

A Novel Architecture and Analysis of Challenges for Combining Text and Image for Medical Image Retrieval

Pradnya Kulkarni, Siddhivinayak Kulkarni, Andrew Stranieri
*Centre for Informatics and Applied Optimization
Federation University, Victoria, Australia*

Abstract

The introduction of digital information storage and internet technologies has changed diagnostic radiology practices significantly. Radiological support systems including PACS (Picture Archiving and Communication System), SRS (Structured Reporting System) and image and text retrieval systems have been developed to assist radiologists. In this article we analyse the integration of image and text retrieval is promising for enhancing radiological teaching, research and diagnosis. Presentation of similar images as diagnostic aids has proven to be helpful in domains of radiology such as chest radiographs, lung CT and mammograms. We propose a diagnostic support system based on a combination of image and text retrieval. Our approach has the potential to increase the radiologist's confidence in his/her diagnosis which could translate to greater accuracy and reduced health care costs.

1. Introduction

Different radiological support systems such as PACS, Structured Reporting Systems (SRS) and Content Based Image Retrieval (CBIR) have been developed in order to improve assessment accuracies and facilitate radiological assessments. Most PACS implementations provide basic search functionality for retrieving images based on a patient's identifier. However, the retrieval of medical images based solely on the patient's identifier is not sufficient for diagnostic applications as radiologists often prefer to retrieve images based on a specific disease in order to compare and contrast cases.

It is evident from the various studies that radiologists report benefits from the presentation of images of lesions with known pathology that are similar to a new unknown lesion, currently under investigation. The presentation of similar images as a diagnostic aid has proven to be helpful for chest radiographs. Presentation of images of lesions that were actually malignant though they looked benign, or images that were actually benign though they looked malignant caused confusion and diminished the usefulness of the retrieval of similar images.

This article provides an overview of the integration of image retrieval with text retrieval is more likely to be aligned with reasoning processes so

can be expected to be more useful to support radiologists when making assessments and may also enhance teaching, research, and diagnostic applications. However, significant challenges in the integration of text and image retrieval exist. We conclude by proposing a novel image retrieval system based on combining content based features of the image and the associated text.

Rest of the paper is organized as follows, Section 2 provides brief description regarding text retrieval, Section 3 details medical image retrieval, while Section 4 describes combination of text and image retrieval, proposed architecture is described in Section 5, challenges for image and text retrieval are described in Section 6 and finally concluded in Section 7.

2. Text retrieval

Various attempts have been made in order to solve the problem of retrieving images related to past cases based on text that makes up a radiology report. The RadMiner [1] is another such recent system which makes use of NLP techniques to parse the free text radiology reports.

The limitation of text retrieval approaches is that if the physician does not accurately record the features of interest when storing an image, the search for the image using keywords will not succeed. Although captions and full-text excerpts contain descriptions of figures/illustrations, they cannot effectively represent the semantic information in medical images, which are better perceived visually by human experts.

3. Medical image retrieval

Content based image retrieval (CBIR) is based on retrieving similar images by using image features such as colours, textures, shapes and objects. CBIR helps users to find similar images in multimedia applications. CBIR techniques have been successfully applied in various real world applications such as security, arts, entertainment, geographical information systems. This saves users' time considerably in browsing unstructured data and image related information. Radiology images pose challenges compared to multimedia images as they contain rich information, specific features that need to be carefully recognized for medical image

analysis. For example, Cysts and solid nodules usually have uniform internal density, whereas complex lesions have heterogeneous characteristics. This is why texture features are helpful in analysing and distinguishing this type of information. Various texture feature extraction techniques have been tried such as wavelets, Gabor filters and Co-occurrence matrices. Shape is another strongest factor in detecting a certain disease or lesion. Objects of interest such as anatomical structures or lesions are embedded in complex backgrounds and so automatic segmentation presents a great challenge. Single object based shape description has been successfully employed in several CAD applications such as CT colonoscopy and lung nodule detection.

In [2] training images are partitioned into 8×8 equal grid of 64 non-overlapping regions. Colour features composed of mean and standard deviation of each channel in RGB colour space and texture features composed of second order texture moments are used. Two different global thesauruses in the form of similarity and a correlation matrix where each element of the matrices defines concept similarities or co-occurrence relationships respectively are utilized in a quadratic form of distance measure to compare a query and database images. Results show that concept based feature representation performs much better compared to Colour layout descriptor (CLD) and Edge histogram descriptor (EHD). The quadratic similarity matching performs better when compared to Euclidean and Cosine similarity matching. However, a limitation of this system is the construction of the global matrix which is prohibitively difficult for large collections. Arakeri et al. [3] classify brain tumours based on the rotation invariant shape features (circularity and irregularity) and texture features from tumour region into benign and malignant classes and then use the wavelet based Fourier descriptors and local binary patterns for checking similarity. Shape similarity is calculated using Euclidean distance and the texture similarity is calculated using Chi-square distance.

In [4] developed a CT lung image retrieval system for assisting differential diagnosis. They found that most important feature of lung cancer is density (texture). This system uses fixed size rectangles to mark the Region of Interest (ROI) area. A three dimensional lung-based coordinate system, which was suggested by the radiologist, was constructed to describe the position information of selected blocks in order to eliminate the variance inherent in lung's size and shape. Major 2D FFT coefficients representing texture features of each marked region are then extracted. The centre information of the marked region is used for position matching. The experimental results show that in simple query, 96.6% of images can be correctly retrieved with the displacement up to 22% of the block size.

ASSERT proposes physician-in-the-loop approach in which the human delineates the pathology bearing regions (PBR) and a set of anatomical landmarks in the image when the image is entered into the database. Features such as Texture (grey level co-occurrence matrix), Shape features (longer axis, shorter axis Fourier descriptors and moments) and the position information of a PBR are used. Indexing is based on hashing function based on feature grouping. Results show that the retrieval efficiency is quite good [5].

medGIFT considers image as a document containing several visual words. The image is described in terms of a set of words by mapping the visual features derived from local and global colour or texture cues to these keywords. Thus the image retrieval problem is converted into a standard text based information retrieval. Around 1000-3000 features per image are extracted such as global colour histograms colour blocks, histogram of Gabor filters. Limitation of this system is it relies on the annotation tool. Lim and Chevallet [6] adopt a content description approach similar to medGIFT. A statistical learning framework is used to extract vocabularies of meaningful medical terms (VisMed terms) associated with visual appearance of image samples. Manually cropped image regions are used as a training set to statistically learn the terms Image retrieval in medical applications.

IRMA [7] as shown in Figure 1 aims at classifying the medical images according to image modality, body orientation, anatomic region and biological region and uses about 1024 Edge based and texture based features. Similarity is calculated for the whole region and hence similarity based on ROI is not possible.

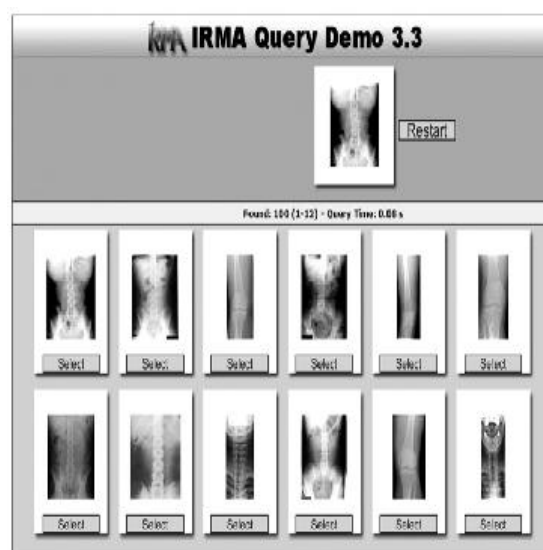


Figure 1: Image Retrieval for Medical Applications (IRMA)

Tao et al [8] developed a medical image search engine for diffuse parenchymal lung disease. The system combines offline learning aspect of classification based CAD systems with online learning aspect of CBIR systems. The Local Binary Patterns (LBP) algorithm is used to extract volumetric texture-based features. 3D bullae size histogram is also used. Patient-specific disease-specific features are used as supplement to cross-patient disease-specific features extracted offline to achieve better results.

Muller et al [9] integrated an image retrieval system into a radiology teaching file system, and the performance of the retrieval system was evaluated, with use of query topics that represent the teaching database well, against a standard of reference generated by a radiologist. The results of this evaluation indicate that content-based image retrieval has the potential to become an important technology for the field of radiology, not only in research, but in teaching and diagnostics as well.

CBIR have been deployed in teaching contexts by exporting interesting cases from PACS or viewing station. Casimage [10], myPACS are some of the example systems that offer maximum flexibility. Further a content based search could be used to compliment text-based methods of accessing data. Students could be able to browse the data in an easy way by using such systems.

4. Text and image retrieval

The problem with a text based image retrieval approach is that, different people can describe an image using different words or use the same word to describe different concepts. The problem with a CBIR approach is that, visually similar images may not be semantically similar. Moreover, in many situations; image content becomes meaningless without its associated textual description. Therefore it is beneficial to combine the content and text associated with it to retrieve meaningful images. Various approaches for combining text with the content of the image are taken in order to rank and retrieve more relevant images in multimedia. These are feature level fusion (Early fusion), Fusion based on similarity scores (Late fusion) and transmedia fusion where the idea is to use either image or text first to gather the relevant images and then to use the other (text representations of visually nearest neighbors).

Images in the radiology database are associated with certain cases or patients. Meta information such as date, imaging method, body part etc. is stored within an image according to the DICOM standard. Apart from metadata there is other textual information such as the symptoms, diagnosis, recommendations from expert, treatment history associated with the image which might be stored separately as a report.

Integrating this information with the actual content information for retrieval is likely to result in more accurate retrieval of the desired images. "Whereas initial image retrieval applications focussed on the image, modern applications are increasingly integrating the medical context into retrieval process such as mixing several images for cases and reports in the medical literature" [11].

A research project by Weyand [12] tries to add textual information retrieval to FIRE (Flexible information retrieval system) and achieve better retrieval results. Web MIRS [13] has targeted retrieval of X-ray images of human cervical and lumbar spines. An integrated image retrieval system to incorporate external knowledge encoded in lexicons, thesauri and ontologies is suggested [14]. Annotation and image markup (AIM) project aims at making the key semantic contents of images machine-readable using controlled terminologies.

Depeursinge et al [15] explain a system for case based retrieval for interstitial lung disease in high resolution CT images. A 3D map of lung tissue patterns with quantification of diseases based on texture analysis of lung parenchyma. Based on the portions of abnormal and normal lung tissue and clinical data of patient similar cases are retrieved. A system for retrieval of liver lesions which uses content based techniques and the annotations done by expert radiologist has worked well on a small dataset [16]. However, resources required for manual annotation is a limitation of this system. medGIFT combines the text case description with the content based part for better results. Due to vast amount of images generated, and the time constraint of radiologist, there is a potential for computer assisted interpretation and decision making.

4.1. Integration for teaching purposes

Case-based teaching files of actual patients are useful in diagnostic radiology education to illustrate pertinent teaching points (radiology teaching files). An CBIR-Text integrated approach would make it easier to search relevant images and cases quickly without having to browse through the huge database and the web based search approach would help study in self-learning mode and available at any place.

4.2. Integration for research purposes

By combining text-based and visual retrieval methods, researchers will conceivably have more options for choice of cases to include into research and studies. Using this approach, researchers would be able to find new correlations between visual information from an image and its diagnosis or textual description. It could also lead to discovery of new knowledge.

4.3. Integration for diagnostic purposes

Radiologists often like to take a case-based reasoning approach where they want to look at similar images to compare and contrast cases which help them in diagnosis. The current PACS implementations offer searching based on patient information or modality only. Integration of text and image retrieval will allow radiologists to search the vast image database based on a particular diagnosis, symptoms, recommendations, modality, organ, and content based similarity for example a cyst, nodule of particular shape, size, texture and position. A system integrating the content based image retrieval and text retrieval can save radiologists' time considerably in contrast to tedious, unstructured browsing.

Breast cancers are misinterpreted as benign and many patients with benign lesions are misinterpreted as being malignant and are sent for biopsy. By looking at similar cases, radiologist can be more confident about the diagnosis or recommendations he/she will make. Thus it can improve the diagnostic accuracy and hence may result in saving of health care cost and reducing patients' anxiety and morbidity.

More importantly radiologists can retrieve images if they are not confident about the diagnosis in a particular case. So this type of system is not something that they have to use but can become handy in complex cases.

Radiologists are trained to differentiate between the intensity variations in the image, but it is possible to have perception errors due to repeated exposure to these images. If they are presented with similar images based on an integration approach, they might be able to better differentiate malignant and benign regions.

As opposed to CAD systems which have to be designed separately for different domains of radiology, an integrated text and image retrieval system could be designed for broad usage thus reducing implementation and maintenance costs. Moreover, these systems are very easy to use and could be widely accepted as they will not require exhaustive training.

The search systems could incorporate a database of normal organs. In this type of approach search based on distance to normality could be used to decide whether the case is pathological or not.

5. Proposed System Architecture

Images from the PACS database will be exported and stored in our image database. The corresponding radiology reports associated with the images would also be exported from the radiology information system (RIS) and stored in our reports database.

Manual annotations of images are requested from the radiologist and are stored in our database.

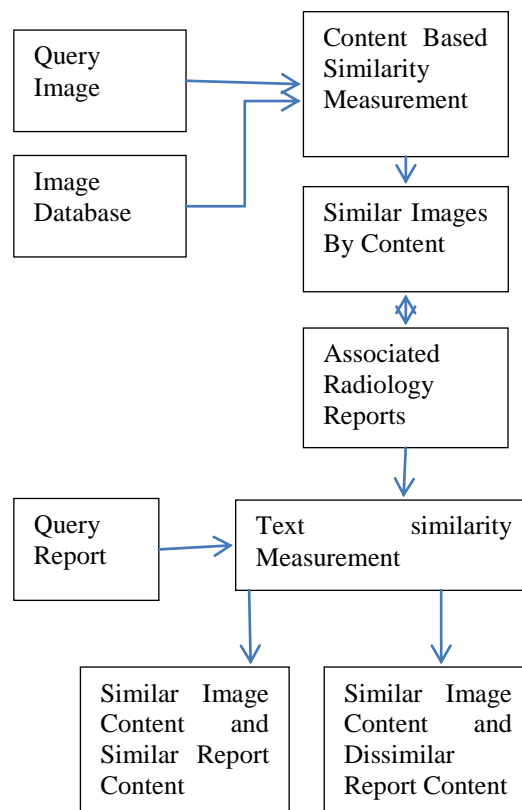


Figure 2: Block Diagram of the Proposed Integration of Image and Text retrieval

5.1 Image feature extraction

Various features from the region of interests (ROI) such as texture shape, boundaries will be extracted from every image in the database and will be stored in the image feature database.

5.2 Text feature extraction

Text features such as stop words, verbs, nouns, are extracted using NLP techniques for all the radiology reports in the database and stored in the text feature database.

5.3 Matching and presentation of results

When a radiologist is interpreting an image he will initially write a report based on his or her clinical judgement. The system will then extract the features from the region of interest from the query image and use these features to calculate similarity between the query image and the images in the

database. This step will provide us with the set of images similar in content with the query image.

The system will then parse the report written by the radiologist to get the query text features. In the second step the reports associated with the image results of the first step are taken into consideration. Query text features are compared with text features of the reports associated with the image results of the first step. This will provide us with the images which are similar in content with the query image and similar in reports to the query reports.

The system will then present the retrieved images and the associated reports in two groups. In each group the cases will be arranged in descending order of similarity.

1. Similar Cases:

The images which are similar in content with the query image and the reports are similar to the query report.

2. Dissimilar Cases:

The images which are similar in content with the query image but the reports are dissimilar to the query report.

We think that by looking at similar and dissimilar cases the radiologists' confidence will most likely be increased thus reducing chances of misinterpretation of images. We think that our system will also improve the quality of radiology reports.

We plan to use the CT lung images for our experiments, but the system could work in other domains of radiology with little or no modification. We intend to use a web based user interface so that it would be available from any place with internet access.

6. Challenges: image and text integration

In medical domain, we see that the image data becomes meaningless without its associated text. Research work in text/image information retrieval has shown that combining text and image information even with simple fusion strategies increase multimedia retrieval results [17]. The combination of textual with visual features or content and context of the images does have the greatest potential to lead to good results [13].

However, there are many challenges for the integration. Images are represented with low-level features while texts are represented by high level features. Mapping both type of information is challenging.

Content based image retrieval is computationally expensive. The Region of Interest needs to be calculated accurately by using segmentation algorithms or they have to be manually marked up by an expert. Having manual annotations is time consuming and getting the segmentation algorithm to appropriately extract ROI is again a challenging task.

The text associated with the image could be in different languages. Dealing with more than one language is another challenge in combining text and image retrieval.

Section 4 demonstrates that various systems have taken different approaches for integrating image and text retrieval (low level fusion, late fusion, transmedia fusion). In [18], it has been shown that either late fusion or transmedia fusion performs better than early fusion technique.

We propose a radiology support system for diagnostic purposes. Our system will present similar and dissimilar cases to the radiologists based on both the image content and the associated text in the radiology report. We intend to use the transmedia fusion technique for this system. This system is likely to assist radiologists to have more confidence in his or her diagnosis.

7. Conclusion

In this paper we have overviewed content based medical image retrieval, text retrieval and systems combining content based and text based retrieval. The image retrieval systems combining text and content of the image are much easier to use, can increase radiologist's confidence in diagnosis, and can save time when a radiologists need to view previous similar cases. These type of systems can also be used for research and teaching. Finally, we proposed a hybrid system combining image content and text from associated radiology reports which presents similar and dissimilar cases to help radiological assessment. But there are many challenges to integrate text and image for retrieval. These challenges will be overcome in the proposed architecture. Future research will incorporate use of radiologist's input as knowledge and skills for checking the working of prototype system.

8. References

- [1] Gerstmair, A., Daumke, P., Simon, K., Langer, M. and Kotter, E., (2012). Intelligent image retrieval based on radiology reports. *European radiology*, 22(12), pp. 2750.
- [2] Rahman, M., Antani, S. and Thoma, G. (2010). Local concept based medical image retrieval with correlation enhanced similarity matching based on global analysis, *Computer Vision and Pattern Recognition Workshops, IEEE Computer Society Conference, San Francisco, CA*, 87 – 94.
- [3] Arakeri, M. and Reddy, R., (2012). Medical image retrieval system for diagnosis of brain tumor based on classification and content similarity, *India Conference (INDICON), 2012 Annual IEEE 2012, Kochi, India* 416-421.

- [4] Liu, C., Tai, P., Chen, A., Peng, C., Lee, T., Wang, J. (2001). A content – based CT lung image retrieval system for assisting differential diagnosis image collection, IEEE International Conference on Multimedia and Expo, Tokyo, Japan. 174 – 177.
- [5] Shyu, C. R., Brodley, C., Kak, A., Kosaka, A., Aisen, A. and Broderick, L. (1999) ASSERT, A physician-in-the-loop content-based image retrieval system for HRCT image databases”, Computer Vision and Image Understanding, 75(½) 111-132
- [6] Lim J, Chevallet J. (2005). Vismed. A visual vocabulary approach for medical image indexing and retrieval. Information Retrieval Technology, Lecture notes in computer Science, Vol. 3689, pp. 84-96.
- [7] Lehmann, T., Guld, M., Thies, C., Plodowski, B. Keyers, D., Ott, B., Schubert, H. (2004). IRMA--content-based image retrieval in medical applications. Stud Health Technol Inform 107(2). 842-846.
- [8] Tao Y, Zhou X, Bi, J., Jerebkoa, A., Wolf, M., Salganicoff, M., Krishnana A. (2009). An adaptive knowledge-driven medical image search engine for interactive diffuse parenchymal lung disease quantification. Medical imaging. computer-aided diagnosis, vol 7260. SPIE, p 726007.
- [9] Muller, H., RosstE, A., Garcia, A., Vall_ee, J. and Geissbuhler, A. (2005). Benefits from content based visual data access in radiology. RadioGraphics, 25, pp. 849-858.
- [10] Muller, H., Rosset, A., Vallee, J. and Geissbuhler, A. (2003). Integrating content-based visual access methods into a medical case database, Medical Informatics Europe conference 2003, pp. 480-485.
- [11] Müller, H., Greenspan, H. (2012). Overview of the Third Workshop on Medical Content-Based Retrieval for Clinical Decision Support. MCBR-CDS 2012. 1-9
- [12] Weyand, T. (2005). Combining content-based image retrieval with textual information retrieval. RWTH Aachen.
- [13] Antani, S., Long, L. and Thoma, G. (2002). A biomedical information system for combined content-based retrieval of spine X-ray images and associated text information. Third Indian Conference on Computer Vision, Graphics and Image Processing
- [14] Zhou, X., Zillner, S., Moeller, M., Sintek, M., Zhan, Y., Krishnan, A. and Gupta, A. (2008). Semantics and CBIR. a medical imaging perspective, International conference on content-based image and video retrieval 2008, pp. 571-580.
- [15] Depeursinge, A., Vargas, A., Gaillard, F., Planton, A., Geissbuhler, A., Poletti, P., Muller, H. (2011). Case-based lung image categorization and retrieval for interstitial lung disease. Clinical workflows, International Journal of Computer Assisted Radiology and Surgery, 7(1), 97-110
- [16] Napel, S., Beaulieu, C., Rodriguez, C., Cui, J., Xu, J., Gupta, A., Korenblum, D., Greenspan, H. and Rubin, D. (2010). Automated retrieval of CT images of liver lesions on basis of image similarity. Method and preliminary results. Radiology, 256(1).
- [17] Clinchant, S., Ah-pine, J. and Csurka, G. (2011). Semantic combination of textual and visual information in multimedia retrieval. 1st ACM International Conference on Multimedia Retrieval. .
- [18] Müller, H. and Depeursinge, A. (2010). Fusion Techniques for Combining Textual and Visual Information Retrieval. In. H. Müller, P. Clough, T. Deselaers and B. Caputo, eds, ImageCLEF The information Retrieval Series. 32 edn. Springer, pp. 95-114.