NEURAL-NETWORK MODEL FOR CHARACTERIZING STOCHASTIC DYNAMIC VARIABILITY OF CELLULAR PARTICLES

M Doggaz, Anne Brelot, Jean-Christophe Olivo-Marin, Thibault Lagache, Giacomo Nardi

NEURAL-NETWORK MODEL FOR CHARACTERIZING STOCHASTIC DYNAMIC VARIABILITY OF CELLULAR PARTICLES.

M. Doggaz⋆, A. Brelot†, J.-C. Olivo-Marin⋆, T. Lagache⋆, G. Nardi⋆

⋆ Institut Pasteur, Universite de Paris Cité, CNRS UMR 3691, BioImage Analysis Unit, Paris, France

† Institut Pasteur, Université de Paris Cité, CNRS UMR 3691, Dynamics of Host-Pathogen Interactions Unit, Paris, France

Corresponding author: giacomo.nardi@pasteur.fr

ABSTRACT

Particle dynamics characterization is fundamental for understanding the biophysical laws orchestrating cellular processes. To classify the dynamic behaviors governing biological particles, we develop a neural network model built on geometric descriptors of trajectories. The model infers the stochastic laws governing the trajectory, enabling the detection of a large family of dynamic behaviors, especially within the subdiffusive regime that characterizes cell signaling processes. Finally, we propose a framework to robustly detect dynamic changes in composed trajectories based on the variability of prediction scores on successive sub-trajectories. The method is validated on simulated composed trajectories simulating the activation pathway of receptors CCR5.

1. INTRODUCTION

Cell processes are constrained by biological particles' dynamic interactions (e.g., molecules, receptors, viruses), which can be observed by coupling fluorescence microscopy and tracking algorithms. Statistical methods for motion analysis are needed thereafter to characterize particle dynamics and decipher the orchestration of cellular signaling [1]. However, the dynamics of particles at the subcellular scale reveal a complex landscape [2] highlighting two main challenges for classification algorithms: inferring a large variety of dynamic behaviors and detecting dynamic changes along trajectories.

The dynamic of biological particles is driven by several biophysical laws inducing confined evolutions towards equilibrium points or trapping phenomena due to particle interactions [3]. Thus, appropriate approaches are needed to distinguish consistent dynamics and establish the link to biophysical constraints. A first improvement in this direction has been made in [4], which develops a machine-learning approach, based on geometric descriptors, to infer the stochastic laws governing trajectories.

However, at a larger temporal scale, the particle dynamics are driven by recurrent dynamic changes. This is due to particle interactions or cell environmental constraints, and motion change detection highlights the role of such constraints in cellular processes. This is the case, for instance, of receptor-ligand-based signaling systems where motion switching governs the activation pathway [3, 1]. Nevertheless, detecting motion switching remains a difficult and little-explored problem and the few existing works are limited to the diffusion framework [5, 6].

This work tackles the previous challenges all at once. We redesign the method proposed in [4] considering an additional motion class and obtaining a 6-class classification algorithm (BM, OU, DIR, FBM sub- and super-diffusive, CTRW) based on four geometric features computed on entire trajectories. Moreover, we replace the Random Forest supervised approach, used in [4], with a neural network model resulting in a more accurate inference of motions. A main asset of neural network models is their ability to output prediction scores accounting for the confidence in process prediction. Accordingly, we define a novel framework to detect motion changes based on prediction scores computed on sub-trajectories. In particular, we prove that the minimum score robustly highlights switching times. The approach is validated on simulated composed trajectories showing the robust detection of motion changes and pointing out the usefulness of neural network approaches for motion analysis.

2. METHODS

2.1. Simulating cellular dynamics

A discrete trajectory is a set of successive positions over time

$$\mathbf{X} = (X_{t_1}, ..., X_{t_N})$$

where $X_i \in \mathbb{R}^2$ with independent components and the time interval between successive positions is constant. Discretizing 2-dimensional continuous stochastic processes generates examples of trajectories with different dynamic and diffusion behaviors. We consider the main stochastic process involved in cellular dynamics [2] and simulate extended datasets of related trajectories.

Brownian motion (BM). The Brownian process B_t^σ describes particles freely moving and verifying $(B_t^\sigma - B_s^\sigma) \sim$

Process	Parameters
BM	$\sigma \in [0.1, 10]$
OU	$\sigma \in [1, 10], \theta = (0, 0), \lambda = r\sigma\,,\ r \in [0.2, 1]$
DIR	$\sigma \in [1, 10], \mu = r\sigma\,,\ r \in [0.2, 1]$
FBMsub	$\mathrm{H} \in [0.15, 0.3]$
FBMsuper	$\mathrm{H} \in [0.7, 0.85]$
CTRW	$\sigma \in [0.1, 10]\,,\ \gamma \in [0, 1]$

Table 1. Parameters used for trajectory simulation used in this work for training and validation steps (1000 trajectories are simulated for each process).

$\mathcal{N}(0, \sigma^2(t-s)\mathbf{I}_2)$ for every $t > s$ where $\mathbf{I}_2$ denotes the 2-dimensional identity matrix, and the so-called diffusion coefficient σ accounts for displacement amplitude.

Directed Brownian motion (DIR). This process describes particles driven by a constant drift component $\mu \in \mathbb{R}^2$ giving an oriented input to each random displacement:

$$dX_t = \mu dt + dB_t^\sigma\,.$$

Ornstein-Uhlenbeck process (OU). This process describes confined motion due to restoring forces and it satisfies

$$dX_t = -\lambda(X_t - \theta)dt + dB_t^\sigma \tag{1}$$

where θ is the equilibrium point and λ weights the drift term.

Fractional Brownian Motion (FBM). It governs particles evolving in constrained or crowded environments resulting in dependent displacements verifying

$$E[(X_t - X_s)^2] = |t-s|^{2H}$$

where H is the Hurst parameter. For $H = 1/2$, this process reduces to Brownian motion, while for $H < 1/2$ and $H > 1/2$ it exhibits a sub- (**FBMsub**) and super-diffusive (**FBMsuper**) behavior, respectively.

Continuous-Time Random Walk (CTRW). This process describes trapping phenomena resulting in trajectories alternating jumps, following a Gaussian distribution $\mathcal{N}(0, \sigma^2)$, and waiting times, modeled in this work by a power law distribution on $\psi(t) = t^{-\gamma-1}$ with $\gamma \in [0, 1]$ and $t \geq 1$.

Simulated trajectories. We perform path simulation for the six processes presented above (BM, OU, DIR, FBMsub, FBMsuper, CTRW) generating 1000 trajectories of a given length N for each process using different parameters (see Table 1). Several datasets are generated for the training and validation steps performed in the rest of the paper.

2.2. Motion descriptors

The proposed method builds on four geometric features to distinguish the process governing the trajectory. This section details the two groups of features used in this work that generalize the features introduced in [4].

Directional Persistence. We consider the distribution of angles between successive points $(X_{t_i}, X_{t_{i+1}}, X_{t_{i+2}})$ defined as

$$\theta_{\mathbf{X}}(t_i) = \pi - \angle X_{t_i} X_{t_{i+1}} X_{t_{i+2}}$$

where counterclockwise (with respect to $X_{t_{i+1}} - X_{t_i}$) angles are considered as positive [7]. As shown in [4], the histogram of angles $\{\theta_{\mathbf{X}}\}$ exhibits different properties depending on the process: BM holds a uniform-like distribution of angles, CTRW has a majority of null angles because of the waiting times, confined motions (OU, FBMsub) show more large angles. The analytical properties of angle distributions can be described by two features. The first one accounts for the coefficient a of a convex fitting of the distribution:

$$p_{\theta_{\mathbf{X}}}(x) \propto ax^2\,. \tag{2}$$

The second one, accounting for a preferred motion direction along the trajectory, considers the index of directionality $P_d(\mathbf{X})$ defined as:

$$P_d(\mathbf{X}) = \mathbb{P}(|\theta_{\mathbf{X}}| < \pi/2) - \mathbb{P}(|\theta_{\mathbf{X}}| \geq \pi/2)\,. \tag{3}$$

Spreading trend. The second set of features describes the spreading properties of trajectories reflecting how the particle unfurls into space. We consider the Ripley's index K_r in a ball $B(X_{t_1}, r)$ of radius r centered at the starting point:

$$K_r = |\{X_{t_i} \in \mathbf{X} \,|\, X_{t_i} \in B(X_{t_1}, r)\}| \,/\, N\,,$$

which allows the definition of the following vector:

$$K_{\mathbf{X}} = (K_R, ..., K_{N*R})\,,\ R = \frac{1}{N}\sum_{i=1}^{N-1} \|X_{t_{i+1}} - X_{t_i}\|$$

describing the evolution of the trajectory through balls of increasing radius. The vector $K_{\mathbf{X}}$ exhibits an increasing behavior converging to 1 for large K [4] justifying the introduction of the feature b defined by the following fitting:

$$K_{\mathbf{X}} \propto 1 - e^{-br}\,. \tag{4}$$

Finally, to account for (local) confinement resulting in constant Ripley's index on successive concentric balls, we estimate the probability that the discrete derivative $K'_{\mathbf{X}}$ (computed by finite differences) is zero:

$$P_p(\mathbf{X}) = \mathbb{P}(K'_{\mathbf{X}} \neq 0)\,. \tag{5}$$

2.3. Neural-network classification model

The presented method builds on a neural-network architecture, using as inputs the four features previously introduced (see equations (2), (3), (4), and (5)) computed on entire trajectories, and learning the related process label (BM, OU, DIR, FBMsub, FBMsuper, CTRW). The architecture π used for the model (see Fig. 1) contains a first entry-level whose

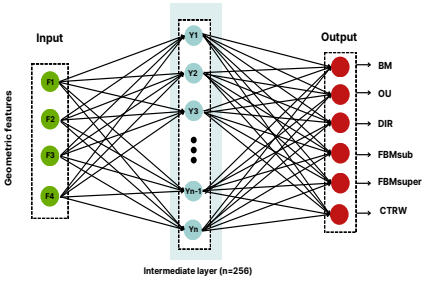

Fig. 1. Neural-network architecture π:4x256x6 used to learn processes based on geometric features.

neurons correspond to the four features, an intermediate layer followed by a ReLu activation function, and a classification layer composed of six neurons corresponding to the six motion classes and followed by a SoftMax function. Training and validation are performed in PyTorch on simulated trajectories with length $N = 100$, and training uses the Adam minimization algorithm with an EarlyStopping to optimize the number of needed epochs. The selected model (π: 4x256x6) guarantees the best accuracy of 95.5% compared to two more elaborate architectures (4x512x256x128x6 and 4x256x128x64x32x6 give an accuracy of 94.7 % and 94.9%, respectively). This proves, in particular, the pertinence of selected features to distinguish the studied processes which reduces the need of complex neural settings.

2.4. Model calibration

The model π is calibrated to ensure that the predicted probabilities correctly estimate the rate of correct predictions [8]. Formally, π associates each feature vector x_i to a process label $c_i \in C$ ($|C| = 6$) based on a prediction score p_i, $\pi(x_i) = (c_i, p_i)$. The prediction score is computed using the SoftMax function and the network logits z_i:

$$p_i = \max_{c=1,\dots,6} \sigma_{SM}(z_i)^c \,,\ \sigma_{SM}(z_i)^c = \frac{e^{z_i^c}}{\sum_{k=1}^{6} e^{z_i^k}}.$$

Given a correctly labeled dataset, composed of feature vectors and process labels $\{(x_i, c_i)\}$ corresponding to simulated trajectories, a calibrated network should verify:

$$\mathbb{P}(c = c_i \,|\, x = x_i) = p \quad \forall p \in [0, 1]\,.$$

Calibration is performed by Temperature Scaling which consists of scaling the logits by an optimal parameter T:

$$\hat{p}_i = \max_{k=1,\dots,6} \sigma_{SM}(z_i/T)^k$$

computed by optimizing the Negative Log Likelihood on correctly annotated dataset $\{(x_i, c_i)\}$:

$$T^* = \underset{T>0}{\operatorname{argmin}} - \sum_{k=1}^{6} \log \sigma_{SM}(z_i/T)^k\,.$$

The optimal T is thus computed in a post-processing mode by training the network with respect to T on a correctly annotated dataset $\{(x_i, c_i)\}$. We point out that the Temperature Scaling approach enables adjusting the logit vectors z_i without affecting the model accuracy.

3. RESULTS

3.1. Method Evaluation

The model π selected in Section 2.3 is validated using training and validation sets of trajectories defined according to Table 1.

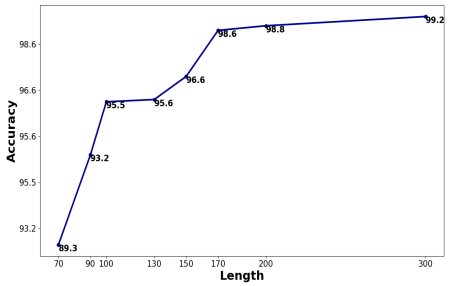

Fig. 2. Accuracy of model π for trajectories with different lengths. An accurate model needs at least paths composed of 100 time points.

Fig. 2 displays the accuracy for trajectories with different lengths, pointing out that the intrinsic properties of motion, described by the geometric features, need time to take place and be robustly distinguishable by statistical methods. These results show that we need to work with trajectories with at least $N = 100$ to exceed an accuracy of 95%, reaffirming the difficulty of classifying short paths.

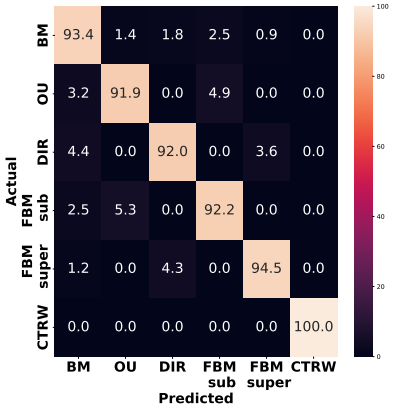

Fig. 3. Confusion matrix for the model π validated on trajectories of length $N = 100$.

The confusion matrix for the model π is shown in Fig. 3,

proving that the proposed approach enables the accurate distinction of different types of motion within the same diffusion class. This is a strong asset for detecting different dynamic behaviors finely and also proves that the presented features are appropriate to capture the dynamic characteristics of different processes.

3.2. Motion switching detection.

In this section, we exploit the prediction scores produced by the model π to detect motion changes along trajectories. We simulated 5000 trajectories by stitching two paths of length 150: the first trajectory $\mathbf{X}^1$ simulates a CTRW process and another process generates the second path $\mathbf{X}^2$. We point out that this dynamic transition is specific to the activation pathway of receptors CCR5 [3, 1]. Each composed trajectory $(X_1^1, ..., X_{150}^1, X_1^2, ..., X_{150}^2)$ is analyzed using the sliding window technique enabling the inference of the process and prediction score based on model π for every window. The window size is set equal to 100 to guarantee the high accuracy of the model π trained with $N = 100$. For windows starting after X_{50}^1 the model infers paths composed of two dynamics: a lower prediction score is expected with a minimum value reached for the windows starting at X_{100}^1 containing both dynamics equally. Fig. 4 shows the starting times distribution for windows reaching the lowest prediction scores. This result indicates that unconfident predictions robustly correlate with motion composition, confirming that the minimum prediction score within a sliding window approach is a reliable parameter to detect the composed nature of trajectories and the related switching times.

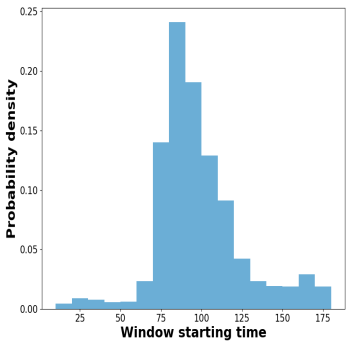

Fig. 4. Distribution of starting times of sub-trajectories with lowest prediction scores. Starting at 100 composition is at 50% explaining low scores.

4. CONCLUSION

The developed method infers the stochastic laws governing particle motion based on geometric descriptors of trajectories. The model builds on a neural network architecture resulting in an accurate classifier of six classes of motion that is a breakthrough in the field. Moreover, we propose a framework to highlight motion composition based on the prediction scores provided after network calibration. In particular, the analysis of prediction score variability on sub-trajectories enables robust detection of switching times, showing the potential of neural network approaches in motion analysis.

5. REFERENCES

[1] D. Calebiro, F. Rieken, J. Wagner, T. Sungkaworn, U. Zabel, A. Borzi, E. Cocucci, A. Zürn, and M. J. Lohse, "Single-molecule analysis of fluorescently labeled g-protein–coupled receptors reveals complexes with distinct dynamics and organization," *Proceedings of the National Academy of Sciences*, vol. 110, no. 2, pp. 743–748, 2013.

[2] V. Briane, M. Vimond, and C. Kervrann, "An overview of diffusion models for intracellular dynamics analysis," *Briefings in Bioinformatics*, vol. 21, no. 4, pp. 1136–1150, 2020.

[3] F. Momboisse, G. Nardi, P. Colin, M. Hery, N. Cordeiro, S. Blachier, O. Schwartz, F. Arenzana-Seisdedos, N. Sauvonnet, J.-C. Olivo-Marin, *et al.*, "Tracking receptor motions at the plasma membrane reveals distinct effects of ligands on ccr5 dynamics depending on its dimerization status," *Elife*, vol. 11, p. e76281, 2022.

[4] M. S. Sano, A. Brelot, J.-C. Olivo-Marin, T. Lagache, and G. Nardi, "Motion classification based on geometrical features of trajectories," in *2024 IEEE International Symposium on Biomedical Imaging (ISBI)*, pp. 1–4, IEEE, 2024.

[5] V. Briane, M. Vimond, C. A. Valades-Cruz, A. Salomon, C. Wunder, and C. Kervrann, "A sequential algorithm to detect diffusion switching along intracellular particle trajectories," *Bioinformatics*, vol. 36, no. 1, pp. 317–329, 2020.

[6] K. Hubicka and J. Janczura, "Time-dependent classification of protein diffusion types: A statistical detection of mean-squared-displacement exponent transitions," *Physical Review E*, vol. 101, no. 2, p. 022107, 2020.

[7] N. Gal, D. Lechtman-Goldstein, and D. Weihs, "Particle tracking in living cells: a review of the mean square displacement method and beyond," *Rheologica Acta*, vol. 52, pp. 425–443, 2013.

[8] C. Guo, G. Pleiss, Y. Sun, and K. Q. Weinberger, "On calibration of modern neural networks," in *International conference on machine learning*, pp. 1321–1330, PMLR, 2017.