PARTICLE DETECTION BASED ON FEW SHOT LEARNING IN 3D FLUORESCENCE MICROSCOPY

Luc Vedrenne *Étienne Baudrier* *Denis Fortun*

ICube, UMR 7357, CNRS, University of Strasbourg, France

ABSTRACT

The observation of small biological particles in 3D fluorescence microscopy faces challenges due to resolution anisotropy and partial fluorescent labeling. A solution to address these issues consists in combining multiple acquisitions of individual particles. The detection of these single particles within each acquired volume is of utmost importance. A significant drawback of deep learning methods in this context is the substantial requirement for manually annotated training data. In response, we propose a particle detection pipeline for 3D data based on few-shot learning. It generalizes a 2D method based on the principle of positive-unlabeled learning. Our experiments with synthetic and real fluorescence-microscopy data demonstrate that our approach achieves high detection rates even with a limited amount of training data.

***Index Terms*—** Particle picking, detection, few shot learning, 3D fluorescence microscopy

1. INTRODUCTION

Fluorescence microscopy is a fundamental observation tool for understanding cellular mechanisms. 3D fluorescence images are obtained by exciting fluorophores that bind around a protein of interest. However, the resolution is highly anisotropic, and the fluorophores do not completely and uniformly cover the targeted particles. A single image gives only a partial view of the actual structure of cellular objects. Single particle reconstruction relies on the acquisition of images containing a large number of isolated, randomly oriented copies of the same particle, which are detected and then combined to reconstruct a 3D model. This method greatly improves resolution and compensates for the inhomogeneity of fluorophore distribution in the input data. The first step in a reconstruction is the detection of isolated particles in the acquired images, also called "particle picking" in microscopy (Fig. 1). The reconstruction quality is highly dependent on the success of the detection step. Noise and parasitic structures in the images make conventional detection methods ineffective. Only few learning methods dedicated to this task in fluorescence microscopy, mainly in 2D or requiring numerous training examples [1, 2, 3]. In practice, the detection of hundreds of particles in 3D images is performed manually, a tedious task prone to bias, limiting the number of particles picked.

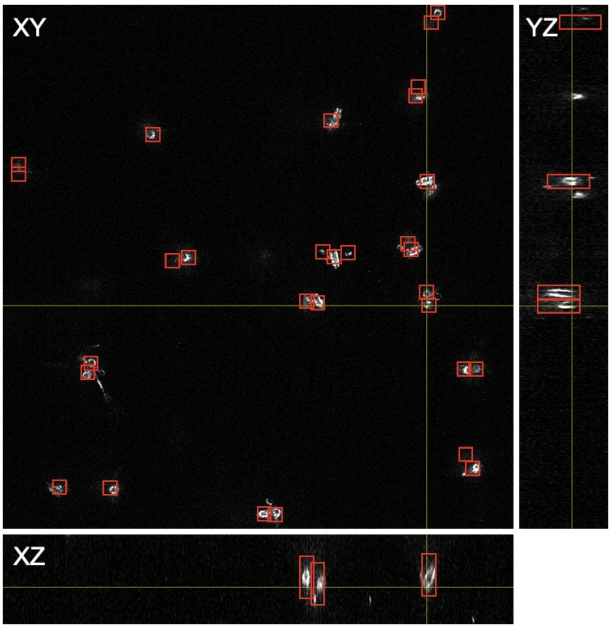


Fig. 1. Detection of single particles by 3D fluorescence microscopy on data described in [4].

Automatic deep learning detection methods have recently been applied to fluorescence microscopy images. In [1, 2], the authors employ 2D detection techniques using Convolutional Neural Networks (CNNs). Consequently, these approaches overlook the rich volumetric information present in 3D microscopy images. In contrast, Spilger et al. propose a fully 3D detection method in [3], but it relies on numerous weak annotations, making it not suitable for noisy situations and more complex objects like centrioles which are often observed in pairs. Deep learning detection methods have also been developed for another acquisition modality called cryo-electron microscopy [5], but most of them require the creation of a large labeled dataset, which limits their use in practice. In 2019, Bepler et al. [6] proposed a new method enabling

*Corresponding author vedrenne@unistra.fr

weakly supervised deep learning for particle detection based on the principle of *Positive-Unlabelled learning*. It is specific to 2D images and requires hundreds of training examples. A straightforward implementation of Bepler's method in 3D fails since the positive label proportion is significantly lower in 3D than in 2D. In this paper, we propose a complete and robust processing pipeline for particle detection in 3D fluorescence imaging. In addition to the idea of [6], we introduce three main steps necessary to process 3D fluorescence data. Firstly, image regions of interest are segmented. Training on these regions of interest is then stabilized by stochastic weight averaging. Finally, a post-processing that improves on the usual *non-maximum suppression* yields centered predictions, with no false positives. To achieve an F1-score of 0.85, conventional training requires 500 examples, while our model requires 10. The code will soon be available on a Gitlab repository .

2. METHOD

2.1. Positive-Unlabelled learning by batch

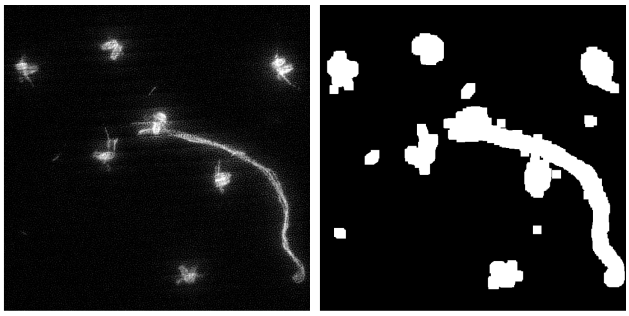
Positive-Unlabelled (PU) learning relies on a priori knowledge of the proportion of positive elements in the whole image to guide detection. Let us denote $\mathcal{P}$ the small portion of the image labeled as positive, $\mathcal{U}$ the rest of the unlabeled image and π the proportion of positive elements in $\mathcal{U}$. So we are looking for an estimator g minimizing $\mathbb{E}_{x\sim\mathcal{P}}(\mathcal{L}(g(x),1))$ under the constraint $\mathbb{E}_{x\sim\mathcal{U}}(g(x)) = \pi$ (where $\mathcal{L}$ is the cross entropy) [7]. As the estimator here is a neural network trained by stochastic gradient descent, the cost function will not be calculated exactly, but estimated from a subset (batch) of data. The distribution of k positive elements in a batch follows a binomial distribution $p(k)$ and the g estimator must then also describe a binomial distribution $q(k)$ over these k positive elements. To impose the constraint $\mathbb{E}_{x\sim\mathcal{U}}(g(x)) = \pi$, we can therefore seek to minimize the cross-entropy between the $q(k)$ and $p(k)$ distributions, so that the cost function becomes

$$\mathbb{E}_{x\sim\mathcal{P}}(L(g(x),1) + \lambda \sum_{k=1}^{n} q(k)\log(p(k)) \qquad (1)$$

This function allows us to apply the principle of Positive-Unlabelled learning in the context of stochastic gradient descent. It is the one proposed by [6] and which we will use.

2.2. Particle detection method

We perform detection by classifying a 3D sliding window that moves over the image with a certain step size. Indeed, while current detection methods are based on neural networks processing the whole image, this approach is not applicable to large 3D images from fluorescence microscopy (typically $100 \times 1000 \times 1000$ voxels).

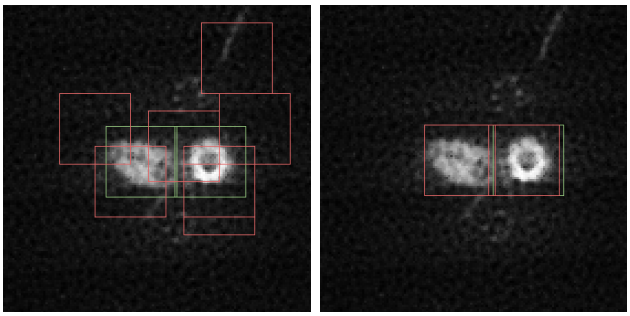


(a) (b)

Fig. 2. Example of segmentation of regions of interest. (a) Real data. (b) Segmented region of interest (white).

Region of interest selection In single particle reconstruction, we know the size of the particle we want to detect. Moreover, for a given acquisition protocol, it is possible make a coarse estimation of the total number of particles in each image, so that the proportion of positive voxels π_{ratio} in each image is approximately known, which enables the use of PU *learning*. For an image with $D = d \times h \times w$ voxels and comprising k particles of radius r, $\pi_{ratio} = \frac{k}{D}\frac{4}{3}\pi r^3$. To guarantee this ratio while having at least one positive element in a mini-batch, we need a minimum mini-batch size of $\lceil 1/\pi_{ratio} \rceil$. In practice, this minimum size is of the order of 10^3 or even 10^4, which requires a very large memory capacity and prevents the method from being implemented. Since 3D fluorescence microscopy images of single particles are predominantly made up of background (see Fig. 1), we propose to restrict the search to a region of interest, sufficiently selective to significantly increase the π_{ratio} value and therefore reduce the size of the mini-batchs, while retaining the particles. In the context of noisy 3D images with numerous artefacts, elementary segmentation methods, typically using simple thresholding, fail to achieve such a trade-off. We achieve this by correcting the histogram, followed by estimating a mixture of two Gaussians. We then erode the resulting segmentation map to eliminate artifacts, and apply a morphological opening to preserve an extended area around the particles (see Fig. 2). This procedure results in a region of interest sufficiently restricted to obtain a correct mini-batch size (typically of the order of 10^2). This procedure is applied both before training and during prediction.

Stochastic weight averaging When a real distribution is estimated from few examples, the estimator's ability to generalize depends strongly on the choice of examples. Thus, after selecting the region of interest of an image and carrying out training with few examples, we observed a high variance in performance: for 10 annotated particles, the F1-score varied between 0.6 and 0.9 (see Fig. 5). To counter this phenomenon, it is useful to average the estimates. For this reason, we apply stochastic weight averaging (SWA) during training



(a) (b)

Fig. 3. Illustration of the post-processing procedure for 3D predictions (here projected in 2D for visualization). (a) Predictions after NMS. (b) Predictions after full post-processing.

[8]. Two models are considered: the first (w_{SWA}) stores the current weight averaging and will be the final model used for predictions, the second (w) explores the weight space with a cyclic learning rate, so that each cycle results in a unique local minimum. At the end of each cycle, the weights are updated according to an exponential moving averaging scheme. The resulting model represents a flatter error minimum, and is therefore more amenable to generalization. For 10 particles, with SWA, the F1-score is between 0.85 and 0.88.

Post-processing Sliding window detection results in a large number of predictions for the same particle. Such predictions are usually filtered using the *non-maximum suppression* (NMS) algorithm, but the latter is not very effective in the case of noisy 3D data, with close objects and lots of artifacts. Although false positives after NMS correspond mainly to predictions close to particles, they are nonetheless troublesome for reconstruction. Since the predictions obtained after NMS are globally centered around a particle, we propose to spatially average overlapping predictions, then recenter these overlapping predictions on their center of mass. Finally, due to noise and numerous artifacts, some predictions are far from particles, but these correspond to low-intensity areas in the image and can therefore be eliminated by filtering them via the Otsu method after locally correcting the image histogram (*contrast limited adaptive histogram equalization*, CLAHE).

Quantitatively, for 60 particles on real data, our method reduces the number of false positives from 200 to 5, without introducing any false negatives (see Fig. 3).

2.3. Implementation details

We consider a standard learning model. The estimator is a convolutional neural network *EfficientNet-b0* [9]. The optimization algorithm is a *Stochastic Gradient Descent with momentum*. The learning rate follows a sinusoidal decay. The error function is that introduced in equation (1). For the reconstruction step, the windows classified as positive by the model must contain a single particle. Since particles can be close to each other, the window size must be approximately equal to that of a single particle. This small window size imposes a shallow network depth, since the spatial dimension of the image decreases as one progresses through its layers, to the point where for windows that are too small, the network becomes so shallow that it is no longer able to classify correctly. In the case of the *EfficientNet* architecture, this spatial reduction is due to the pitch of the convolutions. We change this pitch from 2 to 1 in certain convolutional layers. In practice, for a window size of 24^3, two layers in the first block are modified. Finally, when training on such a large dataset, overlearning is inevitable. We introduce several regularization elements to counteract this. The data are augmented (random composition of translations, rotations and dilations). Network weights are regularized using *drop connect* and *weight decay*. In addition to these standard methods, we apply regularization using an additional reconstruction task, learned in conjunction with classification, as proposed by [6]. The weighting between classification and reconstruction errors is also learned during learning [10].

3. EXPERIMENTS

To study our method, and in particular to compare it with fully supervised training, we need a large annotated set. To achieve this, we generate synthetic data: we use a model of a biological particle called a centriole described in [11], and apply to this model a series of random transformations to generate rotated particles corrupted by the main defects encountered in fluorescence microscopy acquisitions (anisotropy, noise, artifacts). This process results in the creation of several thousand examples of particles. To asses performance stability, each training run for a given number of examples is performed 40 times, randomly selecting a set of examples at each time. Finally, to validate performance on real data, 200 particles were manually annotated on images from [4].

3.1. Impact of number of annotations

We use the same model (same network, same hyperparameters, same regularizations, same pre- and post-processing), for two different error functions: a classical cross-entropy (fully supervised training) and the error function introduced by [6]. The results, summarized in Fig. 4, are as follows: when the number of annotations is sufficiently large ($\sim 10^3$), fully supervised training performs similarly to PU *learning*; when the number of annotations decreases, the performance of fully supervised training drops, while our method maintains a stable high F1-Score. Thus, we observe that for 10 annotated particles our model achieves a satisfactory F1-score of 0.85, which is reached with 500 examples for classical training. In experiments on real data we observe the same phenomenon: for 10 annotated particles our method achieves a F1-score of 0.78,

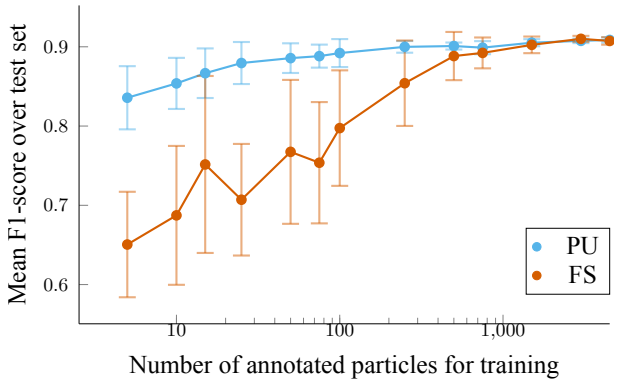


Fig. 4. Comparison of Positive-Unlabelled (PU) and Fully-Supervised (FS) training. Statistics on 40 training sessions.

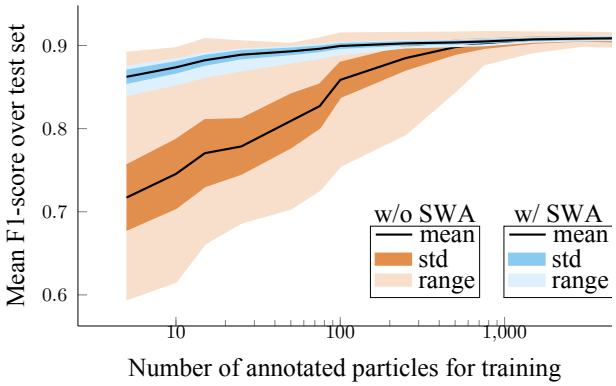


Fig. 5. Impact of number of annotations. Comparison with and without stochastic weight averaging (SWA). Statistics on 40 training sessions

.

compared with 0.56 for a conventional model.

3.2. Impact of the stochastic weight averaging

Training runs performed on different sets of examples reveal the impact of stochastic weight averaging (SWA). We compare two configurations, one with SWA and the other without, while keeping all other parameters identical. The number of examples varies between 5 and 3500. Fig. 5 provides a summary of the results. The overall performance is enhanced with SWA: although the results are stable when the number of examples is in the thousand range, we observe that for ten examples, the variance and amplitude of the results go from 30% without SWA to less than 5% with SWA. Therefore, the inclusion of SWA in the model significantly improves performance stability.

4. CONCLUSION

We have presented a 3D particle detection method for fluorescence microscopy data, leveraging successful application of positive-unlabeled learning. The method relies on a first segmentation step that provides regions of interest, which is crucial to reduce the proportion of background pixels and enable the use of PU learning. The proposed post-processing of predictions then leads to well-centered detection around the particles, suitable for subsequent reconstruction. Furthermore, stochastic weight averaging improves the stability of the results in the context of few-shot learning. Experimental tests show our model outperforms conventional training, particularly with fewer than a few hundred annotated particles, in both synthetic and real datasets.

5. ACKNOWLEDGMENTS

This work was supported by the French National ResearchAgency (ANR) through the SP-Fluo project (ANR-20-CE45-0007).

6. COMPLIANCE WITH ETHICAL STANDARDS

The authors declare that the study does not involve human- or animal-derived samples that require ethical approval. This article does not contain any studies involving human participants performed by any of the authors.

7. REFERENCES

[1] J. M. Newby et al., "Convolutional neural networks automate detection for tracking of submicron-scale particles in 2d and 3d," *PNAS USA*, vol. 115, no. 36, pp. 9026–9031, 2018.

[2] T. Wollmann et al., "Detnet: Deep neural network for particle detection in fluorescence microscopy images," in *Proc. IEEE ISBI*, 2019, pp. 517–520.

[3] R. Spilger, V. O. Chagin, C. Bold, L. Schermelleh, U. C. Müller, M. C. Cardoso, and K. Rohr, "Deep neural network for 3d particle detection in 3d fluorescence microscopy images via density map regression," in *Proc. IEEE ISBI*, 2022, pp. 1–4.

[4] D. Mahecic, D. Gambarotto, K.M. Douglass, D. Fortun, N. Banterle, K.A. Ibrahim, M. Le Guennec, P. Gönczy, V. Hamel, P. Guichard, and Suliana Manley, "Homogeneous multifocal excitation for high-throughput super-resolution imaging," *Nat. Methods*, 2020.

[5] T. Wagner, F. Merino, M. Stabrin, T. Moriya, C. Antoni, A. Apelbaum, P. Hagel, O. Sitsel, T. Raisch, D. Prumbaum, D. Quentin, D. Roderer, S. Tacke, B. Siebolds, E. Schubert, T.R. Shaikh, P. Lill, C. Gatsogiannis, and S. Raunser, "Sphire-cryolo is a fast and accurate fully automated particle picker for cryo-em," *Commun. Biol.*, 2019.

[6] T. Bepler, A. Morin, M. Rapp, J. Brasch, L. Shapiro, A.J. Noble, and B. Berger, "Positive-unlabeled convolutional neural networks for particle picking in cryo-electron micrographs," *Nat. Methods*, 2019.

[7] Gideon S. Mann and Andrew McCallum, "Generalized expectation criteria for semi-supervised learning with weakly labeled data," *J. Mach. Learn. Res.*, 2010.

[8] P. Izmailov, D. Podoprikhin, T. Garipov, D. P. Vetrov, and A. G. Wilson, "Averaging weights leads to wider optima and better generalization," in *Proc. Conf. Uncertain. Artif. Intell.*, 2018.

[9] M. Tan and Q.V. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *ICML*, 2020.

[10] R. Cipolla, Y. Gal, and A. Kendall, "Multi-task learning using uncertainty to weigh losses for scene geometry and semantics," in *CVPR*, 2018.

[11] M. Le Guennec, N. Klena, D. Gambarotto, M.H. Laporte, A.M. Tassin, H. Van den Hoek, P.S. Erdmann, M. Schaffer, L. Kovacik, S. Borgers, K.N. Goldie, H. Stahlberg, M. Bornens, J. Azimzadeh, B.D. Engel, V. Hamel, and P. Guichard, "A helical inner scaffold provides a structural basis for centriole cohesion," *Sci. Adv.*, 2020.