Deep Learning-Based Image Recognition of Herbaceous Tibetan Medicine Slices in Complex Backgrounds

WenPeng

Second School of Clinical Medicine, Yunnan University of Chinese Medicine, Kunming, China

ARTICLE HISTORY
Received: 02 April 2026
Accepted: 29 May 2026

KEYWORDS
Tibetan medicine slices; Complex background; Image recognition; Deep learning; Multi-feature fusion

ABSTRACT

This study proposes a novel method based on deep learning and multi-feature fusion for the recognition of herbaceous Tibetan medicine slice images in complex backgrounds. A dataset comprising 3,200 images of 32 types of Tibetan medicines was constructed and expanded to 22,400 images through data augmentation techniques. For feature extraction, the method integrates RGB color features, improved HOG (GHOG) shape features, and LBP texture features. Deep learning was conducted using an improved AlexNet model (M-AlexNet) with an attention mechanism, and a multi-feature fusion strategy was employed to optimize recognition performance. Experimental results demonstrate that the proposed method achieves a recognition accuracy of 92.31% in complex backgrounds, significantly outperforming other comparative methods. The introduction of data augmentation and GHOG features improved recognition accuracy by 1.94% and 3.42%, respectively. This study provides an effective solution for the automatic recognition of Tibetan medicine slices in complex environments, holding significant implications for research and applications in Tibetan medicine.

I. Introduction

With the advancement of Tibetan medicine, the identification and classification of Tibetan medicinal materials, as crucial components, have become increasingly important for medical research and application. Herbaceous Tibetan medicines, in particular, present significant challenges to traditional manual identification methods due to their diverse species and varied morphologies. The emergence of deep learning technologies has introduced a new solution for the automatic recognition of Tibetan medicine slices using computer vision techniques. However, the high difficulty of image recognition for Tibetan medicine slices in complex backgrounds has led researchers to turn their attention to the application of deep learning models to improve recognition accuracy and efficiency[1, 2].

Deep learning has made remarkable progress in the field of image recognition, especially with the introduction of Convolutional Neural Networks (CNNs), which have significantly enhanced the accuracy and efficiency of image classification and recognition. LeNet, proposed by LeCun et al. [3] was one of the early deep learning networks, initially used for handwritten digit recognition. Subsequently, AlexNet, introduced by Krizhevsky et al. [4] achieved breakthrough results in the ImageNet challenge, drawing widespread attention to the application of deep learning in image recognition. Since then, various deep learning models such as VGGNet, GoogLeNet, and ResNet have been proposed, demonstrating powerful performance in various image recognition tasks.

With technological advancements, deep learning has gradually been applied to medical image processing, including CT scans, X-rays, and MRI [5]. However, for image recognition of herbaceous Tibetan medicines, the complex structure, especially in slice images often accompanied by complex backgrounds and uneven lighting, poses challenges to the direct application of traditional image classification models. Consequently, researchers have begun to explore more targeted deep learning methods to address these specific issues [6, 7].

Herbaceous Tibetan medicines encompass a wide variety of species with diverse morphological characteristics, particularly in their sliced state, where texture, shape, and other features are notably complex. Additionally, during the image acquisition process, slice backgrounds are often varied and complex, potentially containing fragments of other medicinal materials, hair, dust, and other impurities.

These interfering factors increase the difficulty of image recognition [8, 9].
Furthermore, herbaceous Tibetan medicines exhibit morphological diversity within species. Even for the same medicinal material, factors such as growth environment and harvesting time can lead to significant differences in the appearance of slices. This variability makes it challenging for traditional methods based on shallow features to accurately extract useful information. Therefore, developing an automated system capable of accurately recognizing herbaceous Tibetan medicine slices in complex backgrounds has become a focal point of current research.
In recent years, substantial progress has been made in the automated recognition of herbal medicines. Some researchers have begun to utilize convolutional neural networks (CNNs) from deep learning for the classification and recognition of Chinese herbal medicine slices. For instance, Zhou et al. [10] proposed a CNN-based recognition system for Chinese herbal medicine slice images, achieving favorable results through automatic feature extraction and classification. This system employs multiple convolutional layers to gradually extract fine features of medicinal slices, combined with pooling layers and fully connected layers for classification[11, 12].
For the recognition of Tibetan medicine slices, researchers are actively exploring deep learning-based solutions. For example, Li et al. [13] proposed a Tibetan medicine slice image recognition method based on an improved ResNet model. By optimizing the ResNet network structure, they addressed the challenge of feature extraction in complex backgrounds during the recognition process. Additionally, the study incorporated data augmentation techniques such as rotation, translation, and cropping to enhance model robustness[14, 15].
Beyond convolutional neural networks, other deep learning methods such as Generative Adversarial Networks (GANs) and attention mechanisms have been applied to Tibetan medicine recognition research. For instance, Chen et al. [16]proposed a GAN-based image generation and recognition method, improving recognition accuracy in complex backgrounds by generating clear Tibetan medicine slice images. Attention mechanisms, on the other hand, adaptively focus on important regions in the image, reducing the impact of background interference on recognition results.
In identifying Tibetan medicinal material slices against complex backgrounds, preprocessing and feature extraction are crucial alongside the selection of deep learning models. Common preprocessing techniques include image denoising, background segmentation, and color space conversion. For instance, utilizing image segmentation techniques to separate slices from backgrounds effectively reduces background interference. Additionally, techniques such as illumination correction and contrast enhancement contribute to improving slice identifiability [17].
Regarding the selection of deep learning models, lightweight network architectures such as ResNet, DenseNet, and MobileNet have become predominant choices. Compared to traditional CNN models, these networks maintain high recognition accuracy while requiring lower computational complexity, making them suitable for real-time recognition on resource-constrained devices. Furthermore, researchers have introduced transfer learning techniques, applying pre-trained models from other domains to Tibetan medicine identification tasks to enhance training efficiency and recognition accuracy[18, 19].
Tibetan medicinal material slice images present significant challenges due to complex backgrounds and variable stacking formations, making it difficult to distinguish them from similar regions in the background. Moreover, environmental factors such as lighting variations and noise interference pose additional challenges to ensuring model robustness in these complex scenarios. This paper primarily investigates the multi-feature fusion of texture, shape, and color characteristics of Tibetan medicinal slices, integrating deep learning networks and attention mechanisms to apply multi-feature fusion for slice image recognition in complex backgrounds.

II. Materials and Methods

2.1 Herbal Tibetan Medical Slices in Complex Backgrounds

In the field of Tibetan medicinal slice image recognition, there is a notable absence of a publicly accessible standardized dataset, which complicates the comparative evaluation of different research methodologies under unified criteria. Previous studies have predominantly focused on datasets of individual medicinal slices obtained under controlled environments. However, in practical scenarios, medicinal slices often exist in cluttered backgrounds or stacked arrangements, limiting the applicability of previous research findings in natural conditions. To enhance the practical utility of Tibetan medicinal slice image recognition technology, this study has compiled a dataset of slice images in complex backgrounds. Figure 1.1 illustrates the comparison between images captured under controlled environments and complex backgrounds.

Figure 1.1 Comparison of Images Under Controlled Environment and Complex Backgrounds

2.2 Construction of Herbal Tibetan Medical Slice Image Dataset

During the research process, 32 types of processed root and rhizome slices from plant-based Tibetan medicines, including Rhodiola, Tibetan Gastrodia, and Tibetan Saussurea, were selected as research subjects with reference to the "Complete Collection of Chinese Tibetan Medicines." These medicinal materials hold significant importance in Tibetan medical research. For data collection, 1,000 images of Tibetan medicinal slices were captured in natural settings at the Tibet Natural Museum (Medical Hall) and specialty stores in Lhasa. Subsequently, these images underwent manual annotation to identify the category information of each slice. Due to the large original image sizes, they were standardized to 512×512 pixels using image processing software to facilitate storage and processing.

2.3 Herbal Tibetan Medical Slice Image Acquisition

For image acquisition, a Selenium-based approach was adopted, offering significant advantages over manual image downloading. Given Bing's anti-crawling mechanisms that render conventional web scraping techniques inadequate, we implemented a Selenium-based image acquisition method inspired by automated testing practices, which simulates human browsing behavior for data collection. This method was primarily employed to gather slice images with complex backgrounds from major Tibetan medicine websites and Bing Images.

The web-scraped images contained considerable erroneous data. For instance, searching for "Rhodiola slice images" yielded numerous irrelevant results. Training models on such unfiltered data would significantly impair model performance. Therefore, guided by the "human-in-the-loop" principle, we implemented an automated annotation method. This approach centered on utilizing a pre-trained ImageNet network to identify misclassified categories, followed by manual selection of 50 correct instances per category for model training. Subsequently, the trained model was used to recognize unlabeled data, with confidence threshold evaluation in the human-machine loop identifying inaccurately labeled samples for manual re-annotation.

After screening, blurred or incorrectly labeled images were removed, and all images were standardized to 512×512 pixels using image processing software. The study compiled images of 32 types of Tibetan medicinal slices, totaling 3,200 images.

2.4 Data Augmentation Methods for Herbal Tibetan Medical Slice Images

Data augmentation, a common technique in deep learning, aims to enhance model generalization by increasing training data diversity. Given the limited number of directly collected or web-scraped Tibetan medicine images, which constrains model training effectiveness and accuracy, this study implemented several innovative data augmentation methods, including brightness adjustment, random rotation, mirror flipping, and rotation with noise addition.

Through these augmentation techniques, the original 3,200 slice images were effectively expanded to 22,400 images, significantly enriching training sample diversity. To validate the augmentation effects, models were trained on both pre- and post-augmented datasets and evaluated on unseen test sets. The original dataset was randomly split into training and test sets at a 9:1 ratio to ensure experimental reliability and generalizability.

2.5 Feature Extraction Methods for Herbal Tibetan Medical Slice Images

Tibetan medicinal materials are widely distributed and diverse. Variations in growing regions, climate conditions, and processing methods result in differences in color, shape, and texture features of medicinal slices, providing crucial basis for classification and identification.
Compared to grayscale images, human visual perception is richer for color images. For scientific color information processing, quantitative methods for describing image color information are necessary through color spaces (color models). While acquired images are typically described in RGB space, HSI space better aligns with human visual perception. The HSI color space features two characteristics: 1) Hue (H) and Saturation (S) components closely correlate with human color perception; 2) Intensity (I) component is independent of color information. The HSI color space components correspond to three elements of human visual characteristics, with each channel operating independently.
As a crucial element in Tibetan medicinal slice identification, this chapter first extracts color features using RGB color feature methods. Through RGB encoding, each color's intensity is represented by red, green, and blue variables. After encoding, the RGB color space is converted to HSI color space for feature vector extraction, followed by normalization of the obtained color feature vectors.
The HOG (Histogram of Oriented Gradient) algorithm, which forms features by calculating and statistically analyzing gradient direction histograms in local image regions, is widely used for shape feature extraction. An improved HOG method (GHOG) is employed for shape feature extraction, processing images in blocks to include more information in the extracted feature vectors.
The LBP (Local Binary Patterns) texture analysis operator, first proposed by Ojala, is widely applied in texture feature extraction. For Tibetan medicinal slice images in complex backgrounds affected by lighting, angles, and mutual stacking occlusion, the LBP algorithm effectively extracts texture features, enhancing the robustness and accuracy of slice image recognition. The process begins with standardizing image sizes before texture extraction.

2.6 Complex Background Tibetan Medicinal Slice Recognition Method Based on Multi-Feature Fusion

The recognition of Tibetan medicinal slice images involves comprehensive analysis of color, shape, and texture features. Since single feature extraction methods struggle to capture these attributes comprehensively, this study proposes a multi-feature fusion strategy combining RGB, HOG, and LBP algorithms for more comprehensive image feature description. This method optimizes feature fusion by adjusting weights of different features, calculated as follows:

$$F = aF_{RGB} + bF_{HOG} + cF_{LBP}$$

where F represents the fused features, F_{RGB} represents color features, F_{IHOG} represents shape features, F_{LBP} represents texture features, and a, b, c represent their respective weight coefficients.
To determine optimal weights, a values were chosen from 0.1, 0.2, 0.3; b values from 0.3, 0.4, 0.5; and c values from 0.3, 0.4, 0.5, using orthogonal analysis to enhance experimental efficiency, as shown in Table 1.1.
Table 1.1 Optimal Weight Table

A	b	c
0.1	0.3	0.3
0.1	0.4	0.5
0.1	0.5	0.4
0.2	0.3	0.5
0.2	0.4	0.4
0.2	0.5	0.3
0.3	0.3	0.4
0.3	0.4	0.3
0.3	0.5	0.5

Regarding deep learning model selection, AlexNet was chosen over more complex networks like ResNet and GoogLeNet due to its simpler structure and fewer parameters for Tibetan medicinal slice image recognition. AlexNet's architecture consists of eight layers: five convolutional layers followed by three fully connected layers. The network employs ReLU activation functions and incorporates Dropout layers in fully connected layers to prevent overfitting. Additionally, maximum pooling and local response normalization layers enhance model generalization capability.
To further improve recognition accuracy in complex backgrounds, this study introduces an attention mechanism. This mechanism simulates human visual processing, enabling the network to focus on key information while ignoring irrelevant parts. Incorporating attention units into AlexNet (M-AlexNet),

particularly in crucial texture feature regions, effectively extracts precise feature information. The attention units primarily comprise 1×1×C convolution filters and global maximum pooling layers, facilitating feature map conversion into attention maps.

2.7 Evaluation Metrics

To comprehensively evaluate the proposed model's performance, accuracy serves as the primary evaluation metric. Accuracy is a crucial parameter measuring the model's recognition capability, reflecting its ability to correctly identify samples. The accuracy calculation formula is:

$$\mathrm{AC} = \frac{TT + TF}{TT + TF + FT + FF}$$

where TT represents true positives (correctly identified positive samples), TF represents true negatives (correctly identified negative samples), FT represents false positives (negative samples incorrectly identified as positive), and FF represents false negatives (positive samples incorrectly identified as negative).

III. Results and Discussions

3.1 Ablation Experiments

During training, the network model's recognition accuracy gradually improved and stabilized with increasing iterations. Notably, the AlexNet network model integrated with attention mechanism demonstrated higher recognition accuracy. Overall, models employing feature extraction techniques surpassed those using only deep learning networks. Specifically, the standard AlexNet network model achieved 72.40% recognition accuracy. The multi-feature fusion method achieved superior performance with 92.94% recognition accuracy. The combination of multi-feature fusion strategy with deep learning network model (M-AlexNet) demonstrated more robust performance in complex background Tibetan medicinal slice image recognition, with accuracy degradation of only approximately 1%, compared to about 3% degradation for both standard AlexNet and attention mechanism-enhanced networks. These results convincingly demonstrate the effectiveness of multi-feature fusion technology in extracting texture information from Tibetan medicinal slices in complex backgrounds.

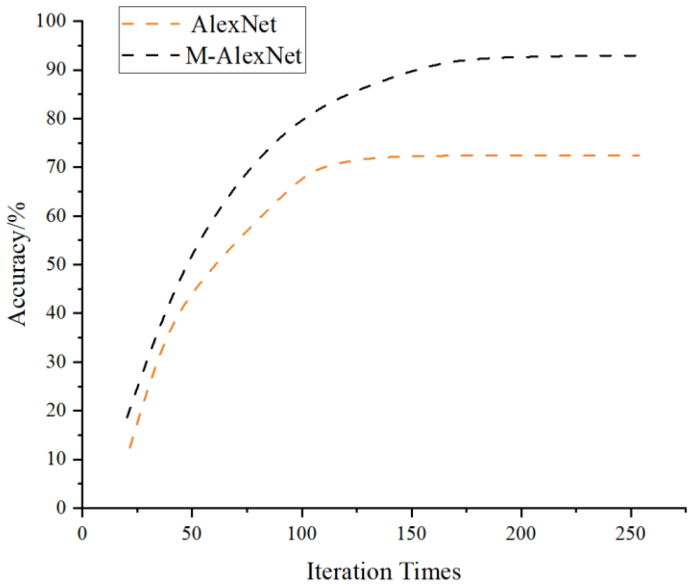

Figure 3.1 Recognition Accuracy Comparison between Multi-Feature Fusion Method and Single Methods

3.2 Feature Weight Determination

In this study, we investigated the impact of different feature weight combinations on the recognition accuracy of Tibetan medicinal slice images. Through a series of experiments, we compared recognition performance under various weight configurations and determined the optimal feature fusion weights. The experimental results are shown in Table 3.1.

Table 3.1 Optimal Weight Results

a	b	c	Accuracy
0.1	0.3	0.3	82.34%
0.1	0.4	0.5	87.67%
0.1	0.5	0.4	88.23%
0.2	0.3	0.5	91.35%
0.2	0.4	0.4	88.29%
0.2	0.5	0.3	82.17%
0.3	0.3	0.4	85.29%
0.3	0.4	0.3	83.24%
0.3	0.5	0.5	81.29%

Table 3.1 clearly demonstrates that when feature weights are set to a=0.2, b=0.3, c=0.5, the system achieves maximum recognition accuracy for Tibetan medicinal slice images under complex background conditions. This finding indicates that appropriate fusion of RGB, HOG, and LBP features enables more comprehensive image feature description. Specifically, in this optimal weight configuration, RGB features are weighted at 0.2, HOG features at 0.3, and LBP features at 0.5, highlighting the importance of texture and shape features in expressing Tibetan medicinal slice image content.

Further analysis reveals that while using RGB, HOG, or LBP features independently can provide certain image information, single features alone cannot comprehensively reflect multi-dimensional image attributes. By combining these three features with different weights, we can more precisely describe and recognize Tibetan medicinal slice images. This multi-feature fusion method not only improves recognition accuracy but also enhances model robustness in complex background image recognition.

Furthermore, this study incorporates an attention mechanism that simulates human visual processing, enabling the network to focus on key information while ignoring irrelevant parts during image processing. By incorporating attention units into AlexNet (M-AlexNet), particularly in crucial texture feature regions, we can more effectively extract precise feature information. The attention units, primarily comprising 1×1×C convolution filters and global maximum pooling layers, facilitate feature map conversion into attention maps, thereby enhancing overall model performance.

3.3 Comparison between HOG and GHOG Methods

The HOG algorithm is a classical image feature extraction method that forms features by calculating gradient direction histograms in local image regions, widely applied in shape feature extraction. To further enhance shape feature expression capability, this paper introduces an improved HOG method, termed GHOG. In the GHOG method, we process images in blocks to ensure extracted feature vectors contain more information. This approach better captures local details and shape features in images. To validate GHOG method effectiveness, we conducted a series of experiments applying both original HOG and improved GHOG methods using optimal weights. The experimental results are shown in Figure 3.2.

The results demonstrate that compared to the original HOG method, the GHOG method, when fused with other features, significantly improved recognition accuracy of Tibetan medicinal slice images in complex backgrounds by 3.42%. This result indicates that the GHOG method has clear advantages in extracting richer shape features, providing a more effective solution for image recognition tasks in complex backgrounds.

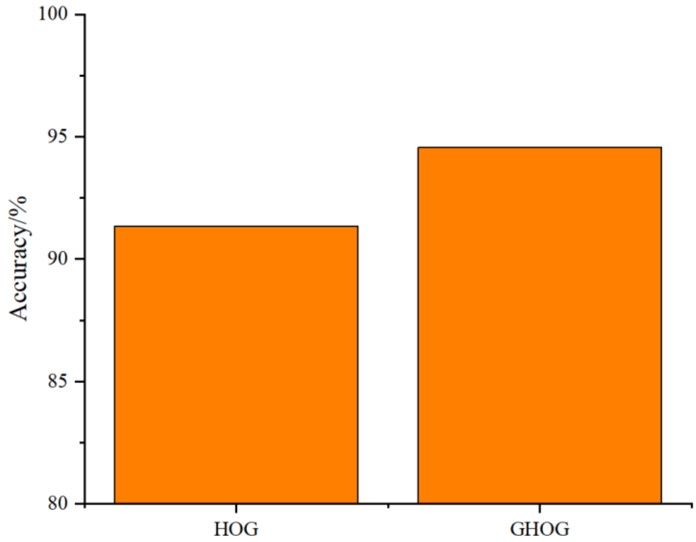

Figure 3.2 Comparison Results between HOG and GHOG Methods

3.4 Data Augmentation

This experiment validated the effectiveness of data augmentation by training the proposed model on an augmented dataset and testing it using original data images. As shown in Figure 3.3, data augmentation improved recognition accuracy of Tibetan medicinal slice images in complex backgrounds by 1.94%. This result indicates that data augmentation can reduce network overfitting to some extent while enhancing model generalization ability and robustness.

Furthermore, data augmentation not only increased the quantity of training samples but also introduced more variations and noise, enabling the model to better adapt to different input scenarios. This improvement is particularly crucial when processing Tibetan medicinal images with complex backgrounds, which typically contain various interfering factors such as lighting variations and background clutter. Through data augmentation, the model can learn from more diverse training environments, thereby improving its robustness and accuracy in practical applications.

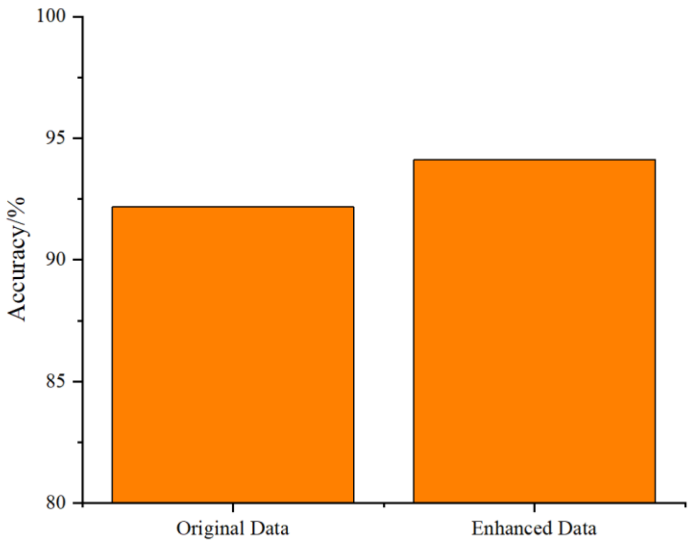

Figure 3.3 Comparison Results between Original and Augmented Data

3.5 Complex Background Tibetan Medicinal Slice Recognition

The performance of our proposed model was evaluated and analyzed through comparison with other models. These experiments aimed to compare performance differences between our model and various advanced network methods currently used for complex background medicinal slice recognition. Experimental results are shown in Figure 2.4, clearly demonstrating our model's superior accuracy compared to other algorithms. Specifically, the RGB-SVM model [20] performed worst among all compared models, with only 62.12% accuracy. This poor performance can be attributed to the model's sole reliance on image RGB features for classification. However, in complex background images, RGB features are easily affected by background variations, lacking stability and reliability. Consequently, the RGB+SVM model performed unsatisfactorily when processing Tibetan medicinal slice images with complex backgrounds. The MobileNet V2 model [21] achieved 67.12% accuracy, slightly higher than RGB+SVM but still suboptimal. This may be due to MobileNet V2's heavy dependence on optimization strategies, which limited its performance in the complex background dataset.

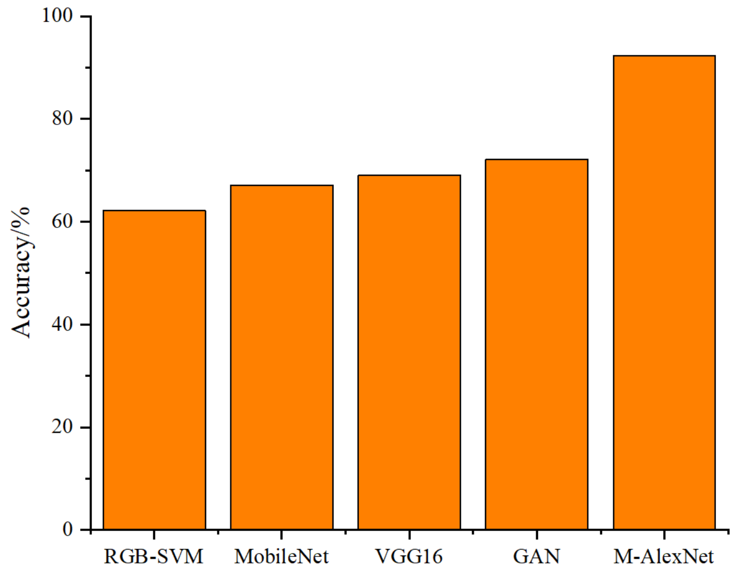

Figure .4 Recognition Accuracy of Different Models

Other models such as VGG16[22] and GAN[23], despite achieving high accuracy in identifying Chinese medicine slices under ideal conditions, failed to meet expectations under this study's complex background conditions. This indicates these models' limitations in handling challenging complex backgrounds.

In contrast, our proposed model demonstrated the highest image recognition accuracy at 92.31% in comparative experiments. By comprehensively considering multiple features and leveraging deep learning capabilities, our model can more accurately distinguish and recognize Tibetan medicinal slices in complex backgrounds, demonstrating significant value for practical applications.

V. Conclusion

This study proposed a novel method based on deep learning and multi-feature fusion for herbal Tibetan medicinal slice image recognition in complex backgrounds. Through a series of experiments and analyses, we drew the following conclusions:

1. Effectiveness of Multi-Feature Fusion Strategy: The multi-feature fusion method combining RGB color features, improved HOG (GHOG) shape features, and LBP texture features significantly improved recognition accuracy. Experimental results showed optimal performance with feature weights of RGB:HOG:LBP = 0.2:0.3:0.5, achieving 91.35% recognition accuracy under complex background conditions.
2. Deep Learning Model Improvements: The modified AlexNet (M-AlexNet) with attention mechanism further enhanced recognition performance. Compared to standard AlexNet, M-AlexNet demonstrated stronger robustness in complex backgrounds, with accuracy degradation of only about 1% versus 3% for standard AlexNet.
3. Importance of Data Augmentation: Various data augmentation techniques expanded the original dataset from 3,200 to 22,400 images, improving model generalization ability. Results showed a 1.94% accuracy improvement after data augmentation, confirming its effectiveness in reducing overfitting and enhancing model robustness.
4. GHOG Feature Advantages: The improved HOG method (GHOG) excelled in shape feature extraction. Compared to original HOG, GHOG method improved recognition accuracy by 3.42% after feature fusion, demonstrating superiority in capturing local details and shape features.
5. Comparison with Other Methods: Our proposed method achieved 92.31% accuracy in complex background Tibetan medicinal slice image recognition, significantly outperforming other comparative methods such as RGB-SVM (62.12%) and MobileNet V2 (67.12%).

In conclusion, our proposed method based on multi-feature fusion and improved deep learning models achieved significant results in herbal Tibetan medicinal slice image recognition under complex backgrounds. This method not only improved recognition accuracy but also enhanced model robustness to complex backgrounds, providing strong technical support for Tibetan medicine research and applications.

Disclosure Statement

No potential conflict of interest was reported by the author.

Data Availability Statement

Data available on request from the author.

References

[1] DENG J, CHEN Y, ZHANG X, et al. Intelligent supervision of PIVAS drug dispensing based on image recognition technology. %J PloS one [J]. 2024, 19(4): e0298109-e.
[2] HUSSAENI K A, KARAMITSOS I, ADEWUMI E, et al. CNN-Based Pill Image Recognition for Retrieval Systems %J Applied Sciences [J]. 2023, 13(8).
[3] LECUN Y, BOTTOU L, BENGIO Y, et al. Gradient-based learning applied to document recognition [J]. Proceedings of the IEEE, 1998, 86(11): 2278-324.
[4] KRIZHEVSKY A, SUTSKEVER I, HINTON G. ImageNet Classification with Deep Convolutional Neural Networks [J]. Neural Information Processing Systems, 2012, 25.
[5] LEE H, PARK D-S, CHOI M-H. A Development of Hybrid Drug Information System Using Image Recognition %J Symmetry [J]. 2015, 7(2): 376-82.
[6] LIYUAN Z, HONGMEI G, DINGGUO G, et al. Recognition of Ellipsoid-like Herbaceous Tibetan Medicinal Materials Using DenseNet with Attention and ILBP-Encoded Gabor Features. %J Entropy (Basel, Switzerland) [J]. 2023, 25(6).
[7] NANDYAL S S, ANAMI B S, GOVARDHAN A, et al. A method for identification and classification of medicinal plant images based on level set segmentation and SVM classification %J Int. J. of Computational Vision and Robotics [J]. 2012, 3(1/2): 96-114.
[8] NATIONAL LIBRARY OF MEDICINE N I O H, BETHESDA, MD, USA., TAJ TECHNOLOGIES INC. M H, MN, USA., NATIONAL LIBRARY OF MEDICINE N I O H, BETHESDA, MD, USA., et al. The National Library of Medicine Pill Image Recognition Challenge: An Initial Report. %J IEEE Applied Imagery Pattern Recognition Workshop : [proceedings]. IEEE Applied Imagery Pattern Recognition Workshop [J]. 2016, 2016.
[9] QIU Q, QIN Q, SHANSHAN C, et al. Pointer instrument image recognition system for veterinary drug production based on RFID and deep learning %J Journal of Physics: Conference Series [J]. 2020, 1650(3): 032139-.
[10] ZHOU G, GUO Z, SUN S, et al. A CNN-BiGRU-AM neural network for AI applications in shale oil production prediction [J]. Applied Energy, 2023, 344: 121249.
[11] Technology - Materials Research; Cheng Shiu University Reports Findings in Materials Research (Development of Smart Pillbox Using 3d Printing Technology and Convolutional Neural Network Image Recognition) %J Journal of Engineering [J]. 2020.
[12] SHIN K, LEE T, KIM J, et al. Deep Learning for Herbal Medicine Image Recognition: Case Study on Four-herb Product %J 한국자원식물학회 학술심포지엄 [J]. 2019: 87-.
[13] LI Y, ZHANG Y, XU Y, et al. Does ResNet Learn Good General Purpose Features?; proceedings of the AIACT '17, F, 2017 [C].
[14] XIAOBO Z, XIAOGUANG G, YAN J, et al. [Application of image recognition technology in census of national traditional Chinese medicine resources]. %J Zhongguo Zhong yao za zhi = Zhongguo zhongyao zazhi = China journal of Chinese materia medica [J]. 2017, 42(22): 4266-70.
[15] XIN S, HUINAN Q. Chinese Herbal Medicine Image Recognition and Retrieval by Convolutional Neural Network. %J PloS one [J]. 2016, 11(6): e0156327.
[16] JAIN H, DHUPPER R, SHRIVASTAVA A, et al. AI-enabled strategies for climate change adaptation: protecting communities, infrastructure, and businesses from the impacts of climate change [J]. Computational Urban Science, 2023, 3(1): 25.
[17] YIYI M, YANMEI Z, QIN S, et al. Study on identification algorithm of traditional Chinese medicinals microscopic image based on convolutional neural network. %J Medicine [J]. 2023, 102(25): e34085-e.
[18] XIN S, HUINAN Q, YILIANG X, et al. Deep learning-enabled mobile application for efficient and robust herb image recognition %J Scientific Reports [J]. 2022, 12(1): 6579-.
[19] XUE S. Bar Image Identification Based on Laplacian Eigenmap and Fuzzy SVM %J International Journal of e-Collaboration (IJeC) [J]. 2023, 19(5): 1-13.

[20] AGRAWAL S, VERMA N K, TAMRAKAR P, et al. Content Based Color Image Classification using SVM; proceedings of the 2011 Eighth International Conference on Information Technology: New Generations, F 11-13 April 2011, 2011 [C].
[21] OJALA T, PIETIKAINEN M, MAENPAA T. Multiresolution gray-scale and rotation invariant texture classification with local binary patterns [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2002, 24(7): 971-87.
[22] FATANIYA B, JOSHI M J, MODI U, et al. Automatic Identification of Licorice and Rhubarb by Microscopic Image Processing [J]. 2015, 58: 723-30.
[23] TU Z, TALEBI H, ZHANG H, et al. MaxViT: Multi-Axis Vision Transformer; proceedings of the European Conference on Computer Vision, F, 2022 [C].