# Unveiling Hidden Anomalies
## A Hybrid Approach for Surface Mounted Electronics

Ghezeljehmeidan, Amir Ghorbani; van Driel, Willem Dirk; Dauwels, Justin

**DOI**
10.1109/INDIN58382.2024.10774393

**Publication date**
2024

**Document Version**
Final published version

**Published in**
Proceedings - 2024 IEEE 22nd International Conference on Industrial Informatics, INDIN 2024

# Unveiling Hidden Anomalies: A Hybrid Approach for Surface Mounted Electronics

1st Amir Ghorbani Ghezeljehmeidan
*Department of Microelectronics*
*TU Delft, The Nederlands*
amir.ghorbani@tudelft.nl

2nd Willem Dirk van Driel
*Department of Microelectronics*
*TU Delft, The Nederlands*
willem.van.driel@signify.com

3rd Justin Dauwels
*Department of Microelectronics*
*TU Delft, The Nederlands*
j.h.g.dauwels@tudelft.nl

2024 IEEE 22nd International Conference on Industrial Informatics (INDIN) | 979-8-3315-2747-1/24/$31.00 ©2024 IEEE | DOI: 10.1109/INDIN58382.2024.10774393

***Abstract*—Industrial assembly lines are the heartbeat of modern manufacturing, where precision and efficiency are paramount. This paper introduces a novel hybrid Explainable artificial intelligence (XAI) approach to enhance monitoring and analysis in industrial assembly. By fusing the power of vision anomaly detection models with the clarity of the gradient tree boosting algorithm, this framework not only boosts defect detection accuracy but also provides transparent, actionable insights. This synergy transforms how operators and engineers interact with AI, fostering trust and enhancing operational excellence.**

***Index Terms*—Vision Anomaly Detection, Gradient Tree Boosting, Explainable Artificial Intelligence.**

## I. Introduction

In the Fourth Industrial Revolution, there is a change in the understanding of quality and Quality Management (QM), emphasizing individualized service reliability, design and safety [1]. As entering the Industry 5.0, characterized by even greater automation and interconnected systems, human-machine collaboration takes center stage. While AI excels at data analysis and pattern recognition, achieving true process optimization requires explainable AI. By understanding the rationale behind model's recommendations, human experts can maintain oversight, ensure alignment with quality objectives, and continuously improve the AI models for robust anomaly detection in complex manufacturing environments [2].

### *A. The Pulse Of Modern Manufacturing*

Consumers' relentless demand for exceptional product quality and reliability is driving a clear and steady trend towards smart industry. The electronics sector exemplifies this perfectly. The manufacturing process begins with receiving printed circuit boards (PCBs) and necessitates rigorous quality control procedures throughout. These PCBs then undergo intricate processes like punching and assembly, often facilitated by sophisticated surface mount technology.

Ensuring consistent quality in smart manufacturing environments necessitates a robust and scalable solution for defect detection. Visual inspection machines are crucial for assessing optical attributes and ensuring product quality in many industrial assembly lines. However traditional visual quality inspection faces limitations in high-volume manufacturing due to human limitations, component characteristics, production volume, and potential flaw variety [3]. Automatic Optical Inspection (AOI) emerges as a non-destructive solution to these limitations, particularly for anomaly detection in high-volume manufacturing. However, as AOI systems become more complex with Deep Learning (DL) algorithms, ensuring interpretability through explainable AI becomes critical [4].

This work is supported by the project EXPLAIN, the Netherlands Enterprise Agency RVO under grant AI212001.

### *B. Current Monitoring Paradigms And Their Pitfalls*

The improvements in AI and DL algorithms have led to a noticeable change in focus towards AOI. The advent of AOI as a non-destructive quality assessment method is crucial in surmounting the constraints presented by conventional inspection methods. This robust technique proves particularity advantageous by alleviating the complexities associated with manually examining products, which often depend on human operators relying solely on their unaided vision, that are prone to errors and time consuming. A fully automated optical inspection system comprises both hardware and software parts. The hardware configuration, encompassing image sensor and illumination settings, is responsible for capturing digital images. Simultaneously, the software component executes an inspection algorithm to extract image features and categorize them according to user-defined criteria. Nevertheless, with advances in manufacturing technologies and increasing complexity, utilizing user-defined criteria leads to an increase in false alarms [5]. In response to the evolving demands of quality assessment, the utilization of AI has become instrumental in automating inspection and evaluation processes. These technological strides not only address current challenges but also align with the dynamic nature of production capabilities, reflecting a growing imperative for enhanced efficiency. The integration of these cutting-edge capabilities seamlessly complements the established AOI methodologies, marking a paradigm shift in the landscape of quality control across various industries [6].

### *C. The Promise Of Explainable AI*

While advancements in DL models yield significant progress in anomaly detection for various industrial applications, a critical challenge remains: ensuring interpretability of the results. Traditional anomaly detection methods often function as black boxes, providing anomaly scores without offering clear distinctions between normal and abnormal samples, the reason behind the particular decision, or further information. This lack of transparency hinders user trust and limits the ability to refine the models for optimal performance. Explainable AI emerges as a transformative approach, combining high performance with interpretability [7]. By leveraging XAI techniques, we can gain a richer understanding of the data used in anomaly detection, as it is implemented in real-world applications for healthcare [8], [9] and finance [10]. This multifaceted approach empowers human experts to not only identify anomalies but also grasp the underlying rationale behind the model's decisions, faster issue resolution, and ultimately leading to more informed decision-making and continuous improvement of the industrial anomaly detection system [11], [12].

### *D. Hybrid Solution For The Future*

The field of anomaly detection is rapidly evolving. Cutting-edge vision models excel at capturing anomalies in 2D images with high precision, even in zero/few-shot scenarios where training data for specific defects might be limited. However, these models can sometimes miss anomalies not readily visible on the surface, or explain the true/complete reason behind the decision. Here's where the power of more interpretable models, like decision trees, comes in. By combining the precision of vision models with the rich insights gleaned from well-structured data, we can create a more comprehensive approach. The hybrid model leverages the strengths of both – the exceptional accuracy of vision for surface-level anomalies and the explanatory power of tabular data for deeper understanding of potential root causes. This synergy not only pinpoints anomalies but also helps us understand "why" and "where" they occur, leading to more effective preventive measures and targeted interventions.

## II. Methodology

Cutting-edge deep learning models excel at capturing anomalies in 2D images used in vision inspection. While traditionally supervised learning approaches dominated the filed, by leveraging the large dataset of labelled anomalies. Supervised models can be trained to effectively identify specific types of defects with high precision [6], [13], [14]. However, the challenge lies in acquiring sufficient labeled data, especially for rare anomalies. Techniques like data augmentation and transfer learning can be employed to mitigate this challenge to some extent. However, their reliance on vast amounts of labeled data and potential for missing subtle defects necessitates a more comprehensive approach.

Recent advancements in unsupervised deep learning offer promising alternatives [15], [16]. Most of these techniques let to learn the underlying distribution of normal data. Deviations from this learned distribution can then be flagged as potential anomalies.

### *A. Building A Rich Dataset*

In the SMT-PCB sector, data acquisition and labeling present a significant challenge for deploying DL models. This is due to 2 major factors: firstly, the low frequency of faults, and secondly, the vast variety of potential defect types. Consequently, DL models for SMT-PCB fault detection should be adept at handling limited data, particularly rare or anomalous examples, while still delivering comprehensive and informative outputs. Due to the inherent challenges of data acquisition in the SMT-PCB sector, a dummy board have been manufactured that includes the most common defect types in surface mounted electronics. Additionally, we adopted a training strategy utilizing a limited dataset of normal images. This approach leverages a technique that enables feature augmentation solely on normal data.

### *B. Deep Learning With Vision Transformer*

In computer vision, convolutional neural networks (CNNs) have long been the dominant architecture due to their effectiveness in tasks like image classification. However, the recent success of vision transformers (ViTs) in natural language processing (NLP), achieving state-of-the-art results, has motivated their exploration in computer vision. Building upon the successes of computer vision, industrial vision tailors these algorithms for real-time anomaly detection and process control within industrial environments. Representation-based methods have emerged the leading approach for industrial anomaly detection and localization. These methods compress normal image features into embedding space, where anomalous features deviate from the normal data. Typically these method utilize CNN-baed networks that are pre-trained on ImageNet. Nevertheless images in industrial domain are significantly different from the images found in ImageNet. This mismatch prevents the direct use of extracted features [17].

CNNs are heavily influenced by notion of locality, where features are extracted based on the spatial relationships between neighbouring pixels. This prioritize local information processing. Additionally CNNs exhibit translation equivalence, meaning the network response remains consistent under image shift. This is advantageous where localization is irrelevant. However ViTs rely on self-attention mechanisms that enable global information processing, which is useful for object localization tasks that are crucial in industrial anomaly detection [18].

### *C. Harnessing Decision Tree For Clarity*

Decision trees are a class of machine learning models renowned for their high degree of explainability. This interpretability stems from their structure [19], where a series of simple if questions (based on features) guide the classification process. Additionally, decision trees offer the advantage of being computationally efficient, requiring fewer resources compared to more complex models to reach a prediction specially on certain data types such as tabular data [20].

While decision trees offer the benefit of interpretability and efficiency, their susceptibility to instability can be a concern. This is where Gradient Boosting Decision Trees (GBDTs) appear to deliver a more stable model while still retaining some level of interpretability. XGBoost is a powerful implementation of GBDT that addresses the instability issue of decision trees. It combines the boosting method with decision trees. In essence, XGBoost builds an ensemble of weak decision trees sequentially. Each new tree focuses on correcting the residuals of the previous ones, leading to a more robust and accurate final model. To further enhance performance and avoid overfitting, XGBoost incorporates several key features such as Regularization, a penalty term added to the loss function, discouraging overly complex trees that might overfit the data and Second-Order Taylor Expansion which allows more precise adjustments during the boosting process, leading to improved accuracy and robustness [21].

## III. Experiment

### A. Comprehensive Experiment Description

The data for this study was collected using an industrial AOI inspection machine (MEK ISO-Spector M2) commonly employed in the final stage of the manufacturing process line for electronics inspection. This machine is equipped with top camera positioned to capture images from top-bottom perspective. Additionally, the system offers the flexibility to utilize diverse lighting conditions and positions to provide extra information for the inspected boards. A crucial challenge in industrial anomaly detection is lighting variations that can obscure defects. To address this, we leverage the machine's capability to capture images under different lighting conditions. This approach aims to mitigate the impact of lighting inconsistencies and improve the overall robustness of the anomaly detection system.

Since real-world datasets with a significant number of defective samples are often scarce, a data augmentation strategy is employed. A dedicated dummy board was intentionally fabricated to include a variety of representative defects. This approach helped us enrich the dataset with a broader range of anomaly examples, enhancing the model's ability to generalize and detect anomalies effectively.

*1) Definition of Hybrid framework:* In this work we used the SA-Patchcore [22] with the pre-trained ViT [18], instead of original model that uses ResNet-based models as backbone, trained through a self-distillation method called DINO [23], on ImageNet dataset, for the better feature extraction capability. As stated earlier in III-A1 Unlike CNNs that primarily focus on local features, transformers excel at capturing global dependencies within the image. This is achieved through their self-attention mechanism, which allows the model to analyze relationships between distant pixels. This focus on global features makes transformers well-suited for tasks like object localization, where understanding the overall context of the image is crucial [24]. The mentioned model consist of 11 transformer blocks. In essence, the depth of the hierarchy directly influences the extent of global feature map comprehension, catering to specialized learning tasks. However, directly using features learned on generic objects might not be optimal for industrial anomaly detection. To address this, the model leverages features from an intermediate blocks as input to the self-attention module. This intermediate blocks capture balance between local and global features, providing a foundation for identifying co-occurrence relationships indicative of anomalies.

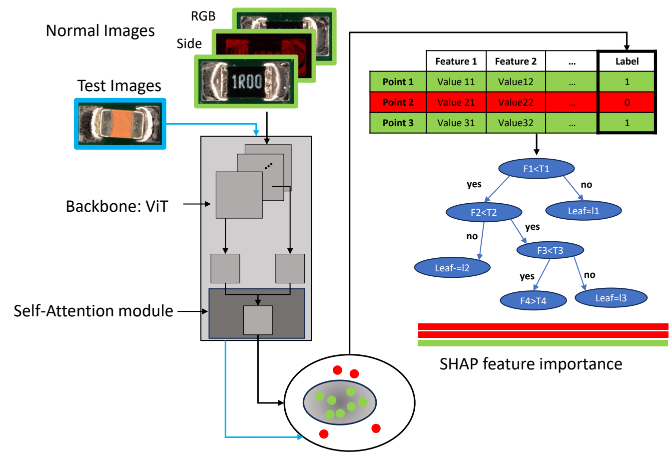


Fig. 1. Introduced framework integrates the high accuracy of industrial anomaly detection models with the explainability power of XGBoost. The industrial model identifies potential defects, while XGBoost provides insights into the features contributing to those detected.

TABLE I
DATASETS PARTITIONING AND NUMBER OF DATA POINTS

| Component | Dataset | Normal | Defected | Lighting |
|---|---|---|---|---|
| R0805 | Train | 180 | - | 2 |
| | Test | 58 | 18 | 2 |

For the vision model, we utilize Region Of Interest (ROI) images extracted from the captured data under various lighting conditions. These ROIs focus on specific components of the board. Two distinct lighting conditions are employed: a primary light source from top normal to the surface capturing colored real-world images, and a secondary light source from the side perpendicular to the board with a red filter. This multifaceted illumination approach aims to enhance the model's generalizability to real-world scenarios where lighting conditions may vary significantly [25]. The datasets used for training, feature bank construction, and testing are detailed in Table I. This comprehensive evaluation allows us to assess both the model's accuracy and its ability to provide informative explanations. In conjunction with the vision model, an XGBoost model is employed. This model leverages a separate data set containing the measured features of the ROI, and extra information (detailed in Table II) of the boards Fig 3. This table includes features derived from the 3D data, such as depth measurements or surface variations, which can complement the visual information for anomaly detection. Utilizing such approach the inspection machine is able to construct over 200 features from constructed images but only 40 most informative features are chosen to have deeper understanding of the defect types the model identifies and significantly enhance the model explainability.

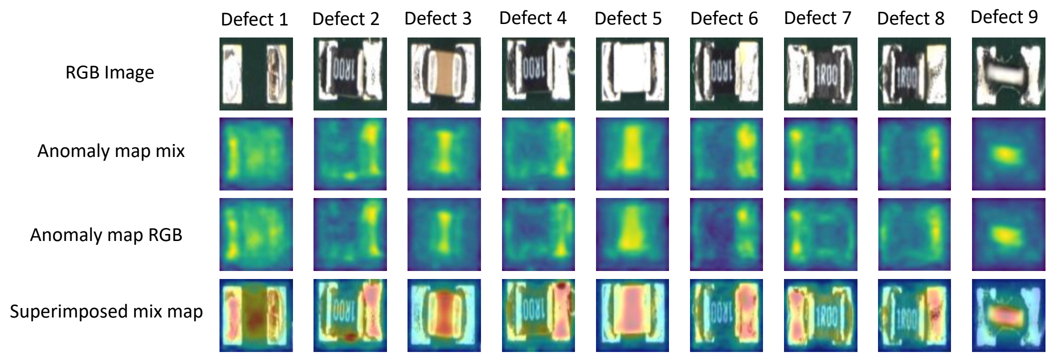


Fig. 2. Visualization of various defect types in SMD-PCBs using a vision model. Each image shows the original ROI with a corresponding attention map highlighting the detected defect. In Defect 1 the component is entirely absent from the designated placement area, in Defect 2 exhibits Pseudo solder, in Defect 3 an incorrect component is present, in Defect 4, 6 and 7 the components are misaligned, deviating from their intended placement positions, in Defect 5 the component lacks the expected identification label, in Defect 8 the component stands upright, resembling a tombstone and in Defect 9 the component is rolled.

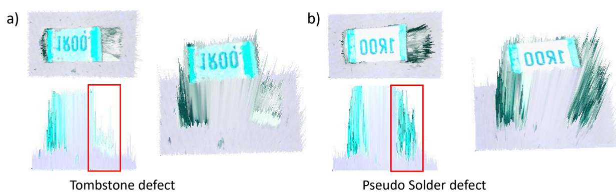


Fig. 3. Combination of surface texture information from the original image with a height indication to create 3D representations of the components. These initial 3D images are typically noisy and require further processing before being directly fed into vision models. However, they provide valuable insights into potential anomaly types and assure higher explainability. Features utilized by the XGBoost model are then extracted from these pre-processed 3D representations. a) shows the Defect 8 that has the Tombstone defect while b) shows the Defect 2 that is Pseudo Solder defect. As shown both top views look very similar.

TABLE II
DESCRIPTION OF DIFFERENT FEATURES EXTRACTED FROM GENERATED 3D IMAGES.

| Feature number | Description |
|---|---|
| 0,1 | Solder height and length |
| 2,3 | Solder angle(close to surface and close to Pad) |
| 4,6,7 | Solder volume in different sections |
| 5,8 | Lead height |
| 9,10,16-19 | Area of different solder slices |
| 11-14 | Solder height in different slices |
| 15,20,35 | Solder shape (Convexity) |
| 21-29 | Center of gravity for different slice of solder |
| 30-39 | Length and width of different slices of solder |

## IV. METRICS OF EVALUATION

To comprehensively evaluate the performance of the vision model we employed Area Under the Curve (AUC) metric which is robust particularly when dealing with imbalanced datasets. The Fig 4 shows the performance of the vision model. The AUC value is the area under the receiver operating characteristic (ROC) curve, which shows the true positive (tp) and false positive (fp) rate at various thresholds. A perfect classifier would have an ROC curve that result in an AUC of 1. Conversely, a random classifier would have an AUC of 0.5, which indicates no better performance than chance.

As demonstrated in Fig 4, the trained model achieves a high AUC, signifying its effectiveness in distinguishing between normal and abnormal images. This suggests that the model can accurately identify defects while minimizing false alarms.

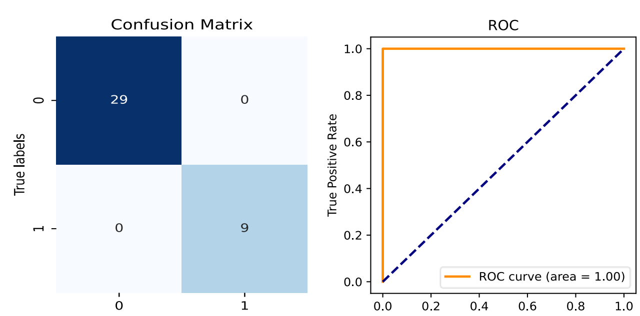


Fig. 4. Evaluation metrics for the vision model. Right figure, shows the ROC and corresponding AUC score as 1. The final performance is checked on main lighting scenario. On the left figure, confusion matrix provides a detailed breakdown of the model's classification results. It shows the number of correctly classified normal and defective ROIs, along with any misclassifications.

## V. RESULTS

As highlighted earlier, XAI empowers humans to understand the reasoning behind a model's decisions. This interpretability is crucial for building trust and ensuring responsible AI deployment, particularly in critical domains such as industrial

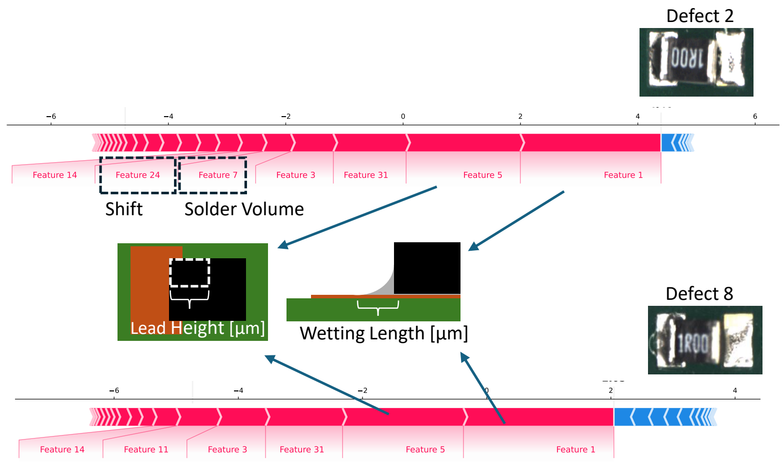


Fig. 5. SHAP feature attributions for defect detection in SMD-PCBs. Both Defect 2 and 8 share similar feature contributions, primarily focusing on right pad connection absence and slight component lift. However, Defect 2 exhibits additional features contributing (highlighted with dashed box), suggesting a potential component shift and excessive solder that makes it different from Tombstone defect type which is not possible to capture from solely top-bottom anomaly maps.

anomaly detection. The following will detail the the effectiveness of our framework in explaining the model's decisions for anomaly detection.

To further analyze the anomalies, we extract explanations and utilize attention maps to pinpoint the specific areas of deviation. Our approach to generating anomaly maps involves a two-step process. First, we extract anomaly scores at the pixel level by interpolating scores obtained from analyzing smaller image patches. This allows for a more granular localization of anomalies.

Second, we leverage the model's ability to handle different lighting conditions. We predict separate anomaly maps for each lighting configuration (e.g., real-world lighting and red filter lighting) used during training. These individual maps are then combined to create a comprehensive, lighting-agnostic anomaly map. This combined map incorporates insights from various lighting conditions, potentially leading to a more robust and generalizable representation of anomalies. Fig 2 illustrates the generated anomaly maps. The figure compares the results obtained using a single lighting condition (real-world lighting) with those achieved by combining anomaly maps from both lighting configurations. Notably, the figure reveals subtle variations in the anomaly map for certain defect types. These variations highlight the influence of lighting conditions on the model's ability to detect deviations from normal images. This underscores the importance of considering multiple lighting scenarios during model training to enhance its generalizability in real-world applications.

While the visual anomaly maps generated by vision model effectively indicated certain defects clearly specially in cases that there exist an isolated defect such as missing label on the component or incorrect component types, limitations arise when dealing with more complex scenarios. For instance:

- When multiple defects are present simultaneously, the combined anomaly score might not clearly reveal the presence and type of each individual defect.
- If distinct anomalies occur in the same or close locations, such as insufficient of excessive solder amount, the current visual map might not effectively distinguish between them.
- Defects that are not fully visible from the top-down image perspective might be missed or misinterpreted. Additional information are necessary to capture true anomalies, for instance the defects such as pseudo solder or tombstone might not be fully captured in anomaly maps.

By combining complementary features with the visual information captured in the anomaly maps, we create a more comprehensive representation of the defect. We employ SHAP (SHapley Additive exPlanations). SHAP provides valuable insights into the relative importance of different features in the model's decision-making process. This allows us to understand how different features contribute to the final anomaly detection and classification results. Presented in Fig 5, Defect 8 and Defect 2 are two extreme scenarios that the components are lifted from the soldering pads. As shown in the Fig 2 the anomaly map is correctly capturing the shift in both components, but it is failing to represent other roots of the anomaly since both anomaly representations look very similar. While using the XGBoost and SHAP we have shown that the components are lifted. Further by analyzing the feature importance scores provided by SHAP, we can potentially differentiate that the lift in Defect in 2 is not as severe in Defect 8. This enrichment process leads to a more robust and interpretable model that can effectively differentiate between various defect types, even in complex situations.

## VI. Discussion and Conclusion

Empowering the industrial anomaly detection with advanced AI models, have shown promises in achieving high accuracy. While recently advancements in explainability techniques make them more interpretable and transparent and closer for real-world applications in industrial domain. Nevertheless significant challenge lies in the lack of rich datasets. Not all scenarios will have readily available, high-quality tabular data alongside the visual data. This highlights the need for a hybrid framework that can leverage the strengths of both approaches while acknowledging the limitations in data availability and potential misclassification by the vision model. Such a framework holds immense potential for the future of anomaly detection.

In this work, we presented a novel framework for explainable anomaly detection within the context of Industry 4.0 and 5.0 applications. Our framework leverages the strengths of a vision model for capturing visual anomalies and integrates them with complementary features from tabular data. This hybrid approach overcomes the limitations of relying solely on visual anomaly maps, particularly in cases with combined, overlapping, or partially hidden defects. By incorporating additional information from historical records, sensor readings, or component specifications, we create a more comprehensive representation of the anomaly, leading to a more robust and interpretable model.

Furthermore, the utilization of SHAP provides valuable insights into the relative importance of different features used by the model. This allows us to understand the reasoning behind the model's decisions and facilitates the communication of these explanations to human inspectors. This enhanced explainability is crucial for building trust in AI-powered anomaly detection systems and ensuring their responsible deployment in real-world industrial settings.

As an important future step, this framework will be implemented in a real-world production line. Operator feedback will be collected to evaluate the model's effectiveness and understand its decision-making process. As a crucial aspect of application, operators will receive proper training on how to utilize these technologies effectively. This training will cover the framework's functionalities, interpreting its outputs, and providing valuable feedback for further model refinement. This real-world deployment will also capture a wider range of data, allowing us to further enhance the mode's robustness, especially against new defect samples.

## REFERENCES


[1] S. Hyun Park, W. Seon Shin, Y. Hyun Park, and Y. Lee, "Building a new culture for quality management in the era of the Fourth Industrial Revolution," *Total Quality Management and Business Excellence*, vol. 28, no. 9-10, 2017.

[2] X. Xu, Y. Lu, B. Vogel-Heuser, and L. Wang, "Industry 4.0 and Industry 5.0—Inception, conception and perception," *Journal of Manufacturing Systems*, vol. 61, pp. 530–535, 10 2021.

[3] R. Gupta, "Failure Mode and Effects Analysis of PCB for Quality Control Process," *Mapan - Journal of Metrology Society of India*, vol. 38, no. 2, 2023.

[4] J. Zipfel, F. Verworner, M. Fischer, U. Wieland, M. Kraus, and P. Zschech, "Anomaly detection for industrial quality assurance: A comparative evaluation of unsupervised deep learning models," *Computers and Industrial Engineering*, vol. 177, 2023.

[5] A. A. R. M. Ebayyeh and A. Mousavi, "A Review and Analysis of Automatic Optical Inspection and Quality Monitoring Methods in Electronics Industry," 2020.

[6] W. Dai, A. Mujeeb, M. Erdt, and A. Sourin, "Soldering defect detection in automatic optical inspection," *Advanced Engineering Informatics*, vol. 43, 2020.

[7] J. Senoner, T. Netland, and S. Feuerriegel, "Using Explainable Artificial Intelligence to Improve Process Quality: Evidence from Semiconductor Manufacturing," *Management Science*, vol. 68, no. 8, pp. 5704–5723, 8 2022.

[8] M. Hehr, A. Sadafi, C. Matek, P. Lienemann, C. Pohlkamp, T. Haferlach, K. Spiekermann, and C. Marr, "Explainable AI identifies diagnostic cells of genetic AML subtypes," *PLOS Digital Health*, vol. 2, no. 3, 2023.

[9] F. Shi, J. Wang, J. Shi, Z. Wu, Q. Wang, Z. Tang, K. He, Y. Shi, and D. Shen, "Review of Artificial Intelligence Techniques in Imaging Data Acquisition, Segmentation, and Diagnosis for COVID-19," 2021.

[10] J. Černevičienė and A. Kabašinskas, "Review of Multi-Criteria Decision-Making Methods in Finance Using Explainable Artificial Intelligence," 2022.

[11] H. Hagras, "Toward Human-Understandable, Explainable AI," *Computer*, vol. 51, no. 9, 2018.

[12] I. Ahmed, G. Jeon, and F. Piccialli, "From Artificial Intelligence to Explainable Artificial Intelligence in Industry 4.0: A Survey on What, How, and Where," *IEEE Transactions on Industrial Informatics*, vol. 18, no. 8, 2022.

[13] Y. A. Karayiannis, R. Stojanovic, P. Mitropoulos, C. Koulamas, T. Stouraitis, S. Koubias, and G. Papadopoulos, "Defect detection and classification on web textile fabric using multiresolution decomposition and neural networks," in *Proceedings of the IEEE International Conference on Electronics, Circuits, and Systems*, vol. 2, 1999.

[14] G. Acciani, G. Brunetti, and G. Fornarelli, "Application of neural networks in optical inspection and classification of solder joints in surface mount technology," *IEEE Transactions on Industrial Informatics*, vol. 2, no. 3, 2006.

[15] V. Zavrtanik, M. Kristan, and D. Skočaj, "DRÆM - A discriminatively trained reconstruction embedding for surface anomaly detection," in *Proceedings of the IEEE International Conference on Computer Vision*, 2021.

[16] Z. Li, Y. Zhu, and M. Van Leeuwen, "A Survey on Explainable Anomaly Detection," *ACM Transactions on Knowledge Discovery from Data*, vol. 18, no. 1, 2023.

[17] Z. Liu, Y. Zhou, Y. Xu, and Z. Wang, "SimpleNet: A Simple Network for Image Anomaly Detection and Localization," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2023-June, 2023.

[18] A. Dosovitskiy, L. Beyer, A. Kolesnikov, D. Weissenborn, X. Zhai, T. Unterthiner, M. Dehghani, M. Minderer, G. Heigold, S. Gelly, J. Uszkoreit, and N. Houlsby, "AN IMAGE IS WORTH 16X16 WORDS: TRANSFORMERS FOR IMAGE RECOGNITION AT SCALE," in *ICLR 2021 - 9th International Conference on Learning Representations*, 2021.

[19] G. Nuti, L. A. Jiménez Rugama, and A. I. Cross, "An Explainable Bayesian Decision Tree Algorithm," *Frontiers in Applied Mathematics and Statistics*, vol. 7, 2021.

[20] C. Rudin, "Stop explaining black box machine learning models for high stakes decisions and use interpretable models instead," 2019.

[21] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, vol. 13-17-August-2016, 2016.

[22] K. Ishida, Y. Takena, Y. Nota, R. Mochizuki, I. Matsumura, and G. Ohashi, "SA-PatchCore: Anomaly Detection in Dataset With Co-Occurrence Relationships Using Self-Attention," *IEEE Access*, vol. 11, 2023.

[23] M. Caron, H. Touvron, I. Misra, H. Jegou, J. Mairal, P. Bojanowski, and A. Joulin, "Emerging Properties in Self-Supervised Vision Transformers," in *Proceedings of the IEEE International Conference on Computer Vision*, 2021.

[24] Z. Liu, Y. Lin, Y. Cao, H. Hu, Y. Wei, Z. Zhang, S. Lin, and B. Guo, "Swin Transformer: Hierarchical Vision Transformer using Shifted Windows," in *Proceedings of the IEEE International Conference on Computer Vision*, 2021.

[25] Z. Liu, Q. Wang, R. Nestler, and G. Notni, "Investigation on automated visual SMD-PCB inspection based on multimodal one-class novelty detection," in *SPIE Optical Metrology*, 2023.