Fusion Techniques in Biomedical Information Retrieval

Alba García Seco de Herrera and Henning Müller

Abstract For difficult cases clinicians usually use their experience and also the information found in textbooks to determine a diagnosis. Computer tools can help them supply the relevant information now that much medical knowledge is available in digital form. A biomedical search system such as developed in the Khresmoi project (that this chapter partially reuses) has the goal to fulfil information needs of physicians. This chapter concentrates on information needs for medical cases that contain a large variety of data, from free text, structured data to images. Fusion techniques will be compared to combine the various information sources to supply cases similar to an example case given. This can supply physicians with answers to problems similar to the one they are analyzing and can help in diagnosis and treatment planning.

1 Introduction

Clinicians generally base their decisions for diagnosis and treatment planning on a mixture of acquired textbook knowledge and experience acquired through real–life clinical cases [39]. Therefore, in the medical field, two knowledge types are generally available [32]:

- **explicit knowledge**: to the already well established and formalized domain knowledge, e.g., textbooks or clinical guidelines;
- **implicit knowledge**: individual expertise, organizational practices and past cases.

Alba García Seco de Herrera
University of Applied Sciences Western Switzerland (HES–SO)
TechnoArk 3, 3960 Sierre, Switzerland, e-mail: alba.garcia@hevs.ch

Henning Müller
University of Applied Sciences Western Switzerland (HES–SO)
TechnoArk 3, 3960 Sierre, Switzerland, e-mail: henning.mueller@hevs.ch

When working on a new case that includes images, clinicians analyse a series of images together with contextual information, such as the patient age, gender and medical history as these data can have an impact on the visa appearance of the images. Since related problems may have similar solutions, clinicians use past situations similar to the current one to determine the diagnosis and potential treatment options, information that is also transmitted in teaching, where typical or interesting case are discussed, and used for research [32, 52]. Thus, the goal of a clinician is often to solve a new problem by making use of previous similar situations and by reusing information and knowledge [1], also called case–based reasoning. The problem can be defined in four steps, known as the four 'res' [16, 32]:

1. Retrieve the most similar case(s) from the collection;
2. Reuse them, and more precisely their solutions, to solve the problem;
3. Revise the proposed solution;
4. Retain the current case in the collection for further use.

In this chapter, we focus on the retrieval step because the retrieval of similar cases from a database can help clinicians to find the needed information [39, 45]. In the retrieval step a search over the documents in the database is performed using the formulation of the information need that can include text and images or image regions. Relevant documents are ranked depending on the degree of similarity to a given query case or the similarity to the information need. The most relevant cases are then proposed on the top of the list and can be used to solve the current problem [4].

Text analysis and retrieval has been successfully used in various medical fields from lung disease, through cardiology, eating disorders, to diabetes and Alzheimer's disease [25]. Text in the anamnesis are often the first data available, and based on the initial analysis other exams are ordered.

In addition to the text in the anamnesis, another initial data source for diagnosis are the images [52]. Visual retrieval has become an important research area over the past more than 15 years also for medical applications [45]. In the past, the most common visual descriptors used for visual retrieval systems were the color histograms, texture features such as Gabor filters and simple shape measures [45]. In recent years visual words have had most often the best results in object recognition or image retrieval benchmarks [18] and have become the main way of describing images with a variety of basic features such as SIFT (Scale Invariant Feature Transform) [28] and also texture or color measures.

In terms of medical cases, images are always associated with either text or structured data and this can then be used in additional to the visual content analysis for retrieval. Most often text retrieval has much better performance than visual retrieval, describing the context in which the images were taken. Furthermore, there is an evidence that the combination or fusion of information from textual and visual sources can improve the overall retrieval quality [17, 27]. Whereas visual retrieval usually has good early precision and low recall, text retrieval generally has a high recall.

Combination of image and text search can be done as follows [15]:

- Combine results (ranked lists) of visual and text retrieval for the final results;
- Use visual retrieval to rerank results lists of text retrieval;

- Use text retrieval to rerank results lists of visual retrieval;
- Use image analysis and classification to extract relevant information from the images (such as modality types, anatomic regions or the recognition of specific objects in the images such as arrows) to filter results lists or rerank them.

In 2013, the Center of Informatics and Information Technology group CITI presented the Nova MedSearch [1] as a medical multimodal (text and image) search engine that can retrieve either similar images or related medical cases [33]. Case–based retrieval taking into account several images and potentially other data of the case has also been proposed by other authors over the past 7 years [37, 52]. Due to the many challenges in biomedical retrieval, research has been attracting increasing attention, and many approaches have been proposed [27].

The remainder of the chapter is organized as follows. Section 2 describes the text and visual retrieval and discusses several fusion approaches. The biomedical task and a evaluation framework are presented in Section 3. In Section 4 the Khresmoi system is presented as well as the experiments carried out on existing fusion techniques to combine multiple sources. Finally, conclusions are given in Section 5.

2 Visual and text information retrieval

To search through the large amount of data available there is a need for tools and techniques that effectively filter and automatically extract information from text and visual information. Text–based and visual–based methods [22] can in our scenario be used for the retrieval.

2.1 Text retrieval

Most biomedical search engines, also systems searching for images, have been based on text retrieval, only. Sources of biomedical information can be scientific articles and also reports from the patient record [47]. The various parts of the text such as title, abstract, figure captions can then be indexed separately. Some examples for general search tools that have also been used in the biomedical domain are the Lucene, Essie or Terrier information retrieval (IR) libraries. Lucene[2] is a open source full–text search engine. The advantage of Lucene is its simplicity and high performance [31]. Lucene was chosen for the experiments shown in Section 4 because it is fast and easy to install and use. Essie [23] is a phrase–based search engine with term and concept query expansion and probabilistic relevancy ranking. It was also designed to use terms from the Unified Medical Language System

[1] http://medical.novasearch.org/

[2] http://lucene.apache.org/

(UMLS).Terrier[3] is also an open source platform for research and experimentation in text retrieval developed at the University of Glasgow. It supports most state of the art retrieval models such as Dirichlet prior language models, divergence from randomness (DFR) models or Okapi BM25.

2.2 Visual retrieval

Users of biomedical sources are also often interested in images for biomedical research or medical practice [38], as the images carry an important part of the information in articles. Rather than using text queries, in content–based image retrieval systems, images are indexed and retrieved based on their visual content (image features) such as color, texture, shape and spatial location of image elements. This allows to use visual information to find images in a database similar to examples given or with similar regions of interest. Figure 1 shows examples of the visual information that can be extracted from the images.

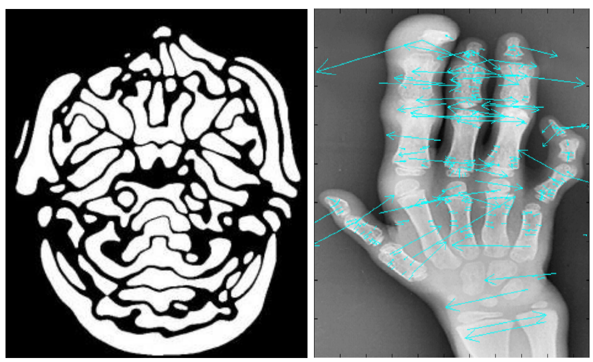


Fig. 1 Shape and SIFT (Scale Invariant Feature Transform) information can be extracted from the visual content of the images. In the left, the regions detected by a key–region detector are shown. In the right, the arrows represent the center, scale and orientation of the keypoints detected by the SIFT algorithm.

The most commonly used features for visual retrieval can be grouped into the following types [22]:

- **Color**: Several color image descriptors have been proposed [5] such as simple color histograms, a color extension to the Scale Invariant Feature Transform (SIFT) [42] or the Bag–of–Colors [19];
- **Texture**: Texture features have been used to study the spacial organization of pixel values of an image like first order statistics, second order statistics, higher order statistics and multiresolution techniques such as wavelet transform [43].

[3] http://terrier.org/

- **Shape**: Various features have been used to describe shape information, including moments, curvature or spectral features [53].

Some systems, such as img(Anaktisi)[4], FIRE[5] (Flexible Image Retrieval Engine) or LIRE[6] (Lucene Image REtrieval), allow content–based image retrieval by various visual descriptors and various combinations of descriptors.

In the following we present some of the processing steps that can potentially improve the retrieval quality of images from the biomedical literature when fusing them with text and/or visual retrieval, particularly for retrieval from the biomedical literature: region–of–interest (ROI) identification, image classification and multi–panel figure separation methods.

2.2.1 Region–of–interest identification

Annotations in images such as arrows are frequently used in images in the biomedical literature (see Figure 2). If the marked regions can then be linked with text describing the images, this can be used for retrieval of focused parts of image [40], so retrieving the regions of interest and not entire images. Several approaches have been used in the literature. For instance, Cheng et al. [10] segmented arrow candidates by a global thresholding–based method followed by edge detection. Also Seo et al. [44] developed a semantic ROI segmentation. An attention window is created and a quad–tree based ROI segmentation is also applied to remove meaningless regions.

2.2.2 Image categorization

In the biomedical literature images can be of several types, some of which correspond to medical imaging modalities such as ultrasound, magnetic resonance imaging (MRI), X–ray and computer tomography (CT) (see examples in Figure 3). Detecting the image type automatically can help in the retrieval process to focus for example on one modality or to remove non clinical images entirely from the retrieval. Image categories can be integrated into any retrieval system to enhance or filter its results [49], improving the precision of the search [24] and reducing the search space to a set of relevant categories [41]. Furthermore, classification methods can be used to offer adaptive search methods [51]. Using image types as a filter is often requested by clinicians as an important functionality of a retrieval system [30]. Some web–accessible retrieval systems such as Goldminer[7] or Yottalook[8] allow users to filter the search results by modality [36].

[4] http://orpheus.ee.duth.gr/anaktisi/
[5] http://thomas.deselaers.de/fire/
[6] http://www.lire-project.net/
[7] http://goldminer.arrs.org/
[8] http://www.yottalook.com/

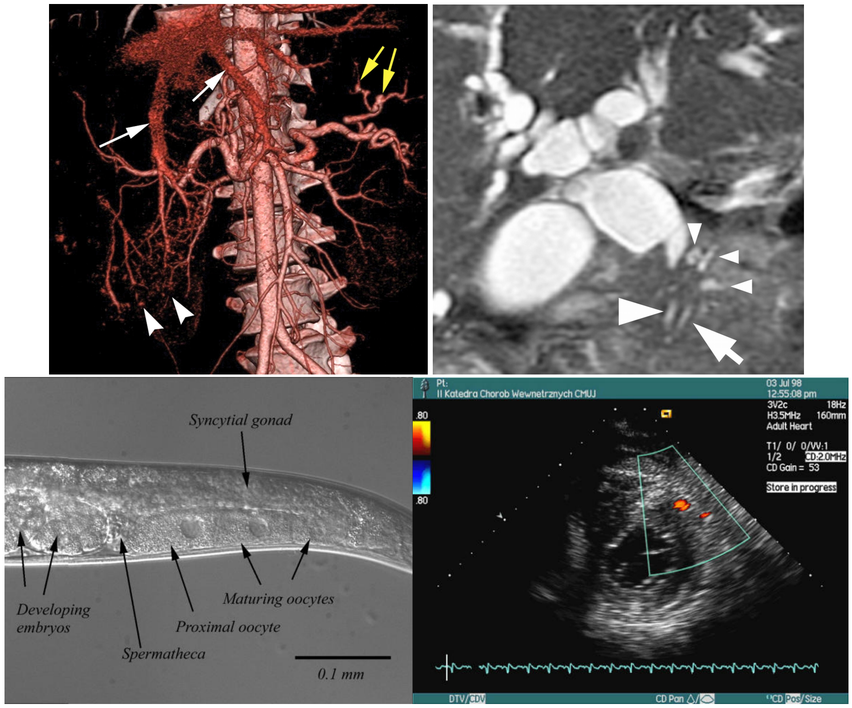
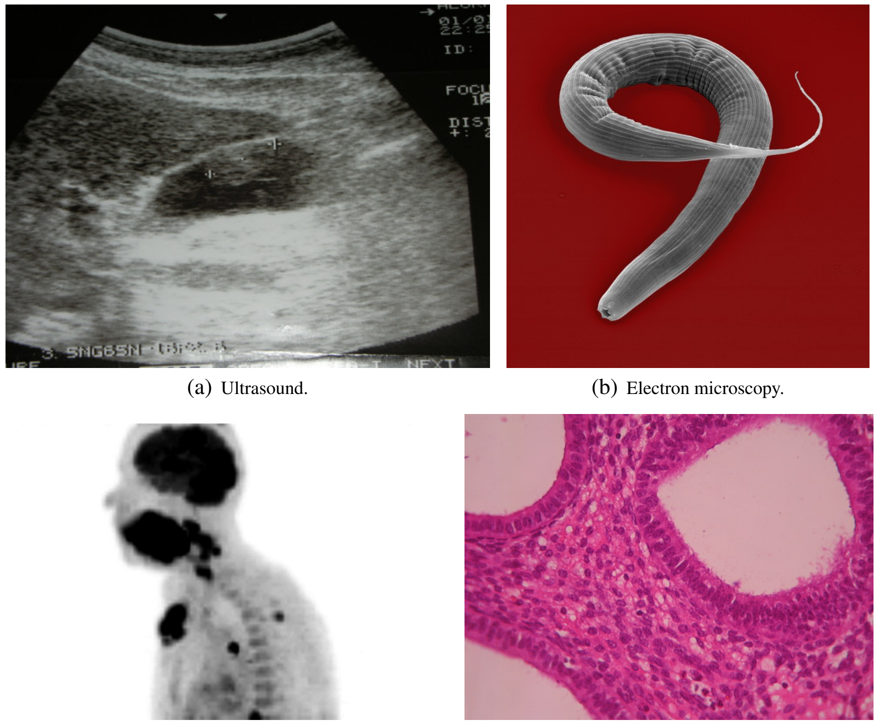


Fig. 2 Examples of images with annotations found in the ImageCLEFmed database. Annotations emphasize specific regions of the image according to special attributes of the highlight region such as lesions or structures important for the case.

ImageCLEF[9] proposes a hierarchy of image types for document images occurring in the biomedical open access literature [17], Figure 4 shows the proposed hierarchy. For more details on the ImageCLEF campaign see Section 3.2. Once the image type information is extracted, the predicted types can be integrated into the search results to generate a final result list. Information on image types can be used in various ways in the retrieval. The following approaches have been used to integrate the the classification into the results [49]:

- **Filtering**: Discarding the images of which the predicted type is different to the query. Thus, when filtering using the image type only potentially relevant results are considered;
- **Reranking**: Reranking the initial results with the image type information. The goal is to to improve the retrieval ranking by moving relevant documents towards the top of the list based on the categorization;
- **Score fusion**: Fusing a preliminary retrieval score S_R with an image classification score S_M using a weighted sum: $\alpha \cdot S_T + (1-\alpha) \cdot S_M$, where S_R and S_T are

[9] http://imageclef.org/



(a) Ultrasound.

(b) Electron microscopy.

(c) Positron emission tomography (PET).

(d) Light microscopy.

Fig. 3 Examples of images of various types that can be found in the biomedical literature.

normalized. This approach allows to adjust the parameter α to emphasize the retrieval score or the categorization results.

2.2.3 Compound figure separation

Compound or multi–panel figures (figures consisting of several sub figures) constitute a very large proportion of the images found in the biomedical literature. Image retrieval systems should be capable of distinguishing the parts of compound figures that are relevant to a given query. Compound figure separation is therefore a required first step to retrieving focused figures [17]. Figure 5 contains several examples of compound figures.

Several approaches have been published for separating figures from text in scanned documents [12] and specifically, for separating compound figures in the biomedical literature [2, 9, 11]. Chhatkuli et al. [11] proposed a compound figure separation technique based on systematic detection and analysis of uniform space gaps. Demner–Fushman et al. [13] determined if an image contains homogeneous regions that cross the entire image. An hybrid clustering algorithm based on particle

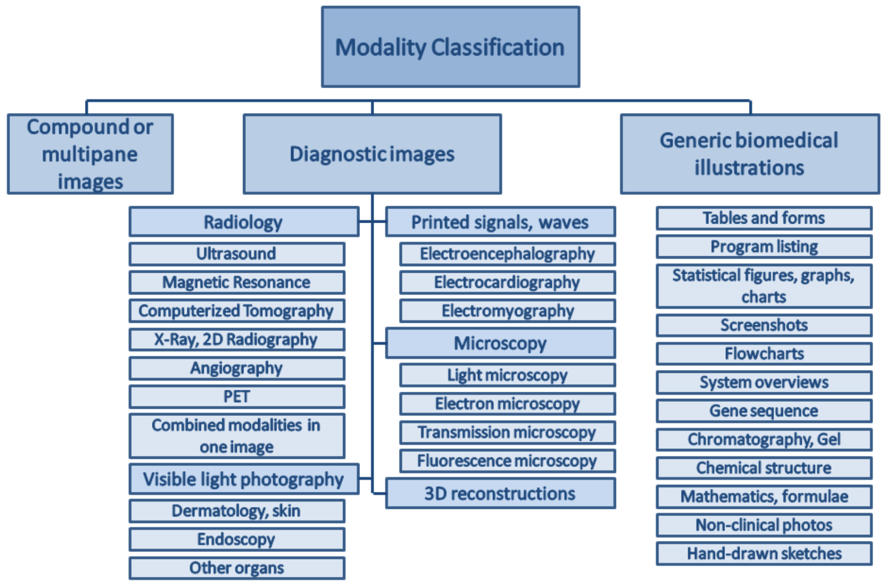


Fig. 4 The image class hierarchy proposed in ImageCLEFmed 2013 campaign for the image classification task.

swarm optimization with a fuzzy logic controller was presented by Cheng et al. [9] to locate related figure components. Using a figure and its associated caption, Apostolova et al. [2] determined if the figure consisted of multiple panels to then separate the panels and the corresponding caption part.

2.3 Fusion of multimodal features

To combine visual and text search several fusion techniques can be used. Such combinations can lead to better results than single modalities. Text retrieval often has much better performance than visual retrieval in medical retrieval [17], therefore the right combination strategies need to be chosen to really improve performance. This section describes several approaches for information fusion that have been used in the past [14]. To combine the results/features of multiple query images into a single ranked list two main fusion strategies were used depending on how the multiple results from the feature extraction are integrated: early and late fusion. Early fusion integrates unimodal features before making any decision (see Figure 6). Since the decision is then based on all information sources, it enables a truly multimodal feature representation [48]. Unimodal feature vectors are concatenated into one vector

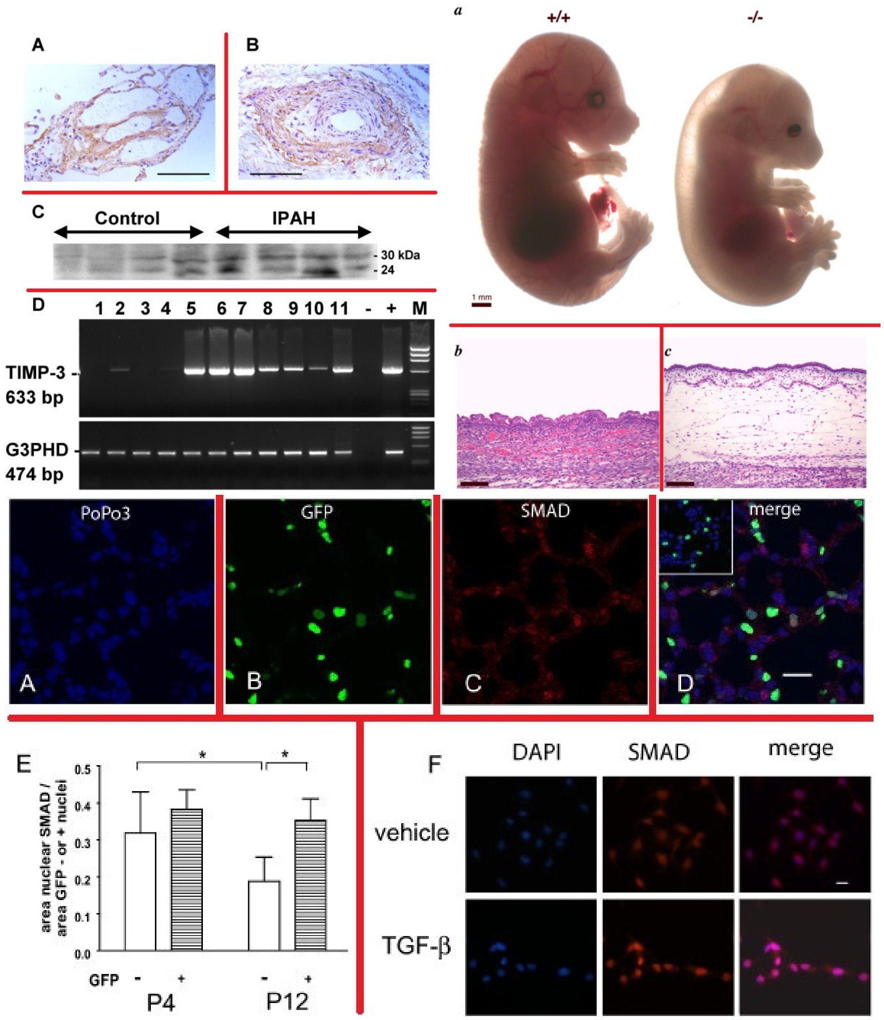


Fig. 5 Examples of compound figures found in the ImageCLEFmed database. These examples show mixed modalities in a single figure and several images from the same modality in the same figure. Red lines separate the subfigures.

using a weighting scheme. Rocchio's algorithm can also be applied to merge the vectors of the same feature spaces into a single vector.

$$\mathbf{q}_m = \alpha \mathbf{q}_o + \beta \frac{1}{|I_r|} \sum_{\mathbf{i}_j \in I_r} \mathbf{i}_j - \gamma \frac{1}{|I_{nr}|} \sum_{\mathbf{i}_j \in I_{nr}} \mathbf{i}_j$$

where α, β and γ are weights, $\mathbf{i}_m$ is the modified query, $\mathbf{i}_o$ is the original query, I_r is the set of relevant documents/images and I_{nr} is the set of non–relevant documents/images. Only the second term of the right part of the equation is used to merge vectors when non–relevant documents/images are done [18].

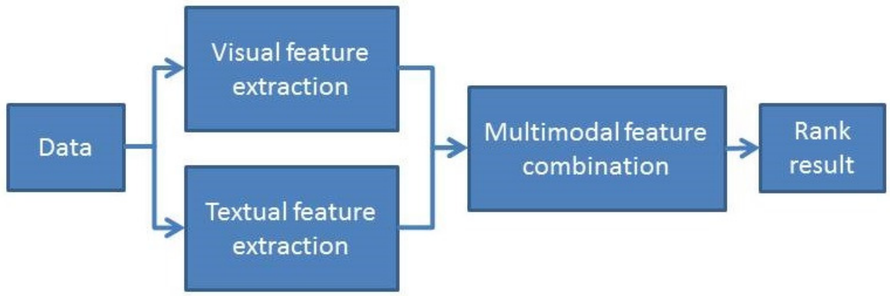


Fig. 6 General scheme for early fusion.

Late fusion consists of a combination of independent results from various approaches, e.g., text and visual approaches. The ranked lists of retrieval results are fused and not the features (see Figure 7).

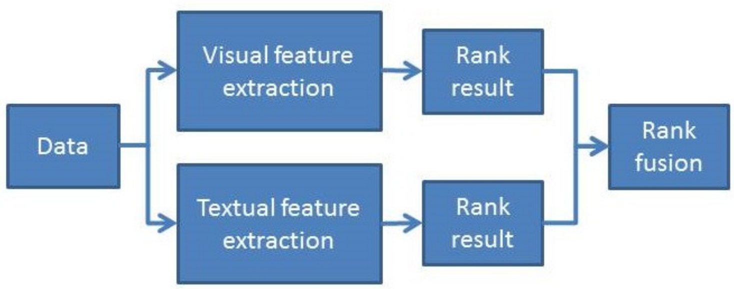


Fig. 7 General scheme for late fusion.

Two main categories of late fusion techniques exist based on which information is used, namely score–based and rank–based methods. In order to obtain a final ranking of a document d fusion techniques are required to reorder documents based on various descriptor lists. An overview of fusion techniques commonly used for the biomedical domain is given below:

- **Score–based methods**:
  - Linear combination

$$LN(d) = \alpha S_t(d) + \beta S_v(d)$$

where S_t and S_v are the textual and visual scores of the document d;

– combSUM

$$combSUM(d) = \sum_{j=1}^{N_j} S_j(d)$$

with N_j being the number of descriptors to be combined and $S(i)$ is the score assigned to document d;

– combMNZ

$$combMNZ(d) = F(d) * combSUM(d)$$

where $F(d)$ is the frequency of document d being returned by one input system with a non–zero score;

– combMAX

$$combMAX(d) = \arg\max_{j=1:N_j}(S_j(d))$$

– combMIN

$$combMIN(d) = \arg\min_{j=1:N_j}(S_j(d))$$

– combPROD

$$combPROD(d) = \prod_{j=1}^{N_j} S_j(d)$$

- **Rank–based methods**:
  – Reciprocal rank fusion:

$$RRF\texttt{score}(d) = \sum_{r \in R} \frac{1}{k + r(d)}$$

where R is the set of rankings assigned to the documents;

– Borda

$$Borda(d) = \sum_{r \in R} r(d)$$

3 Biomedical retrieval

In this section, a biomedical retrieval scenario is investigated. An evaluation framework for biomedical retrieval systems is proposed by ImageCLEFmed and this chapter uses the same framework to make results comparable with the state of the art.

3.1 Medical scenario

A biomedical retrieval system should correspond to real practical informations and be evaluated based on a corresponding scenario. Several user surveys and analyses of search log files have been done for obtaining the place of text and visual information in retrieval, mainly in radiology [34, 35, 50]. Based on these user analyses the tasks in ImageCLEFmed were developed. Particularly radiologists frequently search for images and have a need to search for visual abnormalities linked to specific pathologies. Usually not the entire image is of interest but rather small regions of interest [46]. In the past text retrieval for these tasks has obtained much better information that visual retrieval [17], but combinations can profit from the advantages of the two.

Currently, the used of only visual information still achieves low retrieval performance in this task and the combination of text and visual search is improving and seems promising [17].

3.2 ImageCLEFmed: an evaluation framework

ImageCLEF[10] is the image retrieval track of the Cross Language Evaluation Forum (CLEF)[11] [6]. One of the main goals of the medical task of ImageCLEF (ImageCLEFmed) [17] is to investigate the effectiveness of combining text and images for medical image– and case–based retrieval [14]. Several tasks have been proposed over the years since 2004, always in a very end user oriented way based on surveys or log files analyses. In 2013, four tasks were organized:

- Image–based retrieval;
- Case–based retrieval;
- Modality classification;
- Compound figure separation.

The image–based retrieval task has been running since 2004 with changing databases. The goal of this task is to retrieve images for a precise information need expressed through text and example images. Figure 8 shows one of the 35 topics distributed to the participants in 2013.

The case–based task was first introduced in 2009. This task aims to retrieve cases that are similar to the query case and are useful in differential diagnosis. Each topic consists of a case description with patient demographics, limited symptoms and test results including imaging studies (but not the final diagnosis). An example of a topic can be seen in Figure 9.

Since 2010, the modality classification task has been running. The goal of this task is to classify images into image types that can be medical modalities or other

[10] http://imageclef.org/

[11] http://www.clef-initiative.eu/

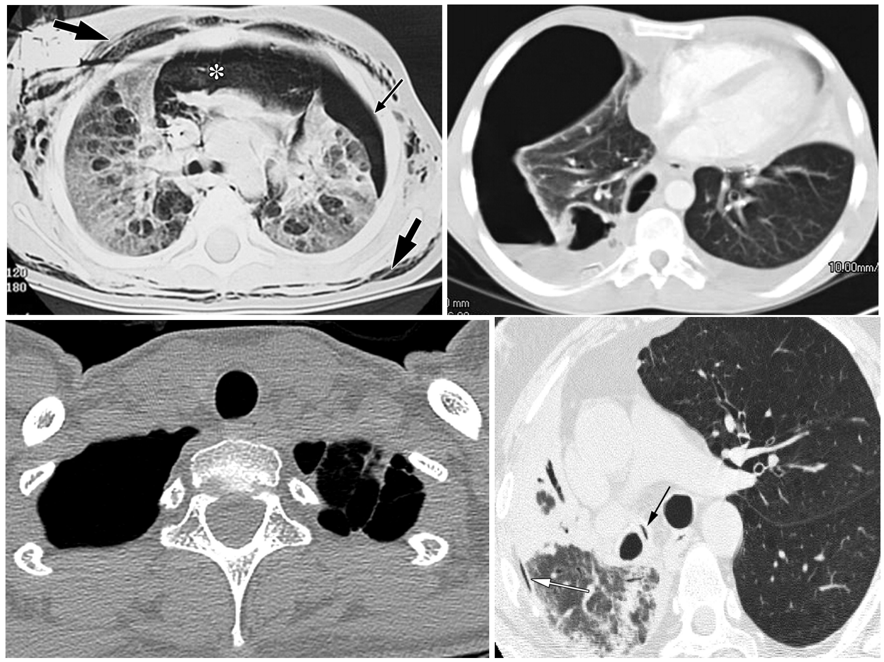


Fig. 8 Images from one of the topics in the image–based retrieval task of ImageCLEFmed 2013. They correspond to the textual query "pneumothorax CT images" that is also expressed in French, German and Spanish.

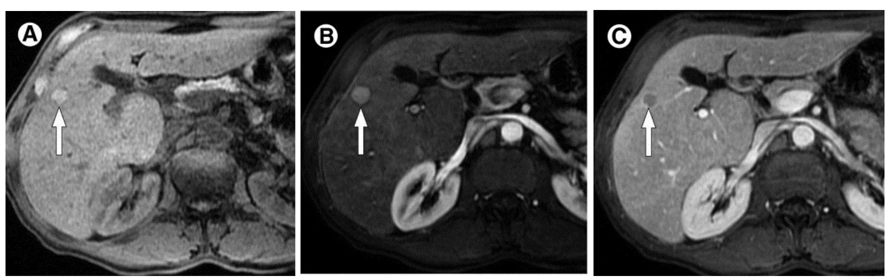


Fig. 9 Images from one of the topics in the case–based retrieval task of ImageCLEFmed 2013. They correspond to the textual query "A 56–year–old woman with Hepatitis C, now with abdominal pain and jaundice. Abdominal MRI shows T1 and T2 hyperintense mass in the left lobe of the liver which is enhanced in the arterial phase".

types occurring in the biomedical literature. More information can be found in Section 2.2.2.

In 2013, a compound figure separation task was added as a large portion of images in the literature turn out to be compound figures.

The ImageCLEFmed evaluation framework gives access to the tools developed for the described tasks including databases and ground truth. These tools were used to conduct the fusion experiments presented in Section 4.

4 Khresmoi and evaluation of fusion techniques

To search through large amounts of biomedical data, the Khresmoi[12] project is developing a multilingual multimodal search and access system for medical and health information and documents [3].

In this section, the work on text and visual fusion as part of Khresmoi is presented. More on the employed fusion techniques can also be found in [21]. The experiments use the ImageCLEFmed 2013 database of the case–based task described in Section 3.2.

For text retrieval, Apache Lucene was used (see Section 2.1). The results achieved with this approach on the case–based task of ImageCLEF is shown in Table 1.

Table 1 Results of the approaches at the ImageCLEF case–based retrieval task when using only text retrieval.

Run	MAP	Bpref	P10	P30
Best textual ImageCLEF run	0.2429	0.2417	0.2657	0.1981
Textual	0.1791	0.1630	0.2143	0.1581

For visual retrieval, a combination of the following descriptors were extracted to incorporate color and texture information from the images [20]:

- color and edge directivity descriptor (CEDD) [7];
- bag of visual words using SIFT, Scale Invariant Feature Transform, (BoVW) [28];
- fuzzy color and texture histogram (FCTH) [8];
- bag of colors (BoC) [19];
- BoVW with a spatial pyramid matching [26] (BoVW–SPM);
- BoC with $n \times n$ spatial grid (Grid BoC).

To enhance visual retrieval several fusion strategies described in Section 2.3 were tested to combine results of each of the query images and of several visual descriptors of the same image. Table 2 shows the results of the visual retrieval using this combination of fusion rules.

[12] http://www.khresmoi.eu/

Table 2 Results of the approaches for the ImageCLEF case–based retrieval task when using various fusion strategies for visual retrieval. Query and descriptor fusion is combined.

Queries f.	Descriptors f.	MAP	Bpref	P10	P30
Best visual ImageCLEF run		0.0281	0.0335	0.0429	0.0238
Rocchio	Borda	0.0004	0.0092	0	0
Rocchio	combMAX	0.0004	0.0096	0	0.0029
Rocchio	combMIN	0.0002	0.0093	0	0.0019
Rocchio	combMNZ	0.0008	0.0084	0.0029	0.0048
Rocchio	combSUM	0.0006	0.0084	0.0029	0.0038
Rocchio	RRF	0.0005	0.0085	0	0.0038
Borda	Borda	0.0005	0.0060	0	0.0019
Borda	combMAX	0.0004	0.0066	0	0.0019
Borda	combMIN	0.0002	0.0124	0	0
Borda	combMNZ	0.0009	0.0055	0.0029	0.0038
Borda	combSUM	0.0005	0.0060	0.0029	0.0029
Borda	RRF	0.0012	0.0061	0.0086	0.0057
combMAX	Borda	0.0006	0.0062	0.0066	0.0019
combMAX	combMAX	0.0006	0.0089	0.0057	0.0029
combMAX	combMIN	0.0003	**0.0156**	0	0.0019
combMAX	combMNZ	0.0036	0.0077	**0.0114**	0.0057
combMAX	combSUM	0.0021	0.0077	0.0086	**0.0067**
combMAX	RRF	0.0013	0.0066	0.0086	0.0048
combMIN	Borda	0.0005	0.0077	0.0029	0.0029
combMIN	combMAX	0.0006	0.0091	0.0086	0.0038
combMIN	combMIN	0.0003	0.0172	0	0.0019
combMIN	combMNZ	0.0032	0.008	0.0086	0.0057
combMIN	combSUM	0.0015	0.0079	0.0057	0.0057
combMIN	RRF	0.0011	0.0060	0.0086	**0.0067**
combMNZ	Borda	0.0005	0.0061	0.0029	0.001
combMNZ	combMAX	0.0004	0.0077	0	0.0038
combMNZ	combMIN	0.0001	0.0111	0	0.001
combMNZ	combMNZ	0.0029	0.0058	0.0086	**0.0067**
combMNZ	combSUM	0.0011	0.0053	0.0057	0.0057
combMNZ	RRF	0.0008	0.0055	0.0029	0.0038
combSUM	Borda	0.0005	0.006	0.0029	0.0019
combSUM	combMAX	0.0005	0.0084	0.0057	0.0038
combSUM	combMIN	0.0002	0.0127	0	0.0019
combSUM	combMNZ	0.0033	0.0075	0.0086	0.0076
combSUM	combSUM	0.0014	0.0067	0.0086	**0.0067**
combSUM	RRF	0.0009	0.0051	0.0029	0.0048
RRF	Borda	0.0005	0.0057	0	0.0019
RRF	combMAX	0.0004	0.0070	0	0.0038
RRF	combMIN	0.0002	0.0121	0	0
RRF	combMNZ	**0.0037**	0.0129	0.0086	**0.0067**
RRF	combSUM	0.0011	0.0060	0.0086	**0.0067**
RRF	RRF	0.0010	0.0047	0.0029	0.0057

The results of the combination of text and visual approach are shown in Table 3. The visual approach selected for these combination used RRF for query fusion and combSUM for the fusion of the descriptors, obtaining the best results in terms of MAP (MAP=0.0037) and P30 (P30=0.0067)(see Table 2).

Although in previous ImageCLEF campaigns, the mixed submissions sometimes achieved worse performance than the textual runs, the best result among all the experiments carried out on this chapter was obtained using a linear combination of text and visual search (MAP=0.1795). The weight of each rank was defined by a function of their performance in terms of MAP, where the best MAP scores obtained using text ($MAP(T) = 0.1293$) and visual ($MAP(V) = 0.0204$) search in ImageCLEFmed 2011 [29] were employed. Despite the lower performance of the the visual and textual approaches compared with the runs submitted to ImageCLEFmed 2013, the fusion of visual and text retrieval outperform the best multimodal approach submitted to ImageCLEFmed 2013. This shows the importance of the multimodal fusion.

Table 3 Results of the approaches for the ImageCLEF case–based retrieval task when using various fusion strategies to combine visual and textual information.

Visual+textual f.	MAP	Bpref	P10	P30
Best ImageCLEF run	0.1608	0.1426	0.1800	0.1257
Borda	0.1302	0.1230	0.1371	0.1105
combMAX	0.1770	0.1625	0.2143	**0.1571**
combMIN	0.1505	0.157	0.2171	0.1438
combMNZ	0.1197	0.1257	0.1714	0.1133
combSUM	0.1741	0.1609	**0.2229**	0.161
RRF	0.1084	0.1011	0.1543	0.1114
LN	**0.1795**	**0.1627**	0.2086	**0.1571**

Sections 2.2.2 and 2.2.3 describe the modality classification and compound figure separation tasks. Both can be fused with text and visual retrieval to improve the quality of the systems but this has not yet been implemented in our approach. The modality classification of Khresmoi approached achieved an accuracy of 69.63%. Moreover, the compound figure separation approach obtained the best accuracy of all the ImageCLEF 2013 participants (84.64%) [20]. There seems to be potential for improving performance including these techniques into the retrieval process.

5 Conclusions

In their practical work, clinicians have information needs when taking informed decisions. Their work sometimes involves search for similar past cases.To help clinicians in their daily routine, several information retrieval approaches involving visual and text retrieval are proposed.

This chapter describes the methods to combine both visual and textual information in a biomedical retrieval system. In the context of the Khresmoi project, experiments on text and visual fusion were done using the ImageCLEFmed 2013 database. Despite the low performance of the visual approaches on the case–based task, the fusion of text and visual techniques are improving the quality of the retrieval. Applying weighted linear combination of text and visual retrieval ranks, results outperform the best multimodal runs submitted at ImageCLEFmed 2013 with a MAP of 0.1795. It demonstrates the effectiveness of proposed the multimodal framework. Moreover, image analysis can be applied to enhance the quality of retrieval system. Image classification and compound figure separation are common techniques that can be integrated into a retrieval systems to improve the performance of a simple text or visual retrieval. The system achieved and accuracy of 69.63% at the ImageCLEFmed 2013 modality classification task and 84.64% at the compound figure separation task. Future work will focus on integration the modality classification and compound figure separation into the retrieval system to show that they can contribute to improve the retrieval.

Acknowledgments

This work was partly supported by the EU 7^{th} Framework Program in the context of the Khresmoi project (FP7–257528).

References

1. Aamodt, A., Plaza, E.: Case–based reasoning: Foundational issues, methodological variations, and systems approaches. AIC **7**(1), 39–59 (1994)
2. Apostolova, E., You, D., Xue, Z., Antani, S., Demner-Fushman, D., Thoma, G.R.: Image retrieval from scientific publications: text and image content processing to separate multi–panel figures. Journal of the American Society for Information Science (2013)
3. Aswani, N., Beckers, T., Birngruber, E., Boyer, C., Burner, A., BystroN, J., Choukri, K., Cruchet, S., Cunningham, H., Dedek, J., Dolamic, L., Donner, R., Dungs, S., Eggel, I., Foncubierta-Rodríguez, A., Fuhr, N., Funk, A., García Seco de Herrera, A., Gaudinat, A., Georgiev, G., Gobeill, J., Goeuriot, L., Gómez, P., Greenwood, M., Gschwandtner, M., Hanbury, A., Hajic, J., Hlavácová, J., Holzer, M., Jones, G., Jordan, B., Jordan, M., Kaderk, K., Kainberger, F., Kelly, L., Mriewel, S., Kritz, M., Langs, G., Lawson, N., Markonis, D., Martinez, I., Momtchev, V., Masselot, A., Mazo, H., Müller, H., Pecina, P., Pentchev, K., Peychev, D., Pletneva, N., Pottecherc, D., Roberts, A., Ruch, P., Samwald, M., Schneller, P., Stefanov, V., Tinte, M.A., Uresová, Z., Vargas, A., Vishnyakova, D.: Khresmoi: Multimodal multilingual medical information search. In: Proceedings of the 24th International Conference of the European Federation for Medical Informatics (2012)
4. Begum, S., Ahmed, M.U., Funk, P., Xiong, N., Folke, M.: Case-based reasoning systems in the health sciences: A survey of recent trends and developments. IEEE Transactions on Systems, Man and Cybernetics **41**(4), 421–434 (2011)
5. Burghouts, G.J., Geusebroek, J.M.: Performance evaluation of local colour invariants. Compututer Vision and Image Understanding **113**(1), 48–62 (2009)

6. Caputo, B., Müller, H., Thomee, B., Villegas, M., Paredes, R., Zellhofer, D., Goeau, H., Joly, A., Bonnet, P., Martinez Gomez, J., Garcia Varea, I., Cazorla, C.: Imageclef 2013: the vision, the data and the open challenges. In: Working Notes of CLEF 2013 (Cross Language Evaluation Forum) (2013)
7. Chatzichristofis, S.A., Boutalis, Y.S.: CEDD: Color and edge directivity descriptor: A compact descriptor for image indexing and retrieval. In: Lecture notes in Computer Sciences, vol. 5008, pp. 312–322 (2008)
8. Chatzichristofis, S.A., Boutalis, Y.S.: FCTH: Fuzzy color and texture histogram: A low level feature for accurate image retrieval. In: Proceedings of the 9th International Workshop on Image Analysis for Multimedia Interactive Service, pp. 191–196 (2008)
9. Cheng, B., Sameer, A., Stanley, R.J., Thoma, G.R.: Automatic segmentation of subfigure image panels for multimodal biomedical document retrieval. In: G. Agam, C. Viard-Gaudin (eds.) Document Recognition and Retrieval, *SPIE Proceedings*, vol. 7874, pp. 1–10. SPIE (2011)
10. Cheng, B., Stanley, R.J., De, S., Antani, S., Thoma, G.R.: Automatic detection of arrow annotation overlays in biomedical images. International Journal of Healthcare Information Systems and Informatics **6**(4), 23–41 (2011)
11. Chhatkuli, A., Markonis, D., Foncubierta-Rodríguez, A., Meriaudeau, F., Müller, H.: Separating compound figures in journal articles to allow for subfigure classification. In: SPIE Medical Imaging (2013)
12. Chiu, P., Chen, F., Denoue, L.: Picture detection in document page images. In: Proceedings of the 10th ACM symposium on Document engineering, pp. 211–214. ACM (2010)
13. Demner-Fushman, D., Antani, S., Simpson, M.S., Thoma, G.R.: Design and development of a multimodal biomedical information retrieval system. Journal of Computing Science and Engineering **6**(2), 168–177 (2012)
14. Depeursinge, A., Müller, H.: Fusion techniques for combining textual and visual information retrieval. In: H. Müller, P. Clough, T. Deselaers, B. Caputo (eds.) ImageCLEF, *The Springer International Series On Information Retrieval*, vol. 32, pp. 95–114. Springer Berlin Heidelberg (2010)
15. Glasgow, J., Jurisica, I.: Integration of case–based and image–based reasoning. In: AAAI Workshop on Case–Based Reasoning Integrations, pp. 67–74. AAAI Press, Menlo Park, California (1998)
16. Gu, M., Aamodt, A., Tong, X.: Component retrieval using conversational case–based reasoning. In: Intelligent information processing II, pp. 259–271. Springer-Verlag, London, UK (2005)
17. García Seco de Herrera, A., Kalpathy-Cramer, J., Demner Fushman, D., Antani, S., Müller, H.: Overview of the ImageCLEF 2013 medical tasks. In: Working Notes of CLEF 2013 (Cross Language Evaluation Forum) (2013)
18. García Seco de Herrera, A., Markonis, D., Eggel, I., Müller, H.: The medGIFT group in ImageCLEFmed 2012. In: Working Notes of CLEF 2012 (2012)
19. García Seco de Herrera, A., Markonis, D., Müller, H.: Bag of colors for biomedical document image classification. In: H. Greenspan, H. Müller (eds.) Medical Content–based Retrieval for Clinical Decision Support, MCBR–CDS 2012, pp. 110–121. Lecture Notes in Computer Sciences (LNCS) (2013)
20. García Seco de Herrera, A., Markonis, D., Schaer, R., Eggel, I., Müller, H.: The medGIFT group in ImageCLEFmed 2013. In: Working Notes of CLEF 2013 (Cross Language Evaluation Forum) (2013)
21. García Seco de Herrera, A., Schaer, R., Müller, H.: Comparing fusion techniques for the imageclef 2013 medical case retrieval task. Computerized Medical Imaging and Graphics (submitted)
22. Hwang Kyung, H., Lee, H., Choi, D.: Medical image retrieval: Past and present. Health Information Research **18**(1), 3–9 (2012)
23. Ide, N.C., Loane, R.F., Demner-Fushman, D.: Essie: A concept–based search engine for structured biomedical text. Journal of the American Medical Informatics Association **14**(3), 253–263 (2007)

24. Kalpathy-Cramer, J., Hersh, W.: Multimodal medical image retrieval: image categorization to improve search precision. In: Proceedings of the international conference on Multimedia information retrieval, MIR '10, pp. 165–174. ACM, New York, NY, USA (2010)
25. Kwiatkowska, M., Atkins, S.: Case representation and retrieval in the diagnosis and treatment of obstructive sleep apnea: A semiofuzzy approach. In: Proceedings European Case Based Reasoning Conference, ECCBR'04 (2004)
26. Lazebnik, S., Schmid, C., Ponce, J.: Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In: Proceedings of the 2006 IEEE Conference on Computer Vision and Pattern Recognition, CVPR, pp. 2169–2178. IEEE Computer Society, Washington, DC, USA (2006)
27. Li, Y., Shi, N., D.Frank, H.: Fusion analysis of information retrieval models on biomedical collections. In: Proceedings of the 14th International Conference onInformation Fusion. IEEE Computer Society (2011)
28. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision **60**(2), 91–110 (2004)
29. Markonis, D., García Seco de Herrera, A., Eggel, I., Müller, H.: The medGIFT group in ImageCLEFmed 2011. In: Working Notes of CLEF 2011 (2011)
30. Markonis, D., Holzer, M., Dung, S., Vargas, A., Langs, G., Kriewel, S., Müller, H.: A survey on visual information search behavior and requirements of radiologists. Methods of Information in Medicine **51**(6), 539–548 (2012)
31. McCandless, M., Hatcher, E., Gospodnetic, O.: Lucene in Action, Second Edition: Covers Apache Lucene 3.0. Manning Publications Co., Greenwich, CT, USA (2010)
32. Montani, S., Bellazzi, R.: Supporting decisions in medical applications: the knowledge management perspective. International Journal of Medical Informatics **68**, 79–90 (2002)
33. Mourão, A., Martins, F.: NovaMedSearch: a multimodal search engine for medical case-based retrieval. In: Proceedings of the 10th Conference on Open Research Areas in Information Retrieval, OAIR'13, pp. 223–224 (2013)
34. Müller, H., Boyer, C., Gaudinat, A., Hersh, W., Geissbuhler, A.: Analyzing web log files of the Health On the Net HONmedia search engine to define typical image search tasks for image retrieval evaluation. In: MedInfo 2007, *IOS press, Studies in Health Technology and Informatics*, vol. 12, pp. 1319–1323. Brisbane, Australia (2007)
35. Müller, H., Despont-Gros, C., Hersh, W., Jensen, J., Lovis, C., Geissbuhler, A.: Health care professionals' image use and search behaviour. In: Proceedings of the Medical Informatics Europe Conference (MIE 2006), IOS Press, Studies in Health Technology and Informatics, pp. 24–32. Maastricht, The Netherlands (2006)
36. Müller, H., García Seco de Herrera, A., Kalpathy-Cramer, J., Demner Fushman, D., Antani, S., Eggel, I.: Overview of the ImageCLEF 2012 medical image retrieval and classification tasks. In: Working Notes of CLEF 2012 (Cross Language Evaluation Forum) (2012)
37. Müller, H., Zhou, X., Depeursinge, A., Pitkanen, M., Iavindrasana, J., Geissbuhler, A.: Medical visual information retrieval: State of the art and challenges ahead. In: Proceedings of the 2007 IEEE International Conference on Multimedia and Expo, ICME'07, pp. 683–686. IEEE (2007)
38. Philip, A., Afolabi, B., Oluwaranti, A., Oluwatolani, O.: Development of an image retrieval model for biomedical image databases. In: C. Jao (ed.) Efficient Decision Support Systems — Practice and Challenges in Biomedical Related Domain, pp. 311–329. Intech (2010)
39. Quellec, G., Lamard, M., Bekri, L., Cazuguel, G., Roux, C., Cochener, B.: Medical case retrieval from a committee of decision trees. IEEE Transactions on Information Technology in Biomedicine **14**(5), 1227–1235 (2010)
40. Rahman, M.M., You, D., Simpson, M.S., Antani, S., Demner-Fushman, D., Thoma, G.R.: An interactive image retrieval framework for biomedical articles based on visual region–of–interest (roi) identification and classification. In: Proceedings of the IEEE Second International Conference on Healthcare Informatics, Imaging and Systems Biology, HISB (2012)
41. Rahman, M.M., You, D., Simpson, M.S., Antani, S.K., Demner-Fushman, D., Thoma, G.R.: Multimodal biomedical image retrieval using hierarchical classification and modality fusion. International Journal of Multimedia Information Retrieval **2**(3), 159–173 (2013)

42. van de Sande, K.E.A., Gevers, T., Smeulders, A.W.M.: The university of amsterdam's concept detection system at ImageCLEF 2009. Lecture Notes in Computer Science **6242**, 261–268 (2010)
43. Selvarajah, S., Kodituwakku, S.R.: Analysis and comparison of texture features for content based image retrieval. International Journal of Latest Trends in Computing **2**, 108–113 (2011)
44. Seo, M., Ko, B., Chung, H., Nam, J.: ROI–based medical image retrieval using human-perception and MPEG–7 visual descriptors. In: Proceedings of the 5th International Conference on Image and Video Retrieval, CIVR'06, pp. 231–240. Springer-Verlag, Berlin, Heidelberg (2006)
45. Shapiro, L.G., Atmosukarto, I., Cho, H., Lin, H.J., Ruiz-Correa, S., Yuen, J.: Similarity-based retrieval for biomedical applications. In: Case–Based Reasoning on Images and Signals, *Studies in Computational Intelligence*, vol. 73, pp. 355–387. Springer (2008)
46. Simonyan, K., Modat, M., Ourselin, S., Criminisi, A., Zisserman, A.: Immediate roi search for 3–d medical images. In: H. Greenspan, H. Müller (eds.) Medical Content–based Retrieval for Clinical Decision Support, MCBR–CDS 2012. Lecture Notes in Computer Sciences (LNCS) (2013)
47. Simpson, M.S., Demner-Fushman, D.: Biomedical text mining: A survey of recent progress. In: Mining Text Data, pp. 465–517. Springer (2012)
48. Snoek, C.G.M., Worring, M., Smeulders, A.W.M.: Early versus late fusion in semantic video analysis. In: MULTIMEDIA '05: Proceedings of the 13th annual ACM international conference on Multimedia, pp. 399–402. ACM, New York, NY, USA (2005)
49. Tirilly, P., Lu, K., Mu, X., Zhao, T., Cao, Y.: On modality classification and its use in text-based image retrieval in medical databases. In: 9th International Workshop on Content–Based Multimedia Indexing (2011)
50. Tsikrika, T., Müller, H., Kahn Jr., C.E.: Log analysis to understand medical professionals' image searching behaviour. In: Proceedings of the 24th European Medical Informatics Conference, MIE'2012 (2012)
51. Wang, J.Z.: Region–based retrieval of biomedical images. In: Proceedings of the ACM Multimedia Conference, pp. 511–512 (2000)
52. Welter, P., Deserno, T.M., Fischer, B., Günther, R.W., Spreckelsen, C.: Towards case–based medical learning in radiological decision making using content–based image retrieval. BMC Medical Informatics and decision Making **11**(68) (2011)
53. Zhang, D., Lu, G.: Review of shape representation and description techniques. Pattern Recognition **37**(1), 1–19 (2004)