






Meta Learning-Driven Iterative Refinement for Robust Anomaly Detection in Industrial Inspection

Muhammad Aqeel^(✉), Shakiba Sharifi, Marco Cristani,
and Francesco Setti

Department of Engineering for Innovation Medicine, University of Verona,
Strada le Grazie 15, Verona, Italy
muhammad.aqeel@univr.it
<https://www.dimi.univr.it/>

Abstract. This study investigates the performance of robust anomaly detection models in industrial inspection, focusing particularly on their ability to handle noisy data. We propose to leverage the adaptation ability of meta-learning approaches to identify and reject noisy training data to improve the learning process. In our model, we employ Model Agnostic Meta-Learning (MAML) and an iterative refinement process through an Inter-Quartile Range rejection scheme to enhance their adaptability and robustness. This approach significantly improves the models' capability to distinguish between normal and defective conditions. Our results of experiments conducted on well-known MVTec and KSDD2 datasets demonstrate that the proposed method not only excels in environments with substantial noise but can also contribute in case of a clear training set, isolating those samples that are relatively out of distribution, thus offering significant improvements over traditional models.

Keywords: Meta Learning · Iterative Refinement · Robust Anomaly Detection · Unsupervised Learning · Industrial Inspection

1 Introduction

Anomaly detection is pivotal in industrial settings, particularly for ensuring rigorous quality control and effective defect detection in manufacturing processes. Traditional methods, while foundational, often struggle with the localization and precise identification of anomalies within images, which is paramount for modern industrial contexts such as automated inspection systems.

Recent advancements in anomaly detection have increasingly utilized deep learning techniques that offer superior performance over traditional image-based methods. These new approaches are capable of detecting, localizing, and classifying anomalies with high precision, even at the pixel level. This shift is largely due to the development of convolutional neural networks (CNNs) and other deep

learning frameworks that have been specifically adapted for anomaly detection in image data, demonstrating remarkable success in identifying subtle and complex anomalies that were previously challenging to detect with conventional methods [5, 9].

The rarity and diversity of anomalies pose significant challenges as they require detection systems to operate with minimal human-labeled data. This scarcity makes the collection and annotation of a comprehensive dataset both expensive and time-consuming. To address these challenges, recent methods have employed unsupervised and semi-supervised learning techniques, where the models are trained predominantly on normal data and anomalies are identified as deviations from this learned norm [8].

Techniques such as autoencoders and generative adversarial networks (GANs) have been pivotal in this respect. Autoencoders, for instance, learn to reconstruct normal images and can highlight anomalies through reconstruction errors. Similarly, GANs have been adapted to generate normal data distributions, with anomalies identified by their failure to conform to these distributions [1, 19].

Furthermore, the integration of transfer learning and meta-learning into anomaly detection systems has shown promising results in reducing the need for extensive labeled datasets. These methods enable systems to adapt quickly to new types of anomalies or changes in data distribution, enhancing their flexibility and utility in dynamic industrial environments [12, 13].

In anomaly detection, the quality of training images plays a pivotal role in the effectiveness of the resulting model. In Fig. 1 visually demonstrates how data quality impacts model performance, comparing clean, corrupted, and iteratively refined training scenarios. Figures 1a and 1b contrast clean and corrupted training scenarios, highlighting the challenges introduced by noise and anomalies in training data. Subsequently, Fig. 1c introduces our method of iteratively refined training, designed to enhance data quality and robustness against such corruptions. The models trained on clean and corrupted data exhibit varying levels of accuracy and sensitivity, as shown in Figs. 1d and 1e, respectively. Our model, demonstrated in Fig. 1f, is tested under conditions reflecting the iterative training process, showcasing its superior capability to maintain high performance even in the presence of data imperfections. These visual comparisons underline the necessity and effectiveness of our proposed meta-learning driven iterative refinement approach in handling real-world industrial data complexities.

Addressing these challenges, our research introduces a robust, unsupervised anomaly detection framework that employs meta-learning with iterative refinement, uniquely tailored to address the complexity and variability inherent in industrial data settings, enhancing accuracy and reducing false alarms. This novel approach significantly improves the model's ability to adapt and recognize subtle, evolving anomaly patterns, crucial for detecting rare and nuanced anomalies. Our main contributions include:

- We introduce a meta-learning framework that dynamically adapts to evolving industrial data characteristics, effectively overcoming the limitations of traditional statistical and machine learning methods.
- Through iterative refinement, the model continuously updates and fine-tunes its detection algorithms, leveraging new anomaly data to refine its understanding and improve accuracy.
- We specifically enhance the framework’s adaptability to environments characterized by significant background noise and operational variability. This targeted adaptation focuses on maintaining high detection performance in noisy conditions, showcasing the framework’s practical effectiveness and reliability in challenging industrial settings.
- We report extensive results on public benchmarks to prove the effectiveness of our approach to isolate problematic samples at training time and train a model on the surviving ones.

This paper is structured as follows: a review of related work in industrial anomaly detection is presented in Sect. 2, we then detail the proposed methodology in Sect. 3, and we demonstrate the approach’s effectiveness in Sect. 4 and Sect. 5. Finally, in Sect. 6 we summarize our findings and discuss some future research directions.

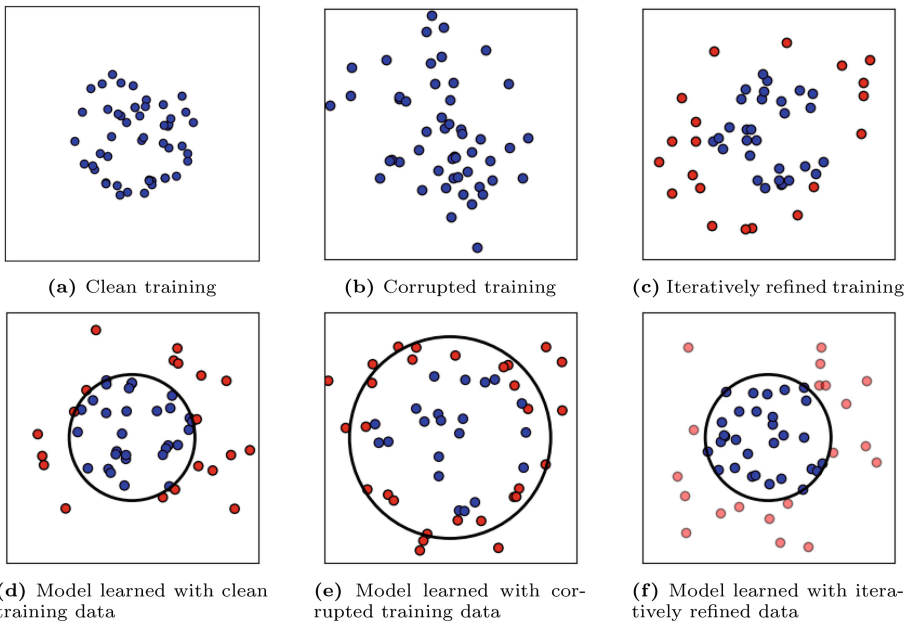


Fig. 1. Comparison of anomaly detection models under various training conditions: Clean, Corrupted, and Iteratively Refined. This sequence demonstrates the effectiveness of our metalearning-driven iterative refinement process in improving the robustness and accuracy of anomaly detection in industrial applications.

2 Related Work

This section reviews recent advancements in image-based anomaly detection, highlighting the transition from traditional statistical methods to sophisticated deep learning models that handle complex datasets. Modern approaches, specifically tailored for handling high-dimensional and complex image datasets, have revolutionized anomaly detection capabilities in industrial settings. This section reviews these developments, placing a strong emphasis on methods that excel in identifying subtle and intricate anomalies in visual data.

Convolutional Neural Networks (CNNs) have become the cornerstone of modern image-based anomaly detection due to their ability to learn hierarchical features from raw images automatically. CNNs are particularly effective in industrial inspection tasks, such as detecting defects in manufacturing processes, demonstrating superior performance over traditional methods [4]. However, their effectiveness can be limited by the quality and volume of training data. Generative Adversarial Networks (GANs) further enhance these capabilities by generating realistic images to identify anomalies as deviations that the generator network fails to reproduce accurately [19].

The integration of unsupervised learning techniques, such as Autoencoders and Isolation Forest, has broadened the scope of anomaly detection. Autoencoders excel by reconstructing normal data patterns and flagging significant reconstruction errors as anomalies, making them effective in noisy environments [18]. Isolation Forest uses random forests to isolate outliers, offering robust solutions where anomalies are subtle [14].

The use of semi-supervised methods like Deep SVDD, which learns a hypersphere enclosing normal data points with anomalies lying outside this boundary, represents a significant advancement in utilizing limited labeled data to enhance detection performance [17]. Similarly, self-supervised anomaly detection techniques that leverage anomalous training data via iterative latent token masking have shown promise in enhancing model adaptability and performance [15]. Informative knowledge distillation techniques are also being explored to improve deep anomaly detection models by preserving essential characteristics during model training, thus addressing some of the limitations inherent in traditional and older machine learning methods [7].

While newer machine learning methods have taken center stage, traditional techniques such as K-means clustering and Principal Component Analysis (PCA) continue to provide a foundational understanding, albeit with limitations in complex image-based scenarios. K-means and PCA often struggle with the high dimensionality typical of industrial image data, highlighting the need for more adaptable and robust solutions [11, 20].

Building on the foundational work reviewed here, our research integrates robust capabilities of deep learning with meta-learning and iterative refinement techniques, tackling challenges of noise and variability in image data while adapting dynamically to new anomaly patterns. By leveraging the strengths of both traditional and modern techniques, our framework sets a new benchmark for reliability and efficiency in industrial anomaly detection, ensuring high adaptability

and performance even in environments characterized by significant background noise and operational variability.

3 Method

Our methodology leverages the strengths of meta-learning, specifically Model Agnostic Meta Learning (MAML) [10], to enhance the robustness of unsupervised anomaly detection in environments laden with noisy data. Meta-learning is particularly suited for these environments as it prepares the model to efficiently handle diverse datasets and adapt dynamically to unseen data conditions without extensive retraining. MAML achieves this by initializing model parameters that are predisposed to adjust to high variability and noise, a common challenge in real-world datasets.

This approach is meticulously outlined in Algorithm 1, which details the steps for continuous adaptation and improvement. The algorithm, as outlined, systematically enhances detection capabilities by recalibrating model parameters in response to real-time data inputs and outcomes, thereby progressively enhancing the model’s sensitivity to subtle and complex anomalies. By leveraging meta-learning, we ensure that our system not only starts with a robust foundation but also continues to adapt and improve, maintaining high accuracy and reliability. Our experiments, which harness the challenging conditions of the KSDD2 and MVTEC AD dataset, demonstrate substantial improvements in detection performance over traditional methods, confirming the effectiveness of our meta-learning-enhanced iterative refinement process in practical, noisy scenarios.

Algorithm 1. Iterative Refinement for Anomaly Detection

- 1: Initialize model parameters θ with MAML
 - 2: **repeat**
 - 3: **for** each mini-batch B **do**
 - 4: Compute feature extraction $f_{\text{ex}}(x; \phi)$ for $x \in B$
 - 5: Transform features with f_{ML} to compute scores $S(x)$
 - 6: Compute IQR and adjust threshold T
 - 7: Classify data points in B based on threshold T
 - 8: Update model parameters θ using MAML
 - 9: **end for**
 - 10: Evaluate model on validation set
 - 11: **until** performance on validation set converges
-

3.1 Model Agnostic Meta Learning

We implement Model Agnostic Meta Learning (MAML) to optimize both the initialization and continuous adaptation of model parameters. This dual approach not only prepares the system for initial exposure to operational data but

also facilitates ongoing learning and adaptation based on evolving data patterns and operational feedback. This method is crucial in environments characterized by significant noise levels, where traditional learning algorithms might fail. The inner loop updates the model parameters on a small subset of data, while the outer loop optimizes these updates across multiple subsets, ensuring generalized performance. We implement Model-Agnostic Meta-Learning (MAML) to address the challenge of rapid adaptation under such noisy conditions. The core principle of MAML is to discover an optimal set of initial model parameters, θ , which can be efficiently fine-tuned through a minimal number of gradient updates to enhance performance on new tasks.

The implementation begins with the inner loop optimization, where the model parameters are adjusted specifically for each noisy dataset. This adjustment involves a gradient descent update defined as:

$$\theta' = \theta - \alpha \nabla_{\theta} L(S, \theta) \quad (1)$$

where α is the learning rate, and $L(S, \theta)$ represents the loss computed on a support set S that includes noise-distorted data. Here, θ' denotes the updated parameters after one gradient step, aimed at improving the model's performance specifically for the noise characteristics of S .

Following the inner loop, the adapted parameters θ' are evaluated on a corresponding query set Q to calculate the task-specific loss. This step is crucial as it assesses the effectiveness of the inner loop adaptations against new samples of noisy data, providing a feedback mechanism to refine the model adaptation. The task-specific loss is referred as $L(Q, \theta')$.

In the outer loop optimization, the initial model parameters θ are updated across multiple tasks. This stage aims to optimize a meta-objective that reduces the expected loss on new noisy datasets, thereby enhancing the model's generalizability and robustness to variations in noise. The meta-objective is mathematically expressed as:

$$\theta = \theta - \beta \nabla_{\theta} \mathbb{E}_{(S, Q) \sim p(T)} L(Q, \theta') \quad (2)$$

where β is the meta-learning rate, and $p(T)$ denotes the distribution of tasks. This equation highlights the process of updating the initial parameters θ based on the aggregate feedback from multiple tasks, ultimately steering θ towards values that offer better general performance on noisy data. These theoretical enhancements are best visualized through direct comparison of optimization landscapes before and after its application, as illustrated in Figs. 2 and 3.

To better visualize the process described, Fig. 2 provides a schematic overview of the MAML framework, illustrating both the inner and outer optimization loops. This graphical representation underscores how MAML effectively tailors the learning process to maximize performance in dynamic, noisy environments.

Figure 3 displays the loss landscapes before and after the application of MAML adaptations, showing the complex, rugged landscape prior to adaptation and a significantly smoothed landscape post-adaptation. This visual comparison

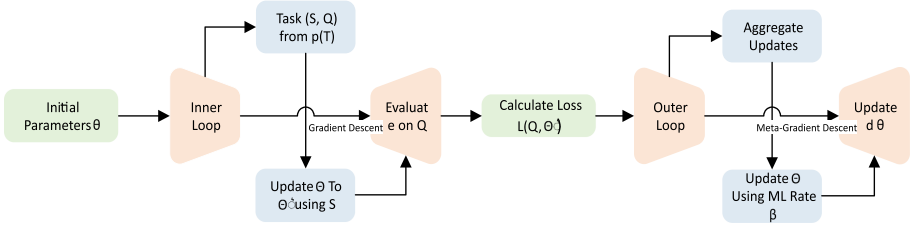


Fig. 2. A schematic representation of the Model-Agnostic Meta-Learning (MAML) process, illustrating the inner and outer loop optimizations and highlighting adaptive adjustments to the model parameters θ .

underscores the practical benefits of employing such sophisticated meta-learning techniques in high-noise environments.

Through this methodology, our implementation leverages MAML’s framework to minimize performance degradation in noisy conditions, ensuring rapid adaptation and high accuracy across various adversarial environments. This approach demonstrates how initial parameter settings can be fine-tuned to accommodate noise challenges, making the model more flexible and resilient.

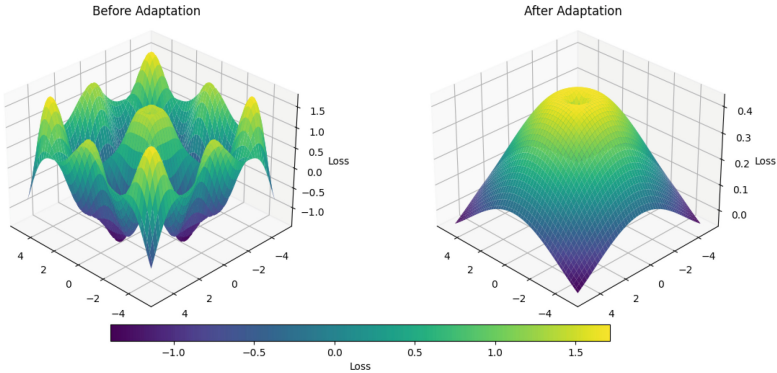


Fig. 3. Comparison of loss landscapes before and after MAML adaptations, clearly showing how MAML smooths the optimization landscape, facilitating more effective learning and generalization.

3.2 Iterative Refinement

Our system adopts an integrated approach to anomaly detection, which not only leverages advanced scoring mechanisms and outlier removal techniques but also incorporates a meta-learning framework that enhances model adaptability and

robustness. This iterative refinement process ensures precise anomaly identification and contributes significantly to the reliability and accuracy of our data analysis.

Refinement Loop: The iterative refinement process is outlined in Algorithm 1, which systematically recalibrates model parameters in response to real-time data inputs and outcomes. This process progressively enhances the model’s sensitivity to subtle and complex anomalies, ensuring that our system maintains high accuracy and reliability. Our experiments with the KSDD2 and MVTec AD datasets, known for their challenging conditions, demonstrate substantial improvements in detection performance over traditional methods. Inspired by the foundational model that uses a pre-trained feature extractor [16], our approach extends its adaptability, applying learned insights across multiple data scenarios without additional optimization of the feature extractor.

Score Computation: Initially, anomaly scores are computed using a hybrid metric, which integrates the feature distribution from the pretrained extractor and outputs from the meta-learning model. We define the score computation as follows:

$$S(x) = \mathbb{E}_{\phi}[-\log p_Z(f_{\text{ML}}(f_{\text{ex}}(x; \phi)))] \quad (3)$$

Here, p_Z represents the probability density function in the latent space Z , f_{ex} denotes the feature extraction from a pretrained model, f_{ML} is the transformation function from the meta-learning model, and ϕ are the parameters optimized by MAML. This scoring method captures deviations from a predefined distribution, essential for the preliminary identification of potential outliers.

Integration with Density Estimation: Our anomaly detection framework further integrates a scoring function based on density estimation from a normalizing flow transformation, aligning with traditional models. The scoring function for an image x subjected to transformations $S_i \in S$ is defined as:

$$\tau(x) = \mathbb{E}_{S_i \in S}[-\log p_Z(f_{\text{NF}}(f_{\text{ex}}(S_i(x))))] \quad (4)$$

This formulation ensures the model captures anomalies by evaluating the negative log-likelihoods across multiple transformations (e.g., rotations, brightness, and contrast adjustments), thus providing a robust mechanism to detect anomalies effectively.

Dynamic Thresholding: After initial scoring, we use the Interquartile Range (IQR) for further data refinement. The IQR helps in filtering out noise and less relevant features through dynamic thresholding. In the thresholding process, we employ a scaling factor k to adjust the sensitivity of our anomaly detection according to the specific characteristics and variability of the data observed. The

scaling factor k is adjusted to optimize the balance between false positives and false negatives, crucial for the effectiveness of anomaly detection systems. The upper threshold is formulated as follows:

$$T = Q_3 + k \times \text{IQR} \quad (5)$$

where Q_3 is the third quartile and $\text{IQR} = Q_3 - Q_1$ is the interquartile range. This dynamically calculated threshold adapts to changes in anomaly score distributions, enhancing the model’s capability to differentiate effectively between normal and anomalous data.

4 Experiments

This section outlines the experimental results achieved from applying our methodology to anomaly detection tasks in the presence of noisy training data. The experiments aimed to assess the enhancement in detection precision and stability across different industrial scenarios.

We began our experiments with the MVTec-AD dataset [4], a well-established benchmark featuring 15 distinct industrial product types. This dataset is particularly notable for its diversity, including more than 70 different defect types each meticulously identified with specific labels and segmentation masks.

Subsequently, we utilized the KSDD2 dataset [6], sourced from the Kolektor Group d.o.o. This dataset includes RGB images of defective production items, characterized by a broad range of defects varying in size, shape, and color—from minor scratches to severe surface damage.

For both datasets, we conducted experiments using proportionally sampled training and testing sets. The testing set consisted of a balanced distribution of nominal and anomalous samples. Each test was repeated multiple times to ensure statistical reliability, and average results were reported. To simulate noisy conditions inherent in unsupervised learning environments, the proportion of challenging data within the training set was incrementally adjusted. This adjustment was carried out at noise levels from 0% to 50%, progressively increasing the complexity of the training data to test the robustness of our detection algorithm.

We implemented our code using PyTorch framework. We resized images to 448×448 pixels and applied standardized preprocessing techniques, including optional rotations and normalization. The network architecture incorporates three scales of input resolution, includes 8 coupling blocks in the normalizing flow model, and features scale-translation networks with fully connected layers, each comprising 2048 neurons, without dropout. We set the training parameters with a learning rate of 2×10^{-4} and conducted training over 240 epochs with a batch size of 96.

We compared our proposed methodology with existing models including DifferNet [16], One Shot Removal (OSR) [3], and Iterative Refinement Process (IRP) [2]. DifferNet leverages a multi-scale feature extraction process to assign meaningful likelihoods to images, thereby facilitating defect localization. DifferNet represents also the model for feature extraction used in our procedure.

OSR enhances the robustness of surface defect detection models through a novel training pipeline, which involves initial training, anomalous sample removal, and model fine-tuning. The IRP strategy systematically refines the training dataset through iterative cycles, enhancing detection accuracy and model robustness, especially in noisy environments.

We tested our model’s capability to handle varying degrees of complexity in the training data, ranging from 0% to 50% noise levels. These levels represent the proportion of challenging samples included to simulate noisy conditions typically encountered in unsupervised learning scenarios. The model’s robustness was evaluated by measuring its performance using the AUROC score across these noise levels. The AUROC score provides a reliable indicator of the model’s ability to differentiate between nominal and anomalous behaviors under increasing noise conditions. Each experiment was repeated three times to confirm the results’ consistency and to comprehensively assess the impact of noise on detection accuracy.

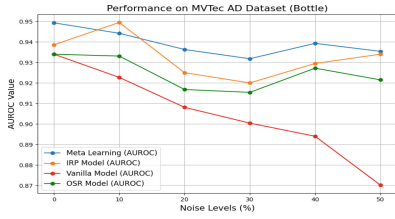
5 Results

This section presents the detailed evaluation results of four different anomaly detection models applied to the MVTec AD dataset across various classes, followed by an overall performance analysis that includes results from both the MVTec AD and KSDD2 datasets. The effectiveness of each model is assessed primarily using the AUROC metric, which measures their capability to differentiate between normal and defective samples under varying levels of noise.

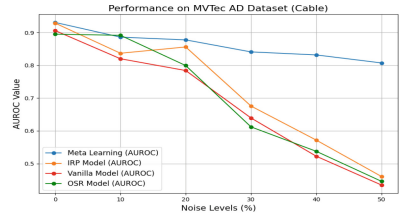
5.1 Performance in MVTec AD Classes

We conducted a comparative analysis of four distinct anomaly detection models: the Meta Learning Model, IPR Model, Vanilla Model, and OSR Model, across ten classes of the MVTec AD dataset: Bottle, Cable, Capsule, Carpet, Leather, MetalNut, Pill, Screw, Tile, and Zipper. The performance of these models was evaluated under different noise conditions to understand each model’s resilience to noise interference and its capability to handle diverse industrial defects. The comparative performance across these classes is demonstrated in Fig. 4 showcasing the AUROC results for each model under various noise scenarios.

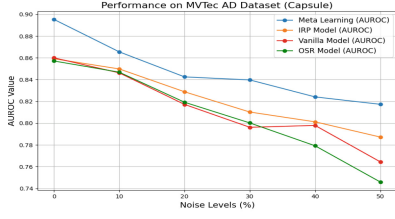
The Meta Learning Model exhibited exceptional performance across all classes, maintaining high AUROC values even in high noise environments, showcasing its robustness and reliability for industrial applications. Notably, in low noise settings, this model performed outstandingly well, highlighting its potential in precision-critical applications. In particular, we can appreciate how our model, proposed to handle noisy data, outperforms the baseline method also in case of no noisy data, *i.e.* noise level 0%. We hypothesize that this effect is due to the ability of our method to effectively isolate those samples that, despite being labeled as nominal, present some aspect typical of anomalous data, pushing the



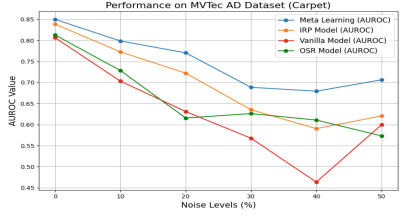
(a) Bottle



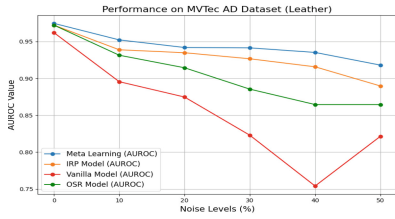
(b) Cable



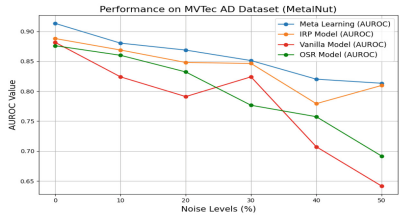
(c) Capsule



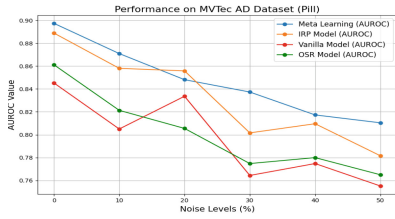
(d) Carpet



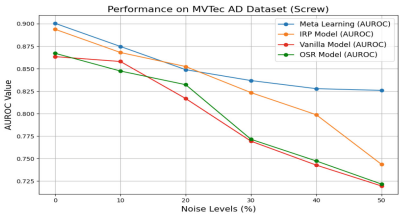
(e) Leather



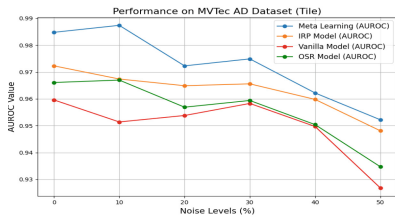
(f) MetalNut



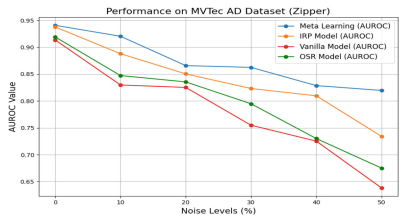
(g) Pill



(h) Screw



(i) Tile



(j) Zipper

Fig. 4. AUROC performance across various classes of the MVTec-AD dataset.

boundaries of the learned model towards the anomalous region (see Fig. 1) and thus generating more false negatives at testing time.

In contrast, the IPR Model showed strong initial performance at low noise levels but experienced a significant drop in effectiveness as noise increased. This trend was particularly marked in classes such as Cable and Capsule. The Vanilla Model, though stable initially, demonstrated a sharper decline in AUROC with increased noise levels, particularly struggling with complex defect patterns in the MetalNut and Zipper classes. The OSR Model, initially outperforming the Vanilla Model, also showed a marked decrease in performance with rising noise levels, indicating its lesser robustness to data variability. The detailed AUROC scores for each class at various noise levels are summarized in Table 1. This table provides a comprehensive statistical breakdown, including standard deviation metrics that offer further insights into the variability and reliability of each model under different testing conditions.

Table 1. Detailed AUROC scores for each model and class at varying noise levels.

Class	Noise Level (%)					
	0%	10%	20%	30%	40%	50%
Bottle	0.9493	0.9442	0.9363	0.9318	0.9393	0.9354
Cable	0.9309	0.8860	0.8776	0.8409	0.8317	0.8072
Capsule	0.8953	0.8655	0.8425	0.8397	0.8241	0.8173
Carpet	0.8499	0.7986	0.7705	0.6886	0.6794	0.7064
Leather	0.9746	0.9522	0.9419	0.9414	0.9352	0.9180
MetalNut	0.9133	0.8803	0.8687	0.8512	0.8201	0.8133
Pill	0.8973	0.8709	0.8481	0.8373	0.8173	0.8104
Screw	0.9004	0.8747	0.8489	0.8368	0.8278	0.8259
Tile	0.9848	0.9874	0.9723	0.9749	0.9622	0.9522
Zipper	0.9408	0.9205	0.8658	0.8625	0.8286	0.8196
MVTec-AD	0.9237	0.8980	0.8773	0.8605	0.8466	0.8406
KSDD2	0.9428	0.9340	0.9160	0.9182	0.9108	0.9027
Standard Deviation	± 0.0386	± 0.0512	± 0.0566	± 0.0750	± 0.0779	± 0.0779

5.2 Overall Performance Analysis

Figure 5a aggregates the findings from the MVTEC AD dataset, presenting the overall effectiveness of each model. The Meta Learning Model stands out with consistently high AUROC scores, reaffirming its robustness and adaptability in diverse conditions. This Fig. 5b illustrates the performance results on the KSDD2 dataset, further demonstrating the models' effectiveness across additional industrial anomalies.

The overall results on the MVTec-AD and KSDD2 datasets, as depicted in Fig. 5, confirm the Meta Learning Model as a superior choice for robust anomaly detection, suitable for industrial applications where both precision and adaptability are crucial. The examination of AUROC performance across different classes and noise levels highlights the importance of model selection in industrial anomaly detection. While each model exhibits unique strengths, the consistently high performance of the Meta Learning Model across various conditions and datasets recommends it as the optimal choice for practical implementations where reliability and accuracy are paramount.

This comprehensive evaluation benchmarks the capabilities of these models and provides a valuable framework for further refinement of anomaly detection technologies tailored to specific industrial requirements.

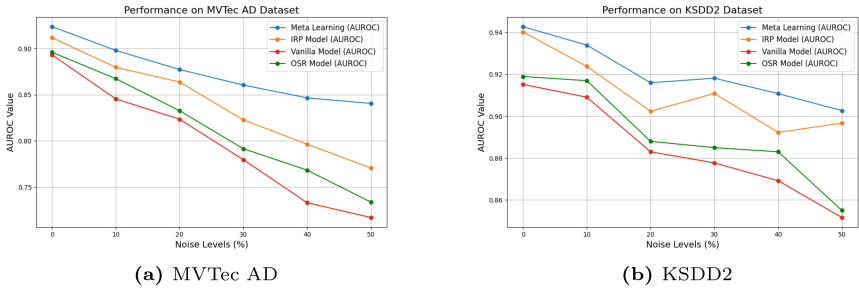


Fig. 5. Overall results on MVTec-AD and KSDD2 datasets.

5.3 Quantitative Analysis of Deleted Samples

In this analysis, we explore the number of samples removed during the iterative refinement process with the KSDD2 and MVTec-AD datasets, as presented in Table 2. The table details the count of “Good” and “Bad” images deleted at various noise levels (0% to 50%) across different classes such as Bottle, Cable, Capsule, Carpet, Leather, MetalNut, Pill, Screw, Tile, and Zipper. At a 0% noise level, the data shows that only “Good” images are removed. This pattern suggests that these non-defective samples might possess characteristics that cause them to be initially misidentified as defective. These images can be associated with samples with low likelihood in a Gaussian distribution. While this is not a problem in an infinite population, in case of the limited amount of data, they can easily lead to overestimating the standard deviation of the distribution thus producing classification boundaries that are prone to generate false negatives, *i.e.* anomalous samples not recognized as that.

As the noise increases, more “Bad” images are deleted across all classes, demonstrating the model’s effectiveness in identifying true defects under challenging conditions. It is worth noting that at 10% noise level, over 9 “Bad” images

are deleted on average, which means that only less than 10% of the anomalous images survive the filtering process. For example, the Carpet class starts with 24 good images being removed at 0% noise, with a progressive increase in defective image removal as noise rises, reaching 30 by 50% noise. This trend is consistent across other categories, indicating robust model performance in various noise environments. This data provides insights into the model’s capabilities to distinguish between defective and non-defective samples effectively, ensuring reliability and accuracy in practical applications.

Table 2. Number of images deleted at various noise levels for KSDD2 and MVTeC-AD datasets, segmented by product categories. The table shows the counts of deleted images categorized as “Good” and “Bad” for each noise level across different classes.

Class	0%		10%		20%		30%		40%		50%	
	Good	Bad	Good	Bad	Good	Bad	Good	Bad	Good	Bad	Good	Bad
Bottle	23	0	21	10	9	18	9	21	6	25	15	26
Cable	9	0	6	9	10	11	4	12	3	18	4	12
Capsule	14	0	10	8	11	13	8	19	8	23	7	25
Carpet	24	0	28	10	25	19	19	29	17	37	12	30
Leather	15	0	18	9	21	14	9	19	12	13	8	19
MetalNut	18	0	16	8	11	15	12	17	4	24	6	18
Pill	22	0	13	9	7	19	8	25	5	28	12	28
Screw	20	0	13	10	10	9	7	22	6	21	8	32
Tile	19	0	16	9	9	17	9	10	7	14	10	13
Zipper	21	0	17	10	14	14	9	22	8	20	5	19
KSDD2	25	0	22	10	13	16	12	18	13	21	16	26

6 Conclusion

In this paper, we presented a novel framework for anomaly detection that, leveraging the adaptation power of meta-learning and outlier rejection of statistical tools, can improve the performance of existing unsupervised anomaly detection methods by pruning out anomalous samples at training time and then fitting a model on the subset of remaining samples. Experiments on two popular image anomaly detection datasets demonstrate the substantial benefits of advanced meta-learning techniques for anomaly detection in industrial settings, outperforming state-of-the-art approaches across various noise levels and industrial contexts, as evidenced by its consistently high AUROC scores. This model not only proved robust in high-noise environments but also excelled in low-noise settings, highlighting its potential for precision-critical applications.

Moving forward, further research could explore the integration of additional adaptive algorithms to improve the granularity of anomaly detection. Moreover, expanding the dataset diversity and including real-time deployment scenarios may provide deeper insights into the operational effectiveness of these models, ensuring they meet the evolving demands of industrial applications.

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: *Computer Vision-ACCV 2018: 14th Asian Conference on Computer Vision, Perth, Australia, 2–6 December 2018, Revised Selected Papers, Part III 14*, pp. 622–637. Springer (2019)
2. Aqeel, M., Sharifi, S., Cristani, M., Setti, F.: Self-supervised iterative refinement for anomaly detection in industrial quality control. arXiv preprint [arXiv:2408.11561](https://arxiv.org/abs/2408.11561) (2024)
3. Aqeel, M., Sharifi, S., Cristani, M., Setti, F.: Self-supervised learning for robust surface defect detection. In: *Deep Learning Theory and Applications*, pp. 164–177. Springer (2024)
4. Bergmann, P., Fauser, M., Sattlegger, D., Steger, C.: Mytec ad—a comprehensive real-world dataset for unsupervised anomaly detection. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 9592–9600 (2019)
5. Bergmann, P., Fauser, M., Sattlegger, D., Steger, C.: Uninformed students: student-teacher anomaly detection with discriminative latent embeddings. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pp. 4183–4192 (2020)
6. Božič, J., Tabernik, D., Skočaj, D.: Mixed supervision for surface-defect detection: from weakly to fully supervised learning. *Comput. Ind.* (2021)
7. Cao, Y., Wan, Q., Shen, W., Gao, L.: Informative knowledge distillation for image anomaly segmentation. *Knowl.-Based Syst.* **248**, 108846 (2022)
8. Chandola, V., Banerjee, A., Kumar, V.: Anomaly detection: a survey. *ACM Comput. Surv. (CSUR)* **41**(3), 1–58 (2009)
9. Chow, J.K., Su, Z., Wu, J., Tan, P.S., Mao, X., Wang, Y.H.: Anomaly detection of defects on concrete structures with the convolutional autoencoder. *Adv. Eng. Inform.* **45**, 101105 (2020)
10. Finn, C., Abbeel, P., Levine, S.: Model-agnostic meta-learning for fast adaptation of deep networks. In: *International Conference on Machine Learning (ICML)*, pp. 1126–1135. PMLR (2017)
11. Hodge, V., Austin, J.: A survey of outlier detection methodologies. *Artif. Intell. Rev.* **22**, 85–126 (2004)
12. Hu, C., Wang, Y.: An efficient convolutional neural network model based on object-level attention mechanism for casting defect detection on radiography images. *IEEE Trans. Industr. Electron.* **67**(12), 10922–10930 (2020)

13. Kang, L., Liu, S., Zhang, H., Gong, D.: Person anomaly detection-based videos surveillance system in urban integrated pipe gallery. *Build. Res. Inf.* **49**(1), 55–68 (2021)
14. Liu, F.T., Ting, K.M., Zhou, Z.H.: Isolation forest. In: 2008 Eighth IEEE International Conference on Data Mining, pp. 413–422. IEEE (2008)
15. Patel, A., et al.: Self-supervised anomaly detection from anomalous training data via iterative latent token masking. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 2402–2410 (2023)
16. Rudolph, M., Wandt, B., Rosenhahn, B.: Same same but different: semi-supervised defect detection with normalizing flows. In: Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision, pp. 1907–1916 (2021)
17. Ruff, L., et al.: Deep one-class classification. In: International Conference on Machine Learning, pp. 4393–4402. PMLR (2018)
18. Sakurada, M., Yairi, T.: Anomaly detection using autoencoders with nonlinear dimensionality reduction. In: Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis, pp. 4–11 (2014)
19. Schlegl, T., Seeböck, P., Waldstein, S.M., Schmidt-Erfurth, U., Langs, G.: Unsupervised anomaly detection with generative adversarial networks to guide marker discovery. In: International Conference on Information Processing in Medical Imaging, pp. 146–157. Springer (2017)
20. Shyu, M.L., Chen, S.C., Sarinapakorn, K., Chang, L.: A novel anomaly detection scheme based on principal component classifier. In: Proceedings of the IEEE Foundations and New Directions of Data Mining Workshop, pp. 172–179. IEEE Press, Piscataway (2003)