

SPOT DETECTION METHODS IN FLUORESCENCE MICROSCOPY IMAGING: A REVIEW

MATSILELE MABASO^{✉,1,2}, DANIEL WITHEY¹, BHEKISIPHO TWALA²

^{1,2}MDS(MIAS), Council for Scientific and Industrial Research, Pretoria, South Africa; ²Department of Electrical and Mining Engineering, University of South Africa, South Africa.

e-mail: MMabaso@csir.co.za, DWithey@csir.co.za, Twalab@unisa.ac.za

(Received January 4, 2017; revised October 10, 2018; accepted October 13, 2018)

ABSTRACT

Fluorescence microscopy imaging has become one of the essential tools used by biologists to visualize and study intracellular particles within a cell. Studying these particles is a long-term research effort in the field of microscopy image analysis, consisting of discovering the relationship between the dynamics of particles and their functions. However, biologists are faced with challenges such as the counting and tracking of these intracellular particles. To overcome the issues faced by biologists, tools which can extract the location and motion of these particles are essential. One of the most important steps in these analyses is to accurately detect particle positions in an image, termed spot detection. The detection of spots in microscopy imaging is seen as a critical step for further quantitative analysis. However, the evaluation of these microscopic images is mainly conducted manually, with automated methods becoming popular. This work presents some advances in fluorescence microscopy image analysis, focusing on the detection methods needed for quantifying the location of these spots. We review several existing detection methods in microscopy imaging, along with existing synthetic benchmark datasets and evaluation metrics.

Keywords: fluorescence microscopy; microscopy image analysis; spot detection; supervised; unsupervised

INTRODUCTION

Advances in fluorescence microscopy have made it possible to visualize and study the interaction of sub-cellular particles, important because the fundamental functions of biological processes occur at the subcellular level (Kervram, 2016). Cells contain subcellular structures such as nucleic acids, proteins and organelles. Monitoring these structures at a subcellular level can help in answering open questions in cell biology (Rezatofghi, 2015). The initial key step in understanding certain biological concepts is to detect and study the interaction of these subcellular structures. To do this, these structures are specifically labelled with fluorescent probes, which are then observed under fluorescence microscopy as bright spots superimposed on a dark or uneven background, as shown in Fig. 1. The term ‘spot’ in our study, refers to a local intensity maximum in the microscopy image whose intensity is significantly different from its neighbourhood. Studying the location of these spots is an important task for further image analysis (Basset, 2015; Kervram, 2016) including spot counting (Byun, 2006; Raj, 2008), spot tracking (Genovesio, 2006; Chenouard, 2014) and spot

pattern recognition (Jackson, 2011).

All of these analyses require a reliable and accurate spot detection method and, as a result, the detection of spots in microscopy imaging is generally accepted as a key step to gain information about certain aspects of cell functions and to assist further analysis (Genovesio, 2006; Ruusuvoori, 2010; Smal, 2010; Basset, 2015). The task of spot detection in microscopy images is to determine if a given image contains one or more groups of connected pixels that display high-intensity values compared to their background and to calculate their positions in the image. These high-intensity spots represent fluorescent molecules within a cell. However there are several issues which complicate quantitative analysis of microscopy data. The image development process in fluorescence microscopy can be modelled mathematically by convolving a function of a sample being imaged with a point spread function (PSF) of a microscope and corrupting the image with a noise model as described in Eq. 1:

$$Image = PSF * Object\ function + Noise, \quad (1)$$

where “*” signifies convolution.

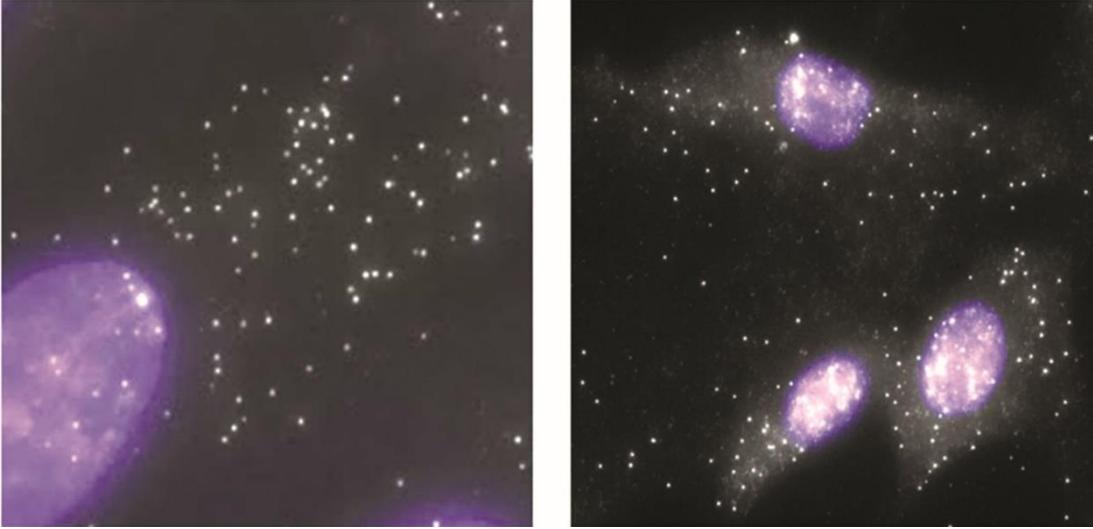


Fig. 1. An illustration of microscopy images showing multiple mRNA localization spots visualized through fluorescence microscopy (Raj, 2016).

The PSF describes the way information on the object function is spread as a result of recording the data, modelled as:

$$PSF(r) = \left(\frac{2J_1(r\alpha)}{r} \right)^2, \quad (2)$$

where,

$$\alpha = 2\pi NA / \lambda, \quad (3)$$

and NA is the numerical aperture, λ is the wavelength of light, r the distance to origin, and J_1 is the Bessel function of order 1 and the object function describes the object being imaged and the way light is emitted from the object to the imaging instrument.

The main issue which complicate data analysis in microscope images is noise. Noise is a nondeterministic function which is caused by unwanted external disturbances which occur during the image acquisition process in fluorescence microscopy. Noise refers to random variations in pixel intensities, caused by sensors. There exists a variety of noise types in micro-scopy imaging, the most common ones are readout, dark and photon noise. The photon noise in micro-scopy images is caused by stochastic nature of photons emitted. Photon noise has a Poisson distribution and is sometimes referred to as Poisson noise. This kind of noise is unavoidable and is always present in an optical image. The statistical distribution of photon noise is described by the Poisson model,

$$P(x; \lambda) = \frac{\lambda^x e^{-\lambda}}{x!} \quad \text{for } x = 0, 1, 2, \dots, \quad (4)$$

where $\lambda > 0$ is the mean and the variance of the intensity.

Dark noise is another type of noise in microscopy images, which is formed as a result of thermally-generated electrons in charge-coupled devices (CCD). Readout noise in microscopy images is introduced during the conversion from analogue to digital. The influence of such noise can be measured and be provided by the instrument. Readout noise follows a normal distribution and modelled as,

$$N(\mu, \sigma^2; x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad (5)$$

where μ , is the mean and σ^2 , is the variance.

Other issues which hinder the performance of many spot detection methods include non-uniformity in microscopy images and overlapping spots. These issues should be taken into account when developing an automated detection algorithm. There exist two ways to detect spots in microscopy imaging, manual detection, and automated detection. Manual detection requires the operator to manually record the coordinates of each spot in a given image sequence. This process can be time-consuming given large data sets, prone to errors, and very laborious. Other disadvantages of manual detection include that it is user dependent, inaccurate and its results cannot be repeated. To overcome the challenges of manual detection, automatic spot detection has drawn much attention from researchers in bio-imaging. These methods have an advantage over manual detection in the sense that they are much quicker and produce reproducible results. These methods have been previously classified as either supervised or unsupervised methods (Smal, 2010), the difference between the two are described in Section 0. Extensive efforts to develop automatic spot detection methods to

minimize the need for manual detection have been ongoing. Previous comparative study (Smal, 2010) has shown that the supervised methods perform best compared to unsupervised methods on low-quality synthetic images, (signal-to-noise ratio, SNR ≈ 2). However, as SNR increases (SNR > 5) the difference in performance of the considered methods was negligible. A number of different automated approaches were developed for the extraction of spots locations in microscopy images; these approaches included the morphological-based methods (Smal, 2008), wavelet-based methods (Olivo-Marin, 2002) and supervised based methods (Jiang, 2007; Ram, 2012). Recently, two detailed quantitative comparison studies of different spot detection methods in microscopy imaging were provided in (Ruusuvaori, 2010; Smal, 2010). The study by Smal et al. (2010) included seven unsupervised and two supervised detection methods in microscopy images. The experiments conducted included the use of synthetic images as well as on real images of fluorescence microscopy. Both studies emphasized the need for a good automated spot detection method to overcome the limitations encountered by some of the existing methods.

SPOT DETECTION

Automated Detection

The typical goal of automated detection methods is to find an association function from input patterns to an output value. In this case, we have images of a particle of interest from an input data and correct labels as corresponding output data. A typical spot detection usually consists of three steps (Smal, 2010): (i) pre-processing, (ii) signal enhancement, and (iii) object extraction, however, the procedure at which these steps are implemented can vary with signal enhancement step being the most important one.

The pre-processing step first enhances the input image, reduces image noise and suppresses the intensity of background structures (*e.g.*, cell background) using denoising algorithms resulting in a denoised image. Techniques for pre-processing can vary from basic filtering to more sophisticated methods (Dabov, 2007). A simple known approach to pre-process the image is to use linear filters. The most common linear filter is the Gaussian filter. Given an input image $I(x, y)$, the filtered image $I_f(x, y)$ is given as:

$$I_f(x, y) = G_\sigma * I(x, y), \quad (6)$$

where, G_σ , the Gaussian kernel is defined as:

$$G_\sigma(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}}, \quad (7)$$

with, σ being the width of the kernel.

Conducting a convolution with kernel enhances the regions that resemble a Gaussian kernel. The appearance of spots resembles a Gaussian kernel, so image regions resembling spots will be enhanced. In general, the Gaussian filter works on images corrupted by random noise. The Gaussian filter blurs the image by attenuating high frequencies while passing the low frequencies. Additional methods for pre-processing the images includes the non-linear filters such as median filters. The median filter replaces the target pixels with the median of neighbouring pixels.

In the signal enhancement, the goal is to find regions in an image that correspond to a spot by applying image processing techniques on the noise filtered image. The signal enhancement is the most important step in a spot detection method. There exist a number of methods for signal enhancement such as, for example, H-Dome, wavelet, kernel density estimation, a detailed discussion of these methods are listed in Section 0. A simple technique is to apply a threshold to the pre-processed image. This will result in pixel values which are above the threshold are assumed to represent a spot. This procedure depends on the threshold selected, thus setting a threshold too low will result in detecting more false positives and too high, false negatives will arise. The output of this step is a binary image which presents the likelihood of spots.

Object extraction, the main aim of this step is to derive the descriptors of each detected spot in an image. These descriptors include, for example, the coordinates, mean intensity, and size of each detected spot in an image. In order to compute such descriptors on the enhanced image, a decision threshold is employed to obtain a binary image. In order to calculate such descriptors, a connected component labelling is employed to identify set of adjacent pixels. Each connected component will represent one spot then spot properties can be computed.

After performing the above steps, biological objects are localized within an image, and their corresponding features can be measured using a statistical classifier in order for such object to be classified.

Various quantitative properties are computed such as:

- Area, this computes the total number of pixels in a detected object. It also describes the size of the detected object.
- Centroid, this estimate the centre of mass of the detected object.

- Local maximum, this describes voxel with the highest intensity in a specific region.
- Other properties include measures of dispersion and central tendency of all objects with voxel intensity. Also, the deepness of the detected object can be computed.

These computed properties can help in making final decisions of whether the detected object should be classified as a spot or not. This information may help in confirming or rejecting the presence of a disease in histopathology image analysis and improve the understanding of biological mechanisms within a cell. A number of automated spot detection methods exist in the literature; examples include (Olivo-Marin, 2002; Jiang, 2007; Kimori, 2010; Ram, 2012; Reza-tofighi, 2012; Basset, 2015; Jaiswal, 2015). A broad evaluation of various automated spot detection methods in microscopy imaging was reported in (Ruusuuvuori, 2010; Smal, 2010).

Categorization of various spot detection methods

According to the quantitative survey by Smal (2010), the existing spot detection methods can be categorized

into two groups, ‘supervised’ and ‘unsuper-vised’ methods as shown in Fig. 2. In supervised methods, a model is prepared via a training process where it is required to make predictions and corrected when those predictions are wrong. The training task continues until the model achieves the desired level of accuracy on the training data. The supervised methods are known to perform better than the unsupervised methods; however, their performance depends on the training step which is subject to having reliable ground truth data. They are usually used when a large training dataset is available. The disadvantage of the supervised methods is that they require re-training whenever the characteristics of the data change. The un-supervised methods refer to non-supervised where user-specified values are used for either the parametric template or model used for noise-suppression/spot-enhancement as well as for the threshold for distinguishing between spots and background. Typically, these methods do not require training, and usually assume that the spots are isolated, and rely on simple image processing techniques and don’t require the availability of labelled training data nor a learning framework. In an unsupervised method, a model is prepared by deducing structures present in the input data. An example of such a method is provided in (Olivo-Marin, 2002).

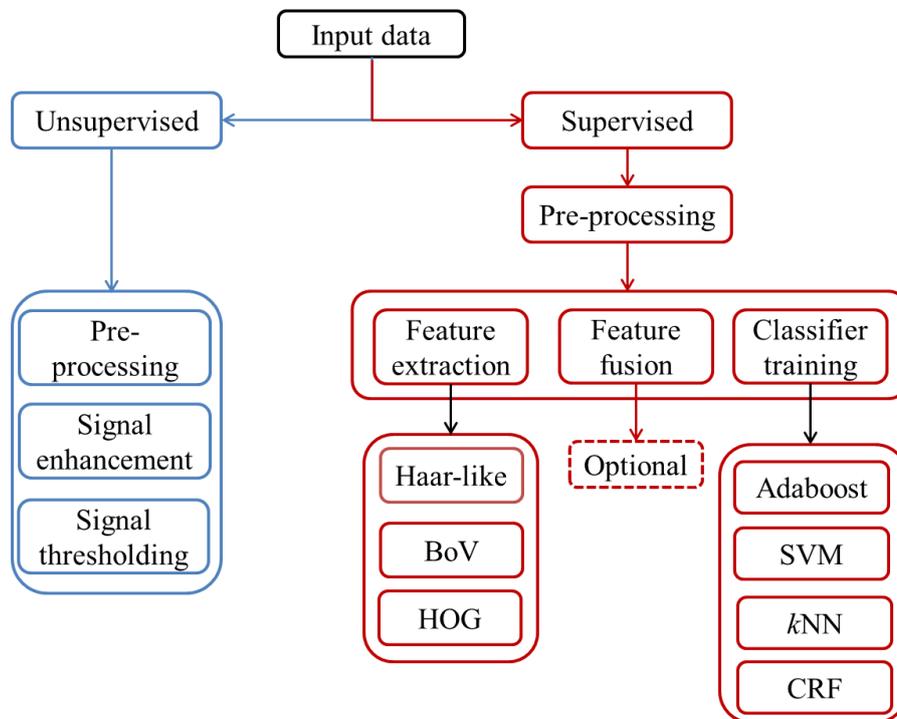


Fig. 2. Steps involved in a spot detection method (supervised and un-supervised). The abbreviations stand for Histogram of orientated gradients (HOG), Bag-of-Visual-words (BoV), support vector machine (SVM), k-nearest neighbor (kNN) and conditional random fields (CRF).

We have followed similar categorization scheme (*i.e.*, supervised and un-supervised) as proposed by Smal (2010). Figure 2 shows the various steps involved in detection methods. In this review, we included most of the popular spot detection methods in microscopy imaging. We aimed at including methods which are publicly available, commonly used, and which cover the basic spot detection algorithm classes. The following subsections provide descriptions of the algorithms, including methods. The key steps of each method are explained, and the listings of their adjustable parameters are shown with brief descriptions.

Un-supervised methods

Intensity thresholding

A detection method based on thresholding is one of the easy and widely implemented algorithms. Thresholding works by grouping pixels in an image into two classes, dark and light, depending on whether the pixel is below or above a predefined intensity threshold. The threshold parameter can be obtained using techniques such as Otsu's method (Otsu, 1979). Thresholding is the creation of a binary image from a grey-scale image by turning all pixels which are less than a given value to zero and all pixels above a given threshold to one. A thresholded image, $h(x, y)$ of $f(x, y)$ is given by

$$h(x, y) = \begin{cases} 1, & f(x, y) \geq T \\ 0, & \text{Otherwise} \end{cases} \quad (8)$$

where T is the predefined threshold.

The problem with thresholding is that it only considers the pixel's intensity and not any relationship between the pixels. Also, these approaches cannot handle complex images in which the contribution of background intensities exceeds that of a spot. In this case, the algorithm will result in detecting more false positives.

Kozubek

Kozubek (1999) presented a method for spot detection named "gradual thresholding" which is based on the watershed technique for the detection of spots in 2D and 3D FISH-stained interphase nuclei images. During the pre-processing, noise is suppressed using a Gaussian filter, followed by a series of thresholds starting from the maximum threshold (maximum intensity within the cell nucleus) to the lowest one (the mean background intensity). At each intensity level, a segmentation is applied and compared with segmentation obtained for the previous (higher) threshold. For every new spot which appears in the segmented image, two possibilities are considered:

- Either the newly detected spot corresponds to a new spot, which has not been identified so far. Or,

- the newly detected spot is a result of intensity fluctuation of the previously detected spot which is located nearby.

The differentiation of these two cases is based on the difference of intensities between the newly detected spot, and the nearby spot of which it may potentially be a part. The maximum allowed intensity difference and distance are specified based on the expected image properties. If the new object segmented at the threshold level t_{j-1} is found to be a part of a spot that was already present at t_j , the properties of the spot are updated accordingly. During the final step, the candidate spots are filtered based on size and intensity criteria.

The free parameters of the method are the following:

- Gaussian kernel width σ ,
- distance for intensity fluctuations,
- allowable intensity difference,
- spot size.

Netten

An approach similar to the Kozubek (1999) method was proposed by Netten (1997) termed "dot label". The detection method is made up of three steps. The first step consists of defining the region-of-interest (ROI) containing the nucleus. Then an ISO-DATA thresholding (Velasco, 1980) is applied within an ROI to distinguish between the nucleus and back-ground, followed by grey-value opening to remove the dots. After the nuclei segmentation, features which describe size, shape and intensity are measured and used to classify the segmented nuclei and reject debris.

Secondly, for each segmented nucleus, spots are detected by applying a 3D morphological top-hat transformation method with a structuring element larger than the expected size of a spot followed by a constant threshold, where the threshold level is defined as:

$$\theta_{threshold} = \mu_{background} + k \times \sigma_{background}, \quad (9)$$

where $\mu_{background}$ and $\sigma_{background}$ are the mean and standard deviation of the background.

These parameters, mean and standard deviation, are estimated using pixels which are less than the 90 percentile of the top-hat image. After applying a 3D top-hat transform, most spots will be detected but some will appear to be merged. A non-linear Laplacian filter is applied to split the touching dots. The Laplacian filter is sensitive to noise and will result in detecting false spots. Thirdly, to overcome these issues, a 3D

top-hat is applied with a variable threshold to label the spots. The threshold level, θ_{seed} , is swept through the image intensities starting at the maximum intensity of the image and progressing towards background intensity. When new spots are encountered at the level corresponding to the current threshold level, they are either treated as a part of an already detected spot nearby, or as a seed of a new spot. This decision is based on whether the detected spot is linked to the existing spot by a list of pixels whose intensities lie within the threshold band.

The free parameters of the method are the following:

- Radius of structuring element, r which determines the size of the detected spot
- Background level percentile (lowest intensity for the sweeping threshold band)

Seed growing

Seed growing based on local thresholding for spot detection method was suggested by Gue et al. (2005) for the extraction of spots locations in micro-scopic images. The first step of the method suppresses noise by applying a median filter to the input image. Then a 3D morphological top-hat filter is performed to further reduce the amount of noise and enhance spots, followed by the detection step. The spot detection step determines the pixels with intensity values greater than the intensity threshold $P\%$ of the histogram. These are regarded as critical pixels of each spot. Then a local threshold is computed, depending on the mean and standard deviation of the intensities at (x, y, z) lines passing through the centre of the detected spot.

Following the top-hat filtering is the segmentation of spots. The pixels with intensity in the $P\%$ of the histogram are classified as seeds where a spot will be segmented. The seed selection and growing are applied until no unsegmented pixels with the intensity in the top $P\%$ of the histogram remain.

As the final step, morphological closing and opening are applied to the spot masks, thus compacting the shapes and removing holes and objects that are too small to be considered.

The free parameters of the method are the following:

- median neighbourhood size,
- radius of the structuring element (SE) for top-hat filter,
- intensity threshold P ,
- radius of the SE for the closing and opening.

Raj

The Raj detection method (Raj, 2008) is based on multiple thresholding of an image. First, the image is filtered with a Laplacian of Gaussian (*LoG*) defined in Eq. 19 to reduce the influence of noise and enhance spot structures. Then thresholding is performed on a filtered image. Then the number of spots found upon thresholding is plotted as a function of threshold. The presence of a plateau indicates that there is a region in which the detected spots are insensitive to a threshold value. The dotted line indicates the selected threshold within a plateau in which the numbers of detected spots are determined. However, the threshold selection in this algorithm is user selected.

The free parameter for the method

- User threshold.

EMax

The EMax detection method was introduced by Matula (2010) based on 3D morphological extended maxima. The method was used for the detection of endoplasmic reticulum exit sites (ERES) in 3D confocal microscopy images. ERES are subcellular structures in a cell responsible for collecting newly formed carrier molecules to transport them to Golgi apparatus. The method starts by convolving a 3D Gaussian kernel with the original image to suppress noise and enhance spots. Then, a HMax transform is applied, defined as,

$$HMax_h(I) = R_I(I - h), \quad (10)$$

where $R_I(I - h)$, denotes the morphological reconstruction of image, I .

The HMax transform will suppress all intensity maxima whose height is less than or equal to h . The EMax transform is defined as regional maxima of the HMax image,

$$EMax_h(I) = RMax(HMax_h(I)), \quad (11)$$

where, $RMax$ is the regional maxima, which are defined by subtracting the h -maxima from image I with $h = 1$;

$$RMax(I) = I - HMax_h(I). \quad (12)$$

After the EMax transform is computed those connected components of the resulting image whose size is lower than the specified limit for fluorescence spots are accepted as spots. When adjusting the height threshold h , the number of detected spots varies.

The free parameters for the method

- Spot height threshold, h

- Maximum allowed spot size
- Gaussian kernel width, σ

The H-Dome

The HDome method was presented by Smal (Smal, et al., 2010) based on the morphological HDome transform of Vincent (Vincent, 1993). The method first uses a Gaussian filter at a scale that corresponds to the smallest spot to suppress noise in the input image. Then, the HDome transform is applied to the Gaussian filtered image as described in (Smal, 2010), defined as,

$$Hdome(I(x, y)) = I(x, y) - \rho_I(I(x, y) - h), \quad (13)$$

where $(I(x, y) - h)$ denotes the results of subtracting a constant, h , from a gray-scale image $I(x, y)$, and $\rho_I(I(x, y) - h)$ is the morphological reconstruction of the gray-scale image, $I(x, y)$ from $(I(x, y) - h)$. The parameter h corresponds to the height of the structures to be extracted. The HDome transform result in local maximum of a a the input image whose height is greater than h . This will yield regions which correspond to spots or background structures with high intensity, referred to as domes. These domes are then raised to exponent, s , and the resulting image is the sampling step. The number of samples (highest) is placed at the location where the intensity is highest. Then, these samples are divided into clusters by the mean shift algorithm (Fukunaga & Hostetler, 1975). Two criteria are defined for each cluster to be considered as a spot,

- The number of cluster samples must be high enough.
- The determinant of the covariance must be within a specified range, identifying between spots and artefacts.

Then, the position of each spot is calculated as the centre of mass of the corresponding cluster.

The free parameters for the method

- HDome height, h
- Gaussian kernel width, σ
- Sample number, n
- Kernel radius, r
- σ_M threshold for the covariance matrix
- Exponent, s , used for importance sampling

Rezatofighi

Rezatofighi (2012) proposed a detection method termed, maximum possible height dome (MPHD) for locating spots in fluorescence microscopy images. The

proposed method overcomes the limitations of the H-Dome method in (Smal, 2010), which is more sensitive to parameter, h and has a tendency of mer-ging neighbouring bright spots and sometimes misses less bright spots, due to the fact that spots do not have the same magnitude, h . The method uses a Gaussian filter to suppress the effect of background noise with the standard deviation set to be the size of the smallest object in the image. The H-Dome transform is performed on the filtered image. However, instead of choosing the threshold value of, h , by trial and error, their method computes a suitable threshold value, h , for each particle based on local information and designs a mask to enhance those particles appropriately. Two criterias are defined when designing an adaptive mask;

- First, for each local maximum in a filtered image, an optimal point is defined by using a line segment, $l_s(\theta, r)$, with θ being the angle and r the radius. The optimal point is then defined as the nearest local minimum, x_l^θ within a search area.
- The optimal point, x_l^* , has the maximum intensity,

$$x_l^* = \underset{x \in L_x}{\operatorname{argmax}}(I(x)).$$

The mask, M_a is obtained by centering peaks within the positions of local maxima with an intensity that corresponds to $I(x_l^*(i))$ instead of using, $I - h$, as in (Smal, 2010).

The maximum possible height dome, H is then defined as:

$$H(I, M_a) = I - R(I, M_a). \quad (14)$$

The image, H contains all maximal structures such as, noise, background structures and spots. The spots will appear as domes with intensity higher compared to the regional background. Then applying intensity threshold, T will suppress background structures. The spots positions are then calculated using intensity-weighted centroid of the domes.

The free parameters are:

- Gaussian filter standard deviation, σ
- Dome radius, r
- Intensity threshold, T

Kimori

The top hat rotational morphological processing (RMP) algorithm is a morphological processing algorithm presented by Kimori (2010) for the detection of spots from electron microscopy images. The proposed algorithm consists of three steps, noise reduction, spot

extraction, and binarization. The first step, noise reduction, suppresses noise in an image by using a Gaussian filter resulting in a filtered image, $I_f(x, y)$.

The second step, spot extraction, this step involves the rotation of the filtered image, $I_f(x, y)$ in a clockwise direction with respect to the centre of the image frame. Assuming a half circle (π [rad]) divided into N equiangles, the rotated images by the angle $\theta_i = \pi i/N$ with i is the number of rotations in range $[0, 1, \dots, N - 1]$ is denoted as, $I_i(x, y)$ which is then subject to opening and closing operation. The opening operation of a rotated image with a structuring element, B are represented as $\gamma_B(I_i(x, y))$. The structuring element B is chosen to be larger than the noise width in an image. These opened images are then rotated i times in counter-clockwise denoted as $R_i(x, y)$. Finally, all the rotated images, $R_i(x, y)$ are then unified and images having the maximum intensity for the same pixel are considered in generating the entire image. The unified images are subject to opening by top-hat RMP with a structuring element, B . These unified opened images are denoted as $O'_B(x, y)$.

$$O'_B(x, y) = \max_{i \in (0, 1, \dots, N-1)} R_i(x, y). \quad (15)$$

Then the top hat transformation based on RMP is defined by subtracting the unified opened images $O'_B(x, y)$ from original noise filtered image $I_f(x, y)$ which is given as:

$$TH_B(x, y) = I_f(x, y) - O'_B(x, y). \quad (16)$$

The last step involves binarizing the top hat transformed image, $TH_B(I(x, y))$ using a thresholding method in equation (17).

$$H(x, y) = \begin{cases} 1, & TH_B(x, y) > 0 \\ 0, & \text{Otherwise} \end{cases}. \quad (17)$$

Then by computing the intensity-weighted centroid of the binarized image, the spots positions can be estimated.

The free parameters are:

- Length of structuring element, SE

Spot enhancing filter (SE)

The spot enhancing filter (SE) method was proposed by (Sage, 2005) for detecting spots in microscopy images. The algorithm enhances spots in images while reducing background structures and suppressing noise. The initial step convolves the original image, $I(x, y)$ with a Laplacian-of-Gaussian (LoG) filter,

$$I_f(x, y) = LoG(x, y, \sigma) * I(x, y). \quad (18)$$

The kernel is defined as,

$$LoG(x, y, \sigma) = \frac{x^2 + y^2 - 2\sigma^2}{2\pi\sigma^2} e^{-\frac{(x^2 + y^2)}{2\sigma^2}}, \quad (19)$$

where, σ , is the Gaussian kernel width chosen based on the size of the particles. The second step of the algorithm consists of intensity thresholding and a connected-component labelling algorithm. The intensity threshold, T , for the filtered image is computed by incorporating the mean, μ_{int} of the filtered image, user defined factor, k times the standard deviation, σ_{int} of the filtered image intensities.

$$T = \mu_{int} + k\sigma_{int}. \quad (20)$$

After thresholding, a connected component labelling is performed to identify spots and each spot location is determined by computing the intensity-weighted centre of mass.

The free parameters are:

- LoG threshold
- Gaussian kernel width, σ

Wavelet Multiscale Product (WMP)

A wavelet multiscale product-based spot detection method was presented by (Olivo-Marin, 2002) for the detection of spots in microscopy images. The method is based on the assumption that spots will be present at each scale of wavelet decomposition and thus will appear in the multiscale product. The method first applies an *à trous* wavelet transform on the original image, $I(x, y)$ to filter out unwanted signals resulting in a smoothed image, $I_i(x, y)$. The so-called wavelet plane, W at level i is then computed by subtracting the smoothed image at level of previous level, $i - 1$, defined as,

$$W_i(x, y) = I_{i-1}(x, y) - I_i(x, y), \quad (21)$$

where

$$I_i(x, y) = I_{i-1} * h \quad 1 \leq i \leq J. \quad (22)$$

With h being the smoothing kernel and I_0 , the original image, and J the number of scales used. Then a thresholding is performed to reduce the number of noisy wavelet coefficient from each scale. At, each scale, W_i , the wavelet transform results correspond to the image matching that scale. This may include both real spots and unwanted noise and background structures. Then the information contributed by each of the scales, W_i , is combined as,

$$P_j(x, y) = \prod_{i=1}^J W_i(x, y). \quad (23)$$

The multiscale product P_j , provides significant values where an important features is present and low values for all other feature. Then, P_j , is binarized using a threshold, to yield in connected components.

The free parameter for the method is,

- Wavelet scales

Jaiswal

Jaiswal (2015) proposed a detection method similar to (Sage, 2005) termed ‘multi-scale spot enhancing filter’ (MSSEF). The proposed method was used for the detection of avian leukosis virus particles in confocal fluorescence microscopy images. The approach works by, first multiplying the input image, $I(x, y)$ with a binary mask $b(x, y, \sigma^{(s-1)})$ obtained at scale $\sigma^{(s-1)}$ then convolving with a Laplacian of Gaussian filter (LoG). The mask $b(x, y, \sigma^{(s-1)})$ is obtained as a result of thresholding the image, $I_f(x, y, \sigma^{(s)})$ with a threshold $T^{(s-1)}$. At each scale, the threshold $T^{(s-1)}$ is calculated for the image as,

$$T^{(s-1)} = \mu_{int}^{(s-1)} + k\sigma_{int}^{(s-1)}, \quad (24)$$

where k is constant, defined by the user.

The filtered image $I_f(x, y, \sigma^{(s)})$ at scale $\sigma^{(s)}$ is as a result of s recursive convolution steps of the original image, $I(x, y)$ with a LoG filter, $LoG(x, y, \sigma^{(s-1)})$ denoted as,

$$I_f(x, y, \sigma^{(s)}) = LoG(x, y, \sigma^{(s)}) * (b(x, y, \sigma^{(s-1)})I(x, y)). \quad (25)$$

The result of multiplying the image with a binary mask sets the noisy background to zero while the foreground intensity pixels remain unchanged. Spots are identified by applying a connected components labeling and the positions are estimated as an intensity-weighted centre of mass.

The free parameter for the method is,

- Constant k
- Gaussian kernel width, σ

Basset

Another method based on optimal scale selection was proposed by (Basset, 2015) for the detection of vesicles in microscopy imaging. The method was named ‘adaptive thresholding of Laplacian of Gaussian images with auto selected scale (ATLAS)’. The algorithm consists of selecting an optimal scale that corresponds to the spot size in the image. The method

mainly consists of two steps, optimal scale selection and adaptive segmentation.

The scale selection step is interested in determining the representative scale of an image which is shared by most of the spots. The scale-space representation is based on the framework proposed by Lindeberg (1998), which is based on using Gaussian kernels to build a representation scales. The scale-space representation, $\{S_t\}$ of an image I is defined as,

$$\forall t \in \mathbb{R}, \quad S_t = G_t * I, \quad (26)$$

where G_t is the 2D Gaussian convolution kernel of variance t and, $*$ is the convolution parameter. Then a scale normalised Laplacian operator is applied in order to suppress noise and enhance the spots. The isotropic Laplacian kernel, ∇^2 for 2D images is defined as,

$$\nabla^2 = \begin{pmatrix} 1/6 & 2/3 & 1/6 \\ 2/3 & -10/3 & 2/3 \\ 1/6 & 2/3 & 1/6 \end{pmatrix}. \quad (27)$$

To reduce the computation, the Gaussian and Laplacian filters are combined using a single normalised LoG kernel h_t , $\exists t\nabla^2 S_t = h_t * I$. Then a multi-scale LoG is defined as,

$$\forall t \in \mathbb{R}, \quad H_t = h_t * I. \quad (28)$$

The LoG filter gives negatives values in the presence of bright spots, the scale is selected based on negative extremes of the LoG filter, referred to as negative blobs. These blobs are mostly located in two specific areas:

- the centre of Gaussian spots corresponding to the vesicles;
- at the bright pixels which are caused by noise.

These blobs are then used to select a LoG scale, a scale in which the number of blobs is the highest is considered as the best scale. The scale selection criteria proposed for the detection of spots is denoted as,

$$C_R: s^* = \operatorname{argmax}_{s \in S} (\rho_s(I) - \rho_s(\varepsilon)), \quad (29)$$

the parameters, $\rho_s(I)$ and $\rho_s(\varepsilon)$ are the blob densities in I and additive Gaussian noise ε respectively at scale s . The selected scale is the one with the maximum number of blobs at which the dissimilarity between $\rho_s(I)$ and $\rho_s(\varepsilon)$ is highest.

After the scale is selected, the second step is the segmentation step, which is based on the local thresholding of the Laplacian of Gaussian (LoG) of the intensity image. The local threshold is calculated based

on the p-value, P at every point, p from the local image statistics,

$$T(p) = \Phi^{-1}(P) \times \sigma^2(p) + \mu(p), \quad (30)$$

where the parameter, Φ denotes the Gaussian distribution function. Thus, thresholding the LoG-filtered image result in connected components. Then, spot locations are estimated by computing the centroid of each of the connected component resulting in a set of locations (r). Then a spot of a ground truth is correctly detected if an only if, its nearest neighbour in the set of detected centroids is closer than 4 pixels away and also the nearest neighbour is in the ground truth set of locations.

The free parameters for the method are,

- The probability of false alarm, P
- Standard deviation of the Gaussian window

C-Craft

The method of conditional random fields for protein transport carrier's segmentation (C-CRAFT) was proposed by (Pécot, 2015) to estimate the back-ground and segment spots in 2D+t and 3D+t fluorescence microscopy images. The detection of spots and the estimation of background are constructed as a global energy minimization problem in the conditional random fields framework. A patch-based image representation is used to detect special irregularities in the image and an iterative scheme based on graph algorithm is proposed for energy minimization. The first step of the algorithm applies a variance stabilizing technique (Boulanger, 2010) to convert the Poisson-Gaussian noise to white noise. Suppose $y_k = \{y_k^i\}_{i \in S}$ be a given data from the input time sequence with y_k^i being the grey value at site i and time k and let $x_k = \{x_k^i\}_{i \in S}$ be the binary map which indicates if a spot is present ($x_k^i = 1$) or not ($x_k^i = -1$) in the image at time, k .

The global an energy $E(x_k, b_k, y_k)$ is then given as,

$$E(x_k, b_k, y_k) = \sum_{i \in S} (H_D(x_k^i, y_k) + \beta H_B(b_k^i, x_k^i, y_k^i)) + \alpha \sum_{\langle i, j \rangle} H_V(x_k^i, x_k^j, \hat{x}_{k-1}^j). \quad (31)$$

The parameter, $H_D(x_k^i, y_k)$ is a discriminative potential for spot detection, $H_B(b_k^i, x_k^i, y_k^i)$ a potential used for determining the difference of y_k^i and background b_k^i , $H_V(x_k^i, x_k^j, \hat{x}_{k-1}^j)$ is the Ising model (Tanaka, 2003) and $\langle i, j \rangle$ denotes a sets of cliques. The values β and α are positive constants used to

balance the energy terms. The energy function $E(x_k, b_k, y_k)$ is minimized based on min-cut/max-flow algorithm (Boykov and Kolmogorov, 2004).

The second step of the algorithm is based on patch approach. The idea in this step is to combine Markov random fields with the patch-based approach and perform a pairwise comparison of n-dimensional patches. This involves the following measurement

$$y_k^i = \sum_{r \in N_i} \frac{\|P_k(i) - P_k(j)\|^2}{4\sigma^2} - \left(\frac{n}{2} - 1\right) \log(\|P_k(i) - P_k(j)\|), \quad (32)$$

where $P_k(i)$ is the patch of $\sqrt{n} \times \sqrt{n}$ centred at site i at time k and σ^2 is the noise variance. N_i . This measure is used to distinguish between the background and vesicle class and it takes high values at the vesicles locations and small ones in the background. Then a threshold, T_k is defined based on Chebyshev inequality (Jiajun, 2016) to discriminate between the two classes. Thus after thresholding, a connected component is performed to identify spots in a binary image, and then their locations are estimated as mass centres.

The free parameters for the method are,

- The p-value for hypothesis
- Spatial regularization weight
- β and α

Worz

Worz (2010) introduced an automatic approach method for the quantification and localization of telomeres and promyelocytic leukemia (PML) bodies. The approach is divided into two steps. The initial step is a 3D spot detection step. In this step, a variety of 3D filtering and smoothing methods are performed to an input image to get coarse spots. Then, a 3D Gaussian filter with σ_f proportional to the size of the desired spot width is applied to the image. After smoothing the image, the image intensities are clipped to suppress the background based on the threshold value, $t_{clip} = \mu + c \times \sigma_\mu$, in which μ and σ_μ are the mean and standard deviation of the histogram respectively and c a user defined parameter. After, this the local maxima are identified in the image. The second step is spot quantification. In this step, each of the detected maxima is fitted to a 3D Gaussian model g_m using a least square model by minimizing,

$$\sum_{x \in ROI} (g_m(x, p) - g(x))^2, \quad (33)$$

with, p , representing the vector containing the spot model parameters. The minimization of the objective function is based on the method of Marquardt (1963).

Spot positions can be obtained based on sub-voxel estimates.

The free parameters of the method are the following

- Gaussian kernel width, σ
- Threshold coefficient, c
- Parameter ranges for spot model

Feature point detection (FPD)

The feature point detection method was developed by Sbalzarini and Koumoutsakos(2005) for the detection of spots in microscopy images. The detection of spots in this algorithm is based on locating local intensity maxima followed by a weighted intensity centroid calculation. The method is made of four steps, the first step, image restoration, suppresses the imperfections in the input image, $I(x, y)$ by convolving it with a Gaussian kernel as described in Eq. 7.

The second step, estimating particle location. This step locates local intensity maxima in the filtered image, $I_f(x, y)$. A pixel is considered to be local maxima if no other pixels within a radius w around it is brighter, and its intensity is in the upper p -percentile. These local intensity maxima are identified using grayscale dilation followed by selecting all pixels that have the same values before and after dilation. The Third and fourth step refine particle location and perform non-particle discrimination. The detected intensity maxima are filtered to reduce the number false positives. These false positives are identified as maxima with intensity moments of order 0 and 2 that are significantly different from all other maxima. Clustering is performed on all points in the intensity moment space and those with density smaller than a given threshold are discarded. Then for every retained point, its position is estimated as the intensity-weighted centroid within the radius w

The free parameters are

- Gaussian kernel width, σ
- P-percentile: threshold for intensity moment space
- Cut-off; size of the structuring element for dilation

Wilson

Wilson (2016) presented a method for the detection of EGFP-labelled large dense-cored vesicles (LDCVs) in rat pheochromocytoma (P12) cells in images obtained via fluorescence microscopy. The presented method is based on the concept of particle probability image (PPI) (Yang, 2010) which computes useful features of spots in a statistical manner. The method

consists of two stages: particle enhancing filter and particle segmentation. The first step, particle enhancing filter is constructed in three stages. (1) The initial step constructs a PP image of the original greyscale image. The construction of PP image uses Haar-like features which is computed for each pixel, $p = (x, y)$ in the original greyscale image at a different scale, s to suppress the amount of background noise in a greyscale image. At a given scale, s , the Haar-like features, H_j^s for ($j = 1, 2, 3$) are defined by:

$$H_j(p) = \max_s(H_j^s(p)). \quad (34)$$

Following the HLF computation, these HLFs are linearly combined as:

$$H(p) = \sum_{j=1}^3 c_j H_j, \quad (35)$$

where, c_j is the normalised weights for each of the Haar-like features, then a weak threshold, $H(p) \geq \lambda$, is introduced on the Haar-like features in order to classify each of the pixels, p to either (particle or background). A particle probability image (PPI) is then defined based on the ratio of the number of spatially connected particle pixels to the total number of pixels in a small region of a particle size.

$$P(p) = \left(\frac{\Delta N}{N_{total}} \right)_{A_p}, \quad (36)$$

where, N_{total} is the number of pixels within a given area A_p centred at p and ΔN is the number of pixels in A_p that satisfies $H(p)$.

The second stage of the method is the particle segmentation which is implemented based on the estimation of the particle existing regions (PERs) of particles and their corresponding markers in the refined PP image. The marker-controlled watershed method is applied to accurately segment the particle regions from the original grayscale image. The estimation of PERs allows the extraction of particle positions at a subpixel level and accurate estimation of particle topologies such as size and intensity. The spots are then localized by computing the intensity-weighted centre of mass of the particle existing regions.

The algorithm free parameters are

- λ , the threshold to classify pixels into two groups
- s the scale of the haar

Supervised methods

Supervised methods recently received much attention in the field of microscopy image analysis for the detection of spots. The main steps involved in these methods as illustrated in Fig. 2 are feature extraction

and classifier training and in some cases a feature fusion. These steps play an important role in the performance of supervised spot-detection methods.

Feature extraction: this step consists of selecting features of interest from a given raw image. These discriminative features are used to build a high dimensional feature space. This step is critical to constructing high-performance spot detection methods. There exists a variety of methods to extract features from a given dataset. Examples include bag-of-words and haar-like features. Haar-like features were introduced by Viola and Jones (Viola & Jones, 2001) for the detection of faces and now introduced to the field of spot detection in microscopy images. Once features are extracted, a classifier can be trained with an idea of minimizing the misclassification error on the training dataset.

Classifier training phase: The training involves two main stages. The first part is the learning stage, in which the classifier adapts to the expected images, followed by the test stage, applying the trained classifier to the actual data. In the training phase, the classifier is given a set, τ , of τ training samples with m labelled observations defined as,

$$\tau = (a_i, b_i), \quad i = 1, 2, \dots, m, \quad a_i \in \mathbb{R}^d. \quad (37)$$

Each training sample consists of d -dimensional input variable a_i which is referred to as input features and b_i is the output which can be referred to as a class label. The set needs to be constructed carefully so that it contains both positive and negative examples covering the whole range of the expected inputs. The main idea of the training process is based on learning the data and construct a prediction model \hat{h} of b_i from the training data set, $\hat{h}(a_i) = b_i + \varepsilon$, where ε is the random error value.

Testing stage, in the testing step, the constructed model in the training phase is used to predict class label b_{new} for a new sample a_{new} .

$$b_{new} = \hat{h}(a_{new}). \quad (38)$$

The problem which arises with supervise-detection algorithms is overtraining. Discussion about this problem and how to avoid can be found in (Kleinberg, 1996). The frequently used classifiers are support vector machine (SVM), AdaBoost, nearest neighbour and decision trees.

The support vector machine (SVM) is the well-known and most popular method among the machine learning algorithms for solving classification problems. The main idea of SVM is to transform the training data into a higher dimensional feature space and search for an optimal decision boundary.

AdaBoost is a widely-used, supervise-method with the idea of combining many weak classifiers to form a strong classifier based on adjusting the weights of training samples. Given the training data τ , the weights are initialized as, $w_i = 1/m$ for all the training examples, then a weak classifier h_i is trained based on weighted least square fitting. Then the weights w_i are updated as $w_i = w_i \times \exp(-b_i h_i(a_i))/L$ with L being the normalization factor. The strong classifier H is calculated as a linear combination of all weak classifiers, defined as,

$$H = \sum_{i=1}^j w_i h_i. \quad (39)$$

Jiang

Jiang (2007) presented an approach for the detection of spots in microscopy images based on the idea proposed by Viola and Jones (2001). The proposed method is based on using of Haar-like features and AdaBoost to detect the spot positions in the images, where a classifier is trained to identify spots in image patches based on haar like features, and features which give the best detection results are then iteratively selected with the boosting algorithm. The number of selected features is fixed by the user. Once the classifier is trained, the detection process is done by sliding a window of a given size through an image, and these windows are then passed into a trained classifier to check if that region contains a spot or not, using a user-defined threshold. Windows which contain spots are subject to further processing to compute additional statistical properties such as, centroid, contrast, angular moment and etc.

The output of AdaBoost is a single strong classifier which is a linear combination of the set of weak classifiers.

The free parameters for the method are:

- Feature set
- User threshold

Ram

The proposed method by Ram (2012) is conducted in two steps: spot segmentation and detection. The segmentation step is based on the unsupervised approaches and the detection utilizes supervised methods. The segmentation consists of applying edge-enhancing diffusion (Gerig, 1992) to sharpen the spot edges while reducing the effect of the noise, then a 3D morphological top-hat filter is utilized to enhance the quality of spots, with the size of the structuring element (SE) selected according to the size of the expected spot. Then, the top-hat filtered image is thresh-

olded using unimodal thresholding to obtain the spot locations, and the candidate spots are segmented using a region growing.

The 2nd stage of the method then performs the classification of the segmented spot candidates using Bayesian classifier which labels the candidates spots as actual spots, or false detections. The spot features used to train the classifier include the average intensity and its standard deviation, volume, shape, convexity, contrast, and the intensity gradient in the interior of the object and at its edge.

The free parameters of the unsupervised stage are the following:

- Edge enhancing diffusion parameters (k and a)
- Number of diffusion iterations
- Radius of the structuring element for the top-hat
- Threshold adjustment
- Bayesian training

Fisher discriminant analysis (FDA)

The Fisher Discriminant Analysis (FDA) (McLachlan, 2004; Smal, 2010) method is a pattern recognition technique that aims at obtaining a combination of variables that separate the two classes. The idea of FDA is to find a projection to a line such that samples of different classes are well separated. In this case of spot detection, the idea is to find a projection in which the separation between spots and background is maximized considering the mean and standard deviation of each class. The mean and the covariance are calculated using the features obtained from the training samples. The computation of FDA relies on the maximization of $J(w)$,

$$J(w) = \frac{(w^T(m_1 - m_2))^2}{w^T(s_1 + s_2)w}, \quad (40)$$

where m is the mean, s the standard deviation between two classes (spot and background) and T is the transpose. The indexes 1,2 represents the two classes. The main idea is to find w , that maximizes $J(w)$, then this will guarantee that the classes are well separable. The solution for the scale factor, w which maximizes $J(w)$ is then defined as,

$$w = (s_1 + s_2)^{-1}(m_1 - m_2). \quad (41)$$

The detection process involves applying a sliding window onto an image to extract patches of a given size. Then, these patches are passed onto an FDA classifier to check if they contain a spot or not. A user threshold is defined to separate between a spot or nonspot. Patches containing a spot are subject to further processing.

The free parameters for the method are:

- Patch size
- User threshold

Logistic regression with Markov random field (LR-MRF)

The method of logistic regression with Markov random field (LR-MRF) was proposed by Ruusuvoori (2012) based on selecting significant features from a set of candidate features. The approach was used for the detection of subcellular structures in microscopy images. The proposed methodology consists of three steps. The first step of the algorithm consists of creating a pool of candidate feature sets using unsupervised detection methods, wavelet decomposition, morphological top-hat, the edge enhancing and Gaussian low-pass filter. Secondly, these features are weighted by the logistic regression in which useful features are preserved and non-informative ones are excluded in the model. In this step, the logistic regression is combined with a least shrinkage and selection operator (LASSO) (Tibshirani, 1994). The LASSO method is based on the least square method but introduces a penalty factor, the l_1 - norm which penalizes the error function. Thus a logistic classifier with LASSO model is designed as,

$$p(c_i|x_i) = 1 / (1 + \exp(\beta_0 + \beta^T x_i)), \quad (42)$$

where x_i , is the feature vector and the model parameters β_0 and $\beta = (\beta_1, \beta_2, \dots, \beta_N)^T$ are estimated by maximizing the log-likelihood

$$\sum_{x_i \in F} \log p(c_i|x_i) + \sum_{x_i \in B} \log(1 - p(c_i|x_i)) - \lambda \|\beta\|_1, \quad (43)$$

using the LASSO, and $p(c_i|x_i)$, the probability of finding pixel label c_i belonging to the foreground given feature vector, x_i . where F and B are foreground and background pixels of training set and λ is the regularization parameter selected by cross-validation.

Then an Ising model (Tanaka, 2003) is used to extract the spots in an image.

The free parameters for the model are:

- Regularization parameter, λ
- Feature sets

SUMMARY

The above section studied the state-of-the-art methods in microscopy images, and each method was described in detail, and its tunable parameters were given. The discussed methods were categorized into two types,

unsupervised and supervised. The supervised methods were defined as those that require training, and the unsupervised methods do not require training. The strength of the unsupervised methods is the flexibility involved in incorporation of shape, geometry and no training needed makes the unsupervised methods a multi-source capable. However, the solution to the full automation of segmentation process is still missing. The expert knowledge for how to define the classification rules is still subjective. In supervised methods, the spot model can be established automatically via machine learning technique. The detection system is scalable and compatible, have high detection accuracy. The main limitation of the supervised methods, they require a lot of training samples of spots and non-spots to teach classifiers; as a result, the detection accuracy depends on the training samples used to train the classifier. In biological imaging and spot detection, the datasets used in the testing of all referenced methods remain limited in terms of context and challenges. Real data are far more complex than training images,

especially the SNR may be too low in real images, and the spots of interests vary in size and intensity. Additionally, more realistic and challenging datasets with ground truth to quantitatively evaluate the referenced methods are under study (Mabaso, 2016).

Table 1. Number of main parameters to be set by the user for the reviewed methods. The aforementioned methods contain one or more parameter related to spot size need to be tuned by the user. Thus setting this parameter too small will result in over detection of spots due to noise, and if the value is too high, the true spots will be smoothed out. These parameters offer ways to adjust a particular method to your specific need. These methods provide an easy-to-use tool for biologists to interpret microscopy images effectively. The number of tunable parameters plays a crucial role in making an easy-to-use image analysis tool for biologists. The greater the number of parameters which need to be tuned becomes a painstaking process for the end user, as results, fewer tunable parameters are recommended.

Table 1. Overview of the spot detection methods with the approach used and number of user-set parameters.

Method	Approach	Number of user tunable parameters
UNSUPERVISED METHODS		
Kozubek (Kozubek, 1999)	Adaptive thresholding.	4
Netten (Netten, 1997)	Morphological top-hat transform with adaptive thresholding.	2
Seed growing (Gue, 2005)	Morphological top-hat transform with adaptive thresholding.	4
Raj (Raj, 2008)	Multiple thresholding of a Laplacian of Gaussian filtered image.	1
EMax (Matula, 2010)	Morphological H-maxima transform.	3
HDome (Vincent, 1993; Smal, 2010)	Morphological HDome transform with the mean-shift algorithm.	6
Rezatofighi (Rezatofighi, 2012)	HDome transform with automatic threshold selection.	3
Kimori (Kimori, 2010)	Morphology top-hat with rotational morphological processing.	1
SEF (Sage, 2005)	Convolution of Laplacian of Gaussian filter followed by intensity thresholding.	2
WMP (Olivo-Marin, 2002)	Multiscale product of wavelet coefficients.	1
Jaiswal (Jaiswal, 2015)	Recursive convolution of LoG kernel applied to masked image	2
Basset (Basset, 2015)	Adaptive thresholding.	2
C-Craft (Pécot, 2015)	Conditional random fields and patch-based image representation.	4
Worz (Worz, 2010)	3D Gaussian intensity model and 3D least-squares fitting.	3
FPD (Sbalzarini & Koumoutsakos, 2005)	Percentile detection and non-particle discrimination.	3
Wilson (Wilson, 2016)	Particle probability image using Haar-like features and marker-controlled watershed.	2
SUPERVISED METHODS		
Jiang (Jiang, 2007)	Haar-like features and Adaboosting.	2
Ram (Ram, 2012)	Top-hat transform with adaptive threshold and Bayesian classifier.	5
FDA (McLachlan, 2004; Smal, 2010)	A spot enhancing filter is learned in the training data with a threshold value for detection. The filter coefficients are computed from intra- and inter-class mean and variance in patches.	2
LR-MRF (Ruusuuvuori, 2012)	A feature set is built using a number of spot detection methods then LR-MRF method is applied to select significant features.	2

EVALUATION METRICS AND DATASETS

Over the past years, efforts have been made in the development of various detection methods in microscopy imaging. In order to compare the results of the existing methods, it is critical to introduce some freely available datasets and evaluation metrics.

DATASETS

There are various datasets and tools publicly available which are successfully used in testing the performance of spot detection methods. Due to the lack of ground truth in real fluorescence microscopy images, the evaluation or testing the performance of any spot detector relies on the use of synthetic images. Therefore, this study will focus on the synthetic datasets.

SIMCEP (Lehmussola, 2007; Ruusuvuori, 2008) is a benchmark dataset of synthetic cell population images with ground truth consisting of cell populations; clustering with increasing probability, cells with nuclei, cytoplasm and subcellular objects, and cells from two populations. The SIMCEP simulator can be tuned to simulate images for a specific research problem.

Synthetic data generator (Smal, 2009). This tool is used to generate different types of synthetic images with round and elongated spots with various backgrounds. The tool was used in a number of studies but it does not create a time sequence images, and as a result, the tool is applicable for the creation of synthetic images for the evaluation of spot detection methods.

Another tool which offers more flexibility termed ‘particle tracking benchmark generator’ was proposed by Chenouard (2016).

Particle tracking benchmark generator (Chenouard, 2016) provides a powerful tool for the creation of synthetic image sequences, for both 2D+t and 3D+t. The tool is designed to create synthetic images for testing both detection and tracking algorithms.

Realistic synthetic datasets (Mabaso, 2016) framework is based on the tool proposed by Chenouard. The addition made into this tool is the option of using the real background in the creation of synthetic image sequences. As a result, the datasets created by this framework offer an excellent resource for testing the performance of various spot detection algorithms.

EVALUATION METRICS

The well-known measures for evaluating various spot

detection methods in microscopy images are F-measure, precision and recall. Parameters involved in the computations include; TP , FP and FN which denote the number of true positives (number of detected spots that corresponds to the ground-truth), number of false positives (number of detected spots which do not correspond to the ground-truth) and number of false negatives (number of missed ground-truth spots), respectively. A detection result is labelled as TP if the overlap region, O_r between the detection and ground-truth exceeds a predefined threshold T ,

$$O_r = \frac{\text{area}(\text{detection} \cap \text{groundtruth})}{\text{area}(\text{detection} \cup \text{groundtruth})} > T. \quad (44)$$

Otherwise, the detection will be classified as a false positive.

Precision measures the fraction of correctly detected spots among all ground truth spots while recall measures the fraction of correctly detected spots. The precision and recall are computed as,

$$\text{Precision} = TP / (TP + FP), \quad (45)$$

$$\text{Recall} = TP / (TP + FN). \quad (46)$$

The F-measure combines both precision and recall into a single measure weighted by the factor, α^2

$$F_\alpha = (1 + \alpha^2) \times \frac{\text{Precision} \times \text{Recall}}{(\alpha^2 \times \text{Precision}) + \text{Recall}}. \quad (47)$$

The parameter, α^2 is set depending on the importance of precision or recall; thus if $\alpha^2 < 1$ then precision is of importance and vice versa. If neither of precision or recall is preferred then $\alpha^2 = 1$. The above three measures have been successfully used in the comparison of various spot detection methods (Smal, 2010).

CONCLUSIONS

Detection of spots in microscopy images has always been a fundamental but a challenging issue in the field of automated image analysis. Over the past years, many efforts have been made to develop various automated spot detection method. In this work, a review of existing detection methods has been presented. The reviewed methods included both supervised and unsupervised detection methods. We categorized the detection methods into two groups, unsupervised and supervised based detection methods and reviewed them exhaustively. We also summarized publicly available data sets, tools, and three standard evaluation metrics.

Even though a number of automated methods have

been developed, researchers in the field of bio-image analysis are still faced with many challenges which need improving for future methods.

- Dealing with the effect of noise on data:
- Accurate spot detection: the images obtained via fluorescence microscopy imaging may be blurred and noisy which makes it difficult to analyse them. The presence of noise in these images limits the performance of any automated spot detection methods.

This review has highlighted the increasing importance of automated based spot detection algorithms to overcome the challenges of manual detection. The aim of this study was not to determine the overall performance of the methods but rather to present a comprehensive catalogue of spot detection methods in microscopy images. All the reviewed methods contain parameters that need to be tuned for better performance. The fewer the parameters, the simpler the method is to implement. However, the choice of the method should be based on the type of input images, and different methods need to be tested on the same dataset to evaluate which method performs best. Recent advances in machine learning namely, deep learning, has showed remarkable results within the task of image classification. As part of future work we propose the use of convolutional neural network (convnet) (LeCun, et al., 2015) for the task of spot detection in microscopy images. Preliminary results using conv-nets for spot detection showed good performance but still requires some improvements.

ACKNOWLEDGEMENTS

This work was carried out with the financial support of the Council for Scientific and Industrial Research (CSIR) and the National Research Foundation (NRF), South Africa.

REFERENCES

- Basset A, Boulanger J, Salamero J, Bouthemy P, Kervrann C (2015). Adaptive spot detection with optimal scale selection in fluorescence microscopy images. *IEEE Transactions on Image Processing* 24(11):4512–27.
- Boulanger J, Kervrann C, Bouthemy P, Elbau P, Sibarita JB, Salamero J (2010). Patch-based nonlocal functional for denoising fluorescence microscopy image sequences. *IEEE Transactions on Medical Imaging* 29(2): 442–54.
- Boykov Y, Kolmogorov V (2004). An experimental comparison of min-cut/max-flow algorithms for energy minimization in vision. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26(9):1124–37.
- Byun J, Verardo MR, Sumengen B, Lewis GP, Manjunath BS, Fisher SK (2006). Automated tool for the detection of cell nuclei in digital microscopic images: Application to retinal images. *Molecular Vision* 12(1): 949–60.
- Chenouard N (2016). Particle tracking benchmark generator. [Online] Available at: http://icy.bioimage-analysis.org/plugin/Particle_tracking_benchmark_generator [Accessed 1 September 2016].
- Chenouard N, Smal I, de Chaumont F, Maška M, Sbalzarini IF, Gong Y, (2014). Objective comparison of particle tracking methods. *Nature Methods* 11(3):281–90.
- Dabov K, Foi A, Katkovnik V, Egiazarian K (2007). Image denoising by sparse 3D transform-domain collaborative filtering. *IEEE Transactions on Image Processing* 16:1–16.
- Fukunaga K, Hostetler LD (1975). The estimation of the gradient of a density function, with applications in pattern recognition. *IEEE Transactions on Information Theory* 2(1):32–40.
- Genovesio A, Liedl T, Emiliani V, Parak WJ, Coppey-Moisand M, Olivo-Marin JC (2006). Multiple particle tracking in 3d+t microscopy: Method and application to the tracking of endocytosed quantum dots. *IEEE Transactions on Image Processing* 15(5):1062–70.
- Gerig G, Kubler O, Kikinis R, Jolesz FA (1992). Nonlinear anisotropic filtering of MRI data. *IEEE Transactions on Medical Imaging*, 11(2):221–32.
- Gue M, Messaoudi C, Sun JS, Boudier T (2005). Smart 3d-fish: automation of distance analysis in nuclei of interphase cells by image processing. *Cytometry Part A* 67A(1):18–26.
- Jackson C, Glory-Afshar E, Murphy RF, Kovačević J (2011). Model building and intelligent acquisition with application to protein subcellular location classification. *Bioinformatics* 27(13):1854–9.
- Jaiswal A, Godinez WJ, Eils R, Lehmann MJ, Rohr K (2015). Tracking virus particles in fluorescence microscopy images using multi-scale detection and multi-frame association. *IEEE Transactions on Image Processing* 24(11):4122–36.
- Jiajun D, Zhongtian C, Xiongxiang H, Yizhao Z (2016). Clustering by finding density peaks based on Chebyshev's inequality. In: *Proceedings of the 35th Chinese Control Conference* 7169–72.
- Jiang, S, Zhou X, Kirchhausen T, Wong SC (2007). Detection of molecular particles in live cells via machine learning. *Cytometry Part A* 71A(8):563–75.
- Kervrann C, Sorzano CÓS, Acton ST, Olivo-Marin JC, Unser Michael (2016). A guided tour of selected image processing and analysis methods for fluorescence and electron microscopy. *IEEE Journal of Selected Topics in Signal Processing* 10(1):6–30.
- Kimori Y, Baba, N, Morone N (2010). Extended morphological processing: a practical method for automatic spot detection of biological markers from microscopic images. *BMC Bioinformatics* 11(373):1–13.

- Kleinberg E (1996). An overtraining-resistant stochastic modelling method for pattern recognition. *Annals of Statistics* 24(6):2319–49.
- Kozubek M, Kozubek S, Lukášová E, Marecková A, Bártošová E, Skalníková M, Jergová A (1999). High-resolution cytometry of fish dots in interphase cell nuclei. *Cytometry* 36(4):279–93.
- LeCun Y, Bengio Y, Hinton G (2015). Deep learning. *Nature* 521(7553):436–44.
- Lehmussola A, Ruusuvoori P, Selinummi J, Huttunen H, Yli-Harja O (2007). Computational framework for simulating fluorescence microscopy images with cell population. *IEEE Transactions on Medical Imaging* 26(7):1010–6.
- Lindeberg T (1998). Feature detection with automatic scale selection. *International Journal of Computer Vision* 30(2):79–116.
- Mabaso M, Withey D, Twala B (2016). A framework for creating realistic synthetic fluorescence microscopy image sequences. In: *Proceedings of the 3rd International Conference on Bioimaging* 85–92.
- Marquardt DW (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial and Applied Mathematics* 11(2):431–41.
- Matula P, Verissimo F, Wörz S, Eils R, Pepperkok R, Rohr K (2010). Quantification of fluorescence spots in time series of 3-D confocal microscopy images of endoplasmic reticulum exit sites based on the HMAX transform. In: *proc SPIE* 7626:1–7
- McLachlan GJ (2004). *Discriminant analysis and statistical pattern recognition*. John Wiley & Sons.
- Netten H, Young IT, van Vliet LJ, Tanke HJ, Vrolijk H, Sloos WC Netten (1997). Fish and chips: automated of fluorescent dot counting in interphase cell nuclei. *Cytometry* 28(1):1–10.
- Olivo-Marin J-C (2002). Extraction of spots in biological images using multiscale products. *Pattern Recognition* 35(9):1989–96.
- Otsu N (1979). A threshold selection method from gray-level histograms. *IEEE Transactions on Systems, Man, and Cybernetics* 9(1):62–6.
- Pécot T, Bouthemy P, Boulanger J, Chessel A, Bardin S, Salamero J, (2015). Background fluorescence estimation and vesicle segmentation in live cell imaging with conditional random fields. *IEEE Transactions on Image Processing* 24(2):667–80.
- Raj A (2016). Raj laboratory for system biology. [Online] Available at: <http://rajlab.seas.upenn.edu>. [Accessed 20 August 2016].
- Raj A, van den Bogaard P, Rifkin SA, van Oudenaarden A, Tyagi S (2008). Imaging individual mRNA molecules using multiple singly labeled probes. *Nature Methods* 5(10):877–9.
- Ram S, Rodriguez JJ, Bosco G (2012). Segmentation and detection of fluorescence 3d spots. *Cytometry*, 81A: 198–212.
- Rezatofighi SH, Gould S, Vo BT, Vo BN, Mele K, Hartley R (2015). Multi-target tracking with time-varying clutter rate and detection profile: application to time-lapse cell microscopy sequences. *IEEE Transactions on Medical Imaging* 34(6):1–14.
- Rezatofighi SH, Hartley R, Hughes WE (2012). A new approach for spot detection in total internal reflection fluorescence microscopy. In: *Proceedings of the 9th IEEE Int Symp on Biomedical Imaging (ISBI)* 860–3.
- Ruusuvuori P, Aijö T, Chowdhury S, Garmendia-Torres C, Selinummi J, Birbaumer M, (2010). Evaluation of methods for detection of fluorescence labeled subcellular objects in microscope images. *BMC Bioinformatics* 11:1–17.
- Ruusuvuori P, Lehmussola A, Selinummi J, Rajala T, Huttunen H, Yli-Harja O (2008). Benchmark set of synthetic images for validating cell image analysis algorithms. In: *Proceedings of the 16th European Signal Processing Conference (EUSIPCO)*, 1–5
- Ruusuvuori P, Manninen T, Huttunen H (2012). Image segmentation using sparse logistic regression with spatial prior. In: *Proceedings of the 16th European Signal Processing Conference (EUSIPCO)*, 2253–7
- Sage D, Neumann FR, Hediger F, Gasser SM, Unser M (2005). Automatic Tracking of Individual Fluorescence Particles: Application to the Study of Chromosome Dynamics. *IEEE Transactions on Image Processing* 14(9):1372–83.
- Sbalzarini IF, Koumoutsakos P (2005). Feature Point Tracking and Trajectory Analysis for Video Imaging in Cell Biology. *Journal of Structural Biology* 151(2): 182–95.
- Smal I (2009). [Online] Available at: <http://smal.ws/wp/software/synthetic-data-generator/>. [Accessed 8 November 2016].
- Smal I, Loog M, Niessen W, Meijering E (2010). Quantitative comparison of spot detection methods in fluorescence microscopy. *IEEE Transactions on Medical Imaging* 29(2):282–301.
- Smal I, Meijering E, Draegestein K, Galjart N, Grigoriev I, Akhmanova A, (2008). Multiple object tracking in molecular bioimaging by Rao-Blackwellized marginal particle filtering. *Medical Image Analysis* 12(6):764–77.
- Tanaka K, Inoue J-I, Titterton DM (2003). Probabilistic image processing by means of the Bethe approximation for the Q-Ising model. *Journal of Physics*, 36:1–15.
- Tibshirani R (1994). Regression shrinkage and selection via the Lasso. *Journal of the Royal Statistical Society, Series B*, 58:267–88.
- Velasco DF (1980). Thresholding using the ISODATA clustering algorithm. *IEEE Transactions on Systems, Man, and Cybernetics* 10:771–4.
- Vincent L (1993). Morphological grayscale reconstruction in image analysis: Applications and efficient algorithms. *IEEE Transactions on Image Processing* 2:176–201.

- Viola P, Jones M (2001). Rapid object detection using a boosted cascade of simple features. In: Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR), 511–8.
- Wilson RS, Yang L, Dun A, Smyth AM, Duncan RR, Rickman C, (2016). Automated single particle detection and tracking for large microscopy datasets. *Royal Society Open Science* 3(5):1–13.
- Worz S, Sander P, Pfannmoller M, Rieker RJ, Joos S, Mechtersheimer G, (2010). 3D geometry-based quantification of colocalizations in multichannel 3d microscopy images of human soft tissue tumors. *IEEE Transactions on Medical Imaging* 29(8):1474–84.
- Yang L, Parton R, Ball G, Qiu Z, Greenaway AH, Davis I, Lu W (2010). An adaptive non-local means filter for denoising live-cell images and improving particle detection. *Journal of Structural Biology* 172(3):233–43.