SEGMENTATION OF OVERLAPPING MACROPHAGES USING *ANGLEGRAM* ANALYSIS

José Alonso Solís-Lemus[1], Brian Stramer[2], Greg Slabaugh[1], and Constantino Carlos Reyes-Aldasoro[1]

[1] School of Mathematics, Computer Science and Engineering, City, University of London, UK
[2] Randall Division of Cell & Molecular Biophysics, King's College London, UK

Abstract. This paper describes the automatic segmentation of overlapping cells through different algorithms. As the first step, the algorithm detects junctions between the boundaries of overlapping objects based on the angles between points of the overlapping boundary. For this purpose, a novel 2D matrix with multiscale angle variation is introduced, i.e *anglegram*. The anglegram is used to find junctions of overlapping cells. The algorithm to retrieve junctions from the boundary was tested and validated with synthetic data and fluorescently labelled macrophages observed on embryos of *Drosophila melanogaster*. Then, four different segmentation techniques were evaluated: (i) a Voronoi partition based on the nuclei positions, (ii) a slicing method, which joined the clumps together (junction *slicing*), (iii) a partition based on the following of the edges from the junctions (*edge following*), and (iv) a custom self-organising map to fit to the area of overlap between the cells. Only (ii)-(iv) were based on the junctions. The segmentation results were compared based on precision, recall and Jaccard similarity. The algorithm that reported the best segmentation was the junction slicing.

Keywords: Segmentation, overlapping objects, macrophages, self-organising maps

1 INTRODUCTION

The migration of cells is of great importance in many biological processes, one of them is within the immune system. Macrophages are one of the cells of the immune system that settle in lymphoid tissues and the liver, which serve as filters for trapping microbes and foreign particles [1]. Cell migration is an essential biological process that ensures homeostasis in adults, where an unbalanced migratory response results in human disease [2]. The model organism *Drosophila melanogaster* can offer complementary insights into how macrophages integrate cues to migration [3]. It has been shown that interactions amongst the cells' structures appear to anticipate the direction of migration [4], thus, accurate cell segmentation could provide information for specific cells for biological studies.

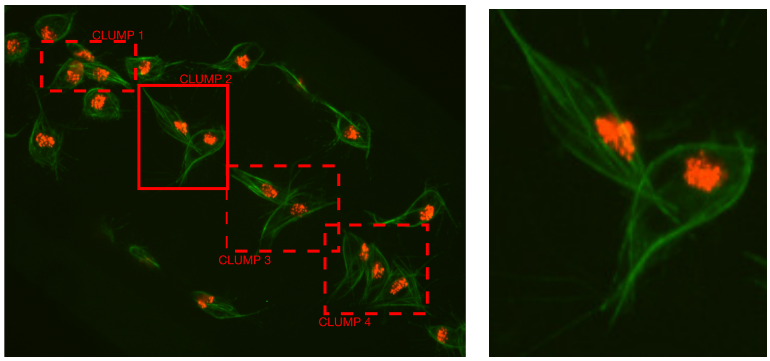
Segmentation of cells in fluorescence microscopy is a widely studied area [5], with many approaches ranging from thresholding techniques [6], to active surfaces [7]. In recent years, techniques like adaptive active physical models [8] and multilevel sets [9] have been used to address the problem of cells that overlap in cervical cancer images. Other techniques like self-organising maps (SOM) have also been used for biomedical image segmentation [10].

Junctions are commonly acquired by looking for extrema in the curvature of the image gradient [11,12]. In this work, a novel approach to find junctions is proposed for boundaries of overlapping objects, whose intersections would correspond to the junctions acquired. The junctions detected would later be used as the basis for completing a segmentation of the overlapping cells. Four methods are presented, of which three use the information from the junctions detected.

2 MATERIALS

2.1 Macrophages embryos

Fluorescently labelled macrophages were observed in embryos of the model organism *Drosophila melanogaster*. The nuclei were labelled with GFP-Moesin, which appeared red, whilst the microtubules were labelled with a green microtubule probe (Clip-GFP) [4]. RGB images of dimensions $(n_h, n_w, n_d) = (512, 672, 3)$ and $t = 541$ time frames were acquired. The images have two layers of fluorescence. Figure 1 shows one representative time frame. The green channel illustrates overlap that makes an accurate segmentation of the cells complicated.



(a) Full frame. (b) CLUMP 2.

Fig. 1: Example of cell overlapping in a single frame. (a) Presents the full frame with (red) squares highlighting all regions where instances of overlapping cells (clumps) are shown and labelled for easy reference. (b) Detail of CLUMP 2, present in (a).

2.2 Synthetic data

In order to assess the limitations of the junction detection methodology, images of pairs of synthetic overlapping ellipses with varying angles and separation distances were generated ($n = 142$). Let $\mathcal{E}(\phi, \mathbf{x}_0) = \{\mathbf{x}_{\phi,\mathbf{x}_0}(t) : t \in [0, 2\pi]\}$, be the ellipse defined by the equation $\mathbf{x}_{\phi,\mathbf{x}_0}(t) = R(\phi)(a\cos(t), b\sin(t))^T + \mathbf{x}_0$, which is rotated with respect to the x-axis by ϕ degrees and whose centre is located on position $\mathbf{x}_0$. The pairs of ellipses constructed in this work differ both in angle and position with three conventions taken into consideration: (i) presetting the values of the axes (a, b); (ii) defining a *central* ellipse $\mathcal{E}_0 = \mathcal{E}(0, \mathbf{x}_0)$, common to all pairs; and (iii) the difference in position would only be made by moving the ellipses in the x-axis. Thus, the pairs of ellipses can be defined in terms of the differences to $\mathcal{E}_0$, namely the angle and distance from the centre (ϕ, Δ). A set of pairs of ellipses was generated in MATLAB® to test the method at different values of (ϕ, Δ) ranging ϕ from 0 to 90 degrees and Δ from 0 to 160 pixels with increments of 10. Images of size $(n_h, n_w) = (256, 512)$ were generated with $\mathbf{x}_0 = (128, 128)^T$ and axes $(a, b) = (120, 53)$ that contained an overlapping of $\mathcal{E}_0$ and $\mathcal{E}_{\phi,\Delta}$. Disregarding the images where there was no overlap present in the generated ellipses, a total of 142 images was generated. Fig. 2 contains a subset of the ellipses tested. Cases where there was no overlap were ignored from the analysis.

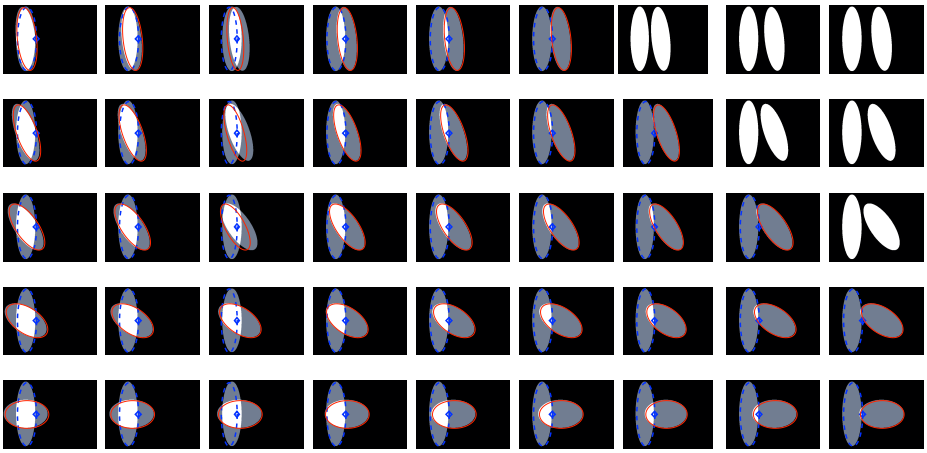


Fig. 2: Overview of the range of pairs of ellipses investigated. The pairs presented on this image represent a sample of the ellipses that were tested by the method presented. The boundary of the central ellipse $\mathcal{E}_0$ is highlighted in blue while the second ellipse's boundary is presented in red.

3 METHODS

In this work, a *clump* will be understood as a cluster of two or more overlapped objects. A clump was detected when two or more nuclei on the red channel were detected within a single region on the green channel. Segmentation of the green channel was performed by low-pass filtering with a 5×5 Gaussian filter, following with a hysteresis thresholding technique [6]. A morphological opening with a disk structural element (r=3) was performed to smooth the edges and remove noise. Segmentation of the red channel followed the same methodology. Then, the number of nuclei per region was counted to determine the presence of *clumps*.

3.1 Junction detection

Let $\mathcal{B}$ define the boundary of a clump, then for each of the ordered points $\mathbf{p}_i = \mathbf{x}_i \in \mathcal{B}$, the inner angle of the point is defined as follows,

Definition 1 (Inner angle of a point) *The inner angle of a point $\mathbf{p}_i \in \mathcal{B}$ in the boundary is the angle $\theta_{i,j}$ adjacent to the point, and measured from the jth previous point $\mathbf{p}_{i-j}$ to the following jth position $\mathbf{p}_{i+j}$.*

Fig. 3 shows examples of the calculation of an *inner angle* for a given point in the boundary. By visual inspection, it can be noticed that the inner angle of a junction would be greater than 180 degrees for a number of separations j. This number of separations will be referred to as the *depth* of the junction. Thus, the method consists of computing the inner angle $\theta_{i,j}$ at every point $\mathbf{p}_i \in \mathcal{B}$, and on every separation j. The **anglegram matrix** $\Theta = ((\theta_{i,j}))$ is defined as the values of the inner angles of each point i and per separation j, Fig. 3 (c).

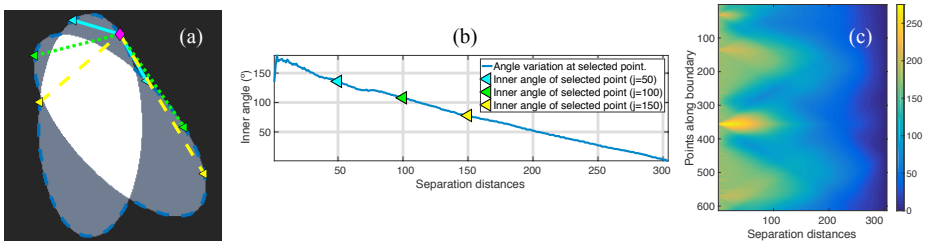
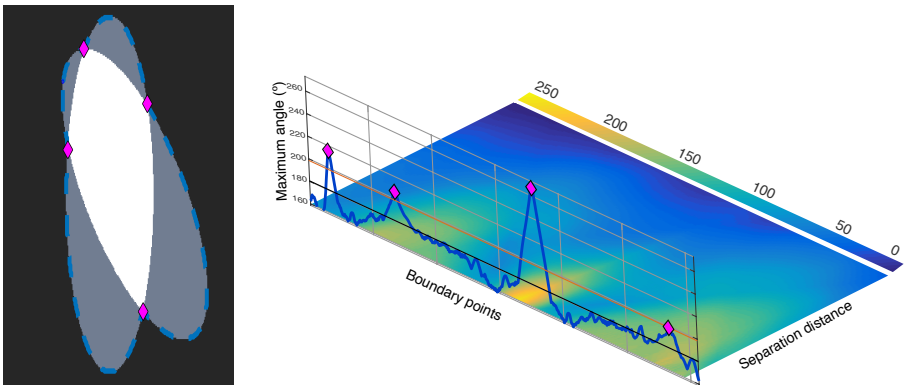


Fig. 3: Representation of *inner point angle* calculation and generation of *anglegram matrix*. (a) Represents a synthetic clump with its boundary outlined (blue, dotted), where a point (magenta $\diamond$) in the boundary will have various inner point angles per separation j. All the inner point angles for the highlighted point are displayed in (b). (c) Shows the *anglegram matrix*, where each row represents the graph displayed in (b) for each boundary point.

The local maxima on a projection over the horizontal dimension of the anglegram is related to the position of the junctions of the boundary and the

depth of the junction. Each row, $\Theta(i,:)$, corresponds to the inner angles of point $\mathbf{p}_i$, therefore taking a summary of the rows would yield a measurement of the general inner angles of each point. For this work, the maximum intensity projection $\hat{\theta}_{\max}$, Fig. 4(b), was compared with mean, median and area under the curve, but maximum provided the best results (data not shown due to space limitations). To account for quantisation errors in the boundaries extracted from the clumps, an averaging filter of size 5×5 was applied to the anglegram matrix, Θ, before the calculation of $\hat{\theta}_{\max}$. The local maxima of the 1D projection were found by using the function `findpeaks` from MATLAB®, which identifies local maxima of the input vector by choosing points of which its two neighbours have a lower value. Due to quantisation noise in $\hat{\theta}_{\max}$, the parameters `MinPeakDistance` and `MinPeakHeight` were set to empirically consistent values. First, `MinPeakDistance`, which restricts the function to find local maxima with a minimum separation, was set to 25. Furthermore, the parameter `MinPeakHeight` was set to $\text{mean}(\hat{\theta}_{\max}) + 0.75 \times \text{std}(\hat{\theta}_{\max})$.



(a) Junctions detected. (b) Anglegram matrix and maximum intensity projection.

Fig. 4: Junction detection on overlapping objects through the maximum intensity projection of the anglegram matrix. The junctions detected on a synthetic pair or ellipses is shown in (a), where the boundary of the clump is represented as a dotted line (blue) as well as the junctions (magenta $\diamond$). The definition of $\hat{\theta}_{\max}$ is represented in (b), where the anglegram matrix Θ is displayed in a plane and $\hat{\theta}_{\max}$ is represented along the boundary points. Detection of junctions are shown with $\diamond$ markers (magenta).

3.2 Segmentation of overlapping regions

This section describes the comparison of methodologies to segment clumps into overlapping cells. Initially, as a benchmark, a simple partition of the clump based on Voronoi partitioning [13] was developed. Then, the three methods, which incorporate the information from the junctions into a segmentation output were

used. The methods differed in the way the junctions' information was incorporated into a complete segmentation. Junction Slicing (JS) and Edge Following (EF) involved the explicit use of the junctions' position, while the proposed self-organising map (SOM) fitting involved the information of junctions into creating a custom SOM that adapts to the overlapping section of the data. In this work, only the cases where two junctions were found were examined in detail. A diagram showing all methods presented and the data flow is presented in Fig. 5.

Voronoi partition This method was included as a lower-bound benchmark for comparisons against with the proposed methods. The results should be the worst as no information from the green channel is used. The image area was partitioned, using Voronoi tesselations [13]. The partition of the space was based on the centroids of the detected nuclei from the image's red channel. The clumps detected on the green channel were divided based on the Voronoi partition.

Junction Slicing (JS) This method partitioned the clump with the line that joined two junctions. For each junction detected, each of the two adjacent segments of the boundary of the clump would correspond to one of the different objects within the clump. Since the points in the boundary are ordered, starting at one point $\mathbf{p}_1$ and moving alongside $\mathcal{B}$ in a clockwise manner, then the segment that appeared before a detected junction would correspond to one cell, whereas the segment that appeared after the junction would correspond to the other cell. For cases where only two junctions were found, the problem of selecting which pair of junctions will be joined becomes trivial. However, considering a case like the one presented on Figure 4(a), where four junctions would appear, different combinations of the boundary segments could yield different *candidates* of segments.

Edge Following (EF) In order to obtain the edge information, the Canny algorithm [14] was used on the green channel of the image. The algorithm consists of finding the local maxima of the image gradient. In this work, the parameter of the standard deviation was set to $\sigma = 1$. The trend of the two adjacent segments leading to the junction was defined by approximating the tangent line of the boundary at the junction point. The definition of the tangent line was taking an average slope of the secant lines leading up to the detected junction. The tangent line was extended, and a region of interest (ROI) was defined by a triangle where the approximated tangent line goes along the vertex and the adjacent angle corresponds to 20 degrees to each side of the tangent line (Figure 5). The ROI defined for each of the adjacent line segments was then intersected with the edge information of the image, resulting in a set of binary line segments, which were labelled. Labelling of the binary line segments allowed for individual analysis of each line. Each line detected was analysed in terms of its orientation and size, preserving the one that has the most similar orientation to the extended

line segment. Binary line segments with a change in direction were split by removing the strongest corners, detected through the corner detection algorithm by Harris [11]. The lines found by both ROIs on each junction were then used as new coordinates to add to the boundary of the corresponding cell.

Self-organising Maps (SOM) Fitting This work proposes an alternative implementation of the self-organising maps [15] that adapts itself to the overlapped area. For this SOM, a custom network was defined, as well as the input data and additional rules to the definition of the step-size parameter, α. Let a Network $\mathcal{N} = (\mathcal{V}, \mathcal{L})$, where $\mathcal{V} = \left\{\mathbf{m}_i = (x_i, y_i) \in \mathbb{R}^2 : i = 1, \cdots, n_v\right\}$ are nodes assigned to positions in the plane and $\mathcal{L}$ are some edges linking the some of the nodes in $\mathcal{V}$. Each node $\mathbf{m}_i \in \mathcal{V}$ has an identifier, position, and a *speed* parameter, related to the movement of each node. The input data was determined by the positions and normalised intensity values of the image, i.e. $(\mathbf{x}_t, I(\mathbf{x}_t))$. Values in $I(\mathbf{x}_t)$ that were selected by an Otsu's threshold [16] and were located within a bounding box that contains the junctions. Given an input, the algorithm proposed by Kohonen [15] follows two basic steps: identifying the closest node in the network to the input, shown in Equation (1), and update the positions of the nodes inside a neighbourhood, determined by a distance n_e to the winner node $\mathbf{m}_c$, (2),

$$\mathbf{m}_c(t) = \underset{c \in \{1,\cdots,n_v\} S}{\arg\min} \|\mathbf{x_t} - \mathbf{m}_i(t)\|_2^2 \tag{1}$$

$$\mathbf{m}_i(t+1) = \begin{cases} \mathbf{m}_i(t) + \alpha_t \left(\mathbf{x}_t - \mathbf{m}_c(t)\right) & , (i,c) \in \mathcal{L} \text{ and } \operatorname{dist}(\mathbf{m}_i, \mathbf{m}_c) \leqslant n_e \\ \mathbf{m}_i(t) & , \text{otherwise} \end{cases}, \tag{2}$$

where $\operatorname{dist}(\mathbf{m}_i, \mathbf{m}_j)$ refers to the distance from node i to node j in the shortest path determined by the edges $\mathcal{L}$. In this work, the parameter α_t was determined the intensity level of the image, I, and the speed parameter of the node. The proposed formula for the parameter α_t is shown in equation (3),

$$\alpha_{t,i} = \alpha_0 \times (0.2 + I(\mathbf{x}_t))^4 \times \operatorname{speed}(\mathbf{m}_i), \tag{3}$$

where $\operatorname{speed}(\mathbf{m}_i)$ is 0.1, or 1, depending on where the node resides in the topology. The network was defined by taking a subset of the boundary points in $\mathcal{B}$ in a ring topology, and then adding two networks in a grid topology to each side of the line joining two junctions. The three networks are independent from each other. Thus, $\operatorname{speed}(\mathbf{m}_i) = 0.1$, if $\mathbf{m}_i$ was located in the boundary of the clump, and $\operatorname{speed}(\mathbf{m}_i) = 1$, if it was one of the grid networks. The assumption is that the network taken from $\mathcal{B}$ would be closer to the actual cell, and therefore it should not move abruptly, whereas the networks inside the clump will adjust and adapt to the shape of the overlapping area between the cells. In order to finalise the network final state into a segmentation, the external network was taken as a new clump and it was partitioned by the same line used in the junction slicing (JS) method. Finally, the area formed by the inner network that adapted to the overlapping section of the cell was dilated with a 5×5 square element and then attached to both partitions of the new clump. The right column of Figure 5 displays the main steps of the SOM fitting method described.

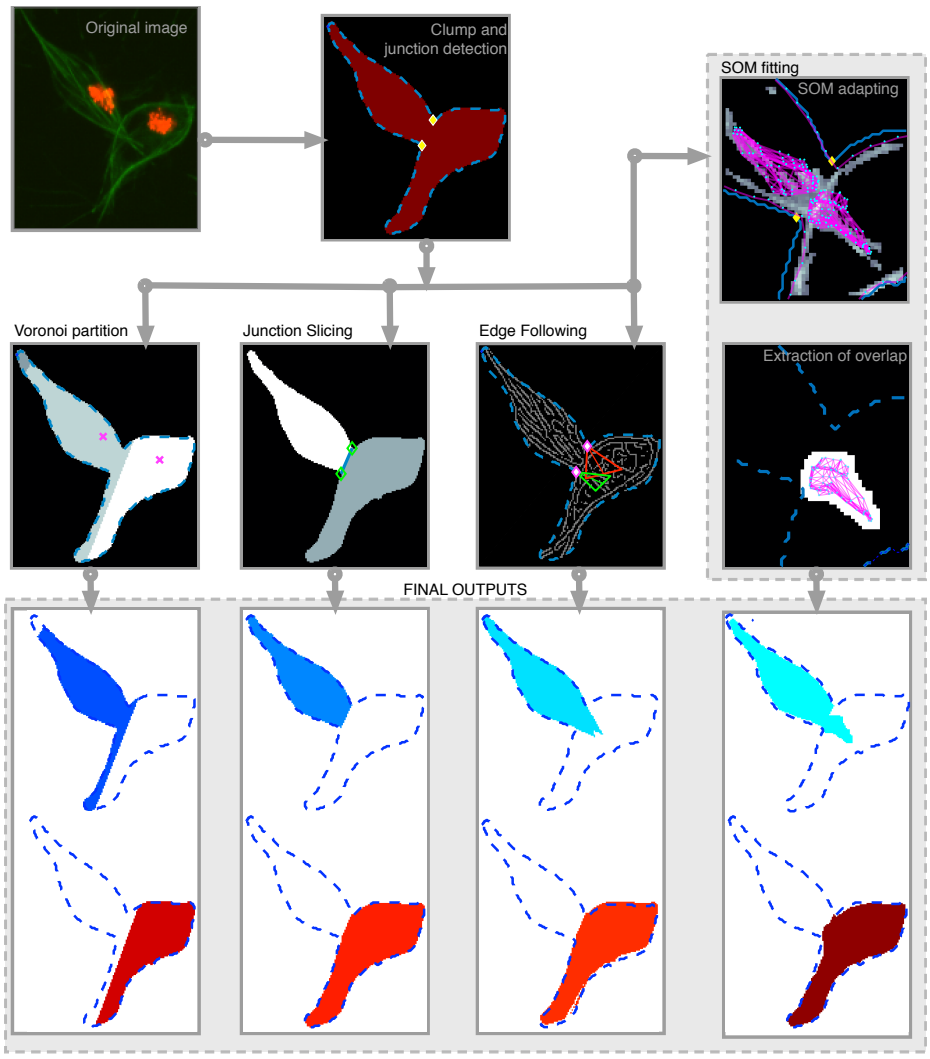


Fig. 5: Illustration of all the methods developed and the workflow to obtain results. Top left shows the detail of CLUMP 2 in the original frame. Clumps are detected and the boundary was extracted. With the boundary information, the *anglegram* was calculated and the junctions were detected (top, middle). On the second row, a diagram to the methods were presented. From left to right, the Voronoi partition, Junction Slicing (JS), Edge Following (EF) and SOM fitting. Bottom row shows the outputs from each method for both cells within the detected clump.

4 RESULTS

The junctions that were correctly detected on the synthetic data had a range of angles (based on the corresponding value in vector $\hat{\theta}_{\max}$) [188.64 - 328.4] degrees; whilst the missed junctions had a range of [162 - 191.96] degrees. This indicates that very wide angles, close to a straight line are easy to miss. The overlap region of [188 - 192] deserves a further investigation outside the scope of this paper. For overlapping clumps, junctions detected by the anglegram algorithm were compared qualitatively against the Harris corner detector [11], Fig. 6.

The results obtained for one frame of overlapping cells examined in detail are shown in Fig. 7, which are clumps that look similar throughout the images. Finally, a comparison with manually segmented ground truth was performed. In order to have the best results shown for each of the methods, the input clumps used were taken from the ground truth images. The Jaccard Similarity Index [17], recall and precision [18] statistics were computed for both clumps on all the frames and all the methods described, box plots of the results are shown in Fig. 8 and summarised in Table 1. Qualitative comparisons are also provided for some examples of the ten images and some of the clumps not analysed in detail, as well as for the SOM outputs segmenting the overlapped area in the clumps.

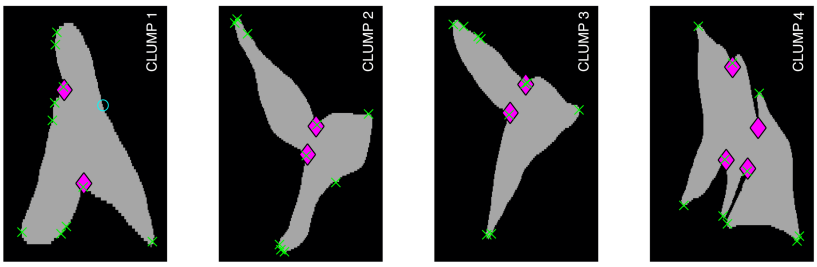


Fig. 6: Qualitative comparison of junction detection via anglegram (magenta ⋄) versus the Harris corner detector (green ×). The strongest 10 corners from the Harris detector per clump are displayed. Only CLUMP 1 has a missing junction (cyan ○), it should be noticed how difficult detection of the junction would be.

Table 1: Comparison of mean values of Precision, Recall and Jaccard Index for clumps 2 and 3 over 10 frames. This table summarises the results in Fig. 7. Highest results are highlighted.

	CLUMP 2			CLUMP 3		
	Precision	Recall	Jaccard Index	Precision	Recall	Jaccard Index
Voronoi	0.906	0.925	0.843	0.872	0.868	0.771
JS	**0.970**	0.953	0.926	**0.974**	**0.948**	**0.925**
EF	0.964	**0.983**	**0.948**	0.938	0.950	0.896
SOM	0.965	0.951	0.919	0.973	**0.948**	0.923

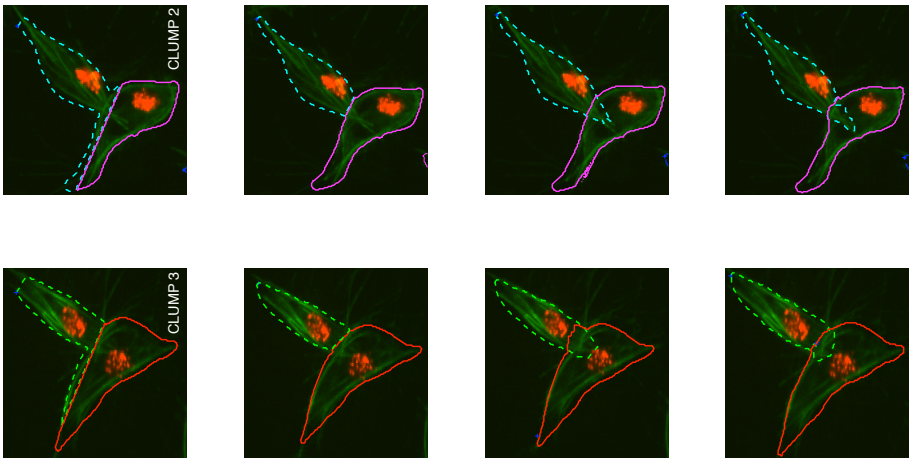
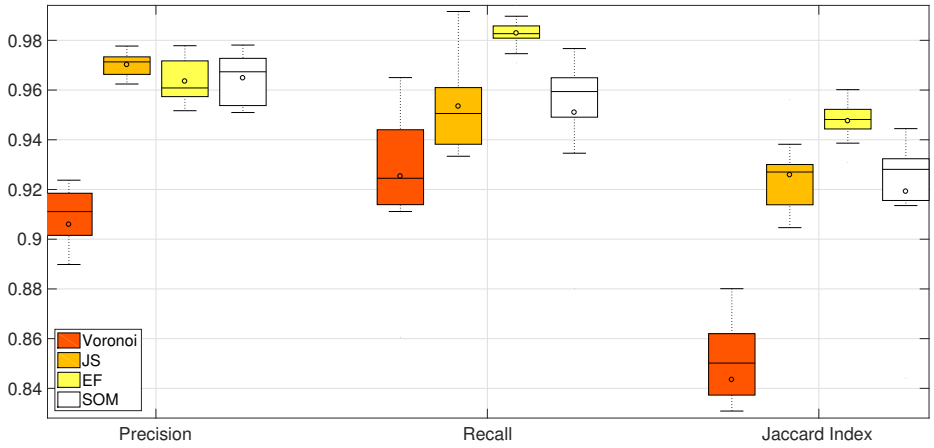


Fig. 7: Qualitative comparison of different segmentation methods in one of the frames. From left to right, the segmentation results for the Voronoi method, Junction Slicing (JS), Edge Following (EF) and SOM fitting are presented. Top and Bottom rows represent the results for CLUMP 2 and CLUMP 3 respectively.

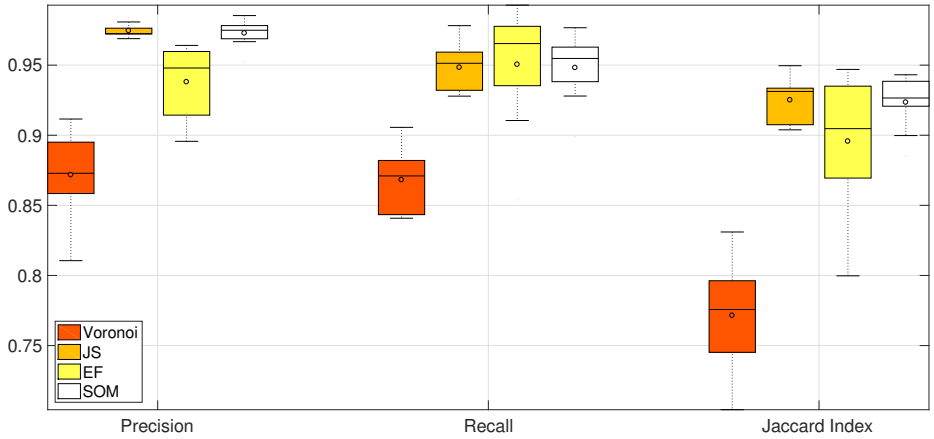
5 DISCUSSION

Preliminary work using thresholding techniques [6], active contours [19] or multilevel set methods [9] did not provide satisfactory segmentation of the overlapping cells, as these techniques could only detect clusters of overlapping objects without distinction between each of the individual objects (data not shown). In this paper, a method to segment overlapping cells through the analysis of the boundary of the clump was proposed. Its main advantage is to present a way to find relevant junctions from a boundary. Fig. 6 shows the junctions detected by the anglegram would not require further processing to select the useful corners, unlike the Harris algorithm outputs. Consistent with synthetic tests, limitations were observed in CLUMP 1, where a junction was missed by both methods. This limitation depends on cell positions and is transferred to the underlying segmentation methods. Table 1 shows a better performance from all three junction-based methods compared to the Voronoi partition. Furthermore, the percentile box sizes in Fig. 8 show that the EF method (yellow) is less consistent than the SOM method (white).

Current experimental results demonstrate the promise of this method to produce correct segmentations of overlapping cells, Fig. 7. Further experimentation is ongoing applying these techniques to cases where there are more than two main junctions detected, such as the one presented in Fig. 4(a) and cases where more than two cells are present in the clump. Future work will consider the extension of the anglegram matrix as a prior to a probabilistic modelling of the position of the junctions.



(a) CLUMP 2. Y-axis ranges from $(0.82, 1)$



(b) CLUMP 3. Y-axis ranges from $(0.7, 1)$

Fig. 8: Comparison of Precision, Recall and Jaccard Index for all methods of segmentation of overlapping in clumps 2 and 3. Horizontal axis correspond to the box plots from the different methods and their summarised performance in the metrics computed. Three groups corresponding to Precision, Recall and Jaccard Index contain four box plots; which, from left to right, correspond to Voronoi, JS, EF and SOM methods. Table 1 summarises the information on this image.

References

1. Martinez, F., Sica, A., Mantovani, A., Locati, M.: Macrophage activation and polarization. Front Biosci 13, 453—461 (2008)
2. Pocha, S.M., Montell, D.J.: Cellular and molecular mechanisms of single and collective cell migrations in drosophila: themes and variations. Annual review of genetics 48, 295–318 (2014)
3. Wood, W., Martin, P.: Macrophage functions in tissue patterning and disease: New insights from the fly. Developmental Cell 40(3), 221–233 (2017)
4. Stramer, B., Moreira, S., Millard, T., Evans, I., Huang, C.Y., Sabet, O., Milner, M., Dunn, G., Martin, P., Wood, W.: Clasp-mediated microtubule bundling regulates persistent motility and contact repulsion in Drosophila macrophages in vivo. J Cell Biology 189(4), 681–9 (2010)
5. Maška, M., Ulman, V., Svoboda, D., Matula, P., Matula, P., Ederra, C., Urbiola, A., España, T., Venkatesan, S., Balak, D.M., et al.: A benchmark for comparison of cell tracking algorithms. Bioinformatics 30(11), 1609–1617 (2014)
6. Henry, K., Pase, L., Ramos-Lopez, C.F., Lieschke, G.J., Renshaw, S.a., Reyes-Aldasoro, C.C.: PhagoSight: An Open-Source MATLAB® Package for the Analysis of Fluorescent Neutrophil and Macrophage Migration in a Zebrafish Model. PLoS ONE 8(8), e72636 (2013)
7. Dufour, A., Shinin, V., Tajbakhsh, S., Guillén-Aghion, N., Olivo-Marin, J.C., Zimmer, C.: Segmenting and tracking fluorescent cells in dynamic 3-d microscopy with coupled active surfaces. IEEE Trans Imag Proc 14(9), 1396–1410 (2005)
8. Plissiti, M.E., Nikou, C.: Overlapping cell nuclei segmentation using a spatially adaptive active physical model. IEEE Trans Imag Proc 21(11), 4568–4580 (2012)
9. Lu, Z., Carneiro, G., Bradley, A.P.: An improved joint optimization of multiple level set functions for the segmentation of overlapping cervical cells. IEEE Trans Imag Proc 24(4), 1261–1272 (2015)
10. Reyes-Aldasoro, C.C., Aldeco, A.L.: Image segmentation and compression using neural networks. In: Advances in Artificial Perception and Robotics, CIMAT. pp. 23–25 (2000)
11. Harris, C., Stephens, M.: A combined corner and edge detector. In: Proc 4th Alvey Vision Conf. p. pp. 147–151. Alvety Vision Club (1988)
12. Lindeberg, T.: Junction detection with automatic selection of detection scales and localization scales. In: ICIP. vol. 1, pp. 924–928 vol.1 (1994)
13. Okabe, A., Boots, B., Sugihara, K., Chiu, S.N., Kendall, D.G.: Spatial Tessellations Concepts and Applications of Voronoi Diagrams. Wiley (2008)
14. Canny, J.: A computational approach to edge detection. IEEE TPAMI 8(6), 679–698 (1986)
15. Kohonen, T.: The self-organizing map. Neurocomputing 21(1), 1–6 (1998)
16. Hannah, I., Patel, D., Davies, R.: The use of variance and entropic thresholding methods for image segmentation. Patt Rec 28(8), 1135–1143 (1995)
17. Jaccard, P.: Étude comparative de la distribution florale dans une portion des Alpes et des Jura. Bull Soc vaud Sci nat 37, 547–579 (1901)
18. Fawcett, T.: An introduction to roc analysis. Patt Rec Lett 27(8), 861–874 (2006)
19. Caselles, V., Kimmel, R., Sapiro, G.: Geodesic Active Contours. Int J Comp Vis 22(1), 61–79 (1997)