scientific reports

OPEN

Accelerating the characterization of dynamic DNA origami devices with deep neural networks

Yuchen Wang[1✉], Xin Jin[2] & Carlos Castro[1✉]

Mechanical characterization of dynamic DNA nanodevices is essential to facilitate their use in applications like molecular diagnostics, force sensing, and nanorobotics that rely on device reconfiguration and interactions with other materials. A common approach to evaluate the mechanical properties of dynamic DNA nanodevices is by quantifying conformational distributions, where the magnitude of fluctuations correlates to the stiffness. This is generally carried out through manual measurement from experimental images, which is a tedious process and a critical bottleneck in the characterization pipeline. While many tools support the analysis of static molecular structures, there is a need for tools to facilitate the rapid characterization of dynamic DNA devices that undergo large conformational fluctuations. Here, we develop a data processing pipeline based on Deep Neural Networks (DNNs) to address this problem. The YOLOv5 and Resnet50 network architecture were used for the two key subtasks: particle detection and pose (i.e. conformation) estimation. We demonstrate effective network performance (F1 score 0.85 in particle detection) and good agreement with experimental distributions with limited user input and small training sets (~ 5 to 10 images). We also demonstrate this pipeline can be applied to multiple nanodevices, providing a robust approach for the rapid characterization of dynamic DNA devices.

Structural DNA nanotechnology is a rapidly growing field that has shown great utility in the bottom-up fabrication of devices and materials with applications spanning areas like nanofabrication[1], nanophotonics[2], molecular computation[3], bioimaging[4], and nanotherapeutics[5]. Over the last several years there has been a surge of interest in dynamic DNA nanotechnology, since the ability to design reconfigurable DNA nanodevices, combined with the ability to interface DNA with a wide range of biomolecules or nanomaterials, is highly attractive for the development of sensors[6], nanorobots[7], tunable plasmonic devices[8], and biophysical measurement tools[9]. Since many of these applications rely on physical interactions with other molecules or materials, understanding the mechanical properties of dynamic DNA structures is crucial to quantitively describe their functions. The most common approach to characterize the mechanical properties of dynamic DNA devices is through imaging (transmission electron microscopy (TEM) or atomic force microscopy (AFM)) to visualize conformational fluctuations. The magnitude of these fluctuations is related to the structure stiffness. However, quantifying these conformations is typically done through manual measurement, which is tedious and often the major bottleneck limiting the characterization pipeline and slowing down the experimentation and overall design and test cycle. Hence, there is a critical need for approaches that facilitate, and ideally automate, rapid characterization of structure conformations for a variety of dynamic DNA nanodevices.

Machine learning is a type of artificial intelligence that enables machines to learn to identify patterns from data and improve their performance on a specific task without being explicitly programmed. It involves the use of algorithms that can be trained (i.e., learn) to make predictions based on current observations. Among several different algorithms, the use of deep neural networks (DNN) has been the dominant approach for a wide range of data problems[10]. Specifically, efforts in both academia and industry[11] are using DNNs to solve challenging real-world problems such as autopilot[12-14], robotics[15-17], speech recognition[18-20], predictive analytics[21-24], and computer vision[25,26]. For example, to date, AlphaFold, which is an algorithm based on DNN, has already provided over 200 million protein structures with high accuracy[27]. In contrast, from traditional X-ray crystallography or cryo-EM, there are only ~ 200 thousand protein structures shared in the protein data bank (PDB). Specific to DNA nanotechnology, recent studies[28-30] showed the DNN is also a great solution for automatically recognizing nanostructure in atomic force microscopy or fluorescence microscopy with high accuracy, which provided a

[1]Department of Mechanical and Aerospace Engineering, The Ohio State University, Columbus, OH 43210, USA. [2]Department of Computer Science and Engineering, The Ohio State University, Columbus, OH 43210, USA. ✉email: wang.9298@osu.edu; castro.39@osu.edu

foundation to solving the DNA nanostructure identification problem. However, these works only demonstrated the feasibility of identifying static nanostructures from images and did not address the need for automated property characterization, which is necessary to overcome the characterization bottleneck for dynamic DNA nanodevices.

Here, we demonstrate a DNN pipeline that can accelerate the analysis of mechanical properties (i.e., flexibility) of dynamic DNA origami nanodevices[31,32]. The pipeline implements two DNNs to facilitate the sampling (i.e., nanostructure identification) and quantification (i.e., conformation measurement) steps. We first establish the approach using a 'Hinge' nanostructure, which is representative of dynamic nanodevices that are widely used for biophysical measurements[9], biosensing[33], and controlling biomolecular interactions[34]. Secondly, we demonstrated the robustness and versatility of this pipeline by applying it to other dynamic DNA origami device characterizations including a 'Hinge-Nucleosome' system[9] and a three-arm device designed to exhibit steric interactions between the arms[35]. Our results suggest that the DNN algorithm can be used to overcome the bottlenecks that require excessive labor work for post-processing of micrographs to characterize dynamic DNA nanodevices, which can greatly facilitate the design, experiment, and development cycle.

We will open the source of the dataset and the code of our pipeline after this paper being published.

Results

Dynamic DNA origami structure analysis workflow.

We selected a DNA origami hinge structure[9] as the basis to develop our DNN pipeline, since hinges are simple dynamic devices that are widely used. In particular, we used a hinge that was recently demonstrated as a useful assay to measure the dynamic properties of biological samples[9] and apply high forces on nanometer scale[36]. This hinge structure (Fig. 1) consists of two arms that are connected by several short single-stranded DNA (ssDNA) connections that form a hinge vertex. The two arms (~70 nm in length) are highly stiff and can be regarded as solid bodies. The vertex is designed to be much more flexible, allowing for rotational motion of the two arms, which is primarily constrained to one degree of rotational freedom. The hinge exhibits preferred angular conformations; hence, it can be regarded effectively as a torsional spring. In order to apply these hinge devices (e.g., to apply forces to biomolecules[9], detect biomolecules[33], or control enzyme interactions[34]), it is critical to understand their mechanical properties. The mechanical property most relevant to the function of the hinge is the torsional stiffness. The torsional properties are typically considered in terms of a rotational free energy landscape, which can be determined from angular conformation distributions through Boltzmann inversion. Hence, a key step to characterizing the mechanical properties of hinge devices is measuring angular distributions.

The most common approach to visualize the conformations of dynamic DNA origami devices is transmission electron microscopy (TEM) imaging[37,38]. Most studies implement negative stain TEM, where dynamic nanodevices are deposited on a surface followed by imaging. The data analysis process to determine the conformational free energy landscape is typically carried out in multiple steps: (1) sampling many hinges (hundreds to thousands) by manual selection from images (i.e. clicking to identify a properly folded and isolated hinge nanostructure)[39]; (2) manually measuring their angles using tools like ImageJ[39]; (3) determining an angular probability distribution from the manual measurement of hundreds to thousands of hinges; and (4) applying Boltzmann inversion to determine the angular free energy landscape[40]. This experiment and characterization pipeline is illustrated in Fig. 1A. The manual sampling and measuring processes take a significant portion of time. For example, to determine one condition of free energy landscape, it includes ~ 30 min of sample preparation, ~ 30
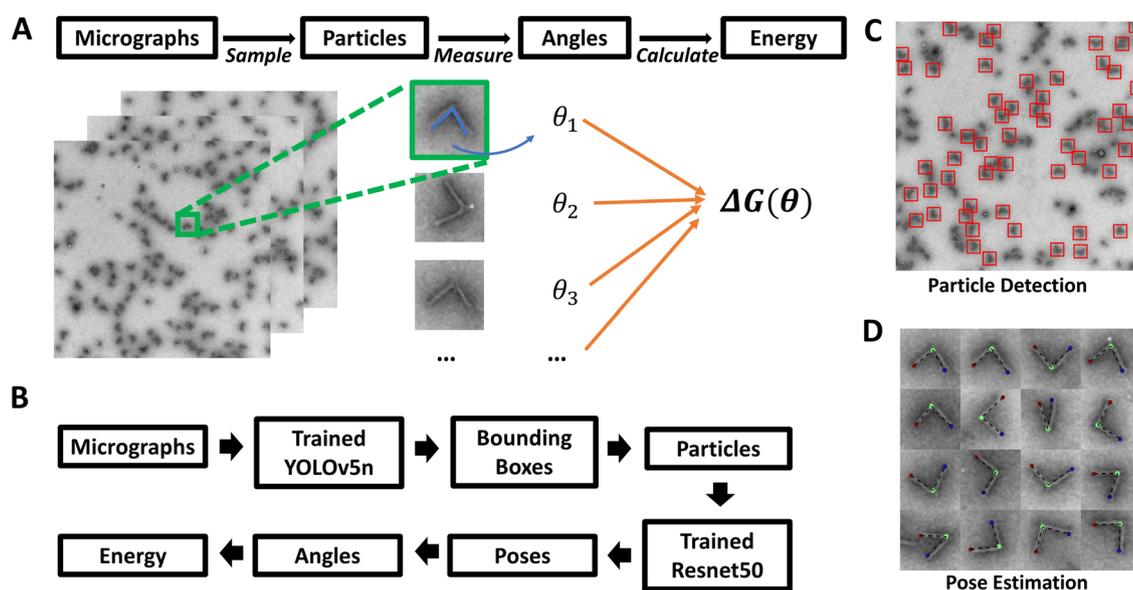

Figure 1. (**A**) The traditional pipeline for analyzing dynamic DNA origami device free energy landscapes with manual particle sampling and measurement. (**B**) The DNN-based pipeline. (**C**) Example micrograph illustrating particle detection. (**D**) Example particle montage illustrating pose estimation.

min of imaging, but ~ 2 h for the manual characterization. Additionally, this amount of work can be easily scaled up by experiment iteration and repetition. Therefore, there is a clear need for a fully automated approach to accelerate the experiment cycle.

Hence, we introduce modified characterization pipeline that leverages DNNs to automate both the structure sampling and the conformation measurement. In our DNN-based pipeline, we utilize YOLOv5 network for sampling, where the DNN provides the center location and size (width and height) of a bounding box containing an isolated folded DNA origami hinge particle. Individual particle images were generated by cropping raw micrographs according to bounding boxes. Secondly, we utilize Resnet50 for angle measurements where it provides three critical positions with two hinge tips and one vertex points, as shown in Fig. 1B. We refer to the first step as a 'particle detection problem' (Fig. 1C) and the second step as a 'pose estimation problem' (Fig. 1D).

Particle detection problem. We employed the DNN YOLOv5[41] to solve our particle detection problem. There are several versions of the YOLOv5 network with various architectures and different complexities. Here, we used the smallest network YOLOv5nano (YOLOv5n. We also tested larger YOLOv5 networks, but they did not demonstrate significantly better performance, see Supplementary Fig. S2). In order to train the network, we first manually labeled (~ 2 to 3 h) the square bounding boxes from TEM micrographs for a total of 1257 individual particles from a total of 49 images as a ground truth reference. Based on the image index number, we then split the Raw TEM Image Dataset (images + corresponding bounding box labels) into a training set (9 images), a validation set (10 images), and a test set (30 images). The motivation for the small training set is that we aimed to minimize the annotation work and computational cost to facilitate rapid training for future applications. The training set was used for fitting the YOLOv5n model. The validation set was used for tuning the network such as its architecture and hyperparameters to avoid overfitting. The test set was then used to evaluate the final performance of the trained network (see "Methods" for specific details).

As shown in Fig. 2A, the raw TEM images have several different major features: (1) isolated hinges with two clear arms (target particles for sampling), (2) hinges with a different orientation (i.e., vertical orientation) where the hinge angle cannot be observed due to the TEM grid deposition, (3) local aggregation where two or more hinges are touching or very near each other, and (4) image background. The trained YOLOv5n was able to effectively identify the target hinges among all these features. In Fig. 2A, each identified particle is labeled with its predicted bounding box. In the particle detection task with only single class (e.g., here is only hinge), the F1 score is typically used as the evaluation matrix. The F1 score is a balanced approach of precision value and recall value. Generally, the precision value is a measurement of false positive over all positive and the recall
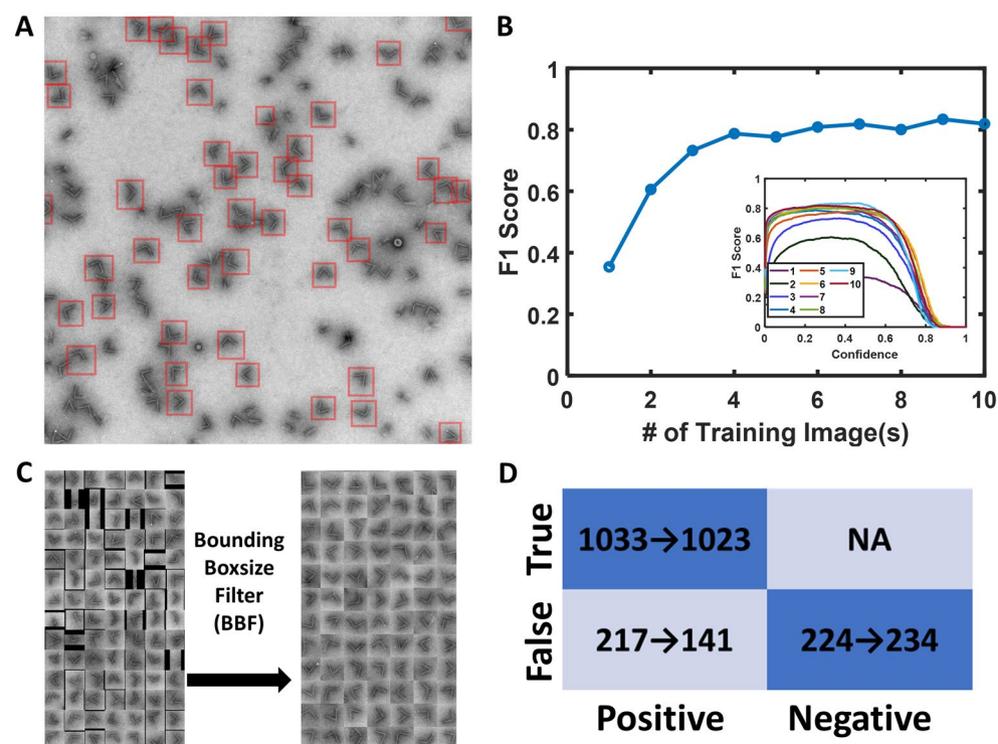

Figure 2. Particle detection performance. (**A**) An example of TEM image with YOLOv5 bounding box on the top. (**B**) Training data size sensitivity with F1 score as a function of number of training images. The number of particles per image ranges from ~ 20 to 30 and the total number of particles for 10 images is 259. Inset, F1 score with different confidence value for different training image numbers (**C**) The bounding boxsize filter removed particles with aspect ratio > 1.5. (**D**) Confusion matrix. The arrow represents the results after applying the filter.

value is a measurement of how much of true sample are missed during prediction (see detailed definition in "Method" section).

Since we aimed to minimize the annotation labor work, we conducted training size sensitivity experiments by using a different number of TEM images for our training set. We found that only 6–10 images (~ 30 target hinges per image) led to good network performance as indicated by the F1 score of ~ 0.8 (Fig. 2B). We selected nine images since it gives the highest F1 score and comparable labor work than 4–5 images.

To further improve performance, we also revisited the bounding box aspect ratio from prediction and found that many of the bounding box aspect ratios were not close to 1 as expected, especially for boxes on the image boundary. One potential way to mitigate this large aspect ratios would be to use a higher penalty value for the width and height loss, but this could influence the balance of bounding box position accuracy without careful tuning. By comparing with our ground truth annotation, we found these predicted boxes on the boundary with large aspect ratio contributed to a significant portion of false positives (Fig. 2C). Therefore, we developed a bounding box-size filter (BBF) that (1) removes all boxes with greater than 1.5 aspect ratio; (2) re-defined the height and width of the bounding boxes to both be 50 pixels, which was the case for annotation. By doing so, we observed that the false positive value decreased from 217 to 141 (Fig. 2D). This increased our precision from 0.83 to 0.88. We also noticed applying the BBF removed a minor fraction of true positives from 1033 to 1023, which reduced recall value from 0.82 to 0.81. Overall, the BFF increased the F1 score from 0.82 to 0.85 (summary of BFF effects shown in Fig. 2D). Once the bounding boxes are defined by the network, all the predicted particles can rapidly be cropped out from the raw micrographs by image processing software such as MATLAB or Python.

Pose estimation problem. To quantify the hinge angle conformation, we employed the Resnet50 neural network that was streamlined by DeepLabCut from Mathis Group[42]. To get a sufficiently large dataset for a size sensitivity test, instead of only using the 1257 ground truth from labeled data in previous section, we also collected 5115 hinge particles from ref[9]. We manually annotated each particle with three critical points that define the hinge angle (two hinge tips A, B and one vertex that fit the inner lines of arms). Additionally, the second person annotated a subset of the data to evaluate potential annotation bias that could result to the limited TEM resolution (See Supplementary Fig. S9). We split the whole Image Particle Dataset into training (107, 269, 644, 1343, 2686, 5103 image particles), validation (269 image particles), and test set (1000 image particles).

To evaluate the sensitivity of training to the Image Particle Dataset size, we quantified the mean angle error for the Resnet50 model trained for different numbers of particles (Fig. 3B). The annotation angle error was estimated from two different annotators in a smaller sub-dataset (616 image particles). We found the prediction error converged to below 4 deg when the training size was 644 particles or more, and even a training dataset size of ~ 100 particles leads to lower than 8 deg angle error.

For all evaluations in this section, we selected to use the network with training size 644. Figure 3C shows the spatial prediction error of the model compared with ground truth.

The hinge in TEM has no preference for orientation, and there is no specific feature difference between each arm that would allow identification of the corresponding tip. In other words, it is equivalent to arbitrarily flip Tip A and Tip B labels in the ground truth tip coordinate annotation. Therefore, it is seemingly impossible to consistently classify these datasets in the same order, and we likely have ~ 50% of cases with flipped tip prediction due to randomness. If the misclassification happens, the spatial error would be significantly larger. but in fact we only have 3% flipped (Supplementary Figure S10). We assumed Resnet50 captured the spatial labeling pattern from the annotator, e.g., in Fig. 3A, the blue tips are the first point, and the right tips are the third point during annotation. The blue points are higher than red points since the annotation was generally carried out from top to bottom on an individual particle.

We found the error distributions of tip A and tip B are larger than the vertex. We reasoned this was due to the different structural characteristics of the vertex and the tips. The vertex point was manually selected at the visually identified intersection of two lines along the inside of the arms; while tip A or tip B points were selected by visually identifying the tip along the inside of the arms, which is not as well-defined likely due to fraying of the ends. In practice, the tip position can be identified along the inner arm, but the distance away from the vertex is harder to define (see supplement Figure S11 for radial length error).

To further evaluate the capacity for the Resnet50 DNN to quantify mechanical properties, we converted the hinge angle measurements into angular probability distributions, and then calculated the free energy landscape from the probability distributions through Boltzmann inversion[40]. The free energy landscape gives a useful overall depiction of the relevant mechanical properties, and the applied torques and forces can be directly calculated from the free energy landscape as we have previously demonstrated[9]. We compared both the hinge angle probability distribution and angular free energy landscape predictions of the Resnet50 DNN to the experimental results (i.e., probability distributions and free energy landscapes calculated from a full dataset annotated manually). Figure 3D shows a comparison of the angle distributions (top) and the angular free energy landscapes (bottom), illustrating the good agreement between the model prediction and the experimental results. Additionally, the two angle distributions passed the Two-sample Kolmogorov–Smirnov test (see "Method" for detail).

Ideally, the Resnet50 DNN pose estimation could be applied to a variety of dynamic DNA origami devices. To test the robustness of the approach, we applied the same Resnet50 architecture for two other previously obtained particle image datasets: (1) a dataset of hinge devices with incorporated nucleosome (i.e. DNA wrapped around a histone protein core, which is the base packing unit of genomic DNA in eukaryotes) where the nucleosome position is of interest[9], and (2) a dynamic devices with two fluctuating arms where the angular conformation of both arms are of interest[35]. Using the same training protocol, we trained Resnet50 models to predict specified critical features of the device conformations. In the first example, we are interested in quantifying the hinge angle, similar to the free hinge, and the nucleosome position as one additional coordinate point. Correlating
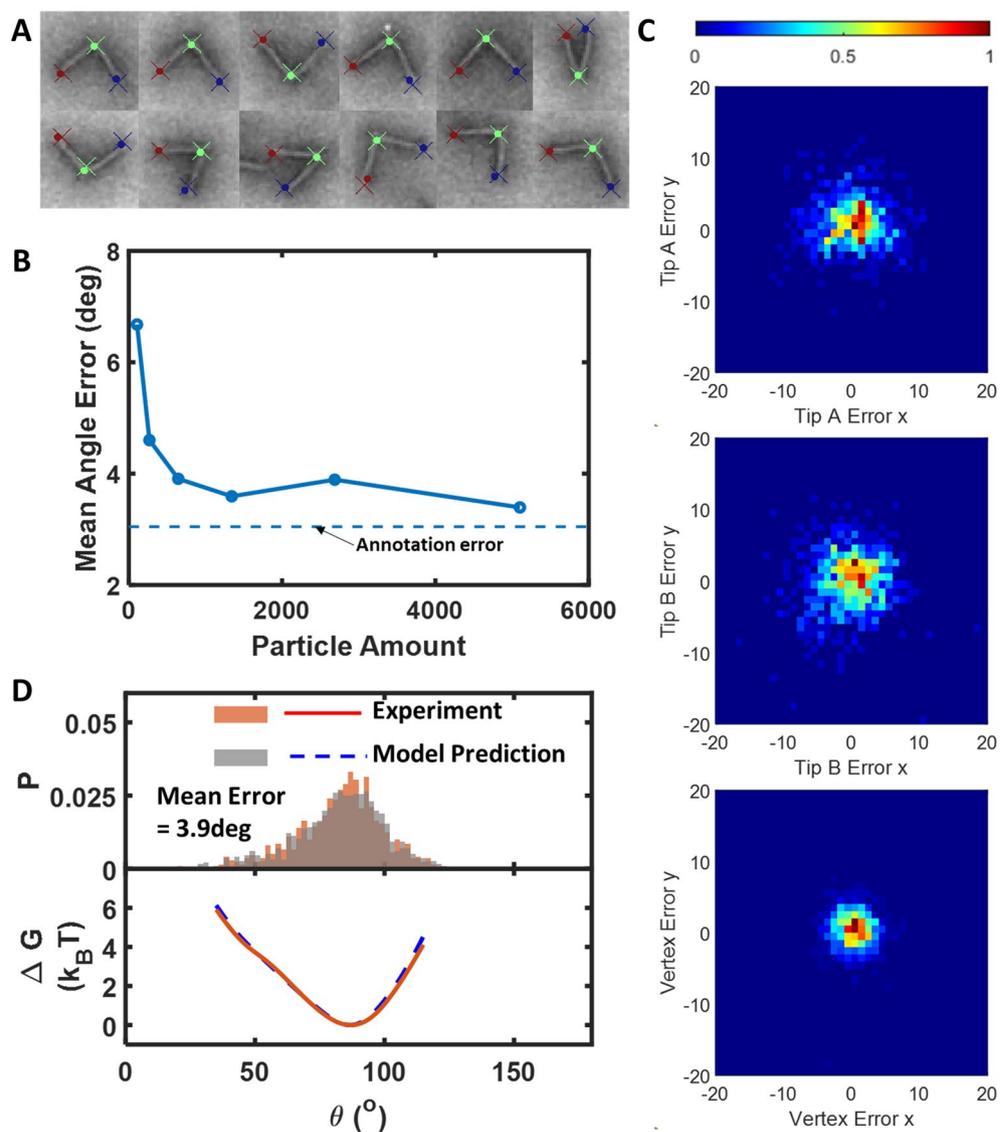
Figure 3. Pose estimation performance. (**A**) Random selected hinge particles with labels on. Cross: experiment. Dot: prediction. (**B**) Training Data size test, the reference dash line is the annotation error (**C**) Spatial error distribution between prediction and experiment in the unit of nm, the probability is normalized from 0 to 1 as shown in color bar. (**D**) Downstream data test with hinge angle distribution and torque distribution.

the hinge angle and nucleosome position can be useful to study the wrapping/unwrapping of nucleosomes[9]. We labeled 321 image particles with 301 as training data. We found that the Resnet50 DNN can successfully recognize nucleosomes linked with hinge even in the presence of background noise such as free nucleosomes (white dots in Fig. 4A). Figure 4B shows the error between the Resnet50 predicted nucleosome position and ground truth, by using. yielding 3.3 nm two-dimensional standard deviation for nucleosome. In the second example we used a dynamic device with two fluctuating arms, which we refer to as the SteriDyn[35] due to the steric dynamic interactions of the arms. In this example, we are interested in the conformation of two moving components on a single device (i.e., the angle of each arm relative to the base platform). We manually annotated 4469 particles with 3500 as training dataset from TEM images using four points to define the two hinge angles. Figure 4C shows Resnet50 DNN can successfully recognize the orientation of the structure even though the left arm and right arm are very similar. The error distributions for the SteriDyn points are shown in Fig. 4D. The two-dimensional standard deviation are 4.6, 4.5, 5.5, 4.8 nm for tip L, tip R, vertex L, vertex R, respectively.

Discussion

Here, we demonstrated a DNN-based pipeline for quantifying the structure and mechanical properties (i.e., free energy landscape) of dynamic DNA origami devices using DNNs to automate two key steps of the characterization, namely particle detection and pose estimation (i.e., conformation measurement). We divided this workflow into two parts: (1) using a YOLOv5 DNN to detect the DNA origami structures from raw TEM images, and (2)
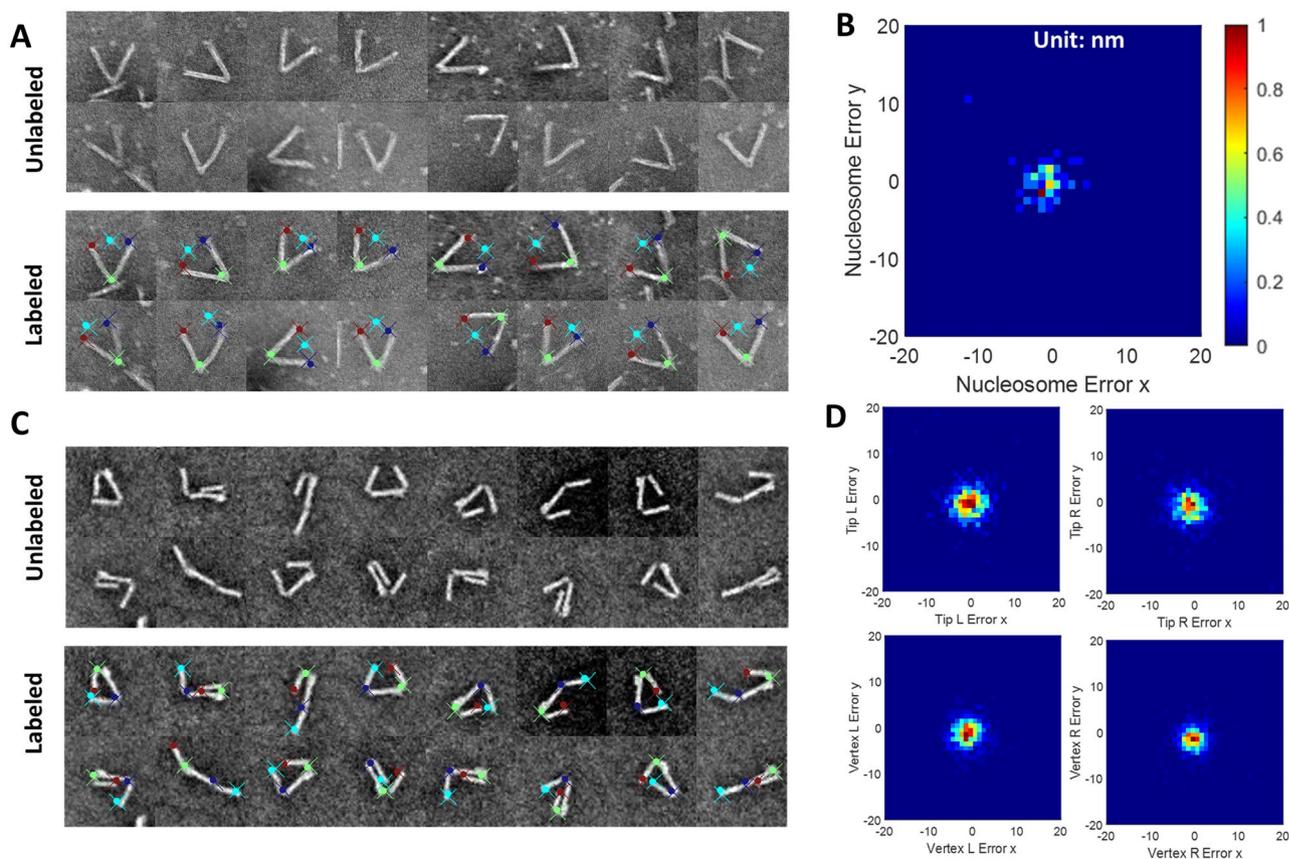
Figure 4. Pose estimation performance for Hinge-Nucleosome and SteriDyn datasets. Representative particles with annotation and plot showing error distribution for the hinge-nucleosome (**A,B**) and SteriDyn (**C,D**) datasets. Color scale for (**D**) is the same as (**B**). The probability is normalized from 0 to 1.

then using a Resnet50 DNN to detect the conformation from individual particle images. In the particle detection process, we used a small training set (9 TEM images led to the highest F1 score in the current study) and simple network complexity. We further used BBF filters to remove a set of false positives that mostly happened at the image boundary and increased F1 score from 0.82 to 0.85. This aspect ratio filter effectively eliminated incomplete particles near the border. However, the appropriate filtering approach after the initial particle selection may depend on the structure and can be considered on a case-by-case basis.

In the pose estimation process, we used 644 particles as a training set and achieved 3.9 deg hinge angle mean error. Furthermore, we demonstrated that our pose estimation process worked well for two other dynamic DNA nanodevices: Hinge-Nucleosome and SteriDyn, both of which agreed well with experimental data. The Hinge-Nucleosome example shows the pose estimation can be useful for devices applied to probe a molecule or interaction, and the SteriDyn example shows the pose estimation can be useful for systems containing more than one dynamic component/feature of interest.

The DNNs presented here were developed as a toolbox for finding patterns and quantifying specific features from experimental data. While the DNNs can perform forward evaluation within seconds, which is much faster than the manual approach, it is important to note there is significant manual effort required for the annotation of datasets (training, validation, and test datasets). Limiting this laborious process is the motivation for using small datasets to train the network. Therefore, the time saved for the entire experiment labor work would be dependent on the target throughput. For example, it typically takes more than 300 particles for sampling dynamic structures with a single condition[9,35,36]. Doing the manual annotation for a dataset of ~300 to 500 particles to directly provide the ground truth could be completed within several hours, which is comparable to the time required to obtain datasets for training and testing DNN performance. Hence, developing an automated pipeline may not be worthwhile for a single dataset. However, in practice, dynamic DNA origami designs often take multiple iterations for optimization, or it may be of interest to design several versions with distinct properties (i.e., distinct conformational probability distributions and free energy landscapes). This often leads to anywhere from several to tens of datasets that need to be analyzed where the automated pipeline can make a major difference. As long as these cases involve the same basic structure with the same basic features, it should be possible to apply the same DNN pipeline, saving days or even weeks' worth of manual annotation. To illustrate the benefits of accelerating data analysis, consider a comparison of applying our pipeline versus completely manual analysis for the development of the hinge as a nanomechanical device (based on studies presented in ref[9]). To perform the entire image data analysis for all devices used in that study, we estimated a workload of 2.6 h using a deep learning pipeline instead of 46.25 h with a completely manual analysis (see Supplementary material for detailed calculations).

Rather than pursuing higher model precision, we sought to balance the amount of annotation work while providing effective results for the data analysis. For example, our results showed our pipeline led to 3.9° errors angle measurements and passed the Two-sample Kolmogorov–Smirnov test. Nevertheless, there are multiple ways to improve the network performance in terms of precision such as hyperparameter tunning, dataset redistribution, network architecture modification. For example, users can simply add more training images for better performance. However, the labor work could significantly increase, and this may conflict with the purpose of using DNN in this work.

More broadly, this general workflow can not only solve the mechanical properties of dynamic DNA nanodevices, but also could be suitable for non-DNA materials such as antibodies or other protein complex structures that undergo significant thermal fluctuations or conformational changes.

Method

Preparation of the DNA origami structures.
The DNA structures used in this work are based on scaffold DNA origami, which consists of a long single-stranded DNA (ssDNA) scaffold (M13MP18 bacteriophage virus prepared in our laboratory as described in[43]) and ~ 200 short ssDNA staples. Based on Watson–Crick base-pairing rules, the design of staples sequences determines the assembly of hinge and SteriDyn structures, which are both previously reported[9,35]. All staples were ordered from a commercial vendor (IDT, Coralville, IA). In the experiment, a final concentration of 20 nM scaffold, 200 nM of each staple strand,5 mM Tris, 5 mM NaCl, 1 mM EDTA, and 18 mM MgCl2, at pH 8.0 in aqueous solutions were made and then subjected to thermal annealing in a thermal cycler (Bio-Rad, Hercules, CA) for self-assembly. After that, the excess staples were purified by centrifugation in a polyethylene glycol (PEG) solution[44]. The remaining structures were resuspended with buffer (0.5× TBE with net 10 mM MgCl2) and quantified by NanoDrop (NanoDrop 2000C Spectrophotometer, Thermo Scientific). The structures were diluted to 1 nM for downstream imaging.

TEM imaging.
Transmission electron microscopy (TEM) was used to visualize the structure with nanometer resolution. Specifically, 4 μL of sample droplet was deposited on Formvar-coated copper TEM grids, stabilized with evaporated carbon film (Ted Pella; Redding, CA) for 4 min. The droplet was wicked away by filter paper and then stained by applying 7 μL 2% uranyl formate (SPI, West Chester, PA) and wicked away twice for 2 s and 15 s, respectively. TEM imaging was carried out at the OSU Campus Microscopy and Imaging Facility on an FEI Tecnai G2 Spirit TEM at an acceleration voltage of 80 kV at a magnification of 45,000×, with an 1824 by 1824 pixel size.

Particle detection process with YOLOv5.
The 50 raw TEM images that we used for training, validation, and testing were resized to 960×960 pixel jpeg files and further split into a training set (10 or less, we selected 9 images as the final training set), a validation set (10 images), and a test set (30 images). The manual annotation work was conducted in Roboflow[45] and modified by custom MATLAB scripts for adjusting all bounding boxes with a 50×50 pixel size. The prepared Raw TEM Image Dataset was augmented stochastically by using the YOLOv5 default value(hyp.scratch-low.yaml). The neural network started from a pre-trained model (yolov5n) and it took ~ 5 min on a machine with a RTX3060 graphics card RTX3060 (~ 12 min on 1080Ti) for 500 epochs with the converged loss function. Specifically, YOLOv5 provided the x-position, y-position, width, height, and confidence of the predicted bounding box. The F1 score value was evaluated with the validation set for determining the optimal network parameters. After YOLOv5n network was selected, the test set was used for evaluating F1 score. The confidence threshold and Intersection over Union (IoU) are selected to 0.47, and 0.3, respectively. The bound boxsize filter (BBF) was then employed for improve F1 score by removing bounding box with lower than 1.5 aspect ratio.

The precision (Pr), recall (Re), and F1 score are defined as:

$$Pr = \frac{TP}{TP + FP}$$

$$Re = \frac{TP}{TP + FN}$$

$$F1 = 2 \times \frac{Pr \times Re}{Pr + Re}$$

where TP, FP, FN represent true positive, false positive, false negative, respectively.

Pose estimation process with Resnet50.
All particles in this work were resized to 200×200 pixel jpeg files and split into a training set (644 particles), a validation set (4728 particles), and a test set (1000 particles). The manual annotation work was conducted in ImageJ[39]. Specifically, we used the 'angle tool' to mark two hinge tips and one vertex for each particle and the encrypted ROI data was parsed to xy position csv files using custom MATLAB scripts. The Image Particle Dataset was processed by the pose-estimation tool DeepLabCut[42]. By default, we used 'imgaug' image augmenter and the Resnet50 network. This neural network was finetuned from a pre-trained model 'ImageNet' and it took ~ 25 min (on RTX3060) for 50k iterations to converge. The confidence threshold was selected to 0.92 to eliminate the majority of higher pixel errors.

The performance was evaluated by spatial error, mean angle error, and hinge energy comparison. For spatial error, we calculated the x and y arithmetical difference for each critical hinge point between ground truth and network prediction and plotted the two dimensional histogram as a heatmap by using 'pcolor' function in MATLAB. For mean angle error, we calculated the absolute angle difference between ground truth and network prediction for each hinge as a distribution. We then take the mean value of this distribution and finally, we apply the same process for models that come from different training sizes. For hinge energy, we applied Boltzmann inversion to convert the angular probability distribution to a free energy landscape[40]. To quantify the agreement of angle distribution between experiment and network prediction, we gave a Kolmogorov–Smirnov (KS) test. The null hypothesis is that two angle arrays are from the same continuous distribution. We tested these two angle arrays with build-in 'kstest2' function in MATLAB[46]. The p value[47] is 0.48 and is much greater than the significance level is 0.05. Therefore, this indicates a failure to reject the null hypothesis, which indicates that network predictions are in a great agreement with experiment.

For Hinge-Nucleosome, the training size is 304 and the test size is 76. For SteriDyn, the training size is 3500 and the test size is 875. All other methods were identical compared with hinge structure.

Data availability

Transmission electron microscopy images used to develop the deep neural network analysis pipeline presented here are available on the open science framework (https://doi.org/10.17605/OSF.IO/BMXHF). All devices presented here are previously reported and the design details including sequences are available in prior publications (hinge design presented in[9], SteriDyn design presented in[35]).

References

1. Shen, J., Sun, W., Liu, D., Schaus, T. & Yin, P. Three-dimensional nanolithography guided by DNA modular epitaxy. *Nat. Mater.* **20**(5), 5. https://doi.org/10.1038/s41563-021-00930-7 (2021).
2. Gopinath, A., Miyazono, E., Faraon, A. & Rothemund, P. W. K. Engineering and mapping nanocavity emission via precision placement of DNA origami. *Nature* **535**(7612), 7612. https://doi.org/10.1038/nature18287 (2016).
3. Woods, D. *et al.* Diverse and robust molecular algorithms using reprogrammable DNA self-assembly. *Nature* **567**(7748), 7748. https://doi.org/10.1038/s41586-019-1014-9 (2019).
4. Jungmann, R. *et al.* Single-molecule kinetics and super-resolution microscopy by fluorescence imaging of transient binding on DNA origami. *Nano Lett.* **10**(11), 4756–4761. https://doi.org/10.1021/nl103427w (2010).
5. Weiden, J. & Bastings, M. M. C. DNA origami nanostructures for controlled therapeutic drug delivery. *Curr. Opin. Colloid Interface Sci.* **52**, 101411. https://doi.org/10.1016/j.cocis.2020.101411 (2021).
6. Selnihhin, D., Sparvath, S. M., Preus, S., Birkedal, V. & Andersen, E. S. Multifluorophore DNA origami beacon as a biosensing platform. *ACS Nano* **12**(6), 5699–5708. https://doi.org/10.1021/acsnano.8b01510 (2018).
7. Liu, F., Liu, X., Huang, Q. & Arai, T. Recent progress of magnetically actuated DNA micro/nanorobots. *Cyborg Bionic Syst.* https://doi.org/10.34133/2022/9758460 (2022).
8. Liu, Q., Kuzyk, A., Endo, M. & Smalyukh, I. I. Colloidal plasmonic DNA-origami with photo-switchable chirality in liquid crystals. *Opt. Lett. OL* **44**(11), 2831–2834. https://doi.org/10.1364/OL.44.002831 (2019).
9. Wang, Y. *et al.* A nanoscale DNA force spectrometer capable of applying tension and compression on biomolecules. *Nucleic Acids Res.* **49**(15), 8987–8999. https://doi.org/10.1093/nar/gkab656 (2021).
10. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**(7553), 7553. https://doi.org/10.1038/nature14539 (2015).
11. Datta, S. & Davim, J. P. *Machine Learning in Industry*. (Springer, 2021).
12. Erjavec, J. & Thompson, R. *Automotive Technology: A Systems Approach*. (Cengage Learning, 2014).
13. Nieuwenhuis, P. & Wells, P. *The Automotive Industry and the Environment*. (Woodhead Publishing, 2003).
14. Singh, K. B. & Arat, M. A. Deep learning in the automotive industry: Recent advances and application examples. *arXiv* https://doi.org/10.48550/arXiv.1906.08834 (2019).
15. Gu, S., Holly, E., Lillicrap, T. & Levine, S. Deep reinforcement learning for robotic manipulation with asynchronous off-policy updates. In *2017 IEEE International Conference on Robotics and Automation (ICRA)*, May 2017. 3389–3396. https://doi.org/10.1109/ICRA.2017.7989385 (2017).
16. Nagabandi, A., Kahn, G. Fearing, R.S. & Levine, S. Neural network dynamics for model-based deep reinforcement learning with model-free fine-tuning. In *2018 IEEE International Conference on Robotics and Automation (ICRA)*, May 2018. 7559–7566. https://doi.org/10.1109/ICRA.2018.8463189 (2018).
17. Luo, J. *et al.* Reinforcement learning on variable impedance controller for high-precision robotic assembly. In *2019 International Conference on Robotics and Automation (ICRA)*, May 2019. 3080–3087. https://doi.org/10.1109/ICRA.2019.8793506 (2019).
18. Deng, L. & Platt, J. Ensemble deep learning for speech recognition. In *Presented at the Proceedings Interspeech, Sep 2014*. https://www.microsoft.com/en-us/research/publication/ensemble-deep-learning-for-speech-recognition/. Accessed 28 Feb 2023 (online) (2023).
19. Kamath, U., Liu, J. & Whitaker, J. *Deep Learning for NLP and Speech Recognition*. https://doi.org/10.1007/978-3-030-14596-5 (Springer, 2019).
20. Zhang, Z. *et al.* Deep learning for environmentally robust speech recognition: An overview of recent developments. *ACM Trans. Intell. Syst. Technol.* **9**(5), 1–49. https://doi.org/10.1145/3178115 (2018).
21. Gamboa, J. C. B. Deep learning for time-series analysis. *arXiv*. https://doi.org/10.48550/arXiv.1701.01887 (2017).
22. Jin, X., Pei, K., Won, J. Y. & Lin, Z. SymLM: Predicting function names in stripped binaries via context-sensitive execution-aware code embeddings. In: *Proceedings of the 2022 ACM SIGSAC Conference on Computer and Communications Security*, Los Angeles CA USA. 1631–1645. https://doi.org/10.1145/3548606.3560612 (ACM, 2022).
23. Zhang, D., Yin, C., Zeng, J., Yuan, X. & Zhang, P. Combining structured and unstructured data for predictive models: A deep learning approach. *BMC Med. Inform. Decis. Mak.* **20**(1), 280. https://doi.org/10.1186/s12911-020-01297-6 (2020).
24. Trask, N., Patel, R. G., Gross, B. J. & Atzberger, P. J. GMLS-Nets: A framework for learning from unstructured data. *arXiv*. https://doi.org/10.48550/arXiv.1909.05371 (2019).
25. Fang, H.-S., Xie, S., Tai, Y.-W. & Lu, C. RMPE: Regional multi-person pose estimation. In *Presented at the Proceedings of the IEEE International Conference on Computer Vision*. 2334–2343. https://openaccess.thecvf.com/content_iccv_2017/html/Fang_RMPE_Regional_Multi-Person_ICCV_2017_paper.html. Accessed 28 Feb 2023 (2017).

26. Ciregan, D., Meier, U. & Schmidhuber, J. Multi-column deep neural networks for image classification. In *2012 IEEE Conference on Computer Vision and Pattern Recognition*. 3642–3649. https://doi.org/10.1109/CVPR.2012.6248110 (2012).
27. Jumper, J. *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature* **596**(7873), 7873. https://doi.org/10.1038/s41586-021-03819-2 (2021).
28. Chiriboga, M. *et al.* Rapid DNA origami nanostructure detection and classification using the YOLOv5 deep convolutional neural network. *Sci. Rep.* **12**(1), 1. https://doi.org/10.1038/s41598-022-07759-3 (2022).
29. Wanninger, S. *et al.* Deep-learning assisted, single-molecule imaging analysis (deep-LASI) of multi-color DNA origami structures. *bioRxiv*. 2023.01.31.526220. https://doi.org/10.1101/2023.01.31.526220 (2023).
30. Chen, C., Nie, J., Ma, M. & Shi, X. DNA origami nanostructure detection and yield estimation using deep learning. *ACS Synth. Biol.* **12**(2), 524–532. https://doi.org/10.1021/acssynbio.2c00533 (2023).
31. Rothemund, P. W. K. Folding DNA to create nanoscale shapes and patterns. *Nature* **440**(7082), 297–302. https://doi.org/10.1038/nature04586 (2006).
32. DeLuca, M., Shi, Z., Castro, C. E. & Arya, G. Dynamic DNA nanotechnology: toward functional nanoscale devices. *Nanoscale Horizons* **5**(2), 182–201. https://doi.org/10.1039/C9NH00529C (2020).
33. Le, J. V. *et al.* Probing nucleosome stability with a DNA origami nanocaliper. *ACS Nano* **10**(7), 7073–7084. https://doi.org/10.1021/acsnano.6b03218 (2016).
34. Liu, M. *et al.* A DNA tweezer-actuated enzyme nanoreactor. *Nat. Commun.* **4**, 1–5. https://doi.org/10.1038/ncomms3127 (2013).
35. Wang, Y. *et al.* Steric communication between dynamic components on DNA nanodevices. *ACS Nano* **17**(9), 8271–8280. https://doi.org/10.1021/acsnano.2c12455 (2023).
36. Darcy, M. *et al.* High-force application by a nanoscale DNA force spectrometer. *ACS Nano* **16**(4), 5682–5695. https://doi.org/10.1021/acsnano.1c10698 (2022).
37. Castro, C. E. *et al.* A primer to scaffolded DNA origami. *Nat. Methods* **8**(3), 221–229. https://doi.org/10.1038/nmeth.1570 (2011).
38. Castro, C. E., Su, H. J., Marras, A. E., Zhou, L. & Johnson, J. Mechanical design of DNA nanostructures. *Nanoscale* **7**(14), 5913–5921. https://doi.org/10.1039/c4nr07153k (2015).
39. Abramoff, M. D., Magalhães, P. J. & Ram, S. J. Image processing with ImageJ. *Biophoton. Int.* **11**(7), 36–42 (2004).
40. Marras, A. E., Zhou, L., Su, H. J. & Castro, C. E. Programmable motion of DNA origami mechanisms. *Proc. Natl. Acad. Sci. U.S.A.* **112**(3), 713–718. https://doi.org/10.1073/pnas.1408869112 (2015).
41. Jocher, G. *et al.* ultralytics/yolov5: v7.0—YOLOv5 SOTA realtime instance segmentation. *Zenodo*. https://doi.org/10.5281/zenodo.7347926 (2022).
42. Lauer, J. *et al.* Multi-animal pose estimation, identification and tracking with DeepLabCut. *Nat. Methods* **19**(4), 4. https://doi.org/10.1038/s41592-022-01443-0 (2022).
43. Douglas, S. M. *et al.* Self-assembly of DNA into nanoscale three-dimensional shapes. *Nature* **459**(7245), 414–418. https://doi.org/10.1038/nature08016 (2009).
44. Stahl, E., Martin, T. G., Praetorius, F. & Dietz, H. Facile and scalable preparation of pure and dense DNA origami solutions. *Angew. Chem. Int. Ed.* **53**(47), 12735–12740. https://doi.org/10.1002/anie.201405991 (2014).
45. Roboflow: Give your software the power to see objects in images and video. https://roboflow.com/. Accessed 28 Feb 2023 (2023).
46. Two-sample Kolmogorov–Smirnov test—MATLAB kstest2. https://www.mathworks.com/help/stats/kstest2.html. Accessed 25 Apr 2023 (2023).
47. Winkler, J. R. Numerical recipes in C: The art of scientific computing, second edition. *Endeavour* **17**(4), 201. https://doi.org/10.1016/0160-9327(93)90069-F (1993).

Acknowledgements

Author contributions

Y.W and C.E.C. conceptualized the project, guided the execution of experiments, and supervised the research team. Y.W. led the execution of experiments including dataset preparation, network training, network evaluation. Y.W. led the MATLAB and Python code that were written for implementation. X. J. supported the experimental work and provided the critical network feedback. Y.W. and X. J. led the initial drafting with all authors providing editing and feedback. C.E.C. acquired the funding for this work. All authors have read and agreed to the published version of the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1038/s41598-023-41459-w.

Correspondence and requests for materials should be addressed to Y.W. or C.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.