Detecting Differences of Fluorescent Markers Distribution in Single Cell Microscopy: Textural or Pointillist Feature Space?

Ali Ahmad[1,2], Carole Frindel[2] and David Rousseau[1]*

[1] *Laboratoire Angevin de Recherche en Ingénierie des Systèmes, UMR INRAE IRHS, Université d'Angers, Angers, France,*
[2] *Centre de Recherche en Acquisition et Traitement de l'Image pour la Santé, CNRS UMR 5220-INSERM U1206, Université Lyon 1, INSA de Lyon, Lyon, France*

We consider the detection of change in spatial distribution of fluorescent markers inside cells imaged by single cell microscopy. Such problems are important in bioimaging since the density of these markers can reflect the healthy or pathological state of cells, the spatial organization of DNA, or cell cycle stage. With the new super-resolved microscopes and associated microfluidic devices, bio-markers can be detected in single cells individually or collectively as a texture depending on the quality of the microscope impulse response. In this work, we propose, via numerical simulations, to address detection of changes in spatial density or in spatial clustering with an individual (pointillist) or collective (textural) approach by comparing their performances according to the size of the impulse response of the microscope. Pointillist approaches show good performances for small impulse response sizes only, while all textural approaches are found to overcome pointillist approaches with small as well as with large impulse response sizes. These results are validated with real fluorescence microscopy images with conventional resolution. This, a priori non-intuitive result in the perspective of the quest of super-resolution, demonstrates that, for difference detection tasks in single cell microscopy, super-resolved microscopes may not be mandatory and that lower cost, sub-resolved, microscopes can be sufficient.

Keywords: microscopy, fluorescence, classification, texture, spot detection, point spread function

1. INTRODUCTION

Over the last two decades, microscopy benefited from several scientific revolutions. For instance, innovations in chemistry via the production of new fluorescent markers, in optics with lasers tunable both in wavelength and impulse duration, or innovations in microfluidic bringing *in vitro* samples under the microscope automatically. These revolutions enabled the advent of intermediate to super-resolution microscopy techniques, such as lattice light sheet fluorescence microscopy (LLSFM), structured illumination microscopy (SIM), stimulated emission depletion microscopy (STED), or single molecule localization microscopy (PALM/STORM) techniques (Betzig et al., 2006; Rust et al., 2006; Schermelleh et al., 2010; Stelzer, 2015; Cremer et al., 2017). It is now possible to observe in 2D or 3D sub-cellular items inside single cells with resolutions which goes below the Rayleigh criterion for a classical microscope (Lakadamyali and Cosma, 2015; Ryabichko et al., 2017). These super-resolved systems are still not fully transferred in industrial applications or even

in microscopy platforms open to users that would not be expert in instrumentation. One reason for this translation delay is that super-resolution comes with the price of constraints in terms of micro-positioning which are more demanding as the size of the point spread function of the microscope is smaller. One way to relax such constraints consists in coupling the choice of the optical elements, i.e., designing the point spread function, jointly with the biological question raised and the associated image processing pipelines. We propose such an approach in this work.

In this article, we consider images of single cells observed with a microscope in which fluorescent markers have been activated (see **Figure 1**). We consider informative tasks that consist in detecting differences in the spatial organization of these fluorescent markers. Such differences could be either in terms of density or in terms of clustering. Detecting changes in spatial organization tasks are important issues in numerous biological contexts. For instance, distinct epigenetic states are associated with specific chromatin spatial modifications and compactions. Hence, defining the 3D-organization of cancer-associated chromatin domains could represent a new frontier to decipher tumor heterogeneity during tumor progression and metastasis formation (Kundu et al., 2017; Boettiger et al., 2016; Stevens et al., 2017). In another instance, the detection of nucleoids distribution changes is an important issue for the study of mitochondrial defect under various stresses. For example, disturbance in nucleoids components and mutations in mtDNA were identified as significant in various diseases, like carcinogenesis (Lee and Han, 2017) and neurodegenerative diseases (Chevrollier et al., 2012). These two use-cases focused on chromatin or on nucleoids of mitochondria are illustrated in **Figure 2**. The binary classification task here corresponds to a detection between healthy and unhealthy from the observation of the fluorescent markers inside individual cells. The use-cases illustrated in **Figure 2** were produced with a sub-resolved

microscope in panel A where markers appear as a texture and are not distinguishable from one another while it was produced with a super-resolved microscope in panel B where markers can all be located individually. However, to achieve the global characterization of a cell, it might not be necessary to locate individually each of these markers and thus, there is no guaranty that super-resolution is indeed mandatory. Such considerations are very important in practice because super-resolved microscopes are much more costly than sub-resolved microscopes in terms of optics, acquisition procedure, or numerical memory load. It would therefore be very useful to be able to determine a priori what would be the best resolution for a given task in order to choose the most appropriated microscope or design an optimal point spread function (PSF). In practice, for experimental optical acquisition, several acquisition conditions could be tested offering various PSF sizes. In simulation, a continuous set of PSF can be tested freely offering a complete view of how the cell detection would behave and enable to envision what would be the good range of PSF before real implementation.

Two main families of approaches are proposed in the literature to address the characterization of cells in sub-resolved microscopy (Kleppe et al., 2018; Paunovic et al., 2019) and super-resolved microscopy Griffié et al. (2016), Xu and Liu (2019). For the sub-resolved case, where cells are determined by more or less organized global patterns, one can use a textural approach. While for the super-resolved case, fluorescent markers distributions are classically studied with a pointillist approach. We propose to compare, for the first time to the best of our knowledge, the performances of algorithms based on the identification of each markers (pointillist approach) or the characterization of the texture created by these markers (textural approach). We compare the performances of these algorithms for different sizes of point spread function of a microscope and specially focus on the situations where the optical systems passes from
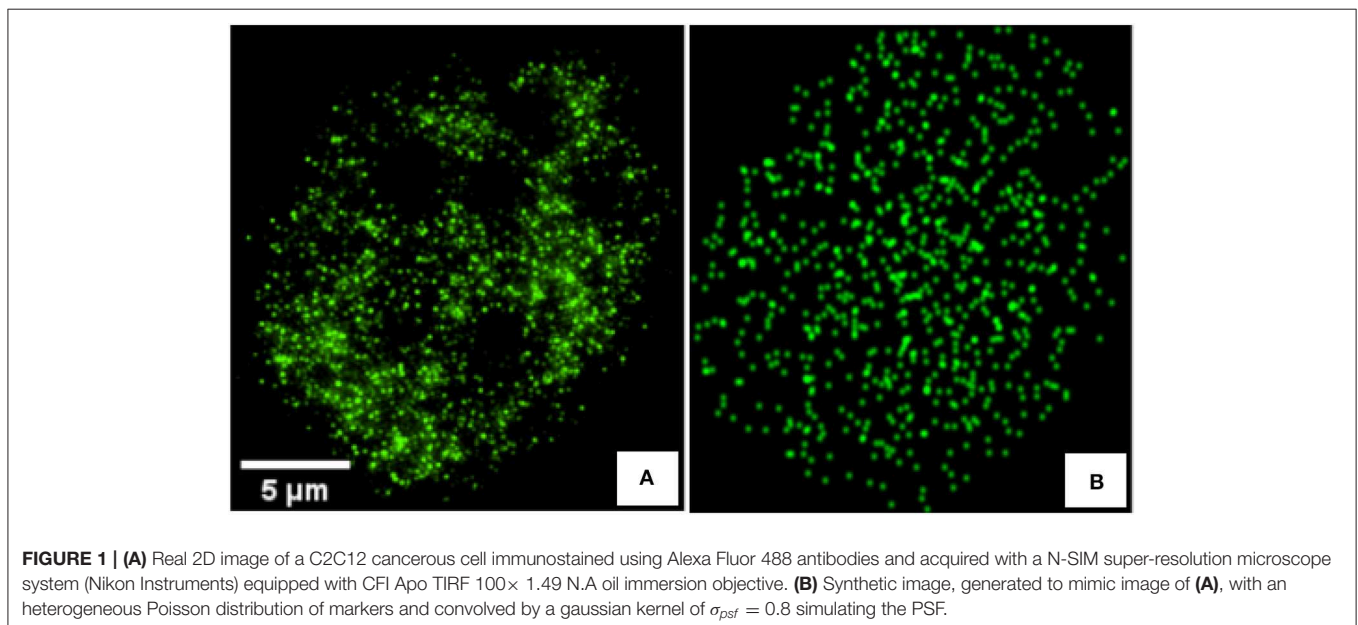

FIGURE 1 | (A) Real 2D image of a C2C12 cancerous cell immunostained using Alexa Fluor 488 antibodies and acquired with a N-SIM super-resolution microscope system (Nikon Instruments) equipped with CFI Apo TIRF 100× 1.49 N.A oil immersion objective. **(B)** Synthetic image, generated to mimic image of **(A)**, with an heterogeneous Poisson distribution of markers and convolved by a gaussian kernel of $\sigma_{psf} = 0.8$ simulating the PSF.

FIGURE 2 | Real 2D images of healthy and unhealthy nuclei of cells from different microscopy techniques. **(A)** Nuclei of breast cells images acquired by an Aurox Clarity structured illumination/spinning disk laser-free confocal microscopy system. The lens used was a 63× 1.4NA Oil objective, with for a 583 nm excitation and 631 nm emission wavelengths, and looking at mCherry labeled histone-H2B. **(B)** Nucleoids of skin fibroblast mitochondria stained with fluorescein isothiocyanate (FITC) and acquired with super-resolved stochastic optical reconstruction microscopy technique.

super-resolved to sub-resolved. Such an experiment would be very time consuming to be undertaken with real microscopes and can benefit from a simulation scheme as proposed here and also as current practice in the literature (Lehmussola et al., 2007; Rubin-Delanchy et al., 2015; Gazagnes et al., 2017; Samacoits et al., 2018; Ma et al., 2019). Simulated images are produced with the help of simulated point spread functions which realistically mimics real fluorescent images as shown in **Figure 1B**. An example of real images is also provided to validate the result obtained on simulation.
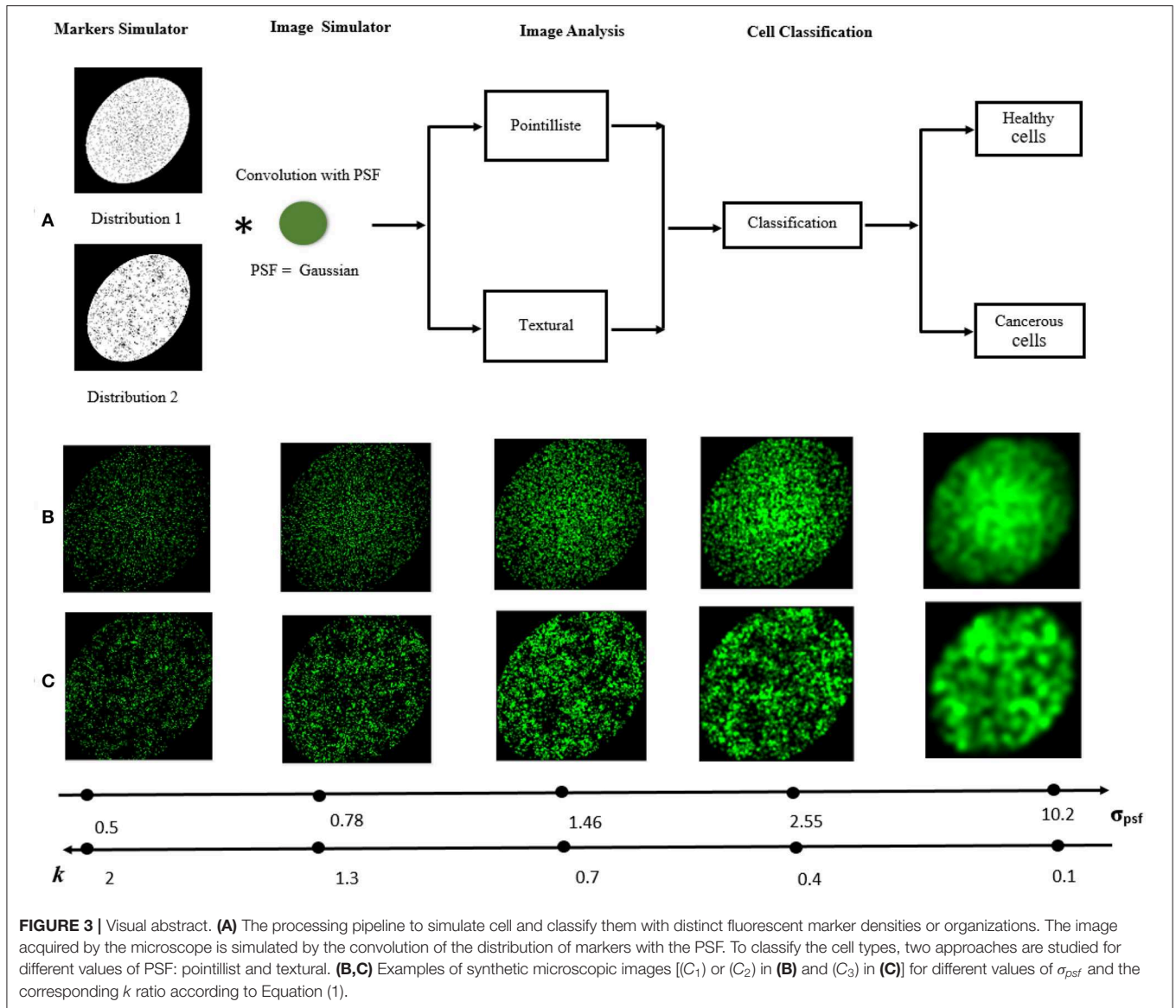
The article is organized in the following way. The process for the simulation of the images is first given. Then tools used for the characterization with textural or pointillist approaches are described. The comparison of these feature spaces reduced to the same dimension and applied to the same classifier is then produced before discussion and conclusion.

2. SIMULATING FLUORESCENCE MICROSCOPY IMAGES OF SINGLE CELLS

Two simulations were realized to investigate two distinct binary classifications with a difference in fluorescent markers density or a difference in spatial clustering of fluorescent markers. This corresponds to the practical situation of cases illustrated in **Figure 2**.

First, we generated two populations of cells (C_1) and (C_2) with a difference of marker density. The coordinates (x, y) of each fluorescent marker were picked randomly according to independent and identically distributed Gaussian distributions on horizontal and vertical dimensions of an image of $M \times N$ pixels, respectively, where $x \in \{1, 2, 3, ..., M = 256\}$ and $y \in \{1, 2, 3, ..., N = 256\}$. The two classes of cells, mimicking healthy (C_1) and pathological (C_2), were generated with a difference of standard deviation in their distributions (**Figures 3A,B**). The parameters were empirically adjusted to mimic observations on the real cells of **Figure 1A**, with for healthy cells (C_1) : $N_x(126, 100), N_y(126, 100)$, a total amount of markers of $3,000$, the area of the cell is $100 \times 100 = 10,000$ pixels, and a resulting density of markers $\frac{number\ of\ markers}{area\ of\ the\ cell} = 0.3$. For the pathological cells (C_2) we have $N_x(126, 100), N_y(126, 90)$, a total number of markers of $3,000$, an area of $100 \times 90 = 9,000$ pixels, and a resulting density of markers of 0.33. The difference of marker density between classes is of 0.03. We used here this difference between (C_1) and (C_2) to compare the pointillist and textural approaches.

Second, we generated populations of cells with same density of markers but with different spatial organization. We considered the task to detect the difference between cells with heterogeneous Poisson processes markers like (C_1) and cells with clustered markers that we coined (C_3). Clustered fluorescent markers

FIGURE 3 | Visual abstract. **(A)** The processing pipeline to simulate cell and classify them with distinct fluorescent marker densities or organizations. The image acquired by the microscope is simulated by the convolution of the distribution of markers with the PSF. To classify the cell types, two approaches are studied for different values of PSF: pointillist and textural. **(B,C)** Examples of synthetic microscopic images [(C_1) or (C_2) in **(B)** and (C_3) in **(C)**] for different values of σ_{psf} and the corresponding k ratio according to Equation (1).

were generated with a two steps process (i) 300 seeds were, independently from each other, distributed uniformly on horizontal and vertical space of the image then (ii) around each generated seed $s \in \{1, 2, 3, ..., 300\}$, 10 markers were generated with a distance to the seed which follows an exponential distribution $D_s\left(0, \lambda_D = \frac{1}{\mu_D}\right)$, where λ_D is the rate parameter of this distribution and μ_D is the mean of the distribution which was set to 35 (Diggle, 1983). The overall number of markers generated in each image was 3, 000. An instance of this cell class C_3 is shown in **Figures 3A,C**. Clustered (C_3) and heterogeneous Poisson processes markers (C_1) organization was used to compare the pointillist and textural approach. A crop inside the simulated cells is performed to avoid any issue concerning on the boundary of the simulated cells.

The PSF of the microscope was modeled by a convolving kernel here taken for illustration as a Gaussian kernel with size

σ_{psf}. The simulated images were then simply the convolution of the randomly and clustered positioned fluorescent dots with the PSF (see **Figure 3**). Each image was 256×256 pixels and the realism can visually be appreciated in **Figures 1A,B**. The performance of the classification between two types of cells ((C_1) (C_2)) or ((C_1) (C_3)) was investigated as a function of the size of the PSF governed by σ_{psf}. This situation corresponds to the practical use-case in instrumentation where ones seeks to design the PSF of a microscope for a given informative task. Here, the objective was to find the condition of PSF which enables to obtain the best binary classification performance.

The range of exploration of the size of the PSF σ_{psf}, was adimensioned by the distance between markers

$$k = \left(\frac{d_{minC_1C_2}}{\sigma_{psf}}\right) \tag{1}$$

where $d_{minC_1C_2}$ is the smallest value of the minimal distances between markers calculated for C_1 and C_2. This distance computed for the simulated data set is $d_{minC_1C_2} = 1.02$. We have explored the values of σ_{psf} around $k = 1$ which intuitively corresponds to the switch between the super-resolved regime ($k > 1$) to the sub-resolved ($k < 1$) as shown in **Figure 3C**.

As a complement to simulation, the same approach was applied to a real data set consisting of healthy and unhealthy (cancerous) breast single cells. These reals images shown in **Figure 2A** can be considered to be in the sub-resolved regime due to the estimated large experimental PSF of the microscope. The total amount of the real data set is 907 cancerous cell images and 1, 007 healthy cell images.

3. POINTILLIST FEATURE SPACES

A first step before studying fluorescent markers spatial distribution, is to localize them. In the literature, several algorithms for the localization of fluorescent markers have been developed (Holden et al., 2011; Ovesný et al., 2014; Gazagnes et al., 2017). For our study, we used UNLOC (*Unsupervised particle localization*), the state of the art method recently introduced in Mailfert et al. (2018).

UNLOC is a fast algorithm free of parameter that provides a list of coordinates and associated parameters for each detected particle for *a posteriori* quantification and image reconstruction. The algorithm is based on the decision theory without the need of initialization of any parameters relative to the data (SNR, particle density, background level). Only parameters relative to the optical setup must be provided like the PSF size (σ_{psf}) of the microscope to perform a PSF-deconvolution step. UNLOC has been shown in Mailfert et al. (2018) to approach the Cramér-Rao bound for the detection of particles in high density and without prior knowledge of their intensity. We applied UNLOC to the simulated images in sub-resolved and super-resolved regimes. Markers detection performance is presented in **Figure 4**. A uniform increase of detection performance for both randomly distributed markers with heterogeneous Poisson processes and clustered markers organization occurs for a $k = [0.5\ 1.3]$ which corresponds to a PSF size $\sigma_{psf} < 2$. A maximum performance of around 80% of detected markers occurs for $k = 1.3$ corresponding to $\sigma_{psf} = 0.8$ the range where UNLOC achieve a maximum performance of detection as found in Mailfert et al. (2018) where the minimum inter-marker distance is $> 1.23\sigma_{psf}$ ($d_{minC_1C_2} = 1.02 > 1.23\sigma_{psf}$).

After extracting the markers, it is necessary to characterize their spatial organization statistically. In the following section, we describe the proposed statistical descriptors computed in the pointillist approach and introduced to detect density and organization differences of fluorescent markers.

3.1. Distribution Analysis

This method was applied for marker density detection problem. We analyzed statistically the markers detected by UNLOC by computing the distance between markers, the distance of markers to the mean markers position and the distance between each
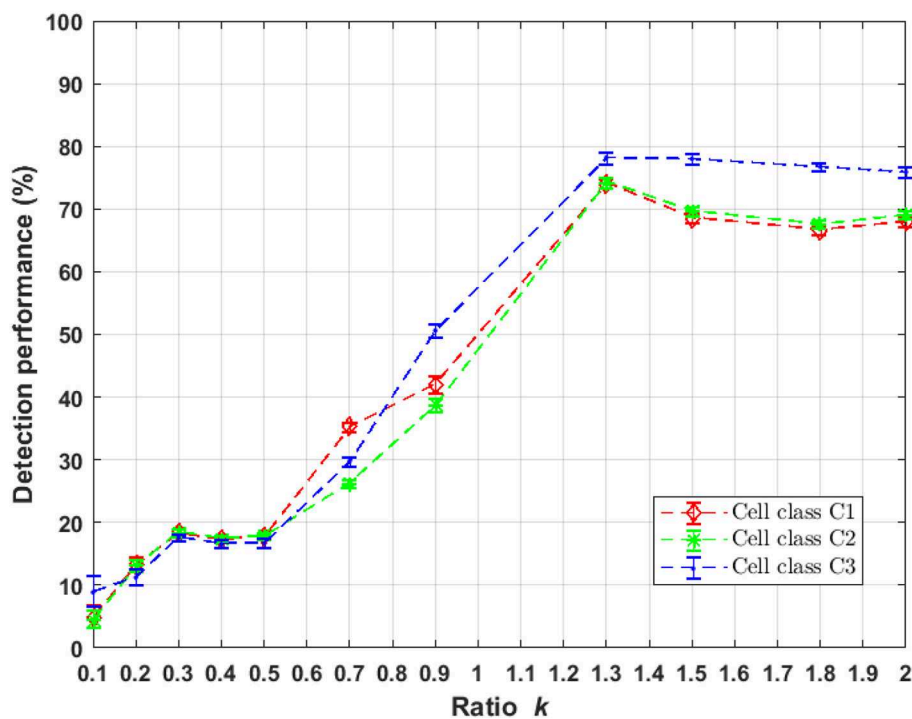

FIGURE 4 | Performance of markers localization using UNLOC in randomly and clustered distribution as a function of parameter k of Equation (1). UNLOC is more efficient in the range $k = [1.3\ 2]$.

marker and its nearest neighbor. From these distances, a large set of 16 common distributions of the literature were tested as proposed in Aminov (2019). As a tradeoff between the quality of the fit and the number of parameters in the distribution, the best distribution among all the one tested, was selected to minimize the bayesian inference criterion (*BIC*) (Schwarz, 1978; Neath and Cavanaugh, 2012) expressed as

$$BIC = -2ln(L) + p.ln(nb) \quad (2)$$

where L is the likelihood of the model, nb the number of observations in the sample and p the number of parameters in the model. In our study, the selected model for each calculated distance according to *BIC* was found to be Rayleigh $R(0, scale\ \sigma_R)$ for the distance between markers, exponential $exp(0, rate\ \lambda_E)$ for the distance to the mean and generalized extreme value $GEV(shape\ \xi_G, scale\ \sigma_G, mean\ \mu_G)$ for the distance of markers to the closest neighbor. Illustration of quality of these fits for the distribution parameters are given in **Figure 5**. These five statistical parameters $(\sigma_R, \lambda_E, \xi_G, \sigma_G, \mu_G)$ of the distance distributions were then used as features for the classification between cells.
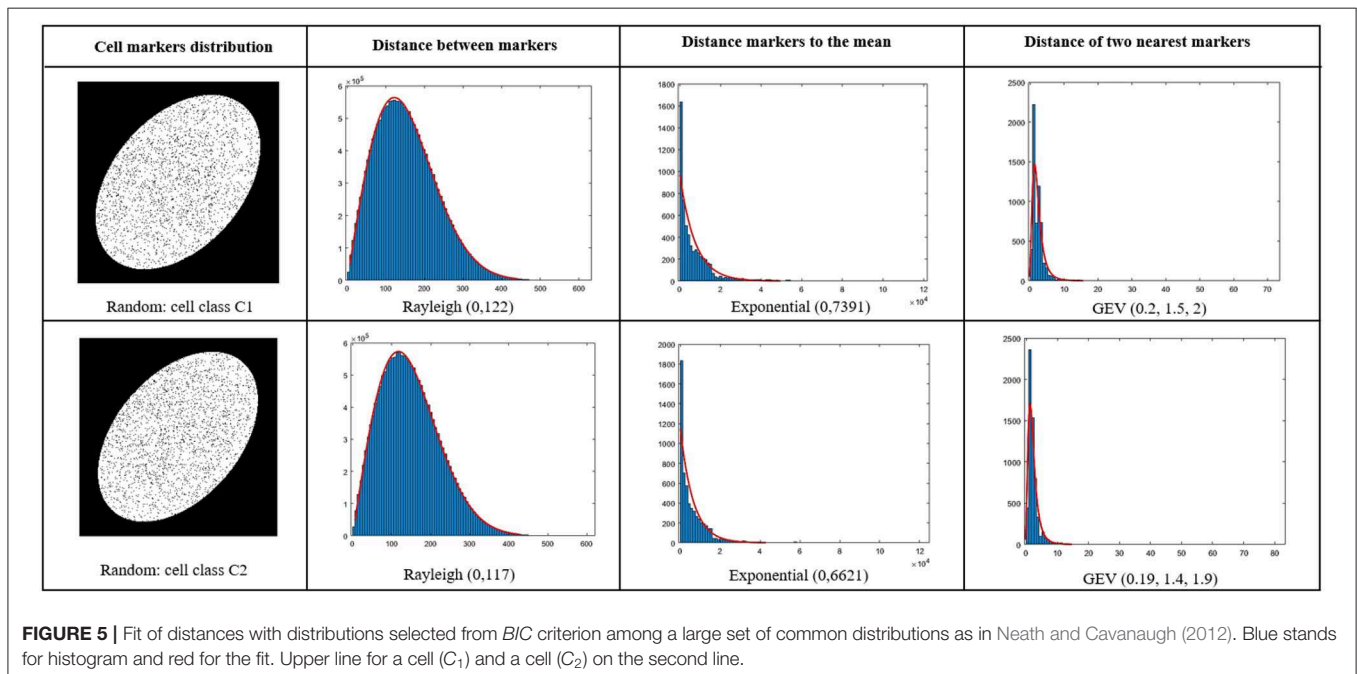
3.2. Spatial Pattern Analysis

In this work, we are interested in detecting the differences between classes of cells based on the change in spatial organization of fluorescent markers. For this aim, we work with Ripley's K-function that is classically used in microscopy (Hansson et al., 2013; Amgad et al., 2015; Samacoits et al., 2018) to summarize completely spatial randomness or clustering behavior of fluorescent markers and estimate markers organization parameters. But, in our work, we used this function as a descriptor to detect change between heterogeneous Poisson

processes and clustered distributed markers. By definition, Ripley's K-function is a spatial analysis to describe how point patterns occur over a given area of interest (circle of radius r) and whose standard expression is

$$K(r, n) = \frac{|\Omega|}{n(n-1)} \sum_{x \neq y} L_{\{|x-y| \leq r\}} f(x, y) \quad (3)$$

where n is the total number of events within the given field of view $|\Omega|$, $L_{\{|x-y| \leq r\}}$ is an indicator function equal to 1 if the distance between markers located in x and y is smaller than the radius r, and equal to 0 otherwise. $f(x, y)$ is a boundary correction term that prevents a bias in $K(r, n)$ at large values of r due to the finite size of $|\Omega|$. Indeed, some pairs of markers closer than r can fall outside the observation window $|\Omega|$, leading to an underestimation of K. Multiple edge correction methods have been devised for Ripley's K-function. The most widely used boundary correction is the Ripley's correction $f(x, y) = \frac{1}{2}(P(x, y) + P(y, x))$, where $P(x, y) = \frac{|\partial b(x, |x-y|)|}{|\partial b(x, |x-y|) \cap \Omega|}$. It consists of dividing the number of events at a certain distance from the central event by the proportion of the circumference of a circle $\partial b(x, |x-y|)$ that is included within the field of view $|\Omega|$ (Lagache et al., 2013). So with this boundary correction and under the hypothesis of completely random process, the expectation $\mathbb{E}[K(r, n)] = \pi r^2$ (Ripley, 1991, page 39). One problem with the original Ripely's K-function is that it is not centered and normalized which complicates its quantitative interpretation. In our work, we used the estimation version proposed by Besag (1977) given as

$$\hat{K}(r, n) = \sqrt{\frac{K(r, n)}{\pi}} - r. \quad (4)$$

FIGURE 5 | Fit of distances with distributions selected from *BIC* criterion among a large set of common distributions as in Neath and Cavanaugh (2012). Blue stands for histogram and red for the fit. Upper line for a cell (C_1) and a cell (C_2) on the second line.

In this work, we exploited $\hat{K}$ curve differences between the studied cases (see **Figure 6**) for the detection tasks between class of cells with markers density differences (C_1 and C_2) and spatial distribution differences (C_1 and C_3). We extracted from the $\hat{K}$ curves five features: the maximum $\hat{K}$ value, the maximum gradient [0, *max*], the minimum gradient [*max*, *end*], the radius corresponding to the maximum $\hat{K}$ value and Spearman correlation between $\hat{K}$ and the radius r similarly to what was proposed in Samacoits et al. (2018).

4. TEXTURAL FEATURE SPACES

In this section, we describe the textural methods used to detect a difference of fluorescent markers density or a difference of spatial clustering of fluorescent markers. There is a wide range of methods (Mirmehdi, 2008) and there is no proof of optimality for any tool. We do not intend to be exhaustive and pick up a selection of classical methods. For a fair comparison of all tested textural and pointillist methods, the feature spaces produced by each textural method was reduced to the dimension of the method with the smallest feature space with a principal component analysis.

4.1. Auto-Correlation

A standard tool to characterize the second-order statistics of a texture consists in analyzing the spatial Fourier transform of the auto-correlation of an image. This was for instance used to characterize the arrangement of fluorescent markers in microscopy (Kolin and Wiseman, 2007; Robertson and George, 2012). By definition, auto-correlation is equivalent to comparing all possible pixel pairs and reporting the likelihood that both will be bright as a function of the distance and direction of separation. Mathematically, auto-correlation is the convolution of a function with itself. For a microscopy image I of size $M \times N$, auto-correlation function $G(a, b)$ is calculated by the following equation

$$G(a, b) = \sum_{x}^{M} \sum_{y}^{N} i(x, y) * i(x - a, y - b) \quad (5)$$

where $i(x, y)$ is the image intensity at position (x, y) and a and b represent the shift from the initial position x and y. Practically, auto-correlations is calculated more efficiently and in a quicker way via fast Fourier transforms using the Wiener-Khinchin theorem stating that the auto correlation of an image is equal to the Fourier transform (F) of the power spectrum of this image. In our study, the used auto-correlation computation method was as the following

$$G(i) = F^{-1}[PS(i)] \quad (6)$$

where $PS(i) = |F[i(x, y)]|^2$ is the power spectrum of the image. The shape of the auto-correlation can be summarized with various features. In our study we computed five features for our
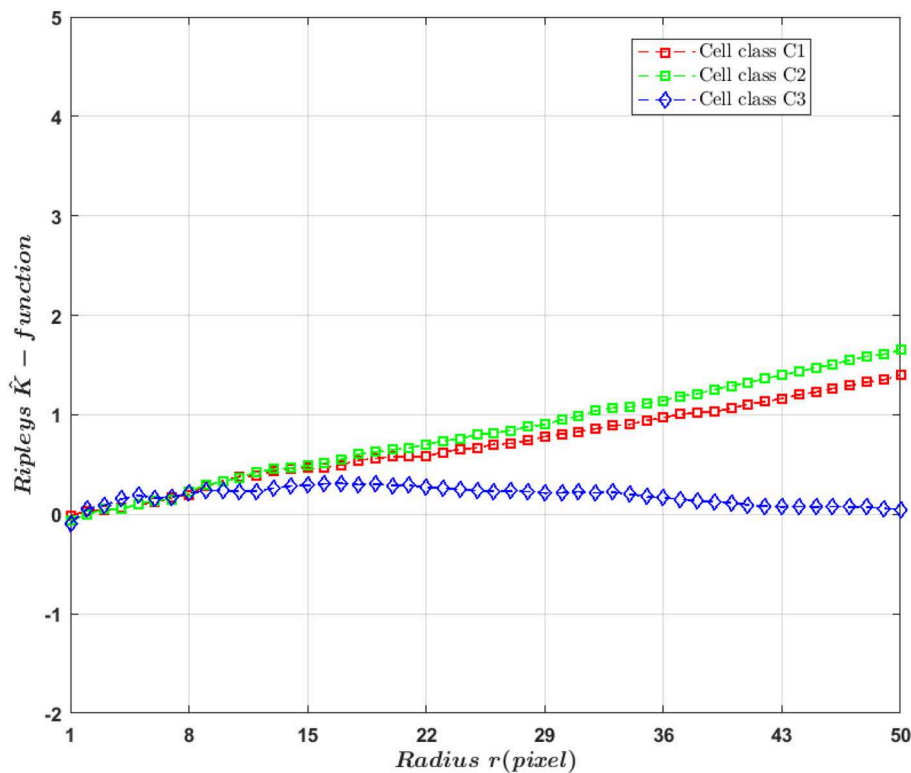

FIGURE 6 | $\hat{K}(r, n)$ curve for raw markers (i.e., non-convolved with a PSF) distribution according to radius r.

classification tasks: maximum auto-correlation value, full width at half maximum (FWHM), maximum and minimum gradient and the variance the remaining portion of the autocorrelation functions profile after removing the central peaks.

4.2. Gray Level Co-occurrence Matrix (GLCM)

Another classical statistical approach that can well-describe second-order statistics of a texture image is provided by the so-called gray level co-occurrence matrix (GLCM). GLCM was firstly introduced by Haralick et al. (1973) and is essentially a two-dimensional histogram in which the $(i, j)th$ element is the frequency of pixel intensity i co-occurring with pixel intensity j. A co-occurrence matrix is specified by the relative frequencies $C(i, j, d, \theta)$ in which two pixels, separated by a distance d, occurs in a direction specified by the angle θ, one with gray level i and the other with gray level j. A set of 14 Haralick coefficients summarizing the GLCM is then computed. In our study, since we expected no specific orientation a priori, and, as a trade off to respect this isotropy and limit the computation time, we included four directions for θ: 0, 45, 90, and 135°. The size of the neighborhood was chosen to be a multiple of the maximum value of σ_{psf}, and found optimal at 72×72 pixels (see **Supplementary Material** section 2). A principal component analysis was then applied to select only the five first significant components from the 14 Haralick coefficients. The five most significant features for the whole range of tested PSF size were found to be contrast, variance, sum variance, difference variance, and sum average.

4.3. Local Binary Patterns (LBP)

Local binary patterns are also among the most used texture descriptors in classification tasks (Ojala et al., 2002). In our study, the LBP was computed by dividing each original microscopy image in regions of 72×72 pixels similarly to the scale chosen with the GLCM. For each central pixel position coordinate (x, y) of these regions, local binary pattern (LBP) indicates a sequential set of the binary comparison of its value with the eight neighbors. So that the LBP assign to each neighbor the value 0 or 1, if its value is smaller or greater than the pixel placed at the center, respectively. The resulting decimal value of the generated binary number replaces the central pixel value and can be expressed as follows

$$LBP(x, y) = \sum_{n=0}^{7} 2^n b\left(i_n - i_{x,y}\right) \qquad (7)$$

where $i_{x,y}$ is the gray value of the central pixel and i_n denotes the nth neighboring one. Besides, the function $b(z)$ is defined as follows

$$b(z) = \begin{cases} 1, & \text{if } z \geq 0 \\ 0, & \text{if } z < 0 \, . \end{cases} \qquad (8)$$

The frequency of occurrences of each decimal code was then calculated over each region and used as a texture descriptor. A principal component analysis was finally applied to reduce the total number of descriptors per image to 5 as in the GLCM approach.

5. CLASSIFICATION

Classification tasks were addressed to discriminate between cells populations $((C_1)$ and $(C_2))$ with a small difference of fluorescence marker density and between cells $((C_1)$ and $(C_3))$ with different spatial marker organizations. For fair comparison all features spaces either pointillist and textural was set to 5. These features spaces were applied to the same simple support vector machine with linear kernel. Comparison with other classical classifiers (decision tree, logistic regression classifier, and K-nearest neighbors) are also provided in the **Supplementary Material**. The classification performance was tested for 11 different values of σ_{psf} ranging from sub-resolved to super-resolved regimes. For each value of σ_{psf} the simulated data set was composed of $4,000$ images for training, with $2,000$ for each class, and 500 images for test (respectively 250 images for each class). Classification was also performed on the real data set of **Figure 2A**. Standard deviation of performances were computed using 10-folds cross-validation method.

6. RESULTS

6.1. Difference of Density

Classification performances between populations of cells with a small fluorescent marker density differences $((C_1),(C_2))$ as a function of σ_{psf} are presented in **Figure 7**. The performance of the textural approach overpasses everywhere the pointillist ones either in sub-resolved and also super-resolved regimes. Among the textural features spaces, auto-correlation textural approach shows the most stable classification performance for all PSF sizes. Distance distribution among the pointillist approaches shows good performances for the detection of difference in the super-resolved regime only. Since we did not test all existing methods for the pointillist and the textural approaches, comparison is not exhaustive. Nevertheless, one should here recall and underline the specific choices for the tested methods (Haralick coefficient, Local binary pattern and Auto-correlation) which all constitute very basic methods for the textural approach. Therefore, other textural approaches could surely provide even better results while the UNLOC method for dense detection of fluorescence has been shown to be the current reference for the state of the art (Mailfert et al., 2018). We can thus conclude that, globally, for the considered classification task, feature spaces based on textural approaches outperform a pointillist-based feature space and this in both sub-resolved and super-resolved regimes.

It should be noticed that there is not only a difference in density between (C_1) and (C_2) but also in spatial organization because of the different standard deviations used for the simulation. This is why a differences between cell classes C_1 and (C_2) was found in Ripely's K-function (see **Figure 6**). If the only difference between classes were in the density, then theoretically they would be no difference in K-function. It is indeed a basic design property of this descriptor to capture only second-order characteristics of a point process, and to be invariant to changes in density. A variant of the simulation of classes (C_1) and (C_2) with differences only based on density without the diffraction effect of the microscope were conducted in the **Supplementary Material**. As expected the detection of
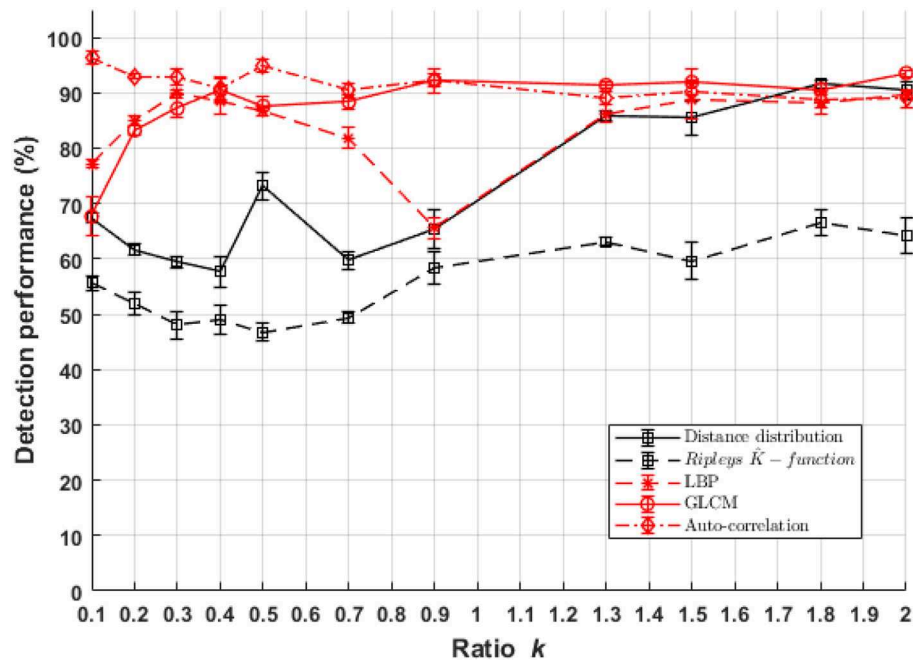
FIGURE 7 | Performance of binary classification for marker density difference between cell classes C_1 and C_2 as a function of parameter k of Equation (1) for both textural and pointillist approaches. The higher k the smaller the PSF size. Standard deviation of performances are computed using 10-folds cross-validation method. In red, methods following the textural approach, and in black, methods following a pointillist approach.

difference of density with Ripley's k- function is pointless for the super-resolved cases. However, due to the convolution process and to the instability of the UNLOCK detection which strongly depends on σ_{psf} some discriminant effect can occur between cell classes (C_1) and (C_2) in the sub-resolved cases. The simulation details and results for this experiment are described in **Supplementary Material**, section 3.

6.2. Difference of Spatial Organization

Classification performances between populations of cells with a difference of fluorescent marker spatial distribution $((C_1),(C_3))$ as a function of σ_{psf} are presented in **Figure 8**. The performance of the textural and pointillist classification approaches are found to be very high and very close to each other almost everywhere in terms of σ_{psf} and, remarkably in the sub-resolved regimes. The sub-resolved regime, not surprisingly, is the place where the performance of the pointillist approach drops. Some markers are detected as artifacts but the spatial organization of the two populations of cells becomes very close so that the discrimination between them drops. This is illustrated in **Figure 9** with the global evolution of the Ripley's K-functions when plotted for various σ_{psf}.

6.3. Test on Real Data

Classification performances between populations of healthy and cancerous cells of real data set from sub-resolved microscopy technique of **Figure 2A** are presented in **Table 1**. As a global comparison with the results of synthetic data, performances behaviors are globally similar. Auto correlation

approach of textural feature space shows the best classification performance with an important gap with other proposed textural and pointillist feature spaces. By contrast with what was found in simulation, LBP performance is less than the GLCM performance. This may come from the fact that LBP is sensitive to the noise, such as the thermal noise of the camera. Such noise was not taken into account in the simulation. Other derivative of LBP were proposed in the literature like median binary pattern (MBP) (Hafiane et al., 2007), local ternary patterns (LTP) (Tan and Triggs, 2010), and improved LTP (ILTP) (Kylberg and Sintorn, 2013). These could be tested to circumvent this noise sensitivity problem. Another important point to assess the value of a feature space lay in its computational time. They are presented in **Table 2**. A tradeoff between performances and computational time is found with the correlation based textural feature space.

7. CONCLUSION

In this study, we have simulated two classes of images of cells (healthy and pathological) with fluorescent markers having either a weak difference of density or a difference of spatial organization. We have then simulated different size of microscope PSF around the switch of regime between super-resolved and sub-resolved of the markers. These synthetic data sets served to compare the detection performance both with a textural and a pointillist approaches. We found
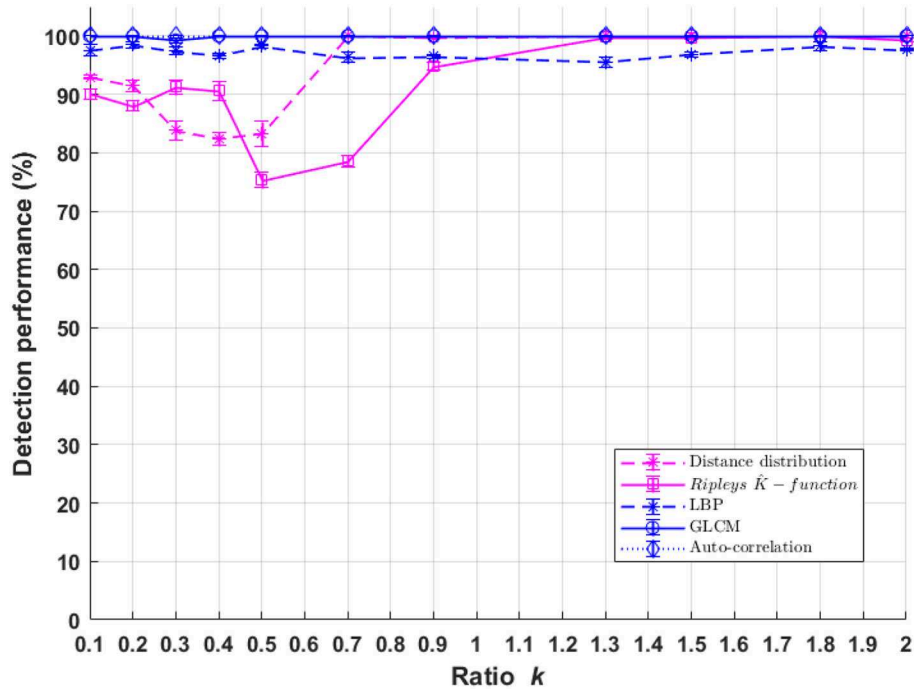
FIGURE 8 | Performance of binary classification for markers spatial organization difference for cell classes C_1 and C_3 as a function of parameter k of Equation (1) for both textural and pointillist approach. The higher k the smaller the PSF size. Standard deviation of performances are computed using 10-folds cross-validation method. In blue, methods following the textural approach, and in pink, methods following a pointillist approach.
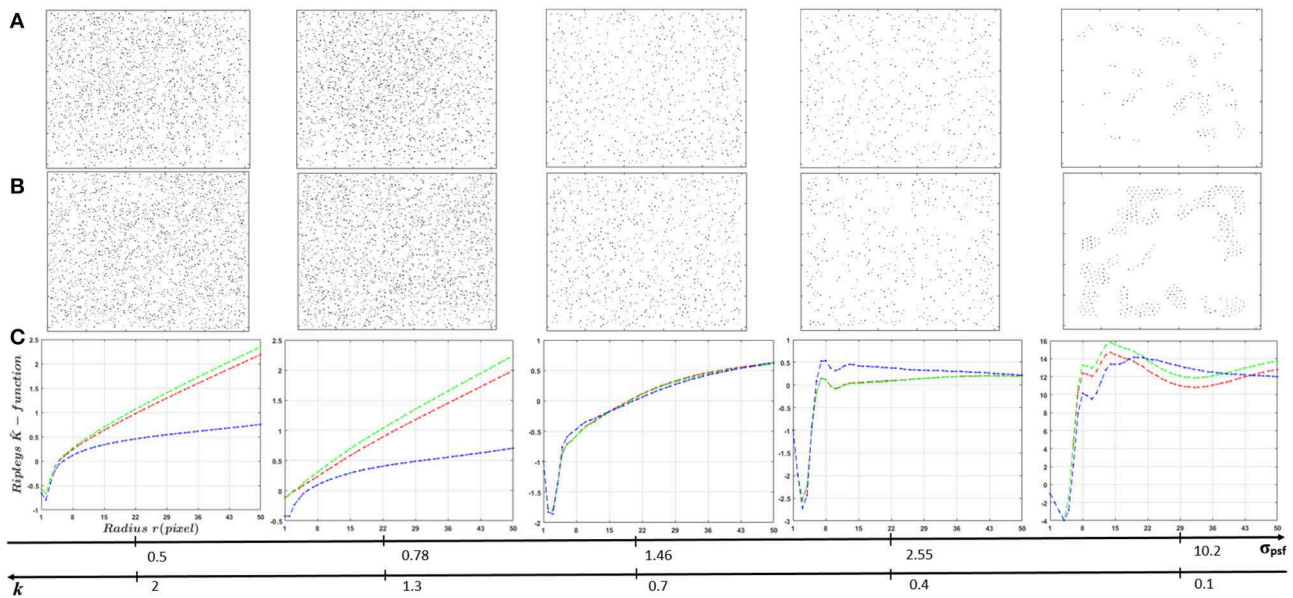

FIGURE 9 | Influence of PSF size (ratio k) on Ripley's K-functions $\hat{K}(r, n)$ of localized markers using UNLOC. **(A)** Visualization of localized markers in random distribution case. **(B)** Visualization of localized markers in clustered distribution case. **(C)** $\hat{K}(r, n)$ curves. Spatial organization of fluorescent markers changes according to the PSF size. Globally, for small PSF sizes (super-resolution images), markers organizations shows a similar organization as the raw data (**Figure 6**). Switching to sub-resolution images (large PSF size), spatial organization of detected markers changes due to miss detection caused by UNLOC sensitivity drop.

that the textural approach reaches better performances in all regimes sub-resolved as well as super-resolved. We also tested a real data set acquired with sub resolved microscopy. In accordance with the result on synthetic data, results from this real data set showed that the classification performance when using auto correlation textural approach

TABLE 1 | Classification results of the studied methods applied to real data set of **Figure 2A**.

Feature spaces	Real data
Auto-correlation	94.5 ± 0.18 %
GLCM	70 ± 1.4 %
LBP	64 ± 1.4 %
Distance distribution	78.8 ± 0.7 %
Ripley's $\hat{K}$-function	69 ± 1.7 %

TABLE 2 | Features space computational time averaged on 100 real images.

Feature spaces	Auto-correlation	GLCM	LBP	Distance distribution	Ripley's $\hat{K}$-function
Computation time (ms)	0.66	0.12	6.3	340	68, 790

overcome GLCM and LBP textural approaches as well as pointillist approaches.

Somehow counter-intuitive when considering the current quest for super-resolution, this strong and practically important results demonstrates that it may not be necessary to systematically search for expansive super-resolution techniques or perform time-consuming deconvolution when gazing at collective spatial organization of fluorescent markers in single-cell microscopy. This result is more common in signal processing. Indeed, an analogy is found in kernel-based density estimation methods where the kernel (the PSF in the case studied here) spread the information contained in discrete points to a larger area and thus contribute to create its continuous representation. This representation is easier to handle than the discrete one. Back in microscopy, when detection or classification is targeted sub-resolved images can carry sufficient information to enable high performances. This was obtained here with simulated images where the ground truth was established automatically and with one real data set. This exactly corresponds to the situation where cells can be sorted automatically based on a biological experimental plan or using standard flow cytometry. If such ground truth cannot be established, it might be the case that, similarly to what happens in histology, only a visual human inspection of the cell can enable to constitute a reference on which supervised learning can be trained. In this case, even for classification tasks, super-resolution coupled with sub-resolution may be necessary. However, if such pairs are constituted during the training, then only sub-resolution images can be used for classification during the testing as shown in this work.

REFERENCES

Amgad, M., Itoh, A., and Tsui, M. M. K. (2015). Extending Ripley's k-function to quantify aggregation in 2-d grayscale images. *PLoS ONE* 10:e0144404. doi: 10.1371/journal.pone.0144404

Aminov, Y. (2019). *"Find the Best Distribution" Tool (FBD)*. Cambridge: MATLAB Central File Exchange.

Further investigations could be undertaken in at least two directions. First, in this article the PSF of the microscope was purposely naïve under the form of a simple Gaussian 2D kernel. The proposed methodology could easily be translated without any difficulty to any type of more realistic PSF and can even be extended in 3D. The global methodology could thus be used in instrumentation to validate the quality of a PSF for a given informative task. The realism of the simulator could also be enhanced to account for non-spatial invariance of the PSF due to the sample (Cuplov et al., 2014) or the non-uniformity of fluorescence intensity of the markers. Second, only binary classification tasks were considered in this article and it could be interesting to consider if other informative tasks, such as regression could benefit from the proposed approach.

DATA AVAILABILITY STATEMENT

The simulator developed for this article will be available on request after acceptance of this article.

AUTHOR CONTRIBUTIONS

The experiment conceptualized, curation, formal analysis, funding acquisition, investigation, methodology, project administration, resources, software, supervision, validation, visualization, writing review, and editing have been performed equally by AA, CF, and DR.

FUNDING

ACKNOWLEDGMENTS

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/frobt.2020.00039/full#supplementary-material

Besag, J. (1977). Comments on Ripley's paper. *J R. Stat. Soc.* 39, 193–195. doi: 10.1111/j.2517-6161.1977.tb01607.x

Betzig, E., Patterson, G. H., Sougrat, R., Lindwasser, O. W., Olenych, S., Bonifacino, J. S., et al. (2006). Imaging intracellular fluorescent proteins at nanometer resolution. *Science* 313, 1642–1645. doi: 10.1126/science.1127344

Boettiger, A. N., Bintu, B., Moffitt, J. R., Wang, S., Beliveau, B. J., Fudenberg, G., et al. (2016). Super-resolution imaging reveals distinct chromatin folding

for different epigenetic states. *Nature* 529, 418–422. doi: 10.1038/nature 16496

Chevrollier, A., Cassereau, J., Ferré, M., Alban, J., Desquiret-Dumas, V., Gueguen, N., et al. (2012). Standardized mitochondrial analysis gives new insights into mitochondrial dynamics and opa1 function. *Int. J. Biochem. Cell Biol.* 44, 980–988. doi: 10.1016/j.biocel.2012.03.006

Cremer, C., Szczurek, A., Schock, F., Gourram, A., and Birk, U. (2017). Super-resolution microscopy approaches to nuclear nanostructure imaging. *Methods* 123, 11–32. doi: 10.1016/j.ymeth.2017.03.019

Cuplov, V., Buvat, I., Pain, F., and Jan, S. (2014). Extension of the gate monte-carlo simulation package to model bioluminescence and fluorescence imaging. *J. Biomed. Opt.* 19:026004. doi: 10.1117/1.JBO.19.2. 026004

Diggle, P. J. (1983). *Statistical Analysis of Spatial Point Patterns*. London: Academic Press.

Gazagnes, S., Soubies, E., and Blanc-Feraud, L. (2017). "High density molecule localization for super-resolution microscopy using cel0 based sparse approximation," in *2017 IEEE 14th International Symposium on Biomedical Imaging* (Melbourne, VIC), 28–31.

Griffié, J., Shannon, M., Bromley, C. L., Boelen, L., Burn, G. L., Williamson, D. J., et al. (2016). A bayesian cluster analysis method for single-molecule localization microscopy data. *Nat. Protoc.* 11, 2499–2514. doi: 10.1038/nprot.2016.149

Hafiane, A., Seetharaman, G., and Zavidovique, B. (2007). "Median binary pattern for textures classification," in *International Conference Image Analysis and Recognition* (Berlin: Springer), 387–398.

Hansson, K., Jafari-Mamaghani, M., and Krieger, P. (2013). Ripleygui: software for analyzing spatial patterns in 3d cell distributions. *Front. Neuroinform.* 7, 5–14. doi: 10.3389/fninf.2013.00005

Haralick, R. M., Shanmugam, K., and Dinstein, I. (1973). Textural features for image classification. *IEEE Trans. Syst. Man Cybern.* 3, 610–621. doi: 10.1109/TSMC.1973.4309314

Holden, S. J., Uphoff, S., and Kapanidis, A. N. (2011). Daostorm: an algorithm for high-density super-resolution microscopy. *Nat. Methods* 8, 279–280. doi: 10.1038/nmeth0411-279

Kleppe, A., Albregtsen, F., Vlatkovic, L., Pradhan, M., Nielsen, B., Hveem, T. S., et al. (2018). Chromatin organisation and cancer prognosis: a pan-cancer study. *Lancet Oncol.* 19, 356–369. doi: 10.1016/S1470-2045(17)30899-9

Kolin, D. L., and Wiseman, P. (2007). Advances in image correlation spectroscopy: measuring number densities, aggregation states, and dynamics of fluorescently labeled macromolecules in cells. *Cell Biochem. Biophys.* 49, 141–164. doi: 10.1007/s12013-007-9000-5

Kundu, S., Ji, F., Sunwoo, H., Jain, G., Lee, J. T., Sadreyev, R. I., et al. (2017). Polycomb repressive complex 1 generates discrete compacted domains that change during differentiation. *Mol. cell.* 65, 432–446. doi: 10.1016/j.molcel.2017.01.009

Kylberg, G., and Sintorn, I.-M. (2013). Evaluation of noise robustness for local binary pattern descriptors in texture classification. *EURASIP J. Image Video Process.* 2013:17. doi: 10.1186/1687-5281-2013-17

Lagache, T., Lang, G., Sauvonnet, N., and Olivo-Marin, J.-C. (2013). Analysis of the spatial organization of molecules with robust statistics. *PLoS ONE* 8:e80914. doi: 10.1371/journal.pone.0080914

Lakadamyali, M., and Cosma, M. P. (2015). Advanced microscopy methods for visualizing chromatin structure. *FEBS Lett.* 589, 3023–3030. doi: 10.1016/j.febslet.2015.04.012

Lee, S. R., and Han, J. (2017). Mitochondrial nucleoid: shield and switch of the mitochondrial genome. *Oxidat. Med. Cell. Longev.* 2017, 1–15. doi: 10.1155/2017/8060949

Lehmussola, A., Ruusuvuori, P., Selinummi, J., Huttunen, H., and Yli-Harja, O. (2007). Computational framework for simulating fluorescence microscope images with cell populations. *IEEE Trans. Med. Imaging* 26, 1010–1016. doi: 10.1109/TMI.2007.896925

Ma, H., Xu, J., and Liu, Y. (2019). Windstorm: robust online image processing for high-throughput nanoscopy. *Sci. Adv.* 5:eaaw0683. doi: 10.1126/sciadv.aaw0683

Mailfert, S., Touvier, J., Benyoussef, L., Fabre, R., Rabaoui, A., Blache, M.-C., et al. (2018). A theoretical high-density nanoscopy study leads to the design of unloc, a parameter-free algorithm. *Biophys. J.* 115, 565–576. doi: 10.1016/j.bpj.2018.06.024

Mirmehdi, M. (2008). *Handbook of Texture Analysis*. London: Imperial College Press.

Neath, A. A., and Cavanaugh, J. E. (2012). The bayesian information criterion: background, derivation, and applications. *Wiley Interdiscipl. Rev. Comput. Stat.* 4, 199–203. doi: 10.1002/wics.199

Ojala, T., Pietikäinen, M., and Mäenpää, T. (2002). Multiresolution gray-scale and rotation invariant texture classification with local binary patterns. *IEEE Trans. Pattern Anal. Mach. Intell.* 24, 971–987. doi: 10.1109/TPAMI.2002.1017623

Ovesný, M., Křížek, P., Borkovec, J., Švindrych, Z., and Hagen, G. M. (2014). Thunderstorm: a comprehensive imagej plug-in for palm and storm data analysis and super-resolution imaging. *Bioinformatics* 30, 2389–2390. doi: 10.1093/bioinformatics/btu202

Paunovic, J., Vucevic, D., Radosavljevic, T., Pantic, S., Veskovic, M., and Pantic, I. (2019). Gray-level co-occurrence matrix analysis of chromatin architecture in periportal and perivenous hepatocytes. *Histochem. Cell Biol.* 151, 75–83. doi: 10.1007/s00418-018-1714-5

Ripley, B. D. (1991). *Statistical Inference for Spatial Processes*. Cambridge: Cambridge University Press.

Robertson, C., and George, S. C. (2012). Theory and practical recommendations for autocorrelation-based image correlation spectroscopy. *J. Biomed. Opt.* 17:080801. doi: 10.1117/1.JBO.17.8.080801

Rubin-Delanchy, P., Burn, G. L., Griffié, J., Williamson, D. J., Heard, N. A., Cope, A. P., et al. (2015). Bayesian cluster identification in single-molecule localization microscopy data. *Nat. Methods* 12, 1072–1076. doi: 10.1038/nmeth.3612

Rust, M. J., Bates, M., and Zhuang, X. (2006). Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (storm). *Nat. Methods* 3:793. doi: 10.1038/nmeth929

Ryabichko, S., Ibragimov, A., Lebedeva, L., Kozlov, E., and Shidlovskii, Y. (2017). Super-resolution microscopy in studying the structure and function of the cell nucleus. *Acta Nat.* 9, 42–51. doi: 10.32607/2075-8251-2017-9-4-42-51

Samacoits, A., Chouaib, R., Safieddine, A., Traboulsi, A.-M., Ouyang, W., Zimmer, C., et al. (2018). A computational framework to study sub-cellular RNA localization. *Nat. Commun.* 9:4584. doi: 10.1038/s41467-018-06868-w

Schermelleh, L., Rainer, H., and Leonhardt, H. (2010). A guide to super-resolution fluorescence microscopy. *J. Cell Biol.* 190, 165–175. doi: 10.1083/jcb.201002018

Schwarz, G. (1978). Estimating the dimension of a model. *Ann. Stat.* 6, 461–464. doi: 10.1214/aos/1176344136

Stelzer, E. H. (2015). Light-sheet fluorescence microscopy for quantitative biology. *Nat. Methods* 12, 23–26. doi: 10.1038/nmeth.3219

Stevens, T. J., Lando, D., Basu, S., Atkinson, L. P., Cao, Y., Lee, S. F., et al. (2017). 3D structures of individual mammalian genomes studied by single-cell Hi-C. *Nature* 544, 59–64. doi: 10.1038/nature21429

Tan, X., and Triggs, B. (2010). Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE Trans. Image Process.* 19, 1635–1650. doi: 10.1109/TIP.2010.2042645

Xu, J., and Liu, Y. (2019). A guide to visualizing the spatial epigenome with super-resolution microscopy. *FEBS J.* 286, 3095–3109. doi: 10.1111/febs.14938