A novel generic dictionary-based denoising method for improving noisy and densely packed nuclei segmentation in 3D time-lapse fluorescence microscopy images

Lamees Nasser[1,2*] **and Thomas Boudier**[1,3]

[1]Sorbonne Université, UPMC Univ Paris 06, UJF, CNRS, IMT, NUS, Image and Pervasive Access Lab (IPAL), 138632, Singapore.
[2]BioInformatics Institute (BII), Agency for Science, Technology and Research (A*STAR), Singapore 138671, Singapore
[3]Walter and Eliza Hall Institute of Medical Research, Melbourne, Australia
[*]lameesm@bii.a-star.edu.sg

ABSTRACT

Time-lapse fluorescence microscopy is an essential technique for quantifying various characteristics of cellular processes, i.e. cell survival, migration, and differentiation. To perform high-throughput quantification of cellular processes, nuclei segmentation and tracking should be performed in an automated manner. Nevertheless, nuclei segmentation and tracking are challenging tasks due to embedded noise, intensity inhomogeneity, shape variation as well as a weak boundary of nuclei. Several nuclei segmentation approaches have been reported in the literature. However, most of them require fine-tuning of many parameters. Furthermore, they can only be applied to particular models or acquisition systems. As a result, there is a crucial need for a generic method to segment noisy and densely packed nuclei in microscopy images. In this paper, we present a sparse representation method for denoising of microscopy images. The proposed method works properly in the context of most challenging cases. In addition, it provides a denoised image and simultaneously the potential locations of nuclei. To evaluate the performance of the proposed method, we tested our method on two datasets from cell tracking challenge. Across all datasets, the proposed method achieved satisfactory results with 96.96% recall for *C.elegans* dataset. Besides, in *Drosophila* dataset, our method achieved very high recall (99.3%).

Introduction

In cellular and molecular biology research, automatic segmentation and tracking of biological structures, i.e. cells or their nuclei are important tasks for further understanding of cellular processes. Time-lapse fluorescence microscopy (TLFM) is one of the most appreciated imaging techniques which can be used to quantify various characteristics of cellular processes, such as cell survival[1], proliferation[2], migration[3], and differentiation[4]. The quantification of these processes plays a significant role in studying embryogenesis, cancer cells, stem cells, and other applications in the fields of molecular and developmental biology.

In TLFM imaging, not only spatial information is acquired but also temporal information as well as spectral information, that produces up to five-dimensional images. Typically, the generated datasets consist of several (hundreds or thousands) images, each containing hundreds to thousands of objects to be analyzed[5]. These large volumes of data cannot easily be parsed and processed, via visual inspection or manual processing within any reasonable time.

Nowadays, there is a growing consensus that automated cell segmentation methods are necessary to manage time issue and provide a level of reliability and validity. Accordingly, the implementation of automated high-throughput cell nuclei detection and segmentation techniques may be able to improve the clinical diagnosis, predict the treatment outcome, and help to enhance therapy planning.

Although, TLFM imaging is a very powerful technique and can capture valuable information from cellular structures, it still poses several limitations that can be summarized as follows: 1) non-uniform background illumination because of the fluorescence in cytoplasm and mounting medium; 2) low contrast and weak boundaries of non-obvious nuclei; 3) degradation of image intensity over time due to photobleaching of fluorophores[5]. As a result of these limitations, obtained images become very noisy and difficult to interpret, which might lead to false detection and false segmentation results.

Over the last few years, several methods have been proposed for filtering and segmenting of cell nuclei microscopy images. Regarding the existing filtering literature, Darbon *et al.*[6] have used a Non-Local Mean approach to reduce the noise in

microscopy images. This method relied on replacing the intensity value of a pixel with the average intensity values of the most similar pixels in the image. The algorithm demonstrated good performance in enhancing particles contrast and reducing the noise in electron cryomicroscopy images. For Transmission Electron Microscopy (TEM) images, several digital filters have been introduced by Kushwaha *et al.*[7] such as median or Wiener filter. A similar work proposed by Sim *et al.*[8] and Aguirre[9] based on employing an adaptive Wiener filter to enhance the effectiveness of classical Wiener filter by considering variance of the noise. Luisier *et al.*[10] have suggested a *Poisson Unbiased Risk Estimation-Linear Expansion of Thresholds (PURE-LET)* technique for denoising images corrupted with Poisson noise. The method is based on three criteria: 1) minimizing of an unbiased estimate of Mean Square Error (MSE) for Poisson noise, 2) linear parametrization of the denoising process 3) preserving of Poisson statistics across scales. This algorithm had a minimum time complexity and limited system requirements. Moreover, it is particularly promising in case of large datasets and images with low signal to noise ratio.

For cell nuclei segmentation literature, the reported approaches are classified into two categories: simple approaches such as thresholding method[11–14], edge detection[15] and shape matching[16–18] and more sophisticated approaches like region growing[19–22], energy minimization[23] and machine learning[24–27].

In the first category, R. Bise *et al.*[11], Arteta *et al.*[12], Liao *et al.*[13] and Gul-Mohammed *et al.*[14] applied a thresholding-based approach to segment cell nuclei. This approach assumes that the cell is usually brighter than its surrounding areas and there often exists an optimal threshold where individual cells can be segmented as separate objects. This assumption is not applicable in the challenging regions, because it is impossible to find a suitable threshold to separate all touching cells. Wählbyedge *et al.*[15] suggested to use an edge detection approach, in which an edge filter is applied to the image and therefore pixels are classified as edge or non-edge. These edges are usually detected by the first or second order derivatives method. Nevertheless, this method fails to detect non-obvious cells boundaries. Cicconet *et al.*[16] and Türetken et al.[17] proposed to use a shape matching-based approach. This approach depends on the assumption that cells, particularly nuclei, have round shapes. Hence, multi-scale blob detection[28] can be employed to detect and segment cells.

In the second category, Cliffe *et al*[19], Liu *et al.*[20], Tonti *et al.*[21] and Gul-Mohammed and Boudier[22] employed region-based segmentation techniques in which the basic idea of these approaches is to combine the neighboring pixels of initial seed points which have similar properties to form individual cells.

On the other hand, machine learning-based approaches can be implemented for cell segmentation. To give some examples, Ronneberger *et al.*[24] proposed U-Net convolutional networks to segment cells by assigning class labels to every pixel in the image. A Convolutional neural network presented by Akram *et al*[25] to predict a regression on the cell bounding box and to retrieve the proposal from a cell gallery by representing each image patch with a descriptor. Liu *et al.*[26] reported the use of a convolutional encoder-decoder network to segment the cell nuclei by learning stochastic characteristics of noise and shape of nuclei. Thus, segmented cell nuclei images can be generated. Moreover, Sadanandan *et al*[27] advised to use the Deep Convolutional Neural Networks (DCNNs) to segment cells in fluorescence microscopy images. The aforementioned filtering and segmentation methods have achieved good results in microscopy images. However, most of these methods required a predefined set of parameters. Moreover, they may be used effectively only for specific applications. As a result, identifying a proper and robust approach for various datasets regarding high variation in cell nuclei volume, shape, and stain distribution as well as high cell density has become a major challenge in image analysis.

The contribution of this paper is to design a generic method for denoising of 3D cell nuclei images based on a sparse representation model. Consequently, a classical segmentation method can be used to segment cell nuclei without the need for more complicated segmentation approaches. The proposed model is able to handle the inhomogeneous intensity, shape variation and weak boundary of cell nuclei. Additionally, it can provide the denoised image and simultaneously a sparse representations map, that indicates the potential locations of cell nuclei. Afterward, the denoised image and the sparse representations map will be used for the cell nuclei detection and segmentation process.

Results

Datasets description

Since the main motivation of our work is to automate the detection and segmentation of cell nuclei in the developmental biology of *C.elegans* where an only limited number of existing methods had provided satisfactory results, the proposed framework is extensively tested on three real datasets for embryonic cells and one dataset of synthetic images with different values for the noise variance and object size.

Synthetic dataset.

In order to measure the robustness of the proposed method, we generated a synthetic images of size (XYZ) equal $(100 \times 100 \times 20)$ voxels containing spheres of two radii: 7 and 9 voxels. As it is common in fluorescence microscopy images to have low contrast and low signal to noise ratio (SNR) as a consequence of weak fluorescent staining or microscope properties. The images are distorted with different levels of Poisson-Gaussian noise, resulting in SNRs of 2 dB, -1 dB, -5 dB and -7 dB, respectively.

Furthermore, the images include touching spheres where these conditions simulate the same characteristics exist in the real datasets as shown in Supplementary Fig 1.

Real dataset.

The first dataset comes from the work of Gul-Mohammed[14] and it is named as *CE-UPMC*. The other two datasets come from the cell tracking challenge[29,30], and namely *Fluo-N3DH-CE* and *Fluo-N3DH-DRO*. The last two datasets are proven to be the hardest to be fully segmented automatically[29]. Each dataset from cell tracking challenge contains 2 sequences. For all datasets, cell nuclei are manually annotated by a specialist and hence taken as ground truth (GT). The datasets are described as follows:

CE-UPMC dataset. It involves *C. elegans* embryonic cells. The size (XYZT) of dataset is $512 \times 512 \times 31 \times 160$. The cells were acquired every 1 minute using a spinning-disk confocal microscope. This dataset is very challenging as images intensity were decay over time because of the labeling technique and acquisition system, thus the acquired images quality were low.

Fluo-N3DH-CE dataset. It includes *C. elegans* embryonic cells. The size (XYZT) of the first sequence is $708 \times 512 \times 35 \times 250$ and of the second sequence is $712 \times 512 \times 31 \times 250$. Both sequences are $8-$bit images with cells imaged every 1.5 minutes. The cells are acquired using a Zeiss LSM 510 Meta Confocal Microscope. This dataset is a challenging as well since it has low signal to noise ratio (SNR = 6.74 dB) in addition the fluorescence can fade when the cells divide. Furthermore, the cells become smaller and smaller over time.

Fluo-N3DH-DRO dataset. It contains *Drosophila melanogaster* embryonic cells. The size (XYZT) of each sequence is $1272 \times 603 \times 125 \times 49$. Both sequences are $8-$bit images with cells imaged every 30 second. The cells are acquired using SIMView light-sheet microscope. This dataset is very challenging as it has a large number of densely packed cells in addition to has low signal to noise (SNR = 2.46 dB).

Experimental setup and suitable parameters selection.

To study the effect of parameters selection on cell nuclei detection and segmentation results, the proposed method is extensively analyzed on synthetic dataset. Our approach is based on the building of a dictionary (small patches of the image) that will be eventually used for denoising and detection of cell nuclei. We investigated the optimal size of the patches and the number of patches (called atoms) in the dictionary. Then, the randomly created dictionary is updated, and so we investigated the number of iteration for the update. Finally, we investigated the sparsity level (i.e., the number of used atoms) for the reconstruction of the denoised image and sparse representation map. To start with, we tested several values for patch size ($p = 5 \times 5 \times 5$, $10 \times 10 \times 5$, $15 \times 15 \times 5$ and $20 \times 20 \times 5$), dictionary size ($K = 64$, 128, 256 and 512), sparsity level ($L = 3$, 6 and 9) and number of iterations ($N = 5$, 10, 15, 20 , 25 and 30) at different noise levels (SNR $= 2$, -1, -5 and -7 dB). Since the coefficient of variation (CV) is a useful statistic for comparing the degree of variation from one data series to another one. The CV is defined as the ratio of the standard deviation to the mean[31]. Thus, we employed the CV to measure the effect of changing the parameters on the result. The average CV from the patch size, dictionary size, sparsity level and the number of iterations over the four noise levels are approximately 15, 2, 2 and 2 % respectively.

In all aforementioned parameters, we observed that the patch size is considered as the dominant parameter where the change in this value has a major impact on the segmentation results as shown in Supplementary Fig 2 (a). On the contrary, changes in the other parameters i.e dictionary size, sparsity level and the number of iterations often achieve very close results as shown in Supplementary Fig 2 (b), (c) and (d). Therefore, we fixed all parameters while tuned the patch size according to the object's size present in the images.

In order to confirm the importance of patch size tuning, we conducted more analysis in term of cell nuclei detection as shown in Supplementary Fig 3. For example, at the first three noise levels 2, -1 and -5 dB all cell nuclei are correctly detected for the different patch size values. However, at noise level equal -7 dB, many objects are falsely detected with patch size equivalent to $p = 5 \times 5 \times 5$ and one nucleus is not detected at patch size equal $20 \times 20 \times 5$. Though, for patch size equal $10 \times 10 \times 5$ and $15 \times 15 \times 5$, all cell nuclei are correctly detected.

For all previously mentioned patch size values, the average of recall, precision, F-measure, and Jaccard index with different noise levels are presented in Supplementary Fig. 3 (b), where these measures are high when the patch size values equal $p = 10 \times 10 \times 5$ and $15 \times 15 \times 5$ compared to the measures of the other two values .

Following the above experiment, we observed that, for robust detection and segmentation results, the patch size should not be less than 25% and not more than 100 % of the average cell nuclei volume in the images.

Denoising of 3D cell nuclei images.

In this work, a sparse representation model[32,33] is employed to get the denoised images. Our method is compared with the PURE-LET[10], which is one of the most efficient, fast and automatic methods for denoising of multidimensional fluorescence microscopy images. The main motivation behind the need for cell nuclei denoising is assisting better segmentation of cell nuclei images. Therefore, the comparison between the denoising methods is performed in context of improving segmentation results. For instance, the results in Supplementary Fig 4 (first row), Supplementary Fig 5 (a, c) and Supplementary Fig 6 (a, c)

show that our method able to reduce, and almost remove the noise as well as enhance the contrast of cell nuclei. We have also noticed a better contrast than PURE-LET results as shown in Supplementary Fig 4 (third row), Supplementary Fig 5 (b, d) and Supplementary Fig 6 (b, d).

For further assessment, thresholding based approach is applied to the denoised images to obtain the initial segmentation results. It can be noted from Supplementary Fig 4 (second row), Supplementary Fig 5 (e) and Supplementary Fig 6 (e) that our method succeeded to segment all nuclei in comparison with the other method which failed to segment some nuclei as demonstrated in Supplementary Fig 6 (f). Even though, PURE-LET method able to segment all cell nuclei shown in Supplementary Fig 5 (f), the size of segmented nuclei are smaller than their original size. Unfortunately, this method can not segment any cell nuclei at very low signal to noise ratios (-5 dB and -7 dB,) as presented in Supplementary Fig 4 (fourth row)

Segmentation of 3D cell nuclei images.

Following the denoising step, a local adaptive thresholding[34] is applied to the denoised image to get an initial segmentation (Fig. 2 (c, d)) of candidate regions. In order to obtain candidates location of cell nuclei centers, we used a novel representation called the sparse representation map. Each voxel in this representation map is given by the summation of its patch coefficients. Thus, the maximum response image obtained by combining the denoised image and sparse representation map is used to detect the local maxima (Fig. 1). Afterward, the obtained local maxima are employed as an input for a 3D marker-controlled watershed segmentation of the cell nuclei (Fig. 2). The number of detected marker points represents the total objects that are extracted by the watershed algorithm.

In the synthetic dataset, a great performance is observed on very low signal to noise ratios (2 dB, -1 dB, -5 dB and -7 dB), in which our method is capable of correctly identify and segment all cell nuclei at the various noise levels as presented in Supplementary Fig 7 .

Regarding the *CE-UPMC*, there is an intensity decay over time owing to the labeling technique and acquisition system. As a result, the acquired image quality is low. Table 3 and Supplementary Fig 8, illustrate the results obtained at certain time points. For instance, at early time points (40, 60, 80 and 100) all cell nuclei are correctly detected. In addition, a small number of false positive are also detected (2 objects out of 119 cell nuclei). Even though the image quality at advanced time points (120, 140 and 160) is low, only 11 cell nuclei out of 247 are not detected, also there exists a small number of false positives (3 objects out of 247 cell nuclei which is displayed as yellow arrows in Supplementary Fig 8 (f). Typically, the reason for missing cell nuclei is the detection of clustered cell nuclei (indicated by red arrows in Supplementary Fig 8 (f)) rather than detecting them separately. Supplementary Fig 9 shows the segmentation results of our method and the results from the original paper[14] of *CE-UPMC* dataset.

In the *Fluo-N3DH-CE*, the proposed approach is able to identify and segment correctly more than 96% of total cell nuclei. Furthermore, it detects a small number of false positives (9 objects out of 876 cell nuclei) as well as a small number of false negative (29 cell nuclei). The achieved F-measure of approximately 97.8%, which is comparable or better than the competing algorithm, i.e., KTH[29,30] (Table 2, Supplementary Fig 10 and Supplementary Fig 11).

For *Fluo-N3DH-DRO*, despite our method succeeds to detect more cell nuclei (99%recall), it has low precision (3%) due to the annotated ground truth, which considered only the cell nuclei located in the early nervous system and all other nuclei are deemed as false positives. As a result, the obtained F-measure is low with an approximate value of 6.4%. In addition, it has low segmentation accuracy compared to the KTH competing approach[29,30] (Table 2, Supplementary Fig 12 and Supplementary Fig 13).

The results achieved by our method for the two datasets obtained from cell tracking challenge are compared with the top-ranked KTH algorithm[29,30]. KTH algorithm is chosen for the reason that, it presented the best overall performance in the challenge. This algorithm is mainly based on adopting the band-pass filter to detect and segment cell nuclei.

For KTH algorithm, some detected objects are actually noise and some cell nuclei are not detected. This is because it detected clustered cell nuclei instead of detecting them separately. For example, at time point 28 from *Fluo-N3DH-CE* (seq1) and at time point 106 from *Fluo-N3DH-CE* (seq2), KTH algorithm failed to resolve the fusion of two nuclei (as presented by the red arrows in Supplementary Fig 11 (b). In contrast, our method succeeds to identify and segment each nucleus individually as shown in Supplementary Fig 11 (a) and (c). For *Fluo-N3DH-DRO*, although our method succeeds to detect more cell nuclei than KTH approach (Table 2 and Supplementary Fig 13), the evaluation method considered those cell nuclei as false positives, due to the annotation method.

The range of parameters utilized for the sparse representation model and the segmentation are the fundamental parameters that required to be tuned for denoising and segmentation of cell nuclei. For all the above-mentioned datasets, the set of empirically determined parameter values being applied to the datasets are listed in Table 1. All other parameters are being fixed for different datasets and experiments with the subsequent values number of atoms (K) = 64, sparsity level (L) = 3 and number of iterations (N) = 15.

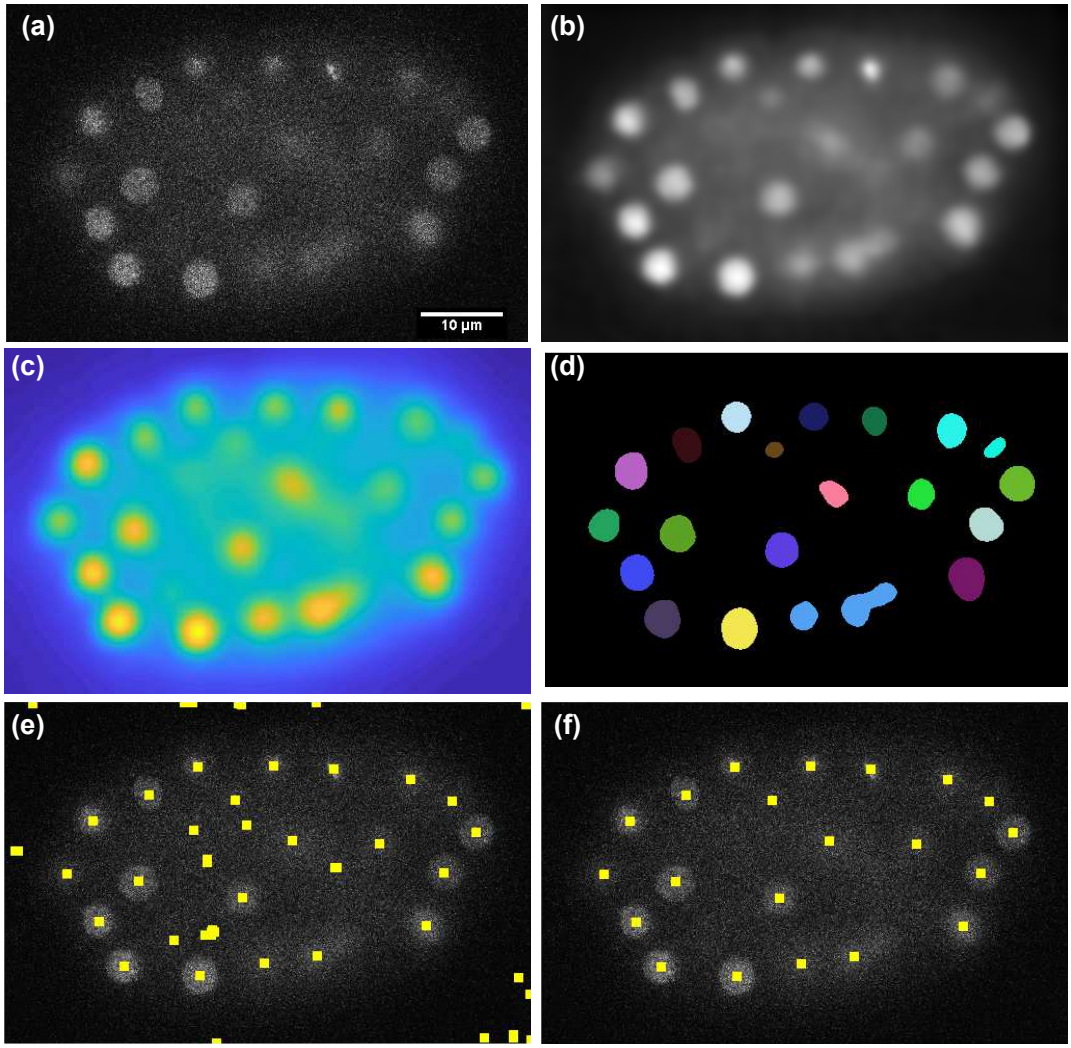

Figure 1. Marker point detection with sparse representation model. (a) A single plane (Z = 15) of time point (T = 100) from *CE-UPMC*. (b) The denoised image obtained by applying the sparse representation model to the image in (a). (c) Sparse representation map resulted from applying the sparse representation model to the image in (a). (d) The obtained cell nuclei after initial segmentation. (e) Marker points detected from image(d) by the local maxima search displayed as yellow squares overlaid on the raw image. (f) Detection's in the background region are discarded by multiplying the detected maxima image with the initial segmentation result. Note that, the marker point detection here is performed in two dimensions for the purpose of explanation and visualization. However, in the framework it is applied in three dimensions.

Table 1. Denoising and segmentation parameters. When the values of parameters differ between the first and the advanced time points, the value for the advanced time points is given in round brackets.

	Datasets				
Parameters	**Synthetic**	***CE-UPMC***	***Fluo-N3DH-CE (seq1)***	***Fluo-N3DH-CE (seq2)***	***Fluo-N3DH-DRO(seq1 & seq2)***
Patch size	[15 15 5]	[20 20 5]	[25 25 5]	[25 25 5]	[10 10 5]
Threshold	global Otsu's[35]	local adaptive[34]	local adaptive[34]	local adaptive[34]	local adaptive[34]
SensitivityFactor	-	0.56 (0.58)	0.5	0.5	0.5
MinNucleiVolume	20	5000 (1000)	10,000 (3000)	10,000 (3000)	-
NucleiDilationShape	sphere	round	round	round	round
NucleiSeedDilation	3	5	10 (5)	20 (5)	5

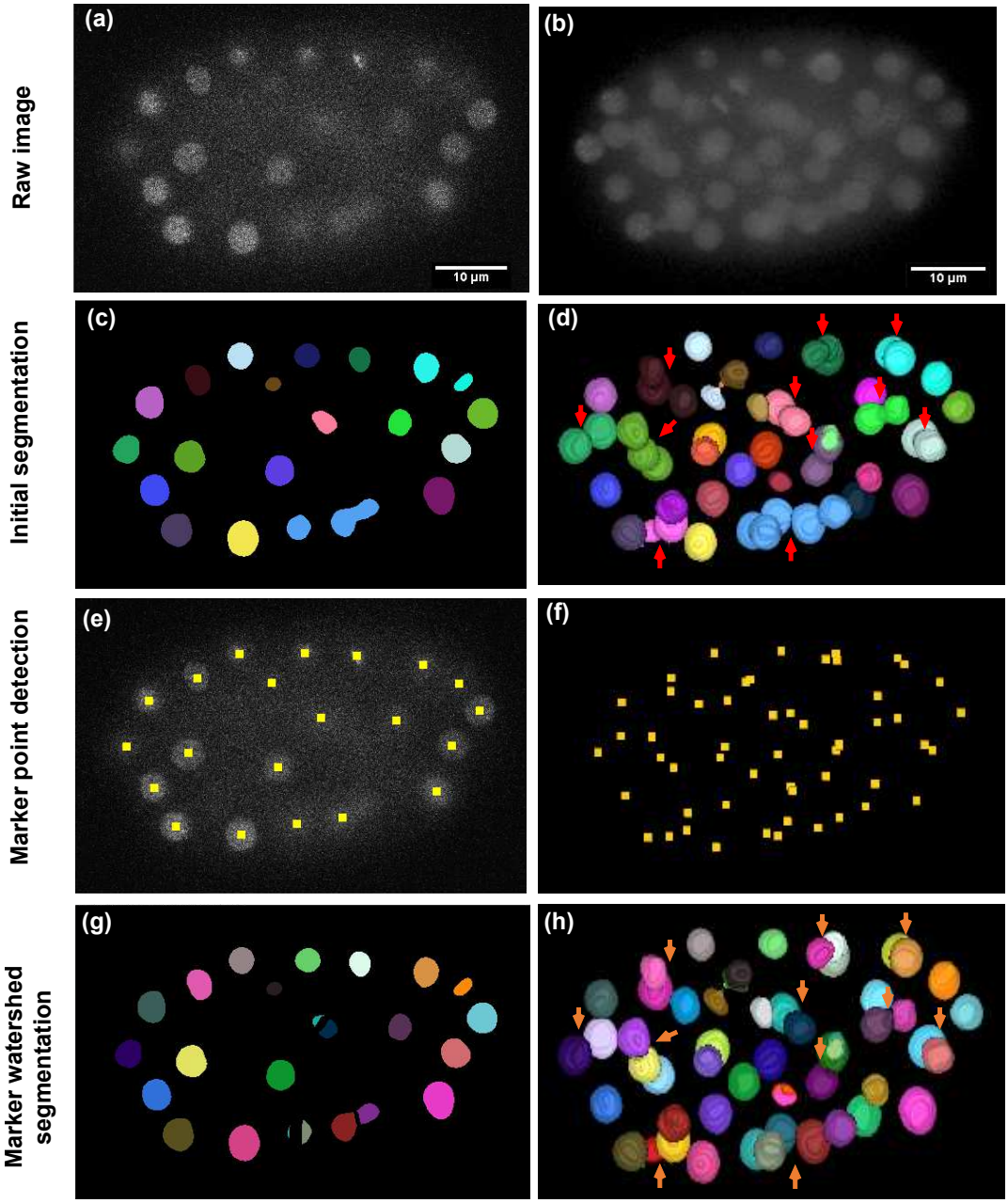

Figure 2. An overview of cell nuclei segmentation steps. First column: shows a single plane (Z = 15) of time point (T = 100) from *CE-UPMC*. Second column: shows a three-dimensional view of the same time point. (a, b) The raw images. (c, d) The initial segmentation, which identifies the cell nuclei (presented as colored) in the image, but fails to separate apparently touching cell nuclei (shown as red arrows). (e, f) Marker points (indicated by yellow squares) are obtained from the sparse representation model. (g, h) Marker-controller watershed segmentation that succeeds to separate apparently touching cell nuclei (orange arrows). Note that, different colors represent individual components. The marker points detection at (e) is performed in two dimensions for illustration process. However, in the framework it is applied in three dimensions.

Methods

This section introduces a novel method for denoising and detection of cell nuclei in 3D TLFM images based on a sparse representation approach[32,33]. The use of sparse signal representation is becoming popular in several fields such as face recognition[36], image denoising[33] and inpainting[37], and image classification[38]. Indeed, natural images represent very sparse data, especially in biology where numerous instances of the same structure, i.e. cell or nucleus, are present in the image.

Table 2. Segmentation performance of our method (SRS) and KTH[29,30], for datasets from cell tracking challenge. The values shown in bold represent the highest performance. GT, number of cell nuclei in the ground truth; SN, number of cell nuclei determined by the segmentation; TP, true positives; FN, false negatives; FP, false positives.

Dataset	Algorithm	GT	SN	TP	FN	FP	Recall (%)	Precision (%)	F-measure (%)	Jaccard (%)
Fluo-N3DH-CE_01	**SRS**	478	464	459	19	5	96.03	98.9	97.44	66
Fluo-N3DH-CE_02		398	392	388	10	4	97.49	98.98	98.23	70
Total		876	856	847	29	9	**96.69**	**98.94**	**97.84**	**68**
Fluo-N3DH-CE_01	**KTH**	478	463	459	19	4	96.03	99.1	97.54	64
Fluo-N3DH-CE_02		398	394	386	12	8	96.98	97.97	97.47	59
Total		876	857	845	31	12	96.5	98.6	97.55	61.5
Fluo-N3DL-DRO_01	**SRS**	3792	122193	3757	35	118436	99.08	3.07	6	60
Fluo-N3DL-DRO_02		4097	120494	4082	15	116412	99.63	3.39	6.56	73
Total		7889	242687	7839	50	234848	**99.37**	3.32	6.43	66.5
Fluo-N3DL-DRO_01	**KTH**	3792	112535	3733	59	108802	98.44	3.32	6.4	62
Fluo-N3DL-DRO_02		4097	104655	4037	60	100618	98.54	3.86	7.4	78
Total		7889	217190	7770	119	209420	98.49	**3.58**	**6.9**	**70**

Table 3. Segmentation performance of our method (SRS) for *CE-UPMC* dataset.

Time	GT	SN	TP	FN	FP	Recall (%)	Precision (%)	F-measure (%)
40	14	14	14	0	0	100	100	100
60	24	24	24	0	0	100	100	100
80	28	29	28	0	1	100	96.55	98.2447
100	51	52	51	0	1	100	98.08	99.03
120	57	53	53	4	0	92.98	100	96.3623
140	91	93	88	3	2	96.70	97.78	97.2370
160	104	101	100	4	1	96.15	99.01	97.5590

Moreover, a dictionary-based approach is usually linked to unsupervised learning approach since the data itself can be used to learn the basis vectors to build a sparse representation matrix.

An overview of the proposed method is shown in Fig. 3 and it can be explained as follows. Firstly, an initial dictionary is constructed from the raw image. This process is performed by selecting random patches from the raw image, in particular patches whose intensity is greater than the average intensities of all extracted patches to assure the presence of cell nuclei patches. Secondly, a technique based on K-clustering with singular value decomposition (K-SVD)[32] is implemented to update and obtain the final dictionary. Thirdly, the updated dictionary is used to reconstruct the denoised image as well as th sparse representation map that will be used for detection of cell nuclei. Fourthly, the maximum response image, which is a combination of the denoised image and sparse representation map is employed to detect the potential location of cell nuclei. Fifthly, a thresholding-based approach[34,35] is proposed to get the initial segmentation result. Finally, a marker-controlled watershed approach[39] is used to obtain the final cell nuclei segmentation result.

An introduction to sparse representation

The idea of sparse representation is to obtain an efficient representation of a signal as a linear combination of few atoms chosen from a dictionary. Given a dictionary $D \in R^{n\times K}$ that contains K atoms as column vectors $d_j \in R^n, j = 1,2,\ldots,K$. The sparse representation problem of a signal $x \in R^n$ can be described as finding the sparsest vector $\alpha \in R^K$ where $X \simeq D\alpha$. The problem can be formulated as an energy optimization problem as follows:

$$\hat{\alpha} = \underset{\alpha}{\operatorname{argmin}} \|x - D\alpha\|_2 \quad s.t. \quad \|\alpha\|_0 \leq L \tag{1}$$

Where x is the signal, α denotes the sparse representation coefficients, $\|\alpha\|_0$ is the L_0 pseudo-norm that counts the number of non-zeros of α and L is a predetermined sparsity threshold.

Solving the previous optimization problem is NP-hard and numerically intractable, thus several methods have been developed to get an approximate solution for this particular problem. The first type of methods uses L_0-norm minimization such as matching pursuit (MP)[40] or orthogonal matching pursuit (OMP)[41]. The second type of method uses L_1-norm for optimization. The objective of L_1-norm is to make the optimization problem convex, which can be addressed efficiently using basis pursuit (BP)[42].

The crucial issue in practical applications is to select the dictionary D. Basically, dictionaries are of two types: (1) fixed dictionaries, (2) adaptive dictionaries. The fixed dictionary as is the case for curvelet, discrete cosine, wavelet, ridgelet, or bandlet which use pre-defined and fixed atoms. This type of dictionary may not ensure a well-defined representation of all

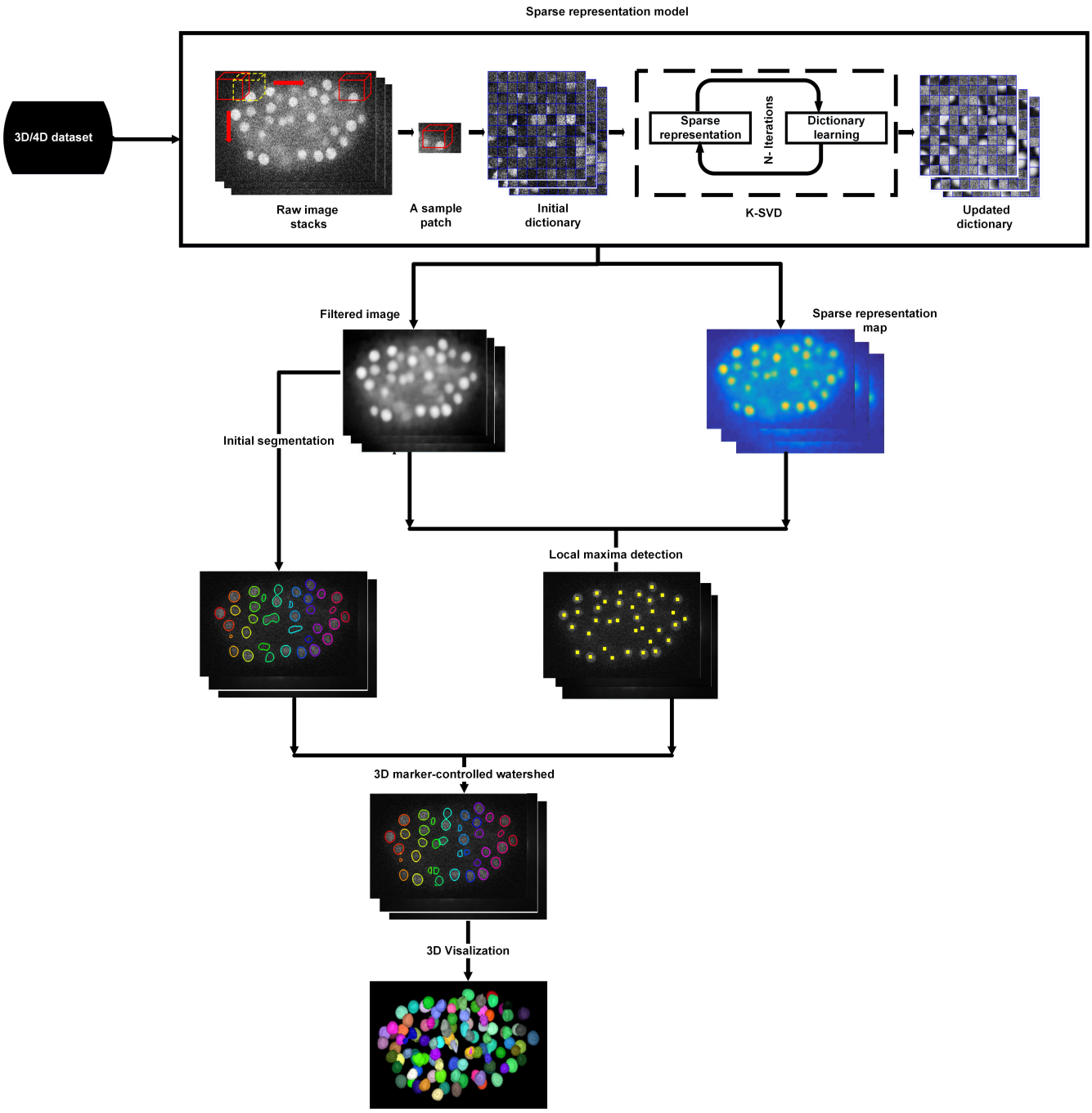

Figure 3. General representation of the proposed framework for denoising and segmentation of cell nuclei in 3D time-lapse fluorescence microscopy images. The proposed pipeline consists of data preprocessing, initial cell nuclei segmentation, cell nuclei detection, final segmentation as well as 3D visualization. In the preprocessing step, an initial dictionary is constructed by selecting random patches from the raw image as well as a K-SVD technique is implemented to update the dictionary and obtain the final one. Then, the maximum response image is employed to detect the potential location of cell nuclei. Furthermore, a thresholding-based approach is proposed to get the initial segmentation result. At last, a marker-controlled watershed approach is used to get the final cell nuclei segmentation result and hence cell nuclei are displayed in a 3D view.

given signals. As a result, it is more appealing to use an adaptive dictionary approach to learn the dictionary directly from the data itself.

Learning the dictionary requires two-step, the first step is to compute an initial dictionary. It is usually computed by taking random patches directly from the raw image. To assure the presence of patches containing nuclei in the initial dictionary beside background patches, we select the patches whose intensity is greater than the average intensities of all patches extracted from the image. The second step is to improve this initial dictionary using an updating method. For updating the initial dictionary, we use the K-SVD algorithm[32]. This algorithm is a standard unsupervised adaptive dictionary learning algorithm that generalizes the well-known K-means clustering approach. It jointly learns a dictionary $D = [d_1, d_2, \ldots, d_K]$ and a related sparse representation matrix $\alpha = [\alpha_1, \alpha_2, \ldots, \alpha_m]$, $\alpha \in R^K$, from a set of training signals $X = [x_1, x_2, \ldots, x_m]$, where each $x_i \in R^n$ by solving the following problem:

$$\underset{D,\alpha}{\operatorname{argmin}} \|X - D\alpha\|_2 \quad s.t. \quad \|\alpha\|_0 \leq L \tag{2}$$

This technique solves the optimization problem by alternating between finding the sparse representation coefficients α and the dictionary D using an iterative approach. Assuming that D is known, the best sparse representation matrix is constructed by solving equation 2 using an orthogonal matching pursuit algorithm (OMP). Following the sparse representation stage, the representation vectors (α) are assumed to be fixed. Subsequently, the best dictionary is computed. Since finding the whole dictionary at the same time is impractical, the dictionary is updated atom by atom. Once the best dictionary and sparse representation coefficients are obtained, the denoised image and sparse representation map can be constructed.

Images with sparse representation.

In this section, we present the reconstruction of the denoised image and sparse representation map which will be used later in the detection and segmentation of cell nuclei

- **Denoised image reconstruction.** The denoised image with dictionary learning is formed by solving the following problem:

$$E(X) = \underset{X,\alpha,D}{\operatorname{argmin}} \ \lambda \ \|X - Y\|_2 + \|X - D\alpha\|_2 \quad s.t. \quad \|\alpha\|_0 \leq L \tag{3}$$

  Where X is the ideal image, Y is the noisy image and λ is the regularization parameter.

- **Sparse representation map reconstruction.** In point of fact, the denoised image does not have sufficient contrast to completely separate touching nuclei. In order to improve cell nuclei detection, a sparse representation map image that indicates the potential locations of cell nuclei will be built. The construction of this image is based on the computation of the sparse coefficients (α_i) of each image patch. It can be obtained by:

$$P_i = \frac{1}{C} \sum_{k=1}^{K} \alpha_i^k \quad i = 1, 2, 3, \ldots, m. \tag{4}$$

  Where, C is a normalization term, P_i is the probability value corresponding to the $i-th$ patch and α_i^k denotes the $k-th$ element of α_i. Notably, voxels within the center of the nucleus have very high sparse coefficients values, in contrast to voxels far away from the center have low values. Consequently, the p_i value of the patches containing nuclei tend to be large compared with the p_i value of background patches. As a result, dictionary learning technique with sparse representation can capture strong structures of biological images as well as restrain the noise.

Cell nuclei segmentation.

- **Initial cell nuclei segmentation.** A local adaptive thresholding approach[34] is applied to the denoised image. The general concept of the algorithm is that every image's voxel threshold is determined by the following equation:

$$\mathrm{T_{local}} = \mathrm{mean_{local}} * (1 - \mathrm{SensitivityFactor}) \tag{5}$$

  Where, $\mathrm{mean_{local}}$ is the mean intensity value in the neighborhood of each voxel and the SensitivityFactor is a scalar value within a range from zero to one which controls sensitivity towards thresholding more voxels as foreground. Accordingly, voxels with intensity values larger than $\mathrm{T_{local}}$ are set to 1, all others are set to 0. Small regions which detected as foreground and those smaller than a predefined volume denoted by (MinNucleiVolume) are discarded. This threshold corresponds to the volume of the smallest cell nucleus and is determined prior to the segmentation step.

- **Marker points detection.** For splitting of touching cell nuclei, we employed the marker-controlled watershed technique. The marker points are obtained as follow: First, the denoised image is multiplied by the sparse representation map to provide a maximum response image. Second, The maximum response image is processed with a maximum transform to detect the local maxima (Fig. 4). The obtained binary image is multiplied by the initial segmentation to discard local maxima detected in the background. Third, a morphological dilation operator with a flat round shape structuring element denoted by (NucleiDilationShape) of certain radius denoted by (NucleiSeedDilation) is employed to avoid detection of multiple local maxima for the same object by merging those maxima that were in close proximity to each other. Finally, the modified image determining the marker points is fed to the subsequent watershed algorithm.

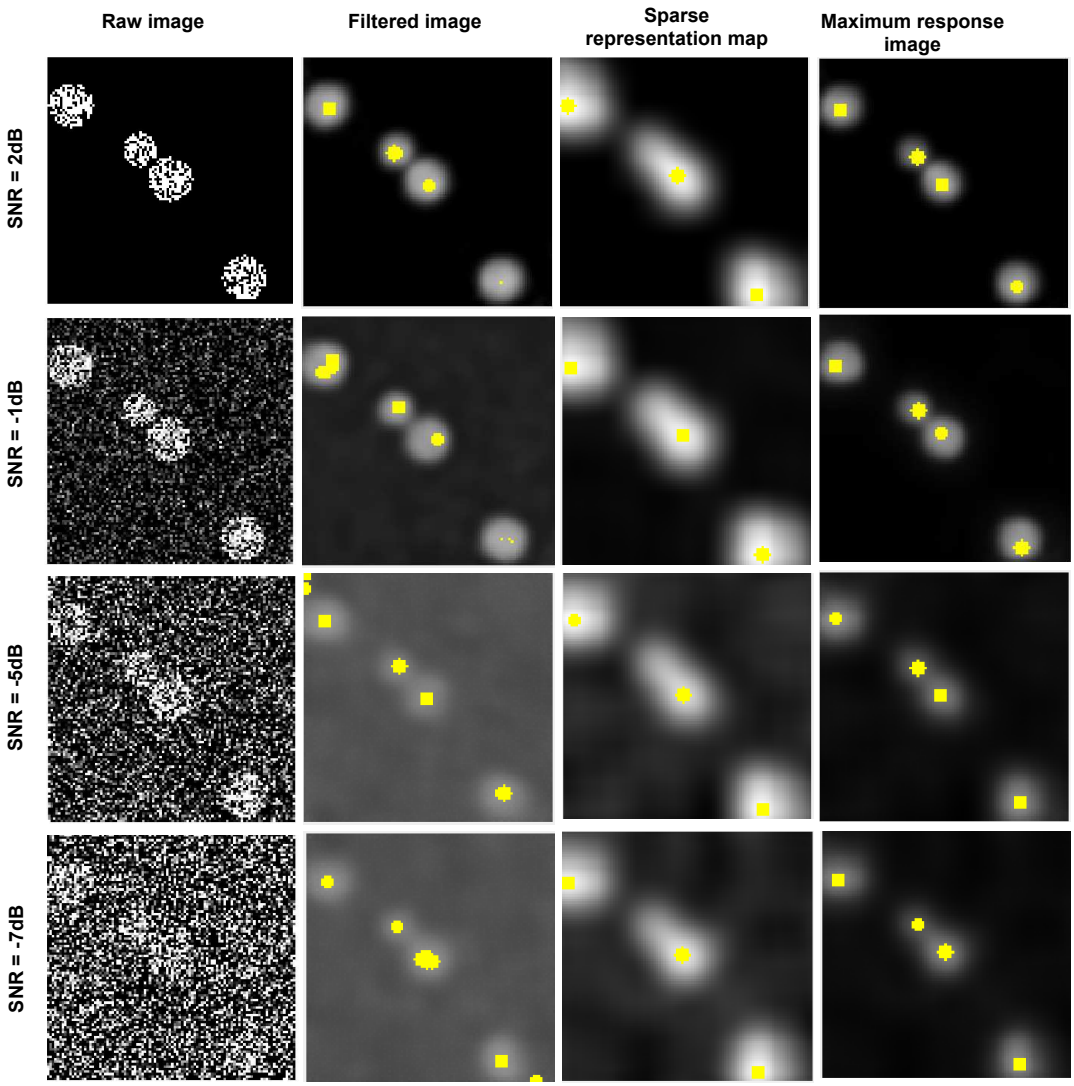

Figure 4. A comparison of marker points detection at various noise level from different images. First column: representative single plane (Z = 10) of the raw image. Second column: the results of marker points detection from the denoised image. Third column: the results of marker points detection from the sparse representation map. Fourth column: the results of marker points detection from the maximum response image. Note that, for all images marker points depicted by yellow markers

- **3D marker-controlled watershed segmentation of overlapped cell nuclei.** Marker-controlled watershed segmentation is presented to separate the connected cell nuclei clusters. The basic principle of watershed approach is to flood the denoised image, which contains the merged objects starting at the marker points as sources. Sometimes the flooding process is not stopped at the border of a cell nucleus, therefore the denoised image is multiplied by the initial segmentation result prior to the flood. Eventually, watersheds are built when different sources meet during the flooding process. This approach allows splitting clusters of apparently touching cell nuclei whereas very densely clustered cells growing on top

of each other cannot be resolved.

Evaluation method and metrics.

To assess the performance of the proposed algorithm, three metrics are employed. The first two metrics are the recall and precision of object detection. The recall is the proportion of the number of relevant detected cell nuclei to the total number of relevant cell nuclei in the ground truth. Precision is the proportion of the number of relevant detected cell nuclei to the total number of irrelevant and relevant detected cell nuclei. These parameters are defined as follows:

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{6}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{7}$$

$$\text{F-measure} = 2 * \frac{\text{Sensitivity} * \text{Precision}}{\text{Sensitivity} + \text{Precision}} \tag{8}$$

Where, True Positive (TP) represents the total number of correctly detected nuclei, False Negative (FN) represents the number of undetected nuclei and False Positive (FP) represents the number of falsely detected nuclei. To compute these values, we used the following steps: First, we calculated the distance between the centroids of ground truth nuclei and centroids of segmented objects. Second, a weight is assigned to each pair of segmented and ground truth objects, equal to the distance between them. Third, Hungarian algorithm is used to solve this assignment problem. Objects with no match to any other object are considered as FP, objects absent in ground truth but they appear in segmentation result are deemed FP, and FN were the objects absent in segmentation result despite these objects appear in ground truth.

The third metric is the Jaccard index that measures the segmentation accuracy of the segmented objects. The Jaccard index for each set of segmented (A) and ground truth (B) objects is defined as the intersection between them divided by their union.

$$J(A,B) = \frac{A \cap B}{A \cup B} \tag{9}$$

The final measure is then the average of the Jaccard indices of the matched pairs.

Implementation details.

The image analysis framework is developed using MATLAB (R2017b) on a Windows-based computer (Intel Core i7, 3.07 GHz, and 16 GB RAM). Furthermore, 3D viewer plugin[43] along with tools from the 3D ImageJ suite[44] is used for three-dimensional visualization of the final segmentation result. The source code is available upon request. Data processing using the complete framework took 5 mins for the synthetic dataset, 11 mins for the *CE-UPMC* dataset, 35 mins for *Fluo-N3DH-CE* dataset and 48 mins for *Fluo-N3DH-DRO* dataset to process only one 3D image from the complete dataset. Regarding cell tracking challenge, the web site (http://www.codesolorzano.com/celltrackingchallenge) provides access to the datasets with the ground truth. In addition, it provides access to Windows and Linux executable files for the evaluation software as well as an executable program includes the process description for KTH work.

Conclusion

In this paper, we have presented a novel generic method for denoising of 3D cell nuclei images based on a sparse representation model. Consequently, straightforward methods can be used to detect and segment cell nuclei in 3D time-lapse fluorescence microscopy images. Over the last few years, deep learning approaches have achieved promising results in several domains. However, they introduce some limitations e.g., they require ground truth labels as well as a huge amount of training data. Furthermore, deep learning approaches are usually implemented to solve a specific problem in which every new dataset will require new training.

In contrary to previous approaches, although our method is also based on a learning technique, the learning is completely unsupervised and hence can be performed for any new data without any change. Moreover, the proposed method can handle the most challenging cases involving noisy, densely packed and multiple touching cell nuclei. In addition, it can produce the denoised image and simultaneously the potential locations of cell nuclei. The proposed method is adapted to the segmentation of cell nuclei in 3D time-lapse fluorescence microscopy images, nevertheless, it can be employed to detect and segment the nearly interacting intracellular organelles, including the endosomes, lysosomes, and lipid droplets. Our method is successfully

evaluated on two embryo models, the *C. elegans*, and the *Drosophila* datasets. The overall detection and segmentation results are comparable or better than the existing methods, which is a good starting point for automated cell nuclei tracking process.

The advantage of the proposed model is that it does not require well-defined many parameters. The sole parameter required is the average of cell nuclei volume present in the image. Furthermore, the training step is usually linked to unsupervised learning approach and the dictionary can be directly learned from the image itself. We believe that no similar studies have been reported in existing literature for denoising and simultaneously predicting objects location in images. As a future work, we will investigate an online learning method to handle the time issue to reduce the processing time needed for the dictionary learning.

References

1. Payne, S. L. *et al.* In vitro maturation of human ipsc-derived neuroepithelial cells influences transplant survival in the stroke-injured rat brain. *Tissue Eng. Part A* **24**, 351–360 (2018).

2. Rapoport, D. H., Becker, T., Mamlouk, A. M., Schicktanz, S. & Kruse, C. A novel validation algorithm allows for automated cell tracking and the extraction of biologically meaningful parameters. *PloS one* **6**, e27315 (2011).

3. Bise, R., Kanade, T., Yin, Z. & Huh, S.-i. Automatic cell tracking applied to analysis of cell migration in wound healing assay. In *Engineering in Medicine and Biology Society, EMBC, 2011 Annual International Conference of the IEEE*, 6174–6179 (IEEE, 2011).

4. Zhang, J., Moradi, E., Somekh, M. G. & Mather, M. L. Label-free, high resolution, multi-modal light microscopy for discrimination of live stem cell differentiation status. *Sci. reports* **8**, 697 (2018).

5. Meijering, E., Smal, I., Dzyubachyk, O. & Olivo-Marin, J.-C. Time-lapse imaging. *Microsc. Image Process.* 401–440 (2008).

6. Darbon, J., Cunha, A., Chan, T. F., Osher, S. & Jensen, G. J. Fast nonlocal filtering applied to electron cryomicroscopy. In *2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro*, 1331–1334 (2008). DOI 10.1109/ISBI.2008.4541250.

7. Kushwaha, H. S., Tanwar, S., Rathore, K. S. & Srivastava, S. Denoising filters for TEM (transmission electron microscopy) image of nanomaterials. In *2012 Second International Conference on Advanced Computing Communication Technologies*, 276–281 (2012). DOI 10.1109/ACCT.2012.41.

8. Sim, K. S., Teh, V. & Nia, M. E. Adaptive noise wiener filter for scanning electron microscope imaging system. *Scanning* **38**, 148–163 (2016). URL http://dx.doi.org/10.1002/sca.21250. DOI 10.1002/sca.21250.

9. Aguirre, P. Image processing of microscopic cellular samples. In *2012 16th IEEE Mediterranean Electrotechnical Conference*, 366–369 (2012). DOI 10.1109/MELCON.2012.6196450.

10. Luisier, F., Vonesch, C., Blu, T. & Unser, M. Fast interscale wavelet denoising of poisson-corrupted images. *Signal Process.* **90**, 415 – 427 (2010). URL http://www.sciencedirect.com/science/article/pii/S0165168409003016. DOI http://dx.doi.org/10.1016/j.sigpro.2009.07.009.

11. Bise, R. & Sato, Y. Cell detection from redundant candidate regions under nonoverlapping constraints. *IEEE Transactions on Med. Imaging* **34**, 1417–1427 (2015). DOI 10.1109/TMI.2015.2391095.

12. Arteta, C., Lempitsky, V. S., Noble, J. A. & Zisserman, A. Detecting overlapping instances in microscopy images using extremal region trees. *Med. image analysis* **27**, 3–16 (2016).

13. Liao, M. *et al.* Automatic segmentation for cell images based on bottleneck detection and ellipse fitting. *Neurocomputing* **173**, 615–622 (2016).

14. Gul-Mohammed, J., Arganda-Carreras, I., Andrey, P., Galy, V. & Boudier, T. A generic classification-based method for segmentation of nuclei in 3D images of early embryos. *BMC Bioinforma.* **15**, 9 (2014). URL http://dx.doi.org/10.1186/1471-2105-15-9. DOI 10.1186/1471-2105-15-9.

15. WÃ„HLBY, C., SINTORN, I.-M., ERLANDSSON, F., BORGEFORS, G. & BENGTSSON, E. Combining intensity, edge and shape information for 2D and 3D segmentation of cell nuclei in tissue sections. *J. Microsc.* **215**, 67–76 (2004). URL http://dx.doi.org/10.1111/j.0022-2720.2004.01338.x. DOI 10.1111/j.0022-2720.2004.01338.x.

16. Cicconet, M., Geiger, D. & Gunsalus, K. C. Wavelet-based circular hough transform and its application in embryo development analysis. In *VISAPP* (2013).

17. TÃ¼retken, E., Wang, X., Becker, C. J., Haubold, C. & Fua, P. Network flow integer programming to track elliptical cells in time-lapse sequences. *IEEE Transactions on Med. Imaging* **36**, 942–951 (2017). DOI 10.1109/TMI.2016.2640859.

18. Türetken, E., Wang, X., Becker, C. J. & Fua, P. Detecting and tracking cells using network flow programming. *CoRR* **abs/1501.05499** (2015). URL http://arxiv.org/abs/1501.05499. 1501.05499.

19. Cliffe, A. *et al.* Quantitative 3D analysis of complex single border cell behaviors in coordinated collective cell migration. *Nat. Commun.* **8**, 14905 (2017).

20. Liu, K. *et al.* Optical flow guided cell segmentation and tracking in developing tissue. In *2014 IEEE 11th International Symposium on Biomedical Imaging (ISBI)*, 298–301 (2014). DOI 10.1109/ISBI.2014.6867868.

21. Tonti, S., Cataldo, S. D., Bottino, A. & Ficarra, E. An automated approach to the segmentation of hep-2 cells for the indirect immunofluorescence ana test. *Comput. Med. Imaging Graph.* **40**, 62 – 69 (2015). URL http://www.sciencedirect.com/science/article/pii/S089561111400202X. DOI https://doi.org/10.1016/j.compmedimag.2014.12.005.

22. Gul-Mohammed, J. & Boudier, T. Classified region growing for 3D segmentation of packed nuclei. In *Biomedical Imaging (ISBI), 2014 IEEE 11th International Symposium on*, 842–845 (IEEE, 2014).

23. Tarnawski, W. *et al.* A robust algorithm for segmenting and tracking clustered cells in time-lapse fluorescent microscopy. *IEEE J. Biomed. Heal. Informatics* **17**, 862–869 (2013). DOI 10.1109/JBHI.2013.2262233.

24. Ronneberger, O., Fischer, P. & Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Navab, N., Hornegger, J., Wells, W. M. & Frangi, A. F. (eds.) *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, 234–241 (Springer International Publishing, Cham, 2015).

25. Akram, S. U., Kannala, J., Eklund, L. & Heikkilä, J. Cell segmentation proposal network for microscopy image analysis. In Carneiro, G. *et al.* (eds.) *Deep Learning and Data Labeling for Medical Applications*, 21–29 (Springer International Publishing, Cham, 2016).

26. Liu, Z. *et al.* Nucleinet: A convolutional encoder-decoder network for bio-image denoising. In *Engineering in Medicine and Biology Society (EMBC), 2017 39th Annual International Conference of the IEEE*, 1986–1989 (IEEE, 2017).

27. Sadanandan, S. K., Ranefall, P., Le Guyader, S. & Wählby, C. Automated training of deep convolutional neural networks for cell segmentation. *Sci. reports* **7**, 7860 (2017).

28. Akram, S. U., Kannala, J., Eklund, L. & Heikki, J. Joint cell segmentation and tracking using cell proposals. In *2016 IEEE 13th International Symposium on Biomedical Imaging (ISBI)*, 920–924 (2016). DOI 10.1109/ISBI.2016.7493415.

29. Ulman, V. *et al.* An objective comparison of cell-tracking algorithms. *Nat. Methods* (2017). URL http://lmb.informatik.uni-freiburg.de/Publications/2017/RBB17.

30. MaÅ¡ka, M. *et al.* A benchmark for comparison of cell tracking algorithms. *Bioinforma.* **30**, 1609–1617 (2014). URL http://dx.doi.org/10.1093/bioinformatics/btu080. DOI 10.1093/bioinformatics/btu080. /oup/backfile/content_public/journal/bioinformatics/30/11/10.1093_bioinformatics_btu080/2/btu080.pdf.

31. Kesteven, G. The coefficient of variation. *Nat.* **158**, 520 (1946).

32. Aharon, M., Elad, M. & Bruckstein, A. k -svd: An algorithm for designing overcomplete dictionaries for sparse representation. *IEEE Transactions on Signal Process.* **54**, 4311–4322 (2006). DOI 10.1109/TSP.2006.881199.

33. Elad, M. & Aharon, M. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions on Image Process.* **15**, 3736–3745 (2006). DOI 10.1109/TIP.2006.881969.

34. Bradley, D. & Roth, G. Adaptive thresholding using the integral image. *J. Graph. Tools* **12**, 13–21 (2007). URL https://doi.org/10.1080/2151237X.2007.10129236. DOI 10.1080/2151237X.2007.10129236. https://doi.org/10.1080/2151237X.2007.10129236.

35. Otsu, N. A threshold selection method from gray-level histograms. *IEEE transactions on systems, man, cybernetics* **9**, 62–66 (1979).

36. Wright, J., Yang, A. Y., Ganesh, A., Sastry, S. S. & Ma, Y. Robust face recognition via sparse representation. *IEEE Transactions on Pattern Analysis Mach. Intell.* **31**, 210–227 (2009). DOI 10.1109/TPAMI.2008.79.

37. Mairal, J., Elad, M. & Sapiro, G. Sparse representation for color image restoration. *IEEE Transactions on Image Process.* **17**, 53–69 (2008). DOI 10.1109/TIP.2007.911828.

38. Bao, C., He, L. & Wang, Y. Linear spatial pyramid matching using non-convex and non-negative sparse coding for image classification. *CoRR* **abs/1504.06897** (2015). URL http://arxiv.org/abs/1504.06897.

39. Parvati, K., Rao, P. & Mariya Das, M. Image segmentation using gray-scale morphology and marker-controlled watershed transformation. *Discret. Dyn. Nat. Soc.* **2008** (2008).

40. Mallat, S. G. & Zhang, Z. Matching pursuits with time-frequency dictionaries. *IEEE Transactions on Signal Process.* **41**, 3397–3415 (1993). DOI 10.1109/78.258082.

41. Mallat, S. & Zhang, Z. Adaptive time-frequency transform. In *1993 IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 3, 241–244 vol.3 (1993). DOI 10.1109/ICASSP.1993.319480.

42. Chen, S. S., Donoho, D. L. & Saunders, M. A. Atomic decomposition by basis pursuit. *SIAM J. on Sci. Comput.* **20**, 33–61 (1998). URL http://dx.doi.org/10.1137/S1064827596304010. DOI 10.1137/S1064827596304010. http://dx.doi.org/10.1137/S1064827596304010.

43. Ollion, J., Cochennec, J., Loll, F., Escude, C. & Boudier, T. Tango: a generic tool for high-throughput 3D image analysis for studying nuclear organization. *Bioinforma.* **29**, 1840–1841 (2013).

44. Schmid, B., Schindelin, J., Cardona, A., Longair, M. & Heisenberg, M. A high-level 3D visualization api for java and ImageJ. *BMC bioinformatics* **11**, 274 (2010).

Author contributions statement

L.N. wrote the manuscript text, analyzed the data and developed the proposed algorithm; T.B. contributed to the conception and reviewed the manuscript.

Additional information

Competing financial interests: The authors declare no competing for financial interests.