Letter

Intelligent Small Sample Defect Detection of Concrete Surface Using Novel Deep Learning Integrating Improved YOLOv5

Yongming Han, *Member, IEEE*, Lei Wang, Youqing Wang, *Senior Member, IEEE*, and Zhiqiang Geng

Dear Editor,

This letter presents an intelligent small sample defect detection of concrete surface using novel deep learning integrating the improved YOLOv5 based on the Wasserstein GAN (WGAN) enhancement algorithm. The proposed method is capable of producing top-notch data sets to address the issues of insufficient samples and substandard quality. Moreover, the proposed method can efficiently detect numerous minor flaws present in real concrete structures, thereby compensating for the drawbacks of current techniques in terms of real-time performance, practicality, and precision. The study findings reveal a noteworthy increase in the precision of the suggested approach when compared to other methods, reaching 86.2%.

Defects in concrete structures significantly impact their durability, service life and safety. The primary challenge lies in the fact that many surface defect detection methods necessitate predetermined inspection targets and parameters, which are often difficult to meet in practice. Numerous techniques are available; but they lack practicality. Due to the presence of numerous small defects within concrete, conventional inspection methods require high levels of accuracy, which can result in inaccurate results. Consequently, a more flexible and precise approach is urgently required for detecting concrete defects. The proposed method outlined in this letter offers a viable solution to this challenge.

Related work: To tackle the drawbacks of manually inspecting concrete buildings for defects, researchers have put a lot of effort into recent years and have developed four main types of methods: manual-based, ultrasonic-based, image processing technologies (IPTs) and deep learning-based ones.

Manual-based defect detection methods is the most traditional detection method. However, detecting concrete defects using human perception is time-consuming, labour-intensive and subjective, making it unsuitable for detecting concrete defects [1].

Ultrasonic-based methods for defect detection are primarily utilized to detect internal defects. Zhao *et al.* [2] proposed an ultrasonic echo detection method capable of detecting defects in insulation pull rods of high voltage circuit breakers. The defects were identified by introducing a range of common defects in the simulation model and detecting them through the analysis of echo propagation time and amplitude. Ultrasound is capable of identifying internal defects but struggles with surface defects, particularly those with 90° cracks.

IPTs can be overcome by using image processing-based techniques if the parameters are set beforehand. Dhule *et al.* [3] applied the edge detection with the Soble operator to identify cracks on walls. However, IPTs are tailored to a particular setup or database. Alterations to the settings or databases often cause the methods to malfunction.

The application of deep learning in defect detection has become increasingly popular in recent years and has yielded noteworthy outcomes due to technological advancements. Lin *et al.* [4] proposed a data-driven semantic segmentation network, known as WCU-Net, to selectively gather more information about fine cracks for detecting cracks in wood. Li and Xu [5] combined with a lightweight mobile convolutional mobile network feature extraction network to solve the problem of appearance defect detection in electronic products.

Previous related research suggests that detection speed and efficiency are progressing from manual detection methods to deep learning methods, particularly when combined with convolutional neural networks, leading to improved detection ability. However, the proposed YOLOv5 algorithm for machine modification not only addresses the issue of limited samples but also enhances detection of minor targets in concrete building surface defect features, producing better detection outcomes.

Problem statement: Defects in concrete buildings often have varying shapes and sizes, and measuring them using uniform standards can be challenging. Hence, to identify these challenging-to-standardize defects, deep learning is currently the most recommended approach.

To achieve high-precision defect detection with deep learning methods, it is necessary to prepare the dataset and train the defect detection model in advance. The quality of the defect detection model depends on the quality of the dataset. When there is an insufficient amount or quality of dataset available, the WGAN generation algorithm can produce enough high-quality samples to ensure the accuracy of the defect detection model.

However, identifying concrete surface defects accurately through ordinary deep learning models is problematic due to their small size. Our proposed solution is an integrated and improved YOLOv5 model, which not only generates high-quality datasets but also detects small-sample defects on concrete surfaces with high precision, thus enhancing the accuracy and efficiency of concrete surface defect detection [6].

Basic concepts:

1) UnSharp mask (USM): It is a linear filtering algorithm used to enhance image details [7].

2) WGAN: It is an enhanced algorithm for generative adversarial networks, which addresses the training instability issue in traditional GANs by introducing the Wasserstein distance metric [8].

3) General image augmentation (GIA): A series of common data augmentation methods including random cropping, random rotation, random scaling and random fusion splicing.

4) YOLOv5: YOLOv5 comprises a backbone network, a detection neck and three detection heads. After mosaic processing, training images are fed into the network. Image features are extracted at different scales using the backbone network. In the feature fusion, the detection neck connected to the backbone plays a role. For the prediction of smaller and larger targets, respectively, three types of feature maps are used. Finally, the feature maps are gridded and detected by the head [9].

Proposed integrated improved model: For the detection of concrete surface defects, an integrated and improved YOLOV5 model, which can effectively detect small targets, as shown in Fig. 1. The improved YOLOV5 model training consists of four stages: Stages 1−3 generate high-quality samples, perform image enhancement, and feed the dataset into the improved model for training. In the fourth stage, we verify the validity of the experiment by cross-comparing the detection results with several models and datasets. A detailed analysis of the four stages is shown as follows:

Corresponding authors: Youqing Wang and Zhiqiang Geng.

Citation: Y. Han, L. Wang, Y. Wang, and Z. Geng, "Intelligent small sample defect detection of concrete surface using novel deep learning integrating improved YOLOv5," *IEEE/CAA J. Autom. Sinica*, vol. 11, no. 2, pp. 545–547, Feb. 2024.

The authors are with the College of Information Science and Technology, Beijing University of Chemical Technology, Beijing 100029, and also with the Engineering Research Center of Intelligent PSE, Ministry of Education in China, Beijing 100029, China (e-mail: hanym@mail.buct.edu.cn; 2023200829@buct.edu.cn; wang_youqing@mail.buct.edu.cn; gengzhiqiang@mail.buct.edu.cn).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/JAS.2023.124035

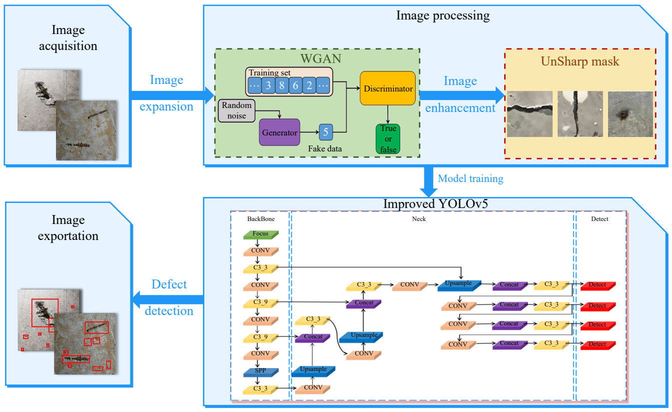

Fig. 1. The process of the proposed defect detection method.

Stage 1: The WGAN generative network is used to augment the limited data set. The generator and discriminator networks are first constructed using the PyTorch framework. A small number of prepared datasets are fed into the generator, and the generative power of the generator is improved by alternately training the generator and discriminator networks. Moreover, the networks are trained, and in each training iteration, real and generated images are fed separately into the discriminator network, the loss function is computed, and the parameters of the discriminator are updated by the backpropagation. Random noise is then fed into the generator network to produce dummy images, which are fed into the discriminator network, and again the loss function is calculated and the parameters of the generator are updated. The Wasserstein distance measure of the difference between the generator and the discriminator is then introduced to optimise the training process, and finally random noise is fed into the generator network, using the generator network to generate images similar to the training data. The image generated here is the output of Stage 1.

Stage 2: The USM is applied to image created in Step 1. The USM algorithm is a linear filtering technique used to enhance image details, commonly used in digital image processing. The principle of the USM algorithm is: first, Gaussian blur the image, then subtract the blurred image from the original image, and the result is the sharpened image.

$$y = \frac{x - w \times x_g}{1 - w} \tag{1}$$

where y represents the enhanced image, x represents the source image, w represents the weight (0.1 to 0.9), and x_g represents the source image after Gaussian blurring, and y is the output of Stage 2. The Gaussian kernel in this letter has a size of 2 and a standard deviation of 1.5.

Stage 3: The feature extraction module of the original YOLOv5 model fails to effectively leverage the shallow feature extraction module due to the presence of numerous small-sized targets in the concrete surface defects. The addition of a shallow feature extraction module is proposed to the original YOLOv5 model for detecting smaller targets. The default loss function for the rectangular box is CIoU Loss.

Beyond the 17th layer, the feature map undergoes further processing, including up-sampling, resulting in its expansion. Simultaneously, at the 20th layer, the obtained 160×160 feature map is concatenated with the feature map of the 2nd layer in the backbone network to obtain larger feature maps for detecting small targets. A small target detection layer is added at layer 31, with a total of four layers being employed for detection. The addition of a shallow feature extraction layer enables full fusion of front layer features with the original back layer features, improving the effectiveness in accurately detecting small targets and ultimately enhancing the detection performance of the model. The structure is shown in Fig. 2.

Stage 4: In order to facilitate the quantification of experimental results, this letter uses evaluation indicators of commonly used models in object detection, including precision, recall, mean average precision (mAP). Multiple datasets were inputted into the proposed improved model and compared against the single shot multibox detector (SSD) [10], YOLOv3 [11], and YOLOv5 models.

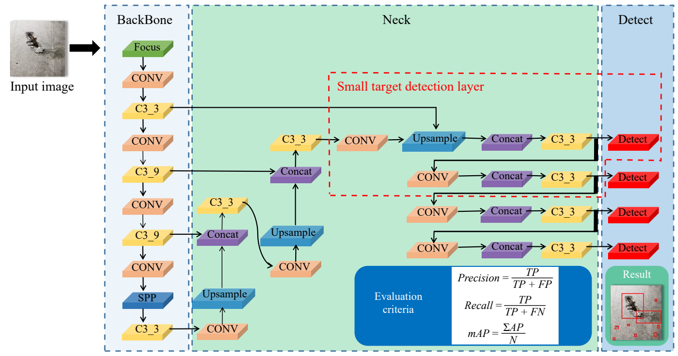

Fig. 2. The structure of the improved YOLOv5.

$$Precision = \frac{TP}{TP+FP},\ Recall = \frac{TP}{TP+FN},\ mAP = \frac{\sum AP}{N} \tag{2}$$

where TP refers to the positive class detected as positive; FP refers to the negative class detected as positive; FN refers to a positive class that is detected as negative. mean average precision (mAP) represents the average value of all categories of AP

Experiments: Our experiments use an Intel Corei7-9700 CPU, NVIDIA GeForce GTX1600 GPU, 16G RAM and Windows 10 64-bit. Programming languages with Python 3.7, OpenCV 4.5.1 and Pytorch 1.7.1 are used.

To exhibit the efficacy of the proposed approach in this letter, we develope six datasets. The specifics of six datasets are detailed in Table 1. DataOR denotes the uncompromised dataset captured initially. Meanwhile, DataCA is the dataset generated after the GIA, and DataWG is the dataset expanded using DataOR, both of which are variations of the original data. DataOR*, DataCA*, and DataWG* are the altered datasets generated from DataOR, DataCA, and DataWG, respectively, through the USM enhancement. The training parameters in this experiment are shown in Table 2.

Table 1. Summary of Six Datasets

Dataset	Number of images		
	Total	Training set	Test set
DataOR	200	160	40
DataCA	1000	800	200
DataWG	1000	800	200
DataOR*	200	160	40
DataCA*	1000	800	200
DataWG*	1000	800	200

Table 2. Experimental Parameters

Parameter	Meaning	Value
Image_size	Resolution of the training image	640×640
Epochs	Number of iterations	500
Batch_size	Size of each training batch	16

To illustrate the importance of each step in the proposed concrete defect detection method, comparative experiments are carried out. Table 3 presents the experimental results. In the experiment, each component of the proposed defect detection in this letter is varied to develop different defect detection methods named M1 to M6 as presented in Table 4. The 6th method (M6) in this letter introduces the proposed defect detection method. Samples of the six datasets are shown in Fig. 3.

Tables 3 and 4 demonstrate that M6 exhibits the greatest performance across all three assessment metrics. The evaluation metrics for M1, M2 and M4 are compared with those of M3, M5 and M6 and the

Table 3. Experimental Results of Models on DataOR and DataWG*

Model	Dataset	Precision	Recall	mAP@0.5
SSD	DataOR	0.493	0.271	0.302
	DataWG*	0.562	0.404	0.502
YOLOv3	DataOR	0.525	0.250	0.288
	DataWG*	0.698	0.496	0.688
YOLOv5	DataOR	0.552	0.306	0.487
	DataWG*	0.743	0.612	0.742
Improved YOLOv5	DataOR	0.562	0.417	0.485
	DataWG*	**0.862**	**0.737**	**0.805**

Table 4. Six Different Methods and Experimental Results for Detecting Concrete Defects

Method	Processing steps				Dataset	Experimental results		
	Data expansion (WGAN)	Data expansion (traditional)	Data enhancement	Defect detection		Precision	Recall	mAP@0.5
M1				√	DataOR	0.562	0.417	0.485
M2	√			√	DataWG	0.733	0.614	0.690
M3			√	√	DataOR*	0.609	0.486	0.503
M4		√		√	DataCA	0.684	0.463	0.601
M5		√	√	√	DataCA*	0.778	0.656	0.712
M6	√		√	√	**DataWG***	**0.862**	**0.737**	**0.805**

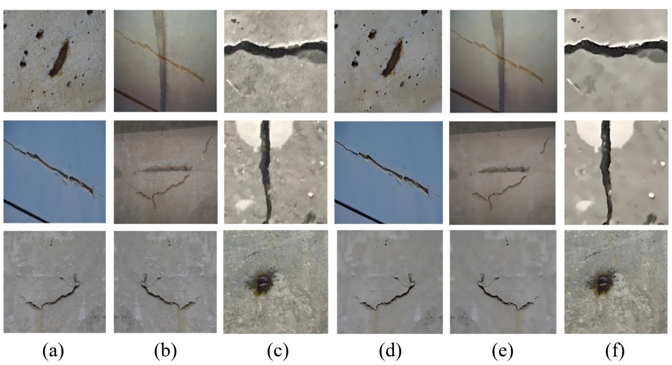

Fig. 3. Image samples in datasets. (a) DataOR; (b) DataCA; (c) DataWG; (d) DataOR*; (e) DataCA*; (f) DataWG*.

use of the USM is shown to considerably improve the model detection accuracy. A comprehensive analysis is conducted after comparing the assessment metrics of M3, M5 and M6. The results indicate that M6 successfully tackles the challenge of defect detection in limited sample sizes, thereby improving the overall detection effectiveness of the defect detection model. The WGAN+USM approach offered in this research surpasses traditional data expansion methods.

Conclusions: In this letter, we propose a novel intelligent small sample defect detection method combining the WGAN and the improved YOLOv5 algorithm to overcome the limitations of concrete surface defect detection methods in terms of real-time performance, practicality, and accuracy. The proposed method employs the WGAN and UnSharp mask techniques to refine the dataset and enhance the YOLOv5 model with an additional target detection layer. Meanwhile, the proposed method can address imprecise detection owing to inadequacies in the dataset or defects of insufficient size. Eevidence suggests that the proposed method can enhance the accuracy of detecting defects on concrete surfaces.

The proposed method's drawback is that its model parameters rely on empirical adjustment, which is not sufficiently convenient. In forthcoming research, we will adopt an adaptive neural network approach to ascertain the necessary model parameters and improve the outcome.

Acknowledgments: This work was supported by the National Natural Science Foundation of China (21978013) and the Fundamental Research Funds for the Central in China (XK1802-4).

References

[1] L. Yang, B. Li, W. Li, H. Brand, B. Jiang, and J. Xiao, "Concrete defects inspection and 3D mapping using CityFlyer quadrotor robot," *IEEE/CAA J. Autom. Sinica*, vol. 7, no. 4, pp. 991–1002, 2020.

[2] R. Y. Zhao, J. Li, Y. Chen, *et al.*, "Defects detection of high voltage insulation pull rod based on ultrasonic wave method," in *Proc. IEEE Int. Conf. High Voltage Engineering and Application*, 2020, pp. 1–4. DOI: 10.1109/ICHVE49031.2020.9279505.

[3] J. J. Dhule, N. B. Dhurpate, S. S. Gonge, and G. M. Kandalkar, "Edge detection technique used for identification of cracks on vertical walls of the building, " in *Proc. Int. Conf. Computing and Network Communi.*, 2015, pp. 263–268. DOI: 10.1109/CoCoNet.l2015.7411196.

[4] Y. Lin, Z. Xu, D. Chen, Z. Ai, Y. Qiu, and Y. Yuan, "Wood crack detection based on data-driven semantic segmentation network," *IEEE/CAA J. Autom. Sinica*, vol. 10, no. 6, pp. 1510–1512, 2023.

[5] Y. Li and J. Xu, "Electronic product surface defect detection based on a MSSD network," in *Proc. 4th IEEE Information Technology, Networking, Electronic and Automation Control Conf.*, 2020, pp. 773–777. DOI: 10.1109/ITNEC48623.2020.9084756.

[6] L. Zhu, J. Zhang, and C. Jia, "An improved YOLOv5-based method for surface defect detection of steel plate," in *Proc. China Automation Congr.*, 2022, pp. 2233–2238. DOI: 10.1109/CAC57257.2022.1005 5055.

[7] A. Zaafouri, M. Sayadi, and F. Fnaiech, "A developed unsharp masking method for images contrast enhancement," in *Proc. Int. Multi-Conf. Systems, Signals & Devices, Sousse*, 2011, pp. 1–6. DOI: 10.1109/ SSD.2011.5767378.

[8] J. Engelmann and S. Lessmann, "Conditional Wasserstein GAN-based oversampling of tabular data for imbalanced learning," *Expert Systems Applications*, vol. 174, p. 114582, Jan. 2021.

[9] Z. Wang, L. Jin, S. Wang, and H. Xu, "Apple STEM/CALYX real-time recognition using YOLO-v5 algorithm for fruit automatic loading system," *Postharvest Biology Technology*, vol. 185, p. 111808, 2022.

[10] A. Del Signore, A. J. Hendriks, H. R. Lenders, R. S. Leuven, and A. M. Breure, "Development and application of the SSD approach in scientific case studies for ecological risk assessment," *Environmental Toxicology and Chemistry*, vol. 35, no. 9, pp. 2149–2161, May 2016.

[11] J. Redmon and A. Farhadi, "YOLOv3: An incremental improvement," arXiv preprint arXiv: 1804.02767, Apr. 2018.