N} I^n\right) - I^0 \\ I^i = -\sum_{c=1}^{C} p_c \, log p_c, \quad i = 0..N \end{cases} \tag{1}$$

where p_c is the percentage of cases with label c (c = 1..C) in a node, I^0 is the entropy in the parent node (before dividing it) and I^n (n = 1..N) is the entropy in the nth child node. This measure characterizes the purity of the segmentation. DTs can process any feature, so long as we provide a way to cluster cases according to that feature. Since each test is performed on a single feature, DTs are well suited to process heterogeneous cases. Moreover, DTs can manage missing information. It is supposed that the value of a feature f, tested at a node v_0, is missing for some case. Then this case is assigned to each child v_i of v_0 with a weight w_{v_i}, $0 \leq w_{v_i} \leq 1$. w_{v_i} is the percentage of samples, whose value for f is known, assigned to v_i.

C. Multimodal Decision Tree Based Indexing

At the end of the learning step, each supervision example i is assigned to each leaf j (j = 1..N) with a weight w_{ij} (w_{ij}=0 or 1 if every tested attribute is known for i, 0 $\leq w_{ij} \leq$1 otherwise,). Similarly, when a new case *q* is presented to the system, we can assign it to each leaf with a weight w_{qj} . To derive a retrieval system from a DT, we apply the following method:

1) The similarity measure S_{qi} between q and each supervision example i is initially set to 0.
2) If q and some example i fall in the same leaf j, their similarity measure are increased according to their assignment weight to j, namely by $w_{qj}.w_{ij}$. In other words, $S_{qi} = \sum_{j=1}^{N} w_{qj} . w_{ij}$.
3) Examples i are ordered by decreasing order of S_{qi}.

A similar retrieval system can be derived from several trees simultaneously: the similarity measure S_{qi} is then simply computed over every leaf of the set of trees. The performances as classifier are usually better than that of single DTs.

D. Calibration Procedure

Class labels for supervision examples are assigned to evaluate the discrimination ability of each attribute (1). The disease severity level was used in that purpose.

To learn DTs, cases are divided into:

i. a learning set (the supervision examples), used to find the most discriminative attributes at each node,
ii. a validation set, used to determine when we should stop dividing nodes,
iii. a test set to evaluate the efficiency of the system

The efficiency of the system is defined as the mean sensitivity over the test set: the sensitivity is the percentage of retrieved cases whose label is identical to the query cases. To improve the system, only the "best" generated trees are used for retrieval.

E. Application

DTs can handle missing values and to avoid over learning. The latter property makes this method well suited to process databases with few cases such as the DR database. This stands to reason since an image alone is generally not sufficient for experts to correctly diagnose the disease severity level of a patient. Using images series without contextual information, instead of single images, increases by itself the sensitivity by a factor of 144.7%. Besides, this non-linear retrieval method is 152.0% more sensitive than a simple linear combination of heterogeneous distances on the DR database.

III. Description of the Bayesian Network and Dezert Smarandache Theory

A novel content-based heterogeneous information retrieval framework, particularly well-suited to browse medical databases and support new generation Computer Aided Diagnosis (CADx) systems, is presented in [9] to retrieve possibly incomplete documents, consisting of several images and semantic information, from a database. The proposed retrieval method relies on image processing, in order to characterize each individual image in a document by their digital content, and information fusion. Two novel information fusion methods are proposed in [9]. In the first method, the degrees of match are fused by the Bayesian network itself. In the second method, they are fused by the Dezert-Smarandache theory. The proposed methods were applied to two heterogeneous medical databases, a diabetic retinopathy database and a mammography screening database, for computer aided diagnosis. Precisions at five of 0.809 ± 0.158 and 0.821 ± 0.177, respectively, were obtained for these two databases.

A. Bayesian Network Theory

A Bayesian network [10] is a probabilistic graphical model that represents a set of variables and their probabilistic dependencies. It is a directed acyclic graph whose nodes represent variables, and whose edges encode conditional independencies between the variables. Information coming from each attribute is then used to derive an estimation of the degree of match between a query document and a reference document in the database. Then, these estimations are fused.

1. Learning Bayesian Network from Data

A Bayesian network is defined by a structure and the conditional probability of each node given its parents in that structure (or its prior probability if it does not have any parent). These parameters can be learned automatically from data.

Defining the structure consists in finding pairs of nodes (*X,Y*) directly dependent, i.e. such that:

1) *X* and *Y* are not independent $(P(X,Y) \neq P(X)P(Y))$
2) There is no node set *Z* such that *X* and *Y* are independent given $Z\ (P(X,Y\ |Z) \neq P(X|Z)P(Y|Z))$

2. Including Images in Bayesian Network

Contextual information is included as usual in a Bayesian network: a variable with a finite set of states, one for each possible attribute value, is defined for each field. To include images in a Bayesian network, we first define a variable for each image in the sequence. For each "image variable", the usual steps of Content-Based Image Retrieval (CBIR) [11] are followed, i.e.:

1) building a signature for each image (i.e. extracting a feature vector summarizing their digital content), and
2) defining a distance measure between two signatures.

Thus, measuring the distance between two images comes down to measuring the distance between two signatures. Similarly, in a Bayesian network, defining states for an "image variable" comes down to defining states for the signature of the corresponding images.

3. Computing the Conditional Probabilities

To compute the conditional probability first estimation of the probability that a reference document with full membership to a particular state of particular attribute is relevant. It can then be computed using Bayes' theorem. The prior probability is required; it can be estimated by the probability that two documents belong to the same class.

4. Retrieval Process

The different reference documents in the database are processed sequentially. To process a document, every available attribute for that document is processed as evidence and Lauritzen and Spiegelhalter's [12], inference algorithm is used to compute the posterior probability of each variable. The reference documents in the database are then ranked in decreasing order of the computed posterior probability.

B. Dezert-Smarandache Theory

The Dezert-Smarandache Theory (DSmT) of plausible and paradoxical reasoning, proposed in recent years [13], helps to combine any types of independent sources of information represented in term of belief functions. It generalizes the theory of belief functions (Dempster-Shafer Theory - DST) [14], which itself generalizes the Bayesian theory, used in the system above. DSmT is mainly focused on the fusion of uncertain, highly conflicting and imprecise sources of evidence.

To process a reference document, every available attribute for that reference document *is* processed as evidence and Lauritzen and Spiegelhalter's inference algorithm is used. If a particular attribute of the reference document is available, the degree of match is computed accordingly and the belief masses are computed. The sources available for that document are then fused. Once the sources available for the document are fused by the proposed rule of combination, the pignistic probability is computed. The reference documents are then ranked in decreasing order of probability.

C. Application

The proposed method has been applied to CADx on two heterogeneous databases. First, it has been applied to diabetic retinopathy severity assessment on a dataset (DRD) built at the Inserm U650 laboratory, in collaboration with ophthalmologists of Brest University Hospital. Then, it has been applied to breast cancer screening on a public access database (DDSM).

1. Diabetic Retinopathy Database (DRD)

The diabetic retinopathy database contains retinal images of diabetic patients, with associated anonymized information on the pathology. The database consists of 67 patient files containing 1112 photographs altogether. Images have a definition of 1280 pixels/line for 1008 lines/image. They are lossless compressed images. The contextual information available is the age and sex of the patient, as well as structured medical information. Patient's records consisted of at most 10 images per eye and 13 contextual attributes.

2. Digital Database for Screening Mammography (DDSM)

The DDSM project [15] is a mammographic image database for research on breast cancer screening. It consists of 2277 patient files. Each of them includes two images of each breast, associated with patient information (age at time of study, subtlety rating for abnormalities, American College of Radiology breast density rating and keyword description of abnormalities) and image information (scanner, spatial resolution, ...). The following contextual attributes are used in this study:

i. the age at time of study
ii. the breast density rating

Images have a varying definition, of about 2000 pixels/line for 5000 lines/image. There is no missing information in DDSM. Each patient file has been graded by a physician. Patients are then classified in three groups: normal, benign and cancer.

IV. Description of Time Series Analysis

In [16] authors describe an approach to Case Based Reasoning (CBR) for image categorisation. The technique is founded on a time series analysis mechanism whereby images are represented as time series and compared using time series similarity techniques. This paper explores two mechanisms where images can be represented as time series (curves). The first takes into account the entire image while the second is directed at some specific feature within the image. The choice of images is partially dependent on the nature of the application. If the content of the entire image is important or if there is no single defining feature, then the first should be adopted. The second approach is only applicable if there is some feature that exists across the image set that is significant with respect to a particular application. Once a time series representation has been adopted some similarity checking mechanisms are required. This procedure involves some form of curve comparison. The technique promoted in this paper is Dynamic Time Warping (DTW). This was selected because it operates on curves that are not necessarily of the same unit length.

A. DTW

Dynamic Time Warping (DTW) is a technique whereby two time series can be compared. The techniques does not require the two curves to be of the same length and takes in to account a certain amount "skew" to obtain a best fit. Given two time series: Q = $\{q_1, q_2, \ldots, q_i, \ldots q_n\}$ and C = $\{c_1, c_2, \ldots, c_j, \ldots c_n\}$, these can be aligned using DTW by constructing a n by m grid (matrix) such that the value for element (i; j) is the squared Euclidean distance from point c_j on curve C (a comparator sequence) to point q_i on curve Q (The query sequence, i.e. a sequence we wish to compare to C with the aim (say) of categorizing Q). The best match between the two sequences Q and C is the warping path that minimizes the total cumulative distance from grid element (0,0) to (n,m). A warping path is any contiguous set of matrix elements from (0,0) to (n,m). The warping cost associated with a particular path is its cumulative distance.

B. Essential Theory

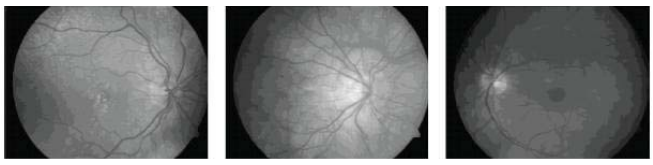
Time Series Analysis (TSA) is concerned with the study of data that can be represented as one or more curves with a view to extracting knowledge. The fundamental issues of TSA are: (i) how to measure similarity between time series and (ii) how to compress time series while maintaining discriminatory power. Similarity can be measured in terms of ([17]): (i) similarity in time, (ii) similarity in shape and (iii) similarity in change.

C. Application

Two specific applications are considered: (i) the screening of retina images for Age-related Macular Degeneration (AMD), and (ii) the categorization of Magnetic Resonance Imaging (MRI) brain scans. AMD is an eye condition that affects the macula, the central portion of the retina. It is the leading causes of irreversible blindness in the elderly and is a growing healthcare challenge due to our ageing population; early detection may offer timely preventive treatment to inhibit the progress of the condition. A good way of identifying the early onset of AMD is through the identification of "fatty deposits" (called drusen) and pigment abnormality in the retina. The screening of retina images or AMD requires the entire image to be taken into consideration. The second application, the categorisation of MRI brain scans is directed at a particular region within such scans, namely the corpus callosum. The corpus callosum connects the two hemispheres of the brain. It is conjectured that the size and shape of the corpus callosum dictates certain human abilities (such as mathematical or musical abilities) and that it characterizes certain medical conditions such as epilepsy.

1. AMD Screening

The objective is to detect the presence of AMD in retina images collected as part of a screening programme. Thus retina images are categorised or classified as positive (evidence of AMD detected) or negative (normal). Three example images are presented in Fig. 1. The image on the rightmost (Fig. 1 (c)) is from a normal eye, while the other two Figs. 1 (a) and (b) are images from eyes of AMD that contain drusen (light coloured flecks scattered across the image) and other pathological features. The optic disc (OD), a bright coloured disc featured in the retina images in Fig. 1 from which all blood vessels emanate, connects the retina to the "optic nerve". The macula, a dark colored region as is shown in Fig. 1 (c), acts as a light detector and provides us the central vision essential for seeing fine details (the macular is obscured by drusen in Figs. 1 (a) and (b).

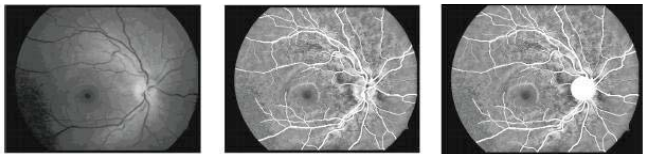


(a) (b) (c)

Fig. 1 Example Retina Images, (a) and (b) feature AMD, while (c) does not

- *Time Series CBR for AMD Screening*

Images are represented in terms of pixel values using the Red-Green-Blue (RGB) color model and the Hue Saturation-Intensity (HSI) representation of the RGB model. As such each image can be represented as a sequence of histograms, with length M, to which a curve can easily be fitted. Each histogram is represented as $h_i(m) = \beta$ *where* $0 \leq m < M$ *and* β is the number of occurrences of intensity value *m* in image i.



(a) (b) (c)

Fig. 2 Image Enhancement

Prior to translating the images into time series some enhancement to the histograms was undertaken. The enhancement was done by applying a Contrast Limited Adaptive Histogram Equalisation (CLAHE) [18] technique. CLAHE computes histogram for each different parts of an image and equalise each histogram separately. This image enhancement process increased the visibility of edges in the retinal images, as shown in Fig. 2 (a). The technique described here thus considers only the green and saturation channels.

It was also found that the removal of pixels representing blood vessels enhanced the categorisation process. The identified blood vessels pixels were replaced by null values and consequently omitted from the histogram generation process. Fig. 2 (b) gives an example of a retina image with blood vessel pixels removed by applying this process to the image given in Fig. 2 (a). The optic disc can obscure the presence of drusen. It is technically possible to remove the pixels representing the optic disc in the same way that blood

vessel pixels were removed. The optic disc pixels value was then replaced with null values. Fig. 2 (c) shows the retinal image given in Fig. 2 (b) with the optic disc removed.

However, the routine removal of the optic disc can result in the removal of pixels representing drusen; especially where the drusen are close to or superimposed on, the optic disc. A two stage CBR approach was thus adopted comprising two Case Bases (CBs), the primary CB and the secondary CB. The primary CB comprised the green and saturation histograms of labelled retina images (positive and negative) that included the optic disc but with blood vessels pixels removed, the secondary CB comprised the similar histograms but with both the blood vessels and optic disc removed.

Given a new image we attempt to categorize this with reference to the primary CB first. The green channel histogram of the new image is compared to each of the green channel histograms in CB1 by means of computing the similarity measure between two histograms using DTW. A similar approach is also applied to the saturation histograms. These processes will generate the preliminary results comprising distance values between the green and saturation histograms of the new image, with the green and saturation histograms of each image in primary CB1. The similarity between the new image and each case in CB1 is then calculated by taking the average of each case green and saturation histograms similarity values. If there exist only one "most similar" case, or there exist a number of most similar cases but all with the same label, the preliminary results will be taken as the final categorisation result and consequently the new image will be labelled as AMD or normal according to the label of the most similar image in CB1. If no clear result is obtained (i.e. there are a number of most similar cases with contradicting labels) the pixels representing the optic disc in the new image are removed and a similar CBR process is followed but with the secondary CB.

2. MRI Scan Categorisation

The second application considered in this paper is the categorisation of MRI brain scan according to a single feature within those scans, namely the Corpus callosum as shown in Fig. 3. The objective of the study was to investigate the application of time series CBR in the context of a Region of Interest (ROI) contextualisation. The size and shape of the corpus callosum has been shown to be correlated to gender, age, neurodegenerative diseases and various lateralised behaviour in people.

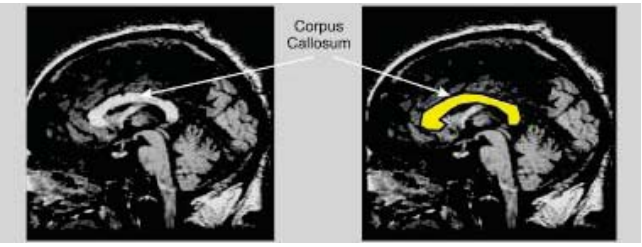


Fig. 3 Midsagital MRI brain scan slice showing the corpus callosum (highlighted in the right-hand image)

- *Time Series CBR for MRI Scan Categorisation*

In order to categorise images the first issue is to identify and isolate the feature of interest. A segmentation algorithm to identify the corpus callosum pixels as the position of the Corpus Callosum with respect to the boundaries of an MRI brain scan is roughly known. For this the efficient graph-based segmentation algorithm [19] is used. This method is based on minimum spanning trees (MST). All pixels of the original image are viewed as separate components. Two components are merged if the external variation between the components is small compared to the internal variations of them in successive way. The segmentation can be problematic as a related tissue structure; the Fornix is often included together with some other spurious pixel clusters. Some data cleaning like a smoothing technique is first applied to the MRI scans before the application of segmentation to preserves the boundaries between regions. This smoothing operation had the overall effect of bringing points in a cluster closer together.

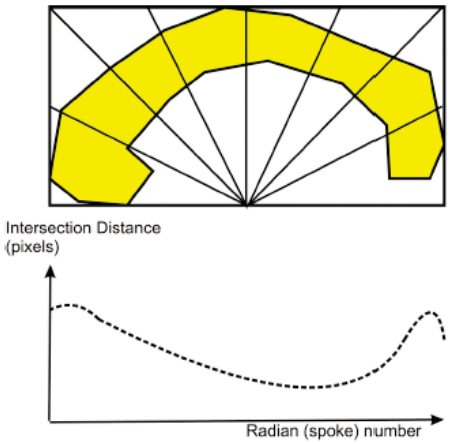


Fig. 4 Corpus callosum time series generation

Once the corpus callosum was identified it is represented as a time series so that the proposed time series CBR technique could be applied. The adopted time series generation approach is illustrated in Fig. 4. A series of "spokes" are radiated out from the mid-point of the base of the Minimum Bounding Rectangle (MBR) surrounding a detected a corpus callosum. The interval between spokes was one pixel measured along the edge of the MBR, consequently the number of spokes used to encode a corpus callosum varied from image to image. For each spoke the distance *Di* (where i is the spoke identification number) over which the spoke intersects with a sequence of corpus callosum pixels was recorded. The midpoint along the base of the MBR was chosen as this would ensure that there was only one intersection per spoke. The result is a time series with the spoke number i representing time and the value *Di*, for each spoke point, the magnitude. By plotting the *Di* against i a time series may be derived (as shown in Fig. 4).

To categorise new MRI brain scans, according to the nature of the corpus callosum, an appropriate Case Base (CB) was constructed comprising labelled curves generated in the

manner described above. A new case could then be compared, using DTW, to identify the most similar curve(s) in the CB.

V. Proposed Algorithm

Diabetic Retinopathy (DR) is a leading cause of loss of vision in patients afflicted with Diabetes. One of the earliest symptoms associated with DR is the presence of hard exudates (Figs. 5 (a) and (b)) which are yellow spots seen in the retina, usually in the posterior pole near the macula, caused by a breakdown of lipid products. A timely detection of hard exudates can help in therapy planning for arresting the proliferation of DR.

(a) (b)

Fig. 5 Example of retina; (a) normal, (b) with hard exudate

Our propsed algorithm (as shown in Fig. 6) involves a two stage process consisting of an extractiion of hard exudates from retina fundus images using the Fuzzy C-Means (FCM) clustering algorithm in the first stage. Then, in the second stage, these exudates are then compared with a retina image database containing DR images with hard exudates and a similarity measure is computed. The procedure is described as follows:

A. Image Feature Extraction

RGB images of the retina are first preprocessed (e.g. Contrast enhancement and color normalization) enhanced through the green channel. The green channel was selected due to its ability to show the greatest contrast compared to other colour channels, seen as essential for retinal object identification. Subsequently, these images are segmented into five classes using the Fuzzy C-means (FCM)[20] . each image can be represented as a sequence of histograms. Histogram for each different parts of an image is computed and each histogram is equalized separately. This image enhancement process increased the visibility of edges in the retinal images. First, the image is preprocessed and segmented. Then segmented image along with Optic Disc (OD) is chosen. Feature vector based on color and texture is extracted from the selected segment. The selected feature vectors are then classified as exudates and non-exudates [21]. Using the morphological erosion operation the exudates are removed from this segmented image and the resulting image is labeled as the OD image. Then a subtraction operation is performed to remove the optic disc from the segmented image using the OD image.

B. Classification of Images

The resulting segmented images are compared with the image database and the similarities are computed using the Euclidean distance measure and the distance measure given in Section II following [4].

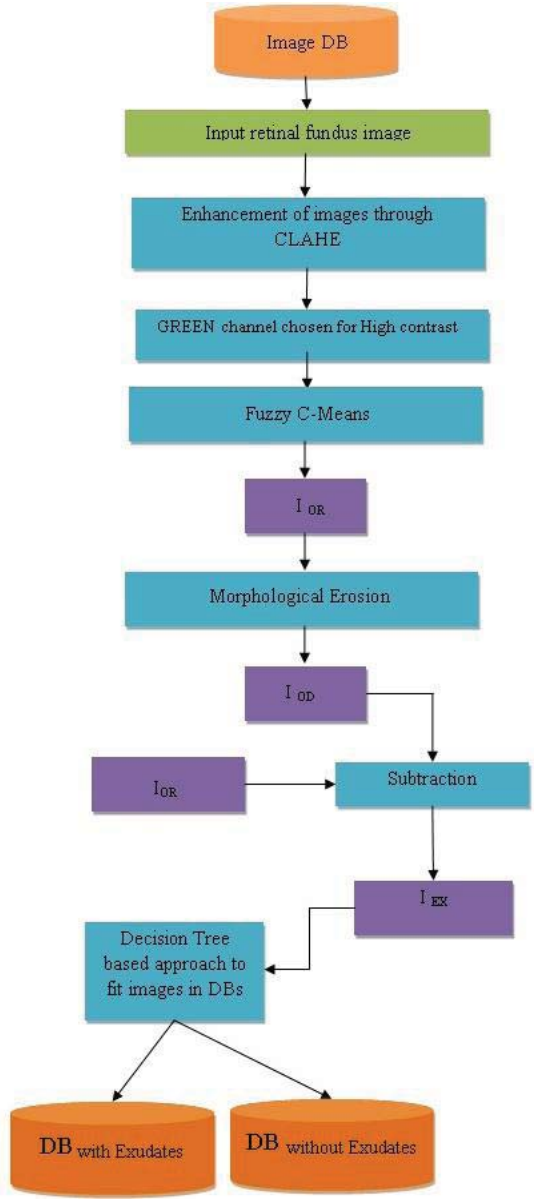


Fig. 6 Flowchart of proposed algorithm

VI. Conclusion and Future Work

In this paper, we have presented a survey of similarity based image retrieval or CBIR based techniques coupled with CBR for medical image analysis that are especially relevant to our proposed algorithm. We have also outlined our proposed algorithm. We have chosen a color histogram based segmentation technique due to its simplicity and robustness. Our similarity matching technique is based on the decision tree approach, because it is well suited to handle incomplete information. We chose the Euclidean distance measure (EM)

on account of its simplicity. We hope to use the distance measure defined in [4], [5] and compare its efficacy with EM based on a expert's (doctor) opinion. In this context, some other similarity based measures will also be tried and the best measure will be adopted in our algorithm.

ACKNOWLEDGMENT

The authors would like to acknowledge a grant from Department of Biotechnology, Government of India (No. BT/PR4256/BID/7/393/2012 dated 02.08.2012) for supporting this research.

REFERENCES

[1] A. Aamodt, "Case-based reasoning: Foundational issues, methodological variations, and system approaches," AI Communications, vol. 7, no. 1, pp. 39–59, March 1994.
[2] A. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 22, no.12, pp.1349–1380, December 2000.
[3] L. G. Shapiro , I. Atmosukarto , H. Cho , H. J. Lin , S. Ruiz-Correa, and J. Yuen, "Similarity-Based Retrieval for Biomedical Applications", in Case Based Reasoning in Images and Signals, by Petra Perner, Springer-Verlag, Berlin, Heidelberg, Studies in Computational Intelligence, Vol. 73 (2008)
[4] Gw´enol´e Quellec, Mathieu Lamard, Lynda Bekri, Guy Cazuguel, B´eatrice Cochener, Christian Roux, "Multimedia medical case retrieval using decision trees", 29th Annual International Conference of the IEEE EMBS , Lyon, France, August 23-26, 2007.
[5] M. Lamard, W. Daccache, G. Cazuguel, C. Roux, and B. Cochener, "Use of jpeg-2000 wavelet compression scheme for content-based ophtalmologic retinal retrieval," in Proceedings of the 27th annual international conference of IEEE engineering in medecine and biology society, september 2005.
[6] J. Bezdek, "Fuzzy mathemathics in pattern classification," Ph.D. dissertation, Applied Math. Center, Cornell University, Ithaca, 1973.
[7] J. Quinlan, C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers, 1993.
[8] L. Breiman, J. Friedman, R. Olshen, , and C. Stone, "Classication and regression trees," Wadsworth, Belmont, Ca., 1984.
[9] Gw´enol´e Quellec, Mathieu Lamard, Guy Cazuguel, *Member, IEEE*, Christian Roux, *Fellow Member, IEEE* and B´eatrice Cochener, "Case Retrieval in Medical Databases by Fusing Heterogeneous Information", IEEE Transaction on medical imaging 2010, pp.1-11.
[10] J. Pearl, *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann, 1988.
[11] A. W. M. Smeulders, M. Worring, S. Santini, A. Gupta, and R. Jain, "Content-based image retrieval at the end of the early years," *IEEE Trans Pattern Anal Mach Intell*, vol. 22, no. 12, pp. 1349–1380, 2000.
[12] S. L. Lauritzen and D. J. Spiegelhalter, "Local computations with probabilities on graphical structures and their application to expert systems," *J R Stat Soc*, vol. 50, no. 2, pp. 157–224, 1988.
[13] F. Smarandache and J. Dezert, *Advances and Applications of DSmT for Information Fusion I.* American Research Press Rehoboth, 2004, http://fs.gallup.unm.edu/DSmT-book1.pdf.
[14] G. Shafer, *A Mathematical Theory of Evidence*. Princeton University Press, 1976.
[15] M. Heath, K. Bowyer, D. Kopans, R. Moore, and W. P. Kegelmeyer, "The digital database for screening mammography," in *Proceedings of the Fifth International Workshop on Digital Mammography, Toronto, Canada*, 2000, pp. 212–218.
[16] Ashraf Elsayed, Mohd Hanafi Ahmad Hijazi, Frans Coenen, Marta Garcia-Finana, Vanessa Sluming, Yalin Zheng, "Image Categorisation Using Time Series Case Based Reasoning", Internatinal conference on case based reasoning (ICCBR 2011), London, UK, Springer 2011, pp. 423-436.
[17] Bagnall, A., Janacek, G.: Clustering Time Series with Clipped Data. Machine Learning, vol. 58, pp. 151-17 (2005)
[18] Zuiderveld, K.: Contrast Limited Adaptive Histogram Equalization. Academic Press Graphics Gems Series, pp. 474-485 (2001)
[19] Felzenszwalb, P., Huttenlocher, D.: E_cient Graph-based Image Segmentation. Int. journal of Computer Vision, vol. 59(2), pp. 167-181 (2004)
[20] Akara Sopharak,Bunyarit Uyyanonvara 1 and Sarah Barman, "Automatic Exudate Detection from Non-dilated Diabetic Retinopathy Retinal Images Using Fuzzy C-means Clustering", *Sensors* 2009, pp.2148-2161.
[21] B. Ramasubramanian, G. Prabhakar, "An Early Screening System for the detection of Diabetic Retinopathy using image processing", International Journal of Computer Applications (0975 – 8887) Volume 61– No.15, January 2013.