

# Segmenting Compound Biomedical Figures into Their Constituent Panels

Pengyuan Li, Xiangying Jiang, Chandra Kambhamettu, and Hagit Shatkay

Department of Computer and Information Sciences,  
University of Delaware,  
Newark, DE, USA  
{pengyuan, shatkay}@udel.edu

**Abstract.** Many of the figures in biomedical publications are compound figures consisting of multiple panels. Segmenting such figures into constituent panels is an essential first step for harvesting the visual information within the biomedical documents. Current figure separation methods are based primarily on gap-detection and suffer from over- and under-segmentation. In this paper, we propose a new compound figure segmentation scheme based on Connected Component Analysis. To overcome shortcomings typically manifested by existing methods, we develop a quality assessment step for evaluating and modifying segmentations. Two methods are proposed to re-segment the images if the initial segmentations are inaccurate. Experiments and results comparing the performance of our method to that of other top methods demonstrate the effectiveness of our approach.

**Keywords:** Compound image separation; Biomedical image; Connected component analysis

## 1 Introduction

A fundamental task in biomedical informatics is to make information within documents available to researchers. Images convey essential information in biomedical publications. A few recent efforts started exploring the use of image information within biomedical documents [1,2]. However, many of figures within biomedical documents are compound images consisting of multiple panels, where each panel potentially carries a different type of information. To obtain the information embedded within each part of the image, it is essential to first segment each compound image into its constituent panels.

Current compound image segmentation methods are primarily based on finding gaps between panels [2–7]. The gaps, which are solid (typically white or black) bands in compound images, are commonly detected and used as panel separators. However, due to inconsistency in image quality gaps can be hard to detect, which leads to *under-segmentation*, that is, parts of the image may not be correctly segmented into individual panels. To overcome this issue, the image can be transformed, for instance via edge-detection [5,7], so that gaps are

more readily detected. Notably, some white/black bands occurring in images are not necessarily panel separators. Still, gap-based segmentation methods tend to interpret all solid bands as gaps, and as a result, erroneously split images into too many panels, to which we refer as *over-segmentation*. To address under- and over-segmentation, captions and image labels have been used to estimate the number of panels in compound images and to identify true gaps of separation [2,4,5]. However, such methods are not always effective, and may not even be applicable, when captions and labels are not available. Additionally, extracting labels from images requires optical character recognition – a time consuming operation. An alternative approach [6], applies several rules, eliminating gaps that are not panel separators aiming to avoid over-segmentation. While this method does not require processing image captions or labels, it is still time consuming. Furthermore, its separation accuracy leaves much room for improvement.

Unlike the above methods that segment images through gap detection, Shatkey *et al.* [1] proposed a method based on first identifying connected contents within individual panels. They used Connected Components Analysis (CCA) to detect individual panels in images. Lopez *et al.* [9] and Kim *et al.* [8] also used the same method for panel separation. Similar to the gap-based approach discussed earlier, CCA can also suffer from over-segmentation; unconnected small objects may be detected as individual panels and segmented off the main image-panel. Aiming to address a different task, namely the *identification of multi-paneled images*, Wang *et al.* [10] used a post-processing step by setting a threshold on panel-size to avoid fragmentation into very small panels. However, their work was not applied to the image-segmentation task, but rather aimed only to identify whether an image is compound or not. Notably, none of the above methods can segment *stitched compound images* whose panels are not separated by visible gaps. Santosh *et al.* [11] first proposed a method to separate stitched compound images based on straight lines detected in the images. Their method is applicable only to stitched compound images and as such relies on a manual selection step in which such images are identified within the dataset.

In this paper, we present a new CCA-based scheme for separating compound figures, including stitched compound images. To do this, we first introduce a preprocessing step to broaden and un-blur gaps in images. We then present the CCA method for segmenting images into panels. To avoid over- and under-segmentation, we extend our method by adding an assessment step to detect, evaluate and modify segmentation errors, and re-separate some of the images accordingly. The rest of the paper is organized as follows: Section 2 describes the complete framework of our method; in Section 3 we discuss experiments used to assess performance and present related results; Section 4 concludes and outlines directions for future work.

## 2 Methods

Our goal is to segment compound images appearing in biomedical documents. As noted above, compound images consist of several panels, typically separated

by gaps, which appear as vertical or horizontal light/dark bands; such gaps may be blurry or too thin to recognize. We first preprocess compound images by resizing, adjusting, and cropping them to make the gaps in the images clearer and broader. We then apply Connected Component Analysis (CCA) to segment compound images into individual panels. This approach eliminates small objects and keeps only the main components as individual panels. We assess separation quality of the extracted panels, and modify them if the image segmentation quality appears to be low.

We note that CCA may not correctly segment panels whose contents are not well-connected, highly blurred images, and stitched compound images. We thus introduce specific methods for handling blurry and fragmented images as well as stitched images. We assess the segmentation quality of the panels obtained, and modify the segmentation if needed. The complete framework is shown in Fig. 1. The rest of this section introduces these methods.

**Image Preprocessing.** Gaps in compound images typically separate panels into distinct individual components. However, some panels may be positioned too close to one another, or a thin gap may be noisy or blurred, making separation hard. To address this issue we apply the bicubic interpolation [12] to the image  $I$ , of size  $m \times n$ ; this scales up the image ( $2m \times 2n$ ) and enhances contrast between image regions and gaps. The gaps in the scaled image,  $I_{resized}$ , thus become broader and clearer.

Notably, the separating gaps are not always white or black, that is, the intensity of pixels in gaps can be non-binary. To improve gap clarity and detectability we adjust the intensity of images by mapping pixel intensities whose values are in the interval  $[T_{low}, T_{high}]$  to the entire intensity interval  $[0, 1]$  using *linear mapping*. This mapping enhances contrast within the image so that gaps, which are the lightest or the darkest bands in compound images, become clearer. In the experiments described here, we set  $T_{low}$  to 0.05 and  $T_{high}$  to 0.95.

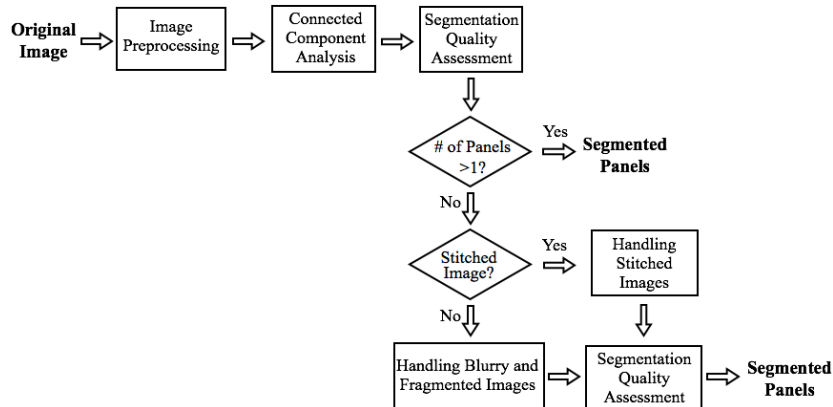


Figure 1: Our framework for compound image segmentation.

We also note that it is hard to distinguish between the external boundary of the image as a whole and the boundaries of individual panels. To disambiguate image-boundaries, we crop the image borders by removing rows and columns of pixels whose maximum gradient value is 0. We denote the image obtained by applying all these preprocessing steps by  $I_{processed}$ .

**Connected Component Analysis (CCA).** To segment a preprocessed image, we first detect connected components within it. We assume that gaps among image-panels are white (which can be reversed later by inverting pixel values). To identify gaps among panels, a binary mask  $M$  is generated as:

$$M_{(x,y)} = \begin{cases} 1 & \text{if } I_{processed}(x,y) \leq t ; \\ 0 & \text{if } I_{processed}(x,y) > t , \end{cases} \quad (1)$$

where  $I_{processed}(x,y)$  denotes the pixel at row  $x$  and column  $y$  in the preprocessed image  $I_{processed}$ . By setting the threshold  $t$ , each pixel  $I_{processed}(x,y)$  in the preprocessed image is labeled as *background*  $M(x,y) = 0$  if  $I_{processed}(x,y) > t$  and as *foreground*  $M(x,y) = 1$  otherwise. In our experiments the threshold  $t$  is set to 0.95. Based on the mask  $M$  we detect connected components by applying the *Connected Component Labeling* method [13]. This method works by scanning the mask  $M$  and assigning labels to pixels. Adjacent pixels sharing the same pixel intensity are assigned the same label. A connected component is a set of pixels that have the same label value. In this paper we set the connectivity to 4, which means we count pixels above and below the central pixel, as well as those to the left and right of the central pixel as the adjacent pixels.

Using CCA may give rise to many small connected components due to small and unconnected objects in the image, such as text. A panel bounding box is set around the smallest rectangle that contains all pixels in each connected component. To initially eliminate connected components covered by bounding boxes of very small box-height or box-width, we thus set two thresholds:  $t_{height} = height/20$ ,  $t_{width} = width/20$ , where  $width$  and  $height$  are the total figure width and height. The relatively large bounding boxes, which typically correspond to the main components of the image, are kept and viewed as the main segmented panels within the compound image.

Fig. 2 illustrates the way our CCA method proceeds. Fig. 2 (a) is a preprocessed image  $I_{processed}$ . Fig. 2 (b) is the binary mask generated according to Eq. 1. By using the Connected Component Labeling method, we obtain connected components, indicated as bounding boxes and shown as textured rectangles in Fig. 2 (c). We then extract only the main components that are covered by large bounding boxes as the output of the CCA method, as shown in Fig. 2 (d).

**Segmentation quality assessment.** After the segmentation method is applied, some pieces of the original image may not be covered by the segmented panels, or the original image may be over-segmented. A quality assessment step is thus added here to assess and adapt segmentation results in order to address

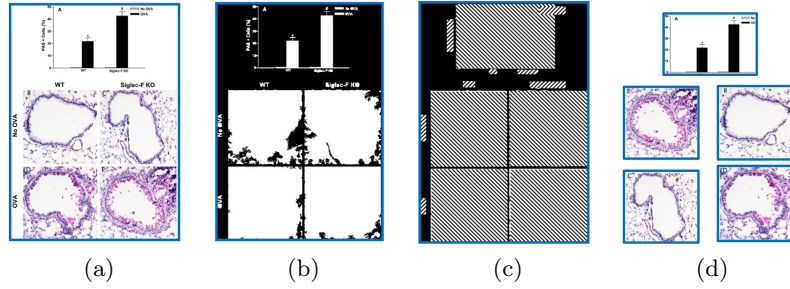


Figure 2: Steps in Connected Components Analysis. The original image is Figure 2, in Publication PMID: 21040544. (a) The preprocessed image. (b) The binary mask generated according to Eq. 1. (c) The Connected Component Labeling result. (d) The segmented image resulting from CCA.

these shortcomings. We assess segmentation quality by employing the five steps described below. Steps 1-3 are used to evaluate and modify individual panels obtained by the segmentation methods discussed here, while steps 4 and 5 are used to assess and adapt the overall segmentation result.

*1. Merge overlapping panels:* Components within a panel may be erroneously detected by CCA as individual panels. As the largest connected component within a panel is typically indicative of the panel's boundary, the bounding boxes of smaller components within the same panel will typically overlap with the bounding box of the largest component. For example, the bounding box of legends may overlap the bounding box of corresponding line graph. We thus compute the ratio between the intersection area and the area of the minimum intersecting bounding box, and merge two bounding boxes when their overlap ratio exceeds 0.1.

*2. Temporarily eliminate small components:* Similar to the elimination step in CCA, we eliminate bounding boxes that are small (less than 1/5 in height or width) compared to the largest bounding box, thus reducing noise.

*3. Recover missing panels:* Due to blurred or disconnected contents in compound figures, some panels may be omitted in the initial segmentation process. We thus introduce a recovery step, in which missing panels are detected and recovered. We assume that the missing panel is similar in size and symmetric in position to present panels. We thus check for each panel whether there is enough space for another bounding box to its left, right, as well as above or below it. The space available for a bounding box next to a present panel indicates the position of a candidate panel. The candidate panels are expected to have similar content area, calculated as the number of non-white pixels within it, as that of the present panel and the same intensity values of all boundary pixels.

*4. Check segmentation area:* To detect incorrect segmentation, we compute the ratio between the sum of the areas of segmented panels and the area of the original image. If this ratio is below 0.5, we consider the segmentation to

be incorrect. Incorrect segmentations are discarded leaving the image contents unsegmented.

5. *Recover small components:* During the elimination of small bounding boxes, some essential parts, such as the text and legend may also be erroneously eliminated. To re-adopt these small components into the panels, we merge eliminated small bounding boxes into their nearest bounding box. To avoid merging bounding boxes that are not part of the same panel during the recovery process, we employ several rules:

- If merging changes both height and width of a qualified bounding box - do not merge.
- If merging changes more than 20% of the height or the width of a qualified bounding box - do not merge.
- If the change of height or width for a qualified bounding box is more than 20% through the small components recovery step, this qualified bounding box keeps its original size.
- An eliminated small bounding box is merged at most once.

**Handling Blurry and Fragmented images.** Through the steps above, several cases may not correctly be segmented, namely: very blurry images, fragmented images that have components with very low internal connectivity, and stitched images. Notably, stitched compound images are different from the other two kinds, which consist of panels that are separated by gaps. To segment these images, we employ a classifier to distinguish stitched images from the other two kinds of images. We define a *gap* as a row or a column whose minimum gray value is above 0.95. If a gap is found in a compound image, the image is classified as a compound image with gaps; otherwise, it is labeled as stitched.

To handle blurry images and fragmented images, we apply an edge detector, which sharpens blurry components in the  $I_{processed}$ . The corresponding edge image, denoted  $I_{edge}$ , may still have poor connectivity. To enhance connectivity of components in  $I_{edge}$ , we dilate the connected regions within the edge image using the minimum gap-width in the image as the dilation factor. After dilation, the connectivity within the dilated edge image is increased. We then apply the CCA method on the dilated edge image again to obtain the segmentation.

Fig. 3 illustrates the handling of blurry and fragmented images. Fig. 3 (a) is a blurry compound image containing many small pieces. By applying an edge detector, we unblur the blurry components, as shown in Fig. 3 (b). We then find the gaps in the edge image along the horizontal and the vertical directions and use the width of the thinnest gap as the dilation factor. Fig. 3 (c) is the dilated edge image, while Fig. 3 (d) shows the segmentation result obtained by using CCA on the dilated edge image. Three panels are detected and highlighted by bounding boxes in Fig. 3 (d). The augmented method thus correctly handles this blurry and fragmented image and identifies the segments within it.

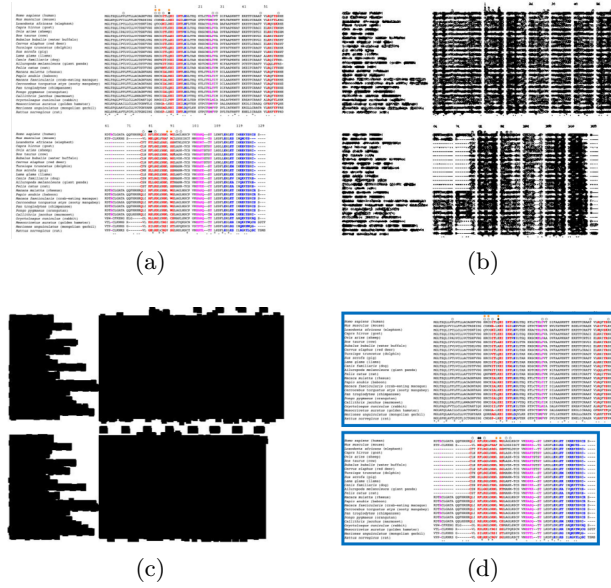


Figure 3: Example in which we handle blurry and fragmented images. (a) The compound figure is Figure 2, in Publication PMID: 20649995. (b) The edge image of the figure shown in (a). (c) The dilated edge image. (d) The result of CCA applied to dilated image.

**Handling Stitched Images.** Stitched compound images do not contain any gap between panels, and as such, cannot be directly segmented by the CCA method. Identifying panel boundaries in such images is thus the main challenge.

Edge detection is applied to identify pixels whose neighbors' intensity sharply changes. Using edge detection, we can create an edge image that clearly shows the boundaries between panels. The edge detector is applied to the preprocessed image to generate a binary edge image in which the boundaries between panels are intensified (see e.g. Fig. 4 (b)). The objective thus becomes that of detecting boundaries in the resulting edge image  $I_{edge}$ . Given an edge image  $I_{edge}$ , if pixel  $(x, y)$  is detected as a pixel along an edge we set  $I_{edge}(x, y) = 1$ . Summing the pixel value along the horizontal and the vertical directions gives rise to two projections:  $Proj_{horizontal}$  and  $Proj_{vertical}$ , which are calculated as:

$$\begin{aligned}
 Proj_{horizontal} &= I_{edge}(x, y), y \in 1 \dots 2n \ ; \\
 Proj_{vertical} &= I_{edge}(x, y), x \in 1 \dots 2m \ .
 \end{aligned}
 \tag{2}$$

The values  $2n$  and  $2m$  are the width and height of the edge image, respectively. The panel segmentation takes place along the horizontal or the vertical line that goes through the highest projection position. For images with complex layout,

the boundary between panels may not cross the whole image; in such cases we recursively segment the image along one direction at a time, where the projection peak value is at least 0.7 of the height or the width of the region currently considered for segmentation.

Fig. 4 shows an example of the steps applied for handling stitched images. Fig. 4 (a) is the original stitched compound image; Fig. 4 (b) shows the edge image obtained by applying an edge detector. Panel boundaries are observed as straight black lines in the image; Fig. 4 (c) shows the horizontal projection plot  $Proj_{horizontal}$  and vertical projection plot  $Proj_{vertical}$  of Fig. 4 (b). By recursively choosing the peak position along the horizontal projection and vertical projection as panel separators, we segment the image into individual panels shown in Fig. 4 (d).

### 3 Experiments and Results

#### 3.1 Experiments

To evaluate our method we conducted two sets of experiments using datasets from the Figure Separation task in the ImageCLEF Medical tasks. In the first experiment, we assess the separation accuracies obtained by the different steps of our segmentation method. We use the training and test datasets of ImageCLEF'16 [16] to train our system and test its performance.

In the second experiment, we compare the separation accuracy of our comprehensive method against that of state-of-the-art systems using test datasets from ImageCLEF'13, '15 and '16 [14–16]. Additionally, to demonstrate the general applicability of our method, we test our method, trained over ImageCLEF'15 dataset, on the ImageCLEF'13 test dataset. For selecting an edge-detector, we experimented with several methods, and decided to use the SUSAN edge detector [17] as it has demonstrated the best performance in this context.

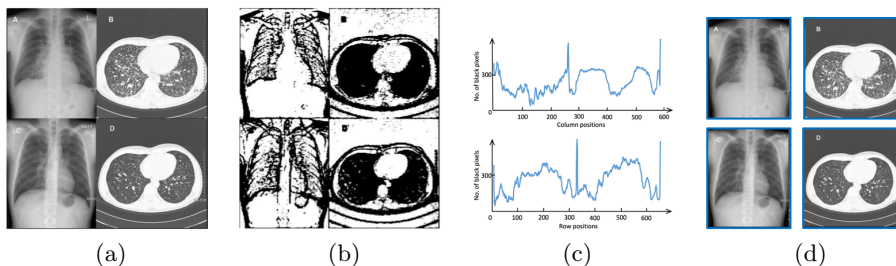


Figure 4: Example of the steps applied for handling stitched images. (a) The stitched compound image taken from Figure 1, Publication PMID: 16480497. (b) The edge image of the figure shown in (a). (c) Horizontal projection (top plot) and vertical projection (bottom plot) calculated according to Eq. 2, based on the edge image. (d) A complete segmentation result by our method.



### 3.2 Datasets and evaluation

We used five imageCLEF datasets in this study, two for training (ImageCLEF'15 and '16) and three for testing (ImageCLEF'13, '15 and '16). The images in the datasets are first extracted from the biomedical publications stored in PubMed Central and then identified as compound images through manual classification.

The ground truth tagging pertaining to the five datasets used in our experiments was provided by ImageCLEF organizers. To evaluate our image separation performance, we use the tool provided by ImageCLEF Medical [14]. This tool computes the accuracy of the separation result for a compound image  $I_i$  as:

$$Accuracy_i = \frac{C}{\max(N_G, N_D)} ,$$

where  $C$  is the number of detected panels that overlap with at least 2/3 of the area of the ground-truth panel,  $N_G$  is the true number of panels in the image, and  $N_D$  is the number of panels we detected. The overall accuracy for the dataset as a whole is then calculated by averaging the accuracies of all separations.

### 3.3 Results

Table 1 shows the separation accuracies obtained in our first set of experiments using different combination of steps within our method over the imageCLEF'16 test dataset. The dataset contains 1615 compound figures, which include 8528 individual panels. The CCA method alone achieves 73.57% accuracy, where 162 images remain unsegmented. Proceeding the CCA method by a preprocessing step leads to an increase of 0.73% in accuracy. Table 1 also shows that 16 additional images are segmented when the preprocessing step is added. By combining the segmentation-quality-assessment step and the CCA method, 40 fewer images are separated compared to CCA-alone, but the separation accuracy increases by 1.7% (compared to the first row in the table). Thus, the segmentation-quality-assessment step improves the correctness of separation result. Combining the image preprocessing step and the segmentation-quality-assessment step with the CCA method (Row 4 in the table), the overall accuracy reaches 74.38%, but 243 images remain unsegmented.

To reduce the number of images that remain unsegmented, we utilize additional steps, as described in Section 2. Applying the step for handling blurry and fragmented images, the accuracy reaches 81.23% and the number of compound images that remain unsegmented decreases to 85. Similarly, applying the step for handling stitched images, the accuracy over the whole dataset reaches 84.03% and the number of compound images that remain unsegmented decreases to 13. Combining all the steps leads to the highest accuracy of 84.03% on the ImageCLEF'16 test dataset while only 9 figures remain unsegmented.

Fig. 5 shows several examples of successful compound figure separation results. Our method not only correctly segments figures containing a single type of image type such as microscopy, graphs, or medical images, but also images containing multiple types of panels.

Table 1: Segmentation accuracies obtained and numbers of images that remain unsegmented by employing different combinations of steps within our method.

Methods used	Separation accuracy	# of unsegmented images
CCA-alone	73.57%	162
Preprocessing+CCA	74.30%	146
CCA+Segmentation quality assessment	75.27%	202
Preprocessing+CCA+Segmentation quality assessment	74.38%	243
Preprocessing+CCA+Segmentation quality assessment+Handling blurry and fragmented images	81.23%	85
Preprocessing+CCA+Segmentation quality assessment+Handling stitched images	84.03%	13
The combination of all methods	<b>84.43%</b>	9

In the second set of experiments, we compare the results obtained by our comprehensive method with those of other systems submitted to ImageCLEF’15 Medical, using the 2015 test dataset. Santosh *et al.*’s method [18] (based on their previous work [2,11]) achieves an accuracy of 84.64%, while Taschwer *et al.* [7] report an accuracy of 84.90%. Our method performs significantly better than all other systems with an accuracy of 90.65%.

To demonstrate the general applicability of our method, we used parameters obtained by training over one of the datasets (ImageCLEF’15) to segment images provided in another dataset (ImageCLEF’13); Our result shows 84.47% accuracy. The other three top performers reported accuracy of 68.59% [21], 69.27% [20], and 84.64% [19]. We note that while the performance of our method is slightly lower than that reported by de Herrera *et al.* [19] using method proposed by Chhatkuli *et al.* [6], the average time required to process one image by our system is 0.74 seconds (Wall-clock), which is much faster than that reported by de Herrera *et al.* [19], i.e. 2.4 seconds.

For ImageCLEF2016 [16], as the only team participating in the Figure Separation task, we achieved 84.43% accuracy on the Figure Separation task test dataset. The segmentation accuracy is similar to the best result obtained in ImageCLEF’15. This result is particularly noteworthy, given that the difficulty of the Figure Separation task was increased in 2016 by adding more stitched compound images and compound images containing multiple types of panels, as indicated in the task description [16].

## 4 Conclusion

We have presented a new scheme for segmenting compound figures, including stitched compound images. We first proposed a preprocessing step to make gaps

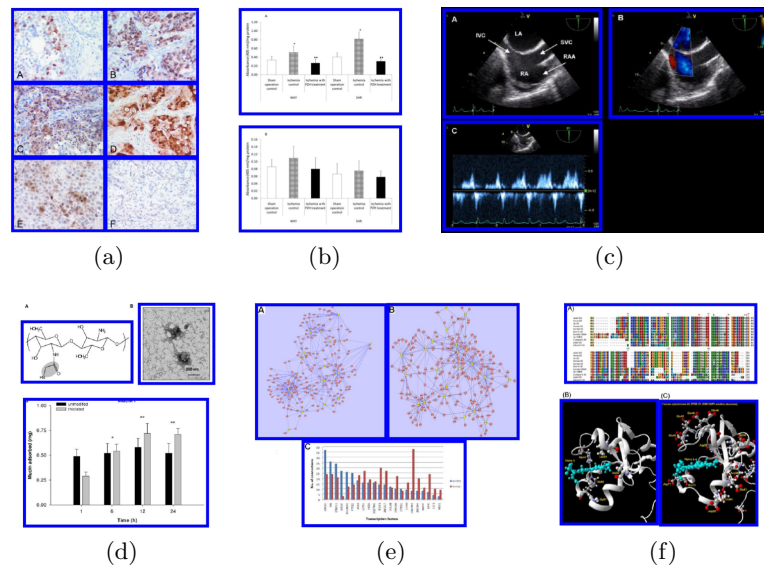


Figure 5: Examples of successful segmentation obtained by our method. The original images of (a)-(f) are taken from: Publication PMID: 18282279, Fig. 6; PMID: 20955558, Fig. 1; PMID: 19439081, Fig. 14; PMID: 16930490, Fig. 1; PMID: 21073692, Fig. 6; PMID: 21129218, Fig. 1, respectively.

clearer so that more images are segmented. We then introduced a method based on Connected Components Analysis to segment images into panels. Segmentation errors were addressed through a step of segmentation quality assessment. Notably, this step evaluates separation quality, back-tracks separation errors, and ensures that only panels that are likely to be correct are extracted from images. Error stemming from over- and under-segmentation in very blurry images, fragmented images, and stitched images are more difficult to address. As such, we proposed two advanced methods to directly handle blurry and fragmented images, and stitched images accordingly. The results demonstrate that our comprehensive method improves upon the panel-segmentation performance of state-of-the-art methods.

While our method achieves a high accuracy in segmenting compound images, there are still challenging cases that are not perfectly addressed. For example, compound images in which both panels and gaps vary in size are hard to segment accurately. We plan to develop methods to identify such cases and address them.

## Acknowledgments.

This work was partially supported by NIH grant R56LM011354A.

## References

1. Shatkay, H., Chen, N., Blostein, D.: Integrating Image Data into Biomedical Text Categorization. *Bioinformatics*, 22(14), e446-e453 (2006)
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 Multipanel Figures. *Journal of the American Society for Information Science and Technology*, 64(5), 893-908 (2013)
3. Murphy, R.F., Velliste, M., Yao, J., Porreca, G.: Searching Online Journals for Fluorescence Microscope Images Depicting Protein Subcellular Location Patterns. *Proc. of the IEEE Int. Symp. on Bioinformatics and Bioengineering*, 119-128 (2001)
4. Antani, S., Demner-Fushman, D., Li, J., Srinivasan, B.V., Thoma, G.R.: Exploring Use of Images in Clinical Articles for Decision Support in Evidence-based Medicine. *Proc. of SPIE Document Recognition and Retrieval XV*, 68150Q (2008)
5. Cheng, B., Antani, S., Stanley, R.J., Thoma, G.R.: Automatic Segmentation of Subfigure Image Panels for Multimodal Biomedical Document Retrieval. *Proc. of SPIE Document Recognition and Retrieval XVIII*, 78740Z (2011)
6. Chhatkuli, A., Foncubierta-Rodriguez, A., Markonis, D., Meriaudeau, F., Müller, H.: Separating Compound Figures in Journal Articles to Allow for Subfigure Classification. *Proc. of SPIE Medical Imaging 2013*, 86740J (2013)
7. Taschwer, M., Marques, O.: Compound Figure Separation Combining Edge & Band Separator Detection. *Proc. of Int. Conf. on Multimedia Modeling*, 162-173 (2016)
8. Kim, D., Ramesh, B.P., Yu, H.: Automatic Figure Classification in Bioscience Literature. *Journal of Biomedical Informatics*, 44(5), 848-858 (2011)
9. Lopez, L.D., Yu, J., Arighi, C., Tudor, C.O., Torii, M., Huang, H., Vijay-Shanker, K., Wu, C.: A Framework for Biomedical Figure Segmentation Towards Image-based Document Retrieval. *BMC systems biology*, 7(4), 1 (2013)
10. Wang, X., Jiang, X., Kolagunda, A., Shatkay, H., Kambhamettu, C.: CIS UDEL Working Notes on Image-CLEF 2015. *CLEF Working Notes* (2015)
11. Santosh, K.C., Antani, S., Thoma, G.: Stitched Multipanel Biomedical Figure Separation. *Proc. of IEEE Int. Symp. on Computer-Based Medical Systems*, 54-59 (2015)
12. Keys, R.: Cubic Convolution Interpolation for Digital Image Processing. *IEEE Trans. on Acoustics, Speech, and Signal Processing*, 29(6), 1153-1160 (1981)
13. Gonzalez, R.C., Woods, R.E.: *Digital Image Processing*. Prentice Hall. (2002)
14. de Herrera, A.G.S., Kalpathy-Cramer, J., Demner-Fushman, D., Antani, S., Müller, H.: Overview of the ImageCLEF 2013 Medical Tasks. *CLEF Working Notes* (2013)
15. de Herrera, A.G.S., Mller, H., Bromuri, S.: Overview of the ImageCLEF 2015 Medical Classification Task. *CLEF Working Notes* (2015)
16. de Herrera, A.G.S., Schaer, R., Bromuri, S., Müller, H.: Overview of the ImageCLEF 2016 Medical Tasks. *CLEF Working Notes* (2016)
17. Smith, S.M., Brady, J.M.: SUSAN-A New Approach to Low Level Image Processing. *Int. Journal of Computer Vision*, 23(1), 45-78 (1997)
18. Santosh, K.C., Xue, Z., Antani, S., Thoma, G.: NLM at ImageCLEF 2015: Biomedical Multipanel Figure Separation. *CLEF Working Notes* (2015)
19. de Herrera, S., Garcia, A., Markonis, D., Schaer, R., Eggel, I., Müller, H.: The medGIFT group in ImageCLEFmed 2013. *CLEF Working Notes* (2013)
20. Simpson, M.S., You, D., Rahman, M.M., Demner-Fushman, D., Antani, S., Thoma, G.R.: ITI's Participation in the 2013 Medical Track of ImageCLEF. *CLEF Working Notes* (2013)
21. Kitanovski, I., Dimitrovski, I., Loskovska, S.: FCSE at Medical Tasks of ImageCLEF 2013. *CLEF Working Notes* (2013)