

# Unsupervised Mitochondria Segmentation using Recursive Spectral Clustering and Adaptive Similarity Models

Julia Dietlmeier<sup>\*,a</sup>, Ovidiu Ghita<sup>a</sup>, Heiko Duesmann<sup>b</sup>, Jochen H. M. Prehn<sup>b</sup>, Paul F. Whelan<sup>a</sup>

<sup>a</sup>*Centre for Image Processing and Analysis, Dublin City University, Dublin, Glasnevin 9, Ireland*

<sup>b</sup>*Royal College of Surgeons in Ireland, Dublin, Ireland*

---

## Abstract

The unsupervised segmentation method proposed in the current study follows the evolutionary ability of human vision to extrapolate significant structures in an image. In this work we adopt the perceptual grouping strategy by selecting the spectral clustering framework, which is known to capture perceptual organization features, as well as by developing similarity models according to Gestaltic laws of visual segregation. Our proposed framework applies but is not limited to the detection of cells and organelles in microscopic images and attempts to provide an effective alternative to presently dominating manual segmentation and tissue classification practice. The main theoretical contribution of our work resides in the formulation of robust similarity models which automatically adapt to the statistical structure of the biological domain and return optimal performance in pixel classification tasks under the wide variety of distributional assumptions.

*Key words:* Molecular imaging, Perceptual organization, Spectral clustering, Learning models

---

## 1. Introduction

The most prominent challenge in the present state of the art in human- and computer-based visual inspection of micrographs is the complexity and variety of biological structures at the micro- and nano-scales. The image characteristics of cellular and subcellular environments differ widely according to the organism, tissue type, anatomical sample preparation practices, staining protocols and not at least the applied imaging modality. A recent review on segmentation tools in electron tomography [2] emphasizes the need for more objective segmentation tools. Manual segmentation is the present clinical practice [17] and to illustrate the time complexity of this process we refer to the study published in [14]. Here, the authors report that image

---

\*Corresponding author.

*Email address:* julia.dietlmeier@ieee.org (Julia Dietlmeier)

acquisition and the interpretation of the complete Golgi ribbon required approximately 9 months of manual segmentation followed by ~ 3 months of additional editing for detailed analysis. Noske et al. in [16] report a considerable throughput improvement as compared to the previous study and remark that by using a manual segmentation software IMOD [13] they were able to mark-up all of the mitochondria in less than 15 hours.

A highly active research direction towards the automation pursuit in biological image analysis constitutes statistical machine learning within the class of *semi-supervised* methods [25]. The guiding idea behind these approaches is that given manually labeled images, the goal is to learn models that can identify novel instances of the regions in test datasets [15]. Within this field, a commercial software package to support the automatic segmentation of subcellular structures such as mitochondria, for example, is yet to become available and up to date only a few applied solutions have been reported. Due to their central role in cellular bioenergetics, apoptosis and autophagy, identification and classification of mitochondria is an area of increasing biomedical importance. Narasimha et al. in [15], for example, propose an automatic texture-based joint classification and segmentation of mitochondria by utilizing  $k$ -nearest neighbor ( $k$ -NN) classifier, support vector machines (SVMs), adaptive boosting (AdaBoost) and histogram matching. Another texture-based approach is reported in [23] where authors consider the detection of mitochondria in TEM images of brain tissue and apply the Gentle-Boost classifier. A timely review of training-based pattern recognition techniques with application to biological image analysis is provided in [21] where the authors also emphasize the major limitation of the framework as being the considerable amount of data required during the training process.

Fully *unsupervised* learning methods, on the other hand, skip the training stage and do not necessarily require a-priori knowledge about the structure of an object. Here, the partitioning of an image into a set of discrete clusters is implemented with dimensionality reduction techniques such as Principal Component Analysis (PCA), its non-linear extension to kernel-PCA, Linear Discriminative Analysis (LDA), Independent Component Analysis (ICA) and many others [12]. Particularly, spectral methods, which consider eigenvector-based analysis of features [1, 9] are known to capture perceptual organization features and to efficiently separate important (foreground) from non-important (background) information [1, 18]. This ability is based on the fact that the leading eigenvectors of a similarity matrix account for the most variability in the data and contain the information about the most dissimilar regions of an image. A similarity model, which serves as an input to a spectral clustering algorithm, encompasses various image features such as texture, color, motion, brightness or higher level primitives such as extracted line segments and contours.

In this work we utilize the perceptual grouping strategy firstly by selecting the spectral clustering framework and secondly by developing our similarity models according to Gestaltic laws of visual segregation

[3, 10], such as *Closure*, *Proximity*, *Similarity* and *Figure-Ground*. Within this framework, we view a mitochondrion as a closed membrane organism and specifically target the detection of membranes. One immediate asset of our approach is that the outer membrane is the only descriptive feature of a mitochondrion which can be considered invariant under the apoptotic (pathogenic) changes of inner mitochondrial morphology [22]. Another advantage is based on the fact that perceptual systems reflect scale-invariance of the environment provided they are adapted to its statistical structure [5]. This inference is of special importance when working with large scale microscopic images and allows us to apply low-resolution processing with the objective to reduce computational load. At last, the manually annotated data used in our study for algorithm validation build an additional interface between artificial and human vision.

The major contribution of our paper is the development of an *unsupervised* segmentation framework based on the combination of principles governing the perceptual organization and the unsupervised machine learning. In Section 3 we propose the concept of adaptive similarity mixtures which eliminates the need for user intervention towards the selection of parameters. In Section 4 we systematically validate our algorithm on two datasets which have been specifically selected to test the sensitivity of detection and exhibit the worst case image characteristics such as low contrast, non-uniform illumination, speckle noise, absence of staining, varying image sizes and resolution. Finally, we conclude our study in Section 5.

## 2. Recursive spectral clustering framework

In order to compare  $n$  image features we construct the corresponding  $n \times n$  similarity or *affinity* matrix  $\mathbf{S}$  and the corresponding dissimilarity matrix  $\mathbf{A}$ . Although the spectral methods differ widely in the type of normalization of the matrix used for diagonalization, in this work we adopt a clustering framework based on *kernel*-PCA, Euclidean embedding and  $k$ -means which is detailed in [20]. The option of non-linear *kernelization* of the feature space allows us to employ a clustering algorithm with the linear cut capability, such as  $k$ -means. We briefly review the main processing path of *kernel*-PCA and introduce the notion of a centralized dot product matrix  $\mathbf{C} = -0.5\mathbf{Q}\mathbf{A}\mathbf{Q}$ , with  $\mathbf{Q} = \mathbf{I}_n - \frac{1}{n}\mathbf{1}_n\mathbf{1}_n^T$ . Here,  $\mathbf{I}_n$  denotes a  $n \times n$  identity matrix,  $\mathbf{1}_n = (1, 1, \dots, 1)^T$  is an  $n$  vector of ones and  $\mathbf{Q}$  is the projection matrix on the orthogonal complement of  $\mathbf{1}_n$ . Reconstructed data  $\mathbf{X}$  are obtained through spectral decomposition of  $\mathbf{C} = \mathbf{U}\mathbf{\Lambda}\mathbf{U}^T$  and the subsequent re-scaling by applying  $\mathbf{X} = \mathbf{U}\sqrt{\mathbf{\Lambda}}$  [20]. The rows of  $\mathbf{X}$  correspond to the multidimensional vectors  $\mathbf{x}_i, i = 1, \dots, n$ . The dimensionality reduction step is carried out by taking the first  $L$  columns of  $\mathbf{X}_L = \mathbf{U}_L\sqrt{\mathbf{\Lambda}_L}$ . We retain the leading component  $X_1$  and apply  $k$ -means in the feature space to obtain the set of discrete clusters.

Table 1: Notations.

---



---

$n$	number of pixels in an image
$\mathbf{S}$	similarity matrix
$\mathbf{A}$	kernel matrix (dissimilarity matrix)
$\mathbf{C}$	double-centered kernel matrix
$\mathbf{C}$	$n$ vector of ones
$\mathbf{I}$	$n \times n$ identity matrix
$\mathbf{Q}$	projection matrix on the orthogonal complement of $\mathbf{1}_n$
$\mathbf{U}$	matrix containing eigenvectors
$\mathbf{\Lambda}$	matrix containing eigenvalues on the main diagonal
$\mathbf{X}$	kernel-PCA based feature vectors
$L$	number of dimensions
$k$	number of $k$ -means clusters
$c$	image downscaling factor
$f$	Gaussian filter size
$G$	Gaussian-filtered image
$\bar{G}$	inverted Gaussian-filtered image (membranes are in bright pixels)
$\nabla$	image gradient
$I$	image intensity
$F$	intensity-based similarity nominator
$D$	distance-based dissimilarity denominator
$\mathbf{Z}$	Z-score based matrix
$\alpha, \beta, \gamma$	adaptive similarity parameters (5)

---

The complete processing path is shown in Fig. 1 and we note that our algorithm is designed to work with tiled and further downscaled (by a factor of  $c$ ) images. In the first stage of clustering our objective is to close and to segment dark membranes. For this purpose we create a filtered image  $G$  by convolving the original input image with a designed Gaussian-based filter, multiplying the result with the original image and further

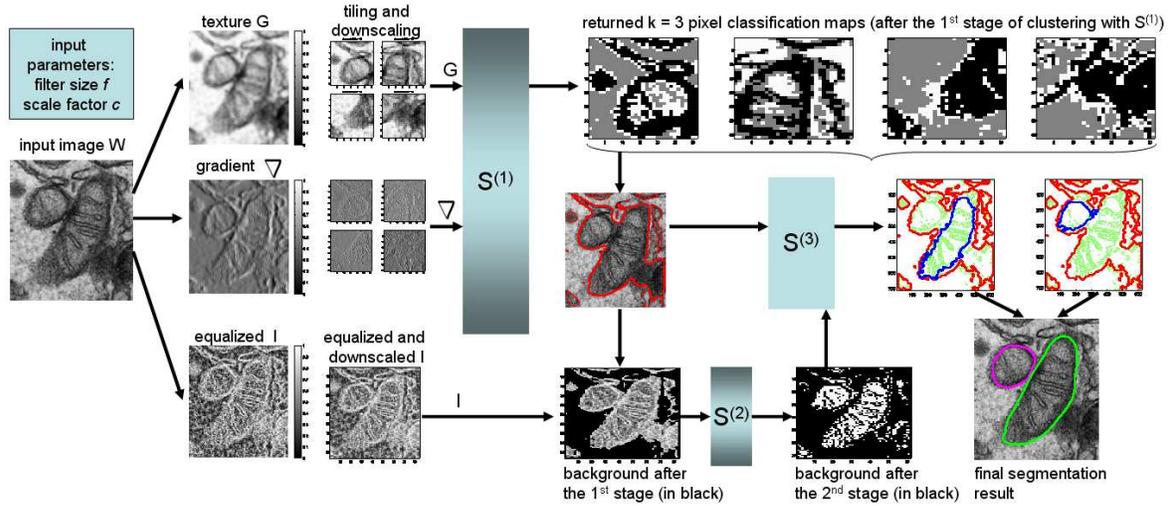


Figure 1: Proposed processing pipeline showing the progression from the input electron microscopy (EM) image on the far left to the set of detected organelles in the far right images. The three clustering stages  $m = \{1, 2, 3\}$  are denoted by  $S^{(m)}$ . Diagrams are best viewed in color

re-normalizing the final product to the  $[0, 1]$  range. The degree of *Closure* is controlled by the scale of the Gaussian filter. Separate classification maps for each tile are further combined to yield the binary result. A postprocessing step based on mathematical morphology performs the flood-fill of the inside of mitochondria and the final outer contour (shown in red) can be seen in Fig. 1. In the second recursive clustering stage, we recluster the obtained foreground with a different similarity model  $S^{(2)}$  where the objective is to classify the inner part of a mitochondrion. The third stage of clustering  $S^{(3)}$  involves a contour processing step and the computation of a minimum distance between the set of extracted inner contours  $S^{(2)}$  and the outer contour  $S^{(1)}$ . The final segmentation result in Fig. 1 shows the identification of adjacent organelles.

### 3. Proposed adaptive similarity models

In this paper we connect the notion of adaptivity with non-parametric learning models [3, p. 32] which can adjust to the biological domain without user intervention and perform equally well under a wide variety of distributional assumptions. In many studies however, the selection of parameters to accompany a similarity model is a sensitive issue and is primarily data-driven [4, 18]. One example of a state-of-the-art similarity

model is given in (1) [4, 18] and involves squared difference of intensities  $I$  between two pixels  $p_1$  and  $p_2$ :

$$F_{\Delta I}(p_1, p_2) = \exp[-(I(p_1) - I(p_2))^2/a] \quad (1)$$

where  $a$  is a penalization parameter. The goal of (1) is to discover regions with small variations in intensities and thus small intensity differences  $\Delta I$  representing similarly bright pixels are being weighted equally as similarly dark pixels of the same  $\Delta I$ . Such a model, therefore, is inherently inefficient in capturing structures represented by bright pixels. This limitation is of special importance for the selected perceptual grouping strategy by considering the fact, that in a complex visual arrangement of gray and white pixels the human attention is easily drawn to brighter areas - a phenomenon based on the Gestalt law of *Figure-Ground* [10].

In this work and in the first clustering stage, we introduce an additional intensity-based term  $F_1^{(1)}$  which weights similarly bright pixels higher than similarly dark pixels in the inverted filtered image  $\tilde{G}$  where the salient membrane features are given by bright pixels. In order to weigh the pixels near the occurring gradient edge higher, we additionally utilize the inverse exponentially mapped and normalized gradient  $\nabla$  of the input image. Our final intensity-based function  $F^{(1)}$  is a maximization between the brightness criterion represented by  $F_1^{(1)}$  and the region-based contrast criterion represented by  $F_2^{(1)}$ . The weighting constants  $k_1$  and  $k_2$  provide an option to balance the influence of an either term as shown in (2):

$$\begin{aligned} F_1^{(1)}(p_1, p_2) &= k_1 \cdot [\tilde{G}(p_1) \cdot e^{-\nabla(p_1)} + \tilde{G}(p_2) \cdot e^{-\nabla(p_2)}], \\ F_2^{(1)}(p_1, p_2) &= k_2 \cdot [\max(\tilde{G} \cdot e^{-\nabla}) - |\tilde{G}(p_1) \cdot e^{-\nabla(p_1)} - \tilde{G}(p_2) \cdot e^{-\nabla(p_2)}|], \\ F^{(1)}(p_1, p_2) &= \max(F_1^{(1)}, F_2^{(1)}), \\ F^{(2)}(p_1, p_2) &= \exp[-|I(p_1) - I(p_2)|], \end{aligned} \quad (2)$$

The last equation in (2) shows the intensity-based model  $F^{(2)}$  for the second recursive stage of spectral clustering where the objective is set to recluster the foreground extracted after  $S^{(1)}$  (please refer to Fig. 1).

In order to include the Gestalt law of *Proximity* into our model we consider the squared Euclidean distance  $D$  between pixels  $p_1$  and  $p_2$ . In the final similarity model, our objective is to let  $F$  to decay with increasing  $D$  in a ratio relationship. In order to alleviate the problem that coordinates and intensity are measured on different scales and in different quantities we propose the following three normalization procedures (see also Fig.3).

First, contrary to formulations presented, for example, in [4, 18], we avoid user-based selection of a distance penalization parameter  $\sigma_D^2$  by defining  $D$  in terms of Z-score normalized coordinates  $z_x$  and  $z_y$ :

$$D^{(1)}(p_1, p_2) = D^{(2)}(p_1, p_2) = \exp[(z_{x1} - z_{x2})^2 + (z_{y1} - z_{y2})^2] \quad (3)$$

Second, in every clustering stage we construct two matrices  $\mathbf{F}$  and  $\mathbf{D}$  according to pairwise similarities (2) and pairwise distances (3) and define the final similarity matrix  $\mathbf{S}$  as a pairwise ratio of  $\mathbf{F}$ - and  $\mathbf{D}$ -based Z-scores, denoted  $\mathbf{Z}_F$  and  $\mathbf{Z}_D$  where higher Z-scores correspond to higher probability of non-accidental occurrence. Standardization is attained with arithmetic mean  $\mu$  and standard deviation  $\sigma$ , both being computed over all matrix entries [6].

Third, in order to control the degree of mixing  $\mathbf{Z}_F$  and  $\mathbf{Z}_D$ , we introduce two constants  $\alpha \in \mathbb{R}^+$  and  $\beta \in \mathbb{R}$  in the denominator as can be seen in (5). In this formulation, the parameter  $\alpha$  can be thought of as the standard deviation of the new distribution

$$\mathbf{Z}'_D = \alpha \cdot (\mathbf{Z}_D + \beta) \quad (4)$$

and the parameter  $\beta$  can be thought of as the mean of  $\mathbf{Z}'_D$ . Thus,  $\alpha$  controls the increase or decrease in the variability of the  $\mathbf{Z}'_D$  distribution and  $\beta$  controls the shift of  $\mathbf{Z}'_D$  to the left for  $\beta < 0$  or to the right for  $\beta > 0$  while the goal is to align the right tail of  $\mathbf{Z}_F$  and the left tail of  $\mathbf{Z}_D$ . In every clustering stage  $m$ , the final similarity matrix  $\mathbf{S}^{(m)}$  is obtained according to:

$$\begin{aligned} \mathbf{S}^{(m)} : s_{ij}^{(m)} &= z_{Fij}^{(m)} / [\alpha^{(m)} \cdot (z_{Dij}^{(m)} + \beta^{(m)})], \\ \beta^{(m)} &= \gamma^{(m)} + |\min[\mathbf{Z}_D^{(m)}]|, \\ \gamma_{opt}^{(m)} &= \max(\mathbf{Z}_F^{(m)}), \\ \alpha_{opt}^{(m)} &= 1 \end{aligned} \quad (5)$$

For a case of an approximately *Gaussian* probability density function (pdf) of  $\mathbf{Z}_F$  one possible scenario of optimal sampling is to position the left end of  $\mathbf{Z}_D$  at  $3\sigma_{Z_F} = 3$  which captures 99.7% of the random variation in  $\mathbf{Z}_F$ . Applications seeking to increase the influence of the distance term may scale  $\mathbf{Z}_D$  by  $\alpha < 1$ .

For a general case of a *non-Gaussian* pdf of  $\mathbf{Z}_F$  we propose to align the right end of  $\mathbf{Z}_F$  with the left end of  $\mathbf{Z}_D$  (5) and by doing so to completely separate the two distributions. Two examples in Fig. 2 are provided to illustrate that by varying  $\alpha$  and  $\gamma$  we can control the accuracy of membrane detection in terms of thickness and gap completion. Specifically by comparing columns (c) and (e) in Fig. 2 it can be seen that moving  $\mathbf{Z}_D$  away from  $\mathbf{Z}_F$  (compare Fig. 2(b) and Fig. 2(d)) reduces the amount of identified false positives.

The final step in data preparation before proceeding with spectral decomposition is the dissimilarity mapping and kernelization. These procedures are necessary in order to transform the final similarity matrix  $\mathbf{S}$  into a centralized pseudo-covariance matrix  $\mathbf{C}$  which will further be used as an input to kernel-PCA (explained for example in [20]). For our two spectral clustering stages we applied the linear dissimilarity

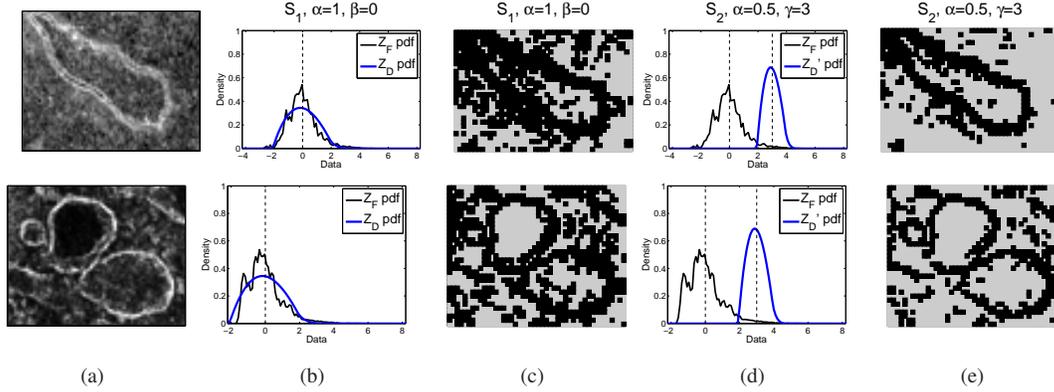


Figure 2: Concept of adaptive similarity mixtures  $\mathbf{S}$  where the nominator  $\mathbf{F}$  and the denominator  $\mathbf{D}$  are modeled as probability density functions  $\mathbf{Z}_F$  and  $\mathbf{Z}_D$  through the introduced standardization of matrix elements. The bright membrane structures (here in inverted EM images) are captured in the right half of  $\mathbf{Z}_F$

mapping and for the first stage  $\mathbf{C}^{(1)}$  we additionally applied the square root kernel:

$$\begin{aligned} \mathbf{C}^{(1)} &= -\frac{1}{2}\mathbf{Q}\sqrt{(\max(\mathbf{S}) - \mathbf{S})\mathbf{Q}}, L^{(1)} = 1, \quad k^{(1)} = 3 \\ \mathbf{C}^{(2)} &= -\frac{1}{2}\mathbf{Q}(\max(\mathbf{S}) - \mathbf{S})\mathbf{Q}, L^{(2)} = 1, \quad k^{(2)} = 2 \end{aligned} \quad (6)$$

After the diagonalization of  $\mathbf{C}$ , the final set of discrete clusters is obtained with  $k$ -means clustering algorithm on the retained first ( $L = 1$ ) kernel-PCA output feature vector  $X_1$ .

As can be seen in (6), the first clustering stage returns 3 clusters  $\{k_1, k_2, k_3\}$ , where the decision on the foreground is made in the image domain according to  $\mu(W(k_1)) < \mu(W(k_2)) < \mu(W(k_3))$  and thus we select the cluster with the minimal average brightness. Our choice of the 3-class partition was motivated by the assumption that with  $k > 2$  we can segment thinner structures while closing the membrane gap with the width  $d$  and at the same time maintaining the separation of close lying organelles with the same  $d$ . The above mentioned effect of  $k = 3$  can, for example, be observed in the returned pixel classification maps in Fig. 1.

The third clustering stage  $S^{(3)}$  is outside the spectral clustering framework. For the number of extracted inner contours obtained after  $S^{(2)}$  we compute and evaluate Euclidean distances between the inner contour and the outer contour obtained after  $S^{(1)}$  with the objective to retain only those outer contour points with the minimal distance to the currently analyzed inner contour.

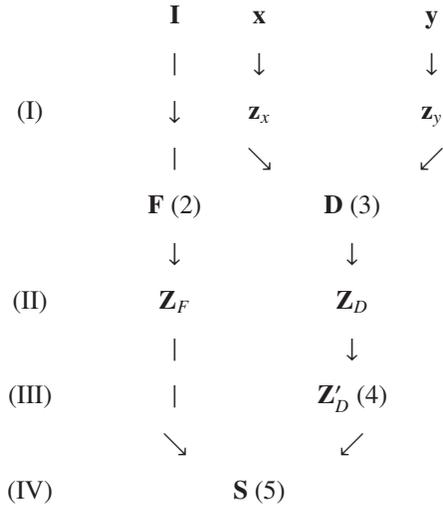


Figure 3: Proposed four-stage cascaded raw data normalization and data fusion. Stage (I) consists of standardization of the raw coordinates  $x$  and  $y$ . Stage (II) consists of the computation of the Z-score matrices  $\mathbf{Z}_F$  and  $\mathbf{Z}_D$ . Stage (III) determines the optimal shift of  $\mathbf{Z}_D$ . Stage (IV) determines the final data fusion.

#### 4. Experimental results

We applied our model to two datasets each consisting of 17 EM images provided by the American Society of Cell Biology. The images contain mitochondria from the *ductulus efferens* of the ground squirrel, *Citellus lateralis* [11]. All mitochondria of this type have an elongated profile, however some organelles appear spherical or elliptical with the "washed-out" cristae and/or outer membrane and some organelles contain the shadow characteristics in the outer region as can be seen in Fig. 4(a). These problems are related to the section thickness in typical EM studies ranging from 50-90nm and not being thick enough to include the whole mitochondria [11]. Thus, the shadow artefacts mentioned above are the result of tangentially oriented membrane within the section thickness. Therefore, based on the suggestion made by our biomedical partners, the data annotation was carried out as a smooth closed contour extrapolation.

In order to construct the first Dataset I, we extracted the regions of interest at different magnifications. Thus, images in the first dataset have different resolution and therefore the thickness of a membrane is assumed to be unknown. In the acquisition of the second dataset we applied the equal magnification while extracting the regions of interest. All images have different sizes and generally the image quality and the contrast are extremely low, which on the other hand allows us to test the sensitivity of our detection algo-

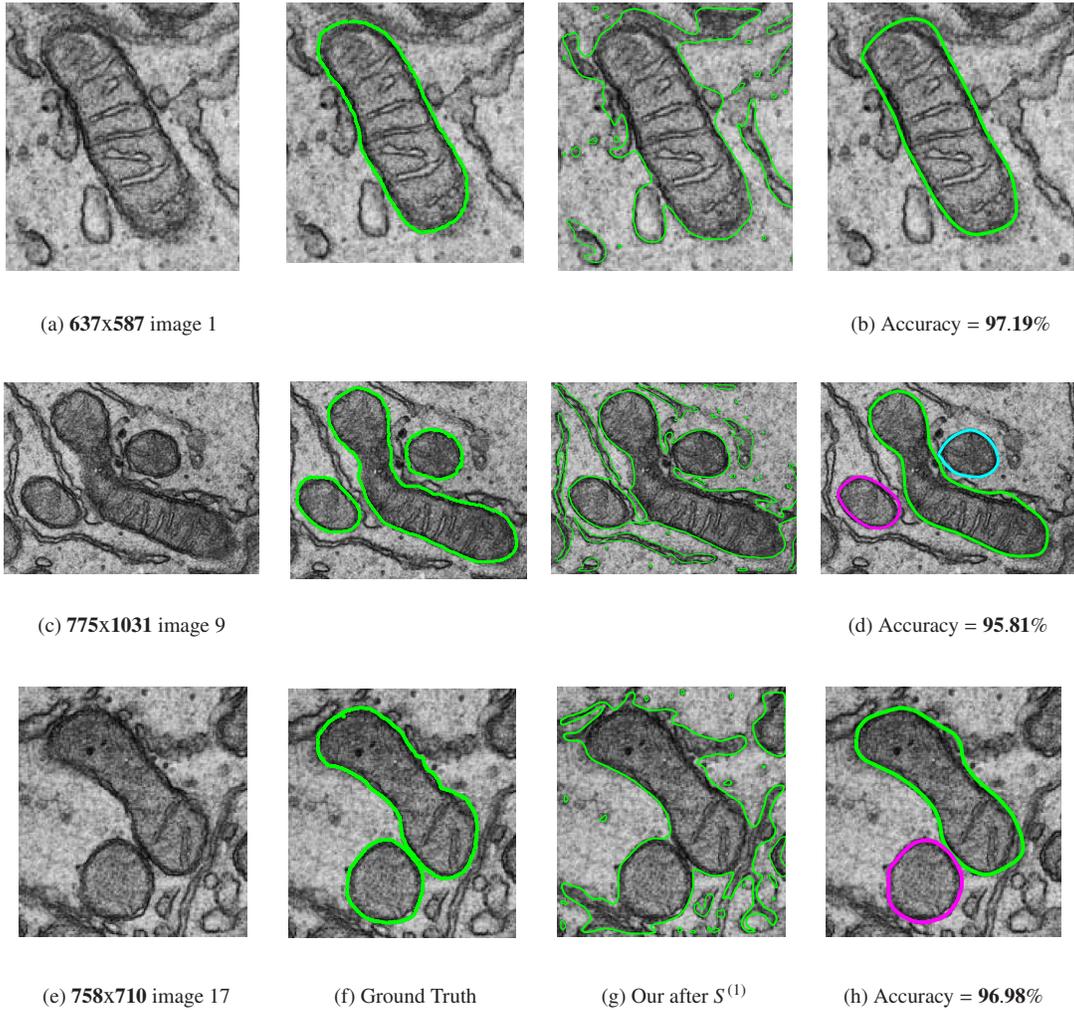


Figure 4: Selected qualitative results for the Dataset1 ( $c = 20, f = 5$ ). (a,c,e): original EM images; Column (f): clinically validated Ground Truth segmentation; Column (g): Our intermediate segmentation after the 1st clustering stage; **(b,d,h): Our final segmentation result**. Results using a user-interactive semi-automated software "Livewire" [24] are shown in Fig.5. Diagrams are best viewed in color.

rithm. The weighting constants  $k_1$  and  $k_2$  as in (2) are selected according to the robust Qn spread estimator of the input data  $G$  [6] such that  $k_1 = \sigma_{Qn}^2(G)$  and  $k_2 = 1 - \sigma_{Qn}^2(G)$ .

Selected qualitative segmentation results are provided in Fig. 4 where the comparison is given between the original image, segmentation achieved by using 2D user-interactive semi-automated tracing software

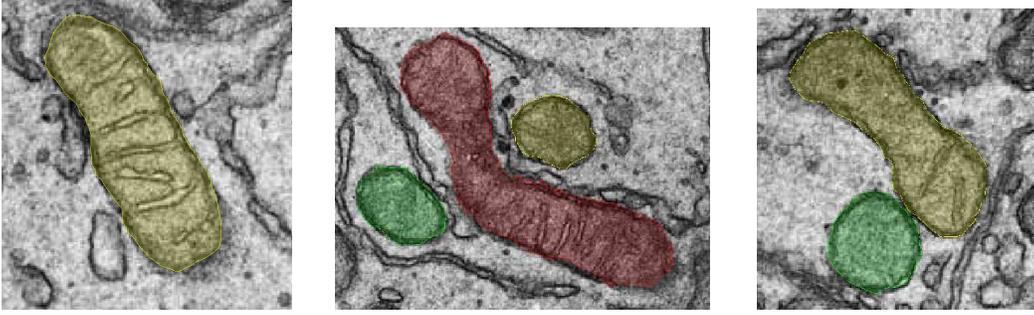


Figure 5: Selected segmentation results using the **user-interactive semi-automated tracing software "Livewire"** [24]. Diagrams are best viewed in color.

"Livewire" (Fig.5) which is a popular medical segmentation tool [24], our segmentation results and the manually annotated (Adobe Photoshop 7.0) and clinically validated ground truth data which has been independently generated for each dataset. Specifically, Fig. 5 demonstrates the high accuracy in semi-automated segmentation using "Livewire" and shows its high performance when applied to the segmentation of mitochondria and compared to the Ground Truth data in Fig.4 column (f).

For quantitative assessment the result of segmentation and the ground truth data are given by *filled* closed segmented contours, such that the foreground object (mitochondria) is 1 (white) and the background is 0 (black). In order to evaluate quality of binary segmentation we selected Precision, Recall, Accuracy [19] and Dice coefficient [26] performance measures which are based on the combination of the number of true positives (TP), true negatives (TN), false positives (FP) and false negatives (FN) classified pixels.

$$Precision = \frac{TP}{TP + FP} \quad (7)$$

$$Recall = \frac{TP}{TP + FN} \quad (8)$$

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \quad (9)$$

$$Dice = \frac{2TP}{TP + FP + TP + FN} \quad (10)$$

A true positive (TP) is 1 when output of the segmentation is 1 and the ground truth is 1, the true negative (TN) is 1 when the result of segmentation is 0 and the ground truth is 0. A false positive (FP) is 1 when the result of segmentation is 1 and ground truth is 0. A false negative (FN) is 1 when the result of segmentation is 0 and the ground truth data is 1. For the perfect segmentation the Precision and Recall will be 1. The

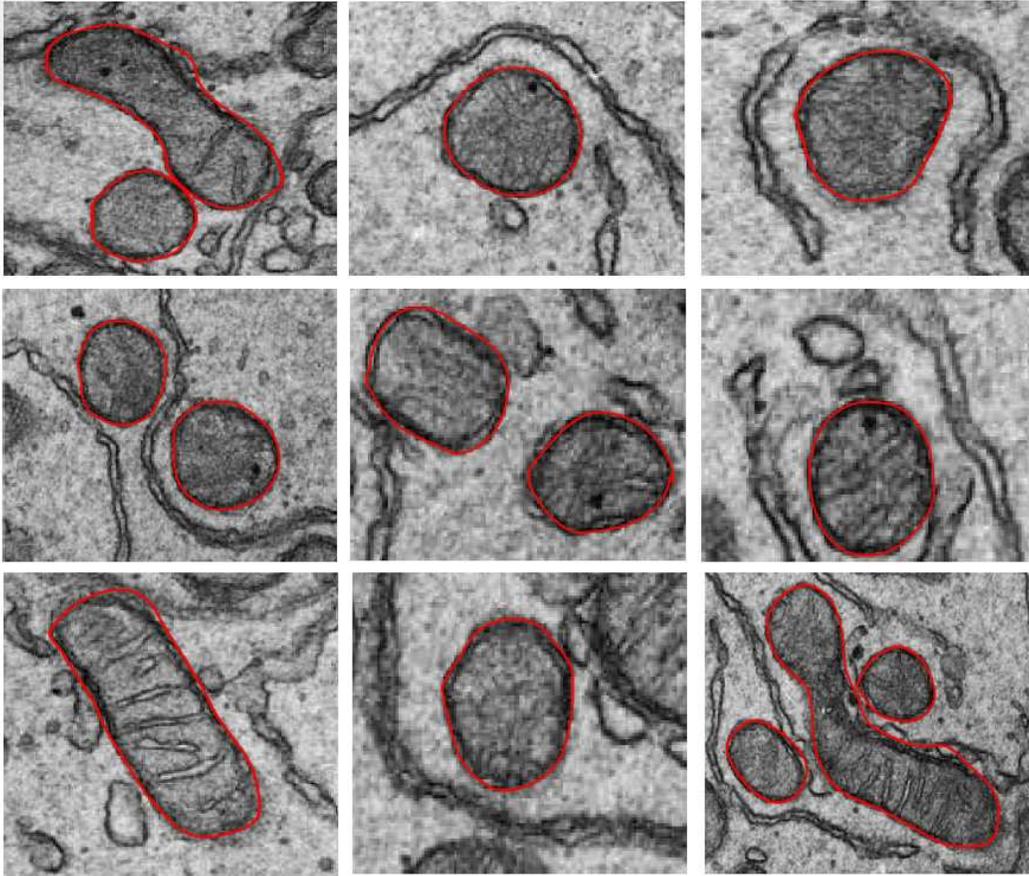


Figure 6: More segmentation results using our developed segmentation approach [7]. Quantitative performance evaluation is given in Fig.7. Diagrams are best viewed in color.

oversegmentation is the case when Precision is low and the undersegmentation is the case of low Recall. Accuracy is the ratio of the sum of TN and TP to the total number of pixels. The Dice coefficient is a simple and useful measure of spatial overlap which is applied to assess the accuracy in image segmentation. The value of Dice coefficient ranges from 0, indicating no spatial overlap between two sets of binary segmentation results, to 1, indicating complete overlap [26].

The quantitative results are shown in Fig. 7 where the Precision, Recall, Dice and Accuracy measures have been calculated with respect to the accuracy of membrane points in the segmentation and ground truth data. The quantitative results in Fig. 7 show that the average accuracy of our approach is above 90% except

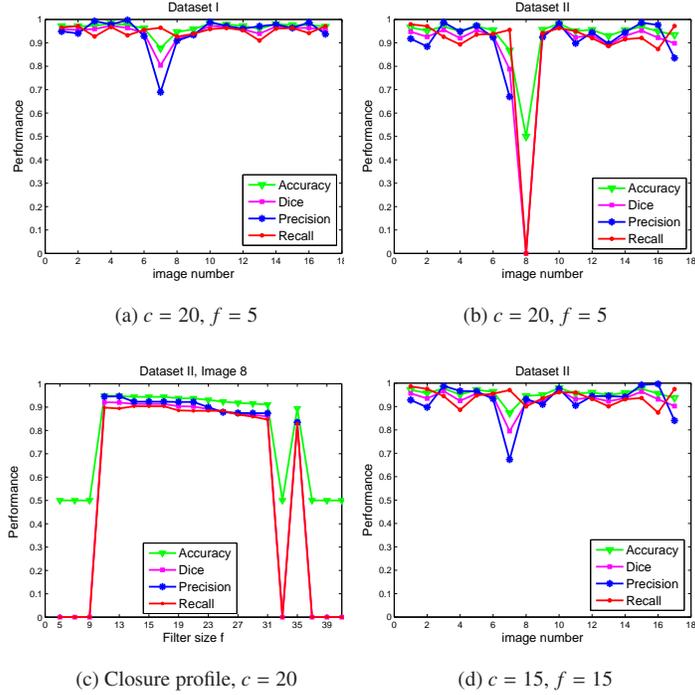


Figure 7: Quantitative results. (a) Dataset I (DI). The drop in performance for the image 7 is due to an additional mitochondrion which has been detected by our algorithm but omitted in the ground truth segmentation. (b) Dataset II (DII) with the parameters optimized for DI. (c) Closure profile for the image 8. (d) Performance graph for DII with optimized  $c$  and  $f$ . Diagrams are best viewed in color.

for the image 7 where the segmentation accuracy drops to 87.72% (Fig. 7a) and 87.19% (Fig. 7d). The performance graph for the second dataset DII in Fig. 7(b) shows one current limitation of our approach such as the automatic selection of the *Closure* related filter size  $f$  corresponding to the downscaling factor  $c$ . The closure profile of the image 8 is shown in Fig. 7(c) and suggests the appropriate filter size ranging from 9 to 31. The results with optimized  $c$  and  $f$  in Fig. 7(d) show the improved performance as compared to Fig. 7(b).

## 5. Conclusions

This paper presented a new approach to the unsupervised segmentation of mitochondria in EM images based on the combination of principles governing the perceptual organization and the unsupervised machine learning. The major theoretical novelty of our paper lies primarily in the development of adaptive similarity models which eliminate the need for user intervention towards the selection of penalization parameters. In

our experimental section, we demonstrated the successful separation of adjacent organelles while maintaining the segmentation accuracy of above 90% when working with low contrast large scale EM images. The selected results based on the approach detailed in this paper have been presented at the Biophotonics and Imaging conference (BioPic) [7]. Our future work is focused on efficient computational approaches that permit processing of very large scale images [8].

### **Acknowledgments**

This research was supported by the National Biophotonics and Imaging Platform Ireland funded under the HEA PRTL Cycle 4, co-funded by the Irish Government and the European Union - *Investing in your future*. The authors would like to thank the American Society for Cell Biology (ASCB) for provision of EM images.

### **References**

- [1] Alzate,C., and Suykens,J. A. K.: Multiway Spectral Clustering with Out-of-Sample Extensions through Weighted Kernel PCA. *IEEE TPAMI*, 32(2):335347, (2010).
- [2] Barcena, M., and Koster,A.J.: Electron tomography in life science. *Seminars in Cell and Developmental Biology*, 20:920930, (2009).
- [3] Boyer, K.L., and Sarkar,S.: *Perceptual Organization for Artificial Vision Systems*. Kluwer, (2000).
- [4] Boykov,J. and Funka-Lea,G.: Graph Cuts and Efficient N-D Image Segmentation. *IJCV*, 70(2):109131, (2006).
- [5] Chater,N. and Brown,G.: Scale-invariance as a unifying psychological principle. *Cognition*, 69:B17 B24, (1999).
- [6] Daszykowski,M. and et al.: Robust statistics in data analysis - A review. basic concepts. *Chemometrics and Intelligent Laboratory Systems*, 85:203219, (2007).
- [7] Dietlmeier,J. and et al.: Unsupervised segmentation of mitochondria using model-based spectral clustering, *BioPIC 2013 BioPhotonics and Imaging Conference 25th-27th March 2013, Castleknock Country Club, Dublin, Ireland (2013)*. Available online at [www.cipa.dcu.ie/](http://www.cipa.dcu.ie/)

- [8] Dietlmeier,J. and et al.: "A New Anticorrelation-based Spectral Clustering Formulation", ACIVS 2011-Advanced Concepts for Intelligent Vision Systems, Aug. 22-25 2011, Het Pand, Ghent, Belgium,(2011). Published in Lecture Notes on Computer Science (LNCS), Springer.
- [9] Donoser,M. and et al.: Saliency Driven Total Variation Segmentation. ICCV, (2009).
- [10] Ehrenstein,W.,H., and et al.: Gestalt issues in modern neuroscience. *Axiomathes*, 13(3):433458, (2003).
- [11] Fawcett,D.,W.: *The Cell*. <http://www.ascb.org>, (1981).
- [12] Jolliffe,I.,T.: *Principal Component Analysis*, 2nd ed. Springer-Verlag, (2002).
- [13] Kremer,J.R., and et al.: Computer visualization of three-dimensional image data using IMOD. *Journal of Structural Biology*, 116:7176, (1996).
- [14] Marsh,B.J., and et al.: Organellar relationships in the golgi region of the pancreatic beta cell line, HIT15, visualized by high resolution electron tomography. *PNAS USA*, 98:23992406, (2008).
- [15] Narasimha,R., and et al.: Automatic joint classification and segmentation of whole cell 3D images. *Pattern Recognition*, 42:10671079, (2009).
- [16] Noske,A.,B., and et al.: Expedited approaches to whole cell electron tomography and organelle mark-up in situ in high-pressure frozen pancreatic islets. *Journal of Structural Biology*, 161:298313, (2008).
- [17] Perkins,G.A., and et al.: Correlated light and electron microscopy /electron tomography of mitochondria in situ. *Methods in Enzymology*, 456:2952, (2009).
- [18] Perona,P., and Freeman,W.T.: Factorization Approach to Grouping. *ECCV*, pages 655670, (1998).
- [19] Powers,D.M.W.: Evaluation: From Precision, Recall and F-measure to ROC, informedness, markedness and correlation. *Journal of Machine Learning Technologies*, 2(1):3763, (2011).
- [20] Roth,V., Laub,J., Kawanabe,M., and Buhmann,J.: Optimal Cluster Preserving Embedding of Non-metric Proximity Data. *IEEE TPAMI*, 25(12):15401551, (2003).
- [21] Shamir,L., Delaney,J.D., Orlov,N., Eckley,D.M., and Goldberg,I.G.: Pattern recognition software and techniques for biological image analysis. *PLoS Computational Biology*, 6(11):110, (2010).

- [22] Sun,M.G., and et al.: Correlated three-dimensional light and electron microscopy reveals transformation of mitochondria during apoptosis. *Nature Cell Biology*, 9(9):10571065, (2007).
- [23] Vitaladevuni,S., and et al.: Mitochondria detection in electron microscopy images. MIAAB, New York, USA, (2008).
- [24] Suzuki, C.T.N. and Falcão, A.X. Livewire: 2D interactive tool for medical image segmentation. Available online at <http://www.ic.unicamp.br/~afalcao/livewire/>.
- [25] Zheng,N., and Xue,J.: *Statistical Learning and Pattern Analysis for Image and Video Processing*. Springer, (2009).
- [26] Zou,K.H., and et al.: Statistical validation of image segmentation quality based on a spatial overlap index. *Acad. Radiology*, 11(2):178189, (2004).