Semantic segmentation of anomalous diffusion using deep convolutional networks

Xiang Qu [●],[1] Yi Hu,[2] Wenjie Cai [●],[1] Yang Xu,[2] Hu Ke,[2] Guolong Zhu,[1] and Zihan Huang [●][1,*]

[1]*School of Physics and Electronics, Hunan University, Changsha 410082, China*
[2]*Hubei Medical Devices Quality Supervision and Test Institute, Wuhan 430075, China*

Heterogeneous dynamics commonly emerges in anomalous diffusion with intermittent transitions of diffusion states but proves challenging to identify using conventional statistical methods. To effectively capture these transient changes of diffusion states, we propose a deep learning model (U-AnDi) for the semantic segmentation of anomalous diffusion trajectories. This model is developed with the dilated causal convolution (DCC), gated activation unit (GAU), and U-Net architecture. The study addresses two key subtasks related to trajectory segmentation and changepoint detection, concentrating on variations in diffusion exponents and dynamic models. Additionally, extended analyses are conducted on the segmentation of single-model trajectories, multistate biological trajectories, and anomalous diffusion with added correlation functions. By rationally designing comparative models and evaluating the performance of U-AnDi against these models, we discover that U-AnDi consistently outperforms other models across all segmentation tasks, thereby affirming its superiority in the field. This performance edge also sheds light on the interpretability of U-AnDi's core components: DCC, GAU, and U-Net. The clarity with which these components contribute to U-AnDi's success underscores their congruence with the intrinsic physics underlying anomalous diffusion. Furthermore, our model is examined using real-world anomalous diffusion data: the diffusion of transmembrane proteins on cell membrane surfaces, and the segmentation results are highly consistent with experimental observations. Our findings could offer a heuristic deep learning solution for the detection of heterogeneous dynamics in single-molecule/particle tracking experiments, and have the potential to be generalized as a universal scheme for time-series segmentation.

I. INTRODUCTION

Anomalous diffusion [1,2], characterized by the nonlinear transport phenomena that deviate from standard Brownian motion, has drawn significant attention due to its widespread occurrence in various scientific fields, such as physics [3–7], chemistry [8–11], biology [12–16], and finance [17–19]. Since anomalous diffusion can provide valuable insights into the underlying mechanisms of complex systems, it is of great importance to perform accurate and reliable analysis of this phenomenon [20–43]. However, due to the heterogeneous dynamics with intermittent transitions of diffusion states [44–47], conventional statistical methods usually fail to precisely describe the anomalous diffusion processes in real-world experimental environments. This emphasizes the necessity of exploring effective techniques for the semantic segmentation of anomalous diffusion trajectories, which could offer a more fine-grained approach to detect state transitions and distinguish heterogeneous behaviors in diffusion dynamics.

*huangzih@hnu.edu.cn

On the other hand, as a pivotal technique in computer vision, semantic segmentation makes pixelwise predictions in an image to enable the identification and delineation of distinct regions [48]. With the advent of deep learning techniques, semantic segmentation methods have experienced a rapid development in recent years, significantly outperforming traditional approaches that rely on handcrafted features [49–56]. These deep-learning-based methods have demonstrated remarkable success in a wide array of applications, including object recognition [49–52], scene understanding [53], and medical image analysis [54–56]. In particular, convolutional neural networks (CNNs) have shown unparalleled capabilities in capturing complex patterns and hierarchical representations from raw data, leading to state-of-the-art results in various semantic segmentation tasks [48]. One representative example is the U-Net architecture [54], which is built on the basis of an encoder-decoder structure and the rational use of skip connections. U-Net has emerged as one of the most popular and effective frameworks for semantic segmentation and proven to be highly effective in diverse segmentation tasks [56].

In this paper, motivated by the successful applications of CNNs in semantic segmentation, we explore the potential of employing deep convolutional networks for the semantic segmentation of anomalous diffusion. We introduce a deep learning model, U-AnDi, which synergizes the dilated causal convolution (DCC), gated activation unit (GAU), and U-Net architecture. The model's capabilities are validated across two

pivotal subtasks of trajectory segmentation, based on variable diffusion exponents and different dynamic models, respectively. Moreover, the adaptability of U-AnDi is demonstrated by thoroughly analyzing both single-model trajectories and multistate biological trajectories, as well as diffusion exhibiting extended correlations. The results manifest that U-AnDi excels in all these segmentation tasks, outperforming other comparative models. The superior performance of U-AnDi underscores its effectiveness and the interpretability of its core components: DCC, GAU, and U-Net, in reflecting the model's alignment with the nature of anomalous diffusion. Furthermore, the generalization ability of U-AnDi for real-world data is evaluated through experimental observations of the diffusion of transmembrane proteins on cell membrane surfaces [57]. The segmentation results exhibit a high degree of consistency with experimental observations, highlighting the applicability of our approach to characterizing anomalous diffusion in real complex systems.

The rest of this paper is organized as follows. Section II describes the segmentation tasks for anomalous diffusion, methods for generating simulated trajectories, and main evaluation metrics used in this work. Section III presents the architecture and core components of U-AnDi and its design principles. Section IV provides an analysis of U-AnDi's efficacy on simulated trajectories and compares its performance with that of other models. The interpretability of U-AnDi's core components is consolidated in Sec. V. In Sec. VI, we showcase the implementation of our method on experimental trajectories of transmembrane protein on membrane surfaces. Finally, we engage in a discussion and draw our conclusions in Sec. VII.

II. TASKS AND DATASETS

In this section, we present the details of two subtasks for the semantic segmentation of anomalous diffusion and corresponding simulated datasets used for model training and evaluation. These tasks aim to effectively address the challenge of capturing the intermittent transitions among different diffusion states. Data generation for both tasks is facilitated by utilizing the open-source Python package, *andi-datasets*, which is specifically designed for generating, managing, and analyzing anomalous diffusion trajectories [25,58]. In particular, we primarily focus on 2D trajectory data in this paper, as most experimental observations of anomalous diffusion are performed in 2D or quasi-2D environments.

A. Subtask 1: Segmentation of trajectories with varying diffusion exponents

The first subtask focuses on the segmentation of trajectories in which the subtrajectories exhibit varying diffusion exponents. Here, the diffusion exponent α quantifies the relation between mean squared displacement (MSD) and time t, written as

$$\text{MSD} \sim t^\alpha. \tag{1}$$

This exponent equals 1 in the description of standard Brownian motion. For anomalous diffusion, it can be either less than or greater than 1, corresponding to subdiffusion and superdif-

fusion, respectively. In real-world scenarios, the diffusion exponent may not remain constant due to inevitable changes of interactions between the random walker and its surroundings [45]. However, traditional statistical methods, which are based on the calculation of MSD, require sufficiently long trajectories to accurately determine the diffusion exponent. As a result, for anomalous diffusion trajectories with limited length scales, these conventional approaches face substantial difficulties when attempting to distinguish segments with different diffusion exponents.

To address this challenge, subtask 1 is designed to empower our model with the capability to identify these segments and predict their corresponding diffusion exponents. In detail, simulated segments are generated by the *andi-datasets* package within the theoretical framework of fractional Brownian motion (FBM) [59]. This diffusion model describes an ergodic diffusion process that is driven by a fractional Gaussian noise, where the diffusion exponent α satisfies $0 < \alpha < 2$. When generating segments, their diffusion coefficients are fixed at 1.0, while diffusion exponents α are drawn uniformly from the interval [0.05, 2). Here, to quantify the difference in diffusion exponents between two adjacent segments, we introduce $\delta\alpha = \alpha_2 - \alpha_1$, where α_1 and α_2 represent the exponents of front and rear segments, respectively. In particular, a minimum value for $|\delta\alpha|$ is set as $\delta\alpha_{\min}$ to ensure a distinct transition between diffusion states. The default value of $\delta\alpha_{\min}$ is 0.5 in this paper.

After that, M segments are combined to form a trajectory of length L. Here, M is randomly selected from the values 2, 3, 4, and 5. The length of each segment T is determined from a uniform distribution, with its bounds set between a minimum length $T_{\min} = 10$ and a maximum length $T_{\max} = L - (M - 1)T_{\min}$. Representative example of a single simulated trajectory with $L = 500$ and $M = 5$ is illustrated in Fig. 1(a), where different segments are marked in distinct colors. Time evolution of this trajectory is depicted in Fig. 1(b) for both the x and y dimensions, indicating that the trajectory data exhibits typical time-series characteristics. We assign the diffusion exponents of segments to each point as pointwise labels, leading to a pointwise regression task for our model. As shown in Fig. 1(c), these labels enable the clear identification of transition points (changepoints) among different diffusion states.

B. Subtask 2: Segmentation of trajectories with different dynamics models

The second subtask aims to segment trajectories composed of segments, each with dynamics originating from a distinct diffusion model. It has been proven that the dynamics of anomalous diffusion is multifaceted and can be described by a variety of theoretical stochastic process models. Similar to the varying diffusion exponents in subtask 1, the anomalous diffusion dynamics observed in the real world often undergoes transitions due to changes or fluctuations in the surrounding media [44,57]. For instance, the diffusion of transmembrane proteins on the cell membrane surface [57,60] exhibits a duality of dynamics, encompassing both FBM and continuous-time random walk [61]. However, traditional statistical methods are not particularly adept at identifying the
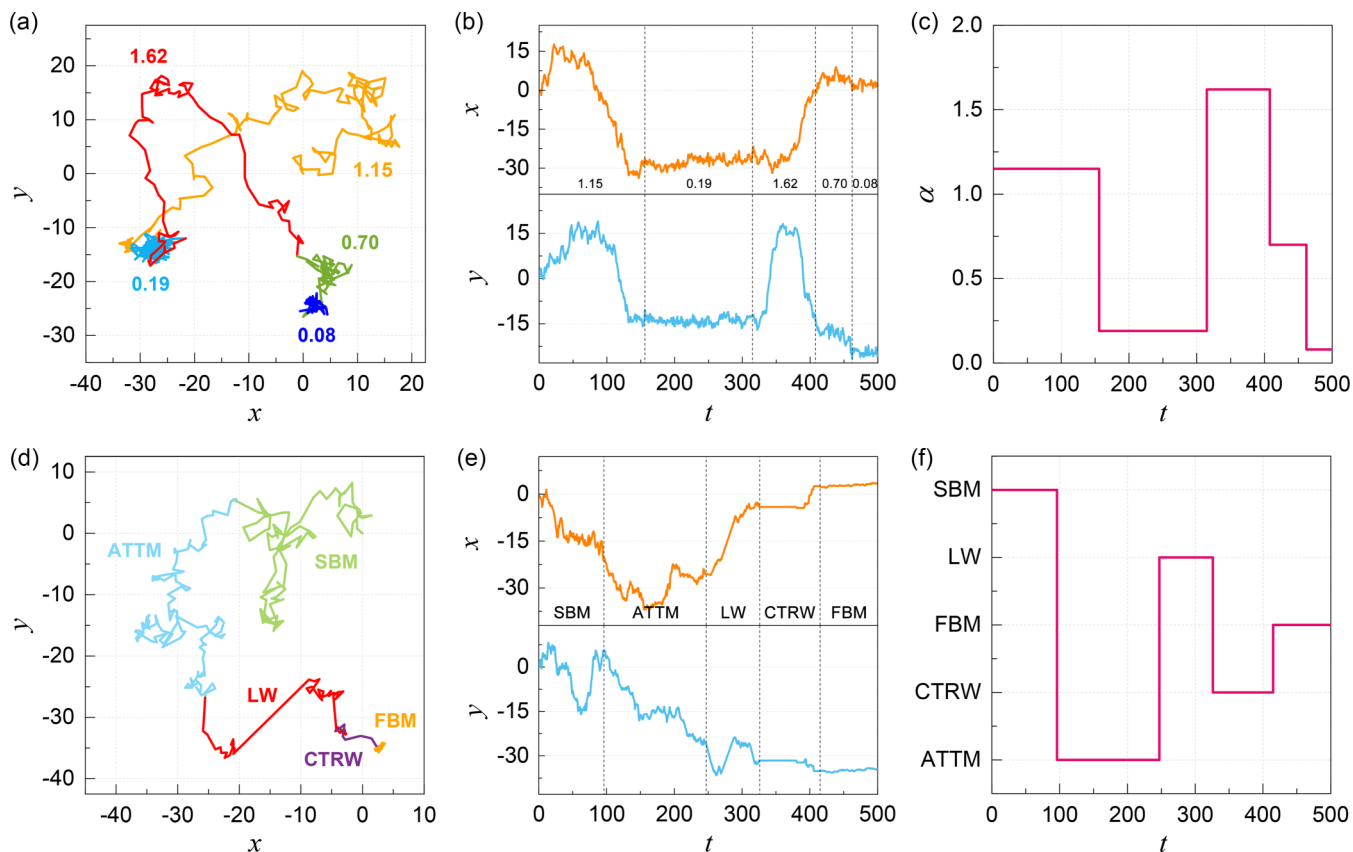
FIG. 1. [(a), (d)] Representative examples of trajectories with varying diffusion exponents (a) and different dynamics models (d). Different segments are marked in distinct colors. [(b), (e)] Time evolutions of trajectories in (a) and (d) for both x and y dimensions, respectively. Changepoints of diffusion states are indicated by dashed lines. [(c), (f)] Pointwise labels of trajectories in (a) and (d), respectively.

dynamics model from raw trajectory data, and the detection of transitions among these dynamics proves even more challenging. Therefore, developing effective methods to address this issue bears considerable significance for the analysis of anomalous diffusion behaviors and the elucidation of underlying physical mechanisms.

For that purpose, we design the subtask 2 with the goal of enabling our model to detect state changepoints in the diffusion process and accurately identify the dynamics models of segments. Five theoretical diffusion models are considered:

(1) Annealed transient time motion (ATTM) [62]. Brownian motion with a diffusion coefficient that varies randomly in either time or space ($0.05 \leqslant \alpha \leqslant 1$).

(2) Continuous-time random walk (CTRW) [61]. The waiting time between two consecutive steps is irregular and randomly chosen ($0.05 \leqslant \alpha \leqslant 1$).

(3) Fractional Brownian motion (FBM) [59]. Diffusion process driven by a power-law correlated fractional Gaussian noise ($0.05 \leqslant \alpha < 2$).

(4) Lévy walk (LW) [63]. The waiting time between subsequent steps is irregular, while the step length is not Gaussian distributed ($1 \leqslant \alpha \leqslant 2$).

(5) Scaled Brownian motion (SBM) [64]. Brownian motion with a deterministically time-dependent diffusion coefficient ($0.05 \leqslant \alpha \leqslant 2$).

Under the guidance of these five models, segments are generated using the *andi-datasets* package. The procedure for

forming a trajectory in this subtask is the same as in subtask 1, i.e., M segments are joined to form a trajectory of length L. The parameter values are also consistent with those in subtask 1. As an illustrative example, we present a typical sample of a trajectory with $L = 500$ and $M = 5$ in Fig. 1(d) and its corresponding time evolution for both x and y dimensions in Fig. 1(e). This trajectory with different dynamics models is also labeled in a pointwise manner, as depicted in Fig. 1(f).

C. Evaluation metrics

The main evaluation metric for subtask 1 is the mean absolute error (MAE),

$$\text{MAE} = \frac{1}{NL} \sum_{i=1}^{N} \sum_{t=1}^{L} \left| \alpha_{i,t}^{\text{Pred}} - \alpha_{i,t}^{\text{GT}} \right|. \qquad (2)$$

Here, N is the number of trajectories to be evaluated. For the ith trajectory at time step t, $\alpha_{i,t}^{\text{Pred}}$ and $\alpha_{i,t}^{\text{GT}}$ denote the predicted and ground truth values of the diffusion exponent, respectively.

For subtask 2, we select the mean Dice coefficient (mDice) as the main evaluation metric. The Dice coefficient, also known as the Sørensen-Dice coefficient, can measure the similarity of two samples and is commonly utilized in segmentation tasks [65]. This metric can be calculated in subtask

2 as

$$\text{mDice} = \frac{1}{N} \sum_{i=1}^{N} \frac{2 \sum \left(\mathbf{X}_i^{\text{Pred}} \otimes \mathbf{X}_i^{\text{GT}} \right)}{\sum \mathbf{X}_i^{\text{Pred}} + \sum \mathbf{X}_i^{\text{GT}}}. \tag{3}$$

Here, $\mathbf{X}_i^{\text{Pred}}$ is the predicted pointwise label and $\mathbf{X}_i^{\text{GT}}$ refers to the ground truth of the ith trajectory. Both are converted into matrices of dimensions $[L, n_c]$ by one-hot encoding for the Dice coefficient calculation. This encoding means each category is represented by a vector where a specific position is marked as 1 to indicate the presence of that category, and all other positions are set to 0. Here, n_c is the total number of diffusion-model types. The symbol $\otimes$ denotes the elementwise multiplication operator, and the small $\sum$ in the fraction represents the summation of all elements in the matrix. The value of the mean Dice coefficient ranges from 0 to 1, with a higher value indicating a better segmentation performance.

On the other hand, pinpointing the changepoints among diffusion states is of significant importance with respect to the segmentation of anomalous diffusion. However, when performing the changepoint detection, the model may generate some nonexistent changepoints (false positives, FPs), while potentially missing genuine changepoints (false negatives, FNs). This inevitably leads to errors in the identification of state transitions. Hence, to systematically quantify the model's performance in detecting these transitions, we define δc as

$$\delta c = c^{\text{Pred}} - c^{\text{GT}}, \tag{4}$$

where c^{Pred} and c^{GT} are the predicted and ground truth changepoints, respectively. For the model-predicted changepoints, we categorize those with $|\delta c| < T_{\min}$ as true positives (TPs) and those with $|\delta c| \geqslant T_{\min}$ as false positives (FPs). The number of FNs can be obtained by subtracting the count of true positives (TPs) from the total number of actual changepoints. After that, we employ the metrics below to further evaluate model's performance in changepoint detection for both subtasks [34],

$$\text{precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{5}$$

$$\text{recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{6}$$

$$\text{F1-score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}. \tag{7}$$

III. MODEL STRUCTURE AND DESIGN PRINCIPLES

The U-AnDi model proposed in this work is constructed by integrating the WADNet encoder [34] into the U-Net architecture. This integration is primarily driven by the aim to effectively capture the intrinsic dynamics of anomalous diffusion, particularly the long-time correlations, with the help of dilated causal convolution (DCC) and gated activation unit (GAU). Another motivation is the utilization of U-Net's adeptness at capturing local features and nuances for semantic segmentation tasks, which complements the WADNet encoder's focus on long-range dependencies. In the following parts, we will introduce the design principles of U-AnDi in detail, and outline their connections with the underlying

physics of anomalous diffusion. The detailed implementation of U-AnDi can be found in Appendix A, while the associated training scheme is outlined in Appendix B.

A. Dilated causal convolution

Considering that the motion of a random walker is solely determined by its current and past states, preserving causality to prevent using future information when processing the anomalous diffusion data is not just a technical requirement but a fundamental necessity. This crucial aspect is highlighted by Verdier *et al.* in their recent paper [35], where the causality is maintained by selecting incoming edges of each node originate only from nodes in the past in a graph neural network (GNN). Unlike the specific wiring scheme used in the GNN model, U-AnDi employs causal convolution to preserve the causality inherent in anomalous diffusion. Specifically, when applying the causal convolution to calculate the output at time step t, only the data from the preceding t steps in the previous layer will be used. This can be achieved by applying padding on both sides of the sequence.

On the other hand, as the long-time correlation is one of the essential properties of anomalous diffusion, our design should empower U-AnDi with the capability to capture these extended dependencies in trajectory data. However, the receptive field of a standard causal convolution is limited, posing challenges in capturing the long-range dependencies in longer trajectories. To address this issue, dilated convolution is concurrently employed with causal convolution within the WADNet encoder. As highlighted in Ref. [49], dilated convolutions can systematically aggregate multiscale contextual information without compromising resolution by harnessing the exponentially expanding receptive fields. This is particularly crucial when dealing with anomalous diffusion data, where discerning long-time tails and correlations is of utmost importance. A schematic representation of a dilated causal 1D convolution is given in Fig. 2(a), where the kernel size (ks) is set as 3. Notably, by doubling the dilation factor layer by layer, after applying convolutions for d (dilation depth) times, a single node can cover the information of $2^{d+1} - 1$ nodes in the input layer. The exponentially expanding receptive fields enable our model to encapsulate more long-range correlation information within the feature map. Consequently, this power-law expansion assists the model in appropriately balancing the sampling for measurements associated with long-time tails, characteristic of anomalous diffusion, where power-law correlation rather than simple exponential behavior is often observed.

B. Gated activation unit

While the exponentially expanding receptive field enhances U-AnDi's ability in collecting long-range correlation information, effectively processing this information and subsequently inferring the intrinsic properties of anomalous diffusion remains a challenge. Recognizing that anomalous diffusion often exhibits correlations where the influence of a past event decays slowly over time, the gated activation unit [66,67,69], with its capacity to maintain memory and manage information flow, is well suited to capture these slowly
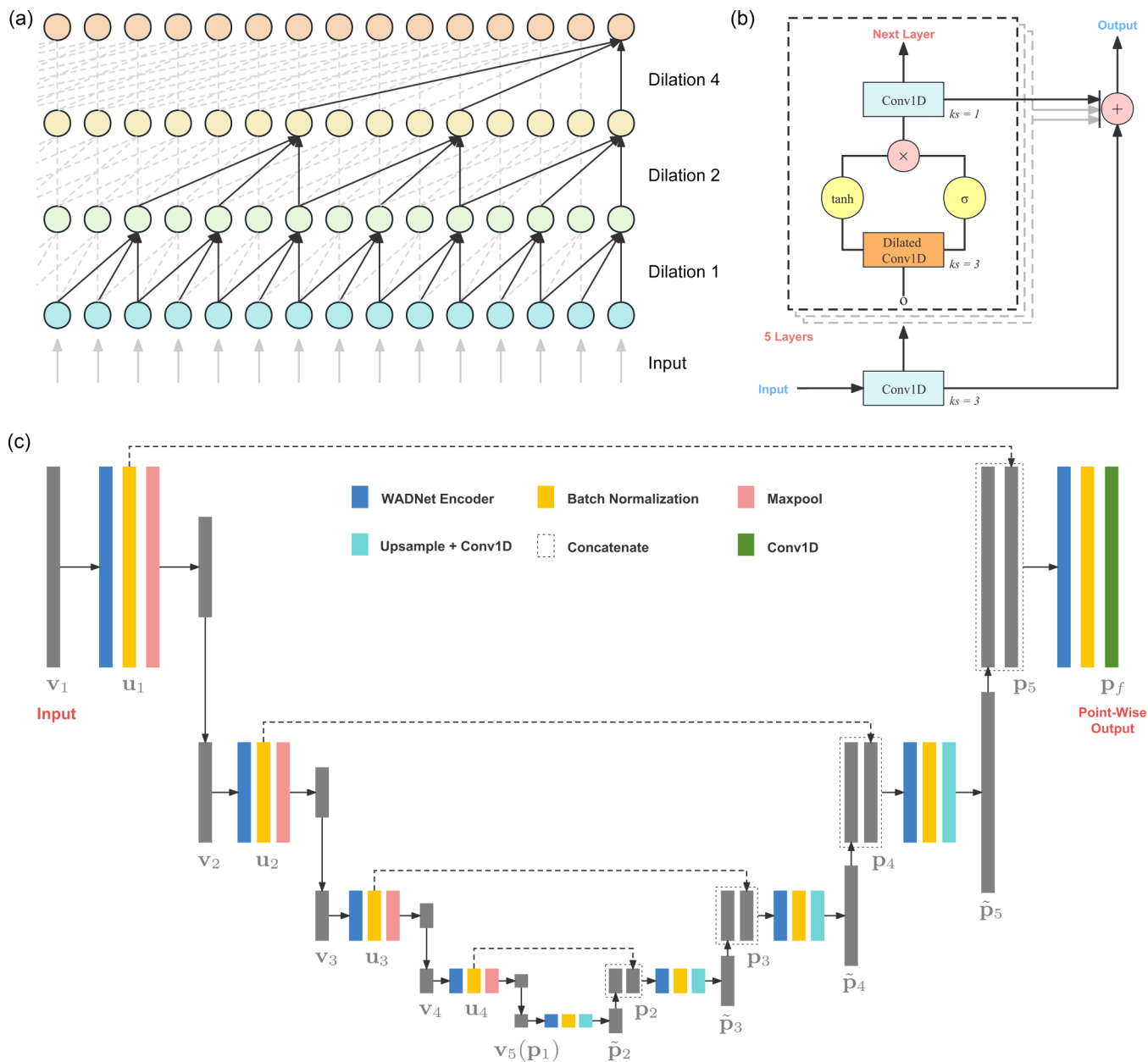
FIG. 2. (a) Schematic diagram of the dilated causal 1D convolution with a kernel size of 3. The dilation factor of convolution is doubled layer by layer. (b) Detailed structure of the WADNet encoder. (c) The architecture of our U-AnDi model. Skip connections are highlighted by the dashed lines.

decaying dependencies. Here, the gating mechanism within the GAU plays a pivotal role. Essentially, it acts as a regulatory system, determining which information should be allowed to pass through and which should be retained or discarded. Given the nature of anomalous diffusion, where certain events or states have prolonged effects, it is crucial for a model to discern which pieces of information are pertinent over extended periods. The gating mechanism allows the model to "remember" significant events from the past and "forget" or downplay less relevant ones. This selective retention on certain data points aligns perfectly with the characteristics of anomalous diffusion, ensuring that the model remains sensitive to long-time correlations. Therefore, to adeptly process the long-range information gathered by the dilated causal con-volution, we integrate the GAU into our network encoder. The detailed structure of this encoder is depicted in Fig. 2(b), with a comprehensive description in a mathematical form provided in Appendix A.

C. U-Net architecture

For semantic segmentation tasks, the precise delineation of boundaries between distinct objects or states necessitates a model's capability to effectively capture local variations. This aspect is as crucial as apprehending long-time correlations when it comes to the semantic segmentation and transition detection in anomalous diffusion. The reason is that transi-tions in diffusion states typically do not manifest as slow

TABLE I. Overview of comparative models, their respective abbreviations, and the presence (highlighted by ✓) or absence of key components: dilated causal convolution (DCC), gated activation unit (GAU), and U-Net architecture.

Model	Abbr.	DCC	GAU	U-Net
Long short-term memory [66]	LSTM		✓	
Gated recurrent unit [67]	GRU		✓	
Transformer [68]	TFM			
WADNet encoder [34]	WE	✓	✓	
U-CNN	U-CNN	✓		✓

and continuous processes but rather occur within short-time intervals.

Therefore, to enhance the model's ability in detecting these local changes, we integrate the WADNet encoder into the U-Net architecture. Such an architecture, originally designed for biomedical image segmentation, has garnered significant attention due to its exceptional performance in various segmentation tasks. As illustrated in Fig. 2(c), U-Net utilizes a symmetric encoder-decoder architecture, ensuring that spatial information diminished during encoding is effectively reinstated during decoding. Moreover, skip connections in U-Net transfer fine-grained details from the encoder directly to the decoder, ensuring that local features and nuances are preserved and enhanced in the output. That is, U-Net captures fine details through skip connections and grasps the broader context with its encoder-decoder structure. This dual capability makes it well suited for pinpointing state transition points in anomalous diffusion trajectories, where subtle local variations and long-range dependencies are closely interwoven. Details of constructing the U-Net architecture in this work are also provided in Appendix A.

IV. ANALYSIS OF MODEL PERFORMANCE ON SIMULATED TRAJECTORIES

In this section, we present the performance of the U-AnDi model in segmenting simulated trajectories of varying lengths. The results are based on evaluations conducted on the validation sets for both subtasks, each comprising 200 000 trajectories. Considering the "black box" nature of machine learning models, it is imperative to provide controls and calibrations analogous to standard tests in traditional analytical methods. For this reason, and to gain a deeper understanding of the individual contributions of U-AnDi's core components: DCC, GAU, and U-Net architecture, we also evaluate the performance of several comparative models. Our choice of comparative models is both strategic and comprehensive, encompassing the models outlined in Table I, with detailed specifications provided in Appendix C. By contrasting U-AnDi with these baseline models, we can distinctly discern the individual and synergistic contributions of DCC, GAU, and U-Net to the semantic segmentation of anomalous diffusion.

A. Performance of U-AnDi on subtask 1

To provide a more intuitive demonstration of U-AnDi's performance on subtask 1, we present four representative
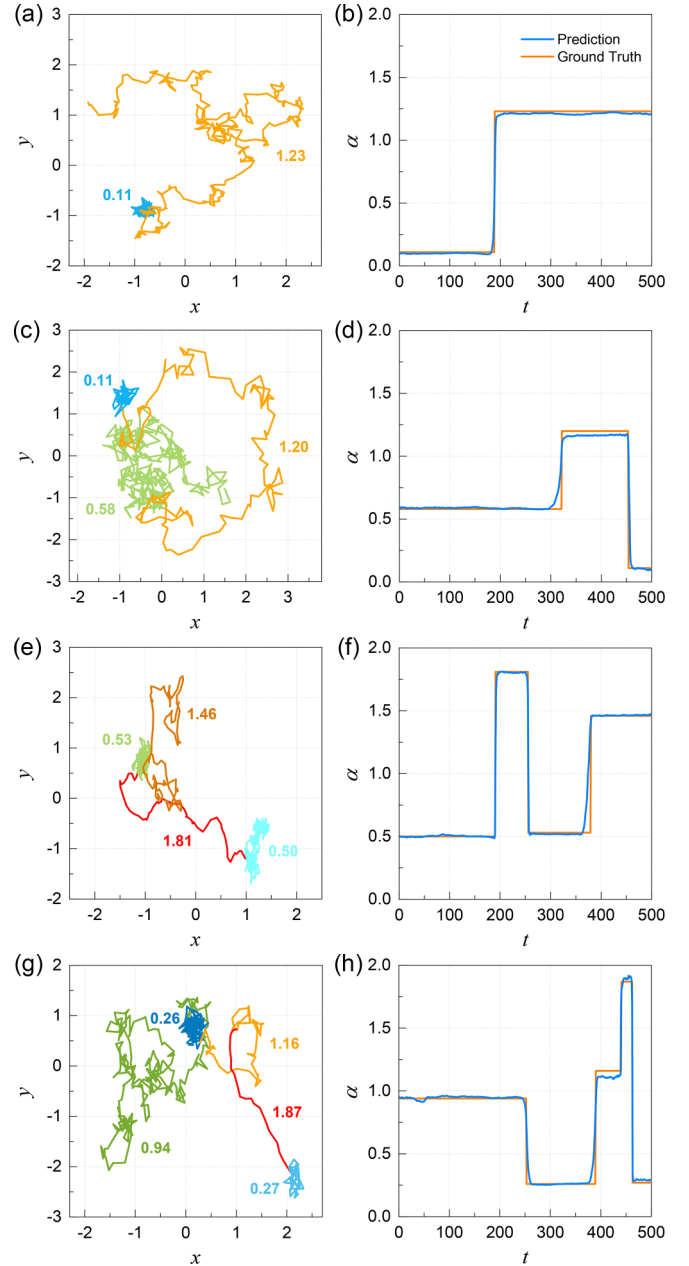

FIG. 3. Representative examples with $L = 500$ demonstrating U-AnDi's performance on subtask 1 for various number of segments M: (a)–(b) $M = 2$, (c)–(d) $M = 3$, (e)–(f) $M = 4$, (g)–(h) $M = 5$. Left panel: Trajectory visualizations, with segments of different diffusion exponents marked in distinct colors. Right panel: Comparisons between the predicted (blue) and ground truth (orange) diffusion exponents.

trajectories of length $L = 500$ and number of segments $M = 2, 3, 4, 5$ in the left panel of Fig. 3. The right panel compares their predicted values with the ground truth values. It can be observed that the predicted diffusion exponents are quite close to the ground truth values, with only minor differences noticeable in a few short intervals. Quantitatively, we show the MAE scores of U-AnDi on the validation set for various trajectory lengths spanning from 50 to 900 in Fig. 4(a). The performances of other
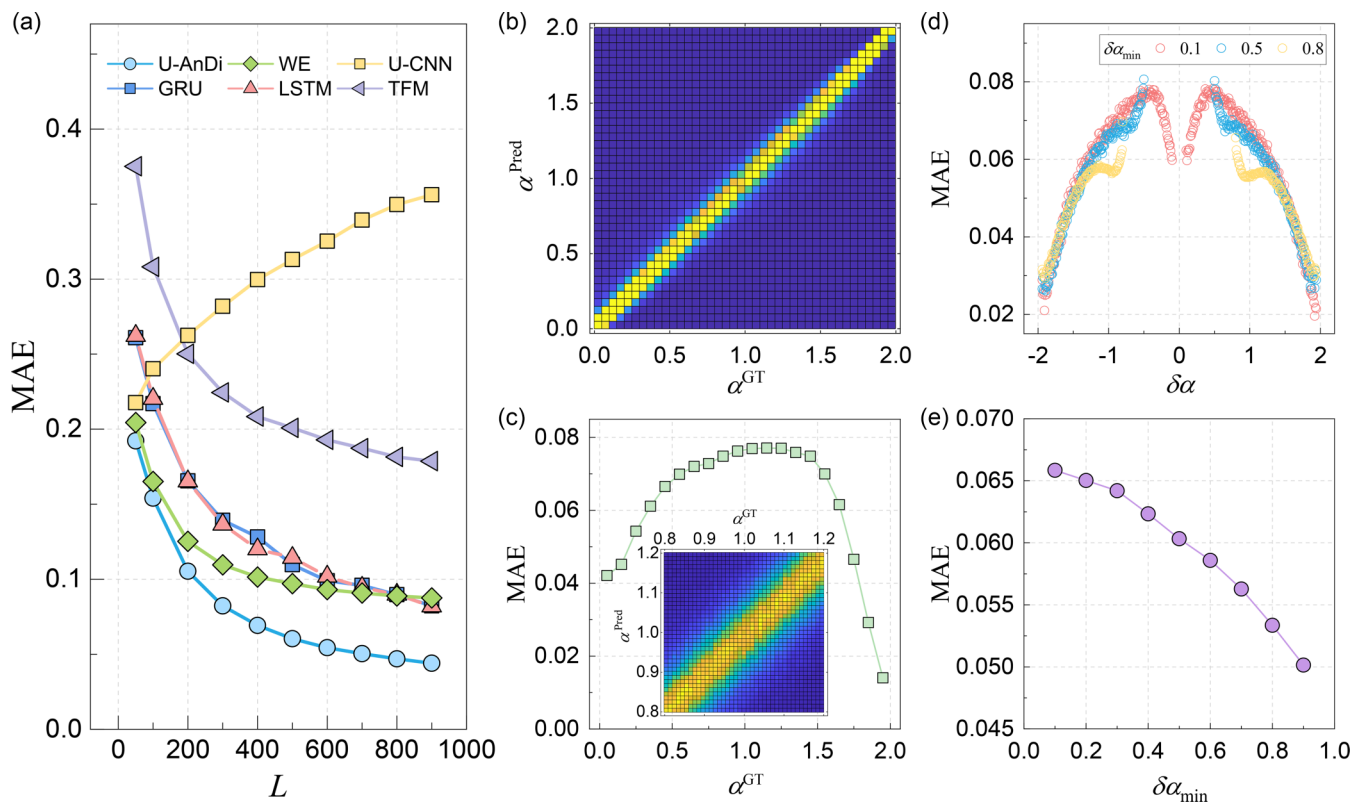
FIG. 4. (a) MAE scores of U-AnDi on the validation set for various trajectory lengths from 50 to 900, with performances of other comparative models presented for reference. (b) 2D histogram of prediction α^{Pred} vs ground truth α^{GT}. (c) Distribution of MAE for α^{GT} values ranging from 0.05 to 2. (Inset) Higher-resolution 2D histogram of α^{Pred} versus α^{GT} spanning from 0.8 to 1.2. (d) MAE variations corresponding to different $\delta\alpha$ for three scenarios: $\delta\alpha_{\min} = 0.1, 0.5,$ and 0.8. (e) Model performance, as characterized by MAE, across various values of $\delta\alpha_{\min}$. Here, (b)–(e) display the results for trajectories of length 500.

comparative models are also presented for reference. Evidently, across all trajectory lengths, U-AnDi consistently achieves the lowest MAE. This result underscores U-AnDi's robust capability, which is augmented by the synergistic effects of DCC, GAU, and U-Net, in discerning segments with different diffusion exponents. Interestingly, as the trajectory length increases, the MAE for all models, with the exception of U-CNN, consistently decreases. This enhancement in segmentation performance with longer trajectories can be attributed to the extended average length of individual segments. Such elongation offers the model a richer feature information for analysis.

Upon further analysis, models that integrate the GAU (U-AnDi, WE, LSTM, and GRU) outperform those without it (TFM and U-CNN). This pronounced difference highlights the pivotal role and efficacy of GAU in dealing with long-time correlations inherent in anomalous diffusion. Moreover, models equipped with DCC (U-AnDi and WE) exhibit an additional boost in performance, emphasizing the significance of DCC in capturing long-range dependence information for anomalous diffusion segmentation. However, the U-CNN, which integrates DCC but lacks GAU, experiences a decline in segmentation performance as trajectory length increases. This suggests that, in the absence of GAU, the DCC's ability to process long-time correlations is somewhat lacking. Additionally, it implies that the U-Net architecture, being more attuned

to local variations, might misinterpret features without the ability to effectively handle long-time correlations, especially as trajectory lengths grow.

Next, we use trajectories of length 500 to provide a detailed analysis of U-AnDi's segmentation performance on subtask 1. For a fine-grained comparison between predictions and ground truths of diffusion exponents, we present the 2D histogram of α^{Pred} versus α^{GT} in Fig. 4(b). As expected, the high-frequency pairs are predominantly located near the line $\alpha^{\text{Pred}} = \alpha^{\text{GT}}$, suggesting that the prediction errors of are small for all α^{GT} values ranging from 0.05 to 2. However, it should be noted that the prediction errors are not entirely uniform across different α^{GT} values. As displayed in Fig. 4(c), for α^{GT} values close to 0 or 2, the errors are relatively smaller, while for values around 1, the errors tend to be slightly larger. This result can be attributed to the characteristics of FBM, which approaches standard Brownian motion without long-time correlations when $\alpha \approx 1$ [25]. Compared to scenarios with pronounced long-range dependencies (where α is close to 0 or 2), its identification becomes relatively more challenging. Considering that $\alpha \approx 1$ might typify many experimental situations, we provide a higher-resolution histogram of α^{Pred} versus α^{GT} for α^{GT} ranging from 0.8 to 1.2 in the inset of Fig. 4(c). As observed, even though the MAE is generally larger around $\alpha^{\text{GT}} \approx 1$, the error distribution of U-AnDi's predictions remains relatively uniform without any
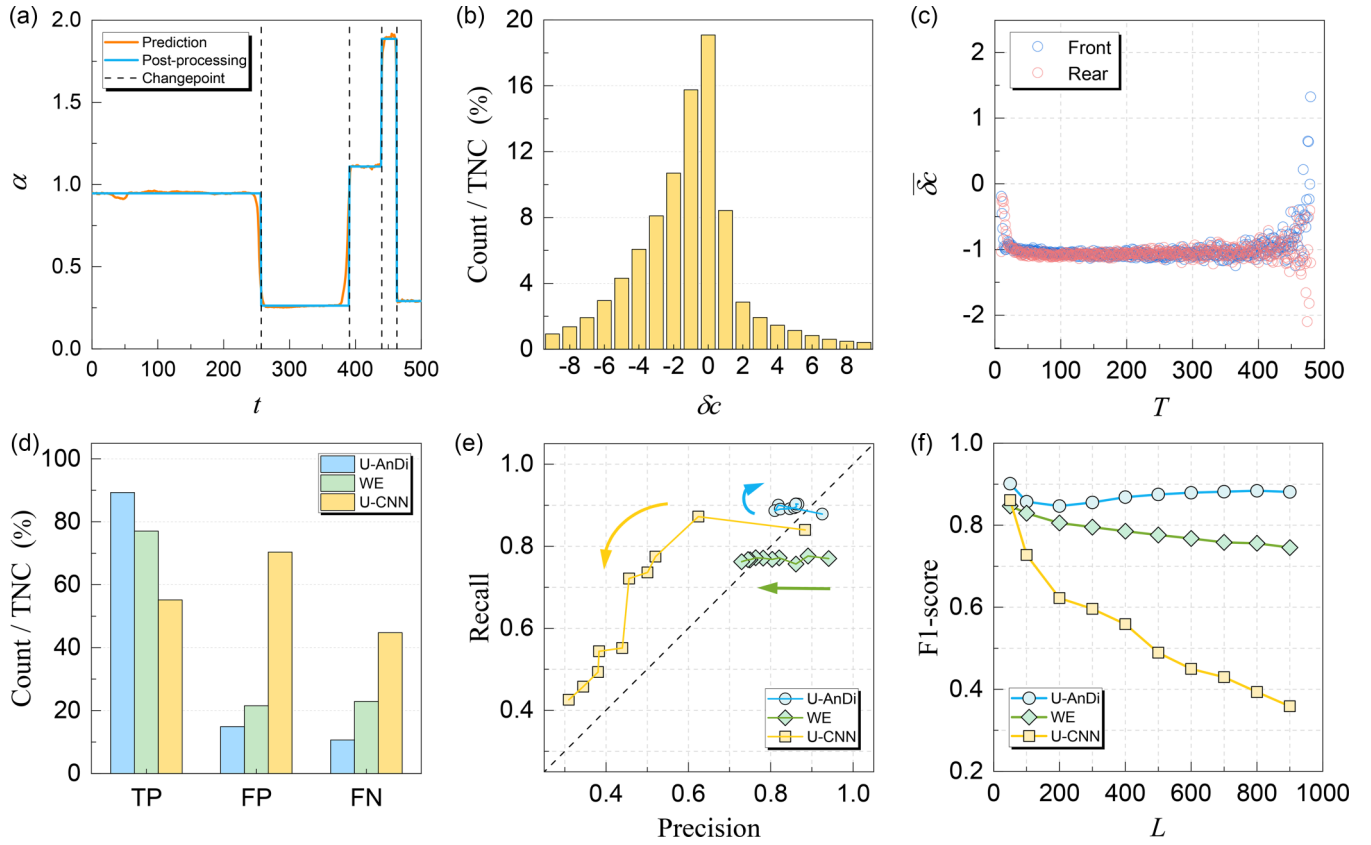
FIG. 5. (a) Representative example of the postprocessing technique in subtask 1, which converts the continuous pointwise predictions into discrete segmented predictions. The dashed lines denote the predicted locations of changepoints. (b) Distribution of δc as a percentage of TNC, spanning from –9 to 9. (c) Mean values of δc ($\overline{\delta c}$) corresponding to various lengths of the front and rear segments adjacent to changepoints. (d) Distribution of TP, FP, and FN as a percentage of TNC, juxtaposed with results from WE and U-CNN. Here, (b)–(d) show the results for trajectories of length 500. (e) Plots of recall versus precision for U-AnDi, WE, and U-CNN across different trajectory lengths, with arrows indicating the direction of increasing trajectory length. (f) F1-score variations across trajectory lengths for the three models.

pronounced bias, highlighting the stability of model's prediction for $\alpha \approx 1$.

Given that the variation in diffusion states between two adjacent segments can be characterized by the difference in their diffusion exponents, denoted as $\delta\alpha$, we investigate the influences of $\delta\alpha$ on U-AnDi's segmentation performance to quantify the effects of state variations. MAEs corresponding to different values of $\delta\alpha$ are presented in Fig. 4(d) for three scenarios: $\delta\alpha_{\min} = 0.1, 0.5$, and 0.8. A conspicuous feature of the MAE versus $\delta\alpha$ plots is their clear symmetry, indicating that whether the exponent α increases or decreases does not significantly affect U-AnDi's predictions. Moreover, in most cases, a larger magnitude of $|\delta\alpha|$ correlates with a lower MAE, suggesting that more pronounced differences in diffusion states facilitate the model's ability to discern transitions. However, an exception arises when $|\delta\alpha| \approx 0$, where this trend reverses. This is because a negligible $|\delta\alpha|$ might result in indistinct changes in diffusion states, prompting the model to treat two adjacent segments as a single diffusion state, thereby yielding errors commensurate with the magnitude of $|\delta\alpha|$. In addition, we showcase the model's overall performance across different $\delta\alpha_{\min}$ values in Fig. 4(e). Notably, a larger $\delta\alpha_{\min}$, which guarantees more significant diffusion state variations, results in a smaller MAE, aligning well with our expectations.

Besides the accuracy of predicting diffusion exponents, another crucial aspect is the identification of changepoints among different diffusion states. As shown in the right panel of Fig. 3, the U-AnDi model, which generates continuous pointwise values as its output, enables an approximate visual determination of the changepoint ranges. However, this output cannot provide precise demarcations between successive distinct diffusion states. To address this issue, we introduce a postprocessing technique that converts the continuous pointwise predictions into discrete segmented predictions (see Appendix D for details). An example of this process is illustrated in Fig. 5(a). After applying the post-processing, the predictions for all points within each segment become consistent, allowing for the identification of changepoints through the abrupt changes in diffusion exponents, as denoted by the dashed lines in Fig. 5(a). Based on this technique, we can further quantitatively investigate U-AnDi's performance on the detection of changepoints.

We define the total number of changepoints in the validation set as TNC and display the distribution of δc as a percentage of TNC in Fig. 5(b), spanning from –9 to 9. The majority of δc values are clustered around 0, signifying a relatively small error in the prediction of changepoints. Interestingly, the distribution demonstrates a mild leftward skew,
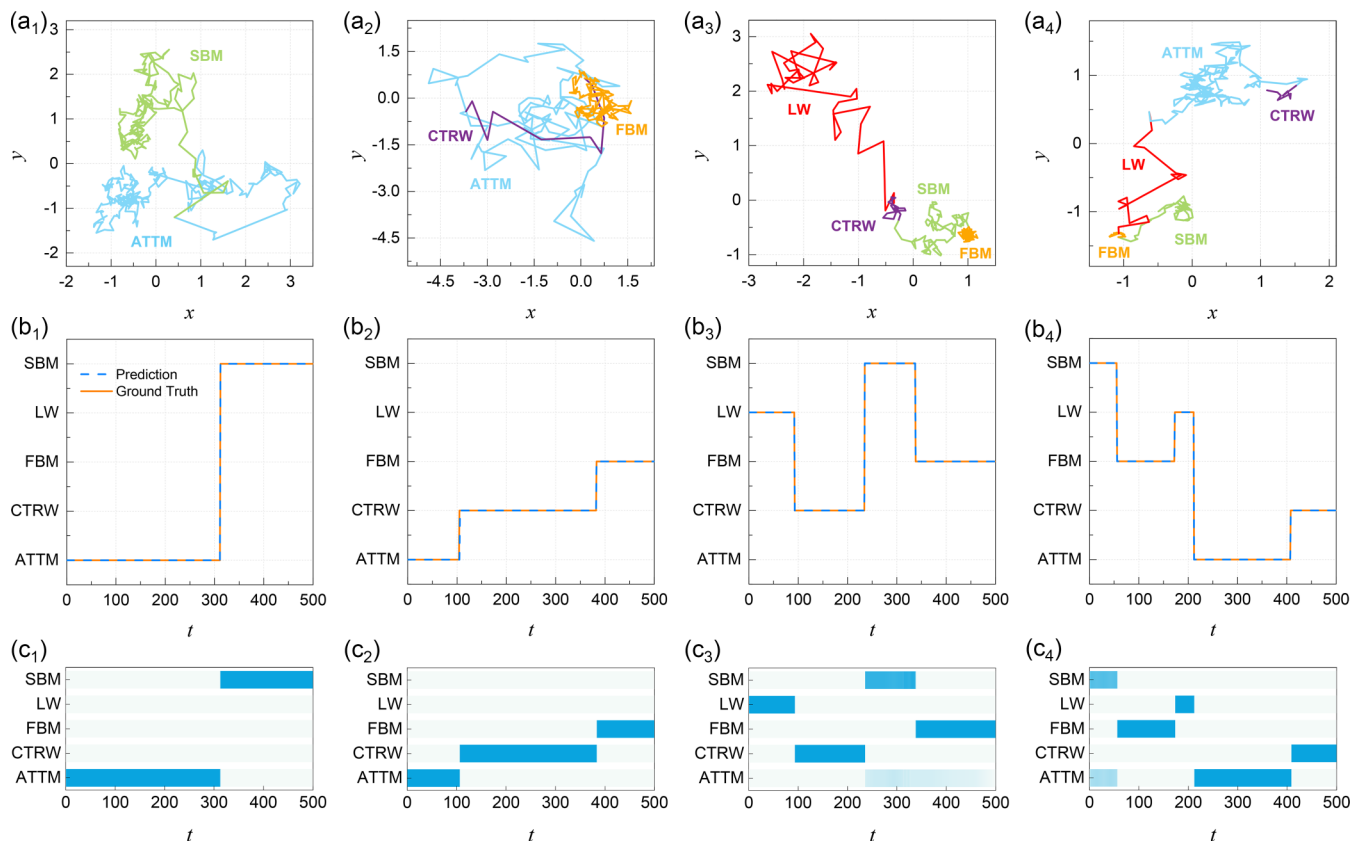
FIG. 6. Visualization of U-AnDi's performance on the segmentation task with different diffusion models. [(a_1)–(a_4)] Illustrations of four representative trajectories of length 500 with varying number of segments ($M = 2, 3, 4, 5$), where segments with different diffusion models are marked in distinct colors. [(b_1)–(b_4)] Comparisons of predictions (blue dashed lines) and ground truths (orange solid lines) for each trajectory. [(c_1)–(c_4)] Heatmaps of pointwise probability distribution for each trajectory, with darker colors indicating higher probabilities.

suggesting that the model has a slight inclination towards predicting changepoints earlier than their true positions. In addition, the relationship between δc and segment length T is also examined. As demonstrated in Fig. 5(c), we present the mean values of δc, denoted as $\overline{\delta c}$, corresponding to various lengths of the front and rear segments adjacent to the changepoints. It is evident that for the vast majority of segment lengths, $\overline{\delta c}$ consistently aligns close to -1. This trend reinforces our previous observation that U-AnDi has a propensity to slightly anticipate changepoints. In particular, when the segment length approaches the trajectory length, the prediction error increases significantly. Such an increase occurs when one segment dominates the majority of the trajectory, relegating other segments to near-minimal lengths. These short segments provide limited feature information, leading the model to yield less accurate predictions.

Expanding our analysis on the changepoint detection, we probe the distribution of TP, FP, and FN as a percentage of TNC and compare these metrics with the results from WE and U-CNN, as illustrated in Fig. 5(d). Obviously, U-AnDi demonstrates a higher count of TP while exhibiting fewer instances of FP and FN, highlighting its superior precision and recall. To validate the generality of this observation, we explore the distribution of precision and recall for the three models across different trajectory lengths, with results summarized in Fig. 5(e). Unsurprisingly, U-AnDi consistently outperforms both WE and U-CNN in terms of precision and

recall across all lengths. Moreover, as indicated by the arrows, which denote the direction of increasing trajectory length, U-AnDi exhibits remarkably consistent changepoint detection performance, maintaining stable precision and recall across varying lengths. In contrast, as trajectory length increases, WE shows a slight decline in precision while its recall remains stable; U-CNN, devoid of GAU, witnesses a more pronounced drop in both precision and recall. The F1-score variations across trajectory lengths for the three models are illustrated in Fig. 5(f). Mirroring the MAE trends seen in Fig. 4(a), U-CNN, without GAU, significantly underperforms in changepoint detection for longer trajectories compared to GAU-equipped U-AnDi and WE. These findings reemphasize the pivotal role of GAU in effectively managing the long-time correlations inherent in anomalous diffusion.

B. Performance of U-AnDi on subtask 2

In a manner consistent with subtask 1, we initially seek to visually demonstrate U-AnDi's performance on the segmentation of trajectories with different dynamics models. For this purpose, the segmentation results of four representative trajectories of length 500 with number of segments $M = 2, 3, 4, 5$ are summarized in Fig. 6. In more detail, Fig. 6 is organized into three rows for a coherent presentation. The first row [Figs. 6(a_1)–6(a_4)] presents the visualizations of these four trajectories, using unique colors to differentiate segments
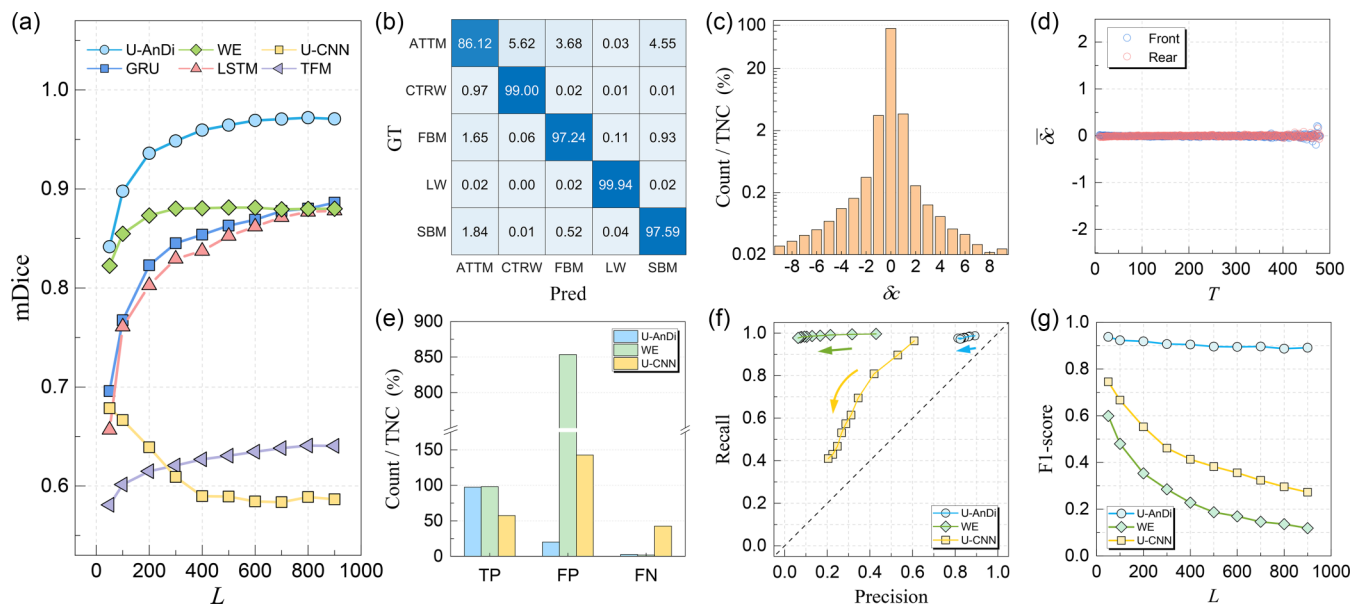
FIG. 7. (a) mDice scores on the validation set of U-AnDi and comparative models across trajectory lengths from 50 to 900. (b) Confusion matrix of the pointwise classification results of U-AnDi, where horizontal and vertical coordinates denote predicted and ground truth labels, respectively. (c) Distribution of δc spanning from –9 to 9 as a percentage of TNC, which is plotted on a logarithmic scale. (d) $\overline{\delta c}$ as a function of the lengths of the front and rear segments adjacent to changepoints. (e) Comparative distribution of TP, FP, and FN values for U-AnDi, WE, and U-CNN, expressed as a percentage of TNC. Here, (b)–(e) illustrate the findings for trajectories of length 500. (f) Recall against precision for U-AnDi, WE, and U-CNN across varying trajectory lengths, where arrows denote the direction of increasing trajectory length. (g) Dependence of F1-score on trajectory length for the three models.

associated with distinct diffusion models. Subsequently, comparisons of predictions with ground truths for each of the four trajectories are displayed in the second row [Figs. 6(b₁)– 6(b₄)]. It is evident that the lines depicting the predicted values (blue) closely align with those representing the ground truth values (orange). This highlights the robust performance of U-AnDi in segmenting trajectories with different diffusion models. Moreover, to provide a more comprehensive understanding of the model predictions, we show the heatmaps of predicted pointwise probability distribution in the third row [Figs. 6(c₁)–6(c₄)], with darker colors representing values closer to 1. As illustrated, in the majority of instances, the model is capable of predicting the specific model at a given step with a high degree of confidence (probability).

Building on the insights from subtask 1, we further quantitatively evaluate U-AnDi's performance on subtask 2. As depicted in Fig. 7(a), the mDice scores of U-AnDi and other comparative models across different trajectory lengths ranging from 50 to 900 align with the findings from subtask 1. This consistency not only underscores U-AnDi's robustness across different segmentation tasks of anomalous diffusion, but also reinforces our prior analysis about the functionalities of U-AnDi's core components: DCC, GAU, and U-Net. Further, we continue to use trajectories of length 500 as our primary example. To gain a deep insight into the error origins regarding this pointwise classification task, the confusion matrix of classification results is displayed in Fig. 7(b). As observed, while the classification accuracy for the CTRW, FBM, LW, and SBM models exceeds 97%, the accuracy for the ATTM model is merely around 86%. This discrepancy can be attributed to the machine-learning features of ATTM extracted by the WADNet encoder, which bear resemblance

to CTRW, FBM, or SBM, as elucidated in Ref. [34]. Consequently, this leads U-AnDi to occasionally misclassify ATTM steps as CTRW (5.62%), FBM (3.68%), or SBM (4.55%).

On the other hand, detecting changepoints among diffusion states holds significant importance in subtask 2 as well. In contrast to subtask 1, no additional postprocessing techniques are required for the detection in this task. We can efficiently determine the positions of changepoints by examining abrupt changes in the predicted labels. In Fig. 7(c), the distribution of δc values ranging from –9 to 9 is presented as a percentage of TNC. A distinction from subtask 1 is that 91.1% of the cases have δc values of 0, indicating a significant portion of the predicted changepoints aligning perfectly with the ground truths. Given the highly imbalanced distribution of δc values, a logarithmic scale is employed in Fig. 7(c) to enhance visualization. Moreover, $\overline{\delta c}$ as a function of the lengths of segments adjacent to changepoints is explored, with the results depicted in Fig. 7(d). Differing from subtask 1, the majority of data points for both front and rear segments are concentrated around 0. Even as the segment length nears the trajectory length, there is no significant escalation in error. This suggests that U-AnDi exhibits commendable stability in segmenting trajectories with varying diffusion models, even when dealing with very short segments.

Similar to subtask 1, regarding U-AnDi, WE, and U-CNN, the distribution of TP, FP, FN for $L = 500$, along with the plots of recall versus precision across varying lengths and the F1-scores for different lengths are illustrated in Figs. 7(e), 7(f), and 7(g) respectively. While the results largely mirror those from subtask 1, a stark deviation is observed in the FP count of WE, which have increased dramatically [Fig. 7(f)]. This surge adversely impacts the WE's precision, rendering its F1-score

lower than both U-AnDi and U-CNN [Fig. 7(g)]. Considering that WE lacks the U-Net architecture, it can be deduced that such an architecture plays a crucial role in adeptly processing intricate local fluctuations intertwined with long-time correlations for the segmentation of anomalous diffusion.

Building on insights from both subtasks, we further investigate the case where the diffusion exponent and the dynamic model concurrently vary within a single trajectory. Detailed findings of this analysis are presented in Appendix E, which confirms U-AnDi's adaptability and robustness in the face of such complex conditions.

C. Extended analysis on the segmentation of anomalous diffusion

While U-AnDi has demonstrated commendable segmentation performance in discerning changes in diffusion exponents and transitions among diffusion models, it remains imperative to assess its capabilities in more intricate and broadly relevant anomalous diffusion or single-particle tracking scenarios. The robustness and widespread applicability of U-AnDi in diverse contexts are pivotal for its integration into practical applications, offering significant utility to potential researchers in the anomalous diffusion field who may not be well versed in machine learning.

For that purpose, we expand our analysis to encompass three other distinct segmentation tasks in this section. Each task is designed to mirror the complexities often encountered in the domain of anomalous diffusion, serving as a basis for a rigorous examination of U-AnDi's versatility in handling complex anomalous diffusion scenarios.

1. Semantic segmentation of single-model trajectories with constant parameters

In the landscape of anomalous diffusion research, while transitions between diffusion states are frequently observed, it is equally common to encounter scenarios where a random walker moves exclusively within a framework of a single diffusion model [1]. These single-model trajectories, devoid of state transitions, represent a fundamental and ubiquitous aspect of experimental studies. In our previous discussions, we center on trajectories that exhibit state transitions (with $M \geqslant 2$). This emphasis naturally leads to considerations about U-AnDi's performance on single-model trajectories. Such an evaluation is crucial as it offers potential users insights into U-AnDi's stability and reliability across an expansive range of diffusion scenarios.

For this investigation, we generate trajectories of length 500 using six diffusion models: ATTM, CTRW, FBM, LW, SBM, and the Brownian diffusion (BD). BD is included because it represents the most fundamental diffusion phenomenon in experimental observations. Here, it should be noted that BD is equivalent to the scenario of FBM and SBM when the diffusion exponent $\alpha = 1$. To ensure the generality of our results, we extensively evaluate U-AnDi across a broad range of core parameters for the diffusion models. For the first five models, the selected parameter is α, with its range consistent with the description in Sec. II B for each model. For BD, the parameter is the diffusion coefficient D, which spans a range from 0.5 to 10. Further, to guarantee statistical
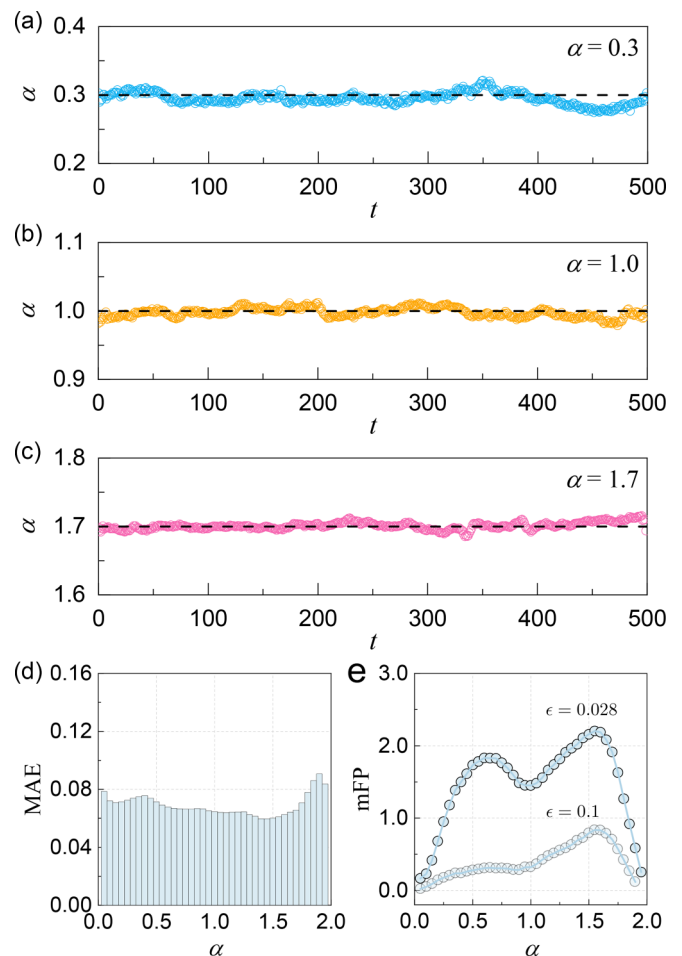

FIG. 8. [(a)–(c)] Predictions of U-AnDi-Exponent for FBM trajectories with diffusion exponents $\alpha = 0.3$, 1.0, and 1.7, representing subdiffusion, Brownian diffusion, and superdiffusion, respectively. (d) MAE scores for various α values. (e) Distributions of the mFP for different α with threshold $\epsilon = 0.028$ and 0.1, respectively.

robustness, we generate 10 000 trajectories for each parameter value for testing. For convenience, we introduce two nomenclatures in this section: "U-AnDi-Exponent" refers to the model from subtask 1 that segments based on diffusion exponent, while "U-AnDi-Model" pertains to the model from subtask 2 that segments based on the diffusion model.

We first assess the performance of U-AnDi-Exponent on FBM trajectories with fixed diffusion exponents α. To visualize the variation between the predicted values and the fixed exponent, we display the predictions of U-AnDi-Exponent for three representative trajectories with $\alpha = 0.3$ (subdiffusion), 1.0 (Brownian diffusion), and 1.7 (superdiffusion) in Figs. 8(a)–8(c), respectively. Across these scenarios, the model's predictions closely cluster around the true values, indicating only slight variation. Quantitatively, we present the MAE scores for various α values in Fig. 8(d). It is evident that MAE values are all approximately around 0.07, suggesting consistently minor fluctuations.

Furthermore, we apply the postprocessing technique described in Appendix D to identify changepoints. Given that these trajectories inherently lack changepoints, our analysis
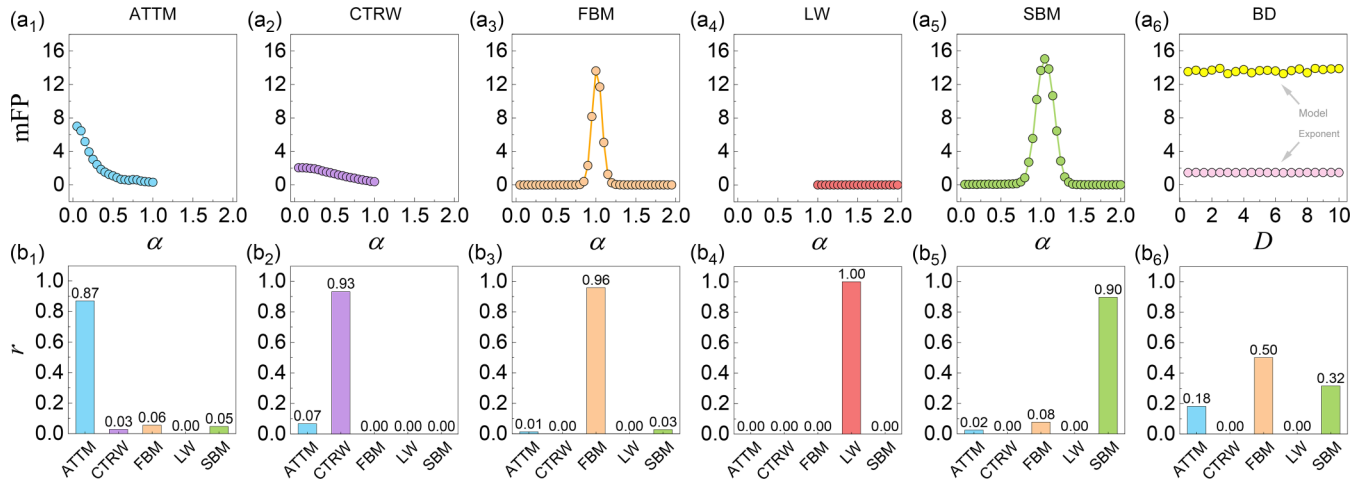
FIG. 9. (a_1)–(a_6) Distribution of mFP predicted by U-AnDi-Model for trajectories generated by six different diffusion models: ATTM, CTRW, FBM, LW, SBM, and BD, respectively. The performance of U-AnDi-Exponent is also given in (a_6) for BD trajectories for comparison. (b_1)–(b_6) The mean proportion of predicted diffusion models within each individual trajectory for ATTM, CTRW, FBM, LW, SBM, and BD, respectively.

concentrates on the FPs in the predictions. We define the mean number of FPs per trajectory within the dataset as mFP and illustrate the distribution of mFP for different α in Fig. 8(e). Despite the fluctuation of mFP in response to changes in α, the values consistently remain low. Notably, by increasing the threshold ϵ in the postprocessing technique, the model's precision can be further improved, as evidenced by a reduction in mFP [Fig. 8(e)]. Overall, U-AnDi-Exponent demonstrates a remarkable stability in dealing with FBM trajectories with constant parameters. In particular, the flexibility of our approach allows users to refine the precision of changepoint detection through threshold customization, thereby tailoring the model to their specific needs.

On the other hand, we employ U-AnDi-Model to test single-model trajectories generated by the six diffusion models. In addition to utilizing mFP as the evaluation metric [Figs. 9(a_1)–9(a_6)], the mean proportion of predicted diffusion models within each individual trajectory, denoted as r, is displayed in Figs. 9(b_1)–9(b_6) to manifest the variation between predicted and true values. It is noteworthy that for ATTM, CTRW, FBM, LW, and SBM, the true model accounts for at least 87% of the predicted composition, underscoring the robustness of our method in this task. Since U-AnDi-Model primarily relies on the distinctive features of diffusion models to determine changepoints, the magnitude of mFP on single-model trajectories is largely influenced by the uniqueness of the diffusion model relative to others. For instance, the characteristics of LW are notably distinct from the other five models (Ref. [34]), allowing for a stable identification of LW without predicting state transitions, i.e., mFP = 0 for all α [Fig. 9(a_4)]. Conversely, for FBM and SBM, as α approaches 1, the trajectories of these two models converge towards the characteristics of Brownian diffusion. This convergence results in an overlap that challenges U-AnDi's discriminative power, leading to an increased incidence of FPs [Figs. 9(a_3) and 9(a_5)]. This phenomenon is also evident when U-AnDi-Model is tasked with analyzing BD trajectories, where the mFPs closely mirror those observed for FBM and SBM with $\alpha = 1$

and are considerably higher than the forecasts by U-AnDi-Exponent [Fig. 9(a_6)]. As shown in Fig. 9(b_6), predictions for BD trajectories show a significant proportion attributed to both FBM and SBM, emphasizing the indistinguishability between these two at $\alpha = 1$.

2. Segmentation of simulated multistate trajectories in biological experiment scenarios

In the realm of single-particle tracking experiments within biological systems [70–75], the microdynamics of the observed entities are prone to transitions due to environmental influences, necessitating effective methods for segmenting the trajectory data. However, in such experimental systems, the diffusion dynamics often observed are not limited to specific anomalous diffusion models but include transitions among Brownian, confined, directed, and immobile states. These states correspond to

(1) Brownian diffusion (BD): Free movement characterized by random walks, reflecting the absence of constraints in an open environment.

(2) Confined diffusion (CD): Restricted movement within certain boundaries, such as within cellular compartments or membranes.

(3) Directed Motion (DM): Motion driven by the external force or energy, such as bacteria propelled by flagellar motion or cells migrating in response to chemical signals.

(4) Immobile state (IS): A state where particles are trapped or bound to structures within the cell, exhibiting little to no movement over time.

For instance, as described in Ref. [38], molecules like CD44 in macrophages can exhibit these states. CD44 molecules display confined diffusion when interacting with the actin cortex, free diffusion when such interactions are inhibited, and immobility possibly when tethered to the cortex.

Hence, to validate the effectiveness of U-AnDi in such biological experimental scenarios, we primarily adopt the simulation methods from Ref. [38] to generate segments
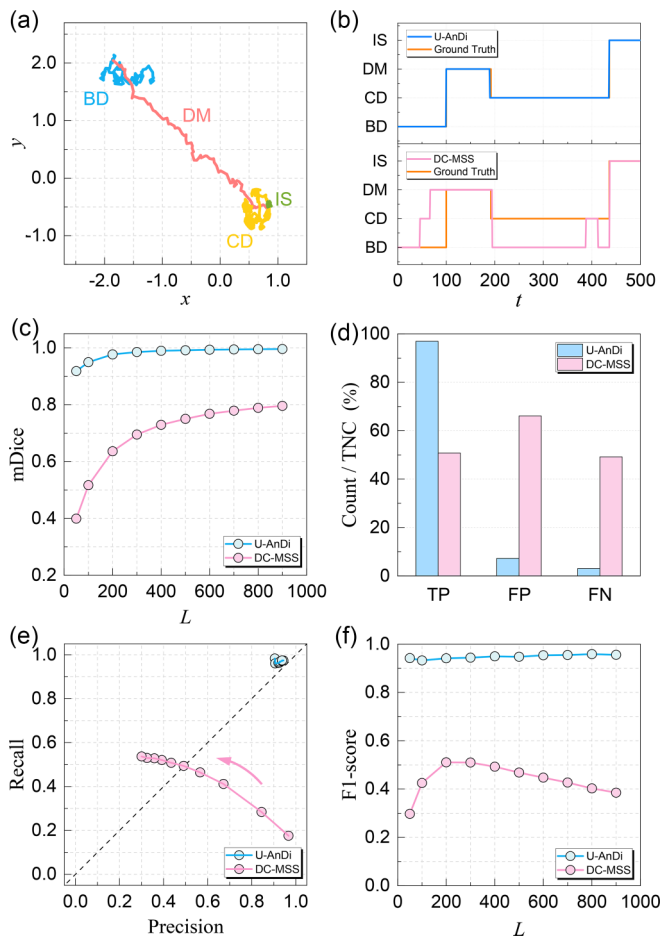
FIG. 10. (a) Representative example of a simulated composite trajectory of length 500, containing the four dynamic states in biological experiment scenarios: Brownian diffusion (BD), confined diffusion (CD), directed motion (DM), and immobile state (IS). (b) Comparative segmentation results of U-AnDi and DC-MSS on the composite trajectory in (a). (c) mDice score comparisons between U-AnDi and DC-MSS for trajectories of varying lengths. [(d)–(f)] Analysis of changepoint detection for U-AnDi and DC-MSS. (d) Distribution of TP, FP, and FN as a percentage of TNC where $L = 500$. (e) Recall vs precision for different trajectory lengths. (f) The variation of the F1-score with trajectory length.

representing these four dynamic states. We then combine these segments using the same approach as in subtask 2 to create composite trajectories for the validation, with a representative example of length $L = 500$ illustrated in Fig. 10(a). The simulation methods for these segments are detailed in Appendix F. The training of U-AnDi follows the same protocol as in subtask 2.

Furthermore, to observe the potential advancements of our model over conventional approaches, we benchmark the performance of U-AnDi against the traditional divide-and-conquer moment scaling spectrum (DC-MSS) analysis method [38]. A brief introduction to DC-MSS is provided in Appendix C. As displayed in Fig. 10(b), we present the segmentation results of both methods on the trajectory shown in Fig. 10(a). It is evident that U-AnDi's predictions align almost perfectly with the ground truth, whereas DC-MSS exhibits

numerous misjudgments, including incorrect state identification and changepoint detection errors. Quantitatively, we compare the mDice scores of both models for trajectories of varying lengths, ranging from 50 to 900, and summarize the results in Fig. 10(c). Notably, U-AnDi decisively outperforms DC-MSS across all lengths with mDice close to 1, indicating a significantly higher accuracy in state identification.

Similarly, U-AnDi's performance in detecting state transitions is markedly superior to that of DC-MSS. As with subtask 2, Figs. 10(d)–10(f) detail the distribution of TP, FP, and FN as a percentage of TNC for $L = 500$, plots of recall versus precision for different trajectory lengths, and the variation of the F1-score with trajectory length, respectively. Across various lengths, U-AnDi demonstrates high precision, recall, and F1-scores, while DC-MSS lags in performance. Additionally, as trajectory length increases, DC-MSS shows a gradual increase in recall but a decrease in precision, revealing a trade-off between the two metrics [Fig. 10(e)]. This trade-off results in uniformly low F1-scores for all lengths, which is evident from Fig. 10(f).

These findings underscore U-AnDi's exceptional segmentation capabilities for anomalous diffusion trajectories in biological experimental scenarios. They also highlight the significant performance gap between unsupervised algorithms trained with unlabeled data like DC-MSS and supervised approaches dealing with labeled data, reinforcing the superiority of the latter in trajectory segmentation tasks.

3. Evaluating U-AnDi with other correlations

Considering that long-time correlations are a fundamental property inherent to anomalous diffusion behaviors, we aim to further validate the versatility and adaptability of U-AnDi in segmenting trajectories characterized by more correlations. To achieve this, we generate trajectories where the velocity autocorrelation functions (VACFs) conform to four distinct correlations. Subsequently, U-AnDi is employed to segment these trajectories. The chosen correlation functions for this evaluation are listed in the following paragraph. To maintain clarity and without loss of generality, we present the mathematical expressions in a one-dimensional framework in this section.

We denote $C(t) = \langle v(0)v(t) \rangle$ as the VACF. Four correlation functions are detailed below:

(1) Exponential decay (exp):

$$C(t) = Ae^{-t/\tau}, \tag{8}$$

where A is the amplitude and τ is the time constant.

(2) Multi-exponential decay (multi-exp):

$$C(t) = A[\phi e^{-t/\tau_1} + (1 - \phi)e^{-t/\tau_2}]. \tag{9}$$

Here, A is the amplitude, ϕ is the weighting factor. τ_1 and τ_2 are the time constants of two exponential decays, respectively.

(3) Exponentially damped cosine wave (exp-cos):

$$C(t) = Ae^{-t/\tau} \cos \omega t, \tag{10}$$

where A is the amplitude, τ is the time constant, and ω is the angular frequency.

TABLE II. Summary of parameters, labels, and corresponding equations to calculate labels for segmentation tasks associated with the four correlation functions.

Correlation type	Parameter	Label	Label range	Label calculation
exp	change A, set $\tau = 10$	D	[0.05,2)	$D = A\tau$
multi-exp	change A, set $\tau_1 = 10, \tau_2 = 20, \phi = 0.5$	D	[0.05,2)	$D = A[\phi\tau_1 + (1 - \phi)\tau_2]$
exp-cos	change A, set $\tau = 10, \omega = 0.1$	D	[0.05,2)	$D = A\tau/(1 + \tau^2\omega^2)$
M-L	change λ, set $\tau = 10, A = 1$	α	[0.05,2)	$\alpha = 2 - \lambda$

(4) Mittag-Leffler decay (M-L) [76]:

$$C(t) = \frac{A}{\tau^\lambda} E_\lambda[-(t/\tau)^\lambda],\qquad(11)$$

where

$$E_\lambda(z) = \sum_{k=0}^{\infty} \frac{z^k}{\Gamma(\lambda k + 1)}.\qquad(12)$$

Here, τ acts as a characteristic memory time and A is the amplitude coefficient. λ is the exponent satisfying $\lambda > 0$. $\Gamma(\cdot)$ denotes the Gamma function.

We generate these trajectories using a modified version of the "fbm" Python package [77,78], which employs the Davies-Harte algorithm [79]. Associated codes are available in our GitHub repository [80]. The effectiveness of this generation method is validated in Appendix G.

Labels for the trajectory segmentation are determined based on the diffusion behaviors corresponding to their correlation functions. Using the equation below [81]

$$\text{MSD} = 2\int_0^t (t - s)C(s)ds,\qquad(13)$$

it is evident that trajectories exhibiting exp, multi-exp, and exp-cos correlations manifest long-time Fickian behaviors (see Appendix G), i.e., MSD $\sim t$. Consequently, for these three types of correlations, we utilize the long-time diffusion coefficient D as the target for trajectory segmentation. On the other hand, as outlined in Ref. [82], the diffusion exponent α of trajectories characterized by the M-L correlation equals $2 - \lambda$. Therefore, for trajectories with the M-L correlation, we choose α as the segmentation label. In Table II, we provide a summary of the parameters, labels, and corresponding equations to calculate segmentation labels (see Appendix G for details) with respect to the four correlations.

Trajectory segments are combined to form a composite trajectory using the same method as in subtask 1. Additionally, the model and training approach remain consistent with those applied in that subtask. We evaluate U-AnDi's segmentation performance using MAE and F1-score, and compare it with WE and U-CNN. The results, illustrated in Fig. 11, encompass not only the performance of U-AnDi for the four distinct correlations but also its evaluation scores in subtask 1, which pertains to the power-law correlation. A clear observation from the results is the superior performance of U-AnDi in comparison to both WE and U-CNN across all five types of correlations, whether evaluated based on MAE or F1-score. Such consistent superiority across diverse correlation types highlights U-AnDi's robust adaptability, emphasizing its

potential to serve as a universal scheme for the segmentation of time-series data.

Interestingly, as quantified by the MAE, both U-AnDi and WE exhibit a notably better performance in predicting the diffusion exponent compared to the diffusion coefficient. However, this distinction is not evident in the case of U-CNN. A potential reason for this difference can be attributed to the inherent characteristics of diffusion coefficient and diffusion exponent. While the diffusion coefficient primarily captures the "speed" of diffusion, the diffusion exponent provides deeper insights into the underlying mechanisms of diffusion process, inherently reflecting the correlation information. Consequently, the diffusion exponent can be more readily predicted by models like U-AnDi and WE, both of which are adept at processing long-time correlations with the help of GAU. In contrast, U-CNN, which lacks the GAU, does not exhibit a similar advantage.

V. INTERPRETABILITY OF U-AnDi'S CORE COMPONENTS

The quest for interpretability in machine learning, particularly within the realms of physical sciences, is not merely a pursuit of operational transparency but a foundational
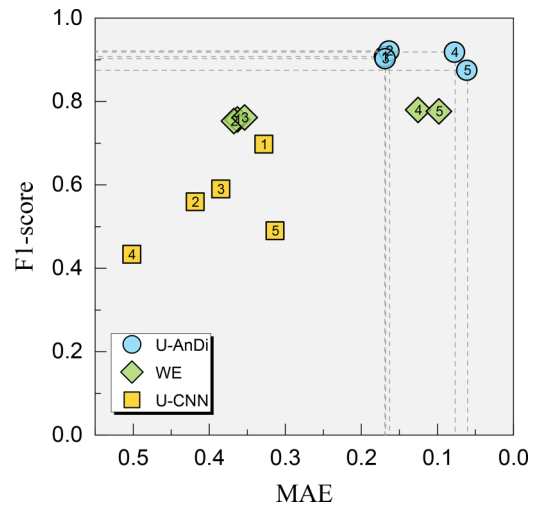

FIG. 11. Comparative evaluation of U-AnDi, WE, and U-CNN using MAE and F1-score across five distinct correlations: (1) exponential decay; (2) multi-exponential decay; (3) exponentially damped cosine wave; (4) Mittag-Leffler decay; (5) power-law decay. The dashed lines represent the best score for each correlation type, and the numbers on the dots correspond to the aforementioned correlation types.

requirement for scientific validity. In this paper, the comparative evaluation of U-AnDi against other models partially diminishes the "black box" associated with our machine learning model, shedding light on the inner working of U-AnDi's core components: DCC, GAU, and U-Net. As long-time correlation is one of the fundamental properties of anomalous diffusion, these components faithfully encapsulate this underlying nature and are in strict accordance with our design principles. Within this section, we consolidate the interpretability insights that has been gained from the comparison among the performance of U-AnDi and its comparative models in the aforementioned tasks.

(1) *Dilated causal convolution: Capturing long-time correlations.* To effectively capture the pivotal long-time correlations inherent in anomalous diffusion, DCC, with its exponentially expanding receptive fields, is designed to encompass the full spectrum of these extensive dependencies. This enables DCC to acquire a comprehensive integration of temporal information across scales, ensuring that the long-time correlations are accurately represented within the model's predictions. Such an advantage is reflected in the enhanced performance of U-AnDi as well as WE's competitive superiority over other counterparts, where both models incorporate the DCC.

(2) *Gated activation unit: Processing long-range dependencies.* To adeptly navigate the intricate long-time correlations in anomalous diffusion, the integration of GAU is engineered to leverage its gating mechanism to selectively emphasize salient temporal features. This selective filtration and prioritization by GAU are not merely structural benefits but are directly responsive to the characteristic memory effects and nonlocal interactions of anomalous diffusion. By modulating the flow of information, the GAU ensures that the model's focus is attuned to the most critical aspects of the data, which are indicative of the underlying long-time correlations. The superior performance of models equipped with GAU across all subtasks robustly demonstrates the proficiency of GAU in effectively processing long-time correlation information.

(3) *U-Net architecture: Enhancing detection of local changes.* When segmenting the anomalous diffusion trajectories, the abrupt transitions and heterogeneous local dynamics demand a model's ability not only in emphasizing long-time correlations but also in precisely delineating local variations. The U-Net architecture excels in this regard, capturing fine details through skip connections, while concurrently grasping the broader context with its encoder-decoder framework. This dual capability makes U-Net exceptionally well equipped to identify state transition points in anomalous diffusion. In this paper, the integration of WE within the U-Net framework further refines the model's segmentation performance, covering all metrics. This result reinforces the utility of U-Net in the segmentation and changepoint detection of anomalous diffusion.

The interpretability of U-AnDi's core components is a tangible reflection of the model's alignment with the complex nature of anomalous diffusion. By capturing long-time correlations and enhancing the detection of local changes, these core components collectively contribute to the model's robust performance across varied segmentation tasks, bridging the gap between advanced machine learning techniques and the fundamental principles of anomalous diffusion.

VI. IMPLEMENTATION OF U-AnDi ON REAL-WORLD ANOMALOUS DIFFUSION DATA

In this section, we apply the U-AnDi model to real-world anomalous diffusion data and examine its performance on the semantic segmentation of experimental trajectories. The selected real-world system is the diffusion of transmembrane proteins on cell membrane surfaces, as illustrated by the schematic diagram in Fig. 12(a). The pertinent data are obtained through single-particle tracking techniques [70–75] from experimental observations, and are provided by N. Granik *et al.* in Ref. [57]. The content of this section is structured as follows. Section VI A provides an overview of the raw experimental trajectory dataset, and describes the segmentation tasks along with the approach for preparing the validation set. The training strategy is outlined in Sec. VI B. In Sec. VI C, we demonstrate and discuss the performance of the U-AnDi model on interpreting the anomalous diffusion dynamics of transmembrane proteins.

A. Dataset description and segmentation tasks

The dataset with respect to the diffusion of transmembrane proteins on membrane surfaces comprises a total of 70 221 2D trajectories, in which the x and y pixel locations are provided. The distribution of trajectory lengths is displayed in Fig. 12(b), with the counts plotted on a logarithmic scale to enhance clarity in visualization. An evident long-tail distribution can be identified. In detail, trajectories of lengths no greater than 100 account for 98.6% (69 259) of the entire dataset, while the maximum length in this dataset reaches 1406. Diffusion models of these trajectories have been previously assessed by machine learning approaches in Ref. [57], with N. Granik *et al.* proposing that the predominant dynamics consist of both CTRW and FBM. Therefore, this dataset offers us an excellent opportunity to test the segmentation performance of U-AnDi on experimental data.

Two specific segmentation tasks are addressed here: one involving labeled data and the other utilizing unlabeled data. In the task with labeled data, the validation set is prepared as below. Trajectories to be evaluated are generated by combining short trajectories with lengths not exceeding 100 from the experimental dataset. Diffusion-model labels of these short trajectories are assigned by WADNet, a deep learning model capable of providing the probability of the diffusion model to which the input trajectory is most likely to belong. Further details regarding WADNet can be found in our previous paper [34]. Leveraging the model probabilities provided by WADNet for these short trajectories, we carefully select CTRW and FBM instances that exhibit high confidence levels (with probabilities of CTRW or FBM exceeding 0.9) to generate extended trajectories for further evaluation. Representative samples of these selected short trajectories with different lengths are illustrated in the inset of Fig. 12(b). The quantities of short CTRW and FBM trajectories spanning various lengths from 10 to 100 are presented in Fig. 12(c). By setting the overall trajectory length L and the number of segments
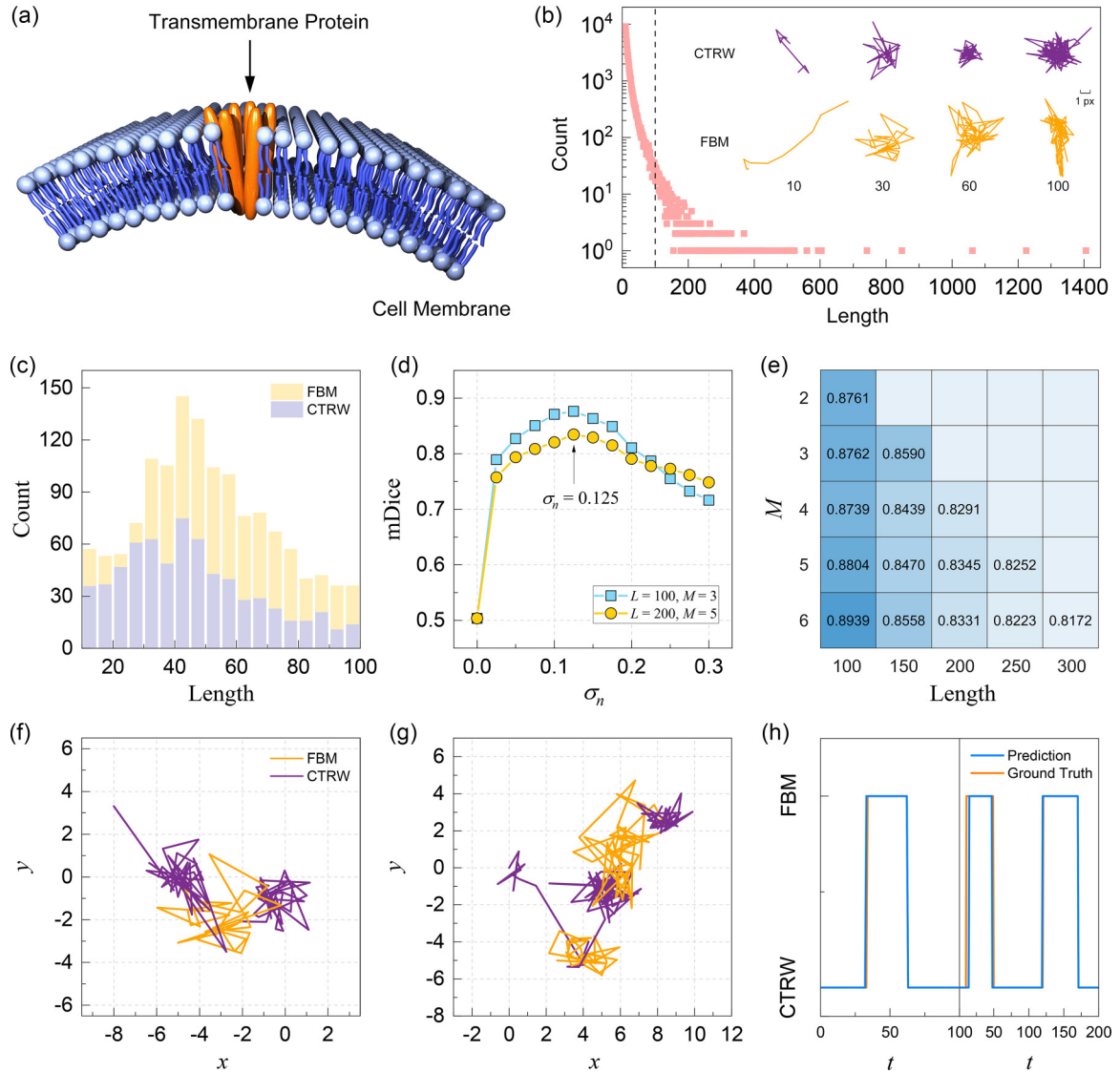
FIG. 12. (a) Schematic diagram of the selected real-world anomalous diffusion system: Diffusion of transmembrane proteins on cell membrane surfaces. (b) Logarithmically plotted distribution of trajectory lengths in the experimental dataset, with the dashed line denoting the length at 100. (Inset) Representative instances of selected short trajectories of CTRW and FBM with different lengths, with labels assigned by WADNet. (c) The quantity distribution of short CTRW and FBM trajectories spanning various lengths from 10 to 100. For clarity, the FBM bars are measured from the top of the CTRW bars. (d) Segmentation performance of U-AnDi at different noise standard deviations σ_n. (e) mDice scores of U-AnDi on validation sets under diverse L and M parameters. [(f), (g)] Trajectories from the validation sets with parameters $L = 100, M = 3$ in (f) and $L = 200, M = 5$ in (g). (h) Comparisons between the predictions (blue) and ground truths (orange) for the trajectory in (f) (left) and the trajectory in (g) (right), underscoring a high-level consistency.

M, we can utilize these short trajectories to create the validation set, thereby facilitating the evaluation of U-AnDi's performance. In particular, we establish a rule specifying that the average length of segments (L/M) should not exceed 50. This constraint ensures the prevention of trajectory generation errors or the dominance of overly long segments.

On the other hand, in the task concerning unlabeled data, we employ the U-AnDi model to segment those long trajectories with lengths exceeding 100. To examine the performance of our model, the segments distinguished by U-AnDi are classified by WADNet to identify their corresponding diffusion models. Due to the absence of ground truth labels, we cannot use a conventional evaluation metric like the mDice to assess

the segmentation results. Instead, we rely on the confidence levels (probabilities) provided by WADNet to evaluate the effectiveness of our model.

B. Training scheme

The architecture of the U-AnDi model and training methods in this section remain almost identical to those in subtask 2. One deviation is a reduction in the filter number of the last 1D convolutional layer from 5 to 2. This modification is due to the task at hand, which requires semantic segmentations of just two diffusion models, CTRW and FBM. Another update is the initialization of model weights. To accelerate the

training process, we adopt a transfer learning strategy where the optimal model weights from subtask 2 are employed to initialize the model. This strategy allows the model to reach optimal performance within approximately 10 epochs, significantly reducing the training time.

In addition, taking the inevitable errors and noise during experimental observations into account, Gaussian white noise with a zero mean and a standard deviation of σ_n is added when generating simulated trajectories for training. In order to pinpoint the optimal σ_n and better mimic the circumstances of real experimental data, we select two validation sets comprising of experimental trajectories to evaluate the performance of U-AnDi at different σ_n. The parameters for these two validation sets are $L = 100$, $M = 3$ and $L = 200$, $M = 5$, respectively, with each set consisting of 1000 trajectories. As demonstrated in Fig. 12(d), for both validation sets, U-AnDi achieves the best performance (the highest mDice score) when $\sigma_n = 0.125$. Consequently, simulated trajectories with the standard deviation of noise $\sigma_n = 0.125$ are chosen to train the U-AnDi model in this section.

C. Performance of U-AnDi on two tasks regarding experimental trajectories

For the task involving labeled data, we investigate the performance of U-AnDi on validation sets with different parameters, as quantified by the mDice scores. The results are summarized in Fig. 12(e), where each validation set contains 1000 trajectories. It is observed that all mDice scores are above 0.8 for trajectory lengths ranging from 100 to 300, indicating an exceptional capability of U-AnDi in segmenting experimental trajectories. We also observe that segmentation performance decreases as trajectory length increases. This trend may be due to the fact that trajectories with larger L are composed of longer segments, which introduce more noise and thus can mislead the model's classification accuracy. Moreover, to illustrate the segmentation results intuitively, we present two representative samples from the validation sets, along with their corresponding predicted labels, in Figs. 12(f) and 12(g). Comparisons between the predictions and ground truths of these two samples are presented in the left and right panels of Fig. 12(h), respectively, with a high degree of consistency being observed as expected.

Next, in the task with unlabeled data, we utilize the trained U-AnDi model to segment natural trajectories in the experimental dataset with lengths exceeding 100. In the absence of ground truth labels, we leverage WADNet to obtain probabilities of diffusion models for the segments identified by U-AnDi. To provide a concrete illustration of the model's effectiveness, three representative trajectories and their corresponding results are displayed in Fig. 13. As distinguished by the distinct colors, trajectories in Figs. 13(a)–13(c) are segmented by U-AnDi into 2, 3, and 4 segments, respectively. The corresponding probabilities of diffusion models for these segments are presented in Figs. 13(d)–13(f). Observations suggest that, for the majority of segments, the diffusion models with the highest confidence levels (probabilities) remain CTRW and FBM. Nevertheless, for the second segment of the trajectory in Fig. 13(c), SBM is identified by WADNet as the most probable diffusion model. The result derived from
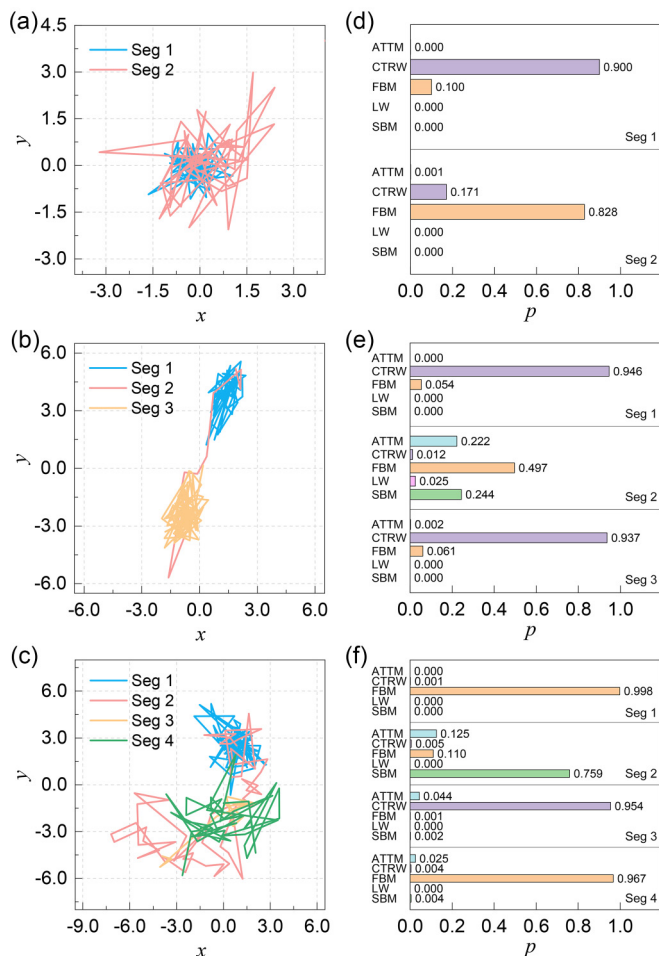

FIG. 13. [(a)–(c)] Representative natural trajectories from the experimental dataset with lengths surpassing 100. Segments identified by U-AnDi are distinguished by distinct colors. [(d)–(f)] Probabilities of diffusion models provided by WADNet for segments in (a)–(c), respectively.

our analysis on the task involving unlabeled data affirms the conclusion of N. Granik *et al.* that the diffusion of transmembrane proteins on cell membranes primarily exhibits CTRW and FBM characteristics [57]. However, it also hints that this conclusion may not be a comprehensive description of the entire diffusion process, which could potentially entail a higher degree of complexity.

VII. CONCLUSIONS

In summary, we have introduced U-AnDi, a deep learning model for the semantic segmentation of anomalous diffusion trajectories, which integrates the strengths of the dilated causal convolution (DCC), gated activation unit (GAU), and U-Net architecture. A variety of segmentation tasks have been addressed, demonstrating a superior performance of U-AnDi against other methods. Comparative evaluations underscore the interpretability of U-AnDi's core components and, to a certain extent, mitigate the "black box" nature of deep learning models, aligning them with the complex physics of anomalous diffusion. Further, U-AnDi also exhibits an

excellent capability in dealing with the real-world anomalous diffusion data, which is examined using the experimental dataset regarding the diffusion of transmembrane proteins on cell membrane surfaces. The supporting codes in this paper are accessible at our GitHub repository [80].

Despite the promising performance of our model, there still exist opportunities for further refinements. The first is optimizing the detection of changepoints among different diffusion states, where reducing FP and FN can be a critical task for future training and finetuning. Secondly, the implementation of U-AnDi is trained to recognize only five diffusion models in this study. Extending the model's ability to identify a larger variety of diffusion models would enhance the utility of U-AnDi in future research. Such an expansion would enable researchers to apply U-AnDi more effectively in the segmentation and property prediction of experimentally observed heterogeneous diffusion dynamics.

Additionally, considering that U-AnDi does not require specific prior knowledge about anomalous diffusion, this model has the potential to be trained and perform segmentation tasks on other time-series datasets instead of anomalous diffusion. In other words, our method not only provides a powerful and versatile tool for better understanding the dynamics of anomalous diffusion, but it also holds promise to be generalized as a universal scheme for the segmentation of time-series data.

APPENDIX A: IMPLEMENTATION DETAILS OF THE U-AnDi MODEL

The U-AnDi model is constructed by integrating the WAD-Net encoder into the U-Net architecture. The detailed structure of the WADNet encoder is depicted in Fig. 2(b). The dilation depth d within the encoder is 5, and the filter number f is consistent across all convolutional layers within a single encoder. The input tensor $\mathbf{x}$ initially passes through a standard causal convolution with a kernel size (ks) of 3. After that, the output $\mathbf{x}_1$ is processed by the subsequent layers, which consist of the gated activation unit (GAU) and 1D convolutional layer with $ks = 1$, as highlighted by the rectangular dashed box in Fig. 2(b). Let $\mathbf{x}_k$ represent the input of the kth layer where k ranges from 1 to d. The output of the kth layer $\mathbf{h}_k$ can be written as

$$\mathbf{h}_k = W^1 * [\tanh(W_{f,k} * \mathbf{x}_k) \otimes \sigma(W_{g,k} * \mathbf{x}_k)]. \quad \text{(A1)}$$

Here, $*$ denotes the convolution operation, $\tanh(\cdot)$ and $\sigma(\cdot)$ refer to the hyperbolic tangent function and sigmoid function, respectively. W^1 represents the 1D convolution with $ks = 1$. $W_{f,k}$ and $W_{g,k}$ are independent dilated causal 1D convolutions with $ks = 3$ and dilation $= 2^{k-1}$. $\mathbf{h}_k$ is directly used as the input of next ($k + 1$)th layer, meaning $\mathbf{x}_{k+1} = \mathbf{h}_k$. The output $\mathbf{z}$ of an encoder can be finally obtained by $\mathbf{z} = \mathbf{x}_1 + \sum_{k=1}^{d} \mathbf{h}_k$.

For convenience, we denote the operation of WADNet encoder as $\mathcal{W}(\cdot)$, i.e., $\mathbf{z} = \mathcal{W}(\mathbf{x})$, in the following discussion. In particular, the encoder in the down-sampling block of U-Net is denoted as $\mathcal{W}^{\mathrm{d}}$, while the one in the up-sampling block is represented as $\mathcal{W}^{\mathrm{u}}$.

By incorporating this encoder into the U-Net structure, we construct the U-AnDi model, with its architecture displayed in Fig. 2(c). The down-sampling block consists of a WADNet encoder, a batch normalization layer, and a max-pooling layer. We employ four consecutive down-sampling blocks to perform feature extraction on the input tensor. In the mth block, the input $\mathbf{v}_m$ is processed as governed by

$$\mathbf{u}_m = \mathrm{BN}\big[\mathcal{W}_m^{\mathrm{d}}(\mathbf{v}_m)\big], \quad \text{(A2)}$$

$$\mathbf{v}_{m+1} = \mathrm{MaxPool}(\mathbf{u}_m), \quad \text{(A3)}$$

where $m = 1, 2, 3, 4$. BN and MaxPool represent the batch normalization [83] and max-pooling operators, respectively. The filter number f_m in encoder $\mathcal{W}_m^{\mathrm{d}}$ is $64 \times 2^{m-1}$ for subtask 1 and $128 \times 2^{m-1}$ for subtask 2. The pooling size of max-pooling layer is set to 2.

Next, the output from the fourth down-sampling block $\mathbf{v}_5$ is processed through five consecutive up-sampling blocks to generate full-resolution feature maps, which are then used for pointwise predictions. The up-sampling block is comprised of a WADNet encoder, a batch normalization layer, an up-sampling layer using the nearest neighbor algorithm, and a 1D convolutional layer with $ks = 1$. In particular, skip connections are applied to get the input $\mathbf{p}_n$ of the nth up-sampling block when $n \geqslant 2$, as displayed by the dashed lines in Fig. 2(c). These connections serve to reduce the loss of fine-grained details during the up-sampling process and can be mathematically described as

$$\tilde{\mathbf{p}}_n = W_{n-1}^1 * \mathrm{UP}\big\{\mathrm{BN}\big[\mathcal{W}_{n-1}^{\mathrm{u}}(\mathbf{p}_{n-1})\big]\big\}, \quad \text{(A4)}$$

$$\mathbf{p}_n = \mathrm{Concat}(\tilde{\mathbf{p}}_n, \mathbf{u}_{6-n}), \quad \text{(A5)}$$

where $n = 2, 3, 4, 5$ and $\mathbf{p}_1 = \mathbf{v}_5$. UP denotes the up-sampling operator, with a scale factor set as 2. Concat is the concatenation operator, which is applied in the channel dimension. The filter number f_n in encoder $\mathcal{W}_n^{\mathrm{u}}$ is $64 \times 2^{5-n}$ for subtask 1 and $128 \times 2^{5-n}$ for subtask 2. The filter number of the 1D convolutional layer W_n^1 is half that of in encoder $\mathcal{W}_n^{\mathrm{u}}$. Note that when performing the skip connection in Eq. (A5), the shapes of $\tilde{\mathbf{p}}_n$ and $\mathbf{u}_{6-n}$ might not match. In such cases, we apply padding to $\tilde{\mathbf{p}}_n$ to make it compatible with $\mathbf{u}_{6-n}$ for concatenation. In the last up-sampling block, the up-sampling layer is removed, and the filter number of the convolutional layer is adjusted (1 for subtask 1 and 5 for subtask 2). This yields a final output $\mathbf{p}_f$ of the U-AnDi model, written as

$$\mathbf{p}_f = W_5^1 * \mathrm{BN}\big[\mathcal{W}_5^{\mathrm{u}}(\mathbf{p}_5)\big]. \quad \text{(A6)}$$

To illustrate the details of the U-AnDi architecture more clearly, we take a trajectory with $L = 500$ as the input and present the shapes of intermediate tensors for subtask 2 in Table III. The shape follows the format [length, channel].

The postprocessing techniques used to obtain the final predictions for both subtasks are summarized below. In subtask 1, which is a pointwise regression task, the output $\mathbf{p}_f$ from a trained U-AnDi model serves as the predicted diffusion

TABLE III. The shapes of intermediate tensors for subtask 2 when length of input trajectory is 500.

Tensor	Shape	Tensor	Shape
$\mathbf{v}_1$	[500, 2]	$\tilde{\mathbf{p}}_2$	[62, 1024]
$\mathbf{v}_2$	[250, 128]	$\mathbf{p}_2$	[62, 2048]
$\mathbf{v}_3$	[125, 256]	$\tilde{\mathbf{p}}_3$	[124, 512]
$\mathbf{v}_4$	[62, 512]	$\mathbf{p}_3$	[125, 1024]
$\mathbf{v}_5(\mathbf{p}_1)$	[31, 1024]	$\tilde{\mathbf{p}}_4$	[250, 256]
$\mathbf{u}_1$	[500, 128]	$\mathbf{p}_4$	[250, 512]
$\mathbf{u}_2$	[250, 256]	$\tilde{\mathbf{p}}_5$	[500, 128]
$\mathbf{u}_3$	[125, 512]	$\mathbf{p}_5$	[500, 256]
$\mathbf{u}_4$	[62, 1024]	$\mathbf{p}_f$	[500, 5]

exponents. To ensure a reasonable prediction with exponents ranging from 0.05 to 2, values outside the interval are clipped to the interval edges in $\mathbf{p}_f$. On the other hand, $\mathbf{p}_f$ is transformed using the following equations in subtask 2 to address this pointwise classification task,

$$\mathbf{X}_i^{\text{Prob}} = \text{Softmax}(\mathbf{p}_{f,i}), \tag{A7}$$

$$\mathbf{X}_i^{\text{Pred}} = \arg\max\left(\mathbf{X}_i^{\text{Prob}}\right). \tag{A8}$$

Here, i denotes the ith trajectory. Softmax represents the softmax activation function. $\mathbf{X}_i^{\text{Prob}}$ and $\mathbf{X}_i^{\text{Pred}}$ are the predicted probability and label, respectively.

APPENDIX B: DATA PREPROCESSING AND TRAINING SCHEME OF THE U-AnDi MODEL

To train our deep learning model, 1 000 000 trajectories are generated as the dataset for a fixed length parameter L. This dataset is randomly split into two parts: 80% for the training set and 20% for the validation set. The raw trajectory data is independently normalized along the x and y dimensions, achieving a mean of 0 and a standard deviation of 1 for each dimension. Then, the normalized data from both dimensions, $[x(1), x(2), x(3), \ldots, x(L)]$ and $[y(1), y(2), y(3), \ldots, y(L)]$, are concatenated to form a sequence:

$$\begin{cases} [x(1), x(2), x(3), \ldots, x(L)] \\ [y(1), y(2), y(3), \ldots, y(L)] \end{cases}, \tag{B1}$$

which is utilized as the input of the U-AnDi model.

The selected loss functions are mean squared error (MSE) for subtask 1 and cross entropy (CE) for subtask 2. The model is trained using the backpropagation algorithm with a batch size of 512 on a single NVIDIA A100 GPU. The learning rate is 0.0002, and the optimizer is Adam. In particular, an L2 penalty is applied for subtask 2 with a weight decay of 2×10^{-5}. The training process involves 100 epochs for subtask 1 and 200 epochs for subtask 2. The model performance on the validation set is evaluated every 1000 steps.

We use the trajectory data with $L = 500$ as the example to illustrate the evolution of loss function during training, as shown in Fig. 14(a) for subtask 1 and Fig. 14(b) for subtask 2. A rapid convergence of the train loss can be observed for both tasks, with the valid loss exhibiting a similar trend. This indicates that our model is fully capable of accomplishing
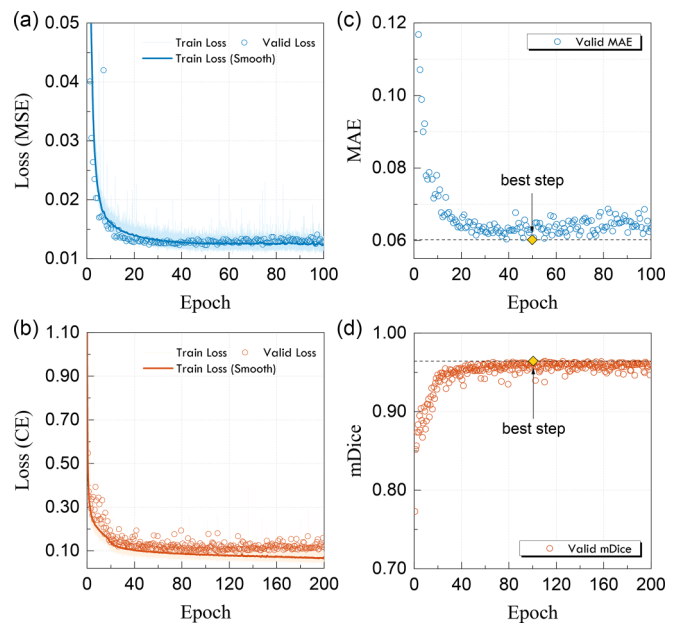

FIG. 14. [(a), (b)] Evolution of the loss function during training for subtask 1 (a) and subtask 2 (b). The thick line represents a smoothed version of the original train loss to provide better visual guidance. [(c), (d)] Epoch dependence of the valid metric for subtask 1 (c) and 2 (d). The dashed line denotes the value of the best metric, and the yellow dot highlights the location of the best step.

both tasks without encountering over-fitting issues. Furthermore, the optimal weight of the U-AnDi model is determined based on the best metric achieved on the validation set. The epoch dependencies of valid metrics for subtask 1 and 2 are presented in Figs. 14(c) and 14(d), respectively. As indicated by the yellow dot, the selected weight corresponds to the step where the best metric is reached.

APPENDIX C: DETAILS OF COMPARATIVE MODELS

In this Appendix, we provide a comprehensive description of the models utilized for comparative analysis in our study. These models are specifically chosen and designed to enable a robust comparative evaluation against our proposed U-AnDi model regarding the semantic segmentation of anomalous diffusion trajectories.

(1) Long short-term memory (LSTM) [66]: The LSTM model we use in this paper is structured as a three-layered stack, where the dimensionality of hidden state is set as 64.

(2) Gated recurrent unit (GRU) [67]: In this paper, the GRU model is also architecturally configured as a three-layered stack, wherein each layer possesses a hidden state with a dimensionality of 64.

(3) Transformer (TFM) [68]: Analogous to the LSTM and GRU, the TFM model is structured with three layers of Transformer encoders, where each encoder employs an embedding dimensionality of 128.

(4) WADNet encoder (WE): The WE model is derived from the encoder part of our previously proposed WADNet structure [34]. The model's hyperparameters align with the WADNet encoder in the first down-sampling block of U-AnDi.

(5) U-CNN: The U-CNN model represents a modified architecture developed from U-AnDi, wherein WADNet encoders are substituted with simpler dilated causal 1D CNN layers. Specifically, the input and output channels, kernel size, and dilation rate of a CNN layer are maintained consistent with the corresponding WADNet encoder.

In addition, we provide a concise overview of the divide-and-conquer moment scaling spectrum (DC-MSS) here as described in Ref. [38]. DC-MSS is a technique designed for analyzing single-particle tracking data. It works by dividing trajectories into shorter segments and computing their moment scaling spectrum. Through detecting shifts in the scaling exponents of these segments, DC-MSS pinpoints changes in diffusion behavior. Notably, it excels at distinguishing transitions among different modes of motion, including Brownian diffusion, confined diffusion, directed motion, and the immobile state. In this paper, we employ the DC-MSS technique with the default hyper-parameters as outlined in Ref. [38].

APPENDIX D: POSTPROCESSING TECHNIQUE IN SUBTASK 1

Given a continuous pointwise prediction of diffusion exponents, denoted as

$$\alpha_1, \alpha_2, \alpha_3, \ldots, \alpha_L, \tag{D1}$$

for a trajectory of length L, we first calculate the differences between adjacent elements and get the sequence

$$\beta_1, \beta_2, \beta_3, \ldots, \beta_{L-1}, \tag{D2}$$

where $\beta_i = \alpha_{i+1} - \alpha_i$. Next, we set a threshold value ϵ and identify all indices i satisfying $\beta_i > \epsilon$. These indices form a new sequence

$$i_1, i_2, i_3, \ldots, i_k. \tag{D3}$$

If two consecutive indices i_m and i_{m+1} in this sequence meet $i_{m+1} - i_m \geqslant 10$, we denote $i_m + 1$ as the starting point of the current segment, and i_{m+1} as the endpoint of this segment. Then, the changepoint between two successive predicted segments $[i_a + 1, i_{a+1}]$ and $[i_b + 1, i_{b+1}]$ can be identified in the interval $[i_{a+1} + 1, i_b]$. We select the point within this interval that exhibits the maximum absolute value of the derivative as the changepoint. Applying this rule, we can obtain the sequence of changepoints, given by

$$c_1, c_2, c_3, \ldots, c_n. \tag{D4}$$

Here, $c_2, c_3, \ldots, c_{n-1}$ are the predicted changepoints, and c_1 and c_n are set as 1 and L, respectively. Based on the detection result of changepoints, the predicted diffusion exponent $\tilde{\alpha}_i$ of the ith segment can be given as

$$\tilde{\alpha}_i = \frac{1}{c_{i+1} - c_i} \sum_{j=c_i}^{c_{i+1}} \alpha_j, \tag{D5}$$
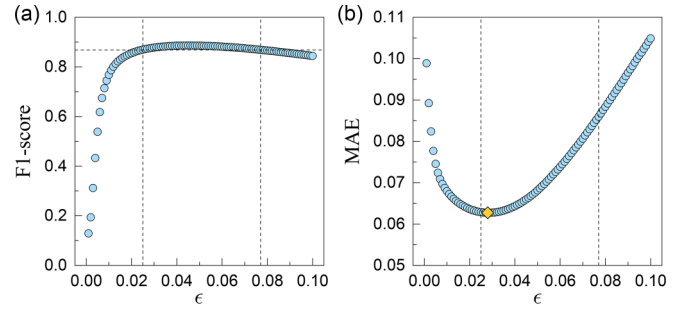
where $i = 1, 2, \ldots, n - 1$.

FIG. 15. (a) F1-score for changepoint detection as a function of ϵ. The horizontal dashed line signifies 98% of the maximum F1-score, and the vertical dashed lines indicate the candidate interval for ϵ, within which the F1-score retains at least 98% of its peak value. (b) MAE as a function of ϵ. The optimal ϵ is selected to minimize the MAE within the candidate interval (vertical dashed lines). The yellow dot highlights the chosen optimal ϵ value of 0.028 for this paper.

In particular, the method for determining the threshold value ϵ is summarized here. A larger ϵ tends to yield higher precision in detecting changepoints, albeit with a lower recall, while a smaller ϵ may compromise the precision but enhance recall. Therefore, as ϵ increases, the F1-score for changepoint detection initially rises and subsequently diminishes, as depicted in Fig. 15(a). Concurrently, alterations in ϵ also induce variation in the model's MAE, as illustrated in Fig. 15(b). Under these circumstances, the selection of ϵ emerges as a result of balancing a trade-off between MAE and F1-score on the validation set. In more detail, we select the optimal ϵ from a candidate interval where the F1-score is at least 98% of its maximum value, as indicated by the vertical dashed lines in Fig. 15(a). Within this interval, we seek the ϵ that minimizes the MAE value. In this work, 0.028 is identified as the optimal ϵ, as highlighted by the yellow dot in Fig. 15(b).

APPENDIX E: ADDRESSING CONCURRENT EXPONENT AND MODEL VARIATIONS USING MULTITASK LEARNING

The versatility of U-AnDi has been previously highlighted for its effectiveness at segmenting trajectories based on variations in either the diffusion exponent or the dynamic model. However, real-world scenarios often present more intricate challenges where both exponent and model might concurrently vary within the same trajectory. To address this issue, we have further enhanced U-AnDi's capabilities through the integration of a multitask learning strategy in this section.

The method for simulating trajectories with concurrent variations in both exponent and model is an extension of the approach used in subtask 2. We combine segments from different diffusion models to generate a composite trajectory, as illustrated in Fig. 16(a). A notable distinction in this enhanced simulation is the allowance for variations of the diffusion exponent within a segment characterized by a single dynamic model, as depicted in Fig. 16(b). To enrich the diversity of simulated trajectories, we expand the number of segments, M, from the initial range of 2 to 5 to an extended range of 3 to 8.
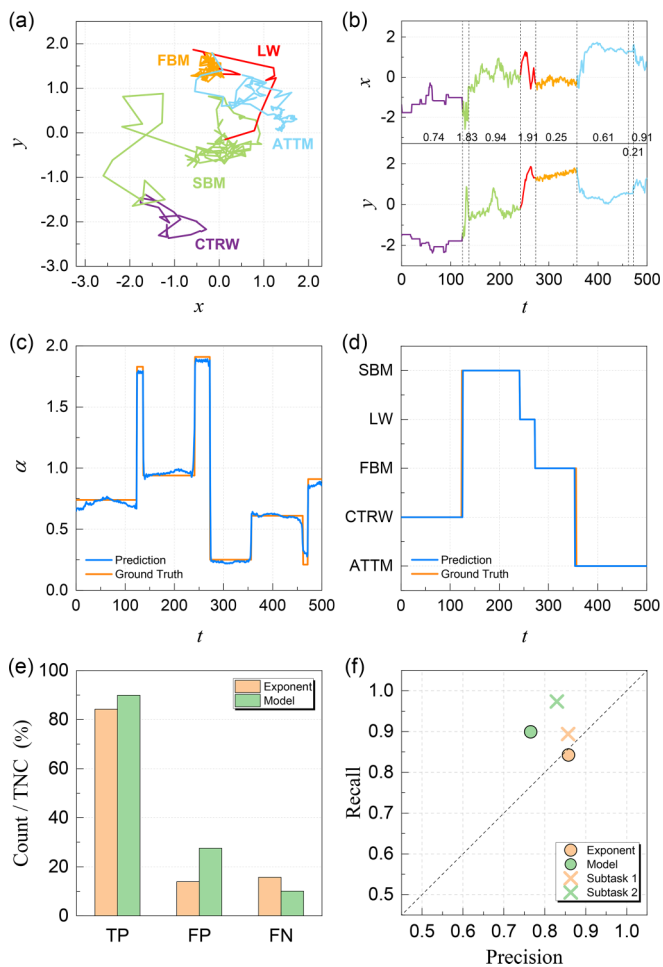
FIG. 16. (a) Illustration of a composite trajectory of length 500 with different diffusion models. (b) Time evolution of trajectories in (a) for both x and y dimensions, showing variations of the diffusion exponent within a single model segment. [(c)–(d)] Segmentation results for the trajectory in (a) based on the diffusion exponent (c) and the diffusion model (d). (e) Distribution of TP, FP, and FN as a percentage of TNC for both segmentation tasks. (f) Comparative analysis regarding precision and recall of changepoint detection results with those from subtasks 1 and 2.

Subsequently, we employ a multitask learning approach to enable U-AnDi to effectively segment these trajectories. In detail, we utilize the intermediate output of U-AnDi, just prior to the final 1D CNN layer, which is denoted as $\mathbf{p}_o$ and expressed as $\mathbf{p}_o = \mathrm{BN}[\mathcal{W}_5^u(\mathbf{p}_5)]$. This intermediate output is directed into two separate 1D CNN head layers, $W_{f_1}^1$ and $W_{f_2}^1$. The layer $W_{f_1}^1$, with an output $\mathbf{p}_{f_1} = W_{f_1}^1 * \mathbf{p}_o$, is dedicated to a pointwise regression task, aiming to segment based on the diffusion exponent. In contrast, the other layer $W_{f_2}^1$ focuses on a pointwise classification task, targeting the segmentation of the diffusion model and leading to an output $\mathbf{p}_{f_2} = W_{f_2}^1 * \mathbf{p}_o$. After that, we compute the loss functions $\mathcal{L}_1$ and $\mathcal{L}_2$ for $\mathbf{p}_{f_1}$ and $\mathbf{p}_{f_2}$ using MSE and CE respectively. The final loss function $\mathcal{L}$ for this task is given as $\mathcal{L} = \mathcal{L}_1 + \mathcal{L}_2$. The data size and training method are consistent with those in subtask 2.

In line with the presentation style of subtask 1 and 2, we demonstrate U-AnDi's segmentation performance on trajectories of length 500 here. To visualize this intuitively, segmentation results for the diffusion exponent and dynamic model of the trajectory in Fig. 16(a) are illustrated in Figs. 16(c) and 16(d), respectively. As expected, both predictions display a high degree of alignment with the ground truth. The results on the validation set yield an MAE of 0.1376 and an mDice of 0.9474, indicating that U-AnDi can effectively segment trajectories where the exponent and model concurrently vary.

On the other hand, when it comes to the detection of changepoints, Fig. 16(e) reveals that TP consistently ranks highest for both segmentation tasks, whether based on the exponent or the model. F1-scores achieved for these two detections are 0.8501 and 0.8269, respectively, underscoring a robust capability of U-AnDi in identifying changepoints. Furthermore, as depicted in Fig. 16(f), we compare the detection results with those from subtasks 1 and 2. For both exponent and model changepoint detections, there is a decline in recall, with the latter also experiencing a noticeable decrease in precision. The common reduction in recall can be attributed to the increased complexity of the task. The diminished precision in detecting model changepoints might stem from variations of diffusion exponent within a single model segment. These variations potentially result in occasional misinterpretations and consequently lead the model to detect more transition points than actual.

APPENDIX F: SIMULATION OF MULTISTATE TRAJECTORIES IN BIOLOGICAL EXPERIMENT SCENARIOS

In this section, we introduce the simulation techniques in detail for generating trajectories that represent different states commonly observed in biological experiments. These techniques are primarily adapted from the procedures described in Ref. [38]. While the simulation parameters for each individual trajectory remain constant, they differ across the various trajectories. The chosen parameters align closely with those in Ref. [38] to ensure an equitable evaluation of the DC-MSS method's segmentation performance.

(1) Brownian diffusion (BD): Simulating the free Brownian diffusion requires a series of x and y displacements. Each displacement is drawn from a Gaussian distribution $\mathcal{N}(0, \sqrt{2D\delta t})$, where D is the diffusion coefficient and δt is the time interval. We set $\delta t = 1$ and select D uniformly from the range 1.0 to 3.0.

(2) Confined diffusion (CD): The generation of confined diffusion trajectories is akin to that of free Brownian diffusion trajectories. However, a reflection occurs once the subsequent coordinate of the random walker exceeds the boundary of the confined region. This confined region is defined as a circle with a radius R. Here, the radius R ranges from 3.0 to 5.0, while the diffusion coefficient D varies between 1.0 and 3.0.

(3) Directed motion (DM): Each displacement in a trajectory of the directed motion consists of two components: a Brownian diffusive component with a diffusion coefficient $D \in [1.0, 3.0]$ and a directed component. The directed component is defined by $v\cos\theta \cdot \delta t$ for x displacement and $v\sin\theta \cdot$
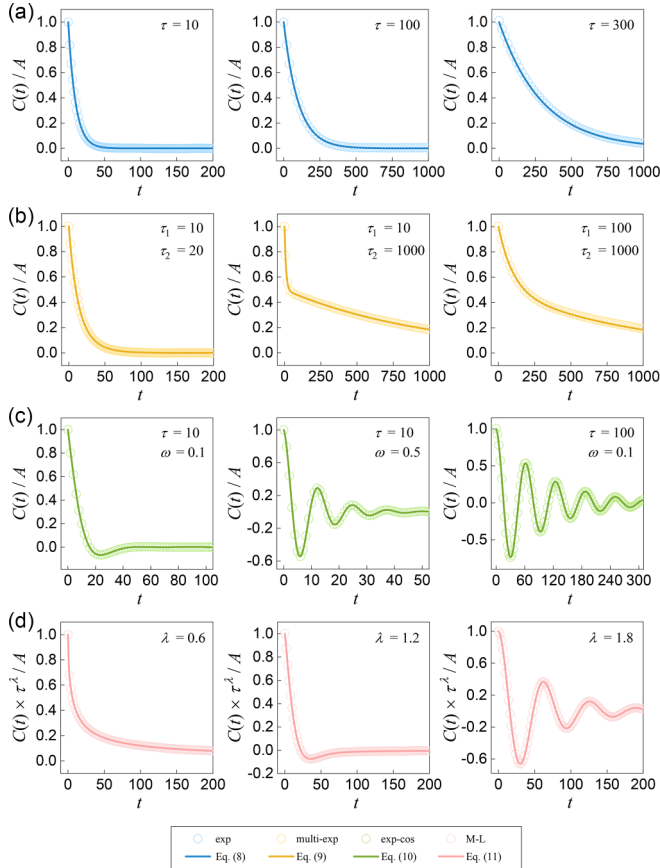
FIG. 17. Comparisons of the velocity autocorrelation functions $C(t)$ of simulated trajectories (dots) with their theoretical values (lines) from Eqs. (8)–(11) under various parameters for the four types of correlations: (a) exponential decay (exp); (b) multi-exponential decay (multi-exp) where $\phi = 0.5$; (c) exponentially damped cosine wave (exp-cos); (d) Mittag-Leffler decay (M-L) where $\tau = 10$. Here, A in the y axis label refers to the amplitude used in Eqs. (8)–(11).

δt for y displacement. The drift velocity v varies between 2.0 and 3.0, while the drift angle θ spans from 0 to 2π.

(4) Immobile state (IS): The simulation of immobile state is achieved by generating x and y coordinates from a Gaussian distribution $\mathcal{N}(0, \sigma_p)$. Here, σ_p represents the positional localization error, with values ranging from 0.8 to 1.6.

Leveraging these simulation techniques, we employ the same approach as in subtask 2 to combine trajectory segments of four diffusion states, ultimately producing the multistate trajectories in biological experiment scenarios. The codes for trajectory generation can be found in our GitHub repository [80].
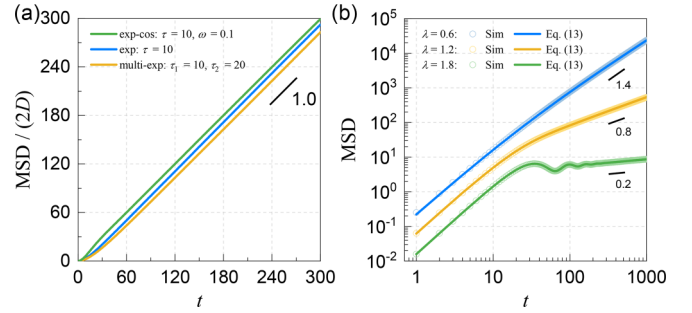

FIG. 18. (a) Relationship between MSD/(2D) and time t for simulated trajectories associated with exp, multi-exp, and exp-cos correlations. (b) Log-log plot of MSDs of simulated trajectories governed by a M-L correlation for specific λ values of 0.6, 1.2, and 1.8 (represented by dots). The corresponding theoretical predictions, derived from Eq. (13), are also depicted as lines.

APPENDIX G: VALIDATING THE METHOD FOR GENERATING TRAJECTORIES WITH DEFINED CORRELATION FUNCTIONS

Within this Appendix, we will validate the effectiveness of our trajectory generation method based on the Davies-Harte algorithm. For that purpose, we initially compare the velocity autocorrelation functions [VACFs, $C(t)$] of trajectories generated using this method under various parameters with their theoretical values. The results of this comparison are summarized in Fig. 17. As observed, for the four types of correlations, i.e., exp, multi-exp, exp-cos, and M-L, VACFs of simulated trajectories (represented by dots) align closely with their respective theoretical values (depicted by lines). This consistency strongly attests to the validity of our trajectory generation method.

On the other hand, from the perspective of MSD, we further validate the effectiveness of this method. Using the equation MSD $= 2 \int_0^t (t - s)C(s)ds$, we derive the theoretical expressions for exp, multi-exp, and exp-cos correlations, as detailed in Table IV. From the long-time behavior of MSD, we can discern that the trajectories associated with exp, multi-exp, and exp-cos correlations exhibit long-time Fickian diffusion theoretically. This diffusion type is typically represented by MSD $\approx 2Dt$, where D denotes the long-time diffusion coefficient. To substantiate this, we present the relationship between MSD/(2D) and time t for simulated trajectories of these three types of correlations in Fig. 18(a). The slope of 1.0 across these plots confirms that MSDs for these trajectories are consistent with their theoretical descriptions. For the M-L correlation, the theoretical MSD value is determined via numerical integration. In Fig. 18(b), we present the MSDs of simulated trajectories for specific λ values of 0.6, 1.2, and 1.8 in a log-log form. These are juxtaposed with their

TABLE IV. Theoretical expressions of MSD and long-time MSD corresponding to the exp, multi-exp, and exp-cos correlations.

Correlation type	Theoretical expression of MSD	Long-time MSD
exp	$2A\tau[t + \tau(e^{-t/\tau} - 1)]$	$2A\tau t$
multi-exp	$2A[\phi\tau_1 t + \phi\tau_1^2(e^{-t/\tau_1} - 1) + (1-\phi)\tau_2 t + (1-\phi)\tau_2^2(e^{-t/\tau_2} - 1)]$	$2A[\phi\tau_1 + (1-\phi)\tau_2]t$
exp-cos	$\frac{2A\tau}{(1+\tau^2\omega^2)^2}[(1 + \tau^2\omega^2)t - \tau(1 - \tau^2\omega^2) + \tau e^{-t/\tau}[(1 - \tau^2\omega^2)\cos\omega t - 2\tau\omega\sin\omega t]]$	$\frac{2A\tau}{1+\tau^2\omega^2}t$

theoretical predictions, underscoring a pronounced congruence between them. Furthermore, as indicated by the slope, the long-time diffusion exponent α derived from our measurements relates to λ following the relationship $\alpha = 2 - \lambda$. This observation agrees well with the descriptions provided in Ref. [82].

In summary, the alignment of the VACF and MSD with their theoretical values underscores the effectiveness of our trajectory generation method. This method offers a solid framework for studies focusing on segmenting anomalous diffusion trajectories that exhibit defined long-time correlations.

[1] R. Metzler, J.-H. Jeon, A. G. Cherstvya, and E. Barkaid, Anomalous diffusion models and their properties: Non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking, Phys. Chem. Chem. Phys. **16**, 24128 (2014).

[2] J. Klafter and I. M. Sokolov, Anomalous diffusion spreads its wings, Phys. World **18**, 29 (2005).

[3] T. G. Mason and D. A. Weitz, Optical measurements of frequency-dependent linear viscoelastic moduli of complex fluids, Phys. Rev. Lett. **74**, 1250 (1995).

[4] F. D. A. Aarão Reis, Scaling relations in the diffusive infiltration in fractals, Phys. Rev. E **94**, 052124 (2016).

[5] G. Volpe and G. Volpe, The topography of the environment alters the optimal search strategy for active particles, Proc. Natl. Acad. Sci. USA **114**, 11350 (2017).

[6] Y. Sagi, M. Brook, I. Almog, and N. Davidson, Observation of anomalous diffusion and fractional self-similarity in one dimension, Phys. Rev. Lett. **108**, 093002 (2012).

[7] D. Hartich and A. Godec, Thermodynamic uncertainty relation bounds the extent of anomalous diffusion, Phys. Rev. Lett. **127**, 080601 (2021).

[8] E. Barkai, Y. Garini, and R. Metzler, Strange kinetics of single molecules in living cells, Phys. Today **65**, 29 (2012).

[9] Y. Ding, A. A. Hassanali, and M. Parrinello, Anomalous water diffusion in salt solutions, Proc. Natl. Acad. Sci. USA **111**, 3310 (2014).

[10] J.-J. Song, R. Bhattacharya, H. Kim, J. Chang, T.-Y. Tang, H. Guo, S. K. Ghosh, Y. Yang, Z. Jiang, H. Kim, T. P. Russell, G. Arya, S. Narayanan, and S. K. Sinha, One-dimensional anomalous diffusion of gold nanoparticles in a polymer melt, Phys. Rev. Lett. **122**, 107802 (2019).

[11] Z. Xu, X. Dai, X. Bu, Y. Yang, X. Zhang, X. Man, X. Zhang, M. Doi, and L.-T. Yan, Enhanced heterogeneous diffusion of nanoparticles in semiflexible networks, ACS Nano **15**, 4608 (2021).

[12] F. Höfling and T. Franosch, Anomalous transport in the crowded world of biological cells, Rep. Prog. Phys. **76**, 046602 (2013).

[13] X.-L. Wu and A. Libchaber, Particle diffusion in a quasi-two-dimensional bacterial bath, Phys. Rev. Lett. **84**, 3017 (2000).

[14] M. C. González, C. A. Hidalgo, and A.-L. Barabási, Understanding individual human mobility patterns, Nature (London) **453**, 779 (2008).

[15] B. Wang, J. Kuo, and S. Granick, Bursts of active transport in living cells, Phys. Rev. Lett. **111**, 208102 (2013).

[16] P. Chen, Z. Huang, J. Liang, T. Cui, X. Zhang, B. Miao, and L.-T. Yan, Diffusion and directionality of charged nanoparticles on lipid bilayer membrane, ACS Nano **10**, 11541 (2016).

[17] V. Plerou, P. Gopikrishnan, L. A. Nunes Amaral, X. Gabaix, and H. E. Stanley, Economic fluctuations and anomalous diffusion, Phys. Rev. E **62**, R3023(R) (2000).

[18] J. Masoliver, M. Montero, and G. H. Weiss, Continuous-time random-walk model for financial distributions, Phys. Rev. E **67**, 021112 (2003).

[19] Z.-Q. Jiang, W.-J. Xie, W.-X. Zhou, and D. Sornette, Multifractal analysis of financial markets: A review, Rep. Prog. Phys. **82**, 125901 (2019).

[20] V. Sposini, D. Krapf, E. Marinari, R. Sunyer, F. Ritort, F. Taheri, C. Selhuber-Unkel, R. Benelli, M. Weiss, R. Metzler *et al.*, Towards a robust criterion of anomalous diffusion, Commun. Phys. **5**, 305 (2022).

[21] O. Vilk, E. Aghion, T. Avgar, C. Beta, O. Nagel, A. Sabri, R. Sarfati, D. K. Schwartz, M. Weiss, D. Krapf, R. Nathan, R. Metzler, and M. Assaf, Unravelling the origins of anomalous diffusion: From molecules to migrating storks, Phys. Rev. Res. **4**, 033055 (2022).

[22] W. Wang, A. G. Cherstvy, R. Metzler, and I. M. Sokolov, Restoring ergodicity of stochastically reset anomalous-diffusion processes, Phys. Rev. Res. **4**, 013161 (2022).

[23] H. Seckler and R. Metzler, Bayesian deep learning for error estimation in the analysis of anomalous diffusion, Nat. Commun. **13**, 6717 (2022).

[24] B. Requena, S. Masó, J. Bertran, M. Lewenstein, C. Manzo, and G. Muñoz-Gil, Inferring pointwise diffusion properties of single trajectories with deep learning, Biophys. J. **122**, 4360 (2023).

[25] G. Muñoz-Gil, G. Volpe, M. A. Garcia-March, E. Aghion, A. Argun, C. B. Hong, T. Bland, S. Bo, J. A. Conejero, N. Firbas *et al.*, Objective comparison of methods to decode anomalous diffusion, Nat. Commun. **12**, 6253 (2021).

[26] C. Manzo, G. Muñoz-Gil, G. Volpe, M. A. Garcia-March, M. Lewenstein, and R. Metzler, Preface: Characterisation of physical processes from anomalous diffusion data, J. Phys. A: Math. Theor. **56**, 010401 (2023).

[27] S. Bo, F. Schmidt, R. Eichhorn, and G. Volpe, Measurement of anomalous diffusion using recurrent neural networks, Phys. Rev. E **100**, 010102(R) (2019).

[28] P. Kowalek, H. Loch-Olszewska, and J. Szwabiński, Classification of diffusion modes in single-particle tracking data: Feature-based versus deep-learning approach, Phys. Rev. E **100**, 032410 (2019).

[29] N. Firbas, Òscar Garibo-i Orts, M. Ángel Garcia-March, and J. A. Conejero, Characterization of anomalous diffusion through convolutional transformers, J. Phys. A: Math. Theor. **56**, 014001 (2023).

[30] G. Muñoz-Gil, G. G. i Corominas, and M. Lewenstein, Unsupervised learning of anomalous diffusion data: An anomaly detection approach, J. Phys. A: Math. Theor. **54**, 504001 (2021).

[31] A. Gentili and G. Volpe, Characterization of anomalous diffusion classical statistics powered by deep learning (condor), J. Phys. A: Math. Theor. **54**, 314003 (2021).

[32] A. Argun, G. Volpe, and S. Bo, Classification, inference and segmentation of anomalous diffusion with recurrent neural networks, J. Phys. A: Math. Theor. **54**, 294003 (2021).

[33] H. Verdier, M. Duval, F. Laurent, A. Cassé, C. L. Vestergaard, and J.-B. Masson, Learning physical properties of anomalous random walks using graph neural networks, J. Phys. A: Math. Theor. **54**, 234001 (2021).

[34] D. Li, Q. Yao, and Z. Huang, WaveNet-based deep neural networks for the characterization of anomalous diffusion (WADNet), J. Phys. A: Math. Theor. **54**, 404003 (2021).

[35] H. Verdier, F. Laurent, A. Cassé, C. L. Vestergaard, and J.-B. Masson, Variational inference of fractional Brownian motion with linear computational complexity, Phys. Rev. E **106**, 055311 (2022).

[36] H. Verdier, F. Laurent, A. Cassé, C. L. Vestergaard, C. G. Specht, and J.-B. Masson, Simulation-based inference for non-parametric statistical comparison of biomolecule dynamics, PLoS Comput. Biol. **19**, e1010088 (2023).

[37] H. Kabbech and I. Smal, Identification of diffusive states in tracking applications using unsupervised deep learning methods, in *2022 IEEE 19th International Symposium on Biomedical Imaging (ISBI)* (Kolkata, India, 2022), pp. 1–4.

[38] A. R. Vega, S. A. Freeman, S. Grinstein, and K. Jaqaman, Multistep track segmentation and motion classification for transient mobility analysis, Biophys. J. **114**, 1018 (2018).

[39] M. Arts, I. Smal, M. W. Paul, C. Wyman, and E. Meijering, Particle mobility analysis using deep learning and the moment scaling spectrum, Sci. Rep. **9**, 17160 (2019).

[40] H. Seckler, J. Szwabinński, and R. Metzler, Machine-learning solutions for the analysis of single-particle diffusion trajectories, Phys. Chem. Chem. Phys. **14**, 7910 (2023).

[41] E. Meijering, I. Smal, O. Dzyubachyk, and J.-C. Olivo-Marin, in *Microscope Image Processing*, 2nd ed., edited by F. A. Merchant and K. R. Castleman (Academic Press, New York, 2023), Chap. 14, pp. 393–430.

[42] J. Pineda, B. Midtvedt, H. Bachimanchi, S. Noé, D. Midtvedt, G. Volpe, and C. Manzo, Geometric deep learning reveals the spatiotemporal features of microscopic motion, Nat. Mach. Intell. **5**, 71 (2023).

[43] Y. Zhang, F. Ge, X. Lin, J. Xue, Y. Song, H. Xie, and Y. He, Extract latent features of single-particle trajectories with historical experience learning, Biophys. J. **122**, 4451 (2023).

[44] K. Chen, B. Wang, and S. Granick, Memoryless self-reinforcing directionality in endosomal active transport within living cells, Nat. Mater. **14**, 589 (2015).

[45] K. Chen, B. Wang, J. Guan, and S. Granick, Diagnosing heterogeneous dynamics in single-molecule/particle trajectories with multiscale wavelets, ACS Nano **7**, 8634 (2013).

[46] F. Persson, M. Lindén, C. Unoson, and J. Elf, Extracting intracellular diffusive states and transition rates from single-molecule tracking data, Nat. Methods **10**, 265 (2013).

[47] N. Monnier, Z. Barry, H. Y. Park, K.-C. Su, Z. Katz, B. P. English, A. Dey, K. Pan, I. M. Cheeseman, R. H. Singer *et al.*, Inferring transient particle transport dynamics in live cells, Nat. Methods **12**, 838 (2015).

[48] Y. Mo, Y. Wu, X. Yang, F. Liu, and Y. Liao, Review the state-of-the-art technologies of semantic segmentation based on deep learning, Neurocomputing **493**, 626 (2022).

[49] F. Yu and V. Koltun, Multi-scale context aggregation by dilated convolutions, in *Proceedings of the 4th International Conference on Learning Representations, ICLR 2016* (San Juan, Puerto Rico, 2016).

[50] K. He, G. Gkioxari, P. Dollor, and R. Girshick, Mask R-CNN, in *Proceedings of the IEEE International Conference on Computer Vision (ICCV)* (IEEE Computer Society, Venice, 2017), pp. 2961–2969.

[51] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, Feature pyramid networks for object detection, in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (IEEE Computer Society, Honolulu, Hawaii, 2017), pp. 2117–2125.

[52] V. Badrinarayanan, A. Kendall, and R. Cipolla, SegNet: A deep convolutional encoder-decoder architecture for image segmentation, IEEE Trans. Pattern Anal. Mach. Intell. **39**, 2481 (2017).

[53] L.-J. Li, R. Socher, and F.-F. Li, Towards total scene understanding: Classification, annotation and segmentation in an automatic framework, in *2009 IEEE Conference on Computer Vision and Pattern Recognition* (IEEE Computer Society, Miami, Florida, 2009), pp. 2036–2043.

[54] O. Ronneberger, P. Fischer, and T. Brox, U-Net: Convolutional networks for biomedical image segmentation, in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* (Springer International Publishing, Cham, 2015), pp. 234–241.

[55] M. Antonelli, A. Reinke, S. Bakas, K. Farahani, A. Kopp-Schneider, B. A. Landman, G. Litjens, B. Menze, O. Ronneberger, R. M. Summers *et al.*, The medical segmentation decathlon, Nat. Commun. **13**, 4128 (2022).

[56] G. Du, X. Cao, J. Liang, X. Chen, and Y. Zhan, Medical image segmentation based on U-Net: A review, J. Imaging Sci. Technol. **64**, 020508-1 (2020).

[57] N. Granik, L. E. Weiss, E. Nehme, M. Levin, M. Chein, E. Perlson, Y. Roichman, and Y. Shechtman, Single-particle diffusion characterization by deep learning, Biophys. J. **117**, 185 (2019).

[58] G. Muñoz-Gil, B. Requena, G. Volpe, M. A. Garcia-March, and C. Manzo, Andichallenge/andi_datasets: Challenge 2020 release (v.1.0). zenodo. (2021).

[59] B. B. Mandelbrot and J. W. V. Ness, Fractional Brownian motions, fractional noises and applications, SIAM Rev. **10**, 422 (1968).

[60] M. Chein, E. Perlson, and Y. Roichman, Flow arrest in the plasma membrane, Biophys. J. **117**, 810 (2019).

[61] H. Scher and E. W. Montroll, Anomalous transit-time dispersion in amorphous solids, Phys. Rev. B **12**, 2455 (1975).

[62] P. Massignan, C. Manzo, J. A. Torreno-Pina, M. F. García-Parajo, M. Lewenstein, and G. J. Lapeyre, Nonergodic subdiffusion from Brownian motion in an inhomogeneous medium, Phys. Rev. Lett. **112**, 150603 (2014).

[63] J. Klafter and G. Zumofen, Lévy statistics in a Hamiltonian system, Phys. Rev. E **49**, 4873 (1994).

[64] S. C. Lim and S. V. Muniandy, Self-similar Gaussian processes for modeling anomalous diffusion, Phys. Rev. E **66**, 021114 (2002).

[65] L. R. Dice, Measures of the amount of ecologic association between species, Ecology **26**, 297 (1945).

[66] S. Hochreiter and J. Schmidhuber, Long short-term memory, Neural Comput. **9**, 1735 (1997).

[67] K. Cho, B. van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio, Learning phrase representations using RNN encoder–decoder for statistical machine translation, in *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)* (Association for Computational Linguistics, Doha, Qatar, 2014), pp. 1724–1734.

[68] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, Attention is all you need, in *Proceedings of the 31st International Conference on Neural Information Processing Systems*, Vol. 30 (Curran Associates Inc., Red Hook, NY, 2017), pp. 6000–6010.

[69] A. van den Oord, S. Dieleman, H. Zen, K. Simonyan, O. Vinyals, A. Graves, N. Kalchbrenner, A. Senior, and K. Kavukcuoglu, WaveNet: A generative model for raw audio, arXiv:1609.03499.

[70] C. Manzo and M. F. Garcia-Parajo, A review of progress in single particle tracking: From methods to biophysical insights, Rep. Prog. Phys. **78**, 124601 (2015).

[71] H. Shen, L. J. Tauzin, R. Baiyasi, W. Wang, N. Moringo, B. Shuang, and C. F. Landes, Single particle tracking: From theory to biophysical applications, Chem. Rev. **117**, 7331 (2017).

[72] H. Qian, M. Sheetz, and E. Elson, Single particle tracking. Analysis of diffusion and flow in two-dimensional systems, Biophys. J. **60**, 910 (1991).

[73] M. J. Saxton, Single-particle tracking: Connecting the dots, Nat. Methods **5**, 671 (2008).

[74] J. A. Torreno-Pina, C. Manzo, and M. F. Garcia-Parajo, Uncovering homo-and hetero-interactions on the cell membrane using single particle tracking approaches, J. Phys. D: Appl. Phys. **49**, 104002 (2016).

[75] F. Laurent, H. Verdier, M. Duval, A. Serov, C. L. Vestergaard, and J.-B. Masson, TRamWay: Mapping physical properties of individual biomolecule random motion in large-scale single-particle tracking experiments, Bioinformatics **38**, 3149 (2022).

[76] A. D. Viñales and M. A. Despósito, Anomalous diffusion induced by a Mittag-Leffler correlated noise, Phys. Rev. E **75**, 042102 (2007).

[77] GitHub, fbm, https://github.com/crflynn/fbm, 2019.

[78] GitHub, mittag-leffler, https://github.com/khinsen/mittag-leffler, 2017.

[79] R. B. Davies and D. S. Harte, Tests for Hurst effect, Biometrika **74**, 95 (1987).

[80] GitHub, U-AnDi, https://github.com/huangzih/U-AnDi, 2023.

[81] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Clarendon, Oxford, 1987).

[82] A. D. Viñales and G. H. Paissan, Velocity autocorrelation of a free particle driven by a Mittag-Leffler noise: Fractional dynamics and temporal behaviors, Phys. Rev. E **90**, 062103 (2014).

[83] S. Ioffe and C. Szegedy, Batch normalization: Accelerating deep network training by reducing internal covariate shift, in *Proceedings of the 32nd International Conference on International Conference on Machine Learning - Volume 37* (JMLR.org, Lille, France, 2015), pp. 448–456.