PAPER

WaveNet-based deep neural networks for the characterization of anomalous diffusion (WADNet)

View the article online for updates and enhancements.

WaveNet-based deep neural networks for the characterization of anomalous diffusion (WADNet)

Dezhong Li, Qiujin Yao and Zihan Huang[*]

School of Physics and Electronics, Hunan University, Changsha 410082, People's Republic of China

E-mail: huangzih@hnu.edu.cn

Abstract

Anomalous diffusion, which shows a deviation of transport dynamics from the framework of standard Brownian motion, is involved in the evolution of various physical, chemical, biological, and economic systems. The study of such random processes is of fundamental importance in unveiling the physical properties of random walkers and complex systems. However, classical methods to characterize anomalous diffusion are often disqualified for individual short trajectories, leading to the launch of the anomalous diffusion (AnDi) challenge. This challenge aims at objectively assessing and comparing new approaches for single trajectory characterization, with respect to three different aspects: the inference of the anomalous diffusion exponent; the classification of the diffusion model; and the segmentation of trajectories. In this article, to address the inference and classification tasks in the challenge, we develop a WaveNet-based deep neural network (WADNet) by combining a modified WaveNet encoder with long short-term memory networks, without any prior knowledge of anomalous diffusion. As the performance of our model has surpassed the current 1st places in the challenge leaderboard on both two tasks for all dimensions (6 subtasks), WADNet could be part of state-of-the-art techniques to decode the AnDi database. Our method presents a benchmark for future research, and could accelerate the development of a versatile tool for the characterization of anomalous diffusion.

Keywords: anomalous diffusion, single trajectory characterization, anomalous diffusion challenge, machine learning, deep neural network, WaveNet

(Some figures may appear in colour only in the online journal)

[*]Author to whom any correspondence should be addressed.

1. Introduction

A random walker undergoes anomalous diffusion when its transport dynamics deviates from the principle of standard Brownian motion [1, 2]. Close attention is paid to such diffusion process in a variety of scientific fields, such as physics [3–6], chemistry [7], biology [8–12], economics [13–15], and social science [16]. The raw information of anomalous diffusion dynamics can be recorded as trajectories of random walkers via techniques like single particle tracking [17–21]. Since these walkers directly probe their surroundings [22–27], an in-depth characterization of their trajectory data is of significant importance in not only understanding the underlying mechanisms of anomalous diffusion, but also uncovering the intrinsic properties of both random walkers and complex environments.

Classical methods to analyze these trajectories are mostly based on mathematical statistics. A representative example is the calculation of MSD, which is quantified by the time-average or ensemble-average of squared displacements during a time interval [28, 29]. These approaches can be valid to detect the diffusion properties when the trajectory is sufficiently long or a quantity of trajectories is available. However, as a consequence of the limitations of experimental techniques, one can usually obtain few, short, and noisy trajectories when recording a diffusion process [30–32]. These trajectories suffer the undesirable variations introduced by noise and a lack of significant statistics. Thus, mining information of anomalous diffusion from such a dataset is rather difficult and challenging for standard statistics-based approaches.

Therefore, alternative methods with improved performance on the characterization of anomalous diffusion are in great demand. This urgent issue leads to the launch of the anomalous diffusion (AnDi) challenge (https://andi-challenge.org) which is spearheaded by a team of notable scientists [33]. The AnDi challenge aims at assessing and comparing new approaches that go beyond classical methods for single trajectory characterization, regarding three tasks: the inference of the anomalous diffusion exponent; the classification of the diffusion model; and the segmentation of trajectories. Each task is split into 3 subtasks to deal with trajectories in different spatial dimensions (1D, 2D, and 3D). A challenge leaderboard is available by scoring the predictions of participating teams on a common test dataset. The rank of a team in the leaderboard can be an intuitive and objective criterion for the performance of the corresponding method.

On the other hand, an unprecedented revolution of machine learning technologies has been witnessed in the past few years [34–39]. In particular, deep learning models that process sequential (or time-series) data are recently undergoing a rapid development [36–39], from traditional recurrent neural networks (RNNs) [36, 37] to the attention-based transformer [39]. These models can automatically learn the rules to extract useful information from sequences without any prior knowledge. Since trajectories of random walkers are actually sequences, these deep networks are highly expected to be qualified for the characterization of anomalous diffusion [40–46]. In this article, as a response to the AnDi challenge, we develop a WaveNet-based deep neural network (WADNet) by combining a modified WaveNet encoder [38] with long short-term memory (LSTM) networks [36], to address two tasks in the challenge: the inference of the anomalous diffusion exponent, and the classification of the diffusion model. Despite that no prior knowledge of anomalous diffusion is provided for WADNet, the performance of our model has surpassed the current 1st places in the challenge leaderboard on both two tasks for all dimensions (6 subtasks). Such a result suggests that WADNet could be part of state-of-the-art techniques to decode the AnDi database.

2. Methods

2.1. Tasks, evaluation metrics, and datasets

In this work, we focus on two tasks in the AnDi challenge: the inference of the anomalous diffusion exponent (the inference task), and the classification of the diffusion model (the classification task). Brief introductions to these two tasks and corresponding evaluation metrics are given as follows:

- **Inference of the anomalous diffusion exponent**. One of the most common behaviors in anomalous diffusion dynamics is the nonlinear growth of MSD versus time, given by:

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

Here, the anomalous diffusion exponent α is a real positive number. For standard Brownian diffusion we have $\alpha = 1$, while for superdiffusion (subdiffusion) we have $\alpha > 1 (\alpha < 1)$. Representative 1D trajectories (part) of these three diffusion patterns are demonstrated in figure 1(a) where $\alpha = 0.1, 1.0$, and 1.8 respectively. To identify the exponent α of a single trajectory $\mathbf{x}(t)$, traditional statistics-based method is calculating the time-averaged mean squared displacement (TA-MSD), written as [41]:

$$\text{TA-MSD}(t) = \frac{1}{t_{\max} - t} \int_0^{t_{\max} - t} \|\mathbf{x}(\tau + t) - \mathbf{x}(\tau)\|_2^2 \mathrm{d}\tau. \tag{2}$$

For example, time evolutions of TA-MSD corresponding to three trajectories in figure 1(a) are given in figure 1(b). As shown by the triangles, the exponent α can be determined by fitting the slope of the log–log plot of TA-MSD versus time. However, such an approach relies heavily on a sufficient long trajectory, and the result can usually be sensitive to the manual selection of fitting range. Hence, in this task, AnDi challenge calls for improved methods to infer the anomalous diffusion exponent of a single short trajectory. The evaluation metric for this regression problem is the mean absolute error (MAE):

$$\text{MAE} = \frac{1}{N} \sum_{i=1}^{N} |\alpha_{i,\mathrm{p}} - \alpha_{i,\mathrm{GT}}|, \tag{3}$$

where N is the number of trajectories in the test dataset, $\alpha_{i,\mathrm{p}}$ and $\alpha_{i,\mathrm{GT}}$ are the predicted and ground truth values of anomalous diffusion exponent of the i-th trajectory, respectively.

- **Classification of the diffusion model**. It is well known that the properties of standard Brownian motion can be captured by the Wiener process. Analogously, the dynamics of anomalous diffusion can also be described by various theoretical stochastic process models. However, confident identification of these models from trajectory data is quite challenging for classical methods. Thus, to address this issue, the participants in this task are requested to develop valid approaches to classify the diffusion model of a single short trajectory. Five typical diffusion models are included in the challenge. Here, we give a brief introduction to these models and show the range of corresponding α in the AnDi database:

    * Annealed transient time motion (ATTM) [47], a non-ergodic model that describes a Brownian motion with the diffusion coefficient randomly varying either in time or space ($0.05 \leqslant \alpha \leqslant 1$).
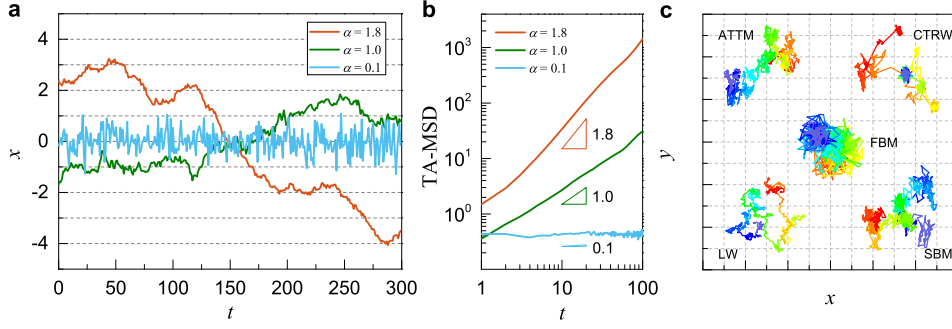
Figure 1. Schematic illustrations of the inference and classification tasks in the AnDi challenge. (a) Representative 1D trajectories (part) of subdiffusion ($\alpha = 0.1$), Brownian diffusion ($\alpha = 1.0$), and superdiffusion ($\alpha = 1.8$), respectively. (b) Time evolutions of TA-MSD corresponding to three trajectories in (a). The exponent α can be determined by fitting the slope (shown as the triangle) of the log–log plot of TA-MSD versus time. (c) Representative trajectories of five diffusion models in the AnDi database: annealed transient time motion (ATTM), continuous-time random walk (CTRW), fractional Brownian motion (FBM), Lévy walk (LW), and scaled Brownian motion (SBM). The color on the trajectory denotes the time.

* Continuous-time random walk (CTRW) [48], a non-ergodic model where the waiting time of a random walker between two subsequent steps is irregular and randomly chosen ($0.05 \leqslant \alpha \leqslant 1$).
* Fractional Brownian motion (FBM) [49], an ergodic diffusion process that is driven by a fractional Gaussian noise. Such noise is normally distributed but has a power-law correlation ($0.05 \leqslant \alpha < 2$).
* Lévy walk (LW) [50], a non-ergodic model where the waiting time between subsequent steps is also irregular as CTRW but the step length is not Gaussian distributed ($1 \leqslant \alpha \leqslant 2$).
* Scaled Brownian motion (SBM) [51], a non-ergodic model that describes a Brownian motion with a deterministically time-dependent diffusion coefficient ($0.05 \leqslant \alpha \leqslant 2$).

Detailed descriptions regarding the mechanisms of these theoretical models can be found in the instruction of AnDi challenge [33]. Representative trajectories of these five models are shown in figure 1(c), where color on the trajectory denotes the time. The evaluation metric in this classification task is the 'micro' version of F1-score, expressed as:

$$\text{F1} = 2 \cdot \frac{\text{precision} \cdot \text{recall}}{\text{precision} + \text{recall}}. \tag{4}$$

Here, precision $= \frac{\text{TP}}{\text{TP+FP}}$ and recall $= \frac{\text{TP}}{\text{TP+FN}}$, where TP, FP and TN represent the *total* number of true positives, false positives and false negatives in the classification result, respectively.

The datasets in the AnDi challenge are composed of simulated trajectories that are generated by an open-source Python package *andi-dataset* (https://github.com/AnDiChallenge/ANDI_datasets) [52]. The working principle of *andi-dataset* is briefly summarized here. Trajectory generation is based on the theoretical framework of five diffusion models in the database, where time steps are uniformly sampled with a unitary time interval. The ground truth value of anomalous diffusion exponent α_{GT} ranges from 0.05 to

Table 1. List of specific lengths and corresponding trajectory numbers in training set.

Length	Number	Length	Number	Length	Number
10	5000 000	105	2000 000	400	1000 000
15	5000 000	115	2000 000	425	1000 000
20	5000 000	120	2000 000	450	1000 000
25	5000 000	125	2000 000	475	1000 000
30	5000 000	150	2000 000	500	1000 000
40	5000 000	175	2000 000	550	1000 000
45	5000 000	200	2000 000	600	1000 000
50	5000 000	225	2000 000	650	1000 000
55	5000 000	250	2000 000	700	1000 000
60	5000 000	250	2000 000	750	1000 000
70	5000 000	275	1500 000	800	1000 000
80	5000 000	300	1500 000	850	1000 000
90	5000 000	325	1500 000	900	1000 000
100	5000 000	350	1500 000	950	1000 000
105	2000 000	375	1500 000		

2.0 with a step 0.05. Trajectory length varies from 10 to 999 points, where one point denotes the spatial coordinate of the random walker at one time step. In particular, to mimic real experimental data, simulated trajectories are corrupted with zero-mean Gaussian noise. The signal-to-noise ratio (SNR) of a trajectory can be regulated through changing the standard deviation σ_n of noise distribution (SNR $= 1/\sigma_n$). In the *andi-dataset*, σ_n is set as 1.0, 0.5, or 0.1, resulting in SNR $= 1$, 2, or 10.

On the basis of *andi-dataset*, a test dataset consisting of 10 000 trajectories is generated by organizers as the common dataset for leaderboard scoring. The training dataset we utilized in this work is also generated by the package *andi-dataset*. For both the inference and classification tasks, we choose 43 specific lengths and separately generate the training subsets at these lengths. The values of these specific lengths and corresponding trajectory numbers are listed in table 1. Before training or inference, trajectory data is normalized to ensure its position's average and standard deviation in each dimension are 0 and 1 respectively.

2.2. Workflow of WADNet

The overview of the entire workflow of WADNet is demonstrated in figure 2, including three main parts: *input*, *WaveNet encoder*, and *RNN & MLP*. As shown in the *input* part, input data can be obtained through a simple preprocessing of the raw trajectory, which is valid for trajectories in arbitrary dimensions: (i) we separate the trajectory in each dimension into 1D time-series data; (ii) we stack the 1D time-series data of all dimensions to compose a multi-channel tensor as the input of WADNet. For instance, the transformation of a 3D trajectory $\mathbf{x}(t) = [x_1(t), x_2(t), x_3(t)]^{\mathrm{T}}$ with a length T to an input tensor $\bar{\mathbf{x}}$ can be expressed as:

$$\left\{ \begin{bmatrix} x_1(1) \\ x_2(1) \\ x_3(1) \end{bmatrix}, \begin{bmatrix} x_1(2) \\ x_2(2) \\ x_3(2) \end{bmatrix}, \ldots, \begin{bmatrix} x_1(T) \\ x_2(T) \\ x_3(T) \end{bmatrix} \right\} \rightarrow \left\{ \begin{bmatrix} [x_1(1), x_1(2), \ldots, x_1(T)] \\ [x_2(1), x_2(2), \ldots, x_2(T)] \\ [x_3(1), x_3(2), \ldots, x_3(T)] \end{bmatrix} \right\} \quad (5)$$

After the preprocessing, the input tensor is going to pass through a modified WaveNet encoder. WaveNet is a deep neural network designed by Google Deepmind for raw audio generation [38]. The encoder of WaveNet based on dilated causal convolutions, gated activation
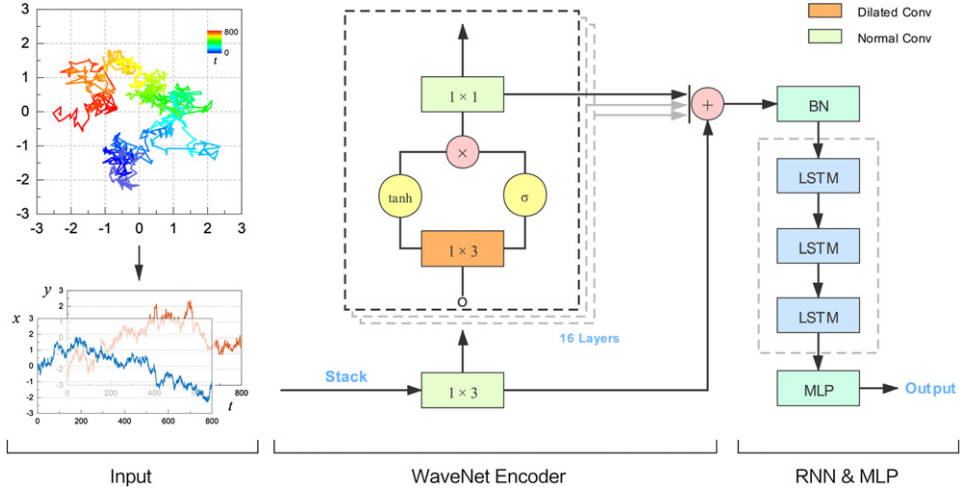
Figure 2. Overview of the workflow of WADNet. The workflow of WADNet consists of three main parts: *input*, *WaveNet encoder*, and *RNN & MLP*. The *input* part describes the preprocessing method of raw trajectories. The *WaveNet encoder* part shows the detailed architecture of the modified WaveNet encoder. The output of this encoder is utilized as the input of the following *RNN & MLP* part, in which the inference and classification tasks are accomplished.

units, and skip connections has been proved to be a powerful tool for extracting information from sequential data [53, 54]. Here, a dilated convolution with dilation m is a convolution operation where the filter skips the input values with a step $m-1$. Such a structure allows the filter to deal with a larger area than its size. For example, a 3×3 filter with dilation 4 can be directly applied on a 9×9 area. Thus, by simply stacking the dilated convolutions, networks are enabled to have very large receptive fields with just a few layers. In the *WaveNet encoder* part of figure 2, we show the detailed architecture of the modified WaveNet encoder utilized in this work, where the dilated convolution part is marked in orange. The input tensor $\bar{\mathbf{x}}$ is firstly filtered by a normal convolutional layer with filter size 1×3 and filter number 32. Corresponding output $\tilde{\mathbf{x}}_1$ is subsequently processed by the stacked dilated convolutions with a dilation depth $d = 16$. In this work, the dilation is doubled for every layer, which is the same as the original WaveNet [38].

Next, we apply the gated activation unit in each layer, which is firstly proposed in the gated PixelCNN [55] to introduce multiplicative units and model more complex interactions. For that purpose, the dilated convolution part in the modified WaveNet encoder is composed of two parallel convolution operators with the same structures, named as 'filter' and 'gate'. The outputs of 'filter' and 'gate' are activated by the hyperbolic tangent and sigmoid functions respectively, leading to an output $\mathbf{z}_k$ of the gated activation unit:

$$\mathbf{z}_k = \tanh\left(W_{f,k} * \tilde{\mathbf{x}}_k\right) \otimes \sigma\left(W_{g,k} * \tilde{\mathbf{x}}_k\right). \tag{6}$$

Here, k is the layer index, $\tilde{\mathbf{x}}_k$ is the input tensor of the kth layer, W represents the dilated convolution filter with filter size 1×3 and filter number 32, f and g denote the 'filter' and 'gate', $*$ denotes the convolution operator, $\tanh(\cdot)$ is the hyperbolic tangent activation function, $\sigma(\cdot)$ is the sigmoid activation function, $\otimes$ denotes the element-wise multiplication operator. After that, a 1×1 convolution filter W_1 with filter number 32 is applied on $\mathbf{z}_k$ to obtain the

input of next layer, written as:

$$\tilde{\mathbf{x}}_{k+1} = W_1 * \mathbf{z}_k, \quad k = 1, 2, 3, \ldots, d. \tag{7}$$

In particular, this operation is different from that in the original WaveNet encoder, where residual [35] is used in the layer, i.e. $\tilde{\mathbf{x}}_{k+1} = W_1 * \mathbf{z}_k + \tilde{\mathbf{x}}_k$. Skip connections are then utilized to obtain the output tensor $\mathbf{u}$ of the WaveNet encoder, given by:

$$\mathbf{u} = \mathrm{BN}\left(\sum_{k=1}^{d+1} \tilde{\mathbf{x}}_k\right), \tag{8}$$

Here, $\mathrm{BN}(\cdot)$ represents the 1D batch normalization which is achieved through a normalization step that is restrained to each mini-batch [56].

At last, $\mathbf{u}$ is used as the input of the following RNN module. In this work, LSTM is chosen as the building block of this module. As one of the most common RNN architectures, LSTM is developed to deal with the vanishing gradient problem that can be encountered when training traditional RNNs [36]. As shown in the *RNN & MLP* part of figure 2, the RNN module is made up of 3 stacked LSTM units, whose dimension of hidden layers is 64. The output $\mathbf{v}$ of RNN module is the feature vector of original trajectory learned by WADNet. To accomplish the challenge task, feature $\mathbf{v}$ is processed by a multilayer perceptron (MLP) network. For the inference task, the output dimension of MLP is 1 and no activation function is utilized; for the classification task, the output dimension of MLP is 5 and the activation function is the softmax function.

2.3. Training procedure and inference strategy

Models for each specific length are trained separately. 80% of training data is used for training, while the other is used for validation. The model is trained using back-propagation with a batch size 512, where loss function is the mean squared error (MSE) and cross entropy for the inference and classification tasks, respectively. The optimizer is Adam with an initial learning rate $l = 0.001$. During the training process, the learning rate is changed as $l_{\mathrm{new}} = l_{\mathrm{old}}/5$ if the valid loss does not decrease after 2 epochs. When the number of such learning-rate changes exceeds 2, the training process is early stopped to save training time and avoid overfitting.

To demonstrate the convergence of our method, we take the 1D trajectory at length 300 as the example, and show corresponding evolutions of train loss, valid loss, and valid metric in figure 3(a) (the inference task) and 3(b) (the classification task). Since train and valid losses can both be effectively reduced during training, an excellent convergence of WADNet can be identified. In particular, we determine the best model weight for a specific length by the best valid metric, as depicted by the arrows in figures 3(a) and (b). Moreover, to examine the model performance in the training process, we show the best valid metrics of all specific lengths in all dimensions in figure 3(c) (the inference task) and 3(d) (the classification task). For both two tasks, the performance of WADNet is better for higher spatial dimensions, and is gradually boosted as the trajectory length increases. This result arises from the fact that a longer trajectory with a higher dimension can provide more data points and hidden information.

Next, we pay attention to the inference of test dataset. Note that the length of a trajectory in the test dataset varies from 10 to 999 points. Since we only have model weights for 43 specific lengths, our model cannot be directly applied for inference. To address this issue, the inference of test dataset is guided by the following strategy:

- If the original length of a trajectory belongs to these 43 specific lengths, the data of this trajectory will be directly used for inference.
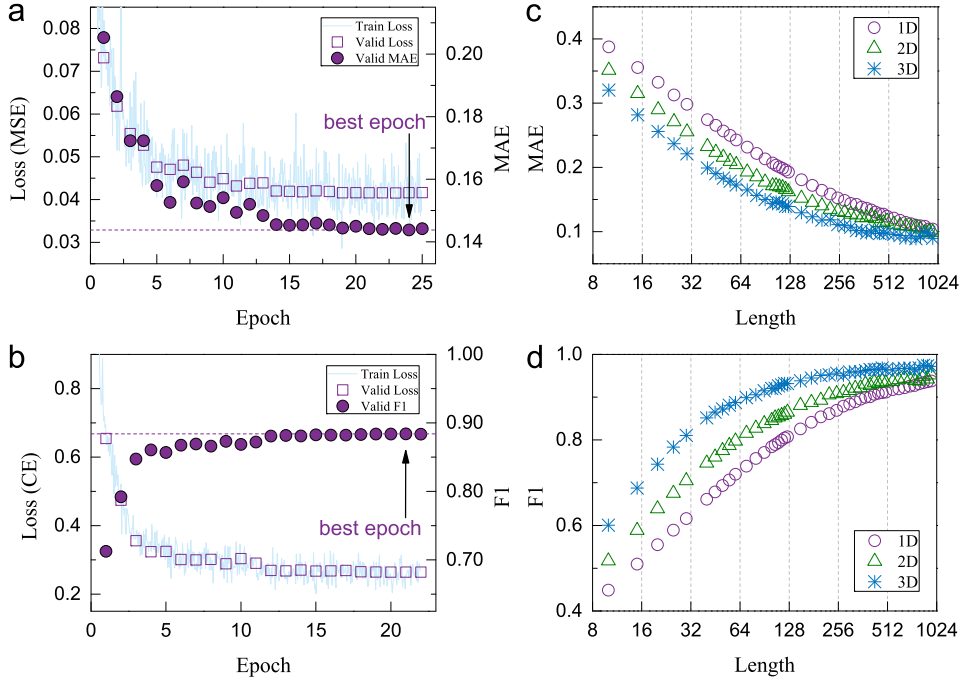
Figure 3. Model performance in the training process. (a) and (b) Evolutions of train loss, valid loss, and valid metric for the 1D trajectory at length 300, regarding the inference task (a) and the classification task (b). The arrow denotes the best model weight for a specific length, which is determined by the best valid metric. (c) and (d) The best valid metrics of all specific lengths in all dimensions for the inference task (c) and the classification task (d).

- Otherwise, a new length of this trajectory will be set as the closest smaller specific length. For instance, the new length of a trajectory with an original length 49 should be 45. The trajectory data is subsequently transformed into 2 sequences. For clarity, we set the trajectory as $\mathbf{x} = [x_1, x_2, \ldots, x_T]$, where T is the original length. We denote T_n as the new length. Note that $T_n < T$, two sequences $\mathbf{x}_1 = [x_1, x_2, \ldots, x_{T_n}]$ and $\mathbf{x}_2 = [x_{T-T_n+1}, x_{T-T_n+2}, \ldots, x_T]$ with length T_n can be obtained. Such two sequences are both used for inference, where corresponding feature vectors learned by WADNet are $\mathbf{v}_1$ and $\mathbf{v}_2$ respectively. After that, we choose the mean vector of these two vectors as the approximation of the feature $\mathbf{v}$ of the original trajectory, i.e. $\mathbf{v} \approx (\mathbf{v}_1 + \mathbf{v}_2)/2$. The final prediction is made based on $\mathbf{v}$ by the MLP.

The codes of WADNet for training and inference are freely available online [57], where the neural networks are implemented by PyTorch 1.6.0.

In addition, to further improve the performance of WADNet, K-fold cross validation technique [58] is utilized for both two tasks. In detail, the training dataset is randomly split into K folds where $K = 5$ in this work. When training a model, one fold is utilized as the validation data while other folds are used as training data. Such process is then repeated K times, with each of the folds used exactly once as the validation data. Thus, the prediction of a sample can be obtained by averaging the results of K single-fold models. On the other hand, for the inference of the anomalous diffusion exponent, predictions of test dataset are all rectified by

multiplying a constant w and are subsequently clipped to ensure $\alpha \in [0.05, 2.0]$. The constant w is determined by analyzing an external validation dataset containing 100 000 trajectories, and is identified to be 1.0110, 1.0129, and 1.0083 for 1D, 2D, and 3D trajectories, respectively.

3. Results and discussions

3.1. Performance of WADNet in the challenge leaderboard

To examine objectively the performance of WADNet in the AnDi challenge, we compare our scores of test dataset with the top 4 ranks in the challenge leaderboard. The results of comparisons for both the inference and classification tasks in all dimensions (6 subtasks) are summarized in figure 4, where the best scores are marked in different colors for clarity. It can be found that the scores of our method have surpassed the current 1st places in the leaderboard in all 6 subtasks. In particular, without the improvement of K-fold cross validation, the scores of a single-fold model (WADNet-SF) have also surpassed the 1st places in 5 subtasks. This result illustrates that the ability of WADNet in the characterization of anomalous diffusion primarily arises from the model itself. Such ability could be mainly attributed to the WaveNet encoder, which is initially designed to extract features from acoustic waves and has demonstrated excellent performance on processing time-series data [38, 53, 54]. Since trajectories are also time-series data, WaveNet encoder shall be qualified to deal with raw trajectory data and is expected to improve the performance of the baseline model. Therefore, WADNet could be part of state-of-the-art techniques for the AnDi database. Note that there are only 43 specific lengths in the training dataset. That is, by a simple expansion of training data, the performance of WADNet can be further improved. Analysis in the following sections is based on the best performance of WADNet that is improved by the K-fold cross validation.

3.2. Analysis of the inference of the anomalous diffusion exponent

In this section, we provide a detailed analysis of the performance of WADNet on the inference task. Firstly, as the spatial dimension of trajectory increases, MAE value of test dataset is reduced from 0.150 69 (1D) to 0.132 00 (2D) to 0.112 11 (3D). This result can be attributed to the extra information given by a higher dimension. Moreover, to intuitively demonstrate the differences between the predicted values α_p and ground truth values α_{GT}, we plot the 2D histograms of α_p versus α_{GT} for all dimensions in figures 5(a)–(c). It can be found that 1D and 2D trajectories with $\alpha_{GT} \approx 1.0$ are more likely to be predicted as subdiffusion ($\alpha_p < 1$), while such a deviation is slightly eliminated in 3D space. This bias can also be captured by the plots of MAE versus α_{GT} (figures 5(d)–(f)), where MAE values at $\alpha_{GT} \approx 1.0$ are relatively higher than that of neighborhoods. Meanwhile, significant errors can be identified at $\alpha_{GT} \to 0$. The reason is that a random walker in this case is nearly motionless and corresponding trajectory is more like a record of localization noise.

Next, we explore the dependence of model performance on different diffusion models, and show the results in figure 6(a). As expected, FBM that is ergodic has the lowest MAE values in all dimensions. Interestingly, for those non-ergodic models, WADNet shows a better performance on CTRW and LW than on ATTM and SBM, indicating that trajectories of CTRW and LW could provide more confidential information than the other two. Additionally, MAE values of CTRW and LW are nearly independent of the trajectory dimension, while those of ATTM, FBM, and SBM decrease as the dimension increases. This result suggests that the quality of features of CTRW and LW learned by WADNet is less correlated with the spatial dimension.

On the other hand, since experimental trajectories can usually be noisy, it is of crucial importance to examine the performance of WADNet on noisy trajectories. For that purpose, we
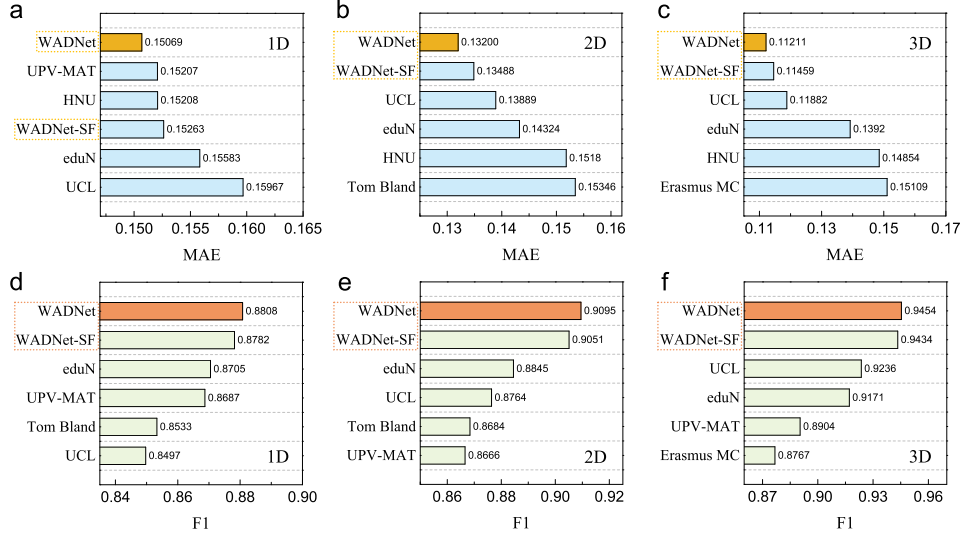
J. Phys. A: Math. Theor. **54** (2021) 404003

D Li *et al*

Figure 4. Performance of WADNet in the challenge leaderboard. Results of comparisons between WADNet and the top 4 ranks in the challenge leaderboard. Label 'WADNet' denotes the score of our method improved by the *K*-fold cross validation, and label 'WADNet-SF' denotes the score of a single-fold WADNet model. The best score in each subtask is marked in different color for clarity. (a)–(c) MAE scores in the inference task for 1D (a), 2D (b), and 3D (c) trajectories. (d)–(f) F1-scores in the classification task for 1D (d), 2D (e), and 3D (f) trajectories.

investigate the effects of SNR on the performance of our model. As shown in figure 6(b), MAE value can be slightly reduced by the improvement of SNR for the 1D trajectory. However, a surprising result is that MAE values are nearly identical for all SNR in 2D and 3D spaces. Such a result implies that WADNet has the ability to remove the undesirable effects introduced by localization noise for 2D and 3D trajectories, even for trajectories where the standard deviation of noise has the same amplitude as that of displacement (i.e. SNR = 1). Therefore, our method is of high potential to be utilized for the analysis of noisy experimental trajectory data. At last, the influence of trajectory length is demonstrated in figure 6(c). As expected, MAE rapidly decreases with the increase of trajectory length, and reaches a relatively steady value near the length 600.

3.3. Analysis of the classification of the diffusion model

In this section, we focus attention on the performance of WADNet on the classification task. At first, F1-score of test dataset is boosted from 0.8808 (1D) to 0.9095 (2D) to 0.9454 (3D), as a consequence of the information gain given by a higher spatial dimension. To examine the classification performance of WADNet in detail, we show the confusion matrixes for 1D, 2D and 3D trajectories in figures 7(a)–(c) respectively. For clarity, we use the percentage instead of the sample number as the element of confusion matrix. Thus, the diagonal element of the matrix is actually the classification accuracy of WADNet for each model. As characterized by the accuracy, the performance on CTRW and LW is relatively better (accuracy >98% in all dimensions) and almost independent on the spatial dimension, while the performance on ATTM, FBM, and SBM is gradually improved as the dimension increases.
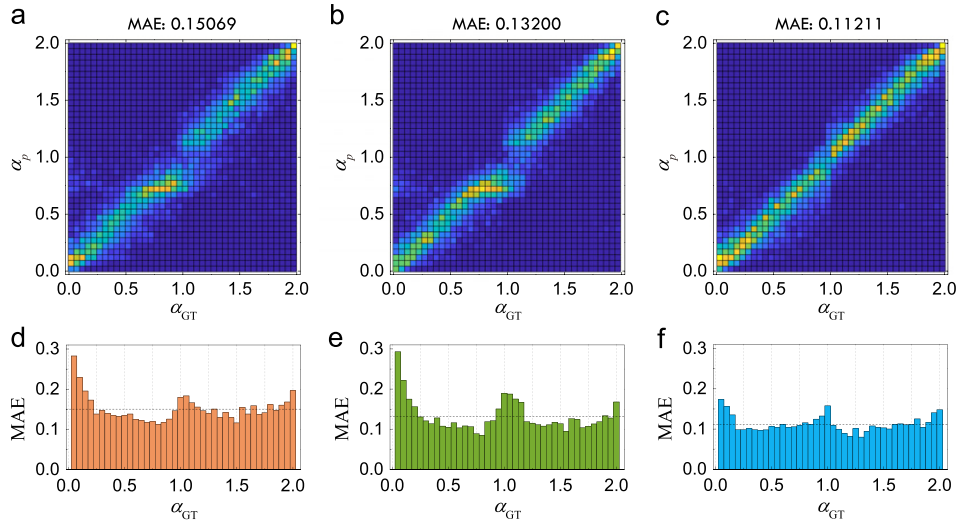
Figure 5. Differences between the predicted values and ground truth values of the anomalous diffusion exponent. (a)–(c) 2D histograms of α_p versus α_{GT} for 1D (a), 2D (b), and 3D (c) trajectories. (d)–(f) Plots of MAE versus α_{GT} for 1D (d), 2D (e), and 3D (f) trajectories. The dashed lines denote the mean values of MAE.
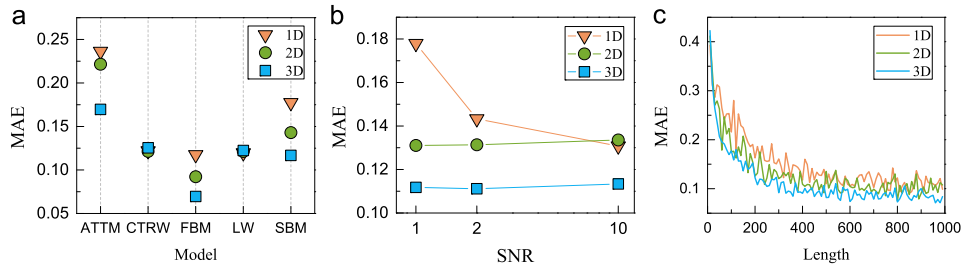

Figure 6. Influence of the diffusion model, SNR, and trajectory length on the inference task. (a) MAE of 5 diffusion models. (b) Plots of MAE versus SNR. (c) Dependence of MAE on the trajectory length.

From another perspective, the classification ability of WADNet is further analyzed by calculating the precision and recall scores of each diffusion model. The plots of precision versus recall for all dimensions are given in figures 7(d)–(f) as dots, where the diffusion models are distinguished by different colors. A general phenomenon in all dimensions is that precision and recall are approximately equivalent for LW and SBM, but show non-negligible differences for the other 3 models. In more detail, the recall is smaller than precision for ATTM, and undergoes a reverse for FBM and CTRW. That is, regarding the error of classifying a trajectory, ATTM is more likely to be mistaken, while FBM and CTRW are more preferred to be the targets.

Next, similar to the inference task, the performance of WADNet on noisy trajectories is also examined by studying the effects of SNR. Corresponding results in all dimensions are plotted in figures 7(d)–(f) as squares marked by SNR values. The colored region is a guide of square locations to the eye, and its area reports the stability of our method on noisy trajectories for each model. As characterized by the small area of colored regions, insensitivity of F1-score to
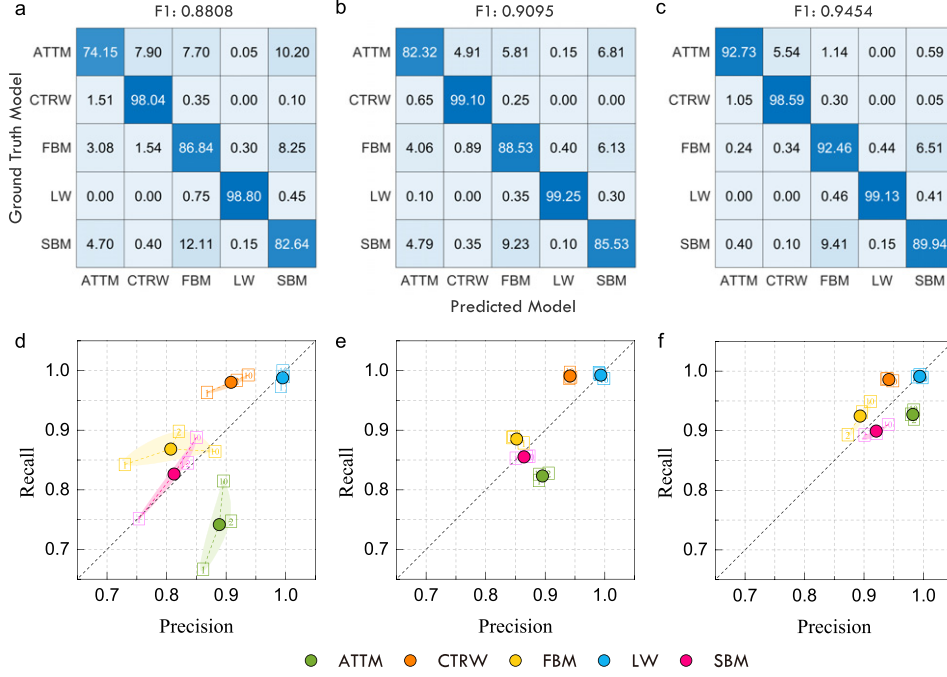
Figure 7. Detailed analysis of the performance of WADNet on the classification task. (a)–(c) Confusion matrixes of classification results for 1D (a), 2D (b) and 3D (c) trajectories. The element of confusion matrix is the percentage, and the diagonal element of matrix is the classification accuracy for each model. (d)–(f) Plots of precision versus recall for 1D (d), 2D (e) and 3D (f) trajectories. The dots denote the overall scores, and the squares denote the results for different SNR. The colored region is a guide of square locations to the eye, and its area reports the stability of WADNet on noisy trajectories for each model.

SNR can be identified for LW and CTRW in all dimensions. In contrast, F1-scores of ATTM, FBM, and SBM in 1D space can be slightly improved by the increase of SNR. However, for 2D and 3D trajectories, F1-scores of these 3 models are nearly independent of SNR, as the same with LW and CTRW. Thus, negative influences of localization noise on the classification performance could be vastly eliminated by WADNet.

Moreover, to investigate the influence of anomalous diffusion exponent on the classification result, we show the F1-score as a function of α_{GT} in figure 8(a). Obviously, the worst performance of WADNet can be identified for trajectories with $\alpha_{GT} \approx 1.0$ in all dimensions. The possible reason is that the dynamics of these trajectories is close to Brownian motion, which leads to similar statistical properties regardless of the specific diffusion model. In addition, the classification performance is also poor at $\alpha_{GT} \to 0$ in 1D and 2D spaces. This result can be attributed to the fact that the transport of a random walker with an extremely small α is nearly immobile and dominated by random noise. This phenomenon vanishes for 3D trajectories, probably owing to the information gain. In addition, the dependence of F1-score on trajectory length is given in figure 8(b). As was anticipated, F1-score undergoes a rapid growth as the trajectory length increases, and reaches a relatively steady value near the length 600.

At last, to gain a more in-depth insight into the classification ability of our method, we present a visualization of feature vector **v** extracted from the penultimate layer of WADNet.
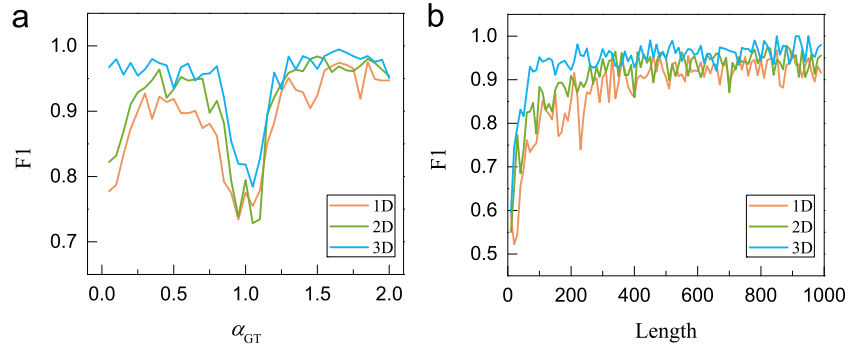
Figure 8. Influence of anomalous diffusion exponent and trajectory length on the classification task. (a) Plots of F1-score as a function of α_{GT}. (b) Plots of F1-score versus trajectory length.
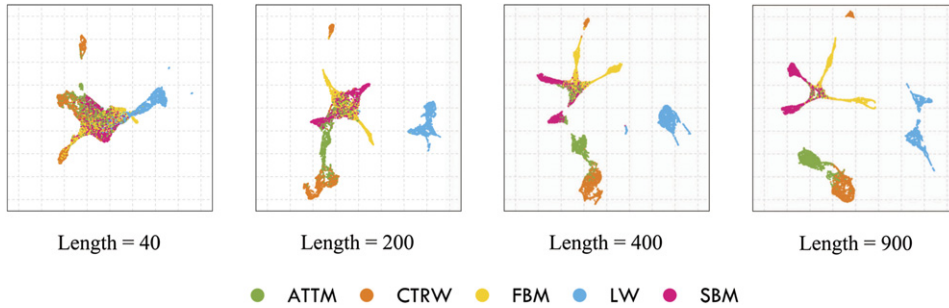

Figure 9. Visualization of learned features. Uniform manifold approximation and projection for dimension reduction (UMAP) projections of feature vectors of 1D trajectories with length 40, 200, 400, and 900, where each length has 10 000 trajectories. Feature vectors are extracted from the penultimate layer of WADNet.

Figure 9 shows the uniform manifold approximation and projection for dimension reduction (UMAP) [59] projections of feature vectors of 1D trajectories with length 40, 200, 400, and 900, where each length has 10 000 trajectories. Projections of features are marked by different colors according to corresponding diffusion models. It can be found that features are mixed for short trajectories, but can be clearly distinguished for long trajectories. In more detail, as the trajectory length increases, the features of LW can be firstly separated (length 200), and two main isolated clusters can be identified next (length 400 and 900). One cluster is primarily composed of FBM and SBM, and the other is a combination of CTRW and ATTM. Therefore, from the perspective of feature similarity, the underlying dynamics of these diffusion models is not completely uncorrelated. The overlap between FBM and SBM suggests that the diffusion driven by a fractional Gaussian noise can sometimes be similar to the Brownian motion with a time-dependent diffusivity. Meanwhile, the effects of an irregular waiting time might be approximated by a random variation of diffusivity at times, as indicated by the adjacency between CTRW and ATTM. For future research, these features learned by WADNet have the potential to be used as quantitative representations of trajectories with multiple and mixed diffusion dynamics.

J. Phys. A: Math. Theor. **54** (2021) 404003

D Li *et al*

4. Conclusions

In conclusion, by combining a modified WaveNet encoder with LSTM networks, we have developed a WaveNet-based deep neural network (WADNet) for the characterization of anomalous diffusion in response to the AnDi challenge. Without any prior knowledge of anomalous diffusion, the performance of WADNet has surpassed the current 1st places in the challenge leaderboard for all dimensions, with respect to the inference of the anomalous diffusion exponent and the classification of the diffusion model. That is, WADNet could be part of state-of-the-art techniques to decode the AnDi database. Moreover, the results show that WADNet can effectively extract information from noisy trajectories and vastly eliminate the undesirable influences of localization noise. The features learned by WADNet have the potential to be utilized as the quantitative representation of complex anomalous diffusion dynamics. In particular, the main architectures of WADNet for different tasks are nearly the same, where modifications are only applied on the input and output dimensions of the network. Thus, our findings bring new insights into the generalization of anomalous diffusion, and could accelerate the design of a versatile strategy for single trajectory characterization.

Acknowledgments

Data availability statement

The data that support the findings of this study are openly available at the following URL/DOI: https://github.com/huangzih/WADNet.

ORCID iDs

Zihan Huang ⓘ https://orcid.org/0000-0003-3094-9017

References

[1] Metzler R, Jeon J-H, Cherstvy A G and Barkai E 2014 Anomalous diffusion models and their properties: non-stationarity, non-ergodicity, and ageing at the centenary of single particle tracking *Phys. Chem. Chem. Phys.* **16** 24128–64
[2] Klafter J and Sokolov I M 2005 Anomalous diffusion spreads its wings *Phys. World* **18** 29
[3] Mason T G and Weitz D A 1995 Optical measurements of frequency-dependent linear viscoelastic moduli of complex fluids *Phys. Rev. Lett.* **74** 1250
[4] Aarão Reis F D A 2016 Scaling relations in the diffusive infiltration in fractals *Phys. Rev.* E **94** 052124
[5] Volpe G and Volpe G 2017 The topography of the environment alters the optimal search strategy for active particles *Proc. Natl Acad. Sci. USA* **114** 11350–5
[6] Xu Z, Dai X, Bu X, Yang Y, Zhang X, Man X, Zhang X, Doi M and Yan L-T 2021 Enhanced heterogeneous diffusion of nanoparticles in semiflexible networks *ACS Nano* **15** 4608–16
[7] Barkai E, Garini Y and Metzler R 2012 Strange kinetics of single molecules in living cells *Phys. Today* **65** 29

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

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

[10] González M C, Hidalgo C A and Barabási A-L 2008 Understanding individual human mobility patterns *Nature* **453** 779–82

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

[12] Chen P, Huang Z, Liang J, Cui T, Zhang X, Miao B and Yan L-T 2016 Diffusion and directionality of charged nanoparticles on lipid bilayer membrane *ACS Nano* **10** 11541–7

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

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

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

[16] Castellano C, Fortunato S and Loreto V 2009 Statistical physics of social dynamics *Rev. Mod. Phys.* **81** 591

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

[18] Shen H, Tauzin L J, Baiyasi R, Wang W, Moringo N, Shuang B and Landes C F 2017 Single particle tracking: from theory to biophysical applications *Chem. Rev.* **117** 7331–76

[19] Qian H, Sheetz M P and Elson E L 1991 Single particle tracking. Analysis of diffusion and flow in two-dimensional systems *Biophys. J.* **60** 910–21

[20] Saxton M J 2008 Single-particle tracking: connecting the dots *Nat. Methods* **5** 671–2

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

[22] Wong I Y, Gardel M L, Reichman D R, Weeks E R, Valentine M T, Bausch A R and Weitz D A 2004 Anomalous diffusion probes microstructure dynamics of entangled F-actin networks *Phys. Rev. Lett.* **92** 178101

[23] Banks D S and Fradin C 2005 Anomalous diffusion of proteins due to molecular crowding *Biophys. J.* **89** 2960–71

[24] Chen D T N, Lau A W C, Hough L A, Islam M F, Goulian M, Lubensky T C and Yodh A G 2007 Fluctuations and rheology in active bacterial suspensions *Phys. Rev. Lett.* **99** 148302

[25] Huang Z, Chen P, Zhu G, Yang Y, Xu Z and Yan L-T 2018 Bacteria-activated Janus particles driven by chemotaxis *ACS Nano* **12** 6725–33

[26] Chen P, Yue H, Zhai X, Huang Z, Ma G-H, Wei W and Yan L-T 2019 Transport of a graphene nanosheet sandwiched inside cell membranes *Sci. Adv.* **5** 3192

[27] Lozano C, Gomez-Solano J R and Bechinger C 2019 Active particles sense micromechanical properties of glasses *Nat. Mater.* **18** 1118–23

[28] Wang B, Anthony S M, Bae S C and Granick S 2009 Anomalous yet Brownian *Proc. Natl Acad. Sci.* **106** 15160–4

[29] Kim J, Kim C and Sung B J 2013 Simulation study of seemingly Fickian but heterogeneous dynamics of two dimensional colloids *Phys. Rev. Lett.* **110** 047801

[30] Weron A, Janczura J, Boryczka E, Sungkaworn T and Calebiro D 2019 Statistical testing approach for fractional anomalous diffusion classification *Phys. Rev.* E **99** 042149

[31] Jeon J-H and Metzler R 2010 Analysis of short subdiffusive time series: scatter of the time-averaged mean-squared displacement *J. Phys. A: Math. Theor.* **43** 252001

[32] Akin E J, Solé L, Johnson B, Beheiry M e, Masson J-B, Krapf D and Tamkun M M 2016 Single-molecule imaging of Na$_v$1.6 on the surface of hippocampal neurons reveals somatic nanoclusters *Biophys. J.* **111** 1235–47

[33] Muñoz-Gil G, Volpe G, Garcia-March M A, Metzler R, Lewenstein M and Manzo C 2020 AnDi: the anomalous diffusion challenge (arXiv:2003.12036)

[34] Silver D *et al* 2016 Mastering the game of Go with deep neural networks and tree search *Nature* **529** 484–9

[35] He K, Zhang X, Ren S and Sun J 2016 Deep residual learning for image recognition *Proc. IEEE Conf. Computer Vision and Pattern Recognition* pp 770–8

[36] Hochreiter S and Schmidhuber J 1997 Long short-term memory *Neural Comput.* **9** 1735–80

[37] Cho K, van Merriënboer B, Gulcehre C, Bahdanau D, Bougres F, Schwenk H and Bengio Y 2014 Learning phrase representations using RNN encoder–decoder for statistical machine translation *Proc. Conf. Empirical Methods in Natural Language Processing* pp 1724–34

[38] van den Oord A, Dieleman S, Zen H, Simonyan K, Vinyals O, Graves A, Kalchbrenner N, Senior A and Kavukcuoglu K 2016 WaveNet: a generative model for raw audio (arXiv:1609.03499)

[39] Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A N, Kaiser L and Polosukhin I 2017 Attention is all you need *Proc. Int. Conf. Neural Information Processing Systems* pp 6000–10

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

[41] Muñoz-Gil G, Garcia-March M A, Manzo C, Martín-Guerrero J D and Lewenstein M 2020 Single trajectory characterization via machine learning *New J. Phys.* **22** 013010

[42] Muñoz-Gil G *et al* 2021 Objective comparison of methods to decode anomalous diffusion (arXiv:2105.06766)

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

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

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

[46] Manzo C 2021 Extreme learning machine for the characterization of anomalous diffusion from single trajectories (AnDi-ELM) *J. Phys. A: Math. Theor.* **54** 334002

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

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

[49] Mandelbrot B B and Ness J W V 2006 Fractional Brownian motions, fractional noises and applications *SIAM Rev.* **10** 422–37

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

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

[52] Muñoz-Gil G, Manzo C, Volpe G, Garcia-March M A, Metzler R and Lewenstein M 2020 The anomalous diffusion challenge dataset https://zenodo.org/record/3707702

[53] Kleijn W B, Lim F S C, Luebs A, Skoglund J, Stimberg F, Wang Q and Walters T C 2018 Wavenet based low rate speech coding *Proc. IEEE Int. Conf. Acoustics, Speech and Signal Processing* pp 676–80

[54] Chorowski J, Weiss R J, Bengio S and van den Oord A 2019 Unsupervised speech representation learning using WaveNet autoencoders *IEEE/ACM Trans. Audio Speech Lang. Process.* **27** 2041–53

[55] van den Oord A, Kalchbrenner N, Vinyals O, Espeholt L, Graves A and Kavukcuoglu K 2016 Conditional image generation with PixelCNN decoders *Proc. Int. Conf. Neural Information Processing Systems* pp 4797–805

[56] Ioffe S and Szegedy C 2015 Batch normalization: accelerating deep network training by reducing internal covariate shift *Proc. Int. Conf. Machine Learning* 37 pp 448–56

[57] Li D, Yao Q and Huang Z 2021 WADNet https://github.com/huangzih/WADNet

[58] Stone M 1974 Cross-validatory choice and assessment of statistical predictions *J. R. Stat. Soc.* B **36** 111–33

[59] Mclnnes L, Healy J and Melville J 2020 UMAP: uniform manifold approximation and projection for dimension reduction (arXiv:1802.03426)