



Article

Deep Learning-Based Model for Effective Classification of *Ziziphus jujuba* Using RGB Images

Yu-Jin Jeon ¹, So Jin Park ¹, Hyein Lee ¹, Ho-Youn Kim ² and Dae-Hyun Jung ^{1,*}

¹ Department of Smart Farm Science, Kyung Hee University, Yongin 17104, Republic of Korea; yjnara95@gmail.com (Y.-J.J.); morningsojin@khu.ac.kr (S.J.P.); lee.hyein@khu.ac.kr (H.L.)

² Smart Farm Research Center, Korea Institute of Science and Technology (KIST), Gangneung-si 25451, Republic of Korea; hykim@kist.re.kr

* Correspondence: daehyun@khu.ac.kr; Tel.: +82-31-201-2657

Abstract: Ensuring the quality of medicinal herbs in the herbal market is crucial. However, the genetic and physical similarities among medicinal materials have led to issues of mixing and counterfeit distribution, posing significant challenges to quality assurance. Recent advancements in deep learning technology, widely applied in the field of computer vision, have demonstrated the potential to classify images quickly and accurately, even those that can only be distinguished by experts. This study aimed to develop a classification model based on deep learning technology to distinguish RGB images of seeds from *Ziziphus jujuba* Mill. var. *spinosa*, *Ziziphus mauritiana* Lam., and *Hovenia dulcis* Thunb. Using three advanced convolutional neural network (CNN) architectures—ResNet-50, Inception-v3, and DenseNet-121—all models demonstrated a classification performance above 98% on the test set, with classification times as low as 23 ms. These results validate that the models and methods developed in this study can effectively distinguish *Z. jujuba* seeds from morphologically similar species. Furthermore, the strong performance and speed of these models make them suitable for practical use in quality inspection settings.



Citation: Jeon, Y.-J.; Park, S.J.; Lee, H.; Kim, H.-Y.; Jung, D.-H. Deep Learning-Based Model for Effective Classification of *Ziziphus jujuba* Using RGB Images. *AgriEngineering* **2024**, *6*, 4604–4619. <https://doi.org/10.3390/agriengineering6040263>

Academic Editors: Yee Siang Gan, Sze-Teng Liang and Zhengjun Qiu

Received: 28 October 2024

Revised: 23 November 2024

Accepted: 29 November 2024

Published: 3 December 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: Suan Zao Ren; medicinal plants; non-destructive testing; convolutional neural networks (CNNs); ResNet; Inception; DenseNet

1. Introduction

Medicinal plants have been collected and utilized around the world throughout the development of human civilization. Even after the advancement of synthetic drugs, medicinal plants are recognized for their medical value and they have fewer side effects on the human body compared to synthetic drugs [1]. The international trade of medicinal plants and their products is estimated to be approximately \$60 billion, with projections reaching around \$5 trillion by 2050 [2]. Meanwhile, the composition and concentration of active ingredients in medicinal plants can vary significantly even within the same family and species [3], and the types of medicinal plants and the varieties used as ingredients also vary by country. Therefore, identifying and classifying species of medicinal plants during the distribution process in the global medicinal market is crucial for ensuring the quality of medicinal materials and food safety [4].

Suan Zao Ren, also known as sour jujube, is a medicinal plant consisting of the seeds of *Ziziphus jujuba* Mill. var. *spinosa* Hu ex H. F. Chou. This variety is primarily cultivated in the northwestern regions of China, especially around the Yellow River Basin [5]. Suan Zao Ren is known in traditional Chinese and Korean medicine for its use in promoting sleep and treating sleep disorders [6]. This medicinal material is rich in potassium and magnesium and is also used to treat various symptoms such as appetite loss, fatigue, anxiety, migraines, and irritable bowel syndrome [7,8]. While the value of Suan Zao Ren is recognized globally and demand remains steady, the issue of similar-looking seeds from different plants being mixed during distribution has persisted [9]. This problem affects the

quality of the medicinal materials and the industry, necessitating classification techniques for medicinal plant species and their parts.

Traditionally, classification was performed by sensory panel experts using sight, smell, and touch, leveraging their extensive knowledge and experience with medicinal plants. Advances in chemical analysis, genomics, and molecular marker technologies have also been applied to classify Suan Zao Ren and other mixed species. For instance, Li et al. (2010) [10] used sequence-related amplified polymorphism (SRAP) technology to analyze genetic relationships among 14 species of *Ziziphus*, including 11 varieties of *Z. jujuba* Mill. and one outgroup. Sun et al. (2014) [11] utilized high-performance thin-layer chromatography (HPLC) and principal component analysis (PCA) to analyze trace elements and main bioactive compounds in Suan Zao Ren to classify impurity plant materials. Furthermore, Du et al. (2023) [5] investigated the plastome variation pattern of sour jujube, constructed a phylogenomic tree, and studied molecular markers that could distinguish Suan Zao Ren. However, these technologies are limited by their high cost and the time required to complete analyses, making them challenging to apply in distribution settings.

Recent advancements in computer vision and deep learning technologies have led to the development of models that can quickly and cost-effectively identify targets without destruction [12,13]. Deep learning involves Artificial Neural Networks (ANNs) consisting of input, hidden, and output layers, where complex features in data are learned through multiple deep hidden layers, enabling the networks to perform predictions or classifications [14]. Specifically, Convolutional Neural Networks (CNNs) comprise five key components: the input layer, convolutional layers, pooling layers, fully connected layers, and output layers [15]. The convolutional layers use small matrix filters or kernels to move across the image pixels, extracting local features, and this process is repeated across various layers, allowing the model to learn intricate characteristics of the targets [16].

Various classification models based on CNNs with such characteristics have been developed, and they have been shown to achieve higher classification accuracy compared to traditional machine learning-based classification models in most cases [17]. Table 1 presents the research status of deep learning-based models for the non-destructive classification of plant varieties using RGB images. These deep learning-based models are capable of recognizing and classifying features from RGB images of plant parts, such as seeds or leaves, for each plant variety. Moreover, their performance demonstrates an accuracy of over 90%, indicating their potential to replace human-based classification.

Table 1. Research progress on deep learning approaches for classifying plant varieties via RGB imaging.

No.	Number of Varieties and Categories	Model	Accuracy (%)	Reference
1	5 varieties of peaches	CNN	94.70	Rong [18]
2	4 varieties of sunflowers	GoogleNet	95.00	Kurtulmuş [19]
3	29 varieties of wheat	ResNet-18	94.64	Xing [20]
4	27 varieties of grapes	ResNet-50	91.30	Nart [21]
		Inception-v3	99.70	
5	79 species of plants	DenseNet-161	97.30	Shelke and Mehendale [22]
6	8 species of weeds and 2 species of crops	VGG16	92.60	Mesías-Ruiz [23]
		ResNet-152	95.60	

However, in the medicinal plant sector, there is a scarcity of studies applying deep learning to classify species of medicinal materials in distribution, including seeds like Suan Zao Ren. No reported studies have explored the classification of *Ziziphus mauritiana* Lam. and *Hovenia dulcis* Thunb. seeds, which are visually similar and often mixed with Suan Zao Ren seeds.

Therefore, the objective of this study is to investigate whether deep learning can learn the differences between seeds of the medically recognized *Ziziphus jujuba* Mill. var. *spinosa* and the non-medically recognized *Ziziphus mauritiana* Lam. and *Hovenia dulcis* Thunb. and classify them rapidly and accurately. The detailed objectives are as follows: (1) to collect seeds of *Z. jujuba* Mill., *Z. mauritiana* Lam., and *H. dulcis* Thunb., which are practically distributed as mixtures, and obtain their RGB images; (2) to preprocess these images to build an input dataset for deep learning models; and (3) to design deep learning-based models, train them using the dataset, and evaluate their classification performance. The deep learning-based model developed through this study can rapidly and accurately classify *Ziziphus jujuba* Mill. seeds and mixed seeds, serving as an effective tool for quality verification in the distribution of *Ziziphus jujuba* seeds. Furthermore, the proposed approach in this study can be applied to other medicinal materials with significant mixing issues, contributing to the improvement of safety and reliability in the medicinal materials industry.

2. Materials and Methods

2.1. Sample Collection and Image Acquisition

As shown in Figure 1, this study used seeds from three species as samples: *Ziziphus jujuba* Mill. var. *spinosa* (Figure 1a), *Ziziphus mauritiana* Lam. (Figure 1b), and *Hovenia dulcis* Thunb. (Figure 1c). The National Development Institute of Korean Medicine (94 Hwarang-ro, Gyeongsan-si, Gyeongsangbuk-do, Republi of Korea) collected and verified the seeds produced and distributed from various regions. Through the collection and verification process, samples representing natural variations in size, color, and shape for each species were gathered.

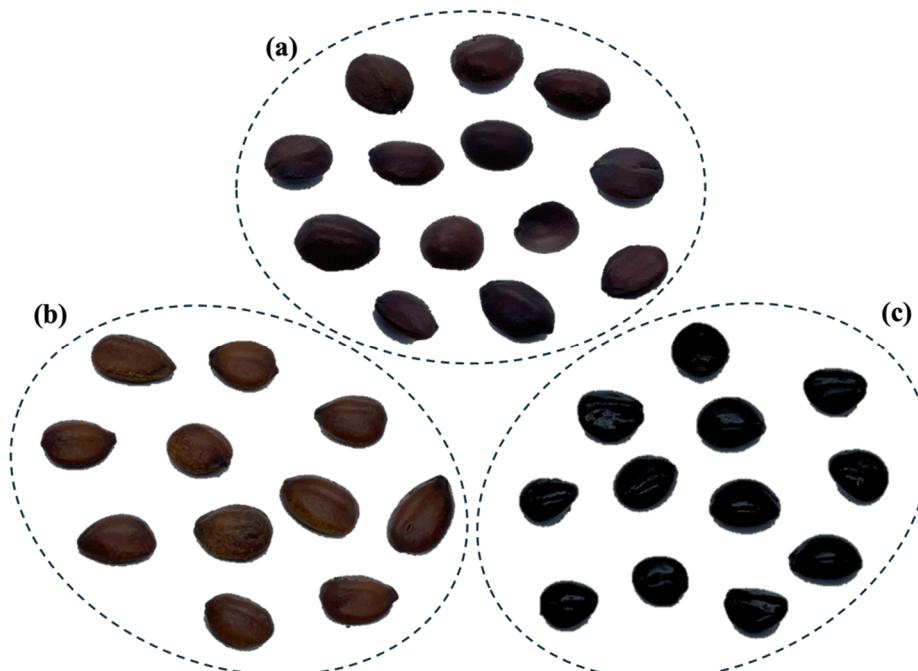


Figure 1. Sample images of three species: (a) *Ziziphus jujuba* Mill., (b) *Ziziphus mauritiana* Lam., and (c) *Hovenia dulcis* Thunb.

To minimize noise and artifacts during the image acquisition and to maintain consistency, specific imaging conditions were established. The images were captured in a photo booth (PULUZ photo light box) with uniform lighting providing an illuminance of 25 lux at a distance of 15 cm, a consistent background, and a fixed distance between the light source and the subject. A high-quality megapixel USB camera (ELP-USB16MP02-AF100, Shenzhen Ailipu Technology Co., Ltd., Shenzhen, China) was used for the image acquisition. The total number of images collected was 243 for *Z. jujuba* Mill., 242 for *Z. mauritiana* Lam.,

and 240 for *H. dulcis* Thunb. To maintain uniformity in the training data, 240 images were randomly selected from both *Z. jujuba* Mill. and *Z. mauritiana* Lam., resulting in an equal count of 240 raw images for each species.

2.2. Image Dataset Preparation

Initially, to extract the RGB information from the captured images, the seeds were detected based on their center and then cropped around the detected seed. The object detection method used involved initially converting the original image to grayscale to maximize the contrast between the object and the background, then detecting contours and designating the space within the contours as the object. The images were then cropped to a size of 512×512 (Figure 2a). Next, to extract the RGB area of the seeds from the cropped images, the HSV (Hue, Saturation, Value) color space from the OpenCV library was utilized. The images were first converted from RGB to HSV color space, after which the HSV range corresponding to the background was defined based on the method described in [24]. This range was set to 0, while all other ranges were set to 1 to create a binary mask. This mask was then applied to the original image to extract only the RGB values of the sample (Figure 2b). Furthermore, to enhance the robustness of the deep learning model, image rotation was performed to increase the diversity of the training data (Figure 2c). The rotation was executed using the rotate function of the Pillow library. Each of the original images were rotated six times at 60° increments, increasing the total from 240 to 1440 images per species.

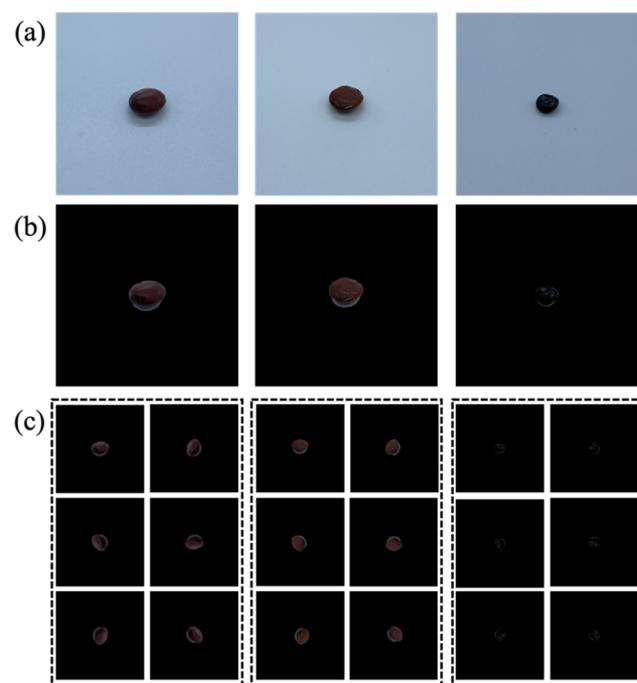


Figure 2. Examples of data from three species: (a) central cropped, (b) background-removed, and (c) images rotated at six different angles.

Additionally, to enhance the deep learning model's ability to learn robust features from the data, the Gaussian noise method was employed to add noise to the training data [25]. Each image was processed with a signal-to-noise ratio (SNR) value of 20 dB. Subsequently, the model's input data was resized to two different sizes: 224×224 and 229×229 . The data was subsequently randomly divided into training, validation, and test sets in a 6:2:2 ratio. The training set was used as input data for training the model, while the validation set served as input data for evaluating the trained model's performance and adjusting hyperparameters. The test set was used solely as input data to assess the

predictive performance of the final constructed model. The configuration of the dataset is as follows, with details provided in Table 2.

Table 2. Dataset distribution.

Cultivar	Raw	Augmentation	Train	Validation	Test
<i>Ziziphus jujuba</i> Mill.	240	1440			
<i>Ziziphus mauritiana</i> Lam.	240	1440	2764	692	864
<i>Hovenia dulcis</i> Thunb.	240	1440			

2.3. Deep Learning Method

In this study, we used CNN-based architectures—ResNet-50, Inception-v3, and DenseNet-121—for deep learning on RGB image datasets of *Z. jujuba* Mill., *Z. mauritiana* Lam., and *H. dulcis* Thunb. ResNet, Inception, and DenseNet were utilized in this study. ResNet, proposed by He et al. (2016) [26], introduced the concept of residual learning to address issues of slowed learning rates and performance degradation in deep neural networks. ResNet-50 incorporates residual blocks, as shown in Figure 3a, which directly connect input information to the output, enabling residual learning. This allows effective training even in deep networks. Inception, proposed by Szegedy et al. (2016) [27], operates by processing various sizes of convolution operations in parallel within a single layer, as illustrated in Figure 3b. This approach effectively extracts features at multiple scales, enabling the network to learn more accurate and comprehensive information. DenseNet, introduced by Huang et al. (2017) [28], features a structure where each layer is directly connected to all previous layers, as depicted in Figure 3c. This design enhances the flow of information and gradients between feature maps, enabling efficient learning and superior feature extraction.

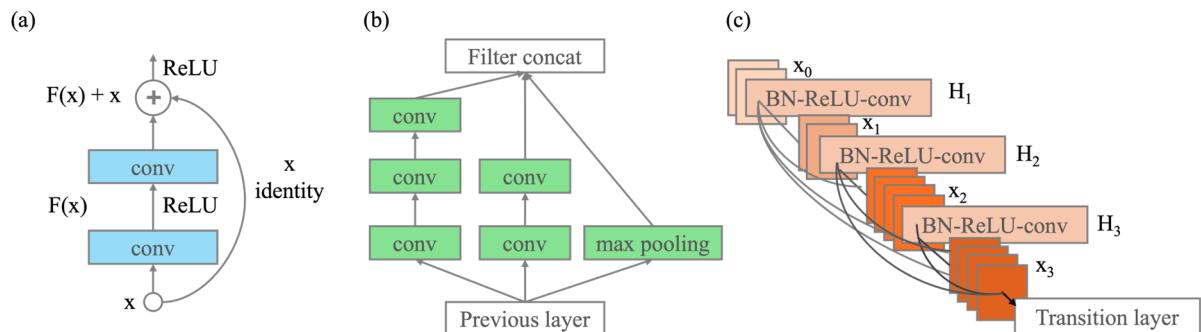


Figure 3. Comparative architectural diagrams of (a) ResNet, (b) Inception, and (c) DenseNet models.

All three models employed transfer learning, as shown in Figure 4. They were pre-trained using the ImageNet dataset, loading the weights and then performing global average pooling at the global average pooling layer to compute the spatial dimensions' averages, thus producing feature vectors composed of each feature map's average value. Subsequently, a dense layer with 1024 nodes was added as the fully connected layer, utilizing the ReLU activation function [29] to perform non-linear transformations. For the output layer, a dense layer with three nodes was added to make predictions for three classes, using the Softmax activation function to output the probabilities for each variety.

The training process incorporated the K-Fold cross-validation method [30] to prevent the models from overfitting to specific training sets. The data was randomly divided into five folds, using four for training and one for validation, with k set to 5. The optimization algorithm used was Adam, with a learning rate of 0.00001 and a loss function of categorical cross-entropy. The batch size was set at 64, with 100 epochs, allowing for 4400 forward and backward passes and weight updates. Additionally, the stability of model training was monitored by tracking the accuracy and loss values for each epoch for the train and

validation sets. As for the experimental setup, the frameworks for ResNet-50, Inception-v3, and DenseNet-121 were implemented in Python 3.11.8 using TensorFlow 2.16.1.

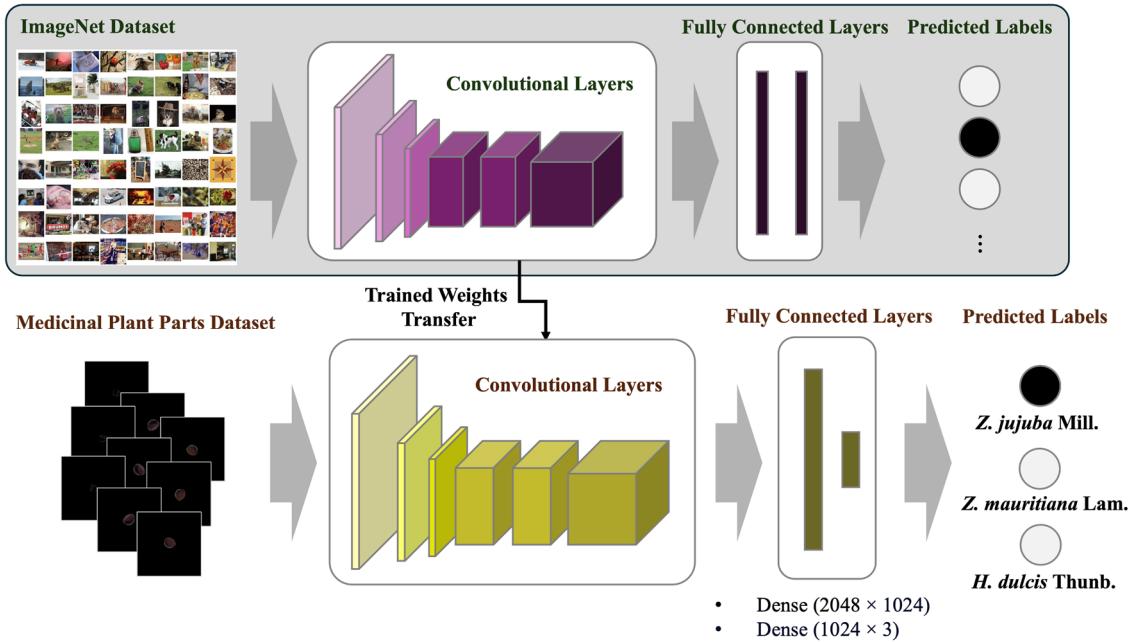


Figure 4. Transfer learning framework using pre-trained convolutional neural networks for distinguishing *Z. jujuba* Mill., *Z. mauritiana* Lam., and *H. dulcis* Thunb.

2.4. Model Performance Evaluation

We employed several metrics to evaluate the performance of the constructed models, including accuracy, precision, recall, F1 score, kappa coefficient, and the accuracy of 5-fold cross-validation as follows in Equations (1)–(6).

Accuracy represents the proportion of samples correctly predicted by the model out of the total dataset. Precision indicates the proportion of actual positives among items predicted as positive by the model, while recall refers to the proportion of actual positives that were correctly identified as positive by the model. Precision is crucial when minimizing false positives is important, and recall is advantageous when identifying all positive samples is vital. The F1 score, a harmonic mean of precision and recall, is useful when accurate identification of positive samples is crucial [31].

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

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \quad (2)$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}} \quad (3)$$

$$\text{F1 score} = 2 \times \frac{\text{Recall} \times \text{Precision}}{\text{Recall} + \text{Precision}} \quad (4)$$

These metrics were calculated using the results of the model classifications on the test set, specifically the counts of true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). The confusion matrix method [32] was used to calculate the TP, TN, FP, and FN values for each species, identifying the occurrence of prediction errors.

The kappa coefficient evaluates the performance of the model by considering the difference between observed accuracy and the accuracy that would occur by chance, accounting for class imbalance [33].

$$\text{Kappa coefficient} = \frac{P_o - P_e}{1 - P_e}, \quad \text{where} \\ P_e = \frac{\sum_{i=1}^k (n_{i+} \times n_{+i})}{n^2} \quad (5)$$

$$\text{Accuracy}_{CV} = \frac{1}{k} \sum_{i=1}^k \text{Accuracy}_i \quad (6)$$

In Equation (5), P_o represents the observed accuracy, which is the proportion of items that the classifier has correctly predicted in the actual data. P_e is the accuracy of random predictions based on the frequencies of each class, calculated using n_{i+} , the total number of actual instances in class i ; n_{+i} , the total number of predicted instances in class i ; and n , the total number of observations. The 5-fold cross-validation accuracy, which averages the accuracies across each fold, serves to assess the model's generalization capability [30]. In Equation (6), K is used as the number of folds, which is 5, and Accuracy_i represents the accuracy in the i -th fold.

As in the ROC-AUC method [34], the true positive rate (TPR) and false positive rate (FPR) were calculated to plot the ROC curve and evaluate the model's performance across various thresholds. Additionally, the area under the ROC curve (AUC) was calculated to assess whether the model consistently performs well across different thresholds.

Additionally, the total number of parameters and trainable parameters for each model, as well as the average time taken to classify test images, were examined to compare the complexity, efficiency, and applicability of the models in applications.

To verify the effective extraction of species features by each model, the t-Distributed Stochastic Neighbor Embedding (t-SNE) technique [35] was used. This technique involves calculating similarity probabilities in a high-dimensional space, reducing dimensions, and minimizing the cost function to map the distribution of samples from each species into a three-dimensional space. Using these techniques, we first reduced the dimensions of the dataset constructed with rotated and noise-added RGB images of *Z. jujuba* Mill., *Z. mauritiana* Lam., and *H. dulcis* Thunb., and mapped the distribution of samples from each species in a three-dimensional space. Next, the output vectors obtained by forwarding the input data to the global average pooling layer in each model were used as feature vectors for each sample, which were then mapped in a three-dimensional space.

3. Results

3.1. Construction of Deep Learning-Based Models for Distinguishing *Z. jujuba* Mill.

Firstly, Figure 5 illustrates the accuracy and loss values across the epochs during training on the training and validation sets for each model. For the ResNet-50 model, Figure 5a shows changes in accuracy and loss values during training. The training accuracy of this model increased sharply early on, nearing 1. The validation accuracy showed a similar trend, maintaining a slightly lower value than the training accuracy and stabilized after epoch 20. The training loss decreased sharply as the epochs progressed, ultimately converging close to 0. However, the validation loss, while initially decreasing significantly, showed some volatility and remained higher compared to the training loss up to epoch 100.

For the model using the Inception-v3 architecture, changes in accuracy and loss values during the training are depicted in Figure 5b. This model's training accuracy remained consistently near 1, and while the validation accuracy was also very high, it was slightly lower and showed some fluctuations. The training loss decreased sharply initially and then remained at a low level with little change thereafter. In contrast, the validation loss decreased rapidly initially and then continued to show minor fluctuations. Compared to the ResNet-50 model, the validation loss for Inception-v3 showed less volatility.

For the model based on the DenseNet-121 architecture, changes in accuracy and loss during training can be observed in Figure 5c. Both the training and validation accuracy of this model quickly rose to high levels and then stabilized. Both the training and validation losses decreased sharply initially and remained low, with the training loss approaching

zero and the validation loss showing minimal fluctuations. When compared to the training process graphs of the ResNet-50 and Inception-v3 models, the DenseNet-121 model showed less variability in both accuracy and loss.

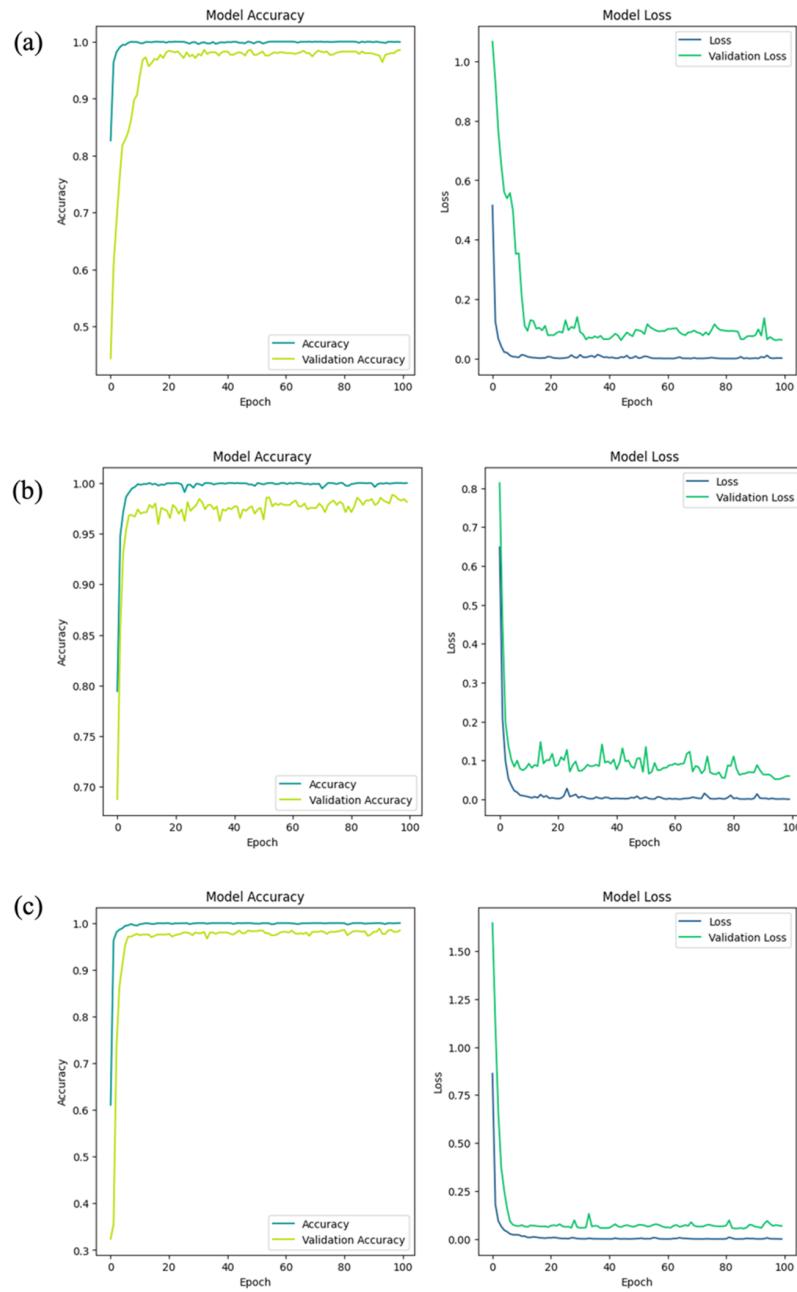


Figure 5. Training accuracy and loss curves for three deep learning models: (a) ResNet-50, (b) Inception-v3, and (c) DenseNet-121.

3.2. Performance Evaluation of Deep Learning-Based Models

The performance metrics for the ResNet-50, Inception-v3, and DenseNet-121 models on the test set, on which the models had not been trained, are summarized in Table 3. All three models demonstrated high accuracy, exceeding 98%, with DenseNet-121 showing a slightly higher accuracy. Precision, recall, and F1 scores were similarly high across all models, with ResNet-50 around 98.1%, Inception-v3 around 98.3%, and DenseNet-121 over 98.6%. The kappa score for DenseNet-121 was also higher than the other two models. In terms of the 5-fold cross-validation (CV) accuracy, all models exceeded 99%, with DenseNet-121, ResNet-50, and Inception-v3 in descending order of performance.

Overall, DenseNet-121 demonstrated a slight edge over the other two models in all metrics according to Table 3.

Table 3. Performance metrics for the test set and 5-fold cross validation accuracy.

Model	Accuracy	Loss	Precision	Recall	F1 Score	Kappa	5-Fold CV
ResNet-50	98.03%	0.097	98.12%	98.16%	98.13%	97.04%	99.97%
Inception-v3	98.26%	0.088	98.34%	98.37%	98.35%	97.39%	99.71%
DenseNet-121	98.61%	0.074	98.67%	98.72%	98.68%	97.91%	99.97%

The results of each model's confusion matrix are shown in Figure 6. In the ResNet-50 confusion matrix (Figure 6a), the *Z. mauritiana* Lam. class was correctly classified with an accuracy of 96.20%, with 3.80% being misclassified as *Z. jujuba* Mill. The *Z. jujuba* Mill. class showed a very high accuracy of 98.29%, with only 1.71% misclassified as *Z. mauritiana* Lam. The *H. dulcis* Thunb. class was perfectly classified with 100% accuracy. In the Inception-v3 confusion matrix (Figure 6b), the *Z. mauritiana* Lam. class was classified with a higher accuracy of 97.15% compared to ResNet-50, with 2.85% misclassified as *Z. jujuba* Mill. The *Z. jujuba* Mill. class had a slightly lower accuracy of 97.95% but was still high, with 2.05% misclassified as *Z. mauritiana* Lam. The *H. dulcis* Thunb. class was also perfectly classified at 100% accuracy in this model. In the DenseNet-121 confusion matrix (Figure 6c), the *Z. mauritiana* Lam. class was classified with an accuracy of 96.84%, with 3.16% misclassified as *Z. jujuba* Mill. The *Z. jujuba* Mill. class had the highest accuracy among the three models at 99.32%, with only 0.68% misclassified as *Z. mauritiana* Lam. The *H. dulcis* Thunb. class was also perfectly classified at 100% accuracy. Thus, all three models perfectly classified the *H. dulcis* Thunb. class, with DenseNet-121 performing best for the *Z. jujuba* Mill. class, and Inception-v3 showing the best performance for *Z. mauritiana* Lam.

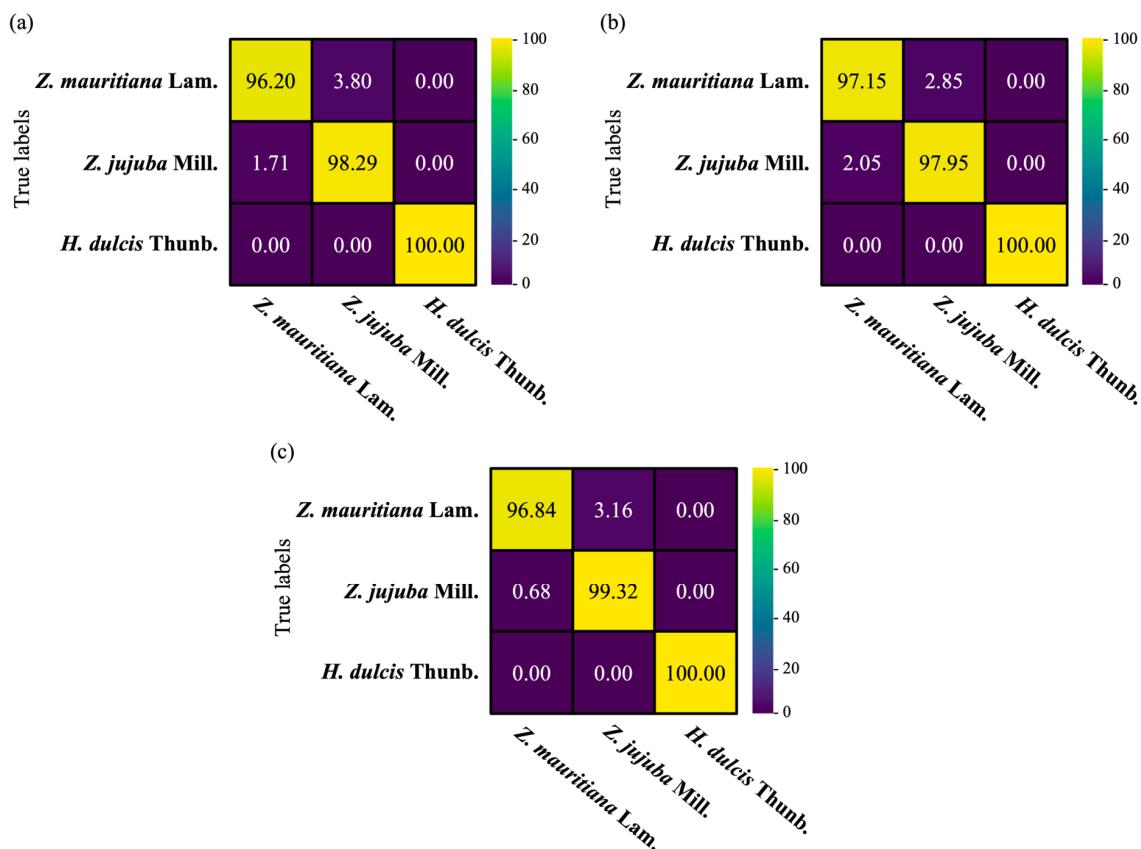


Figure 6. Confusion matrix for three deep learning models: (a) ResNet-50, (b) Inception-v3, and (c) DenseNet-121.

Additionally, the results for each model's ROC curve and AUC values for the test set are presented in Figure 7. All three models achieved very high AUC values exceeding 0.998 for all species. For *H. dulcis* Thunb., all three models achieved an AUC value of 1.0. For *Z. mauritiana* Lam. and *Z. jujuba* Mill., all models achieved AUC values exceeding 0.998, with the Inception-v3 model showing a slightly higher AUC of 0.999 for *Z. mauritiana* Lam., while all models had an AUC of 0.998 for *Z. jujuba* Mill.

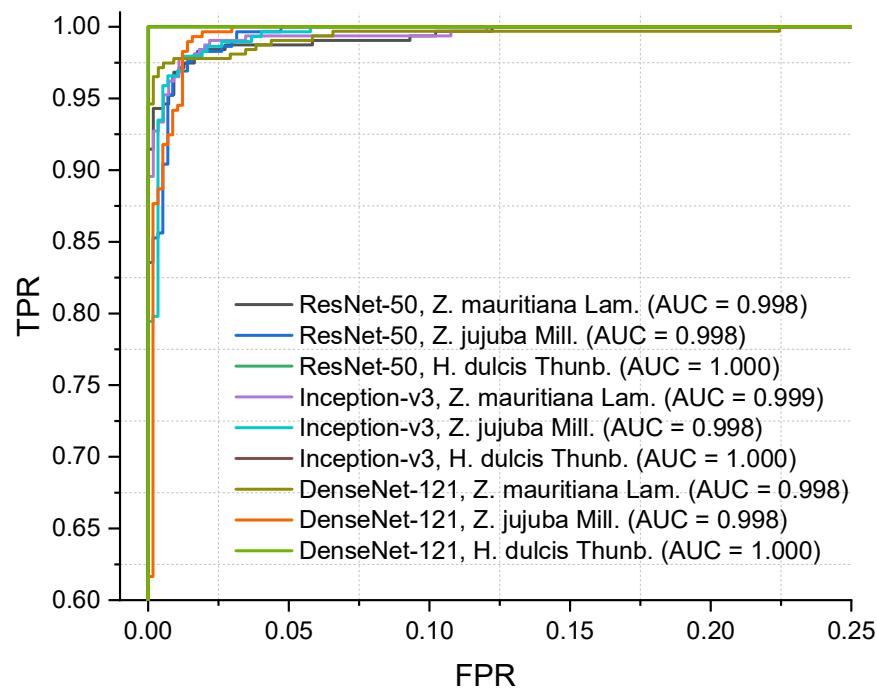


Figure 7. ROC curve with AUC for different species from three deep learning models.

3.3. Visualization of Feature Learning Through Unsupervised Clustering

We employed the t-SNE method, a type of unsupervised clustering analysis, to investigate whether each model effectively learned the characteristics of the seed species from the RGB data (Figure 8). Initially, the distribution of data points in the training dataset appeared somewhat clustered by variety as shown in Figure 8a, yet there was an overlap in the distributions of *Z. jujuba* Mill. and *Z. mauritiana* Lam. across both the plane formed by components 1 and 3, and the plane formed by components 1 and 2. For *H. dulcis* Thunb., more distinct clustering was observed in the plane formed by components 1 and 2, though the cluster appeared somewhat divided into two parts in the plane formed by components 1 and 3.

Subsequently, the results of mapping the feature output vectors from each deep learning-based classification model into a three-dimensional space are shown in Figure 8b, Figure 8c, and Figure 8d for ResNet-50, Inception-v3, and DenseNet-121, respectively. As seen in Figure 8b, the vectors output by ResNet-50 were more densely clustered than those in the training dataset. However, in the plane formed by components 1 and 3, the clusters for *Z. jujuba* Mill. and *Z. mauritiana* Lam. were divided into two, and in the plane formed by components 1 and 2, *Z. jujuba* Mill. formed two close clusters, while *Z. mauritiana* Lam. and *H. dulcis* Thunb. appeared as one elongated elliptical cluster.

The vectors output by the Inception-v3 model, as shown in Figure 8c, formed two clusters for *Z. jujuba* Mill. and *Z. mauritiana* Lam. in the plane of components 1 and 3, and in the plane of components 1 and 2, *Z. jujuba* Mill. and *H. dulcis* Thunb. were more circularly clustered compared to ResNet-50, with *Z. mauritiana* Lam. appearing as two overlapping circles.

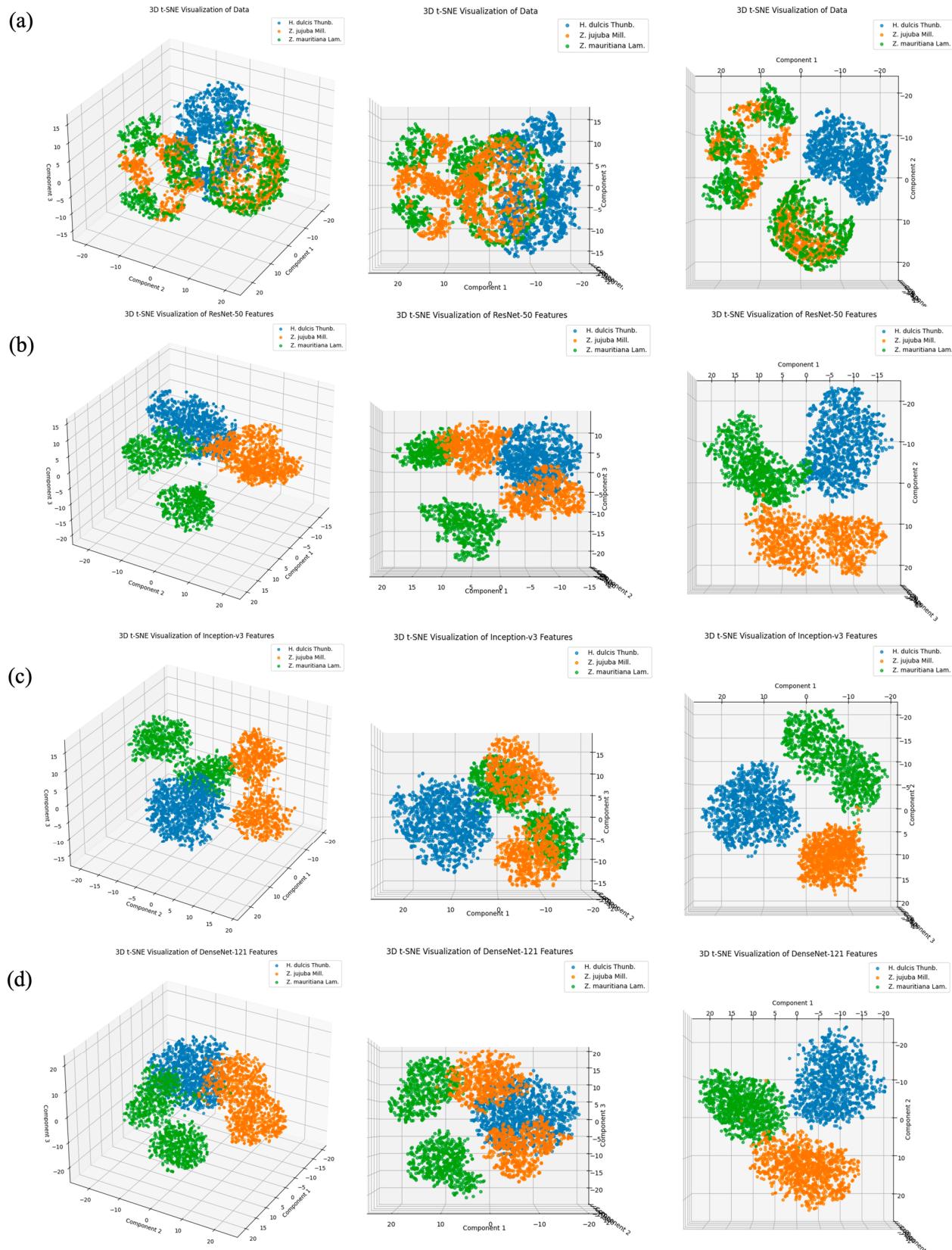


Figure 8. Three-dimensional t-SNE visualization of training data features and features from three models: **(a)** training data, **(b)** ResNet-50, **(c)** Inception-v3, and **(d)** DenseNet-121. In the plots, the colors represent different species: orange for *Ziziphus jujuba* Mill., green for *Ziziphus mauritiana* Lam., and blue for *Hovenia dulcis* Thunb.

As indicated in Figure 8d, DenseNet-121, in the plane formed by components 1 and 3, showed clustering similar to Inception-v3's results for each species, but in the plane formed by components 1 and 2, all species formed more distinct clusters, appearing as a tighter cluster compared to the results from Inception-v3.

3.4. Comparison of Model Parameters and Testing Time Efficiency

The study assessed the number of parameters and the time required for testing each model, as indicated in Table 4. When looking at the number of trainable parameters for the base models, Inception-v3, ResNet-50, and DenseNet-121 have them in decreasing order. The parameter counts for Inception-v3 and ResNet-50 are similar, around 20 million, but DenseNet-121 has significantly fewer, approximately 7 million. Even when comparing the number of parameters added for new layers to perform the classification in this study, the order remains the same: Inception-v3, ResNet-50, and then DenseNet-121, with the DenseNet-121 model having the lowest total optimized parameters during the training process at 6.963 million.

Table 4. Comparison of trainable parameters, average testing time for three models: (a) ResNet-50, (b) Inception-v3, and (c) DenseNet-121.

Model	Base Model Trainable Params (M)	Total Trainable Params (M)	Avg Test Time (ms)
ResNet-50	23.53	25.63	20.67
Inception-v3	23.87	25.97	25.00
DenseNet-121	6.96	6.96	23.00

The investigation into the average time taken to classify test data using the trained models showed that all models required less than 25 ms. The time taken was least for DenseNet-121, followed by ResNet-50, and Inception-v3, with a very small difference between the models, approximately 2–3 ms.

4. Discussion

In this study, we constructed and compared the performance of three CNN-based models for classifying RGB images of three types of seeds: *Z. jujuba* Mill., which is used as a medicinal material, and *Z. mauritiana* Lam. and *H. dulcis* Thunb., which are not recognized as medicinal but are often mixed in distribution.

ResNet-50, Inception-v3, and DenseNet-121 exhibited comparable accuracy and loss metrics on both the training and validation sets (Figure 5), demonstrating no signs of overfitting. Performance metrics for the test set (Table 3) showed very high accuracy values for all models, exceeding 98.03%. These accuracy levels are slightly higher compared to those reported in previous studies that used CNN-based models to classify RGB images of seeds. For instance, Kurtulmuş et al. (2021) [19] achieved a 95.00% accuracy when classifying four varieties of sunflower seeds, while Xing et al. (2023) [20] reported a 94.64% accuracy in classifying 29 varieties of wheat seeds. This suggests that the models developed and selected in this study are highly applicable for quality assurance processes where accuracy is of critical importance.

Among the models, DenseNet-121 demonstrated the best performance across all categories. Its high precision and recall indicate that it most effectively classifies true positive cases while minimizing errors, which is further supported by its highest F1 score, reflecting a strong balance between precision and recall. Additionally, DenseNet-121's high kappa value suggests that its prediction accuracy significantly exceeds random chance. All models, particularly ResNet-50 and DenseNet-121, achieved exceptional generalization capabilities, with a 5-fold cross-validation accuracy exceeding 99.9%.

DenseNet-121, in particular, achieved the highest classification accuracy for *Z. jujuba* Mill., which is recognized as a medicinal material. This superior performance is likely attributed to the architecture of DenseNet-121 (Figure 3c), where each layer is directly

connected to all previous layers, effectively preserving and learning specific features, as observed in Shelke et al. (2023). The t-SNE analysis further demonstrated that DenseNet-121 formed more circular clusters for each species, indicating its superior ability to extract and learn individual characteristics.

Additionally, an analysis of trainable parameters and the time required for classification tests (Table 4) revealed that DenseNet-121 used less memory compared to the other two models. Despite having fewer parameters, DenseNet-121 exhibited relatively fast prediction speeds, suggesting that it possesses a more efficient structure for classifying the three species.

However, the limited diversity of the models used and the challenges observed in individually analyzing the features extracted from the plot data suggest areas for potential improvement. Additionally, while achieving a maximum accuracy of 98.61%, there is still room for performance enhancement. Future research could focus on identifying the characteristics of overlapping regions between classes observed through t-SNE analysis and exploring architectures such as attention mechanisms [36] that assign different weights to individual features. Moreover, employing methods such as LIME (Local Interpretable Model-agnostic Explanations) [37] and Grad-CAM (Gradient-weighted Class Activation Mapping) [38] could provide deeper insights into the features most influential in classifying *Z. jujuba* Mill., allowing for a better understanding of the model's reliance on specific characteristics. These follow-up studies could lead to the development of models optimized for medicinal material classification.

According to the results of this study, the CNN-based model developed here is highly lightweight, fast, and capable of achieving high classification accuracy using only RGB images. Therefore, as illustrated in Figure 9, integrating this model into a compact embedded computing device equipped with a USB-based RGB camera module, a display device, and a box-shaped structure providing a consistent imaging environment with connected LED lights could serve as a portable auxiliary tool for practical use in the medicinal materials market.

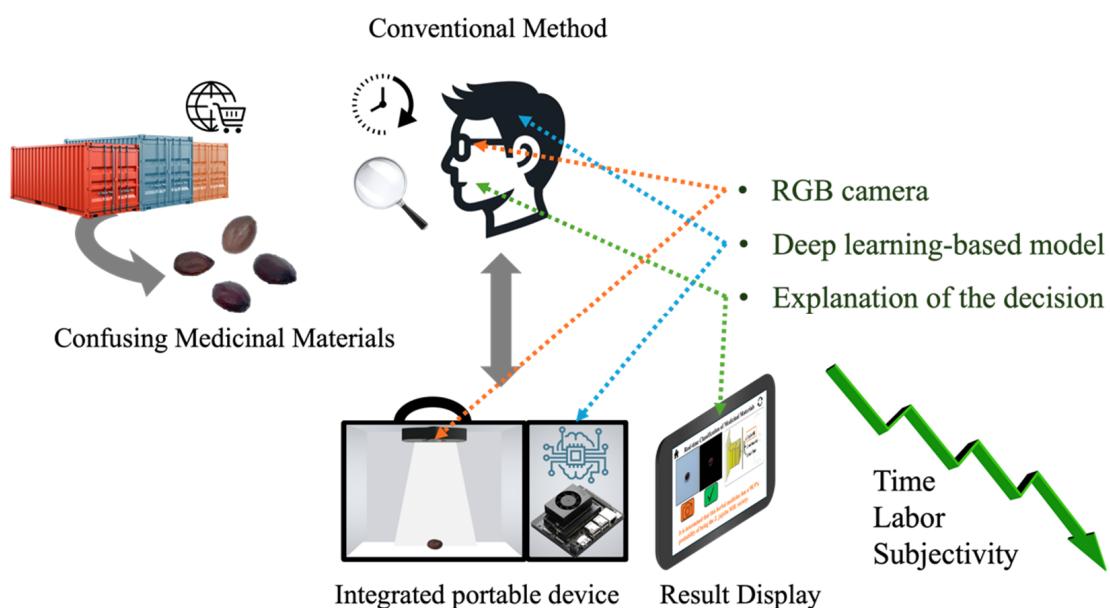


Figure 9. Future work: development of a portable medicinal materials classifier device.

5. Conclusions

This study aimed to explore the potential of deep learning to develop models capable of rapidly and accurately distinguishing seeds of *Ziziphus jujuba* Mill. var. *spinosa*, which is recognized as a medicinal material, from *Ziziphus mauritiana* Lam. and *Hovenia dulcis* Thunb., which are not recognized as medicinal materials. We utilized three advanced deep

learning models—ResNet-50, Inception-v3, and DenseNet-121—to classify RGB images of these seeds and compared their performance using various evaluation metrics.

In this study, the CNN-based models—ResNet-50, Inception-v3, and DenseNet-121—successfully and efficiently discriminated between medicinal and non-medicinal seeds. In particular, DenseNet-121 demonstrated superior performance, achieving the highest classification accuracy and showcasing its robust capability in distinguishing species differences. This model achieved a high true positive (TP) rate for *Z. jujuba* Mill., recording the highest values across all metrics. Furthermore, the t-SNE analysis revealed its exceptional ability to effectively learn and differentiate species-specific characteristics. Additionally, DenseNet-121 required less memory and exhibited faster processing speeds than its counterparts, making it highly suitable for practical applications. However, since the model relies on RGB information, it is expected that factors such as seed conditions, camera quality, and lighting conditions may significantly affect its performance in real-world applications.

To address these limitations, future research could focus on developing a standardized device capable of providing controlled imaging conditions based on the proposed methodology. Including seeds under diverse conditions in the training dataset could further enhance the robustness of the model. Additionally, the application of the model to morphologically similar medicinal materials, such as *Lycium chinense* and *Cornus officinalis*, could be explored. Moreover, incorporating interpretability techniques such as LIME and Grad-CAM would enhance the user-friendliness of the model and device by providing explanations for the model's decisions.

In conclusion, the model and methodology developed in this study offer a practical solution to address issues of subjectivity, time consumption, and labor intensity in the distribution process of *Z. jujuba* Mill., while also helping to manage quality-related challenges. Furthermore, this work is expected to lay a solid foundation for future research on the classification and authentication of a broader range of medicinal materials.

Author Contributions: Conceptualization, Y.-J.J. and D.-H.J.; methodology, Y.-J.J. and D.-H.J.; software, H.L.; validation, Y.-J.J. and S.J.P.; formal analysis, Y.-J.J.; investigation, Y.-J.J.; resources, H.-Y.K.; data curation, S.J.P.; writing—original draft preparation, Y.-J.J.; writing—review and editing, D.-H.J.; visualization, Y.-J.J.; supervision, D.-H.J. and H.-Y.K.; project administration, D.-H.J.; funding acquisition, D.-H.J. and H.-Y.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by a grant from the Ministry of Food and Drug Safety (23192MFDS106) in 2024. Partial support was also provided by the BK21 FOUR program of the Graduate School, Kyung Hee University (GS-1-JO-NON-20240359).

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding authors.

Conflicts of Interest: The authors declare that this research was conducted in the absence of any commercial or financial relationships that could be construed as potential conflicts of interest.

References

- Wang, W.; Xu, J.; Fang, H.; Li, Z.; Li, M. Advances and Challenges in Medicinal Plant Breeding. *Plant Sci.* **2020**, *298*, 110573. [[CrossRef](#)] [[PubMed](#)]
- Zahra, W.; Rai, S.N.; Birla, H.; Singh, S.S.; Rathore, A.S.; Dilnashin, H.; Keswani, C.; Singh, S.P. Economic Importance of Medicinal Plants in Asian Countries. In *Bioeconomy for Sustainable Development*; Keswani, C., Ed.; Springer: Singapore, 2020; pp. 359–377. ISBN 9789811394317.
- Gao, T.; Yao, H.; Song, J.; Liu, C.; Zhu, Y.; Ma, X.; Pang, X.; Xu, H.; Chen, S. Identification of Medicinal Plants in the Family Fabaceae Using a Potential DNA Barcode ITS2. *J. Ethnopharmacol.* **2010**, *130*, 116–121. [[CrossRef](#)] [[PubMed](#)]
- Yu, J.; Wu, X.; Liu, C.; Newmaster, S.; Ragupathy, S.; Kress, W.J. Progress in the Use of DNA Barcodes in the Identification and Classification of Medicinal Plants. *Ecotoxicol. Environ. Saf.* **2021**, *208*, 111691. [[CrossRef](#)] [[PubMed](#)]
- Du, S.; Hu, X.; Guo, Y.; Wang, S.; Yang, X.; Wu, Z.; Huang, Y. A Comparative Plastome Analysis of *Ziziphus Jujuba* Var. *Spinosa* (Bunge) Hu Ex H. F. Chow and Implication of the Origin of Chinese Jujube. *AoB Plants* **2023**, *15*, plad006. [[CrossRef](#)] [[PubMed](#)]
- Ben Mahmoud, K.; Wasli, H.; Ben Mansour, R.; Jemai, N.; Selmi, S.; Jemmali, A.; Ksouri, R. Antidiabetic, Antioxidant and Chemical Functionalities of *Ziziphus Jujuba* (Mill.) and *Moringa Oleifera* (Lam.) Plants Using Multivariate Data Treatment. *S. Afr. J. Bot.* **2022**, *144*, 219–228. [[CrossRef](#)]

7. Zhang, R.; Chen, J.; Shi, Q.; Li, Z.; Peng, Z.; Zheng, L.; Wang, X. Quality Control Method for Commercially Available Wild Jujube Leaf Tea Based on HPLC Characteristic Fingerprint Analysis of Flavonoid Compounds. *J. Sep. Sci.* **2014**, *37*, 45–52. [[CrossRef](#)]
8. Damiano, S.; Forino, M.; De, A.; Vitali, L.A.; Lupidi, G.; Taglialatela-Scafati, O. Antioxidant and Antibiofilm Activities of Secondary Metabolites from *Ziziphus Jujuba* Leaves Used for Infusion Preparation. *Food Chem.* **2017**, *230*, 24–29. [[CrossRef](#)]
9. Liu, S.-J.; Lv, Y.-P.; Tang, Z.-S.; Zhang, Y.; Xu, H.-B.; Zhang, D.-B.; Cui, C.-L.; Liu, H.-B.; Sun, H.-H.; Song, Z.-X.; et al. *Ziziphus Jujuba* Mill., a Plant Used as Medicinal Food: A Review of Its Phytochemistry, Pharmacology, Quality Control and Future Research. *Phytochem. Rev.* **2021**, *20*, 507–541. [[CrossRef](#)]
10. Li, L.; Peng, J.; Bai, R. Analysis of the Genetic Relationships in Chinese *Ziziphus* with SRAP Markers. *Agric. Sci. China* **2010**, *9*, 1278–1284. [[CrossRef](#)]
11. Sun, S.; Liu, H.; Xu, S.; Yan, Y.; Xie, P. Quality Analysis of Commercial Samples of *Ziziphi spinosae* Semen (*Suanzaoren*) by Means of Chromatographic Fingerprinting Assisted by Principal Component Analysis. *J. Pharm. Anal.* **2014**, *4*, 217–222. [[CrossRef](#)]
12. Zhang, Y.; Gao, Z.; Wang, X.; Liu, Q. Image Representations of Numerical Simulations for Training Neural Networks. *Comput. Model. Eng. Sci.* **2022**, *134*, 821–833. [[CrossRef](#)]
13. Xin, J.; Xu, W.; Cao, B.; Wang, T.; Zhang, S. A Deep-Learning-Based MAC for Integrating Channel Access, Rate Adaptation and Channel Switch. *arXiv* **2024**, arXiv:2406.02291. [[CrossRef](#)]
14. LeCun, Y.; Bengio, Y.; Hinton, G. Deep Learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)] [[PubMed](#)]
15. Alzubaidi, L.; Zhang, J.; Humaidi, A.J.; Al-Dujaili, A.; Duan, Y.; Al-Shamma, O.; Santamaría, J.; Fadhel, M.A.; Al-Amidie, M.; Farhan, L. Review of Deep Learning: Concepts, CNN Architectures, Challenges, Applications, Future Directions. *J. Big Data* **2021**, *8*, 53. [[CrossRef](#)]
16. Albawi, S.; Mohammed, T.A.; Al-Zawi, S. Understanding of a Convolutional Neural Network. In Proceedings of the 2017 International Conference on Engineering and Technology (ICET), Antalya, Turkey, 21–23 August 2017; pp. 1–6.
17. Bi, C.; Zhang, S.; Chen, H.; Bi, X.; Liu, J.; Xie, H.; Yu, H.; Song, S.; Shi, L. Non-Destructive Classification of Maize Seeds Based on RGB and Hyperspectral Data with Improved Grey Wolf Optimization Algorithms. *Agronomy* **2024**, *14*, 645. [[CrossRef](#)]
18. Rong, D.; Wang, H.; Ying, Y.; Zhang, Z.; Zhang, Y. Peach Variety Detection Using VIS-NIR Spectroscopy and Deep Learning. *Comput. Electron. Agric.* **2020**, *175*, 105553. [[CrossRef](#)]
19. Kurtulmuş, F. Identification of Sunflower Seeds with Deep Convolutional Neural Networks. *J. Food Meas. Charact.* **2021**, *15*, 1024–1033. [[CrossRef](#)]
20. Xing, X.; Liu, C.; Han, J.; Feng, Q.; Lu, Q.; Feng, Y. Wheat-Seed Variety Recognition Based on the GC_DRNet Model. *Agriculture* **2023**, *13*, 2056. [[CrossRef](#)]
21. Nart, D.D.; Gardiman, M.; Alba, V.; Tarricone, L.; Storchi, P.; Roccatelli, S.; Ammoniaci, M.; Tosi, V.; Perria, R.; Carraro, R. Vine Variety Identification through Leaf Image Classification: A Large-Scale Study on the Robustness of Five Deep Learning Models. *J. Agric. Sci.* **2024**, *162*, 19–32. [[CrossRef](#)]
22. Shelke, A.; Mehendale, N. A CNN-Based Android Application for Plant Leaf Classification at Remote Locations. *Neural Comput. Appl.* **2023**, *35*, 2601–2607. [[CrossRef](#)]
23. Mesías-Ruiz, G.A.; Borra-Serrano, I.; Peña, J.M.; de Castro, A.I.; Fernández-Quintanilla, C.; Dorado, J. Weed Species Classification with UAV Imagery and Standard CNN Models: Assessing the Frontiers of Training and Inference Phases. *Crop Prot.* **2024**, *182*, 106721. [[CrossRef](#)]
24. Mir, M.A.; Qazi, F.; Naseem, M.; Zia, S.S.; Agha, D.-S.; Mubeen, T. Invisibility Cloak Using Color Extraction and Image Segmentation with OpenCV. In Proceedings of the 2022 Global Conference on Wireless and Optical Technologies (GCWOT), Malaga, Spain, 14–17 February 2022; pp. 1–6.
25. Boncelet, C. Chapter 7—Image Noise Models. In *The Essential Guide to Image Processing*; Bovik, A., Ed.; Academic Press: Boston, MA, USA, 2009; pp. 143–167, ISBN 978-0-12-374457-9.
26. 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, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
27. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the Inception Architecture for Computer Vision. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 2818–2826.
28. Huang, G.; Liu, Z.; van der Maaten, L.; Weinberger, K.Q. Densely Connected Convolutional Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 4700–4708.
29. Agarap, A.F. Deep Learning Using Rectified Linear Units (ReLU). *arXiv* **2019**, arXiv:1803.08375.
30. Fushiki, T. Estimation of Prediction Error by Using K-Fold Cross-Validation. *Stat. Comput.* **2011**, *21*, 137–146. [[CrossRef](#)]
31. Yacoubi, R.; Axman, D. Probabilistic Extension of Precision, Recall, and F1 Score for More Thorough Evaluation of Classification Models. In Proceedings of the First Workshop on Evaluation and Comparison of NLP Systems, Online, 20 November 2020; Eger, S., Gao, Y., Peyrard, M., Zhao, W., Hovy, E., Eds.; Association for Computational Linguistics: Stroudsburg, PA, USA, 2020; pp. 79–91.
32. Luque, A.; Carrasco, A.; Martín, A.; de las Heras, A. The Impact of Class Imbalance in Classification Performance Metrics Based on the Binary Confusion Matrix. *Pattern Recognit.* **2019**, *91*, 216–231. [[CrossRef](#)]
33. Kraemer, H.C. Kappa Coefficient. In *Wiley StatsRef: Statistics Reference Online*; John Wiley & Sons, Ltd.: Hoboken, NJ, USA, 2015; pp. 1–4, ISBN 978-1-118-44511-2.

34. Bradley, A.P. The Use of the Area under the ROC Curve in the Evaluation of Machine Learning Algorithms. *Pattern Recognit.* **1997**, *30*, 1145–1159. [[CrossRef](#)]
35. Arora, S.; Hu, W.; Kothari, P.K. An Analysis of the T-SNE Algorithm for Data Visualization. In *31st Conference on Learning Theory*; PMLR: Birmingham, UK, 2018; pp. 1455–1462.
36. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention Is All You Need. In *Advances in Neural Information Processing Systems*; Curran Associates, Inc.: Nice, France, 2017; Volume 30.
37. Di Cicco, V.; Firmani, D.; Koudas, N.; Merialdo, P.; Srivastava, D. Interpreting Deep Learning Models for Entity Resolution: An Experience Report Using LIME. In *Second International Workshop on Exploiting Artificial Intelligence Techniques for Data Management*; Association for Computing Machinery: New York, NY, USA, 2019; pp. 1–4.
38. Selvaraju, R.R.; Cogswell, M.; Das, A.; Vedantam, R.; Parikh, D.; Batra, D. Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization. *Int. J. Comput. Vis.* **2020**, *128*, 336–359. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.