International Journal of Engineering Trends and Technology
ISSN: 2231–5381 / https://doi.org/10.14445/22315381/IJETT-V73I12P116

Volume 73 Issue 12, 194-206, December 2025

Original Article

Evaluating Denoising Models for Feature Enhancement and Improved SVM-Based Classification of Locally Made Earthen Ceramic Pots

Aljon L. Abines[1], Aimee D. Molato[2]

[1,2]*College of Computing Studies, Information Communication Technology, Isabela State University - Cuayan Campus, Cauayan, Isabela City, Philippines.*

[1]*Corresponding Author: aljon.abines_cyn@isu.edu.ph*

Received: 25 July 2025 Revised: 15 November 2025 Accepted: 25 November 2025 Published: 19 December 2025

Abstract *- Manual inspection of locally made earthen ceramic pots suffers from inconsistency and subjectivity, creating problems for quality control in traditional pottery production. This research examines how denoising affects feature extraction in SVM-based ceramic pot classification. The study compares three deep learning denoising architectures: the Denoising Autoencoder with Convolutional Autoencoder (DAE-CAE), Denoising Convolutional Neural Network (DnCNN), and a Generative Adversarial Network (GAN). PSNR, SSIM, and RMSE as metrics were used for performance evaluation. Results show that the DAE-CAE Model outperforms the other architectures, achieving a PSNR of 23.2087, SSIM of 0.4828, and RMSE of 0.0713, while DnCNN reaches 23.0786 PSNR, 0.4742 SSIM, and 0.0725 RMSE, and GAN achieves 23.1815 PSNR, 0.4784 SSIM, and 0.0719 RMSE. Features extracted from DAE-CAE-denoised images are used to train classifiers, and SVM achieves 93.23% accuracy. This exceeds both Random Forest at 90.73% and CNN at 90.20%. The results indicate that denoising improves classification generalizability, precision, and reliability. The DAE-CAE-enhanced SVM framework proves most effective for this task. Combining deep learning denoising with SVM provides a practical automated alternative to manual inspection, offering both improved accuracy and potential for scaling quality assessment in traditional ceramic production.*

Keywords - *Deep Learning, Earthen ceramic pot classification, Feature enhancement, Image denoising models, Support Vector Machine.*

1. Introduction

Earthen ceramic pots are among the oldest and most common forms of material culture and decorative art. These earthen ceramic pots remain popular, but the natural variation in hand-crafted production means quality assessment is always necessary. Manual inspection is still the standard approach, though it takes considerable time and produces inconsistent results. Machine learning and artificial intelligence now enable automated visual inspection, improving both speed and accuracy in ceramic quality evaluation [5, 6, 9]. The Support Vector Machines (SVMs) have proven effective at classifying the complex textures and patterns found in earthen ceramic pot materials [4, 21].

Previous studies show that computer vision methods can classify or detect defects in ceramic products. Artificial vision systems using SVM and K-Nearest Neighbor (KNN) have delivered promising results with ceramic tiles [14, 19]. Variational Autoencoders (VAEs) have clustered Roman potsherds in an unsupervised manner [21]. Bag-of-Features pipelines have identified defects and irregularities in ceramic images [3], and real-time detection systems based on deep learning have highlighted the value of AI for quality assessment [6, 9, 20, 32]. Convolutional Neural Network (CNN) classifiers for tiles have also achieved strong accuracy across multiple defect types, although many rely on large pretrained models and omit feature-reduction or denoising stages, increasing the risk of overfitting and limited domain transfer [26].

Despite these advances, a key practical challenge remains underexplored: locally crafted earthen pots exhibit substantial variation in materials, firing conditions, and craftsmanship. These factors introduce noise and textural irregularities that degrade feature quality and reduce the consistency of downstream classification. While denoising has been proposed as a pre-processing step [25], the comparative effectiveness of different denoising architectures, evaluated specifically for improving SVM-based classification of artisanal, locally produced pottery, has not been systematically examined. Research gap: existing work evaluates ceramic classification and defect detection broadly,

but does not establish which denoising models best enhance features for an SVM classifier in the context of locally made earthen pots with high intra-class variability. Beyond accuracy, recent work urges reporting computational budgets and efficiency metrics to support practical, sustainable deployment [31].

This study addresses that gap by comparing three widely used denoising models, such as Denoising Autoencoder with Convolutional Autoencoder (DAE-CAE) [28], Denoising CNN (DnCNN) [26], and a Generative Adversarial Network (GAN)-based denoiser [30], as feature-enhancement stages prior to SVM classification of locally made earthen ceramic pots. The focus is on methods that balance effectiveness, efficiency, and deployability in resource-constrained settings deployability in resource-constrained settings [31] typical of local pottery industries typical of local pottery industries. Contemporary models, such as SwinIR, can yield high reconstruction quality; however, their computational demands and longer training times make them less suitable for practical deployment in these environments [15].

Methodologically, the denoising fidelity using Peak Signal-to-Noise Ratio (PSNR) [23], Structural Similarity Index (SSIM) [18, 22], and Root Mean Squared Error (RMSE) [8], will be assessed first [18, 24]. Then extract features from the best performing denoised images and will be trained by SVM to classify pots into standard and substandard quality. To measure the model's effectiveness, accuracy, and confusion matrix analysis are utilized. The two-stage designs allow the study to quantify how each denoiser preserves informative structure while suppressing noise in earthen ceramic pots, and finally, determine whether those improvements translate into more consistent SVM classification of artisanal earthen ceramic pottery images.

The research makes three essential contributions. Provides a direct comparison of the three models: the DAE-CAE, the DnCNN, and the GAN denoisers, specifically for SVM classification of locally made earthen ceramic pots. The application includes a non-uniform production and complex textures from industrial tile datasets that are usually utilized in prior work [14, 19, 26]. It connects reconstruction metrices like PSNR, SSIM, and RMSE in classifying outcomes of the different denoising, and it clarifies which denoising characteristics matter most for reliable, quality decisions and assessment. Then, it prioritizes a feasible deployment by focusing on models and settings that operate under modest hardware constraints, typical of local potteries.

The traditional earthen pottery remains important not only for cooking but also for cultural practices in many communities [10]. By integrating denoising that enhances the feature preparation with SVM classification, which is tailored to this context, the study pursues a practical framework for a consistent and automated quality assessment. This approach reduces reliance on manual inspection, thereby aligning it with the operational realities of potter's production and, in return, supporting not only cultural heritage practices but also their quality assurance, whether in small or large-scale manufacturing.

1.1. Objective of the Study

This study aims to assess the effectiveness of denoising techniques in improving feature extraction for SVM classification of locally made earthen ceramic pots. This study seeks specifically to:

- Evaluate three different denoising models, Denoising Autoencoder with Convolutional Autoencoder (DAE-CAE), Denoising Convolutional Neural Network (DnCNN), and Generative Adversarial Network (GAN), using PSNR, SSIM, and RMSE.
- Determine the best-performing denoising model that preserves crucial image details while minimizing noise.
- Use the extracted features from the best model to train SVM, Random Forest, and CNN classifiers, and identify the most effective classification model relative to accuracy, precision, recall, and F1-score.

2. Related Work

In replacing traditional manual inspection with automated approaches, machine learning has been increasingly applied in the classification of ceramics. Highlighting the role of artificial intelligence in quality assessment, Vision-based platforms using algorithms like Support Vector Machines (SVM) and K-Nearest Neighbors (KNN) have proven effective in identifying defects in ceramic tiles [21]. Also, [9] proposed an automated defective ceramic tile classification using an image processing technique, which equates to a promising result in industrial sectors. The Variational Autoencoders (VAEs), which are an unsupervised clustering technique, have been applied to archaeological ceramics like Roman potsherds, showing the potential of deep learning in analyzing pottery fragments [19].

There are different methods used in this study, where anomaly detection methods, including Bag of Features, have been introduced for defect identification, which confirms the viability of an automated system for ceramic quality assessment [32]. It also presented a real-time system that automatically detects defects in a ceramic tile using deep learning [6]. To emphasize the growing importance of artificial intelligence in ceramic quality assessment [5] and demonstrate an embedded digital quality control of intelligent ceramic-tile manufacturing [5]. All these reported a significant improvement in ceramic production efficiency. However, these methods focus only on industrial ceramics with controlled production environments and yet fail to address an end-to-end implementation framework suited to a specialized domain such as ceramic assessment. While this optimization improved the classification metrics, it did not

provide insights into memory consumption, which limits scalability in a resource-constrained setting [4, 16]. Furthermore, most of the studies focus on general-purpose industrial ceramics and do not explore classification challenges in handcrafted pottery materials in an indigenous context. Parallel to that, there is a growing focus on computing efficiency alongside its accuracy in real-world readiness [31] and on using standard mobile inference indicators [32].

Significantly, the quality of extracted figures influences their classification model performance across diverse machine learning applications. The conventional feature engineers rely on manual design extraction methods that leverage domain expertise and identify discriminative patterns, particularly for defective detection tasks. In contrast, the deep convolutional neural networks have demonstrated the capability for autonomous feature learning and eliminate the need for manual feature design [13]. These achieve superior classification performance in a hierarchical representation learning that is applied to visually informed data.

Despite the effectiveness of these methods, deep learning approaches often require substantial computational resources and may suffer from overfitting in training the data, which are limited to challenges particularly relevant for a resource-constrained deployment scheme. Thus, to address data scarcity, the augmentation strategies have proven their importance in enhancing model generalization using stratified sampling techniques, which ensure balanced representation during their training [1, 25, 29].

In the SVM-based classification, selection, and quality of input features directly impact decision boundary effectiveness [16]. While Nalipa and Kawulok [16] emphasize that appropriate training set selection and feature quality are critical determinants of the SVM performance. However, the existing literature predominantly focuses on feature extraction and selection methodologies, but it pays limited attention to pre-processing strategies, which only a denoising model can enhance feature discriminability.

The gap is particularly relevant for a traditional ceramic quality assessment, where noise from varied capture conditions degraded its feature quality and subsequent classification reliability, in integrating denoising-based feature enhancement with traditional machine learning classifiers that represent an underexplored approach which may offer practical advantages for pottery production. The computational constraints and limited training data present unique challenges not sufficiently addressed by pure deep-learning solutions.

The techniques to address noise in image processing remain a critical challenge in most machine learning applications [27]. The study of stacked denoising autoencoders [28], demonstrating its effectiveness in learning useful representations through a local denoising criterion that prevents co-adaptation of features and improving generalization across various tasks [28], provided a comprehensive survey on image denoising [27]. Upon using deep learning, it highlighted the evolution of denoising architectures. Both the Denoising Convolutional Neural Networks (DnCNN) and Generative Adversarial Networks (GANs) employ residual learning against complex noise patterns in various image restoration tasks and exploring image denoising with optimization algorithms [8, 30]. The introduction of an adaptive non-local mean filter in eliminating salt-and-pepper noise [32] addresses specific noise patterns commonly found in an image processing model. Recently, transformer-based models like SwinIR [15] outperform conventional architectures in restoration tasks and require large datasets with higher computational resources. This also limits the practical deployment in resource-constrained industries. Finally, in terms of classification, the SVMs remain a reliable choice since they are readily available for separating complex boundaries and handling small datasets effectively [4]. It also provides a comprehensive survey on SVM classification across various domains while emphasizing its accuracy in handling high-dimensional feature spaces.

In a comprehensive performance evaluation, the multiple metrics beyond its basic accuracy are essential, with a provision on practical guidance for evaluation metrics offering insights into PSNR, SSIM, and image quality assessment. However, despite the advances in denoising and classification, independently limited work has explored its integration in the context of locally made ceramic pots [18, 22].

While denoising techniques can mitigate image degradation and feature inconsistency when applied during preprocessing stages, most denoising studies focus on general-purpose datasets and do not explore denoising's interaction with feature extraction in traditional pottery or handcrafted ceramic contexts. Moreover, while SVMs have been successfully applied to industrial ceramic classification, their performance with denoised features from traditional pottery remains unexplored. Therefore, the systematic evaluation of denoising models as preprocessing steps for SVM-based classification in artisanal ceramics represents a significant research opportunity, which this study uniquely addresses.

3. Materials and Methods

This study takes a systematic approach by applying three denoising models, such as DAE, DnCNN, and GAN, to determine which model best improves feature extraction that enhances the SVM classification of earthen ceramic pots [14]. The diagram in Figure 1 shows a process for classifying ceramic pots. Standard and substandard images are augmented, then denoised using DAE-CAE, DnCNN, or GAN models. The best denoising model is chosen based on image quality metrics (PSNR, SSIM, RMSE). Then, Features are

extracted and fed into SVM, Random Forest, and CNN classifiers. The results are compared using classification metrics to identify the most accurate model for quality assessment.

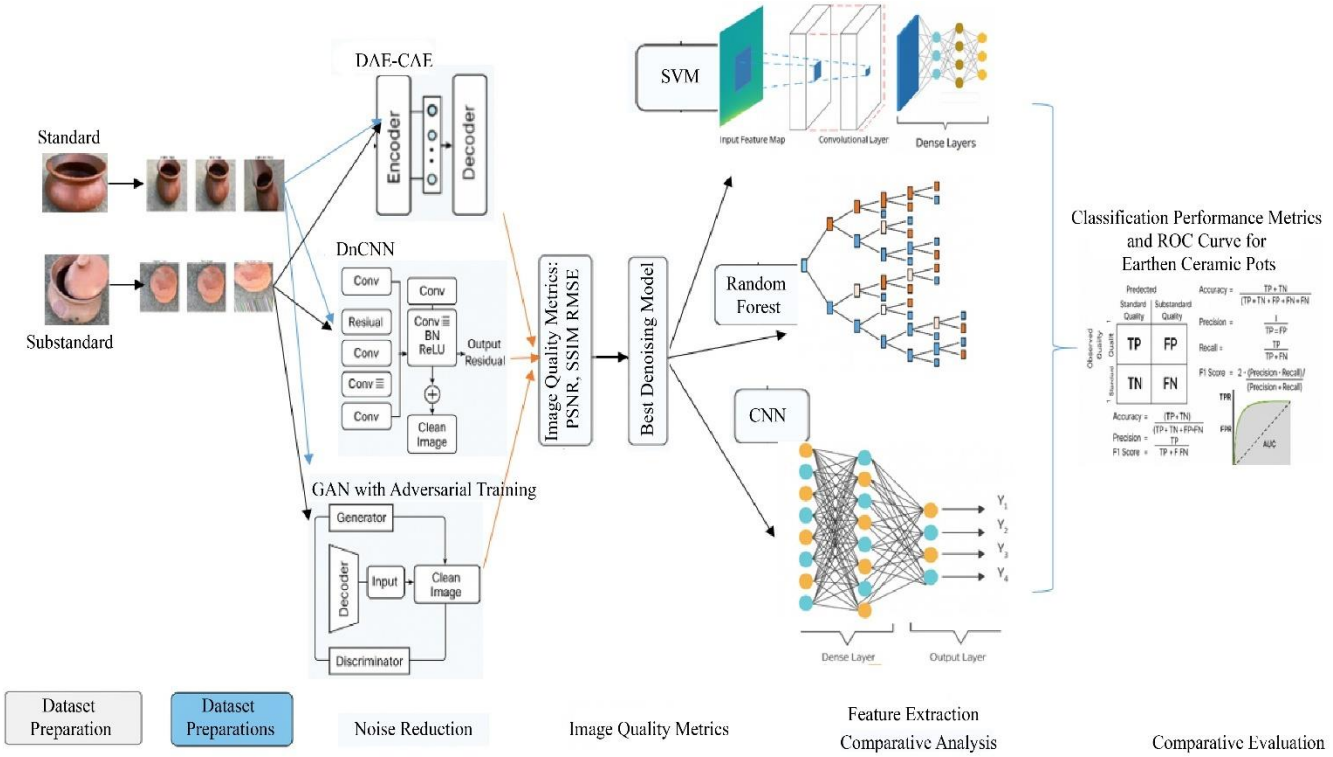

Fig. 1 Proposed pipeline for noise reduction, feature extraction, and classification of earthen ceramic pots

3.1. Data Preparation

The dataset includes pictures of locally produced earthen ceramic pots that have been categorized into two quality classes: Substandard (500 photos; faulty/inferior) and Standard (507 photos; good quality). To maintain class proportions in both categories, a stratified 80/20 split was used, allowing for objective assessment and shielding the measures from imbalance effects [10, 29] a stratified 80/20 split. For training, this produced 406 Standard images and 400 Substandard images, and for testing, it produced 101 Standard images and 100 Substandard images.

The images were downsized to 224 x 224 pixels and then transformed into tensors and normalized to a [0,1] consistency and stable optimization. It still maintains a standard input scale that improves numerical stability, reduces gradient issues associated with raw pixel ranges, and accelerates convergence during training [13]. To guarantee reproducibility across runs, a fixed random seed was employed for the split. At this point, no more augmentations were used.

3.2. Data Augmentation and Noise Injection

To strengthen generalization and match real workshop variability, we applied structured augmentation [2] to the training split, only leaving the test set untouched, including controlled rotations, horizontal flips, rescaling, brightness adjustments, and slight viewpoint shifts to mimic pose, distance, lighting, and framing changes typical in pottery settings [17].

Then, they performed controlled noise injection to create paired data for denoising, treating each augmented image as the clean reference and generating a degraded counterpart with (a) additive Gaussian noise to approximate sensor/thermal and electronic readout noise and (b) salt-and-pepper noise to model impulse artifacts from transmission errors or stuck/dead pixels [25]. This sequence augments first, then injects noise, encouraging robust reconstruction over memorization, and expands the effective training set [16] for optimization.

It strictly preserves the original test split to prevent leakage and ensure unbiased evaluation. All transforms were executed with fixed random seeds and bounded parameters for reproducibility across runs.

3.3. Model Training

Three distinct denoising models were trained: Denoising Autoencoder with Convolutional Autoencoder (DAE-CAE), Denoising Convolutional Neural Network (DnCNN), and Generative Adversarial Network (GAN)-based generator were developed and trained to reconstruct clean ceramic images from their noise-corrupted counterparts.

The training optimization employed the Adam algorithm, an adaptive learning rate method recognized for computational efficiency and robust performance when handling noisy gradients and sparse data distributions [7]. The Adam optimizer updates model parameters θ based on adaptive moment estimates according to (1):

$$\theta_t = \theta_{\{t-1\}} - \alpha \backslash frac\{\{m\}_t\}\left\{\sqrt{\{\{v\}_t\}} + \epsilon\right\} \quad (1)$$

θ_t represents the parameters at step t, with α the learning rate and $\hat{m}_t$ and $\hat{v}_t$ the bias-corrected first and second moment estimates. ε is included to prevent division by very small values. The model was trained using Mean Squared Error (MSE) as the loss, defined in (2).

$$MSE = frac1n \sum_{i=1}^{n} = (y_i - \hat{y_i})^2 \quad (2)$$

Where n is the total number of pixels, y_i is the pixel values of the original clean image, and Þ$_i$ is the pixel values that were reconstructed. Greater sensitivity to big mistakes is a beneficial property for enhancing image quality in denoising tasks, as indicated by the MSE [8], which calculates the average squared discrepancies between actual and predicted values [12]. There was no improvement in validation loss during the 150 epochs of training with early stopping set up to cease optimization after 20 consecutive epochs, which helps avoid overfitting.

In order to maintain steady convergence and prevent poor local minima during optimization, the ReduceLROnPlateau scheduler was utilized to lower the learning rate when validation performance ceased to improve [7]. A fair comparison based on architectural differences rather than variances in training processes is made possible by this training methodology, which guarantees that all three denoising architectures go through the same optimization process.

3.4. Model Evaluation

To thoroughly analyze the trained models' image reconstruction ability, three essential quantitative metrics were used: Root Mean Squared Error (RMSE) [8], Structural Similarity Index (SSIM) [18, 22], and Peak Signal-to-Noise Ratio (PSNR) [24]. Better image quality is indicated by higher PSNR values, which calculate the ratio between the greatest potential signal value and the image's noise level. It has the definition found in (3).

$$PSNR = 10 \cdot log10(MSEMAXI2) \quad (3)$$

Here, $MAXI$ denotes the maximum possible pixel intensity (e.g., 255 for 8-bit images), and MSE is the mean squared error. The SSIM [22] evaluates structural and perceptual similarity between the original and reconstructed images by comparing luminance, contrast, and structure. Its formula is given in (4).

$$SSIM(x, y) = (\mu x2 + \mu y2 + C1)(\sigma x2 + \sigma y2 + C2)(2\mu x\mu y + C1)(2\sigma xy + C2) \quad (4)$$

Where μx and μy are the mean pixel values of the original and reconstructed images, and are their variances, σxy is the covariance between the images, and $C1$and $C2$ are small constants to prevent instability in low-contrast areas. Higher SSIM values (closer to 1) indicate superior structural and perceptual quality [18, 28]. The RMSE measures the average magnitude of the reconstruction error by calculating the square root of the mean squared differences between predicted and actual pixel values. It is defined as in (5).

$$RMSE = \sqrt{frac1n \sum_{i=1}^{n}(y_i - \hat{y_i})^2} \quad (5)$$

Where y_i is the actual pixel value, y_i is the predicted pixel value, and n is the total number of pixels. Lower RMSE values signify improved reconstruction accuracy [8].

These metrics are used to assess the visual and structural quality of reconstructed images by comparing them to original, clean versions of the earthen ceramic pot. The Higher PSNR and SSIM scores, combined with lower RMSE values, indicate a better performance and effective noise suppression while preserving key image features. This evaluation identifies the best-performing model for further analysis [25].

3.5. Computational Efficiency & Benchmarking

The study prioritizes efficiency and reports compute-related metrics [23] while the latency was measured as single-image inference time or batch = 1 with warm-up runs and fixed input size, following mobile/edge benchmarking conventions [11]. Metrics include inference time, the model size, and parameter count.

3.6. Brief Comparison with Recent Denoising Architectures or Model Selection

This study specifically focuses on three widely recognized denoising models that will undergo a comparative evaluation these are Denoising Autoencoder with Convolutional Autoencoder (DAE-CAE), Denoising Convolutional Neural Network (DnCNN), and Generative Adversarial Network (GAN) based approaches due to their proven effectiveness, interpretability and relevance in practical image denoising tasks [25] focusing on their balance between efficiency, performance and the ease of integration into a classification pipeline or architecture. The models representing a diverse architectural paradigm, such as DAE-CAE, leverage unsupervised representation learning, while DnCNN applies a deep residual learning with a convolutional layer. The GANs introduce adversarial training to improve perceptual quality. Recently, transformer-based models such as SwinIR have demonstrated excellent performance in image restoration tasks. Resulting in a significantly larger dataset, a higher computational resource, and longer training times [15].

Meanwhile, the selected denoising models in this study offer a more computationally efficient and accessible solution for the classification, making it more suitable for real-world applications with limited resources.

3.7. Comparative Evaluation of Classification Models Using Extracted Features

After knowing the appropriate denoising model using evaluation metrics, a feature extraction is performed on both training and testing datasets using batch processing for computational efficiency. These features serve as inputs in training an SVM classifier with a linear kernel and are chosen because they perform well in high-dimensional feature spaces, even with a limited training sample. Beyond these matrices, several metrics are computed to assess each classifier's performance in detail. These metrics include precision, recall, and F1-score, which demonstrate how well the classifier manages false positive and false negative predictions.

These metrics are crucial in ceramic quality assessment, where errors can result in actual material and financial costs. The same characteristics from the top denoising model will be used to train two more models, Random Forest and Convolutional Neural Network (CNN) [26], in order to evaluate the SVM methodology and assess how well it performs in comparison to alternative classification techniques. This comparison will show whether traditional machine learning or deep learning methods work better with the improved features from denoising.

4. Results and Discussion

This section presents the findings for three denoising models these are the Denoising Autoencoder with Convolutional layers or the DAE-CAE, Deep Convolutional Neural Network as DnCNN, and a Generative Adversarial Network the GAN with emphasis on their noise-reduction effectiveness and ability to preserve structural details in earthen ceramic pots images their performance was measured using Peak Signal-to-Noise Ratio (PSNR), Structural Similarity Index (SSIM), and Root Mean Squared Error (RMSE) to determine the best preprocessing approach for classification. As presented in Table 1, the dataset encompasses 507 images representing standard-quality locally made earthen ceramic pots and 500 images depicting substandard specimens, creating a nearly balanced binary classification problem. To improve model robustness against variations in pottery craftsmanship, lighting conditions, and surface textures inherent in traditional ceramic production, data augmentation techniques were applied. An augmentation generated 1,000 synthetic variations per class through rotation, scaling, horizontal flipping, and brightness adjustment, yielding 2,768 total images. The dataset is split by 80/20, for training 80 and 20 for testing, using stratified sampling to preserve class distributions, resulting in 2,214 training images and 554 testing images. The augmentation strategy addresses the limited labelled data common in a specialized domain like pottery while simulating the natural variability in real-world ceramic inspection.

Table 1. Dataset distribution summary

Datasets Name	Number of datasets
Standard	507
Substandard	500
Augmented (num_aug=1000)	2768
Training dataset size	2214
Test dataset size	554

4.1. Visualization of Augmented Images and Noise Injection

The visual representations, as shown in Figures 2 and 3, present representative images from both standard and substandard quality categories following augmentation and controlled noise injection. The visualizations demonstrate the application of Gaussian noise and salt-and-pepper noise to simulate real-world image degradation sources such as sensor noise, compression artifacts, and environmental interference during image capture in pottery workshops. These pre-processed images reveal how augmentation preserves class-distinguishing characteristics such as surface uniformity, glaze consistency, and structural integrity while introducing variability that prevents model overfitting to specific imaging conditions. Noise injection creates a challenging learning environment that prevents models from simply memorizing clean image patterns, but instead forces them to develop robust reconstruction capabilities for the different images. Analysing these modified images reveals how each denoising architecture handles different corruption types and recovers diagnostic features needed for accurate quality classification of earthen ceramic pots.

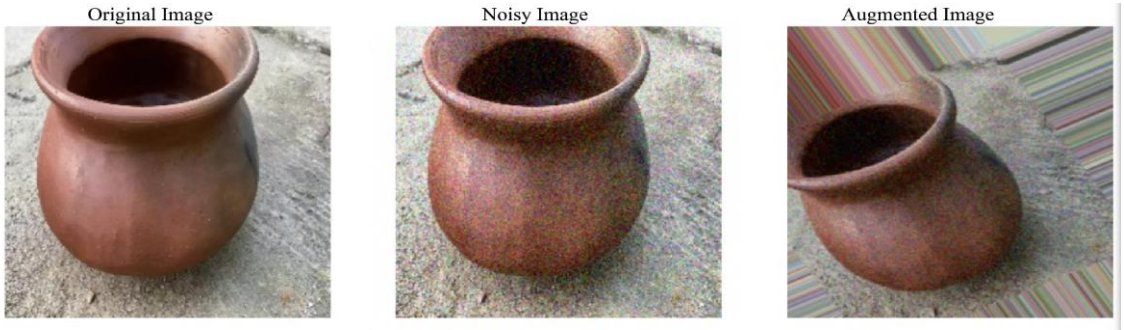

Fig. 2 Standard

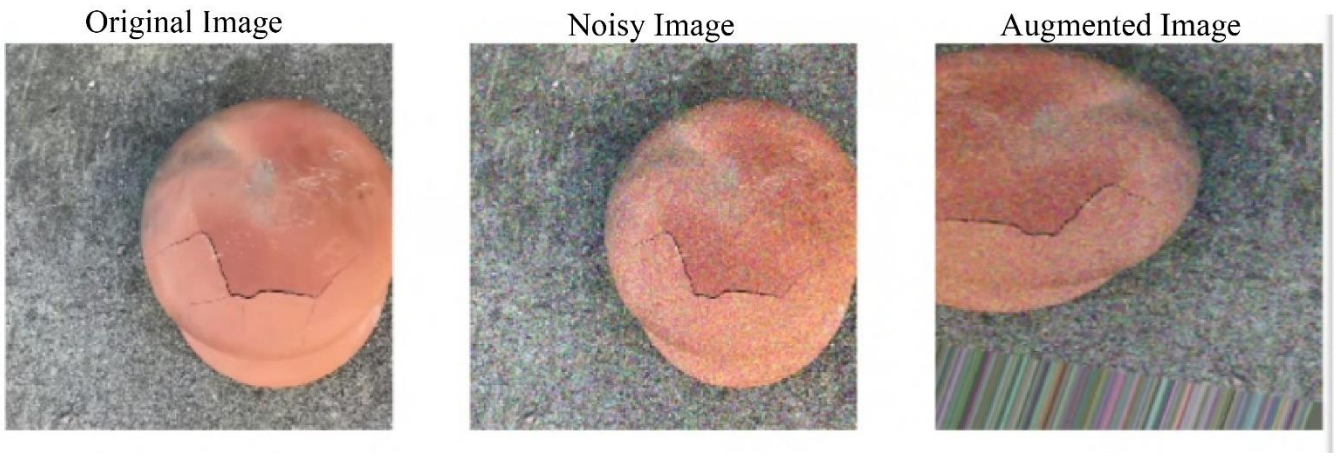

Fig. 3 Substandard

4.2. Training Dynamics and Convergence Analysis

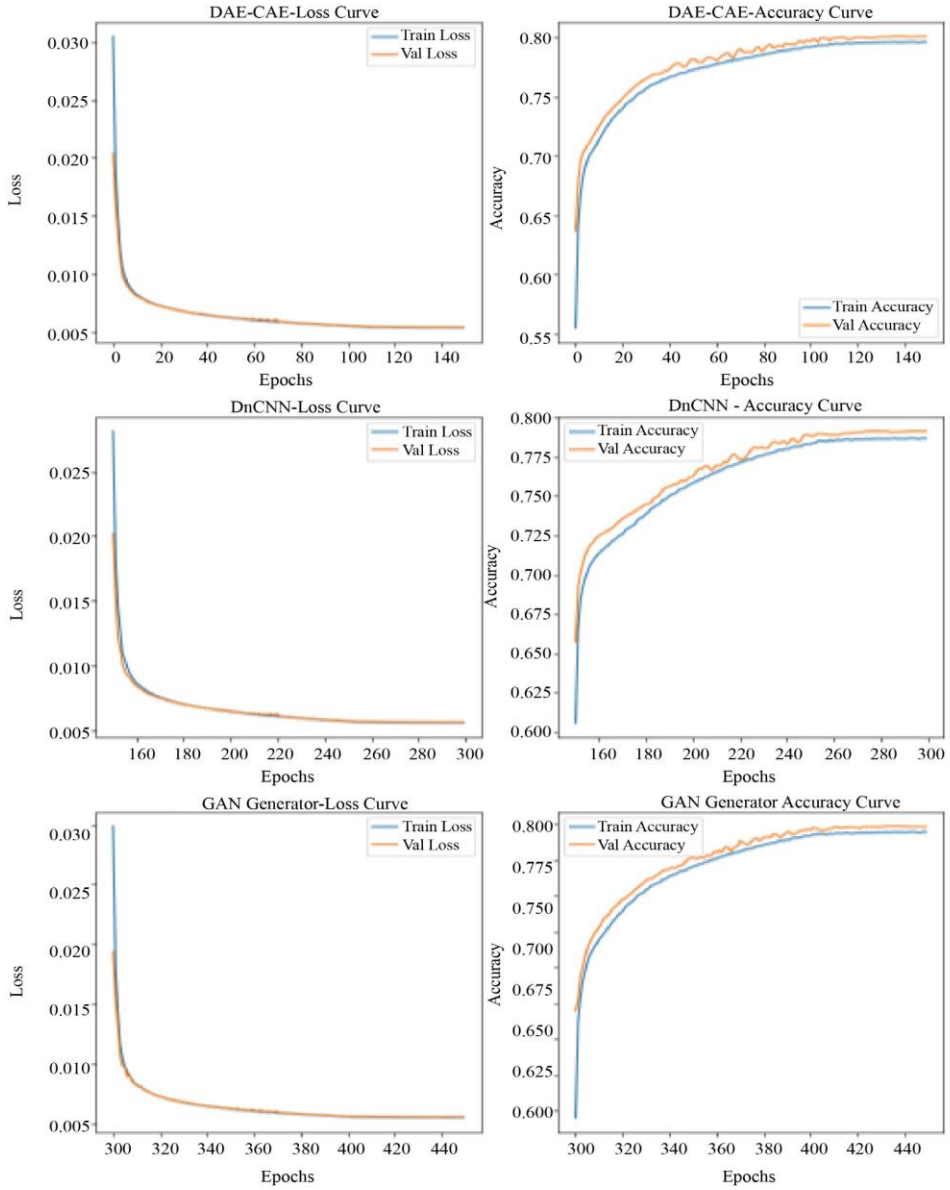

Fig. 4 Training loss and accuracy curves of DAE-CAE, DnCNN, and GAN generator

Figure 4 shows the loss and accuracy progression of the three denoising models: DAE-CAE, DnCNN, and GAN Generator, which were trained for over 150 epochs. All three models exhibited a consistent downward trend in both training and validation loss this confirms the effectiveness of error minimization in reconstructing clean images for the DAE-CAE it reached convergence earliest with its curves stabilizing before the 100th epoch indicating efficient learning and faster adaptation to the noise distribution in earthen ceramic pot images while the DnCNN required more training epochs to approach stable performance though its validation accuracy continued to improve steadily in later training stages. The GAN Generator displayed the slowest convergence rate, with its curves flattening only after extended epochs, reflecting the inherently more complex adversarial optimization process.

Comparing the accuracy of the three models, the DAE-CAE consistently maintained a slight edge, achieving earlier stabilization around 0.80, while DnCNN and the GAN Generator approached similar levels at around 0.79 more gradually. The close alignment of training and validation curves across all models suggests minimal overfitting and strong generalization. While all three models denoise effectively on images of earthen ceramic pots, the DAE-CAE demonstrates the most efficient training dynamics, whereas DnCNN and GAN require longer epochs to reach similar stability and accuracy levels.

4.3. Quantitative Evaluation of Denoising Model and Best Model Identification

Table 2. Comparative denoising performance metrics

Model	PSNR	SSIM	RMSE	Overall Rank
DAE-CAE	23.2069	0.4828	0.0713	1
GAN Generator	23.1203	0.4769	0.0719	2
DnCNN	23.0786	0.4742	0.0722	3

The quantitative results in Table 2 confirm the DAE-CAE as the top denoising model across all metrics. It achieves the highest PSNR of 23.2069 and SSIM of 0.4828 with the lowest RMSE of 0.0713, indicating strong noise suppression and better structural preservation than the alternatives. The GAN generator ranks second, having the PSNR = 23.1203, SSIM = 0.4769, and RMSE = 0.0719, reflecting effective but slightly weaker reconstruction quality. DnCNN ranks third with PSNR = 23.0786, SSIM = 0.4742, and RMSE = 0.0722. The result shows comparatively less success in preserving detail and reducing error. The per-metric rankings of the DAE-CAE = 1, GAN = 2, DnCNN = 3 for PSNR, SSIM, and RMSE that aligned with the overall ranking that establishes DAE-CAE as the best model for denoising earthen ceramic pots images, achieving the most favorable balance of noise removal and feature fidelity.

4.4. Visualization of Model Performance Metrics

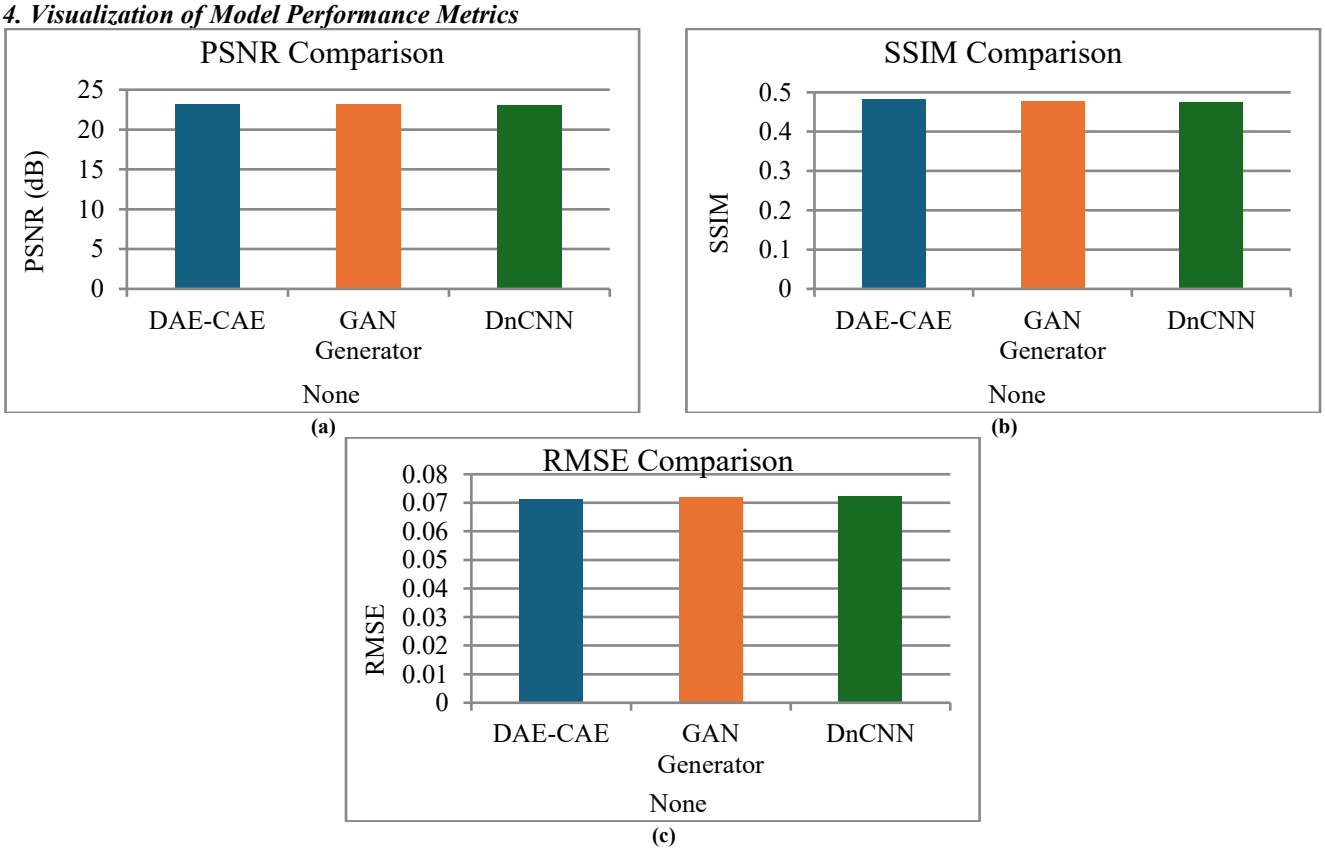

Fig. 5 Bar chart comparison of PSNR, SSIM, and RMSE across models

Figure 5 provides a clear visual comparison of the three denoising architectures' performance differences. The bar charts validate DAE-CAE's quantitative superiority by showing its small but consistent advantage in PSNR and SSIM while maintaining the lowest RMSE. All three models achieved PSNR values clustered around 23 and SSIM values near 0.48, which suggests that absolute differences between them are minimal.

However, even small gains in these metrices, which translate into a meaningful practical benefit for the classification of the accuracy of the earthen ceramic pots. The close competition among models confirms that all of them are effective denoising techniques, but DAE-CAE consistently leads across multiple independent metrices that indicate systematic advantages rather than random performance variations.

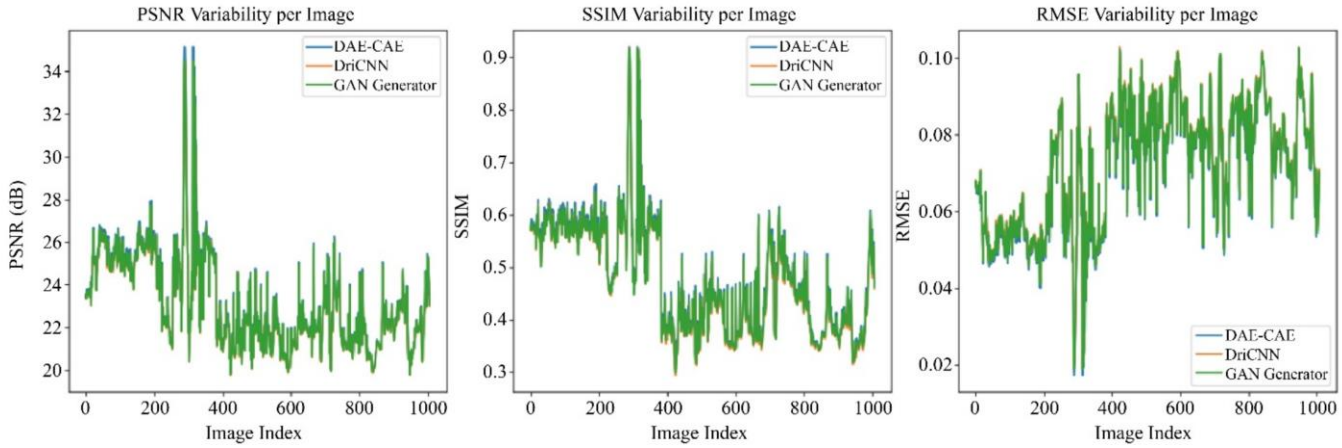

Fig. 6 Variability matrix across test set

This figure tracks how the PSNR, SSIM, and RMSE vary across the individual test images for the DAE-CAE, DnCNN, and GAN Generator. The DAE-CAE maintains the most stable performance across the three metrices. The PSNR holds consistently in the 22-26 dB range with minimal deviations, matching the DnCNN's stability, while the GAN generator shows an erratic spike above 34 db. Both SSIM and DAE-CAE stay within a controlled 0.4-0.6 band, demonstrating a reliable structural recovery across the different ceramic samples. Consequently, the DnCNN behaves similarly, while the GAN generator fluctuates widely between 0.3 and 0.9. The RMSE further confirms that the DAE-CAE's consistency with small but controlled fluctuations compared to the GAN Generator's range of 0.02 to 0.10. Since the DAE-CAE's performance across the dataset is stable, it is the most suitable for an automated defect inspection where consistent results matter more than occasional high-quality outputs. This model maintains a tight control over reconstruction error throughout the earthen ceramic pot samples, which delivers the predictable quality needed for a productive environment.

This figure presents a heatmap of the encoding model of its performance through color intensity across PSNR, SSIM, and RMSE metrices. Ultimately, the DAE-CAE model achieves the highest PSNR, which is at 23.21, outperforming the DnCNN, which is 23.08, and the GAN Generator, which is roughly at 23.12 with a 0.13 dB. The SSIM and DAE-CAE scores 0.48, matching the GAN generator and slightly exceeding DnCNN at 0.47. The visible cool blue shading across all models indicates moderate structural similarity, with DAE-CAE maintaining the strongest structural preservation, while the RMSE provides the clearest differentiation. The DAE-CAE achieves 0.071 compared to 0.072 for both DnCNN and the GAN Generator. The darker blue shading in this column, where lower value indicate better performance, visually confirms that DAW-CAE's advantage in minimizing reconstruction error. The heatmap shows that the DAE-CAE model consistently leads across all three metrices, which either matches or exceeds competing models in every measure. Thus, establishes it as the most balanced and effective architecture for denoising earthen ceramic pots.

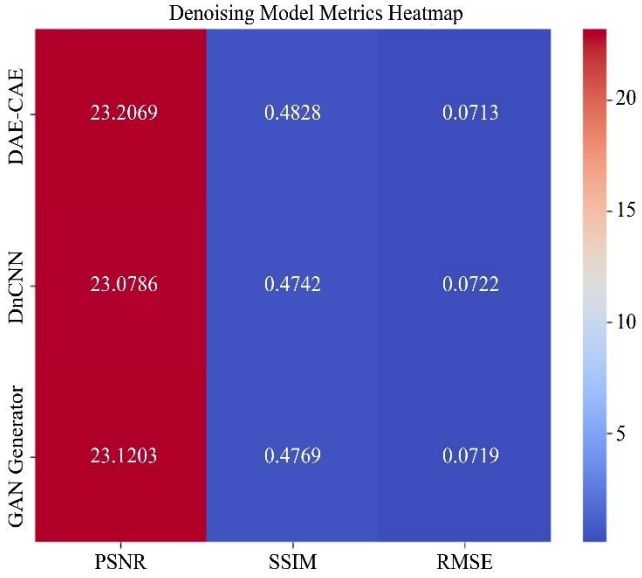

Fig. 7 Performance metrics of heatmap visualization

4.5. Feature Extraction and Classification Performance Analysis

After evaluating the three denoising models, DAE-CAE is selected as the optimal pre-processing approach based on the results of feature extraction, and was performed on the denoised earthen ceramic pots images. These features were then used to train and evaluate three classification algorithms: Support Vector Machine (SVM), Random Forest, and Convolutional Neural Network (CNN). This comparison result determines which algorithm best exploits the enhanced feature representations produced through denoising preprocessing.

Table 3. Comparative classification performance

Matrices	SVM	Random Forest	CNN Classifier
Accuracy	93.23%	90.73%	90.20%
Precision	0.9323	90.73%	0.9052
Recall	0.9323	90.73%	0.9020
F1-score	0.9323	90.73%	0.9018
True Positives (TP)	262 (correctly predicted class 1)	249	240
False Positives (FP)	20 (class 0 incorrectly predicted as class 1)	21	15
False Negatives (FN)	18 (class 1 incorrectly predicted as class 0)	31	40

The classification results, as shown in Table 3, among the three models, the SVM achieves the most reliable classification performance. This demonstrates SVM's superior performance in exploiting the DAE-CAE extracted features for earthen ceramic pots quality assessment, with 93.23% accuracy and perfectly balanced precision, recall, and F1-score metrics, all achieved a 0.9323. The confusion matrix reveals an optimal error distribution with 20 false positives only, as incorrectly classifying substandard in earthen ceramic pots as standard and 18 false negatives, incorrectly rejecting standard quality earthen ceramic pots, a near-equal balance suggesting unbiased classification without systematic preference for either error type. This balanced error profile is particularly valuable in earthen ceramic pots production contexts where both error types carry consequences. False positives compromise product quality by allowing defective items into standard inventory, while false negatives waste resources by rejecting acceptable pottery. SVM treats both error types equally, making it appropriate for quality control where maintaining standards matters, but excessive rejection also costs money.

Meanwhile, the Random Forest achieved 90.73% accuracy, produced a higher false negative rate at 31 instances compared to 21 false positives. The asymmetry shows the ensemble classifier tilts toward a conservative prediction and rejects borderline standard-quality ceramics more frequently than it accepts borderline substandard ones. This conservative tendency might fit applications requiring strict quality enforcement despite higher rejection rates, but it represents suboptimal performance for balanced quality control. CNN achieved the lowest accuracy at 90.20% despite its deep learning architecture. It generated the fewest false positives at 15, showing strong specificity in identifying standard-quality ceramics. This came at the expense of substantially higher false negatives at 40, the most among tested classifiers. This imbalance suggests CNN learned overly restrictive decision boundaries, labelling many acceptable ceramics as defective. The gap between CNN's architectural capacity and its weaker performance likely stems from limited training data. With 2,214 images, traditional machine learning approaches like SVM better exploit extracted features without needing the large datasets CNNs typically require to reach full potential.

Table 4. SVM's leading performance

Classifier	Accuracy	Precision	Recall
SVM	93.23%	0.9323	0.9323
Random Forest	90.73%	0.9078	0.9073
CNN	90.20%	0.9052	0.9020

Table 4 confirms SVM's leading performance across all metrics. Precision, recall, and F1-score align at 0.9323, indicating a well-calibrated decision boundary that avoids being either too lenient or too strict. As seen, the balance, combined with the highest accuracy, establishes SVM as the most dependable classifier in a practical and quality assessment for a ceramic pot. The results validate the integrated accuracy of DAE-CAE in denoising, followed by the SVM classification, which balances and efficiently solves for an automated quality control.

The 93.23% accuracy reflects near-human expert performance in creating a quality pot, providing a significant advantage not only in consistency and objectivity but also in scalability. These factors matter most in industrial deployment, specifically in traditional pottery production, where manual inspections remain widespread despite these obvious shortcomings.

4.6. Computational Efficiency and Runtime Performance of Denoising Models

The computational efficiency in assessing denoising quality and classification accuracy, this study analyzed computational demands to determine deployment viability. Parameter count, model size, and inference time reveal efficiency trade-offs across architectures. These metrics matter for resource-limited environments like small pottery workshops or mobile inspection devices.

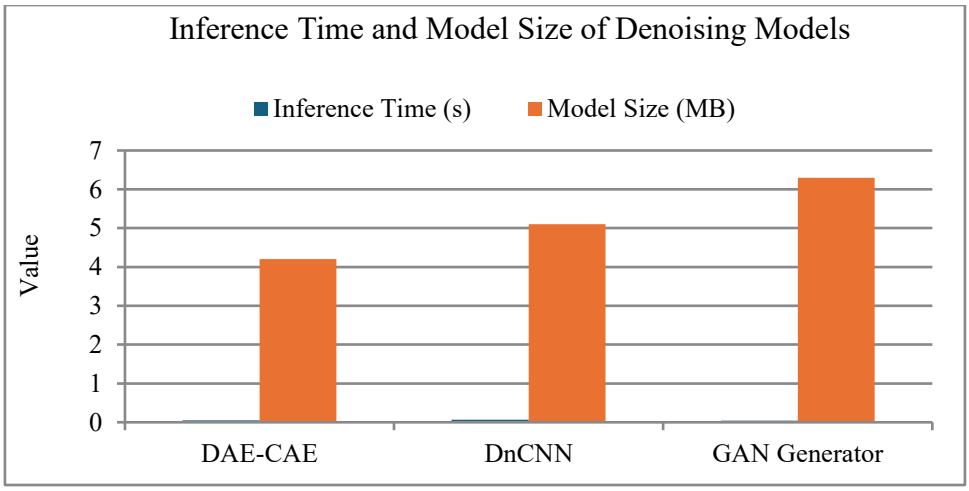

Fig. 8 Computational efficiency comparison of denoising models

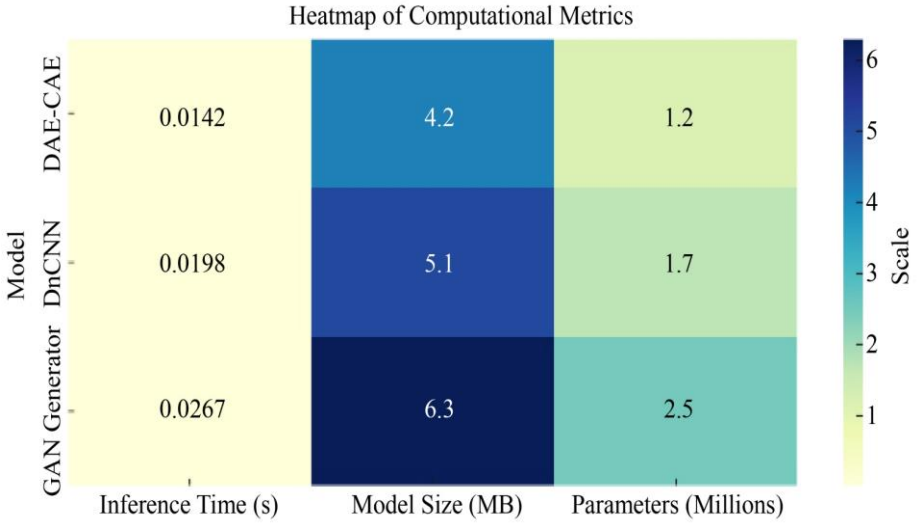

Fig. 9 Computational metrics heatmap

As illustrated in Figure 8, inference time and model size across the three architectures are shown. DAE-CAE shows the best efficiency: shortest inference time and smallest model size. This profile is suitable for real-time inspection systems or edge devices with constrained resources. DnCNN falls in the middle, with moderately higher inference time and model size. The increased cost reflects DnCNN's deeper residual architecture with extra convolutional layers that capture complex noise patterns. This design works when offline batch processing is feasible. The GAN Generator shows the highest computational demands, the longest inference time, and the largest model size. It shows architectural complexity as required for adversarial training. The GAN is suited for high-performance computing setups, even offline scenarios where visual quality outweighs speed. The differences become significant on a large scale. In DAE-CAE, processing 1,000 daily images takes 14 seconds, while it takes 27 seconds for the GAN Generator. In a high-end facility, it handles tens of thousands of images; the differences show whether it is a real-time inspection of earthen ceramic pots, thus, it remains more practical.

This heatmap compares the computational efficiency not only in inference time and model size but also in parameter counts. The darker the colors, the lower the efficiency, while for lighter ones, it indicates higher efficiency. With this, the DAE-CAE achieves a 0.0142-second inference time of about 4.2 MB size and 1.2 million parameters, making it the most efficient architectural method. At the same time, the low parameter count directly translates to reducing memory bandwidth and fewer operations, which matters for a resource-constrained deployment in all ceramic inspection systems. Finally, the DnCNN requires a 0.0198-second inference time of about 5.1 MB and 1.7 million parameters. The 42% parameter increase over the DA-CAE stems from its deeper residual architecture method, thus producing a moderate computational overhead. In the positions of DnCNN between its minimal and high-resource scenarios, the GAN generator

needs a 0.0267-second inference time with a 6.3 MB model and 2.5 million parameters. More than doubled DAE-CAE's count, reflecting the architectural complexity of adversarial training and thus limiting GANs to environments where computational resources are limited.

4.7. Performance, Quality, and Efficiency Convergence

DAE-CAE ranks best among the evaluated denoisers across independent criteria. It achieves consistently higher PSNR and SSIM values with lower RMSE compared to competing models. Classification using DAE-CAE denoised features yields the strongest downstream performance when combined with SVM, reaching 93.23% accuracy, outpacing training on raw or noise-corrupted inputs. The model also proves most efficient by inference time, model size, and parameter count, meaning performance gains come without increased computational cost. These outcomes indicate that DAE-CAE is the most effective preprocessing choice for earthen ceramic pots quality assessment.The gain matters operationally for small workshops where capture conditions vary, and defects are subtle relative to sensor noise. Denoising stabilizes the feature distribution, allowing a conventional SVM to achieve high accuracy without the large labelled datasets that end-to-end deep learning typically requires. This aligns with prior evidence that targeted preprocessing raises classical model performance. The pipeline addresses common constraints in artisanal production: limited training data and tight compute budgets, making deployment feasible on modest hardware.

5. Conclusion

This study evaluated how learning-based denoising affects feature quality and downstream classification for earthen ceramic pots. Three architectures were compared: DAE-CAE, DnCNN, and a GAN-based denoiser. DAE-CAE delivered the strongest image quality, reaching PSNR 23.21, SSIM 0.48, and RMSE 0.071. Using DAE-CAE features, SVM achieved an accuracy of 93.23%, outperforming Random Forest at 90.73% and CNN at 90.20%. The denoiser processed images in 0.0142 seconds with a 4.2 MB model and 1.2 million parameters, supporting real-time use under limited compute resources.

The findings met the stated objectives, which are a comparative evaluation of denoising architectures, identification that the DAE-CAE is the most effective pre-processing stage for earthen ceramic pots features, and the validation of SVM trained on denoising image features, which exceeds the alternative classifiers. Combining an effective denoising with a conventional machine learning yielded an accurate and scalable quality assessment, compared to manual inspection of an earthen ceramic pot, as the workflow provides an objective.

The repeatable decisions at real-time speeds reduce the operator's burden and preserve expert oversight. This system serves as the decision support for potters rather than replacing their domain skills. Methodologically, this study provides a reproducible protocol for evaluation and pre-processing settings with scarce labels and even constrained hardware. The same analysis applies not only to agricultural grading and textile inspection but also to other craft manufacturing tasks where automated visual inspection raises reliability and throughput of the ceramic pot. Finally, a high-quality denoising method enables a traditional model to reach strong and most accurate results without large datasets. These matters for the production of pots, where extensive labelling is impractical.

References

[1] Khaled Alomar, Halil Ibrahim Aysel, and Xiaohao Cai, "Data Augmentation in Classification and Segmentation: A Survey and New Strategies," *Journal of Imaging*, vol. 9, no. 2, pp. 1-26, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[2] Markus Bayer, Marc-André Kaufhold, and Christian Reuter, "A Survey on Data Augmentation for Text Classification," *ACM Computing Surveys*, vol. 55, no. 7, pp. 1-39, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[3] Tom Brown et al., "Language Models are Few-Shot Learners," *34th Conference on Neural Information Processing Systems*, Vancouver, Canada, pp. 1877-1901, 2020. [Google Scholar] [Publisher Link]

[4] Jair Cervantes et al., "A Comprehensive Survey on Support Vector Machine Classification: Applications, Challenges, and Trends," *Neurocomputing*, vol. 408, pp. 189-215, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[5] Huseyin Coskun, Tuncay Yiğit, and İsmail Serkan Üncü, "Integration of Digital Quality Control for Intelligent Manufacturing of Industrial Ceramic Tiles," *Ceramics International*, vol. 48, no. 23, pp. 34210-34233, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[6] Esteban Cumbajin et al., "A Real-Time Automated Defect Detection System for Ceramic Pieces Manufacturing Process based on Computer Vision with Deep Learning," *Sensors*, vol. 24, no. 1, pp. 1-22, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[7] Steffen Dereich, and Arnulf Jentzen, "Convergence Rates for the Adam Optimizer," *arXiv Preprint*, pp. 1-43, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[8] Timothy O. Hodson, "Root Mean Square Error (RMSE) or Mean Absolute Error (MAE): When to Use them or Not," *Geoscientific Model Development Discussions*, vol. 15, no. 14, pp. 1-10, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[9] Shafaf Ibrahim et al., "Automated Defective Ceramic Tiles Classification using Image Processing Techniques," *Journal of Advanced Research in Applied Sciences and Engineering Technology*, vol. 32, no. 3, pp. 355-365, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[10] Imran, Naeem Iqbal, and Do-Hyeun Kim, "Intelligent Material Data Preparation Mechanism based on Ensemble Learning for AI-Based Ceramic Material Analysis," *Advanced Theory and Simulations*, vol. 5, no. 11, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[11] Vijay Janapa Reddi et al., "MLPerf Mobile Inference Benchmark: An Industry-Standard Open-Source Machine Learning Benchmark for On-Device AI," *Proceedings of the 5th MLSys Conference, Santa Clara*, CA, USA, pp. 352-369, 2022. [Google Scholar] [Publisher Link]

[12] Taehyeon Kim et al., "Comparing Kullback-Leibler Divergence and Mean Squared Error Loss in Knowledge Distillation," *arXiv Preprint*, pp. 1-11, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[13] See Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton, *From Photographic Image to Computer Vision*, Monitoring Laws Profiling and Identity in the World State, Cambridge University Press, pp. 135-157, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[14] Yuh-Jye Lee, and O.L. Mangasarian, "SSVM: A Smooth Support Vector Machine for Classification," *Computational Optimization and Applications*, vol. 20, no. 1, pp. 5-22, 2001. [CrossRef] [Google Scholar] [Publisher Link]

[15] Jingyun Liang et al., "Swinir: Image Restoration using Swin Transformer," *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV) Workshops*, pp. 1833-1844, 2021. [Google Scholar] [Publisher Link]

[16] Jakub Nalepa, and Michal Kawulok, "Selecting Training Sets for Support Vector Machines: A Review," *Artificial Intelligence Review*, vol. 52, no. 2, pp. 857- 900, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[17] K. Nanthini et al., "A Survey on Data Augmentation Techniques," *2023 7th International Conference on Computing Methodologies and Communication (ICCMC)*, Erode, India, pp. 913-920, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[18] Jim Nilsson, and Tomas Akenine-Möller, "Understanding SSIM," *arXiv Preprint*, pp. 1-8, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[19] Simone Parisotto et al., "Unsupervised Clustering of Roman Potsherds via Variational Autoencoders," *arXiv Preprint*, pp. 1-16, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[20] Derek A. Pisner, and David M. Schnyer, *Support Vector Machine*, Machine Learning: Methods and Applications to Brain Disorders, pp. 101-121, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[21] Edisson Pugo-Mendez, and Luis Serpa-Andrade, "Development of a Platform based on Artificial Vision with SVM and KNN Algorithms for the Identification and Classification of Ceramic Tiles," *Artificial Intelligence and Social Computing*, vol. 28, no. 28, pp. 173-181, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[22] Yuriy Reznik, "Another Look at SSIM Image Quality Metric," *Electronic Imaging*, vol. 35, pp. 1-7, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[23] Roy Schwartz et al., "Green AI," *Communications of the ACM*, vol. 63, no. 12, pp. 54-63, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[24] De Rosal Igantius Moses Setiadi, "PSNR vs SSIM: Imperceptibility Quality Assessment for Image Steganography," *Multimedia Tools and Applications*, vol. 80, no. 6, pp. 8423-8444, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[25] Connor Shorten, and Taghi M. Khoshgoftaar, "A Survey on Image Data Augmentation for Deep Learning," *Journal of Big Data*, vol. 6, no. 1, pp. 1-48, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[26] Okeke Stephen, Uchenna Joseph Maduh, and Mangal Sain, "A Machine Learning Method for Detection of Surface Defects on Ceramic Tiles Using Convolutional Neural Networks," *Electronics*, vol. 11, no, 1, pp. 1-22, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[27] Chunwei Tian et al., "Deep Learning on Image Denoising: An Overview," *Neural Networks*, vol. 131, pp. 251-275, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[28] Pascal Vincent et al., "Stacked Denoising Autoencoders: Learning Useful Representations in a Deep Network with a Local Denoising Criterion," *Journal of Machine Learning Research*, vol. 11, no. 12, pp. 3371-3408, 2010. [Google Scholar] [Publisher Link]

[29] Slamet Widodo, Herlambang Brawijaya, and Samudi Samudi, "Stratified K-Fold Cross Validation Optimization on Machine Learning for Prediction," *Sinkron: Jurnal Dan Penelitian Teknik Informatika*, vol. 6, no. 4, pp. 2407-2414, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[30] Min-Ling Zhu, Liang-Liang Zhao, and Li Xiao, "Image Denoising based on GAN with Optimization Algorithm," *Electronics*, vol. 11, no. 15, pp. 1-12, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[31] Bo Zhang et al., "Signal Data Augmentation Algorithm Research based on Fractal Theory," *2024 IEEE 7th International Conference on Information Systems and Computer Aided Education (ICISCAE)*, Dalian, China, pp. 1023-1026, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[32] Houwang Zhang, Yuan Zhu, and Hanying Zheng, "NAMF: A Non-Local Adaptive mean Filter for Salt-and-Pepper Noise Removal," *arXiv Preprint*, pp. 1-9, 2020. [CrossRef] [Google Scholar] [Publisher Link]