

Full length article

Inverse design of energy-dissipating steel plate walls based on self-supervised diffusion models

Tianyang Zhang, Yinxiao Lv, Tong Liu, Weizhi Xu*, Shuguang Wang, Dongsheng Du, Aiguo Zhao

College of Civil Engineering, Nanjing Tech University, Nanjing, China

ARTICLE INFO

Dataset link: <https://github.com/AshenOneme/SSIDM>

Keywords:

Data-driven
Deep learning
Diffusion model
Seismic resistance
Inverse design

ABSTRACT

This research proposes a novel energy-dissipating system integrating non-structural infill walls and energy-dissipating steel plate walls (EDWs) to improve the seismic performance of frame-infill wall structures. Based on the diffusion model, a self-supervised inverse design model (SSIDM) is developed to enhance the design accuracy of complex EDWs. SSIDM is developed by constructing a pre-trained hysteresis performance prediction network (HPPN) and a microstructure generation network (DiffEDW). During the microstructure generation process, a discriminator is employed to predict the overall mechanical performance of the EDW. In cases where the mechanical performance does not meet the expected demands, additional Gaussian random noise is introduced, and the generation process is repeated until an EDW structure that satisfies the hysteresis performance target is obtained. SSIDM learns the conditional distribution of microstructures corresponding to a given complete hysteresis performance, enabling a one-to-many mapping from attributes to geometry. Ablation experiments demonstrate that the self-supervised diffusion model with the HPPN generates EDWs with smaller errors in the finite element analysis (FEA) results and mechanical performance targets compared to the model without the HPPN. This study reveals that the self-supervised method offers significant advantages in generating periodic large-scale complex topologies and demonstrates potential for accelerating multi-scale structure generation. The dataset and related code are available at <https://github.com/AshenOneme/SSIDM>.

1. Introduction

Earthquakes are characterized by a sudden onset, high destructiveness, and severe secondary disasters. Field investigations of earthquake damage [1–5] have shown that frame-infill wall structures often fail to form a unified energy-dissipating and seismic resistance mechanism. The deformation pattern is difficult to control, and the damage mechanism is suboptimal, resulting in structural component failure and severe damage to non-structural components under rare seismic events.

Under seismic loading, the deformation and damage of frame structures and the components are typically concentrated at specific floors, while other floors and components may only exhibit minor damage. To enhance the seismic performance of structures, energy-dissipating and damping technologies can be implemented. By installing energy-dissipating devices, the structure can achieve additional stiffness [6–10], load-bearing capacity, and energy-dissipation capacity [11–15], thereby effectively controlling the seismic response.

The current development trend of energy-dissipating devices is moving towards multi-stage designs with small space requirements. Researchers aim to optimize the performance of existing damping devices under both frequent and rare seismic events, leading to innovations in damper structures. Most of these innovations involve the use of series and parallel arrangements of metallic and friction components to design multi-stage dampers [16–18]. Multi-stage energy-dissipating dampers are beneficial for enhancing the seismic redundancy of structures and enabling coordinated control of structural responses under different seismic levels. However, such energy-dissipating devices are often large in volume, resulting in low spatial efficiency. Additionally, existing dampers are typically designed solely for structural components, without considering more complex seismic damage scenarios.

Infill walls are essential components of building structures, and due to their large quantity and widespread presence, they account for a significant portion of the overall investment in the building. In past seismic events, infill walls have experienced considerable damage

* Corresponding author.

E-mail address: xuwz@njtech.edu.cn (W. Xu).

[19–22]. Traditional energy-dissipating technologies, such as dampers, can effectively address damage to structural components under rare seismic events; however, dampers do not provide protection for non-structural infill walls. For the reinforcement or repair of masonry infill walls, external reinforcement components are commonly used. Currently, textile-reinforced concrete [23–27] or metal components [28–31] are frequently employed to enhance the in-plane or out-of-plane seismic performance of infill walls. However, the commonly used methods for enhancing the seismic performance of masonry infill walls primarily focus on improving the load-bearing capacity, integrity, and ductility of the walls [27]. Although these approaches can effectively prevent damage during seismic events, they do not provide sufficient energy dissipation [28], and therefore, the overall structural response is not reduced.

In response to the aforementioned issues, specifically the fact that traditional energy-dissipating and vibration control techniques do not account for the protection of non-structural components and that conventional reinforcement methods for non-structural components lack inherent energy-dissipating capabilities, recent advancements have introduced the concept of integrating energy-dissipating mechanisms with non-structural components. The damped infill wall, which integrates both energy-dissipating and infill wall functions, is designed by introducing a damping layer with low shear stiffness between subwalls in a partition wall system. Under seismic excitation, relative motion occurs between adjacent subwalls, causing shear deformation in the damping layer, which dissipates energy through viscoelastic mechanisms or metallic yielding. Lu et al. [32] employed shear-type metallic dampers in the central region of vertically segmented masonry walls. Under low-cycle reversed loading, the damage development and strength degradation of the infill wall were effectively suppressed, and its energy-dissipating performance was significantly enhanced. Zhou et al. [33–36] proposed a method for the lateral subdivision of masonry infill walls and the incorporation of a damping layer within the wall. Research findings indicate that the use of an asphalt membrane damping layer effectively concentrates the shear deformation of the entire wall at the location of the damping layer, resulting in improved energy-dissipating capacity of the infill wall and reduced damage.

Although the integration of energy-dissipating devices with infill walls can reduce seismic damage to the walls and enhance the seismic performance of the structure, the design methods for these energy-dissipating parts are not sufficiently straightforward. When the infill wall and energy-dissipating part are combined into a single unit, both components become relatively complex, leading to challenges in the design of the energy-dissipating and load-bearing capacity. In recent years, with the development of artificial intelligence (AI) technologies, particularly the introduction of AI-generated content (AIGC) [37–40], the reverse design of energy-dissipating devices has become more convenient. Vlassis [41] employed a denoising diffusion algorithm to identify microstructures with nonlinear fine-tuning characteristics. This algorithm not only enables the reverse design of nonlinear microstructures but also facilitates the learning of nonlinear structure-property mapping relationships. Bastek et al. [42] deviated from the common strategy of directly learning mappings from properties to design microstructures, extending the framework to incorporate the expected deformation paths and full-field stress distributions of the structure, in agreement with finite element simulations. Wang et al. [43] proposed a metamaterial inverse design diffusion model, which can be customized based on the desired stress-strain curve to generate microstructures. Although the aforementioned studies indicate that energy-absorbing microstructures can be generated through diffusion models, the microstructures generated in these studies are relatively simple, and their application is mostly limited to impact energy dissipation, where the structures are designed to withstand unidirectional loading. This study, addressing seismic loading, aims to achieve full hysteresis energy dissipation as the mechanical performance target and proposes a self-supervised diffusion model to generate larger-scale

energy-dissipating steel plate wall (EDW) structures.

2. The deficiencies in enhancing energy-dissipating capacity of building structures

For frame structures, the overall energy-dissipating capacity is often insufficient, requiring energy-dissipating enhancement to improve seismic performance. Traditional methods for enhancing energy dissipation typically involve the addition of dampers to the structure. However, in frame structures, infill walls perform important functional roles, such as space partitioning and housing embedded utilities. As a result, the installation of either metal or viscous dampers can hinder the design of infill walls, particularly masonry infill walls, where large-volume dampers are difficult to conceal.

Therefore, for structural systems with infill walls, particularly in this study, a novel energy-dissipating system (NEDS) is proposed, as shown in Fig. 1. This system integrates the infill wall with energy-dissipating elements, where the steel EDWs provide both load-bearing and energy-dissipating capacities, while the masonry infill wall primarily serves as a partition. The advantage of this system lies in its ability to enhance the seismic performance of the structure while effectively restraining the out-of-plane deformation of the masonry infill wall, resulting in a uniform improvement in the seismic resilience of both structural and non-structural components. Although the NEDS integrates non-structural infill walls and EDWs without occupying the usable space of the building, the energy-dissipating and load-bearing capacities of the structure require specialized design. For an EDW with the same thickness, its yield load capacity, energy-dissipating capacity, and residual deformation can only be adjusted through the design of its geometric configuration. However, the design of the EDW is relatively complex, and the finite element analysis (FEA) process is time-consuming. Therefore, the application of traditional design methods to NEDS is not feasible.

The structural inverse design algorithm based on generative AI models can address the issue of inefficiency in traditional structural configuration design methods. However, for larger and more complex structures, results generated by previous generative models often fail to accurately meet the established mechanical performance objectives, and the generated structures may lack regularity. To address the aforementioned issues, this research proposes a self-supervised inverse design model (SSIDM) for energy-dissipating structures, which can more accurately meet the established mechanical performance targets for EDWs.

3. Methodology

The diffusion model is a type of generative model that models the distribution $p(x)$, of high-dimensional data. Unlike likelihood-based models, which directly estimate $p(x)$, diffusion models attempt to estimate the score function $\nabla_x \log p(x)$. When using a diffusion model to generate samples, random noise is initially input, and the noise is progressively removed. A commonly used denoising method is Stochastic Gradient Langevin Dynamics (SGLD).

3.1. DiffEDW

From a methodological perspective, there exist two approaches to conditional control generation, namely classifier-guidance [44] and classifier-free scheme [45]. Classifier-guidance introduces an additional classifier and uses gradient-based guidance for conditional generation, which incurs a high computational cost. Therefore, the SSIDM proposed in this study adopts a classifier-free scheme for the diffusion model component.

In conditional-free diffusion model for EDW (DiffEDW) generation, an additional mechanical property y is introduced. The objective is to model the conditional probability distribution $p(x_t|y)$. Using Bayes'

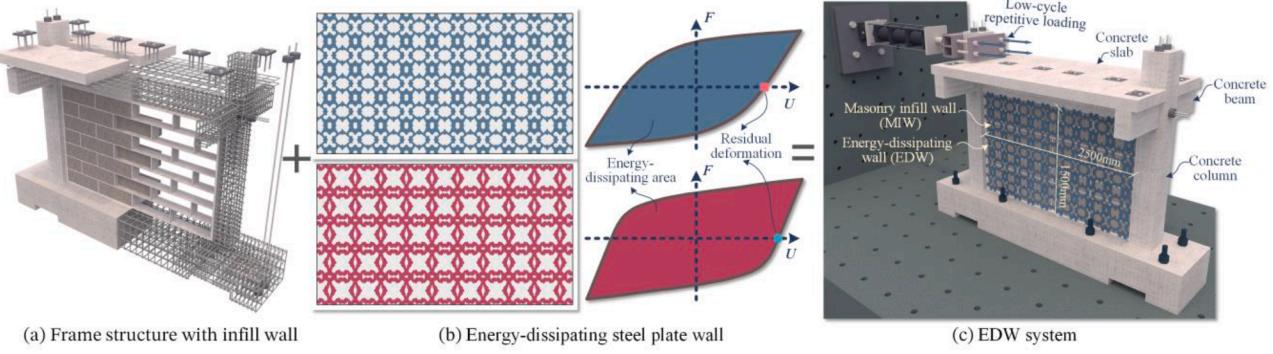


Fig. 1. The energy-dissipating system.

theorem and the score function, the gradient expression of the conditional distribution can be derived [46]. Specifically, the gradient of the conditional distribution $p(x_t|y)$ can be computed as follows:

$$\nabla_{x_t} \log p(x_t|y) = \nabla_{x_t} \log p(x_t) + \nabla_{x_t} \log p(y|x_t) \quad (1)$$

where $\nabla_{x_t} \log p(x_t)$ is the unconditional score, which is typically associated with the diffusion model and is used to capture the distributional characteristics of the data. $\nabla_{x_t} \log p(y|x_t)$ is the adversarial gradient of the conditional probability. In practical applications, an additional parameter, referred to as the guidance scale s , is introduced to adjust the strength of the control. The final result is the weighted combination, with s serving as the weighting factor:

$$\begin{aligned} \nabla_{x_t} \log p(x_t|y) &= \nabla_{x_t} \log p(x_t) + s \nabla_{x_t} \log p(y|x_t) \\ &= \nabla_{x_t} \log p(x_t) + s(\nabla_{x_t} \log p(x_t|y) - \nabla_{x_t} \log p(x_t)) \\ &= (1-s)\nabla_{x_t} \log p(x_t) + s\nabla_{x_t} \log p(x_t|y) \end{aligned} \quad (2)$$

The joint probability of the reverse process for the original unconditional diffusion model is:

$$p_\theta(x_{0:T}) = p(x_T) \prod_{t=1}^T p_\theta(x_{t-1}|x_t) \quad (3)$$

After the addition of the supplementary mechanical performance guidance condition y , the joint probability of the reverse process for DiffEDW is:

$$p_\theta(x_{0:T}|y) = p(x_T) \prod_{t=1}^T p_\theta(x_{t-1}|x_t, y) \quad (4)$$

Algorithm 1

Training and sampling with DiffEDW.

Input: $D = \{(x_i, y_i)\}_{i=1}^N$ (training dataset), η (learning rate), T (epochs), β_t (noise schedule)
Output: θ (trained model parameters), \hat{x}_0 (generated samples)

1. **Forward process:**
 - 1.1 Initialize the denoising model $e_\theta(x_t, y)$, and set parameters θ
 - 1.2 **For** $t = 1$ to T **do:**
 - Sample paired data: $(x_0, y) \sim D$
 - Sample a noise level: $t \sim \text{Uniform}(1, T)$
 - Generate noisy data: $x_t = \sqrt{\alpha_t}x_0 + \sqrt{1-\alpha_t}\epsilon$, $\epsilon \sim \mathcal{N}(0, I)$
 - Compute loss: $L = \mathbb{E}_{x_0, y, t, \epsilon} [\|\epsilon - e_\theta(x_t, y)\|_2^2]$
 - Update parameters: $\hat{\theta} \leftarrow \theta - \eta \cdot \nabla_\theta L$
2. **Reverse process:**
 - 2.1 Input: y (Target hysteresis performance), s (DiffEDW scale)
 - 2.2 Initialize: $x_T \sim \mathcal{N}(0, I)$ (sample Gaussian noise)
 - 2.3 **For** $t = T$ to 1 **do:**
 - a. Predict noise with condition: $\epsilon_{cond} = e_\theta(x_t, y)$
 - b. Predict noise without condition: $\epsilon_{uncond} = e_\theta(x_t, \emptyset)$
 - c. Combine predictions using DiffEDW: $\epsilon_{final} = (1+s)\cdot\epsilon_{cond} - s\cdot\epsilon_{uncond}$
 - d. Denoise: $x_{t-1} = \frac{1}{\sqrt{\alpha_t}}(x_t - \sqrt{1-\alpha_t}\cdot\epsilon_{final}) + \sigma_t z$, $z \sim \mathcal{N}(0, I)$
3. **Return:** $\hat{x}_0 = x_0$ (final denoised sample)

$$p_\theta(x_{t-1}|x_t, y) = \mathcal{N}(x_{t-1}; \mu_\theta(x_t, y, t), \sigma_\theta(x_t, y, t)) \quad (5)$$

where $\mu_\theta(x_t, y, t)$ and $\sigma_\theta(x_t, y, t)$ represent the mean and standard deviation of the conditional Gaussian distribution, respectively. These are typically the outputs of a neural network and are computed based on the current time step t , the condition y , and the current variable x_t . μ_θ represents the mean of the reverse process, typically predicted by a neural network. The model, through learning, attempts to regress from the current state x_t to the previous state x_{t-1} . σ_θ represents the standard deviation of the reverse process, typically describing the uncertainty of the model. It is a predefined value that is associated with the time step t . The training and sampling procedure of DiffEDW is presented in [Algorithm 1](#).

3.2. SSIDM

Although DiffEDW can generate geometric configurations that approximately meet the target mechanical performance criteria, it often lacks accuracy in generating fine-grained structures. This study proposes an innovative SSIDM framework, in which a pre-trained mechanical performance prediction model is utilized as a discriminator to predict the mechanical properties of EDWs during the generation process. The training procedure of the hysteresis performance prediction network (HPPN) is presented in [Algorithm 2](#). These predicted properties are then compared with the target mechanical performance. In cases where the error is significant, noise is reintroduced into the generated EDWs, guiding the mechanical properties of the resulting structures to approach the target performance, as shown in [Fig. 2](#).

Algorithm 2

Training HPPN.

Input: $D = \{(x_i, z_i)\}_{i=1}^N$ (training dataset), η (learning rate), T (epochs), B (batch size)

Output: $\hat{\theta}$ (trained model parameters)

1. Initialization:

1.1 Define HPPN: $f(x; \theta)$ 1.2 Initialize parameters: θ 1.3 Define MSE loss function: $L(z, \hat{z}) = \frac{1}{n} \sum_{i=1}^n (z_i - \hat{z}_i)^2$

1.4 Set up optimizer: AdamW

2. Training process:

For $t = 1$ to T do:2.1 For each mini-batch (x_{batch}, z_{batch}) in D :a. Forward pass: $\hat{z}_{batch} = f(x_{batch}; \theta)$ b. Computer loss: $L_{batch} = \text{MSELOSS}(z_{batch}, \hat{z}_{batch})$

c. Backward pass:

i. Zero the gradient buffers in the optimizer

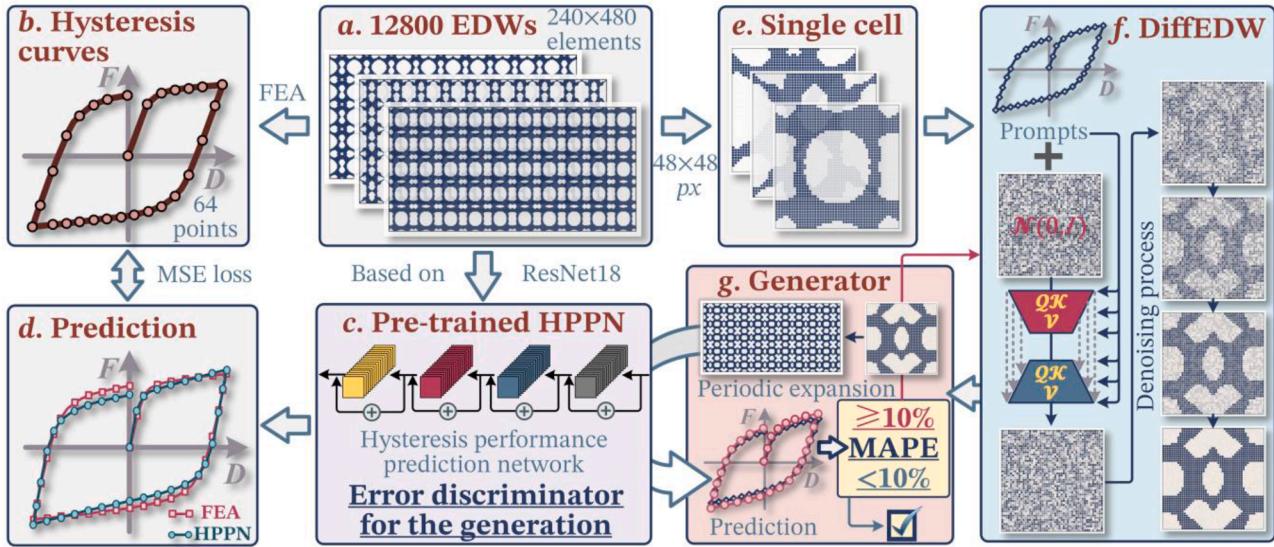
ii. Compute the gradients of the loss w.r.t. the HPPN parameters: $\nabla_{\theta} L_{batch}$ d. Update parameters: $\hat{\theta} \leftarrow \theta - \eta \cdot \nabla_{\theta} L_{batch}$ 3. Return: $\hat{\theta}$ 

Fig. 2. The structure of the self-supervised inverse design model.

After training a relatively accurate discriminator, the integration of HPPN and DiffEDW models in a complete self-supervised diffusion model framework, as illustrated in [Algorithm 3](#). In this study, MPAE is employed to quantify the discrepancy between the hysteresis performance predicted by HPPN and the target mechanical properties of the configurations generated by DiffEDW. The MPAE threshold is set to 10 %, such that configurations with an MPAE below 10 % are output. Notably, the results generated by DiffEDW are 48×48 -dimensional

tensors, which are periodically replicated to form 240×480 -dimensional tensors before being input into HPPN for prediction, as shown in [Fig. 2](#).

4. Dataset construction**4.1. EDW structure**

Based on Gaussian random field sampling [42], a threshold was set

Algorithm 3

Generating with SSIDM.

Input: y (Target), s (DiffEDW scale), τ (MAPE threshold), N_{max} (maximum attempts)

Output: \hat{x} (accepted sample)

1. Initialization:

1.1 Define generator: DiffEDW

1.2 Define discriminator: HPPN

1.3 Set Gaussian noise: $e \sim \mathcal{N}(0, \sigma^2)$

2. Generation and evaluation loop:

For $n = 1$ to N_{max} do:2.1 Generate EDW: $\hat{x} \leftarrow \text{DiffEDW}(y, s) + e$ 2.2 Hysteresis performance prediction: $\hat{y} \leftarrow \text{HPPN}(\hat{x})$ 2.3 Compute the MAPE: $\text{MAPE} \leftarrow \frac{1}{M} \sum_{i=1}^M \left| \frac{y_i - \hat{y}_i}{y_i} \right| \times 100\%$ 3. Return: if $\text{MAPE} < \tau$, return \hat{x}

for the random field. Surfaces that lay above the threshold were designated as the desired portions, while surfaces below the threshold were discarded. This method was employed to construct the basic unit cell structure of the wall, ensuring that the opposite boundaries of the unit cells were connected. Subsequently, the pattern was mirrored sequentially along the two edges of the unit cell to generate a complex periodic wall structure. The use of Gaussian random fields to generate the unit cells provided the advantage of obtaining a large number of EDW structures while maintaining both the connectivity and spatial distribution of the unit cells. By adjusting the threshold, the volume fraction could be controlled to approximately 50 %, as shown in Fig. 3. Each unit cell was designed to be 48×48 pixels, and by replicating the pattern five times in one direction and ten times in the other, the overall dimensions of the EDWs reached 240×480 pixels. Using this approach, a total of 12,800 random steel EDW structures were generated, with the steel material proportion ranging between 40 % and 60 % of the total volume.

4.2. The mechanical properties of the EDWs

After generating the structure of the EDWs, the corresponding pattern pixels were mapped to the finite element mesh, resulting in an EDW comprising 115,200 elements. Shell elements were used to model the EDWs, with a thickness of 5 mm and a material with a nominal yield strength of 235 MPa. A cyclic shear hysteresis deformation was applied to the shell elements, with a maximum drift set to 2 %, as shown in Fig. 4. FEA was performed for all EDWs to obtain the corresponding structures and hysteretic behavior at a 2 % drift. The hysteresis curve data consisted of 64 coordinate points, from which the mechanical properties of the EDWs, such as yield displacement, yield load capacity, energy-dissipating capacity, and stiffness were calculated.

The volume fraction of the EDWs and the distribution of the mechanical properties are shown in Fig. 5. The median volume fraction is 47.22 %, which is in close agreement with the setting defined during the dataset construction process, thereby demonstrating the feasibility of the dataset generation method. A linear regression analysis was performed on 12,800 sets of EDWs to examine the variation trends of mechanical properties under different volume fractions. The positive and negative correlations between hysteretic performance and volume fraction were determined based on the weight coefficients derived from the linear regression model. In general, the yield displacement of EDWs exhibits a negative correlation with the volume fraction, while the yield force shows a positive correlation. The stiffness of EDWs (K_1) is positively correlated with the volume fraction; however, the distribution of stiffness after yielding (K_2) is nearly independent of the volume fraction. This suggests that it is challenging to accurately define the mechanical behavior of EDWs, particularly the hysteresis behavior, solely by controlling the volume fraction.

Even under the same volume fraction, the distributions of indicators such as yield and stiffness are highly scattered, presenting challenges for

the generation algorithm. Since the structure generated by DiffEDW represents a single unit cell of the EDW, it is related to the overall EDW but cannot accurately characterize its overall performance. Therefore, after combining the generated structures into the complete EDW, it is necessary to introduce a discriminator for further supervised generation.

5. Discriminator establishment and inference

5.1. HPPN pretraining process

The discriminator is used to supervise the generation of EDWs, and thus, it is necessary to balance the prediction accuracy and prediction speed of the mechanical properties of EDWs. In this study, eight representative deep learning models were pre-trained using 90 % of the data from the dataset for training, with 10 % used for testing. The specific training process is outlined in Algorithm 2. The errors of the eight representative networks on the training and test sets during the training process are shown in Fig. 6. All eight models were trained for 200 epochs with a batch size of 32. The AdamW [47] optimizer was used, with an initial learning rate of 1×10^{-3} and an exponential learning rate decay with a decay factor of 0.98. An L2 regularization term was added to the loss function to mitigate overfitting, with the regularization coefficient set to 1×10^{-4} .

Floating point operations (Flops) is an important metric for measuring the computational complexity of a model, representing the number of floating point operations required for a single forward inference. It is used to assess the inference efficiency of the model. Transformer-based models such as ViT and SwinT have a significantly large number of learnable parameters, resulting in a high Flops value. However, the RMSE on the test set is relatively high, indicating that these transformer-based deep learning models exhibit poor adaptability to EDWs. VGG16, as a classical convolutional neural network, achieves the lowest RMSE. However, considering that the model contains over 100 million learnable parameters and has a high memory usage during both training and inference, it is not suitable as a discriminator in terms of computational efficiency. ResNet18, ResNet50, and ResNet101, which are based on residual connections, exhibit similar RMSE values. Notably, ResNet18 has the smallest number of parameters among all eight models and also the lowest Flops, as shown in Table 1. The number of parameters in ResNet18 is 8.18 % of VGG16 and 26.55 % of ResNet101, yet it achieves relatively high accuracy, second only to VGG16. Therefore, in this study, the pre-trained ResNet18-based HPPN was selected as the discriminator component of the SSIDM.

5.2. HPPN prediction results

The prediction results of certain EDWs on the test set using the ResNet18-trained HPPN are shown in Fig. 7. The red curve represents the finite element analysis results, and the blue curve represents the

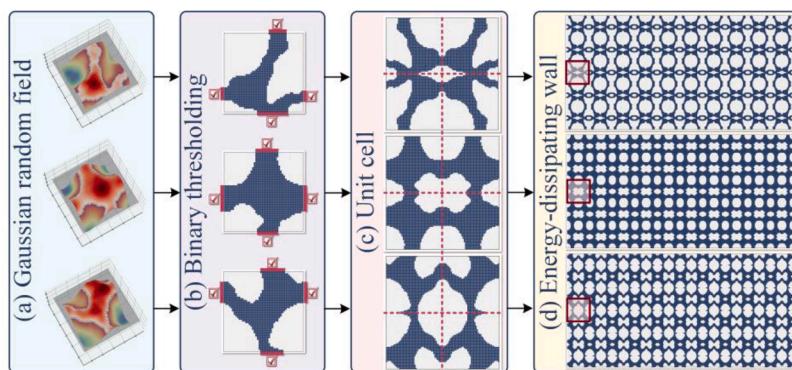


Fig. 3. The process of establishing the EDW structures.

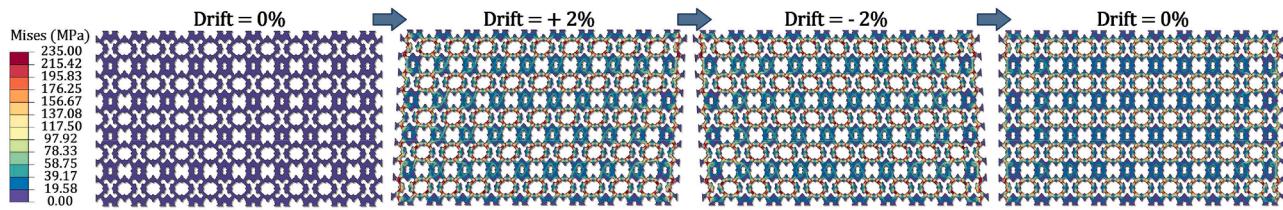


Fig. 4. Low cyclic hysteresis loading.

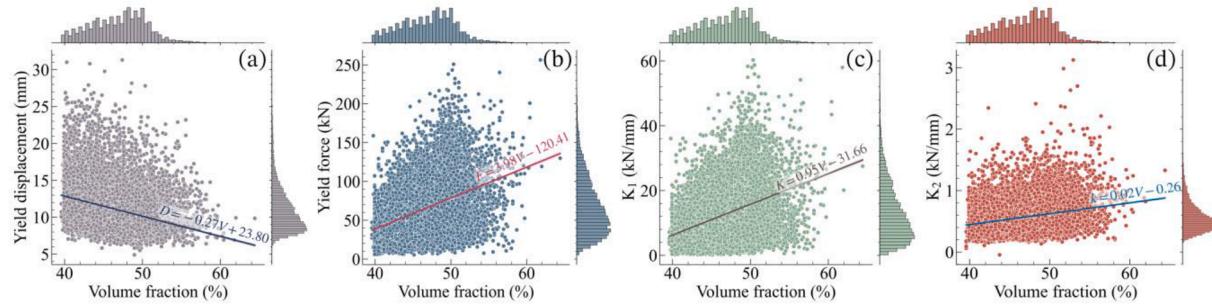


Fig. 5. The distribution of the mechanical properties of EDWs.

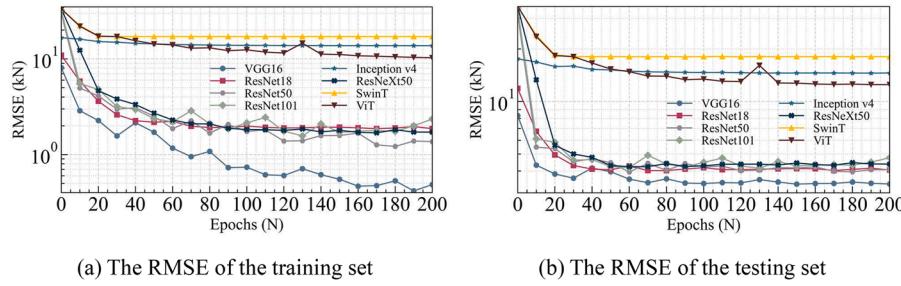


Fig. 6. The variation of the loss function during the training process.

Table 1
Model parameters and test error.

Model	Params	Memory (MB)	Madd (G)	Flops (G)	MemR+W (MB)	RMSE (kN)
VGG16 [48]	138,421,457	250.89	70.49	35.3	1034.24	3.37
ResNet18 [49]	11,318,337	52.45	8.12	4.06	149.79	4.06
ResNet50 [49]	23,634,945	252.61	18.72	9.38	595.69	4.07
ResNet101 [49]	42,627,073	372.14	35.81	17.93	907.20	4.74
Inception v4 [50]	321,473	139.25	16.52	8.26	342.84	14.56
ResNeXt50 [51]	23,106,817	310.32	19.38	9.71	709.10	4.38
SwinT [52]	48,836,513	376.99	486.67	246.27	1331.20	18.06
ViT [53]	85,260,353	554.78	208.98	104.95	651.79	12.51

HPPN predictions. Overall, the structural differences of EDWs are significant, and there are considerable variations in their mechanical properties, particularly in terms of energy-dissipating capacity and load-bearing capacity. As shown in Fig. 7(i) and (k), the volume fractions of the two are 47.05 % and 51.22 %, respectively. However, it is evident that the hysteresis loop in Fig. 7(k) has a smaller energy-dissipating area, and the maximum load-bearing capacity is reduced by 30.48 %. The HPPN provides accurate predictions of the hysteresis curves for EDWs, with the exception of relatively small errors during the unloading phase of certain curves. The loading phase aligns closely with the FEA results. When the load-bearing capacity of an EDW is larger, the RMSE of the predictions is also relatively higher, as the RMSE shares the same units as the load-bearing capacity, leading to a corresponding increase in absolute error. The maximum RMSE in Fig. 7 is 8.81, while the minimum

is 1.29, with an overall RMSE of 4.06 for the test set, all of which are relatively low values. Since RMSE is a dimensional quantity, the relative differences across different EDWs are significant. Therefore, during the self-supervised generation process of SSIDM, the dimensionless MAPE is used as the evaluation metric. Specifically, the HPPN training process employs MSE as the loss function, while the inference process utilizes MAPE to assess the error. Comprehensive predictions were performed on the test set using HPPN, focusing on four key mechanical property metrics: stiffness, energy dissipation, yield force, and maximum force, as illustrated in Fig. 8. With the exception of a few outlier samples where HPPN exhibited slightly larger prediction deviations, the overall prediction accuracy was high in other cases, with R^2 values exceeding 0.9. This suggests that HPPN is capable of functioning as a discriminator for the complete hysteresis behavior of EDW, thereby supervising the

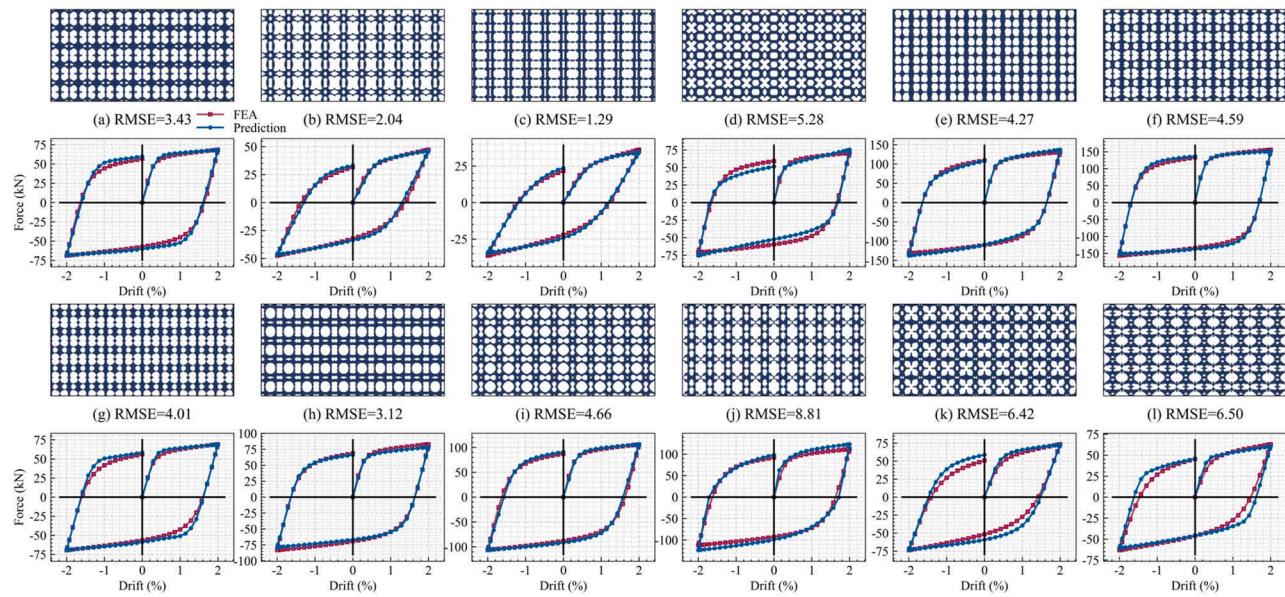


Fig. 7. Hysteresis curve prediction results.

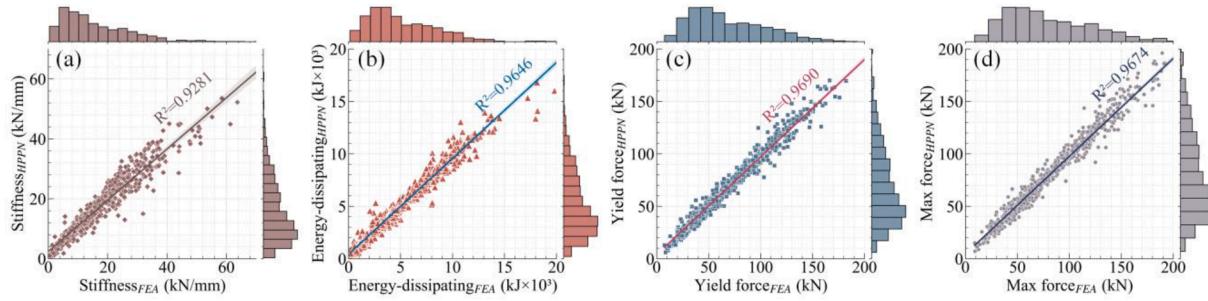


Fig. 8. Comparison of predictive hysteresis performance.

generation of EDW.

Gradient-weighted Class Activation Mapping (Grad-CAM) is a visualization technique used in computer vision tasks [54]. It generates

activation maps for specific target classes by utilizing gradient information, thereby aiding in the understanding of the decision-making process of deep neural networks. By visualizing the regions that the

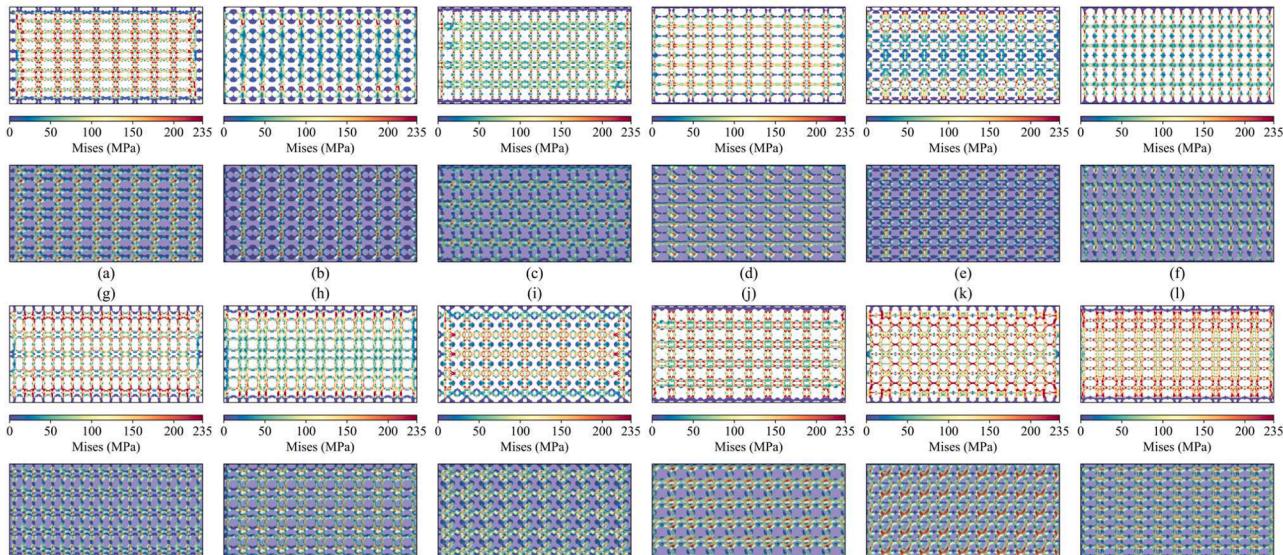


Fig. 9. Stress distribution and Grad-CAM results.

network focuses on, Grad-CAM reveals the image features that the model relies on. The stress distribution states from FEA of 12 different EDW structures, along with the Grad-CAM results from the HPPN prediction process, are illustrated in Fig. 9. This study found that the Grad-CAM results for the EDWs exhibit a certain correlation with the finite element simulation results. As shown in Fig. 9(b), the finite element simulation indicates that large areas of the EDWs have stress values approaching zero, while the Grad-CAM results suggest that the HPPN prediction process rarely utilizes these low-stress regions. Furthermore, as shown in Fig. 9(f), the central region of each unit cell structure experiences relatively low bending moments, resulting in smaller stresses, whereas the edge regions contribute more to the overall load-bearing and energy-dissipating capabilities of the EDW. The trained HPPN primarily focuses on the regions that contribute significantly to the mechanical performance of the EDW, thereby enabling accurate prediction of the hysteresis curve.

The Grad-CAM results indicate that the HPPN is capable of effectively learning the physical principles governing the EDW hysteresis process, thus enabling accurate prediction of the hysteresis curve. Therefore, the HPPN demonstrates feasibility as an SSIDM discriminator from both the perspective of curve values and physical principles. In the subsequent generation process, the use of an HPPN based on ResNet18 to establish a self-supervised generative model is proven to be effective.

6. Generation of EDWs based on the SSIDM

In this research, both the HPPN and DiffEDW employ a distributed training mode, which is followed by integration. The HPPN predicts the complete EDWs, while the DiffEDW generates individual unit cell structures. Although the mechanical properties of the complete EDWs are related to those of the individual unit cells, the mapping relationship of the mechanical properties is highly complex and cannot be described as a simple additive relationship. The mechanical behaviors between unit cells are coupled and mutually influential. Therefore, after DiffEDW generates individual unit cells, the process of assembling them into a complete EDW requires the introduction of the HPPN for prior assessment. The integration of the generator and discriminator in the SSIDM results in more accurate generation outcomes.

6.1. Training and inference of DIFFEDW

A network architecture based on a conditional diffusion model [45] is employed, with the core generator network utilizing an improved two-dimensional conditional UNet [55]. The design of this network aims to achieve efficient denoising and generation through multi-scale feature extraction and conditional modeling. The input to the model is a two-dimensional, single-channel unit cell structure, and the output is the predicted noise component. The entire network architecture can be divided into two parts: the encoder and the decoder, with multi-scale feature transfer facilitated by skip connections. The encoder consists of multiple downsampling blocks, which are responsible for progressively extracting high-level features from the input image. Each down-sampling block is composed of four residual blocks, which increase the feature representation capability through feature channel stacking. During the downsampling process, certain modules incorporate a spatial self-attention mechanism to enhance the network's ability to perceive global features. The decoder reverses the process applied by the encoder, gradually restoring the spatial dimensions of the image through multiple upsampling blocks. The decoder modules correspond to the encoder and gradually reduce the number of feature channels, maintaining a symmetric structure. Similarly, some upsampling modules also introduce the spatial self-attention mechanism. The DiffEDW, implemented based on Algorithm 1, generates unit cell structures through a progressive denoising process. The sampling process starts with randomly initialized samples from a Gaussian noise distribution and, through reverse inference across time steps, gradually removes noise, ultimately producing

samples that approximate the real data distribution. The denoising process during reverse diffusion employs the conditional UNet as the core network of the generative model. At each time step, the network predicts the noise component based on the current state x_t and conditional information, thereby advancing the denoising process, as shown in Fig. 10.

In the training process of DiffEDW, the noise addition process is divided into 1000 steps, with a noise scheduling strategy based on a cosine function. This method smoothly adjusts the noise weight, utilizing the properties of the cosine curve to allow for a more natural distribution of noise during training. DiffEDW has 107,789,313 learnable hyperparameters, with a total of 100 training epochs and a batch size of 128. The initial learning rate is set to 1×10^{-3} , and, similar to the HPPN training process, an exponentially decaying learning rate is used. The denoising method during inference is based on the denoising diffusion implicit model (DDIM), with the denoising process consisting of 100 steps. Notably, in the subsequent integration of HPPN and DiffEDW into SSIDM, unit cell structure prediction and generation begin at step 50, significantly enhancing denoising efficiency.

6.2. The generation process and results

The generation results of SSIDM under the given mechanical property demands are shown in Fig. 11. Four distinct mechanical property demands were randomly selected, each with markedly different characteristics. The mechanical property requirements in Fig. 11(b) and (c) are represented by relatively full hysteresis curves, indicating a high demand for energy dissipation. In contrast, the requirements in Fig. 11(a) and (d) correspond to lower yield strength demands. In each figure, the red line represents the mechanical property requirements, while the other four colored lines correspond to four different EDW structures generated simultaneously. FEA was performed on the generated EDWs to compare the mechanical properties under each mechanical property target.

Sixteen distinct EDW structures were generated based on four mechanical property targets. Since the generation results were in the form of tensors with values between 0 and 1, a binarization process was applied: regions with a value of 0 represent areas without steel material, while regions with a value of 1 indicate the presence of steel material. The generated results exhibited isolated, disconnected regions, as shown in Fig. 11(a). However, these disconnected regions do not affect the mechanical performance of the EDW. Although some of the results are not fully optimized, overall, the generated structures align with the mechanical property targets. The hysteresis curve of the generated EDW shows the largest mechanical property error during the unloading process, which affects the fullness of the hysteresis curve. This results in a discrepancy between the energy dissipation of the generated structure and the target, as illustrated in Fig. 11(c). The generated result exhibits a maximum MAPE of 11.38 % compared to the target, as shown in Fig. 11(d). The MAPE threshold set during the SSIDM generation process was 10 %. Although the actual generated result slightly exceeds this threshold, the remaining MAPE values are all below 10 %. Overall, the self-supervised framework proposed in this study meets the generation demands.

The generation process of the EDWs under three different mechanical property targets is illustrated in Fig. 12(a) to (c). The red lines represent the mechanical property requirements, while the blue lines indicate the FEA results of the intermediate products during the EDW generation process. When the sampling step is 1, the generated result is entirely Gaussian random noise, with large regions of the EDW remaining disconnected. The EDW generated under hysteretic loading exhibits no energy dissipation. When the sampling step is 40, the periodic characteristics of the EDW become apparent. At this point, the generated result exhibits connectivity, and the FEA results indicate that the EDW possesses preliminary energy-dissipating capacity. When the number of sampling steps is 60, the location of maximum stress in the

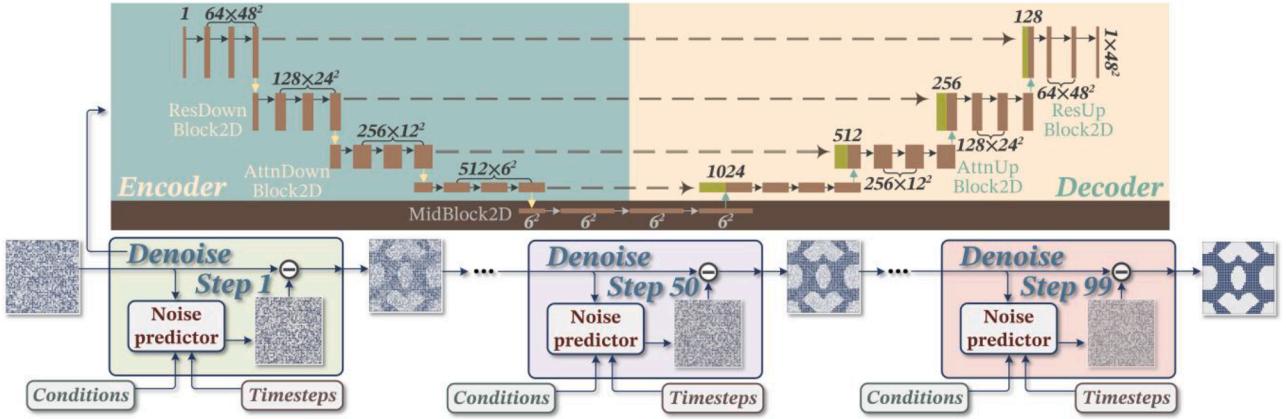


Fig. 10. The generation process of unit cell structures by DiffEDW.

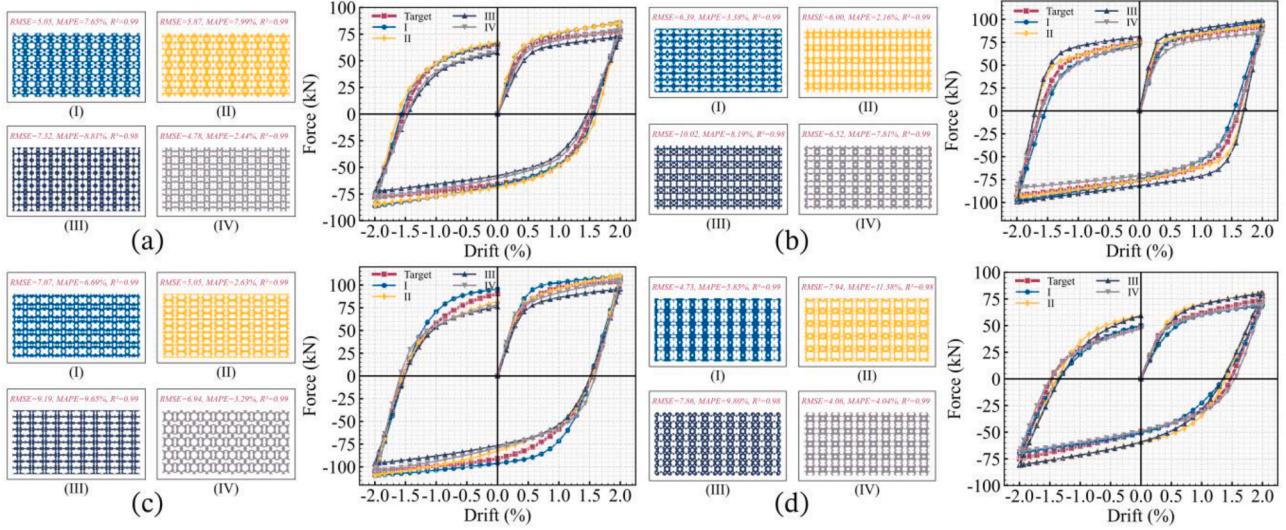


Fig. 11. The generation results of SSIDM.

generated EDWs is already in close agreement with the final result, although the details of the cell structure are not yet fully refined. However, both the energy-dissipating and load-bearing capacity show significant improvement compared to previous steps. When the sampling step is 100, the final generated EDW closely matches the target mechanical properties, with RMSE values of 6.90, 12.33, and 4.49 for the three generated results. The process of generating EDWs using SSIDM shows that with fewer sampling steps, the generated results are less distinct, and the mechanical properties exhibit a larger discrepancy from the target. However, an increase in sampling steps enhances the quality of the generated results, refining the details of the cell structure and reducing the error between the hysteresis curve and the target performance.

6.3. Generated results with and without HPPN

The above results indicate that the SSIDM with the self-supervised module can generate EDW structures with higher accuracy. By removing the HPPN and regenerating the EDW, the hysteresis curves of the EDW generated with and without HPPN are compared, as shown in Figs. 13 and 14. Four EDWs were generated for each group. The blue lines in the figures represent the mean hysteresis curves of the four EDWs, while the yellow areas indicate the coverage of the hysteresis curves.

The generation results indicate that the EDWs produced by the SSIDM without HPPN exhibit more dispersed mechanical properties. Although the mean values are relatively close to the target mechanical properties, the excessive dispersion suggests that the model lacks stability, as shown in Fig. 14(b) and (d). The RMSE and MAPE corresponding to the generation results for the four target mechanical properties in the figures were calculated and are presented in Table 2. After incorporating HPPN as the self-supervised module, the accuracy of SSIDM was significantly improved under different mechanical property requirements. Specifically, the MAPE decreased from 11.28 % to 1.77 %, and the generated results, in terms of both stiffness and energy-dissipating capacity of the overall hysteresis curve, better aligned with the performance demands when HPPN was applied.

7. Discussion

The integration of HPPN into SSIDM not only ensures one-to-many mapping capability from macroscopic hysteresis performance to microscopic topological configurations but also achieves enhanced generative accuracy. Nevertheless, several limitations persist in this study. Firstly, given the substantial dataset scale and computational constraints, necessary simplifications were implemented to ensure numerical convergence. Specifically, material constitutive models were restricted to idealized elastic-plastic formulations, while EDW buckling

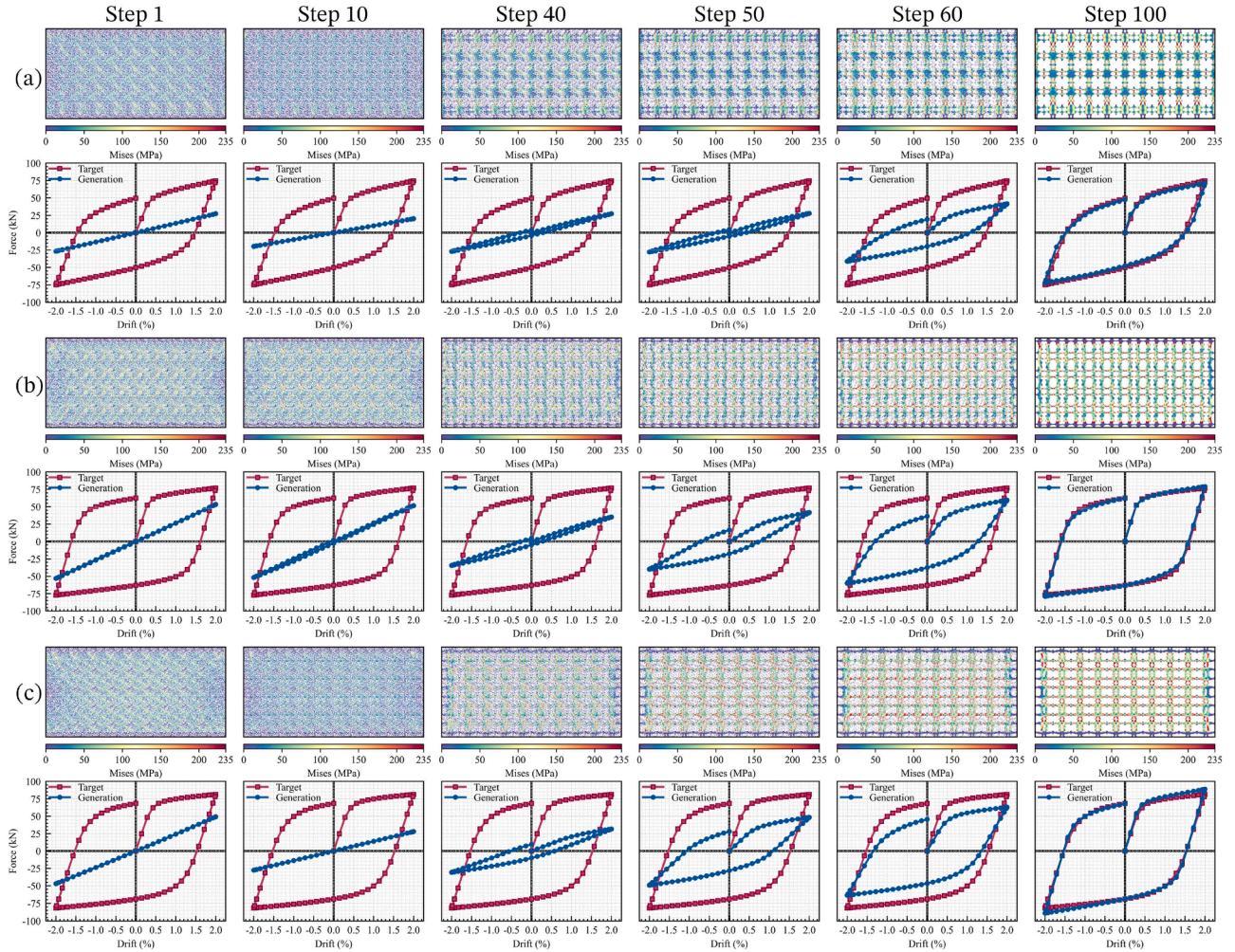


Fig. 12. The generation process of SSIDM.

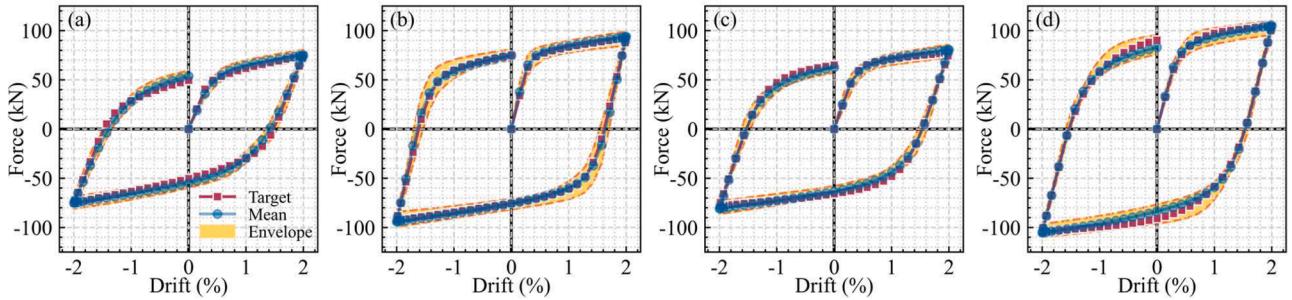


Fig. 13. The generation results with HPPN.

behaviors were excluded. Despite these simplifications, the modeling and computational phases required three consecutive months to complete using an AMD EPYC 9654 96-Core Processor. Secondly, the DifFedDW module employs the established DDPM framework from prior research, inevitably requiring significant computational resources, particularly demanding GPU memory capacity during training phases. Thirdly, the impacts of EDWs on structural dynamic characteristics and their interactions with non-structural infill wall systems have not been thoroughly investigated, which will constitute a primary research focus in subsequent studies to facilitate engineering applications of emerging EDW energy-dissipation systems. Within current hardware limitations, we have maximized dataset diversity and SSIDM parameter

optimization to improve both generative precision and model robustness. Future investigations will explore mechanical performance evolution induced by microstructural scale variations, systematically analyze how complex EDW buckling behaviors influence hysteresis properties, and further refine the existing generative architecture through algorithmic optimizations.

8. Conclusion

This study introduces a novel framework for designing energy-dissipating walls by utilizing an inverse design method based on self-supervised diffusion models. SSIDM is established by integrating a

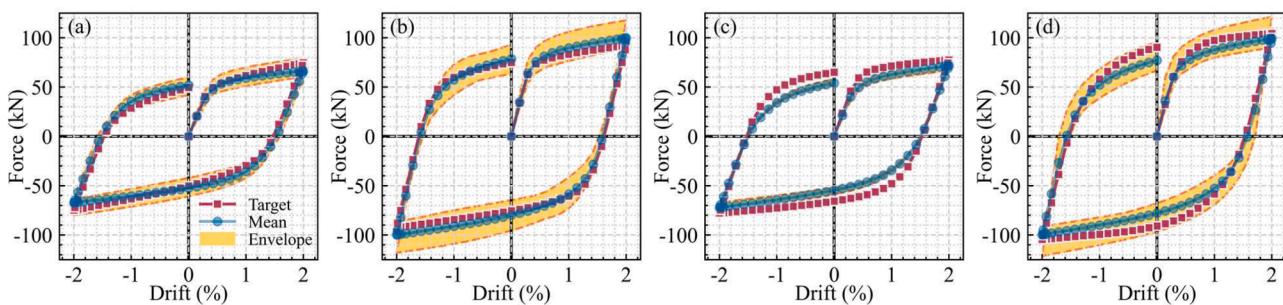


Fig. 14. The generation results without HPPN.

Table 2
Comparison of errors with and without HPPN.

Target	With HPPN		Without HPPN	
	RMSE (kN)	MAPE (%)	RMSE (kN)	MAPE (%)
a	3.38	3.37	5.45	5.72
b	2.06	1.48	5.59	7.08
c	2.48	1.77	8.90	11.28
d	3.43	1.81	8.51	7.61

generator (DiffEDW) and a discriminator (HPPN), which can generate EDW structures that meet the specified mechanical property targets based on the seismic performance demands of the structure. The main conclusions are as follows:

1. A total of 12,800 EDWs were generated based on Gaussian random field sampling, and FEA was conducted to obtain the hysteresis curves of the EDWs. Statistical results show that EDWs with the same volume fraction exhibit significant variability in mechanical performance. However, in general, as the volume fraction increases, the yield displacement decreases and the yield force increases, while the post-yield stiffness shows no correlation with the volume fraction.
2. HPPN was trained using eight classic deep learning models, with the complete EDW geometric information as input to predict its hysteresis behavior. The results indicate that the HPPN, based on ResNet18, can accurately predict the mechanical behavior of EDWs while having a relatively small number of learnable parameters, demonstrating potential as a discriminator in SSIDM.
3. SSIDM used mechanical performance requirements as inputs, and refined unit cell structures were obtained through continuous sampling. By periodically replicating these unit cell structures, the final generated EDW was produced. EDWs generated using SSIDM exhibited a convergence of mechanical performance toward the specified hysteresis target as the number of sampling steps increased.
4. Ablation tests demonstrated that SSIDM with HPPN resulted in an EDW exhibiting a lower error between the mechanical performance obtained through FEA and the specified mechanical performance target compared to the case without HPPN. For the multi-scale generation problem of creating a single unit cell and assembling it into a complete EDW, HPPN significantly enhanced the generation accuracy.

CRediT authorship contribution statement

Tianyang Zhang: Writing – review & editing, Writing – original draft, Software, Methodology, Investigation. **Yinxiao Lv:** Writing – review & editing, Software, Investigation. **Tong Liu:** Software, Investigation, Conceptualization. **Weizhi Xu:** Writing – review & editing, Visualization, Supervision, Resources, Project administration. **Shuguang Wang:** Writing – review & editing, Supervision, Funding acquisition. **Dongsheng Du:** Writing – review & editing, Methodology,

Funding acquisition, Data curation. **Aiguo Zhao:** Writing – review & editing, Supervision, Project administration, Funding acquisition.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

The authors are grateful for financial support from the National Natural Science Foundation of China (52208173) and The Natural Science Foundation of the Jiangsu Higher Education Institutions of China (22KJA560003).

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.tws.2025.113865.

Data availability

<https://github.com/AshenOneme/SSIDM> (Related code is available at)

References

- [1] B. Zhao, F. Tauer, T. Rossetto, Field investigation on the performance of building structures during the 12 May 2008 Wenchuan earthquake in China, Eng. Struct. 31 (2019) 1707–1723.
- [2] K. Sharma, L. Deng, C.C. Noguez, Field investigation on the performance of building structures during the April 25, 2015, Gorkha earthquake in Nepal, Eng. Struct. 121 (2016) 61–74.
- [3] M.L. Ivanov, W.K. Chow, Structural damage observed in reinforced concrete buildings in Adiyaman during the 2023 Turkiye Kahramanmaraş Earthquakes, Structures 58 (2023) 105578.
- [4] O. Ince, Structural damage assessment of reinforced concrete buildings in Adiyaman after Kahramanmaraş (Turkiye) Earthquakes on 6 February 2023, Eng. Fail. Anal. 156 (2024) 107799.
- [5] T.A. Ihsan, A. Celik, A. Kumbasaroglu, H. Yalciner, Assessment of reinforced concrete building damages following the Kahramanmaraş earthquakes in Malatya, Turkey (February 6, 2023), Eng. Sci. Technol. 54 (2024) 101718.
- [6] R. Feng, W. Xu, S. Wang, D. Du, L. Xie, Q. Miao, Seismic response of two-storey steel frame equipped with metallic yielding and viscous dampers under far- and near-field earthquakes: shaking table tests and numerical investigations, Structures 58 (2023) 105538.
- [7] J. Chen, N. Abbas, J. Sun, T. Furuta, Y. Wei, H. Xiong, An innovative energy-dissipation angle bracket for CLT structures: experimental tests and numerical analysis, Eng. Struct. 314 (2024) 118381.
- [8] H. Jiang, L. Xu, X. Xie, G. Zhang, Pseudo-dynamic tests of steel frame with disc spring self-centering energy dissipation braces under doublet earthquake and mainshock-aftershock sequence, Eng. Struct. 316 (2024) 118524.
- [9] X. Jin, W. Liao, Y. Tian, L. Xie, X. Lu, Numerical simulation and seismic performance analysis of sacrificial-energy dissipation beam-column joint, Eng. Struct. 296 (2023) 116979.

- [10] F. Kazemi, N. Asgarkhani, N. Lasowicz, R. Jankowski, Development and experimental validation of a novel double-stage yield steel slit damper-buckling restrained brace, *Eng. Struct.* 315 (2024) 118427.
- [11] Q. Liu, P. Li, C. Yuan, W. Li, S. Ding, Experimental and analytical modelling on a novel self-centering viscous damper, *J. Build. Eng.* 75 (2023) 107020.
- [12] X. Yan, M.S. Alam, G. Shu, Y. Qin, A novel self-centering viscous damper for improving seismic resilience: its development, experimentation, and system response, *Eng. Struct.* 279 (2023) 115632.
- [13] H. Su, L. Zhu, L. Meng, The shaking table test and the numerical analysis on a novel viscous damper installation configuration, *J. Build. Eng.* 83 (2024) 108418.
- [14] B. Rouhani, R. Aghayari, S.A. Mousavi, Fluid viscous dampers in tackle-damper configuration: an experimental study, *Eng. Struct.* 321 (2024) 118927.
- [15] X. Ji, R. Jia, L. Wang, M. Wang, X. Wu, Seismic design and performance assessment of a retrofitted building with tuned viscous mass dampers (TVMD), *Eng. Struct.* 305 (2024) 117688.
- [16] T. Sun, Y. Liu, K. Dai, A. Camara, Y. Lu, L. Wang, Experimental and analytical studies on a novel double-stage coupling damper, *Thin-Walled Struct.* 195 (2024) 111324.
- [17] W. Xu, Y. Cai, T. Zhang, S. Wang, D. Du, L. Wan, Experimental and numerical investigations of a staged energy dissipation device, *Thin-Walled Struct.* 198 (2024) 111721.
- [18] Y. Yang, J. Xue, Z. Luo, R. Liu, Experimental and numerical investigation on a novel dual-stage friction self-centering damper, *Eng. Struct.* 326 (2025) 119606.
- [19] M.S. Ghobadi, R.A. Jazany, H. Farshchi, In situ repair technique of infill masonry walls in steel frames damaged after an earthquake, *Eng. Struct.* 178 (2019) 665–679.
- [20] X. Xie, Z. Qu, H. Fu, L. Zhang, Effect of prior in-plane damage on the out-of-plane behavior of masonry infill walls, *Eng. Struct.* 226 (2021) 111380.
- [21] M. Marinković, C. Butenweg, Numerical analysis of the in-plane behaviour of decoupled masonry infilled RC frames, *Eng. Struct.* 272 (2022) 114959.
- [22] J.K. Bhaskar, D. Bhunia, L. Koutas, In-plane behaviour of masonry infill walls with openings strengthened using textile reinforced mortar, *Structures* 63 (2024) 106439.
- [23] M. Agante, A. Furtado, H. Rodrigues, A. Arêde, P. Fernandes, H. Varum, Experimental characterization of the out-of-plane behaviour of masonry infill walls made of lightweight concrete blocks, *Eng. Struct.* 244 (2021) 112755.
- [24] L. Jing, N. Wang, S. Yin, Shear performance of textile-reinforced concrete (TRC)-strengthened brick masonry walls, *Constr. Build. Mater.* 397 (2023) 132401.
- [25] J.H. Lee, D.Y. Yoo, Enhancement of in-plane shear performance in masonry walls strengthened with high-performance fiber-reinforced cementitious composites, *Constr. Build. Mater.* 409 (2023) 133895.
- [26] B. Wang, S. Yin, F. Qu, F. Wang, Cyclic shear test and finite element analysis of TRC-reinforced damaged confined masonry walls, *Constr. Build. Mater.* 421 (2024) 135766.
- [27] P. Yang, W. Tian, L. Qing, In-plane shear behavior of brick masonry walls strengthened with basalt textile-reinforced concrete, *Constr. Build. Mater.* 401 (2023) 132732.
- [28] M.S. Ghobadi, R.A. Jazany, H. Farshchi, In situ repair technique of infill masonry walls in steel frames damaged after an earthquake, *Eng. Struct.* 178 (2019) 665–679.
- [29] D. Sen, H. Alwashali, M.S. Islam, M. Seki, M. Maeda, Lateral strength evaluation of ferrocement strengthened masonry infilled RC frame based on experimentally observed failure mechanisms, *Structures* 58 (2023) 105428.
- [30] M. Scamardo, S. Cattaneo, P. Crespi, N. Vafa, Cyclic behavior of C-shaped masonry wall retrofitted with twisted bars or bonded rebars, *Constr. Build. Mater.* 443 (2024) 137703.
- [31] M. Vailati, M. Mercuri, A. Gregori, Out-of-plane non-intrusive seismic retrofitting of in-plane damaged masonry infill through 3D printed recycled plastic devices, *Constr. Build. Mater.* 438 (2024) 137256.
- [32] X. Lu, S. Zha, Full-scale experimental investigation of the in-plane seismic performance of a novel resilient infill wall, *Eng. Struct.* 232 (2021) 111826.
- [33] Y. Zhou, Z. Chen, G. Zhong, Investigation on the seismic performance of the masonry infill wall with damping layer joint, *Eng. Struct.* 285 (2023) 115979.
- [34] Z. Chen, Y. Zhou, T. Kinoshita, M.J. DeJong, Distinct element modeling of the in-plane and out-of-plane response of ordinary and innovative masonry infill walls, *Structures* 50 (2023) 1447–1460.
- [35] Z. Chen, Y. Zhou, W. Lie, X. Fang, Macro-modelling method for the in-plane behaviour of the damped masonry infill wall in a frame structure, *J. Build. Eng.* 80 (2023) 108114.
- [36] Z. Chen, Y. Zhou, G. Zhong, Investigation on the influence of previous in-plane loading on the out-of-plane behavior of the innovative damped masonry infill wall, *Eng. Struct.* 308 (2024) 118029.
- [37] D.P. Kingma, M. Welling, Auto-Encoding Variational Bayes, 2013, ab1 arXiv preprint arXiv:1312.6114.
- [38] J. Sohl-Dickstein, E.A. Weiss, N. Maheswaranathan, S. Ganguli, Deep Unsupervised Learning using Nonequilibrium Thermodynamics, 2015, arXiv preprint arXiv: 1503.03585.
- [39] I. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, et al., Generative adversarial networks, *Commun. ACM.* 63 (2020) 139–144.
- [40] A. Nichol, P. Dhariwal, Improved Denoising Diffusion Probabilistic Models, 2021, arXiv preprint arXiv:2102.09672.
- [41] N.N. Vlassis, W. Sun, Denoising diffusion algorithm for inverse design of microstructures with fine-tuned nonlinear material properties, *Comput. Methods Appl. Mech. Engrg.* 413 (2023) 116126.
- [42] J.H. Basteck, D.M. Kochmann, Inverse design of nonlinear mechanical metamaterials via video denoising diffusion models, *Nat. Mach. Intell.* 5 (2023) 1466–1475.
- [43] H. Wang, Z. Du, F. Feng, Z. Kang, S. Tang, X. Guo, DiffMat: data-driven inverse design of energy-absorbing metamaterials using diffusion model, *Comput. Methods Appl. Mech. Engrg.* 432 (2024) 117440.
- [44] S. Mukhopadhyay, M. Gwilliam, V. Agarwal, N. Padmanabhan, A. Swaminathan, S. Hegde, et al., Diffusion Models Beat GANs on Image Classification, 2023, arXiv preprint arXiv:2307.08702.
- [45] J. Ho, T. Salimans, Classifier-Free Diffusion Guidance, 2022, arXiv preprint arXiv: 2207.12598.
- [46] C. Luo, Understanding Diffusion Models: a Unified Perspective, 2022, arXiv preprint arXiv:2208.11970.
- [47] I. Loshchilov, F. Hutter, Decoupled Weight Decay Regularization, 2019, arXiv preprint arXiv: 1711.05101.
- [48] K. Simonyan, A. Zisserman, Very Deep Convolutional Networks for Large-Scale Image Recognition, 2015, arXiv preprint arXiv:1409.556.
- [49] K. He, X. Zhang, S. Ren, J. Sun, Deep Residual Learning for Image Recognition, 2015, arXiv preprint arXiv:1512.03385.
- [50] C. Szegedy, S. Ioffe, V. Vanhoucke, Inception-v4, Inception-ResNet and the Impact of Residual Connections on Learning, 2016, arXiv preprint arXiv:1602.07261.
- [51] S. Xie, R. Girshick, P. Dollar, Z. Tu, K. He, Aggregated Residual Transformations for Deep Neural Networks, 2017, arXiv preprint arXiv:1611.05431.
- [52] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, et al., An Image is Worth 16×16 Words: transformers for Image Recognition at Scale, 2021, arXiv preprint arXiv:2010.11929.
- [53] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, et al., Swin Transformer: hierarchical Vision Transformer using Shifted Windows, 2021, arXiv preprint arXiv:2103.14030.
- [54] R.R. Selvaraju, M. Cogswell, A. Das, R. Vedantam, D. Parikh, D. Batra, Grad-CAM: visual Explanations from Deep Networks via Gradient-based Localization, 2019, arXiv preprint arXiv:1610.02391.
- [55] O. Ronneberger, P. Fischer, T. Brox, U-Net: convolutional Networks for Biomedical Image Segmentation, 2015, arXiv preprint arXiv:1505.04597.