A Survey of Multi-panel Image Segmentation Methods

Farrukh Jabeen, Mushtaq Ali*, Amanullah Baloch, Faqir Gul
Department of Information Technology, Hazara University, Mansehra, KP, Pakistan.

* Corresponding author. Tel.: 0923027047689; email: mushtaqyaqubi@hotmail.com
Manuscript submitted May 23, 2018; accepted August 13, 2018.
doi: 10.17706/jcp.13.12.1395-1402

Abstract: Multi-panel images play a vital role in describing complicated situations; however, their availability degrades the retrieval accuracy of image retrieval systems. For improving the accuracy of image retrieval systems, the researchers proposed different approaches for the segmentation of multi-panel images into their constituent sub-images. In this paper, five multi-panel image segmentation approaches are reviewed: (1) Manual Segmentation, (2) Inter-panel Boundary Detection, (3) Cluster Based Approach, (4) Sub-panel Label Detection, and (5) Panel Boundary Detection. We discuss the performance metrics used for the evaluation of these proposed approaches. An overview of the datasets used in the literature is also provided. The practical aspects of sub-image separation are finally considered.

Key words: Image segmentation, boundary detection, subpanel label detection, OCR.

1. Introduction

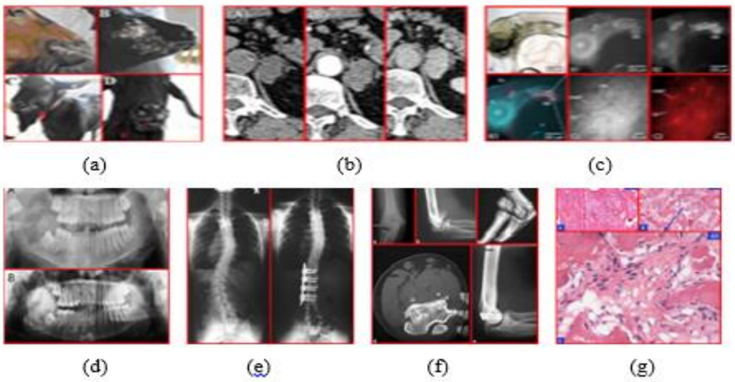

Fig. 1. The multi-panel images (a-g) representing stitched multi-panel figures from image database are shown with and without panel labels as well as irregular panel boxes (taken from [4]).

Presently, people are using multi-panel images for different purposes, for example, for comparing the results of different methods available in the literature [1]. However, their availability creates problems for image retrieval systems. Different researchers proposed various methods [1]-[4] for automatic segmentation of multi-panel images. Manual approach [2] was used to segment multi-panel images. Other proposed approaches compute the locations of the inter-panel borders and then separate the given multi-panel image at one of the computed locations. It is time consuming to separate the images manually due to increasing number of multi-panel images. The results of other approaches are not satisfactory for multi-panel images having dissimilar-sized sub images. In this paper, we review multi-panel image segmentation approaches for evaluation purpose. Five multi-panel image segmentation approaches are reviewed: (1) Manual Segmentation, (2) Inter-panel Boundary Detection, (3) Cluster Based Approach, (4) Sub-panel Label Detection, and (5) Panel Boundary Detection. Effective segmentation of multi-panel images plays a key role in the domain of content based image retrieval, machine vision, medical imaging, object detection, and recognition task. For instance, the results of image retrieval systems are poor for an image database containing multi-panel images like the one shown in Fig. 1.

This is due to the fact that the image retrieval system cannot have access to the sub-images of multi-panel images. Same is the case with other applications' domains. The results of image retrieval systems can be improved by decomposing multi-panel images in an image database into its constituent sub-images like the one shown in Fig. 2.

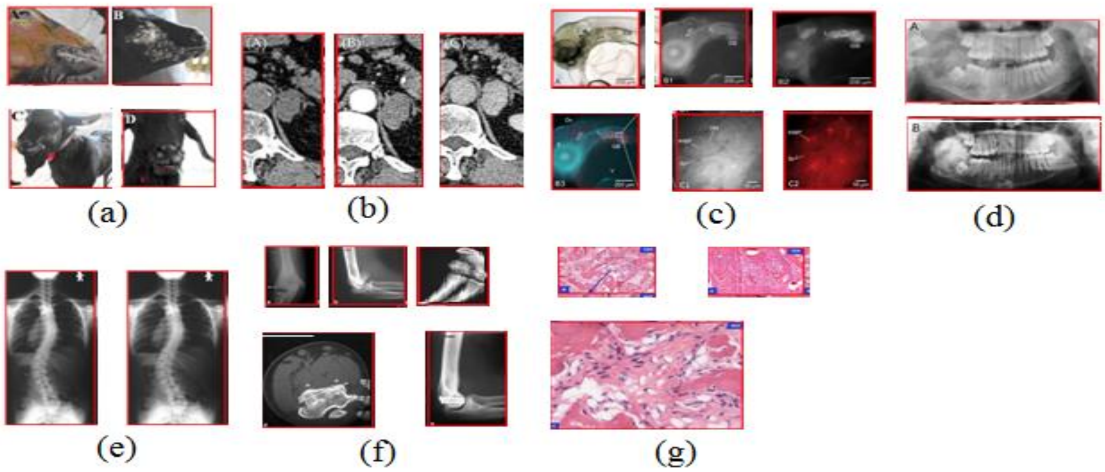

Fig. 2. Subfigures of multi-panel images (a)-(g) shows subfigures of the multi-panel images shown in Fig. 1.

2. Multi-panel Image Segmentation Methods

This section presents the available approaches used for the segmentation of multi-panel images into their constituent sub-images. We found the following five approaches which are explained one by one.

2.1. Manual Approach

In [2], the authors proposed a system for finding most similar images for the given image in support of clinical decision. However, the medical image database contains both single and multi-panel images. As discussed in the introduction section, the multi-panel images degrade the performance of medical image retrieval system because it cannot get access to the sub-images. In order to improve its performance, the authors manually separate the sub-figures of multi-panel images. However, this manual process is time consuming. An automatic method is needed for separating the sub-images of multi-panel images.

2.2. Inter-panel Boundary Detection

In [3], the authors proposed a recursive algorithm for separating the sub-images of compound figure. The proposed algorithm first employs an illustration classifier for finding the class of the given compound figure, i.e. illustration or non-illustration image. If the input compound figure is illustration then the proposed algorithm apply the band- based algorithm for decomposing the given figure into sub-images, otherwise, the edge-based algorithm is used.

The proposed algorithm generates better results for compound images containing white space separator or edges. However, the performance of the proposed algorithm is not satisfactory for multi-panel images containing line separator with color other than white.

Ajad *et. al* [5] proposed automatic figure classification technique for the biomedical images available in the medical journals. A main problem in figure classification is due to the fact that many figures in the biomedical literature are compound figures that mostly consist of more than one type of figures. The proposed algorithm worked through detection and analysis of compound figures for the presence of uniform space gaps between single figures. It works at a large scale on biomedical images present in the biomedical journals. Distinctive feature of proposed technique is that without any prior information about image type, a variety of compound figure types can be detected and analyzed. Furthermore, it effectively utilizes figure's visual information since captions often rely on inaccurate information. Moreover, captions are either not available or incomplete. As a result, sub-images detection techniques cannot rely solely on captions. It learns a set of custom rules from the training image dataset to be used further in detection phase. However, the proposed algorithm fails for images like the one shown in Fig. 3. This is due to the reason that the image in Fig. 3 has no white-space gap between the sub-images.

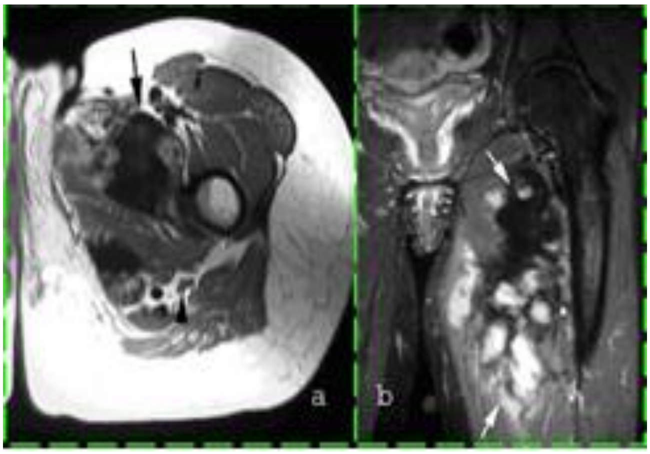

Fig. 3. Sample compound images (taken from ImageCLEF 2015 CFS dataset [6], (a) a vertical edge. Dashed lines show the expected output of CFS.

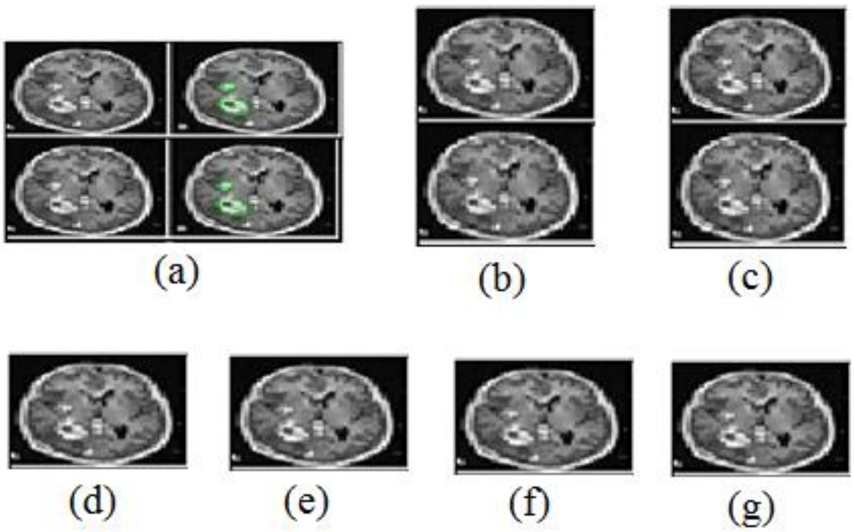

Fig. 4. Instances of multi-panel images are; (a) original multi-panel images (b-c) sub-images separated vertically by using detector separation method, (d-g) sub-images of the original image in (a).

In [1], the proposed framework excludes the boundary of the input multi-panel image and then locates the longest inter-panel border in the resultant image. Next, the original multi-panel image is decomposed into two sub-images at the identified location of the longest inter-panel border. This procedure is demonstrated in Fig. 5. The Fig. 5(b) shows resultant image with no boundary and Fig. 5(c) describes the location of the longest inter-panel border. However, the proposed framework fails to separates the sub-images of multi-panel images containing no clear inter-panel borders.

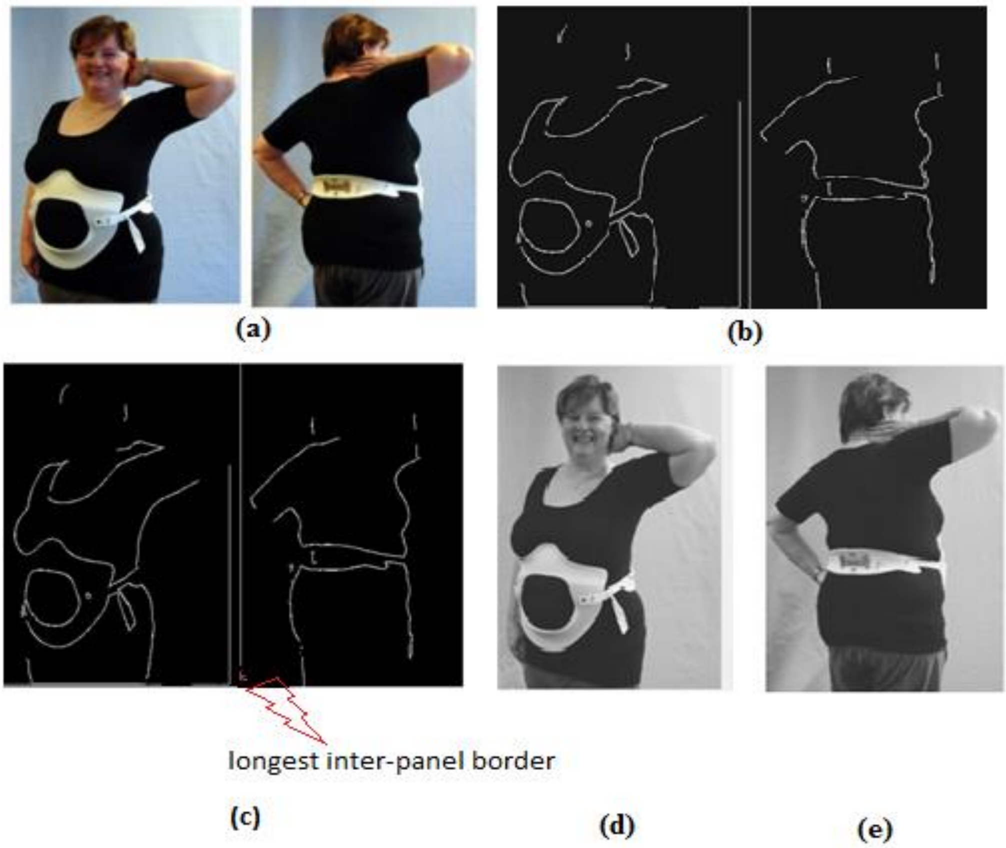

Fig. 5. (a) Original image (b) Image without boundary (c) Longest inter-panel border (d-e) Separated sub-images.

2.3. Cluster Based Separator Line

In [7] the author proposed two techniques (1) Low Variance (2) Particle Swarm Optimization (PSO) clustering. The method in (1) sometimes forms more than one bounding boxes for one sub-image of multi-panel image like the one shown in Fig. 6.

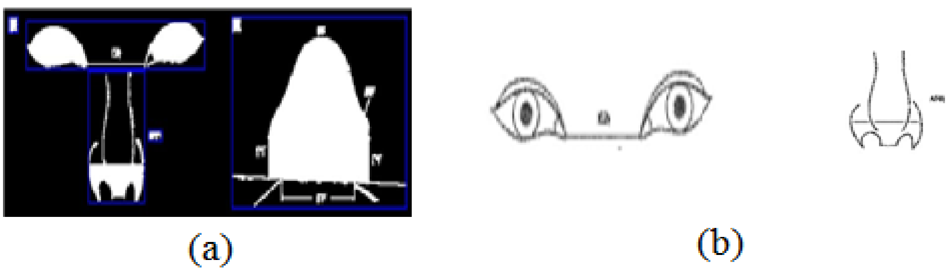

Fig. 6. Incorrect bounding box segmentation. (a) Original image. (b) Bounding box result.

On the other hand, the method (2) may generate poor results for multi-panel images containing sub-images having multi-region with gap between them. Example of such image is show in Fig. 6.

2.4. Subpanel Label Detection

In [8] the authors combined figure captions with pixel-level representation of images. First of all, figure captions were analyzed, to recognize the label style which is used to label panels. After that, pixel-level representation was used to divide a figure into a set of boundary boxes which is converted into text using OCR. Lexical analysis was performed on the text to explore panel labels within the figure that related the result analysis of caption. Finally, optimal panel layout was used to divide the figure. However, the results of the OCR are not satisfactory for images having complicated texture and low contrast labels.

The accurate indexing of images which are available in scientific publication is crucial for efficient and effective retrieval of images. However, indexing of such images is difficult because in scientific literature figure often combine multiple individual sub-images i.e. panels. The proposed technique in [9] automatically finds out panel boundaries, identifies panel labels in the images and changes them into text. Finally, it finds out labels and textual descriptions of each panel from the associated captions. The proposed technique links the output of image content and text processing on caption. Similarly, the technique also decomposes the multi-panel figure into sub-images and corresponding section of the caption is assigned to each subfigure. The results of this technique are poor for multi-panel images containing low contrast panel labels.

2.5. Panel Boundary Detection

The purpose of this work as explained in [4] is to define a new and advanced method to disintegrate subpanels from stitched multi-panel figures. This new technique is usually found in biomedical research articles. The images in such figures have different imaging modalities. For better biomedical Content Based Image Retrieval (CBIR), the most demanding step is to separate them. This technique plays a vital role in the local line segment detection which is usually based on the gray level pixel changes. The second important function of this is to apply a line vectorization process. This process relates the remarkable broken lines with the subfigure boundaries and within the subpanels, unimportant line segments are eliminated. The validity of this approach is checked by applying it on a subset of stitched multi-panel biomedical figures obtained from articles within the open access subset of PubMed central repository, and we got precision and recall of 81.22% and 85.00% respectively.

3. Performance Metrics

In [3], the sensitivity and precision were used for evaluation of separation lines. In [5], manual approach was used for evaluating the results of proposed method. In [7], area under the ROC cure is used for the evaluating the results of proposed method. In [1], the precision, recall and F1 score were used for computing the accuracy of the proposed system. In [4], the precision and recall were used for computing the accuracy of the proposed system.

4. Multi-panel Image Dataset

In [2], the authors used 743 manually annotating images. All these images were taken from the image CLEFmed 2004-2005 dataset. In [8], the authors used number of images extracted from 2,848 articles for testing the accuracy of the proposed approach. Each article in the data set has6 figures and each figure has3 panels on average. In [3], 1237 multi-panel images were selected from Bio medical central journals for testing the performance of the proposed work. The selected multi-panel image data set contains three categories of images; regular images, illustration images and mixed images.In [5], the authors used 515 biomedical multi-panel images. In [7], the authors used ImageCLEF benchmark image dataset for the evaluation of proposed technique. It contains total 300,000 biomedical images taken from scientific literature. This data set is a subset of PubMed Centrally image data set. For evaluation tests, 2982compound images were manually selected. In [1], the authors used 2407, 2340 multi-panel images as a testing and training dataset for checking the accuracy of their system. All these images were taken from the image CLEF Med 2013 dataset. In [4], the authors used 150 biomedical figures extracted from the image CLEF Med 2015 data set for checking the accuracy of their system. The precision and recall were used for computing the accuracy of the proposed system.

5. Subfigures Separation In Practice

This section includes image class identification methods, perfection of automatic methods over manual methods and detail discussion on the available multi-panel image datasets.

5.1. Image Class Identification Methods

The automatic methods for sub-figures separation need image class identification as a first step. For class identification, different methods and approaches are used. In [1], the longest inter-panel border is detected inside the image borders based on dynamic threshold using ratio of length and width of the given image. The image satisfying the threshold is considered as multi-panel.

In [3], [7], the authors proposed approaches for the segmentation of multi-panel images and tested them on a dataset only containing multi-panel images. Hence, no image classification methods were used.

5.2. Perfection of Automatic Methods over Manual Methods

The medical image database contains both single and multi-panel images. The performance of medical image retrieval system is degraded by the availability of multi-panel images as it cannot get access to the sub-images. Manual approach was used for separating the sub-figures of multi-panel images. However, separating multi-panel images manually is time consuming and laborious for dataset containing numerous multi-panel images. In order to overcome these limitations, an automatic method is necessary. The brief description of the available automatic method is given below.

In [3], the authors suggested band based and edge based algorithms for separating the sub-images of compound figure. First, it employs an illustration classifier for finding the class of the given compound figure, i.e. illustration or non-illustration and then applies one of the suggested algorithms based on the class of compound figure.

Two approaches are proposed in [7], low variance and particle swarm optimization (PSO) clustering. The low variance method sometimes generates more than one bounding boxes for one sub-image of multi-panel image. On the other hand, the PSO generates poor result for multi-panel images containing sub-images with gap between them.

In [8], the figure captions are combined with pixel-level representation of images. First of all figure captions were analyzed to identify the label style. Then pixel level representation is used to divide a figure into a set of boundary boxes for detecting the panel labels. Finally, optimal panel layout is used to divide the figure.

The purpose of this work [4] is to define a new method to disintegrate subpanels from stitched multi-panel figures. This approach plays a vital role in the local line segment detection. The second important function is to apply a line vectorization process. It relates the remarkable broken lines with the sub-image boundaries and within the subpanels while removing unimportant line segments.

[1] Mushtaq, A., Dong, L., & Rizwan, A. (2017). Multi-panel medical image segmentation framework for image retrieval system. Springer Science+Business Media, LLC.

[2] Demner-Fushman, D., Sameer, A., George, R., & ThomaLister, H. (2007). Automatically finding images for clinical decision support. *Proceedings of the 7th IEEE International Conference on Data Mining-Workshops,* 0-7695-3019.

[3] Chhatkuli, A., Foncubierta-Rodr´ıguez, A., Markonis, D., Meriaudeau, F., & Mu¨ller, H. (2013). Separating compound figures in journal articles to allow for subfigure classification. *Proceedings of SPIE* (p. 8674).

[4] Santosh, K. C., Sameer, A., & George T. (2015). Stitched multipanel biomedical figure separation K.C. Santosh, Sameer Antani, George Thoma. *Proceeding of IEEE 28th International Symposium on Computer-Based Medical Systems: Vol. 10.* (p. 1109).

[5] Daekeun, Y., Sameer, A., Demner-Fushman, D., Venu, G., & George, R. T. (2011). Detecting figure-panel labels in medical journal articles using MRF. *Proceedings of IEEE International Conference on Document Analysis and Recognition* (pp. 1520-5363).

[6] Garc´ıaSeco de Herrera, A., Mu¨ller, H., & Bromuri, S. (2015). Overview of the image clef2015 medical classification task. *CLEF 2015 Working Notes,* 1391.

[7] Cheng, B., Antani, S., Stanley, R. J., & Thoma, G. R. (2011). Automatic segmentation of subfigure image panels for multimodal biomedical document retrieval. In G. Agam, & C. Viard-Gaudin (Eds.), *Document Recognition and Retrieval XVIII DRR 2011, 18th Document Recognition and Retrieval Conference, Part of the IS&T-SPIE Electronic Imaging Symposium* (pp. 1–10). CA, USA: San Jose.

[8] Luis, D. L., Jingyi Y., Catalina, O. T., Cecilia, N. A., Hongzhan, H. K., Vijay, S., & Cathy, H. W. (2012). Robust segmentation of biomedical figures for image-based document retrieval. *Proceedings of IEEE International Conference on Bioinformatics and Biomedicine.*

[9] Emilia, A., Daekeun, Y., Zhiyun, X., Sameer, A., Dina, D., & George, R. T. (2013). Image *Retrieval From Scientific Publications: Text and Image Content Processing to Separate Multipanel Figures.*

Farrukh Jabeen was born on April 15, 1992 in Haripur. She did her master in computer science, from University of Haripur, KP, Pakistan in 2015. Currently, she is pursuing her MSCS in image processing, from Hazara University, Mansehra, KP, Pakistan. Her area of specialization is image segmentation.

Mushtaq Ali was born on August 1, 1977 in swabi, KPK, Pakistam. He received the M.Sc (computer science) degree from Gomal University, Pakistan and the M.S (computer science) degree from COMSATS University, Pakistan, in 2002 and 2007, respectively. He got the degree of doctor of engineering in computer science and technology from University of Electronic Science and Technology of China (UESTC) in year 2015. Currently, he is working as an assistant professor in Department of Information Technology, Hazara University Mansehra. His research area includes digital image processing, content-based image retrieval system, and computer vision.

Amanullah Baloch was born in DIKhan, on 30th March, 1966. He did his master in computer science, from Gomal University, DIKhan, KP, Pakistan in 1991. He did his MSCS in image processing, from Hazara University, Mansehra, KP, Pakistan in 2018. His area of specialization is image segmentation, feature extraction and image classification. Currently he is pursuing his Ph.D in image processing from Hazara University, Mansehra, KP, Pakistan.

He is working as assistant professor, in the Department of Information Technology, Hazara University, Mansehra since 2003.

Faqir Gul was born on March 01, 1976 in Charsadda, KPK, Pakistan. He got the M.sc (computer science) degree from Gomal University and MSCS degree from Hazara University, KPK, Pakistan in 2002 and 2014 respectively. His area of specialization is image forgery detection. He is working as lecturer in the Department of IT, Hazara University, Pakistan.