Deep Learning for Saffron Adulteration Detection from Stigma Images: A Comparative Study of AlexNet, ResNet, and VGG16

Mitra Gholami
Urmia University
Rahman Farrokhi Teimourlou
r.farrokhi@urmia.ac.ir

Urmia University
farzaneh jannatdoust
Urmia University

Article

Keywords: Saffron, Authentication, Deep learning, Non-destructive quality assessment, Computer vision, Postharvest quality control

Posted Date: July 20th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-10292094/v1

Additional Declarations: No competing interests reported.

Deep Learning for Saffron Adulteration Detection from Stigma Images: A Comparative Study of AlexNet, ResNet, and VGG16

Mitra Gholami[1], Rahman Farrokhi Teimourlou[2*], Farzaneh Jannatdoust[3]

Abstract

Saffron (Crocus sativus L.) is a high-value spice prone to economically motivated adulteration due to its distinctive morphology and premium price. This study investigates the potential of convolutional neural networks (CNNs) as a rapid, non-destructive, and cost-effective tool for distinguishing authentic saffron stigmas from visually counterfeit materials. A curated dataset of 1,020 expert-labeled images (719 authentic, 301 counterfeit) was used to evaluate three CNN architectures—AlexNet, ResNet, and VGG16—through transfer learning. The dataset was split into 80% training and 20% validation, with a separate test set for final evaluation. AlexNet achieved the highest accuracy (96.49%) with perfect sensitivity (100%) and high specificity (93.75%), outperforming ResNet (94.74%) and VGG16 (87.72%). These results demonstrate the feasibility of CNN-based visual authentication for saffron and provide a foundation for practical deployment in postharvest quality control pipelines. Future work should explore larger datasets, domain generalization, and lightweight on-device inference for real-world applications.

Keywords: Saffron; Authentication; Deep learning; Non-destructive quality assessment; Computer vision; Postharvest quality control

1. Introduction

Saffron (*Crocus sativus* L.) is recognized as the world's most valuable spice due to its unique biochemical composition, sensory attributes, and labor-intensive harvesting process. The stigma of saffron contains bioactive compounds such as crocin, picrocrocin, and safranal, which confer its commercial and pharmacological value [1-5]. Iran produces more than 90% of the global saffron supply, and therefore, the authenticity, quality, and postharvest handling of saffron have significant implications for the national economy, export standards, and international trade [6].

Due to its high price and complex supply chain, saffron is particularly susceptible to economically motivated adulteration (EMA). Fraudulent materials—including dyed fibers, foreign plant tissues, synthetic filaments, and low-grade botanical substitutes—are often mixed with authentic stigmas to increase bulk weight or mimic genuine color and morphology [7,8]. Such practices negatively

[1]. Master degree Graduated, Department of Mechanical Engineering of Biosystem, Faculty of Agriculture, Urmia University, Urmia, Iran (m.gholami@urmia.ac.ir)
[2*]. Associate Professor, Department of Mechanical Engineering of Biosystem, Faculty of Agriculture, Urmia University, Urmia, Iran (r.farrokhi@urmia.ac.ir)
[3]. Phd candidate, Department of Mechanical Engineering of Biosystem, Faculty of Agriculture, Urmia University, Urmia, Iran (f.jannatdoust@urmia.ac.ir)

impact postharvest quality, alter chemical and colorimetric properties, weaken consumer confidence, and may violate international food safety regulations [9].

Conventional laboratory techniques such as high-performance liquid chromatography (HPLC), gas chromatography–mass spectrometry (GC–MS), UV–Vis spectroscopy, and FT-IR analysis have demonstrated excellent accuracy for saffron authentication. However, these methods are often destructive, time-consuming, costly, and require specialized infrastructure and skilled personnel—limiting their applicability for rapid screening during postharvest grading, packaging, and export inspections [10,11]. In contrast, nondestructive sensing systems and imaging technologies offer rapid, on-site assessment capabilities suitable for high-value crops [12].

Classical machine-vision techniques have attempted to discriminate saffron quality based on color, texture, shape, and morphological descriptors [13]. However, these approaches depend heavily on handcrafted features that may not generalize well under variable lighting conditions, camera settings, or natural variability in stigma morphology. This limitation has motivated a shift toward deep learning, which enables automated feature extraction and robust classification for agricultural products.

Convolutional neural networks (CNNs) such as AlexNet, VGG16, and ResNet have reshaped the landscape of image-based quality assessment by providing superior performance in defect detection, produce grading, disease diagnosis, and nondestructive authenticity verification [14,15]. Their integration into postharvest systems has been particularly beneficial for high-value crops that require rapid and accurate quality monitoring.

Despite these advances, deep-learning-based saffron authentication remains insufficiently explored in the scientific literature. Many existing studies focus on chemical profiling or small-scale machine-vision datasets, and few have adopted a systematic comparison of multiple CNN architectures under unified experimental conditions. Furthermore, the potential application of AI-driven authentication in real postharvest operations—including sorting, grading, and export-level quality monitoring—has not been adequately addressed [16,17].

To address these gaps, the present study introduces a curated and expert-labeled dataset of authentic and counterfeit saffron stigmas captured under controlled imaging conditions. Three widely used CNN architectures—AlexNet, VGG16, and ResNet—are evaluated using a standardized transfer-learning pipeline to determine their effectiveness in detecting adulteration.

By integrating deep learning into the context of postharvest quality assurance, this study demonstrates how rapid, nondestructive, and scalable AI-based screening systems can be incorporated into saffron supply chains to enhance authenticity verification, reduce fraud risk, and maintain export-grade quality. Through a comprehensive comparison of CNN models, the findings contribute both to the scientific understanding of saffron adulteration and to the development of practical tools aligned with postharvest technology requirements.

2. Materials and Methods

2.1. Dataset and study design

The authentic saffron samples were sourced from major Iranian producers and classified into three commercial grades (Sargol, Negin, Pushal) under the supervision of an expert, reflecting real-world market variability [18,19]. The counterfeit samples were meticulously prepared in the laboratory to mimic common visual adulteration scenarios reported in literature, such as the use of dyed corn silk, safflower stamens, and colored plant roots [7] (Figure 1). A total of 1,020 RGB images were acquired under controlled illumination, comprising 719 authentic samples and 301 counterfeit samples (Figure 2 & Table 1). Each physical sample was photographed only once to prevent data leakage, ensuring that no images from the same specimen appeared in multiple subsets. The dataset was randomly partitioned into an 80% training set (816 images) and a 20% validation set (204 images). Since each physical sample was photographed only once, this partitioning effectively prevents data leakage between the training and validation sets. Additionally, a separate test set consisting of mixed authentic and counterfeit samples was prepared for final model evaluation, as shown in Figure 2.

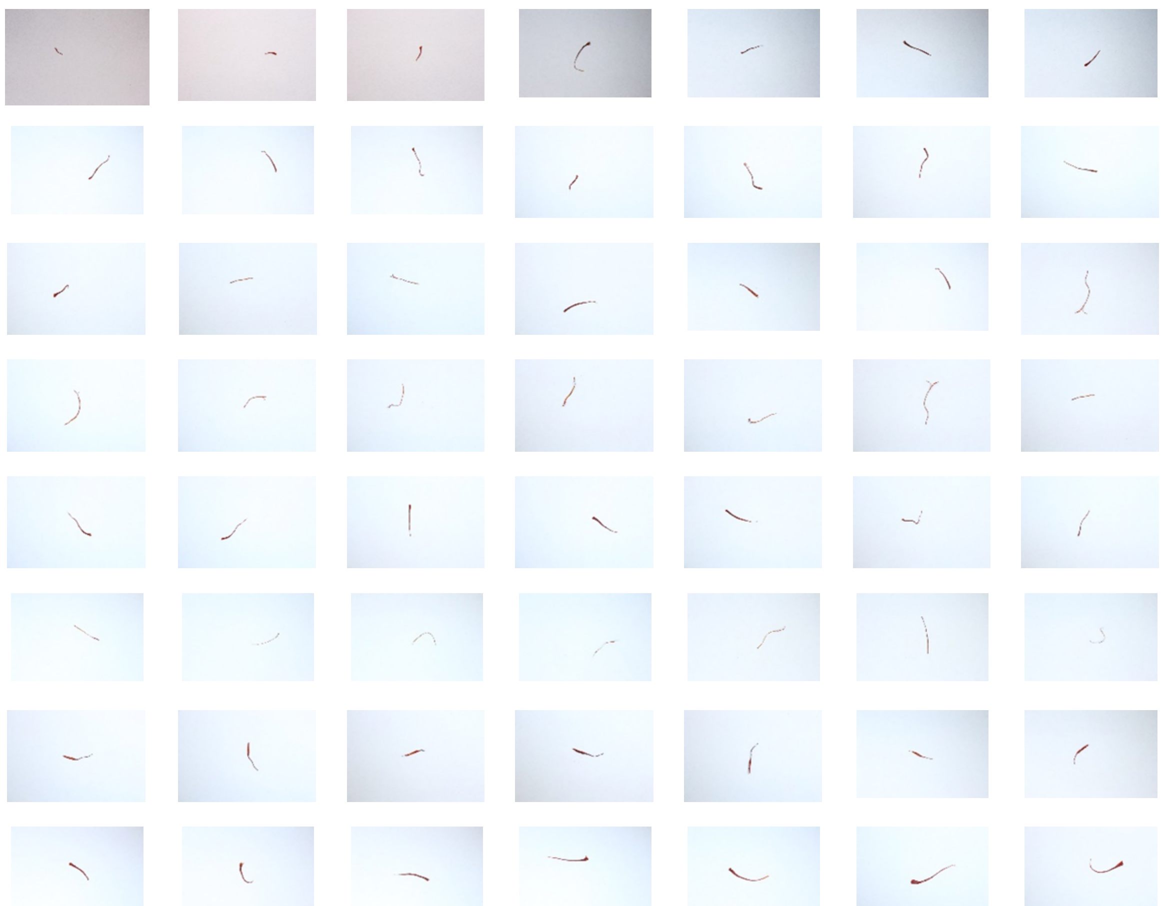

Figure 1. *a) saffron filament base(pushal) type1 b) saffron filament base(pushal) type2 c) saffron filament base(pushal) type3 d) Sargol saffron e) Negin saffron f) colored saffron root g) dyed corn silk h) colored safflower*

Figure 2. *mixture of original and fake samples for network testing*

Table 1. *Dataset composition*

Class	Count	Notes
Authentic (all grades)	719	Three expert-annotated grades
Counterfeit	301	Lab-prepared visual adulterants
Total	1,020	80/20 train/validation split

2.2. Image acquisition and preprocessing

Images were captured using a Canon 700D DSLR camera with a standard 18-55mm lens. Two preprocessing stages were applied to enhance image quality and reduce background artifacts. In the first stage, 10% of the border margins (top, bottom, left, and right) were removed (Figure 3), and a binary mask was applied to isolate the saffron filaments from the background. In the second stage, background homogenization was performed using RGB thresholding, where pixels with red, green, and blue values exceeding 180 were identified as background and replaced with a uniform white background. This ensured consistent lighting conditions across all images and prevented the network from learning spurious background patterns. All images were then resized to the input dimensions required by each CNN backbone (e.g., 227×227 for AlexNet) and normalized channel-wise (Figure 4) [14].

***Figure 3**. Output of first preprocessing: a) input image, b) cropped input image*

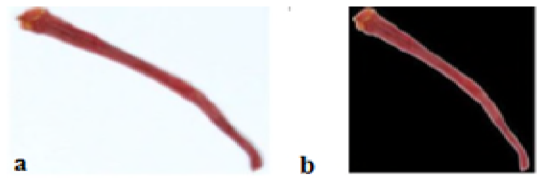

Figure 4. *Output of second preprocessing: a) input image, b) output image*

Figure 5 illustrates nine representative input samples of saffron stigmas after preprocessing, ready to be fed into the AlexNet architecture. Figure 6 demonstrates the manual labeling process for validation images, showing examples of both authentic (**a-c**) and counterfeit (**d**) samples.

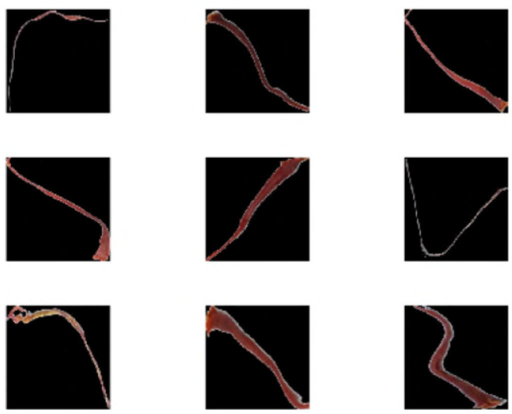

Figure 5 . 9 images of input samples for AlexNet

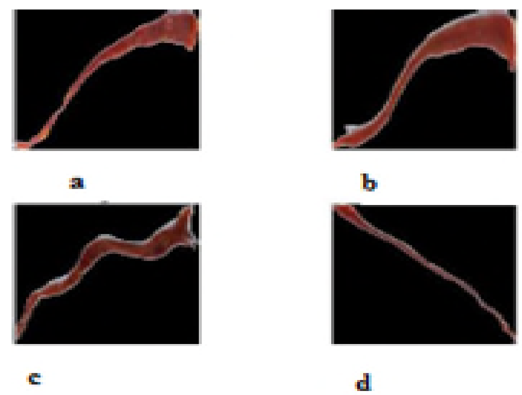

Figure 6 .Test and Labeling the validation images: ***a,b,c)*** *original and* ***d)*** *fake.*

2.3. Data augmentation

To enhance model robustness and invariance to pose and capture variations, an online data augmentation strategy was employed during training. This included random rotations over the full 0–360° range and random horizontal/vertical flips. The augmentation was applied in real-time (online) during the training process, which was particularly suitable for our dataset size and prevented the computational overhead of offline augmentation. This approach effectively increases data diversity and reduces the risk of overfitting, especially for moderate-sized datasets [23]. Validation images were not augmented and were only resized. The separate test set was also presented without augmentation to evaluate real-world performance.

2.4. CNN Architectures and Transfer Learning

We evaluated three canonical pre-trained CNNs: AlexNet [14], VGG16 [20], and ResNet [15]. These architectures were selected to represent a spectrum of model complexity and design philosophies, from the pioneering and efficient AlexNet to the very deep and residual connection-based ResNet. For each network, the final classification layer was replaced with a two-unit fully connected layer followed by a softmax activation function. Earlier layers were initialized with weights pre-trained on the ImageNet dataset and were fine-tuned end-to-end using stochastic gradient descent with momentum, a standard practice in transfer learning for computer vision [21,22]. All models were trained on a CPU-based system (Intel Core i5-6200U, 8GB RAM) due to hardware limitations, which particularly affected the convergence speed of the deeper VGG16 architecture.

2.5. Training protocol and implementation

The hyperparameters used for training are summarized in Table 2. A fixed training schedule of 50 epochs was used with a mini-batch size of 64 and an initial learning rate of 1e-4. Training was performed using stochastic gradient descent with a momentum of 0.9.

Table 2. Hyperparameters of the deep convolutional neural network

Parameter	Value
Mini-batch size	64
Epochs	50
Initial learning rate	0.0001
Momentum	0.9
Number of architectures	3
Optimization function	Stochastic Gradient Descent
Loss Function	Categorical Cross-Entropy
Software	MATLAB R2020a

All models were implemented in MATLAB R2020a (MathWorks, Inc.) using the Deep Learning Toolbox. The loss function optimized during training was categorical cross-entropy. The training was performed on a single CPU (Intel Core i5-6200U, 2.40 GHz) with 8GB RAM and Windows 10 Pro operating system. Model selection was based on the highest validation accuracy observed during training. Due to the stochastic nature of the random data split, the validation results represent performance on a single random partition of the dataset.

2.6. Evaluation Metrics and Statistical Analysis

Performance was evaluated using standard metrics (equations 1-5) for binary classification: Accuracy, Sensitivity (Recall for the counterfeit class), Specificity (Recall for the authentic class), Positive Predictive Value (PPV), and Negative Predictive Value (NPV). These metrics were computed from the confusion matrix derived from the held-out validation set [19].

$$Accuracy = \frac{(TP + TN)}{(TP + TN + FP + FN)} \tag{1}$$

$$Sensitivity = \frac{(TP)}{(TP+FN)} \tag{2}$$

$$Specificity = \frac{(TN)}{(TN+FP)} \tag{3}$$

$$Precision(PPV) = \frac{(TP)}{(TP+FP)} \tag{4}$$

$$Negative\ Predictive\ Value = \frac{(TN)}{(TN+FN)} \tag{5}$$

Therefore, according to Equation (1), higher sensitivity and specificity lead to greater overall accuracy of the proposed model. The results of these five performance parameters, along with the confusion matrix, are reported for all three models. It should be noted that the confusion matrices presented in Figures 7-9 represent a random subset of the validation set for illustrative purposes, while the metrics in Table 3 reflect the complete validation set.

3. Results and Discussion

3.1. Overall classification performance

The performance of the three CNN models is summarized in Table 3 and visualized via their confusion matrices in Figures 7, 8, and 9. Table 3 demonstrates that the adapted CNNs demonstrated a strong capability to discriminate between authentic and counterfeit saffron stigmas. AlexNet delivered the best overall performance with an accuracy of 96.49%, achieving a perfect sensitivity of 100% (correctly identifying all counterfeit samples) and a specificity of 93.75%. This was followed by ResNet (94.74% accuracy) and VGG16 (87.72% accuracy). The high performance of these models, particularly AlexNet, indicates that even relatively compact architectures can effectively capture the salient morphological and chromatic cues that distinguish genuine stigmas from visually crafted adulterants.

Table 3. *Comparative performance of the proposed CNN architectures on the saffron adulteration test set.*

Model	*Accuracy (%)*	*Sensitivity (%)*	*Specificity (%)*	*Precision (%)*	*NPV (%)*
AlexNet	**96.49**	**100**	93.75	92.59	100
ResNet-50	94.74	90.00	**100**	**100**	76.66
VGG16	87.72	**100**	79.41	100	100

The confusion matrix for the AlexNet model (Figure 7) demonstrates that it emerged as the top-performing model, achieving the highest overall accuracy (96.49%) and perfect sensitivity (100%), meaning it correctly identified all adulterated samples in the validation subset. Its slightly lower specificity (93.75%) indicates a very small number of authentic samples were misclassified as adulterated. This high performance can be attributed to its architectural efficiency, which is well-suited for the specific texture and color features of saffron and its adulterants, without being overly complex for the dataset size.

		Actual Values: authentic	Actual Values: adulterated	
Predicted Values	authentic	TP 25	FP 2	PPV 92.59% 7.41%
Predicted Values	adulterated	FN 0	TN 30	NPV 100% 0%
		Sensitivity 100% 0%	Specificity 93.75% 6.25%	Accuracy 96.49% 3.51%

***Figure 7**. Confusion matrix for the AlexNet model.*

The confusion matrix in Figure 8 showed ResNet-50's excellent specificity (100%) and precision (100%), indicating that every sample it predicted as "authentic" was indeed authentic. However, its sensitivity was lower (90%), meaning it misclassified 10% of the adulterated samples as authentic. This could be due to the model's greater depth, which may require more data to fully optimize for this specific task or might be learning more generalized features that are less sensitive to the specific adulterants used.

		Actual Values		
		authentic	adulterated	
Predicted Values	authentic	TP 27	FP 0	PPV 100% 0%
	adulterated	FN 3	TN 27	NPV 90% 10%
		Sensitivity 90% 10%	Specificity 100% 0%	Accuracy 94.74% 5.26%

Figure 8. *Confusion matrix for the ResNet-50 model.*

		Actual Values		
		authentic	adulterated	
Predicted Values	authentic	TP 27	FP 0	PPV 100% 0%
	adulterated	FN 7	TN 23	NPV 76.66% 23.34%
		Sensitivity 100% 0%	Specificity 79.41% 20.59%	Accuracy 87.72% 12.28%

Figure 9. *Confusion matrix for the VGG16 model*

Figure 9 showed that VGG16 yielded the lowest accuracy (87.72%). While it achieved perfect sensitivity (100%), its specificity was significantly lower (79.41%). This suggests that VGG16 struggled the most with correctly identifying authentic samples, often misclassifying them as adulterated. This is likely due to its very deep and simple architecture with a massive number of parameters, making it prone to overfitting, especially on a dataset of this size, and computationally heavier to train effectively on the available hardware.

3.2. Accuracy–efficiency trade-offs and convergence behavior

AlexNet's superior performance in this study (Figure 10) is closely aligned with its favorable optimization dynamics on moderately sized, texture-rich datasets. Its relatively shallow depth and larger early convolutional kernels enable rapid extraction of filament curvature, branching geometry, and terminal morphology, allowing the network to converge quickly without overfitting to background artifacts.

In contrast, VGG16’s substantially larger parameter count and the CPU-bound training environment (due to hardware limitations) limited the number of effective weight-update steps. As a result, the model showed signs of under-convergence, which is reflected in the lower validation accuracy and the persistent train–validation gap observed in the learning curves (Figure 12). This indicates that the performance difference between VGG16 and the other models may be attributed to computational constraints rather than intrinsic architectural limitations.

ResNet, benefiting from residual connections, exhibited more stable gradient flow throughout training (Figure 11). Its learning curves indicate a slower but consistent ascent toward higher accuracy, approaching the performance of AlexNet. With longer training schedules or stronger regularization procedures, ResNet may match or even exceed AlexNet in absolute accuracy. We emphasize that these observations are contingent on the compute budget and dataset scale rather than intrinsic architectural limitations.

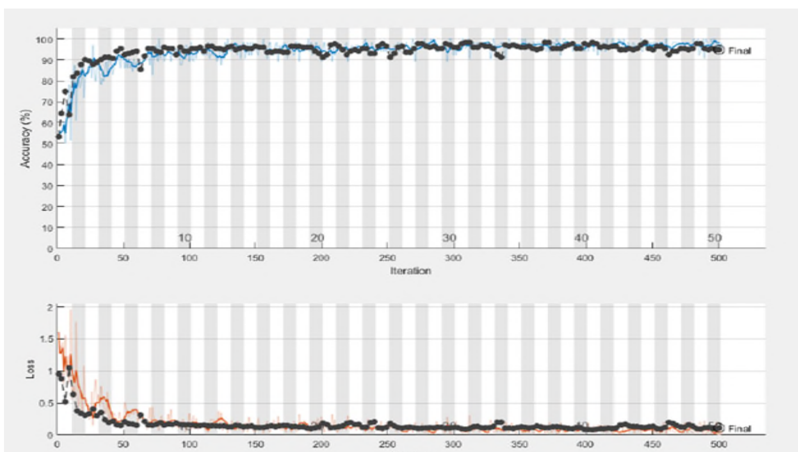

***Figure 10**. Training and validation accuracy per epoch for the AlexNet, illustrating convergence rate and generalization gap.*

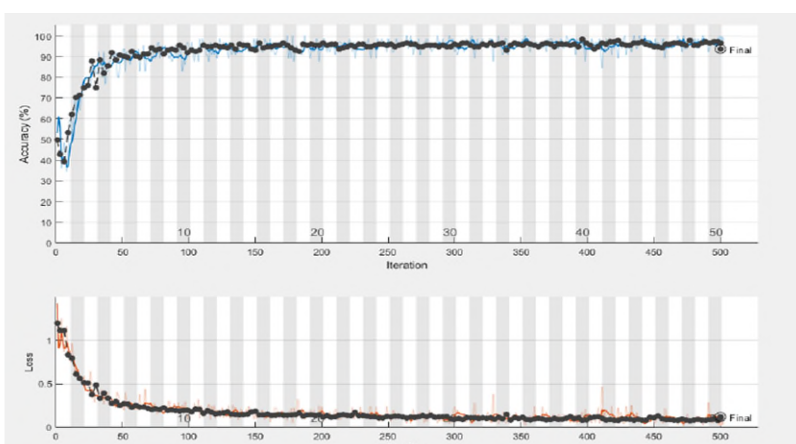

***Figure 11.** Training and validation accuracy per epoch for the ResNet, illustrating convergence rate and generalization gap.*

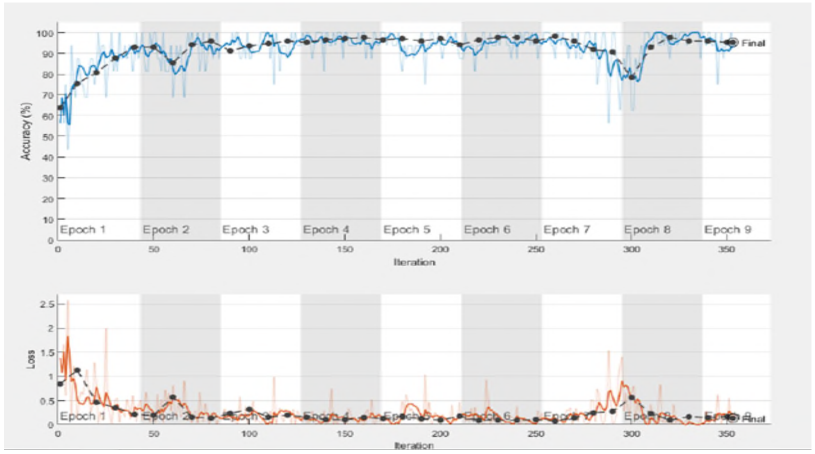

Figure 12. *Training and validation accuracy per epoch for the VGG16, illustrating convergence rate and generalization gap.*

Overall, the training trends in Figures 10–12 illustrate rapid stabilization for AlexNet within the first 15–20 epochs, steady progressive improvement for ResNet, and plateauing for VGG16, indicating capacity underutilization due to computational constraints rather than architectural limitations.

Examples of the output generated by these models are presented in Figures 13, 14, and 15:

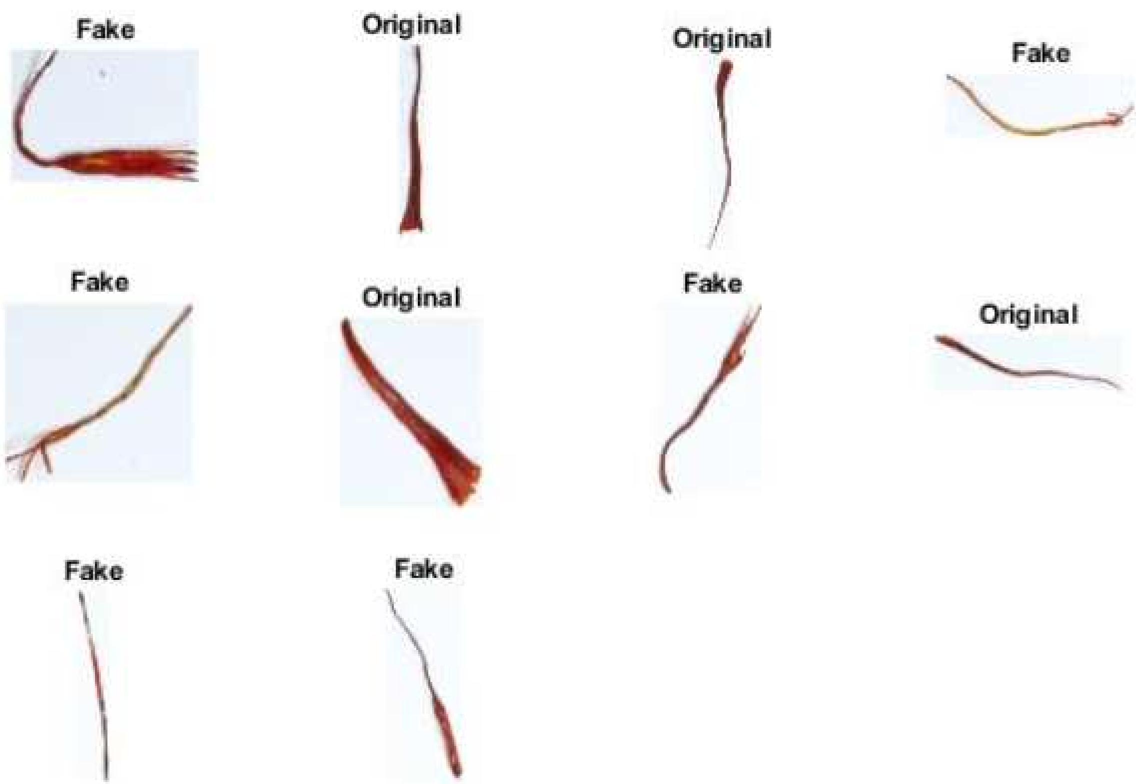

Figure 13. *Test images of the AlexNet architecture.*

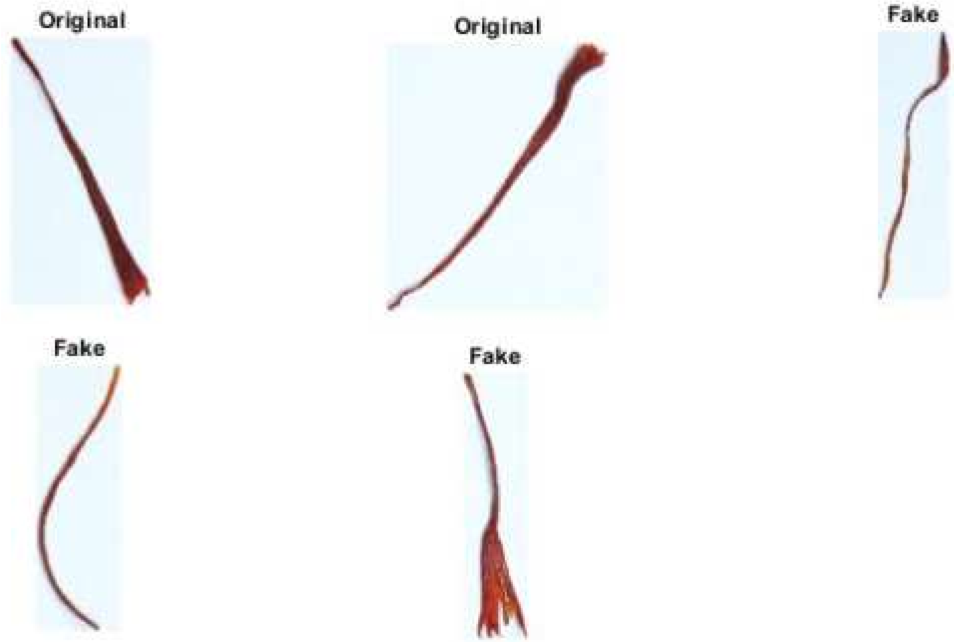

Figure 14. *Test images of the ResNet architecture.*

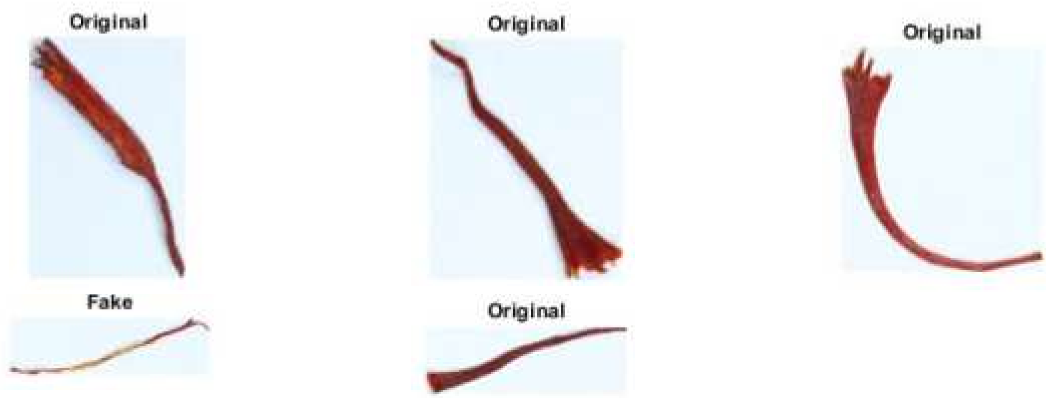

Figure 15*.* *Test images of the VGG architecture.*

3.3. Analysis of Misclassifications and Feature Learning

The confusion matrices shown in Figures 7-9 reveal insightful patterns about model behavior. AlexNet's two false positives (authentic samples misclassified as counterfeit) represent a tolerable error in a screening context, as these samples can be escalated for confirmatory testing without significant economic loss. AlexNet and VGG16 achieved perfect sensitivity (no false negatives), meaning they correctly identified all counterfeit samples. In contrast, ResNet misclassified 3 adulterated samples as authentic (90% sensitivity), which resulted in 3 false negatives. From an operational perspective, models with fewer false negatives are preferred to minimize the risk of undetected adulteration, as missed counterfeits pose a far greater economic and reputational risk to suppliers and exporters [11].

The learned feature representations were notably robust to variations in stigma orientation and minor background artifacts, a strength attributable to the extensive data augmentation and the inherent translation invariance learned by CNN architectures during ImageNet pre-training [14,24].

This suggests that the models focused on intrinsic morphological properties of the stigmas rather than relying on spurious correlations.

The superior performance of AlexNet in our study can be attributed to its architectural efficiency and optimization dynamics on datasets of moderate size [14]. Its larger receptive fields in the initial layers are particularly adept at capturing the macroscopic morphological features—such as filament curvature, branching patterns, and terminal structures—that are crucial for discriminating authentic saffron stigmas from visually crafted adulterants. This makes AlexNet not only accurate but also computationally frugal, a significant advantage for potential deployment in resource-limited settings, such as intake points in export companies or quality control labs [11]. In contrast, the deeper architectures like VGG16, while powerful, were prone to under-convergence in our CPU-bound training environment, highlighting the critical trade-off between model depth and computational feasibility in practical agricultural applications.

3.4. Comparative Analysis with Prior Literature

When contrasted with prior machine vision approaches for saffron quality assessment, our AlexNet-based model (96.49% accuracy) outperforms several reported methods. For instance, Kiani and Minaei (2016)[7] reported 92% accuracy using traditional computer vision and PLS regression for color quality characterization. Another study employing electronic nose technology for adulteration detection reported an average accuracy of 74.02% [1], highlighting the superior discriminative power of deep learning-based visual analysis.

Table 4. *Comparative performance with prior studies on saffron quality assessment.*

Study	*Method*	*Reported Accuracy*
Kiani & Minaei (2016)[7]	Machine Vision + PLS	92.00%
Heidarbeigi et al. (2016)[1]	Electronic Nose	74.02% (Average)
Proposed (VGG16)	Deep Learning (CNN)	87.72%
Proposed (ResNet)	Deep Learning (CNN)	94.74%
Proposed (AlexNet)	Deep Learning (CNN)	96.49%

However, it is important to note that direct comparisons with prior studies are challenging due to differences in dataset composition, adulterant types, imaging conditions, and evaluation protocols. The reported accuracies from previous works are provided solely to contextualize our findings, not to claim absolute superiority. A rigorous comparative study would require a unified dataset and standardized evaluation protocol applied to all methods.

4. Conclusion

This study demonstrates that deep convolutional neural networks (CNNs) can serve as a rapid, non-destructive, and cost-effective tool for authenticating saffron stigmas and detecting visual adulteration. Among the models evaluated, AlexNet provided the best balance of accuracy and computational efficiency, achieving 96.49% accuracy and 100% sensitivity on the validation set. These results indicate that even relatively shallow networks can effectively capture the critical morphological and color features of saffron, enabling reliable discrimination between authentic and counterfeit samples. Deeper architectures, such as ResNet and VGG16, showed comparable performance but may require larger datasets and greater computational resources for optimal convergence.

The findings of this study contribute to both the scientific understanding of saffron adulteration and practical postharvest applications, offering a scalable image-based screening approach suitable for grading, quality control, and export inspections. It should be acknowledged that our imaging setup used controlled illumination and uniform backgrounds, which may not fully represent real-world postharvest screening conditions where bulk samples with overlapping filaments and variable lighting are common. Furthermore, the performance metrics are based on a single random data split without cross-validation; future work should employ k-fold cross-validation or multiple random splits with reported standard deviations to provide a more robust evaluation of model generalization. Future research should also aim to validate these models across larger and more diverse datasets, integrate explainable AI techniques to enhance transparency and trust, and explore hybrid imaging–spectroscopy pipelines to combine visual and chemical information for more robust authentication. Additionally, the development of lightweight, deployable models could facilitate on-site inference in resource-limited environments. For practical deployment, the system should be implemented in real-time at customs and export companies, and evaluated using smartphone cameras to assess consumer-level authentication capability.

3.5. Limitations of the Study

Several limitations should be acknowledged. First, the imaging setup used controlled illumination and uniform backgrounds, which may not fully represent real-world postharvest screening conditions. Second, the performance metrics are based on a single random data split without cross-validation. Third, all models were trained on a CPU-based system, which particularly affected VGG16's convergence. Fourth, the confusion matrices presented in this study represent a random subset of the validation set for illustrative purposes, while the metrics in Table 3 are based on the complete validation set. Future work should address these limitations by employing k-fold cross-validation, GPU-accelerated training, and more challenging imaging conditions.

CRediT authorship contribution statement

Mitra Gholami: Project administration, Conceptualization, Methodology, Investigation, Software, Validation. **Rahman Farrokhi Teimourlou:** Supervision, Investigation, Writing-review & editing. **Farzaneh Jannatdoust:** Project administration, Conceptualization, Methodology, Investigation, Software, Validation, Writing original draft.

Declaration of competing interest

The authors have declared no conflict of interest.

Funding Declaration

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

Data availability

The datasets generated during and/or analysed during the current study are not publicly available due to commercial restrictions but are available from the corresponding author on reasonable request.

References

1. Heidarbeigi, K., Mohtasebi, S. S., Serrano-Diaz, J., Medina-Plaza, C., Ghasemi-Varnamkhasti, M., Alonso, G. L., & Rodriguez-Mendez, M. L. Flavour characteristics of Spanish and Iranian saffron analysed by electronic tongue. Qual. Assur. Saf. Crops Foods 8, 359–368 (2016).

2. Shahdadi, H., Barati, F., Bahador, R. S. & Eteghadi, A. Clinical applications of saffron (Crocus sativus) and its constituents: A literature review. Der Pharm. Lett. 8, 205–209 (2016).

3. Tan, J. & Xu, J. Applications of electronic nose (e-nose) and electronic tongue (e-tongue) in food quality-related properties determination: A review. Artif. Intell. Agric. 4, 104–115 (2020).

4. Zhang, G. & Abdulla, W. On honey authentication and adulterant detection techniques. Food Control 108992 (2022).

5. Çetin, N., Karaman, K., Kavuncuoğlu, E., Yıldırım, B. & Jahanbakhshi, A. Using hyperspectral imaging technology and machine learning algorithms for assessing internal quality parameters of apple fruits. Chemom. Intell. Lab. Syst. 230, 104650 (2022).

6. Rashed-Mohassel, M. H. Saffron from the wild to the field. Acta Hortic. 739, 37–42 (2008).

7. Kiani, S. & Minaei, S. Potential application of machine vision technology to saffron (Crocus sativus L.) quality characterization. Food Chem. 212, 392–394 (2016).

8. Azarabadi, N. & Özdemir, F. Determination of crocin content and volatile components in different qualities of Iranian saffron. GIDA/J. Food 43, 476–489 (2018).

9. Salehi, A., Shariatifar, N., Pirhadi, M. & Zeinali, T. An overview of different detection methods of saffron (Crocus sativus L.) adulterants. J. Food Meas. Charact. 16, 4996–5006 (2022).

10. Masi, E., Taiti, C., Heimler, D., Vignolini, P., Romani, A. & Mancuso, S. PTR-TOF-MS and HPLC analysis in the characterization of saffron (Crocus sativus L.) from Italy and Iran. Food Chem. 192, 75–81 (2016).

11. Minaei, S., Kiani, S., Ayyari, M. & Ghasemi-Varnamkhasti, M. A portable computer-vision-based expert system for saffron color quality characterization. J. Appl. Res. Med. Aromat. Plants 7, 124–130 (2017).

12. Yousefi-Nejad, S., Heidarbeigi, K. & Roushani, M. Applications of electronic tongue system for quantification of safranal concentration in saffron (Crocus sativus L.). J. Food Meas. Charact. 15, 1626–1633 (2021).

13. Pourreza, A., Pourreza, H., Abbaspour-Fard, M.-H. & Sadmia, H. Identification of nine Iranian wheat seed varieties by textural analysis with image processing. Comput. Electron. Agric. 83, 102–108 (2012).

14. Krizhevsky, A., Sutskever, I. & Hinton, G. E. ImageNet classification with deep convolutional neural networks. Adv. Neural Inf. Process. Syst. 1097–1105 (2012).

15. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition 770–778 (2016).

16. Chakravartula, S. S. N., Moscetti, R., Bedini, G., Nardella, M., & Massantini, R. Use of convolutional neural network (CNN) combined with FT-NIR spectroscopy to predict food adulteration: A case study on coffee. Food Control 135, 108816 (2022).

17. Wu, X., Gao, S., Niu, Y., Zhao, Z., Xu, B., Ma, R., & Zhang, Y. Identification of olive oil in vegetable blend oil by a one-dimensional convolutional neural network combined with Raman spectroscopy. J. Food Compos. Anal. 108, 104396 (2022).

18. Kafi, M., Koocheki, A. & Rashed, M. H. Saffron (Crocus sativus): Production and Processing. (Science Publishers, 2006).

19. Momeny, M., Neshat, A. A., Jahanbakhshi, A., Mahmoudi, M., Ampatzidis, Y., & Radeva, P. Grading and fraud detection of saffron via learning-to-augment incorporated Inception-v4 CNN. Food Control 147, 109554 (2023).

20. Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. Preprint at arXiv:1409.1556 (2014).

21. Jahanbakhshi, A., Abbaspour-Gilandeh, Y., Heidarbeigi, K. & Momeny, M. A novel method based on a machine vision system and deep learning to detect fraud in turmeric powder. Comput. Biol. Med. 136, 104728 (2021a).

22. Jahanbakhshi, A., Abbaspour-Gilandeh, Y., Heidarbeigi, K. & Momeny, M. Detection of fraud in ginger powder using an automatic sorting system based on image processing technique and deep learning. Comput. Biol. Med. 136, 104764 (2021b).

23. Momeny, M., Neshat, A. A., Gholizadeh, A., Jafarnezhad, A., Rahmanzadeh, E., Marhamati, M., & Zhang, Y. D. Greedy AutoAugment for classification of Mycobacterium tuberculosis image via generalized deep CNN using mixed pooling based on minimum square rough entropy. Comput. Biol. Med. 141, 105175 (2022).

24. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S. & Fei-Fei, L. ImageNet large scale visual recognition challenge. Int. J. Comput. Vis. 115, 211–252 (2015).