Extraction of Endoscopic Images for Biomedical Figure Classification

Zhiyun Xue, Daekeun You, Suchet Chachra, Sameer Antani, L. Rodney Long, Dina Demner-Fushman, George R. Thoma
National Library of Medicine, NIH, Bethesda, MD

ABSTRACT

Modality filtering is an important feature in biomedical image searching systems and may significantly improve the retrieval performance of the system. This paper presents a new method for extracting endoscopic image figures from photograph images in biomedical literature, which are found to have highly diverse content and large variability in appearance. Our proposed method consists of three main stages: tissue image extraction, endoscopic image candidate extraction, and ophthalmic image filtering. For tissue image extraction we use image patch level clustering and MRF relabeling to detect images containing skin/tissue regions. Next, we find candidate endoscopic images by exploiting the round shape characteristics that commonly appear in these images. However, this step needs to compensate for images where endoscopic regions are not entirely round. In the third step we filter out the ophthalmic images which have shape characteristics very similar to the endoscopic images. We do this by using text information, specifically, anatomy terms, extracted from the figure caption. We tested and evaluated our method on a dataset of 115,370 photograph figures, and achieved promising precision and recall rates of 87% and 84%, respectively.

Keywords: image modality classification, content-based image retrieval, endoscopic image figure extraction, ophthalmic image figure extraction

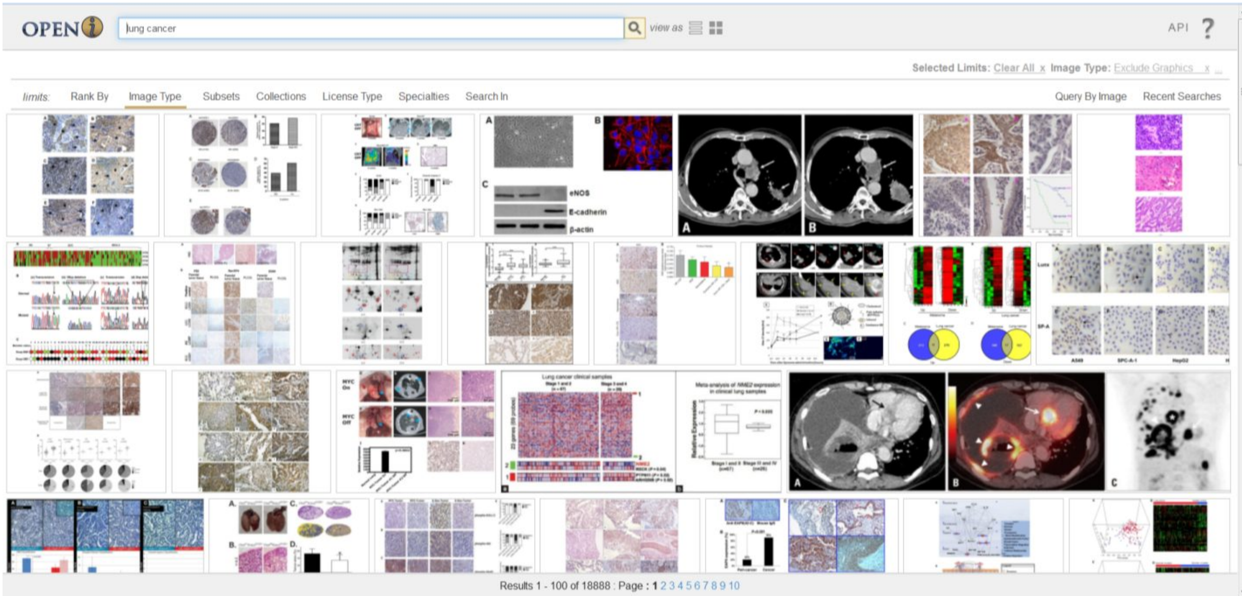
1. INTRODUCTION

OpenI[1] is a biomedical literature searching system developed by the U.S. National Library of Medicine (NLM). It allows users to search figure images from biomedical articles using either keywords or a query image. Currently OpenI provides open access to nearly 1.7 million images from approximately 584,000 open access biomedical research articles obtained from the NLM's PubMed Central (PMC) repository. Each day, there are about 40,000 OpenI visitors and over a million pages are accessed. Figure 1 has screenshots of the system interface, showing "query by keyword" (Figure 1a), and "query by image" (Figure 1b). OpenI also provides various filters (listed on the top region of the interface) to limit the search space category. One such filter is *image type* (*modality*). Successfully filtering the search results by image modality is not only a desirable feature for end users, but also may significantly improve retrieval performance of the system. We have been investigating various approaches for figure modality classification [1-3] and have adopted a hierarchical classification strategy, which has been shown to achieve better performance than a flat classification strategy [2]. Therefore, in OpenI, figures are first classified into regular figures or graphical illustrations (e.g., charts, sketches, graphs). Then the regular figures are classified based on the image modality, such as MRI, CT, X-ray, Ultrasound, or photograph.

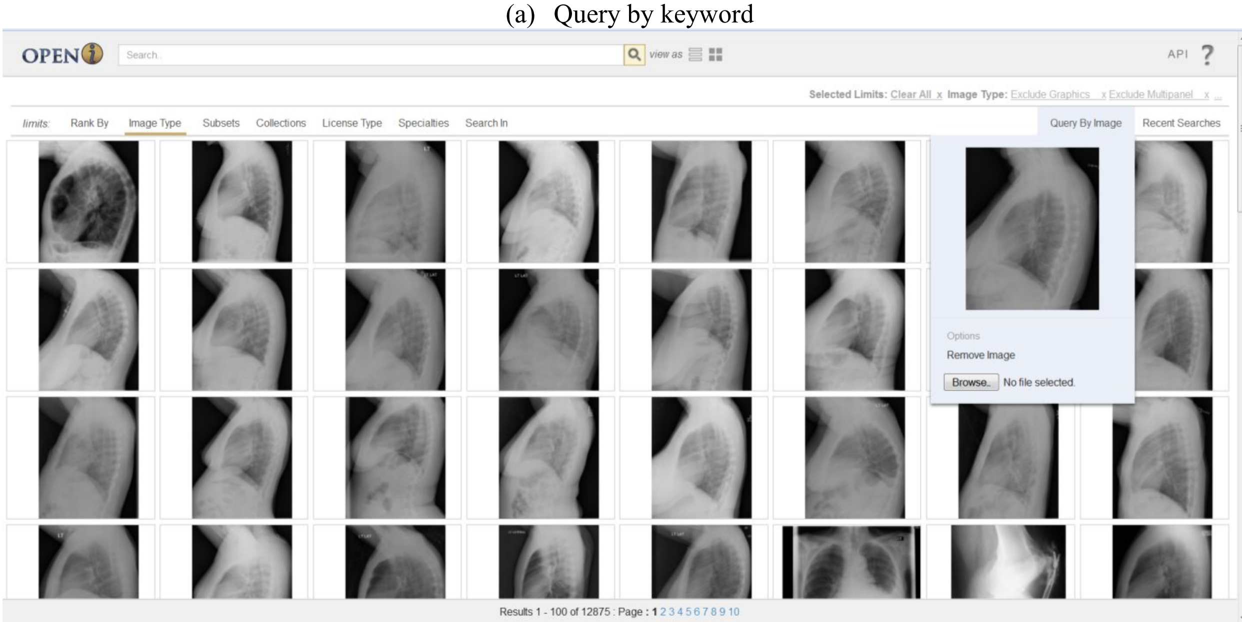
In this paper, we focus on the further classification of the category of *photograph* (visible light image). The content of photographic figures is very diverse and there are many non-clinical photos in the OpenI database. However, there are some medical images among them that we are interested in and hope to be able to extract. Two such medical image types are dermatology (skin) images and endoscopic images, as indicated by the hierarchical image modality classification model proposed by ImageCLEF [4], which is shown in Figure 2. We have been developing a method for extracting images in which a majority of the area in the image is skin-color tissue (we call this type of image a *tissue image*). As demonstrated in Figure 3, it includes various types of skin tissue images, such as images of the human face, mouth, open-wound tissue, skin at various other areas on the body surface, and endoscopic images of inner body surfaces. After carrying out a visual examination of the tissue images that we had automatically extracted from 115,370 photographs, we identified a characteristic common to the majority of the endoscopic images, although the visual content of the

[1] http://openi.nlm.nih.gov

Medical Imaging 2015: PACS and Imaging Informatics: Next Generation and Innovations,
edited by Tessa S. Cook, Jianguo Zhang, Proc. of SPIE Vol. 9418, 94180P
 · doi: 10.1117/12.2081033



(a) Query by keyword



(b) Query by image

Figure 1. OpenI system interface

endoscopic images varies greatly (showing epithelial surfaces of different locations in esophagus, stomach and colon). This common characteristic is shown in Figure 3(b): the endoscopic images are enclosed by an approximately circular shape, or part of such a shape, on a dark background. This seemingly simple geometric feature was the initial focus of our investigation. In spite of its simplicity, however, we learned that developing a high-performing endoscopic image detection method is not as straightforward as it might appear: many difficulties are encountered, due to the intrinsic large variety in endoscopic images in the biomedical literature. For example, Figure 4 shows several sample images where the outline of the endoscopic region is not entirely round. The boundary of the endoscopic region may be blurred. The illumination may not be uniform. Additionally, there are text overlays, arrows, or bright bounding borders in some figures that may interfere with the extraction of the boundary of the endoscopic region. Each of these issues poses a challenge to the task of successfully and reliably extracting the image. In addition, another important type of medical image – ophthalmic photographs – also exhibit the same visual characteristics as endoscopic images; as shown in Figure

3(j), ophthalmic photographs may have roundish outlines, and may be easily misclassified as endoscopic images, since their color is close to epithelial tissue color. To filter out these ophthalmic images, we used text information associated with the figure. Therefore, the side benefit of this work is that this method can also extract ophthalmic images – another important class of medical image modality.

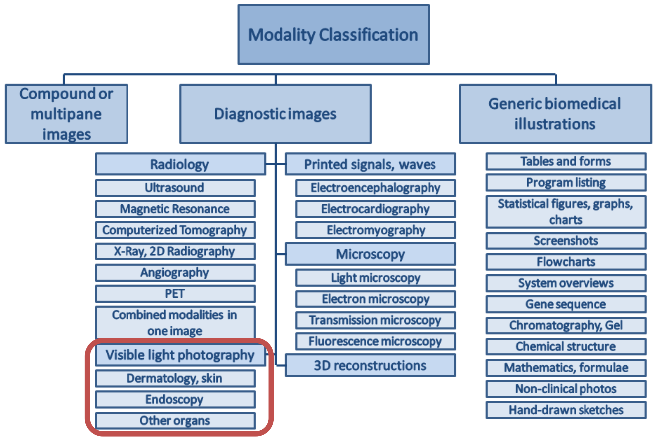


Figure 2. ImageCLEF modality hierarchy [4]

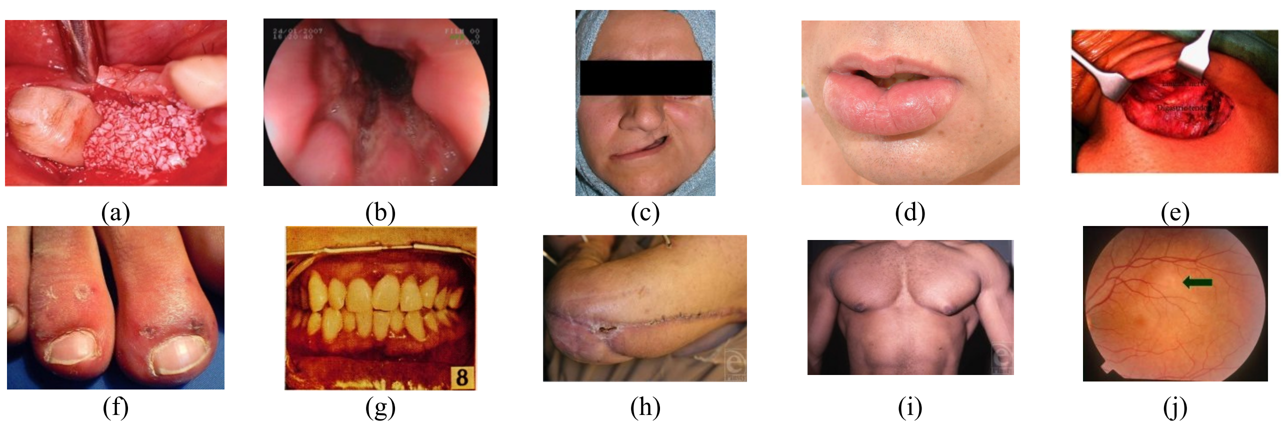


Figure 3. Skin-color tissue images

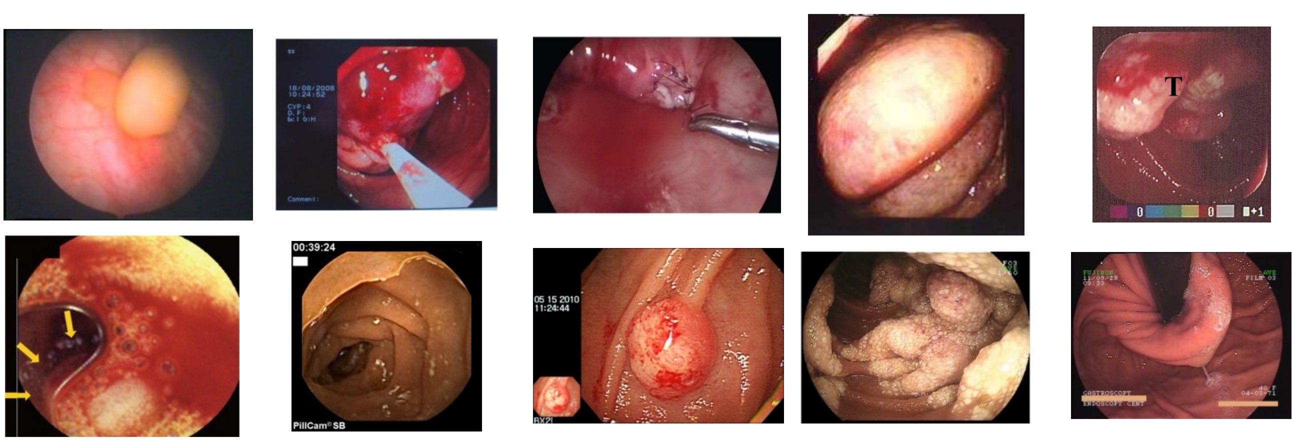


Figure 4. Endoscopic image figures

The rest of the paper is organized as follows: Section 2 describes the method we developed to extract tissue images from photographs. Section 3 presents the method for extracting candidate endoscopic images from these tissue images. Section 4 explains the method for separating ophthalmic images out using textual information extracted from figure captions. Section 5 shows experimental results and discussion. Section 6 concludes the paper with future work.

2. TISSUE IMAGE EXTRACTION

Tissue image in our research denotes a photographic image with tissue regions greater than a pre-determined threshold (e.g., 50%) of the entire image. Hence defining *tissue region* is the first step for separating tissue regions from non-tissue regions in images and further tissue images from non-tissue images. We develop a tissue region detection method based on Markov random field (MRF) theory that is applied to label image patches. Images are split into non-overlapping patches and each patch is assigned one semantic label: *tissue*, *bright*, *dark*, or *photo*. Each label corresponds to a different type of region in the image. The *tissue* region is used to categorize images into two groups, tissue and non-tissue. Our tissue image extraction method has been published in [5]. Please refer to [5] for details of the algorithm. As a summary, each step in the algorithm is briefly described below.

Image resizing and patch creation
Each image is resized by setting its smaller side (either width or height) to 100 pixels while preserving the aspect ratio of the original image. We found resizing to be effective in reducing the computation time in our MRF model. Resized images are then split into non-overlapping patches of size 10×10 pixels.

Initial patch classification
A three dimensional CIELUV color vector is computed from the input RGB color channel values at each patch pixel. These vectors are concatenated to form a 300-dimension patch feature vector. Patches from training samples are clustered into 525 clusters (empirically selected) by K-means clustering with the feature vectors. An unknown patch is classified by finding the nearest cluster center to the patch feature in the feature space, and the numeric cluster label (0-524) is assigned to the patch as its initial label.

MRF model for relabeling patches
Patches are relabeled from their initial numeric labels to one of the four semantic labels using a Markov random field (MRF) model. Two types of compatibility functions are defined to model contextual dependencies that exist among neighboring patches. They are unary and binary compatibility functions that are defined by visual characteristic of a single patch and visual similarity between the patch and its neighboring patches, respectively. An optimal labeling configuration (labels of the patches) is found by belief propagation that passes local messages between patches and updates their labels iteratively based on the propagated messages. Details of our MRF model can be found in [6].

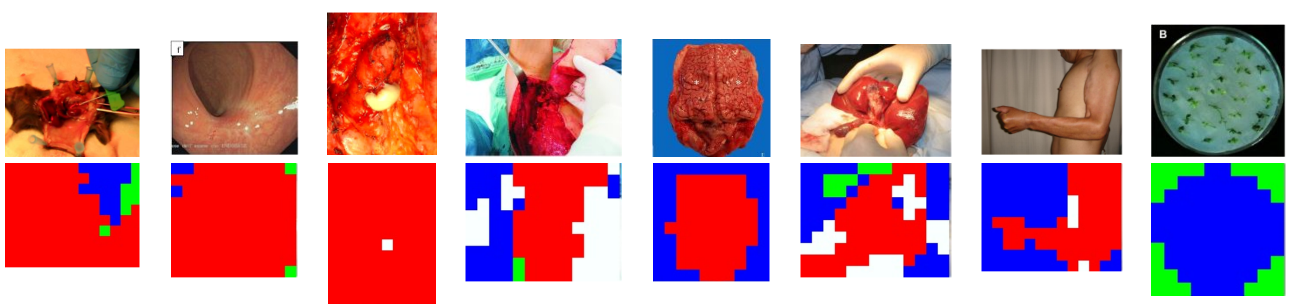


Figure 5. Image segmentation results (red: tissue, blue: photo, white: bright, green: dark)

Figure 5 shows several example results of tissue region detection. Essential conditions that are applied to separate tissue images from non-tissue images are as follows. The percentage threshold is applied to the entire image region.

(a) Condition 1: *tissue* region > 30%
(b) Condition 2: *tissue* region + *bright* region + *dark* region > 80%
(c) Condition 3: *tissue* region > *photo* region

In the condition 2, we add *bright* and *dark* regions to the *tissue* region, since tissue images are frequently accompanied with dark background (e.g., endoscopy images) or bright regions (e.g., teeth in mouth images). We considered two thresholds, 80% and 50%, for condition 2. With 80% threshold, only the left three images in Figure 5 are detected as tissue images. With 50% threshold, the next three images are detected. The last two images are categorized as non-tissue images. As shown in the segmentation results, *tissue* regions are accurately separated from other regions.

3. ENDOSCOPIC IMAGE CANDIDATES EXTRACTION

Once tissue images have been extracted, our goal is to filter out the endoscopic images. The method we propose has two main steps: 1) region-of-interest (ROI) boundary extraction and 2) ROI shape measurement.

3.1 ROI boundary extraction

The performance of the ROI boundary extraction is crucial since it affects the next step, ROI shape measurement. In our dataset of figures from biomedical articles, we have observed a large variation with respect to the size, shape, and color of the endoscopic ROIs. There may be graphical annotations, such as text and arrows, sometimes presented in the images and overlaid on the boundary of the endoscopic ROI. The border of the endoscopic ROI may blend with the background or connect with other objects. The illumination may not be uniform, and the dynamic range may be quite different across images. To address these problems, we combine techniques of contrast enhancement, edge detection, thresholding, and morphological processing. Figure 6 shows a diagram of the procedure.

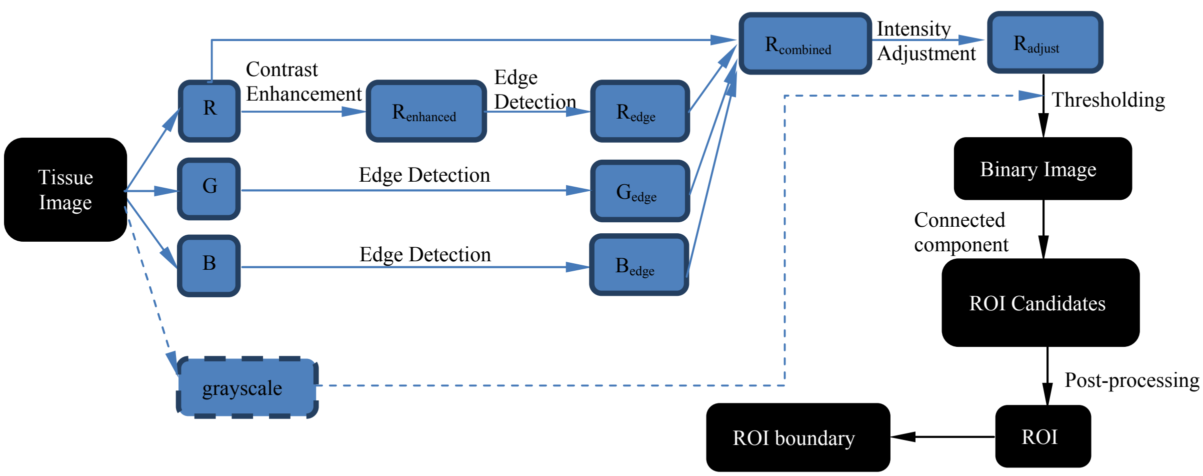
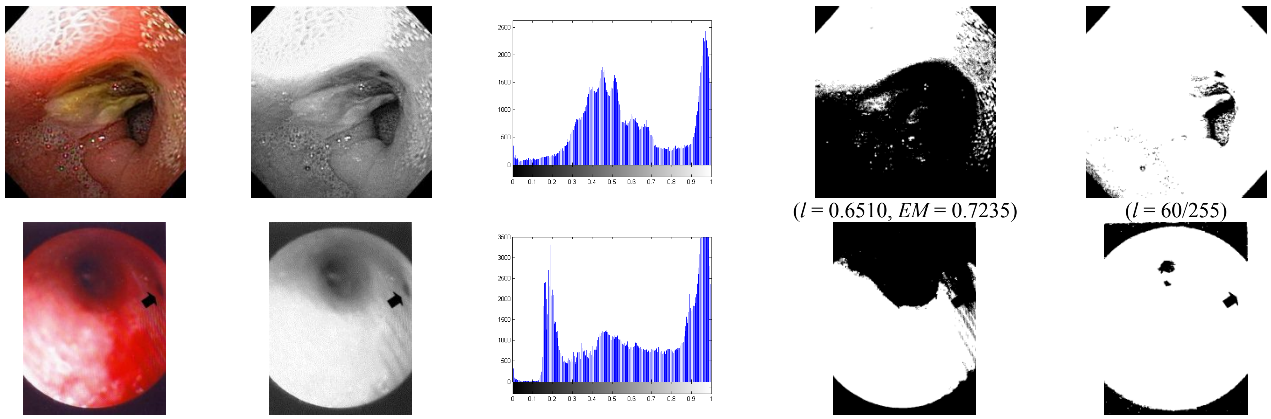


Figure 6. Diagram of the ROI extraction method

We first apply an intensity threshold to convert the input image to binary form, since the background area is dark compared to the endoscopic region. As shown in the diagram in Figure 6, there are two paths (denoted by blue dotted lines and blue solid lines) before applying the threshold. One is to simply convert the input color image into a grayscale image and then apply the threshold (blue dotted lines). The other uses a series of pre-processing steps before the application of the threshold (blue solid lines). These pre-processing steps are designed to process the difficult cases when the simple path fails. We will explain these pre-processing steps shortly after the discussion of the method for applying the threshold. For thresholding, we found that using the global threshold (for example, 60/255) outperformed an adaptive value computed by Otsu's method [7]. That is because the Otsu's method chooses a threshold for each image based on maximizing the measure of class separability in gray levels. When there are small numbers of background pixels, the threshold obtained may put foreground pixels whose intensity is closer to that of the background pixels (and are very often located on or adjacent to the boundary of the endoscopic region) into the same group as the background pixels. Figure 7 shows two examples in which the results of Otsu thresholding (Figure 7(d)) are much less satisfactory than those of constant thresholding (Figure 7(e)), l is the value of the threshold and EM is the value of the effective measure output by Otsu's method.



(l = 0.6510, EM = 0.7235) (l = 60/255)

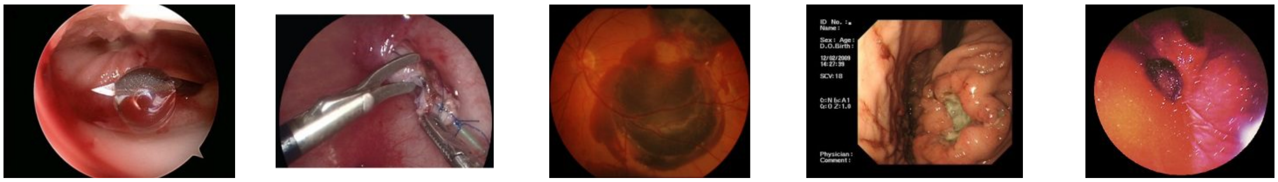
(l = 0.6157, EM = 0.7972) (l = 60/255)

(a) Original image (b) R channel (c) Histogram (d) Otsu (e) Contant

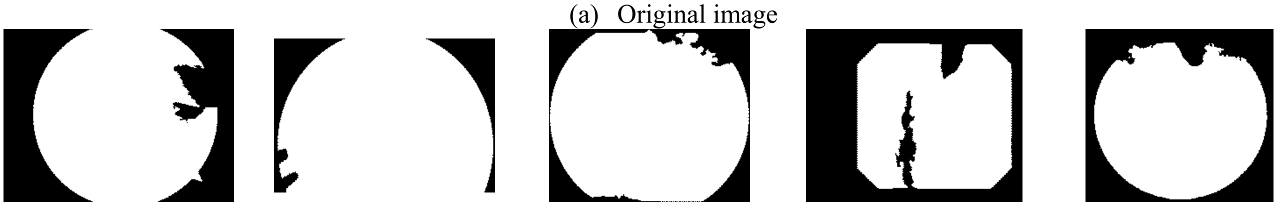
Figure 7. Thresholding results

Although applying a fixed threshold works for many cases, it fails sometimes due to the large variety of the dataset as well as the other challenges discussed previously. To increase the performance of the binarization, we also employed several pre-processing methods as displayed by the blue solid lines in Figure 6. The endoscopic tissue is reddish/pinkish, therefore we used the red channel of the tissue image (R) as the main grayscale image for extracting the ROI; for these images, this channel generally provides better contrast to the background than the other two channels (G and B). This red channel image is then binarized to obtain the candidates of ROIs by thresholding. We first applied Canny edge detector [8] to extract significant edges from each individual channel (R, G, B) grayscale image. Then we added these three edge images to the R image to get image $R_{combined}$. We will discuss the value of this step later. We tested two ways of adding edges. One is to make those edge pixels to be white in the image $R_{combined}$. The other is to set the intensity of each edge pixel to be the complement of the maximum intensity of its neighborhood pixels in the R image. The latter method may be helpful at the ROI binary thresholding step in the cases where the ROI connects to a white background. However, although this latter method helps for these cases, it might worsen the overall ROI extraction performance because of the existence of many edges inside the ROI, which may separate the ROI into multiple regions. Because of this consideration, we adopted the first method. We provide additional explanation in our discussion of the post-processing step to handle bright borders in figures. We applied contrast-limited adaptive histogram equalization (CLAHE) [9] to the R image before edge detection. We then adjusted the intensity values of the grayscale image $R_{combined}$ to create the image R_{adjust} , using a map that takes $R_{combined}$ values to the values in R_{adjust} such that 1% of the pixel data is saturated at low and high intensities of $R_{combined}$. This not only increases the contrast in R_{adjust} but also makes the background area darker and the endoscopic region brighter, which is propitious for the thresholding step to follow. Figure 8 shows several ROI extraction results with fixed value threshold (l = 60/255), where it can be seen that the results with the pre-processing steps are much more precise than those without the pre-processing. As demonstrated by Figure 8, the edge detection step in pre-processing aims to preserve the boundary which may not be separated by fixed value thresholding.

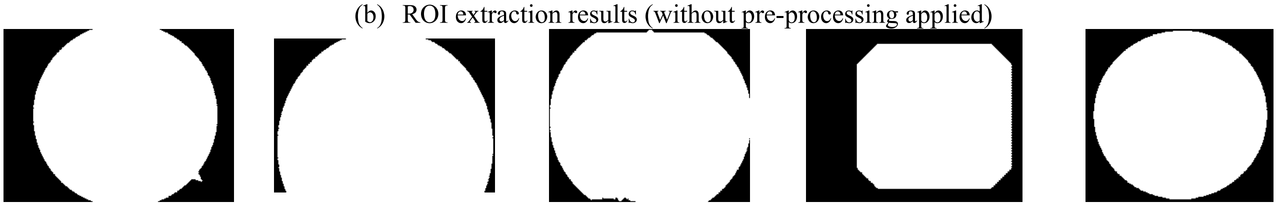
After binarization, the region having the largest area is extracted and regarded as the candidate ROI. Next, two post-processing steps are applied before extracting the ROI boundary. Step 1: we check to see if the binary ROI candidate has a white rectangular border and, if so, we remove it. For figures in the literature, it is not uncommon that there is a bright band on the periphery of the figure, as in the example given in Figure 9(a) (the border of the figure is enclosed with black lines in order to make it easy to recognize the white border area). If these borders are not removed, they will be identified as part of the ROI boundary which will in turn cause failure in the subsequent shape roundness detection. The border removal is done as follows. We first crop the binary image based on the bounding box of the black regions. Then, after cropping off these borders, we check the four corners of the cropped binary image for white pixels; if the corners contain white pixels, we iteratively crop the borders up to a small number (empirically determined) of times/lines. This



(a) Original image

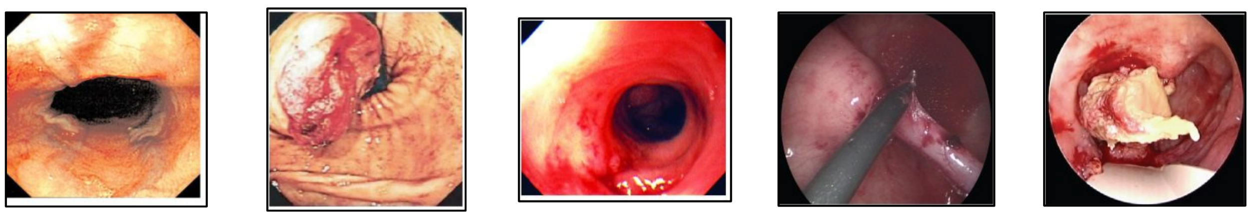


(b) ROI extraction results (without pre-processing applied)

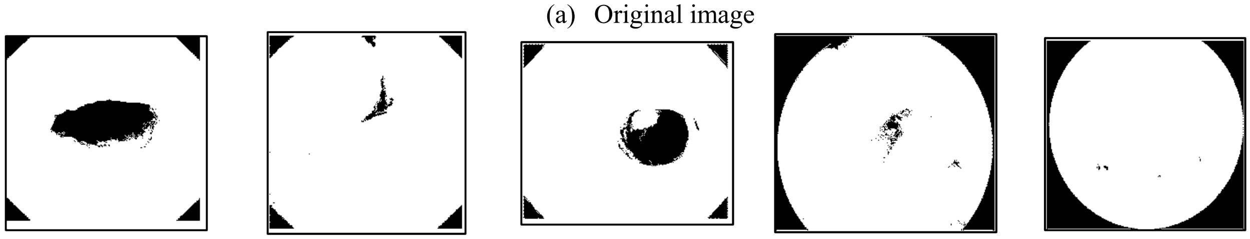


(c) ROI extraction results (with pre-processing applied)

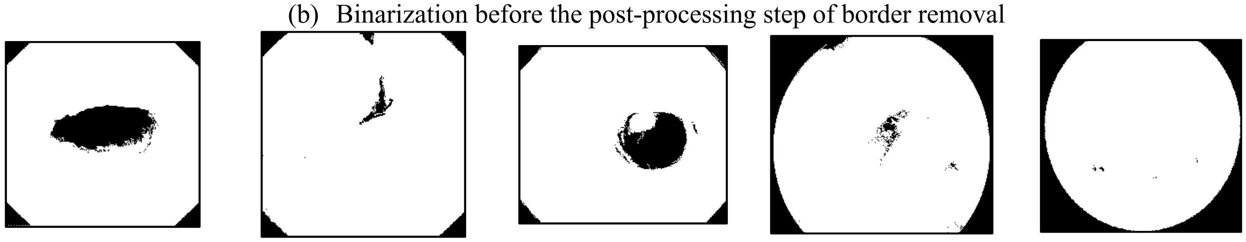
Figure 8. Benefits of pre-processing



(a) Original image



(b) Binarization before the post-processing step of border removal



(c) after the post-processing step of border removal

Figure 9. Removal of bright border lines

procedure is to address the issue that the white border may be a little bit tilted, in which case the method of border removal using the bounding box of the black regions will not remove all the white border pixels. Figure 9(b) and Figure 9(c) show the results of binarization with and without border removal, respectively, for several example images. Step 2: after white border removal, we apply morphological image processing to the ROI. For this processing, instead of applying "hole-filling", we first extract the four background corners by applying the flood fill algorithm and then get the complementary binary image. This aims to deal with the cases in which the dark endoscopic region connects to the image boundary. Two examples are shown in Figure 10. As shown in Figure 10(b), the dark region that connects to the boundary is not filled by using morphological hole filling (as a result, the boundary of the ROI does not enclose the whole endoscopic region), but will be enclosed within the ROI area by using the method of flood filling from the four corners (as shown in Figure 10(c)). We also apply morphological opening to remove thin lines generated by the pre-processing step of edge enhancement.

(a) original image (b) using hole filling (c) boundary of (b) (d) using flood fill from corners (e) boundary of (d)

Figure 10. ROI extraction using flood fill from corners

3.2 ROI shape measurement

As discussed in Section 1, the dataset we work on has high variability, and the endoscopic region inside an image is very often cropped, that is, it does not have an entirely round outline, as demonstrated by Figure 4. Therefore, the roundness measurement needs to accommodate this type of complex input, as well as non-perfect segmentation results. We considered shape measures including the minimum bounding circle, the least square fitting circle, and the maximum enclosing circle of the extracted ROI region. Figure 11 gives several examples, in which the corresponding minimum bounding circle (in red color), least square fitting circle (in blue color), and maximum enclosing circle (in green color), are overlaid on the ROI. Note that, in the second row of Figure 11, the binary ROI was first cropped based on the ROI's bounding box and then was padded on the boundary in order to show the minimum bounding circle (therefore they have different size as their original images). As indicated by Figure 11, the measures based on the minimum bounding circle, the least squares fitting circle, and the maximum enclosing circle, may not be suitable for our application. Cases like those given in Figure 11 tend to be representative in the figures of our dataset, and we may reasonably assume many failures in endoscopic region shape measurement under such measures. Therefore, instead of using those measures, we adopted the following rule-based procedure:

1) Black corner rule
   All the four corners of the cropped binary ROI (based on its bounding box) should be black regions. This is done by checking if the 3 x 3 region in each corner is black.
2) Shape rule
   The roundness of the ROI is measured by
   $$A = 4 \times \pi \times \frac{area}{perimeter^2} > 0.85$$
3) Symmetric rule
   The ROI needs to be symmetric either horizontally or vertically. This is measured by flipping the cropped binary ROI horizontally/vertically along the central line and calculating the area of the un-matching region. The area of the un-matching region must be less than 5% of the total ROI area for the ROI to be considered symmetric.

If the ROI satisfies all above three rules, then the image is flagged as a *candidate endoscopic image*.

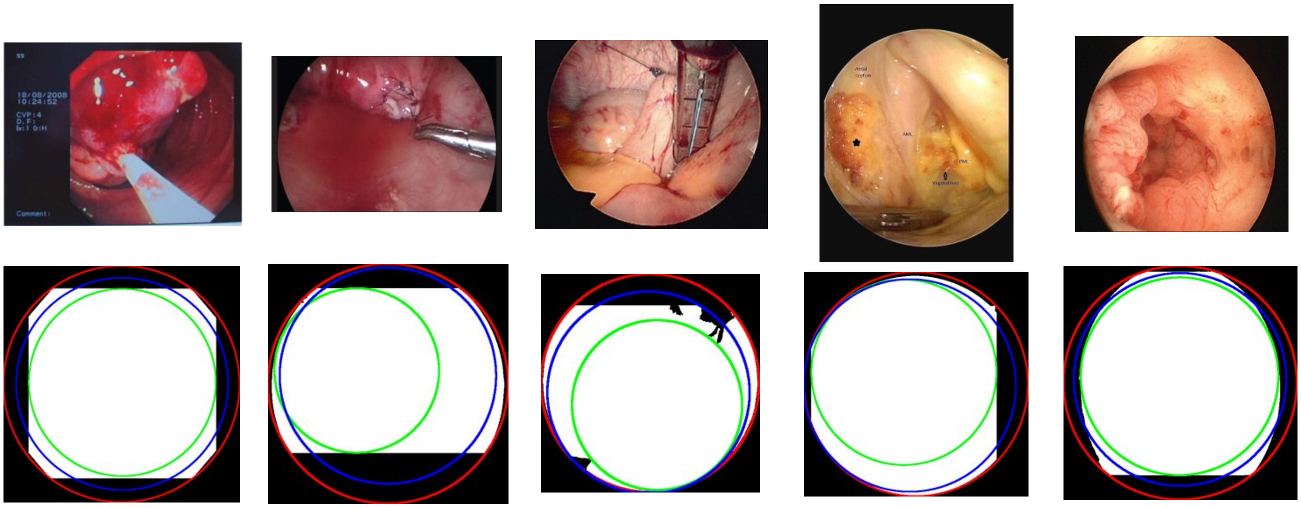


Figure 11. The minimum bounding circle, the least square fitting circle, and the maximum enclosing circle

4. FILTERING OF OPHTHALMIC PHOTOGRAPHS

The method for candidate endoscopic image extraction described in Section 3 is mainly based on the shape information. It does not utilize any text content related to the image. Therefore, the output candidate images contain many ophthalmic photos because they share a shape outline similar to endoscopic images, as shown in Figure 12. To filter out those images, we exploit the text information provided by the figure caption.

The figure caption, wherever present, provides description of the content of the associated biomedical figure. It may also sometimes include additional information which may not be directly inferred via visual analysis of the biomedical figure such as a relevant biomedical concept or information about a certain scientific procedure being photographed. Our textual approach to filter the ophthalmic photographs focuses on identifying key anatomy concepts in the figure captions using the Unified Medical Language System® (UMLS®) Metathesaurus®. The UMLS Metathesaurus is a large, multi-purpose, and multi-lingual thesaurus that contains millions of biomedical and health related concepts, their synonymous names, and their relationships. The Metathesaurus is organized into a set of relational files containing biomedical concepts. These files also link alternative names and views of the same concept from different vocabularies and identify useful relationships between different concepts.

The first step in acquiring anatomy terms consists of analyzing the figure caption and detecting the presence of concepts in the UMLS semantic group Anatomy [10]. This is done by searching the captions text for the anatomical UMLS concepts. This step is similar in part to concept recognition using MetaMap, a tool that maps biomedical text to the UMLS Metathesaurus using a knowledge-based natural language processing approach [11]. Sometimes the concepts extracted from the captions are too specific and result in a sparse concept vocabulary, which makes it difficult to detect and filter out all ophthalmic images. To overcome this sparsity, once the associated figure caption is analyzed and the anatomical concepts are extracted, we expand the extracted concepts to more general anatomical concepts. This is done by utilizing the relationship information provided by the UMLS Metathesaurus, and by registering the ‘is-a’ and ‘part-of’ relation between anatomical concepts. This provides us with not only the concepts that are directly detectable within the figure concept, but also with their parent concepts.

After we have extracted the anatomical concept information from the figure caption, the next step is to determine if the concepts indicate that the image under consideration is an ophthalmic image. We use the following three sets of keywords to filter out the ophthalmic photos:

1) Keyword set I: *eye*, *eyeball*, *retinal*, *retina*, *retinitis*, *iris*, *optic nerve*, *cranial nerve*, *optic neuropathy*, *macula*, *retinopathy*
2) Keyword set II: *funduscopy*, *fundoscopy*, *funduscopic*, *fundoscopic*, *fundus photography*, *retinal image*, *retinal photograph*, *optic disc*
3) Keyword set III: *endoscopic*, *endoscopy*

If the extracted concepts of the figure contains any of the keywords in set I or the figure caption contains any of the keywords in set II but the figure caption contains none of the keywords in set III, then the figure is categorized as an ophthalmic image.

Figure 12. Ophthalmic photographs

5. EXPERIMENTAL RESULTS AND DISCUSSION

5.1 Data

We downloaded 115,370 images from OpenI system. These images are those that are returned by OpenI by selecting "Photographs" and "Exclude Multipanel" in the *Image Type* (i.e. images that are automatically classified into *single-panel photograph* modality by the OpenI system). These images contain not only true photographs but also several types of noisy images, such as radiography, microscopy, and illustration figures. We use a subset of these images to train and test our method for extracting tissue images. Then we apply the tissue image extraction method to the entire dataset of 115,370 photographic images. Finally, we extract the endoscopic images out of those tissue images using the proposed method described in Section 3 and 4. We report the results of tissue image extraction, endoscopic image candidate extraction, and ophthalmic image filtering in the following. All ground truth classifications described below were done by manual inspection by two engineers (non-medical experts) with several years' experience in medical image processing.

5.2 Result of Tissue Image Extraction

We used an image set consisting of 1,997 photograph images as a training dataset for creating patch clusters and training parameters for the MRF model. Out of the 1,997 training images, we manually selected 408 tissue images (most of the image regions in these images are tissue regions). We randomly selected 23,000 out of the 115,370 photographs as a testing dataset for evaluating the performance of our tissue image extraction method. The testing dataset doesn't include any images used in the training set. The proposed method was applied to the testing set to extract tissue image and we then manually examined the results to compute precision of the extraction. The results are listed in Table 1. Two thresholds, 80% and 50% as discussed in the section 2, were used to set the minimum area of tissue region in determination of tissue and non-tissue images. Table 1 lists the number of extracted images, the number of true positive images, the number of false positive images, and the precision rate for each threshold. A recall rate could not be obtained since we did not create ground truth for the entire 23,000 image test set. As shown in the result, a smaller threshold (50%) detected about 40% more tissue images than a larger threshold (80%) with insignificant precision decrease. We believe that the threshold of 50% is suitable for our tissue image detection and classification.

Table 1. Tissue image extraction result

Threshold	Extracted tissue images	True positive images	False positive images	Precision (%)
80%	3,359	2,757	602	82.1%
50%	4,628	3,788	840	81.8%

5.3 Result of Endoscopic Image Extraction

We applied the tissue image extraction method (using the threshold of 50%) to the entire dataset of 115,370 photographs and extracted 14,510 tissue images. Then, we employed the method described in Section 3 to get the candidate endoscopic images out of these tissue images. There are 1242 candidate images. The candidate images are further filtered by separating out ophthalmic photos. To evaluate the performance, we generated ground-truth data by visually examining all the 14,510 tissue images and identifying all the endoscopic images (890 images) that have four dark background corners around the endoscopic region in the image (using both visual and caption information). Therefore, cases such as those shown in Figure 13, which are zoomed in/cropped too much and (almost) do not have background region outside endoscopic region, will not be counted. Table 2 lists the results of endoscopic image detection with and without the step of ophthalmic image filtering, respectively. The precision is increased significantly (from 61.2% to 87.3%) after sorting these photos out with a slightly drop of recall (from 85.4% to 84.7%). This indicates the large quantity of ophthalmic images that have similar shape/color characteristics as endoscopic images and also demonstrates the effectiveness of our method for detecting ophthalmic images.

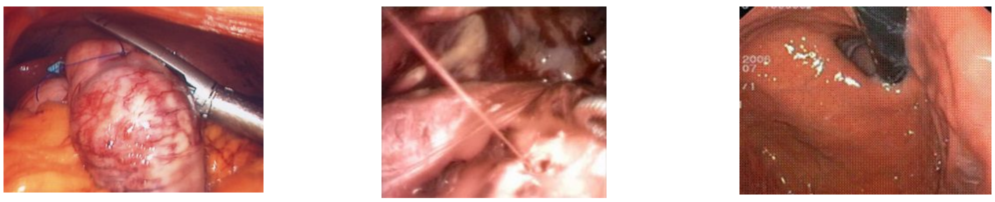


Figure 13. Endsocopic images that are not considered because of unusual zoom or cropping

Table 2. Endoscopic image extraction result

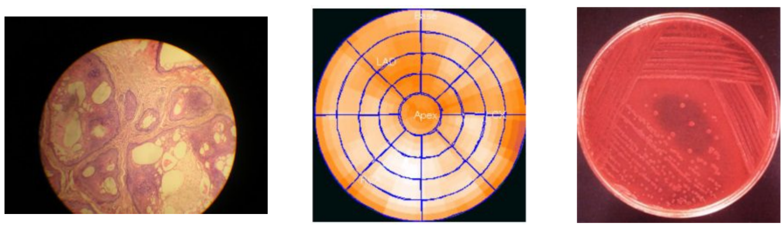
	Extracted images	True positive images	False negative images	Precision	recall
Without ophthalmic photo filtering	1242	760	130	61.2%	85.4%
With ophthalmic photo filtering	864	754	136	87.3%	84.7%

5.4 Discussion

The proposed method exploits color and shape information to extract endoscopic images from figures in biomedical articles. It also uses text information, such as anatomy concepts obtained from figure captions, to filter out unwanted images and boost the algorithm performance. (In fact, we found that this method is also a very effective approach for extracting ophthalmic photos, although for the study of this paper, they are considered unwanted images). The method has been shown promising as demonstrated by the experimental results given in Table 1 and Table 2. For further performance improvement, we examined the false positive and false negative images, several of which are displayed in Figure 14.

The false positive examples are extracted due to the similarity of their shape and color to those of endoscopic images (in this respect, similar to the ophthalmic photos). Filtering out these images likely requires the use of semantic information about the image content. In addition to filtering out unwanted images using additional textual information (like using the keyword *histopathology* for removing the first case in Figure 14(a)), in the future we will explore the option of extracting desired image classes using textual information. We have already carried out one experiment on retrieving endoscopic images by text-only searching with the keywords of *endoscopic* and *endoscopy* in figure captions. Out of 14,510 tissue images, only 486 endoscopic images were retrieved. This is much lower than that of the proposed method which uses the image visual information first. Therefore, we may need to identify more candidates by expanding the keywords with anatomy terms which refer to medical "targets" of the endoscopy, such as esophagus, stomach, or colon and then filter the results by using image visual information.

Our false negative examples are mainly caused by the unsuccessful binarization/segmentation results due to the intensity of part of the endoscopic region boundary area being very close to background intensity. We will explore the idea of resetting the threshold value for each image using the intensity of the background area identified by the first round of ROI extraction if it did not pass the shape measurement criteria; for example, for the second case in Figure 14(b), a lower threshold of 20/255 gives improved results. The method described in this paper uses rules to measure the shape information extracted from the ROI segmentation result in the endoscopic images. We plan to test classification-based methods using shape descriptors extracted from the entire endoscopic image and compare the performance of those methods to that of our algorithm. We have previously developed a contour-based shape descriptor [10], which we adapted from a feature extraction method used in optical character recognition for character shape representation. We have applied this descriptor to radiology images (e.g., MRI, CT, ultrasound, X-ray) to evaluate its usefulness for modality classification. We will apply this shape descriptor and other state-of-the-art shape descriptors in the classification-based method.



(a) False positive cases

(b) False negative cases

Figure 14. False positive and false negative cases

6. CONCLUSIONS

Because of the prevalence and the rich information contained in medical images, their search and retrieval within the peer-reviewed biomedical literature has become an important research topic. For a biomedical literature figure searching system such as OpenI, image modality classification is one essential component for achieving high retrieval performance. This paper focuses on the topic of detecting endoscopic figures, a modality for which, to the best knowledge of the authors, there is no work specifically reported in the literature. The proposed method consists of three major steps: extraction of tissue images based on color information, extraction of endoscopic image candidates based on shape information, and identification and removal of ophthalmic images from the endoscopic image candidates based on textual information. We have combined a set of image processing techniques as well as text information from figure captions to boost the method's performance, aiming to cope with the high content, color, and shape variability of these figures. We evaluated our method on a large dataset of photographs categorized by OpenI and have shown that it achieves very promising results. Future work includes using textual information to augment the detection of endoscopic image candidates and developing a shape feature descriptor based classification method.

ACKNOWLEDGEMENT

This research was supported by the Intramural Research Program of the National Institutes of Health (NIH), National Library of Medicine (NLM), and Lister Hill National Center for Biomedical Communications (LHNCBC).

REFERENCES

[1] Demner-Fushman, D., Antani, S.K., Simpson, M., Thoma, G.R., "Design and development of a multimodal biomedical information retrieval system," JCSE, 6(2), 168-177 (2012).

[2] Simpson, M., You, D., Rahman, M.M., Demner-Fushman, D., Antani, S.K., Thoma, G.R., "ITI's participation in the ImageCLEF 2012 medical retrieval and classification tasks," CLEF 2012 Working Notes, Rome, Italy (2012).

[3] Rahman, M.M., Antani, S. K., Thoma, G. R., "A learning-based similarity fusion and filtering approach for biomedical image retrieval using SVM classification and relevance feedback," IEEE Transactions on Information Technology in Biomedicine, 15(4), 640-646 (2011).

[4] Müller, H., Kalpathy-Cramer, J., Demner-Fushman, D. and Antani, S.K., "Creating a classification of image types in the medical literature for visual categorization," SPIE Medical Imaging, San Diego, USA (2012)

[5] You, D., Antani, S.K., Demner-Fushman, D., and Thoma, G. R., "Biomedical Image Segmentation for Semantic Visual Feature Extraction," Proceedings of the IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Belfast, UK, November 2014

[6] You, D., Antani, S.K., Demner-Fushman, D., and Thoma, G. R., "A MRF model for biomedical image segmentation," Proceedings of the IEEE 27th International Symposium on Computer-Based Medical Systems (2014).

[7] Otsu, N., "A Threshold Selection Method from Gray-Level Histograms," IEEE Transactions on Systems, Man, and Cybernetics, 9(1), 62-66 (1979).

[8] Canny, J., "A Computational Approach to Edge Detection," IEEE Transactions on Pattern Analysis and Machine Intelligence, PAMI-8 (6), 679-698 (1986).

[9] Zuiderveld, K., "Contrast Limited Adaptive Histograph Equalization," Graphic Gems IV. San Diego: Academic Press Professional, 474–485 (1994).

[10] McCray, A.T., Burgun, A., Bodenreider, O., "Aggregating UMLS semantic types for reducing conceptual complexity," Stud Health Technol Inform (Proc Medinfo), 84(Pt 1), 216-220 (2001).

[11] Aronson, A.R., Lang, F.M. "An overview of MetaMap: historical perspective and recent advances", J Am Med Inform Assoc. 17(3), 229-236 (2010).

[12] You, D., Antani, S.K., Demner-Fushman, D., and Thoma, G. R., "A contour-based shape descriptor for biomedical image classification and retrieval," Proceedings of SPIE Document Recognition and Retrieval XXI, Vol. 9021 (2014)