Smart Materials and Structures

PAPER

CrackDiffusion: crack inpainting with denoising diffusion models and crack segmentation perceptual score

To cite this article: Lizhou Chen *et al* 2023 *Smart Mater. Struct.* **32** 054001

IOP Publishing
Smart Materials and Structures
Smart Mater. Struct. 32 (2023) 054001 (10pp)
https://doi.org/10.1088/1361-665X/acc624

CrackDiffusion: crack inpainting with denoising diffusion models and crack segmentation perceptual score

Lizhou Chen[1,2], **Luoyu Zhou**[1,2,*], **Lei Li**[3] **and Mingzhang Luo**[1,2,*]

[1] Electronics and Information School, Yangtze University, Jingzhou, Hubei, People's Republic of China
[2] Institute for Artificial Intelligence, Yangtze University, Jingzhou, Hubei, People's Republic of China
[3] CCDC Drilling and Production Engineering Technology Research Institute, Deyang, Sichuan, People's Republic of China

E-mail: luoyuzh@yangtzeu.edu.cn and lmz@yangtzeu.edu.cn

Received 1 December 2022, revised 15 March 2023
Accepted for publication 21 March 2023
Published 30 March 2023

Abstract

Cracks commonly occur in engineering structures. Imaging method is one of the most effective detection method for crack. However, crack information captured by the imaging sensors is often interfered by noise and the other environmental factors. In this paper, we propose a crack inpainting method that can automatically repair the missing crack information. The inpainting method consists of a denoising diffusion model and a segmentation guidance model. Taking advantages of denoising diffusion model's stability and segmentation guidance model's accuracy, we can achieve coherent inpainting patches as well as accurate crack traces. Furthermore, we propose a fine crack metric—crack segmentation perceptual score to guide high quality crack generation. Experimental results show that our method achieves both high quality and precise crack inpainting results, which is very beneficial to the crack detection and evaluation in structural health monitoring.

Keywords: structural health monitoring, crack inpainting, denoising diffusion models, crack segmentation perceptual score (CSPS), deep learning

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

1. Introduction

Engineering structures are often operated in environments with adverse factors such as vibration [1–3] and corrosion [4, 5], and damages of different degrees often exist in such structures. A large quantity of public resources has been spent in inspection and maintenance of these structures every year. Cracks are one of the most common structural damages. Timely detection and maintenance of these structural damages are critical to the safe operation of the entire structure. Crack detection, as an important research topic in structural health [6, 7], is the first and important step to solve the maintenance problems of systems. Traditionally, crack detection depends on manual inspection which is known to be subjective and labor consuming [8], even sometimes dangerous. With the rapid development in sensors [9–11], communication [12–14], and algorithms [15], recent decades have seen increase development in automatic crack detection. There are two major approaches to automatic structural crack detection: contact based and contactless based. The contact based approach requires the installation of sensors, such as piezoceramic sensors [16–18] and fiber optical sensors [19–21], to measure or estimate the cracks using a variety of algorithms. A disadvantage of the contact based approach is the requirement of installation of sensors, which may not be practical for some

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

applications. This problem can be solved by the contactless approach, which detects structural cracks via imaging methods [22–24]. Another advantage of the contactless method is that it can be easily integrated with unmanned aerial vehicle to automate the inspection process [25–27].

Many types of imaging sensors can be used to capture crack information, however they are inclined to be interfered by noise and complex environmental factors. For example, visual defects often occur due to the bad environment like light insufficient, structural ruin, occlusion, equipment noise and so on [28]. These factors negatively impact crack detection and evaluation. Therefore, it is necessary to develop a method to solve these problems and repair the original information of cracks as much as possible.

Image inpainting techniques aim at reconstructing the missing information of a damaged image. Although crack image segmentation has become a emerging research topic over the past few years, less attention has been paid on crack image inpainting [29]. Due to the vulnerability of information exists everywhere, crack inpainting techniques can play an important role in the structural health monitoring, which is very beneficial to the following crack detection and evaluation.

In this paper, we develop a crack inpainting model (named as CrackDiffusion) with a deep learning and denoising diffusion model-based method. Firstly, we propose and train a denoising diffusion implicit model (DDIM) for generating high quality crack images. Meanwhile, crack segmentation perceptual score (CSPS) is proposed to improve the crack trace accuracy of the generated crack images. Secondly, we adopt the trained DDIM that repairs the missing crack regions. With our proposed guidance model, the denoising diffusion model can generate amazing inpainting results, including both highly coherent background and highly accurate crack trace. Experimental results show that our method can achieve better quantitative and qualitative results compared to the baseline methods. The contributions of the paper are summarized as follows:

- We propose a DDIM that is suitable for crack inpainting. We not only take the advantage of powerful generation ability of denoising diffusion models, but also guide the model to generate more accurate crack trace, which is very important for crack inpainting in structural health monitoring.
- We define CSPS which can facilitate crack inpainting and is a good guidance for diffusion models. CSPS is achieved by a high level feature extractor of crack (named as Crack Unet model).
- The proposed DDIM is guided by CSPS. With the help of CSPS, our model can generate high quality crack inpainting results in terms of subjective vision and objective metrics.

2. Related work

2.1. Image inpainting

The main inspiration of traditional image inpainting methods is to use the texture features to fill the missing regions of an image. Early works attempt to utilize background patches close to the missing regions, gradually padding from the boundary to the center of missing regions [30]. However, these methods assume the lost information is the copy of background, and thus it is hard to reproduce the lost information when encountering complex structures.

Deep learning methods assume that one image is combined by several high-level semantic information. This semantic information can be represented as some vectors in a latent space which is transformed by a neural network. To model these information, deep learning models are trained with millions of images [31]. In order to optimize the process, many search algorithms are used to tune the parameters of the neural network. The most successful search algorithm is gradient descent [32], which can optimize the network to an ideal state. Convolutional neural networks (CNNs) [33, 34] are powerful models which exhibit strong expression ability. CNN-based methods make great progresses in many tasks like image classification, object detection, semantic segmentation, image inpainting and so on.

Generative adversarial network (GAN) is one of most popular deep learning networks in recent years [35, 36]. It is proposed to solve the problem that loss function and evaluation criterion are hard to describe manually. GAN introduces a discriminator network that can work as a complex loss function and makes great successes in image generation. With the success of GAN, a series of methods based on GAN are proposed for image inpainting. Contexture autoencoder [37] first defines an encoder–decoder architecture that can represent the missing region as an embedding. Contextual attention [38] take the missing region boundary and global image information as prior information to repair the missing regions. Their generation results are very coherent and have fewer visual artifacts. Edge connect and gated convolution [39, 40] take edge information as prior knowledge and redesign the convolution kernel to generate sharper inpainting results. Pyramid-context encoder network [41] has developed the encoder–decoder framework, and learns more abundant embeddings. Style transformer [42] uses vision transformers [43] as backbone networks and large mask inpainting [44] introduces auto-regressive model for image inpainting. Although they all achieve excellent inpainting results in certain applications, such as natural images, landscapes and faces, they are not specially designed for crack images.

2.2. Diffusion models

Diffusion models recently achieve competitive performance compared to GANs. It defines a process noted as diffusion, in other words, a process that continuously injects noise into the image and renders it become noise. The model also defines the image generation process as a reverse diffusion process. In the diffusion process, the Markov chain process is generated based on Langevin dynamics and it is more stable than GAN, as it can overcome the impact of large data difference. Moreover, compared with StyleGAN2 [45], denoising diffusion probabilistic models (DDPMs) [46] improve the sample quality of the diffusion process and achieve much higher efficiency. Improved diffusion model [47] proposes a new sample method to improve efficiency, and use precision and recall to compare how well DDPMs and GANs cover the target distribution. Guided diffusion model [48] introduces a classifier to guide the generation process and can produce specific samplers. The model improves GAN in terms of both evaluation metrics and visual effects. Therefore, diffusion denoising models are applied in many other tasks such as super resolution [49] and image inpainting [50].

Furthermore, these inpainting methods are suitable for normal data such as natural scenery, and faces. They are not suitable for crack image and cannot get satisfactory performance, as they do not take crack characteristics especially crack traces into account. The crack samplers generated by denoising diffusion models are very style-similar to the original image, however a lot of crack trace information is lost. This may be due to the fact that the crack regions just account for a small proportion of the image and are easy to be ignored. To overcome this issue, we propose CSPS to better constrain and emphasize the crack part of the whole image.

3. Method

3.1. Crack segmentation perceptual score (CSPS)

In this paper, we use diffusion models as a crack generative method. As show in figure 1, at the training stage of diffusion model, the noise is injected into diffusion process and taken as labels, then the reverse diffusion process can be predicted by the deep learning models [46].

The diffusion model shows that a large total step can get high generation quality, but the inference time increases largely. DDIMs [51] proposes a non-markovian forward process that leads to the more satisfactory surrogate objective function. Simultaneously, DDIM can optimize the generation process of diffusion model.

Although DDIM can generate high quality samples, it originally does not have a configuration of class generation. Therefore, DDIM can be optimized by introducing graceful class-specific sample results and class information gradient.

Furthermore, it is necessary that the crack inpainting model takes the background regions of images as condition information. It is enough to generate style consistent results by using the background regions, as it provides much background information. However, experimental results show that DDIM produces sample cracks which are similar to the dataset only in terms of style. These samples contain less crack traces. This might be that thin cracks are relatively small parts of the whole image, and easy to be ignored. Derived from classifier guidance, we propose CSPS to guide DDIM which can take the gradient of a crack semantic segmentation model as guidance information. As shown in figure 2, we appoint a crack semantic segmentation model to guide the denoising diffusion models during generation process.

Since the introduction of CSPS is to guide the crack inpainting, we should pay more attention to crack traces in engineering practice because it contains important characteristic information of crack. Semantic segmentation aims to consider the position and distribution of objects, and it can learn feature maps for representing the traces of objects. Therefore, crack traces can be learned by a segmentation model.

In this paper, we adopt U-net model pre-trained on crack segmentation dataset (discussed in section 4.1) as a feature extractor, noted as Crack Unet. Its architecture consists of a symmetric multi-scale encoder and decoder, which is one of the most popular segmentation model in deep learning [52]. As shown in figure 3, Crack Unet works as a feature extractor and produces high-level crack information by middle layers.

Based on the above prerequisites, given input image I_i, let φ_i be the ith layer activation feature maps produced by Crack Unet, we define the CSPS by weighted sum of φ_i as

$$S_{\mathrm{csp}} = E\left[\sum\nolimits_i^n \frac{1}{N_i}\|\varphi_i(I_i)\|_1\right] \tag{1}$$

where $\|\varphi_i(I_i)\|_1$ is L1 norm of feature maps $\varphi_i(I_i)$. N_i is the normalization parameter.

For the distance measurement, we follow the perceptual loss method and calculate the L1 norm to measure the feature maps. Experimental results show that the crack inpainting results can be improved with the CSPS, as discussed in section 5.

The main reason of this property CSPS is that the crack segmentation model Crack Unet is trained to predict crack traces with high positive values feature maps. If crack traces are absent in image patches, the feature maps output by the model will be very rare. Thus, CSPS is an excellent measurement of whether crack traces exist in image patches. In next subsection, we combine CSPS and DDIM to build crack inpainting model and overcome the difficulty of crack trace insufficiency of inpainting results generated by diffusion model.

3.2. CrackDiffusion model: crack inpainting model with DDIM and CSPS

In order to simulate the missing areas of an image, we define an mask m. For the mask, we define the missing regions as 1 and background regions as 0. Then, the missing regions of the image can be represented as $m \odot x$, noted as masked region. The background regions can be represented as $(1-m)\odot x$, noted as unmasked region. Then, we give the definition of conditional generation process, shown as:

$$x_{t-1}^{\mathrm{known}} \sim N(\sqrt{\bar{\alpha}_t}x_0, (1-\bar{\alpha}_t)I) \tag{2}$$

$$x_{t-1}^{\mathrm{unknown}} \sim N(\mu_\theta(x_t,t), \sum_\theta(x_t,t)) \tag{3}$$

$$x_{t-1} = m\odot x_{t-1}^{\mathrm{known}} + (1-m)\odot x_{t-1}^{\mathrm{unknown}} \tag{4}$$

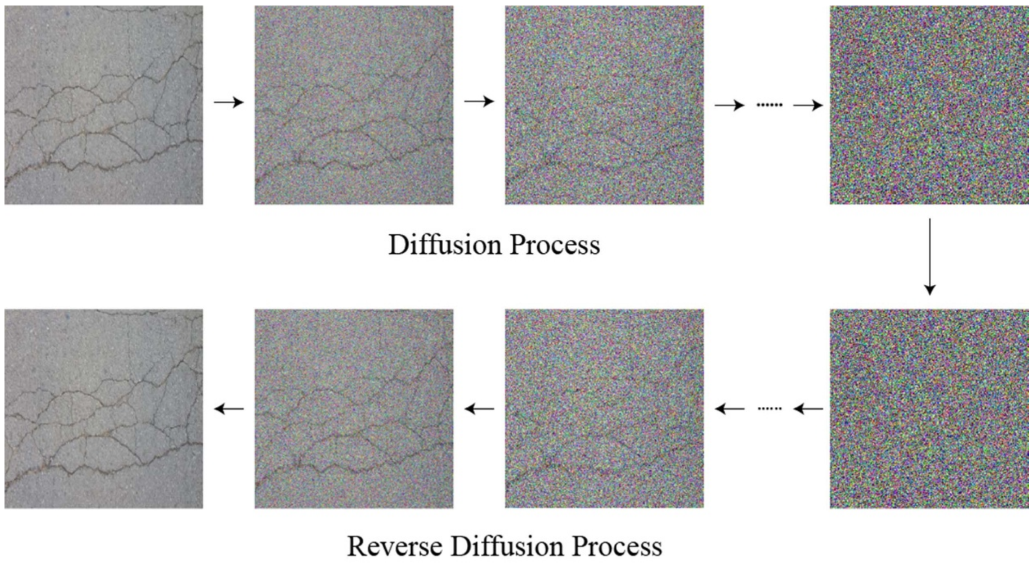

Figure 1. Illustration of denoising diffusion process. The first row we inject noise step by step to the ground truth image, until it becomes absolute noise. The second row we denoise the absolute noise image to a clear result by reverse diffusion.

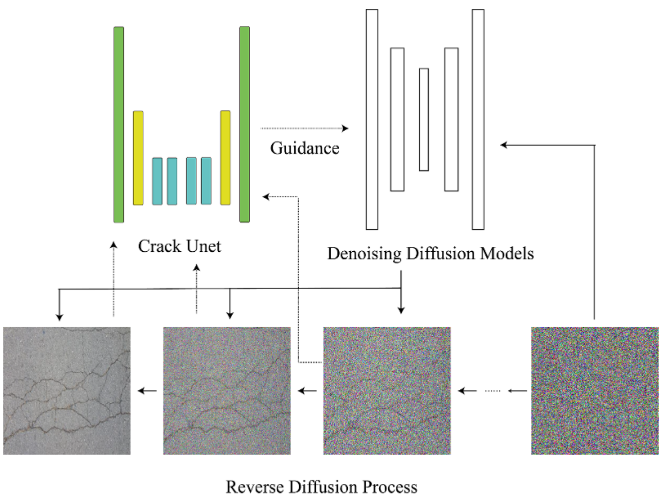

Figure 2. The guidance denoising diffusion process. To produce high quality samples, we appoint a crack semantic segmentation model to guide the denoising diffusion models during generation process.

where x_{t-1}^{known} is sampled using the known pixels in the given image $(1-m)\odot x$, x_{t-1}^{unknown} is sampled by the previous iteration x_t. Then we can acquire x_{t-1} using x_{t-1}^{known} and x_{t-1}^{unknown} by (4).

As shown in figure 4, we can first generate a set of intermediate results of forward diffusion process that only inject noise to the background regions. Then, we acquire background regions with different noise at any time step, which are amazing conditional training labels that can be injected into a reverse diffusion process. In our forward diffusion process, we not only sample the missing region x_{t-n}^{unknown}, and also acquire the corresponding background region samples x_{t-n}^{known}. Finally, they are merged as a complete sample and play a role of conditional constraint information.

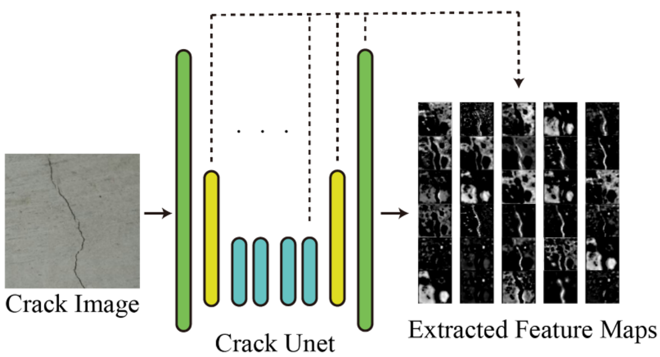

Figure 3. CrackUnet works as a high level feature extractor. Given an input of crack image, feature maps are output from the middle layers and taken as a high-level crack information.

In a word, we propose a crack inpainting model with DDIM and CSPS, named as CrackDiffusion. The proposed model not only generates consistent inpainting results, but also makes the generation results conform to the original characteristics of the crack traces. It first performs a known conditional generation process and gets filled results. Then the model diffuses these results back to the previous steps by injecting noise when the results with low CSPS. The coefficient of this reinjected noise is normalized by the CSPS. For the generation results, the larger the CSPS is, the more sufficient the crack trace information is. Therefore, we hope that the model can generate images with high CSPS, which can reflect that the inpainting results are close to the real information. On the other hand, DDIM trends to generate information that are only highly coherent to the background. CSPS allows DDIM to coordinate whether the crack information is consistent with the background on the premise of generating sufficient crack information, rather than directly generating background patches without crack information.

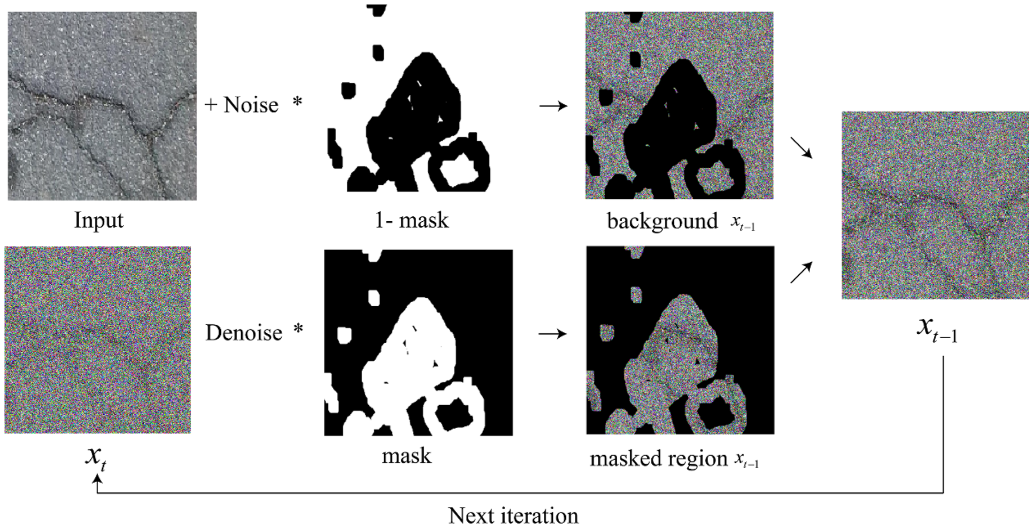

Figure 4. Overview of the denoising diffusion implicit process. The iterative process is divided into two steps: Firstly, we inject noise into the masked region, and we denoise the unmasked region of absolute noise image x_t. Then, we combine background x_{t-1} and masked region x_{t-1} as the final x_{t-1}.

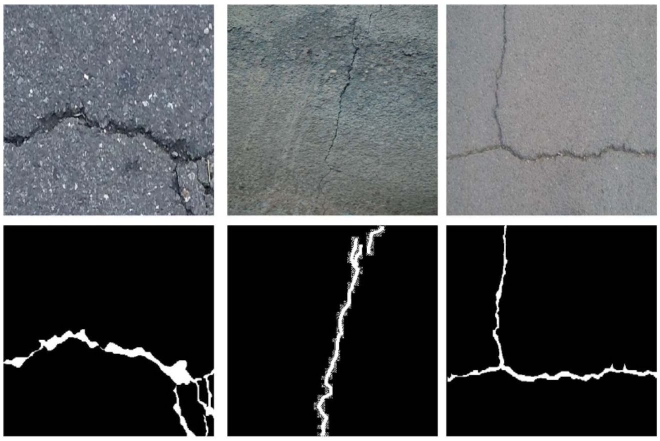

Figure 5. Example from crack segmentation dataset images and masks.

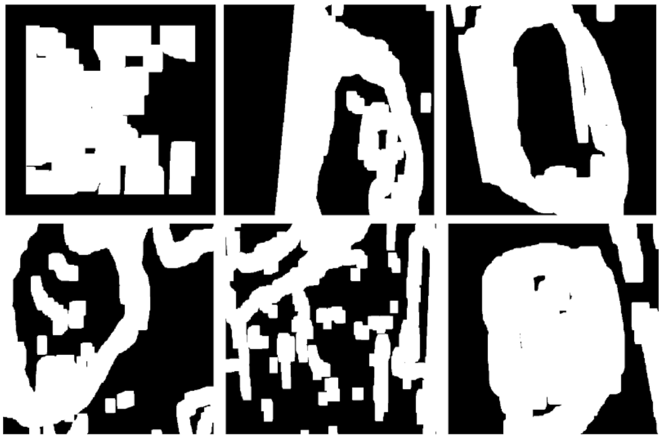

Figure 6. Example from mask dataset–nvidia imd dataset.

4. Experiments

4.1. Dataset

The dataset we used in this paper is crack segmentation dataset. This dataset (www.kaggle.com/lakshaymiddha/crack-segmentation-dataset) contains around 11 300 images. We divide the dataset into training set of 9600 images and validation set of 1700 images. Figure 5 shows some typical examples of the datasets. Crack segmentation dataset provides segmentation labels of all crack images. Our Crack Unet is trained by images and segmentation labels pairs.

In addition, in the test stage, we use different masks which are chosen from nvidia-imd dataset. Nvidia-imd dataset contains 10 000 training images and 2000 evaluation images with different masks, selected mask examples are shown in figure 6.

4.2. Model configuration and training strategy

We adopt the same architecture of DDIM in [48], as it has high generation performance and sampling efficiency. We train DDIM for 18 000 epochs with batch size 2 on a single Nvidia GeForce RTX 3090. We use the image size 256 × 256 and attention resolutions 32, 16 and 8. The diffusion step is set as 1000. For the training of Crack Unet model, we use dice loss, binary cross entropy (BCE) loss and L1 loss with batch size 4 and Adam optimizer learning rate 10^{-3} until converge. The number of training parameters of the models is 554 MB.

4.3. Training and testing details

The training curve and the testing curve are shown in figure 7. Obviously, the training loss and testing loss are nearly close

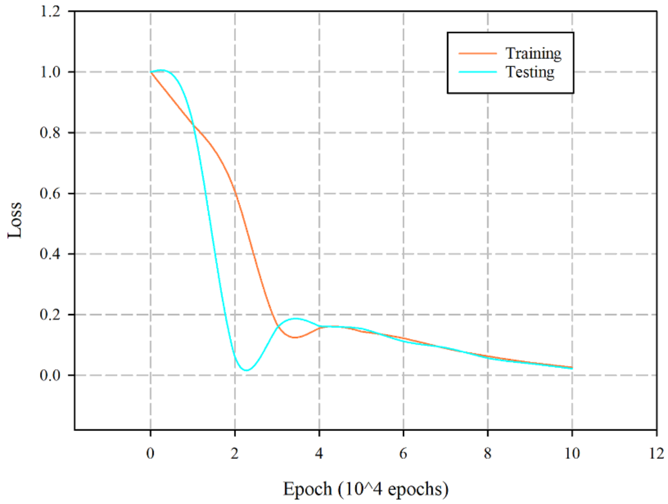

Figure 7. The training and testing curve of our model.

to 0. It implies that our model performs well in training process and can be well generalized to unknown test data.

However, convergence rate of our model is relatively slow. Actually, we train the models on single RTX 3090 for about 2 d, with 110 K iterations. For testing, due to the multiple sampling steps of DDIMs, we feed an image with size of 256×256 into the models and get inference time about 300 s (diffusion steps: 1000). The main reason is that we introduces a sampling guidance method. Although the sampling guidance method can enable diffusion models to generate high-quality crack inpainting results, it is a time-consuming model.

5. Results

5.1. Qualitative comparison

After the training, our network is evaluated on the evaluation dataset. We compare the quantitative and qualitative results with four baseline models, including patch match [53], contextual attention [38], gated convolution [40] and repaint [50]. Since these methods are not tested on the crack image dataset in the original paper, we adopt these models to train on the crack image dataset (until converge) to obtain the results. Qualitative results are shown in figure 8.

Patch match [53], contextual attention [38] and gated convolution [40] only concentrate on the background information and lose a lot of crack information. In contrast, our method obtains satisfactory results with abundant crack information. Especially at the condition where masks are relatively large and only very limited crack information is visible, our method retains crack information to a great extent. Although repaint [50] can also pay more attention to crack information and achieve some notable inpainting results, it still lacks effective guidance. In contrary, we introduce CSPS for guiding DDIMs, which forces the network to consider the information of the crack trace and reconstruct missing crack information. Therefore, by combining CSPS and DDIM, our method produces both precise and high-quality inpainting results. Compared with repaint [50] (figure 8(f)), our method (figure 8(g)) exhibits more crack traces, and the generated crack traces are very consistent with the original cracks (figure 8(a)).

5.2. Quantitative results

For the crack inpainting results, we choose several representative evaluation metrics [54, 55]: (1) peak signal-to-noise ratio (PSNR), (2) structural similarity index (SSIM), (3) mean average error (MAE). (4) Frechet inception score. The evaluation values for crack inpainting results from dataset 1 and dataset 2 are shown in table 1. Obviously, our method can achieve relatively better scores than the other methods, which demonstrates that our proposed method is very successful in crack inpainting.

5.3. Ablation study

For ablation study, we conduct controlled experiments to demonstrate the effect of CSPS. Intuitively, different CSPSs have different effects for DDIM generation process. Low CSPS means that there is no crack in the missing regions. The model only generates background pixels, which will lose ground truth crack information. In contrast, high CSPS indicates that there are plenty of cracks, which may hint model to generate patches full filled with cracks. This is also inconsistent with the actual situation. Thus, suitable CSPS can represent crack information in real environment and generate inpainting results with adequate crack information. Even if there are no cracks, the coordination ability of DDIM can be free from the wrong guidance of CSPS and generates background patches. Therefore, CSPS is the prerequisite to high quality crack inpainting, as shown in figure 9.

In addition, we set CSPS as different values: 4.0 (high value), 0.8 (suitable value) and 0.4 (low value). Results show that these techniques can improve the performance of the crack inpainting. Table 2 shows the quantitative results. We observe

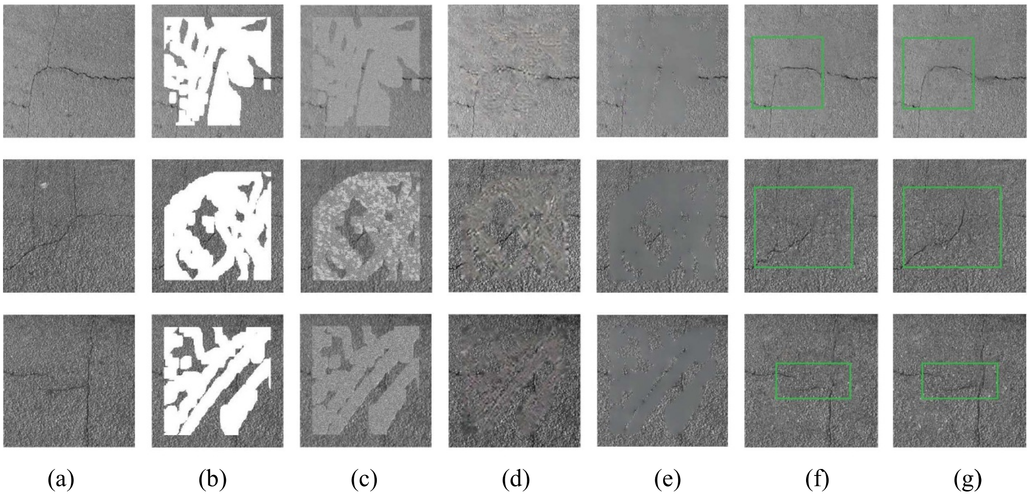

Figure 8. Inpainting results from crack segmentation dataset. (a) Ground truth, (b) ground truth with a mask, (c) patch match [53], (d) contextual attention [38], (e) gated convolution [40], (f) repaint [50], (g) our method.

Table 1. Quantitative evaluation of inpainting results.

Model/metric	PSNR[a]	SSIM[a]	MAE[b]	FID[b]
Patch match [53]	24.30	0.67	0.0304	36.54
Contextual attention [38]	26.45	0.76	0.0210	28.26
Gated convolution [40]	27.65	0.74	0.0203	24.83
Repaint [50]	29.31	0.84	0.0164	7.26
Crack diffusion (ours)	30.56	0.91	0.0152	6.24

[a] Higher is better.
[b] Lower is better.

Figure 9. Effect of CSPS. The first row, ground truth crack image and 4 samples from DDIM without CSPS guidance. The second row, ground truth crack with mask and 4 samples from DDIM with CSPS guidance. It is obvious that we can generate more crack traces information with CSPS.

that with higher CSPS, there are only moderate improvements. The reason might be that if a crack reaches the limit of CSPS, even if we give a higher score, the gradient from guidance model may not changes much. Thus, the sampling mean of the DDIM changes very slightly, resulting in no changes in the inpainting results. In our model, CSPS is computed by equation (1), which is adaptive to image characteristics and guides DDIM to achieve satisfactory inpainting results.

5.4. The effect on real image

In order to illustrate the performance of our method in real scenes, we apply our method on some real crack images, as shown in figure 10. We collect many crack images with leaf occlusion and stone occlusion. This real image dataset is captured by a fixed stand camera with the size of 256. We take photos of ground crack and let stones and leaves occlude some part of crack to imitate information loss. We directly adopt our

Table 2. Quantitative results from different CSPS, 4.0 (high value), 0.8 (suitable value) and 0.4 (low value).

Model/metric	PSNR[a]	SSIM[a]	MAE[b]	FID[b]
CSPS 4.0	30.22	0.88	0.0158	6.92
CSPS 0.8	30.56	0.91	0.0152	6.24
CSPS 0.2	29.52	0.86	0.0162	6.85
Without CSPS	29.31	0.84	0.0164	7.26

[a] Higher is better.
[b] Lower is better.

(a) (b) (c) (d) (e) (f)

Figure 10. The effect on real images. (a) Real image without occlusion, (b) leaf occlusion, (c) inpainting results of (b), (d) real image without occlusion, (e) stone occlusion, (f) inpainting results of (e).

models trained on crack segmentation dataset to these images. It is shown that our method can generate high quality inpainting results. The crack trace information can be repaired well by our method.

6. Conclusions

In this paper, we develop deep learning and powerful denoising diffusion models for crack inpainting. We propose Crack-Diffusion, a kind of DDIMs, first trained to generates coherent crack. Moreover, we propose CSPS, a powerful method to extract high-level crack semantic information by a pre-trained semi-supervised crack segmentation model. With the guidance of CSPS, the DDIMs can generate high quality crack while comparative methods tend to ignore some important crack information. Even in extreme large defect conditions, our method can still restore a lot of crack information. Qualitative results and quantitative comparisons demonstrate that our methods can achieve superior performance.

The limitation of this work is that convergence rate of our model is relatively slow. The future work is mainly to optimize our network structure to improve the convergence rate.

Data availability statement

All data that support the findings of this study are included within the article (and any supplementary files).

Acknowledgments

The authors gratefully acknowledge the financial support of this research comes from the National Natural Science Foundation of China (Grant Nos. 61901059, 51978079).

Conflict of 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.

ORCID iD

Luoyu Zhou https://orcid.org/0000-0003-4417-1250

References

[1] Giagopoulos D, Arailopoulos A, Dertimanis V, Papadimitriou C, Chatzi E and Grompanopoulos K 2019 Structural health monitoring and fatigue damage estimation using vibration measurements and finite element model updating *Struct. Health Monit.* **18** 1189–206
[2] Li N, Wang F and Song G 2020 New entropy-based vibro-acoustic modulation method for metal fatigue crack detection: an exploratory study *Measurement* **150** 107075
[3] Zhang C, Wang R and Song G 2020 Effects of pre-fatigue damage on mechanical properties of Q690 high-strength steel *Constr. Build. Mater.* **252** 118845
[4] Peng J, Hu S, Zhang J, Cai C S and Li L Y 2019 Influence of cracks on chloride diffusivity in concrete: a five-phase mesoscale model approach *Constr. Build. Mater.* **197** 587–96
[5] Yang Y, Peng J, Liu X, Cai S C and Zhang J 2021 Probability analysis of web cracking of corroded prestressed concrete box-girder bridges considering aleatory and epistemic uncertainties *Eng. Struct.* **228** 111486
[6] Kong Q, Robert R, Silva P and Mo Y 2016 Cyclic crack monitoring of a reinforced concrete column under simulated pseudo-dynamic loading using piezoceramic-based smart aggregates *Appl. Sci.* **6** 341
[7] Zhou L, Zheng Y, Ma H and Song G 2021 Attenuation characteristics of stress wave in cracked concrete beam using smart aggregate transducers enabled time-reversal technique *J. Intell. Mater. Syst. Struct.* **32** 473–85
[8] Baker A, Rajic N and Davis C 2009 Towards a practical structural health monitoring technology for patched cracks in aircraft structure *Composites* A **40** 1340–52
[9] Kong Q, Fan S, Bai X, Mo Y L and Song G 2017 A novel embeddable spherical smart aggregate for structural health monitoring: part I. Fabrication and electrical characterization *Smart Mater. Struct.* **26** 095050
[10] Kong Q, Fan S, Mo Y L and Song G 2017 A novel embeddable spherical smart aggregate for structural health monitoring: part II. Numerical and experimental verifications *Smart Mater. Struct.* **26** 095051
[11] Yan S, Ma H, Li P, Song G and Wu J 2017 Development and application of a structural health monitoring system based on wireless smart aggregates *Sensors* **17** 1641
[12] Siu S, Ji Q, Wu W, Song G and Ding Z 2014 Stress wave communication in concrete: I. Characterization of a smart aggregate based concrete channel *Smart Mater. Struct.* **23** 125030
[13] Ji Q, Ho M, Zheng R, Ding Z and Song G 2015 An exploratory study of stress wave communication in concrete structures *Smart Struct. Syst.* **15** 135–50
[14] Ma Y, Ji Q, Chen S and Song G 2017 An experimental study of ultra-low power wireless sensor-based autonomous energy harvesting system *J. Renew. Sustain Energy* **9** 054702
[15] Song H, Xiang M, Lu G and Wang T 2022 Singular spectrum analysis and fuzzy entropy-based damage detection on a thin aluminium plate by using PZTs *Smart Mater. Struct.* **31** 035015
[16] Jiang T, Hong Y, Zheng J, Wang L and Gu H 2019 Crack detection of FRP-reinforced concrete beam using embedded piezoceramic smart aggregates *Sensors* **19** 1979
[17] Chiu C K, Liao W I and Hartono A 2019 Study on the application of post-embedded piezoceramic transducers for crack detection on earthquake-damaged RC columns *Smart Mater. Struct.* **28** 055039
[18] Xu K, Deng Q, Cai L, Ho S and Song G 2018 Damage detection of a concrete column subject to blast loads using embedded piezoceramic transducers *Sensors* **18** 1377
[19] Fischer O, Thoma S and Crepaz S 2019 Distributed fiber optic sensing for crack detection in concrete structures *Civil Eng. Des.* **1** 97–105
[20] Henault J M, Quiertant M, Delepine-Lesoille S, Salin J, Moreau G, Taillade F and Benzarti K 2012 Quantitative strain measurement and crack detection in RC structures using a truly distributed fiber optic sensing system *Constr. Build. Mater.* **37** 916–23
[21] Li W, Xu C, Ho S C M, Wang B and Song G 2017 Monitoring concrete deterioration due to reinforcement corrosion by integrating acoustic emission and FBG strain measurements *Sensors* **17** 657
[22] Mohan A and Poobal S 2018 Crack detection using image processing: a critical review and analysis *Alexandria Eng. J.* **57** 787–98
[23] Xu C, Zhou N, Xie J, Gong X, Chen G and Song G 2016 Investigation on eddy current pulsed thermography to detect hidden cracks on corroded metal surface *NDT&E Int.* **84** 27–35
[24] Munawar H S, Hammad A W, Haddad A, Soares C A P and Waller S T 2021 Image-based crack detection methods: a review *Infrastructures* **6** 115
[25] Lei B, Wang N, Xu P and Song G 2018 New crack detection method for bridge inspection using UAV incorporating image processing *J. Aerosp. Eng.* **31** 04018058
[26] Ellenberg A, Kontsos A, Bartoli I and Pradhan A 2014 Masonry crack detection application of an unmanned aerial vehicle *2014 Int. Conf. on Computing in Civil and Building Engineering* 1788–95
[27] Lei B, Ren Y, Wang N, Huo L and Song G 2020 Design of a new low-cost unmanned aerial vehicle and vision-based concrete crack inspection method *Struct. Health Monit.* **19** 1871–83
[28] Park S, Ahmad S and Yun C B 2006 Multiple crack detection of concrete structures using impedance-based structural health monitoring techniques *Exp. Mech.* **46** 609–18
[29] Yang F, Zhang L, Yu S, Prokhorov D, Mei X and Ling H 2020 Feature pyramid and hierarchical boosting network for pavement crack detection *IEEE Trans. Intell. Transp. Syst.* **21** 1525–35
[30] Cozzolino D, Poggi G and Verdoliva L 2014 Copy-move forgery detection based on PatchMatch *2014 IEEE Int. Conf. on Image Processing* pp 5312–6
[31] Tian Y, Luo M and Zhou L 2021 Fusion network for blur discrimination *J. Electron. Imaging* **30** 033030
[32] Bottou L 2012 *Stochastic Gradient Descent Tricks* (Berlin: Springer) pp 421–36
[33] Wang F, Song G and Mo Y L 2021 Shear loading detection of through bolts in bridge structures using a percussion-based one-dimensional memory-augmented convolutional neural network *Comput.-Aided Civ. Infrastruct. Eng.* **36** 289–301
[34] Guo Y and Zhou L 2022 MEA-net: a lightweight SAR ship detection model for imbalanced datasets *Remote Sens.* **14** 4438
[35] Creswell A, White T, Dumoulin V, Arulkumaran K, Sengupta B and Bharath A A 2018 Generative adversarial networks: an overview *IEEE Signal Process. Mag.* **35** 53–65
[36] Yi X, Walia E and Babyn P 2019 Generative adversarial network in medical imaging: a review *Med. Image Anal.* **58** 101552
[37] Pathak D, Krahenbuhl P, Donahue J, Darrell T and Efros A A 2016 Context encoders: feature learning by inpainting *2016 IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)* p 2536–44
[38] Yu J, Lin Z, Yang J, Shen X, Liu X and Huang T 2018 Generative image inpainting with contextual attention *2018 IEEE/CVF Conf. on Computer Vision and Pattern Recognition* pp 5505–14

[39] Nazeri K, Ng E, Joseph T, Qureshi F and Ebrahimi M 2019 EdgeConnect: structure guided image inpainting using edge prediction *2019 IEEE/CVF Int. Conf. on Computer Vision Workshop (ICCVW)* (https://doi.org/10.1109/ICCVW.2019.00408)
[40] Yu J, Lin Z, Yang J, Shen X, Lu X and Huang T 2019 Free-form image inpainting with gated convolution *2019 IEEE/CVF Int. Conf. on Computer Vision (ICCV)* pp 4471–80
[41] Zeng Y, Fu J, Chao H and Guo B 2019 Learning pyramid-context encoder network for high-quality image inpainting *2019 IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)* pp 1486–94
[42] Hu X, Huang Q, Shi Z, Li S, Gao C, Sun L and Li Q 2022 Style transformer for image inversion and editing *2022 IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)* pp 11337–46
[43] Liu Z, Lin Y, Cao Y, Hu H, Wei Y, Zhang Z, Lin S and Guo B 2021 Swin transformer: hierarchical vision transformer using shifted windows *2021 IEEE/CVF Int. Conf. on Computer Vision (ICCV)* pp 10012–22
[44] Suvorov R, Logacheva E, Mashikhin A, Remizova A, Ashukha A, Silvestrov A, Kong N, Goka H, Park K and Lempitsky V 2022 Resolution-robust large mask inpainting with Fourier convolutions *2022 IEEE/CVF Winter Conf. on Applications of Computer Vision (WACV)* pp 2149–59
[45] Viazovetskyi Y, Ivashkin V and Kashin E 2020 StyleGAN2 distillation for feed-forward image manipulation. *European Conf. on Computer Vision* pp 170–86
[46] Ho J, Jain A and Abbeel P 2020 Denoising diffusion probabilistic models *Advances Neural Information Processing Systems* ***33*** pp 6840–51
[47] Nichol A Q and Dhariwal P 2021 Improved denoising diffusion probabilistic models *Int. Conf. on Machine Learning* pp 8162–71
[48] Dhariwal P and Nichol A 2021 Diffusion models beat GANs on image synthesis *Advances Neural Information Processing Systems* ***34*** pp 8780–94
[49] Saharia C, Ho J, Chan W, Salimans T, Fleet D J and Norouzi M 2022 Image super-resolution via iterative refinement *IEEE Trans. Pattern Anal. Mach. Intell.* **45** 1–14
[50] Lugmayr A, Danelljan M, Romero A, Yu F, Timofte R and Gool L V 2022 RePaint: inpainting using denoising diffusion probabilistic models *2022 IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR)* pp 11461–71
[51] Song J, Meng C and Ermon S 2020 Denoising diffusion implicit models (arXiv:2010.02502)
[52] Ronneberger O, Fischer P and Brox T 2015 U-net: convolutional networks for biomedical image segmentation *Int. Conf. on Medical Image Computing and Computer-assisted Intervention* pp 234–41
[53] Barnes C, Shechtman E, Finkelstein A and Goldman D B 2009 PatchMatch: a randomized correspondence algorithm for structural image editing *ACM Trans. Graph.* **28** 24
[54] Zhou L and Tang J 2017 Fraction-order total variation blind image restoration based on L1-norm *Appl. Math. Model.* **51** 469–76
[55] Heusel M, Ramsauer H, Unterthiner T, Nessler B and Hochreiter S 2017 GANs trained by a two time-scale update rule converge to a local Nash equilibrium *Advances in Neural Information Processing Systems* ***30*** pp 6629–40