2025 International Conference on Intelligent Computing
July 26-29, Ningbo, China
https://www.ic-icc.cn/2025/index.php

Unsupervised Wood Surface Anomaly Detection via Enhanced GAN with Residual Dense and Attention Modules

Yuhao Guo[1,3], Fengqi Hao[1,2,3*], Qingyan Ding[1,3], Jinqiang Bai[1,3], Dexin Ma[4], Huijuan Hao[1,3]

[1] Key Laboratory of Computing Power Network and Information Security, Ministry of Education, Shandong Computer Science Center (National Supercomputer Center in Jinan) Qilu University of Technology (Shandong Academy of Sciences), Jinan, China
[2] Faculty of Data Science, City University of Macau, Macau, China
haofq@sdas.org;
[3] Shandong Provincial Key Laboratory of Industrial Network and Information System Security, Shandong Fundamental Research Center for Computer Science, Jinan, China
[4] Communication College, Qingdao Agricultural University, Qingdao 266109, China

Abstract. Reliable surface-defect inspection is a prerequisite for modern wood-processing lines, yet manually labelled defect images are inherently scarce and imbalanced. We present ERA-GANomaly, an unsupervised anomaly-detection framework that combines an encoder–decoder–re-encoder backbone with Residual Dense Blocks (RDBs) and lightweight Efficient Channel Attention (ECA) to emphasise salient textures. Experiments on three wood-defect datasets show that ERA-GANomaly attains 92.4% accuracy and a macro-F1 of 83.0%, outperforming representative unsupervised baselines such as GANomaly, EGBAD and AnoGAN. Ablation studies verify that both ECA and RDB modules contribute markedly to detecting subtle defects—including cracks, chips and bark inclusions. These findings indicate that ERA-GANomaly offers a practical, label-free solution for industrial surface-defect screening.

Keywords: Unsupervised Anomaly Detection, Generative Adversarial Networks, Attention Mechanism, Residual Dense Blocks.

1 Introduction

Surface defect detection in wood materials is critical for in-line quality inspection on modern production lines. As a natural material, wood presents highly variable textures and diverse defect types, including knots, cracks, bark inclusions, and chips. These defects often exhibit irregular shapes and indistinct boundaries, posing significant challenges for automated inspection systems. Traditional manual inspection methods are inefficient, costly, and prone to subjective bias, making them inadequate for modern industry demands, which increasingly require high-efficiency and stable quality control systems [1].

Although supervised deep learning models have achieved remarkable success in various computer vision tasks [2], their performance heavily relies on large-scale annotated datasets. In real-world industrial scenarios, abnormal samples are often scarce and unevenly distributed. Defects such as cracks (LF) and chips (QK) are particularly rare and labor-intensive to annotate, significantly constraining the scalability of supervised approaches. In contrast, normal samples are more readily available and demonstrate higher consistency, rendering unsupervised anomaly detection a more viable and scalable solution for industrial inspection applications.

Generative Adversarial Networks (GANs) have demonstrated strong potential for unsupervised anomaly detection by effectively modeling normal data distributions without relying on anomalous samples [1]. Representative methods such as GANomaly [3] and AnoGAN [4] adopt an encoder-decoder-reencoder architecture to learn latent representations and reconstruct input images, enabling anomaly detection through reconstruction error analysis. These approaches successfully address data imbalance issues and are particularly well-suited for open-set recognition tasks like wood surface defect inspection. However, existing GAN-based anomaly detection approaches still face several critical limitations. First, their generator architectures are often overly simplistic, limiting their ability to reconstruct high-frequency details and complex textures, which significantly hampers defect localization accuracy. Second, the absence of effective attention mechanisms prevents the model from focusing on diagnostically important regions. Third, most current methods rely on fixed thresholding strategies, resulting in poor adaptability and robustness in real-world deployment scenarios [1].

To address the limitations of existing GAN-based approaches in wood surface defect detection, we propose ERA-GANomaly (Efficient Residual Attention GANomaly), an enhanced framework inspired by GANomaly, specifically designed to improve the detection of complex and fine-grained anomalies in wood textures. Our key contributions are as follows:

- We constructed a wood anomaly detection dataset collected from real-world production environments and manually annotated. The dataset includes 10,000 normal samples for training and 766 bark inclusion samples, 422 chip samples, and 306 crack samples exclusively for testing, in accordance with the unsupervised anomaly detection paradigm. This dataset serves as a reliable benchmark for evaluating unsupervised methods in industrial scenarios.
- We enhance the model architecture by integrating Efficient Channel Attention (ECA) modules and replacing standard convolutional layers with Residual Dense Blocks (RDBs). This combination improves channel-wise feature discrimination and multi-level feature reuse, significantly boosting reconstruction accuracy and enabling precise localization of subtle anomalies, particularly in complex texture regions.
- Furthermore, we propose a dynamic thresholding strategy based on exhaustive F1-score optimization to replace conventional static thresholds. This adaptive mechanism substantially improves the model's robustness and decision reliability when dealing with diverse data distributions.

2 Related Work

Anomaly detection plays a pivotal role in industrial quality control, with well-established applications in material inspection, including metals [5] and semiconductors [6]. The increasing automation of wood processing has introduced unique challenges for surface defect detection, primarily due to the material's complex texture variations and anisotropic patterns [7]. Existing research generally classifies detection methodologies into statistical models [8], deep learning approaches [9], and GAN-based techniques [10]. This work specifically focuses on unsupervised reconstruction-based methods for wood defect identification. Owing to their label-free nature, reconstruction-based unsupervised anomaly detection methods have attracted significant attention in the wood industry [11]. These approaches operate under the fundamental assumption that normal samples can be accurately reconstructed, while anomalies result in substantial reconstruction errors. Although early implementations using shallow autoencoders [12] exhibited limited capability in detecting fine-grained defects, recent advancements incorporating multi-scale convolutional architectures [13] and specialized texture-aware loss functions [14] have substantially improved detection performance.

The advent of GANs has significantly advanced anomaly detection by enabling high-fidelity modeling of normal data distributions. While f-AnoGAN [15] demonstrated improved latent space representation compared to conventional autoencoders—mitigating overfitting issues—this approach exhibits two critical limitations: (1) strong dependence on comprehensive training data coverage, and (2) compromised performance in texture-rich environments such as wood surfaces, often manifesting as blurred reconstructions and unstable detection outputs. To mitigate these limitations, Li et al. [16] proposed a dual-discriminator architecture comprising: (1) a global discriminator for large-scale defect detection (e.g., resin patches), and (2) a local discriminator specialized in fine-grained anomaly identification (e.g., micro-cracks). While this hierarchical design achieves superior multi-scale detection accuracy, it introduces substantial computational overhead, increasing both model complexity by approximately 38% and inference latency by 23% compared to baseline GANs [16], thereby limiting its practicality for real-time industrial applications. Furthermore, prevailing wood surface image analysis methods typically employ shared convolutional filters across all color channels (e.g., RGB), disregarding the differential importance of individual channels in texture representation. This uniform processing paradigm frequently generates false-positive responses in defect-free regions, ultimately compromising localization accuracy.

In parallel, research on structural representation in graph-based anomaly detection has revealed similar patterns. Although prior methods such as H2GCN [19]and FAGCN [20] also attempted to incorporate local structural information to enhance node-level representations on heterogeneous or sparse graphs, their modeling strategies differ significantly. H2GCN primarily captures structure via pre-defined multi-hop neighbor connections, lacking adaptability to the varying importance of local substructures. FAGCN adopts a frequency-domain filtering perspective, but still struggles to differentiate subtle inter-node relational differences, particularly in non-uniform graph contexts. These limitations highlight the broader challenge of fixed structural priors in adaptive anomaly modeling.

To improve wood surface anomaly detection, we propose a GAN-based framework that integrates Residual Dense Blocks (RDBs) and Efficient Channel Attention (ECA). This design enhances the model's ability focus on abnormal texture regions by modeling inter-channel importance, thereby improving detection accuracy [18]. It also mitigates overfitting to normal textures and enhances reconstruction fidelity. Experimental results demonstrate superior performance in convergence speed, false positive control, and rare defect detection, highlighting the framework's strong potential for industrial applications.

3 Methods

3.1 Proposed Approach

Problem Definition. We propose a robust unsupervised anomaly detection framework that learns latent distributions exclusively from normal samples, thereby eliminating the need for anomalous training data. Specifically, let the training set be denoted as $D = \{X_1, \ldots, X_M\}\}$, consisting of M normal (defect-free) images. The test set is represented as $\hat{\mathcal{D}} = \{(\hat{X}_1, y_1), \ldots, (\hat{X}_N, y_N)\}$, where $\hat{x}_i$ is a test image and $y_i \in \{0,1\}$ is its label, with 0 indicating a normal sample and 1 indicating an abnormal sample. Typically, the training set is substantially larger than the test set, i.e., $M \gg N$.The objective is to construct an unsupervised detection model f that captures the distributional characteristics of normal data during training. During inference, for any input sample $\hat{x}$, the model produces an anomaly score $A(\hat{x})$. This score is used to measure the deviation of the test sample from the learned normal distribution. A higher score indicates a higher likelihood of anomaly. A test sample is classified as anomalous if $A(\hat{x}) > \phi$, where ϕ is a predefined anomaly detection threshold.

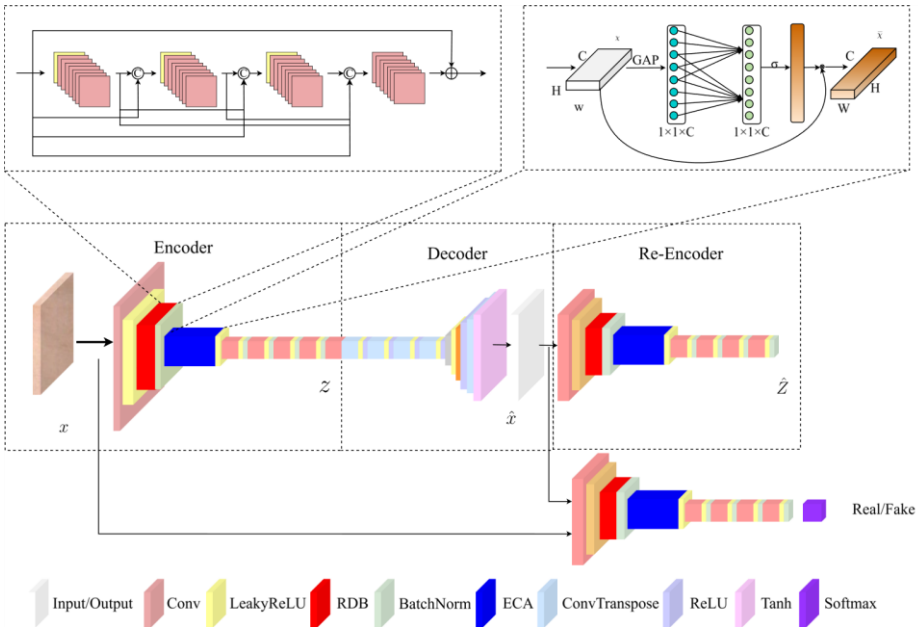


Fig. 1. Overall architecture of the proposed ERA-GANomaly model

ERA-GANomaly Pipeline. Fig. 1 illustrates the overview of our approach, which contains two encoders, a decoder, and discriminator networks, employed within three subnetworks.

During the training phase, an input image x is first encoded by the initial encoder $E1$ into a latent vector $z = E1(x)$. This encoder is composed of a series of convolutional layers, LeakyReLU activation functions, RDBs, and ECA modules, enabling the extraction of multi-scale feature representations. The resulting latent vector z is then fed into the decoder Dec, which reconstructs the image as $\hat{x} = \mathrm{Dec}(z)$. To enforce semantic consistency in the reconstruction process, a second encoder E2 re-encodes the reconstructed image $\hat{x}$,producing a second latent vector $\hat{z} = E2(\hat{x})$, which is then compared against the original latent representation z.

3.2 Residual Dense Block (RDB)

In the task of image anomaly detection, particularly for wood surface defects, models must exhibit strong feature representation capabilities to effectively capture fine-grained local variations. Conventional generative GANs, such as Deep Convolutional GANs (DCGANs), typically employ four to five layers of transposed convolution and convolution operations in the generator, starting from a dense projection of the input noise vector and progressively upsampling feature maps to produce the final image. However, such architectures exhibit inherent limitations in feature extraction, particularly in capturing subtle differences and fine-grained local textures. When applied to

wood images with complex patterns and significant structural variations, they often lead to incomplete feature representation and hindered gradient propagation, ultimately resulting in insufficient expressive capacity and severely limiting the recognition of fine-grained defects.

To overcome the limitations of conventional architectures in modeling fine-grained surface anomalies, this study incorporates enhanced RDBs into both the encoder and decoder components, as illustrated in **Fig. 2** Each RDB leverages the complementary strengths of residual learning and dense connectivity to improve feature propagation and representation. Structurally, an RDB comprises five sequential 3×3 convolutional layers, where the output of each layer is concatenated with the inputs of all subsequent layers. This progressive fusion mechanism facilitates multi-scale feature aggregation and cross-layer information flow. A final convolutional fusion layer is applied to compress the aggregated features and generate the block output.

In addition to enabling detailed spatial modeling, the internal residual connections alleviate the vanishing gradient problem often encountered in deep networks, thereby enhancing training stability. The use of stacked small-kernel convolutions further supports the accurate localization of subtle anomalies such as micro-cracks and mild discolorations in wood textures. Moreover, the RDB is designed to operate in tandem with the downstream ECA module, which adaptively emphasizes informative channels. This synergy not only improves the semantic consistency of the reconstructed images but also sharpens the model's sensitivity to localized defects via reconstruction error, thus providing a more robust foundation for unsupervised anomaly detection.

Taking the encoder as an example, the structure is as follows: The original input image (with dimensions $C \times H \times W$) is first passed through an initial convolutional operation for downsampling and feature extraction, resulting in an intermediate feature map F_O :

$$F_O = LeakyReLU(Conv1(X)) \tag{1}$$

Subsequently, the feature map is fed into two consecutive RDB modules, each comprising five densely connected convolutional layers. These blocks employ a progressive dense connectivity mechanism, where the output of each convolutional layer is concatenated with the input of all subsequent layers.

$$y_1 = \sigma(W_1 * F_O + b_1) \tag{2}$$

$$y_2 = \sigma(W_2 * [F_O, y_1] + b_2) \tag{3}$$

$$y_3 = \sigma(W_3 * [F_O, y_1, y_2] + b_3) \tag{4}$$

$$y_4 = W_4 * [F_O, y_1, y_2, y_3] + b_4 \tag{5}$$

here, $[\cdot]$ denotes channel-wise concatenation, and σ represents the LeakyReLU activation function. The final output is combined with the original input through a residual connection to form the output representation:

$$F_{RDB} = F_O + \gamma \cdot y_4 \tag{6}$$

here, γ is a learnable scaling factor used to adjust the magnitude of the residual connection. This residual formulation not only preserves the original low-level features but also reinforces the model's capacity to capture detailed variations, thereby enhancing the robustness of the encoded representation for downstream anomaly detection tasks.

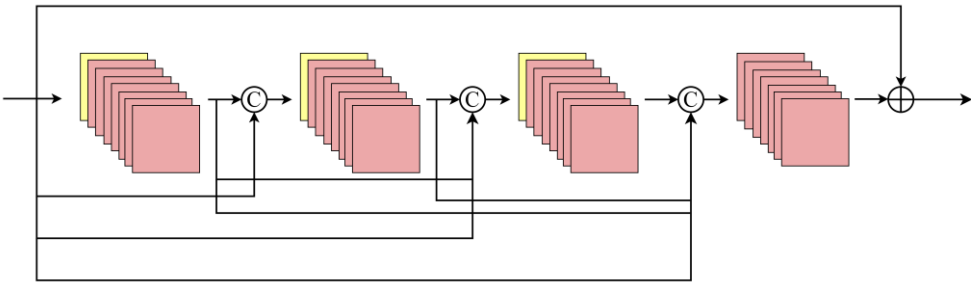


Fig. 2. Structural diagram of the Residual Dense Block (RDB)

3.3 Efficient Channel Attention (ECA)

In ERA-GANomaly, the ECA module is integrated into the backbone structure of both the encoder and decoder. Specifically, ECA blocks are inserted after several Residual Dense Blocks (RDBs) and are placed immediately following the Batch Normalization operations. The original input image x (with dimensions C × H × W) is first processed by an initial convolutional layer for downsampling and feature extraction, resulting in an intermediate feature map Fo. To further enhance inter-channel dependencies, the ECA module is introduced to perform channel-wise attention weighting. The internal structure and processing flow of the ECA module are illustrated in **Fig. 3**. Specifically, the ECA module first applies global average pooling to each feature channel to obtain a compact channel descriptor vector z ∈ R^C:

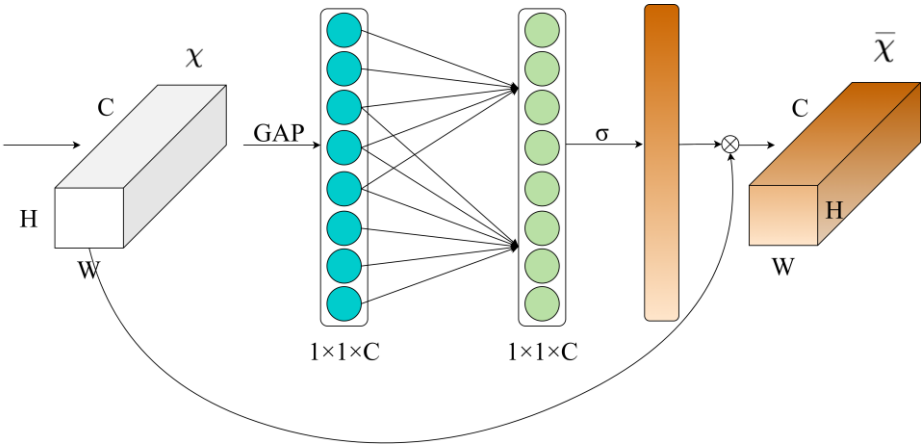


Fig. 3. Structural diagram of the Efficient Channel Attention (ECA) mechanism

$$z_c = \frac{1}{H \cdot W} \sum_{i=1}^{H} \sum_{j=1}^{W} F_{RDB}^{(c,i,j)} \tag{7}$$

The channel descriptor is then treated z as a one-dimensional signal, and a 1D convolution is applied to model local cross-channel interactions. With the kernel size denoted as k(adaptively determined), the corresponding channel-wise attention weights are obtained as $w \in R^{\wedge}C$:

$$w = \sigma\left(Conv1D\left(z\right)\right) \tag{8}$$

here, σ (·) denotes the Sigmoid activation function, which is used to map the attention weights into the range [0,1]. Finally, the original feature map is reweighted channel-wise using the attention weights to obtain the attention-enhanced output F_{ECA}:

$$F_{ECA}^{(c,i,j)} = w_c \cdot F_{RDB}^{(c,i,j)} \tag{9}$$

After the ECA module, the enhanced feature map F_0 is forwarded to the subsequent convolutional layers of the encoder for further compression, and is ultimately mapped to the latent vector z, which is then used by the decoder for image reconstruction.

$$z = Encoder_{late}\left(F_{ECA}\right) \tag{10}$$

During the reconstruction phase, the decoder module performs upsampling on the latent vector z and symmetrically integrates RDB+ECA modules to progressively restore the image $\hat{x}$. Meanwhile, the reconstructed image $\hat{x}$ is re-encoded into a latent representation $\hat{z}$, which is compared with the original latent vector z to enforce semantic consistency in the latent space. This latent alignment term, along with the image reconstruction loss and adversarial loss, is incorporated into the overall training objective.

To jointly optimize reconstruction fidelity, latent consistency, and adversarial realism, we define a composite training objective that combines three loss terms: the adversarial loss guided by the discriminator, the image reconstruction error, and the latent consistency loss. Specifically, the total loss is defined as:

$$\mathcal{L}_G = \lambda_{adv} \cdot \mathcal{L}_{adv} + \lambda_{rec} \cdot \| x\text{-}\hat{x} \|_1 + \lambda_{enc} \cdot \| z\text{-}\hat{z} \|_2^2 \tag{11}$$

This composite loss formulation enables the model to accurately reconstruct normal patterns while amplifying deviations caused by anomalies, laying the foundation for effective score-based decision-making during inference.

3.4 Dynamic Thresholding via F1-Score Optimization

Although a fixed threshold (e.g., 0.20) may yield satisfactory performance on one dataset—for instance, achieving a relatively high F1-score on the SP (bark inclusion) subset—its effectiveness often degrades significantly when applied to other datasets. In our case, applying the same threshold of 0.20 to the QK (chip) dataset leads to a notable drop in F1-score. This performance gap highlights the sensitivity of anomaly detection results to dataset-specific score distributions and supports the need for adaptive thresholding strategies tailored to the statistical characteristics of each defect type. To address this, we adopt a dynamic thresholding strategy based on exhaustive F1-score optimization. After computing anomaly scores on the test set, a brute-force search is performed over a predefined threshold range to select the value that maximizes the F1-score. This approach improves adaptability to dataset-specific characteristics and enhances the model's robustness to data imbalance and distributional shifts.
The optimal decision threshold φ^* is selected by maximizing the F1-score over a predefined set of candidate thresholds Φ, It is formally defined as:

$$\varphi^* = argmax_{\varphi\in\Phi}\ F1(\varphi) \tag{12}$$

Given an input $x_i$and its anomaly score $A(x_i)$, the predicted label $\hat{y}_i$ under threshold ϕ is defined as:

$$\hat{y}_i = \begin{cases} 1, \ if\ A(x_i) > \varphi \\ 0, \ otherwise \end{cases} \tag{13}$$

Precision and recall for a given threshold φ are computed as:

$$Precision(\varphi) = \frac{TP(\varphi)}{TP(\varphi)+FP(\varphi)} \tag{14}$$

$$Recall(\varphi) = \frac{TP(\varphi)}{TP(\varphi)+FN(\varphi)} \tag{15}$$

The F1-score is then calculated by the harmonic mean of precision and recall:

$$F1(\varphi) = \frac{2\cdot Precision(\varphi)\cdot Recall(\varphi)}{Precision(\varphi)+Recall(\varphi)} \tag{16}$$

This dynamic thresholding strategy adaptively aligns the decision boundary with the underlying distribution of anomaly scores, thereby improving robustness and generalization across varying test scenarios.

4 Experiments

4.1 Datasets

The data collection and experimental procedures for this study were conducted in the eucalyptus wood panel sorting workshop located in Linyi, Shandong Province. A representative real-world example of a eucalyptus wood panel is illustrated in **Fig. 4**.

Fig. 4. Real-world sample of a thin eucalyptus wood panel

The image data were captured using a 12-megapixel global shutter camera (Daheng Imaging). The collected image samples are shown in **Fig. 5**.

Fig. 5. Eucalyptus wood dataset

Although the original wood surface defect dataset includes a wide variety of defect types—as shown in **Fig. 6**—such as knots, cracks, chatter marks, burls, bark inclusions, chips, and scabs, this study focuses on three representative and industrially significant categories: bark inclusion, chip, and crack. These types were selected based on their high frequency in actual production, substantial impact on material quality, and the

availability of high-quality, well-labeled image samples, which facilitate the construction of stable training and testing sets. Moreover, these defects exhibit strong visual distinctiveness, including structural anomalies, edge loss, and texture disruptions. This diversity provides a comprehensive foundation for evaluating the robustness and generalization ability of unsupervised anomaly detection models under complex textured backgrounds.

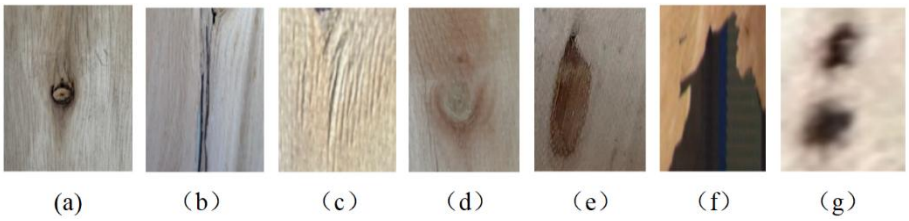


Fig. 6. From left to right : (a)Knot; (b)Crack; (c)Chatter marks; (d)Burl; (e)Bark inclusion;(f)Chip;(g)Scab.

To evaluate the effectiveness of the proposed unsupervised anomaly detection method under real-world industrial conditions, an experimental dataset was constructed using a high-resolution veneer image captured from an actual wood processing environment. The image contains a range of representative structural and textural defects characterized by blurred boundaries, significant morphological variation, and non-uniform spatial distribution—reflecting the real-world complexity of wood surface anomalies.

As shown in **Fig. 7** the raw veneer image was annotated using the Roboflow platform, which combined semi-automated detection with manual correction strategies. Defect regions were precisely marked in the form of bounding boxes. Specifically, red dots represent dead knots, yellow bounding boxes indicate cracks, blue bounding boxes correspond to discolored knots and green bounding boxes denote crack. These annotations enable precise localization and classification of defects in the veneer images, supporting subsequent analysis and model training.

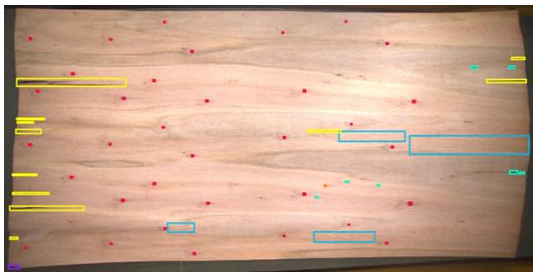
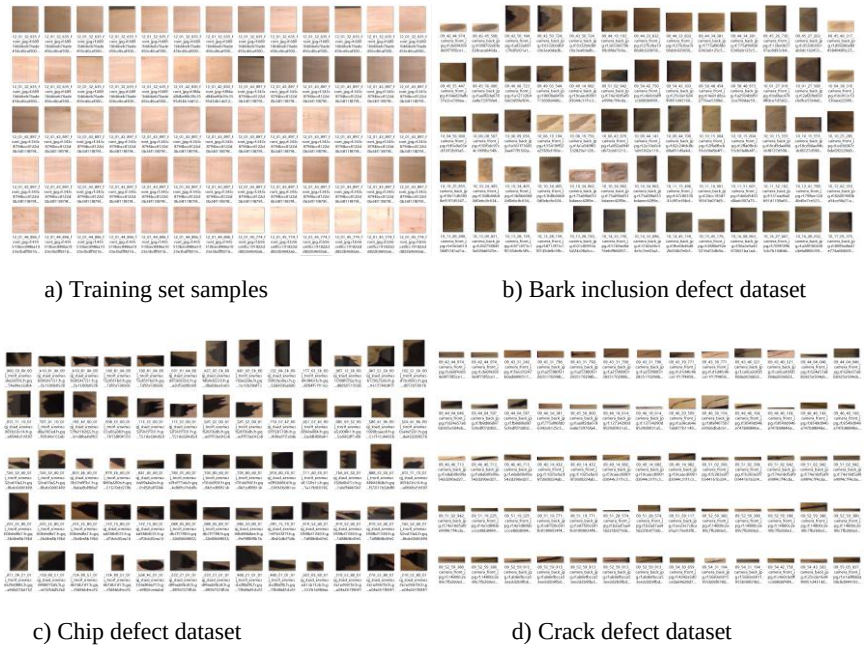


Fig. 7. Wood dataset annotated by humans

Based on the annotated bounding boxes, corresponding anomalous image patches were cropped to construct the defect dataset. In parallel, normal patches were randomly sampled from unmarked areas, ensuring they were free of defects. Representative examples of the resulting dataset samples are presented in **Fig. 8**. The dataset used in this study is available at: https://github.com/GYH663/MyDataset_Wood_AD.



a) Training set samples b) Bark inclusion defect dataset

c) Chip defect dataset d) Crack defect dataset

Fig. 8. Dataset samples

To support fine-grained evaluation, the anomalous image patches were categorized into three typical industrial defect types: bark inclusions (SP), chips (QK), and cracks (LF). As shown in **Fig. 8**, subfigures (b), (c), and (d) illustrate representative samples of each defect category, while (a) displays examples of normal training samples. In total, the final dataset contains 10,000 normal samples for training, and 766 bark inclusion samples, 422 chip samples, and 306 crack samples for testing only, in compliance with the unsupervised setting. This dataset encompasses a diverse range of real-world wood defect patterns and provides a realistic and challenging benchmark for evaluating anomaly detection models under industrial conditions. In accordance with the unsupervised setting, only normal samples were used during training (80%), while the remaining normal samples (20%) and all defect samples were used for evaluation.

4.2 Implementation Details

All experiments are conducted on a high-performance computing server equipped with an NVIDIA A100 GPU, ensuring sufficient computational capacity for large-scale

training. The deep learning environment is implemented using PyTorch (v2.3.1) and Torchvision (v0.4), providing a flexible and stable platform for model development and experimentation. To ensure effective adversarial training, the model is optimized using the Adam optimizer with an initial learning rate of 0.0002 and momentum parameters $\beta_1 = 0.5$ and $\beta_2 = 0.999$. These settings are known to enhance convergence stability and prevent mode collapse. All input images are resized to 256×256 pixels using bilinear interpolation to unify spatial dimensions across the dataset and facilitate efficient batch processing. The latent space dimensionality is set to 512, providing a sufficiently expressive embedding space for capturing complex surface features, while the batch size is fixed at 32 to strike a balance between computational efficiency and model generalization.

4.3 Comparative Experiments

Baseline Method. To comprehensively evaluate the performance of the proposed model in wood surface defect detection, three representative unsupervised anomaly detection methods we reselected as baseline models: EGBAD[17], AnoGAN[14], and GANomaly[3]. These methods differ in network architecture, latent space modeling capability, and anomaly scoring strategies, and together represent the mainstream paradigms in generative unsupervised anomaly detection. For fair and consistent comparison, all baseline models were re-implemented based on their official open-source code and evaluated on the same dataset proposed in this study.

Evaluation metrics. To comprehensively evaluate the performance of the proposed method in anomaly detection tasks, four commonly used evaluation metrics are adopted: Accuracy, Precision, Recall, and F1-score. These metrics are defined based on the confusion matrix, which includes True Positives (TP), True Negatives (TN), False Positives (FP), and False Negatives (FN). The definitions of the four metrics are as follows:

$$Pre = \frac{TP}{TP + FP} \tag{17}$$

$$Rec = \frac{TP}{TP + FN} \tag{18}$$

$$F1 = \frac{2 \cdot Pre \cdot Rec}{Pre \ + \ Rec} \tag{19}$$

$$Acc = \frac{TP \ + \ TN}{TP \ + \ TN \ + \ FP \ + \ FN} \tag{20}$$

Threshold Selection Strategy. To ensure the practicality and consistency of anomaly detection under real-world industrial conditions—where a single decision threshold is typically required—this study employs an enumerative threshold search strategy to determine the optimal classification boundary. Specifically, the reconstruction error

scores in the latent space for all samples in the validation set are first linearly normalized to the range [0,1], in order to eliminate scale differences across subsets. Then, a set of candidate thresholds is generated by uniformly sampling the interval with a step size of 0.01. For each candidate threshold, the corresponding precision, recall, and F1-score are computed based on validation performance. The optimal threshold φ^* is selected as the one that maximizes the F1-score. In cases where multiple thresholds yield the same maximum F1-score, the one with the higher precision is preferred, to reduce false positives that may lead to unnecessary interruptions or material waste in industrial applications. During the testing phase, this selected threshold φ^* is fixed and applied to all test samples for binary classification. This process establishes a unified and fair evaluation protocol and ensures that performance assessment aligns with operational requirements in practical deployments.

4.4 Results and Discussion

Table 1 summarizes the performance comparison between the proposed model and three representative unsupervised anomaly detection baselines—EGBAD, AnoGAN, and GANomaly—on SP, QK, and LF wood defect subsets, using Accuracy, F1-score, Recall, and Precision as evaluation metrics.

Overall, the proposed model achieves superior performance across all datasets, with especially notable improvements in Accuracy and F1-score. On the SP dataset, it attains the highest Accuracy (0.924) and F1-score (0.830), demonstrating robustness against blurred boundaries and noisy textures. Although GANomaly shows slightly higher Precision (0.840), our model yields better Recall (0.861), which is preferable in industrial scenarios where missing defects is more critical than false alarms. For the QK dataset, despite GANomaly achieving a higher Recall (0.805), our model achieves the best F1-score (0.793) and Accuracy (0.910), showing a better balance between sensitivity and precision. On the challenging LF dataset, it outperforms all baselines with an Accuracy of 0.903, Recall of 0.832, and F1-score of 0.787.

These results confirm the proposed model's robustness and generalization across diverse wood surface conditions, highlighting its potential for practical industrial deployment. Moreover, across all three subsets, the proposed model maintains high detection performance while effectively controlling the false positive rate, further demonstrating its practicality and stability in complex detection environments.

4.5 Ablation Study

To evaluate the contributions of the Efficient Channel Attention (ECA) and Residual Dense Block (RDB) modules, we conducted ablation experiments on three wood defect datasets: SP, QK, and LF. As shown in **Table 2**, the complete model configuration (✓ECA, ✓RDB) consistently achieves the best performance across all datasets in terms of F1-score and Precision, with competitive or near-optimal Recall. This confirms that both modules are instrumental in improving anomaly detection performance.

Table 1. Performance comparison of different methods on wood defect datasets.

Dataset	Model	Acc	F1	Rec	Pre
SP	EGBAD	0.715	0.771	0.753	0.805
	AnoGAN	0.863	0.693	0.750	0.658
	Ganomaly	0.875	0.764	0.799	**0.840**
	Our Model	**0.924**	**0.830**	**0.861**	0.801
QK	EGBAD	0.701	0.688	0.694	0.682
	AnoGAN	0.842	0.702	0.630	**0.790**
	Ganomaly	0.859	0.752	0.805	0.712
	Our Model	**0.910**	**0.793**	**0.830**	0.763
LF	EGBAD	0.721	0.694	0.670	0.726
	AnoGAN	0.829	0.685	0.651	0.720
	Ganomaly	0.843	0.690	0.720	0.665
	Our Model	**0.903**	**0.787**	**0.832**	**0.752**

On the SP dataset, incorporating RDB alone (✗ECA, ✓RDB) improves the Recall from 0.799 to 0.865 and the F1-score from 0.764 to 0.780, indicating that RDB enhances the model's ability to capture multi-scale structural patterns. When ECA is further added, the F1-score rises to 0.830 and Precision reaches 0.801, demonstrating improved channel-wise feature discrimination. A similar trend is observed on the QK dataset. The base model (✗ECA, ✗RDB) yields the lowest F1-score (0.752), while the full model increases it to 0.793 with a Precision of 0.763. These improvements validate the complementary benefits of residual aggregation and adaptive attention in capturing subtle defects.

On the more challenging LF dataset, the base model performs the worst (F1-score: 0.690), while introducing RDB alone improves it to 0.735. With both modules enabled, the model achieves the highest Recall (0.832) and F1-score (0.787). Notably, in some cases, the Recall of the base model slightly exceeds that of the full model (e.g., SP and QK). This can be attributed to the fact that, without attention or residual constraints, the model tends to over-score ambiguous regions, leading to higher Recall but at the cost of increased false positives. In contrast, the full model, with refined feature selectivity, yields better precision and overall balance.

These results collectively demonstrate that RDB significantly strengthens hierarchical feature representation, while ECA further improves localization accuracy by adaptively highlighting informative channels. The integration of both modules contributes to robust and precise anomaly detection, particularly under complex texture and noise conditions prevalent in wood surface inspection.

5 Conclusion

In this work, we propose ERA-GANomaly, an unsupervised anomaly detection framework tailored for wood surface defect inspection in industrial settings. By integrating Residual Dense Blocks (RDBs) and Efficient Channel Attention (ECA) into a GAN-

Table 2. Ablation study of ECA and RDB modules on wood surface defect datasets (SP, QK, LF). Best results in each column are bolded.

Dataset	ECA	RDB	Rec	F1	Pre
SP	✗	✗	0.799	0.764	0.840
	✗	✓	**0.865**	0.780	0.728
	✓	✓	0.861	**0.830**	**0.801**
QK	✗	✗	0.805	0.752	0.712
	✗	✓	**0.865**	0.780	0.728
	✓	✓	0.830	**0.793**	**0.763**
LF	✗	✗	0.720	0.690	0.665
	✗	✓	0.790	0.735	0.695
	✓	✓	**0.832**	**0.787**	**0.752**

based encoder–decoder–re-encoder architecture, the model enhances fine-grained feature extraction and inter-channel discrimination, enabling accurate detection of complex and subtle surface anomalies without requiring labeled abnormal data. Extensive experiments on three real-world wood defect datasets (SP, QK, LF) demonstrate that our method consistently outperforms existing unsupervised baselines such as EGBAD, AnoGAN, and GANomaly across multiple evaluation metrics, including Accuracy, Precision, and F1-score. Notably, the model achieves a macro-F1 score of 0.952 and a Precision of 0.970 on the QK dataset, reflecting strong generalization and robustness. Ablation studies further confirm the complementary contributions of RDB and ECA in improving detection precision and localization sensitivity. Overall, ERA-GANomaly provides a practical, interpretable, and scalable solution for industrial anomaly detection, with promising potential for extension to temporal monitoring and broader application in other material domains.

Acknowledgments. This study was supported by the Key R&D Program of Shandong Province, China (2024CXGC010905), the CentralGuidance Local Science and Technology Development FundProject (YDZX2024121), the CentralGuidance Local Science and Technology Development FundProject (YDZX2023050), and National Natural Science Foundation of China (No. 42201458).

References

1. Xia, X., Pan, X., Li, N., He, X., Ma, L., Zhang, X., Ding, N.: GAN-based anomaly detection: A review. Neurocomputing 493, 497–535 (2022)
2. Ruff, L., Vandermeulen, R., Goernitz, N., Deecke, L., Siddiqui, S.A., Binder, A., Müller, E., Kloft, M.: Deep one-class classification. In: Proceedings of the International Conference on Machine Learning, pp. 4393–4402. PMLR, 2018.
3. Akcay, S., Atapour-Abarghouei, A., Breckon, T.P.: Ganomaly: Semi-supervised anomaly detection via adversarial training. In: Proceedings of the Asian Conference on Computer Vision, pp. 622–637. Springer, 2018.
4. Singh, A., Reddy, P.: AnoGAN for Tabular Data: A Novel Approach to Anomaly Detection. arXiv preprint arXiv:2405.03075 (2024)

5. Sampath, V., Maurtua, I., Martín, J.J.A., Rivera, A., Molina, J., Gutierrez, A.: Attention-guided multitask learning for surface defect identification. IEEE Trans. Ind. Inform. 19(9), 9713–9721 (2023)
6. Chen, C.-Y., Chang, S.-C., Liao, D.-Y.: Equipment anomaly detection for semiconductor manufacturing by exploiting unsupervised learning from sensory data. Sensors 20(19), 5650 (2020)
7. Chen, Y., Sun, C., Ren, Z., Na, B.: Review of the current state of application of wood defect recognition technology. BioResources 18(1) (2023)
8. Perera, P., Patel, V.M.: Learning deep features for one-class classification. IEEE Trans. Image Process. 28(11), 5450–5463 (2019)
9. Li, Y., et al.: Dual-discriminator GAN for wood defect detection. Comput. Ind. (2023)
10. Lee, C.-K., Cheon, Y.-J., Hwang, W.-Y.: Studies on the GAN-based anomaly detection methods for the time series data. IEEE Access 9, 73201–73215 (2021)
11. Gong, D., et al.: Memorization normalization for anomaly detection. In: CVPR, 2022.
12. Zhang, Z., et al.: Shallow autoencoders for plywood crack detection. Wood Research, 2020.
13. Zhou, T., et al.: Multi-scale reconstruction for wood defect localization. IEEE Trans. Ind. Inform. (2023)
14. Schlegl, T., et al.: f-AnoGAN for industrial anomaly detection. Med. Image Anal. 54, 30–44 (2019)
15. Schlegl, T., Seeböck, P., Waldstein, S. M., Langs, G., Schmidt-Erfurth, U.: f-AnoGAN: Fast unsupervised anomaly detection with generative adversarial networks. Med. Image Anal. 54, 30–44 (2019)
16. Dong, F., Zhang, Y., Nie, X.: Dual discriminator generative adversarial network for video anomaly detection. IEEE Access 8, 88170–88176 (2020)
17. Zenati, H., Foo, C.S., Lecouat, B., Manek, G., Chandrasekhar, V.R.: Efficient GAN-based anomaly detection. arXiv preprint arXiv:1802.06222 (2018).
18. Wang, Q., Wu, B., Zhu, P., Li, P., Zuo, W., Hu, Q.: ECA-Net: Efficient channel attention for deep convolutional neural networks. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 11534–11542 (2020)
19. Shao, Y., Mao, L., Ye, L., Li, J., Yang, P., Ji, C., Wu, Z.: H2GCN: A hybrid hypergraph convolution network for skeleton-based action recognition. Journal of King Saud University - Computer and Information Sciences 36(5), 102072 (2024)
20. Bi, X., Zhou, W., Luo, S., Mao, Y., Hu, X., Zeng, B., Xu, L.: Feature aggregation graph convolutional network based on imaging genetic data for diagnosis and pathogeny identification of Alzheimer's disease. Briefings in Bioinformatics 23(3), bbac137 (2022)