End-to-end neural reconstruction of DNA structures from single-frame fluorescence images

Fushuai Wang [1 2 3] **Qianxiao Li** [3 4] **Beatrice W. Soh** [1 3]

[1]*Department of Chemical and Biomolecular Engineering, National University of Singapore, Singapore* [2]*Department of Materials Science and Engineering, National University of Singapore, Singapore* [3]*Institute for Functional Intelligent Materials, National University of Singapore, Singapore* [4]*Department of Mathematics, National University of Singapore, Singapore*. Correspondence to: Beatrice W. Soh bwysoh@nus.edu.sg.

1. Introduction

Directly extracting biomolecular structures from noisy fluorescence microscopy images remains a fundamental challenge across the physical and life sciences. Conventional image analysis techniques such as thresholding and skeletonization tend to fail when the signal-to-noise ratio is low. In this work, we propose a deep learning framework for reconstructing the spatial conformations of DNA from single-frame fluorescence microscopy images. Our approach integrates a deep convolutional neural network encoder with a graph-based decoder, explicitly capturing the sequential continuity and geometric constraints of DNA. By incorporating distance-aware structural representations into the learning process, the model robustly maps noisy, resolution-limited fluorescence images to high-fidelity unordered DNA coordinates. An ordering network is then employed to connect and reorder these unordered coordinates, recovering a physically consistent and ordered DNA chain structure, akin to the output from coarse-grained simulations. This enables an end-to-end reconstruction paradigm that directly infers detailed DNA spatial conformations from fluorescence images. Furthermore, the proposed framework can serve as a structure-aware image feature extractor that provides observation spaces and reward signals for reinforcement learning agent, enabling the exploration of structure-guided DNA conformation regulation and shape manipulation tasks. This work offers a new AI-driven solution for DNA structure analysis under single-molecule imaging conditions and lays the foundation for downstream intelligent decision-making applications.

2. Methodology

2.1 Related work

Traditional filament tracing pipelines based on segmentation/skeletonization or active-contour models have been widely used to extract polymer-like structures; however, their performance often degrades under common fluorescence microscopy conditions such as low signal-to-noise ratio, severe blur, and structural overlap [1, 2]. In recent years, researchers have attempted to use deep learning methods to infer DNA conformations from fluorescence images [3]. Nevertheless, existing studies have mostly focused on recovering coarse 2D DNA contours under relatively stretched configurations, and the exploration of complex conformations (e.g., coils, folds, and partially stretched states) or high-precision coordinate recovery remains limited.

2.2 Simulation data generation with PSF

To obtain large-scale paired training data, we construct a physics-grounded simulation-to-image pipeline that generates fluorescence images from simulated DNA conformations.

DNA conformation simulation. The Langevin equation describes the motion of a DNA chain in solution [4]:

$$\frac{\mathrm{d}\mathbf{r}_i}{\mathrm{d}t} = \mathbf{u}(\mathbf{r}_i) + \frac{1}{\zeta}\left(\mathbf{F}^{\mathrm{ev}} + \mathbf{F}^{\mathrm{c}} + \mathbf{F}^{\mathrm{br}}\right), \tag{1}$$

where $\mathbf{r}_i$ is the position of bead i, $\mathbf{u}(\mathbf{r}_i)$ is the background flow velocity, ζ is the drag coefficient of a single bead, and $\mathbf{F}^{\mathrm{ev}}$, $\mathbf{F}^{\mathrm{c}}$, and $\mathbf{F}^{\mathrm{br}}$ denote the excluded volume, constraint, and Brownian forces, respectively.

We employed a predictor–corrector scheme [5] to determine the position of each bead at every time step. The enforcement of rigid-rod constraints leads to a system of nonlinear equations for the rod tensions T_i, which we solved using Newton's method.

PSF-based fluorescence image formation. To synthesize realistic fluorescence microscopy images, we adopt the Gibson–Lanni point spread function (PSF) model, which describes 3D imaging of fluorescent structures under a widefield microscope [6]. To reduce computational cost, we employ a fast approximation of the Gibson–Lanni model by expanding the original integral formulation into a truncated series of Bessel basis functions [7]:

$$\mathrm{PSF}(r, z, z_p, \mathbf{p}) \approx \left| A \sum_{m=1}^{M} c_m(z)\, R_m(r, \mathbf{p}) \right|^2, \tag{2}$$

where r is the radial coordinate in the image plane, z is the axial coordinate of the focal plane, z_p is the axial location of the point source in the specimen layer relative to the coverslip, and $\mathbf{p} = (NA, \mathbf{n}, \mathbf{t})$ denotes the microscope parameter vector. Here A is a normalization constant, $c_m(z)$ are fitted coefficients, and $R_m(r, \mathbf{p})$ are precomputed radial Bessel basis functions.

2.3 Model overview

Let I denote a single-frame fluorescence image. A CNN encoder extracts multi-scale feature maps from I. These features are used to infer an unordered set

of N DNA coordinates $\{\hat{\mathbf{x}}_i\}_{i=1}^{N}$ via a graph-based decoder that performs message passing between predicted nodes. To encode polymer-specific structure, we incorporate distance-aware representations into the graph decoder, allowing the model to emphasize locally consistent connectivity under strong image blur. Finally, an ordering network takes the unordered coordinates as input and predicts an ordered chain sequence, yielding ordered DNA coordinates $\hat{\mathbf{X}} = [\hat{\mathbf{x}}_{(1)}, \ldots, \hat{\mathbf{x}}_{(N)}]$ suitable for downstream physical readouts. The full pipeline is illustrated in Fig. 1.

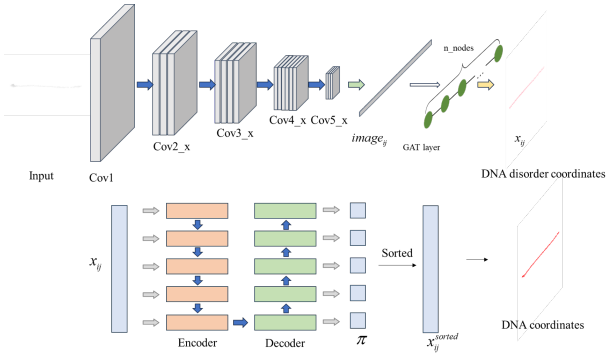

Fig. 1: Overview of the proposed single-frame DNA reconstruction framework

2.4 Results

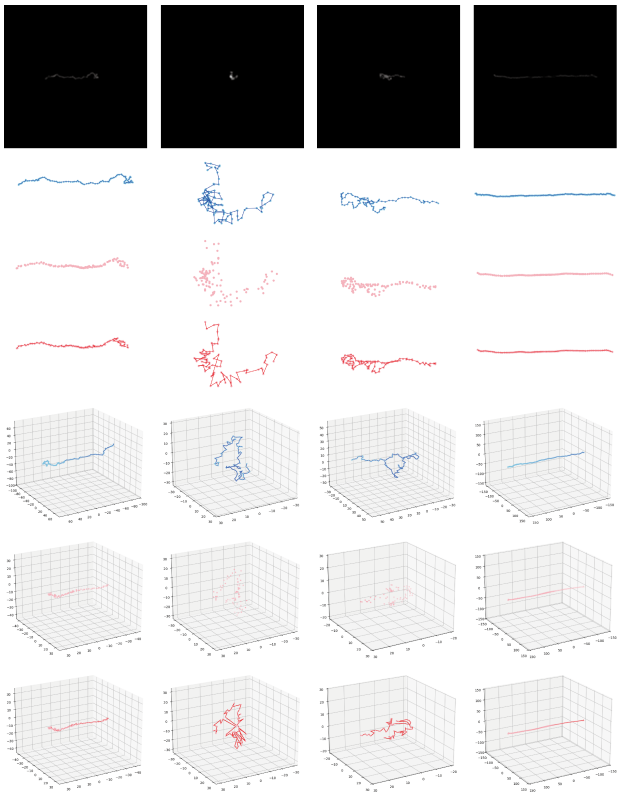

Fig. 2: Qualitative comparison between ground truth and predictions across diverse simulated DNA conformations

We evaluate our model on both synthetic simulation data and real experimental microfluidic data. Fig. 2 presents qualitative comparisons on diverse simulated DNA conformations, where the first row shows the input fluorescence images, blue points denote the ground-truth coordinates, and red points indicate model predictions.

Our method accurately reconstructs DNA contours across stretched, coiled, and intermediate states, demonstrating robust topology-consistent recovery from single fluorescence frames.

We further evaluate the model on experimental fluorescence images of DNA molecules as they stretch under electric fields in microfluidic channels. Fig. 3 shows representative reconstruction results for DNA under stretching conditions. Despite the domain gap between simulation and experiment (e.g., imaging noise, blur, and intensity variation), our method remains effective in reconstructing DNA stretching states from experimental single-frame images, demonstrating strong generalization capability in realistic microfluidic settings.

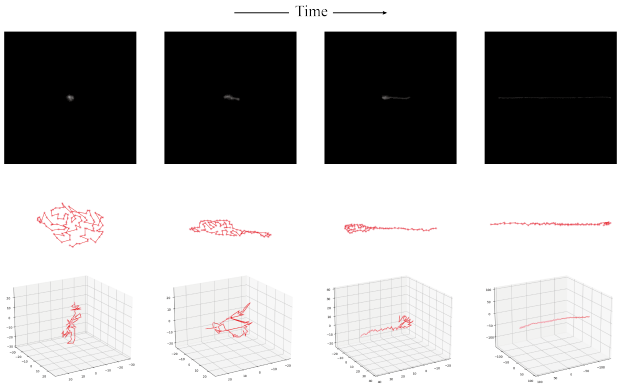

Fig. 3: Reconstruction on real microfluidic fluorescence images demonstrates simulation-to-experiment generalization under stretching conditions

Acknowledgments

The computational work for this article was partially performed on resources of the National Supercomputing Centre, Singapore. Fushuai Wang was supported by a research scholarship from Prof. Kostya Novoselov at the National University of Singapore. The authors acknowledge funding from the National University of Singapore Presidential Young Professorship start-up grant (Grant No. A-0010260-00-00).

References

[1] Matthew B Smith, Hongsheng Li, Tian Shen, Xiaolei Huang, Eddy Yusuf, and Dimitrios Vavylonis. Segmentation and tracking of cytoskeletal filaments using open active contours. *Cytoskeleton*, 67(11):693–705, 2010.

[2] Mathew Schneider, Alaa Al-Shaer, and Nancy R Forde. Autosmartrace: Automated chain tracing

and flexibility analysis of biological filaments. *Biophysical Journal*, 120(13):2599–2608, 2021.

[3] Maged F Serag, Maram Abadi, Hajar Al-Zarah, Omar Ibrahim, and Satoshi Habuchi. Deep learning and single-molecule localization microscopy reveal nanoscopic dynamics of dna entanglement loci. *ACS nano*, 19(6):6236–6249, 2025.

[4] Xiaoli Chen, Beatrice W Soh, Zi-En Ooi, Eleonore Vissol-Gaudin, Haijun Yu, Kostya S Novoselov, Kedar Hippalgaonkar, and Qianxiao Li. Constructing custom thermodynamics using deep learning. *Nature Computational Science*, 4(1):66–85, 2024.

[5] Tony W Liu. Flexible polymer chain dynamics and rheological properties in steady flows. *The Journal of chemical physics*, 90(10):5826–5842, 1989.

[6] Sarah Frisken Gibson and Frederick Lanni. Experimental test of an analytical model of aberration in an oil-immersion objective lens used in three-dimensional light microscopy. *Journal of the optical Society of America A*, 8(10):1601–1613, 1991.

[7] Jizhou Li, Feng Xue, and Thierry Blu. Fast and accurate three-dimensional point spread function computation for fluorescence microscopy. *Journal of the Optical Society of America A*, 34(6):1029–1034, 2017.