DeepCEL0 for 2D Single Molecule Localization in Fluorescence Microscopy

Bioinformatics
doi.10.1093/bioinformatics/xxxxxx
Advance Access Publication Date: Day Month Year
Manuscript Category

OXFORD

Subject Section

DeepCEL0 for 2D Single Molecule Localization in Fluorescence Microscopy

Pasquale Cascarano [1], Maria Colomba Comes [2*], Andrea Sebastiani [1], Arianna Mencattini [2 3], Elena Loli Piccolomini [4], Eugenio Martinelli [2 3]

[1] Department of Mathematics, University of Bologna, Piazza di Porta S. Donato 5, Bologna, 40126, Italy.
[2]Department of Electronic Engineering, University of Tor Vergata, Via del Politecnico 1,00133, Rome, Italy.
[3] Interdisciplinary Center for Advanced Studies on Lab-on-Chip and Organ-on-Chip Applications, Via del Politecnico 1, 00133, Rome, Italy.
[4]Department of Computer Science and Engineering, Mura Anteo Zamboni 7, Bologna, 40126, Italy.

* maria.colomba.comes@uniroma2.it

Associate Editor: XXXXXXX

Received on XXXXX; revised on XXXXX; accepted on XXXXX

Abstract

Motivation: In fluorescence microscopy, Single Molecule Localization Microscopy (SMLM) techniques aim at localizing with high precision high density fluorescent molecules by stochastically activating and imaging small subsets of blinking emitters. Super Resolution (SR) plays an important role in this field since it allows to go beyond the intrinsic light diffraction limit.
Results: In this work, we propose a deep learning-based algorithm for precise molecule localization of high density frames acquired by SMLM techniques whose ℓ_2-based loss function is regularized by non-negative and ℓ_0-based constraints. The ℓ_0 is relaxed through its Continuous Exact ℓ_0 (CEL0) counterpart. The arising approach, named DeepCEL0, is parameter-free, more flexible, faster and provides more precise molecule localization maps if compared to the other state-of-the-art methods. We validate our approach on both simulated and real fluorescence microscopy data.
Availability and implementation: DeepCEL0 code is freely accessible at https://github.com/sedaboni/DeepCEL0
Contact: maria.colomba.comes@uniroma2.it

1 Introduction

The spatial resolution of images acquired by fluorescence microscopy refers to the shortest distance at which two fluorescent entities are perceived separately by the camera system. As a consequence of the light diffraction phenomena, lens with a uniformly illuminated circle aperture generate patterns known as Airy disks (Renz, 2013), namely the fluorescent emitters to be imaged are represented as blobs and not as isolated spots of light. The ability of the microscope to distinguish two relatively close entities is bounded by the well-known diffraction limit (Zheludev, 2008) that represents an intrinsic constraint of the resolution of the optical acquisition device. More precisely, according to the Abbe's Criterion, the smallest resolvable distance by a light microscope corresponds roughly to half the optical wavelength, that is about 200 nm (Abbe, 1873), thus compromising the direct observation of structures at nanoscale (Cascarano *et al.*, 2021) such as proteins, microtubules, mitochondria and less complex molecules.

In the last decade, super-resolution microscopy techniques have revolutionized light microscopy biological imaging allowing biologists to see beyond the diffraction limit (Sahl and Moerner, 2013). Among these techniques, in 2014, the pioneering works on STimulated Emission Depletion (STED) microscopy by Stefan W. Hell, and the development of Single-Molecule Localization Microscopy (SMLM) by Eric Betzig and William E. Moerner, have been awarded the Nobel Prize in Chemistry.

In particular, SMLM techniques, such as Photo-Activated Localisation Microscopy (PALM) (Betzig *et al.*, 2006; Hess *et al.*, 2006) and STochastic Optical Reconstruction Microscopy (STORM) (Rust *et al.*, 2006), provide high-precision molecule localization by sequentially activating and imaging a small percentage of photoswitchable fluorophores (emitters).

These molecules absorb and emit light of a specific wavelength. Therefore, by imaging a sparse set of activated emitters at a time, it is possible to identify their precise location in the Field of View (FOV). First, these SMLM techniques provide a stack of diffraction-limited frames, containing blobs (Airy disks), modelled by Gaussian Point Spread Functions (PSFs) (Rossmann, 1969). Then, each frame is analyzed separately with the aim of providing high precision localization maps of the emitters. After the individual processing is performed, all the frames are re-combined together to finally obtain a unique super-resolved image overcoming the diffraction limit. Finally, from the list of emitter positions, one can quantitatively measure the dimensions and emitter density of structures at nanoscale.

The sparser the set of activated emitters per frame is, the more precise the localization is. However, considering sparse frames takes a longer acquisition time thus limiting the ability to capture fast dynamics within live specimens. Conversely, a high density of activated emitters negatively affects the quality of the super-resolved image in terms of localization precision. Indeed, localization in high-density settings, that are characterized by overlapping PSFs, represents a challenging task for all the existing sophisticated localization software tools (Sage *et al.*, 2015; Bernhem and Brismar, 2018; Davis *et al.*, 2020).

Approaches based on diverse rationale, from blinking statistics (Dertinger *et al.*, 2009; Cox *et al.*, 2012; Gustafsson *et al.*, 2016), standard Gaussian fitting or centroid estimation (Henriques *et al.*, 2010) and subtraction of the model PSF (Gordon *et al.*, 2004; Qu *et al.*, 2004; Sergé *et al.*, 2008) to on-grid (Holden *et al.*, 2011; Min *et al.*, 2014; Hugelier *et al.*, 2016; Gazagnes *et al.*, 2017; Solomon *et al.*, 2019) and off-the-grid (Boyd *et al.*, 2017; Huang *et al.*, 2017) methods with sparsity-promoting priors, have been proposed to handle high-density data.

More specifically, the idea behind on-grid approaches is the creation of a fine-grained grid to model the locations of activated emitters. Among them, we find regularized deconvolution approaches such as CEL0 (Gazagnes *et al.*, 2017). This method solves, through an iterative scheme, a continuous non-convex optimization problem whose objective function is a weighted sum of an ℓ_2 fidelity term and the Continuous Exact ℓ_0 regularizer (Soubies *et al.*, 2015). Furthermore, non-negativity constraints are also added to the model since molecule intensities must be retrieved. Despite being among the most effective methods in the field of SMLM, the accuracy of CEL0 reconstructions strictly depends on the choice of the parameter balancing the ℓ_2 and the sparsity-promoting terms, thus drastically limiting its usage in real experiments. In contrast to on-grid methods, the so-called off-the-grid super resolution methods, such as ADCG (Boyd *et al.*, 2017) and TVSTORM (Huang *et al.*, 2017), work in a greedy way over the space of measures adding new molecules at each iteration and then optimize their positions and/or amplitudes in the continuum.

In the last few years, many deep learning-based approaches exploiting Convolutional Neural Networks (CNNs) have been proposed. The first method belonging to this class is called DeepSTORM (Nehme *et al.*, 2018). DeepSTORM, is trained on synthetic (artificially generated) frames, and tested directly on experimental data, thus avoiding to collect a huge amount of training samples related to the particular experiment under study. Despite being a fast and parameter-free super-resolution microscopy algorithm able to manage high-density data, DeepSTORM is not designed to reconstruct high precision localization maps. Another deep learning-based approach named DECODE (Speiser *et al.*, 2021) has been recently developed. DECODE is able to simultaneously detect and localize the emitters and to predict both the probability of detection and the uncertainity of localization for each emitter in high-density data.

In the present paper, we propose a method tailored to perform the localization, named DeepCEL0, that bridges the gap between deep learning-based and regularized deconvolution approaches. Basically, we exploit the skeleton of the network architecture underneath DeepSTORM using the CEL0 penalty as part of the training loss function. We further add non-negativity constraints through the insertion of a RELU layer in the network architecture. The strength of DeepCEL0 consists in embedding and joining the main advantages of the two standard methods, thus supplying, on the one hand, a quite fast and parameter-free deep-learning algorithm, like DeepSTORM, and, on the other hand, an efficient algorithm providing high precision localization maps, like CEL0. We report quantitative and qualitative molecule localization results on the simulated and real datasets from the 2013 and 2016 SMLM challenge[1]. We also compare the performances achieved by DeepCEL0 with those obtained by the two baseline state-of-the-art methods, CEL0 and DeepSTORM.

2 Mathematical Background

2.1 A mathematical image formation model for PALM/STORM acquisitions

We now recast the problem of finding high precision localization maps as a standard image super resolution inverse problem. The image acquisition model we develop is used in the experimental section to construct reliable synthetic data simulating PALM/STORM diffraction-limited acquisitions. Let $\mathbf{Y} \in \mathbb{R}_+^{M\times M}$ be a frame acquired by PALM/STORM techniques representing a sparse set of activated molecules on a coarse pixel-grid of dimension $M \times M$. Let $\mathbf{X} \in \mathbb{R}_+^{N\times N}$ be the localization map referred to the diffraction-limited PALM/STORM acquisition, that is a HR image defined on a L-time thinner pixel-grid such that $N = LM$, with L a non-negative integer termed as magnification factor.

In the following, for the sake of the readability of the manuscript, given a generic 2D image $\mathbf{Z}$ of dimension $m \times n$ we denote by $\overrightarrow{\mathbf{Z}}$ its vectorized version of dimension $mn \times 1$.

The acquisition of the LR image $\overrightarrow{\mathbf{Y}}$ is modelled through the following discrete degradation process:

$$\overrightarrow{\mathbf{Y}} = \mathcal{N}(\mathbf{S}_L\mathbf{K}\overrightarrow{\mathbf{X}}), \tag{1}$$

where $\mathcal{N}$ models the presence of data dependent Poisson noise and Additive White Gaussian (AWG) noise with mean zero and standard deviation σ_η, which are typical aberrations in fluorescence imaging (Garini *et al.*, 1999; Waters, 2009; Jezierska *et al.*, 2012).
The presence of the typical Airy disk patterns (blobs) in the LR acquisition is modelled through Gaussian PSFs. Therefore, in Eq. (1) we assume the operator $\mathbf{K} \in \mathbb{R}^{N^2\times N^2}$ denotes the discretization of a convolution with a Gaussian kernel $k : \mathbb{R}^2 \to \mathbb{R}$ with mean equal to zero and standard deviation σ_k, which is defined as follows:

$$k(x, y) := \frac{1}{\sigma_k\sqrt{2\pi}}\exp\left(-\frac{x^2+y^2}{2\sigma_k^2}\right). \tag{2}$$

The operator $\mathbf{S}_L \in \mathbb{R}^{M^2\times N^2}$ in Eq. (1) refers to the discretization of the downsampling operator linking the HR localization map represented on a fine pixel grid to the PALM/STORM acquired image defined on a L-time coarser pixel grid.

2.2 A deconvolution approach with a sparsity-promoting prior for single molecule localization

The problem in Eq. (1) is well-known to be ill-posed, meaning that, given $\overrightarrow{\mathbf{Y}}$, it is not feasible to retrieve $\overrightarrow{\mathbf{X}}$ by simply inverting the degradation described by $\mathbf{S}_L\mathbf{K}$, due to the lack of uniqueness and stability of the solution. Since, for each frame, the related localization map is sparse, a quite common approach to overcome this instability, consists in forcing the solution to satisfy some sparsity constraints through the ℓ_0 functional.

[1] https://srm.epfl.ch/Challenge/ChallengeSimulatedData

More precisely, an estimate $\overrightarrow{\mathbf{X}}^*$ of the unknown $\overrightarrow{\mathbf{X}}$ can be viewed as the solution of the following ℓ_0-regularized optimization problem:

$$\overrightarrow{\mathbf{X}}^* \in \underset{\mathbf{x}\in\mathbb{R}^{N^2}}{\arg\min} \frac{1}{2}\|\mathbf{S}_L\mathbf{K}\mathbf{x} - \overrightarrow{\mathbf{Y}}\|_2^2 + \lambda\|\mathbf{x}\|_0 + \mathbf{1}_{\geq 0}(\mathbf{x}), \quad (3)$$

where $\|\cdot\|_0 : \mathbb{R}^{N^2} \to \mathbb{R}$ denotes the ℓ_0 functional defined as:

$$\|\mathbf{x}\|_0 := \sum_{i=1}^{N^2} |\mathbf{x}_i|_0 \quad \text{with} \quad |\mathbf{x}_i|_0 := \begin{cases} 0 & \mathbf{x}_i = 0 \\ 1 & \mathbf{x}_i \neq 0, \end{cases} \quad (4)$$

which counts the number of non-zero elements of $\mathbf{x} \in \mathbb{R}^{N^2}$. Moreover, $\lambda > 0$ is the trade-off parameter measuring the level of sparsity of $\overrightarrow{\mathbf{X}}^*$, whereas $\mathbf{1}_{\geq 0}(\cdot)$ is formally the characteristic function of the positive octant of $\mathbb{R}^{N^2}$ constraining the computed estimation $\overrightarrow{\mathbf{X}}^*$ to have positive entries and $\|\cdot\|_2$ is the ℓ_2-norm.
Gazagnes *et al.* (2017) propose to replace the ℓ_0 penalization (4) with its continuous relaxation $\Phi_{\text{CEL0}} : \mathbb{R}^{N^2} \to \mathbb{R}$, named as CEL0 penalizer (Soubies *et al.*, 2015), which reads:

$$\Phi_{\text{CEL0}}(\mathbf{x}) := \sum_{i=1}^{N^2} \lambda_{\text{CEL0}} - \frac{\|\mathbf{c}_i\|}{2}\left(|\mathbf{x}_i| - \frac{\sqrt{2\lambda_{\text{CEL0}}}}{\|\mathbf{c}_i\|}\right)^2 \mathbf{1}_{V_i}, \quad (5)$$

where by $\mathbf{1}_{V_i}$ we denote the characteristic function of the set $V_i := \{\mathbf{x}_i \in \mathbb{R} \big| |\mathbf{x}_i| < \frac{\sqrt{2\lambda_{\text{CEL0}}}}{\|\mathbf{c}_i\|}\}$; by c_i we denote the i-th column of the matrix $\mathbf{S}_L\mathbf{K}$ for $i = 1 \dots N^2$, whereas the positive scalar λ_{CEL0} balances the strength of the sparsity induced by the CEL0 penalizer. Therefore, given a PALM/STORM acquisition $\overrightarrow{\mathbf{Y}}$, the CEL0-based method estimates the high precision localization map $\overrightarrow{\mathbf{X}}^*$ by solving the following optimization problem:

$$\overrightarrow{\mathbf{X}}^* \in \underset{\mathbf{x}\in\mathbb{R}^{N^2}}{\arg\min} \frac{1}{2}\|\mathbf{S}_L\mathbf{K}\mathbf{x} - \overrightarrow{\mathbf{Y}}\|_2^2 + \Phi_{\text{CEL0}}(\mathbf{x}) + \mathbf{1}_{\geq 0}(\mathbf{x}). \quad (6)$$

For its numerical solution, in (Gazagnes *et al.*, 2017) the authors make use of the iterative reweighted ℓ_1 (IRL1) strategy (Ochs *et al.*, 2015) which is tailored to handle nosmooth noconvex optimization problems and ensures the convergence to a critical point of (6) (Soubies *et al.*, 2015).

Finally, the interest in minimizing (6) comes from the theoretical analysis developed in (Soubies *et al.*, 2015). More precisely, the minimizers of the ℓ_2-ℓ_0 objective (3) are contained in the set of minimizers of the ℓ_2-CEL0 objective (6) and, furthermore, some local minimizers of (3) are not critical points of (6).

2.3 A deep learning based method for single molecule localization

Nehme *et al.* (2018) have introduced a fast and parameter-free deep learning-based method termed as DeepSTORM, which makes use of a CNN to provide a HR counterpart of the LR frames acquired by PALM/STORM techniques. In particular, given an outer training set of K 2D-image pairs $\{(\mathbf{Y}_k, \mathbf{X}_k)\}_{k=1\dots K}$, where $\mathbf{Y}_k \in \mathbb{R}^{M\times M}$ denotes the LR input and $\mathbf{X}_k \in \mathbb{R}^{N\times N}$ denotes its HR counterpart, and a particular encoder-decoder architecture f_θ with weights θ, the DeepSTORM training involves the following ℓ_1-regularized loss function:

$$\frac{1}{K}\left(\sum_{k=1}^{K} \|g * f_\theta(\text{NN}(\mathbf{Y}_k)) - g * \mathbf{X}_k\|_F^2 + \|\overrightarrow{f_\theta(\text{NN}(\mathbf{Y}_k))}\|_1\right), \quad (7)$$

where by NN we denote the coarse Nearest Neighbour super resolution algorithm resampling a LR image on a L-times finer grid by simply replicating the pixel values; by $*$ we refer to the discrete convolution with a 2D Gaussian kernel g whose standard deviation is equal to 1. Moreover, by $\|\cdot\|_F$ and $\|\cdot\|_1$ we denote the Frobenius norm and the ℓ_1-norm, respectively.

Once provided the final set of trained weights θ^* and a PALM/STORM acquisition $\mathbf{Y}$, an approximation $\mathbf{X}^*$ of the HR localization map $\mathbf{X}$ is obtained by computing $f_{\theta^*}(\text{NN}(\mathbf{Y}))$. Roughly speaking, DeepSTORM approximates the map linking the coarse Nearest Neighbour HR approximations of the PALM/STORM LR frames to their high-quality HR counterparts.

As far as the training set is concerned, one of the main novelties of DeepSTORM is that it can provide good performances on real images even if it is trained only on a synthetic dataset created using an ImageJ plug-in called ThunderStorm (Ovesnỳ *et al.*, 2014).

As for standard deconvolution approaches, it is possible to weight the strength of the ℓ_1-regularization term in (7) through a regularization parameter. In (Nehme *et al.*, 2018), the authors fix this hyperparameter equal to 1 in order to keep DeepSTORM parameter-free. However, they claim to have observed high robustness of the resulting HR reconstructions at varying such a parameter.

3 Method

3.1 A CEL0 regularized loss function

Our proposal combines the learning-based approach DeepSTORM and the sparsity-constrained deconvolution approach CEL0. In particular, we aim at providing a method which preserves the main advantages of the both aforementioned approaches, namely the ability of CEL0 to retrieve high precision localization maps and the fast and parameter-free computation provided by DeepSTORM. Therefore, following the idea of (Nehme *et al.*, 2018), we train a CNN architecture f_θ based on the following regularized loss function:

$$\frac{1}{K}\left(\sum_{k=1}^{K} \|g * f_\theta(\text{NN}(\mathbf{Y}_k)) - g * \mathbf{X}_k\|_2^2 + \Phi_{\text{CEL0}}(\overrightarrow{f_\theta(\text{NN}(\mathbf{Y}_k))})\right), \quad (8)$$

where $\{(\mathbf{Y}_k, \mathbf{X}_k)\}_{k=1\dots K}$ is a prefixed training set of K LR-HR 2D-image pairs of size $M \times M$ and $N \times N$, respectively, whilst NN referrers to the Nearest Neighbour upscaling of factor equal to L. Differently from (Nehme *et al.*, 2018), the loss in Eq. (8) presents two main differences. First, ℓ_1-regularizer is replaced by the CEL0 regularizer in order to guarantee better sparse reconstructions. Second, we consider a non-negativity promoting CNN architecture f_θ which is a slightly modified version of the CNN architecture used in (Nehme *et al.*, 2018), in order to force non-negativity constraints on the reconstructed solutions. In the following paragraphs, we describe in detail the skeleton of the architecture and the synthetic training set considered in our experimentation.

3.2 A non-negativity promoting Deep Architecture

The considered CNN architecture, denoted as f_θ in Eq. (8), is a modified version of the one initially proposed in (Nehme *et al.*, 2018). For the sake of brevity, we call convolutional layer the composition of convolutional

filters with a batch normalization layer (Ioffe and Szegedy, 2015) followed by a ReLu non-linearity (Glorot *et al.*, 2011) as activation function. The original architecture is an encoder-decoder and it is composed of seven convolutional layers. Each layer uses 3×3 kernels of different depth equal to 32, 64, 128 and 512, respectively. In the encoder part, the filters' depth increases, and a 2×2 max-pooling is used as downscaling operator to compress the features. In the decoder part, the layers are interleaved with a nearest neighbour upsampling operator and the filters' depth decreases. At the end of the network, another layer is added to compute the pixel-wise prediction. This layer is a 1×1 convolutional filter. In the original implementation, this last layer uses a linear activation function, i.e., the identity. In order to induce non-negativity constraints to the computed solution, we replace the activation layer with ReLu.

3.3 Synthetic training set and implementation notes

In the experimental section, we evaluate how our method performs when dealing with an upsampling factor L equal to 4. According to this choice, we now describe the considered synthetic training set $\{(\mathbf{Y}_k, \mathbf{X}_k)\}_{k=1\dots K}$. In the following, we refer to $\mathbf{Y}_k$ as the input image and to $\mathbf{X}_k$ as the target image, respectively.

As well as for DeepSTORM (Nehme *et al.*, 2018), we generate a synthetic dataset made up of 20 high density images. The emitters are positioned on a FOV of size 64×64 pixels such that each pixel has size of 100 nm. We extract from these high density images, $K = 10000$ patches of size 26×26. By projecting the emitter positions on a 4-times thinner pixel grid, we build the target images $\mathbf{X}_k$ of size 104×104.

The input images $\mathbf{Y}_k$ for $k = 1 \dots K$ are constructed corrupting the synthetic images in order to simulate the experimental conditions. More precisely, we first blur all the 26×26 patches with a discrete Gaussian kernel and add Poisson and Gaussian noisy components. Then, these corrupted patches are upsampled by a factor equals 4 through the NN interpolation algorithm.

We stress that, in real applications, the standard deviation of the PSF considered, if unknown, can be estimated using the Abbe's criterion which requires the light wavelength and the numerical aperture of the optical device used for acquiring the experimental data under study. Furthermore, the amount of noise can be either calculated directly from the microscope and detector characteristics or even through several mathematical techniques (Paul *et al.*, 2010; Mandracchia *et al.*, 2020).

We use this synthetic dataset to train the proposed DeepCEL0 by minimizing the loss function defined in Eq. (8). We train the network for 100 epochs on batches containing 16 samples using Adam optimizer and setting the learning rate equal to 0.001.

In (8), we set g as a fixed Gaussian kernel, whose standard deviation equal to 1. Moreover, we remark that the contribution of the CEL0 regularization term can be weighted using the parameter λ_{CEL0}. It is worth noting that in this deep learning-based framework the parameter λ_{CEL0} has not the same meaning of the regularization parameter in the deconvolution approach. (6). Indeed, in the former case it does not directly correspond to a degree of sparsity of the computed solution. In our experiments, we tested different values for λ_{CEL0} but we observed that setting it to 100 provides outstanding results for all the tests.

The network is implemented in TensorFlow and the code is available at https://github.com/sedaboni/DeepCEL0.

4 Results

4.1 Performance Evaluation and competitors

We assess the quality of the high resolution localization maps provided by our method and the competitors both on synthetic and real PALM/STORM images through quality metrics and visual inspections. When the ground-truth (GT) images are available, the performances are evaluated by pairing the GT molecules with the estimated ones: a match between a GT and an estimated molecule is created when the distance between their localizations is lower than a set tolerance δ, whose standard values, if expressed in terms of pixels, are 2, 4 and 6 (Sage *et al.*, 2015). Such a tolerance δ is chosen lower than the Full-width at Half Maximum (FWHM) of the estimated Gaussian PSF modelling the Airy pattern. In the following, the matched estimated molecules up to the given tolerance are defined as True Positive (TP) molecules; the remaining estimated molecules are referred to as False Positive (FP) molecules; and finally, the GT molecules with no match are categorized as False Negative (FN) molecules. Beyond the estimated molecules, it is important to take into account the pixels of the GT images not corresponding to any molecule, which are labelled as True Negative (TN) molecules. The performances are assessed by computing the following evaluation metrics:

$$\text{Jaccard}(\%) = \frac{\text{TP}}{\text{TP} + \text{FP} + \text{FN}} \times 100, \tag{9}$$

$$\text{Sensitivity}(\%) = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100, \tag{10}$$

$$\text{Specificity}(\%) = \frac{\text{TN}}{\text{TN} + \text{FP}} \times 100. \tag{11}$$

In order to quantify the level of corruption in the simulated data, we consider the signal-to-noise ratio (SNR) given by the following formula:

$$\text{SNR(dB)} := 10 \log_{10} \left(\frac{\|\mathbf{S}\mathbf{H}\vec{\mathbf{X}}\|_2^2}{\|\mathbf{S}\mathbf{H}\vec{\mathbf{X}} - \vec{\mathbf{Y}}\|_2^2} \right) \tag{12}$$

where $\vec{\mathbf{X}}$ is the real vectorized molecule localization map and $\vec{\mathbf{Y}}$ is the LR counterpart.

We compare our approach with the state-of-the-art algorithms CEL0 and DeepSTORM on both high density synthetic and real SMLM PALM/STORM low resolution images whose level of corruption ranges from SNR = 15dB to SNR = 10 dB.

Finally, in order to inspect the role of the non-negativity constraints, we train the non-negativity promoting CNN used in DeepCEL0 by considering the ℓ_1 regularized loss function in (7). In the following, the method is referred to as DeepSTORM-ReLu.

Furthermore, we run DeepSTORM and DeepSTORM-ReLu by selecting the regularization parameters values in the range [1,200]. In brief, DeepSTORM does not achieve any significant improvement in the reconstructions of the localization maps. Therefore, for all the experiments, the regularization parameter is fixed equal to 1, as in its standard formulation (Nehme *et al.*, 2018). Instead, for what concerns DeepSTORM-ReLu, we found the best results by setting the regularization parameter equal to 10. So, in the following sections, only the best results obtained for DeepSTORM-ReLu are reported.

4.2 Localization on synthetic test images with theoretical PSF

DeepSTORM algorithm could provide a fast SR image reconstruction, although it is not tailored to provide high precision localization maps of the emitters. Conversely, CEL0 is designed to localize the molecules but presents two main flaws: the method is strongly dependent on the choice of the regularization parameter and the overall computation process is largely slower than the one provided by DeepSTORM.

Once the network has been trained, the proposed DeepCEL0 and DeepSTORM share a comparable computational time for SR image reconstruction. Moreover, as we demonstrate in the following, DeepCEL0 is also able to effectively localize the molecules.

First of all, to validate the ability of the proposed method to discriminate two or more neighboring emitters, we construct three different synthetic test scenarios as GT images, simulating high resolution fluorescence microscopy localization maps on a FOV of 512×512 pixels of size 25 nm. The two first scenarios, referred as Test 1a and Test 2a, represent two molecules, arranged on the 256th column, at distance of 25 nm and 75 nm, respectively. The other scenario, referred as Test 3a, shows four molecules disposed on a circle of a radius equal to 125 nm. Two molecules are arranged on the 256th column whereas the others are arranged on the 256th row. Once set $L = 4$, by down-sampling the GT images according to the model in Eq. (1), we simulate the acquisition of 128×128 LR diffraction-limited frames, where the PSF is modelled by a Gaussian function whose FWHM is equal to 258.21 nm, that is $\sigma_k \approx 110$ nm.

In the lower panel of Fig. 1, we report the close-ups ($\times 10$ zooming) of the Region of Interest (ROI), namely the central zone of the 512×512 images where the synthetic spots are located, for the GT and Nearest Neighbour (NN), DeepSTORM, DeepCEL0 and CEL0 super resolved reconstructions. All the reconstructions reported have been normalized in the interval [0,1].
The zooms related to these methods are highlighted by blue, purple, green, red and yellow boxes, respectively. It is note worthy that in the purple box the NN $\times 4$ upsampled image is reported in order to visualize the LR image at the same dimensions of the GT image. For the CEL0 method, we depict two different reconstructions obtained by setting two different regularization parameters for each test case.
In the upper panel of Fig. 1, we draw the line profiles corresponding to the 256th column of the 512×512 GT, NN and DeepCEL0 images.

For all the three tests, the two molecules arranged on the 256th column correspond to the two peaks in the line profile of the GT image (*see* the blue line and blue box), whereas we observe they are completely overlapped and indistinguishable in the diffraction limited NN upsampled image's line profile (*see* purple line and purple box).

As qualitative results, in Test 1a, Test 2a and Test 3a, DeepCEL0 is able to well approximate the positions of the two molecules (*see* the red dashed line and red box). In all the three tests considered, the synthetic molecules are placed at a distance which is largely smaller than FWHM. In particular, Test 3a is the most challenging scenario since the number of molecules to estimate is greater than two. DeepCEL0 provides broadly better performances than DeepSTORM and more stable performances with respect to the CEL0 method. Indeed, DeepSTORM (*see* green box) does not separate the two synthetic molecules and provides a large number of FP molecules, thus confirming it is not designed to produce accurate localization maps (Nehme *et al.*, 2018). CEL0 (*see* yellow boxes) , as expected, provides highly accurate localization maps, but, as we can observe, its performances are highly sensitive to the choice of the regularization parameter. Moreover, due to the non-convex nature of (6), the IRL1 (Ochs *et al.*, 2015) algorithm can get stuck in a local minimizer. Indeed, based on the properties of the CEL0 penalty satisfying the Kurdyka-Lojasiewicz property (Soubies *et al.*, 2015), it has been proved that the algorithm described in (Gazagnes *et al.*, 2017) converges to a critical point of the ℓ_2-CEL0 objective which is a local minimizer of the standard ℓ_2-ℓ_0 functional. Furthermore, it is remarkable how different scenarios require different optimal parameters.

On the contrary, the results provided by DeepCEL0 have been computed by using the same architecture trained setting $\lambda_{\text{CEL0}} = 100$. Therefore, these tests prove the effectiveness of our approach retrieving highly precise localization maps and more stability with respect to the choice of the hyperparameter λ_{CEL0}, if compared to the CEL0 method.

4.3 Localization on a realistic dataset with theoretical PSF

We now consider a more realistic dataset provided by the 2013 SMLM challenge. The dataset is a stack of 361 different frames used as GT localization maps in our analysis. A total number of 81049 emitters are counted: for each frame 217 fluorophores are activated on average. The frames simulate realistic high density acquisitions of 8 tubes of diameter size equal to 30 nm depicted on a FOV of 256×256 pixels of size 25 nm. We simulate LR diffraction-limited acquisitions by applying the image formation model in Eq. (1) setting $L = 4$, thus representing the GT scenarios on a coarser grid of 64×64 pixels of size 100 nm. The PSF is modelled by using a Gaussian function of FWHM = 258.21 nm ($\approx \sigma_k = 110$ nm). We further corrupt the LR frames by adding Poisson noise and Gaussian noise with different standard deviations, such that SNR = 10 dB, 12 dB, 15 dB, respectively referred to as Test 2b, Test 3b, Test 4b. Finally, we consider a more realistic scenario, referred to as Test 1b, where the frames are corrupted by Poisson noise and Gaussian noise components of different standard deviations. The localization performances of CEL0, DeepSTORM, DeepSTORM-ReLu and DeepCEL0 are assessed for the four tests by using the evaluation metrics introduced in Section 4.1.
In Table 1, we report the results expressed in terms of Jaccard index for δ set as 2,4,6 pixels, corresponding to a tolerance of 50,100,150 nm, respectively. The sensitivity and specificity values, instead, are computed only with respect to the tolerance $\delta = 2$, which expresses a more faithful compliance of the reconstructions compared to the GT.

DeepSTORM, DeepSTORM-ReLu and DeepCEL0 models are trained on the synthetic images described in Sub-Section 3.3 such that the PSF is set to satisfy the experimental conditions described above and the noise aberration leads to an SNR value equals 15 dB on average. These three trained models are used for all the four tests considered. For the sake of a fair comparison, for each test we estimate the CEL0 regularization parameter by a trial and error procedure sampling 30 values over the range $[1e-3, 1]$ on randomly chosen 8 frames among the stacked ones. In the following, we refer to the best regularization parameter for CEL0 as the one maximazing the Jaccard index with tolerance δ equals 2.

In Table 1, we report for each test (Test 1b, 2b, 3b, 4b) the results of CEL0, DeepSTORM, DeepSTORM-ReLu and DeepCEL0. For what concerns CEL0, in the first row referred to each test, we draw the results obtained with respect to the best regularization parameters, whereas, in the second row, we show the results obtained by choosing as regularization parameter the best one obtained for Test 4b. We stress that for the trained methods the regularization parameter is fixed as described in the previous sections.

We inspect in depth Test 2b, the most challenging one due to a low SNR value. In Fig. 2, we provide the localization maps of the reconstructions for the frame 2. Green circles indicate the TP molecules, whereas red crosses highlight the FP molecules. In Fig. 3, we report the sum of all the stacked reconstructed SR frames provided by the competing methods alongside the stacked GT and LR frames. In particular, for GT and all the methods, above the magenta line we report the activated molecules as white spots, whereas below the magenta line we depict the normalized images. The three close-ups ($\times 4$ zooming) show three region of interest, where the tubulins seems completely overlapped in the LR acquisition because of the diffraction limit.

As a general comment, on the one side, DeepSTORM struggles to provide high precision localization maps: very low Jaccard values for all the tests and tolerances are obtained. For all the tests, the sensitivity reaches the highest possible value but at the expense of a very low specificity, meaning that this result is completely biased since it provides a very huge number of FP molecules (more than 50000). As an example, in Test 2b, this aspect is evident from Fig. 2 and Fig. 3, where we provide the localization maps for a single frame and for the sum of the stacked SR frames.

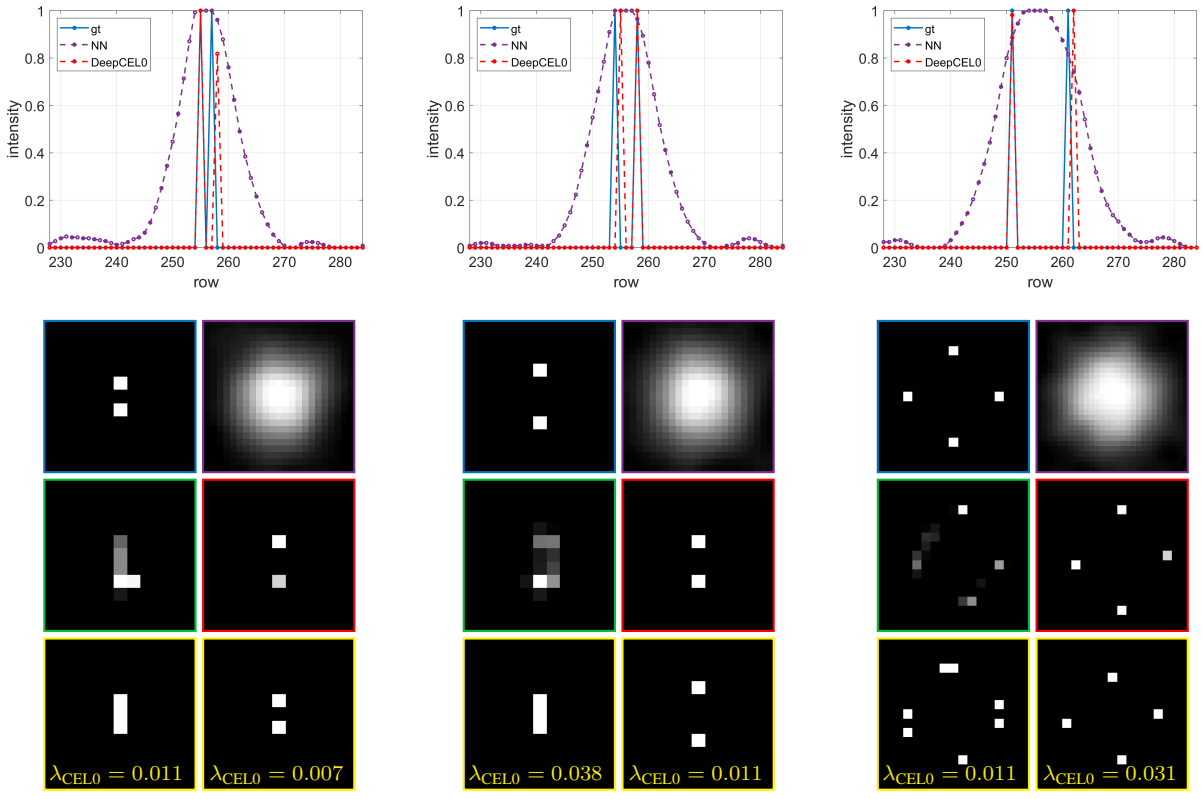

Figure 1: **Localization on synthetic test images.** Line profiles crossing the 256th column of GT, NN and DeepCEL0 images for Test 1a, Test 2a, Test 3a (upper panel from left to right). Close-ups (x10) on the ROI of GT (blue box), NN (purple box), DeepSTORM (green box), DeepCEL0 (red box) and CEL0 with two different regularization parameters (yellow box) for Test 1a, Test 2a, Test 3a (lower panel from left to right).

On the other side, CEL0 reaches the best performances on Test 4b and competing performances on Test 3b, but poor results on Test 1b and Test 2b that represent the most realistic and challenging scenarios.
Conversely to the above mentioned competing methods, DeepCEL0 provides satisfying results for all the tests with respect to the Jaccard, sensitivity and specificity indexes, thus confirming the effectiveness of the novelties introduced, namely the combination of non-negativity constraints and the usage of a CEL0 regularized loss function. In particular, Fig. 2 and Fig. 3 confirm once again how our proposal can better provide more faithful localization maps with less FP molecules if compared to the other competing methods. We investigate the addition of the sole non-negativity constraints by showing the evaluation metrics (Table 1) and the visual reconstructions (Fig. 2 and Fig. 3) obtained by DeepSTORM-ReLu. Our experiments confirm that the non-negativity promoting CNN is the component of DeepCEL0 which contributes the most to an improvement of the evaluation metrics. Indeed, Table 1 highlights how DeepSTORM-ReLu clearly outperforms DeepSTORM, although with globally slightly lower performances than DeepCEL0. Moreover, a visual comparison between the reconstructions of DeepSTORM-ReLu and DeepCEL0 emphasises how the contribution of the CEL0 penalty in the loss of DeepCEL0 allows the method to better induce sparsity, by suppressing false positives, and to retrieve more faithful structures. Finally, an interesting aspect to underline is that, even if DeepCEL0 is trained on images with SNR equals 15 dB on average using for all the experiments the same value of the regularization parameter, the trained models seem stable at varying the level of corruption (Test 1b, Test 2b, Test 3b, Test 4b). This stability is not still valid for CEL0 which shows a strong dependence on the choice of the regularization parameter as highlighted by the results in Table 1.

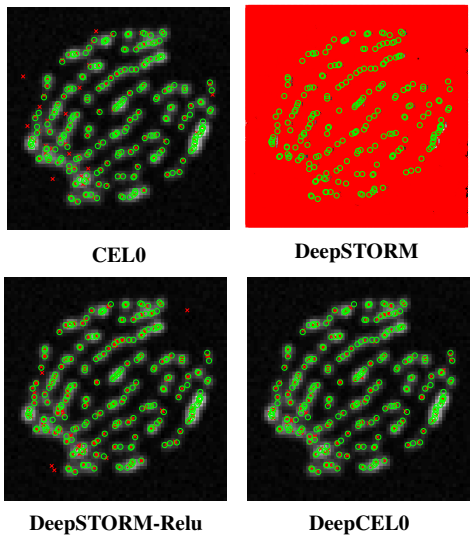

Figure 2: **Localization maps for the second frame of Test 1b.** The green circles indicate the TP molecules; the red cross marks indicate the FP molecules.

4.4 Localization on a realistic dataset without theoretical PSF

In the previous paragraph, we have considered a realistic dataset for which all the simulated acquisitions were constructed applying to the ground

Table 1. Performance evaluation for localization on 2013 SMLM challenge realistic dataset with theoretical PSF in terms of Jaccard Index, sensitivity and specificity. The first and second best Jaccard Index values are highlighted in red and blue, respectively.

Test	**Method**	**Jaccard** (%)			**Sensitivity**(%)	**Specificity** (%)
		$\delta = 2$	$\delta = 4$	$\delta = 6$	$\delta = 2$	$\delta = 2$
1b	CEL0	41.64	49.46	51.65	56.22	99.88
		35.94	43.93	46.70	67.74	99.71
	DeepSTORM	0.38	0.38	0.38	100.00	13.17
	DeepSTORM-ReLu	56.34	67.55	70.50	73.73	99.90
	DeepCEL0	58.96	68.94	71.55	70.89	99.95
2b	CEL0	42.29	50.21	51.48	49.08	99.95
		17.13	28.64	33.16	51.15	99.34
	DeepSTORM	0.38	0.38	0.38	100.00	13.34
	DeepSTORM-ReLu	51.86	65.93	69.70	70.51	99.88
	DeepCEL0	54.83	68.49	71.37	65.44	99.94
3b	CEL0	62.17	63.60	64.32	65.90	99.98
		48.19	50.46	51.38	73.73	99.82
	DeepSTORM	0.38	0.38	0.38	100.00	13.53
	DeepSTORM-ReLu	58.10	67.54	70.08	76.04	99.90
	DeepCEL0	61.13	69.66	71.55	69.59	99.95
4b	CEL0	68.27	70.33	71.02	78.34	99.95
	DeepSTORM	0.38	0.38	0.38	100.00	12.65
	DeepSTORM-ReLu	59.49	69.11	71.76	75.12	99.91
	DeepCEL0	61.48	69.80	72.33	70.12	99.96

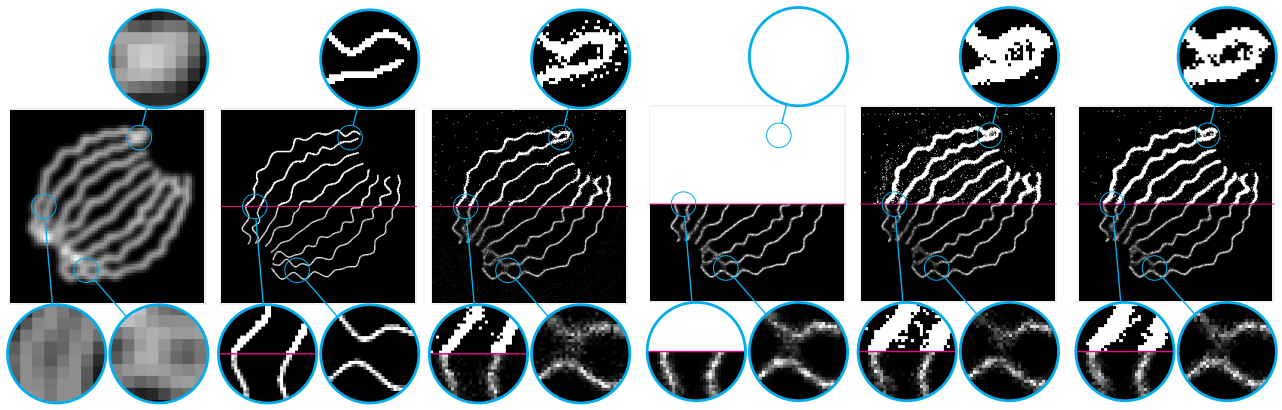

Figure 3: **Image reconstruction results of the whole stack for the 2013 SMLM challenge realistic dataset with theoretical PSF.**The first half of the images shows the binarized versions of the images. The second half of the images shows the normalized counterparts.

truth images the forward model (1), thus the sources of degradation were exactly known, i.e., the PSF and the type of noise (mixed Poisson and Gaussian). Therefore, we now consider the Microtubules dataset from the 2016 SMLM challenge. This experimental datasets consist of sequences of frames from realistic stained biological samples acquired according to a 2D modality. More precisely, these simulated data are created using a sophisticated simulator considering specificities of the fluo dyes and camera. The dataset is a stack of 2500 different frames depicting high density acquisitions (2 emitters per μm^2) of 3 microtubules on a FOV of $6.4 \times 6.4 \times 1.5\ \mu m$ (64×64 pixels of size 100 nm). Since in this case we do not know the theoretical PSF, the FWHM has been estimated through trial and error procedures and set equal to 447,45 nm ($\approx 190nm$).

The regularization parameter used for the CEL0 reconstruction has been heuristically selected in order to maximize the Jaccard ($\delta = 0$) index. In particular, the parameter chosen is equal to 1.85. The other models are trained on a synthetic training set constructed considering the estimated PSF.

In Fig. 4, we depict the LR image alongside the GT image and the reconstructions (setting $L = 4$) obtained by all the considered methods. In particular, below the green line, we show the normalized images, whereas above the green line, we report the maps representing the activated molecules depicted in white. To highlight the differences between the competing methods we consider three close-ups ($\times 4$ zooming). All the considerations raised on the previous subparagraph hold also in this case. The DeepCEL0 reconstruction reaches 39.43 as Jaccard index setting $\delta = 0$, whereas CEL0, DeepSTORM and DeepSTORM-ReLu reconstructions achieve 35.73, 2.32 and 34.38, respectively. This proves once again DeepCEL0 is able to localize the emitters with high precision. Furthermore, the binary maps show that DeepCEL0 preserves better the

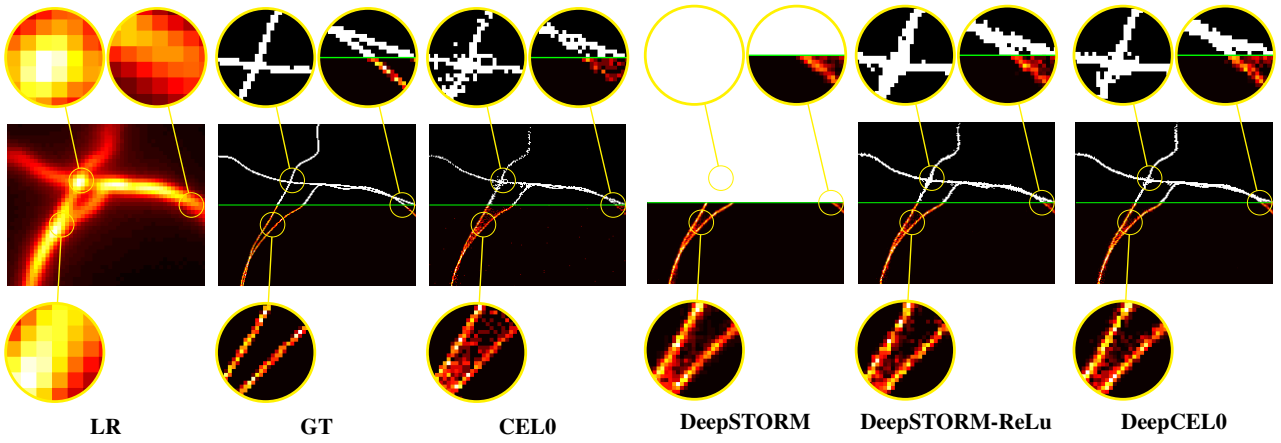

Figure 4: **Image reconstruction results of the whole stack for the 2016 SMLM challenge realistic dataset without theoretical PSF.** The first half of the images shows the binarized versions of the images. The second half of the images shows the normalized counterparts.

shape of the tubulins if compared with CEL0 and DeepSTORM reconstructions. Moreover, by comparing the close-ups of the DeepCEL0 and DeepSTORM-ReLu solutions it is evident how the addition of the CEL0 based regularizer helps to improve the localization and to suppress FP molecules.

4.5 Localization on a real dataset

As final evaluation test, we compare DeepCEL0 with CEL0 and DeepSTORM on a real high density dataset of tubulins that is part of the 2013 IEEE ISBI SMLM challenge. The dataset refers to a real acquisition stack representing a field of view of 128×128 pixels of size 100 nm over 500 different consecutive frames. The PSF is modelled by using a Gaussian function of FWHM = 351.8 nm, that is $\sigma_k \approx 150$ nm (Chahid, 2014). We aim at representing the acquired LR frames on a finer pixel grid of a factor $L = 4$, thus achieving a final resolution of 25 nm as pixel size. CEL0 reconstruction has been obtained by choosing the regularization parameter equal to 0.5 as suggested by (Bechensteen *et al.*, 2019). For all the learning-based models, we trained the networks on a training set of synthetic images described in Sub-section 3.3 by a Gaussian blur modelled with the above underlined PSF. In Fig. 5, we report the normalized LR image, i.e., the sum of all the stacked LR frames. Moreover we depict the reconstructions provided by all the considered competitors. Above the green line we show the image with the activated molecules as white spots, whereas, below the green line the normalized images are reported. Finally, three close-ups are considered to better underline the differences between the competing methods. All the considerations highlighted for the simulated scenarios are still valid in this case. We here appreciate the influence of the L0 regularization both in CEL0 and DeepCEL0 reconstructions. Indeed, they look sharper than the solution provided by DeepSTORM. Finally, if compared with CEL0, DeepCEL0 better separates the diverse tubulins and better preserves the continuous shapes.

5 Conclusion

In this paper, we present a deep learning-based method called DeepCEL0 for precise single molecule localization in high density fluorescence microscopy settings. The proposed method brings together the benefits of two well-known standard methods in the field, i.e., DeepSTORM and CEL0, introducing a network architecture with two main novelties: a continuous ℓ_0-penalized training loss function and the adoption of non-negativity constraints on the solution through a ReLu layer. Compared to the standard methods, numerical results show how DeepCEL0 can provide very high precision localization maps, without detriment to computational cost. Moreover, the method is parameter-free and can be easily tested and applied on real data after a training phase on only synthetic data. The promising results make the methods easy to perform in disparate real applications exploiting fluorescence microscopy.

Funding

P.C. and E.L.P. have been partially supported by the GNCS-INDAM project 2020 *"Ottimizzazione per l'apprendimento automatico e apprendimento automatico per l'ottimizzazione"*.

References

Abbe, E. (1873). Beiträge zur theorie des mikroskops und der mikroskopischen wahrnehmung. *Archiv für mikroskopische Anatomie*, **9**(1), 413–468.

Bechensteen, A., Blanc-Féraud, L., and Aubert, G. (2019). New methods for L 2-L 0 minimization and their applications to 2D Single-Molecule Localization Microscopy. In *2019 IEEE 16th International Symposium on Biomedical Imaging (ISBI 2019)*, pages 1377–1381. IEEE.

Bernhem, K. and Brismar, H. (2018). SMLocalizer, a GPU accelerated ImageJ plugin for single molecule localization microscopy. *Bioinformatics*, **34**(1), 137–138.

Betzig, E., Patterson, G. H., Sougrat, R., Lindwasser, O. W., Olenych, S., Bonifacino, J. S., Davidson, M. W., Lippincott-Schwartz, J., and Hess, H. F. (2006). Imaging intracellular fluorescent proteins at nanometer resolution. *Science*, **313**(5793), 1642–1645.

Boyd, N., Schiebinger, G., and Recht, B. (2017). The alternating descent conditional gradient method for sparse inverse problems. *SIAM Journal on Optimization*, **27**(2), 616–639.

Cascarano, P., Comes, M. C., Mencattini, A., Parrini, M. C., Piccolomini, E. L., and Martinelli, E. (2021). Recursive deep prior video: a super resolution algorithm for time-lapse microscopy of organ-on-chip experiments. *Medical Image Analysis*, page 102124.

Chahid, M. (2014). *Echantillonnage compressif appliqué à la microscopie de fluorescence et à la microscopie de super résolution*. Ph.D. thesis, Bordeaux.

Cox, S., Rosten, E., Monypenny, J., Jovanovic-Talisman, T., Burnette, D. T., Lippincott-Schwartz, J., Jones, G. E., and Heintzmann, R. (2012). Bayesian localization microscopy reveals nanoscale podosome dynamics. *Nature methods*, **9**(2), 195–200.

Davis, J. L., Soetikno, B., Song, K.-H., Zhang, Y., Sun, C., and Zhang, H. F. (2020). RainbowSTORM: an open-source ImageJ plug-in for spectroscopic single-molecule localization microscopy (sSMLM) data analysis and image reconstruction. *Bioinformatics*, **36**(19), 4972–4974.

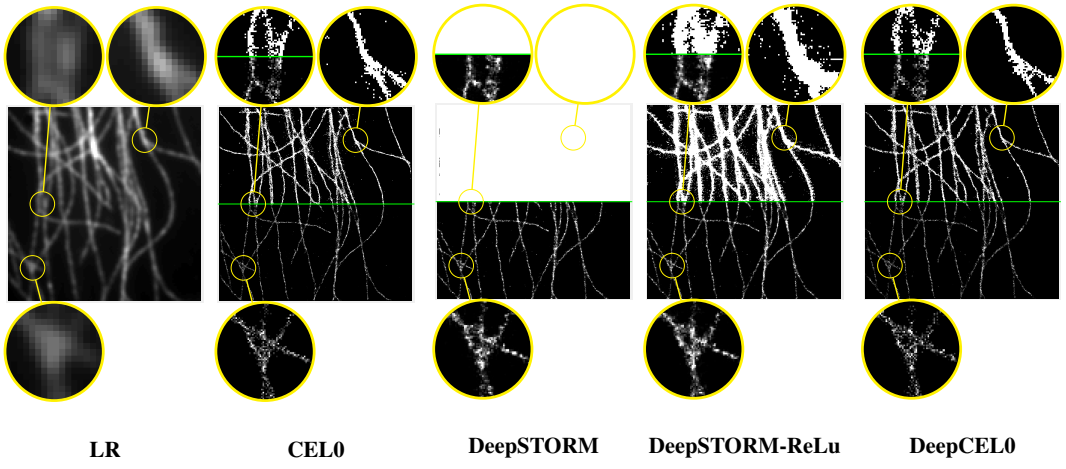

Figure 5: **Image reconstruction results of the whole stack for the IEEE ISBI tubulins dataset.** The first half of the images shows the binarized versions of the images. The second half of the images shows the normalized counterparts.

Dertinger, T., Colyer, R., Iyer, G., Weiss, S., and Enderlein, J. (2009). Fast, background-free, 3D super-resolution optical fluctuation imaging (SOFI). *Proceedings of the National Academy of Sciences*, **106**(52), 22287–22292.

Garini, Y., Gil, A., Bar-Am, I., Cabib, D., and Katzir, N. (1999). Signal to noise analysis of multiple color fluorescence imaging microscopy. *Cytometry: The Journal of the International Society for Analytical Cytology*, **35**(3), 214–226.

Gazagnes, S., Soubies, E., and Blanc-Féraud, L. (2017). High density molecule localization for super-resolution microscopy using CEL0 based sparse approximation. In *2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017)*, pages 28–31. IEEE.

Glorot, X., Bordes, A., and Bengio, Y. (2011). Deep sparse rectifier neural networks. In *Proceedings of the fourteenth international conference on artificial intelligence and statistics*, pages 315–323. JMLR Workshop and Conference Proceedings.

Gordon, M. P., Ha, T., and Selvin, P. R. (2004). Single-molecule high-resolution imaging with photobleaching. *Proceedings of the National Academy of Sciences*, **101**(17), 6462–6465.

Gustafsson, N., Culley, S., Ashdown, G., Owen, D. M., Pereira, P. M., and Henriques, R. (2016). Fast live-cell conventional fluorophore nanoscopy with ImageJ through super-resolution radial fluctuations. *Nature communications*, **7**(1), 1–9.

Henriques, R., Lelek, M., Fornasiero, E. F., Valtorta, F., Zimmer, C., and Mhlanga, M. M. (2010). QuickPALM: 3D real-time photoactivation nanoscopy image processing in ImageJ. *Nature methods*, **7**(5), 339–340.

Hess, S. T., Girirajan, T. P., and Mason, M. D. (2006). Ultra-high resolution imaging by fluorescence photoactivation localization microscopy. *Biophysical journal*, **91**(11), 4258–4272.

Holden, S. J., Uphoff, S., and Kapanidis, A. N. (2011). Daostorm: an algorithm for high-density super-resolution microscopy. *Nature methods*, **8**(4), 279–280.

Huang, J., Sun, M., Ma, J., and Chi, Y. (2017). Super-resolution image reconstruction for high-density three-dimensional single-molecule microscopy. *IEEE Transactions on Computational Imaging*, **3**(4), 763–773.

Hugelier, S., De Rooi, J. J., Bernex, R., Duwé, S., Devos, O., Sliwa, M., Dedecker, P., Eilers, P. H., and Ruckebusch, C. (2016). Sparse deconvolution of high-density super-resolution images. *Scientific reports*, **6**(1), 1–11.

Ioffe, S. and Szegedy, C. (2015). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning*, pages 448–456. PMLR.

Jezierska, A., Talbot, H., Chaux, C., Pesquet, J.-C., and Engler, G. (2012). Poisson-gaussian noise parameter estimation in fluorescence microscopy imaging. In *2012 9th IEEE International Symposium on Biomedical Imaging (ISBI)*, pages 1663–1666. IEEE.

Mandracchia, B., Hua, X., Guo, C., Son, J., Urner, T., and Jia, S. (2020). Fast and accurate scmos noise correction for fluorescence microscopy. *Nature communications*, **11**(1), 1–12.

Min, J., Vonesch, C., Kirshner, H., Carlini, L., Olivier, N., Holden, S., Manley, S., Ye, J. C., and Unser, M. (2014). Falcon: fast and unbiased reconstruction of high-density super-resolution microscopy data. *Scientific reports*, **4**(1), 1–9.

Nehme, E., Weiss, L. E., Michaeli, T., and Shechtman, Y. (2018). Deep-STORM: super-resolution single-molecule microscopy by deep learning. *Optica*, **5**(4), 458–464.

Ochs, P., Dosovitskiy, A., Brox, T., and Pock, T. (2015). On iteratively reweighted algorithms for nonsmooth nonconvex optimization in computer vision. *SIAM Journal on Imaging Sciences*, **8**(1), 331–372.

Ovesnỳ, M., Křížek, P., Borkovec, J., Švindrych, Z., and Hagen, G. M. (2014). ThunderSTORM: a comprehensive ImageJ plug-in for PALM and STORM data analysis and super-resolution imaging. *Bioinformatics*, **30**(16), 2389–2390.

Paul, P., Duessmann, H., Bernas, T., Huber, H., and Kalamatianos, D. (2010). Automatic noise quantification for confocal fluorescence microscopy images. *Computerized Medical Imaging and Graphics*, **34**(6), 426–434.

Qu, X., Wu, D., Mets, L., and Scherer, N. F. (2004). Nanometer-localized multiple single-molecule fluorescence microscopy. *Proceedings of the National Academy of Sciences*, **101**(31), 11298–11303.

Renz, M. (2013). Fluorescence microscopy—A historical and technical perspective. *Cytometry Part A*, **83**(9), 767–779.

Rossmann, K. (1969). Point spread-function, line spread-function, and modulation transfer function: tools for the study of imaging systems. *Radiology*, **93**(2), 257–272.

Rust, M. J., Bates, M., and Zhuang, X. (2006). Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM). *Nature methods*, **3**(10), 793–796.

Sage, D., Kirshner, H., Pengo, T., Stuurman, N., Min, J., Manley, S., and Unser, M. (2015). Quantitative evaluation of software packages for single-molecule localization microscopy. *Nature methods*, **12**(8), 717–724.

Sahl, S. J. and Moerner, W. (2013). Super-resolution fluorescence imaging with single molecules. *Current opinion in structural biology*, **23**(5), 778–787.

Sergé, A., Bertaux, N., Rigneault, H., and Marguet, D. (2008). Dynamic multiple-target tracing to probe spatiotemporal cartography of cell membranes. *Nature methods*, **5**(8), 687–694.

Solomon, O., Eldar, Y. C., Mutzafi, M., and Segev, M. (2019). Sparcom: Sparsity based super-resolution correlation microscopy. *SIAM Journal on Imaging Sciences*, **12**(1), 392–419.

Soubies, E., Blanc-Féraud, L., and Aubert, G. (2015). A continuous exact \ell_0 penalty (CEL0) for least squares regularized problem. *SIAM Journal on Imaging Sciences*, **8**(3), 1607–1639.

Speiser, A., Müller, L.-R., Hoess, P., Matti, U., Obara, C. J., Legant, W. R., Kreshuk, A., Macke, J. H., Ries, J., and Turaga, S. C. (2021). Deep learning enables fast and dense single-molecule localization with high accuracy. *Nature methods*, pages 1–9.

Waters, J. C. (2009). Accuracy and precision in quantitative fluorescence microscopy.

Zheludev, N. I. (2008). What diffraction limit? *Nature materials*, **7**(6), 420–422.