

# ***A Review of AI-Based Stacked Denoising Techniques and Their Comparison with Classical Denoising Algorithms***

**Yuhan Zhao<sup>1,a,\*</sup>**

<sup>1</sup>*Taiyuan University of Technology, Taiyuan, 030000, China*

*a. zhaoyuhan6839@outlook.com*

*\*corresponding author*

**Abstract:** This paper delves into the issue of image noise and its significance in digital image processing. It analyzes the sources of noise, its impact on image quality, and its interference with subsequent processing tasks. Traditional denoising algorithms, such as mean filtering, median filtering, and Gaussian filtering, have limitations when dealing with specific types of noise, particularly in complex noise scenarios where image details are prone to loss. In recent years, the rise of deep learning technology has brought new breakthroughs to image denoising. Specifically, Stacked Denoising Autoencoders have significantly improved denoising effectiveness and image quality by extracting features layer by layer and applying nonlinear transformations, enabling the removal of noise while preserving important details and structural information in the image. However, the training process of SDAE demands high data and computational resources and has a strong dependence on training data, limiting its application in scenarios requiring high real-time performance. This study not only provides a new technological option for image denoising but also offers a scientific basis for algorithm selection and optimization.

**Keywords:** Stacked denoising autoencoder, image denoising, classical denoising algorithms

## **1. Introduction**

The issue of image noise has consistently occupied a central position in the field of digital image processing. As a pivotal factor influencing image quality, noise primarily originates from the processes of image acquisition, data transmission, and storage. The presence of image noise not only diminishes visual presentation effects and reduces overall quality but also affects the accuracy and efficiency of subsequent image processing tasks [1]. In advanced tasks such as image recognition, object detection, and image segmentation, noise interference leads to decreased algorithm performance, increasing the risk of misjudgments and omissions. Therefore, image denoising has become a core aspect of the image processing workflow, crucial for enhancing image quality and optimizing the performance of subsequent tasks.

In recent years, the emergence of artificial intelligence (AI) technology, particularly deep learning, has revolutionized the field of image processing. AI technology, with its powerful data processing and pattern recognition capabilities, has demonstrated exceptional performance in tasks such as image denoising. Traditional denoising methods, such as mean filtering, median filtering, and Gaussian filtering, while effective in removing noise, tend to blur image edge information and detailed features [2]. In contrast, AI-based denoising methods, such as Convolutional Neural Networks (CNNs) and

Generative Adversarial Networks (GANs), by simulating the processing mechanisms of the human visual system, can preserve image edge information and detailed features while denoising, significantly improving image quality and enhancing the performance of subsequent processing tasks.

This study delves into the innovative aspects and performance differences of stacked denoising techniques, not only facilitating a better understanding of the advantages and limitations of this emerging technology but also providing valuable references and guidance for research and applications in the field of image denoising. This will contribute to the continuous advancement and innovation of image processing technologies, making positive contributions to the development of related fields.

## 2. Literature Review

In the field of data processing and analysis, denoising is a crucial task. Noise typically originates from various factors such as sensor errors and environmental interferences, which not only degrade data quality but may also adversely affect subsequent analysis and modeling [3]. To address this challenge, researchers have developed a series of classical denoising algorithms, among which mean filtering, median filtering, and Gaussian low-pass filtering are most widely used in practical applications. Mean filtering is suitable for removing random noise in images, especially when the noise distribution is uniform and the intensity is weak. In fields such as environmental monitoring and digital photography, mean filtering is extensively applied to smooth data and enhance image quality.

Mean filtering is a typical linear filtering algorithm whose basic principle involves replacing each pixel value in the original image with its mean value [4]. For the current pixel  $(x, y)$  to be processed, a template consisting of several neighboring pixels is selected. The mean of all pixels within the template is then calculated and assigned to the current pixel  $(x, y)$  as its gray value in the processed image. Mean filtering reduces noise through smoothing but has the drawback of potentially blurring image details. Its mathematical expression is:

$$f_{\text{mean}}(x, y) = \frac{1}{N} \sum_{(i,j) \in \Omega} f(i, j) \quad (1)$$

where  $f(i, j)$  is the gray value of the original image at pixel  $(i, j)$ ,  $\Omega$  is the neighborhood window centered at  $(x, y)$ , and  $N$  is the total number of pixels within the window.

Median filtering is a nonlinear filtering algorithm that achieves denoising by calculating the median of the gray values of multiple pixels surrounding a given pixel [5]. Median filtering can preserve image edges and details while effectively suppressing noise types such as salt-and-pepper noise [6]. Its basic idea is to use the median of gray values instead of gray value variance for image denoising, determining the optimal smoothing window size and direction through local statistics and analysis. The output of median filtering is:

$$f_{\text{Median}}(x, y) = \text{Median}\{f(i, j) | (i, j) \in \Omega\} \quad (2)$$

where Median denotes the median operation.

Gaussian low-pass filtering employs the properties of the Gaussian function to perform a convolution operation on the image, achieving the purpose of blurring the image and reducing noise [7]. The weight distribution of the Gaussian filter conforms to the Gaussian function, with pixels farther from the center having smaller weights [8]. By adjusting the standard deviation ( $\sigma$ ) of the Gaussian function, the degree of smoothing of the filter can be controlled. Gaussian low-pass filtering is highly effective in filtering out random noise such as Gaussian noise but also has the issue of excessive smoothing leading to loss of image details.

$$f_{\text{Gaussian}}(x, y) = \sum_{(i,j) \in \Omega} f(i, j) \cdot G(i, j) \quad (3)$$

where  $G(i, j)$  is the Gaussian kernel function defined as:

$$G(i, j) = \frac{1}{2\pi\sigma^2} \exp \left( -\frac{(i-x)^2 - (j-y)^2}{2\sigma^2} \right) \quad (4)$$

Classical denoising algorithms such as mean filtering, median filtering, and Gaussian low-pass filtering have significant advantages in processing specific types of noise but also exhibit some common limitations. The main limitation of mean filtering is that it may blur image details, resulting in unclear edges. Additionally, for impulsive noise with strong variability, mean filtering is not ideal. Although median filtering performs well in removing salt-and-pepper noise, its effectiveness in filtering out random noise, such as Gaussian noise is inferior to that of Gaussian low-pass filtering. Furthermore, when noise density is high, median filtering may destroy the detailed structure of the image. The primary limitation of Gaussian low-pass filtering is that excessive smoothing may lead to loss of image details. Moreover, for impulsive noise with strong variability, Gaussian low-pass filtering is not effective. Additionally, designing a Gaussian filter requires selecting an appropriate standard deviation ( $\sigma$ ), which increases algorithm complexity and parameter tuning difficulty. To overcome these limitations, researchers have proposed many improved algorithms and new technologies. For example, image denoising algorithms based on deep learning achieve significant denoising results by training deep neural network models to automatically learn denoising rules. Furthermore, combining the advantages of multiple denoising algorithms for combined denoising is also one of the current research hotspots. By comprehensively applying the strengths of different algorithms, denoising performance can be further enhanced and limitations reduced. Deep learning denoising algorithms possess powerful adaptive and feature learning capabilities, enabling them to handle complex and high-level noise. In addition, these algorithms can retain more detailed information while smoothing images, improving image quality.

### 3. Application of Stacked Denoising Techniques in Image Denoising

#### 3.1. Theoretical Basis and Innovation of Stacked Denoising Autoencoders

Stacked Denoising Autoencoders (SDAE) constitute a learning model with significant advantages in the field of deep learning. They achieve deep feature extraction and efficient representation of complex data by skillfully stacking multiple layers of Denoising Autoencoders (DAE). The core objective of SDAE is to gradually uncover the intrinsic structure and features of data through layered learning and optimization, thereby enhancing the model's generalization ability and the precision of feature extraction. Within the framework of SDAE, each layer of DAE plays a crucial role by extracting information from noisy input data and attempting to recover the original clean data. This process not only aids in removing noise components from the data but also prompts the model to learn more robust and effective data features, laying a solid foundation for subsequent tasks such as classification, regression, clustering, and more.

#### 3.2. Stacked Structure and Feature Learning Capability of SDAE

SDAE constructs a deep network structure by stacking multiple layers of DAE to achieve layered feature extraction and deep representation of data. In each layer, the DAE effectively removes noise from the input data and learns high-order features of the data [8]. Through multi-layer stacking, SDAE can gradually mine deep information in the data, forming more abstract and complex feature representations. These feature representations are highly valuable for subsequent classification, prediction, or dimensionality reduction tasks, as they not only contain core information of the data but also reflect complex relationships and patterns among the data, as shown in Table 1.

Table 1: Main Characteristics and Training Process of SDAE

Item	Simplified Description
SDAE	Stacked Denoising Autoencoder for layer-wise feature extraction
Features	Noise removal and high-level feature learning
Training	Pre-training + Fine-tuning
Pre-training	Layer-wise training for feature learning
Fine-tuning	Connecting classifier and optimizing model

The training process of SDAE comprises two main stages: pre-training and fine-tuning. During the pre-training stage, SDAE trains each DAE layer by layer. Initially, noisy input data is fed into the first DAE to train it to reconstruct the original data from the noisy data. Upon completion of training, the first DAE learns the first-layer feature representation of the input data [9]. Subsequently, the output of the first DAE (i.e., the encoded features) is used as the input for the second DAE, which is similarly trained to reconstruct the feature representation of the previous layer from noisy data. This process is repeated, training multiple DAEs in sequence, ultimately forming a stacked structure. Each layer of DAE is responsible for removing noise from the input data and extracting higher-level features.

After pre-training all layers, SDAE enters the fine-tuning stage. At this stage, the entire stacked denoising autoencoder is connected with a classifier (e.g., Softmax classifier) to form a complete deep learning model. Then, labeled data is used to fine-tune the entire model. The goal of fine-tuning is to adjust the model parameters by minimizing classification loss, enabling the model to perform better on classification tasks. Through fine-tuning, SDAE can further optimize its feature extraction and classification capabilities, achieving accurate processing and efficient representation of complex data.

### 3.3. Application of Stacked Denoising Techniques in Image Denoising

Image denoising is a core problem in image processing, aiming to effectively remove noise from images while preserving important features and details to the greatest extent possible [10]. Stacked Denoising Autoencoders (SDAE), as a deep learning model, exhibit significant advantages in image denoising tasks through layered learning of image feature representations. This section will delve into the application of SDAE in image denoising and analyze its advantages compared to traditional methods.

SDAE is a deep learning model formed by stacking multiple layers of DAE based on the Denoising Autoencoder (DAE). DAE learns robust feature representations of data by adding noise to the input data and training the model to recover the original data from the noisy data. SDAE further leverages this characteristic by stacking multiple DAEs to gradually extract high-level features of images, effectively removing image noise [11]. The structure of SDAE typically includes an input layer, multiple hidden layers (each being a DAE), and an output layer. During training, SDAE encodes and decodes the input image layer by layer, optimizing model parameters by minimizing reconstruction error. As training progresses, SDAE learns the mapping relationship from noisy images to noise-free images, thereby achieving image denoising.

### 3.4. Specific Applications of SDAE in Image Denoising

The application of SDAE in image denoising is extensive, encompassing various scenarios requiring noise removal from images. Firstly, in medical image processing, SDAE can be used to remove noise from medical images, enhancing image clarity and contrast. For instance, in X-ray, CT, and MRI images, SDAE can effectively remove noise, enabling doctors to make more accurate diagnoses.

Secondly, remote sensing images often contain significant noise and interference, affecting image resolution and recognition accuracy. SDAE can be applied to remove noise from remote sensing images, improving image clarity and recognition accuracy, thereby aiding researchers in better analyzing remote sensing data. Additionally, in image inpainting, SDAE can be used to remove damaged parts or noisy areas of images, restoring complete images. For example, in image inpainting tasks, SDAE can learn texture and structural information in images, thereby achieving restoration of damaged parts and noise removal.

#### 4. Comparative Analysis of Stacked Denoising Techniques and Classical Denoising Algorithms

Stacked denoising techniques, particularly Stacked Denoising Autoencoders (SDAEs), as an emerging deep learning approach, exhibit distinct characteristics and advantages compared to classical denoising algorithms across multiple dimensions, as shown in Table 2.

Table 2: Comparison Analysis of Stacked Denoising Techniques and Classical Denoising Algorithms

Dimension	SDAE	Classical Algorithms
Feature Learning	High-level features, handles complex noise	Effective for specific noise types
Denoising Effect	Strong performance, preserves details	May struggle with complex noise
Computational	High training costs, lower inference complexity	Generally efficient, but can be complex for large data
Application	High-quality images/videos, fine detail preservation	Limited resources, real-time requirements
Limitations	Dependence on training data, high costs	Insufficient for complex noise, detail preservation
Improvements	Efficient training, more data	Combine with deep learning, explore new algorithms

Firstly, SDAEs, through multi-layer nonlinear transformations, are capable of learning high-level feature representations of image data, thereby demonstrating robust performance in the denoising process. They can effectively remove common noise types such as Gaussian noise and salt-and-pepper noise, and to some extent, handle more complex, nonlinear noise patterns. In contrast, classical denoising algorithms like Non-Local Means (NLM) and Block-Matching and 3D Filtering (BM3D), while excelling in dealing with specific types of noise, may fall short when confronted with complex noise or in preserving image details. In terms of evaluation metrics, Peak Signal-to-Noise Ratio (PSNR) and Mean Squared Error (MSE) are two crucial indicators for assessing denoising effectiveness. SDAEs typically achieve superior performance on these metrics, particularly at high noise levels, where their advantages are more pronounced. SDAEs better preserve edge and texture information in images, rendering the denoised images more natural and clear.

Secondly, the training process of SDAEs is relatively complex, requiring substantial data and computational resources. Once the model is trained, the computational complexity of its inference process (i.e., the denoising process) is relatively low, but still higher than some classical denoising algorithms. Classical denoising algorithms, such as NLM and BM3D, although intuitively designed, exhibit non-negligible computational complexity when processing large-scale image or video data, especially in scenarios requiring real-time processing. Regarding computational efficiency and feasibility in practical applications, the preprocessing and training phases of SDAEs may require extended periods, limiting their use in certain time-sensitive applications. However, with



advancements in hardware technology and optimizations in deep learning frameworks, the computational efficiency of SDAEs is continually improving. In contrast, classical denoising algorithms, due to their algorithmic simplicity, generally possess higher computational efficiency and better real-time performance.

Thirdly, the advantage of SDAEs in image denoising lies in their powerful feature learning capabilities and ability to handle complex noise patterns. This makes them excel in processing high-quality images, videos, or scenarios requiring fine detail preservation. However, the limitations of SDAEs include their high training costs and strong dependence on training data. If training data is insufficient or the noise type does not match the training data, the denoising effectiveness of SDAEs may be compromised. Classical denoising algorithms are more suitable for scenarios with limited computational resources or high real-time requirements. They typically exhibit lower algorithmic complexity and higher computational efficiency, meeting practical application needs without sacrificing too much denoising performance. However, classical denoising algorithms may perform poorly in handling complex noise or scenarios requiring fine detail preservation.

Fourthly, the primary limitations of SDAEs are their training costs and dependence on training data. To reduce training costs, more efficient training algorithms or methods such as transfer learning can be considered. Simultaneously, increasing the diversity and quantity of training data can enhance SDAEs' adaptability to different noise types. The limitations of classical denoising algorithms lie in their deficiencies in handling complex noise and fine detail preservation. To improve this, combining classical denoising algorithms with deep learning methods, leveraging the feature learning capabilities of deep learning to enhance the denoising performance of classical algorithms, can be considered. Additionally, exploring new denoising algorithms and evaluation metrics can better accommodate the needs of different application scenarios.

## 5. Conclusion

This study delves into the innovative aspects and advantages of stacked denoising autoencoders (SDAEs), conducting a comprehensive comparative analysis with classical denoising algorithms. By learning high-level features of image data through multi-layer nonlinear transformations, SDAEs demonstrate significant advantages in handling complex noise and preserving image details. Their innovation lies in the ability to adaptively learn the intrinsic structure of image data, thereby more effectively removing noise while maintaining the original features of the image. When compared to classical denoising algorithms such as Non-Local Means (NLM) and Block-Matching and 3D Filtering (BM3D), SDAEs typically achieve better performance in evaluation metrics such as Peak Signal-to-Noise Ratio (PSNR) and Mean Squared Error (MSE), especially when dealing with images containing high noise levels or complex noise patterns. However, SDAEs also face limitations such as high training costs and strong dependence on training data. In contrast, classical denoising algorithms offer advantages in computational efficiency and real-time performance.

This study makes substantial contributions to the field of image denoising. Firstly, by comparing and analyzing SDAEs with classical denoising algorithms, it reveals the advantages of SDAEs in denoising performance and their applicable scenarios, providing a new technological option for image denoising. Secondly, this study explores the limitations and directions for improvement of SDAEs, offering valuable references for subsequent research. Lastly, through a comprehensive evaluation of the performance of different denoising methods, this study provides a scientific basis for algorithm selection and optimization in the field of image denoising.

## References

- [1] Kim, J., Zeng, H., Ghadiyaram, D., Lee, S., Zhang, L., & Bovik, A. C. (2017). Deep convolutional neural models for picture-quality prediction: Challenges and solutions to data-driven image quality assessment. *IEEE Signal processing magazine*, 34(6), 130-141.
- [2] Brendlin, A. S., Schmid, U., Plajer, D., Chaika, M., Mader, M., Wrazidlo, R., ... & Tsiflikas, I. (2022). AI denoising improves image quality and radiological workflows in pediatric ultra-low-dose thorax computed tomography scans. *Tomography*, 8(4), 1678-1689.
- [3] Teh, H. Y., Kempa-Liehr, A. W., & Wang, K. I. K. (2020). Sensor data quality: A systematic review. *Journal of Big Data*, 7(1), 11.S
- [4] Chandel, R., & Gupta, G. (2013). Image filtering algorithms and techniques: A review. *International Journal of Advanced Research in Computer Science and Software Engineering*, 3(10).
- [5] Shah, A., Bangash, J. I., Khan, A. W., Ahmed, I., Khan, A., Khan, A., & Khan, A. (2022). Comparative analysis of median filter and its variants for removal of impulse noise from gray scale images. *Journal of King Saud University-Computer and Information Sciences*, 34(3), 505-519.
- [6] Chan, R. H., Ho, C. W., & Nikolova, M. (2005). Salt-and-pepper noise removal by median-type noise detectors and detail-preserving regularization. *IEEE Transactions on image processing*, 14(10), 1479-1485.
- [7] Mafi, M., Martin, H., Cabrerizo, M., Andrian, J., Barreto, A., & Adjouadi, M. (2019). A comprehensive survey on impulse and Gaussian denoising filters for digital images. *Signal Processing*, 157, 236-260.
- [8] Khan, Z. Y., Niu, Z., Sandiwarno, S., & Prince, R. (2021). Deep learning techniques for rating prediction: a survey of the state-of-the-art. *Artificial Intelligence Review*, 54, 95-135.
- [9] Song, S., Yang, J., RASHED, G., Shen, J., Haider, H., Jiang, K., ... & Cao, K. (2023). Refined Deep Transfer Learning with CNN-LSTM and SDAE for Adaptive Assessment of Power System Transient Stability with Time Series Data. *Authorea Preprints*.
- [10] Fan, L., Zhang, F., Fan, H., & Zhang, C. (2019). Brief review of image denoising techniques. *Visual Computing for Industry, Biomedicine, and Art*, 2(1), 7.
- [11] Zhang, L., Wang, J., Chang, R., & Wang, W. (2024). Investigation of the effectiveness of a classification method based on improved DAE feature extraction for hepatitis C prediction. *Scientific Reports*, 14(1), 9143.