Computers and Artificial Intelligence；Vol.3, No.3; 2026
E-ISSN: 3006 -3728 P-ISSN：3006-371X
DOI: https://doi.org/10.70267/cai.26v3n3.114120
Published by: Zeus Press

Empirical Research on the Influence of Image Enhancement Technology on Image Classification Performance Based on Deep Learning

Zhuofan Jia*

School of Physics and Electronic Information, Henan Polytechnic University, Jiaozuo, Henan 454000, China

**Corresponding author: Zhuofan Jia.*

Abstract

To address the problem of insufficient systematic evaluation of image preprocessing technologies, this study adopts the control variable method and quantitative analysis to systematically evaluate the influence of three types of image enhancement technologies (destructive enhancement, conservative enhancement, and combined enhancement) on the performance of an image classification model based on ResNet-50 feature extraction and SVM classification. The CIFAR-10 dataset is used with a sample size of 5,000 images. The experimental results show that destructive enhancement brings an average accuracy decrease of 21.24%, Unsharp Mask in conservative enhancement brings an increase of 3.64%, and Wavelet→Unsharp in combined enhancement brings an increase of 2.26%. The impact of image enhancement on classification performance varies significantly: appropriate enhancement effectively improves model performance, while excessive enhancement impairs it. The findings provide a basis for technology selection in practical applications.

Keywords

image enhancement, deep learning, image classification, ResNet-50, transfer learning

1. Introduction

Image classification is a core task in computer vision, and its performance depends not only on deep learning models but also on constraints imposed by image quality [1]. In complex scenarios such as mine safety monitoring, challenges persist due to environmental factors that degrade image quality [2], and the level of image quality directly affects the accuracy and robustness of classification models [3]. Existing studies have largely focused on the effects of individual models or specific enhancement techniques, without targeted investigations into technologies suited to low-quality small-sized color images in such scenarios. Moreover, they lack quantitative comparisons of multiple techniques under a unified framework, making it difficult to provide clear selection guidelines. This study employs the control variable method to design three sets of experiments, systematically evaluating the differential effects of destructive enhancement, conservative enhancement, and combined enhancement on the performance of an image classification model based on ResNet-50 and SVM. It fills the gap in unified multi-technique evaluations and provides empirical support for technology selection in low-quality image scenarios.

2. Research Methods

2.1 Experimental Design

To accurately evaluate the effects of image enhancement techniques, this study adopts a stratified control variable method and designs three sets of comparative experiments. The selected techniques include both traditional methods such as contrast enhancement and sharpening and newer approaches more compatible with feature extraction, such as wavelet denoising and guided filtering [4]. The experimental workflow consists of four core stages: data preparation, enhancement processing, feature extraction, and classifier training. The first set of experiments examines destructive enhancement techniques by assessing the impact of excessive processing methods such as high contrast enhancement, strong sharpening, color oversaturation, and random erasing. The second set focuses on conservative enhancement techniques, testing mild processing methods including unsharp masking, guided filtering, wavelet denoising, and CLAHE. The third set explores combined enhancement techniques, analyzing the synergistic effects of different methods. Each set of experiments uses the original images as the baseline control group to ensure consistency of experimental conditions.

2.2 Technical Principles

Destructive enhancement disrupts the natural statistical characteristics of images by significantly altering pixel distributions, excessively amplifying high-frequency components, distorting colors, and introducing large-area occlusions, thereby causing pre-trained models to fail in feature extraction. Conservative enhancement, without damaging the original structure, moderately removes noise, enhances edges, and optimizes local contrast; it is compatible with the shallow edge-detection mechanisms of CNNs and is used to identify effective preprocessing approaches. Combined enhancement serially applies techniques according to the logic of “denoising first→then enhancement” or “enhancement first→then smoothing,” thereby avoiding the limitations of single techniques and examining the synergy and conflicts that arise in multi-step processing.

2.3 Data Preparation

The experimental data are drawn from the CIFAR-10 dataset, which comprises 60,000 32×32-pixel color images across 10 categories covering common objects such as airplanes, automobiles, and birds [5]. This study randomly selects 5,000 images from the training set as experimental samples and employs stratified sampling to ensure balanced representation across categories. The dataset is partitioned into a training set (4,000 images) and a test set (1,000 images) at an 8:2 ratio. This division scheme balances data reliability and experimental efficiency. In the data preprocessing stage, all input images are uniformly resized to 224×224 pixels to match the input requirements of the ResNet-50 model. To ensure that every experiment uses identical training and test sets and thereby prevent data leakage from affecting the results, this study employs a fixed random seed (42) and index files for data partitioning.

2.4 Implementation of Enhancement Techniques

The first set of experiments implements destructive enhancement and adopts four representative methods: High contrast enhancement (parameters α=3.0, β=50; based on empirical values in image processing, medium-to-strong settings were chosen to produce clear visual effects. Linear transformation is applied to substantially stretch the pixel-value distribution, resulting in local overexposure or underexposure and thereby testing the model’s tolerance to distorted statistical characteristics). Strong sharpening (based on the Laplacian edge-detection operator and image sharpening principles, using a 3×3 convolution kernel with a center value of 17 to generate pronounced second-order derivative enhancement and excessively amplify high-frequency components). Color oversaturation (the human-perceptible saturation change range is approximately 1–5 times the original value; therefore the saturation enhancement factor is set to 3.0, distorting color distribution in HSV space). Random erasing (erasing probability of 0.8 to ensure statistical significance; erasing area ratios of 10%–60% to simulate scene occlusion).

The second set of experiments implements conservative enhancement and includes: Unsharp masking (intensity parameter set to 0.8, which lies within the literature-recommended range of 0.6–1.1 and enables edge enhancement while avoiding noise amplification and overshoot artifacts [6]). Guided filtering (radius set to 3; relevant studies indicate that for small-scale image processing, the kernel radius is typically 1–3 to balance noise suppression and detail preservation [7]; regularization parameter set to 0.01 to suppress minor noise without blurring edges). Wavelet denoising (db4 wavelet basis selected for its moderate support length, which effectively separates signal from noise and is widely used in signal and image processing; soft thresholding is applied to avoid coefficient discontinuity issues associated with hard thresholding, yielding smoother denoising results). CLAHE (contrast limited adaptive histogram equalization, with clipLimit=2.0 to prevent noise amplification caused by local over-enhancement).

The third set of experiments implements combined enhancement using four logically designed combinations of techniques from the second set: Wavelet→Unsharp (wavelet denoising followed by unsharp masking), Unsharp→Guided (unsharp masking followed by guided filtering), Wavelet→CLAHE (wavelet denoising followed by contrast enhancement), and sequential enhancement (the triple combination of guided filtering → wavelet denoising → unsharp masking).

2.5 Feature Extraction and Classification Evaluation

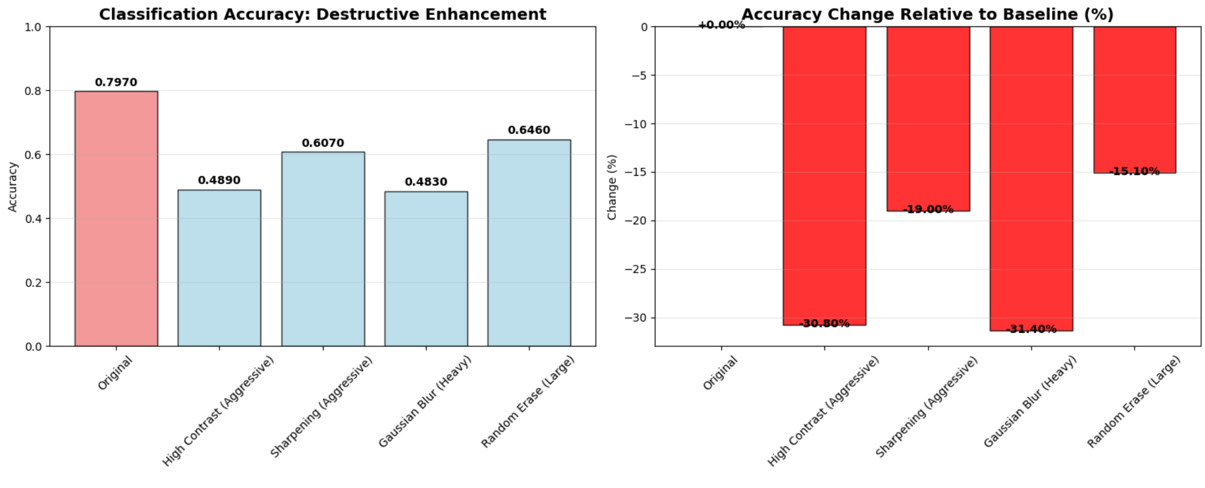
In the feature extraction stage, this study uses the ResNet-50 model pre-trained on ImageNet. The model was chosen for its efficiency in transfer learning and its deep architecture's ability to capture general visual features [8]. Specifically, the final classification layer is removed, and the 2048-dimensional feature vectors preceding the global average pooling layer are extracted as image representations. A linear-kernel SVM is selected as the classifier because of its strong interpretability and high computational efficiency [9]. Of the 5,000 experimental samples, the 4,000 training-set samples are used to train the SVM, and the 1,000 test-set samples are used for evaluation. The core performance metric is classification accuracy. The relative change rate in accuracy is calculated as: Change=(accuracy after enhancement−original accuracy)/original accuracy×100%, where positive values indicate improvement rates and negative values indicate decreases.

3. Experimental Results

3.1 Results of Destructive Enhancement

The experimental results of the destructive enhancement experiments are shown in Figure 1.

Figure 1: Accuracy and relative change percentage



As can be seen from Figure 1, the classification accuracy of the original images (control group) is 0.7970; high contrast enhancement reduces the accuracy significantly to 0.4890, a decrease of 38.64%; sharpening reduces the accuracy to 0.6070, a decrease of 23.84%; color oversaturation reduces the accuracy to 0.7690, a decrease of 3.51%; and random erasing reduces the accuracy to 0.6460, a decrease of 18.95%. The destructive enhancement techniques produce an average decrease of 21.24%, indicating that these techniques impair model performance to varying degrees. The experimental results of destructive enhancement are summarized in Table 1.

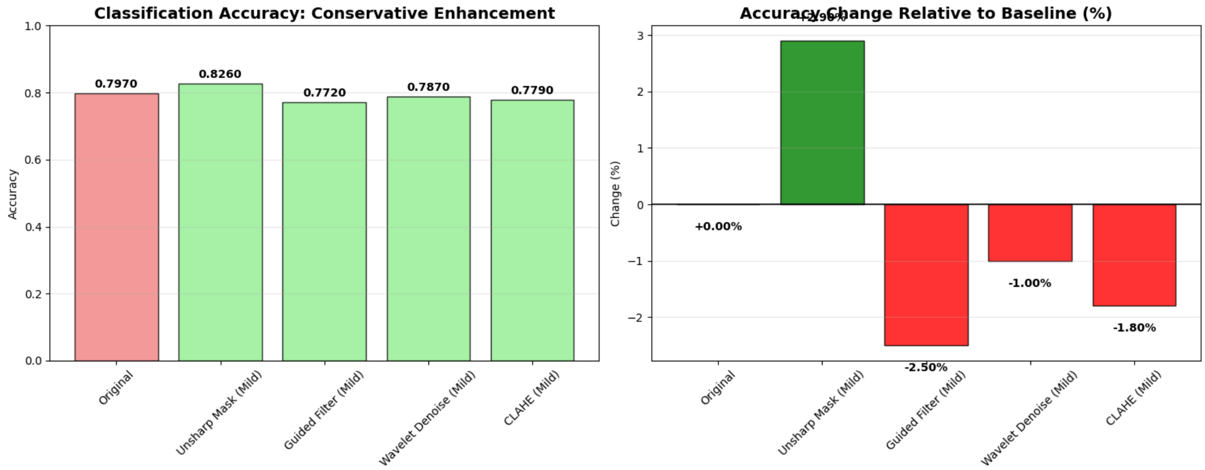
Table 1: Summary of accuracy for destructive enhancement techniques

Method	Accuracy	Change
Original	0.7970	0.00%
High Contrast	0.4890	-38.64%
Sharpening	0.6070	-23.84%
Color Oversaturation	0.7690	-3.51%
Random Erase	0.6460	-18.95%

3.2 Results of Conservative Enhancement

The experimental results of the conservative enhancement experiments are shown in Figure 2.

Figure 2: Accuracy and relative change percentage



As shown in Figure 2, the accuracy of the original images remains 0.7970; Unsharp Mask increases the accuracy to 0.8260, an improvement of 3.64%; guided filtering slightly reduces the accuracy to 0.7720, a decrease of 3.14%; wavelet denoising reduces the accuracy to 0.7870, a decrease of 1.25%; and CLAHE reduces the accuracy to 0.7790, a decrease of 2.26%. Among the conservative enhancement techniques, only Unsharp Mask produces a positive effect, while the other methods show limited impact. The experimental results of conservative enhancement are summarized in Table 2.

Table 2: Summary of accuracy for conservative enhancement techniques

Method	Accuracy	Change
Original	0.7970	0.00%
Unsharp Mask	0.8260	+3.64%
Guided Filter	0.7720	-3.14%
Wavelet Denoise	0.7870	-1.25%
CLAHE	0.7790	-2.26%

3.3 Results of Combined Enhancement

The experimental results of the combined enhancement experiments are shown in Figure 3.

The experimental results indicate that different combinations of enhancement techniques have both positive and negative effects on model performance. The Wavelet→Unsharp combination increases the accuracy to 0.8150, an improvement of 2.26%; the Unsharp→Guided combination increases the accuracy to 0.8050, an improvement of 1.00%; the Wavelet→CLAHE combination performs the worst, reducing the accuracy to 0.7500, a decrease of 5.90%; and sequential enhancement (triple combination) reduces the accuracy to 0.7770, a decrease of 2.51%. The experimental results of combined enhancement are summarized in Table 3.

Figure 3: Accuracy and relative change percentage

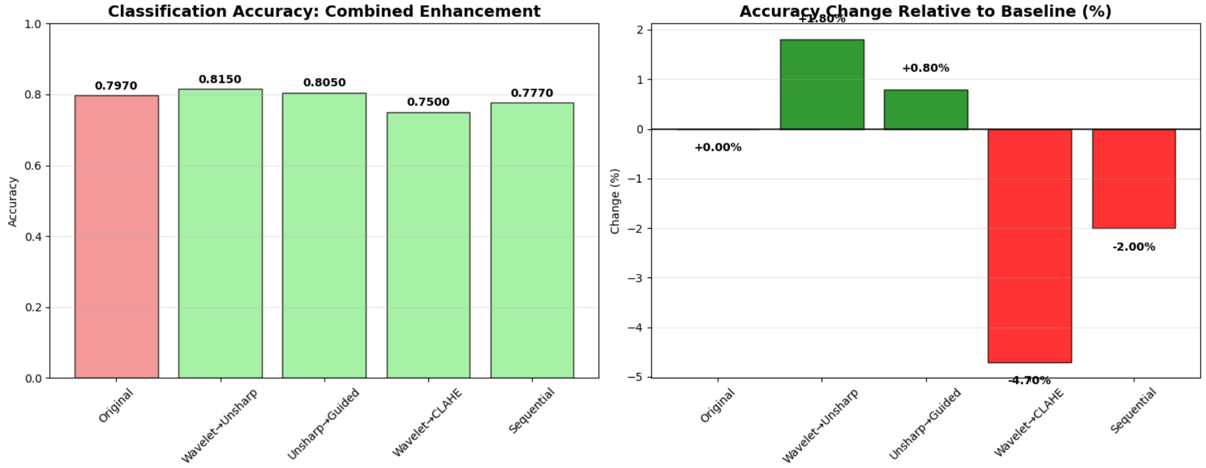


Table 3: Summary of accuracy for combined enhancement techniques

Method	Accuracy	Change
Original	0.7970	0.00%
Wavelet→Unsharp	0.8150	+2.26%
Unsharp→Guided	0.8050	+1.00%
Wavelet→CLAHE	0.7500	-5.90%
Sequential	0.7770	-2.51%

4. Discussion and Analysis

A detailed analysis of destructive enhancement reveals that the performance degradation is closely related to the destruction of the image's statistical characteristics by the enhancement techniques. High contrast enhancement excessively stretches the pixel distribution, severely distorting the image histogram. When the contrast enhancement parameter $\alpha=3.0$ is applied, pixels originally concentrated in the [0,50] and [200,255] intervals are forcibly redistributed, disrupting the inherent statistical patterns of natural images. The convolutional kernel weights in the ResNet-50 feature extractor are optimized for the specific distributions of natural images. Such a strong distributional shift causes abnormal activation patterns in the low-level feature extraction layers, which in turn affects the semantic expressive power of deeper features. Strong sharpening employs a Laplacian kernel that excessively amplifies high-frequency components in the frequency domain, enhancing real edges while also amplifying noise and artifacts. During feature extraction, the first convolutional layer is originally responsible for detecting natural edges, but the artificial edges produced by excessive sharpening differ significantly from natural edges in both direction and intensity. This causes the model to mistakenly identify noise and artifacts as real image edges, thereby interfering with feature aggregation in subsequent layers. Color oversaturation nonlinearly amplifies the saturation component to three times its original value in HSV color space, causing the chromaticity histogram distribution to deviate severely from the statistical characteristics of natural scenes. The pre-trained weights of ResNet-50 are optimized on natural image color distributions; oversaturation abnormally strengthens color features and disrupts the luminance-chrominance balance on which the model relies. Random erasing simulates extreme occlusion in a scene, but its large-area random removal destroys the semantic integrity of the image. From the perspective of the attention mechanism, when key discriminative regions in the image are randomly occluded, the model is forced to learn non-robust correlated features, resulting in reduced generalization performance on the test set. These results demonstrate that when selecting enhancement techniques, full consideration must be given to their compatibility with the pre-trained model.

Further analysis of conservative enhancement shows that the success of Unsharp Mask stems from its alignment with the hierarchical processing characteristics of convolutional neural networks. By enhancing the high-frequency components of the image, it strengthens the model's perception of contour information without introducing destructive artifacts. In the ResNet-50 feature extraction process, shallow layers are specifically responsible for edge and texture detection; moderate edge enhancement complements the function of these layers. Guided filtering exhibits a unique advantage in preserving edge smoothness, yet its limited effectiveness in certain fine-grained classification tasks may arise from excessive smoothing that causes the loss of some subtle texture information of significant value. The multi-resolution analysis property of wavelet denoising aligns strongly with the hierarchical structure of CNNs. The db4 wavelet basis provides good localization in the time-frequency domain, effectively separating signal and noise components in the image; however, the limited improvement may be offset by the model's inherent robustness. CLAHE performs histogram equalization within local regions and limits the extent of contrast enhancement through clipLimit=2.0 to avoid over-enhancement. This locally adaptive processing alleviates uneven illumination to a certain extent, but because it alters the statistical characteristics of local regions, it may lead to inconsistent feature distributions across different areas and thus produce a slight negative effect.

The effectiveness of the combined enhancement techniques in the third group depends heavily on the order and compatibility of the techniques. The Wavelet→Unsharp combination produces a synergistic effect through the logical sequence of "denoising followed by edge enhancement": wavelet denoising first removes noise interference, after which Unsharp Mask performs moderate edge reinforcement on this basis, thereby avoiding excessive noise amplification. The Unsharp→Guided combination employs a "enhancement first, smoothing later" strategy that highlights edges while maintaining overall smoothness. This processing generates more concentrated and stable activation patterns in feature space, which is beneficial for feature discrimination and extraction. In contrast, the accuracy decline caused by Wavelet→CLAHE reveals potential incompatibility between different techniques. Wavelet denoising has already reorganized the image's statistical characteristics, and the subsequent CLAHE processing further alters the local statistical distributions on this basis, resulting in a double distortion of the feature distribution. Such successive nonlinear transformations cause the image's statistical characteristics to deviate severely from those of natural images, thereby undermining compatibility with the pre-trained model. Although the triple combination of guided filtering→wavelet denoising→Unsharp embodies the hierarchical concept of image processing, it also introduces the cumulative risk of multiple nonlinear transformations. The initial smoothing by guided filtering may eliminate some subtle features, affecting the performance of subsequent wavelet denoising, while the final unsharp masking may amplify minor distortions introduced by the first two steps. The experimental results indicate that combined enhancement requires careful design based on the underlying technical principles; indiscriminate combinations may produce negative effects.

5. Conclusions

Through stratified controlled experiments, this study systematically examines the influence of three categories of image enhancement techniques on the performance of deep learning-based image classification and draws the following four core conclusions: different types of image enhancement techniques produce markedly different effects; destructive enhancement techniques, especially high contrast enhancement, severely impair model performance; conservative enhancement techniques such as Unsharp Mask can improve model performance; and the effectiveness of combined enhancement techniques depends heavily on the processing sequence. The experiments show that the impact of different techniques on model performance ranges from a decrease of 38.64% to an improvement of 3.64%. This finding highlights that technology selection is a critical factor—not a secondary one—determining the final performance of the model. Destructive enhancement causes the pre-trained feature extractor to fail because it severely deviates from the statistical distribution of natural images, resulting in an average performance decrease of 21.24%. In contrast, the Unsharp Mask technique within conservative enhancement achieves a 3.64% performance improvement because its moderate mid-to-high-frequency enhancement characteristics align precisely with the edge and texture extraction functions of the shallow layers of CNNs, thereby producing a positive effect. For low-quality image scenarios such as mine monitoring, Unsharp Mask or Wavelet→Unsharp should be prioritized, while high contrast enhancement and Wavelet→CLAHE should be avoided. It is noteworthy that only sequences that conform to the fundamental logic of signal processing—such as Wavelet→Unsharp—can generate positive synergistic effects.

Nevertheless, the generalizability of the conclusions of this study is still limited by the experimental conditions. The main limitations are manifested in three aspects: the relatively limited scale and complexity of the images in the CIFAR-10 dataset; the model architecture being restricted to ResNet-50; and the parameter selection for the enhancement techniques requiring further exploration. On this basis, future research can focus on validating the reliability of these conclusions on larger-scale datasets such as ImageNet, exploring the compatibility of enhancement techniques with more diverse network architectures, and thereby developing image-content-adaptive enhancement selection algorithms.

References

[1] Wu, Chen, & Fu, Y. H. (2025). Research on the application of image preprocessing technology in visual recognition. *China Informatization*, (4), 65-69.

[2] Wang, Z. F., Yang, T., Min, H., et al. (2025). Research on the mine fire image recognition method based on deep learning model. *China Safety Science and Technology*, 21(8), 146-154.

[3] Chen, C. G. (2024). Research on the supervised contrastive learning algorithm based on mixed image features and regional features [Doctoral dissertation, Henan University]. https://doi.org/10.27114/d.cnki.ghnau.2024.002049.

[4] Jia, W. B. (2024). Deep learning feature enhancement technology for blurred image classification [Doctoral dissertation, Lanzhou University]. https://doi.org/10.27204/d.cnki.glzhu.2024.001150.

[5] Liu, P. (2022). Design and implementation of image classification model based on neural network. *Wireless Interconnect Technology*, 19(22), 53-55.

[6] Bai, Y. J., Liu, Y., Zhang, P. C., & Gui, Z. G. (2022). Two-stage unsharp masking method based on adaptive gain coefficient. *Journal of Test Technology*, 36(5), 398-403.

[7] Liu, T., Sui, X. B., Chen, Q., & Chen, X. L. (2025). Multiscale nonuniformity correction algorithm based on rolling guided filtering in the spatio-temporal domain. *Applied Optics*, 46(2), 309-318.

[8] Shi, X. Y., Han, D., Liu, Y. Q., et al. (2025). Research on the application of ResNet50-based image self-recognition technology in urban physical examination. *Building Science*, 41(10), 249-261. https://doi.org/10.13614/j.cnki.11-1962/tu.2025.10.27.

[9] Wu, A. Q., & Han, J. (2024). Analysis and suggestions on computer network image recognition system based on machine learning. *China Science and Technology Information*, (20), 109-111.

Funding

This research received no external funding.

Conflicts of Interest

The authors declare no conflict of interest.

Acknowledgment

This paper is an output of the science project.

Copyrights