A Supervised Learning Approach for Classification of Medical Image Retrieval

Ranjana Battur[*], Jagadisha N

Department of Computer Science Engineering, Canara Engineering College, Mangalore, Karnataka, India

ABSTRACT

The advancement in biomedical engineering has been significant to the medical or healthcare industry. However, it faces issues like how it can be applied to the medicine and biology for healthcare aspects. Recently, quick advances of programming and equipment innovation have made simple, the issue of keeping up beneficial images accumulations. Visual elements like shading, and shape and composition are actualized for image retrieval. Conventional strategies for image indexing have been demonstrated neither reasonable nor effective regarding space and time so it set off the advancement of the new approach. A new concept called Content Based Image Retrieval (CBIR) is beneficial for the different sort of medical images having dissimilar imaging modalities, anatomic areas with diverse directions and biological schemes is projected. Classification of the medical image retrieval is the major concern for group of medical image. Hence, Support Vector Machine (SVM) classifier can be favorable for grouping forecast of query and database images based on similarity matching. It is very difficult to detect the features of the compared images effectively for all the different types of queries. Hence, the proposed SVM-MIR aims to classify and retrieval of biomedical images using SVM classifier method. The SVM-MIR based classification considers numerous groups of medical images for analysis. The outcomes of the proposed SVM-MIR approach achieve better performance compared to the existing approach.

Keywords: Medical image retrieval; SVM; CBIR; Feature extraction; Matching score

INTRODUCTION

The advancement in biomedical engineering has been significant to the medical or healthcare industry. However, it faces issues like how it can be applied to the medicine and biology for healthcare aspects. Recently, quick advances of programming and equipment innovation have made simple, the issue of keeping up beneficial images accumulations [1]. Visual elements as shading, and shape and composition are actualized for retrieval of images. Conventional strategies for image indexing have been demonstrated neither reasonable nor effective [2]. The conventional methods are of two stages process where initial stage performs the separation of image components to a recognizable degree [3]. Further, the components are coordinated which are outwardly comparative is finished. The two retrieval frameworks to be specific, substance and content-based retrieval frameworks vary as in the irreplaceable piece of last framework is human cooperation. Abnormal state highlights as watchwords, content depiction utilizes by people to gauge similitude and image understanding. Image retrieval framework is a viable and productive device for overseeing huge image databases. A CBIR framework permits the client to introduce a query image to recover images put away in the database as indicated by their comparability to the question image.

The CBIR is method used to address the image retrieval issue, like the issue of scanning for advanced images in extensive databases. The extraction of content works on the substance-based image retrieval. In an expansive sense, components may incorporate both content-based elements and visual elements (shading, surface, shape, faces). Also, subsequent to there as of now exists more works are on content based feature extraction framework and data retrieval research groups, hence proposed work limits to the strategies of visual feature extraction [4]. Inside the visual component scope, the elements can be further named general elements and space particular elements. Content-

based in CBIR implies that the inquiry makes utilization of the substance of the images themselves, as opposed to depending on literary comment or human-info metadata [5]. The visual components utilized for indexing and retrieval are grouped in into three classes: Primitive elements that are low level elements, for example, shading, shape and surface; consistent elements that are medium level elements portraying the image by an accumulation of articles and their spatial connections; and dynamic element that are semantic and logical elements. This paper introduces an SVM based medical image retrieval classification for different medical image databases [6]. The paper is specified with different sections: Section-2 (background/review of literature), section-3 (problem description and contribution), section-4 (Proposed SVM based Medical Image Retrieval (SVM-MIR)), section-5 (implementation of SVM algorithm for retrieval), section-6 (results and analysis) and section-7.

LITERATURE REVIEW

The wide range of research works addressing CBIR using supervised learning and fuzzy logic approaches are discussed in this section. A machine learning based approach for MIR is introduced in Rahman et al., considering similarity matching. The approach composed of pre-filtering and statistical distance measurement for similarity matching. Also, Relevance Feedback (RF) approach is implemented to achieve better perception subjectivity for matching functions. The outcomes are analyzed for precision-recall and cross validation accuracy of image categorization. The analysis suggests the improved, effective and efficient retrieval performance [7].

A Region based Image Retrieval (RIL) using Multi-label Neighbourhood Propagation (MNP) is presented in Li et al. The RIL using MNP composed of graph construction for edge detection, multilevel propagation for correlation and SVM for high level categorization of images. The outcomes give significant image retrieval than existing approaches. The disambiguation based Multiple Instance Learning (MIL) approach is described in Li [8]. The disambiguation mechanism helps to identify true positive cases in the positive cases and MIL helps in instance level classification. The given MIL approach offers robust, efficient and accurate classification.

The issues of RF algorithm are addressed in Bian and Tao, by using CBIR based on biased discriminant euclidean embedding. The performance of this technique was compared with traditional RF algorithm for corel image gallery and gives significant performance improvement.

In a work of Blanchart and Datcu, have presented a semi-supervised algorithm for CBIR in large image database [9]. The proposed study detects corresponding image classes with respect to training database.

A similarity fusion and filtering by using SVM classification and RF approach for MIR is presented in Rahman et al. Based on the prediction of a query image category, the system performs individual searches by itself to achieve query-specific outcomes. The experimental results yield efficient computation time and effective precision than existing approach.

A work of Zhang, et al., has mentioned CBIR of histopathological images for disease detection, diagnosis and decision support in medical area. The performance analysis is carried out for image classification and retrieval tests. The results provide higher classification accuracy and time efficiency [10].

A supervised learning approach is introduced to get the similarity between radiological images. This approach consistently improves the performance of CBIR by mean average precision. This CBIR technique can be applied to the systems in different medical problems, and in medical practices.

An Unsupervised Visual Hashing (VSH) approach is presented in for CBIR to boost the significance of visual hashing without any explicit semantic labels. The performance analysis is performed over website images and that offers better performance than existing techniques.

Annotating images with multi-labels are costly and time consuming. Hence, a graph theoretic approach is introduced in to address issues of image retrieval. The analysis is performed on aerial images that signify the proposed method with available CBIR methods [11].

The process of diagnosing cancer in early stage is very much essential. Hence, multi-parametric prostate MRI tool can be useful in prostate cancer diagnosis, Rossi, et al. have interpreted with CBIR. A comparative analysis is conducted with deep learning-based CBIRs and achieved diagnostic medical imaging retrieval.

A neural network based CBIR technique is introduced in for image retrieval [12]. The CBIR technique is applied for complex datasets like Caltech-256, Corel 1000 etc., and achieved better precision, Recall and F-score.

A survey on CBIR using supervised and unsupervised technique like fuzzy, SVM etc., are presented in. The author suggested that this technique is significant for image retrieval through matching of similar features to the input image.

A framework for fuzzy relevance based CBIR is presented in Kim-Hui Yap and Wu. This CBIR technique helps to learn the user information. Further, an efficient gradient descent-based learning mechanism is implemented to to estimate the underlying network parameters and it yields effective image retrieval performance [13].

A prototype of VOI based Functional Image Retrieval System (VOI-FIRS) has been designed in 2006 that performs extraction of multidimensional feature and retrieval. The outcomes of this system better retrieval of matching images having similar VOI features.

A fuzzy membership of descriptor (of type-2 beta) vectors is considered in that manages uncertainty and inaccuracy aspects extracted from the feature of images. The outcomes are obtained by considering three databases and yields better results. The research gap and problem description is discussed in next section.

Problem description and contribution

The domain of biomedical engineering has become a prime topic in the area of healthcare research. Recently, many of the research advancement have been witnessed with medical image analysis by retrieving through image processing. There are many image retrieval techniques have been presented in recent past with the ambition of restoring the radio logical image information and analyze them. From, the research survey it is observed that: (a) CBIR techniques with supervised learning approach are rare and have given less focus on comparative analysis with standard performance matrix, (b) CBIR technique with fuzzy logic are least focused in research area [14]. Hence, there is a need of efficient CBIR technique that can address the above mentioned problem.

Hence, this paper contributes to the medical area by offering a typical CBIR technique for medical images.

- Medical image retrieval for different database by separating and identifying similarity score by using supervised learning approach.
- Classifying similar medical images by using SVM based MIR approach and performs comparative analysis with existing work.

Proposed SVM based Medical Image Retrieval (SVM-MIR)

The structure of introduced SVM-MIR is given in Figure 1, which adapts SVM approach to perform retrieval image classification. The proposed SVM-MIR approach considers stored database of medical images of any type such as chest, hand, ankle, spine, MRI, mammogram images etc. The proposed approach computes image features by using Edge Histogram (EH), Color Layout (CL) descriptors and block division (8 × 8) methods [15]. All the images of the database are trained to extract the features. Next step of proposed SVM-MIR is to collect a query image and to retrieve the suitable matching image. In order to do so all the features of the query image are extracted and classify matching images by using SVM approach. The similarity score of the database is computed for both query and database images. Further, suitable matching is retrieved based on the similarity score and is stored in vector form. Finally, almost 12 similar images are extracted by considering similarity score, retrieve and prepared the matched chart for matched image.

The SVM methods are supervised learning methods and are mainly used to classify the image. Also, it observes the image database as two vector sets of dimension 'n' and performs the construction of separating hyper plane to maximize the margin among the images relevant and not-relevant to the query [16].
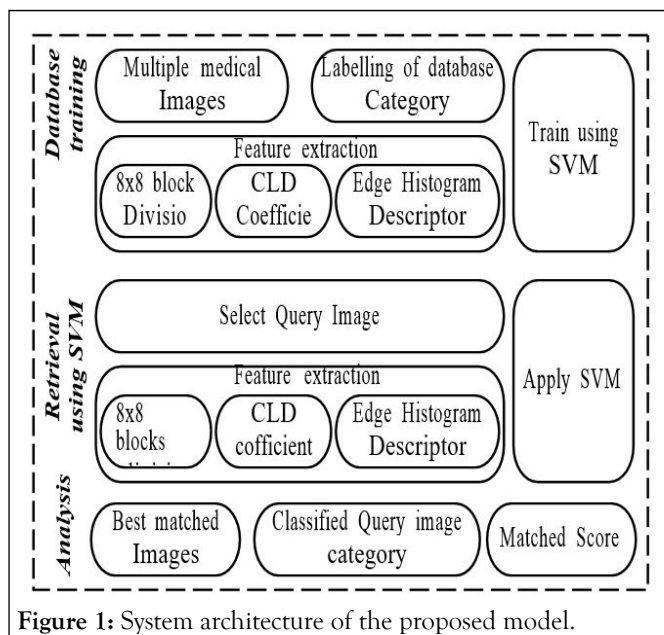

Figure 1: System architecture of the proposed model.

SVM method is a kernel method where it uses kernel function to determine its performance as a maximum margin classifier. These kernel methods maps the data implicitly to a high dimensional kernel space. This margin classifier can be determined in the kernel space and respective decision function in the original space which can be non-linear. The classification of non-linear data the feature space into linear data in kernel space by the SVMs (Figure 2). The SVM approach aims to identify an optimal hyper plane to separate relevant and irrelevant vectors by maximizing the size of the classification margin [17].
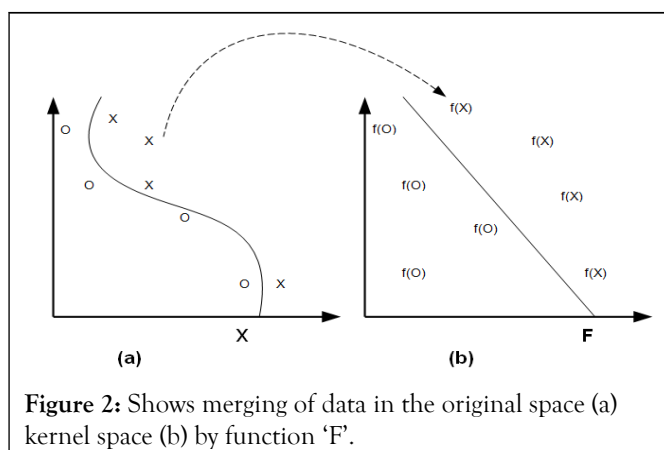

Figure 2: Shows merging of data in the original space (a) kernel space (b) by function 'F'.

The proposed SVM-MIR based image classification is a machine learning approach and it speed-up the image retrieval for large databases and enhances retrieval accuracy. Also, observed that SVM-MIR found more effective in the absence of labelled data, unsupervised clustering. The image clustering depends on a matching or similarity measure and classification through SVM [18].

Implementation of SVM algorithm for retrieval

The implementation of SVM algorithm for classification of retrieval image is given below. The initial step of proposed SVM-MIR approach is to train the database of medical image by using

SVM approach (Step-1). During training step, the number of input medical images is selected [19]. Then, these images are grouped in desired category like brain MRI, chest, mammogram etc. Further, features are extracted from the input images by using 8 × 8 block division, CL descriptor coefficient and Edge Histogram (EH) descriptor.

Algorithm for SVM-MIR approach

Input: Medical images, Query image

Output: Matching score

Start:

Step-1: Database training.

Step-1.1: Select → Medical images.

Step-1.2: Group → Medical image (Ex: brain MRI, chest, Mammogram etc.).

Step-1.3: Extract → Feature (8 × 8 block division, CL coefficient and EH descriptor).

Step-1.4: Apply → SVM (medical images).

Step-1.5: Train ← Medical images.

Step-1.6: Save → Database.

Step-2.1: Select → Query image.

Step-2.2: Extract → Feature (8 × 8 block division, CL coefficient and EH descriptor).

Step-2.3: Retrieve → Matched images and matched score.

End:

The 8 × 8 block division performs the division of image into square blocks that helps in feature extraction from these blocks. The significance of this block division is that it offers high detection accuracy for the textured regions. However, 8 × 8 block division leads to high computational complexity as it searches features in better way (Figure 3).
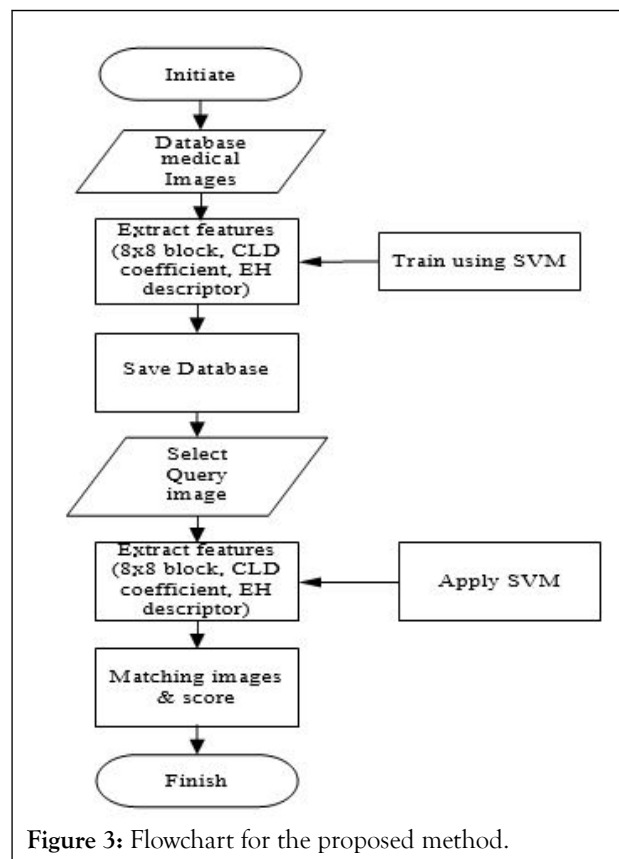

Figure 3: Flowchart for the proposed method.

The CLD mainly indicates the spatial distribution of colors with a few nonlinear quantized DCT coefficients of grid based average colors. Also, Gabor filter works as a band pass filter to distribute the local spatial frequency. The Edge Histogram (EH) descriptor is a texture defining method and is used in image retrieval. The EH descriptor classifies images on local edge distribution with the histogram-based descriptor. Later, the SVM is used to train these images and saved as specific database category.

Further, retrieval process is applied for query image and its features are extracted by using 8 × 8 block division, CL descriptor coefficient and Edge Histogram (EH) descriptor (Step 2). Once the retrieval is done, the matching images of the database are obtained and then matching score of the images are extracted. Flow chart of the proposed SVM-MIR is given Figure 3. The next section describes the results and analysis of proposed SVM-MIR approach [20].
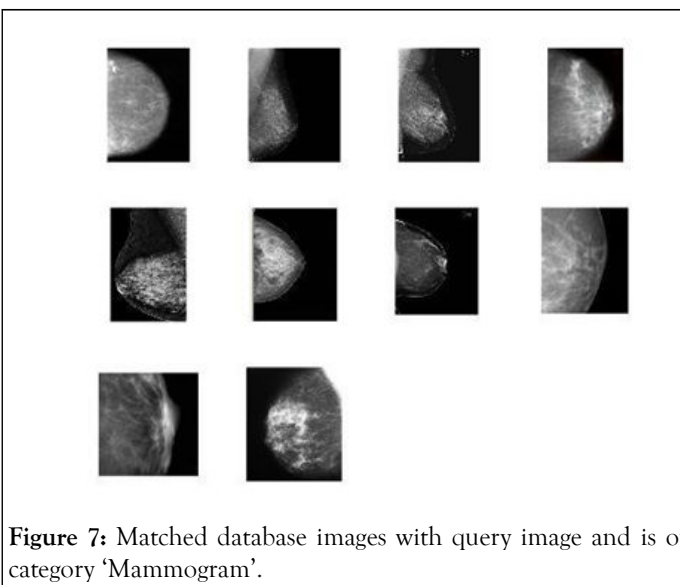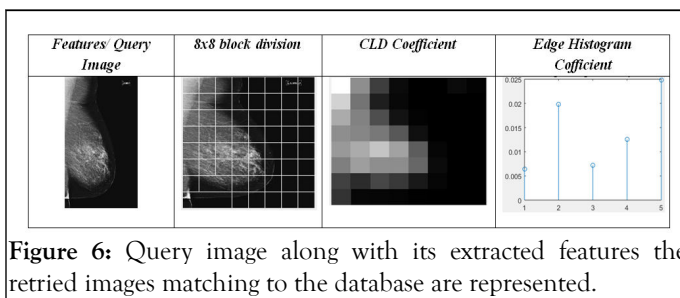
DISCUSSION

The obtained results from proposed SVM-MIR approach are discussed in this section (Figure 4).

Figure 4: Highlighted database images.
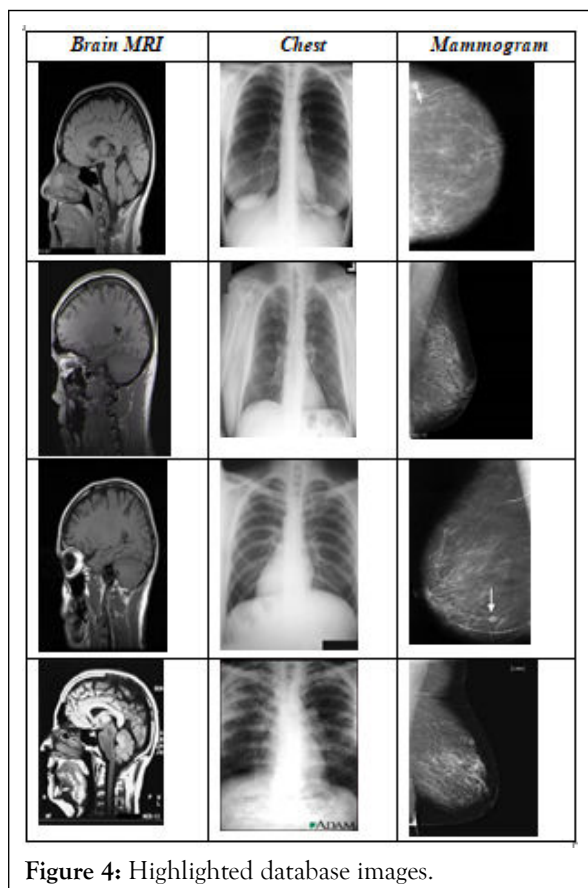
The results are extracted by training the database images and classified by using SVM approach. Considering the database of brain MRI, chest, and mammogram images for the training purpose total 12 images are taken. The database images are shown in Figure 4. (Highlighted only 4 images for representation). The database images are selected randomly and any database can be used to train using SVM-MIR approach and perform the classification of retrieval images of different category. Thus, the proposed approach can be useful for medical research. The next step of the approach is to perform feature extraction of similar contents. The extracted features like 8 × 8 block division, CLD coefficient and edge Histogram coefficient of database image is given in Figure 5.

Figure 5: Extracted Features of database image.
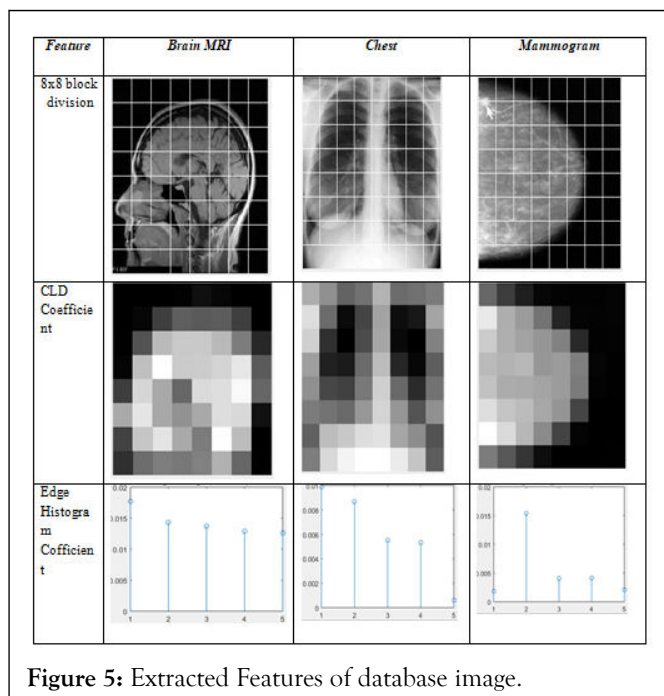
These features are named specific database category and trained then using SVM. Later, the retrieval process is followed by applying SVM and identified the matched images. The Figure 6 gives the query image along with its extracted features (Figure 7).

Figure 6: Query image along with its extracted features the retried images matching to the database are represented.

Figure 7: Matched database images with query image and is of category 'Mammogram'.

The matching score is of the query image with database images with similar texture, shape and color features is given in Figure 8.
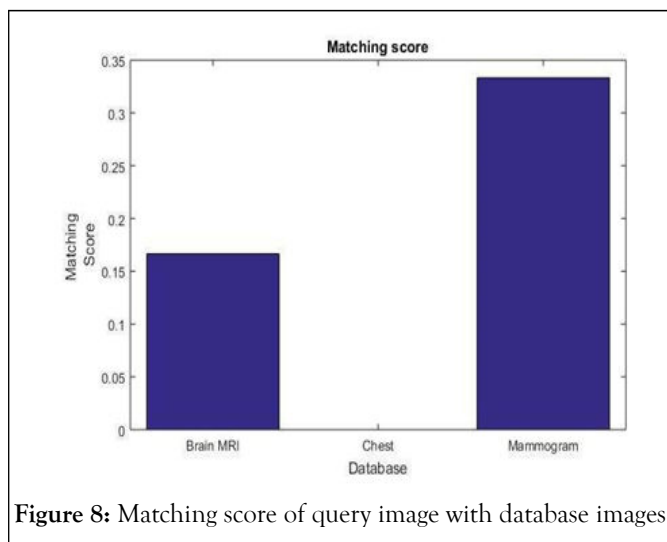
Figure 8: Matching score of query image with database images.

CONCLUSION

The objective of medical image databases is to give ineffective intends to organizing, seeking, and indexing extensive accumulations of retrieval of image. Hence, a SVM based image retrieval of medical image is introduced which connect arrangement to retrieval. In this system, the image classification data is used to sift through unessential images and alter the component weights in a direct blend of closeness coordinating. In general, this retrieval system is helpful as a front end for substantial medical image databases where an inquiry can be performed in different images for instructing, preparing and explore purposes.

ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Not applicable.

CONSENT FOR PUBLICATION

I give my consent for the publication of identifiable details, which can include photograph(s) and/or videos and/or case history and/or details within the text ("Material") to be published in the above journal and article.

AVAILABILITY OF DATA AND MATERIALS

Not applicable competing interests not applicable funding.

AUTHORS' CONTRIBUTIONS

In this work, the whole work has been done by Mrs. Ranjana Battur under the guidance of the guide Mr. Jagadisha N. The designing of the algorithm and the implementation of the algorithm has been done by Mrs. Ranjana Battur. Also the results have been obtained by Mrs. Ranjana Battur. The guide Mr. Jagadisha N has helps Mrs. Ranjana Battur to improve the paper and the results.

REFERENCES

1. Bhat S, Kumar A. Biomaterials and bioengineering tomorrow's healthcare. Biomatter. 2013;3(3):e24717.
2. Abramoff MD, Garvin MK, Sonka M. Retinal imaging and image analysis. IEEE Rev Biomed Eng. 2010;3:169-208.
3. Fan L, Zhang F, Fan H, Zhang C. Brief review of image denoising techniques. Visual Comput Indus Biomed Art. 2019;2(1):1-2.
4. Zhang D, Islam MM, Lu G. A review on automatic image annotation techniques. Pattern Recog. 2012;45(1):346-362.
5. Liu Y, Zhang D, Lu G, Ma WY. A survey of content based image retrieval with high level semantics. Pattern Recog. 2007;40(1):262-282.
6. Smeulders AW, Worring M, Santini S, Gupta A, Jain R. Content-based image retrieval at the end of the early years. IEEE Trans Pattern Analys Mach Intel. 2000;22(12):1349-1380.
7. Lee YH, Rhee SB, Kim B. Content-based image retrieval using wavelet spatial-color and Gabor normalized texture in multi-resolution database. IEEE. 2012;371-377.
8. Alraqibah HA, Ben Ismail MM, Bchir O. Empirical comparison of visual descriptors for content based x-ray image retrieval. Spring Int Publ. 2014;503-510.
9. Oussalah M. Content based image retrieval: Review of state of art and future directions. IEEE. 2008;1-10.
10. Tyagi V, Tyagi V. Content-based image retrieval techniques: A review. Content-based image retr: Idea Influ Curr Trend. 2017:29-48.
11. Rahman MM, Bhattacharya P, Desai BC. A framework for medical image retrieval using machine learning and statistical similarity matching techniques with relevance feedback. IEEE trans InformTechnol Biomed. 2007;11(1):58-69.
12. Li F, Dai Q, Xu W, Er G. Multi-label neighborhood propagation for region-based image retrieval. IEEE Trans Multimedia. 2008;10(8): 1592-604.
13. Li WJ. MILD: Multiple-instance learning *via* disambiguation. IEEE Transac Knowledg Data Eng. 2009;22(1):76-89.
14. Bian W, Tao D. Biased discriminant Euclidean embedding for content-based image retrieval. IEEE Transac Image Proces. 2009;19(2):545-54.
15. Blanchart P, Datcu M. A semi-supervised algorithm for auto-annotation and unknown structures discovery in satellite image databases. IEEE J Sel Top Appl. 2010;3(4):698-717.
16. Rahman MM, Antani SK, Thoma GR. A learning-based similarity fusion and filtering approach for biomedical image retrieval using SVM classification and relevance feedback. IEEE Transactions on Information Technology in Biomedicine. 2011;15(4):640-646.
17. Zhang X, Liu W, Dundar M, Badve S, Zhang S. Towards large-scale histopathological image analysis: Hashing-based image retrieval. IEEE Transactions on Medical Imaging. 2014;34(2):496-506.
18. Ramos J, Kockelkorn TT, Ramos I, Ramos R, Grutters J, Viergever MA, et al. Content-based image retrieval by metric learning from radiology reports: application to interstitial lung diseases. IEEE J Biomed Health Inform. 2014;20(1):281-292.
19. Zhu L, Shen J, Xie L, Cheng Z. Unsupervised visual hashing with semantic assistant for content-based image retrieval. IEEE Trans Knowl Data Eng. 2016;29(2):472-86.
20. Chaudhuri B, Demir B, Chaudhuri S, Bruzzone L. Multilabel remote sensing image retrieval using a semisupervised graph-theoretic method. IEEE Transac Geosci Remote Sen. 2017;56(2):1144-58.