

EXTRACTING AXIAL DEPTH AND TRAJECTORY TREND USING ASTIGMATISM, GAUSSIAN FITTING, AND CNNs FOR PROTEIN TRACKING

*Kristofer delas Peñas^{1,3}, Mariia Dmitrieva¹, Joël Lefebvre¹, Helen Zenner², Edward Allgeyer²
Martin Booth¹, Daniel St Johnston², Jens Rittscher¹*

¹ Department of Engineering Science, University of Oxford, Oxford, UK

² Gurdon Institute, University of Cambridge, Cambridge, UK

³ Department of Computer Science, University of the Philippines, Quezon City, Philippines

ABSTRACT

Accurate analysis of vesicle trafficking in live cells is challenging for a number of reasons: varying appearance, complex protein movement patterns, and imaging conditions. To allow fast image acquisition, we study how employing an astigmatism can be utilized for obtaining additional information that could make tracking more robust. We present two approaches for measuring the z position of individual vesicles. Firstly, Gaussian curve fitting with CNN-based denoising is applied to infer the absolute depth around the focal plane of each localized protein. We demonstrate that adding denoising yields more accurate estimation of depth while preserving the overall structure of the localized proteins. Secondly, we investigate if we can predict using a custom CNN architecture the axial trajectory trend (upwards, downwards, or constant motion) rather than the exact depth. We demonstrate that this method performs well on calibration beads data without the need for denoising. By incorporating the obtained depth information into a trajectory analysis, we demonstrate the potential of improving vesicle tracking.

Index Terms— tracking, biomedical imaging, gaussian fitting, denoising, convolutional neural networks, confocal microscopy

1. INTRODUCTION

The quantitative analysis of the vesicle movement is important for biological studies in cells and movement along z -axis can be essential for understanding biological processes. Several different approaches to achieve three-dimensional (3D) localization had been proposed in recent years [1, 2]. These techniques mainly involve changing the shape of the point

spread function (PSF) used in the imaging setup. Such techniques include astigmatism, double helix, and biplane [3]. Through PSF shape modification, the localization of single fluorescent molecules can be achieved not only in the lateral plane but also axially. Astigmatic localization microscopy is a popular single-molecule localization method, with several software variations and heavy usage in 3D localization competitions. The defining characteristic in images with astigmatism is the elongation of spots in one of the lateral axes as you move from the focal plane towards the objective, and elongation in the other lateral axis as a particle moves away from the focal plane and the objective. The usual way to extract depth with this kind of data is to fit an elliptical Gaussian curve to the localized spots [3]. This Gaussian fitting process is favored because of its simplicity and the speed by which it can be computed. However, due to non-ideal imaging optics and background noise, Gaussian fitting may sometimes fail to obtain good depth estimates.

Localization and tracking algorithms were demonstrated to work really well in several settings. However, these often require the data to be collected over a long period of time to produce images that have high signal-to-noise ratio (SNR). In the case of live imaging and capturing the movement of molecules like proteins, we would often trade off SNR with increased temporal resolution. In such cases, the noise in the collected data poses a challenge in processing and analysis.

In this paper, we utilize two approaches for astigmatic localization: standard Gaussian curve fitting and a CNN-based model to classify the axial trajectory trend of the vesicles. The first approach provides quantitative results in z -localization. It uses the standard Gaussian fitting approach, which is a common solution for depth estimation. We improve on this by applying a denoising step to get good depth estimates even on noisy images. We also present a second approach that doesn't require denoising. We modelled and tested these approaches using astigmatic spinning disk confocal microscopy images of calibration beads. The contributions of the paper are the application of a CNN-based self-supervised denoising step to address the low SNR in confocal imaging to achieve better

K. delas Peñas is supported by doctoral fellowships from the University of the Philippines [FRASDP] and the Philippine Department of Science and Technology [ERDT].

This work was supported by funding from the Engineering and Physical Sciences Research Council (EPSRC) and Medical Research Council (MRC) [EP/L016052/1], and Wellcome Trust [203285/C/16/Z].

depth estimations through Gaussian fitting and the development of a lightweight CNN-based approach to classify the axial trajectory trend. Lastly, we demonstrate how the depth information extracted from the two approaches can be used to improve the association of localized molecules for protein tracking on images of living epithelial cells of *Drosophila* egg chamber as described in [4].

2. METHODOLOGY

2.1. Asymmetric Gaussian fitting

The estimation of the point spread function using Gaussian fitting on astigmatic conditions is powerful enough to estimate depth with 50 nm precision [1, 5]. In general, the standard deviations of a Gaussian curve along the two lateral axes are modelled with the following equation:

$$f(x, y) = A e^{-\left(\frac{(x - x_0)^2}{2\sigma_x^2} + \frac{(y - y_0)^2}{2\sigma_y^2}\right)} + B \quad (1)$$

where A is the intensity or photon count of the Gaussian peak, x_0 and y_0 are the spatial coordinates of the peak, σ_x and σ_y are the standard deviations along the x and y axes, and B is an offset term for the background fluorescence.

In our experiments, we imaged living epithelial cells of *Drosophila* egg chambers using a spinning disk confocal microscope with 0.97 radians peak-to-peak of astigmatism. We adapted the Gaussian fitting method to use the x - y localization obtained from the approach described in [4]. For each x - y localization, we extracted a 16×16 region of interest, with the localized spot in the center, and individually fitted Gaussian curves to obtain the values for σ_x and σ_y .

Next, using the calibration beads data and through nonlinear least square optimization using the Levenberg-Marquardt algorithm as provided in *scipy* [6] optimization library, we model a defocus curve described in [5] as follows:

$$\frac{\sigma_x}{\sigma_y} = \frac{\sqrt{1 + \left(\frac{z-c_x}{d}\right)^2 + A_x \left(\frac{z-c_x}{d}\right)^3 + B_x \left(\frac{z-c_x}{d}\right)^4}}{\sqrt{1 + \left(\frac{z-c_y}{d}\right)^2 + A_y \left(\frac{z-c_y}{d}\right)^3 + B_y \left(\frac{z-c_y}{d}\right)^4}} \quad (2)$$

where d is the depth of focus, c_x and c_y are the lateral offsets, and A_x , A_y , B_x , B_y are coefficients of higher order terms to correct for non-ideal optics.

One challenge in this approach, however, is the amount of noise in regions of interest. Fitting a 2D Gaussian on these noisy images yields inaccurate values of σ_x and σ_y and therefore produces poor estimates of the axial coordinates. To resolve this issue, we employed a deep learning algorithm to denoise the noisy confocal images. Noise2Self [7] is a self-supervised CNN-based algorithm that has been demonstrated to successfully denoise natural and microscopy images. In the absence of ground truth reference data for the confocal

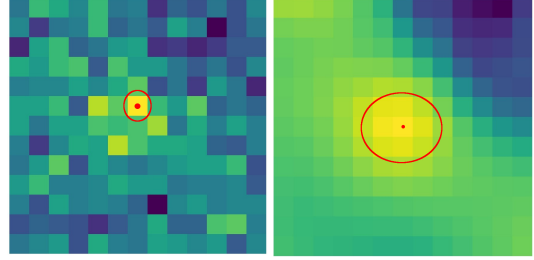


Fig. 1: Using Noise2self. (a) Protein localized with the bright spot at the center. (b) Noise2Self denoising result of the same protein.

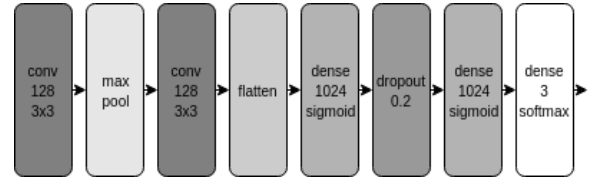


Fig. 2: Custom CNN architecture used for the classification of axial trajectory trend

images, a self-supervised technique like Noise2Self is a suitable method for denoising. This algorithm assumes the statistical independence of the noise across pixels to calibrate the hyperparameters of a median filter or CNNs for denoising. A model consisting of the original implementation of Noise2Self and a Densely Connected Convolutional Network (DCN) [8] as backend was used to suppress noise from the confocal microscopy data prior to Gaussian fitting. This additional step resulted in a more accurate localization of the spots and a better estimation of the Gaussian curves. Figure 1 illustrates one example of a localized protein before and after denoising.

2.2. Classifying protein axial movement trend using CNN

The second approach gives an indirect but simple measure of depth that does not require any kind of denoising. For some purposes, the trend in the axial trajectory (upward, downward, constant) is already sufficient in adding useful information for improving tracking accuracy. With this, we formulate the protein axial tracking as a three-class classification problem. Using the x - y localization from [4], we obtain 16×16 regions of interest, centered on the localized spots. For each localized protein, we stack the ROIs from three successive frames to form a three-channel image that serves as input to the CNN classifier trained for this task.

For the CNN classifier, we utilized a lightweight custom architecture, shown in Figure 2. We trained this network implemented in Keras using the following parameters: Adam optimizer with learning rate ($10e-5$), batch size of 32, 1000 epochs, and categorical cross entropy as loss function.

The training data is obtained from calibration beads

videos acquired with the same spinning disk confocal microscope setup as the vesicle tracking movies. We simulated the upward, downward, and constant movement trend by getting sets of three frames, each from a different depth in the z -stack, and stacked the frames to construct three-channel input. Overall, we have constructed 64,092 three-channel input data and used a 50-50 train-test split. To make the model more robust, we performed data augmentation through addition of Gaussian noise with μ of 5.0 and σ of 10.0.

3. RESULTS AND DISCUSSION

In this section, we evaluated the performance of the two approaches. We used the original calibration beads data imaged with a spinning disk confocal microscope with astigmatism. Z -stacks of the calibration beads were taken with 31 steps and 50 nm step size. Two z -stacks were obtained, one imaged with 1/1000th exposure time relative to the other, yielding one with low SNR and one with high SNR.

3.1. Asymmetric Gaussian fitting

To validate the results for the Gaussian fitting with denoising approach and to ensure that no artifact is introduced in denoising and the measurements obtained are correct, we tested the same denoising model on the low SNR z -scan of calibration beads with known axial coordinates. The calibration beads were imaged with the same conditions as the protein data, with the same camera frame rate to obtain similar level of noise. We observed that the denoising approach preserves the overall structure of the spots. As can be seen in Figure 3, denoising before Gaussian fitting yields better estimated ratio of σ_x/σ_y when compared with noisy data. Without denoising, the x - y standard deviation ratios we obtain from fitting are almost constant, regardless of depth. This indicates that we cannot obtain useful depth information as the noise obscures the fluorescent signal, but with the denoising step, we can decrease the amount of noise that will result to a better fit. Quantitatively, adding the denoising step translated to a ten-fold decrease (from 0.0110 to 0.0013) in the mean squared error (MSE) of the fit with respect to the high SNR data. However, the problem still remains towards the ends of the z range (± 500 nm and further), far away from the focal plane, where we observed an MSE of 0.1361, but this can be attributed to increased loss of signal in depths very far from focus.

3.2. Classifying protein axial movement trend using CNN

Overall, the CNN-based approach to classify the axial trajectory trend provides very encouraging results. With an early stopping criterion set with minimum improvement delta of $10e-7$, training converged early at epoch 40, with 0.0361 loss and 98.91% accuracy. We evaluated the model with the remaining 50% of the data, and found the model performing

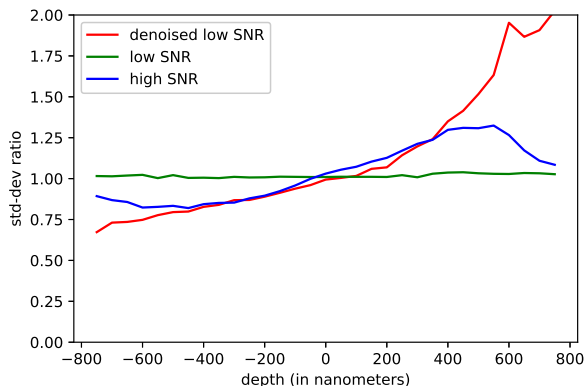


Fig. 3: Gaussian fitting on denoised low SNR data resulted in better depth approximations.

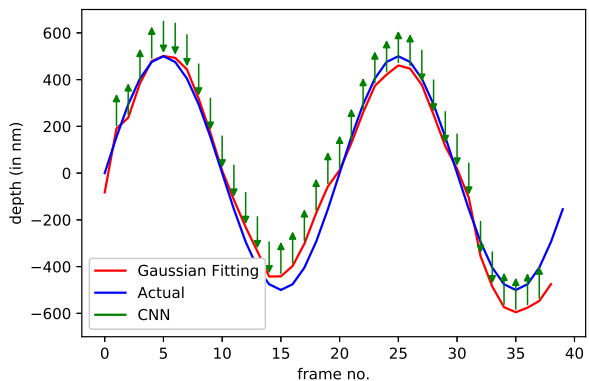


Fig. 4: Test results on calibration beads fixed on x - y with the same stage moved in a sinusoid manner in z .

well at 98.64% test accuracy. To further evaluate this approach, we captured additional calibration beads data but this time, instead of doing a z -scan, the sample stage was moved in a sinusoid manner in the z direction. This experimental data simulates the up and down motion in the z -direction of localized molecules. Testing the CNN model on this data, as shown in Figure 4, the trained model accurately predicted the motion of the sample stage.

3.3. Using extracted depth information in protein tracking

Lastly, we illustrate the advantage of the proposed depth estimators to improve tracking accuracy. Here, we utilize the approach presented by Dmitrieva and collaborators [4] which introduced a two-step process for data association. Firstly, a set of short tracks (tracklets) are formed based on the distance between the detections. Secondly, the tracklets are linked to form final tracks. The linking is based on the

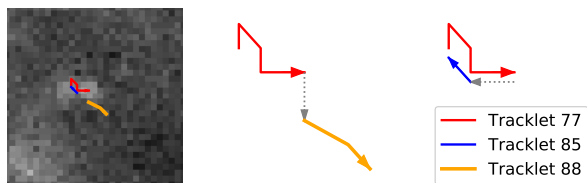


Fig. 5: Some tracklets generated from [4]. Tracklets 85 and 88 are possible connections to track 77. With the Bayesian Network in [4], tracklet 85 is more likely to be connected.

| | TRACKLET 88 | | TRACKLET 85 | |
|-------------|---------------|------------|---------------|---------------|
| | without depth | with depth | without depth | with depth |
| TRACKLET 77 | 0.8536 | 0.8536 | 0.9325 | 0.9625 |

Table 1: Connectivity matrix for tracklet 77 with and without depth depth estimation for tracklets 88 and 85.

tracklets’ parameters and defined by the connectivity score. The score represents a probability of the tracklet pair to be connected and is computed by inference over a Bayesian network (BN).

We applied the tracking approach on the protein data with astigmatism and compared the connectivity scores for tracklets with and without the extracted depth information using the two proposed approaches. Intuitively, we expect an increase in the likelihood of connectivity for tracklets whose ends are more or less at the same depth or following the same trajectory trend. Figure 5 illustrates an example of the track-linking task with three tracklets to be considered. Using the originally proposed tracking solution without depth information, tracklet 77 is linked with tracklet 85 (and not 88) with score of 0.9325. Using the approaches presented in this paper, we estimated tracklet 77 to be moving downwards (647 nm depth) and connecting tracklet 85 continues this trend (679 nm depth) but not tracklet 88 (-274 nm depth). As presented in Table 1, if we include this depth information and modify the BN topology in [4] to incorporate extracted depth information, we still see similar patterns in the connectivity matrix for tracklet 77 but with an increase in the magnitude on the tracklet to be connected. Adding depth information from the proposed approaches strengthens this track linking process by increasing the confidence in connectivity. For cases with close connectivity scores, the depth information may be crucial in correctly linking tracks.

4. CONCLUSION

Two approaches for extracting depth from confocal microscopy data with astigmatism are presented. The first approach uses the standard and widely accepted Gaussian

fitting method. We added a denoising preprocessing step and demonstrated that this additional step resulted in better depth estimates. The second approach was also able to give useful axial information. We have shown that even with a simple architecture, this information can be obtained without the need for denoising.

The CNN-based approach could be extended to finer-grained classes. For example, quantitative evaluation of the movement in z -axis, as an extension of the axial trajectory trend. This would give insight on the speed by which the molecules move in the z direction. For both approaches, their integration with existing 2D localization and tracking algorithms like in [4] and further validation tests may provide better models to understand underlying subcellular mechanisms.

5. REFERENCES

- [1] S. Holden, S. Uphoff, and A. Kapanidis, “Daostorm: An algorithm for high-density super-resolution microscopy,” *Nature methods*, vol. 8, pp. 279–80, 04 2011.
- [2] M. Ovesný, P. Křížek, J. Borkovec, Z. Švindrych, and G. M. Hagen, “ThunderSTORM: a comprehensive ImageJ plug-in for PALM and STORM data analysis and super-resolution imaging,” *Bioinformatics*, vol. 30, pp. 2389–2390, 05 2014.
- [3] D. Sage, T.-a. Pham, H. Babcock, T. Lukes, T. Pengo, J. Chao, R. Velmurugan, A. Herbert, A. Agrawal, S. Colabrese, A. Wheeler, A. Archetti, B. Rieger, R. Ober, G. Hagen, J.-B. Sibarita, J. Ries, R. Henriques, M. Unser, and S. Holden, “Super-resolution fight club: Assessment of 2d 3d single-molecule localization microscopy software,” *Nature Methods*, vol. 16, 05 2019.
- [4] M. Dmitrieva, H. L. Zenner, J. Richens, D. S. Johnston, and J. Rittscher, “Protein tracking by cnn-based candidate pruning and two-step linking with bayesian network,” in *IEEE International Workshop on Machine Learning for Signal Processing*, 2019.
- [5] B. Huang, W. Wang, M. Bates, and X. Zhuang, “Three-dimensional super-resolution imaging by stochastic optical reconstruction microscopy,” *Science (New York, N.Y.)*, vol. 319, pp. 810–3, 03 2008.
- [6] E. Jones, T. Oliphant, P. Peterson, *et al.*, “SciPy: Open source scientific tools for Python,” 2001–. [Online; accessed 7 October 2019].
- [7] J. Batson and L. Royer, “Noise2self: Blind denoising by self-supervision,” *CoRR*, vol. abs/1901.11365, 2019.
- [8] G. Huang, Z. Liu, and K. Q. Weinberger, “Densely connected convolutional networks,” *CoRR*, vol. abs/1608.06993, 2016.