

Article

A Rapid and Highly Efficient Method for the Identification of Soybean Seed Varieties: Hyperspectral Images Combined with Transfer Learning

Shaolong Zhu [†], Jinyu Zhang [†], Maoni Chao, Xinjuan Xu, Puwen Song, Jinlong Zhang and Zhongwen Huang ^{*}

School of Life Science and Technology, Henan Institute of Science and Technology/Henan Collaborative Innovation Center of Modern Biological Breeding, Xinxiang 453003, China; zsl_94121@126.com (S.Z.); zjy891@126.com (J.Z.); chaomaoni@126.com (M.C.); xinjuan_xu@163.com (X.X.); song263549963@126.com (P.S.); zjld163yx@163.com (J.Z.)

^{*} Correspondence: hzw@hist.edu.cn; Tel.: +86-1383-7326-136

[†] These authors contributed equally to this work.

Academic Editors: Heinz Wilhelm Siesler and Hui Yan

Received: 18 November 2019; Accepted: 27 December 2019; Published: 30 December 2019



Abstract: Convolutional neural network (CNN) can be used to quickly identify crop seed varieties. 1200 seeds of ten soybean varieties were selected, hyperspectral images of both the front and the back of the seeds were collected, and the reflectance of soybean was derived from the hyperspectral images. A total of 9600 images were obtained after data augmentation, and the images were divided into a training set, validation set, and test set with a 3:1:1 ratio. Pretrained models (AlexNet, ResNet18, Xception, InceptionV3, DenseNet201, and NASNetLarge) after fine-tuning were used for transfer training. The optimal CNN model for soybean seed variety identification was selected. Furthermore, the traditional machine learning models for soybean seed variety identification were established by using reflectance as input. The results show that the six models all achieved 91% accuracy in the validation set and achieved accuracy values of 90.6%, 94.5%, 95.4%, 95.6%, 96.8%, and 97.2%, respectively, in the test set. This method is better than the identification of soybean seed varieties based on hyperspectral reflectance. The experimental results support a novel method for identifying soybean seeds rapidly and accurately, and this method also provides a good reference for the identification of other crop seeds.

Keywords: hyperspectral image; transfer learning; pretrained network; soybean seed; variety identification

1. Introduction

Accurate identification of soybean seed varieties is of great significance for protecting farmers' interests, ensuring agricultural production, and maintaining order in the seed market. The usual methods for the identification of seed varieties are morphological analysis [1], chemical analysis (random amplified polymorphic DNA (RAPD) and simple sequence repeat (SSR) molecular markers [2,3], protein electrophoresis [4], liquid chromatography [5], spectral analysis [6–8], and image analysis [9]. Morphological analysis requires that the appraiser has extensive experience, and the accuracy is low for two or more varieties with little morphological difference [10]. Although the chemical analysis method has a high identification accuracy, it is destructive to the sample, time consuming, and so difficult to operate that nonprofessionals are not competent to perform it [11].

The spectral analysis method reflects the difference in the internal physical structure and the chemical composition of seed varieties through spectral information, which has the advantages of being fast, accurate, and nondestructive [12]; it has been applied to multiple tasks related to seeds, such as the identification of soybean [13,14], rice [6] and maize [15–17] seed varieties, the identification of nontransgenic and transgenic seeds [18–20], the identification of seed geographical sources [21], the identification of infected germ seeds [22], the identification of infected pest seeds and healthy seeds [23], identification of the year of seeds [24,25], and the determination of tomato [26], soybean [27,28], corn [28], muskmelon [29], cabbage and radish [30] seed vitality. Less research has been performed on image analysis than spectral analysis, and only some plants have been studied through image analysis, such as peppers [31], paddy [32,33], and corn [34]. Some researchers compared spectral analysis and image analysis for the identification of corn seeds, and the results showed that the classification accuracy based on spectral features (95% and 96.2% for each side) is higher than the classification accuracy based on the morphology and texture features (86.8% and 96.6% for each side), while the combination of the two can reach 96.3% and 98.2% discriminant accuracy for each side [7]. Although the spectral analysis and image analysis methods performed well, they still need to go through the processes of reflectance extraction, spectral pretreatment, feature band selection, and morphological and texture feature extraction. In addition, there are problems such as fewer choices of varieties, lower identification accuracy, the reflectance is easily affected by seed relative humidity, and poor portability of the identification models.

Deep learning (DL) is an important artificial intelligence method that enables machines to acquire knowledge from data autonomously [35]. In addition, if the seeds are not severely wet (seed color does not change and seeds do not swell), the relative humidity has no influence on the method of hyperspectral imaging combined deep learning. At present, deep networks have been successfully applied to plant disease identification [36–38], drought monitoring [39], land type classification [40], weed detection [41], and other areas of agriculture. To date, there are few reports on the identification of soybean seed varieties by deep learning, and whether it has advantages that is also unknown. Although some studies used deep learning combined with spectral analysis for seed identification [42–44], these all used one-dimensional spectra as input, while three-dimensional images contain more information. Moreover, in practical applications, neural networks are usually not trained from scratch for a new task: such an operation is obviously very time consuming. In particular, these models have a large number of parameters that need to be trained and require a very large amount of data. When the amount of data required to construct the model cannot be obtained, the model may be overfit or may fall into a local optimal solution [45–48]. Using transfer learning to fine-tune the network is faster and easier than randomly training the weights from scratch and does not require many images.

In this paper, after obtaining the hyperspectral images of 9600 soybean seeds from 10 soybean varieties and taking the images as input, pretrained networks, such as AlexNet and ResNet18, were used to carry out transfer training. It was also compared with the traditional machine learning based on spectral reflectance. The aim of this study was to demonstrate that it is feasible and superior to use a deep learning model for seed variety identification based on images as input and to provide a theoretical basis and practical method for the more rapid, accurate, and nondestructive identification of soybean varieties.

2. Results and Discussion

2.1. Training Progress

The language of deep learning in this article was MATLAB (MATLAB R2019a, The Math Works Inc., Natick, MA, USA), the graphics processing unit (GPU) was a NVIDIA GeForce RTX 2080Ti, and the display memory was 11 GB. Among 0–800 iterations, the accuracy increased and the loss declined rapidly (Figure 1). After 800 iterations, the accuracy of the 6 models all reached 75%, the accuracy increased, and the loss declined slowly. When the training was over, the six models (AlexNet,

ResNet18, Xception, InceptionV3, DenseNet201, and NASNetLarge) achieved accuracy values of 91.6%, 95.6%, 96.6%, 96.7%, 97.5%, and 98.2%, respectively, in the validation set. The training times of the six models were 18 min, 117 min, 84 min, 798 min, 455 min, and 1914 min, respectively, and the results showed that the training time is not only related to the depth of the network but also to the other factors (such as the width of the network). From the curve comparison of the six models, the training accuracy of AlexNet fluctuated widely, while NASNetLarge had a small fluctuation range. From AlexNet to NASNetLarge, the number of network layers increased, indicating that the fluctuation range of training accuracy was inversely proportional to the number of layers. Moreover, the larger the number of network layers, the fewer the number of iterations to achieve a stable training accuracy (with 1400 iterations, the accuracy of NASNetLarge reached a high level, and there was little change after 1400 iterations). Comparing the training curve and the validation curve, the model did not exhibit the phenomena of negative transfer, over-fitting, or under-fitting. The results show that transfer training was very successful.

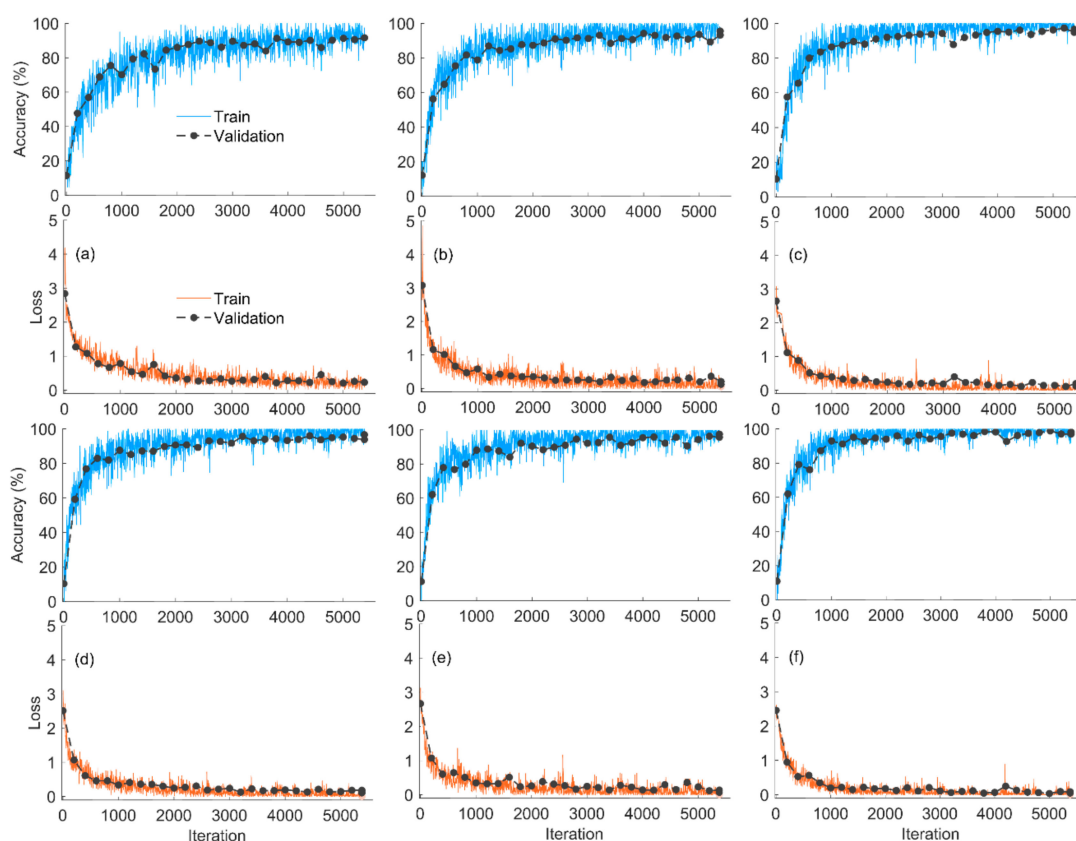


Figure 1. Training progress of the six pretrained models. (a) AlexNet; (b) ResNet18; (c) Xception; (d) InceptionV3; (e) DenseNet201; and (f) NASNetLarge.

2.2. Test Results

Compared with the training accuracy and validation accuracy, the test accuracy was the most important evaluation index. AlexNet, ResNet18, Xception, InceptionV3, DenseNet201, and NASNetLarge achieved accuracy values of 90.6%, 94.5%, 95.4%, 95.6%, 96.8%, and 97.2%, respectively, in the test set. Of the six models, NASNetLarge performed the best with 54 misjudgments (Figure 2), and among them, Nannong 1606 had the most misjudgments while Shangdou 1201 and Zheng 3074 had the least misjudgments. According to the study, the greater the number of deep neural network layers, the better the performance of the identification models, and the conclusion is the same in other fields [49–51]. However, there are also different conclusions [46,52–54], as there are several factors that cause this situation, including dataset, network depth, network width, network

structure, and parameter settings, and overall, the deep network is indeed slightly better than the shallow network.

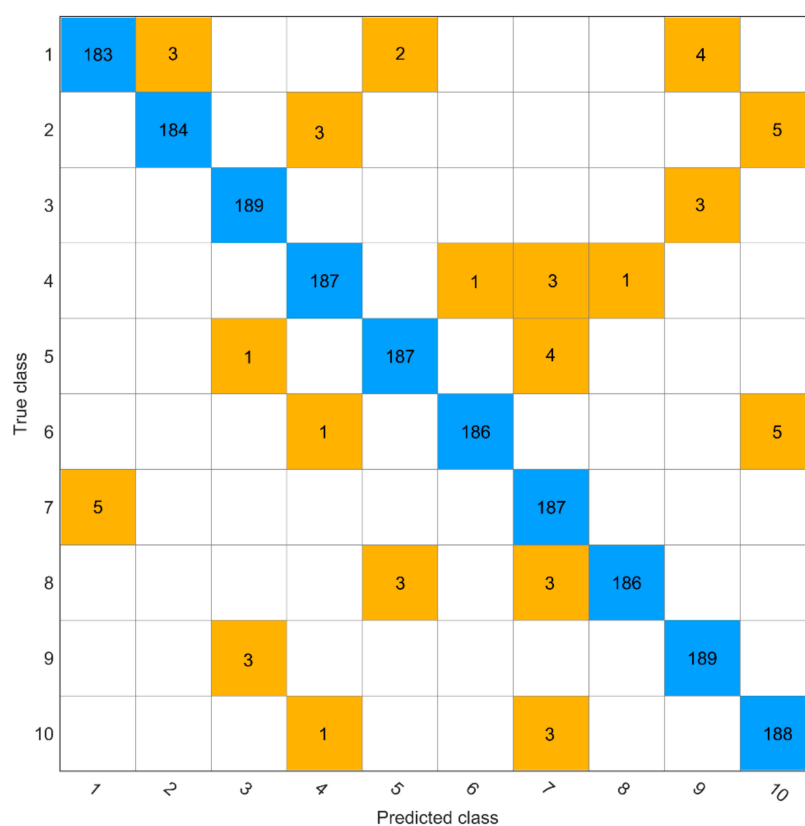


Figure 2. Confusion matrix. (1) Nannong 1606; (2) Shangdou 161; (3) Shangdou 1201; (4) Shangdou 1310; (5) Yudou 18; (6) Yudou 22; (7) Yudou 25; (8) Zheng 196; (9) Zheng 3074; and (10) Zheng 9525.

2.3. Spectral Pretreatment Process

The pretreatment results are shown in Figure 3. The SG (Figure 3b) results were smoother than those in Figure 3a and eliminated the noise of the original spectrum (OS) at 1000 nm. Since it did not involve the average spectra of all samples, the difference between each spectral curve was still large. SNV needs to be calculated based on the spectral average of all wavelength points in a sample. Therefore, the difference was significantly reduced (Figure 3c) between samples after pretreatment compared with Figure 3a. The geometric meaning of the derivative is the tangent slope of the curve at a certain point, so the derivative can magnify the difference. With FD pretreatment, the spectral differences among different soybeans were mainly in the ranges of 623–638 nm, 649–659 nm, and 675–687 nm (Figure 3d), and these different bands were all within the range of bands with large differences in the original spectra, which indicates that derivative transformation highlighted the characteristic wavelengths.

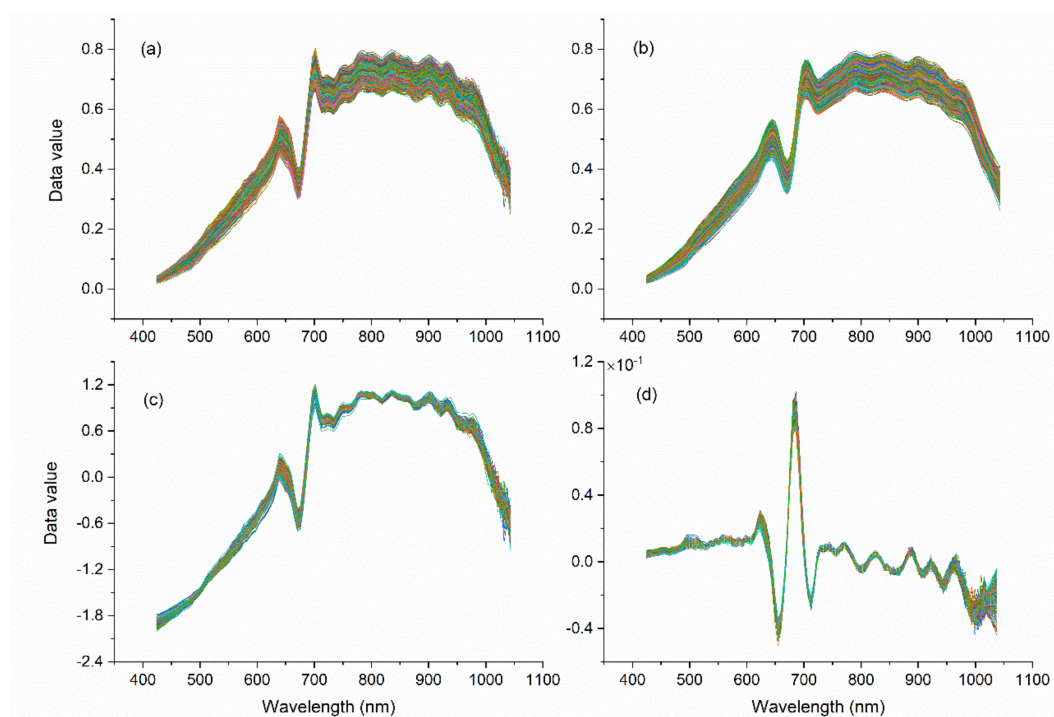


Figure 3. Pretreatment spectrum curves. (a) original spectrum (OS) curve; (b) Savitzky-Golay (SG); (c) standard normal variate (SNV); and (d) first derivative (FD).

2.4. Identification of Models Using Hyperspectral Reflectance

Through PCA, the numbers of principal component factors extracted from the three pretreatments were 4, 7, and 24, and the cumulative loads were 97.3%, 85.0%, and 61.3%, respectively. The model based on original spectrum had achieved the highest training accuracy and test accuracy values of 61.7% and 58.7%, respectively. By comparison, it is useful to preprocess the spectral reflectance, and the FD/GS-SVM combination had the highest identification accuracy, with training accuracy and test accuracy values of 89.8% and 80.4%, respectively (Figure 4). The identification model based on the hyperspectral reflectance was sensitive to the pretreatment method, feature extraction method, and classifier. The results obtained by using different combinations were very different. Other studies also reached this conclusion [55–57]. There are many steps in the spectrum analysis method. Moreover, as the number of steps increases, the uncertainty of the model will also increase. For example, if the optimal combination of varieties A and B is X, the optimal combination of varieties A and C is Y, and the optimal combination of varieties B and C is Z, then the optimal combination of varieties A, B, and C may not be X, Y, or Z. It is difficult to find an optimal combination with a high identification accuracy and strong portability.

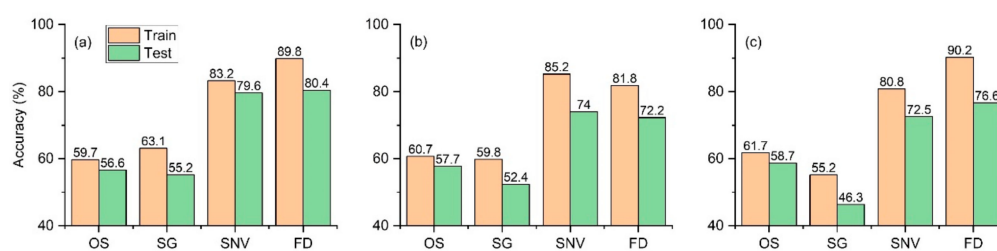


Figure 4. Identification accuracy based on the hyperspectral reflectance of three classifiers with different pretreatments. (a) Grid search optimization support vector machine (GS-SVM) classifier; (b) ensemble learning (EL) classifier; and (c) artificial neural network (ANN) classifier.

2.5. Comparison Analysis

In this study, two identification methods of soybean seed varieties (the hyperspectral image-based deep learning and the hyperspectral reflectance-based machine learning) were compared. The results showed that the worst model (AlexNet with 90.6% test accuracy) of deep learning was better than the best model (FD/GS-SVM with 80.4% test accuracy) of machine learning. In addition, the deep learning method need not to extract the spectral reflectance or perform spectral preprocessing and feature extraction, and there was no need for image cropping or the manual extraction of morphological and texture features of hyperspectral images, which saves a great deal of time. So, the hyperspectral imaging combined deep learning method was superior to the identification model based on hyperspectral reflectance in all respects.

Data diversity is one of the key factors to ensure the generalization ability of the model [37]. Although the study used nearly 10,000 images, the amount of data was still small for 10 soybean varieties. In addition, these models were highly accurate in identifying the 10 soybean varieties, but further verification is needed to determine whether the pretrained models selected for this study will still achieve a high accuracy in the identification of seeds from dozens or hundreds of soybean varieties or whether better CNN models are needed to achieve satisfactory results. The identification model based on hyperspectral images combined with transfer learning is significantly faster than the identification method based on hyperspectral reflectance, but the SOC-710VP imaging spectrometer still needs 35 s to acquire one hyperspectral image. A next step is to determine whether digital cameras or mobile phone cameras can accurately identify images so that this identification method can meet the needs of ordinary people, rather than the needs of professionals. Transfer learning is an important research direction in the field of artificial intelligence in the next few years, and its development provides new research ideas and approaches for the quality and safety detection of agricultural products [58]. The use of machine vision technology combined with deep learning to achieve grain variety detection, grain grading, automated fruit and vegetable variety (origin) detection, and high identification accuracy, etc., and all efforts should be made to increase the research and promotion of this technology.

3. Materials and Methods

3.1. Materials

The choice of seed variety is the primary consideration in the identification of seed varieties. In terms of material selection, this experiment fully tested the feasibility and ability of hyperspectral images combined with a transfer learning model to identify grain varieties based on seed type, luster, hilum color, 100-seed weight, crude protein content, and crude fat content (Table 1). We selected 10 varieties, which all have yellow seed coats, that are grown on a large scale. One hundred twenty complete, undamaged, spotless seeds per variety were selected and divided into a training set, validation set, and test set with a 3:1:1 ratio. The samples were put in a 38 °C oven for 24 h to ensure the same relative humidity of each variety.

Table 1. Soybean seed quality.

Variety	Seed Type	Luster	Hilum Color	100-Seed Weight (g)	Crude Protein (%)	Crude Fat (%)
Nannong 1606	circular	yes	brown	15.4	36.0	19.7
Shangdou 161	circular	yes	brown	21.6	35.6	19.6
Shangdou 1201	oval	yes	brown	19.1	43.1	20.2
Shangdou 1310	oval	weak	pale brown	18.0	42.1	20.5
Yudou 18	circular	yes	brown	16.8	44.5	18.8
Yudou 22	circular	yes	pale brown	19.3	46.5	18.9
Yudou 25	circular	yes	brown	18.4	46.3	17.1
Zheng 196	circular	weak	pale brown	17.4	40.7	19.5
Zheng 3074	flat oval	weak	pale brown	19.7	40.9	17.1
Zheng 9525	circular	yes	pale brown	21.7	45.0	17.7

3.2. Equipment

This study used an SOC-710 portable hyperspectral imaging spectrometer (SOC 710VP, Surface Optics Corporation, San Diego, CA, USA) with a spectral range of 400–1000 nm and a spectral resolution of 4.6875 nm. The light source was two 100 W halogen lamps (Lowel Pro-light, Lower Light Manufacturing Inc., Hauppauge, NY, USA). Other equipment, such as a darkroom, standard gray Spectralon panel and computer, were used. The standard gray Spectralon panel used for reflectance conversion was placed near the seed. The hyperspectral imaging spectrometer lens was placed 30 cm away from the stage, and the incident light of the halogen source was set at an angle of 60° to the stage (Figure 5).

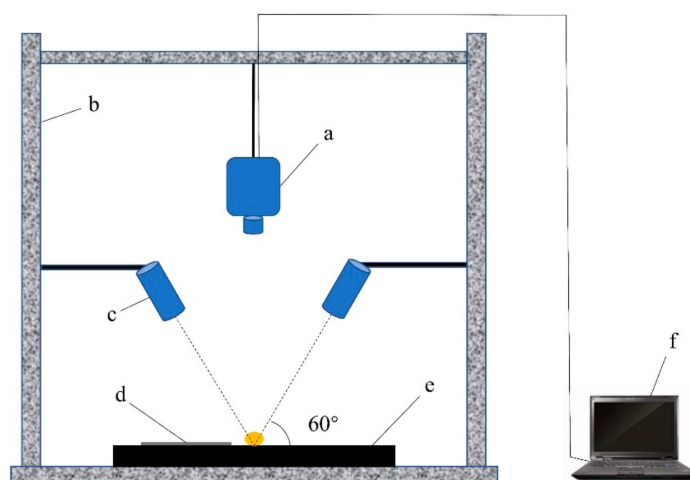


Figure 5. Hyperspectral imaging system. (a) Imaging spectrometer; (b) darkroom; (c) light source; (d) standard gray Spectralon panel; (e) loading stage; and (f) computer.

3.3. Hyperspectral Image Acquisition

SOC 710 Acquisition Software was used to collect images: the integration was set to 20 ms, and the gain was set to 3. First, the lens was covered to obtain black field (dark current) data. Then, the cover was removed to collect the front and back hyperspectral images of the seed (Figure 6). The standard gray Spectralon panel was placed near the seed, which was used for reflectance conversion. Finally, SRAnal 710 (Surface Optics Corporation, San Diego, CA, USA) was used for radiation calibration (the spectrometer manufacturers provided the radiation calibration) and dark current correction.

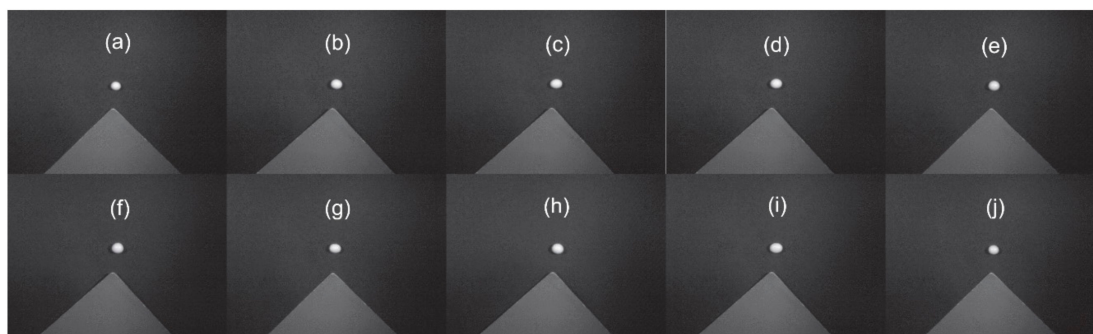


Figure 6. Hyperspectral images of the front of each soybean variety. (a) Nannong 1606; (b) Shangdou 161; (c) Shangdou 1201; (d) Shangdou 1310; (e) Yudou 18; (f) Yudou 22; (g) Yudou 25; (h) Zheng 196; (i) Zheng 3074; and (j) Zheng 9525.

3.4. Image Preprocessing

Because the size of the original images is 696-by-520-by-128 (they are the length, width, and number of bands of the images, respectively), principal component was extracted using principal component analysis (PCA) from the 128 bands. To match the three channels of RGB images, we retained the first three principal components. Blank pixels were added to both sides of the short edge of the images before resizing; otherwise the image underwent deformation. Bicubic interpolation algorithm was used to resize the whole image to match the model input size, as each model requires different input image resolutions (Table 2). Finally, the images were rotated 90°, 180°, and 270°, a total of 9600 images were obtained and the numbers of images in the training set, validation set, and test set were 5760, 1920, and 1920, respectively.

Table 2. The pretrained models with properties.

Network	Image Input Size	Layers	Network	Image Input Size	Layers
AlexNet	227-by-227-by-3	25	InceptionV3	229-by-229-by-3	316
ResNet18	224-by-224-by-3	72	DenseNet201	224-by-224-by-3	709
Xception	229-by-229-by-3	171	NASNetLarge	331-by-331-by-3	1244

3.5. Pretrained Networks

Six pretrained networks were used to transfer the model parameters. All of these networks were trained on more than one million images of 1000 categories from the ImageNet database and achieved good recognition accuracy. AlexNet [59] is the simplest network model in this study (Figure 7). Transfer learning refers to a learning process that applies a model learned in an old domain (source domain) to a new domain (target domain) by the similarity between the two domains [48]. In this study, the source domain was the ImageNet dataset, and the target domain was soybean hyperspectral images. The type of transfer learning was parameter/model based which was the most widely used. The structure of neural network can be transferred directly, and fine-tune was a good embodiment of parameter/model based transfer learning.

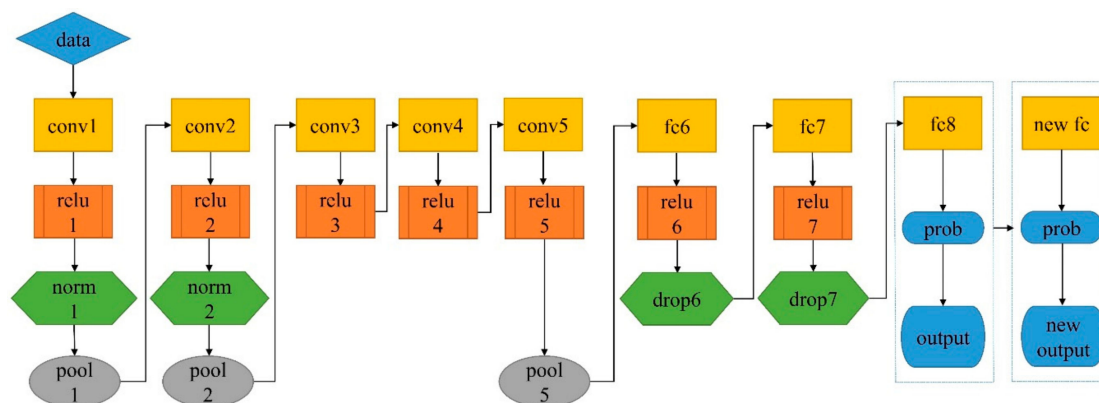


Figure 7. The architecture of AlexNet: conv, relu, norm, pool, fc, drop, and prob are the abbreviations of convolution layer, relu layer, cross channel normalization layer, maxpooling layer, fully connected layer, dropout layer, and softmax layer, respectively.

3.6. Model Parameter Settings

For classification problems, pretrained models generally consist of an input layer, a convolution layer, a rectified linear unit layer, a pooling layer, a fully connected layer, and an output layer. The number of layers in each model is shown in Table 2. When using transfer learning to solve the problem in this study, fine-tuning only the parameters of the last few layers of the model can achieve the goal of transfer learning. As shown in Figure 7, the AlexNet model was applied to soybean seed identification by fine-tuning only the fully connected layers (fc8) and the output parameters. Similarly, for the other five models, the antepenultimate layer was all fully connected layer, and each model had an output size of 1000. We set the output size to 10 to make the network model suitable for the classification of soybean seed hyperspectral images and to debug the parameters, such as the learning rate for the weights (weight learn rate factor). In the process of debugging parameters, different values affect the accuracy of the identification model, and multiple attempts are required to achieve accurate classifications. The last layer of the 6 networks (classification layer) aimed at 1000 category outputs for the ImageNet database, which were replaced with a new classification layer for retraining in this study. The steps mentioned above were completed with MATLAB's deep network designer, which was a visual toolbox. The optimization algorithm was "Adam", which combined Adagrad's ability to deal with sparse gradients and RMSprop's ability to deal with non-stationary targets. Transfer training was carried out directly after adjusting the image data set and fine-tuning the network. The other parameters are shown in Table 3.

Table 3. Parameter settings of all models.

Parameters	Values	Parameters	Values
Momentum	0.9	Max epochs	10
Initial learn rate	0.0001	Mini batch size	10
Initial learn schedule	Piecewise	Shuffle	Every-epoch
Learn rate drop period	10	Validation frequency	200
Learn rate drop factor	0.1	Sequence length	Longest
L2regularization	0.0001	Gradient threshold method	Global-l2norm

3.7. Comparative Experimental Design

3.7.1. Reflectance Conversion

To verify the superiority of deep learning, the traditional identification methods of soybean seed variety based on reflectance were used. First, the environment for visualizing images (ENVI 5.1, Harris Geospatial Solutions, Inc., Boulder, CO, USA) was used to convert the reflectance of soybean seeds.

The region of interest (ROI) of the complete seed image was obtained by an image segmentation algorithm, and the average value of this region was taken as the spectral reflectance. A total of 2400 samples were obtained, and the samples were divided into a training set and test set with a 3:1 ratio. The reflectance of soybean was calculated by using the Equation (1):

$$R = \frac{DN}{DN_N} \times R_N \quad (1)$$

where R is the reflectance of soybean, DN is the digital number of soybean, and the digital number is the brightness value of remote sensing image pixels, DN_N is the digital number of the standard gray Spectralon panel, and R_N is the reflectance of the standard gray Spectralon panel. The R_N was obtained by precalibration in the laboratory. DN and DN_N were measured in this experiment.

3.7.2. Reflectance Preprocessing

The spectral reflectance was pretreated by Savitzky-Golay smoothing (SG), the standard normal variate (SNV), and first derivative (FD). SG performs polynomial least squares fit on the data in a moving window; the number of window points was set to 7, and the polynomial order was set to 2. The SNV algorithm processed each spectrum, and its calculation in essence was a standard normalization of the original spectral data; the SNV was calculated by using the Equation (2):

$$R_{SNV} = \frac{R - \bar{R}}{\sqrt{\frac{\sum_{i=1}^p (R_i - \bar{R})^2}{p-1}}} \quad (2)$$

where R is the original spectrum of a sample, \bar{R} is the spectral average of all the wavelength points in a sample, and $i = 1, 2, \dots, p$, p is the number of wavelength points.

The geometric meaning of the derivative is the slope of the tangent of the curve at a certain point, and the derivative of waveband λ was calculated by using the Equation (3):

$$FD_{(\lambda)} = \frac{R_{(\lambda+1)} - R_{(\lambda-1)}}{(\lambda + 1) - (\lambda - 1)} \quad (3)$$

where $R_{(\lambda+1)}$ is the reflectance at the next waveband of λ , $R_{(\lambda-1)}$ is the reflectance at the last waveband of λ , $(\lambda + 1)$ is the wavelength of the next waveband of λ , $(\lambda - 1)$ is the wavelength of the last waveband of λ .

3.7.3. Principal Component Extraction

The principal component was extracted using PCA. PCA is a statistical analysis method that transforms many variables into a few principal components by dimension reduction technology [21]. It is one of the most common methods to solve redundant or overlapping information. The characteristic values were calculated, and a characteristic value less than 1 indicates that the principal component is not as powerful as the direct use of the original variable. So, principal components with characteristic values less than 1 were removed.

3.7.4. Classifier Selection

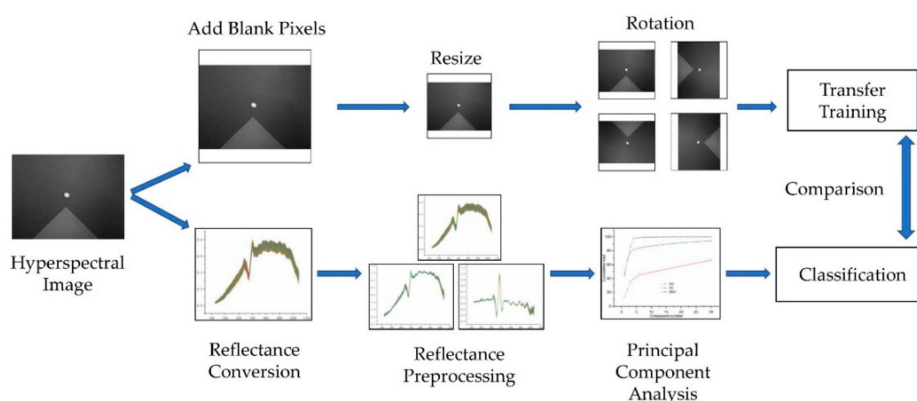
Grid search optimization support vector machine (GS-SVM), ensemble learning (EL), and artificial neural network (ANN) classifiers were used for the identification analysis, and the corresponding parameters are shown in Table 4.

Table 4. Parameter settings of classifiers.

Classifiers	Parameters	Values
GS-SVM	Kernel function	Linear kernel
	Grid c/g bound	−8−8
	Grid c/g step	0.5
EL	Ensemble method	AdaBoost
	Learning rate	0.1
	Number of learners	30
ANN	Type of neural network	Back propagation
	Number of hidden neurons	15
	Training function	Traingdm

3.8. Technical Route

The technical route of this research is shown in Figure 8.

**Figure 8.** Technical route.

4. Conclusions

In this study, 9600 hyperspectral images of 10 soybean varieties were collected, and 6 pretrained networks, such as AlexNet, were used for transfer training to verify the feasibility of the method for identifying soybean seed varieties. The results show that the accurate identification of soybean seed varieties can be realized based on hyperspectral images combined with transfer learning. Moreover, the method based on a hyperspectral image combined with transfer learning has obvious advantages over the method based on the hyperspectral reflectance in terms of the time, operational difficulty, and identification accuracy. In future research, the number of seed varieties and images will be further increased to test the performance of these models, and we hope that these methods can be applied in the seed market.

Author Contributions: Conceptualization, Z.H. and S.Z.; methodology, J.Z. (Jinyu Zhang); software, M.C.; validation, X.X. and J.Z. (Jinlong Zhang); formal analysis, M.C.; investigation, P.S.; resources, M.N.C.; data curation, S.Z.; writing—original draft preparation, S.Z.; writing—review and editing, J.Z. (Jinyu Zhang); visualization, S.Z.; supervision, Z.H.; project administration, Z.H.; funding acquisition, Z.H. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Henan Science and Technology Plan Project (192102110024); the Postgraduate Education Reform and Quality Improvement Project of Henan Province (Yu degree [2018] No. 23); and the National Natural Science Foundation of China (31601347).

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Zhu, D.Z.; Li, Y.F.; Wang, D.C.; Wu, Q.; Zhang, D.Y.; Wang, C. The Identification of Single Soybean Seed Variety by Laser Light Backscattering Imaging. *Sensor. Lett.* **2012**, *10*, 399–404. [[CrossRef](#)]
2. Zhang, C.B.; Peng, B.; Zhang, W.L.; Wang, S.M.; Sun, H.; Dong, Y.S.; Zhao, L.M. Application of ssr Markers for Purity Testing of Commercial Hybrid Soybean (*Glycine max* L.). *J. Agr. Sci. Technol.* **2014**, *16*, 1389–1396.
3. Iqbal, A.; Sadaqat, H.A.; Khan, A.S.; Amjad, M. Identification of Sunflower (*Helianthus annuus*, Asteraceae) Hybrids Using Simple-Sequence Repeat Markers. *Gen. Mol. Res.* **2011**, *10*, 102–106. [[CrossRef](#)] [[PubMed](#)]
4. Rao, P.S.; Bharathi, M.; Reddy, K.B.; Keshavulu, K.; Rao, L.V.S.; Neeraja, C.N. Varietal Identification in Rice (*Oryza sativa*) through Chemical Tests and Gel Electrophoresis of Soluble Seed Proteins. *Indian. J. Agr. Sci.* **2012**, *82*, 304–311.
5. Livaja, M.; Steinemann, S.; Schon, C.C. Application of Denaturing High-Performance Liquid Chromatography for Rice Variety Identification and Seed Purity Assessment. *Mol. Breed.* **2016**, *36*, 1–19.
6. Kong, W.W.; Zhang, C.; Liu, F.; Nie, P.C.; He, Y. Rice Seed Cultivar Identification Using Near-Infrared Hyperspectral Imaging and Multivariate Data Analysis. *Sensors* **2013**, *13*, 8916–8927. [[CrossRef](#)]
7. Yang, X.L.; Hong, H.M.; You, Z.H.; Cheng, F. Spectral and Image Integrated Analysis of Hyperspectral Data for Waxy Corn Seed Variety Classification. *Sensors* **2015**, *15*, 15578–15594. [[CrossRef](#)]
8. Liu, J.J.; Li, Z.; Hu, F.R.; Chen, T.; Zhu, A.J. A Thz Spectroscopy Nondestructive Identification Method for Transgenic Cotton Seed Based on Ga-Svm. *Opt. Quantum Electron.* **2015**, *47*, 313–322. [[CrossRef](#)]
9. Pourreza, A.; Pourreza, H.; Abbaspour-Fard, M.H.; Sadrnia, H. Identification of Nine Iranian Wheat Seed Varieties by Textural Analysis with Image Processing. *Comput. Electron. Agric.* **2012**, *83*, 102–108. [[CrossRef](#)]
10. Boelt, B.; Shrestha, S.; Salimi, Z.; Jorgensen, J.R.; Nicolaisen, M.; Carstensen, J.M. Multispectral imaging—A new tool in seed quality assessment? *Seed Sci. Res.* **2018**, *28*, 222–228. [[CrossRef](#)]
11. Kandala, C.V.K.; Govindarajan, K.N.; Puppala, N.; Settaluri, V.; Reddy, R.S. Identification of Wheat Varieties with a Parallel-Plate Capacitance Sensor Using Fisher’s Linear Discriminant Analysis. *J. Sens.* **2014**, *2014*, 691898. [[CrossRef](#)]
12. Yu, L.N.; Liu, W.J.; Li, W.J.; Qin, H.; Xu, J.; Zuo, M. Non-Destructive Identification of Maize Haploid Seeds Using Nonlinear Analysis Method Based on their Near-Infrared Spectra. *Biosys. Eng.* **2018**, *172*, 144–153. [[CrossRef](#)]
13. Zhu, D.Z.; Wang, K.; Zhou, G.H.; Hou, R.F.; Wang, C. The Nir Spectra Based Variety Discrimination for Single Soybean Seed. *Spectrosc. Spect. Anal.* **2010**, *30*, 3217–3221.
14. Liu, Y.; Wu, T.; Yang, J.J.; Tan, K.Z.; Wang, S.W. Hyperspectral Band Selection for Soybean Classification Based on Information Measure in Frs Theory. *Biosys. Eng.* **2019**, *178*, 219–232. [[CrossRef](#)]
15. Zhang, X.L.; Liu, F.; He, Y.; Li, X.L. Application of Hyperspectral Imaging and Chemometric Calibrations for Variety Discrimination of Maize Seeds. *Sensors* **2012**, *12*, 17234–17246. [[CrossRef](#)]
16. Zhao, Y.Y.; Zhu, S.S.; Zhang, C.; Feng, X.P.; Feng, L.; He, Y. Application of Hyperspectral Imaging and Chemometrics for Variety Classification of Maize Seeds. *RSC Adv.* **2018**, *8*, 1337–1345. [[CrossRef](#)]
17. Huang, M.; He, C.J.; Zhu, Q.B.; Qin, J.W. Maize Seed Variety Classification Using the Integration of Spectral and Image Features Combined with Feature Transformation Based on Hyperspectral Imaging. *Appl. Sci.* **2016**, *6*, 183. [[CrossRef](#)]
18. Wang, H.L.; Yang, X.D.; Zhang, C.; Guo, D.Q.; Bao, Y.D.; He, Y.; Liu, F. Fast Identification of Transgenic Soybean Varieties Based near Infrared Hyperspectral Imaging Technology. *Spectrosc. Spect. Anal.* **2016**, *36*, 1843–1847.
19. Liu, C.H.; Liu, W.; Lu, X.Z.; Chen, W.; Yang, J.B.; Zheng, L. Nondestructive Determination of Transgenic Bacillus Thuringiensis Rice Seeds (*Oryza sativa* L.) Using Multispectral Imaging and Chemometric Methods. *Food Chem.* **2014**, *153*, 87–93. [[CrossRef](#)]
20. Liu, W.; Liu, C.H.; Hu, X.H.; Yang, J.B.; Zheng, L. Application of Terahertz Spectroscopy Imaging for Discrimination of Transgenic Rice Seeds with Chemometrics. *Food Chem.* **2016**, *210*, 415–421. [[CrossRef](#)]
21. Gao, J.F.; Li, X.L.; Zhu, F.; He, Y. Application of Hyperspectral Imaging Technology to Discriminate Different Geographical Origins of *Jatropha curcas* L. Seeds. *Comput. Electron. Agric.* **2013**, *99*, 186–193. [[CrossRef](#)]
22. Baek, I.; Kim, M.S.; Cho, B.K.; Mo, C.; Barnaby, J.Y.; McClung, A.M.; Oh, M. Selection of Optimal Hyperspectral Wavebands for Detection of Discolored, Diseased Rice Seeds. *Appl. Sci.* **2019**, *9*, 1027. [[CrossRef](#)]

23. Chelladurai, V.; Karuppiyah, K.; Jayas, D.S.; Fields, P.G.; White, N.D.G. Detection of *Callosobruchus maculatus* (f.) Infestation in Soybean Using Soft X-ray and Nir Hyperspectral Imaging Techniques. *J. Stored Prod. Res.* **2014**, *57*, 43–48. [[CrossRef](#)]
24. Huang, M.; Tang, J.Y.; Yang, B.; Zhu, Q.B. Classification of Maize Seeds of Different Years Based on Hyperspectral Imaging and Model Updating. *Comput. Electron. Agric.* **2016**, *122*, 139–145. [[CrossRef](#)]
25. He, X.T.; Feng, X.P.; Sun, D.W.; Liu, F.; Bao, Y.D.; He, Y. Rapid and Nondestructive Measurement of Rice Seed Vitality of Different Years Using Near-Infrared Hyperspectral Imaging. *Molecules* **2019**, *24*, 2227. [[CrossRef](#)]
26. Shrestha, S.; Knapic, M.; Zibrat, U.; Deleuran, L.C.; Gislum, R. Single Seed Near-Infrared Hyperspectral Imaging in Determining Tomato (*Solanum lycopersicum* L.) Seed Quality in Association with Multivariate Data Analysis. *Sens. Actuators B Chem.* **2016**, *237*, 1027–1034. [[CrossRef](#)]
27. Baek, I.; Kusumaningrum, D.; Kandpal, L.M.; Lohumi, S.; Mo, C.; Kim, M.S.; Cho, B.K. Rapid Measurement of Soybean Seed Viability Using Kernel-Based Multispectral Image Analysis. *Sensors* **2019**, *19*, 271. [[CrossRef](#)]
28. Agelet, L.E.; Ellis, D.D.; Duvick, S.; Goggi, A.S.; Hurburgh, C.R.; Gardner, C.A. Feasibility of Near Infrared Spectroscopy for Analyzing Corn Kernel Damage and Viability of Soybean and Corn Kernels. *J. Cereal Sci.* **2012**, *55*, 160–165. [[CrossRef](#)]
29. Kandpal, L.M.; Lohumi, S.; Kim, M.S.; Kang, J.S.; Cho, B.K. Near-Infrared Hyperspectral Imaging System Coupled with Multivariate Methods to Predict Viability and Vigor in Muskmelon Seeds. *Sens. Actuators B Chem.* **2016**, *229*, 534–544. [[CrossRef](#)]
30. Shetty, N.; Min, T.G.; Gislum, R.; Olesen, M.H.; Boelt, B. Optimal Sample Size for Predicting Viability of Cabbage and Radish Seeds Based on Near Infrared Spectra of Single Seeds. *J. Near Infrared Spectrosc.* **2011**, *19*, 451–461. [[CrossRef](#)]
31. Kurtulmus, F.; Alibas, I.; Kavdir, I. Classification of Pepper Seeds Using Machine Vision Based on Neural Network. *Int. J. Agric. Biol. Eng.* **2016**, *9*, 51–62.
32. Huang, K.Y.; Chien, M.C. A Novel Method of Identifying Paddy Seed Varieties. *Sensors* **2017**, *17*, 809. [[CrossRef](#)] [[PubMed](#)]
33. Chaugule, A.A.; Mali, S.N. Identification of Paddy Varieties Based on Novel Seed Angle Features. *Comput. Electron. Agric.* **2016**, *123*, 415–422. [[CrossRef](#)]
34. Ran, H.; Cui, Y.J.; Jin, Z.X.; Yan, Y.L.; An, D. Identification of Maize Seed Purity Based on Spectral Images of a Small Amount of Near Infrared Bands. *Spectrosc. Spect. Anal.* **2017**, *37*, 2743–2750.
35. Jordan, M.I.; Mitchell, T.M. Machine learning: Trends, Perspectives, and Prospects. *Science* **2015**, *349*, 255–260. [[CrossRef](#)]
36. Barbedo, J.G.A. Plant Disease Identification from Individual Lesions and Spots Using Deep Learning. *Biosys. Eng.* **2019**, *180*, 96–107. [[CrossRef](#)]
37. Barbedo, J.G.A. Impact of Dataset Size and Variety on the Effectiveness of Deep Learning and Transfer Learning for Plant Disease Classification. *Comput. Electron. Agric.* **2018**, *153*, 46–53. [[CrossRef](#)]
38. DeChant, C.; Wiesner-Hanks, T.; Chen, S.Y.; Stewart, E.L.; Yosinski, J.; Gore, M.A.; Nelson, R.J.; Lipson, H. Automated Identification of Northern Leaf Blight-Infected Maize Plants from Field Imagery Using Deep Learning. *Phytopathology* **2017**, *107*, 1426–1432. [[CrossRef](#)]
39. Shen, R.P.; Huang, A.Q.; Li, B.L.; Guo, J. Construction of a Drought Monitoring Model Using Deep Learning Based on Multi-Source Remote Sensing Data. *Int. J. Appl. Earth Obs.* **2019**, *79*, 48–57. [[CrossRef](#)]
40. Jin, B.X.; Ye, P.; Zhang, X.Y.; Song, W.W.; Li, S.H. Object-Oriented Method Combined with Deep Convolutional Neural Networks for Land-Use-Type Classification of Remote Sensing Images. *J. Indian Soc. Remote Sens.* **2019**, *47*, 951–965. [[CrossRef](#)]
41. Rasti, P.; Ahmad, A.; Samiei, S.; Belin, E.; Rousseau, D. Supervised Image Classification by Scattering Transform with Application to Weed Detection in Culture Crops of High Density. *Remote. Sens.* **2019**, *11*, 249. [[CrossRef](#)]
42. Zhu, S.S.; Zhou, L.; Gao, P.; Bao, Y.D.; He, Y.; Feng, L. Near-Infrared Hyperspectral Imaging Combined with Deep Learning to Identify Cotton Seed Varieties. *Molecules* **2019**, *24*, 3268. [[CrossRef](#)]
43. Wu, N.; Zhang, Y.; Na, R.S.; Mi, C.X.; Zhu, S.S.; He, Y.; Zhang, C. Variety Identification of Oat Seeds Using Hyperspectral Imaging: Investigating the Representation Ability of Deep Convolutional Neural Network. *RSC Adv.* **2019**, *9*, 12635–12644. [[CrossRef](#)]
44. Qiu, Z.J.; Chen, J.; Zhao, Y.Y.; Zhu, S.S.; He, Y.; Zhang, C. Variety Identification of Single Rice Seed Using Hyperspectral Imaging Combined with Convolutional Neural Network. *Appl. Sci.* **2018**, *8*, 212. [[CrossRef](#)]

45. Lu, S.Y.; Lu, Z.H.; Zhang, Y.D. Pathological Brain Detection Based on Alexnet and Transfer Learning. *J. Comput. Sci.* **2019**, *30*, 41–47. [[CrossRef](#)]
46. Suh, H.K.; Ijsselmuiden, J.; Hofstee, J.W.; van Henten, E.J. Transfer Learning for the Classification of Sugar Beet and Volunteer Potato under Field Conditions. *Biosyst. Eng.* **2018**, *174*, 50–65. [[CrossRef](#)]
47. Marmanis, D.; Datcu, M.; Esch, T.; Stilla, U. Deep Learning Earth Observation Classification Using Imagenet Pretrained Networks. *IEEE Geosci. Remote Sens.* **2016**, *13*, 105–109. [[CrossRef](#)]
48. Pan, S.J.; Yang, Q.A. A Survey on Transfer Learning. *IEEE Trans. Knowl. Data Eng.* **2010**, *22*, 1345–1359. [[CrossRef](#)]
49. Cheng, X.; Zhang, Y.H.; Chen, Y.Q.; Wu, Y.Z.; Yue, Y. Pest Identification via Deep Residual Learning in Complex Background. *Comput. Electron. Agric.* **2017**, *141*, 351–356. [[CrossRef](#)]
50. Boulent, J.; Foucher, S.; Theau, J.; St-Charles, P.L. Convolutional Neural Networks for the Automatic Identification of Plant Diseases. *Front. Plant Sci.* **2019**, *10*, 941. [[CrossRef](#)]
51. Motta, D.; Santos, A.A.B.; Winkler, I.; Machado, B.A.S.; Pereira, D.; Cavalcanti, A.M.; Fonseca, E.O.L.; Kirchner, F.; Badaro, R. Application of Convolutional Neural Networks for Classification of Adult Mosquitoes in the Field. *PLoS ONE* **2019**, *14*, e0210829. [[CrossRef](#)] [[PubMed](#)]
52. Milella, A.; Marani, R.; Petitti, A.; Reina, G. In-field High Throughput Grapevine Phenotyping with a Consumer-Grade Depth Camera. *Comput. Electron. Agric.* **2019**, *156*, 293–306. [[CrossRef](#)]
53. Heravi, E.J.; Aghdam, H.H.; Puig, D. An optimized Convolutional Neural Network with Bottleneck and Spatial Pyramid Pooling Layers for Classification of Foods. *Pattern Recognit. Lett.* **2018**, *105*, 50–58. [[CrossRef](#)]
54. Altuntas, Y.; Comert, Z.; Kocamaz, A.F. Identification of Haploid and Diploid Maize Seeds Using Convolutional Neural Networks and a Transfer Learning Approach. *Comput. Electron. Agric.* **2019**, *163*, 104874. [[CrossRef](#)]
55. Feng, L.; Zhu, S.S.; Zhang, C.; Bao, Y.D.; Gao, P.; He, Y. Variety Identification of Raisins Using Near-Infrared Hyperspectral Imaging. *Molecules* **2018**, *23*, 2907. [[CrossRef](#)]
56. Liu, W.J.; Li, W.J.; Li, H.G.; Qin, H.; Ning, X. Research on the Method of Identifying Maize Haploid Based on kpca and Near Infrared. *Spectrosc. Spect. Anal.* **2017**, *37*, 2024–2027.
57. Yang, S.; Zhu, Q.B.; Huang, M.; Qin, J.W. Hyperspectral Image-Based Variety Discrimination of Maize Seeds by Using A Multi-Model Strategy Coupled with Unsupervised Joint Skewness-Based Wavelength Selection Algorithm. *Food Anal. Methods* **2017**, *10*, 424–433. [[CrossRef](#)]
58. Nie, P.C.; Zhang, J.N.; Feng, X.P.; Yu, C.L.; He, Y. Classification of Hybrid Seeds Using Near-Infrared Hyperspectral Imaging Technology Combined with Deep Learning. *Sens. Actuators B Chem.* **2019**, *296*, UNSP 126630. [[CrossRef](#)]
59. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet Classification with Deep Convolutional Neural Networks. *Commun. ACM* **2017**, *60*, 84–90. [[CrossRef](#)]

Sample Availability: Samples of the compounds are available from the authors.



© 2019 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 (<http://creativecommons.org/licenses/by/4.0/>).