Anomaly Detection for Additive Manufacturing with Antenna-Based Embedding Sensors: An Explainable, Instructive Learning Framework

Wenkai Tan[1], **Houbing Song**[1]
PCAM Team[12]
[1]University of Maryland, Baltimore County
Baltimore, MD, USA
[2] Embry-Riddle Aeronautical University
Daytona Beach, FL, USA

Abstract

Additive manufacturing enables the layer-by-layer fabrication of complex and customizable structures but remains prone to internal defects that are difficult to detect during fabrication. This work presents an anomaly detection framework for real-time defect monitoring using a custom-designed wireless antenna sensor. The sensor consists of an aluminum split-ring resonator patterned with gold conductive ink and mounted on a ceramic substrate. It passively responds to structural and thermal changes during the manufacturing process by producing frequency-domain scattering parameter signals. To process these high-dimensional signals without requiring defect-labeled data, we develop a zero-bias deep neural network trained exclusively on normal-class samples. The model constructs a class-structured latent space using cosine similarity to learned templates and applies Mahalanobis distance to define statistical decision boundaries. Anomalies are identified when inputs deviate significantly from the expected distribution of normal data. The framework is designed around two deployment-oriented principles. Explainability enables users and domain experts to understand detection decisions through latent space visualization, feature attribution, and statistical reasoning that connects model outputs to interpretable signal characteristics. Instructibility allows users to tune detection sensitivity through threshold control without retraining, adapting the system to different quality requirements and process tolerances. The system is evaluated using experimental data collected from the antenna sensor under controlled thermal variation. Results demonstrate that the proposed method accurately detects anomalous signal patterns without using defect supervision, while maintaining interpretability and user control. This approach offers a transparent, adaptive, and data-efficient solution for in situ defect monitoring in additive manufacturing environments.

Code — https://colab.research.google.com/drive/1ljGlaFfSnwMrycLCk2k3fM-hB2hRTqSm?usp=sharing

Datasets — https://drive.google.com/drive/folders/1kQet4mDaNtM3hL_8mj0BLa5rJy3aMiJo?usp=sharing

Introduction

Additive manufacturing (AM), commonly referred to as 3D printing, has become a cornerstone of Industry 4.0, enabling layer-by-layer construction of highly customized components with complex internal geometries. This process reduces material waste and allows for greater design flexibility, making it attractive in domains such as aerospace, biomedical implants, and microelectronics fabrication (Vaezi, Seitz, and Yang 2013; Khorasani et al. 2022). However, the intrinsic complexity of AM introduces significant challenges in process stability and defect control. Defects such as incomplete fusion, delamination, or geometric distortion may arise unpredictably during printing, often remaining undetected until post-fabrication testing (Francis et al. 2018).

To address these challenges, real-time monitoring using embedded or external sensors has been proposed. Among various sensing modalities, resonator-based antenna sensors, including microwave and acoustic resonators, have shown promise due to their sensitivity to changes in material properties, temperature, and structural integrity (Mohammadi et al. 2020; Fieber et al. 2020; Peng et al. 2020). These sensors generate scattering parameter data (S-parameters), which capture dynamic electromagnetic responses of the printed material. While rich in information, these signals are typically high-dimensional, nonlinear, and context-dependent, making manual interpretation difficult. Further, variations in geometry, material composition, or antenna placement can result in overlapping signal signatures, even between normal and defective samples. Traditional threshold-based detection methods are often inadequate, as they fail to generalize across different printing conditions or defect types (Kumar et al. 2023).

In this context, machine learning (ML) emerges as a powerful approach to extract latent patterns from sensor data. Techniques such as neural networks and kernel-based classifiers have been explored to detect AM defects or predict resonator responses (Ghosal et al. 2020; Rooney 2023). However, several practical limitations hinder their real-world deployment. First, supervised learning methods typically require large amounts of labeled data, especially defective samples, which are difficult to obtain in low-defect-rate production environments or regulated industries like aerospace and medical devices (Zhang et al. 2023). Second, many ML models operate as black boxes, offering limited interpretability for engineers or operators, which impedes their integration into safety-critical workflows (Jagatheesaperu-

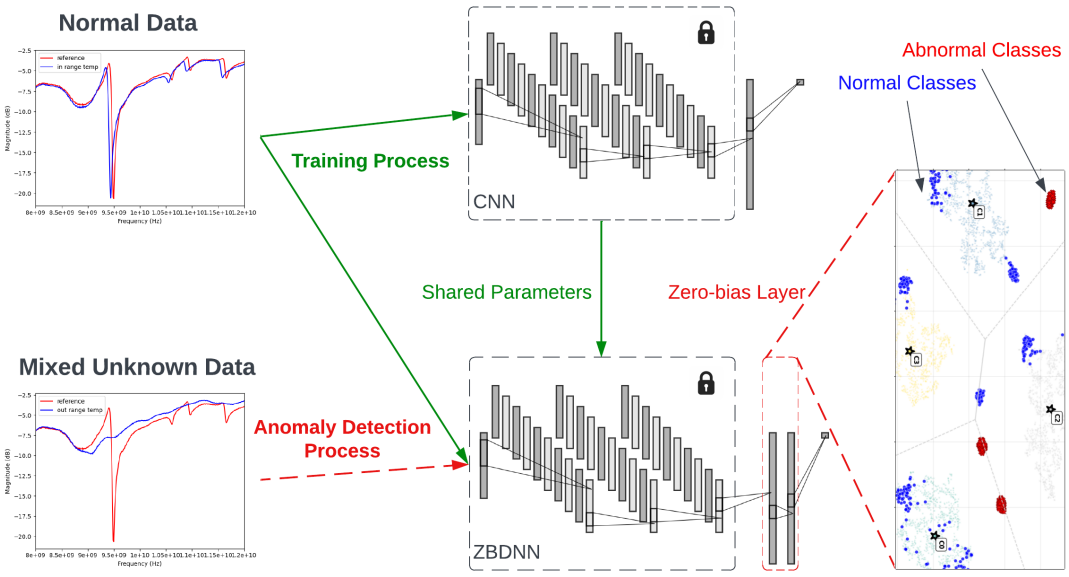


Figure 1: The framework of quick anomaly detection with latent space explanation.

mal et al. 2022; Sharma, Zhang, and Rai 2021; Singh, Rathi, and Antony 2023). Third, many state-of-the-art models are computationally demanding, making them impractical for embedded, in-situ monitoring systems that require fast and lightweight inference (Fan et al. 2023).

To overcome these barriers, we propose an anomaly detection framework for antenna-based defect detection in AM emphasizing explainability and instructibility. Explainability ensures detection decisions can be understood and validated through geometric reasoning in latent space and visual interpretation using dimensionality reduction techniques. Instructibility enables users to adjust latent-space thresholds without retraining, reflecting known tolerances or process-specific characteristics. Together, these principles facilitate transparent operation and enable domain experts to maintain control over detection sensitivity.

Our main contributions are summarized as follows:

- **A Defect Detection Framework Without Defective Data in Antenna Embedded AM**: We develop a classification-based anomaly detection framework using only normal antenna signal data, addressing defect-labeled signal scarcity in additive manufacturing and enabling multiclass discrimination and anomaly identification.
- **Explainable Latent-Space Reasoning with Statistical Decision Boundaries**: We formulate the decision space using cosine similarity and Mahalanobis distance for geometric class separation and deviation detection. This enables interpretability through statistical reasoning and visual analysis, linking decisions to antenna signal characteristics.
- **Controllable Detection Logic via Instructible Decision Thresholds**: Our architecture enables controllable detection sensitivity through latent-space thresholds without retraining, allowing user-defined robustness-sensitivity tradeoffs for diverse AM processes.
- **Antenna Dataset for Frequency-Domain Defect Detection in Manufacturing**: We construct a dataset from S-parameter measurements of flexible slot antennas across multiple AM configurations, supporting multiclass learning and anomaly evaluation.

Background and related work

AM systems increasingly rely on intelligent quality monitoring, where learning-based models must operate under physical constraints, limited supervision, and real-time decision demands. Among various sensing modalities for in-situ monitoring, antenna systems have shown particular promise due to their sensitivity to dielectric and structural variations (Mohammadi et al. 2020; Fieber et al. 2020). By sweeping across frequency ranges and measuring scattering responses, antennas generate S-parameter data that reflect subtle material or geometric changes, making them useful for detecting defects such as deformation, voids, and layer discontinuities. However, these high-dimensional frequency-domain signals are challenging to interpret manually and often require advanced analytical methods. This section reviews the key research areas that form the foundation of our work.

Unsupervised Anomaly Detection

Unsupervised anomaly detection identifies outliers without labeled defect samples, suitable for manufacturing where defect data are scarce. These approaches construct statistical models or decision boundaries from normal data and flag significant deviations.

ECOD (Li et al. 2022) detects outliers by analyzing tail probabilities of empirical marginal distributions. It computes empirical cumulative distribution functions per dimension and identifies samples with low joint tail probabilities

as anomalies. This parameter-free approach is computationally efficient and robust to diverse data characteristics.

SUOD (Zhao et al. 2021) addresses computational challenges of applying multiple detectors to large-scale datasets by approximating base detector outputs through random projection and model selection. It combines algorithms like isolation forests, k-nearest neighbors, and local outlier factor, balancing accuracy with efficiency for real-time industrial monitoring.

Other methods include one-class SVM (Schölkopf et al. 2001), which separates normal data from the origin in high-dimensional space, and autoencoders (Zhou and Paffenroth 2017), which identify anomalies via reconstruction error. These have been applied in fraud detection, network intrusion detection, and industrial quality control.

Deep Learning for High-Dimensional Sensor Signals

DL demonstrates superior capabilities in modeling high-dimensional, nonlinear relationships in antenna-based signal responses. Unlike traditional models requiring manual feature extraction, DNNs directly process raw or minimally preprocessed sensor inputs and learn hierarchical representations capturing local and global patterns.

In AM, these models predict defect types, classify material states, and estimate process parameters (Ghosal et al. 2020; Rooney 2023). However, most DL methods rely on supervised training using labeled defective data, often unavailable in real-world AM (Zhang et al. 2023). Additionally, standard architectures with bias terms and fully connected layers entangle decision logic, hindering feature isolation and unseen anomaly identification.

Explainability and Instructibility in AI Systems

As AI systems are deployed in manufacturing and high-stakes domains, ensuring compatibility with human understanding and operational requirements has become critical, articulated through explainability and instructibility concepts from trustworthy and human-centered AI (Bender and Koller 2020; Gabriel 2020; Bommasani 2021).

Explainability refers to producing outputs whose reasoning can be understood, validated, and traced to meaningful inputs. Feature attribution methods like Grad-CAM (Selvaraju et al. 2017), LayerCAM (Jiang et al. 2021), and integrated gradients identify input regions contributing most to predictions. Dimensionality reduction techniques including PCA (Jolliffe and Cadima 2016) and t-SNE (Maaten and Hinton 2008) enable visualization of learned representations (Tan et al. 2024). Attention mechanisms reveal input focus during decision-making (Vaswani et al. 2017). In manufacturing, explainability validates model decisions against domain knowledge and identifies failure modes (Jagatheesaperumal et al. 2022).

Instructibility emphasizes incorporating human direction and responding to task-specific constraints without retraining. Interactive machine learning research explores adjustable thresholds, calibrated scoring, and user-configurable parameters for adaptability (Das, Adepu, and Zhou 2020; Ruff et al. 2020). In manufacturing, domain knowledge guides preprocessing, feature selection, or regularization. Sharma et al. (Sharma, Zhang, and Rai 2021) review how process information informs representation learning, while Feng et al. (Feng et al. 2020) show engineering heuristics improve deep learning generalization in industrial settings, highlighting the need for human-guided decision logic in environments requiring model auditing and revalidation.

Statistical distance metrics support interpretability in anomaly detection. Mahalanobis distance evaluates deviations from known distributions while accounting for feature correlations (Yan et al. 2023), widely adopted to define class-consistent boundaries and quantify abnormality.

Current Limitations

The reviewed literature highlights several gaps that motivate our approach. First, while antenna-based sensing provides rich frequency-domain signals for AM monitoring, existing methods often rely on manual threshold tuning or require extensive labeled defect data. Second, although unsupervised anomaly detection methods can operate without defect labels, classical approaches typically analyze feature-level statistics without learning hierarchical representations, while deep learning methods often lack transparency and controllability. Third, explainability and instructibility are recognized as important but remain underexplored in the context of antenna-based defect detection for AM. Our work addresses these gaps by developing a zero-bias deep neural network framework that combines several key innovations.

Methodology

Problem Definition

Given antenna signal responses collected exclusively under normal additive manufacturing conditions, the task is to construct a model that classifies signals into predefined categories and identifies defects as anomalies, without using any defect-labeled data. The model must produce decisions that are grounded in signal structure, instructible through human-adjustable thresholds, and aligned with manufacturing objectives such as quality control and defect rejection.

Sensor Design and Experimental Setup

The sensor used in this study is based on a planar split-ring resonator (SRR) fabricated on a ceramic substrate. The resonator itself is made from aluminum, selected for its stable electromagnetic properties and compatibility with high-temperature environments. A gold-based conductive ink is applied to define the SRR pattern on the aluminum surface, forming the resonant structure required for frequency-domain sensing. The substrate underneath the resonator is a ceramic platform that provides mechanical support and electrical isolation. Multiple SRR units with different gap sizes ranging from 100 to 400 μm were fabricated to examine the effect of geometric variation on resonant behavior. A close-up view of the fabricated sensor is shown in Figure 2.

Sensor characterization was conducted using a Keysight PNA network analyzer (model N5227B), which measured

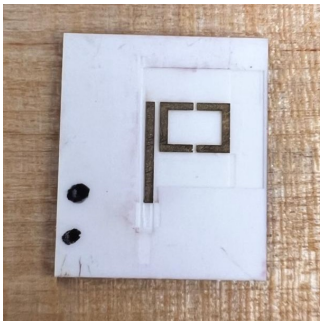


Figure 2: Fabricated split-ring resonator (SRR) on alumina substrate. The SRR was printed using gold-based conductive ink with a slot gap of 100–400 µm.

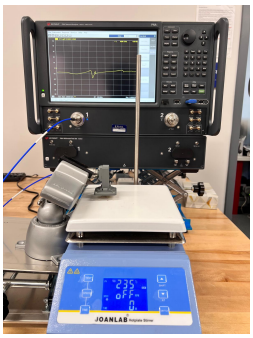
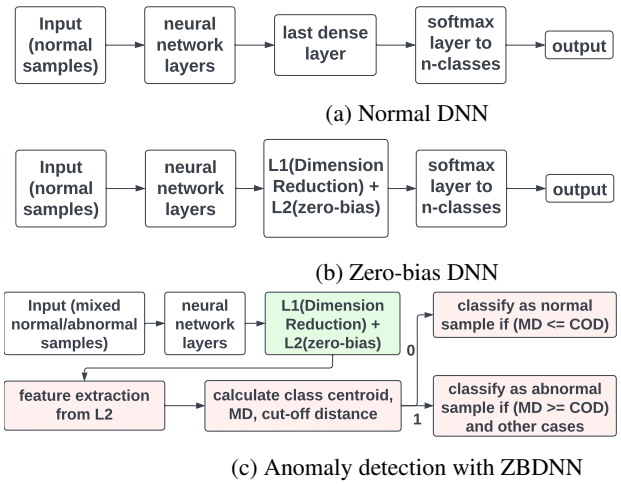


Figure 3: Experimental setup for sensor testing. The SRR sensor is measured using a Keysight N5227B network analyzer and heated on a programmable hot plate.



(a) Normal DNN

(b) Zero-bias DNN

(c) Anomaly detection with ZBDNN

Figure 4: Quick Anomaly Detection with Zero-bias Neural Network

the reflection coefficient (S_{11}) over the frequency range of 8 to 12GHz. A waveguide probe setup enabled non-contact interrogation of the sensor. The device was mounted on a programmable hot plate to enable temperature control during testing. Measurements were recorded at discrete temperature points ranging from 23°C to 200°C. At each step, the S_{11} response was captured and later used to extract the resonant frequency of the SRR. The full experimental setup is shown in Figure 3.

The acquired S-parameter data are used as the input to the anomaly detection framework described in Section .

Zero-bias DNN

DNN models are capable of learning complex patterns from large sets of preprocessed signal response data and can perform accurate classification when provided with sufficient training samples. These models are typically trained on labeled datasets that include examples from all relevant classes. In a typical DNN, shallow layers extract basic features from the input and pass them through intermediate layers, while deeper layers learn to recognize high-level patterns associated with specific categories. In the case of antenna signal responses, this includes learning to distinguish patterns related to the presence of defects. However, traditional deep learning approaches rely on extensive training data that include both normal and defective samples. A major limitation is that such models are unable to detect defects effectively unless defect-labeled training data are available.

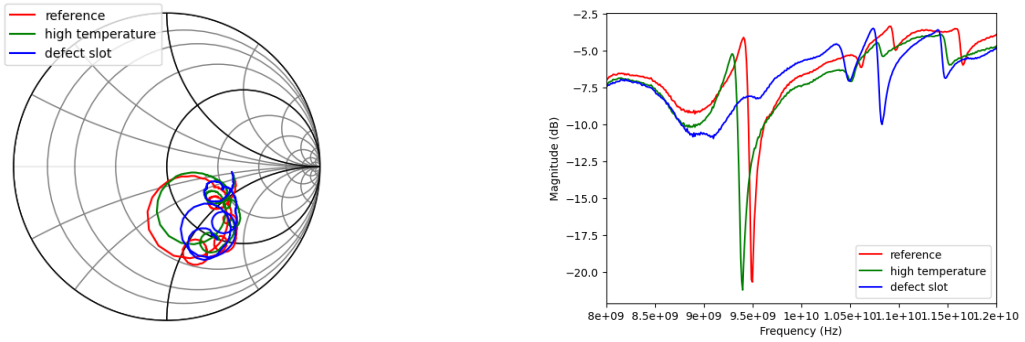
In this study, we adopt a recently proposed Zero-bias DNN (ZBDNN) (Liu et al. 2020) to address the limitations of conventional DNN training in defect detection for AM. The structure of the Zero-bias model, along with its comparison to a traditional DNN, is illustrated in Figures 4a and 4b. A standard DNN for classifying signal response data typically consists of multiple one-dimensional convolutional layers for latent feature extraction, followed by fully connected layers that reduce dimension and generate the final classification output. In contrast, the ZBDNN replaces the final fully connected layer $\mathbf{L}$ with two sequential layers defined as:

$$\mathbf{L} = \mathbf{L}_1(\text{dimension reduction}) + \mathbf{L}_2(\text{Zero-bias matching}). \tag{1}$$

The Zero-bias layer replaces the standard fully connected classifier with a normalized projection that computes the similarity between latent features and learned class templates. By removing bias terms and applying L_2 normalization to both inputs and weights, this layer transforms the classification process into a direction-aware comparison in latent space.

Replacing the final fully connected layer in a traditional DNN with a two-stage Zero-bias structure is mathematically equivalent in terms of representational capacity. Therefore, the ZBDNN maintains classification accuracy comparable to conventional DNNs while fundamentally improving the interpretability and geometric structure of the decision process. In the $\mathbf{L}_2$ stage, cosine similarity is computed between the latent representation of an input and the weight vectors representing each class. This produces a similarity score vector that reflects how closely the input aligns with each class template in latent space. These similarity scores serve as the basis for interpretable and efficient anomaly detection.

To apply the ZBDNN for defect detection, we follow a three-step process. First, a traditional DNN is trained us-



(a) (b)

Figure 5: Smith plot and Magnitude response plot of antenna S-parameter signals under three conditions: room temperature (red), high temperature (green), and defect slot (blue). The trajectory shifts reflect impedance changes caused by thermal and geometric defects.

ing only normal-class antenna data to verify the quality of the preprocessed dataset. Its performance is validated using cross-validation, accuracy metrics, and confusion matrices. Second, the same normal dataset is used to train the ZBDNN. Once the model achieves performance comparable to the traditional DNN, training is stopped to avoid overfitting. A final evaluation is performed to confirm consistency in classification accuracy and stability.

In the final stage, the trained ZBDNN is used to extract latent features and detect potential defects. Each input signal passes through the network, and its corresponding latent representation is computed by the $\mathbf{L}_2$ layer. The $\mathbf{L}_1$ layer performs a standard projection:

$$Y_1 = W_1 X + b_1, \tag{2}$$

where Y_1 is the output of the $\mathbf{L}_1$ layer, W_1 is the trained projection weight, and b_1 is the bias term. This intermediate result is then passed into the Zero-bias $\mathbf{L}_2$ layer, which eliminates bias and computes a final linear transformation:

$$Y_2 = W_2 Y_1 = W_2 W_1 X + W_2 b_1. \tag{3}$$

The output vector Y_2 can be expanded into individual similarity scores for each class:

$$Y_2[y_{1k}] = [w_{21} \cdot y_{1k}, w_{22} \cdot y_{1k}, \ldots, w_{2n} \cdot y_{1k}]. \tag{4}$$

Each term in this vector corresponds to the dot product between the latent feature and a class fingerprint. This can be reformulated as a cosine similarity:

$$w_{2n} \cdot y_{1k} = |w_{2n}| \cdot |y_{1k}| \cdot \cos(w_{2n}, y_{1k}). \tag{5}$$

To quantify the boundary between normal and abnormal samples, we compute the covariance matrix of the latent outputs:

$$\mathrm{Cov}[X] = \mathrm{cov}(Y_2[y_k], Y_2[y_k]). \tag{6}$$

Each class's centroid is defined as the mean of its corresponding latent vectors:

$$C_0^i = \mathrm{mean}(Y_2[y_k]). \tag{7}$$

We then use the Mahalanobis distance to measure the deviation of an input sample from its class centroid:

$$\mathrm{MD} = \sqrt{(Y_2[y_k] - C_0^i)^T \cdot \mathrm{Cov}[X]^{-1} \cdot (Y_2[y_k] - C_0^i)}. \tag{8}$$

A cut-off distance is defined as the largest Mahalanobis distance observed among the training samples for a given class:

$$\mathrm{COD} = \max(\mathrm{MD}(Y_2[y_k], C_0^i)) \tag{9}$$

Based on this threshold, a binary decision rule for anomaly detection is defined:

$$\text{Anomaly} = \begin{cases} 0 & \text{if MD}(Y_2[y_k], C_0^i) \leq \text{COD} \\ 1 & \text{otherwise.} \end{cases} \tag{10}$$

For an unknown dataset that includes both normal and defective samples, each input is passed through the trained ZBDNN. Its feature vector is extracted via the $\mathbf{L}_2$ layer, and its Mahalanobis distance to the class centroid is computed. If this distance exceeds the cut-off threshold, the input is classified as defective. This process is illustrated in Figure 4c, where feature extraction is highlighted in green and the anomaly decision logic in red.

Evaluation and Discussion

Dataset and Pre-processing

To evaluate the proposed anomaly detection framework, we collected scattering parameter data from the fabricated antenna sensor under various thermal and geometric conditions. The measurements include three representative signal groups: (1) a normal baseline collected at standard temperature and gap size, (2) a temperature-induced defect set collected using the same sensor geometry but under elevated thermal conditions that fall outside the process tolerance, and (3) a geometric deviation set collected under the same temperature as the baseline but with a modified resonator gap. These three signal categories are visualized for comparison in both the Smith chart and the magnitude response plot (in decibels), as shown in Figures 5a and 5b, respectively.

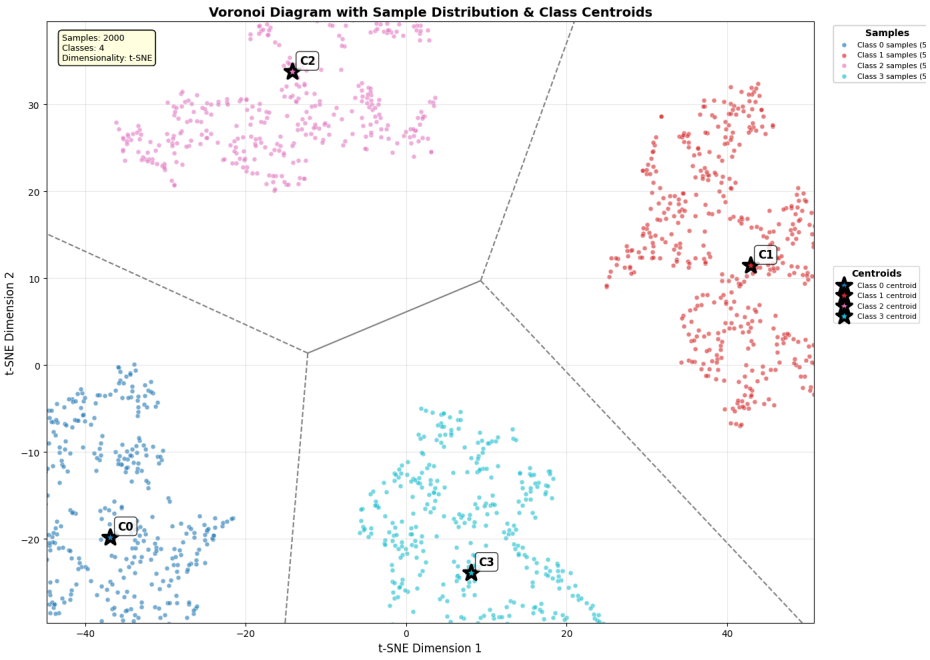


Figure 6: Visualization on latent representation.

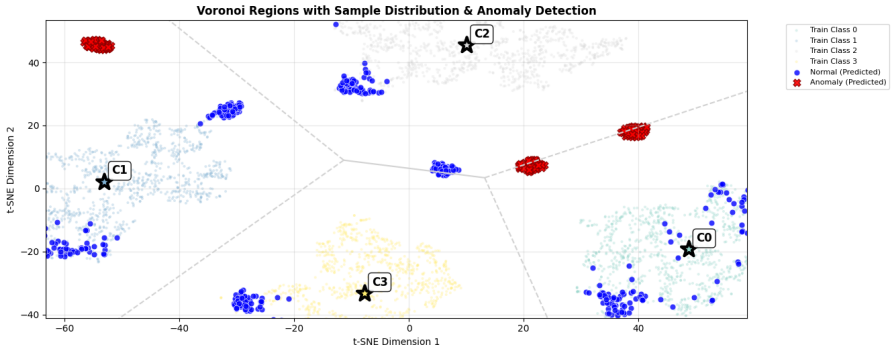


Figure 7: Voronoi diagram of Anomaly Detection with ZBDNN.

Dataset	S-parameter		MNIST	
	F1 score	Training (s)	F1 score	Training (s)
ZBDNN (Ours)	0.9091	29	0.8006	84
ECOD	0.823	286	0.769	>600
SUOD	0.841	332	0.785	>600

Table 1: Performance Comparison

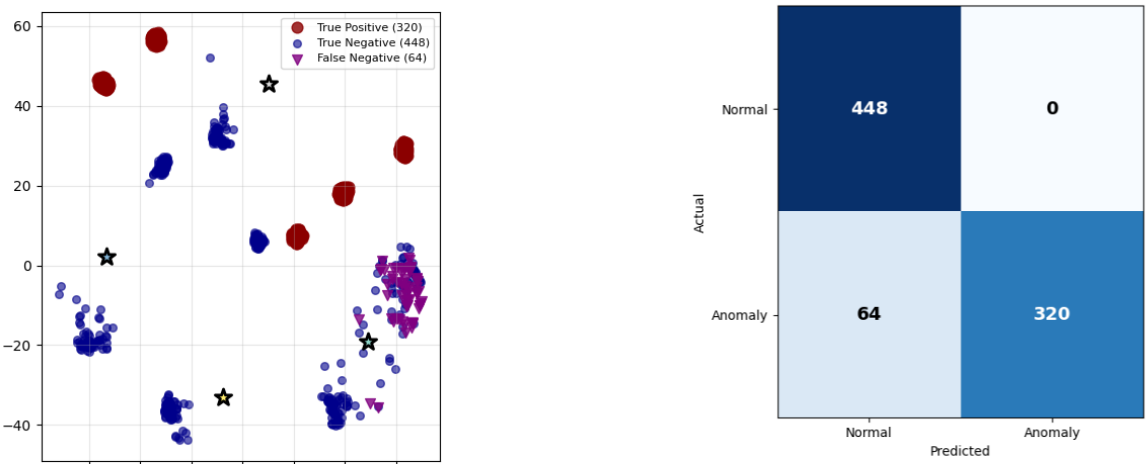
The temperature-induced variation causes a leftward shift in the resonant frequency while preserving the original spectral features, reflecting a change in the effective dielectric environment. The geometric deviation introduces a localized structural disturbance that disrupts the original resonance behavior and results in the emergence of an entirely new reflection pattern.

The normal signal group is used to construct the training dataset for the proposed model. Prior to training, each signal is normalized using zero-mean normalization, where the mean and standard deviation are computed across the training set and stored for later use. This normalization is applied independently to each frequency bin in the magnitude response to eliminate scale-related bias during training.

To evaluate the method's robustness under realistic deployment conditions, we construct an additional testset referred to as the unknown dataset. This dataset is composed of a randomly shuffled mixture of normal and defect samples, including both thermal and geometric variations. To reflect a practical inference scenario, the unknown dataset

Layer	Output Shape	Description
Input	[batch_size, 2, 401]	Two-channel input (real + imaginary)
Conv Block 1	[batch_size, 10, 397]	Conv1D(kernel=5) + BatchNorm + ReLU
Conv Block 2	[batch_size, 10, 397]	Conv1D(kernel=3, padding=1) + ReLU
Conv Block 3	[batch_size, 10, 397]	Conv1D(kernel=3, padding=1) + ReLU
Concatenation	[batch_size, 20, 397]	Skip connection: Block 1 + Block 3
Flatten	[batch_size, 7940]	Flatten for fully connected input
Fully Connected	[batch_size, 4]	Linear layer before fingerprint layer
Zero-Bias Fingerprint	[batch_size, 4]	Cosine-similarity-based projection
Custom Softmax	[batch_size, 4]	Softmax for classification output

Table 2: ZBDNN architecture for S-parameter classification



(a) (b)

Figure 8: Evaluation matrix of proposed anomaly detection pipeline.

is normalized using the stored statistics from the training set without any additional adjustment. This allows the model to operate under the assumption that only normal-class parameters are known a priori.

Experimental Details

The proposed anomaly detection system is implemented using a ZBDNN trained solely on normal sensor data. Each input sample consists of a two-channel representation of the antenna's S-parameter signal, with real and imaginary components sampled across 401 frequency points. The network architecture includes three one-dimensional convolutional layers followed by feature concatenation, flattening, and a dense projection layer. The full architecture is summarized in Table 2, and the final latent vector has a dimensionality of 7940, which feeds into the Zero-bias similarity-based classification structure described in Section 3.3.

Training is performed on a dataset containing four distinct configurations of normal sensor behavior, representing acceptable variations in temperature and antenna gap geometry. The model is optimized using stochastic gradient descent with a cross-entropy loss. To verify its classification capability, we evaluate it on a held-out testset of the normal dataset. The confusion matrix in Figure 9 confirms that the model successfully distinguishes among normal process classes using the learned latent representation.

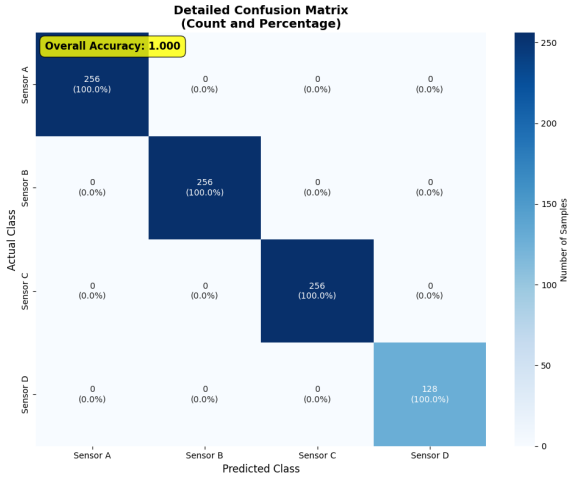


Figure 9: Model evaluation on normal testset.

After training, we extract the latent feature vectors for all normal samples. For each class, we compute the mean fea-

ture vector and define it as the class fingerprint. This fingerprint acts as a geometric reference in latent space. To establish decision boundaries for anomaly detection, the Mahalanobis distance is computed between each training sample and its corresponding class fingerprint. The maximum distance observed for each class is stored as a cut-off threshold, which defines the boundary of normal variation. Figure 6 presents a t-SNE visualization of the latent space, where distinct clusters represent the normal classes, and the fingerprints appear near the center of each cluster. This visualization demonstrates that the model has formed well-separated, class-structured representations suitable for geometric anomaly detection.

To simulate a realistic deployment scenario, we construct an unknown dataset consisting of a randomly mixed collection of normal and defective samples. These include both thermal anomalies and geometry-induced defects. All samples are normalized using the mean and standard deviation computed from prior knowledge. During inference, each test sample is passed through the trained network to obtain its latent feature representation. The cosine similarity and Mahalanobis distance to all class fingerprints are calculated and will be used for anomaly detection.

Results and Discussion

With the ZBDNN trained and class-specific fingerprints constructed, we now evaluate the anomaly detection performance of the proposed framework on an unknown dataset containing both normal and defective samples. The objective is to assess the system's ability to distinguish in-distribution samples from out-of-distribution behavior without requiring any defective data during training.

The evaluation considers multiple defect scenarios, including thermal anomalies and structural perturbations introduced by geometric deviation. These cases simulate realistic variations encountered in sensor-embedded additive manufacturing processes. We report classification results in terms of accuracy, precision, F1 Score, and compare the proposed Zero-bias framework with two representative baselines. In addition to quantitative metrics, we provide visualizations and analysis to highlight the interpretability and stability of the latent decision space.

Figure 7 illustrates the t-SNE projection of latent features extracted from the zero-bias deep neural network, overlaid with Voronoi regions constructed from the class fingerprints. Light-colored clusters represent the training samples from each class, while blue and red points denote predicted normal and anomalous samples, respectively. The black stars (C0–C3) indicate the latent centroids (fingerprints) of the four training classes.

This diagram visualizes how test samples are positioned relative to the class regions in latent space. Samples falling within a region but beyond the cut-off radius from its fingerprint are flagged as anomalies. The observed separation confirms the consistency of the learned latent structure and the effectiveness of the Mahalanobis-based anomaly detection strategy.

The t-SNE projection of test samples with anomaly detection results from the ZBDNN is shown in Figure 8a. True positives (red), true negatives (blue), and false negatives (purple triangles) are plotted with respect to the learned class centroids (black stars). The false negatives are primarily located in the lower right region and correspond to temperature-induced anomalies. Because these thermal conditions did not significantly alter the resonator's spectral response, the resulting latent representations remained close to those of normal data, leading the model to misclassify them as normal.

The corresponding confusion matrix in Figure 8b summarizes the overall classification results. The model correctly identifies 448 normal and 320 anomalous samples, demonstrating high precision and recall. The 64 false negatives align with the thermal anomalies discussed above, confirming the challenge of detecting subtle signal shifts when the antenna response remains largely unchanged.

Table 1 presents a comparison of anomaly detection performance across two datasets using three different methods: the proposed ZBDNN method, ECOD (Li et al. 2022), and SUOD (Zhao et al. 2021). For both the antenna S-parameter dataset and the MNIST benchmark, ZBDNN achieves comparable or superior F1 scores while significantly reducing training time.

Notably, ZBDNN completes training on the S-parameter dataset in only 29 seconds, compared to the 286 seconds and 332 seconds required by ECOD and SUOD, respectively. On the MNIST dataset, ZBDNN maintains reasonable performance while the other two methods require over 600 seconds to complete training. This efficiency arises from the lightweight architecture and the absence of ensemble or iterative model components.

Furthermore, both ECOD and SUOD require the contamination rate (i.e., the proportion of anomalies in the dataset) to be specified as a hyperparameter, which can be difficult to estimate in real-world scenarios. In contrast, ZBDNN operates entirely without this prior knowledge, learning class boundaries from normal data alone and enabling flexible deployment in low-defect-rate industrial settings.

Conclusion

This work presents an explainable and instructible anomaly detection framework for additive manufacturing using antenna sensor feedback and a Zero-bias DNN. The method eliminates dependence on defective training data by constructing a process-consistent latent space for anomaly identification through cosine similarity and Mahalanobis distance-based reasoning.

The framework achieves explainability through latent space visualization, feature attribution, and statistically interpretable scoring, enabling domain experts to validate model decisions. Instructibility is realized via adjustable decision thresholds, allowing users to control detection sensitivity without retraining. Integration into antenna-embedded AM platforms demonstrates practical feasibility, offering real-time defect detection with minimal computational overhead.

Future work will explore multi-modal sensing, domain adaptation across materials and geometries, and enhanced interpretation techniques.

Acknowledgments

This material is based upon work supported by the National Science Foundation under Grant No. 2229155. The opinions, findings, and conclusions, or recommendations expressed are those of the author(s) and do not necessarily reflect the views of the National Science Foundation.

References

Bender, E. M.; and Koller, A. 2020. Climbing towards NLU: On meaning, form, and understanding in the age of data. In *Proceedings of the 58th annual meeting of the association for computational linguistics*, 5185–5198.

Bommasani, R. 2021. On the opportunities and risks of foundation models. *arXiv preprint arXiv:2108.07258*.

Das, T. K.; Adepu, S.; and Zhou, J. 2020. Anomaly detection in industrial control systems using logical analysis of data. *Computers & Security*, 96: 101935.

Fan, Z.; Gao, R. X.; He, Q.; Huang, Y.; Jiang, T.; Peng, Z.; Thévenaz, L.; Xiong, Y.; Zhong, S.; et al. 2023. New Sensing Technologies for Monitoring Machinery, Structures, and Manufacturing Processes. *Journal of Dynamics, Monitoring and Diagnostics*.

Feng, J.; Yao, Y.; Lu, S.; and Liu, Y. 2020. Domain knowledge-based deep-broad learning framework for fault diagnosis. *IEEE Transactions on Industrial Electronics*, 68(4): 3454–3464.

Fieber, L.; Bukhari, S. S.; Wu, Y.; and Grant, P. S. 2020. Inline measurement of the dielectric permittivity of materials during additive manufacturing and 3D data reconstruction. *Additive Manufacturing*, 32: 101010.

Francis, M. P.; Kemper, N.; Maghdouri-White, Y.; and Thayer, N. 2018. Additive manufacturing for biofabricated medical device applications. In *Additive manufacturing*, 311–344. Elsevier.

Gabriel, I. 2020. Artificial intelligence, values, and alignment. *Minds and machines*, 30(3): 411–437.

Ghosal, R.; Maity, S.; Gupta, B.; and Majumder, A. 2020. Analytical prediction of resonant frequencies of annular stacked dielectric resonator antennas. *International Journal of RF and Microwave Computer-Aided Engineering*, 30(8): e22237.

Jagatheesaperumal, S. K.; Pham, Q.-V.; Ruby, R.; Yang, Z.; Xu, C.; and Zhang, Z. 2022. Explainable AI over the Internet of Things (IoT): Overview, state-of-the-art and future directions. *IEEE Open Journal of the Communications Society*.

Jiang, P.-T.; Zhang, C.-B.; Hou, Q.; Cheng, M.-M.; and Wei, Y. 2021. LayerCAM: Exploring Hierarchical Class Activation Maps for Localization. *IEEE Transactions on Image Processing*, 30: 5875–5888.

Jolliffe, I. T.; and Cadima, J. 2016. Principal component analysis: a review and recent developments. *Philosophical transactions of the royal society A: Mathematical, Physical and Engineering Sciences*, 374(2065): 20150202.

Khorasani, M.; Ghasemi, A.; Rolfe, B.; and Gibson, I. 2022. Additive manufacturing a powerful tool for the aerospace industry. *Rapid prototyping journal*, 28(1): 87–100.

Kumar, D.; Liu, Y.; Song, H.; and Namilae, S. 2023. Explainable deep neural network for in-plain defect detection during additive manufacturing. *Rapid Prototyping Journal*.

Li, Z.; Zhao, Y.; Hu, X.; Botta, N.; Ionescu, C.; and Chen, G. H. 2022. Ecod: Unsupervised outlier detection using empirical cumulative distribution functions. *IEEE Transactions on Knowledge and Data Engineering*, 35(12): 12181–12193.

Liu, Y.; Wang, J.; Li, J.; Song, H.; Yang, T.; Niu, S.; and Ming, Z. 2020. Zero-bias deep learning for accurate identification of Internet-of-Things (IoT) devices. *IEEE Internet of Things Journal*, 8(4): 2627–2634.

Maaten, L. v. d.; and Hinton, G. 2008. Visualizing data using t-SNE. *Journal of machine learning research*, 9(Nov): 2579–2605.

Mohammadi, S.; Nadaraja, A. V.; Roberts, D. J.; and Zarifi, M. H. 2020. Real-time and hazard-free water quality monitoring based on microwave planar resonator sensor. *Sensors and Actuators A: Physical*, 303: 111663.

Peng, X.; Kong, L.; Chen, Y.; Shan, Z.; and Qi, L. 2020. Design of a multi-sensor monitoring system for additive manufacturing process. *Nanomanufacturing and Metrology*, 3: 142–150.

Rooney, S. 2023. *AI for State Estimation and Diagnostics of Additive Manufacturing Machinery*. Ph.D. thesis, Stevens Institute of Technology.

Ruff, L.; Vandermeulen, R. A.; Franks, B. J.; Müller, K.-R.; and Kloft, M. 2020. Rethinking assumptions in deep anomaly detection. *arXiv preprint arXiv:2006.00339*.

Schölkopf, B.; Platt, J. C.; Shawe-Taylor, J.; Smola, A. J.; and Williamson, R. C. 2001. Estimating the support of a high-dimensional distribution. *Neural computation*, 13(7): 1443–1471.

Selvaraju, R. R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; and Batra, D. 2017. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. In *2017 IEEE International Conference on Computer Vision (ICCV)*, 618–626.

Sharma, A.; Zhang, Z.; and Rai, R. 2021. The interpretive model of manufacturing: a theoretical framework and research agenda for machine learning in manufacturing. *International Journal of Production Research*, 59(16): 4960–4994.

Singh, M.; Rathi, R.; and Antony, J. 2023. Interpretive model of enablers of Data-Driven Sustainable Quality Management practice in manufacturing industries: ISM approach. *Total Quality Management & Business Excellence*, 34(7-8): 870–893.

Tan, W.; Shivakumar, J. B.; Govindarajan, R. S.; Reed, N.; Kim, D.; Rojas, E.; Liu, Y.; and Song, H. 2024. Explainable Artificial Intelligence for Antenna Sensor Modeling. In *2024 IEEE International Mediterranean Conference on Communications and Networking (MeditCom)*, 559–565. IEEE.

Vaezi, M.; Seitz, H.; and Yang, S. 2013. A review on 3D micro-additive manufacturing technologies. *The International Journal of Advanced Manufacturing Technology*, 67: 1721–1754.

Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A. N.; Kaiser, Ł.; and Polosukhin, I. 2017. Attention is all you need. *Advances in neural information processing systems*, 30.

Yan, W.; Wang, J.; Lu, S.; Zhou, M.; and Peng, X. 2023. A review of real-time fault diagnosis methods for industrial smart manufacturing. *Processes*, 11(2): 369.

Zhang, Y.; Safdar, M.; Xie, J.; Li, J.; Sage, M.; and Zhao, Y. F. 2023. A systematic review on data of additive manufacturing for machine learning applications: the data quality, type, preprocessing, and management. *Journal of Intelligent Manufacturing*, 34(8): 3305–3340.

Zhao, Y.; Hu, X.; Cheng, C.; Wang, C.; Wan, C.; Wang, W.; Yang, J.; Bai, H.; Li, Z.; Xiao, C.; et al. 2021. Suod: Accelerating large-scale unsupervised heterogeneous outlier detection. *Proceedings of Machine Learning and Systems*, 3: 463–478.

Zhou, C.; and Paffenroth, R. C. 2017. Anomaly detection with robust deep autoencoders. In *Proceedings of the 23rd ACM SIGKDD international conference on knowledge discovery and data mining*, 665–674.