

## Research Article

# A Pinning Actor-Critic Structure-Based Algorithm for Sizing Complex-Shaped Depth Profiles in MFL Inspection with High Degree of Freedom

Zhenning Wu <sup>1,2</sup>, Yiming Deng <sup>2</sup>, and Lixing Wang <sup>3</sup>

<sup>1</sup>College of Information Science and Engineering, Northeastern University, Shenyang 110004, China

<sup>2</sup>Department of Electrical and Computational Engineering, Michigan State University, East Lansing, MI 48823, USA

<sup>3</sup>College of Computer Science and Engineering, Northeastern University, Shenyang 110004, China

Correspondence should be addressed to Zhenning Wu; wuzn2003@hotmail.com

Received 14 March 2021; Revised 1 April 2021; Accepted 15 April 2021; Published 24 April 2021

Academic Editor: Rui Wang

Copyright © 2021 Zhenning Wu et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

One of the most efficient nondestructive methods for pipeline in-line inspection is magnetic flux leakage (MFL) inspection. Estimating the size of the defect from MFL signal is one of the key problems of MFL inspection. As the inspection signal is usually contaminated by noise, sizing the defect is an ill-posed inverse problem, especially when sizing the depth as a complex shape. An actor-critic structure-based algorithm is proposed in this paper for sizing complex depth profiles. By learning with more information from the depth profile without knowing the corresponding MFL signal, the algorithm proposed saves computational costs and is robust. A pinning strategy is embedded in the reconstruction process, which highly reduces the dimension of action space. The pinning actor-critic structure (PACS) helps to make the reward for critic network more efficient when reconstructing the depth profiles with high degrees of freedom. A nonlinear FEM model is used to test the effectiveness of algorithm proposed under 20 dB noise. The results show that the algorithm reconstructs the depth profile of defects with good accuracy and is robust against noise.

## 1. Introduction

Magnetic flux leakage (MFL) is one of the most widely used NDT techniques, which has been widely used for inspection of oil and gas pipeline since the 1960s. It is efficient in finding defects caused by corrosion and mechanical damage and other metal loss defects for pipelines and storage tanks [1–3]. It is helpful to give the health condition of the working facilities to the operators, which prevents disasters to environment, industry, and human being due to the leakage of explosive or dangerous chemicals. Estimating the shape of defect is the key problem of inspection. Though MFL is efficient in finding the defects and anomalies, the reconstruction process from inspection signal to defect depth is not an easy task, as it is usually contaminated by sampling noise [4]. Reconstruction results with more details such as the detailed shape of the defect rather than only length,

width, and depth are more helpful to estimate the health condition of the tested material [5]. Among length, width, and depth, depth reconstruction is the most challenging part as it is highly ill-posed. Unfortunately, reconstructing the defect shape with details makes the ill-posed inverse problem even harder to be solved.

The solutions of MFL inverse problem could be classified as either non-model-based methods or model-based methods. Non-model-based methods solve this inverse problem by building a mapping between sampled signal and the shape of defect. Neural networks are usually used to build this mapping [6–10]. The input of this neural network can be the signal of MFL inspection, and the sizing information of the defect is set as the output. These methods are fast but highly rely on the data set used to train the neural network. The accuracy is highly impacted by the quality of training data set.

A forward physical model is involved in the model-based methods. The forward model is used to give simulated signal to a given depth profile. The simulated signal is used for comparison with the reference signal. The residual error between simulated signal and reference signal is used to give information for the iteration strategy. By minimizing the residual error, the size of the defect is repetitively computed [11–16]. Numerical models and analytical models are two categories of methods involved as forward model of MFL. Analytical models are fast but have more limitations as the model is derived with many simplifications, making it less accurate [17, 18]. Numerical model provides accurate results, but it is computationally very costly especially when a fine model is needed. The design of the iteration policy for numerical model is another problem that is hard to design. Classic methods design the policy with gradient information to minimize the residual error [19, 20]. These methods usually have some limitations to assume the shape of the defect a priori. Another kind of solution uses a mapping which is trained to replace the numerical forward model. A novel iterative method of inversion using adaptive wavelets and radial basis function neural network are proposed in [5]. A RBF neural network is used as a forward model in [21]. Heuristic methods are the third kind of solution of designing the iteration policy. Han et al. proposed a particle swarm optimization method to solve this problem [22]. Li et al. proposed a modified harmony search algorithm as the iteration policy [23]. As these heuristic methods are not deterministic, they usually need a vast amount of forward model evaluations.

Considering the state-of-the-art solutions, there are still some common problems in solving the problem of sizing the defects. First, for the non-model-based method, the mapping is trained according to the data without exploration to data not included in the training set. It makes the mapping highly rely on the distribution of the training data set. As the MFL inverse problem is ill-posed, the mapping from signal to defect profiles can also be troubled by the nonuniqueness of the mapping. Second, for the model-based method, the iteration strategy is designed based on the forward model in use and highly relies on it. For numerical model, it has high performance in simulating the inspection signal, but it is hard to build an iteration strategy based on it.

The similarity between the state-of-the-art machine learning technique of game play [24–29] and classic model-based iteration method inspires the study in this paper. Though the RL algorithm is basically a machine learning technique which needs training, the general structure has similarity with the classic iteration method which makes it possible to design an iteration strategy for numerical forward physics model. An actor-critic structure is adopted to design the iteration strategy. The actor network gives the iteration strategy, and the performance is evaluated by the critic network which improves the strategy given by actor network in the coming steps. For the problem with high dimensions in its action space, the “reward” which is used to improve the performance by critic network does not perform as efficient as it does with the problem of lower action space dimensions [30–32]. A pinning-based strategy is given

in this paper to reduce the dimension of action space, which helps to make the critic network more efficient.

The principle of actor-critic based structure is introduced in Section 2 along with the principle of MFL inspection. The detail of PACS algorithm is described in Section 3. Simulated inspection signal from a nonlinear FEM model is used to test the performance of the algorithm proposed under 20 dB noise in Section 4. The conclusion is drawn in Section 5.

## 2. Model and Principle

*2.1. Physics Model.* The principle of MFL inspection is based on electromagnetic theory. By magnetizing the test material into saturation, a magnetic flux leakage can be detected by Hall-effect sensors where a defect is located. Strong permanent magnets are usually used to magnetize the testing material. The Hall-effect sensors are usually located close to the surface of the tested material. The magnetizing and sensing principles are illustrated in Figure 1.

The principle of MFL inspection is magnetic, where Maxwell’s equation can be used to describe its behavior:

$$\nabla \times (\mu^{-1} \nabla \times \mathbf{A}) = \mathbf{J}, \quad (1)$$

where  $\mu$ ,  $\mathbf{J}$ , and  $\mathbf{A}$  represent the permeability of the media, the source current density, and the magnetic vector potential. In (1),  $\mu$  is usually not a constant due to the property of the material and can be described as a function of magnetic flux density  $\mathbf{B}$  as  $\mu = \mu(\mathbf{B})$ . The magnetic flux density which can be collected with the Hall-effect sensors is  $\mathbf{B} = \nabla \times \mathbf{A}$ . For a simple defect, the magnetic flux density  $\mathbf{B}$  is illustrated in Figure 2. The  $x$ -axis component is usually sampled as the inspection signal. There are two ways to solve this Maxwell’s equation formulated as (1). Numerical solution such as finite element method (FEM) is widely used to solve this partial differential equation. Another kind of solution is dipole model, which makes some simplification of the forward model, which gives an analytical solution to this forward model [19, 33, 34].

As the analytical solution cannot provide enough accuracy for complex defects, numerical methods are usually used to get a numerical solution for these problems. FEM is a widely used method to get a numerical solution for partial differential equation. The general process of a FEM solution is as follows. First, the partial differential equation is transformed into corresponding variational functional equation. Then, the domain that needs to be computed is divided into certain number of finite elements. By assembling all the variational functional equations of all the elements within the domain, the solution can be obtained by solving

$$\mathbf{K} \cdot \mathbf{a} = \mathbf{f}, \quad (2)$$

where  $\mathbf{a}$  represents the nodal solutions of elements of discrete approximation in the form of a vector,  $\mathbf{K}$  is the sparse element stiffness matrix, and  $\mathbf{f}$  is the source vector containing the boundary conditions and model inputs.

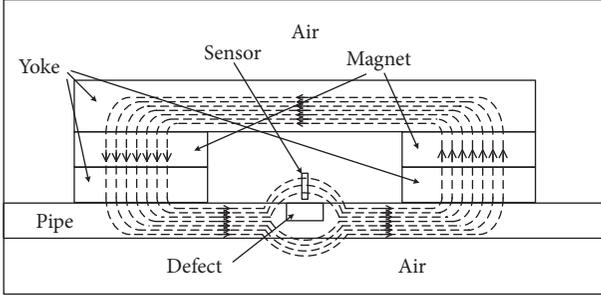


FIGURE 1: Principle of MFL inspection.

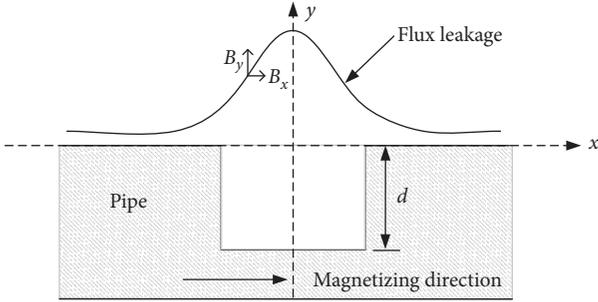


FIGURE 2: Illustration of magnetic flux leakage of a simple defect.

Commonly, the MFL inspection model is built with components as described in Figure 1. As the sensor position is fixed with the magnetizing components, it needs to rebuild the defect location  $n$  times if a sequence of data with  $n$  points are sampled. It means to repeat the forward model  $n$  times, which is computationally very costly. A simplified model is adopted in this paper, which is proposed in [14]. The simplified model is illustrated in Figure 3. As the principle of MFL inspection is magnetizing the material into saturation, a pair of paralleled current layers are adopted to magnetize the testing material. The commonly used permanent magnets and yokes are removed. By adopting this model, the principle and nonlinear character are kept. A sampled signal with  $n$  sampling points only needs to run the forward model once, which saves a lot computationally. The region of interest (ROI) shown in Figure 3 represents the domain where the defect is going to be reconstructed by the algorithm proposed. The depth profile within ROI consists of several subdefects, which make up a complex depth profile.

**2.2. Principle of RL.** Reinforcement learning considers the paradigm of an agent interacting with its environment aiming to learn a behavior which maximizes the reward. The agent consists of an actor network and a critic network. The actor network is trained to decide which action should be taken at current state. The critic network evaluates each action based on its current state with reward and improves the strategy of the actor network. There are four definitions in RL, state  $x_t$ , reward  $r_t$ , action  $a_t$ , and environment  $E$ . An agent takes action  $a_t$  at current state  $x_t$ , where  $t$  is discrete time-step. The action  $a_t$  interacts with certain environment

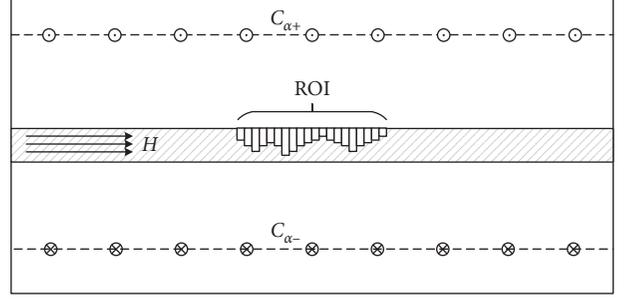


FIGURE 3: Illustration of simplified FEM model used in this paper with complex depth profile.

and obtains a new state  $x_{t+1}$  with a reward  $r_t$ , which evaluates the performance of this action. The action is defined by a deterministic policy  $\pi: \mathcal{S} \rightarrow \mathcal{A}(\mathcal{S})$  which is a mapping from states to actions.  $\mathcal{S}$  represents the state space, and  $\mathcal{A}$  represents the action space. A discounted sum of future rewards is called a return as  $R_t = \sum_{i=t}^T \gamma^{(i-t)} r(\mathbf{s}_i, \mathbf{a}_i)$ , where  $\gamma \in [0, 1]$  is a discounting factor. The agent's goal is to maximize the expected returns  $\mathbb{E}_\pi[R_1]$ . The action-value function is defined as  $Q^\pi(\mathbf{s}_t, \mathbf{a}_t) = \mathbb{E}_\pi[R_t | \mathbf{s}_t, \mathbf{a}_t]$ . It is also called Q-function, which represents the expected return after taking an action  $\mathbf{a}_t$  in state  $\mathbf{s}_t$  and thereafter following the policy  $\pi$ . The critic is updated by minimizing the loss  $L$  as follows:

$$L = \frac{1}{N} \sum_i (y_i - Q(s_i, a_i | \theta^Q))^2, \quad (3)$$

where

$$y_i = r_i + \gamma Q'(s_{i+1}, \pi'(s_{i+1} | \theta_{\pi'})) | \theta^{Q'}. \quad (4)$$

The actor policy is updated using sampled gradient as follows:

$$\nabla_{\theta^\pi} \pi|_{s_i} \approx \frac{1}{N} \sum_i \nabla_a Q(s, a | \theta^Q)|_{s=s_i, a=\pi(s_i)} \nabla_{\theta^\pi} \pi(s | \theta^\pi)|_{s_i}. \quad (5)$$

In the problem of sizing the depth profile of MFL inspection, four parameters are involved in the reconstruction process, the reference depth profile  $\mathbf{d}_{\text{ref}}$ , reconstructed depth profile  $\mathbf{d}_t$  at time-step  $t$ , reference signal  $\mathbf{s}_{\text{ref}}$ , and signal of reconstructed depth profile  $\mathbf{s}_t$  at time-step  $t$ . Reference depth profile is the target of the reconstruction, which is not observable during the entire reconstruction process. The other three parameters are fully observable all the time. During training of non-model-based method, the reference depth profile and corresponding reference signal are used to train the mapping. The model-based method only involves the reference signal during the reconstruction process. In this paper, three parameters, reference depth profile, reference signal, and reconstructed depth profile, are used to train the actor-critic structure-based algorithm proposed. The involvement of reconstructed depth profile gives more information of the depth profile space. The signal of reconstructed signal is not utilized as it is computationally costly.

This paper is inspired by the similarity between the training process of RL and model-based iteration method. The similarity is illustrated as shown in Figure 4. For the iteration structure, it starts with an initial defect depth profile. According to the iteration strategy, a reconstructed depth profile is given. The signal of corresponding depth profile is generated with the forward model. By comparing the signal of reconstructed depth profile with reference signal, a residual is obtained. This process iterates until the residual is smaller than a threshold when a final reconstructed depth profile is obtained. The learning process of the actor-critic structure-based RL method proposed in this paper has similarities with the model-based solution mentioned above. The state can be that the depth profile needs to be reconstructed in a certain form. The agent is the strategy controlling the iteration process. Action is the output of the agent, which controls how to change the state until termination criteria are satisfied. The strategy that controls the iteration process is learned from the data given and generated during the iteration process. It solves the problem that it is hard to design iteration strategy for numerical forward models. The performance of the strategy is evaluated by the critic network with rewards from each step. By involving the data generated during the iteration process, more data is given, which means the training data is not limited to the given data pairs of depth profile and corresponding signal. It improves the problem that the non-model-based solution highly relies on the distribution of the training data.

### 3. Algorithm

In this paper, an actor-critic structure-based RL method for complex depth profile reconstruction is proposed. The algorithm of Deep Deterministic Policy Gradients (DDPG) is adopted to train the actor-critic structure [35–37]. The definitions of parameters for the problem of sizing the depth profile of MFL inspection are described as follows. The state is defined as  $\mathbf{s}_t = (\mathbf{d}_t, \mathbf{s}_{\text{ref}})$ , which consists of two parts, the normalized reconstructed depth profile and the reference signal.  $\mathbf{d}_t$  is the normalized reconstructed depth at time-step  $t$ .  $\mathbf{s}_{\text{ref}}$  is the sampled reference signal. As the signal outside the ROI has less characteristics than signal within the domain of ROI, most of the sampling points are selectively removed to reduce the state dimension. The change of  $\mathbf{d}_t$  at each time-step is taken as action. Different from the model-based method that uses the residual between reference signal and signal of reconstructed depth profile to evaluate the performance, the performance of the actor network is evaluated with reward at each time-step. Reward is designed as minus value of Euclidean distance between reference depth profile and reconstructed depth profile as shown in the following equation:

$$r_t = -\|\mathbf{d}_{t+1} - \mathbf{d}_{\text{ref}}\|. \quad (6)$$

As the target of this MFL inverse problem is sizing the depth profile as precisely as possible, it means the subdefect  $d(i)$  needs to approach its corresponding reference subdefect with small error. Then, the complexity is associated

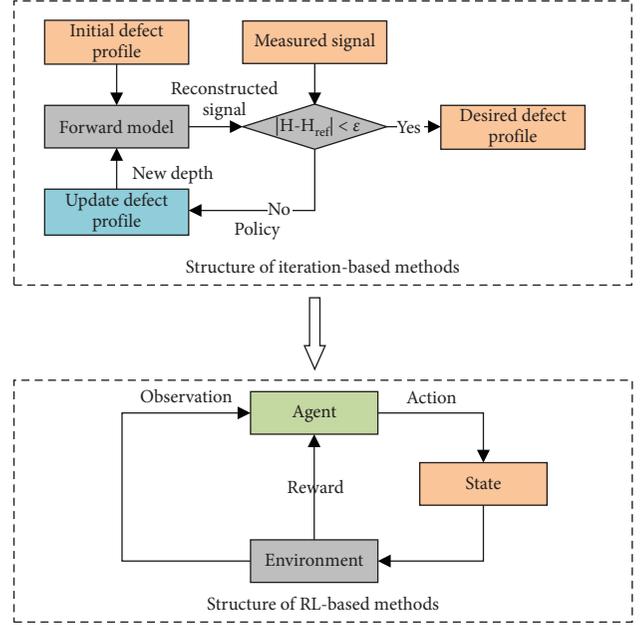


FIGURE 4: Similarity between the learning process of reinforcement learning and classic iteration method.

with the dimension of the degree of freedom of this inverse problem. The problem when encountering a high-dimensional degree of freedom comes that the reward that evaluates the performance of the actor network becomes less efficient. It is because the measurement of distance becomes less efficient in high-dimensional problems [30–32]. In this paper, instead of giving each subdefect an action to control its depth, limited subdefects are selected to accept action given by the actor network. Subdefects accepting control from actor network are called pinning subdefects in this paper. As the defects are usually caused by corrosion or mechanical damage, the difference between adjoining subdefects is usually not sharp. It means that it is possible to use some pinning subdefects, which are controlled by the action to represent the depth profile. Subdefects between two pinning subdefects are interpolated. By adopting this pinning strategy, the dimension of action space is reduced significantly. The reward is still calculated using the full information of depth profile. As the dimension of action space is reduced, the measurement of distance becomes more efficient than using full space. The entire depth profile with pinning subdefects is illustrated in Figure 5. The flowchart of PACS learning process within one episode is illustrated in Figure 6. The entire algorithm including the pinning strategy with the learning process of the actor-critic structure-based reconstruction algorithm is given in Algorithm 1.

From Algorithm 1, it can be seen that, within one episode,  $\mathbf{s}_{\text{ref}}$  as part of the states is combined with many depth profiles  $\mathbf{d}_t$  generated during the iteration process. It means that, despite the relationship between ultimate reconstructed depth profile and reference signal, the function in the depth profile space is also explored by PACS proposed in this paper. It helps to have better

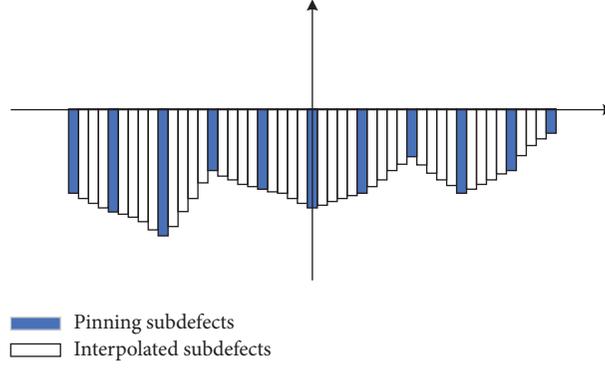


FIGURE 5: Illustration of pinning-based depth reconstruction within the ROI.

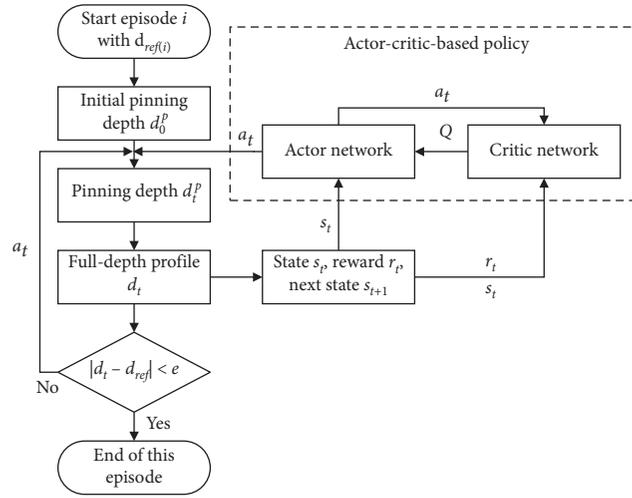


FIGURE 6: Flowchart of PACS learning process within one episode.

- (i) Initialize actor network  $\mu(\mathbf{s}|\theta^\pi)$ , critic network  $Q(\mathbf{s}, \mathbf{a}|\theta^Q)$ , target network  $\pi'$  and  $Q'$ , replay buffer  $R$
- (ii) For episode = 1, M do
- (iii) Initialize pinning subdefects  $\mathbf{d}_0^p$ , interpolate to have the full depth profile
- (iv) Get initial observation state  $\mathbf{s}_1$  from reference signal and depth of sub-defects
- (v) For  $t = 1, T$  do
- (vi) Generate an action  $\mathbf{a}_t = \mu(\mathbf{s}_t|\theta^\pi) + \mathcal{N}_t$  from the output of actor network and exploration noise process
- (vii) Execute action  $\mathbf{a}_t$ , obtain new depth of pinning sub-defects  $\mathbf{d}_t^p$
- (viii) Interpolate to get the full depth profile  $\mathbf{d}_t$  within the ROI, calculate reward  $r_t$  and new state  $\mathbf{s}_{t+1}$
- (ix) Store  $(\mathbf{s}_t, \mathbf{a}_t, r_t, \mathbf{s}_{t+1})$  in  $R$
- (x) If capacity of replay buffer  $R$  is full then
- (xi) Randomly sample  $N$  piece of data  $(\mathbf{s}_t, \mathbf{a}_t, r_t, \mathbf{s}_{t+1})$  from  $R$
- (xii) Update the critic network and actor network with (5) and (3)
- (xiii) Update the target networks:
- (xiv)  $\theta^{Q'} \leftarrow \tau\theta^Q + (1 - \tau)\theta_{Q'}$
- (xiv)  $\theta^{\pi'} \leftarrow \tau\theta^\pi + (1 - \tau)\theta_{\pi'}$
- (xv) end if
- (xvi) If error between each reference subdefect and reconstructed subdefect is less than  $\epsilon$ , then
- (xvii) break
- (xviii) end if
- (xix) end for
- (xx) end for

ALGORITHM 1: PACS.

reconstruction results in an optimal way. As only component  $\mathbf{d}_t$  in state  $\mathbf{s}_t = (\mathbf{d}_t, \mathbf{s}_{\text{ref}})$  is updated, no forward model is called during one episode. It saves a lot as the call of forward model is computationally very costly if a fine model is required.

## 4. Results and Discussion

**4.1. Model and Error Definitions.** To test the accuracy of the algorithm proposed along with robustness, a simplified nonlinear numerical forward model is adopted as in [14]. The detail of the forward model is illustrated in Figure 3 and the dimension of the model can be found in Figure 7. There are 49 subdefects within the ROI. The adjoining subdefects stay tight and the span between centers of adjoining two subdefects is 2 mm. 11 subdefects are selected as pinning subdefects. The position of each pinning subdefect is shown in Figure 5. Subdefects between two pinning subdefects are interpolated with cubic interpolation. The x-component of the signal sampled with lift-off value of 1 mm above the surface is adopted as reference signal. The current density carried in the paralleled layer is  $40000 \text{ kA/m}^2$  with opposite direction. The material is set as 1010 cold rolled steel. The property of the material including the B–H curve can be found in [14].

In order to test the effectiveness of the algorithm proposed in this paper, three error measurements are given. These measurements are root mean squared error (RMSE), peak depth error (PDE), and maximum deviation (MD). These measurements are described in (7)–(9) and illustrated in Figure 8.  $d_{\text{ref}}(i)$  and  $d(i)$  are the  $i$ th depths of subdefect for reference depth profile and reconstructed depth profile, respectively. From (7)–(9) and illustration in Figure 8, the error definitions can be understood in an easy way. MSE is commonly used in error measurements. PDE is the error between the maximum depths of reconstructed depth profile and reference depth profile. The subdefects from the reconstructed subdefect and the reference subdefect may not come from the same location within ROI. The value of 0.1 with regard to PDE means 1 mm error between peak depths if the wall thickness is 10 mm. MD is the maximum error between reconstructed subdefect and reference subdefect with the same location of the subdefects. The subdefects used to calculate MD value may not be the maximum depth of neither reconstructed depth profile nor reference depth profile. The value of 0.1 with regard to MD means 1 mm error if the wall thickness is 10 mm.

$$\text{RMSE}(d) = \sqrt{\frac{1}{N} \sum_{\text{ROI}} (d(i) - d_{\text{ref}}(i))^2}, \quad (7)$$

$$\text{PDE}(\mathbf{d}) = \frac{1}{D} (\max(\mathbf{d}_{\text{ref}}) - \max(\mathbf{d})), \quad (8)$$

$$\text{MD}(\mathbf{d}) = \frac{1}{D} \max(d_{\text{ref}}(i) - d(i)). \quad (9)$$

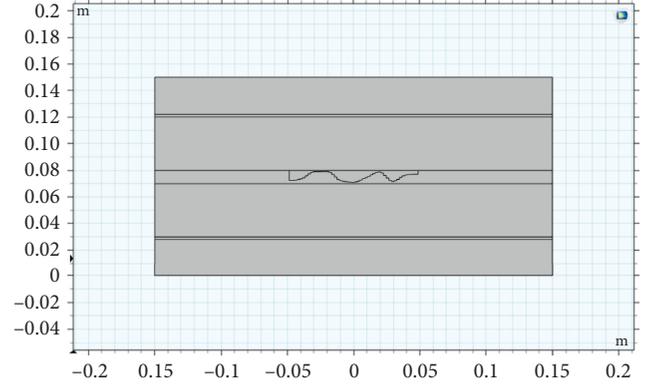


FIGURE 7: Dimension of the forward model with a defect of 49 subdefects.

**4.2. Computing Results.** The structure of actor network and critic network of PACS algorithm is described as follows. There are 82 neutrons in the input layer of actor network: 11 neutrons for the pinning subdefect normalized depth vector and 71 neutrons for the normalized reference signal. The sampling position of signal used as part of the state is illustrated in Figure 9. In the first hidden layer, there are 128 neutrons, and in the second hidden layer, there are 80 neutrons. 11 neutrons are set as the output layer, which controls the action for pinning subdefects. Besides, the activation function of output layer is set as “tanh”, and all the activation functions are set as “ReLU.” The input layer of critic network has two separated parts: one is 82 neutrons as the input layer of actor network and the other is 11 neutrons for the action. There are 128 neutrons, respectively, in the first hidden layer of critic network connecting the corresponding parts. The number of neutrons for the second and third hidden layer is 50, respectively. The output of critic network is Q value with one neutron.

The number of episodes  $M$  is 5000 with 200 time-steps for each episode as  $T$ . The size of replay buffer  $R$  is 1000000. The number of pieces of data of sampling data at each time-step  $N$  is 128. The stop criterion  $\epsilon$  is 0.5. The soft updating parameter  $\tau$  is 0.01. The discounting factor  $\gamma$  is 0.1.

There are 10000 randomly generated complex defects with corresponding sampled signal generated with COMSOL Multiphysics 5.3a with MATLAB. 5000 of the pieces of data are used as training data set and 5000 others as testing data set. The algorithm proposed is coded with Python and TensorFlow 1.15. All the data and algorithm are run on a laptop with Intel i7 10750H processor and 16 GB RAM.

From error definition (7)–(9), the reconstructed results can be shown in different aspects. The results are shown with selected reconstruction results from the different ranking with MD value. The results of the 10%, 30%, 70%, and 90% are shown in Figure 10. The MD value is sorted from the smallest to the largest, which means Figure 10(a) is the best result among all the tested results shown in Figure 10 from the aspect of MD value. The ranking of the results in Figure 10 is tested using signal with SNR = 20 dB. Corresponding reconstructed results with noise-free signal for each depth profile and reference depth profile are also

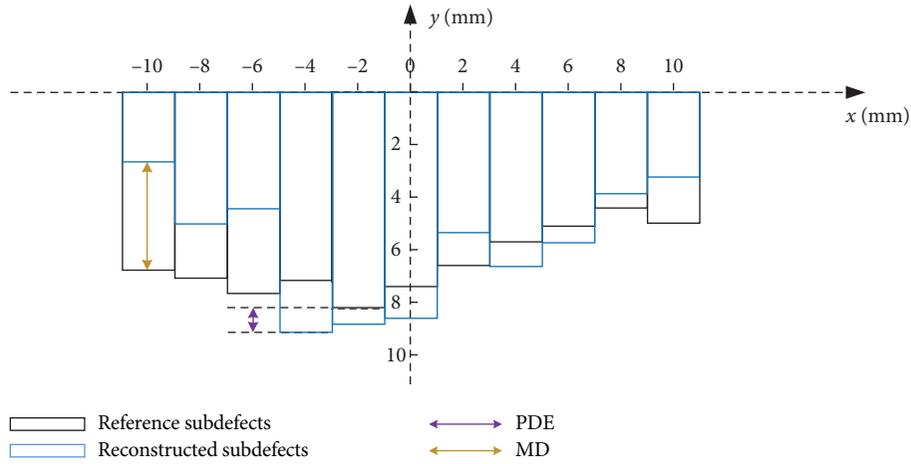


FIGURE 8: Illustration of PDE and MD.

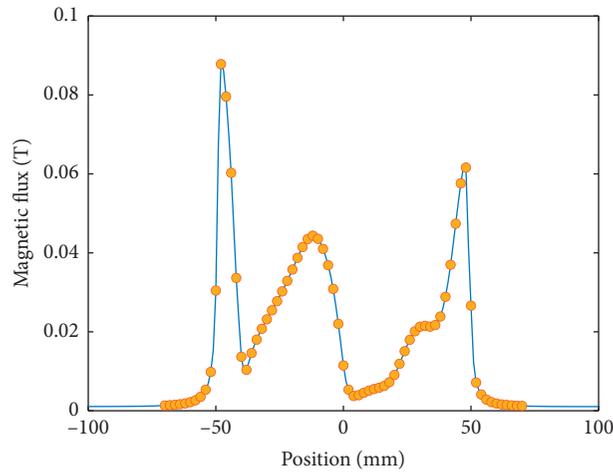


FIGURE 9: Illustration of sampling position of the inspection signal as part of state.

plotted in Figure 10. The corresponding values with ranking of each error definition are listed at the bottom of each subfigure. From Figure 10, it can be seen that all the reconstructed results follow the depth profile well. The worst result from PDE is in Figure 10(a) with the PDE ranking of 80.7% and value of -0.0786. The worst result from RMSE is in Figure 10(d) with the RMSE ranking of 87.2% and value of 0.1461. These are relatively small errors, proving the accuracy of the algorithm proposed in this paper. From the results in Figure 10, it can also be seen that the reconstructed results from signal of 20 dB noise are close to the results reconstructed from noise-free signal, which means the algorithm proposed in this paper is robust against noise.

The signals for corresponding reconstructed depth profile in Figure 10 are shown in Figure 11. The noise-free signal and signal with 20 dB noise are plotted in different color. The values of three error measurements are plotted in Figure 12 for the first 20 time-steps of reconstruction. They are results generated from reconstructing process of results in Figure 10 with noise signal of 20 dB. It can be seen from Figure 12 that all the error measurements converge in less

than 10 time-steps. The results in Figure 12 show that the algorithm proposed in this paper converges fast with limited steps to the final reconstruction results.

To show the robustness of the algorithm proposed in this paper, the algorithm is trained with different size of training data sets. The size of the testing data set is 5000, which is the same as that in Figure 10. The sizes of training data sets are set as 2000, 3000, 4000, and 5000, respectively. A 20 dB noise is also added to the testing data. The results are shown as error distribution in Figure 13. The error distributions of results are shown as histogram. To make the figure clear, they are plotted with markers in different color instead of bars. The  $y$ -axis value of the markers represents the probability that the results fall into corresponding span. The width of the span is equal to that of the span between two adjoining markers from  $x$ -axis direction. The marker is located at the center of its corresponding span. The result is better if the error distributions are more concentrated and closer to zero. From Figure 13, it can be seen that, in contrast to the results from training with 2000 pieces of data, results from 3000, 4000, and 5000 have similar error distributions.

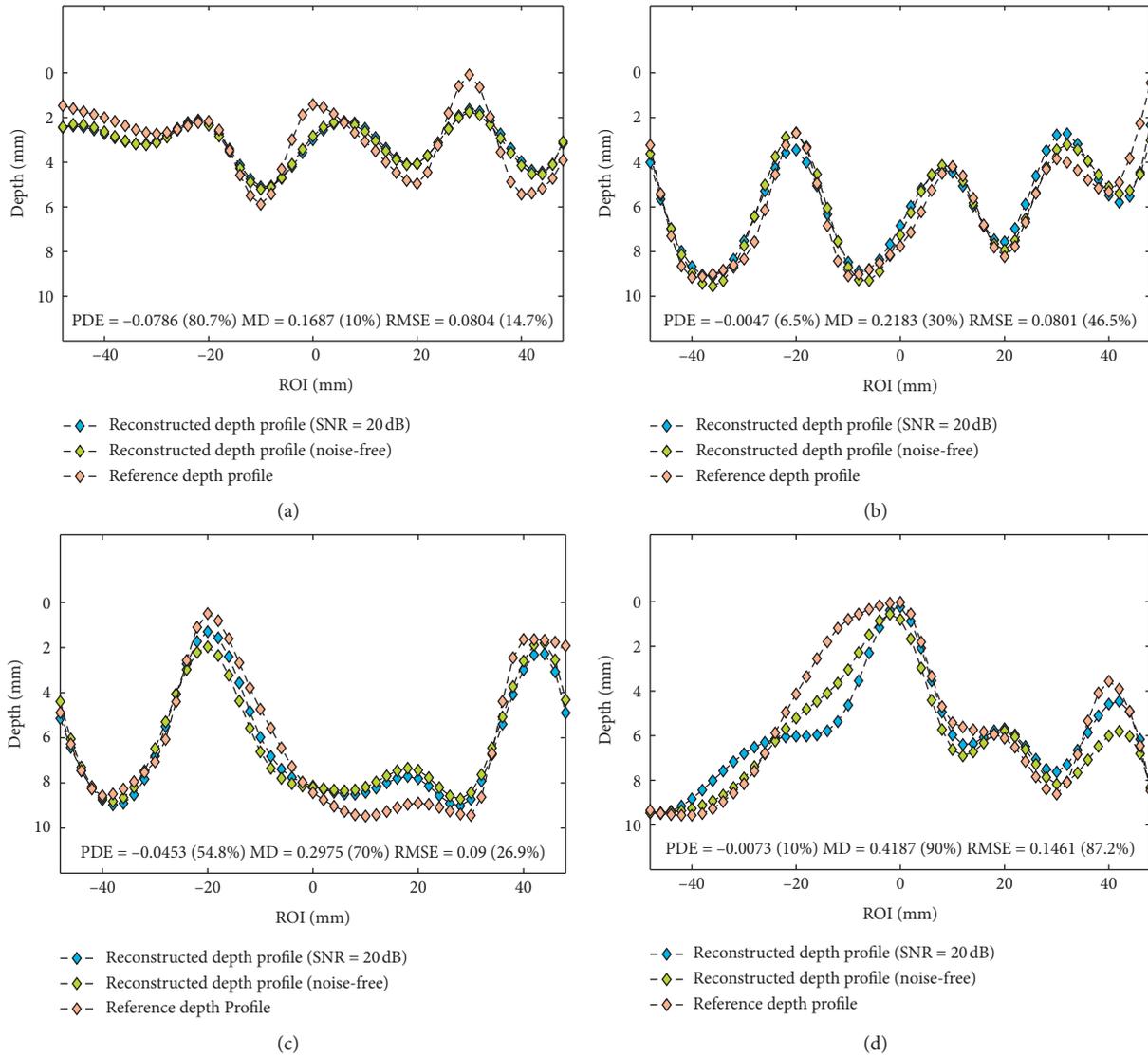


FIGURE 10: Sizing results with different ranking of MD value. (a) 10% ranking. (b) 30% ranking. (c) 70% ranking. (d) 90% ranking.

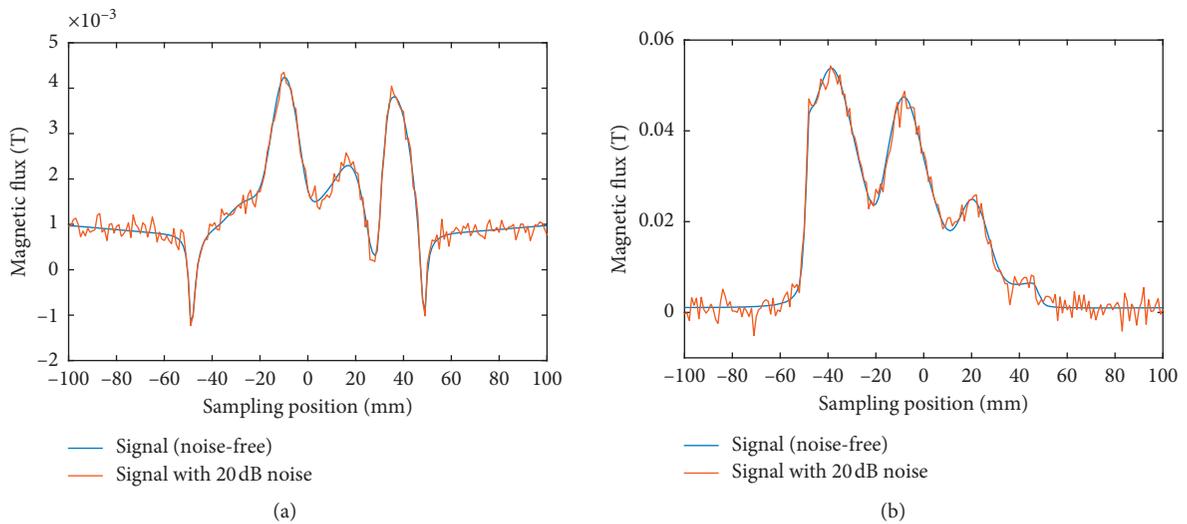


FIGURE 11: Continued.

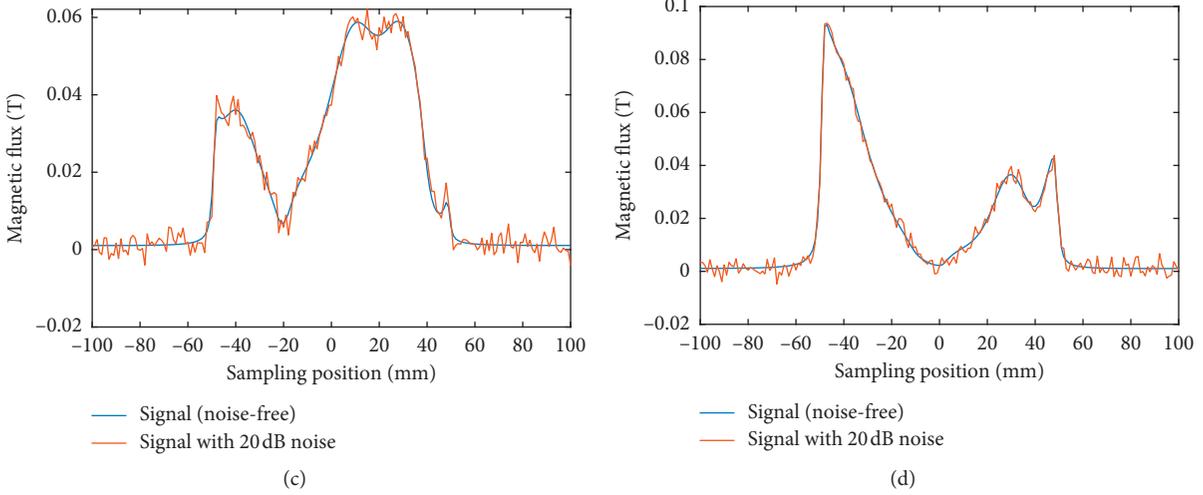


FIGURE 11: Signal used to get the results in Figure 10, including the contaminated signal and corresponding noise-free signal. (a) 10% ranking. (b) 30% ranking. (c) 70% ranking. (d) 90% ranking.

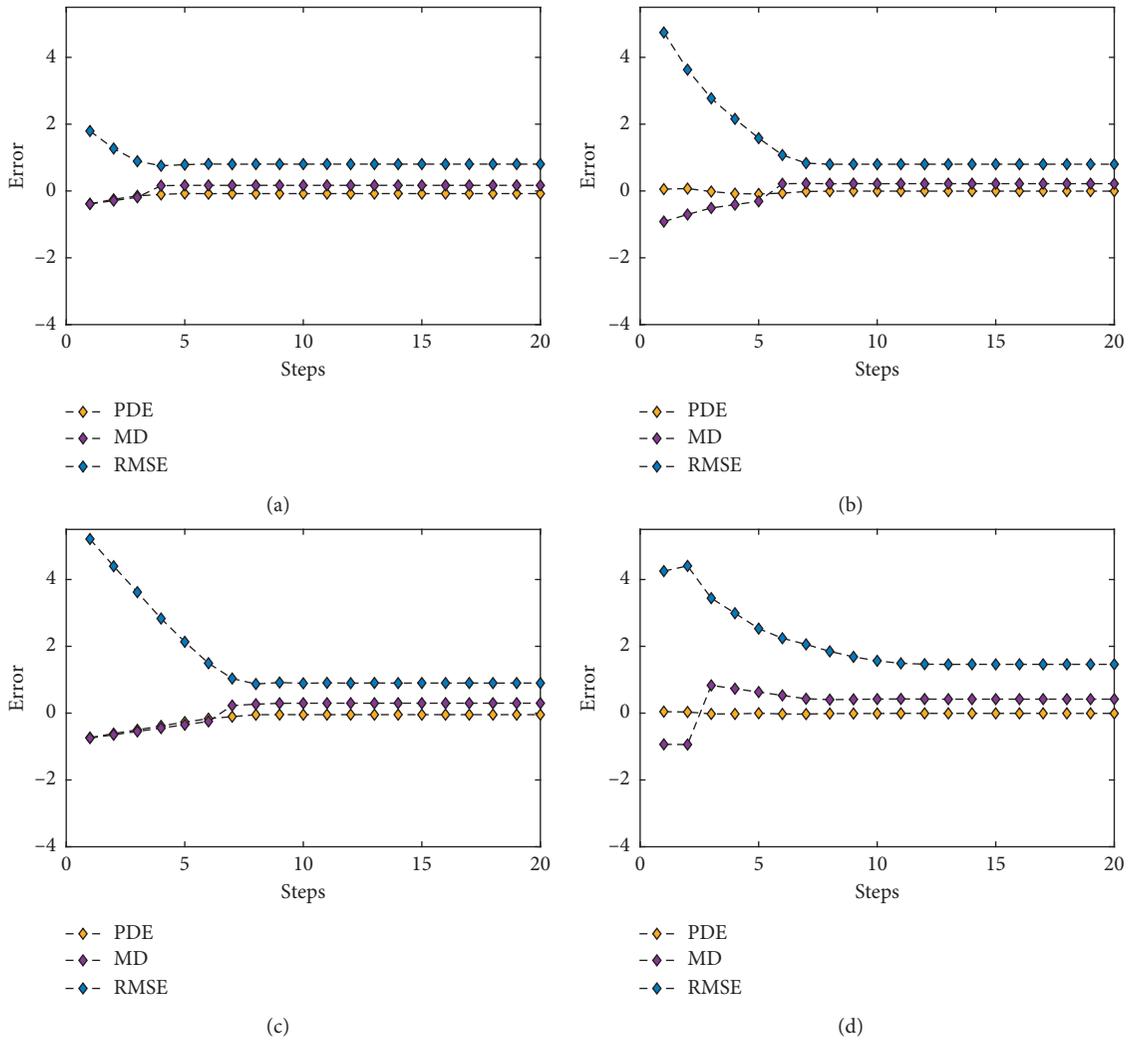


FIGURE 12: Error convergence for corresponding results in Figure 10. (a) 10% ranking. (b) 30% ranking. (c) 70% ranking. (d) 90% ranking.

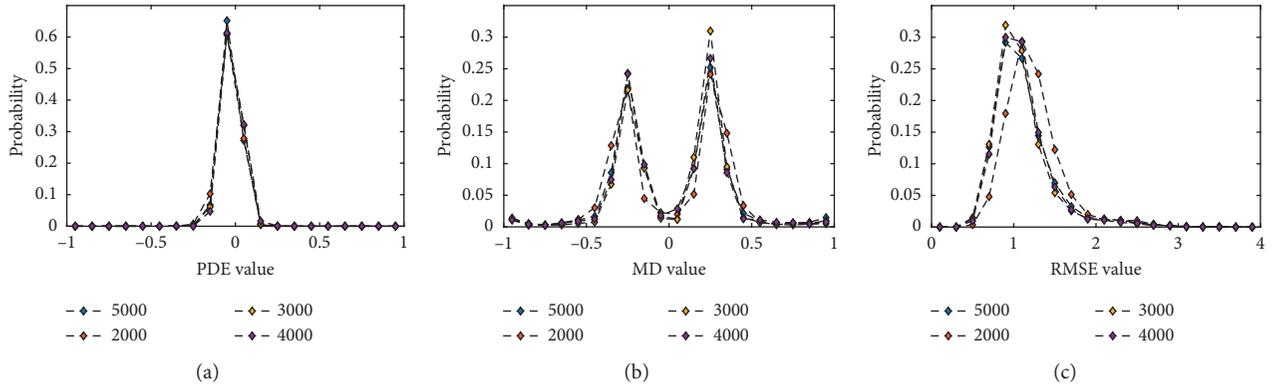


FIGURE 13: Error distribution of results from different size of training data sets. (a) PDE. (b) MD. (c) RMSE.

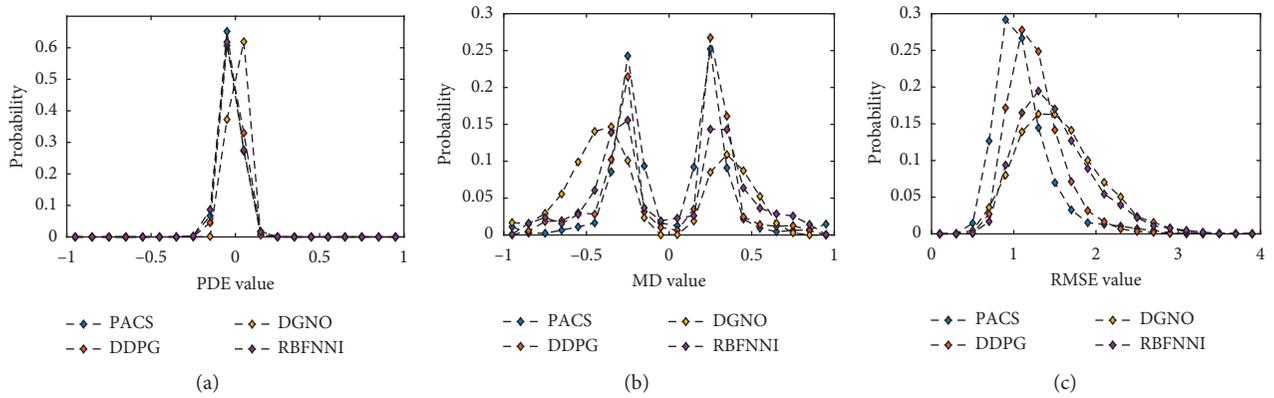


FIGURE 14: Results shown as distribution of error definitions from different reconstruction methods. (a) PDE. (b) MD. (c) RMSE.

From Figures 13(b) and 13(c), the results from 3000 to 4000 are even a little bit better than results from 5000. The results from 2000 have similar error distributions too, but it can also be seen that the performance is obviously not as good as results from larger training data sets from all aspects. The results from Figure 13 show that the quantity requirement of training data is not high and the algorithm is robust.

The actor-critic structured DDPG, the direct Gauss-Newton optimization (DGNO) in [14], and RBF neural network based iteration (RBFNNI) in [21] are selected as representative methods to show the accuracy and robustness of the algorithm proposed. The results are shown in Figure 14 in the form of error distribution too. The means of markers are the same as those in Figure 13. From Figures 14(b) and 14(c), it can be seen that the reconstruction results from PACS have obviously the best performance that the error distributions are more concentrated and closer to zero. From Figure 14(a), results from DDPG and DGNO are slightly better than those from PACS. Considering their performance on MD and RMSE, the results from PACS are still better than those from these methods.

## 5. Conclusion

In this paper, a pinning actor-critic structure-based solution for sizing complex depth profile with high degree of freedom

of MFL inspection is studied. By involving the actor-critic structure, a novel way of utilizing the fine numerical based forward model in reconstructing the depth profile for MFL inspection is given. To solve the problem of the performance of reward deficiency, which is measured as Euclidean distance, a pinning strategy is given. By introducing the pinning subdefects, the action space has less variability than giving every subdefect an action. The robustness of the reconstruction results is improved by involving PACS.

The effectiveness of PACS proposed in this paper is tested with simulation results from nonlinear numerical forward models of MFL inspection with FEM. The results that are shown in a statistic way show the effectiveness of PACS proposed in this paper. The depth profiles reconstructed from signal with 20 dB noise are close to depth profiles reconstructed from noise-free signal, proving the robustness of PACS proposed. The results also show good accuracy compared with representative solutions of depth profile reconstruction.

## Data Availability

No data were used to support this study.

## Conflicts of Interest

The authors declare that they have no conflicts of interest.

## Acknowledgments

This work was supported by the National Natural Science Foundation of China (Grant nos. 61703087 and 71502029).

## References

- [1] R. Palmer-Jones, S. Turner, and P. Hopkins, "A new approach to risk based pipeline integrity management," *International Pipeline Conference*, vol. 426, pp. 811–823, 2006.
- [2] V. E. Loskutov, A. F. Matvienko, B. V. Patramanskii, and V. E. Shcherbinin, "The magnetic method for in-tube non-destructive testing of gas and oil pipelines: the past and the present," *Russian Journal of Nondestructive Testing*, vol. 42, pp. 493–504, 2006.
- [3] Y. Shi, C. Zhang, R. Li, M. Cai, and G. Jia, "Theory and application of magnetic flux leakage pipeline detection," *Sensors*, vol. 15, no. 12, pp. 31036–31055, 2015.
- [4] M. R. Kandroodi, B. N. Araabi, M. M. Bassiri, and M. N. Ahmadabadi, "Estimation of depth and length of defects from magnetic flux leakage measurements: verification with simulations, experiments, and pigging data," *IEEE Transactions on Magnetics*, vol. 53, no. 3, pp. 1–10, 2016.
- [5] A. Joshi, L. Udpa, S. Udpa, and A. Tamburrino, "Adaptive wavelets for characterizing magnetic flux leakage signals from pipeline inspection," *IEEE Transactions on Magnetics*, vol. 42, no. 10, pp. 3168–3170, 2006.
- [6] D. Mukherjee, S. Saha, and S. Mukhopadhyay, "Inverse mapping of magnetic flux leakage signal for defect characterization," *NDT & E International*, vol. 54, pp. 198–208, 2013.
- [7] S. Saha, S. Mukhopadhyay, U. Mahapatra, S. Bhattacharya, and G. P. Srivastava, "Empirical structure for characterizing metal loss defects from radial magnetic flux leakage signal," *NDT & E International*, vol. 43, no. 6, pp. 507–512, 2010.
- [8] M. Hui and G. S. Park, "A study on the estimation of the shapes of axially oriented cracks in cmfl type ndt system," *IEEE Transactions on Magnetics*, vol. 50, no. 2, pp. 109–112, 2014.
- [9] Z. Ma and H. Liu, "Pipeline defect detection and sizing based on mfl data using immune rbf neural networks," in *Proceedings of the 2007 IEEE Congress on Evolutionary Computation*, pp. 3399–3403, IEEE, Singapore, September 2007.
- [10] Y. Li, J. Wilson, and G. Y. Tian, "Experiment and simulation study of 3d magnetic field sensing for magnetic flux leakage defect characterisation," *NDT & e international*, vol. 40, no. 2, pp. 179–184, 2007.
- [11] W. Cheng, "Nondestructive testing of back-side local wall-thinning by means of low strength magnetization and highly sensitive magneto-impedance sensors," *IEEE Sensors Journal*, vol. 16, no. 14, pp. 5548–5556, 2016.
- [12] S. Lu, J. Feng, F. Li, and J. Liu, "Precise inversion for the reconstruction of arbitrary defect profiles considering velocity effect in magnetic flux leakage testing," *IEEE Transactions on Magnetics*, vol. 53, no. 4, pp. 1–12, 2016.
- [13] M. Ravan, R. K. Amineh, S. Koziel, K. Natalia, and J. P. Reilly, "Sizing of 3-d arbitrary defects using magnetic flux leakage measurements," *IEEE Transactions on Magnetics*, vol. 46, no. 4, pp. 1024–1033, 2009.
- [14] R. H. Priedwald, C. Magele, P. D. Ledger, N. R. Pearson, and J. S. D. Mason, "Fast magnetic flux leakage signal inversion for the reconstruction of arbitrary defect profiles in steel using finite elements," *IEEE Transactions on Magnetics*, vol. 49, no. 1, pp. 506–516, 2012.
- [15] K. C. Hari, M. Nabi, and S. V. Kulkarni, "Improved fem model for defect-shape construction from mfl signal by using genetic algorithm," *IET Science, Measurement & Technology*, vol. 1, no. 4, pp. 196–200, 2007.
- [16] J. Feng, F. Li, S. Lu, and J. Liu, "Fast reconstruction of defect profiles from magnetic flux leakage measurements using a rbfnn based error adjustment methodology," *IET Science, Measurement & Technology*, vol. 11, no. 3, pp. 262–269, 2016.
- [17] R. Schifini and A. C. Bruno, "Experimental verification of a finite element model used in a magnetic flux leakage inverse problem," *Journal of Physics D: Applied Physics*, vol. 38, no. 12, p. 1875, 2005.
- [18] C. Edwards and S. B. Palmer, "The magnetic leakage field of surface-breaking cracks," *Journal of Physics D: Applied Physics*, vol. 19, no. 4, p. 657, 1986.
- [19] Y. Zhang, Z. Ye, and C. Wang, "A fast method for rectangular crack sizes reconstruction in magnetic flux leakage testing," *Ndt & E International*, vol. 42, no. 5, pp. 369–375, 2009.
- [20] X. Liu, Y. Deng, Z. Zeng, L. Udpa, and S. S. Udpa, "Model-based inversion technique using element-free galerkin method and state space search," *IEEE Transactions on Magnetics*, vol. 45, no. 3, pp. 1486–1489, 2009.
- [21] J. Chen, S. Huang, and W. Zhao, "Three-dimensional defect inversion from magnetic flux leakage signals using iterative neural network," *IET Science, Measurement & Technology*, vol. 9, no. 4, pp. 418–426, 2015.
- [22] W. Han, J. Xu, P. Wang, and G. Tian, "Defect profile estimation from magnetic flux leakage signal via efficient managing particle swarm optimization," *Sensors*, vol. 14, no. 6, pp. 10361–10380, 2014.
- [23] F. Li, J. Feng, H. Zhang et al., "Quick reconstruction of arbitrary pipeline defect profiles from mfl measurements employing modified harmony search algorithm," *IEEE Transactions on Instrumentation and Measurement*, vol. 67, no. 9, pp. 2200–2213, 2018.
- [24] V. Mnih, K. Kavukcuoglu, D. Silver et al., "Playing atari with deep reinforcement learning," 2013, <http://arxiv.org/abs/1312.5602>.
- [25] V. Mnih, K. Kavukcuoglu, D. Silver et al., "Human-level control through deep reinforcement learning," *Nature*, vol. 518, no. 7540, pp. 529–533, 2015.
- [26] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in Neural Information Processing Systems*, vol. 25, pp. 1097–1105, 2012.
- [27] R. Wang, Q. Sun, W. Hu et al., "Soc-based droop coefficients stability region analysis of the battery for stand-alone supply systems with constant power loads," *IEEE Transactions on Power Electronics*, vol. 36, no. 7, pp. 7866–7879, 2021.
- [28] R. Wang, Q. Sun, D. Ma, and Z. Liu, "The small-signal stability analysis of the droop-controlled converter in electromagnetic timescale," *IEEE Transactions on Sustainable Energy*, vol. 10, no. 3, pp. 1459–1469, 2019.
- [29] W. Rui, S. Qiuye, M. Dazhong, and H. Xuguang, "Line impedance cooperative stability region identification method for grid-tied inverters under weak grids," *IEEE Transactions on Smart Grid*, vol. 11, no. 4, pp. 2856–2866, 2020.
- [30] C. R. Taylor, *Applications of Dynamic Programming to Agricultural Decision Problems*, CRC Press, Boca Raton, FL, USA, 2019.
- [31] V. T. Gerard, "A problem of dimensionality: a simple example," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 3, pp. 306–307, 1979.

- [32] B. Chandrasekaran and A. K. Jain, "Quantization complexity and independent measurements," *IEEE Transactions on Computers*, vol. 100, no. 1, pp. 102–106, 1974.
- [33] S. M. Dutta, F. H. Ghorbel, and R. K. Stanley, "Dipole modeling of magnetic flux leakage," *IEEE Transactions on Magnetics*, vol. 45, no. 4, pp. 1959–1965, 2009.
- [34] D. A. G. Trevino, S. M. Dutta, F. H. Ghorbel, and M. Karkoub, "An improved dipole model of 3-d magnetic flux leakage," *IEEE Transactions on Magnetics*, vol. 52, no. 12, pp. 1–7, 2015.
- [35] T. P. Lillicrap, J. J. Hunt, P. Alexander et al., "Continuous control with deep reinforcement learning," 2015, <http://arxiv.org/abs/1509.02971>.
- [36] D. P. Kingma and B. Jimmy, "Adam: a method for stochastic optimization," 2014, <http://arxiv.org/abs/1412.6980>.
- [37] D. Silver, L. Guy, N. Heess et al., "Deterministic policy gradient algorithms," in *Proceedings of the International Conference on Machine Learning*, pp. 387–395, PMLR, New York, NY, USA, July 2014.