Diffusion-Based Multi-class Defect Detection: A Generative Approach to Industrial Quality Control

Ahmed Mohammed
Machine Learning Institute
Johannes Kepler University Linz
A-4040 Linz, Austria
k12035954@students.jku.at

Abstract

This study presents an innovative approach to anomaly detection using diffusion models, focusing on multi-class defect classification in industrial quality control. The primary objective was to develop a robust system distinguishing defective and non-defective items across multiple defect categories. Our methodology involves fine-tuning a state-of-the-art stable diffusion model on an industrial dataset and optimizing it for defect detection and classification. The approach implements advanced techniques to estimate the likelihood of a sample being normal versus defective, drawing from recent developments in generative models. Our evaluation shows the performance of this diffusion-based anomaly detector and involves a two-stage approach: first, training the model to generate images across all defect classes and the non-defect class; second, during testing, utilizing the model to analyze and classify input images. Results indicate superior performance in detecting subtle and complex defects compared to conventional computer vision techniques. The significance of this work lies in its potential to enhance quality control processes in manufacturing, particularly in scenarios where defects are diverse and challenging to categorize. This project contributes to industrial inspection by demonstrating the effectiveness of generative models in multi-class anomaly detection, paving the way for more accurate and adaptable quality control systems.

Quality control is a crucial aspect of manufacturing, especially in industries where products must meet stringent standards. Automated defect detection, a key component of quality control, has traditionally relied on rule-based methods or discriminative computer vision models to identify defects in manufactured items. However, these approaches often struggle when dealing with complex or subtle defects, and are limited in their ability to generalize to new types of defects not explicitly represented in the training data. As a result, there is a growing need for more adaptable and robust methods in quality control to handle the diversity and complexity of defects that arise in industrial settings.

Recent advancements in generative models, particularly diffusion models, have opened new avenues for image synthesis and anomaly detection. Diffusion models, such as the denoising diffusion probabilistic models (DDPMs) (1) and stable diffusion models (2), are powerful tools for generating high-fidelity images that capture intricate details. These models have demonstrated exceptional performance in various tasks, from inpainting to zero-shot classification. Notably, diffusion models can be conditioned on prompts or class labels to generate images that align closely with specified characteristics. This conditional generative capacity offers a promising approach for defect detection, as the model can be trained to generate images of normal and defective items across multiple defect types, enabling it to classify anomalies in a multi-class setting.

In this study, we initially aimed to present a diffusion-based framework for defect detection in industrial quality control, leveraging the generative capabilities of stable diffusion models. Our approach was designed to address the complexity of detecting diverse types of defects across different regions of a product through a two-stage process. First, we intended to train the diffusion model to generate images representing various defect classes and the non-defective class; second, we planned to use the model's ability to predict defect likelihoods to classify new, unseen items. However, despite our efforts, we were unable to obtain satisfactory results in the first stage, which hindered our progress to the second stage.

Although we were unable to achieve our initial goals, this study highlights the challenges and limitations of applying diffusion models to industrial defect detection. Our experience underscores the need for further research into data preprocessing, model architecture, and training strategies to overcome the hurdles we encountered. As a future direction, we plan to explore alternative solutions, including different data preprocessing techniques and diffusion models, to improve the performance of our approach and ultimately achieve our goal of developing a robust and adaptable defect detection system.

1 Dataset

The dataset used for this project consists of ink-jet printed images of a single template split into three sections as shown in Figure 1. The sections are structured as follows:

- Section One: Contains two features named "distance one" and "edge roughness one"
- Section Two: Contains two features named "edge roughness two" and "edge roughness three"
- Section Three: Contains four features named "edge roughness four", "distance six", "dots", and "angle"

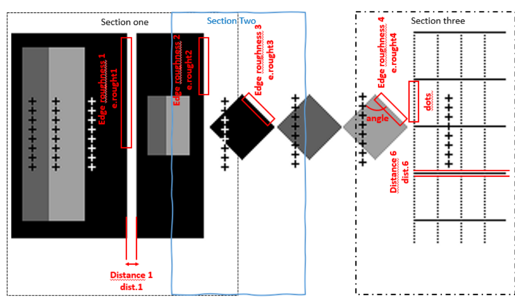

Figure 1: Template layout of the images in the dataset

1.1 Prompt Generation

We developed three versions of prompts with increasing levels of detail and specificity. Here are examples for each version, showing the same feature (distance between parallel lines, type one) in both good and defective conditions:

1.1.1 Version 0 Prompts

These prompts provide a basic description of the feature with or without defects.

Examples for `dist.1`:

- Good: “An inkjet-printed image showing optimal quality in the distance between parallel lines (type one), with no visible defects or irregularities.”
- Defective: “An inkjet-printed image displaying a noticeable defect or irregularity in the distance between parallel lines (type one).”

1.1.2 Version 1 Prompts

These prompts provide more detail about the expected quality and defects.

Examples for `dist.1`:

- Good: "An inkjet-printed image featuring two parallel edges with a consistent gap, representing distance number one. The gap is precisely measured and uniform, emphasizing perfect symmetry in the design."
- Defective: "An inkjet-printed image with a visible defect in two parallel edges with an inconsistent gap, representing distance number one. The gap varies along the length, compromising the symmetry of the design."

1.1.3 Version 2 Prompts

These prompts are the most detailed and provide specific descriptions of the features and defects. Examples for `dist.1`:

- Good: "The horizontal distance between two parallel edges. The gap between these edges is consistently 125.04 pixels across the entire length, indicating a well-defined and defect-free separation."
- Defective: "The horizontal distance between two parallel edges is intended to be 125.04 pixels but displays variation along the length. Gaps range less than 125 pixels, indicating a defect in the spacing."

2 Methodology

2.1 Problem Formulation

We begin with a dataset $D = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$ where each image belongs to one of K classes, classes in our case are the features that we identified previously. $y_i \in \{1, \ldots, K\}$. Our goal is to classify an image by predicting the most probable class assignment assuming a uniform prior over classes:

$$\tilde{y} = \arg\max_{y_k} p(y = y_k|x) = \arg\max_{y_k} p(x|y = y_k) \cdot p(y = y_k) = \arg\max_{y_k} \log p(x|y = y_k) \quad (1)$$

Generative classifiers, as introduced by (3), employ a conditional generative model to approximate the probability distribution $p_\theta(x|y = y_k)$, where θ represents the model's parameters.

2.2 Diffusion Models

The denoising diffusion probabilistic models (DDPMs) (1) introduce a gradual noise addition process. For each timestep t, given an image $\mathbf{x}_{t-1}$, we can generate a slightly noisier version $\mathbf{x}_t$ using:

$$q(\mathbf{x}_t|\mathbf{x}_{t-1}) = \mathcal{N}(\mathbf{x}_t; \sqrt{1-\beta_t}\mathbf{x}_{t-1}, \beta_t\mathbf{I}) \quad (2)$$

where β_t determines the amount of noise added at each step. The full forward process is defined as:

$$q(\mathbf{x}_{1:T}|\mathbf{x}_0) = \prod_{t=1}^{T} q(\mathbf{x}_t|\mathbf{x}_{t-1}) \quad (3)$$

To efficiently compute $\mathbf{x}_t$ for any timestep t given the initial image $\mathbf{x}_0$, we use:

$$q(\mathbf{x}_t|\mathbf{x}_0) = \mathcal{N}(\mathbf{x}_t; \sqrt{\bar{\alpha}_t}\mathbf{x}_0, (1-\bar{\alpha}_t)\mathbf{I}) \quad (4)$$

where $\bar{\alpha}_t = \prod_{i=1}^{T} \alpha_i$ and $\alpha_i = 1 - \beta_i$.

In our implementation, we condition these models on text prompts instead of class labels. We generate these prompts by modifying each label y_k using the method discussed in Section 1.1. These prompts are then used to condition the U-Net during training.

2.3 Diffusion Generative Classifier

Following (3), we generate classification decisions by approximating the conditional log-likelihood $\log p_\theta(x|y = y_k)$ using the diffusion variational lower bound (ELBO).

$$\tilde{y} = \arg\max_{y_k} \log p_\theta(x|y = y_k) \approx \arg\min_{y_k} \mathbb{E}_{\epsilon,t} \left[w_t \|x - \tilde{x}_\theta(x_t, y_k, t)\|_2^2 \right] \quad (5)$$

For Stable Diffusion (SD)(2), x is a latent representation obtained by encoding the image using a VAE. We evaluate the conditional likelihood $p_\theta(x|y = y_k)$ for each class $y_k \in [y_K]$ and assign the class with the highest likelihood obtained Equation1.

This approach, Diffusion Classifier, is theoretically motivated through the variational view of diffusion models and uses the ELBO to approximate $\log p_\theta(x|y_k)$. The method chooses the conditioning y_k that best predicts the noise added to the input image, making it effective for zero-shot classification without additional training.

Our results show that this generative approach attains strong performance on various benchmarks and outperforms alternative methods of extracting knowledge from diffusion models. The Diffusion Classifier also exhibits better multi-modal compositional reasoning abilities and effective robustness to distribution shift compared to discriminative approaches.

3 Results and Discussion

In this section, we examine the performance of our diffusion model in generating images that accurately represent the characteristics of the original dataset. We carefully evaluate the model's ability to generate images of three sections and features, including both normal and defective instances. This analysis highlights both the strengths and limitations of the stable diffusion model.

3.1 Comparative Analysis of Generated and Original Data

To rigorously assess our model's proficiency in recreating key features from the original dataset, we conducted a detailed comparison between generated images and their original counterparts across three distinct sections of the target object.

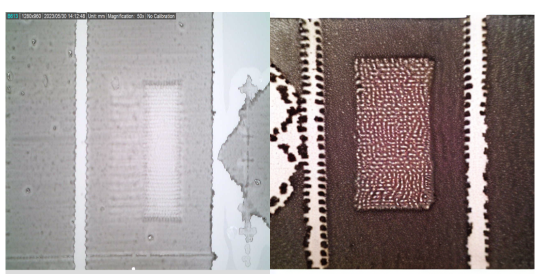

Figure 2: Comparison of an original (left) and generated (right) image from section one

Figure 2 shows the model's ability to create features of section one. This section includes two main features: distance one and edge roughness one. In the original image, distance one is defective while edge roughness one is normal. In the generated image, however, distance one is normal, and edge roughness one shows a defect. This difference shows that the model captures the main characteristics of defects, but sometimes generates variations.

Figure 3 illustrates section two, comparing an original image from the training data with the generated version created by our Stable Diffusion Model (SDM). The original image has a defect in edge roughness two, but edge roughness three is normal. In the generated image, edge roughness two is normal, but edge roughness three is missing. This difference shows that while the model can generate complex features, it sometimes misses details.

Figure 4 compares section three, which includes four different features. In the original image, defects appear in edge roughness four and distance number six, while the dots and angle features are normal.

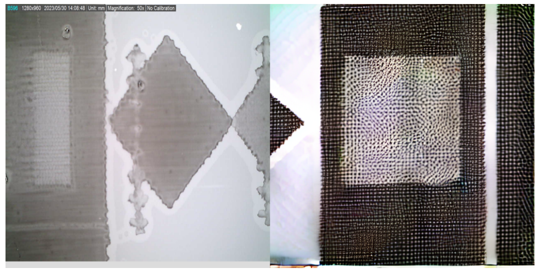

Figure 3: Comparison of an original (left) and generated (right) image from section two

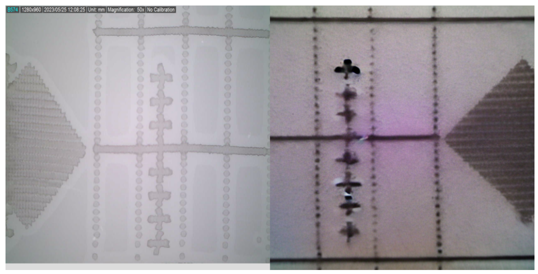

Figure 4: Comparison of an original (left) and generated (right) image from section three

The generated image, however, shows no defects except in the angle feature. This suggests that the model finds it challenging to replicate multiple defects exactly.

3.2 Comparing Generated Images With and Without Defects

To evaluate the model's ability to distinguish between normal and defective cases, we provide examples of generated images with and without defects for different features.

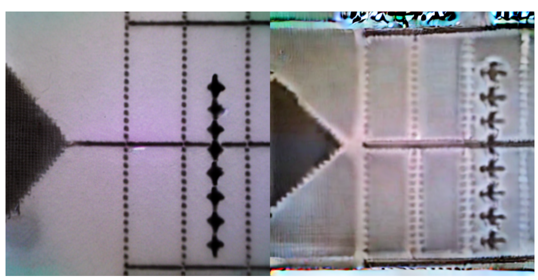

Figure 5: Generated images of feature dots: normal (left) and with defects (right)

Figure 5 shows the model's skill in creating the 'dots' feature both with and without defects. The normal dots are clear, and the defective version accurately shows the defect, which highlights the model's ability to capture small differences in this feature.

Figure 6 demonstrates the model's ability to generate images with different edge roughness levels (two and three). The normal image shows both edge roughness features clearly. In the defective image, however, edge roughness two is missing, which reveals a limitation in generating all defective features reliably.

Figure 7 illustrates the model's performance in generating images of distance one and edge roughness one. The normal image shows no defects, while the defective image accurately represents a defect in distance one, while edge roughness remains normal. This indicates that the model understands how to apply specific defects selectively.

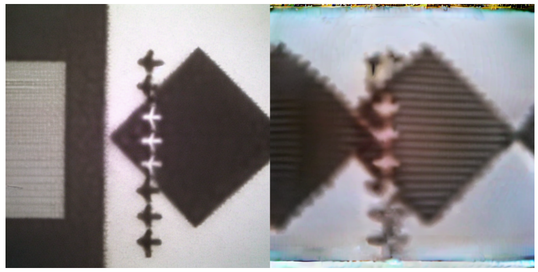

Figure 6: Generated images of edge roughness: normal (left) and with defects (right)

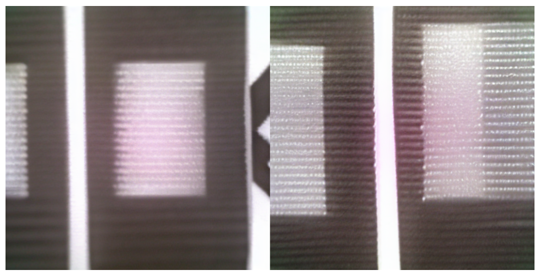

Figure 7: Generated images of feature distance: normal (left) and with defects (right)

4 Conclusion

In this study, we set out to develop a diffusion-based framework for defect detection in industrial quality control. Our initial goal was to leverage the generative capabilities of stable diffusion models to address the complexity of detecting diverse types of defects across different regions of a product. However, we encountered significant challenges in obtaining satisfactory results in the first stage of our proposed two-stage approach.

Despite our efforts, we were unable to achieve the level of performance we had initially anticipated. The difficulties we faced in obtaining good models in the first stage prevented us from progressing to the second stage of our planned classification process. This outcome highlights the complexities involved in applying diffusion models to industrial defect detection and underscores the need for further research and development in this area.

While we were unable to fully realize our initial objectives, this study has provided valuable insights into the challenges of implementing diffusion models for quality control applications. The limitations we encountered suggest that there may be fundamental issues related to data preprocessing, model architecture, or training strategies that need to be addressed to make these models more effective for industrial defect detection.

Looking ahead, our future work will focus on overcoming these challenges. We plan to explore different solutions, including:

1. Significantly increasing the size of our dataset to provide more diverse examples for training and potentially improve model performance.
2. Investigating alternative data preprocessing techniques to enhance the quality and relevance of our training data.
3. Exploring other diffusion model architectures that may be better suited to our specific use case.
4. Developing new strategies to improve the model's ability to learn and reproduce complex defect patterns.

Our ultimate goal remains to develop a robust and adaptable defect detection system that can effectively pass the first stage of our proposed approach and move on to the classification stage. We believe that by addressing the limitations identified in this study, we can make significant progress towards realizing the potential of diffusion models in industrial quality control.

This research, despite its challenges, contributes to the ongoing dialogue about the application of advanced AI techniques in manufacturing. It highlights both the promise and the current limitations of using generative models for defect detection, paving the way for future innovations in this critical area of industrial automation.

References

[1] J. Ho, A. Jain, and P. Abbeel, "Denoising diffusion probabilistic models," *Advances in neural information processing systems*, vol. 33, pp. 6840–6851, 2020.

[2] R. Rombach, A. Blattmann, D. Lorenz, P. Esser, and B. Ommer, "High-resolution image synthesis with latent diffusion models," in *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, 2022, pp. 10 684–10 695.

[3] A. Ng and M. Jordan, "On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes," *Advances in neural information processing systems*, vol. 14, 2001.