



ExDD: Explicit Dual Distribution Learning for Surface Defect Detection via Diffusion Synthesis

Muhammad Aqeel¹(✉) , Federico Leonardi¹, and Francesco Setti^{1,2} 

¹ Department of Engineering for Innovation Medicine, University of Verona, Strada le Grazie 15, Verona, Italy
muhammad.aqeel@univr.it

² Qalyyco S.r.l., Strada le Grazie 15, Verona, Italy

Abstract. Industrial defect detection systems face critical limitations when confined to one-class anomaly detection paradigms, which assume uniform outlier distributions and struggle with data scarcity in real-world manufacturing environments. We present ExDD (**Explicit Dual Distribution**), a novel framework that transcends these limitations by explicitly modeling dual feature distributions. Our approach leverages parallel memory banks that capture the distinct statistical properties of both normality and anomalous patterns, addressing the fundamental flaw of uniform outlier assumptions. To overcome data scarcity, we employ latent diffusion models with domain-specific textual conditioning, generating in-distribution synthetic defects that preserve industrial context. Our neighborhood-aware ratio scoring mechanism elegantly fuses complementary distance metrics, amplifying signals in regions exhibiting both deviation from normality and similarity to known defect patterns. Experimental validation on KSDD2 demonstrates superior performance (94.2% I-AUROC, 97.7% P-AUROC), with optimal augmentation at 100 synthetic samples.

Keywords: Surface Defect Detection · Latent Diffusion Model · Synthetic Images

1 Introduction

Surface defect detection is a cornerstone of industrial quality control, where even microscopic imperfections in materials like copper, steel, or marble can lead to catastrophic failures in downstream applications [5, 17, 20]. Traditional computer vision approaches, struggle with the inherent variability of industrial defects, prompting a shift toward deep learning [6]. However, supervised methods face a critical limitation: the scarcity of annotated defect data due to their rarity in production lines [13]. To address this, recent works like [15] and [18] have popularized one-class anomaly detection, which trains exclusively on normal samples. While effective in controlled settings, these methods implicitly assume anomalies

are uniformly distributed outliers—a flawed premise for structured defects that occupy distinct distributions in feature space [9].

The reliance on one-class paradigms also ignores a key insight: industrial defects often exhibit consistent patterns with distinctive visual characteristics that can be separated from normal textures when properly represented in feature space [19]. Recent attempts to model anomaly distributions, such as patch-level density estimation [10], lack explicit distribution separation, while dual sub-space re-projection [19] oversimplifies defect geometry. Furthermore, synthetic anomaly generation methods like [12] often produce out-of-distribution artifacts due to adversarial training or unrealistic perturbations, as noted by [11]. This misalignment between synthetic and real defects undermines feature learning, particularly for subtle anomalies [8].

Recent advances in diffusion models offer a promising solution. By leveraging text-conditional generation, where defect descriptions in natural language guide the synthesis process, [11] demonstrated that latent diffusion models (LDMs) can synthesize in-distribution defects that preserve the statistical properties of real anomalies. However, their framework treats synthesis as a preprocessing step, decoupling it from the detection pipeline. Meanwhile, self-supervised methods like [3] improve robustness through pretext tasks but fail to explicitly model the defect distribution, resulting in suboptimal separability.

In this paper, we propose *ExDD* (**Explicit Dual Distribution**), a unified framework that bridges explicit dual distribution modeling and diffusion-based defect synthesis for surface inspection. Unlike prior work, ExDD jointly optimizes two memory banks: (1) a normal memory encoding nominal feature distributions and (2) a defect memory populated by diffusion-synthesized anomalies. The synthesis process uses text prompts derived from domain expertise (e.g., “metallic scratches”) to generate defects that align with the true anomaly distribution, as validated by [8]. Crucially, our dual memory architecture enables neighborhood-aware ratio scoring, which amplifies deviations from normality while suppressing false positives caused by normal feature variations a common failure mode in one-class methods [10].

The main contributions of our paper are threefold:

- **Dual Distribution Learning:** We formalize surface defect detection as a dual feature distribution separation problem, explicitly modeling both normal and defect feature distributions via memory banks.
- **Diffusion-Augmented Training:** A text-conditional LDM synthesizes in-distribution defects, expanding the defect memory while preserving geometric fidelity.
- **Ratio Scoring:** A novel scoring mechanism combines distance-to-normal and similarity-to-defect metrics, leveraging the dual memory structure for robust decision boundaries.

2 Related Work

Surface defect detection has advanced through three interconnected research streams: one-class normality modeling, synthetic defect generation, and self-supervised feature learning. One-class anomaly detection dominates industrial applications due to data scarcity, with methods like PatchCore [15] using memory banks of nominal features and PaDiM [10] modeling patch-wise distributions. However, these approaches struggle with structured defects like scratches that occupy distinct feature distributions [9], as highlighted by failures in detecting fine marble cracks [17].

Synthetic data generation addresses annotation scarcity but risks producing unrealistic artifacts. Early methods using random noise [12] often distort defect semantics, while GAN-based approaches improved realism but suffered from mode collapse [1]. Diffusion models offer a robust alternative, with [8] generating defects via text prompts aligned with domain expertise. However, most methods decouple synthesis from detection, preventing joint optimization—a gap addressed by ExDD’s integrated framework.

Self-supervised methods learn features without defect labels through pseudo-label refinement [3] and meta-learning for threshold adaptation [2, 5], though many require partial annotations [7]. Hybrid approaches like DRAEM [18] train discriminative reconstructions but ignore defect feature structures. Recent work with latent diffusion models [11] synthesizes in-distribution defects but decouples synthesis from detection. In contrast, ExDD unifies self-supervised principles with explicit modeling of separate normal and anomalous distributions, ensuring synthetic and real defects form a cohesive anomaly subspace while bridging generation and detection through an integrated framework.

3 ExDD Framework

We propose ExDD, a dual memory bank paradigm that extends memory-based anomaly detection by explicitly modeling both normal and anomalous feature distributions (Fig. 1). In this section, we formalize the problem setup (Sect. 3.1), describe our dual memory bank architecture (Sect. 3.2), detail our diffusion-based synthetic anomaly generation approach (Sect. 3.3), and present our novel anomaly scoring mechanism (Sect. 3.4).

3.1 Problem Formulation

Let \mathcal{X}_N denote the set of nominal images ($\forall x \in \mathcal{X}_N : y_x = 0$) available during training, where $y_x \in \{0, 1\}$ indicates whether an image x is nominal (0) or anomalous (1). Similarly, \mathcal{X}_T represents the test set, with $\forall x \in \mathcal{X}_T : y_x \in \{0, 1\}$. Let \mathcal{X}_A denote the set of anomalous samples, which may be available in limited quantity or generated synthetically.

Traditional anomaly detection operates in a one-class paradigm, modeling only the distribution of normality $P(\mathcal{X}_N)$ and measuring deviations from this

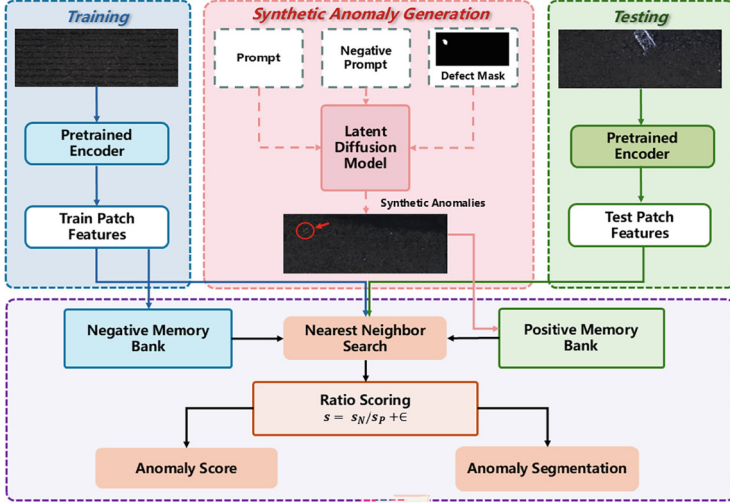


Fig. 1. Overview of the ExDD framework, illustrating the training process with pre-trained encoder and patch feature extraction, synthetic anomaly generation using diffusion models with prompt guidance, testing workflow, and the dual memory bank architecture with ratio-based anomaly scoring mechanism.

established distribution. This approach assumes anomalies are uniformly distributed in the complement space, which is not true for industrial defects with consistent patterns. Our key insight is that industrial anomalies often form distinct distributions in feature space. By explicitly modeling both distributions, we create a more discriminative decision boundary.

Following established protocols [1,2], we use a network ϕ pre-trained on ImageNet as our feature extractor. We denote $\phi_{i,j} = \phi_j(x_i)$ as the features for image $x_i \in \mathcal{X}$ at hierarchy level j of network ϕ , where $j \in \{1, 2, 3, 4\}$ typically indexes feature maps from ResNet architectures.

3.2 Dual Memory Bank Architecture

The core innovation of ExDD is its parallel memory bank architecture, which explicitly models normal and anomalous feature distributions.

Locally Aware Patch Features. To extract an informative description of patches, we employ the local patch descriptors defined in [15]. For a feature map tensor $\phi_{i,j} \in \mathbb{R}^{c^* \times h^* \times w^*}$ with depth c^* , height h^* , and width w^* , we denote $\phi_{i,j}(h, w) = \phi_j(x_i, h, w) \in \mathbb{R}^{c^*}$ as the c^* -dimensional feature slice at position (h, w) . To incorporate local spatial context, we define the neighborhood of position (h, w) with patch size p as:

$$\mathcal{N}_p^{(h,w)} = \{(a, b) | a \in [h - \lfloor p/2 \rfloor, \dots, h + \lfloor p/2 \rfloor], b \in [w - \lfloor p/2 \rfloor, \dots, w + \lfloor p/2 \rfloor]\} \quad (1)$$

The locally aware patch features at position (h, w) are computed as:

$$\phi_{i,j}(\mathcal{N}_p^{(h,w)}) = f_{\text{agg}}(\{\phi_{i,j}(a, b) | (a, b) \in \mathcal{N}_p^{(h,w)}\}) \quad (2)$$

where f_{agg} is an aggregation function, implemented as adaptive average pooling with a 3×3 window size. For a feature map tensor $\phi_{i,j}$, its collection of locally aware patch features is:

$$\mathcal{P}_{s,p}(\phi_{i,j}) = \{\phi_{i,j}(\mathcal{N}_p^{(h,w)}) | h, w \bmod s = 0, h < h^*, w < w^*, h, w \in \mathbb{N}\} \quad (3)$$

where s is a stride parameter (set to 1 in our implementation).

We extract features from both layer 2 and layer 3 of the backbone network. Features from layer 3 are upsampled to match layer 2 dimensions, then concatenated:

$$\mathcal{P}_{s,p}(\phi_{i,\{2,3\}}) = \text{Concat}(\mathcal{P}_{s,p}(\phi_{i,2}), \text{Upsample}(\mathcal{P}_{s,p}(\phi_{i,3}))) \quad (4)$$

Negative and Positive Memory Banks. Unlike traditional one-class methods, ExDD leverages the statistical properties of both normal and anomalous features through parallel memory banks. The Negative Memory Bank (\mathcal{M}_N) stores patch-level features from nominal samples:

$$\mathcal{M}_N = \bigcup_{x_i \in \mathcal{X}_N} \mathcal{P}_{s,p}(\phi_{\{2,3\}}(x_i)) \quad (5)$$

Complementarily, the Positive Memory Bank (\mathcal{M}_P) stores patch-level features from anomalous samples:

$$\mathcal{M}_P = \bigcup_{x_i \in \mathcal{X}_A} \mathcal{P}_{s,p}(\phi_{\{2,3\}}(x_i)) \quad (6)$$

The deliberate separation of memory banks preserves the distinct statistical properties of normal and anomalous feature distributions. Please note that the positive memory bank only accounts for those patches related to the defects. While this would require a pixel-level annotation of all the defects in the case of real images, it comes for free in the case of synthetic images, where the localization of defective patches can be automated by simply computing the difference between the original and generated images.

Dimensionality Reduction and Coreset Subsampling. The concatenated feature vectors have a high dimensionality of 1536 channels. We apply random projection based on the Johnson-Lindenstrauss lemma:

$$\psi : \mathbb{R}^d \rightarrow \mathbb{R}^{d^*} \quad (7)$$

where $d^* = 128 < d = 1536$. The projection matrix is constructed with elements drawn from a standard Gaussian distribution and normalized to unit length.

Even after dimensionality reduction, we employ greedy coreset subsampling:

$$\mathcal{M}_C^* = \arg \min_{\mathcal{M}_C \subset \mathcal{M}} \max_{m \in \mathcal{M}} \min_{n \in \mathcal{M}_C} \|m - n\|_2 \quad (8)$$

This objective ensures that the selected coreset provides optimal coverage of the feature space.

We apply asymmetric subsampling: 2% for the negative memory bank (higher redundancy) and 10% for the positive bank (preserve diverse anomaly representations).

3.3 Diffusion-Based Anomaly Generation

To address the limited availability of anomalous samples, we implement a diffusion-based data augmentation pipeline inspired by DIAG [11].

Theoretical Foundation. DIAG leverages Latent Diffusion Models (LDMs) to generate synthetic anomalies in a lower-dimensional latent space. The data distribution $q(x_0)$ is modeled through a latent variable model $p_\theta(x_0)$:

$$p_\theta(x_0) = \int p_\theta(x_{0:T}) dx_{1:T} \quad (9)$$

$$p_\theta(x_{0:T}) := p_\theta(x_T) \prod_{t=1}^T p_\theta^{(t)}(x_{t-1}|x_t) \quad (10)$$

The parameters θ are learned by maximizing an ELBO of the log evidence:

$$\max_{\theta} \mathbb{E}_{q(x_0)}[\log p_\theta(x_0)] \leq \max_{\theta} \mathbb{E}_{q(x_0, x_1, \dots, x_T)}[\log p_\theta(x_{0:T}) - \log q(x_{1:T}|x_0)] \quad (11)$$

where $q(x_{1:T}|x_0)$ is a fixed inference process defined as a Markov chain.

By conditioning on normal images and anomaly masks derived from real defects, our generative process creates samples within the true distribution of industrial defects. This ensures the positive memory bank captures genuine anomaly patterns rather than artifacts.

Synthetic Anomaly Generation Pipeline. To generate an anomalous image i_a , we start with a triplet (i_n, d_a, m_a) consisting of a nominal image $i_n \in \mathcal{X}_N$, a textual anomaly description d_a , and a binary mask m_a . We utilize Stable Diffusion XL’s inpainting capabilities with prompts like “copper metal scratches” and “white marks on the wall” based on KSDD2 dataset analysis. The pipeline uses inference steps=30, guidance scale=20.0, strength=0.99, and padding mask crop=2, resulting in a robust positive memory bank capturing diverse anomaly patterns.

3.4 Anomaly Detection with ExDD

The key innovation in ExDD’s detection mechanism is its ability to measure both dissimilarity from normality and similarity to anomaly patterns.

Distance Computation. For a test image x_{test} , we compute two complementary distance measures:

1. **Negative Distance** (s_N^*): The maximum minimum Euclidean distance from test patch features to the negative memory bank:

$$m_{\text{test},*}, m^* = \arg \max_{m_{\text{test}} \in \mathcal{P}(x_{\text{test},*})} \arg \min_{m \in \mathcal{M}_N} \|m_{\text{test}} - m\|_2 \quad (12)$$

$$s_N^* = \|m_{\text{test},*} - m^*\|_2 \quad (13)$$

2. **Positive Distance** (s_P^*): The maximum minimum Euclidean distance to the positive memory bank:

$$m_{\text{test},+}, m^+ = \arg \max_{m_{\text{test}} \in \mathcal{P}(x_{\text{test}})} \arg \min_{m \in \mathcal{M}_P} \|m_{\text{test}} - m\|_2 \quad (14)$$

$$s_P^* = \|m_{\text{test},+} - m^+\|_2 \quad (15)$$

Neighborhood-Aware Weighting. To account for the local neighborhood structure in feature space, we incorporate neighborhood-aware weighting based on local density estimation theory:

$$w_N^* = 1 - \frac{e^{-s_N^*/\sqrt{d}}}{\sum_{m \in \mathcal{N}_b(m^*)} e^{-\|m_{\text{test},*} - m\|_2/\sqrt{d}}} \quad (16)$$

This formulation increases the weighting factor when a test patch’s nearest neighbor in the normal memory bank is isolated from other normal features.

For the positive raw anomaly score, we invert the formulation:

$$w_P^* = \frac{e^{-s_P^*/\sqrt{d}}}{\sum_{m \in \mathcal{N}_b(m^+)} e^{-\|m_{\text{test},+} - m\|_2/\sqrt{d}}} \quad (17)$$

The weighted scores are computed as:

$$s_N = w_N^* \cdot s_N^* \quad (18)$$

$$s_P = w_P^* \cdot s_P^* \quad (19)$$

where $\mathcal{N}_b(m^*)$ and $\mathcal{N}_b(m^+)$ represent the b nearest neighbors to m^* and m^+ in their respective memory banks.

Ratio Scoring. We introduce a novel Ratio Scoring method that fuses information from both memory banks:

$$s_{\text{ratio}} = \frac{s_N}{s_P + \epsilon} \quad (20)$$

where ϵ is an arbitrary small constant value added to prevent division by zero.

This ratio amplifies the anomaly signal for regions both dissimilar from normal patterns (high s_N) and similar to known anomaly patterns (low s_P).

Anomaly Localization. For pixel-level anomaly segmentation, we extend our dual memory bank approach to generate spatial anomaly maps:

$$S_N(h, w) = \min_{m \in \mathcal{M}_N} \|\phi_{\text{test}, \{2,3\}}(\mathcal{N}_p^{(h,w)}) - m\|_2 \quad (21)$$

$$S_P(h, w) = \min_{m \in \mathcal{M}_P} \|\phi_{\text{test}, \{2,3\}}(\mathcal{N}_p^{(h,w)}) - m\|_2 \quad (22)$$

After applying neighborhood-aware weighting to obtain $S_N^w(h, w)$ and $S_P^w(h, w)$, we fuse these maps:

$$S_{\text{ratio}}(h, w) = \frac{S_N^w(h, w)}{S_P^w(h, w) + \epsilon} \quad (23)$$

The resulting map is upsampled to match the original image dimensions and smoothed with a Gaussian filter ($\sigma = 2$) to enhance visual clarity while preserving fine details.

4 Experiments

4.1 Dataset

We evaluate our method on the KSDD2 dataset [7], a real-world industrial benchmark for surface defect detection. The dataset contains 2,085 normal and 246 defective training images, along with 894 normal and 110 defective test images. Defects include scratches, spots, and material imperfections, ranging from 0.5 cm to 15 cm, captured under factory conditions. All images are resized to 224×632 pixels to standardize resolution while preserving defect morphology and spatial context.

4.2 Implementation Details

Experiments used an NVIDIA RTX 4090 GPU with a WideResNet50 backbone (ImageNet pretrained). We implemented 3×3 patches, feature hierarchies from ResNet levels 2–3, coreset subsampling (1% negative, 10% positive memory banks), and $k=3$ neighborhood weighting. Synthetic anomalies were generated using SDXL [14] via Diffusers [16], with text prompts “white marks on the wall” and “copper metal scratches” alongside a negative prompt “smooth, plain, black, dark, shadow” to suppress artifacts. Using KSDD2 ground-truth masks, defect-free training images were inpainted to create context-preserving synthetic anomalies, which were combined with the original training set. All models were implemented in PyTorch.

5 Results

We evaluated our ExDD framework on the KSDD2 dataset using both image-level and pixel-level metrics to comprehensively assess detection and localization capabilities. Table 1 presents a comparative analysis with state-of-the-art methods in industrial anomaly detection.

Our experiments reveal that ExDD outperforms most existing methods across both detection and localization tasks. While IRP achieves comparable image-level detection performance (94.0% vs. our 94.2%), it does not provide pixel-wise localization capabilities, which are crucial for practical industrial applications. Both OSR and IRP lack localization ability entirely, as indicated by the missing pixel-wise AUROC values. The ExDD base configuration already surpasses PatchCore by 1.9% in image-level AUROC and 1.1% in pixel-wise AUROC, demonstrating the effectiveness of our dual memory bank architecture. When implemented with the full configuration including synthetic data augmentation, ExDD achieves state-of-the-art performance across both metrics. The 1.1% improvement from base to full configuration highlights the value of our synthetic data approach in enhancing both detection and localization capabilities. Notably, our method substantially outperforms earlier approaches like DRAEM and DSR, which struggle particularly with pixel-wise localization (42.4% and 61.4% respectively, compared to our 97.7%).

Table 1. Anomaly Detection and Localization Performance on KSDD2 Dataset. ExDD (base) denotes our dual memory bank architecture without synthetic data, while ExDD (full) includes diffusion-based synthetic augmentation.

Metric	DRAEM [18]	DSR [19]	PatchCore [15]	OSR [3]	IRP [4]	ExDD (base)	ExDD (full)
I-AUROC (%)	77.8	87.2	91.2	92.1	94.0	93.1	94.2
P-AUROC (%)	42.4	61.4	95.8	–	–	96.9	97.7

5.1 Augmentation Analysis

To isolate and quantify the effect of our synthetic data augmentation strategy within the ExDD framework, we conducted experiments with varying numbers of synthetic samples, as shown in Table 2. With no synthetic samples, the ExDD base configuration relies solely on the limited real defective samples (246) available in the KSDD2 dataset alongside 2,085 normal samples. As synthetic samples are introduced, both detection and localization performance steadily improve. The addition of 100 synthetic samples (50 per text prompt) yields optimal performance across metrics. Interestingly, increasing the synthetic sample count to 150 provides no additional benefits and slightly reduces performance, suggesting a saturation point in the diversity of synthetic defect characteristics. These results validate our ExDD approach of integrating synthetic models for industrial

anomaly detection, demonstrating that carefully generated synthetic defects can effectively supplement limited real-world data while preserving the industrial context necessary for accurate detection and localization.

Table 2. Effect of varying the number of augmented samples on ExDD performance.

Augmented Images	0	50	100	150
I-AUROC (%)	93.1	93.4	94.2	93.5
P-AUROC (%)	96.9	97.1	97.7	97.2

5.2 Qualitative Analysis

The visualization in Fig. 2 demonstrates that ExDD effectively detects and localizes various types of defects in electrical commutators, including subtle scratches, surface anomalies, and material imperfections. When comparing the heatmaps generated by the standard PatchCore approach with those from the ExDD base and full configurations, we observe significantly better alignment with ground truth masks and a reduction in false positives in background regions. The full ExDD implementation produces anomaly maps with sharper boundaries and improved detection of subtle defect patterns, which aligns with the quantitative improvements observed in our experimental results.

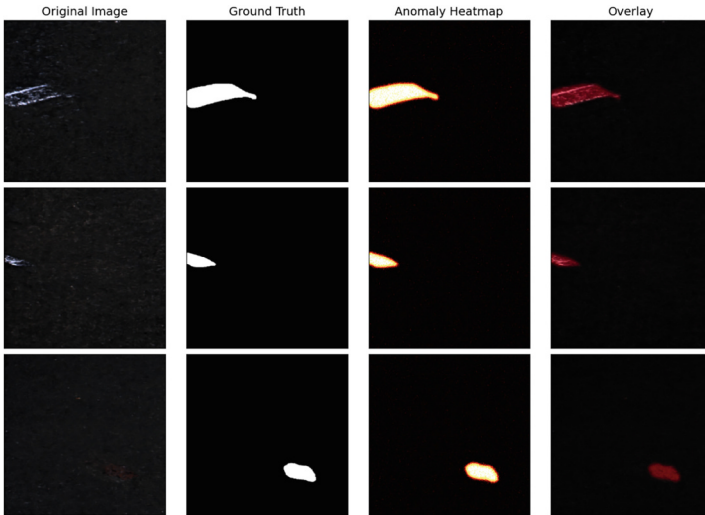


Fig. 2. Qualitative comparison of anomaly localization results on the KSDD2 test set.

6 Conclusion

ExDD represents a significant advancement in industrial anomaly detection by reconceptualizing defects as occupying structured feature distributions rather than arbitrary deviations. The integration of explicit dual-distribution modeling with diffusion-based synthetic defect generation creates a robust framework that leverages limited anomaly data effectively. The empirical performance ceiling observed at 100 synthetic samples suggests an optimal balance between augmentation diversity and potential distribution shift. This work establishes a foundation for future research in adaptive memory dynamics and uncertainty quantification for defect detection in data-constrained industrial environments, particularly for applications requiring precise boundary delineation and reduced false positives.

Acknowledgements. This study was carried out within the PNRR research activities of the consortium iNEST (Interconnected North-Est Innovation Ecosystem) funded by the European Union Next-GenerationEU (Piano Nazionale di Ripresa e Resilienza (PNRR) – Missione 4 Componente 2, Investimento 1.5 – D.D. 1058 23/06/2022, ECS_00000043).

References

1. Akcay, S., Atapour-Abarghouei, A., Breckon, T.P.: GANomaly: semi-supervised anomaly detection via adversarial training. In: Asian Conference on Computer Vision (ACCV) (2019)
2. Aqeel, M., Sharifi, S., Cristani, M., Setti, F.: Meta learning-driven iterative refinement for robust anomaly detection in industrial inspection. In: Del Bue, A., Canton, C., Pont-Tuset, J., Tommasi, T. (eds.) ECCV 2024. LNCS, vol. 15626, pp. 445–460. Springer, Cham (2024). https://doi.org/10.1007/978-3-031-92805-5_28
3. Aqeel, M., Sharifi, S., Cristani, M., Setti, F.: Self-supervised learning for robust surface defect detection. In: International Conference on Deep Learning Theory and Applications (DELTA) (2024)
4. Aqeel, M., Sharifi, S., Cristani, M., Setti, F.: Self-supervised iterative refinement for anomaly detection in industrial quality control. In: International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISAPP) (2025)
5. Aqeel, M., Sharifi, S., Cristani, M., Setti, F.: Towards real unsupervised anomaly detection via confident meta-learning. In: IEEE/CVF International Conference on Computer Vision (ICCV) (2025)
6. Bhatt, P.M., et al.: Image-based surface defect detection using deep learning: a review. *J. Comput. Inf. Sci. Eng.* **21**(4), 040801 (2021)
7. Božič, J., Tabernik, D., Skočaj, D.: Mixed supervision for surface-defect detection: From weakly to fully supervised learning. *Comput. Ind.* **129**, 103459 (2021)
8. Capogrosso, L., et al.: Diffusion-based image generation for in-distribution data augmentation in surface defect detection. In: International Joint Conference on Computer Vision, Imaging and Computer Graphics Theory and Applications (VISAPP) (2024)

9. Chen, Y., Ding, Y., Zhao, F., Zhang, E., Wu, Z., Shao, L.: Surface defect detection methods for industrial products: a review. *Appl. Sci.* **11**(16), 7657 (2021)
10. Defard, T., Setkov, A., Loesch, A., Audigier, R.: PADIM: a patch distribution modeling framework for anomaly detection and localization. In: *International Conference on Pattern Recognition (ICPR)* (2021)
11. Girella, F., Liu, Z., Fummi, F., Setti, F., Cristani, M., Capogrosso, L.: Leveraging latent diffusion models for training-free in-distribution data augmentation for surface defect detection. In: *International Conference on Content-based Multimedia Indexing (CBMI)* (2024)
12. Jain, S., Seth, G., Paruthi, A., Soni, U., Kumar, G.: Synthetic data augmentation for surface defect detection and classification using deep learning. *J. Intell. Manuf.* **33**, 1007–1020 (2022)
13. Jawahar, M., Anbarasi, L.J., Geetha, S.: Vision based leather defect detection: a survey. *Multimedia Tools Appl.* **82**(1), 989–1015 (2023)
14. Podell, D., et al.: SDXL: improving latent diffusion models for high-resolution image synthesis. *arXiv preprint [arXiv:2307.01952](https://arxiv.org/abs/2307.01952)* (2023)
15. Roth, K., Pemula, L., Zepeda, J., Schölkopf, B., Brox, T., Gehler, P.: Towards total recall in industrial anomaly detection. In: *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)* (2022)
16. Von Platen, P., et al.: Diffusers: state-of-the-art diffusion models (2022)
17. Vrochidou, E., et al.: Towards robotic marble resin application: crack detection on marble using deep learning. *Electronics* **11**(20) (2022)
18. Zavrtnik, V., Kristan, M., Skočaj, D.: Draem-a discriminatively trained reconstruction embedding for surface anomaly detection. In: *IEEE/CVF International Conference on Computer Vision (ICCV)* (2021)
19. Zavrtnik, V., Kristan, M., Skočaj, D.: DSR—a dual subspace re-projection network for surface anomaly detection. In: Avidan, S., Brostow, G., Cissé, M., Farinella, G.M., Hassner, T. (eds.) *ECCV 2022*. LNCS, vol. 13691. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-19821-2_31
20. Zhang, S., Zhang, Q., Gu, J., Su, L., Li, K., Pecht, M.: Visual inspection of steel surface defects based on domain adaptation and adaptive convolutional neural network. *Mech. Syst. Signal Process.* **153**, 107541 (2021)