




OPEN


A denoising method based on cyclegan with attention mechanisms for improving the hidden distress features of pavement

Lei Liu, Ligang Cao , Congde Lu, Xingtao Yang, Tongbiao Wei, Xiaocui Li, Hengxin Jiang & Lin Yang

Ground Penetrating Radar (GPR) is one of the most used devices for road structural damages detection. However, due to the different roadbed conditions and various disturbances in the nearby environment during detection, there are great difficulties in interpreting detection images, which also hinders automatic detection based on deep learning. In this work, we design a GPR image denoising method based on Cyclegan. We select the most suitable generator and add different attention mechanisms. After denoising the natural GPR road detection image, using the Yolo (You Only Look Once) to test the accuracy of the original image and the denoised image after adding different attention mechanisms. The detection accuracy is improved by 30%. The results of the detection network and the evaluation of the denoised images by GPR image interpreters indicate that the method has the following advantages: lower requirements for training data sets, a wide range of data sources, low cost, good denoising effect, and automatic detection of GPR images. It is of great help to the automatic detection of GPR images.

The pavement structure is an essential part of the road. Maintaining the stability of the structure helps improve safety and vehicle comfort and reduce fuel consumption. However, pavement structures usually have many defects¹, including poor interlayers, voids, and looseness. As an important tool for detecting subsurface targets, ground penetrating radar (GPR) has been widely used in engineering due to its convenience, personnel safety, and anti-interference^{2,3}. GPR transmits high-frequency electromagnetic waves to the ground and receives reflected waves to bring back objects with different dielectric constants information to achieve the visualization and detection of buried objects. It is a relatively fast technique that provides images of the interior of the overall structure^{4,5}. With the rapid development of deep learning, heavy and complex manual detection has been gradually replaced by automatic recognition^{6–11}. However, there is a large amount of clutter and noise in the image that is unrelated to the primary signal, affecting the judgment.

Due to the influence of clutter, the effectiveness of various detection networks is limited, and the detection accuracy is usually not high. Currently, most of the main denoising methods are based on ground penetrating radar data: (1) traditional signal transformation methods, such as denoising algorithms based on spatial filtering and transform domain, and (2) statistical feature analysis methods, for example, denoising algorithms based on subspace decomposition. The classic algorithms in spatial filtering algorithms include mean filtering and median filtering^{12,13}. After filtering, the image becomes smoother than the original image, which can indeed remove some noise to a certain extent. However, it can also cause signal boundaries to be blurred, resulting in no reduction in detection difficulty; The method of transform domain filtering can more completely separate noise and signals, such as improving continuous wavelet transform and wavelet packet transform from wavelet transform¹⁴. However, due to the fixed basis of wavelet transform, the denoising results cannot actively adapt to the style of the denoised image, resulting in uncontrollable deviations; The denoising algorithms based on subspace decomposition include Component Analysis (PCA)¹⁵ and Independent Component Analysis (ICA)¹⁶,

Key Laboratory of Earth Exploration and Information Techniques of Ministry of Education, Chengdu University of Technology, Chengdu 610059, Sichuan, People's Republic of China.  email: caoligang08@cdut.cn

which decompose the data into clutter subspace and target subspace. The division of the two spaces is limited by human experience and there are no fixed standards, so errors are prone to occur.

We use an image-based method, Generative Adversarial Networks (GAN)¹⁷, to solve problem in denoising. The essence of using CycleGAN to denoise is to transfer the FDTD image style (FDTD image style refer to images without noise) to the measured image to obtain the FDTD style measured image, or the GPR image without noise. Due to the network structure, we do not need paired images and supervised learning. So, much workforce is not required for labeling, and the training process is more straightforward. CycleGAN removes the need for paired training samples by introducing a cycle consistency loss in unsupervised methods^{18,19} that force two mappings to be consistent with each other. Typically, image-to-image translation methods require detecting regions of interest in an input image and learning how to translate the detected regions to the target domain. In unsupervised training, which does not require paired data, it is necessary to learn how to select target regions in an image²⁰. In denoising applications, locating the region of interest is more important. Otherwise, the denoising will be incomplete, or the effective signal will be removed as noise. Adding an attention mechanism can solve this problem to a certain extent^{21–23}. Attention mechanism can effectively improve the effect of noise reduction. Chen et al.³⁴ used extra supervision to train attention networks to improve the quality of overall image conversion. Jie Feng et al.²² proposed a symmetric convolutional GAN based on collaborative learning and attention mechanisms to generate high-quality samples for HSI with complex spatial spectral distributions. Xuran Pan et al.²³ proposed a Generative Addition network (GAN-SCA) with spatial and channel attention mechanism for robust segmentation of buildings in remote sensing images and the results are superior to several state-of-the-art methods. The above work reflects from different aspects that the effect of deep learning networks is improved after the addition of attention mechanisms.

The detection accuracy of this method is relatively high. Our method has the following characteristics: (1) In the training process of this paper, the measured data sets in production are directly used for training. (2) It is considered that the combination of adding channel attention mechanisms and spatial attention mechanisms at the same time is the most suitable for GPR road structural damages image denoising. (3) We chose the Resnet generator that is more suitable for denoising of road structural damages by GPR.

Method

The advantage of the proposed network in this paper is that it directly uses the GPR road measured data. It does not need to pre-produce many paired forward simulation data or make a physical model obtain data for training. We add an attention mechanism to the network according to the data features to enhance the denoising effect, improve the imaging quality, and finally improve the detection accuracy for GPR road structural damages. This is a new data processing scheme in GPR road structural damages detection.

Train generator. The part we need for training is the denoising generator (purple box in Fig. 1), which is used to convert a noisy road GPR image to a noise-free GPR image. The entire training network includes one

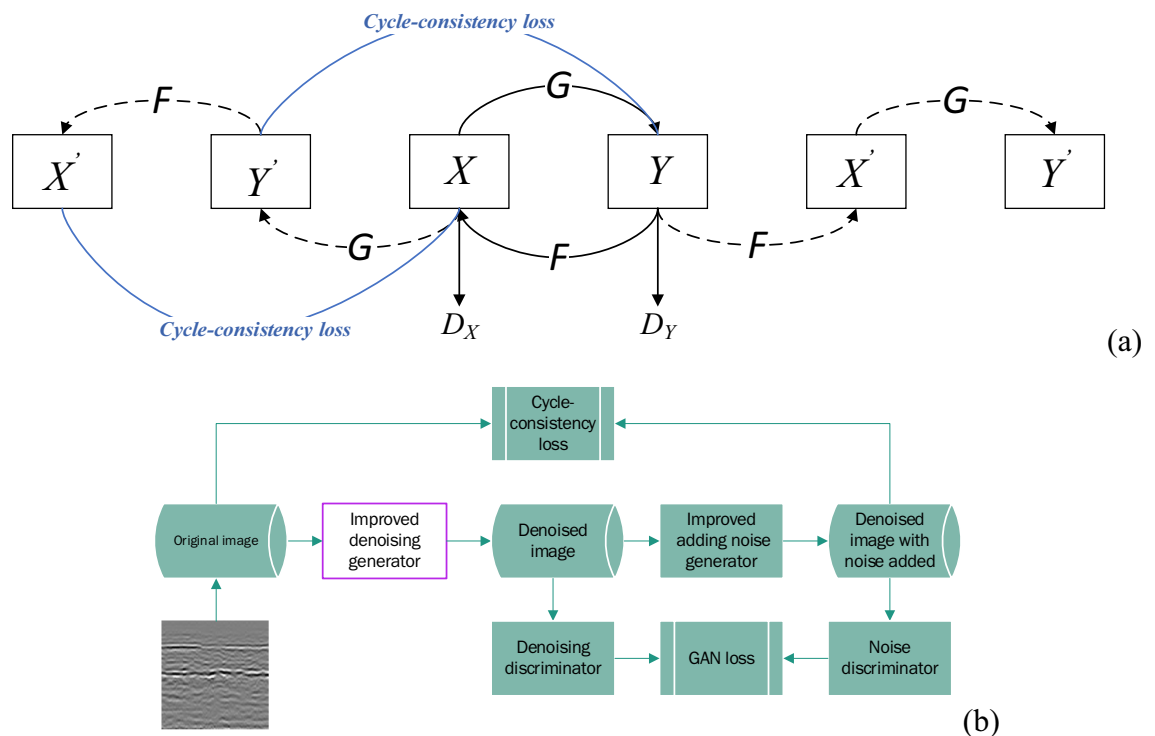


Figure 1. The process of training the denoising generator.

denoising generator G_D , and this generator contains a mapping from a noisy image X to a noise-free (FDTD) image Y ; one adding noise generator G_A , and this generator contains a mapping from a noise-free (FDTD) image Y to a noisy image X ; two discriminators D_X and D_Y , used to identify whether the image contains noise, the goals of the two discriminators are different, and the resulting $Loss_{GAN}$ ensures that the generator and the discriminator evolve, thereby ensuring that the generator can produce more realistic pictures. The sum of the loss functions produced by the two discriminators is as follows:

$$\begin{aligned}
 Loss_{GAN} &= L_{GAN}(G_D, D_Y, X, Y) + L_{GAN}(G_A, D_X, X, Y) \\
 &= \mathbb{E}_{y \sim p_{data}}(y) [\log D_Y(y)] + \mathbb{E}_{x \sim p_{data}}(x) [\log(1 - D_Y(G_D(x)))] \\
 &\quad + \mathbb{E}_{x \sim p_{data}}(x) [\log D_X(x)] + \mathbb{E}_{y \sim p_{data}}(y) [\log(1 - D_X(G_A(y)))]
 \end{aligned}
 \tag{1}$$

$y \sim p_{data}$ and $x \sim p_{data}$ is data distribution. G_D and G_A are the denoising generator and adding noise generator. D_X and D_Y are two discriminators. X is a noisy image. Y is a noise-free (FDTD) image. p_{data} is raw data distribution.

The Fig. 1a shows the network structure, with solid lines representing the path for this training session and dashed lines representing the path for the next training session. Figure 1b shows the overall process. Two generators are trained through the confrontation of two losses, but if we only train with these two losses, the network will not maintain the original image shape and will generate some images similar to the target style; that is, some images will be regenerated. Instead of removing noise based on the original image. Therefore, the two different generators should agree but go in precisely opposite directions. A loss needs to be added to ensure that the image before denoising and the image with noise added after denoising is consistent in the same cycle, so a cycle-consistency loss is added between the two generators to ensure that the output image of the generator and the input image styles different, but the same content. The network can use this to obtain the mapping relationship between the two types of images and establish a mapping between the two types of images. The cycle loss function is as follows:

$$Loss_{cycle} = \mathbb{E}_{x \sim p_{data}}(x) [|F(G(x)) - x|_{L1}] + \mathbb{E}_{y \sim p_{data}}(y) [|G(F(y)) - y|_{L1}]
 \tag{2}$$

Mapping function $F: Y \rightarrow X$. Mapping function $G: X \rightarrow Y$. The L1 distance loss is used to close the reconstructed image to the original input in the L1 sense. The Loss of CycleGAN consists of the above two parts:

$$Loss = Loss_{GAN} + Loss_{cycle}
 \tag{3}$$

Generator structure. The Fig. 2 shows the structure of upsampling and downsampling. First, the output is mirror-padded, convolved, and normalized. Mirror padding can get better convolution results than padding with a fixed value, and then downsampling, downsampling obtains deeper features by continuously compressing shallow features. After the residual block, the compressed features are upscaled to the size of the original image by upsampling.

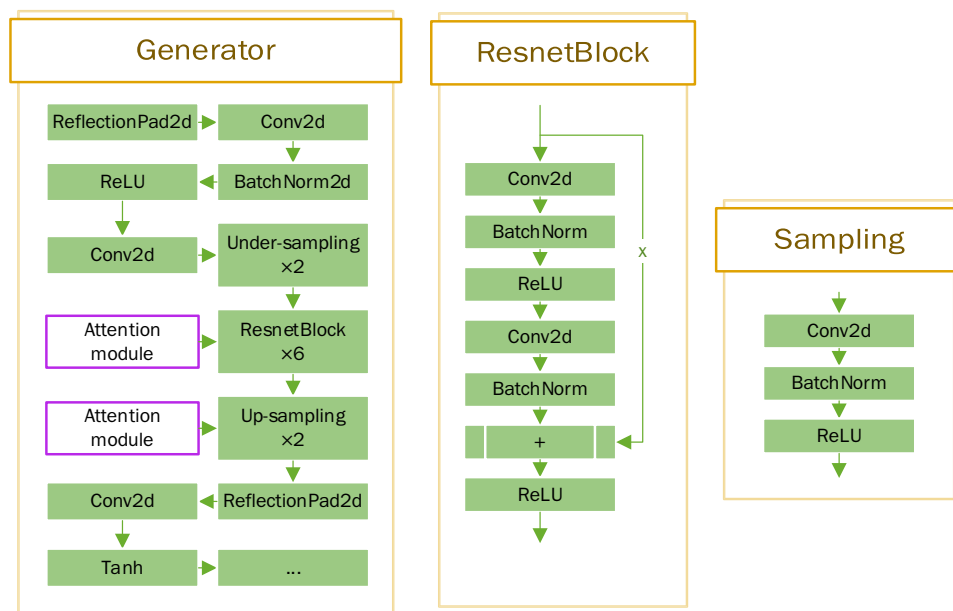


Figure 2. Generator structure (× means there are several identical blocks).

Discriminator structure. Compared with the discriminator of the original GAN, PatchGAN (The structure is shown in Fig. 3) can take into account the influence of different parts and can better represent the situation of the whole picture. Moreover, because its receptive field corresponds to a small area in the input, this training enables the model to pay more attention to image details, so PatchGAN is often used in scenarios such as high resolution.

Attention mechanism. Since the GPR image is a single-channel grayscale image, lacking color information compared with the conventional natural image, the information that the image can contain is reduced, and the difference between the target signal and the noise is not apparent. Therefore, the attention module is executed before and after the target signal connection operation, which effectively increases the weight of information learned by the network and improves the denoising effect. This paper uses two attention modules: Channel Attention and Spatial Attention^{24–26}. At a high level, the attention module calculates the importance or relevance of each input element by comparing it with other elements in the input sequence. It assigns a weight or attention score to each element, indicating its importance in the context of the current task. These attention scores are then used to compute a weighted sum of the input elements, which serves as the output of the attention module. Channel attention is to weight the channel, and spatial attention is to weight the spatial. Both channel attention and spatial attention can be used independently or in combination. They aim to enhance the understanding and expressive capacity of neural networks, enabling them to adapt better to different tasks and contexts. Four combinations of experiments were conducted in this paper. The attention mechanism is structured as Fig. 4.

Experiment

In the GPR data, the structural damages we need to detect are poor interlayers, voids, looseness, etc., all collectively referred to as structural damages in this paper. We need to remove the noise and clutter that are useless to us except for the structural damages detected in the image. The experimental process is shown in Fig. 5.

The ground-penetrating radar images and FDTD images of the highway are trained by the denoising network proposed in this paper to obtain the generator G, and the data containing the structural damages that needs to be denoised is input into G to obtain the denoised data N, and the denoised data N is obtained. The data N is manually marked with the structural damages and input into the evaluation network test to obtain accuracy and rating radar road structural damages detection.

Experimental data. The data used were measured ground penetrating radar images of a certain highway in China and FDTD images produced using GPRMAX²⁷ software (arbitrary and reusable). The measured images were collected using a ground penetrating radar instrument of the MALA GX750 model on board, with a speed of 55 km/h and a center frequency of 750 MHz. The coupling method was air coupling. The processed road measured data images and FDTD simulated images are both cropped to the same size and then input. A total of 100 cropped road measured data images form a noise-containing training set, and a total of 100 FDTD simulated images form a noise-free (FDTD) training set. The original data is directly output by the instrument (The resolution is 1110 × 186). The cropping data is square cropped according to the larger value of the disease length and width in the original image. FDTD images are directly output by GPRMAX software (The resolution is 640 × 640). The sample image is shown in Fig. 6.

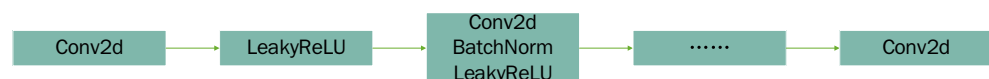


Figure 3. Discriminator structure.

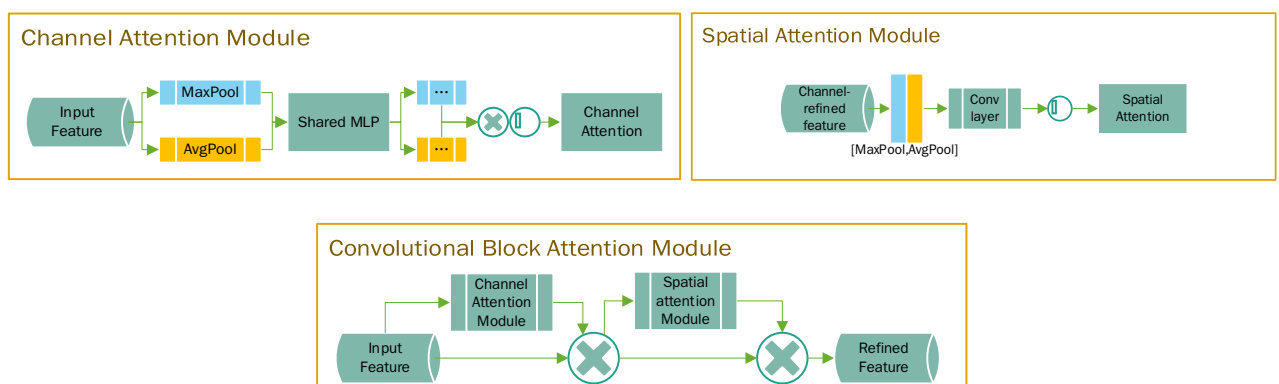


Figure 4. Attention mechanism structure.

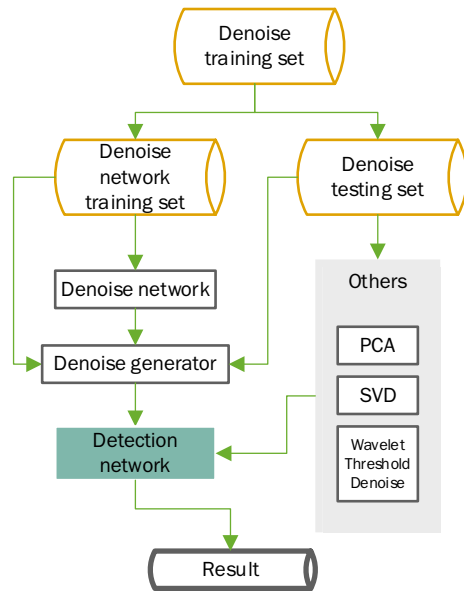


Figure 5. Experimental flow chart.

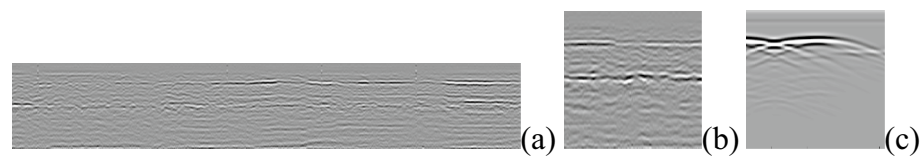


Figure 6. Samples (a) Origin data (b) Cropped data sample (c) No noise (FDTD) sample.

Detection network. The detection network uses Yolov3, which is a fully convolutional network (FCN). Its authors propose a new feature extraction network, Darknet-53. The Darknet-53 network is shown on the following Fig. 7:

Results and discussion

Other denoising methods. Principal components analysis (PCA) can handle the sparse noise problem well, but it is an unsupervised method and cannot use label information to increase the recognition rate. It is difficult to determine the number of principal components that need to be maintained in PCA. This becomes difficult when the transformation of the eigenvalues is gentle. This problem exists for the measured GPR data²⁸. We indirectly select the retained image’s eigenvalues by degree of changing the information. The results are shown in Fig. 8.

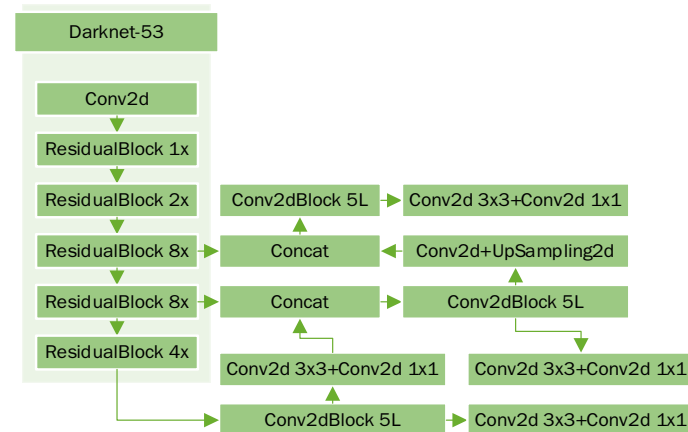


Figure 7. Darknet-53 Structure.

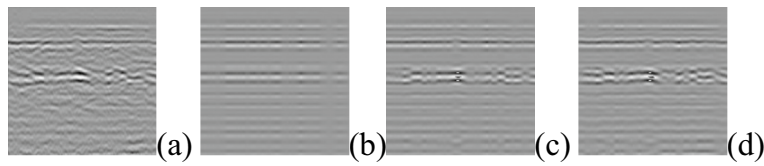


Figure 8. PCA Denoise result (a) Original image (b) DCI=30% (c) DCI=40% (d) DCI=50% (DCI=Degree of changing the information).

Singular value decomposition (SVD) is to decompose the matrix into singular vectors and singular values. Each base contains different discriminative information and reconstruction information, which can be used to extract principal components and remove noise by extracting different components^{29,30}. k is the number of singular values. We have experimented with the singular value taken from 1 to 20. The singular value in the front contains more energy, and the latter contains very little energy. The following shows the three values that the denoising effect and the primary structural damages signal remain relatively balanced. The results are shown in Fig. 9.

The basic idea of wavelet threshold denoising proposed by Donoho³¹ is to select an appropriate threshold to remove the noise with lower wavelet coefficients and retain the relatively large wavelet coefficients signal. Wavelet denoising is a synthesis of feature extraction and low pass filtering. A noisy model can be represented as follows:

$$S(k) = f(k) + \varepsilon \times e(k), \quad k = 0, 1, \dots, n - 1 \quad (4)$$

In the above formula, $f(k)$ is the useful signal, $S(k)$ is the noise-containing signal, $e(k)$ is the noise, and ε is the standard deviation of the noise figure. We choose the threshold with better effect according to the actual situation^{32,33}. The results are shown in Fig. 10.

Unet generator. From the actual effect in Fig. 11, the Unet generator in the first row of pictures does not completely save the signal in the lower part of the image, and the noise suppression in the middle of the image is not as good as Resnet; The second line signal strength is not as good as the Resnet generator. Based on the above experimental results, the main signals of the Unet generator are not completely preserved, which may easily cause the primary signals to be mistakenly removed when removing noise, which reduces the accuracy of structural damages detection. Therefore, it is recommended to use the Resnet generator for road structural damage denoising.

PSNR (Peak Signal-to-Noise Ratio). PSNR is a commonly used indicator for evaluating image quality, which can effectively explain image distortion and is also suitable for evaluating the performance of denoising algorithms. PSNR can evaluate the effectiveness of denoising algorithms by comparing the degree of distortion between the denoised image and the original image. The unit of PSNR is dB. The larger the PSNR value, the less

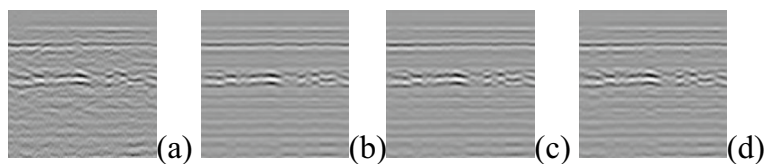


Figure 9. SVD denoise result (a) Original image (b) $k=3$ (c) $k=4$ (d) $k=5$ (k is the number of singular values).

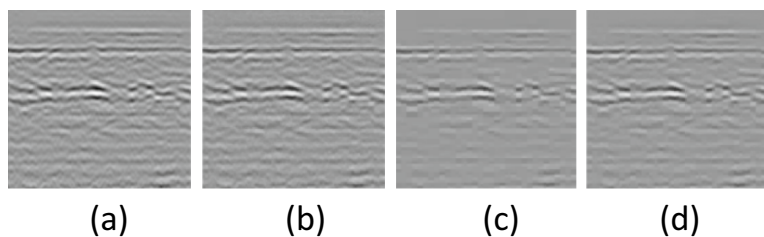


Figure 10. Wavelet threshold denoise result (a) Original image (b) Threshold is Bayes Shrink (c) Threshold is VisuShrink ($\varepsilon = \varepsilon_{est}$) (d) Threshold is VisuShrink ($\varepsilon = \varepsilon_{est}/2$).

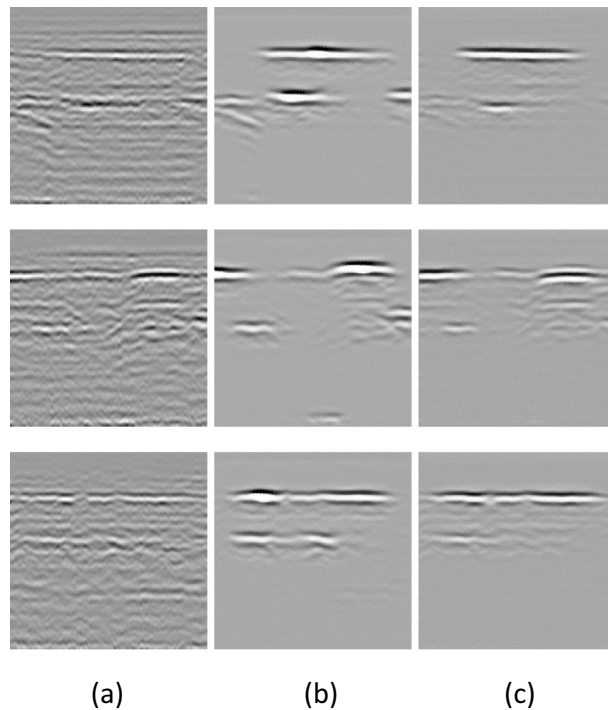


Figure 11. Results of different generators (a) Original image (b) Results of resnet generator (c) Results of unet generator.

the distortion. MSE has its limitations in describing how similar two images are. Its mathematical formula is as follows:

$$MSE = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W (X(i, j) - Y(i, j))^2 \quad (5)$$

$$PSNR = 10 \times \log_{10} \left(\frac{(2^n - 1)^2}{MSE} \right) \quad (6)$$

MSE is the mean square error calculated between the original image and the processed image. X and Y are the target images. H and W are the length and width of X and Y. $2^n - 1$ is the maximum numerical value representing the color of image points. If each sampling point is represented by 8 bits, n is 8.

After calculating the PSNR (Table 1), it can be known that adding the attention mechanism can improve the ratio between the maximum signal and the background noise. That is, the background noise can be removed more cleanly, and the signal we need is more prominent. It can be seen from the above table that the PSNR value is the largest when both attention mechanisms are added. This method should be selected if the pursuit of cleaner noise removal.

Method	PSNR/dB
No attention denoise	19.7035
Channel attention denoise	20.6644
Spatial attention denoise	20.0062
Channel attention and spatial attention denoise	21.0392
PCA	21.1426
SVD	21.3158
Wavelet Threshold	20.3212

Table 1. PSNR of original images and different processed results.

SSIM. Structural Similarity (SSIM) is a measure of the similarity between two images. Among the two images used by SSIM, one is the uncompressed undistorted image x , and the other is the distorted image y . SSIM can be obtained as follows:

$$SSIM(x, y) = \frac{(2\mu_x\mu_y + c_1)(2\sigma_{xy} + c_2)}{(\mu_x^2 + \mu_y^2 + c_1)(\sigma_x^2 + \sigma_y^2 + c_2)}, \quad c_1 = (k_1L)^2, \quad c_2 = (k_2L)^2 \quad (7)$$

μ_x is the mean of x , μ_y is the mean of y , σ_x^2 is the variance of x , σ_y^2 is the variance of y , and σ_{xy} is the covariance of x and y . c_1 and c_2 are constants used to maintain stability. L is the dynamic range of pixel values. $k_1 = 0.01$, $k_2 = 0.03$. When the two images are identical, the value of SSIM is equal to 1. Humans are not sensitive to the absolute brightness or color of pixels, but very sensitive to the position of edges and textures. SSIM mimics human perception by focusing primarily on edge and texture similarity. The results are shown in Table 2.

It can be seen from the table that the SSIM is closest to the original image without the addition of the attention mechanism, and the SSIM is reduced to varying degrees after the attention mechanism is added. Among the several types of added attention mechanisms, the closest to the original image is the one that only added the channel attention mechanism. The reason is that there are many horizontal linear features. Adding a linear attention mechanism is beneficial to maintaining the structural integrity of the graph. The addition of the spatial attention mechanism causes the horizontal linear features to be blurred, resulting in a decrease in the SSIM of the latter two methods. From the denoising point of view, since we remove a lot of structures related to irrelevant signals in the picture, the lower the SSIM, the more signals we remove, and the whole denoising effect is more thorough. We hope that the denoising effect is clean, so choose the method in which both attention mechanisms are added.

Image analysis. The Fig. 12a is the original image. As can be seen from Fig. 12b, the denoising effect is already apparent when the attention mechanism is not added, and most of the background noise in the figure is removed, but there is still small clutter outside the target signal. In the second row, because the background clutter is too chaotic, and some noises appear obvious abrupt, it is already somewhat out of tune with the surrounding pictures. It can be seen from Fig. 12c that after adding the Channel Attention mechanism, the stray clutter is significantly reduced, and there is no other stray clutter in the picture except near the main signal. Compared

Method	SSIM
No attention denoise	0.55927
Channel attention denoise	0.51473
Spatial attention denoise	0.5009
Channel attention and spatial attention denoise	0.50083
PCA	0.6553
SVD	0.69667
Wavelet threshold	0.5997

Table 2. SSIM of original images and different processed results.

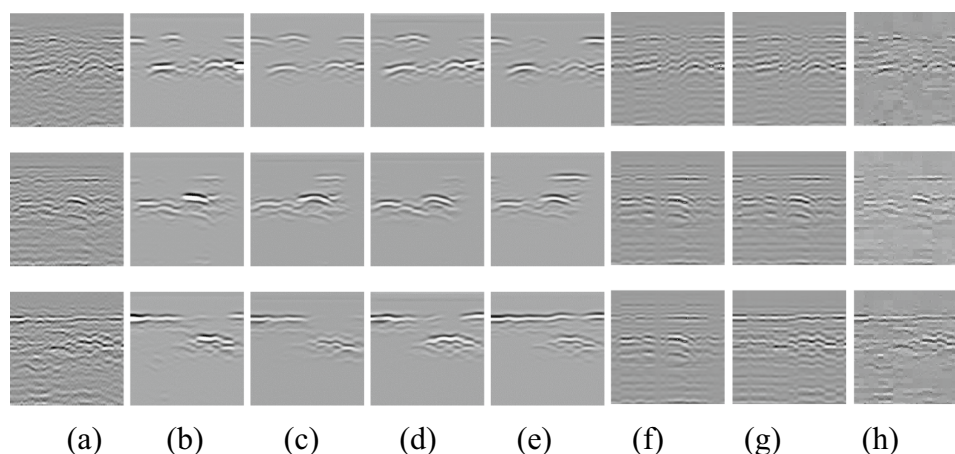


Figure 12. Denoising results of different attention mechanisms (a) Original image (b) Denoising results without attention mechanism (c) Denoising results using channel attention (d) Denoising results using spatial attention (e) Denoising results using channel attention and spatial attention (f) Denoising results of PCA (g) Denoising results of SVD (h) Denoising results of wavelet threshold.

to the denoise without attention mechanisms, it can also be seen that the clutter near the main waveform is also reduced and shrunk to the vicinity of the main waveform. The waveform on the upper right of the second row of graphs is also more apparent, the waveform is not too strong like (b), and the original shape and intensity of the waveform are better maintained. It can be seen from Fig. 12d that the image after adding the Spatial Attention mechanism is relatively compared to Fig. 12b without any attention mechanism added, except that there is basically no other scattered clutter except near the main signal. Compared with adding the Channel Attention mechanism, the waveform is complete and precise, but there are some signals whose strength is not as good as adding the Channel Attention mechanism. The denoising results (Fig. 12e) using Channel Attention and Spatial Attention are cleaner compared to Fig. 12c,d. As can be seen from the first row of images, the signal integrity is better than in Fig. 12c,d. The overall signal strength of the picture is also higher than Fig. 12c, and the surrounding clutter is further shrunk. It can be seen from the second and third row of the picture that the signal in the picture is completely preserved, while the background is cleaner, and the clutter shrinks more thoroughly. The experimental results show that the results of adding Channel Attention and Spatial Attention are the best. Visually speaking, the three columns Fig. 12f,g,h are definitely not as effective as the previous columns.

Use the annotation of the original image for detection. After manually labeling the structural damages with labeling software, a corresponding JSON file is generated for each image for detection. There are 120 images, including 100 training sets and 20 testing sets, and each image has at least one detection target.

First, we use the annotation of the original image to detect the denoised image. Since the boundary after denoising is more obvious, the indistinguishable boundary information before denoising is highlighted, so the size of the denoised image is different from that of the original image, and the accuracy of using annotation of the original image will be much lower than re-labeling. However, using the original image annotation to detect the accuracy still has a role. It can reflect the structural similarity between the image and the original image to a certain extent. The accuracy results are shown in Table 3.

It can be seen from the table that using the annotation of the original image, the accuracy of the image denoised without using the attention mechanism is the highest among all denoising results, indicating that this type of image is the most similar to the original image, followed by the channel attention mechanism. The accuracies are not much different between adding the spatial attention mechanism and adding both. Regarding the accuracy distribution, considering that the image contains a large number of horizontal correlations, the vertical correlation is much less than that of the horizontal, which is probably why the similarity of the spatial attention mechanism is not as high as that of the channel attention mechanism.

But in essence, the target boundary after denoising has changed. For the denoised image, the label of the original image actually has a large error, so AP with proposed method is lower than original image.

Use the relabeling annotation for detection. According to the detection results of the original image label, PSNR, and SSIM, we re-labeled the data in the case of adding two attention mechanisms and re-detected using the detection method described above. The re-labeled labels and detection results are as follows (Annotations are shown in Fig. 13, the accuracy results are shown in Table 4):

Method	AP
Original image	0.579
No attention denoise	0.577
Channel attention denoise	0.642
Spatial attention denoise	0.657
Channel attention and spatial attention denoise	0.617
PCA	0.392
SVD	0.495
Wavelet threshold	0.509

Table 3. The accuracy of using the original image annotation to detect.

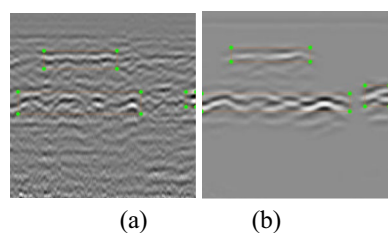


Figure 13. Annotation (a) Original image annotation (b) Re-labeling annotation.

Method	Original image	No attention denoise	Channel attention denoise	Spatial attention denoise	Channel attention and spatial attention denoise
AP	0.617	0.891	0.906	0.758	0.914

Table 4. The accuracy of using the re-labeling annotation to detect.

The actual effect of the detection network before and after denoising is as follows (Fig. 14):

According to the above chart, the image detection accuracy after denoising is significantly improved compared to the original image detection accuracy. The increase is as high as nearly 30% (From 61.7 to 91.4%). From this, it can be seen that the denoising method proposed in this paper has a very obvious effect on improving the accuracy of machine learning detection of structural damages and has considerable practical value.

The limitation of the work. Network structure can be more streamlined. Because the image itself has only single-channel grayscale data, which is smaller than the amount of natural image data (Colorful images with complex line contours), the detection network can choose a network with a lower number of layers and a simple structure that can speed up the training speed. It will not lead to the lack of deep-level information in training, resulting in a decrease in accuracy.

We can also try whether there are more suitable network modules for ground penetrating radar images, including but not limited to attention mechanisms.

Conclusion

We have successfully improved the denoising effect of the network used in this article by changing the network structure, adding attention mechanisms, and changing generators. By comparing with traditional methods, the usability of deep learning methods was demonstrated, providing an optional solution for future ground penetrating radar image denoising.

Our method has those advantages: the data sources are more expansive, the acquisition cost is lower, and the processing is closer to reality. The data is conducive to improving the effect, and the experimental results show that selecting the accurate data set in the actual production process is conducive to improving the accuracy, reducing the simulation data production process, and improving work efficiency. Although the related work of other researchers has added attention mechanisms in CycleGAN, they have not discussed the impact of various attention mechanisms. This paper adds channel attention mechanisms and Spatial attention mechanisms in three different combinations, discusses the influence of different combinations on the results. Comparing the results of different generators, according to visual interpretation, the images generated by the Resnet generator are significantly better than those generated by the Unet generator. The detection results confirmed this, and the method improved AP by nearly 30%. At the same time, the trained model runs fast, and in the future, we will apply it to mobile devices to help on-site construction personnel make quick judgments. The following research direction is to adapt this set of methods to other GPR images.

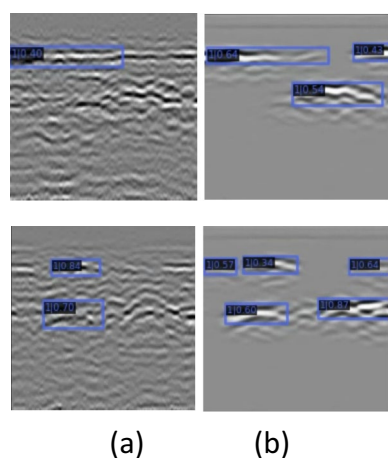


Figure 14. Detection results (a) Before denoising (b) After denoising.

Data availability

The datasets generated and/or analysed during the current study are not publicly available due owned by other companies but are available from the corresponding author on reasonable request.

Received: 29 March 2023; Accepted: 23 August 2023

Published online: 25 August 2023

References

- Guo, S., Xu, Z., Li, X. & Zhu, P. Detection and characterization of cracks in highway pavement with the amplitude variation of GPR Diffracted waves: Insights from forward modeling and field data. *Rem. Sens.* **14**(4), 976 (2022).
- Lai, W. W. L., Dérobert, X. & Annan, P. A review of ground penetrating radar application in civil engineering: A 30-year journey from locating and testing to imaging and diagnosis. *NDT E Int.* **96**, 58–78 (2018).
- Zhao, S. & Al-Qadi, I. L. Development of regularization methods on simulated ground-penetrating radar signals to predict thin asphalt overlay thickness. *Signal Process.* **132**, 261–271 (2017).
- Saarenketo, T. & Scullion, T. Road evaluation with ground penetrating radar. *J. Appl. Geophys.* **43**(2–4), 119–138 (2000).
- Krysiński, L. & Sudyka, J. GPR abilities in investigation of the pavement transversal cracks. *J. Appl. Geophys.* **97**, 27–36 (2013).
- Pham, M. T., & Lefèvre, S. Buried object detection from B-scan ground penetrating radar data using Faster-RCNN. In *IGARSS 2018–2018 IEEE International Geoscience and Remote Sensing Symposium* (pp. 6804–6807). IEEE. (2018).
- Redmon, J., & Farhadi, A. *Yolov: An Incremental Improvement*. arXiv preprint [arXiv:1804.02767](https://arxiv.org/abs/1804.02767). (2018).
- Lin, T. Y., Goyal, P., Girshick, R., He, K., & Dollár, P. Focal loss for dense object detection. In *Proc. of the IEEE International Conf. on Computer Vision* (pp. 2980–2988). (2017).
- Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., Fu, C. Y., & Berg, A. C. Ssd: Single shot multibox detector. In *European Conf. on Computer Vision* (pp. 21–37). Springer. (2016).
- Shen, Z., Liu, Z., Li, J., Jiang, Y. G., Chen, Y., & Xue, X. Dsod: Learning deeply supervised object detectors from scratch. In *Proc. of the IEEE International Conf on Computer Vision* (pp. 1919–1927). (2017).
- Duan, K., Bai, S., Xie, L., Qi, H., Huang, Q., & Tian, Q. Centernet: Keypoint triplets for object detection. In *Proc. of the IEEE/CVF International Conf. on Computer Vision* (pp. 6569–6578) (2019).
- Jeng, Y., Li, Y. W., Chen, C. S. & Chien, H. Y. Adaptive filtering of random noise in near-surface seismic and ground-penetrating radar data. *J. Appl. Geophys.* **68**(1), 36–46 (2009).
- Oskooi, B., Parnow, S., Smirnov, M., Varfinezhad, R. & Yari, M. Attenuation of random noise in GPR data by image processing. *Arab. J. Geosci.* **11**, 1–10 (2018).
- Jianbin, Wu., Hui-Lin, Z. & Mao, T. Based on wavelet packet de-noise method implemented in landmine recognition. *Comput. Eng. Appl.* **18**, 147–149 (2008).
- Chen, G., Fu, L., Chen, K., Boateng, C. D. & Ge, S. Adaptive ground clutter reduction in ground-penetrating radar data based on principal component analysis. *IEEE Trans. Geosci. Rem. Sens.* **57**(6), 3271–3282 (2019).
- Abujarad, F. & Omar, A. Comparison of independent component analysis (ICA) algorithms for GPR detection of non-metallic landmines. *Proc. SPIE* **6365**, 636516 (2006).
- Goodfellow, I. et al. Generative adversarial networks. *Commun. ACM* **63**(11), 139–144 (2014).
- Zhu, J. Y., Park, T., Isola, P., & Efros, A. A. Unpaired image-to-image translation using cycle-consistent adversarial networks. In *Proc. of the IEEE International Conf. on Computer Vision* (pp. 2223–2232). (2017).
- Yi, Z., Zhang, H., Tan, P., & Gong, M. Dualgan: Unsupervised dual learning for image-to-image translation. In *Proc. of the IEEE International Conf. on Computer Vision* (pp. 2849–2857). (2017).
- Emami, H., Aliabadi, M. M., Dong, M. & Chinnam, R. B. Spa-gan: Spatial attention gan for image-to-image translation. *IEEE Trans. Multim.* **23**, 391–401 (2020).
- Chen, X., Xu, C., Yang, X., & Tao, D. Attention-gan for object transfiguration in wild images. In *Proc. of the European Conf. on Computer Vision (ECCV)* pp. 164–180 (2018).
- Feng, J. et al. Generative adversarial networks based on collaborative learning and attention mechanism for hyperspectral image classification. *Rem. Sens.* **12**(7), 1149 (2020).
- Pan, X. et al. Building extraction from high-resolution aerial imagery using a generative adversarial network with spatial and channel attention mechanisms. *Rem. Sens.* **11**(8), 917 (2019).
- Park, J., Woo, S., Lee, J. Y., & Kweon, I. S. Bam: Bottleneck attention module. (2018) arXiv preprint [arXiv:1807.06514](https://arxiv.org/abs/1807.06514).
- Liu, Y., Shao, Z., & Hoffmann, N. *Global Attention Mechanism: Retain Information to Enhance Channel-Spatial Interactions*. (2021) arXiv preprint [arXiv:2112.05561](https://arxiv.org/abs/2112.05561).
- Woo, S., Park, J., Lee, J. Y., & Kweon, I. S. Chbam: Convolutional block attention module. In *Proc of the European Conf on Computer Vision (ECCV)* (pp. 3–19). (2018).
- <http://www.gprmax.com/>.
- Chen, G., Fu, L., Chen, K., Boateng, C. D. & Ge, S. Adaptive ground clutter reduction in ground-penetrating radar data based on principal component analysis. *IEEE Trans. Geosci. Rem. Sens.* **57**(6), 3271–3282 (2019).
- Liu, C., Song, C. & Lu, Q. Random noise de-noising and direct wave eliminating based on SVD method for ground penetrating radar signals. *J. Appl. Geophys.* **144**, 125–133 (2017).
- Downs, C. & Jazayeri, S. Resolution enhancement of deconvolved ground penetrating radar images using singular value decomposition. *J. Appl. Geophys.* **193**, 104401 (2021).
- Donoho, D. L. De-noising by soft-thresholding. *IEEE Trans. Inf. Theory* **41**(3), 613–627 (1995).
- Jansen, M., Malfait, M. & Bultheel, A. Generalized cross validation for wavelet thresholding. *Signal Process.* **56**(1), 33–44 (1997).
- Chang, S. G., Yu, B. & Vetterli, M. Adaptive wavelet thresholding for image denoising and compression. *IEEE Trans. Image Process.* **9**(9), 1532–1546 (2000).
- Chen, X., Xu, C., Yang, X., & Tao, D. Attention-gan for object transfiguration in wild images. In *Proc. of the European Conf. on Computer Vision (ECCV)*, 164–180. (2018).

Author contributions

L.L. wrote the main manuscript text. L.C. and C.L. Guiding the paper. X.Y., T.W. and X.L. prepared figures. All authors reviewed the manuscript.

Funding

The authors are grateful to the key project of National Natural Science Foundation of China (G. No: 41930112).

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to L.C.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023