Measuring the rogue wave pattern triggered from Gaussian perturbations by deep learning

Liwen Zou [1,2] XinHang Luo [2] Delu Zeng [3,*] Liming Ling [2] and Li-Chen Zhao[4,5]

[1]*Department of Mathematics, Nanjing University, Nanjing 210008, China*
[2]*School of Mathematics, South China University of Technology, Guangzhou 510640, China*
[3]*School of Electronics and Information Engineering, South China University of Technology, Guangzhou 510640, China*
[4]*School of Physics, Northwest University, Xi'an 710127, China*
[5]*Peng Huanwu Center for Fundamental Theory, Xi'an 710127, China*

Weak Gaussian perturbations on a plane wave background could trigger lots of rogue waves (RWs), due to modulational instability. Numerical simulations showed that these RWs seemed to have similar unit structure. However, to the best of our knowledge, there are no relative results to prove that these RWs have the similar patterns for different perturbations, partly due to that it is hard to measure the RW pattern automatically. In this work, we address these problems from the perspective of computer vision via using deep neural networks. We propose a rogue wave detection network (RWD-Net) model to automatically and accurately detect RWs in the images, which directly indicates they have the similar computer vision patterns. For this purpose, we herein meanwhile have designed and release the corresponding dataset, termed as rogue wave dataset-10K (RWD-10K), which has 10 191 RW images with bounding box annotations for each RW unit. In our detection experiments, we get 99.29% average precision on the test splits of the proposed dataset. Finally, we derive our metric, termed as the density of RW units, to characterize the evolution of Gaussian perturbations and obtain the statistical results on them.

I. INTRODUCTION

The focusing nonlinear Schrödinger equation (NLSE)

$$iu_t + \tfrac{1}{2}u_{xx} + |u|^2 u = 0 \tag{1}$$

is a universal model in nonlinear sciences, such as the deep water wave [1], nonlinear optics [2–4], Bose-Einstein condensate [5], and even finance [6]. The rogue wave (RW) solution or the Peregrine soliton is a typical exact solution for this model, which is related with the modulational instability (MI) [7,8]. Even though the NLSE is an integrable model and possesses lot of exact solutions (solitons, multisolitons, breathers, rogue waves), which had been observed in the physical experiments. Lots of rogue wave patterns had been discovered and analyzed based on analytic solutions [9,10]. However, it is crucial to consider the general initial data problem for the NLSE, since many different waves can be generated by MI on a plane wave background.

The optical rogue waves are observed in the integrable turbulence for the NLSE [11], which is related to random initial data. Recent studies also show that RW patterns and integrable turbulence are observed in the soliton gas, for which the statistics on the kinetic, potential energies, and other features are performed to characterize integrable turbulence [12]. Starting from the stochastic perturbation on the nonzero background, the maximum peaks and probability density functions on the intensity are analyzed by the numerical method [13], which provides the explanations on the probability of the appearance of rogue waves in a chaotic wave state. In the previous studies,

the turbulence is measured by the statistics. If we depart from arbitrary initial data on the plane wave background, then the chaotic structures can be observed from the numeric simulation due to the development of MI [14–17], in which the weak units similar as the Peregrine solitons can be detected clearly. Thus, it is naturally to measure the RW pattern and compute the density of RW by the RWs number. However, the bijection relation between RWs and the scattering data has not been established, in contrast to soliton density in soliton gas, which is well established by inverse scattering theory [12]. It should be pointed out that inverse scattering method [17,18] can construct the full nonlinear spectrum for the integrable equations by the scattering analysis on the vanishing and nonvanishing backgrounds, but it cannot be utilized to deal with nonintegrable systems. Therefore, universal methods to be utilized to analyze the arbitrary initial data problem for the partial differential equation are highly desirable.

The machine learning method [19] was used to analyze the extreme events in optical fiber MI [20], in which the intensity and spectral intensity are analyzed by the supervised and unsupervised learning method. The machine learning algorithms, which involve the k-nearest neighbors, support vector machine and artificial neural networks, were also used for predicting the amplitude in the chaotic lase pulse [21]. In Ref. [22], the extreme events are predicted by the deep neural networks in a truncated Korteweg-de Vries (tKdV) statistical framework. The neural network uses a relative entropy loss function to calibrate the closeness between the target and the network output as distribution functions, so that the crucial shape of the model solutions is captured instead of a pointwise fitting in the turbulent output field values. An empirical partition function are also used to calibrate the output data under a

---

*dlzeng@scut.edu.cn

combination of empirical partition functions emphasizing the large positive and negative values in the model prediction so that the main features in the solutions are further emphasized. Other relative statistic learning methods to the extreme events are performed by the literature [23–26]. Object detection based on deep learning is widely applied to various scientific fields, e.g., Tao *et al.* combined microscopy observation with artificial neural networks (ANNs) and realized by machine learning algorithms for the study of starch gelatinization [27]. A numerical method based on the neural network can be used to long time simulation of rogue waves and other nonlinear waves having the MI [28]. Recent developed machine learning method provides possibilities to measure the RW pattern and compute the density of RWs.

In this work, we propose our deep neural networks (DNNs) model to automatically and accurately detect the RWs in the images. We have designed the related dataset, which has 10 191 RW images with bounding box annotations for each RW unit. We further derive our metric, the density of RW units (DRW), to characterize the evolution of Gaussian perturbations, and finally give the statistical characters on them.

II. PRELIMINARY

It has been shown that local weak perturbations can generate RWs [29,30]. We therefore consider an initial data $u(x, 0) = 1 + p(x)$ [where $p(x)$ is a white noise or weak localized perturbation] to investigate the evolution patterns containing RWs. By using the fourth-order integrating-factor method for solving the NLSE [31], we see that the initial data can generate lots of RWs with time evolution. Interestingly, each of them has the similar shape as the fundamental RW (Peregrine breather), which has the form [32,33]

$$u_{\mathrm{P}}(x, t) = \left[1 - \frac{4(1 + 2it)}{1 + 4(x^2 + t^2)}\right] e^{it}, \quad (2)$$

whose temporal-spatial pattern is shown in Fig. 1(a). However, there is no systematic evidence to illustrate the results, which is an open problem for us up to now. In this paper, we would like to solve this problem by the artificial neural network. We choose an initial data with usual Gaussian form perturbations [29],

$$u(x, 0) = 1 + 2\varepsilon^{-1} e^{-\frac{x^2}{\mu^2}}, \quad (3)$$

to demonstrate our results, since Gaussian pulse exists widely in optical systems and it is very easy to be operated in experiments [30]. The corresponding Lax spectrum of the initial data was found to locate at the pure imaginary axis [34]. The high-order RWs can be generated by the multi-Gaussian perturbations on a plane wave background [30]. Many fundamental RWs can be excited from one weak Gaussian perturbation, which can be understood by the resonant perturbation in MI regime [29]. But many other waves generated by the nonideal resonant perturbation [17], of which the local structure is similar to the fundamental RW. These waves can be seen as RWs in the local spatial-temporal region. The boundary lines for the RW-like waves region in Fig. 1(a) were explained well by nonlinear stage of MI [17]. But the number of RW or its density feature is still needed to be further investigated. To measure the RW pattern evolved from the
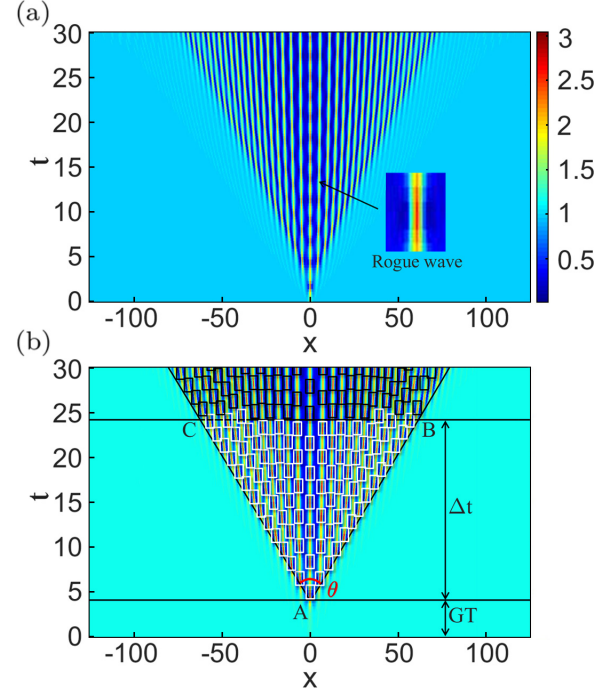

FIG. 1. (a) Rouge waves on rogue wave pattern images. (b) Some terms defined in this paper.

Gaussian perturbation, we define the density of rogue wave (DRW) as follows:

$$\mathrm{DRW} = \frac{N(\Delta t, \varepsilon, \mu)}{S_{\Delta ABC}} = \frac{N(\Delta t, \varepsilon, \mu)\cot\frac{\theta(\varepsilon, \mu)}{2}}{(\Delta t)^2}, \quad (4)$$

where $N(\Delta t, \varepsilon, \mu)$ denotes the number of RWs that appear during the period from $t = GT$ to $t = GT + \Delta t$. We compute N according to the number of RW bounding boxes in ΔABC and incomplete boxes are computed by the proportion of their areas inside ΔABC, for instance, half of a box is computed as 0.5. $S_{\Delta ABC}$ is the area of the triangular ΔABC. It should be pointed out that GT is the time interval between the beginning moment and the time when the first RW peak appears. Figure 1(b) shows the details of some terms defined in this paper. Next, we introduce our DNNs model to detect RW pattern.

III. RWD-Net

Image recognition and object detection by deep learning methods had gotten great development in recent years. Residual network (ResNet) was designed to learn the residual representation between the layers, which not only train the network easier but also improve the accuracy a lot [35]. Even now ResNet is still one of the most fundamental feature extraction networks in the computer vision field. Lin *et al.* proposed the feature pyramid network (FPN) which improves the accuracy of object detection, especially in the detection of weak objects, without adding additional calculations by extracting and fusing multiscale feature map of RW images [36]. Based on the ResNet and FPN, Lin *et al.* proposed their RetinaNet with the focal loss function replacing the traditional cross-entropy loss to solve the unbalance problem between positive and negative objects, which improves both the accuracy and speed [37]. Goldman *et al.* improve the detection
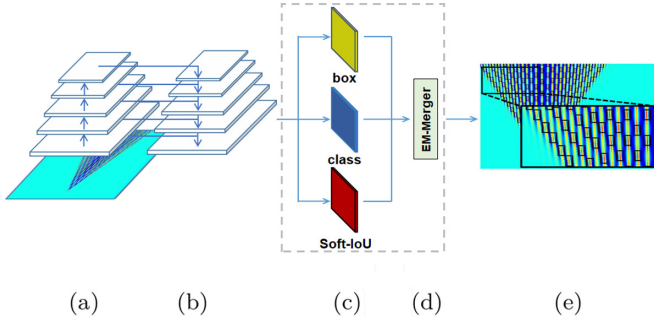
FIG. 2. Structure of our RWD-Net. (a) ResNet-101; (b) FPN; (c) the box, class, and the soft-IoU subnets; (d) EM-Merge unit; (e) detection results.

in densely packed scenes by adding the soft-IoU layer and the expectation-maximization merge (EM-Merge) unit to the RetinaNet [38].

Inspired by the studies above, we propose a different methodology, namely, rogue wave detection network (RWD-Net), developed from RetinaNet improved by Goldman *et al.* [38], which is designed to detect the RW regions in the images and get the number and distribution of them to measure the rogue wave pattern. The architecture of the network is shown in Fig. 2, which mainly consists of three parts: feature extractor or backbone, detector and post-processing unit. The backbone is composed of ResNet and FPN to get five feature maps of RW images under different scales. Each instance in the feature map is covered with nine priori boxes, called anchors, each of which has different sizes and shapes and will be sent to the next three subnets—class, box, and soft-IoU subnet—to be trained. Mathematically, let the pixel values of a RW image be denoted by $I \in \mathbb{R}^{W \times H \times L}, W, H, L \in \mathbb{Z}^{+}$ denotes the width, height, and length of the image, respectively, and the training data is denoted by $\mathcal{D}^{\mathrm{tr}} = \{(I^l, Y^l)\}_{l=1}^{L}$, where $Y^l \in \mathbb{R}^{5 \times N}$ is the set of offset and label $(d_a^x, d_a^y, d_a^w, d_a^h, y_a)$ of each anchor covering I^l and N denoted the number of anchors in image I^l. Our goal is to predict the bounding box matrices $\hat{Y}^l \in \mathbb{R}^{5 \times N}$ of image I^l. A RW image I is first fed into the backbone consists of ResNet and FPN, parameterized by Θ, to produce a feature map $F(I; \Theta) \in \mathbb{R}^{W' \times H' \times L' \times C}$, where C is the number of feature channels. The feature vector $f_a(I; \Theta) \in \mathbb{R}^C$ at each spatial location a on the feature map $F(I; \Theta)$ represents the extracted feature of anchor a on the image I by the ResNet and FPN model. After the feature extracting, there are three subnets of our proposed RWD-Net: the class subnet for RW and background classification of each anchor feature $f_a(I; \Theta)$, the box subnet for the bounding box regression and the soft-IoU subnet for predicting the soft-IoU between the prediction boxes and the ground truth.

In the class subnet, we define the confidence level that an anchor instance $f_a(I; \Theta)$ covers a RW region as

$$p_a(I; \Theta, w_c) = \frac{1}{1 + \exp\left[-w_c^T f_a(I; \Theta)\right]}, \quad (5)$$

where $w_c \in \mathbb{R}^C$ is the parameter of the class subnet, then we minimize the focal loss on the training data with annotations:

$$
\begin{aligned}
l_{\mathrm{class}}&(I^l, Y^l; \Theta, w_c) \\
&= -\sum_{a \in A} \{y_a^l \alpha [1 - p_a(I; \Theta, w_c)^\gamma] \log p_a(I; \Theta, w_c) \\
&\quad + \left(1 - y_a^l\right)(1-\alpha) p_a(I; \Theta, w_c)^\gamma \log[1 - p_a(I; \Theta, w_c)]\},
\end{aligned}
$$
(6)

where y_a^l is the label annotation of anchor a obtained directly from Y^l and A is the set of all anchors covering the image I^l. We set $\alpha = 0.25$ and $\gamma = 2$ in our implementation. And the loss function for training the class subnet over the whole traing set D^{tr} is defined by

$$L_{\mathrm{class}}(D^{\mathrm{tr}}; \Theta, w_c) = \frac{1}{|D^{\mathrm{tr}}|} \Sigma_{(Z^l, Y^l) \in D^{\mathrm{tr}}} l_{\mathrm{class}}(I^l, Y^l; \Theta, w_c).$$
(7)

We now introduce the box subnet module of our RWD-Net. The model predicts the offset between positive anchors and their belonged ground truth as

$$\hat{d}_a = \left(\hat{d}_a^x, \hat{d}_a^y, \hat{d}_a^w, \hat{d}_a^h\right) = w_b^T f_a(I; \Theta), \quad (8)$$

where $w_b \in \mathbb{R}^{C \times 4}$ is the parameter of box subnet. Then we use the smooth $-L_1$ loss function [39] on the image I^l with box annotations to learn the box subnet:

$$l_{\mathrm{box}}(I; \Theta, w_b) = \Sigma_{a \in A} y_a Q_a, \quad (9)$$

where

$$Q_a = \begin{cases} \frac{1}{2}\|\hat{d}_a - d_a\|_2^2, & \text{if } \|\hat{d}_a - d_a\|_2 < 1, \\ \|\hat{d}_a - d_a\|_2 - \frac{1}{2}, & \text{otherwise}, \end{cases} \quad (10)$$

and $d_a = (d_a^x, d_a^y, d_a^w, d_a^h)$ is the annotation of anchors' offset. Now we can write down the loss function on the training data D^{tr}:

$$L_{\mathrm{box}}(D^{\mathrm{tr}}; \Theta, w_b) = \frac{1}{|D^{\mathrm{tr}}|} \sum_{(Z, Y) \in D^{\mathrm{tr}}} l_{\mathrm{box}}(I; \Theta, w_b). \quad (11)$$

The intersection over union (IoU) defined as

$$\mathrm{IoU} = \frac{I(X)}{U(X)} \quad (12)$$

is used as the evaluation metric in the detection, where $I(X)$ and $U(X)$ denote the intersection and union area between the prediction and ground truth respectively. The class subnet gives the confidence and the soft-IoU subnet gives the soft-IoU which is just the IoU predicted by the model and is not the real IoU. The soft-IoU between the prediction boxes and their belonged ground truth is defined as

$$s_a(I; \Theta, w_s) = w_s^T f_a(I; \Theta), \quad (13)$$

where $w_s \in \mathbb{R}^C$ is the parameter of the soft-loU subnet. Then we minimize the cross-entropy loss function on the image I^l:

$$
\begin{aligned}
l_{\mathrm{s\text{-}iou}}(I^l, Y^l; \Theta, w_s) = -\sum_{a \in A} \{\mathrm{IoU}_a \log s_a(I; \Theta, w_s) \\
+ (1 - \mathrm{IoU}_a) \log[1 - s_a(I; \Theta, w_s)]\},
\end{aligned}
$$
(14)

where IoU_a is the IoU ground truth between anchor in location a and their belonged ground truth. And the whole loss in this soft-IoU subnet for training data D^{tr} defined as

$$L_{\text{s-iou}}(D^{\mathrm{tr}}; \Theta, w_s) = \frac{1}{|D^{\mathrm{tr}}|} \sum_{(Z^l, Y^l) \in D^{tr}} l_{\text{s-iou}}(I^l, Y^l; \Theta, w_s). \quad (15)$$

Finally, we write down the overall loss function for training our RWD-Net:

$$\begin{aligned} L_{\text{RWD}}(D^{\mathrm{tr}}; &\Theta, w_c, w_b, w_s) \\ &= L_{\text{class}}(D^{\mathrm{tr}}; \Theta, w_c) + L_{\text{box}}(D^{\mathrm{tr}}; \Theta, w_b) \\ &\quad + L_{\text{s-iou}}(D^{\mathrm{tr}}; \Theta, w_s). \end{aligned} \quad (16)$$

All parameters are jointly optimized during network training. The optimized parameters are obtained by

$$(\Theta^*, w_c^*, w_b^*, w_s^*) = \underset{\Theta, w_c, w_b, w_s}{\arg\min} L_{\text{RWD}}(D^{tr}; \Theta, w_c, w_b, w_s). \quad (17)$$

And we minimize the overall loss function by stochastic gradient descent method.

Given a testing image I, whether its anchor in location a covers RW regions is determined by

$$\hat{y}_a = \begin{cases} 1, & \text{if } p_a(I; \Theta^*, w_c^*) \geqslant 0.5, \\ 0, & \text{otherwise.} \end{cases} \quad (18)$$

If $\hat{y}_a = 1$, then we further need to do the bounding box regression for anchor in a:

$$\hat{d}_a = (w_b^*)^{\mathrm{T}} f_a(I; \Theta^*). \quad (19)$$

With $\hat{y}_a$ and $\hat{d}_a$ above, we obtain the prediction $\hat{Y}$. But it does not end, RWD-Net also predicts the soft-IoU between each positive anchor and the ground truth:

$$\hat{s}_a = (w_s^*)^{\mathrm{T}} f_a(I; \Theta^*), \quad (20)$$

which is used for expectation-maximization-merge (EM-Merge) unit next to obtain the final predicting boxes. The details about the EM-Merge unit can be seen in Ref. [38].

IV. EXPERIMENTAL RESULTS

In this work, we provide our big dataset, termed as rogue wave dataset-10K (RWD-10K) containing 10 191 images of RW pattern. We propose an efficient semi-automatic method, called peak search method, to achieve fast and accurate pre-detection to RWs instead of manual labeling. It is designed to determine the approximate location of the peak for each RW filtering out the local maximum points on the numerical solution matrices such that one peak point on the matrix corresponds to one RW on the image. Then, all of these peak points will be expanded into bounding boxes. It should be noted that the sizes and locations of these bounding boxes will be manually refined due to the errors caused by the dimensional difference between the numerical solution matrices and the image matrices. By peak search and additional minor corrections, we can efficiently build the RWD-10K dataset. The details about peak search are given in the Appendix A.

We variate the parameters μ (from 0.5 to 50 with the interval 0.5) and ε (from 11 to 110 with the interval 1) in
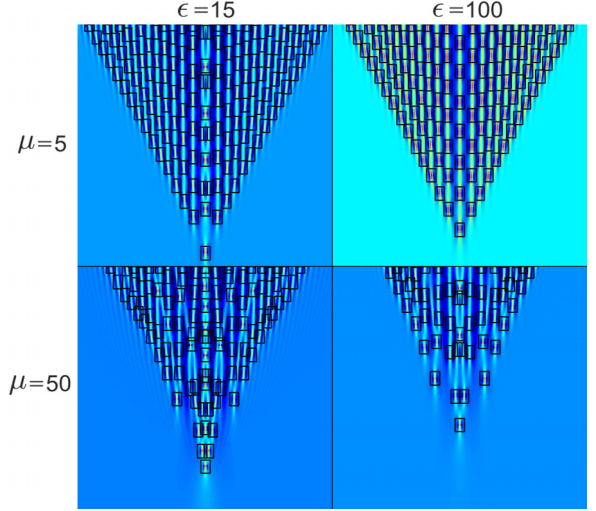

FIG. 3. The detection results of the images we randomly selected from the test splits of RWD-10K using our RWD-Net.

the initial NLSE (1) to generate 10 000 images. Additionally, to study the variation of DRW or other terms for μ and ε, we make other 191 images with larger value range of μ and ε to complete the following statistical experiments. Then we get the pseudo annotations by the peak search algorithm and refine manually so that we assembled our big benchmark RWD-10K. Each image corresponds a parameter two-tuple (ε, μ) in the initial equation. We focus on such settings for two reasons. First, every image in our dataset has its physical meaning. Second, by detecting those images we can capture the pattern similarity of computer vision among them and get certain statistical results about the distribution of these rogue waves [e.g., the evolution that the number of the rogue wave changes with the parameters in the initial data Eq. (1)].

In our detection experiment, the RWD-10K dataset is partitioned into train, validate and test splits. Training consists of 60% of the images (6071 images) and their associated 510 431 bounding boxes; 20% of the images (2024 images), are used for validation (with their 172 881 bounding boxes). The rest 2024 images (175 968 bounding boxes) were used for testing. Images were selected randomly, ensuring that the same RW from the same image does not appear in more than one of these subsets. We present our RWD-10K dataset open for public studies and you can get more details at Ref. [40].

The details about our experiment settings are shown in Appendix C. After 20 epochs training, we get the average precision (AP) of 99.29% in the RW detection experiment, which shows that we successfully capture the RW pattern similarity from the perspective of computer vision. Figure 3 shows the detection results of the images whose parameters are randomly chosen from the test dataset. We can see that we shot almost every RW in the detection. Additionally, it is easy to observe that the distribution of these RW patterns are discriminative under different initial data. The RW patterns depend on initial perturbation forms [41–44], which can be understood well by the full nonlinear spectrum for the integrable equations [13,17]. We can see that the RW patterns in Fig. 3 vary greatly with different amplitudes and widths of the Gaussian perturbations. The Fourier analysis indicates that
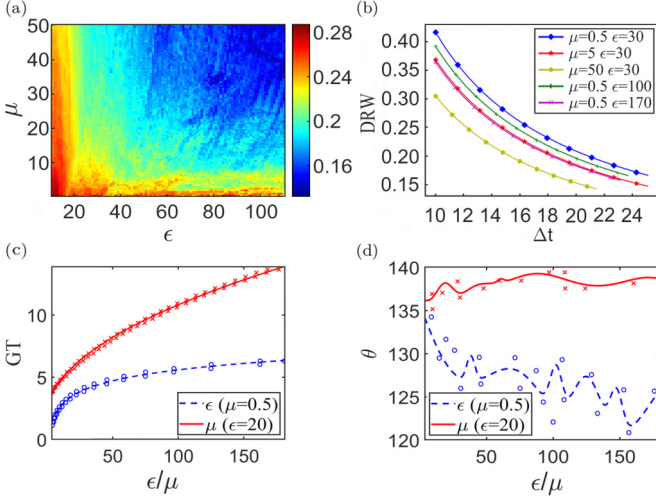
FIG. 4. Results of measuring the rogue wave pattern. (a) The map of DRW with the change of ε and μ when $\Delta t = 15$; (b) the variations of DRW for different Δt; (c) the relation between GT and ε or μ, in which the solid and blue lines are the fitting curves and the cross or circle points are the data points; (d) the θ values with different ε or μ values.

the wave vector components of perturbation depend directly on the perturbation wave form. High-order excitations can emerge for larger width and amplitude cases. The quantitative relation between the patterns and initial conditions is hard to be described analytically, even for integrable models. Many efforts are still paid to address this kind of problems [18]. We would like to characterize some properties of the RW patterns based on the given deep learning techniques.

V. MEASURING THE ROGUE WAVE PATTERN

Now we try to measure the RW pattern under different Gaussian initial data by the trained RWD-Net. We use the DRW defined in Eq. (4) to quantify the RW pattern and we can also get the variation of GT and θ based on the detection results of our RWD-Net model. The results are shown in the Fig. 4 and initial data is given in Eq. (3).

Figure 4(a) is the distribution of the DRW with respect to the parameter ε and μ when Δt in Eq. (4) is fixed as 15. Meanwhile, it is easy to find that the value of DRW decreases from bottom left to top right of the image, which means the value of corresponding DRW will decline in the spatiotemporal regions when ε and μ increase. These characters can be understood by a fact that each localized wave is closer to Peregrine RW for initial condition with larger ε and μ, since initial perturbation with much larger ε and μ approaches more closely to the resonant condition with background [29,45].

In Fig. 4(b), we show the variations of DRW as Δt varies when ε and μ are fixed at different constant values. Since the initial data with different parameters ε and μ will yield different GT, then the range of Δt is different for the fixed image size. According to the figure, it can be seen that when ε and μ are fixed, DRW will decrease as a smooth curve as Δt increase. At the same time, if we only consider the variation of ε for fixed μ, we can see that when ε increases from 30

TABLE I. The parameters of fitting functions Eq. (21).

μ	a	b
10	1.581	-0.804
30	2.136	-0.766
50	2.563	-0.819

to 170, the corresponding DRW value will decrease in turn. Also, if we only consider the variation of μ for the fixed ε, the DRW value will still show a downward trend as the value of μ increases from 0 to 50. Those results are in line with the ones of Fig. 4(a).

In Fig. 4(c), we show the relation between GT and ε (μ), and we respectively use the exponential function with base 0.5 and logarithmic function to fit the curves of GT about the change of ε and μ. In the Appendix B, we compare different fitting functions, and list the results to show that it is more reasonable to use the following two functions. For fixed μ, we have the relation between GT and ε as

$$GT = a \ln(\varepsilon) + b, \tag{21}$$

where a and b are the fitting parameters. For fixed ε, we have the relation between GT and μ as

$$GT = c\sqrt{\mu} + d, \tag{22}$$

where c and d are the fitting parameters. All the corresponding fitting parameters are given in Tables I and II. In Fig. 4(c), we can see that the fitting curves indeed agree well with the numerical results. These simple fitting relations are helpful for further theoretical discussions on the complicated initial problems [46,47]. When the nonlinear spectrum of initial perturbation tends to be degenerate or more complex, the quantitative properties are indeed hard to be explored even for integrable models. The deep learning methods could be used to explore these implicit relations through fitting, which can also be applied to nonintegrable systems. These relations can be checked in real experiments [42–44], and could also provide some hints to develop its theoretical analysis.

Figure 4(d) demonstrates the variation of θ with respect to the parameters ε or μ respectively. Fixed the parameter $\epsilon = 20$, it is shown that the angle θ almost steadily varies between 135° and 140°. But if we fix the parameter $\mu = 0.5$, it is seen that the angle θ will decrease as ϵ increases with the range approximately between 120° and 135°. These results show the amplitude parameter ϵ will effect the angle θ. Additional statistics are given in the Appendix B.

TABLE II. The parameters of fitting functions Eq. (22).

ε	c	d
20	0.682	1.900
50	0.890	2.760
100	1.042	3.382

Algorithm 1: The procedure of peak search algorithm.

Input: The numerical solution matrix, $M = (a_{ij})$;
   The initical location map of the peak points,
   $N = (b_{ij})$ where $b_{ij} = 0$ at first;
Output: The final location map of the peak points,
   $N = (b_{ij})$;
1 If $a_{ij} >= \eta l$, then we set $b_{ij} = 1$, where η represents the peak factor which is a constant we specified in advance (we set $\eta = 1.7$ in our experiment), and l represents the level set value corresponding to the level ground plane.
2 For each $b_{ij} = 1$, if a_{ij} is the highest value point within the distance r, then we set $b_{ij} = 2$, where r denotes the comparison radius we specified in advance (we set $r = 2$ in our experiment).
3 We map the numerical matrix coordinates of each peak point satisfying $b_{ij} = 2$ into a weak bounding box on the image, and then zoom them in size of 20×20;
4 return $N = (b_{ij})$;

VI. CONCLUSIONS

In this paper, we propose an automatic, fast and accurate framework to measure RW pattern via deep learning. Recently, Guo *et al.* [48] described their automated classification and positioning system for identifying localized excitations in atomic Bose-Einstein condensates by deep convolutional neural networks to eliminate the need for human image examination. They implement the detection by CNN-based image classification and least-squares fitting-based position regression. But we focus on the efficient detection of RWs on numerical solutions images of NLSEs entirely by an end-to-end deep learning framework to capture the pattern similarity of computer vision. Besides, we also get the statistics of these RW patterns based on the detection results so that we are able to predict the time interval and region angle in real cases. Our proposed term DRW can intuitively measure the the Gaussian perturbations in our experiments. Our method can be generalized for the other integrable systems with MI [49–61], and nonintegrable systems.

For multi-Gaussian perturbation or other weak localized forms in initial data Eq. (3) (i.e., there are two RW valleys in the images), can we still classify these rogue wave patterns and get their corresponding distribution? We leave this problem to the future work. We also hope that this work will initiate more innovative efforts in this field.

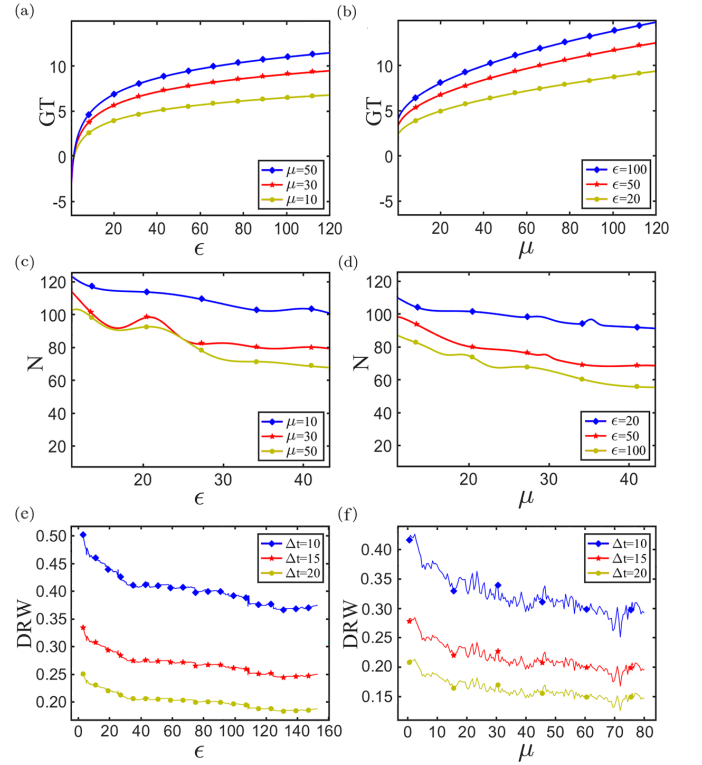FIG. 5. (a, b) Variation of GT for ε or μ when the other is fixed; (c, d) Variation of number N of rogue wave in the images for ε or μ when the other is fixed; (e) Variation of DRW for ε under different Δt; (f) Variation of DRW for μ under different Δt;

APPENDIX

In the Appendix, we give the outline of our proposed peak search method, the description of our RWD-10K dataset and the details about the setting and results of our rogue wave detection experiments.

1. Peak search

Peak search is our proposed fast and cluster-based algorithm to generate the rogue wave images with pseudo annotations. Through our observations, selecting points with larger modulus length from the numerical solution matrix can well filter out the crests of rogue waves. Therefore, we propose the following algorithm in Algorithm 1 to realize the
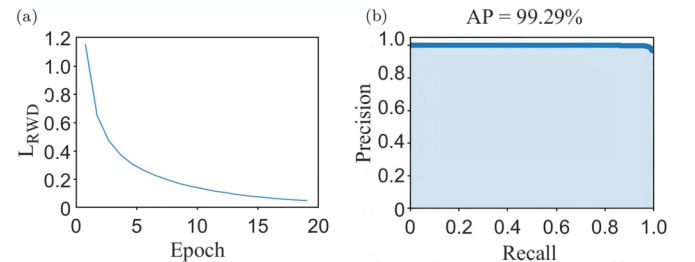

FIG. 6. (a) Evolution of the overall loss L_{RWD}; (b) the precision-recall curve of detection in test splits.

goal to predetect the rogue waves on images. What needs to be pointed out here is that the peak search method only works with rogue wave matrices, which means it cannot replace our RWD-Net to complete the task of detect rogue wave patterns when there are only images without any numerical solution matrices.

2. Additional statistics

We compare different fitting functions of GT, the variation curves are shown in Fig. 5(a) and the fitting parameters are shown in Tables I and II. Besides, we give the curves showing the numbers of rogue waves in the images changing with ε and μ in 5(b). Lastly, the variation of curves for DRW with ε and μ are given in Figs. 5(c) and 5(d) under different Δt.

3. Experimental details

In this part, we show some details and results about our RWD-Net detection experiment. We use the pretraining weight file of ResNet-101 obtained from Ref. [62] as the initial weight. The learning rate is set to 1×10^5, the batch size is set to 4 and the epoch number is set to 20. Our whole training experiment runs on the machine using an NVidia TITAN X GPU with 12 GB of GDDR5 memory.

We record the overall loss changes of the training set during the training process as shown in Fig. 6(a). At the end of training, the overall loss L_{RWD} of the model achieves 0.071 on the train splits. Based on the trained model above, we test it on the test splits. Finally, we get 99.29% AP at the threshold IoU set as 0.5 on the test splits and the Precision-Recall curve is shown in Fig. 6(b).

[1] A. Chabchoub, N. P. Hoffmann, and N. Akhmediev, Phys. Rev. Lett. **106**, 204502 (2011).

[2] D. R. Solli, C. Ropers, P. Koonath, and B. Jalali, Nature (London) **450**, 1054 (2007).

[3] B. Kibler, J. Fatome, C. Finot, G. Millot, F. Dias, G. Genty, N. Akhmediev, and J. M. Dudley, Nat. Phys. **6**, 790 (2010).

[4] J. M. Dudley, F. Dias, M. Erkintalo, and G. Genty, Nat. Photon. **8**, 755 (2014).

[5] Yu. V. Bludov, V. V. Konotop, and N. Akhmediev, Phys. Rev. A **80**, 033610 (2009); Z. Yan, V. V. Konotop, and N. Akhmediev, Phys. Rev. E **82**, 036610 (2010).

[6] V. G. Ivancevic, Cogn. Comput. **2**, 17 (2010); Z. Yan, Commun. Theor. Phys. **54**, 947 (2010); Phys. Lett. A **375**, 4274 (2011).

[7] C. Kharif and E. Pelinovsky, Eur. J. Mech. B/Fluids **22**, 603 (2003); P. K. Shukla, I. Kourakis, B. Eliasson, M. Marklund, L. Stenflo, Phys. Rev. Lett. **97**, 094501 (2006); C. Kharif, E. Pelinovsky, and A. Slunyaev, *Rogue Waves in the Ocean* (Springer, New York, 2009); M. Onorato *et al.*, Phys. Rep. **528**, 47 (2013).

[8] M. Onorato, A. R. Osborne, and M. Serio, Phys. Rev. Lett. **96**, 014503 (2006).

[9] D. J. Kedziora, A. Ankiewicz, and N. Akhmediev, Phys. Rev. E **88**, 013207 (2013).

[10] B. Yang and J. Yang, Physica D **419**, 132850 (2021).

[11] P. Walczak, S. Randoux, and P. Suret, Phys. Rev. Lett. **114**, 143903 (2015).

[12] A. A. Gelash, Phys. Rev. E **97**, 022208 (2018); A. A. Gelash and D. S. Agafontsev, *ibid.* **98**, 042210 (2018); A. Gelash, D. Agafontsev, V. Zakharov, G. El, S. Randoux, and P. Suret, Phys. Rev. Lett. **123**(23), 234102 (2019).

[13] J. M. Soto-Crespo, N. Devine, and N. Akhmediev, Phys. Rev. Lett. **116**(10), 103901 (2016).

[14] T. B. Benjamin, Proc. R. Soc. A **299**, 59 (1967); T. B. Benjamin and J. E. Feir, J. Fluid Mech. **27**, 417 (1967).

[15] F. Baronio, M. Conforti, A. Degasperis, S. Lombardo, M. Onorato, and S. Wabnitz, Phys. Rev. Lett. **113**, 034101 (2014).

[16] V. E. Zakharov and A. A. Gelash, Phys. Rev. Lett. **111**, 054101 (2013).

[17] G. Biondini and D. Mantzavinos, Phys. Rev. Lett. **116**, 043902 (2016).

[18] G. Biondini, S. Li, and D. Mantzavinos, Commun. Math. Phys. **382**, 1495 (2021); G. Biondini, D. Mantzavinos, Commun. Pure Appl. Math. **70**, 2300 (2017).

[19] M. I. Jordan and T. M. Mitchell, Science **349**, 255 (2015).

[20] M. Närhi, L. Salmela, J. Toivonen, C. Billet, J. M. Dudley, and G. Genty, Nat. Commun. **9**, 4923 (2018).

[21] P. Amil, M. C. Soriano, and C. Masoller, Chaos **29**, 113111 (2019).

[22] D. Qi and A. J. Majda, Proc. Natl. Acad. Sci. **117**, 52 (2020).

[23] M. A. Mohamad and T. P. Sapsis, Proc. Natl. Acad. Sci. **115**, 11138 (2018).

[24] G. Dematteis, T. Grafke, and E. Vanden-Eijnden, Proc. Natl. Acad. Sci. **115**, 855 (2018).

[25] A. J. Majda, M. N. J. Moore, and D. Qi, Proc. Natl. Acad. Sci. **116**, 3982 (2019).

[26] L. Salmela, C. Lapre, J. M. Dudley, and G. Genty, Sci. Rep. **10**, 9596 (2020).

[27] J. X. Tao, J. B. Huang, L. Yu, Z. K. Li, H. S. Liu, B. Yuan, and D. L. Zeng, Food Hydrocoll. **74**, 151 (2018).

[28] R. Q. Wang, L. M. Ling, D. L. Zeng, and B. F. Feng, Commun. Nonlinear Sci. Numer. Simul. **101**, 105896 (2021).

[29] L. C. Zhao and L. M. Ling, J. Opt. Soc. Am. B **33**, 850 (2016).

[30] P. Gao, L. C. Zhao, Z. Y. Yang, X. H. Li, and W. L. Yang, Opt. Lett. **45**, 2399 (2020).

[31] J. Yang, *Nonlinear Waves in Integrable and Nonintegrable Systems* (SIAM, 2010).

[32] D. Peregrine and J. Aust, Math. Soc. B, Appl. Math. **25**, 16 (1983).

[33] N. Akhmediev, A. Ankiewicz, and J. M. Soto-Crespo, Phys. Rev. E **80**, 026601 (2009).

[34] G. Biondini and X. Luo, Phys. Lett. A **382**, 2632 (2018).

[35] K. He, X. Zhang, S. Ren, J. Sun and Microsoft Research, IEEE Conf. Comput. Vis. Pattern Recognit. 770 (2016).

[36] T. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, IEEE Conf. Comput. Vis. Pattern Recognit. 2117 (2017).

[37] T. Lin, P. Goyal, R. Girshick, K. He, and P. Dollar, IEEE Conf. Comput. Vis. Pattern Recognit. 2980 (2017).

[38] E. Goldman, R. Herzig, A. Eisenschtat, J. Goldberger, and T. Hassner, IEEE Conf. Comput. Vis. Pattern Recognit. 5227 (2019).

[39] R. Girshick, IEEE Int. Conf. Comput. Vis. 1440 (2015).

[40] https://github.com/ZouLiwen-1999/RogueWave.

[41] R. H. J. Grimshaw and A. Tovbis, Proc. R. Soc. A **469**, 20130094 (2013).

[42] M. Erkintalo, K. Hammani, B. Kibler, C. Finot, N. Akhmediev, J. M. Dudley, and G. Genty, Phys. Rev. Lett. **107**, 253901 (2011).

[43] B. Kibler, A. Chabchoub, A. Gelash, N. Akhmediev, and V. E. Zakharov, Phys. Rev. X **5**, 041026 (2015).

[44] A. Chabchoub, T. Waseda, M. Klein, S. Trillo, and M. Onorato, Phys. Rev. Fluid **5**, 114801 (2020).

[45] L. M. Ling, L. C. Zhao, Z. Y. Yang, and B. Guo, Phys. Rev. E **96**, 022211 (2017).

[46] T. Congy, G. El, and G. Roberti, Phys. Rev. E **103**, 042201 (2021).

[47] M. Bertola, T. Grava, and G. Orsatti, arXiv:2112.05985 (2021).

[48] S. Guo, A. R. Fritsch, C. Greenberg *et al.*, Mach. Learn.: Sci. Technol. **2**, 035020 (2021).

[49] S. V. Manakov, Zh. Eksp. Teor. Fiz. **67**, 543 (1974) [Sov. Phys. JETP **38**, 248 (1974)].

[50] B. L. Guo and L. M. Ling, Chin. Phys. Lett. **28**, 110202 (2011); L. C. Zhao and J. Liu, J. Opt. Soc. Am. B **29**, 3119 (2012); Phys. Rev. E **87**, 013201 (2013).

[51] A. Chabchoub and M. Fink, Phys. Rev. Lett. **112**, 124101 (2014); A. Przadka, S. Feat, P. Petitjeans, V. Pagneux, A. Maurel, and M. Fink, *ibid.* **109**, 064501 (2012).

[52] F. Baronio, A. Degasperis, M. Conforti, and S. Wabnitz, Phys. Rev. Lett. **109**, 044102 (2012); F. Baronio, M. Conforti, A. Degasperis, and S. Lombardo, *ibid.* **111**, 114101 (2013); S. Chen, Y. Ye, J. M. Soto-Crespo, Ph. Grelu, and F. Baronio, *ibid.* **121**, 104101 (2018).

[53] S. Chen and L. Y. Song, Phys. Rev. E **87**, 032910 (2013); S. Chen, Phys. Lett. A **378**, 2851 (2014); S. Chen and D. Mihalache, J. Phys. A **48**, 215202 (2015); S. Chen, X. M. Cai, P. Grelu, J. Soto-Crespo, S. Wabnitz, and F. Baronio, Opt. Express **24**, 5886 (2016).

[54] L. C. Zhao, L. Duan, P. Gao, and Z. Y. Yang, Europhys. Lett. **125**, 40003 (2019).

[55] A. Tikan, C. Billet, G. El, A. Tovbis, M. Bertola, T. Sylvestre, F. Gustave, S. Randoux, G. Genty, P. Suret, and J. M. Dudley, Phys. Rev. Lett. **119**, 033901 (2017).

[56] L. Li, Z. Wu, L. Wang, and J. He, Ann. Phys. **334**, 198 (2013); J. S. He, H. R. Zhang, L. H. Wang, K. Porsezian, and A. S. Fokas, Phys. Rev. E **87**, 052914 (2013).

[57] Y. V. Kartashov, V. V. Konotop, M. Modugno, and E. Ya. Sherman, Phys. Rev. Lett. **122**, 064101 (2019).

[58] G. Q. Zhang, L. M. Ling, and Z. Y. Yan, J. Nonlinear Sci. **31**, 81 (2021).

[59] Y. R. Chen, B. F. Feng, and L. M. Ling, Physica D **424**, 132954 (2021).

[60] B.-F. Feng, L. M. Ling, and D. A. Takahashi, Stud. Appl. Math. **144**, 46 (2020).

[61] Y. F. Mo, L. M. Ling, and D. L. Zeng, Phys. Lett. A **421**, 127739 (2021).

[62] https://github.com/keras-team/kerasapplications/releases/tag/resnet.