

DETAILS ABOUT THE SEGMENTATION ALGORITHM

The proposed algorithm is divided in two steps. In the first step individual cells are manually isolated using a custom-created mask. In the second step a thresholding of the image is performed in order to eliminate noise and atypical intensity values of the image of each cell. The information remaining after this step is related to the labeled organelles. Our algorithm automatically selects a threshold value for the image using its Shannon entropy [Shannon, 1951] to compute the cut percentile in a dynamic manner.

Threshold by entropy

From the point of view of digital image processing, the entropy corresponds to the average of the information contained in the image, and is defined as:

$$E(I) = - \sum_{x=0}^{2^B-1} p_x \log_2 p_x, \quad (\text{S1})$$

where B is the total number of bits of the digitized image I and $p(x) = \frac{K_x}{m \times n}$ is the probability of occurrence of a gray-level value, K_x is the number of times that the pixel with a value x is observed in I , m and n are the numbers of rows and columns of the image respectively, and as consequence $m \times n$ is the total number of pixels in the image. The minimum of sentropy is achieved when the image itself is constant (all pixels have the same value) and in this case $E(I) = \log_2 1 = 0$. The maximum entropy is reached when the pixel's values follow a uniform probability distribution function, this is, if B is the number of bits by pixel, then exists 2^B gray values in the image and the probability of each of them is exactly:

$$p(x_i) = \frac{1}{2^B}, \quad \forall i = 1, 2, \dots, 2^B. \quad (\text{S2})$$

If we consider the definition of entropy (S1) and the condition (S2), then:

$$\begin{aligned} E(I) &= - \sum_{x=0}^{2^B-1} p_x \log_2 p_x = - \sum_{x=0}^{2^B-1} \frac{1}{2^B} \log_2 \frac{1}{2^B} \\ &= - \log_2 \frac{1}{2^B} = - \log_2 2^{-B} = B \end{aligned} \quad (\text{S3})$$

The above information provide the limits for the entropy function of a digital image of B bits per pixel:

$$0 \leq E(I) \leq B \Rightarrow 0 \leq \frac{E(I)}{B} \leq 1 \quad (\text{S4})$$

Considering the entropy function as a measure of the information contained in an image, it is possible to interpret this value as an index that allows to compute the relationship between the signal and the noise in the image. For example, low entropy images, such as those containing a limited number of small biological structures, have very little contrast and a large number of pixels with the same or similar gray values (not much signal). On the other hand, high entropy images such as images with a large number of fluorescent biological objects, have a large contrast between adjacent pixels.

Definition 1 (Entropy Information Criterion (EIC)). *The entropy information criterion is defined as:*

$$EIC_I = 1 - \frac{E(I)}{B},$$

where $E(I)$ denote the entropy of the image I and B is the number of bits per pixels in I . \square

Note that $E(I)/B \in [0, 1]$ (see equation (S4)) and as consequence $EIC_I \in [0, 1]$. The scalar EIC_I could be interpreted as a measure of the percent of redundant (noise) information in the image.

Percentile Computation

For a serie of measurements $\{Y_1, Y_2, \dots, Y_N\}$ denote the data ordered in increasing order. The p -th percentile is a value $Y(p)$, such as $(100 \times p)\%$ of the measurements are lower than it, and at most $(100 \times (1 - p))\%$ are greater [Liu, 2007].

The p -th percentile is obtained by calculating the ordinal rank, and then taking from the ordered list the value that corresponds to that rank. The ordinal rank n is calculated using the formula:

$$n = \left\lceil N \times \frac{p}{100} \right\rceil,$$

where N are the total number of observations, $p \in [0, 1]$ is the percent of the data less than $Y(p)$. The operator $\lceil \cdot \rceil$ is the integer part, and is defined as:

$$\lceil x \rceil = \min \{k \in \mathbb{Z} | x \leq k\}.$$

The value from the ordered list that corresponds to the ordinal rank is the percentile, that is, $Y(p) = Y_n$.

The combination of the EIC (see Definition 1) and the computation of the percentile provide a clear and easy way to determine the threshold of any image.

Definition 2 (Entropy Threshold Criterion (ETC)). *Let I be an image with a colour depth of B bits per pixel, the Entropy Threshold Criterion of I is defined as:*

$$ETC_I = Y(EIC_I),$$

where EIC_I is the “Entropy Information Criterion” defined in Definition 1 and $Y(\cdot)$ is the EIC_I percentile of the image pixels values. \square

The ETC provides a fast, simple and efficient alternative to compute the threshold value for an image of any size and resolution. Also, it can be easily implemented in any programming language and allows to calculate in a dynamic way the threshold value for image sequences.

Applying this procedure to fluorescence imaging, we can automatically select in the image those pixels with the highest intensity values.

VALIDATION OF THE SEGMENTATION ALGORITHM

The algorithm validation was performed using synthetic “ground truth” images, generated taking into account diffraction, white noise, dark-current noise and signal amplification [Sinko et al., 2014, Garcés et al., 2016].

Generation of synthetic microscopy images

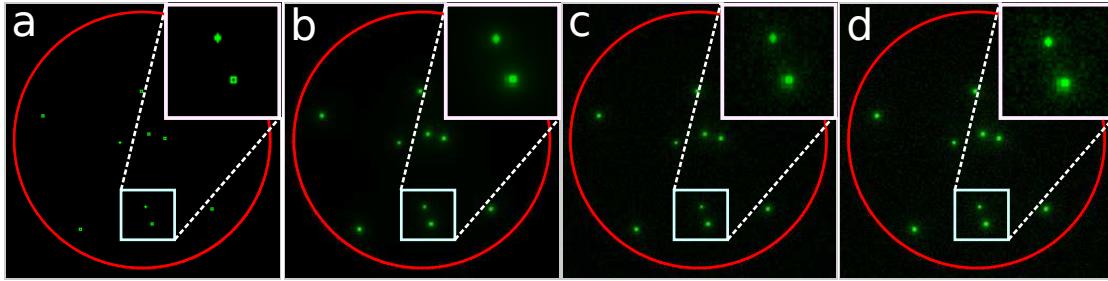
The synthetic images were generated using the approach proposed by Sinko and collaborators in [Sinko et al., 2014]. The next algorithm resume the steps for the synthetic images generation:

Algorithm 1: Synthetic Image Generation

- 1 Creation of image A containing fluorescent organelles (simulated as little ellipses) of biologically-relevant sizes and positions. The number of the organelles, the size and positions are generated randomly using a uniform distribution function.
 - 2 Convolute A using the point spread function (PSF) [Dougherty, 2005]. For this step we used the plug-in “Diffraction PSF-3D” of ImageJ.
 - 3 Add Gaussian noise in order to consider auto-fluorescence and Poisson noise for the purpose of considering electronic noise.
 - 4 Finally, the image is multiplied by $\sqrt{2}$ to consider amplification in a EM-CCD detector [Sinko et al., 2014].
-

In the step 2 of the Algorithm 1 we consider the following parameters for the image convolution: Index of refraction of the media $n = 1.33$, Numerical Aperture $NA = n \sin(\theta) = 1.3$, Wavelength $\lambda = 510\text{ nm}$, Image Pixel Spacing $px = 100\text{ nm}$, Width (pixels) = 512, Height (pixels) = 512. Rayleigh resolution $r = 0.61\lambda/NA = 240\text{ nm}$. For more details of this plug-in see [Dougherty, 2005].

In Figure S1 the steps of the Algorithm 1 for the generation of synthetic images is visually represented. The Figure S1(d) can be considered as an example of the images that was used to the validation of the proposed algorithm.



Supplementary Figure S1. Visual example of the synthetic images generation algorithm. **(a)** Generation of the organelles. **(b)** Convolution of the image (a) using the point spread function. **(c)** Image contamination with Gaussian and Poisson noise. **(d)** Image with the simulated amplification in a EM-CCD detector. The red circumference simulate the cell and only was included for a more accurate visual reference.

Signal/Noise ratios (SNR)

In the analysis of the performance of algorithm, it is important to take into account the relationship between the level of signal versus noise power. According to Gonzalez and Woods [Gonzalez and Woods., 2008], if we denote by $R(x, y)$ the image that only contains the signal and by $C(x, y)$ the corrupted image (containing both signal and noise), then, the signal/noise ratios (SNR) can be computed as:

$$SNR = 10 \log_{10} \left(\frac{mean(Signal^2)}{mean(Noise^2)} \right), \quad (S5)$$

where

$$mean(Signal^2) = \overline{R(x, y)^2} = \frac{\sum_{i=1}^m \sum_{j=1}^n [R(x_i, y_j)]^2}{m \times n} \quad (S6)$$

$$mean(Noise^2) = \overline{[R(x, y) - C(x, y)]^2} = \frac{\sum_{i=1}^m \sum_{j=1}^n [R(x_i, y_j) - C(x_i, y_j)]^2}{m \times n}. \quad (S7)$$

Then, replacing (S6) and (S7) in the equation (S5) we obtain:

$$SNR = 10 \log_{10} \left[\frac{\sum_{i=1}^m \sum_{j=1}^n [R(x_i, y_j)]^2}{\sum_{i=1}^m \sum_{j=1}^n [R(x_i, y_j) - C(x_i, y_j)]^2} \right]. \quad (S8)$$

The images $R(x, y)$ and $C(x, y)$ have a size of $m \times n$ pixels. The SNR is expressed in decibels (dB) and higher numbers correspond to better contrast, since there is more

useful information (the signal) than unwanted data (the noise). For example, when an image have a $SNR = 2dB$, it means that the signal is 2 times higher than the level of the noise. Note that with this definition of SNR it is possible to obtain negative values of SNR, this situation occurs when the image contains more noise than signal.

Accuracy of the segmentation algorithm

The images obtained as a result of the process contain fluorescing organelles of known sizes and positions. This facilitates the analysis of the algorithm's response against several plausible events, hence permitting comparison of the obtained results with "ground truth" organelles considering different signal/noise ratios.

Let $\mathbb{L}_I = \{\mathcal{L}_i, i = 1, 2, \dots, n\}$ and $\mathbb{L}_I^S = \{\mathcal{L}_j^S, j = 1, 2, \dots, k\}$ be the set of organelles that has been generated in the synthetic image I, and the set of organelles obtained as a result of the segmentation algorithm respectively. The organelles in both sets are uniquely identified by their centers (x, y) , and as consequence the sets \mathbb{L}_I and \mathbb{L}_I^S can be compared following the next criteria:

True Positive (TP): The organelle spot $\mathcal{L}_j^S \in \mathbb{L}_I^S$ is a true possitive if:

$$\exists \mathcal{L}_i \in \mathbb{L}_I \mid \sqrt{(x_{\mathcal{L}_i} - x_{\mathcal{L}_j^S})^2 + (y_{\mathcal{L}_i} - y_{\mathcal{L}_j^S})^2} \leq 0.5\mu m.$$

False Positive (FP): The organelle spot $\mathcal{L}_j^S \in \mathbb{L}_I^S$ is a false possitive if:

$$\nexists \mathcal{L}_i \in \mathbb{L}_I \mid \sqrt{(x_{\mathcal{L}_i} - x_{\mathcal{L}_j^S})^2 + (y_{\mathcal{L}_i} - y_{\mathcal{L}_j^S})^2} \leq 0.5\mu m.$$

False Negative (FN): The organelle spot $\mathcal{L}_j \in \mathbb{L}_I$ is a false negative if:

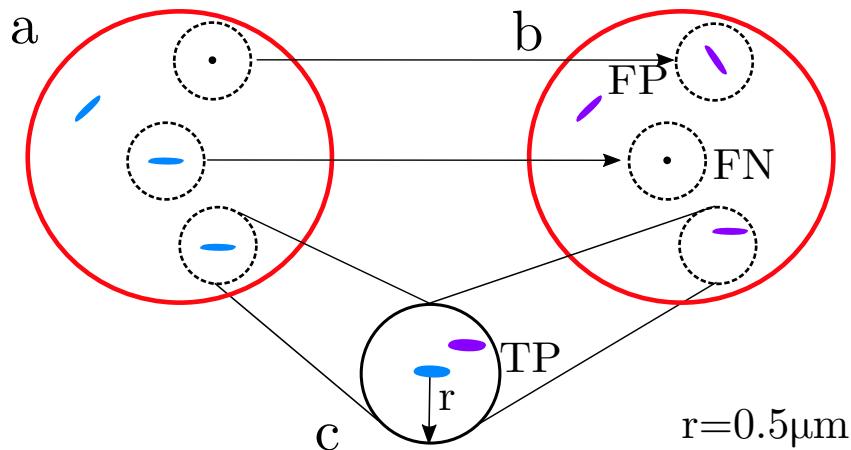
$$\nexists \mathcal{L}_j^S \in \mathbb{L}_I^S \mid \sqrt{(x_{\mathcal{L}_i} - x_{\mathcal{L}_j^S})^2 + (y_{\mathcal{L}_i} - y_{\mathcal{L}_j^S})^2} \leq 0.5\mu m.$$

In each of the previous definitions $(x_{\mathcal{L}_i}, y_{\mathcal{L}_i})$ and $(x_{\mathcal{L}_j^S}, y_{\mathcal{L}_j^S})$ denote the center of mass of the organelles $\mathcal{L}_i \in \mathbb{L}_I$ and $\mathcal{L}_j^S \in \mathbb{L}_I^S$ respectively. Note that in each case $\sqrt{(x_{\mathcal{L}} - x_{\mathcal{L}^S})^2 + (y_{\mathcal{L}} - y_{\mathcal{L}^S})^2}$ is the euclidean distance between the center of mass of the organelles \mathcal{L}_i and \mathcal{L}_j^S .

Similarly to the work of Garcés and collaborators [Garcés et al., 2016], the $0.5\mu m$ threshold was chosen based on the resolution limit of optical microscopes, which in agreement with Rayleigh's criteria, is $r = (0.61\lambda)/NA$, with λ being the emission

wavelength of the fluorophore and NA the numerical aperture of the objective. For the sake of simplicity, we define a $0.5\mu m$ threshold as an approximation to the diameter of the zero-order Airy ring, because in diffraction-limited images it does not make sense to segment structures smaller than $2r$.

The Figure S2 represents the criteria described previously. A TP is returned when an organelle is detected within a distance shorter than $0.5\mu m$ to that of a simulated spot. A FP result is returned when the algorithm detect an organelle that not exist in the “ground truth” image. A FN is obtained when a organelles in the “ground truth” images was not detected by the algorithm.



Supplementary Figure S2. Criteria of True Positive (TP), False Positive (FP) and False Negative (FN) criteria. **(a)** Representation of the “ground truth” image. **(b)** Result of the segmentation algorithm. **(c)** Zoom of a TP classification. The red circumference simulate the cell and all the dashed black circumferences inside the cell represent the ROI and have a radio of $0.5\mu m$.

The criteria presented above allowed to carry up the validation of the algorithm using indexes like the Jaccard index, the Recall and the Precision.

The Jaccard index [Pang-Ning et al., 2005] is a measure of similarity for two sets of data, with a range from 0 to 1. The higher the value, the more similar the two sets are, and as consequence, is very easy to interpret. If we consider the sets \mathbb{L}_I and \mathbb{L}_I^S , the Jaccard index is defined as:

$$J(\mathbb{L}_I, \mathbb{L}_I^S) = \frac{\mathbb{L}_I \cap \mathbb{L}_I^S}{\mathbb{L}_I \cup \mathbb{L}_I^S},$$

in our specific case, and considering the criteria presented before, we obtain that:

$$J(\mathbb{L}_I, \mathbb{L}_I^S) = \frac{\sum TP}{\sum TP + \sum FP + \sum FN},$$

where $\sum TP$, $\sum FP$ and $\sum FN$ indicate the total number of true positive, false positive and false negative respectively.

The Precision [Powers, 2011] (positive predictive value) is the fraction of organelles that were well detected (TP) among all the organelles found by the algorithm (TP+FP). It is defined as:

$$\text{Precision} = \frac{\sum TP}{\sum TP + \sum FP}.$$

The Recall (true positive rate or sensitivity) [Powers, 2011] in the context of our application is the fraction of organelles that have been detected by the algorithm over the real number of organelles in the "ground truth" image. It is defined as:

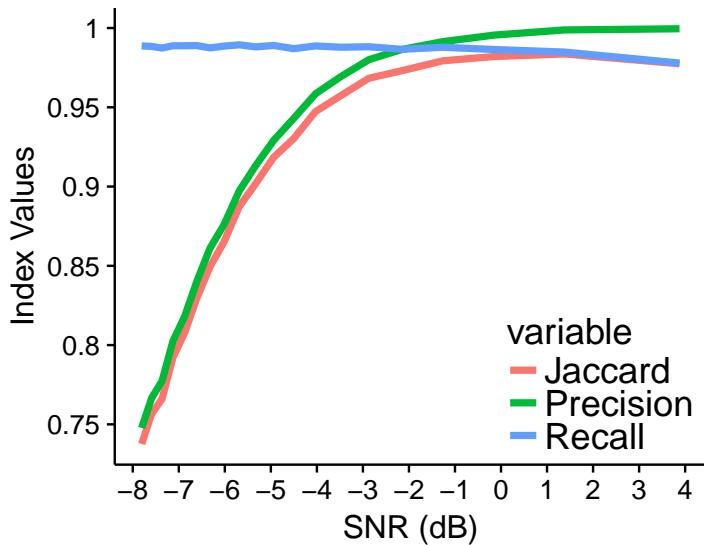
$$\text{Recall} = \frac{\sum TP}{\sum TP + \sum FN}.$$

The Precision and Recall take values in the interval $[0, 1]$. The higher values indicate a better accuracy of the algorithm.

The performance of the algorithm was evaluated over more than 20 000 simulated "ground truth" images generated taking into account different SNR values, number, size and position of the organelles. In the simulated images we included high levels of noise (see step 2 in the Algorithm 1) with the intention to observe the stability and performance of the algorithm in extremely complex scenarios.

Figure S3 shows the obtained results for the Precision, Recall and Jaccard indices as a function of the SNR. The curve of the Jaccard's index show that even for extremely noisy images ($SNR \leq 0dB$) the algorithm have a good performance with values, in all the cases, over 0.7 (over 70% match between the sets \mathbb{L}_I and \mathbb{L}_I^S). A very similar behavior is observed in the case of Precision, where the index value is directly proportional to the SNR, note that even when the $SNR = -8dB$, more than 70% of the organelles detected by the algorithm correspond to organelles in the "ground truth" images. The Recall have values over 0.97 for all the SNR, this means that, approximately 97% of the organelles in the "ground truth" images were detected by the algorithm for all the SNR. In general, the algorithm show a high robustness to noise, even in cases when noise is much higher than the signal in the images ($SNR \leq 0dB$). The values obtained show that our procedure offers a very good detection level of the organelles in the original images.

One of the advantages of the algorithm is the possibility to compute some properties like the area, the mean color intensity, the perimeter and the spatial position of each detected spot corresponding to an organelle. We use the information about the area of each "ground truth" organelle to study the error in the adjustment of each organelle, note that this information is relative to the performance of the algorithm to adjust the real size of a organelle.



Supplementary Figure S3. Plots of Precision, Recall and Jaccard indexes for the detection and adjustment of the organelles considering different levels of SNR in the simulated images.

Note 1. If $\mathcal{L}^S \in \mathbb{L}_I^S$ is a true positive detected organelle, then by definition (see the criteria above):

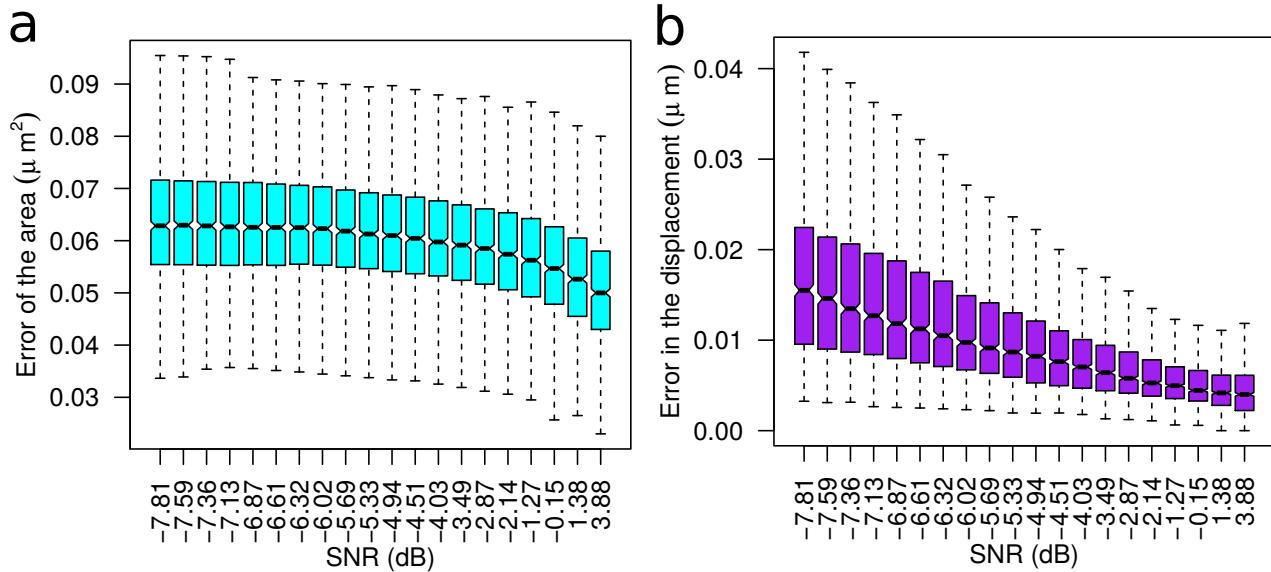
$$\exists \mathcal{L} \in \mathbb{L}_I \mid \sqrt{(x_{\mathcal{L}} - x_{\mathcal{L}^S})^2 + (y_{\mathcal{L}} - y_{\mathcal{L}^S})^2} \leq 0.5\mu m, \quad (\text{S9})$$

then it is possible to compare the area of the organelles \mathcal{L}_i and \mathcal{L}^S . The error in the area of the organelles \mathcal{L} and \mathcal{L}^S is defined as:

$$E_A(\mathcal{L}, \mathcal{L}^S) = |A_{\mathcal{L}} - A_{\mathcal{L}^S}|, \quad (\text{S10})$$

where $|\cdot|$ denote the absolute value. □

Figure S4(a) shows the boxplots for the error in the area adjustment of the organelles (see equation (S10)) in function of the SNR. Even for the images with the highest noise values ($SNR \approx -7.8dB$) the error in the adjustment is less than $0.1\mu m^2$, which is a very small error and constitute a clear evidence that our algorithm have a good performance in the adjustment of the organelles. As expected, the distribution functions of the area error decreases while the level of noise is less in the images, the error in the area approximation is lower (note that when the $SNR = 3.88dB$ the maximum error obtained in our simulation was around $0.08\mu m^2$). Figure S4(b) reveals a very small values for the error of the displacement between the center of the organelles. In this case the difference of the distribution functions between the lowest and highest values of SNR is more pronounced than in Figure S4(a), but for all SNR values the displacement is very small (less than $0.4\mu m$).



Supplementary Figure S4. Boxplot graphs for the error in the area and displacement to the center of the “ground truth” organelles. **(a)** Boxplot of the absolute error in the approximation of the organelles’s area (see Note 1). **(b)** Boxplot for the displacement error in the detection and fitting of the organelles. The displacement error was used to establish the criteria of detection and is defined as the Euclidean distance between the center of mass of the organelles \mathcal{L} and \mathcal{L}^S .

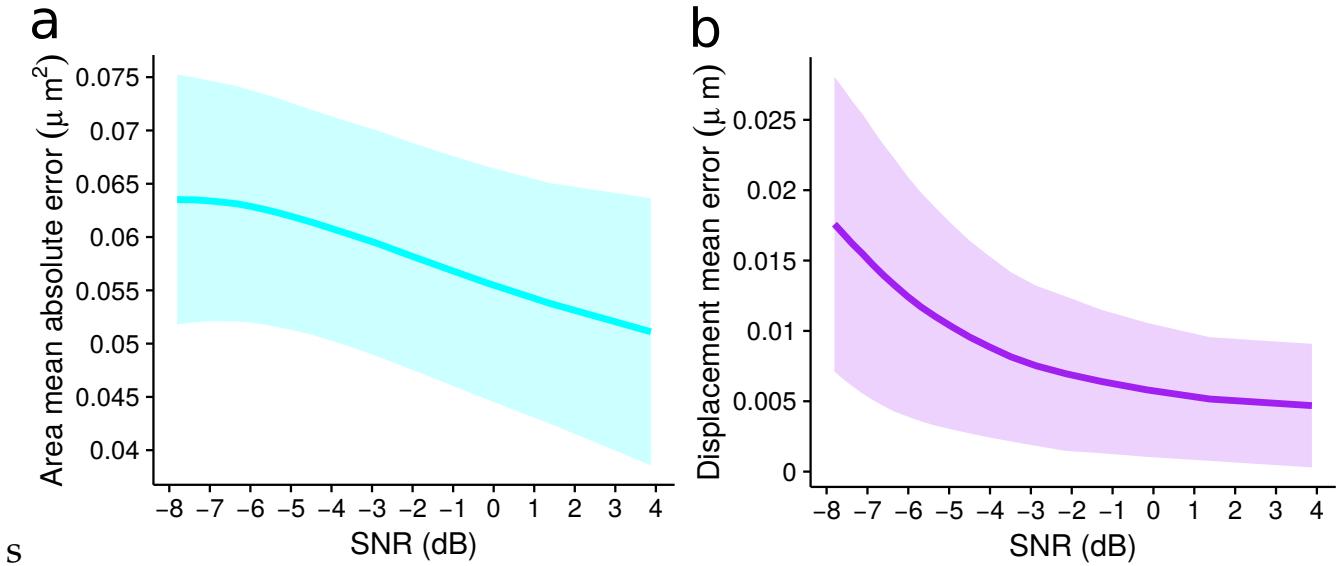
Figure S5 shows the mean values for the area error and for the displacement error as a function of the SNR. In both cases the mean error of the area, and the displacement, decrease when the noise of the image becomes lower. In the Figure S5(a) the upper limit of the confidence interval is always under $0.075\mu\text{m}^2$, and in the case of the Figure S5(b), the displacement error is less than $\approx 0.03\mu\text{m}$ for all the SNR values. These results reflect a very good performance of the algorithm even in images with high noise levels.

The validation of the proposed algorithm illustrates the advantages conferred by this new approach in terms of stability and high robustness to noise. Except for the manual creation of the masks to select individual cells (the first step of the algorithm), this method is automatic, deterministic, have a linear order computational complexity, and can be applied to images of any size and resolution. We consider that this algorithm provides a useful tool for the study of the organelles in each cell and their characterization considering some interesting variables like the area, perimeter, position and the spacial relationship between them.

STATISTICAL ANALYSIS

In our study we analyzed three conditions:

1. “Flow with medium (red and yellow-green emitting puncta)” (GR-Q).



Supplementary Figure S5. Mean error and confidence interval for the area and displacement of the organelles. **(a)** Mean absolute error for the area between the fitted and simulated organelle. **(b)** Mean of the euclidean distance between the centers of the adjusted and simulated organelle. The confidence interval was computed as a $mean \pm sd$, where sd is the standard deviation.

2. “Flow with medium +150 mM D-sorbitol (red and yellow-green emitting puncta) (GR-nQ).
3. “Flow with medium (red-emitting organelles) (R-nQ).

Each experiment consists of a hyperstack (x, y, t, c). The first two are spatial dimensions, and correspond to the x, y coordinates of fluorescent puncta within the image. The temporal dimension is denoted by $t = \{5, 10, 15, \dots, 600\}$ seconds, and c indicate the experimental condition $c = \{GR-Q, GR-nQ, R-nQ\}$. As consequence, for each condition we obtain the following data representation:

$$\begin{aligned}\hat{R}_{GR-Q} &= (\hat{R}_{GR-Q}^5, \hat{R}_{GR-Q}^{10}, \dots, \hat{R}_{GR-Q}^{600}), \\ \hat{R}_{GR-nQ} &= (\hat{R}_{GR-nQ}^5, \hat{R}_{GR-nQ}^{10}, \dots, \hat{R}_{GR-nQ}^{600}), \\ \hat{R}_{R-nQ} &= (\hat{R}_{R-nQ}^5, \hat{R}_{R-nQ}^{10}, \dots, \hat{R}_{R-nQ}^{600}),\end{aligned}$$

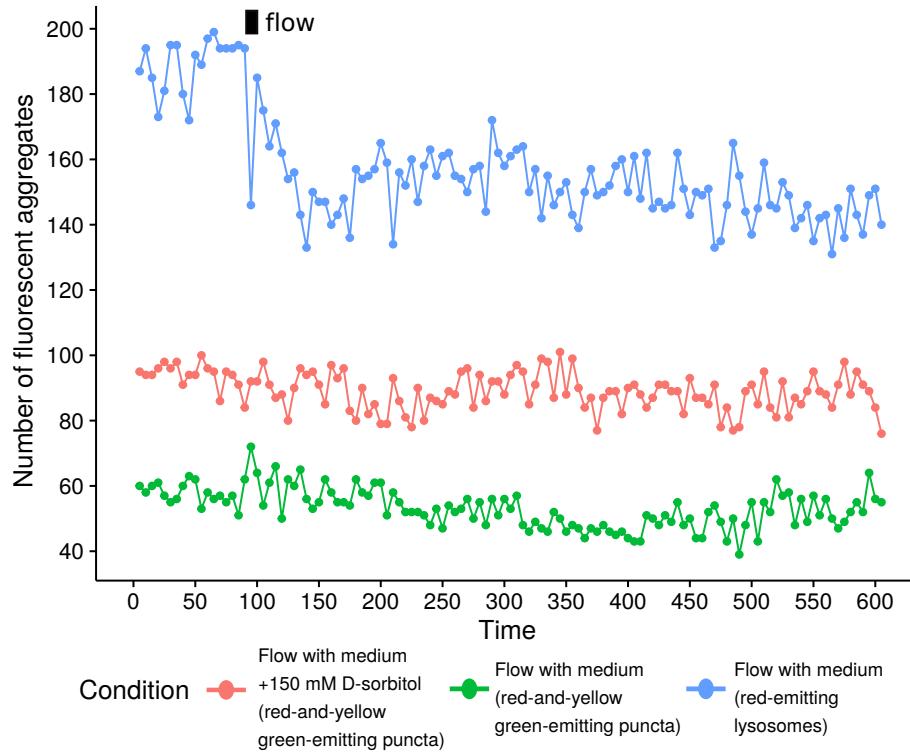
where the elements $\hat{R}_{\text{condition}}^t$ are vectors, in which each component corresponds to the Venus/mCherry ratio for an aggregate at a specific time $t \in \{5, 10, 15, \dots, 600\}$, this is:

$$\hat{R}_{\text{condition}}^t = (\hat{R}_{\text{condition}}^{1,t}, \hat{R}_{\text{condition}}^{2,t}, \dots, \hat{R}_{\text{condition}}^{N,t}) \quad (S11)$$

and

$$\hat{R}_{\text{condition}}^{j,t} = \frac{\text{Venus}_{\text{condition}}^{j,t}}{\text{mCherry}_{\text{condition}}^{j,t}}, \forall j = 1, 2, \dots, N. \quad (\text{S12})$$

N is the number of aggregates detected, which can vary depending on time and the experimental condition (see Figure S6).



Supplementary Figure S6. Number of cytosolic fluorescent aggregates as a function of time, determined with the proposed algorithm of image analysis.

The vectors $\hat{R}_{\text{GR-Q}}^t$, $\hat{R}_{\text{GR-nQ}}^t$, $\hat{R}_{\text{R-nQ}}^t$ can be compared at each time $t = \{5, 10, 15, \dots, 600\}$. The objective of this contrast is to determine whether exist a significant difference in the ratio $\text{Venus} / \text{mCherry}$ between two given experimental conditions in dependence of the time. Then, for each t , it is possible to perform the comparisons:

1. $\hat{R}_{\text{GR-Q}}^t$ vs $\hat{R}_{\text{GR-nQ}}^t$,
2. $\hat{R}_{\text{GR-Q}}^t$ vs $\hat{R}_{\text{R-nQ}}^t$,
3. $\hat{R}_{\text{GR-nQ}}^t$ vs $\hat{R}_{\text{R-nQ}}^t$.

We used a non-parametric test, because there is no evidence that the vectors $\hat{R}_{\text{GR-Q}}^t$, $\hat{R}_{\text{GR-nQ}}^t$, or $\hat{R}_{\text{R-nQ}}^t$ follow a Gaussian distribution or other known parametric distribution

(see section A). The observations in each group are independent, because each observation corresponds to one punctate structure of a specific condition. The variables $\hat{R}_{\text{GR-Q}}^t$, $\hat{R}_{\text{GR-nQ}}^t$, $\hat{R}_{\text{R-nQ}}^t$ are continuos (see equations S11, S12) and have different sizes. For these reasons we select the Mann-Whitney U test to carry up this analysis [Mann and Whitney, 1947]. The null and alternative hypotheses in the Mann-Whitney U test are given as:

$$H_0 : P(X > Y) = P(Y > X),$$

(The distributions of both populations are equal)

$$H_1 : P(X > Y) \neq P(Y > X),$$

(The distribution of both populations are different.)

This study correspond to the Figure 2C of the main document.

A: Normality Test

For each sample $\hat{R}_{\text{GR-Q}}^t$, $\hat{R}_{\text{GR-nQ}}^t$, and $\hat{R}_{\text{R-nQ}}^t$, ($t = (5, 10, 15, \dots, 600)$) we performed a Shapiro-Wilk test to analyze if these samples follow a Gaussian distribution [Shapiro and Wilk, 1965]. The null and alternative hypothesis for this test are:

H_0 : The population is normally distributed

$$(\hat{R}_{\text{Condition}}^t \sim N(\mu, \sigma^2))$$

H_1 : The population is not normally distributed.

$$(\hat{R}_{\text{Condition}}^t \not\sim N(\mu, \sigma^2))$$

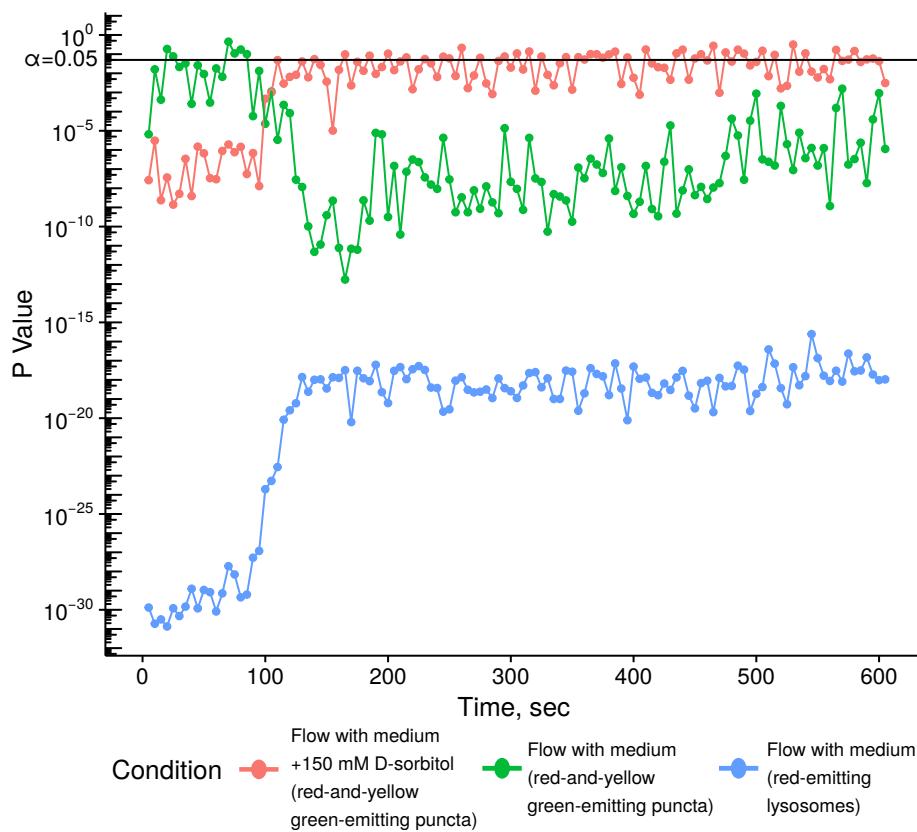
If the p-value (probability of obtaining a result equal to or “more extreme” than what was actually observed, when the null hypothesis is true) is less than $\alpha = 0.05$, then the null hypothesis is rejected and there is evidence that the data tested are not from a normally distributed population.

Figure S7 shows that most of the tests have a value of p less than $\alpha = 0.05$, then the null hypothesis is rejected, that is, the data does not follow a Gaussian distribution. In these type of situations, when there is no information about the distribution of samples, the best option for comparing two or more samples is using non-parametric tests.

REFERENCES

[Dougherty, 2005] Dougherty, B. (2005). Diffraction PSF 3D.

[Garcés et al., 2016] Garcés, Y., Guerrero, A., Hidalgo, P., López, R. E., Wood, C. D., Gonzalez, R. A., and Rendón-Mancha, J. M. (2016). Automatic detection and



Supplementary Figure S7. P-Value of the Shapiro-Wilk test of normality as a function of time.

measurement of viral replication compartments by ellipse adjustment. *Scientific Reports*, 6(36505).

[Gonzalez and Woods., 2008] Gonzalez, R. and Woods., R. (2008). *Digital Image Processing*. Prentice Hall, third edition edition.

[Liu, 2007] Liu, S. T. (2007). Springer handbook of engineering statistics. *Technometrics*, 49(4):494–494.

[Mann and Whitney, 1947] Mann, H. B. and Whitney, D. R. (1947). On a test of whether one of two random variables is stochastically larger than the other. *The Annals of Mathematical Statistics*, 18(1):50–60.

[Pang-Ning et al., 2005] Pang-Ning, T., Steinbach, M., and Kumar, V. (2005). *Introduction to Data Mining*.

[Powers, 2011] Powers, D. M. W. (2011). Evaluation: From precision, recall and f-measure to roc., informedness, markedness & correlation. *Journal of Machine Learning Technologies*, 2(1):37–63.

- [Shannon, 1951] Shannon, C. E. (1951). Prediction and entropy of printed english. *The Bell System Technical Journal*, 30(1):50–64.
- [Shapiro and Wilk, 1965] Shapiro, S. S. and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52(3/4):591–611.
- [Sinko et al., 2014] Sinko, J., Szabó, G., and Erdélyi., M. (2014). Ray tracing analysis of inclined illumination techniques. *Optics Express*, 22:18940–18948.