Available online at www.sciencedirect.com

ScienceDirect

Applied Soft Computing XX (2014) 1–9

Applied Soft Computing

Appendix: Rough-Fuzzy Clustering and Multiresolution Image Analysis for Text-Graphics Segmentation

Pradipta Maji and Shaswati Roy

Machine Intelligence Unit, Indian Statistical Institute, 203 B. T. Road, Kolkata, 700108, West Bengal, India

Biomedical Imaging and Bioinformatics Lab, Indian Statistical Institute, 203 B. T. Road, Kolkata, 700108, West Bengal, India

Abstract

This paper presents a segmentation method, integrating judiciously the merits of rough-fuzzy computing and multiresolution image analysis technique, for documents having both text and graphics regions. It assumes that the text and non-text or graphics regions of a given document are considered to have different textural properties. The M-band wavelet packet analysis and rough-fuzzy-possibilistic c-means are used for text-graphics segmentation problem. The M-band wavelet packet is used to extract the scale-space features, which offers a richer range of possibilities of scale-space features for document image and is able to zoom it onto narrow band high frequency components. A scale-space feature vector is thus derived, taken at different scales for each pixel in an image. However, the decomposition scheme employing M-band wavelet packet leads to a large number of redundant features. In this regard, an unsupervised feature selection method is introduced to select a set of relevant and non-redundant features for text-graphics segmentation. Finally, the rough-fuzzy-possibilistic c-means algorithm is used to address the uncertainty problem of document segmentation. The whole approach is invariant under the font size of text, text line orientation and gray value of the text. The performance of the proposed technique, along with a comparison with related approaches, is demonstrated on a set of real life document images.

1. Qualitative Evaluation of Different Text-Graphics Segmentation Methods

This section presents qualitative performance analysis of different text-graphics segmentation methods applied on more number of document images. The proposed text-graphics segmentation algorithm [1] is compared with the method using dyadic wavelet packets and k-means ($\mathcal{M}1$) [4], the technique using M-band wavelets and k-means ($\mathcal{M}2$) [2] and the method using cubic b-spline wavelets and k-means ($\mathcal{M}3$) [3]. The results are reported in Figs. 1, 2 and 3. It establishes the effectiveness of the proposed algorithm.

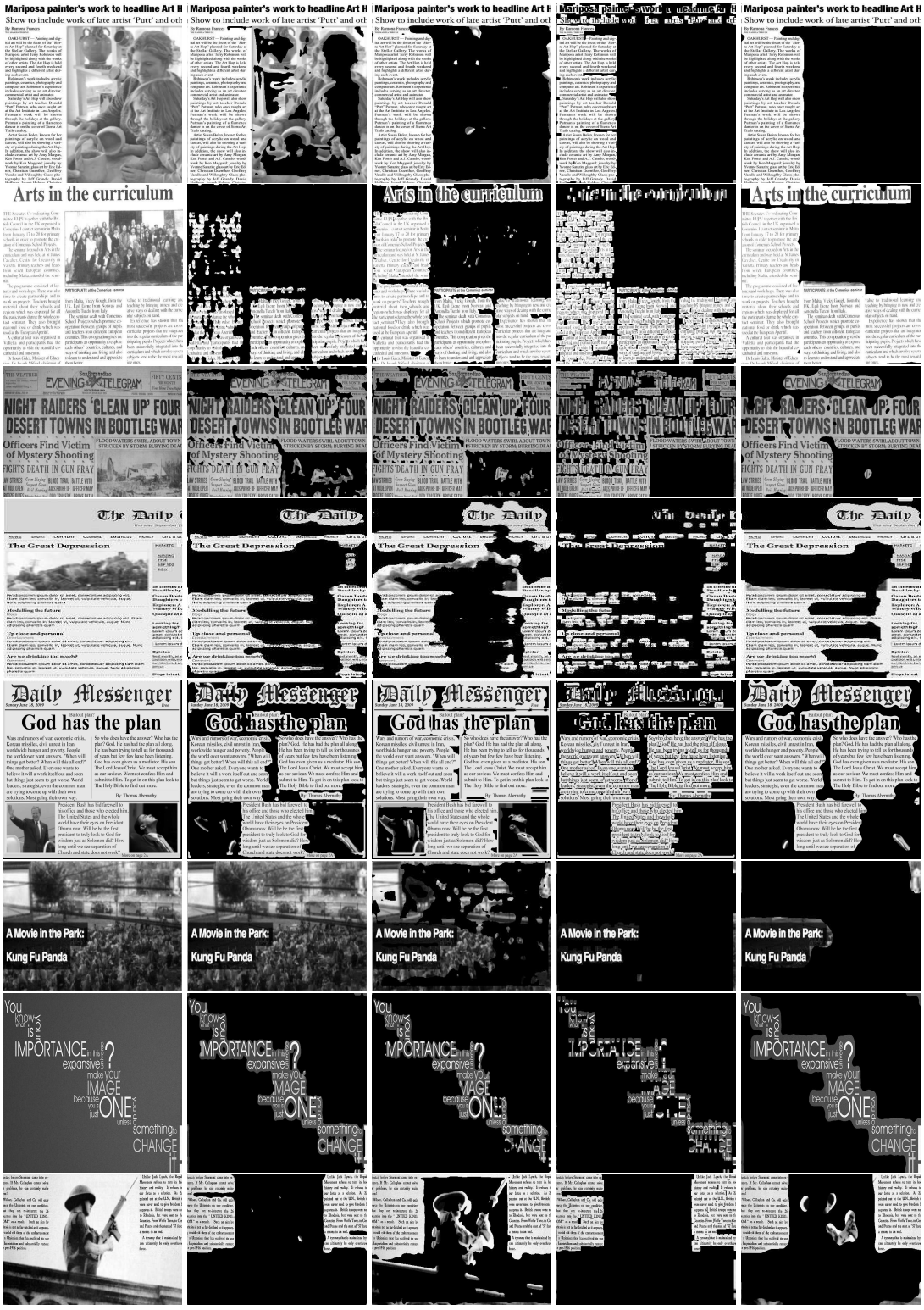
2. Qualitative Analysis for Feature Selection Algorithm

The quantitative performance analysis of feature selection algorithm for different value of feature selection parameter δ with respect to segmentation metrices, namely, Jaccard index, Dice coefficient, sensitivity and specificity is reported in [1]. Fig. 4-13 present the importance of feature selection algorithm qualitatively varying the value of δ. The quality of segmentation improves from left to right in these figures as the increase in the value of δ.

3. Determination for Threshold of Sobel's Operator

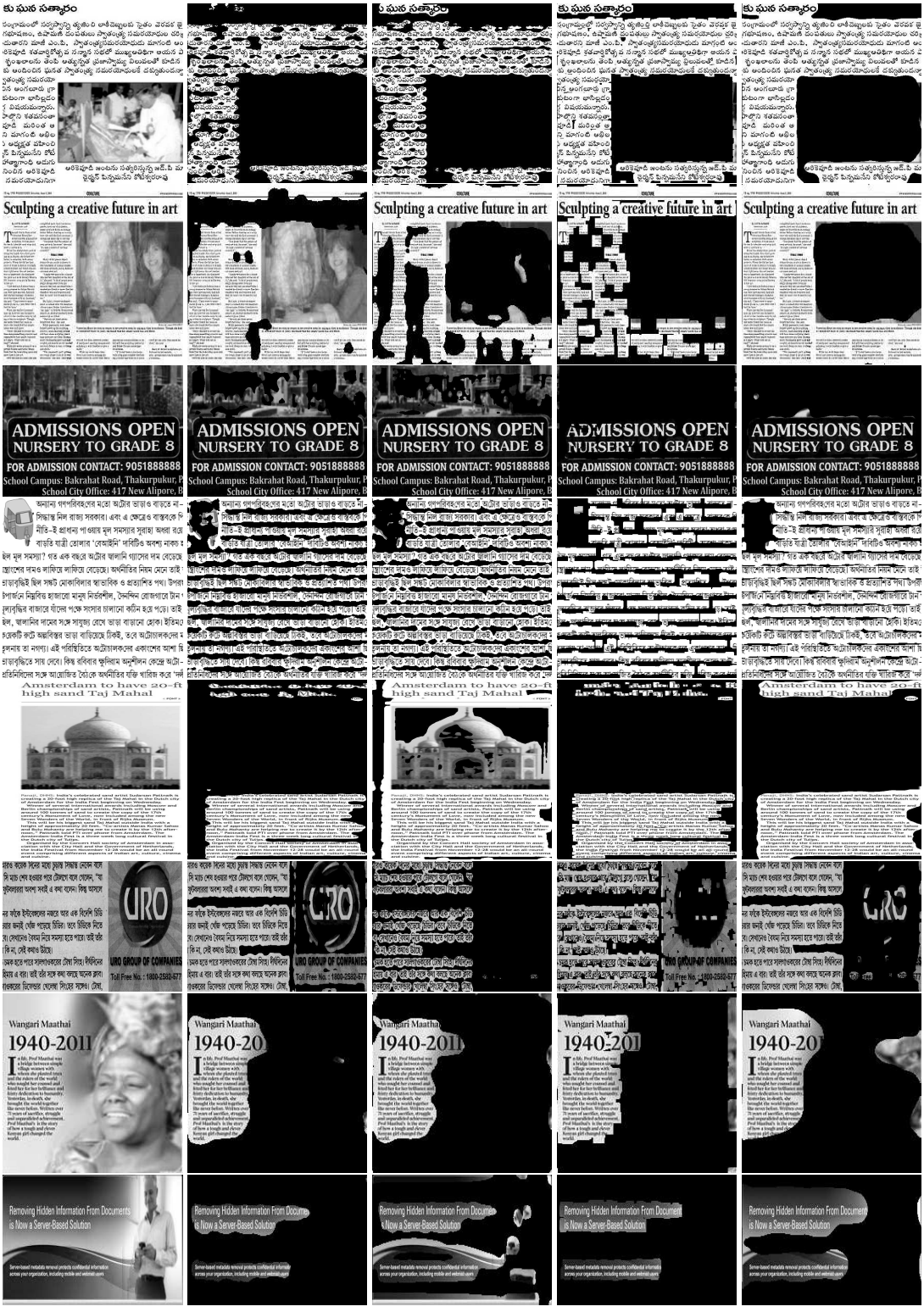
In [1], feature vectors are computed by applying nonlinear function to each of the subbands generated by M-band wavelet packet tranform. The window size for nonlinear energy estimation is calculated based on Sobel's edge operator. The threshold value for Sobel's operator is determined by extensive experiment. Fig. 14 reports the average segmentation accuracy of the dataset used with respect to Jaccard index, Dice coeffcient, sensitivity and specificity for different values of threshold ranging from 0.65 to 0.99. It can be seen from this figure that as the value of Sobel parameter increases, starting from 0.65, the average segmentation accuracy also increases, and it attains highest level accuracy when Sobel parameter is firmed at 0.95 for the fifty benchmark document images used in the experiment.

Email address: {pmaji,shaswatiroy_t}@isical.ac.in (Pradipta Maji and Shaswati Roy)



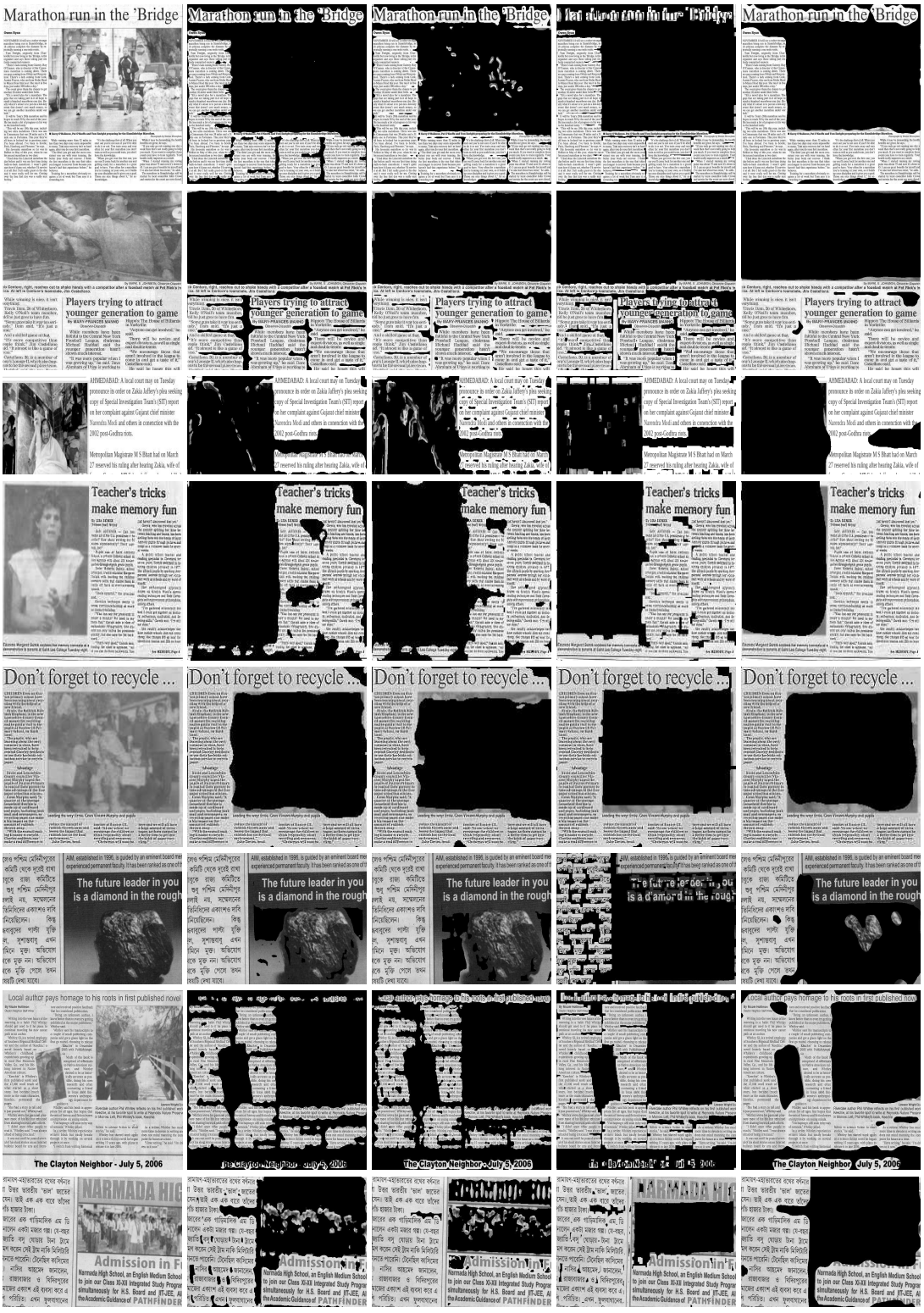
(a) (b) (c) (d) (e)

Figure 1. Text portion obtained using: (a) input image, (b) $\mathcal{M}1$, (c) $\mathcal{M}2$, (d) $\mathcal{M}3$ and (e) proposed



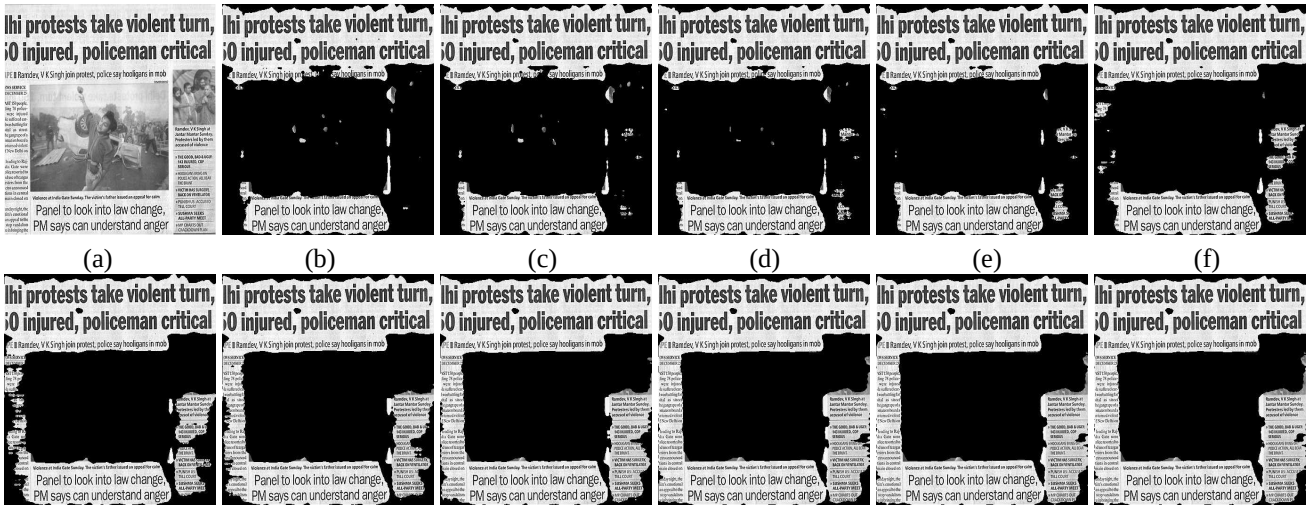
(a) (b) (c) (d) (e)

Figure 2. Text portion obtained using: (a) input image, (b) $\mathcal{M}1$, (c) $\mathcal{M}2$, (d) $\mathcal{M}3$ and (e) proposed



(a) (b) (c) (d) (e)

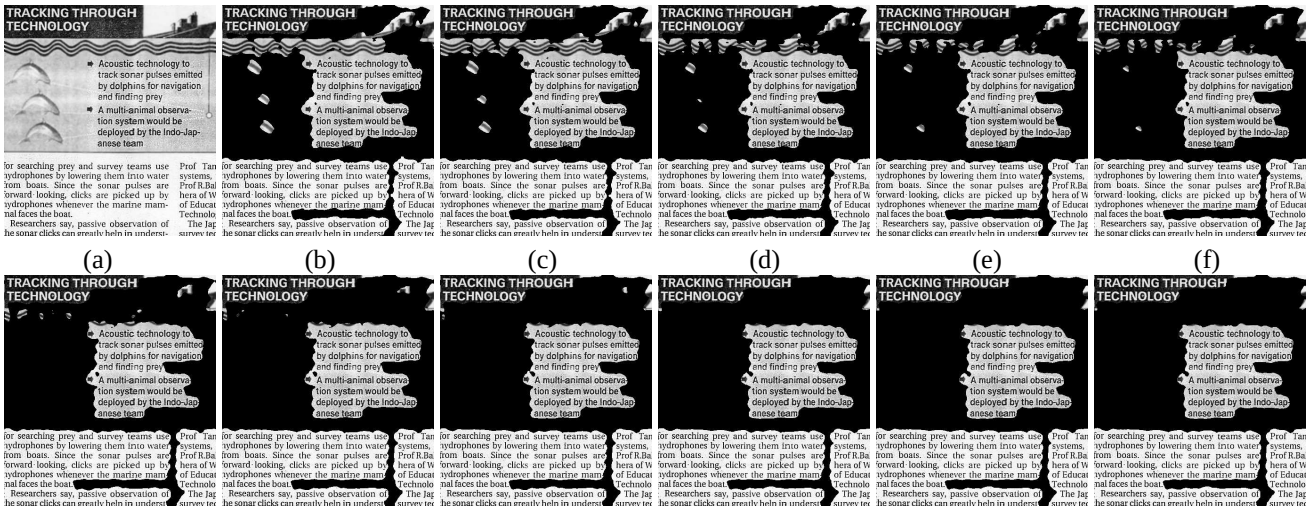
Figure 3. Text portion obtained using: (a) input image, (b) $\mathcal{M}1$, (c) $\mathcal{M}2$, (d) $\mathcal{M}3$ and (e) proposed



(a) (b) (c) (d) (e) (f)

(e) (f) (g) (h) (i) (j)

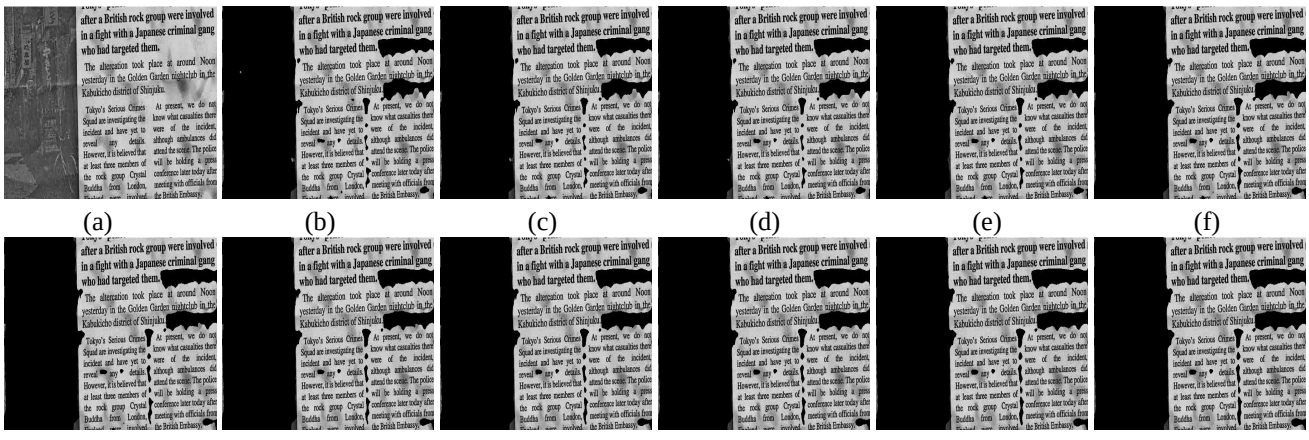
Figure 4. Segmentation results for δ: (a) 0.90 (b) 0.91 (c) 0.92 (d) 0.93 (e) 0.94 (f) 0.95 (g) 0.96 (h) 0.97 (i) 0.98 (j) 0.99



(a) (b) (c) (d) (e) (f)

(e) (f) (g) (h) (i) (j)

Figure 5. Segmentation results for δ: (a) 0.90 (b) 0.91 (c) 0.92 (d) 0.93 (e) 0.94 (f) 0.95 (g) 0.96 (h) 0.97 (i) 0.98 (j) 0.99



(a) (b) (c) (d) (e) (f)

(e) (f) (g) (h) (i) (j)

Figure 6. Segmentation results for δ: (a) 0.90 (b) 0.91 (c) 0.92 (d) 0.93 (e) 0.94 (f) 0.95 (g) 0.96 (h) 0.97 (i) 0.98 (j) 0.99

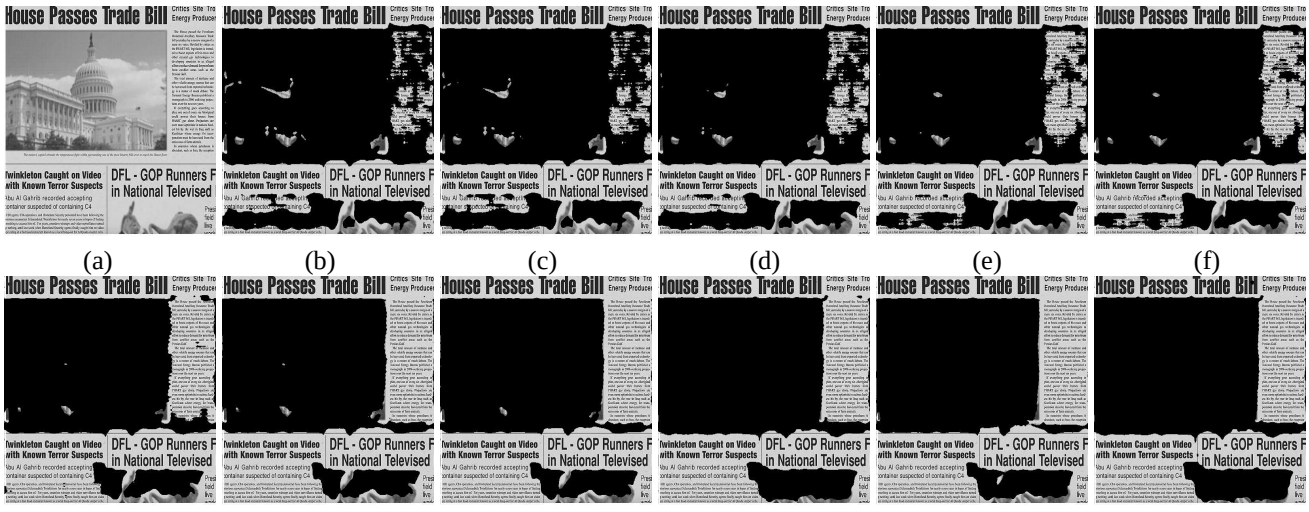


Figure 7. Segmentation results for δ: (a) 0.90 (b) 0.91 (c) 0.92 (d) 0.93 (e) 0.94 (f) 0.95 (g) 0.96 (h) 0.97 (i) 0.98 (j) 0.99

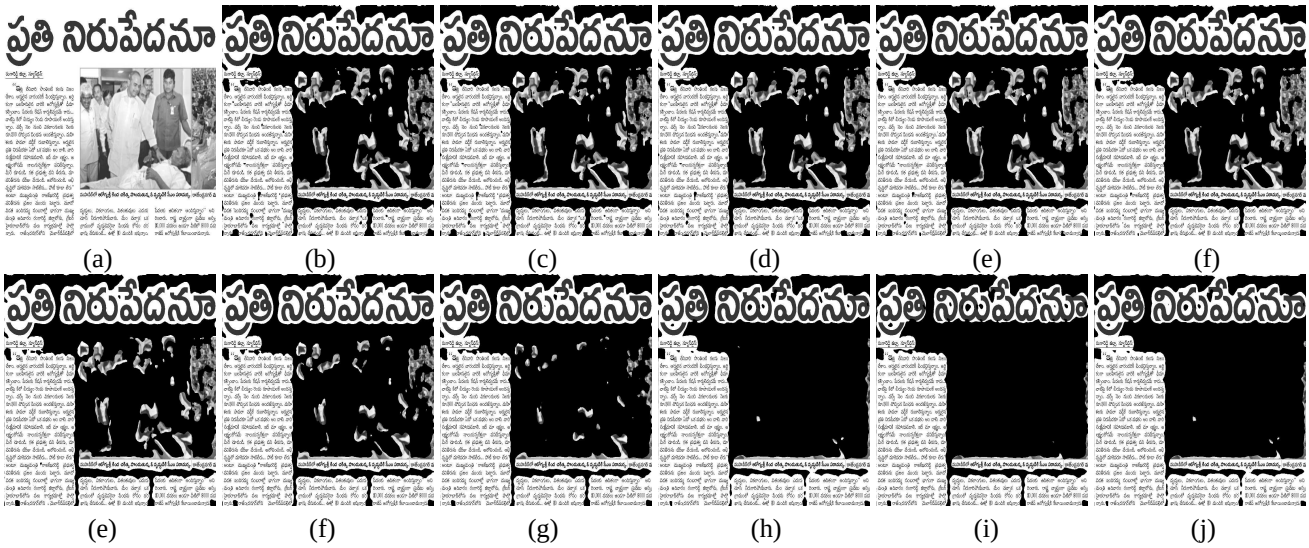


Figure 8. Segmentation results for δ: (a) 0.90 (b) 0.91 (c) 0.92 (d) 0.93 (e) 0.94 (f) 0.95 (g) 0.96 (h) 0.97 (i) 0.98 (j) 0.99

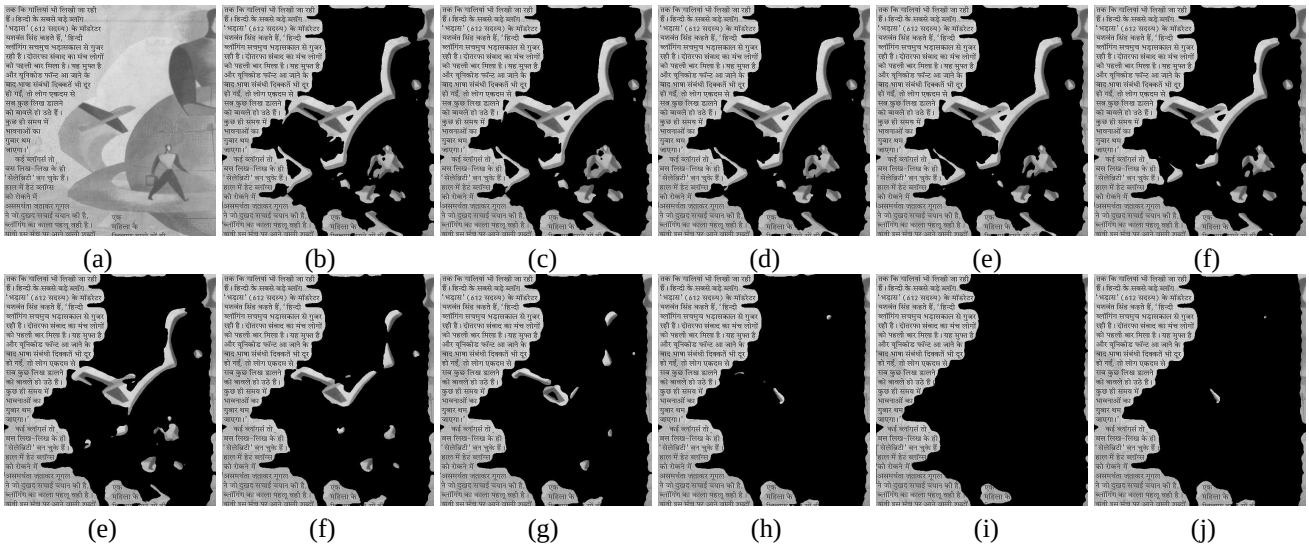


Figure 9. Segmentation results for δ: (a) 0.90 (b) 0.91 (c) 0.92 (d) 0.93 (e) 0.94 (f) 0.95 (g) 0.96 (h) 0.97 (i) 0.98 (j) 0.99

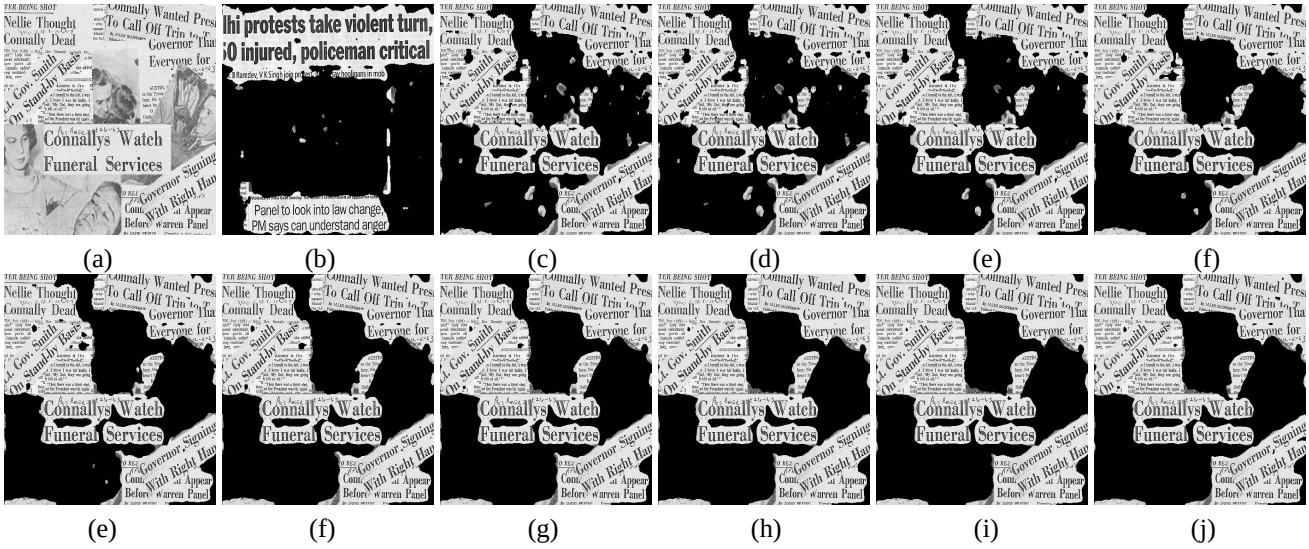


Figure 10. Segmentation results for δ: (a) 0.90 (b) 0.91 (c) 0.92 (d) 0.93 (e) 0.94 (f) 0.95 (g) 0.96 (h) 0.97 (i) 0.98 (j) 0.99

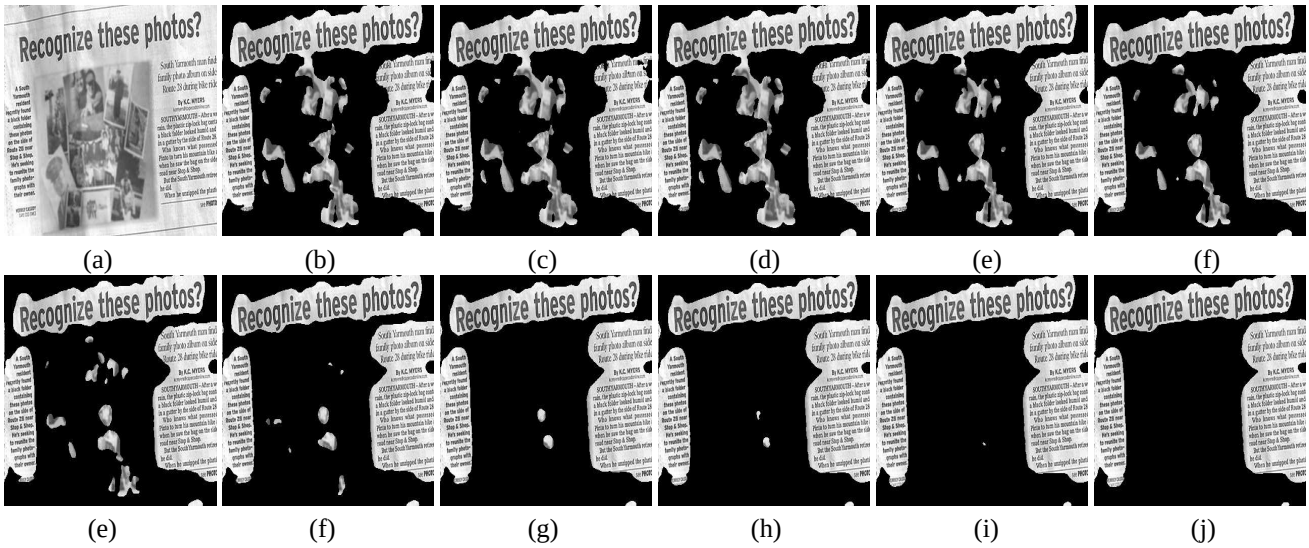


Figure 11. Segmentation results for δ: (a) 0.90 (b) 0.91 (c) 0.92 (d) 0.93 (e) 0.94 (f) 0.95 (g) 0.96 (h) 0.97 (i) 0.98 (j) 0.99

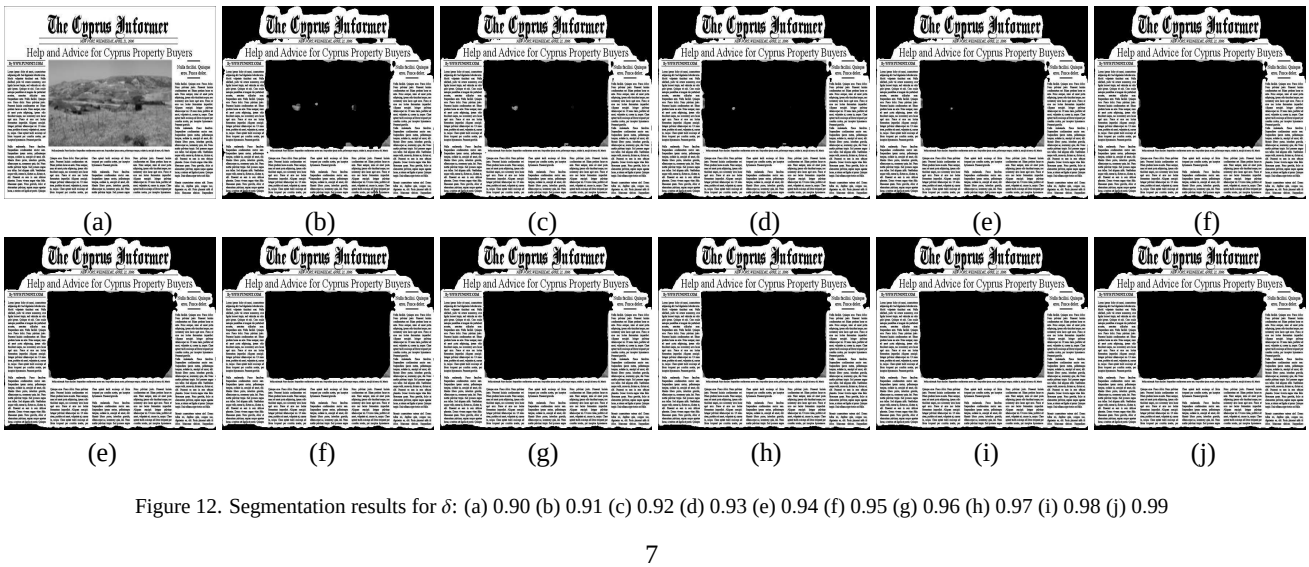


Figure 12. Segmentation results for δ: (a) 0.90 (b) 0.91 (c) 0.92 (d) 0.93 (e) 0.94 (f) 0.95 (g) 0.96 (h) 0.97 (i) 0.98 (j) 0.99

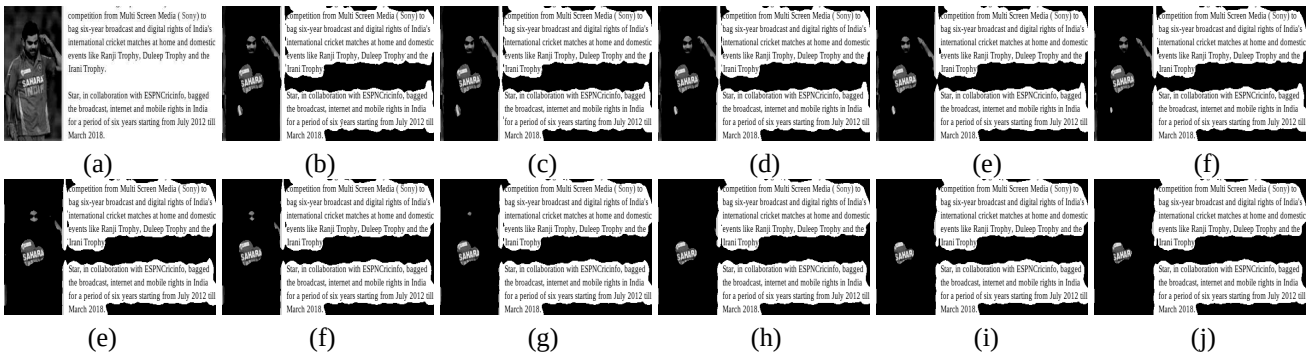


Figure 13. Segmentation results for δ: (a) 0.90 (b) 0.91 (c) 0.92 (d) 0.93 (e) 0.94 (f) 0.95 (g) 0.96 (h) 0.97 (i) 0.98 (j) 0.99

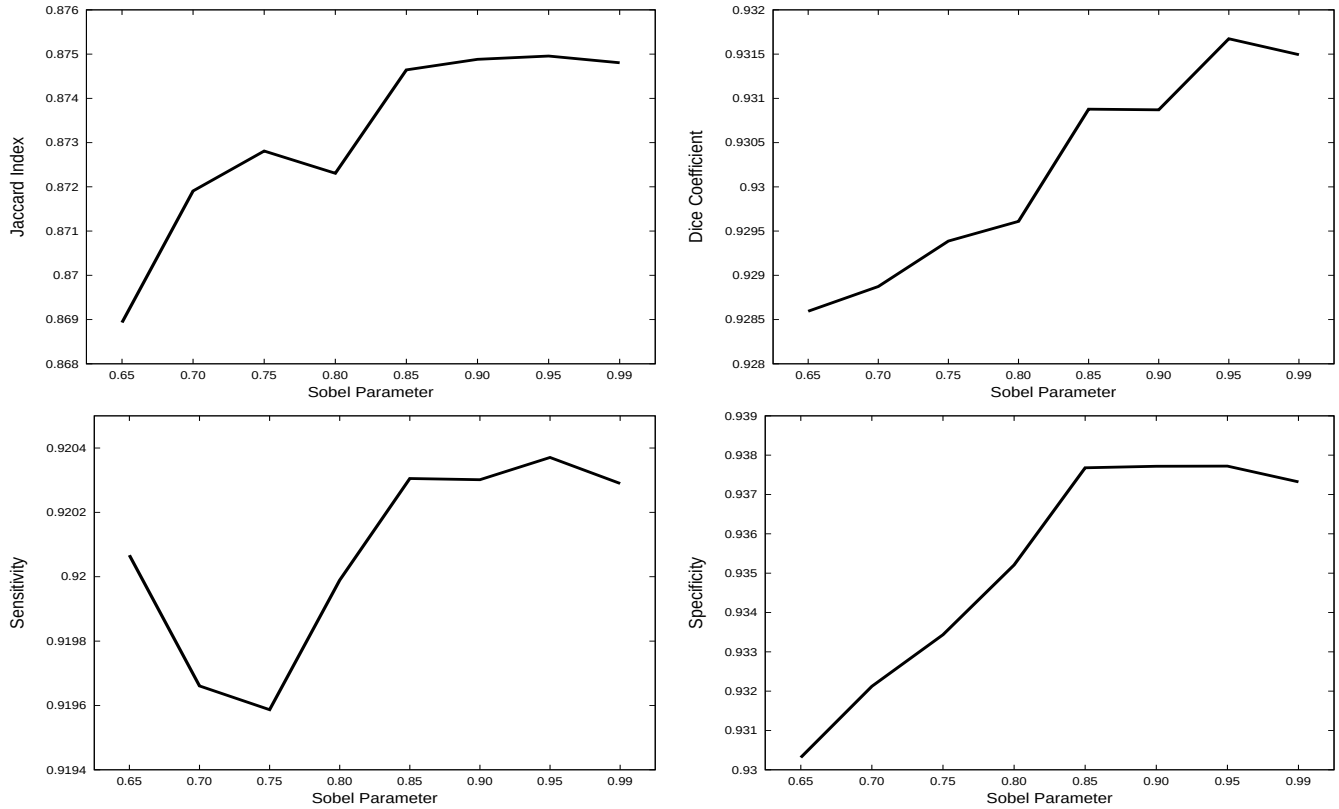


Figure 14. Average segmentation accuracy for different threshold values of Sobel's operator

4. Selection of Clustering Parameters *a* and *w*

Throughout the experiments the parameter values of a and w are made fixed at 0.50 and 0.95 respectively in [1]. Better segmentation results even may be obtained using rough-fuzzy-possibilistic c-means by varying the parameters a and w. In this section, segmentation accuracy is presented by varying the parameters a and w as depicted in Figs 15 to 18 for different images. These figures shows that more desired results may be obtained for other combination of a and w for some document images.

References

[1] P. Maji, S. Roy, Rough-Fuzzy Clustering and Multiresolution Image Analysis for Text-Graphics Segmentation, Applied Soft Computing, Minor Revision.

[2] M. Acharyya, M. K. Kundu, Document Image Segmentation Using Wavelet ScaleSpace Features, IEEE Transactions on Circuits and Systems for Video Technology 12 (12) (2002) 1117–1127.

[3] S. Deng, S. Latifi, E. Regentova, Document segmentation using polynomial spline wavelets, Pattern Recognition 34 (12) (2001) 2533–2545.

[4] G. B. Lee, W. O. Odoyo, J. H. Lee, Y. Chung, B. J. Cho, Two Texture Segmentation of Document Image Using Wavelet Packet Analysis, In Proceedings of the 9th International Conference on Advanced Communication Technology 1 (2007) 395–398.

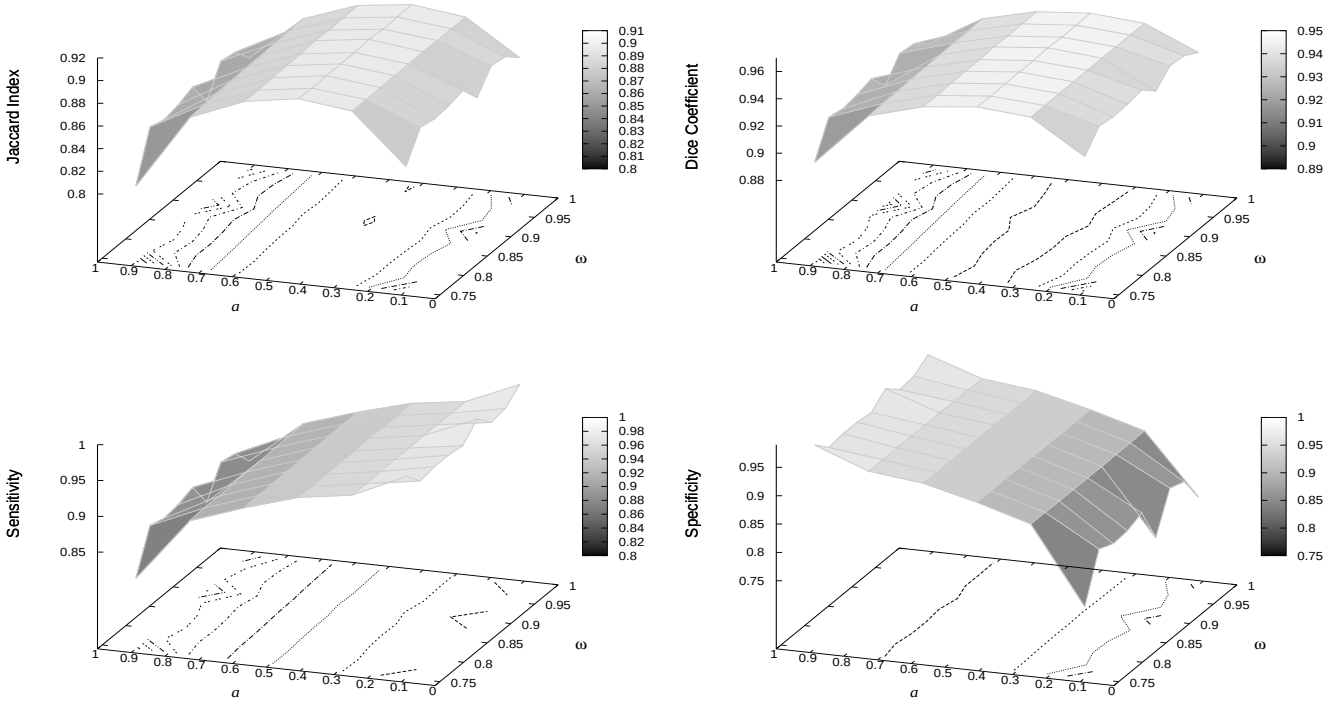


Figure 15. Segmentation accuracy of various combination of *a* and *w*

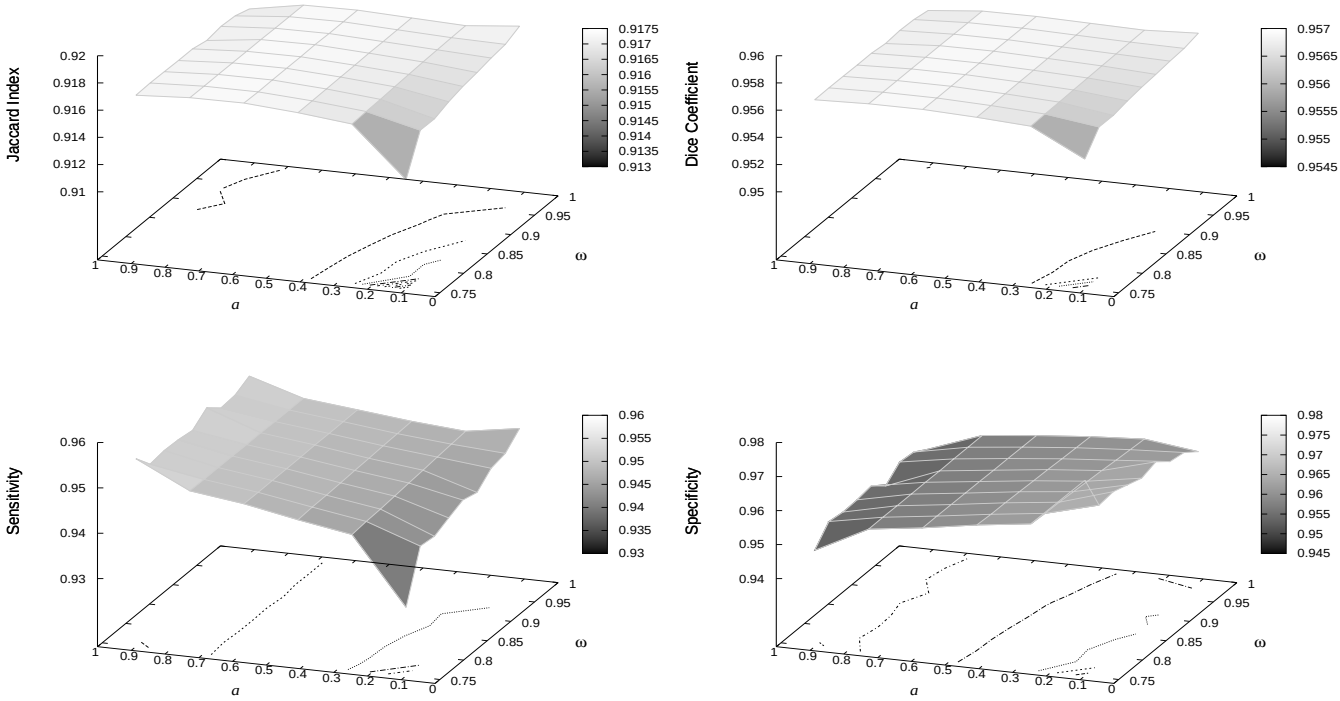


Figure 16. Segmentation accuracy of various combination of *a* and *w*

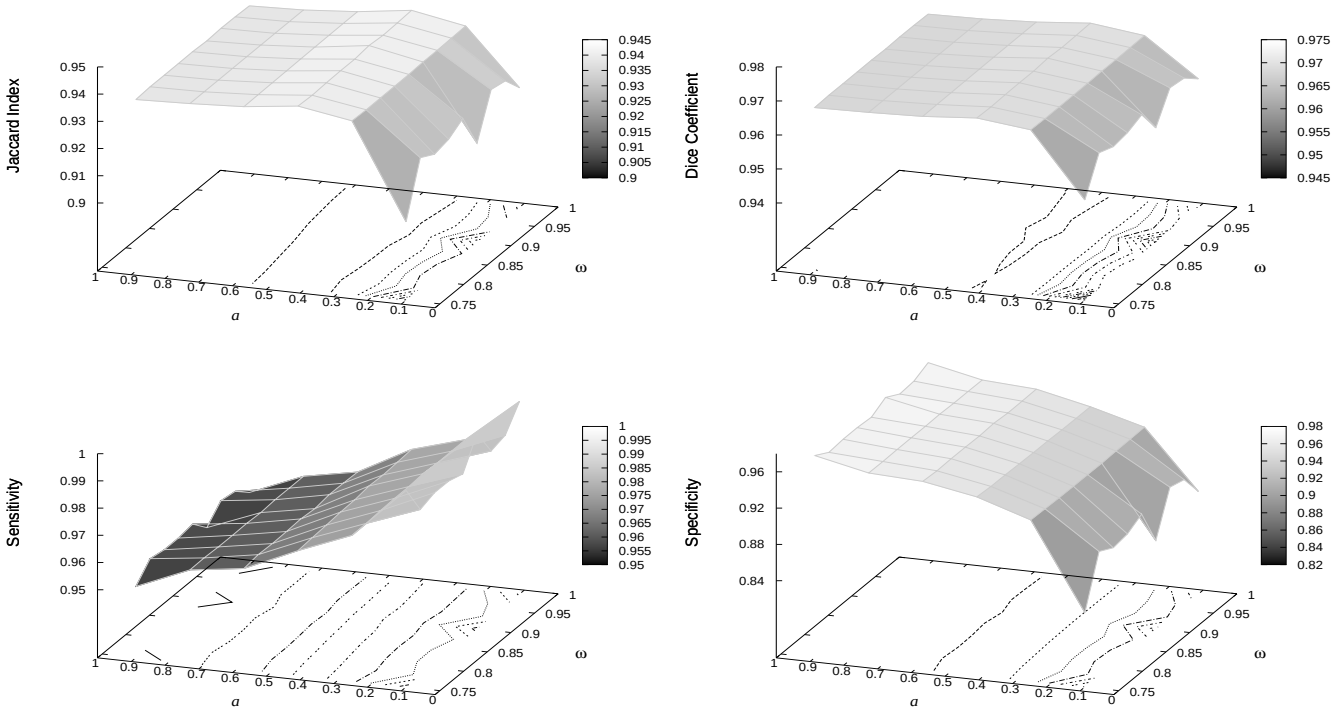


Figure 17. Segmentation accuracy of various combination of *a* and *w*

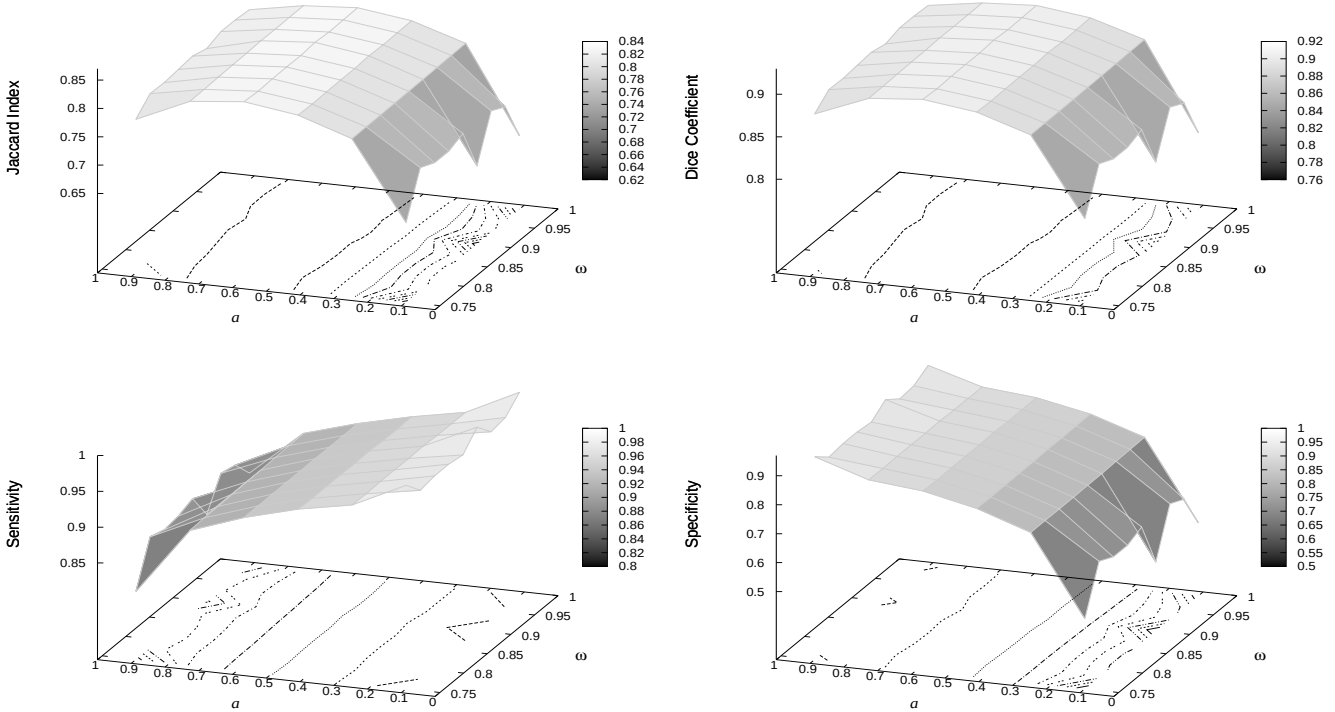


Figure 18. Segmentation accuracy of various combination of *a* and *w*

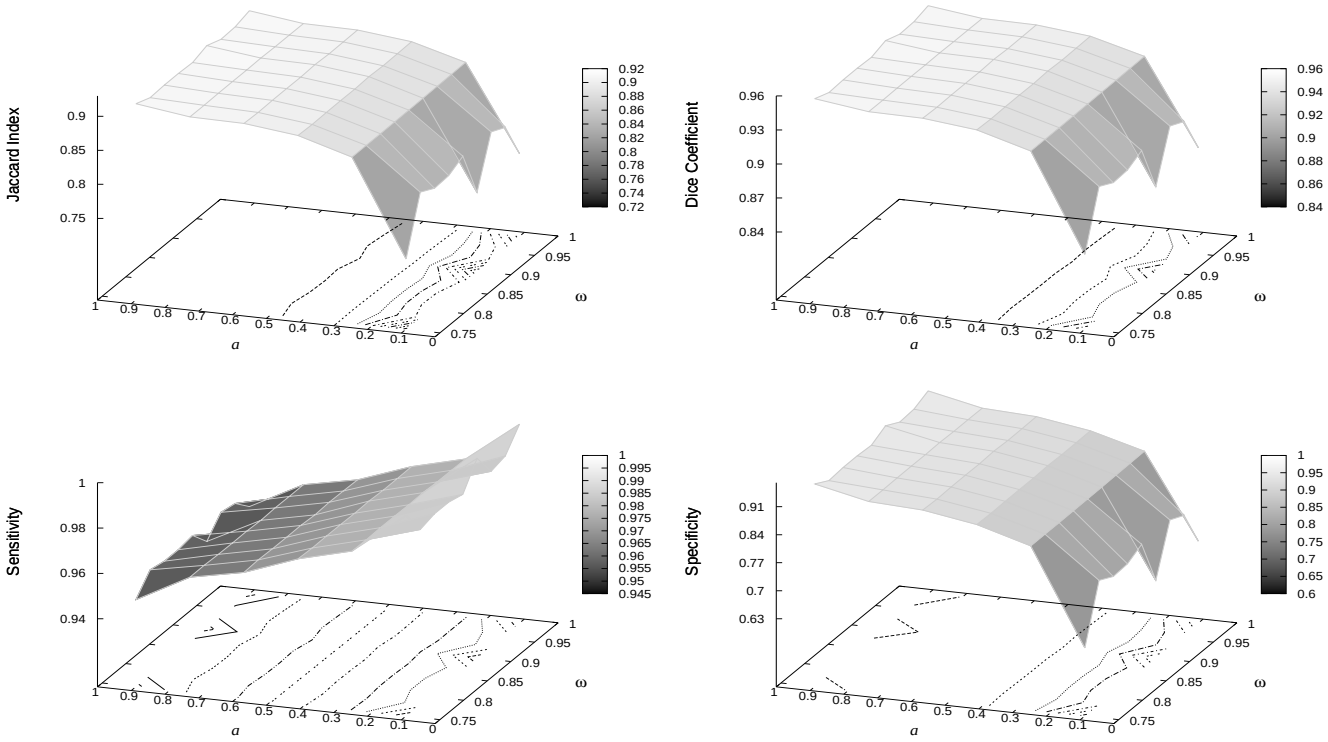


Figure 19. Segmentation accuracy of various combination of *a* and *w*

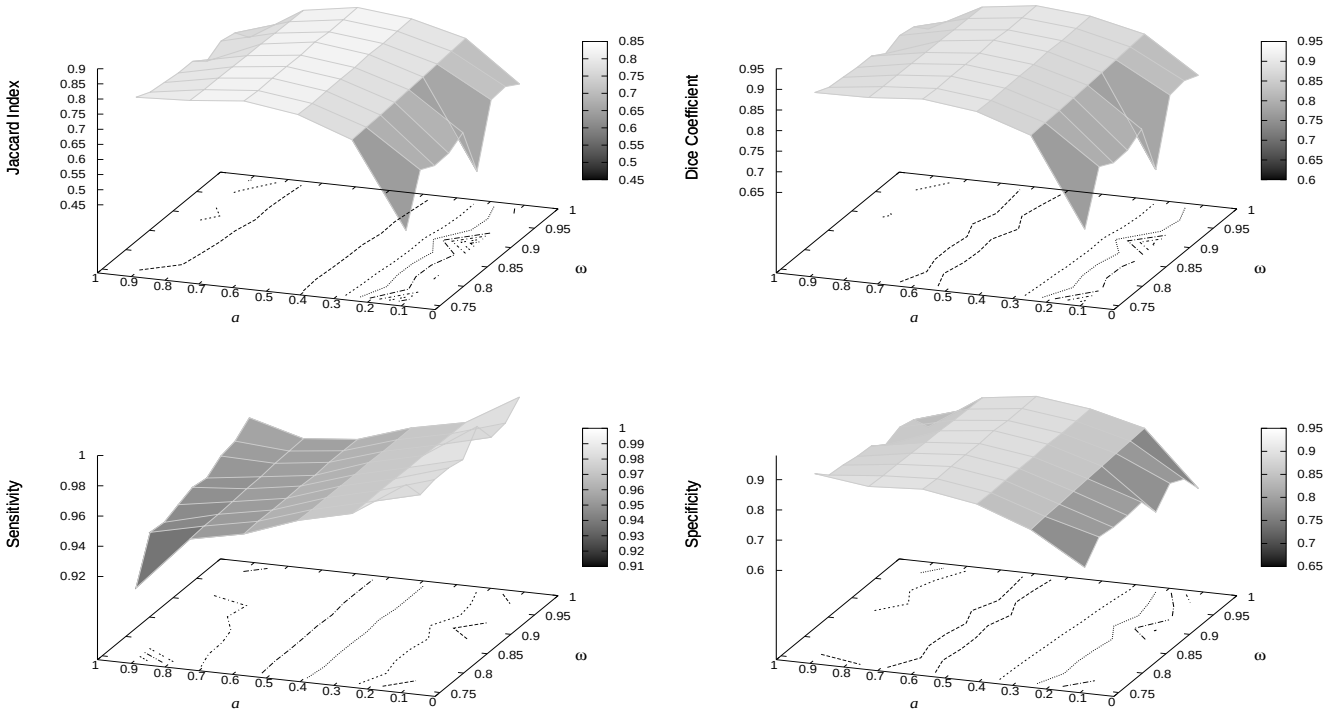


Figure 20. Segmentation accuracy of various combination of *a* and *w*

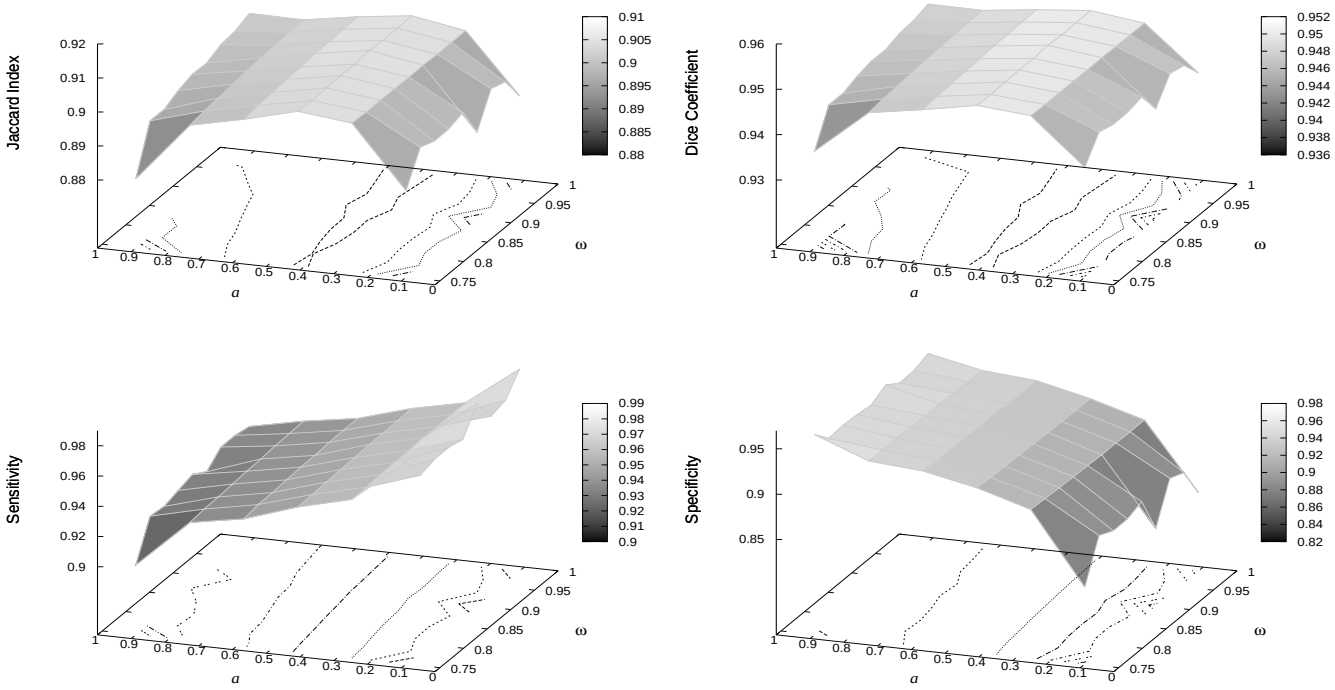


Figure 21. Segmentation accuracy of various combination of *a* and *w*

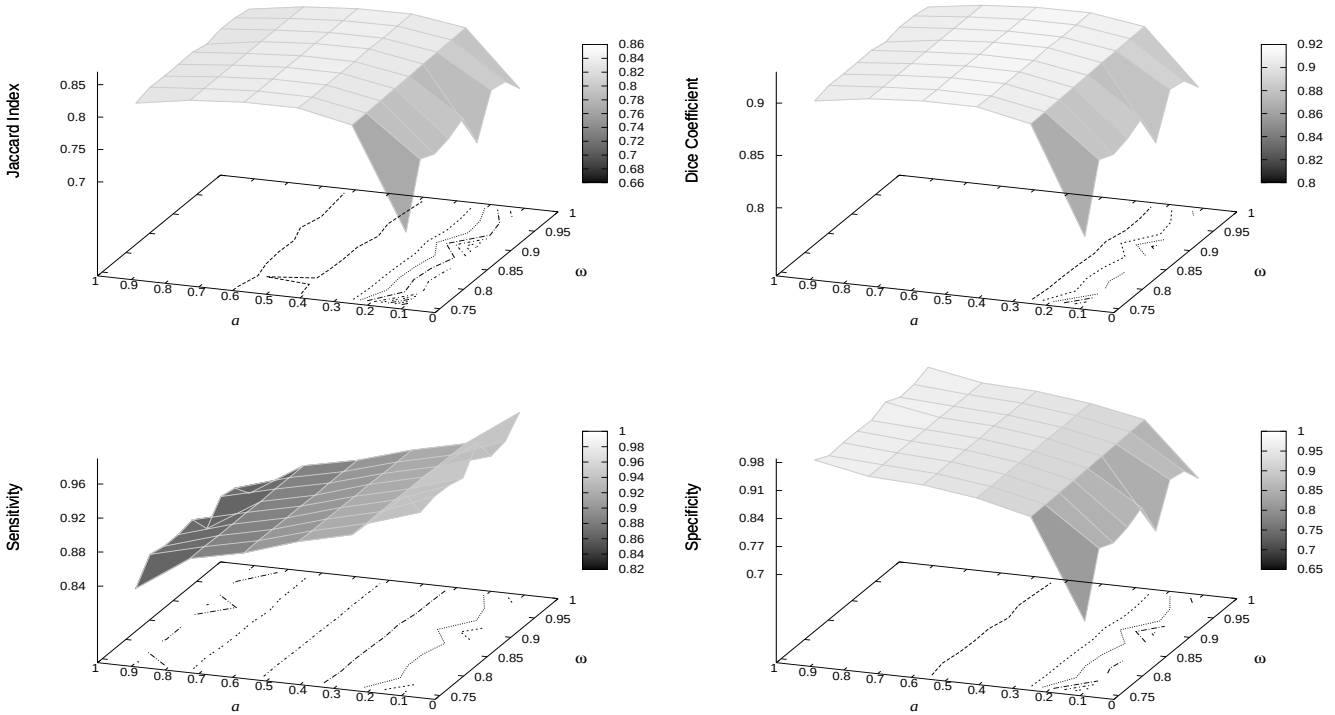


Figure 22. Segmentation accuracy of various combination of *a* and *w*

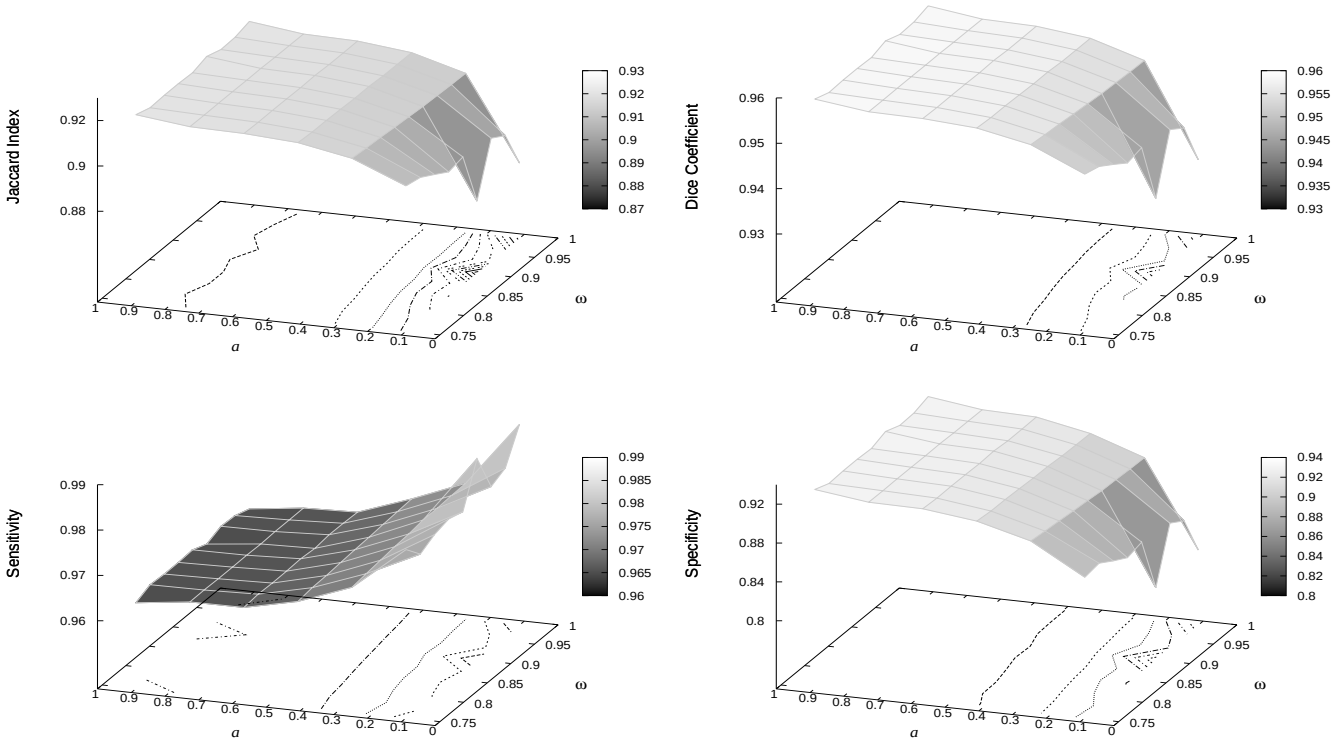


Figure 23. Segmentation accuracy of various combination of *a* and *w*

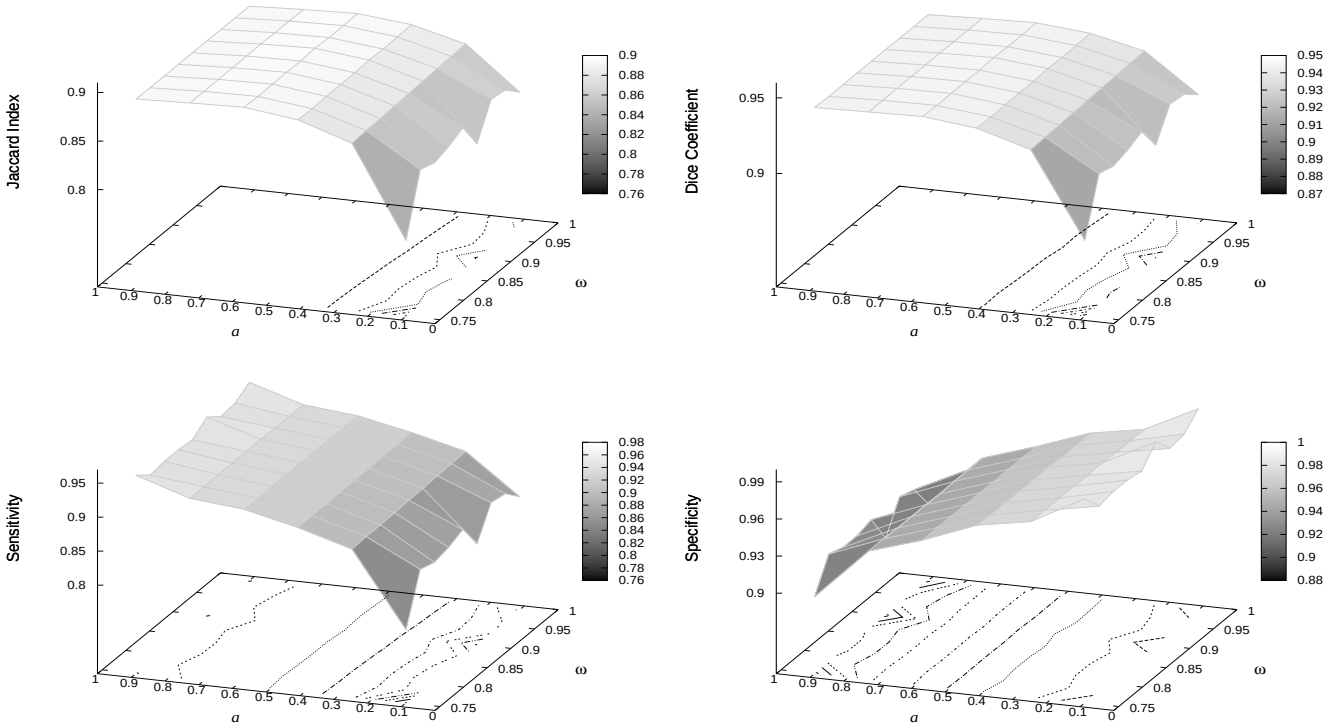


Figure 24. Segmentation accuracy of various combination of *a* and *w*