Proceedings of the 4th International Conference on Software Engineering and Machine Learning
DOI: 10.54254/2755-2721/2026.34447

Research on Generative Models for Few-Shot Industrial Surface Defect Detection

Yizhou Chen[1], Wanli Cheng[2*]

[1]*Changwai Bilingual School, Changzhou, China*
[2]*Leeds College, Southwest Jiaotong University, Chengdu, China*
**Corresponding Author. Email: wanlicheng351@gmail.com*

Abstract. Industrial surface defect detection is essential for manufacturing quality control, but deep learning-based inspection systems often rely on sufficient and well-labelled defect samples. In practical industrial scenarios, defect samples are scarce, imbalanced, and difficult to collect, which limits the robustness and generalization ability of detection networks. This paper reviews how generative models support few-shot industrial defect detection. It focuses on variational autoencoders, generative adversarial networks, and diffusion models, and compares their roles in image-level augmentation, feature-level representation learning, controllable defect synthesis, and image-mask generation. The review shows that variational autoencoder-based methods are useful for stable latent representation and feature generation, generative adversarial network-based methods are better at realistic and controllable image augmentation, and diffusion-based methods have strong potential for high-quality paired image-mask generation. However, the usefulness of generated samples remains limited by data scarcity, distribution bias, imperfect local textures, and inaccurate small-defect masks. This review therefore provides a comparative framework for understanding the advantages and limitations of different generative models in few-shot industrial surface defect detection.

Keywords: industrial surface defect detection, few-shot learning, generative models, data augmentation, diffusion models

1. Introduction

Surface defect inspection is a crucial part of quality control in industrial manufacturing. Quality inspection can account for a considerable proportion of manufacturing revenue, and reliable inspection directly affects product quality, production efficiency, and final output [1]. For a long time, manual visual inspection has been widely used in industrial production. However, manual inspection is easily affected by fatigue, subjective judgment, and differences in inspector experience. Long-term visual inspection may also create physical burden for inspectors and increase operating costs [2]. As a result, traditional inspection methods often suffer from unstable results and limited scalability [3].

Deep learning-based defect detection has been introduced to reduce these limitations. With the help of convolutional feature extraction and data-driven training, such methods have been applied to

DOI: 10.54254/2755-2721/2026.34447

hot-rolled steel strips, printed circuit boards, metal surfaces, steel rails, and glass bottles. Though compared with manual inspection, deep learning methods can achieve the detection enhancement on consistency and efficiency, yet their performance still depends heavily on sufficient labelled data. Meanwhile, in real operation scenarios, where defect data are often scarce, unevenly distributed across categories, it is difficult to collect for new defect types [4].

Therefore, with the purpose of alleviating this problem, generative models is a possible way. These models offer a way to move beyond sole reliance on existing defect samples by synthesizing supplementary images, generating feature representations, or produce paired image-mask samples for downstream detection and segmentation. In few-shot industrial defect detection, this ability is particularly valuable because even a small number of additional reliable samples may improve model training. However, new problems, including unrealistic textures, biased distributions, unstable training, and poor alignment with downstream tasks can also be introduced.

This paper reviews recent applications of three major types of generative models in few-shot industrial surface defect detection: variational autoencoders, generative adversarial networks, and diffusion models. The contribution of this review is threefold. First, it organizes existing methods according to their generation mechanisms and downstream roles. Second, it compares their strengths and limitations through a unified analytical framework. Third, it summarizes remaining challenges and future research directions, especially data scarcity, realism, distribution bias, and small-defect mask alignment.

2. Generative models for few-shot industrial defect detection

2.1. Overall architecture of defect detection technology

According to the taxonomy of deep industrial image anomaly detection, generative-model-assisted few-shot defect detection can be separated into three interacting components: the detection network, the generation model, and the learning paradigm [2]. As the core module, the detection network locates and identifies defects. The generation model provides supplementary data or feature representations, while the learning paradigm determines how generated and real samples are used during training.

As shown in Figure 1, the generation model and the detection network can form a training loop. The generator provides synthetic defect-related information, and the detector uses this information for recognition, localization, or segmentation. The detector can also provide feedback through loss calculation, performance evaluation, and task-specific constraints. This interaction is especially important in few-shot settings, where the detector cannot rely on large numbers of labelled real defects.

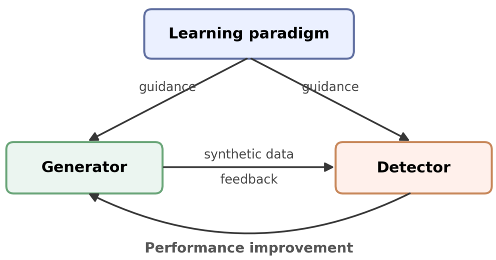

Figure 1. Framework of generative-model-assisted defect detection (picture credit: original)

DOI: 10.54254/2755-2721/2026.34447

2.2. Applications of generative models under few-shot conditions

Defect samples are often scarce and imbalanced in industrial surface inspection. Generative models can alleviate this problem by synthesizing additional defect images, producing useful feature representations, or supporting mask-level annotation. The key question is not only whether a model can generate visually plausible samples, but also whether these samples can improve downstream detection and segmentation.

Industrial defects are usually small, irregular, low-contrast, and strongly dependent on material texture. Therefore, generative models need to capture both global product structure and local defect morphology. The following subsections review variational autoencoder-based methods, generative adversarial network-based methods, and diffusion model-based methods.

2.2.1. Variational autoencoder-based methods

Variational autoencoders (VAEs) were introduced as a variational inference-based generative framework for latent-variable modeling [5]. In industrial defect detection, VAE-based methods can learn compact latent representations and generate supplementary samples or features. Their training process is generally more stable than that of adversarial models, but direct image generation may suffer from blurred details and limited texture realism.

To address imbalanced defect data, Yun et al. applied a conditional variational autoencoder (CVAE) to metal surface defect detection [6]. The method uses category labels together with latent variables, allowing the decoder to generate class-related samples. A convolutional neural network (CNN) classifier was then trained with generated samples. The results showed that data augmentation with the conditional convolutional variational autoencoder improved classification accuracy and F1-score in most settings, indicating that VAE-based augmentation can reduce class imbalance under limited-data conditions [6].

Recent VAE-based studies have also moved from image-level augmentation toward feature-level generation. Zhang et al. proposed a smooth VAE-based model for pipeline digital radiography defect detection under small-sample conditions [7]. The method witnesses the improvement of feature generation by adjustment of the reconstruction loss and introduction of decoupling regularization in the latent space. This design reduces the influence of external interference and helps the model learn more stable feature representations. These studies suggest that VAE-based methods are particularly useful when representation learning and feature robustness are more important than highly realistic image synthesis.

2.2.2. Generative adversarial network-based methods

Generative adversarial networks (GANs) consist of a generator and a discriminator trained in an adversarial manner [8]. The generator synthesizes samples, while the discriminator distinguishes generated samples from real ones. Compared with VAEs, GANs are often better at producing visually realistic images, which makes them suitable for defect image augmentation. However, they may also suffer from unstable training, mode collapse, and distribution bias.

Guo et al. proposed ISU-GAN for unsupervised small-sample defect detection [9]. The method combines a GAN-based framework with a structural similarity index measure (SSIM)-based defect extraction strategy. It also uses cycle-consistency and identity constraints to improve convergence and preserve image details. To enhance defect-related feature learning, the generator adopts a U-Net

DOI: 10.54254/2755-2721/2026.34447

structure and a squeeze-and-excitation module. Experiments on the DAGM 2007 dataset showed strong performance, with an average detection accuracy of 98.43% and an F1-score of 0.9792 [9].

He et al. proposed a defect image generation method based on StyleGAN2-ADA and feature disentanglement [10]. The method learns from defect-free and defect-rich images and uses a defect mask attention module to guide defect generation toward object surfaces rather than irrelevant backgrounds. This design improves the spatial plausibility of generated defects and helps reduce unrealistic generation patterns. The authors reported improved results on the MVTec Anomaly Detection (MVTec AD) and Visual Anomaly (VisA) datasets, with a notable improvement in the kernel inception distance (KID) metric [10].

Yen et al. further explored CycleGAN-based data augmentation for critical-dimension scanning electron microscopy (CD-SEM) images in semiconductor manufacturing [11]. Instead of relying only on generated images, the method filters low-quality samples using SSIM and mean absolute difference metrics. The study also compares different generator structures and shows that a ResNet-based generator is more effective at preserving local structural consistency in CD-SEM images. These works show that GAN-based methods are gradually moving from simple image generation toward quality-controlled and task-oriented defect augmentation.

2.2.3. Diffusion model-based methods

Diffusion models generate images through a gradual denoising process. Noise is first added during the forward process, and the model learns to reverse this process step by step. Compared with VAEs and GANs, diffusion models often provide higher generation quality and more stable training, but they usually require higher computational cost and slower sampling. These properties make diffusion models promising but also challenging for few-shot industrial defect detection.

Hu et al. proposed AnomalyDiffusion for few-shot anomaly image generation [12]. The method generates paired anomaly images and masks, which is important for downstream anomaly localization. It separates anomaly information into anomaly embedding and spatial embedding, and then uses adaptive attention re-weighting to improve the alignment between generated anomalies and masks. On the MVTec dataset, the method produced realistic and diverse defect samples and achieved an average area under the receiver operating characteristic curve (AUROC) of 99.1% and average precision (AP) of 81.4% in downstream anomaly localization [12].

Shi et al. proposed a consistency-modeling-based method for few-shot defect image generation [13]. The method, known as DefectDiffu, uses a text-guided decoupled integration architecture (TDIA) to separate background, defect, and fusion components. This design improves the consistency between generated defects and product backgrounds. It also controls defect severity by adjusting the perturbation scale. The reported Fréchet inception distance (FID) score indicates that this method can generate competitive defect images under few-shot conditions [13].

Wang et al. proposed an uncertainty-guided dual-stage conditional diffusion model (UGDS-CDM) integrated with Vision Mamba-KAN and expert knowledge for small-sample surface defect detection [14]. By using class guidance and uncertainty estimation during generation, the method reduces dependence on a single defect pattern and improves cooperation between generation and downstream detection. This trend suggests that diffusion models are increasingly used not only as independent generators but also as task-oriented modules that work with detection networks.

DOI: 10.54254/2755-2721/2026.34447

2.3. Comparison of generative models

The three types of models differ in generation mechanism, stability, image quality, and downstream usefulness. Table 1 summarizes their main roles in few-shot industrial surface defect detection. This comparison shows that no single model is optimal in all situations. VAE-based methods are useful for stable representation learning, GAN-based methods are suitable for realistic image augmentation, and diffusion-based methods are promising for high-quality image-mask generation.

Table 1. Comparison of VAE, GAN, and diffusion models for few-shot industrial defect detection

Model type	Main role	Strengths	Limitations and suitable use
VAE-based methods	Image-level or feature-level generation	Stable latent representation learning and useful feature augmentation	Direct image realism is limited; suitable when data imbalance and feature robustness are central
GAN-based methods	Realistic defect image generation and controllable augmentation	Strong visual realism and flexible defect synthesis	Training instability and distribution bias remain risks; suitable for controllable image augmentation
Diffusion-based methods	Paired image-mask generation and task-oriented synthesis	High generation quality, stable training, and spatial alignment potential	Generation is slow and computationally expensive; suitable for mask alignment and fine-grained defect control

3. Challenges and future directions

3.1. Data scarcity and category imbalance

Generative models can reduce the pressure caused by scarce defect samples, but they cannot completely remove the need for real defect data. If the original few-shot set is too small or too biased, the generated samples may simply reproduce limited defect patterns. This problem is especially serious as very different textures, shapes, and defect distributions are involved in different product categories. For example, even high-performing diffusion-based methods may show uneven localization performance across categories [12].

Category imbalance is another persistent problem in industrial production. Although some defect categories appear frequently, others are rare but still important for quality control. Meanwhile, generative models can increase sample quantity, but the limitation is that they may fail to capture the true diversity of rare defects. Therefore, future work should pay more attention to category-balanced generation, rare-defect modeling, and evaluation protocols that measure performance separately for different defect types.

3.2. Realism and distribution bias

Realism is not only a visual issue. It is a must for useful generated samples to integrate real defects in local texture, boundary shape, background consistency, and defect-product interaction in industrial defect detection. If generated samples look plausible at a global level but differ from real samples in local details, a distribution gap might be created. Furthermore, such samples can even harm downstream detectors by encouraging them to learn artificial patterns rather than real defect features.

GAN-based methods have improved realism through feature disentanglement, mask attention, and quality filtering [10, 11]. However, the need for these constraints also shows that generated

DOI: 10.54254/2755-2721/2026.34447

samples may still contain unrealistic regions or biased textures. Future studies should therefore evaluate generated data not only with image-quality metrics, but also through downstream detection accuracy, cross-product transfer, and robustness under real production variation.

3.3. Small-defect mask alignment

Small-defect mask alignment remains difficult because industrial defects often have blurred boundaries, irregular shapes, and weak contrast against complex backgrounds. This problem becomes more serious in pixel-level segmentation tasks, where small spatial errors can greatly affect training.

Diffusion-based methods have begun to address this issue by jointly generating anomaly images and masks or by using spatial embeddings and attention re-weighting [12]. Nevertheless, accurate alignment is still difficult for tiny scratches, stains, holes, and texture defects. Future work should combine generative modeling with stronger spatial constraints, boundary-aware losses, and detector-in-the-loop feedback so that generated masks are directly useful for segmentation networks.

3.4. Future research directions

Future research should move from sample-level augmentation toward task-oriented generation. First, controllable generation should allow users to specify defect type, location, severity, and background condition. Second, cross-product transfer should be strengthened so that a model trained on one product category can support defect generation in another category with few real samples. Third, generated data should be optimized together with downstream detection or segmentation networks instead of being evaluated only as independent images. Finally, future methods should report both generation metrics and downstream task metrics to show whether synthetic samples truly improve industrial inspection performance.

4. Conclusion

This paper reviewed the role of generative models in few-shot industrial surface defect detection. The analysis shows that VAE-based, GAN-based, and diffusion-based methods follow different technical routes and are suitable for different needs. VAE-based methods are relatively stable and useful for latent representation learning and feature-level augmentation, but their direct image generation quality is often limited. GAN-based methods are stronger in realistic image synthesis and controllable defect augmentation, especially when defect location and background consistency can be constrained by masks or quality filtering. Diffusion-based methods provide high generation quality and stable training, and they are especially promising for paired image-mask generation and task-oriented defect synthesis. However, their sampling speed and computational cost remain important limitations. Across these methods, the central challenge is no longer simply how to generate more images. The more important question is whether generated samples are realistic, diverse, spatially aligned, and useful for downstream detection or segmentation. Few-shot industrial inspection still faces persistent data scarcity, category imbalance, distribution bias, and small-defect mask alignment problems. Therefore, future studies should focus on controllable generation, cross-product transfer, boundary-aware mask synthesis, and closer integration between generative models and downstream networks. A stronger evaluation framework is also needed. Generated samples should be judged not only by visual quality metrics, but also by whether they improve real inspection performance under limited-data conditions. In this sense, generative models should be viewed as task-supporting components in an industrial inspection system rather than as isolated image generators. Such a view can better connect sample synthesis with practical quality inspection and can guide researchers to design generation strategies that are measurable, controllable, and useful in real industrial scenarios.

Author contribution

All the authors contributed equally and their names were listed in alphabetical order.

References

[1] Ma, Y., Yin, J., Huang, F., & Li, Q. (2024). Surface defect inspection of industrial products with object detection deep networks: a systematic review. Artificial Intelligence Review, 57, Article 333. https: //doi.org/10.1007/s10462-024-10956-3
[2] Liu, J., Xie, G., Wang, J., Li, S., Wang, C., Zheng, F., & Jin, Y. (2024). Deep industrial image anomaly detection: a survey. Machine Intelligence Research, 21(1), 104–135. https: //doi.org/10.1007/s11633-023-1459-z
[3] Ameri, R., Hsu, C.-C., & Band, S. S. (2024). A systematic review of deep learning approaches for surface defect detection in industrial applications. Engineering Applications of Artificial Intelligence, 130, Article 107717. https: //doi.org/10.1016/j.engappai.2023.107717
[4] He, X., Chang, Z., Zhang, L., Xu, H., Chen, H., & Luo, Z. (2022). A survey of defect detection applications based on generative adversarial networks. IEEE Access, 10, 113493–113512. https: //doi.org/10.1109/ACCESS.2022.3217227
[5] Kingma, D. P., & Welling, M. (2013). Auto-encoding variational Bayes. arXiv. https: //doi.org/10.48550/arXiv.1312.6114
[6] Yun, J. P., Shin, W. C., Koo, G., Kim, M. S., Lee, C., & Lee, S. J. (2020). Automated defect inspection system for metal surfaces based on deep learning and data augmentation. Journal of Manufacturing Systems, 55, 317–324. https: //doi.org/10.1016/j.jmsy.2020.03.009
[7] Zhang, T., You, T., Liu, Z., Rehman, S. U., Shi, Y., & Munshi, A. (2025). Small sample pipeline DR defect detection based on smooth variational autoencoder and enhanced detection head faster RCNN. Applied Intelligence, 55(8), Article 716. https: //doi.org/10.1007/s10489-025-06590-3
[8] Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A., & Bengio, Y. (2014). Generative adversarial nets. In Z. Ghahramani, M. Welling, C. Cortes, N. D. Lawrence, & K. Q. Weinberger (Eds.), Advances in Neural Information Processing Systems 27 (pp. 2672–2680). Curran Associates, Inc.
[9] Guo, Y., Zhong, L., Qiu, Y., Wang, H., Gao, F., Wen, Z., & Zhan, C. (2022). Using ISU-GAN for unsupervised small sample defect detection. Scientific Reports, 12, Article 11604. https: //doi.org/10.1038/s41598-022-15855-7
[10] He, Z., Wu, K., & Wen, Y. (2025). Defect image generation through feature disentanglement using StyleGAN2-ADA. Neurocomputing, 647, Article 130455. https: //doi.org/10.1016/j.neucom.2025.130455
[11] Yen, A., Chang, N., Chien, J., Chuang, L., & Lee, E. (2026). CycleGAN-based data augmentation for scanning electron microscope images to enhance integrated circuit manufacturing defect classification. Electronics, 15(4), Article 803. https: //doi.org/10.3390/electronics15040803
[12] Hu, T., Zhang, J., Yi, R., Du, Y., Chen, X., Liu, L., Wang, Y., & Wang, C. (2024). AnomalyDiffusion: Few-shot anomaly image generation with diffusion model. Proceedings of the AAAI Conference on Artificial Intelligence, 38(8), 8526–8534. https: //doi.org/10.1609/aaai.v38i8.28696
[13] Shi, Q., Wei, J., Shen, F., & Zhang, Z. (2025). Few-shot defect image generation based on consistency modeling. In A. Leonardis, E. Ricci, S. Roth, O. Russakovsky, T. Sattler, & G. Varol (Eds.), Computer vision – ECCV 2024: 18th European Conference, Milan, Italy, September 29–October 4, 2024, Proceedings, Part LXXVI (Lecture Notes in Computer Science, Vol. 15134, pp. 360–376). Springer. https: //doi.org/10.1007/978-3-031-73116-7_21
[14] Wang, X., Luo, K., Liu, M., Zheng, R., & Li, Y. (2026). UGDS-CDM: A novel uncertainty-guided dual-stage conditional diffusion model and vision mamba-KAN integrating expert knowledge for surface defect detection under small samples. Advanced Engineering Informatics, 69(Part C), Article 104011. https: //doi.org/10.1016/j.aei.2025.104011