

Text- and Content-based Approaches to Image Modality Classification and Retrieval for the ImageCLEF 2011 Medical Retrieval Track

Matthew Simpson, Md Mahmudur Rahman, Srinivas Phadnis, Emilia Apostolova, Dina Demner-Fushman, Sameer Antani, and George Thoma

Lister Hill National Center for Biomedical Communications
U.S. National Library of Medicine, NIH, Bethesda, MD, USA

Abstract. This article describes the participation of the Communications Engineering Branch (CEB), a division of the Lister Hill National Center for Biomedical Communications, in the ImageCLEF 2011 medical retrieval track. Our methods encompass a variety of techniques relating to text- and content-based image retrieval. Our textual approaches primarily utilize the Unified Medical Language System (UMLS) synonymy to identify concepts in topic descriptions and image-related text, and our visual approaches utilize similarity metrics based on computed “visual concepts” and low-level image features. We also explore mixed approaches that utilize a combination of textual and visual features. In this article we present an overview of the application of our methods to the modality classification, ad-hoc image retrieval, and case-based image retrieval tasks, and we describe our submitted runs and results.

Keywords: Image Retrieval, Case-based Retrieval, Image Modality

1 Introduction

This article describes the participation of the Communications Engineering Branch (CEB), a division of the Lister Hill National Center for Biomedical Communications, in the ImageCLEF 2011 medical retrieval track.

The medical retrieval track [9] of ImageCLEF 2011 consists of an image modality classification task and two retrieval tasks. For the modality classification task, the goal is to classify a given set of medical images according to eighteen modalities (e.g., CT or Histopathology) taken from five classes (e.g., Radiology or Microscopy). In the first retrieval task, a set of ad-hoc information requests is given, and the goal is to retrieve the most relevant images for each topic. Finally, in the second retrieval task, a set of case-based information requests is given, and the goal is to retrieve the most relevant articles describing similar cases.

In the following sections, we describe the textual and visual features that comprise our image and case representations (Sections 2–3) and our methods for the modality classification (Section 4) and medical retrieval tasks (Sections 5–6). Our textual approaches primarily utilize the Unified Medical Language

System (UMLS) [11] synonymy to identify concepts in topic descriptions and image-related text, and our visual approaches rely on similarity metrics based on computed “visual concepts” and other low-level visual features. We also explore mixed approaches for the modality classification and retrieval tasks that utilize a combination of textual and visual features.

In Section 7 we describe our submitted runs, and in Section 8 we present our results. For the modality classification task, our best submission achieved a classification accuracy of 74% and was ranked within the submissions from the top three groups. For the retrieval tasks, our results were lower than expected yet reveal new insights which we anticipate will improve future work. For the modality classification and image retrieval tasks, our best results were obtained using mixed approaches, indicating the importance of both textual and visual features for these tasks.

2 Image Representation

Images contained in biomedical articles can be represented using both textual and visual features. Textual features can include text from an article that pertains to an image, such as image captions and “mentions” (snippets of text within the body of an article that discuss an image), and visual features can include information derived from the content of an image, such as shape, color and texture. We describe the features we use in representing images below.

2.1 Textual Features

We represent each image in the ImageCLEF 2011 medical collection as a structured document of image-related text. Our representation includes the title, abstract, and MeSH terms¹ of the article in which the image appears as well as the image’s caption and mentions. Additionally, we identify within image captions textual Regions of Interest (ROIs). A textual ROI is a noun phrase describing the content of an interesting region of an image and is identified within a caption by a pointer. For example, in the caption “MR image reveals hypointense indeterminate nodule (arrow),” the word *arrow* points to the ROI containing a *hypointense indeterminate nodule*.

The above structured documents may be indexed and searched with a traditional search engine or the underlying term vectors may be exposed and added to a mixed image representation that includes the visual features described in Section 2.2. For the latter approach, the terms in a structured document field D_j (e.g., caption) are commonly represented as an N -dimensional vector

$$\mathbf{f}_j^{\text{term}} = [w_{j1}, w_{j2}, \dots, w_{jN}]^T \quad (1)$$

where w_{j_k} denotes the *tf-idf* weight of term t_k in document field D_j , and N is the size of the vocabulary.

¹ MeSH is a controlled vocabulary created by U.S. National Library of Medicine to index biomedical articles.

2.2 Visual Features

In addition to the above textual features, we also represent the visual content of images using various low-level global image features and a derived feature intended to capture the high-level semantic content of images.

Low-level Global Features We represent the spatial structure and global shape and edge features of images with the Color Layout Descriptor (CLD) and Edge Histogram Descriptor (EHD) of MPEG-7 [2]. We extract the CLD feature as a vector \mathbf{f}^{clD} and the EHD feature as \mathbf{f}^{ehD} . Additionally, we extract the Color and Edge Directivity Descriptor (CEDD) [3] as \mathbf{f}^{cedd} and the Fuzzy Color and Texture Histogram (FCTH) [4] as \mathbf{f}^{fcth} using the Lucene image retrieval (LIRE) library.² Both CEDD and FCTH incorporate color and texture information into single histograms that are suitable for image indexing and retrieval.

Concept Feature In a heterogeneous medical image collection, it is possible to identify specific local patches in images that are perceptually or semantically distinguishable, such as homogeneous texture patterns in gray-level radiological images or differential color and texture structures in microscopic pathology images. The variation in the local patches can be effectively modeled as “visual concepts” [12] using supervised machine learning-based classification techniques.

For the generation of these concepts, we utilize a multi-class Support Vector Machine (SVM) composed of several binary classifiers organized using the *one-against-one* strategy [7]. To train the SVMs, we manually assign a set of L visual concepts $C = \{c_1, \dots, c_i, \dots, c_L\}$ to the color and texture features of each fixed-size patch contained in an image. For a single image, the input to the training process is a set of color and texture feature vectors for all fixed-size patches along with their manually assigned concept labels. We generate the concept feature for each image I_j in the collection by first partitioning I_j into l patches as $\{\mathbf{x}_{1_j}, \dots, \mathbf{x}_{k_j}, \dots, \mathbf{x}_{l_j}\}$, where each $\mathbf{x}_{k_j} \in \mathbb{R}^d$ is a combined color and texture feature vector. Then, for each \mathbf{x}_{k_j} , we determine its concept label by the prediction of the multi-class SVM. Thus, in contrast to the low-level features described above, the concept feature represents an image as a set of high-level “visual concepts.” Based on this encoding scheme, we represent an image I_j as a vector of concepts

$$\mathbf{f}_j^{\text{concept}} = [w_{1_j}, \dots, w_{i_j}, \dots, w_{L_j}]^T \quad (2)$$

where each w_{i_j} denotes the *tf-idf* weight of concept c_i in Image I_j .

Clustered Features In an attempt to avoid the online computational complexity required to calculate visual similarity (described in Section 5.2) using the above features, we create an index of image similarity based on the clustering of feature vectors. For each visual feature described above, we cluster the vectors assigned to all images into $k = d \log |I|$ clusters using the k -means++ [1]

² <http://freshmeat.net/projects/lirecbir/>

algorithm, where d is the number of attributes in each vector and $|I|$ is the total number of images in the collection. We then assign each cluster a unique “word” and represent each image as a sequence of these words. For example, using only MPEG-7 features for simplicity, an image might be represented as the sequence “cld:k1 ehd:k2” if the image’s CLD feature was among the vectors in the first CLD cluster and its EHD feature was among the vectors in the second EHD cluster. The resulting textual interpretation of an image’s visual features may then be indexed and searched using a traditional search engine or added to a mixed image representation that includes the textual features described in Section 2.1.

3 Case Representation

We represent a full-text article as the combination of the textual and visual features of each image appearing in the article. Thus, each article representation consists of an article’s title, abstract, and MeSH terms as well as the caption, mention, textual ROIs, and clustered visual features of each contained image.

4 Modality Classification Task

Owing to their empirical success, we utilize multi-class SVMs for classifying images into eighteen medical image modalities based on their textual and visual features. We compose multi-class SVMs using the *one-against-one* strategy [7] for combining the pairwise classifications of each binary SVM.

Figure 1 describes our textual, visual, and mixed approaches to the modality classification task. Our visual and textual image features (with the text-based features represented as term vectors) can be used individually to produce single-mode classifications, or they may be combined to produce multimodal predictions. For the mixed approaches, the features may be combined into a single feature vector or they may be used independently, with the separate predictions being “fused” to form a single classifier. We fuse the output of multiple classifiers with the popular “Sum” classifier combination technique [6, 10] of Bayes’ theorem.

We utilize the above approach for both flat and hierarchical modality classification. For the former, the system classifies an image’s modality as one of eighteen medical image modalities. For the latter, the system first classifies the image’s modality as belonging to one of five high-level modality classes (i.e., Radiology, Microscopy, Photograph, Graphic, or Other), and then it classifies the image’s modality as one of the original eighteen, given its predicted high-level class. The hierarchical approach requires the training of a single high-level classifier and multiple class-specific classifiers, and an appropriate set of example images must be constructed to train each classifier.

Due to the small number of training examples for several modality classes, we created an extended set of training images from the collection. We accomplished this task by first performing textual image searches, using particular modalities as queries, then by manually inspecting and labeling the retrieved results.

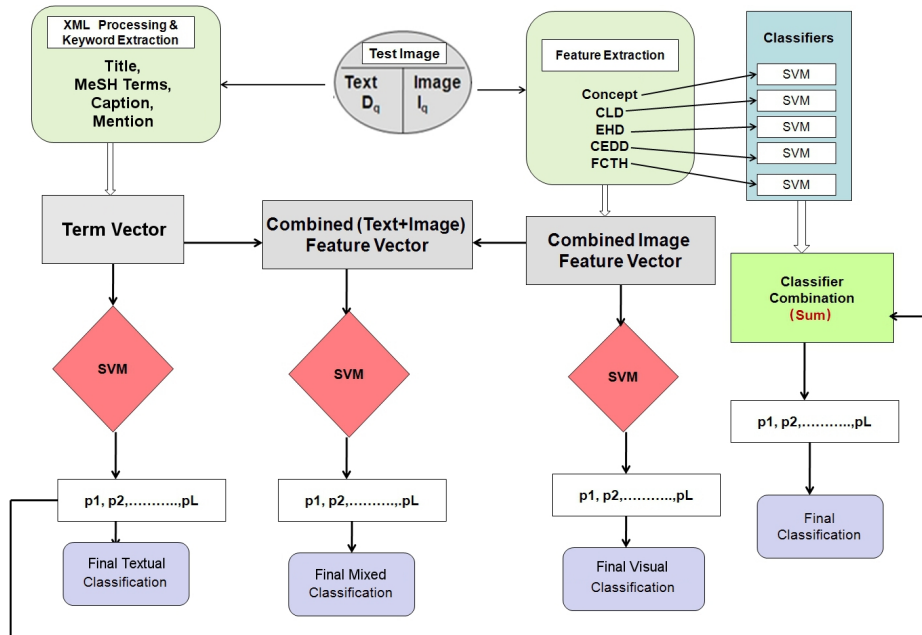


Fig. 1. Process flow diagram for our modality classification approach

5 Ad-Hoc Image Retrieval Task

In this section we describe our textual, visual and mixed approaches to the ad-hoc image retrieval task. Descriptions of the submitted runs that utilize these methods are presented in Section 7.

5.1 Textual Approaches

To allow for efficient retrieval and to compare their relative performance, we index the textual image representations described in Section 2.1 with the Essie [8] and Lucene/SOLR³ search engines. Essie is a search engine developed by the U.S. National Library of Medicine and is particularly well-suited for the medical retrieval task due to its ability to automatically expand query terms using the UMLS synonymy. Lucene/SOLR is a popular search engine developed by the Apache Software Foundation that employs the well-known vector space model of information retrieval and *tf-idf* term weighting. Both Essie and Lucene/SOLR provide the ability to weight term occurrences according to the location in a document in which they occur. For example, we weight term occurrences in image captions higher than those in article abstracts.

³ <http://lucene.apache.org/>

We organize each topic description into the well-formed clinical question (i.e., PICO⁴) framework following the method described by Demner-Fushman and Lin [5]. Extractors identify UMLS concepts related to problems, interventions, age, anatomy, drugs, and image modality. When used as part of an Essie query, each extracted concept is automatically expanded along the synonymy relationships in the UMLS. For approaches that make use of the Lucene/SOLR search engine, we first expand the extracted concepts using Essie’s built-in synonymy and then replace the extracted concepts with their expansions.

To construct a query for each topic, we create and combine several boolean expressions derived from the extracted concepts. First, we create an expression by combining the concepts using the **AND** operator (meaning all of the concepts are required to occur in an image’s textual representation), and then we produce additional expressions by allowing an increasing number of the extracted concepts to be optional. Finally, we combine these expressions using the **OR** operator giving significantly more weight to expressions containing a fewer number of optional concepts. Additionally, we often include the verbatim topic description as a component of a query, but we give minimal weight to this expression compared to those containing the extracted concepts. We use the resulting queries to search the Essie and Lucene/SOLR indices.

5.2 Visual Approaches

Our visual approaches to image retrieval are based on retrieving images that appear visually similar to the given topic images. The similarity between a query image I_q and a target image I_j , based on the visual features described in Section 2.2, is defined by

$$\text{Sim}(I_q, I_j) = \sum_F \alpha^F \text{Sim}^F(I_q, I_j) \quad (3)$$

where $F \in \{\text{Concept, EHD, CLD, CEDD, FCTH}\}$, α^F are feature weights, and Sim^F is Euclidean distance. In the above similarity matching function, the feature weights are determined based on the cross validation accuracies of the feature-specific SVMs trained for the modality classification task. The weights are normalized to $0 \leq \alpha^F \leq 1$ and $\sum \alpha^F = 1$.

In order to avoid the online computation of the above similarity metric, we may utilize clustered visual features, also described in Section 2.2, to retrieve visually similar images. To allow for efficient retrieval, we index the textual interpretations of the images’ clustered visual features using the Essie and Lucene/SOLR search engines. Again, we utilize both search engines in order to compare their relative performance. Retrieval is performed by first extracting a query image’s visual features, then by determining the features’ cluster membership, and finally by combining the unique “words” assigned to the clusters containing the features in order to form a textual query. For a given topic, we combine the textual interpretations of all features for all sample images using the **OR** operator.

⁴ PICO is a mnemonic for structuring clinical questions in evidence-based practice and represents Patient/Population/Problem, Intervention, Comparison, and Outcome.

5.3 Mixed Approaches

Our mixed approaches to image retrieval combine our textual and visual approaches through a process of filtering and re-ranking or by issuing multimodal queries. For the filtering approach, we first filter the image collection by the two most probable modalities of the query images, as indicated by our modality classifier. We then query the remaining images according to a textual approach. For the re-ranking approach, we first query the image collection using a textual approach, and we then re-rank the retrieved images according to their visual similarity with the query images, as indicated by the above similarity metric. Finally, for approaches involving multimodal queries, we utilize Essie and Lucene/SOLR to index images using both their textual features and the textual interpretation of their clustered image features. We construct multimodal queries by combining a query produced by a textual approach with that produced by the visual approach described above that utilizes clustered image features. We join the textual and visual components of the query with the `OR` operator.

6 Case-Based Retrieval Task

Our method for performing case-based retrieval is analogous to our approach for ad-hoc image retrieval. Here, we index the case representations described in Section 3 using the Essie and Lucene/SOLR search engines (for performance comparison). We generate textual and mixed queries appropriate for both search engines according to the approaches described in Sections 5.1 and 5.3.

7 Submitted Runs

In this section we describe each of our submitted runs for the modality classification, ad-hoc image retrieval, and case-based image retrieval tasks. Each run is identified by its `trec_eval` run ID and followed by a submission mode (textual, visual or mixed) and type (automatic, manual or feedback).

7.1 Modality Classification Task

We submitted the following 10 runs for the modality classification task:

1. *image_test_result_original (visual, automatic)*: SVM classification derived from the original set of training images. Each image is represented as a single vector of visual features.
2. *image_test_result_ext (visual, automatic)*: SVM classification like Run 1 but derived from an extended set of training images.
3. *image_text_test_result_original (mixed, automatic)*: SVM classification derived from the original set of training images. Each image is represented as a single vector containing visual features and a subset of textual features (article title, MeSH terms, caption, and mention).

4. *image_text_test_result_ext (mixed, automatic)*: SVM classification like Run 3 but derived from an extended set of training images.
5. *image_text_test_result_sum (mixed, automatic)*: Classifier combination using the “Sum” method of Bayes’ theorem. Each image is represented as a group of vectors for visual and textual features that are individually classified using SVMs derived from the original set of training images.
6. *image_text_test_result_sum_ext (mixed, automatic)*: Classifier combination like Run 5 but using SVMs derived from an extended set of training images.
7. *image_text_test_result_CV (mixed, feedback)*: Linear classifier combination weighting classifiers according to their normalized cross-validation accuracies. Each image is represented as a group of vectors for visual and textual features that are individually classified using SVMs derived from the original set of training images.
8. *image_text_test_result_CV_ext (mixed, automatic)*: Classifier combination like Run 7 but using SVMs derived from an extended set of training images.
9. *image_text_test_result_multilevel (mixed, automatic)*: Hierarchical SVM classification derived from the original set of training images. Each image is represented as a single vector of visual and textual features that is first classified into a top-level modality class and is then further classified using a class-specific SVM.
10. *image_text_test_result_multilevel_ext (mixed, automatic)*: SVM classification like Run 9 but using SVMs derived from an extended set of training images.

7.2 Ad-hoc Image Retrieval Task

We submitted the following 10 runs for the ad-hoc image retrieval task:

1. *iti-essie-baseline+expanded-concepts (textual, automatic)*: Textual search using the Essie search engine. Each image is represented by its textual features, and queries combine the verbatim topic description with extracted concepts and image modalities.
2. *iti-lucene-baseline+expanded-concepts (textual, automatic)*: Textual search using the Lucene/SOLR search engine. Each image is represented as in Run 1, and queries combine the verbatim topic description with extracted concepts and image modalities that are then expanded along synonymy relationships in the UMLS.
3. *iti-lucene-image (visual, automatic)*: Visual search using the Lucene/SOLR search engine. Each image is represented using the textual interpretation of its clustered visual features, and queries combine the visual “words” of each of the sample topic images.
4. *image_fusion_category_weight_filter (visual, automatic)*: Similarity matching over images filtered according to modality. Each image is represented as a subset of visual features (Concept, CLD, and EHD), and similarity scores for each feature are linearly combined and weighted according to modality class.
5. *image_fusion_category_weight_filter_merge (visual, automatic)*: Similarity matching like Run 4, but each image is scored as the sum of its similarity with each of the sample topic images.

6. *image_fusion_category_weight_merge (visual, automatic)*: Similarity matching like Run 5, but the image set is not filtered by modality prior to retrieval.
7. *iti-lucene-baseline+expanded-concepts+image (mixed, automatic)*: Mixed search using the Lucene/SOLR search engine. Image representations and queries combine the textual and visual features of Run 2 and Run 3.
8. *iti-essie-baseline+expanded-concepts+image (mixed, automatic)*: Mixed search using the Essie search engine. Image representations and queries combine the textual and visual features of Run 1 and Run 3.
9. *Multimodal_Rerank_Merge (mixed, automatic)*: Re-ranking of Run 1 according to Run 6.
10. *Multimodal_Rerank_Filter_Merge (mixed, automatic)*: Re-ranking of Run 1 according to Run 5.

7.3 Case-based Retrieval Task

We submitted the following 10 runs for the case-based retrieval task:

1. *iti-essie-manual (textual, manual)*: Textual search using the Essie search engine. Articles are represented by their textual features, and queries were manually generated by a medical doctor with expertise in biomedical informatics.
2. *iti-essie-frames (textual, automatic)*: Textual search using the Essie search engine. Articles are represented by their textual features, and queries combine concepts from automatically generated PICO summary frames.
3. *iti-lucene-frames (textual, automatic)*: Textual search using the Lucene/SOLR search engine. Articles are represented by their textual features, and queries combine concepts from automatically generated PICO summary frames that are then expanded along synonymy relationships in the UMLS.
4. *iti-lucene-baseline (textual, automatic)*: Textual search with the Lucene/SOLR search engine. Articles are represented by their textual features, and queries are the verbatim topic descriptions.
5. *iti-lucene-expanded-concepts (textual, automatic)*: Textual search using the Lucene/SOLR search engine. Articles are represented by their textual features, and queries combine extracted concepts and image modalities that are then expanded along synonymy relationships in the UMLS.
6. *iti-lucene-baseline+expanded-concepts (textual, automatic)*: Textual search like Run 5, but queries also include verbatim topic descriptions.
7. *iti-lucene-baseline+expanded-concepts+cases (textual, automatic)*: Textual search like Run 6, but articles are boosted if their MeSH terms are indicative of case studies or clinical trials.
8. *iti-lucene-expanded-concepts+image (mixed, automatic)*: Mixed search using the Lucene/SOLR search engine. Articles are represented by their textual features and the textual interpretation of the clustered visual features of each contained image. Queries are as in Run 5 but also include the visual “words” of each of the sample topic images.
9. *iti-lucene-baseline+expanded-concepts+image (mixed, automatic)*: Mixed search like Run 8, but queries also include verbatim topic descriptions.

<i>ID</i>	Mode	Type	Accuracy (%)
<i>image_text_test_result_multilevel</i>	Mixed	Automatic	74.12
<i>image_text_test_result_sum_ext</i>	Mixed	Automatic	60.25
<i>image_text_test_result_CV</i>	Mixed	Feedback	59.66
<i>image_text_test_result_multilevel_ext</i>	Mixed	Automatic	59.17
<i>image_text_test_result_sum</i>	Mixed	Automatic	59.17
<i>image_text_test_result_CV_ext</i>	Mixed	Automatic	58.20
<i>image_text_test_result_original</i>	Mixed	Automatic	58.20
<i>image_text_result_original</i>	Visual	Automatic	57.12
<i>image_text_test_result_ext</i>	Mixed	Automatic	54.39
<i>image_test_result_ext</i>	Visual	Automatic	48.53

Table 1. Accuracy results for the modality classification task.

<i>ID</i>	Mode	Type	MAP
<i>iti-lucene-baseline+expanded-concepts+image</i>	Mixed	Automatic	0.1356
<i>iti-lucene-baseline+expanded-concepts</i>	Textual	Automatic	0.1255
<i>iti-essie-baseline+expanded-concepts</i>	Textual	Automatic	0.0966
<i>Multimodal_Rerank_Filter_Merge</i>	Mixed	Automatic	0.0910
<i>Multimodal_Rerank_Merge</i>	Mixed	Automatic	0.0903
<i>iti-essie-baseline+expanded-concepts+image</i>	Mixed	Automatic	0.0843
<i>iti-lucene-image</i>	Visual	Automatic	0.0245
<i>image_fusion_category_weight_filter</i>	Visual	Automatic	0.0221
<i>image_fusion_category_weight_filter_merge</i>	Visual	Automatic	0.0201
<i>image_fusion_category_weight_merge</i>	Visual	Automatic	0.0193

Table 2. Retrieval results for the ad-hoc image retrieval task

- iti-lucene-baseline+expanded-concepts+image+cases (mixed, automatic)*: Mixed search like Run 9, but articles are boosted if their MeSH terms are indicative of case studies or clinical trials.

8 Results

Table 1 presents the classification accuracy of our submitted runs for the modality classification task. *image_text_test_result_multilevel*, a mixed approach, achieved the highest accuracy (74%) of our submitted runs and was ranked within the submissions from the top three groups. This result validates our hierarchical classification approach and, as in our previous experience with image modality classification [14], underscores the benefits of combining textual and visual features. Surprisingly, the use of an extended set of training images did not improve classification accuracy.

Table 2 presents the Mean Average Precision (MAP) of our submitted runs for the ad-hoc image retrieval task. *iti-lucene-baseline+expanded-concepts+image*

<i>ID</i>	Mode	Type	MAP
<i>iti-essie-manual</i>	Textual	Manual	0.0941
<i>iti-lucene-baseline</i>	Textual	Automatic	0.0762
<i>iti-lucene-baseline+expanded-concepts+image</i>	Mixed	Automatic	0.0269
<i>iti-lucene-baseline+expanded-concepts</i>	Textual	Automatic	0.0264
<i>iti-lucene-baseline+expanded-concepts+image+cases</i>	Mixed	Automatic	0.0255
<i>iti-lucene-baseline+expanded-concepts+cases</i>	Textual	Automatic	0.0249
<i>iti-lucene-expanded-concepts+image</i>	Mixed	Automatic	0.0247
<i>iti-lucene-expanded-concepts</i>	Textual	Automatic	0.0243
<i>iti-essie-frames</i>	Textual	Automatic	0.0174
<i>iti-lucene-frames</i>	Textual	Automatic	0.0141

Table 3. Retrieval results for the case-based retrieval task.

achieved the highest MAP (0.14) among our submitted mixed runs, *iti-lucene-baseline+expanded-concepts* achieved the highest MAP (0.13) among our submitted textual runs, and *iti-lucene-image* achieved the highest MAP (0.02) among our submitted visual runs. Although our retrieval results are lower than expected given our previous experience [13, 14], they demonstrate the utility of combining both textual and visual features. In particular, the use of clustered visual features, which can be indexed and searched with a traditional text-based search engine, not only resulted in our best visual approach but, when used in combination with our best textual approach, produced our best overall submission.

Finally, Table 3 presents the MAP of our submitted runs for the case-based retrieval task. *iti-essie-manual* achieved the highest MAP (0.09) among our submitted textual runs, and *iti-lucene-baseline+expanded-concepts+image* achieved the highest MAP (0.03) among our submitted mixed runs. *iti-lucene-baseline*, a textual approach, achieved the highest MAP (0.08) among our submitted automatic runs. Similar to our results for the image retrieval task, our case-based retrieval results are lower than expected given our previous experience. The relatively low MAP for most ImageCLEF 2011 case-based submissions may be due, in part, to the existence in the collection of only a small number of case reports, clinical trials, or other types of documents relevant for case-based topics. Surprisingly, our submissions that utilize extracted concepts and image modalities achieved a lower MAP than our textual baseline, which used the verbatim topic descriptions as queries.

9 Conclusion

This article describes the methods and results of the Communications Engineering Branch, a division of the Lister Hill National Center for Biomedical Communications, for the ImageCLEF 2011 medical retrieval track. We submitted ten runs each for the modality classification task and the ad-hoc image and case-based retrieval tasks. For the modality classification task, our best submission, a mixed

approach, achieved a classification accuracy of 74% and was ranked within the submissions from the top three groups. For the retrieval tasks, our results were lower than expected but reveal the mixed approaches involving clustered visual features to be promising methods for combining textual and visual image features.

References

1. Arthur, D., Vassilvitskii, S.: k-means++: The advantages of careful seeding. In: Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms. pp. 1027–1035. SODA '07 (2007)
2. Chang, S.F., Sikora, T., Puri, A.: Overview of the MPEG-7 standard. *IEEE Transactions on Circuits and Systems for Video Technology* 11(6), 688–695 (2001)
3. Chatzichristofis, S.A., Boutalis, Y.S.: CEDD: Color and edge directivity descriptor: A compact descriptor for image indexing and retrieval. In: Gasteratos, A., Vincze, M., Tsotsos, J.K. (eds.) Proceedings of the 6th International Conference on Computer Vision Systems. Lecture Notes in Computer Science, vol. 5008, pp. 312–322. Springer-Verlag Berlin Heidelberg (2008)
4. 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 Services. pp. 191–196 (2008)
5. Demner-Fushman, D., Lin, J.: Answering clinical questions with knowledge-based and statistical techniques. *Computational Linguistics* 33(1), 63–103 (Mar 2007)
6. Duda, R.O., Hart, P.E., Stork, D.G.: *Pattern Classification*. John Wiley & Sons Ltd. (2001)
7. Hastie, T., Tibshirani, R.: Classification by pairwise coupling. *The Annals of Statistics* 26(2), 451–471 (1998)
8. 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* 1(3), 253–263 (2007)
9. Kalpathy-Cramer, J., Müller, H., Bedrick, S., Eggel, I., de Herrera, A., Tsirikka, T.: The CLEF 2011 medical image retrieval and classification tasks. In: CLEF 2011 Working Notes (2011)
10. Kittler, J., Hatef, M., Duin, R.P.W., Matas, J.: On combining classifiers. *IEEE Transactions on Pattern Analysis* 20(3), 226–2329 (1998)
11. Lindberg, D., Humphreys, B., McCray, A.: The unified medical language system. *Methods of Information in Medicine* 32(4), 281–291 (1993)
12. Rahman, M.M., Antani, S., Thoma, G.: A medical image retrieval framework in correlation enhanced visual concept feature space. In: Proceedings of the 22nd IEEE International Symposium on Computer-Based Medical Systems (2009)
13. Simpson, M., Rahman, M.M., Demner-Fushman, D., Antani, S., Thoma, G.R.: Text- and content-based approaches to image retrieval for the ImageCLEF 2009 medical retrieval track (2009)
14. Simpson, M., Rahman, M.M., Singhal, S., Demner-Fushman, D., Antani, S., Thoma, G.: Text- and content-based approaches to image modality detection and retrieval for the ImageCLEF 2010 medical retrieval track (2010)