Classification of Dianthus Seed Species with Deep Transfer Learning

Esra Sivari[a], Said Akca[b], Recep Eryigit[c], Bulent Tugrul[c,*]

[a]*Department of Computer Engineering, Cankiri Karatekin University, Cankiri, 18100, Turkey*
[b]*Department of Computer Engineering, Sakarya University of Applied Sciences, Sakarya, 54187, Turkey*
[c]*Department of Computer Engineering, Ankara University, Ankara, 06830, Turkey*

Abstract

Since ancient times, the Dianthus has been an essential ornamental flower, most in-demand in the horticultural and floriculture industry and symbolizing various cultures in different societies. Besides its commercial and cultural importance, Dianthus is of medical importance due to its proven pharmacological effects. The Dianthus, a genus of the Caryophyllaceae family, has species with different characteristics such as color, size and fragrance. Classification of seed species is a time-consuming, costly, and vitally important task for plant science and the seed trade, requiring specialized personnel and equipment. In this study, seeds of three different clove species were classified with their new versions obtained by applying deep transfer learning techniques to VGG16, InceptionV3, MobileNet, ResNet152V2, and DenseNet201 models. The new version of the ResNet152V2 model achieved 99.45% accuracy and 0.9996 AUC. The results illustrate that the proposed method is promising for all seed classification tasks in agriculture and botanical science.

Keywords: Deep learning, Transfer Learning, Seed Classification, Dianthus

* Corresponding author at : Department of Computer Engineering, Ankara University, Ankara, 06830, Turkey

Email addresses: esrasivari@karatekin.edu.tr (Esra Sivari), saidakca@subu.edu.tr (Said Akca), reryigit@eng.ankara.edu.tr (Recep Eryigit), btugrul@eng.ankara.edu.tr (Bulent Tugrul)

1. Introduction

For humanity to continue its future, it is vital to provide solutions to the factors that harm the plant ecosystem, such as the continuous increase in the population rate in the world, irregular climate changes, and global warming, and to protect the plant diversity in the ecosystem. Therefore, analyzing them should better understand various physical aspects of complex, multivariate, and unpredictable plant ecosystems. Accumulating data and facilitating access to data with today's technology have enabled the development of artificial intelligence technologies in botanical science. The search for automatic systems that assist experts in analysis has become widespread. Image analysis is an important research area for plant cataloging and conservation in botany. Therefore, intelligent image analysis techniques such as image determination, classification, and anomaly detection are frequently used in various plant analysis applications. Identifying and classifying seed and plant species [1, 2, 3], detecting plant diseases [4, 5], quality analysis [6, 7] and monitoring germination or growth rate [8, 9] are the fundamental problems in which artificial intelligence and image analysis techniques are applied.

The problem of classification of seed species in plant science is essential for increasing productivity and diversity by using appropriate types of seeds. Classification seed species is a time-consuming and costly task requiring genetic examination with expert personnel and equipment. Among the classification methods that do not require genetic analysis, image analysis methods are applied with techniques such as hyperspectral imaging [10] and infrared imaging [11], which require expensive equipment. The classification made with these methods may vary according to the experience and knowledge of the expert who performs the analysis. Thus, there is always the possibility of misidentification of the seed species. The correct classification of seed species is crucial for the seed trade and the scientific sense. Since some species are more valuable in the seed trade, intentional or unintentional misclassification can be made by experts or non-experts at any step of the seed production chain.

In the last decade, the emergence of deep learning, a sub-branch of machine learning, has contributed significantly to the development of artificial intelligence and its application areas. Especially with convolutional neural networks (CNN), which provide significant advantages compared to traditional image processing techniques, successful results were obtained in image analysis problems. Deep learning techniques have been proposed when the literature is examined to classify many different seed species.

Xu et al. [12] proposed the P-ResNet model to classify five different species of maize seeds. In the proposed method, the ResNet model was fine-tuned, and the performance comparison of the P-ResNet model and other pre-trained CNN models was made. The experimental results obtained on 8080 images showed that the P-ResNet model reached 99.70% validation accuracy. Taheri-Garavand et al. [13] proposed the modified VGG16 model to classify four different chickpea species from 400 chickpea seed images. The study performed 5-fold cross-validation and test stages, with 88.41% test accuracy. Erygit and Tugrul [14] classified six grass seed species using 15 different convolutional neural network models and compared the models' performances. In the study using 8654 seed images, the DenseNet201 model reached the highest test accuracy with 99.42%. Sabanci et al. [15] proposed a hybrid model named CNN-SVM-Cubic to classify four species of pepper seeds using 767 images. The proposed method extracted features from pre-trained CNNs were selectively combined and classified with a support vector machine (SVM). As a result of classification, 99.02% accuracy was obtained. Loddo et al. [16] developed the SeedNet model to classify plant seeds. The SeedNet model was trained on two different public datasets containing 29 species of plant seed images, and the accuracy was 95.65% in the first dataset and 97.47% in the second dataset. Gulzar et al. [17] updated the VGG16 model to classify 14 species of plant seeds. In the study using 2733 images, 99% accuracy was obtained. Along with these studies carried out with digital images of seeds, some studies include applications of deep learning and identification and classification of maize [18], cotton [19], and hybrid [20] seeds from hyperspectral images.

This study it is aimed to determine and classify the Dianthus seeds of three different Dianthus species automatically, objectively, quickly, and with high accuracy by using a dataset consisting of digital images. Dianthus, one of the most popular flowers globally, is a genus of various species in the family Caryophyllaceae [21]. Dianthus species have different characteristics, such as colour, size and fragrance. Determining the species from Dianthus seeds is essential for plant science and commercial importance in the horticulture and floriculture industry. At the same time, Dianthus is among the medicinal plants and has been proven to have the high pharmacological potential [22].

Based on the literature study and our research, as far as we know, there is no publicly available dataset containing the species of Dianthus seeds. In addition, no study has been carried out to classify the species of Dianthus seeds using any artificial intelligence technique. One contribution of this study to the literature is introducing a new dataset containing three different Dianthus seeds to the literature. Another contribution is that it is the first study in the literature in which the species of Dianthus seeds is determined and classified by artificial intelligence. In the proposed method for determining and classifying the species of Dianthus seeds, new models were created by applying deep transfer learning techniques to pre-trained CNN models VGG16, DenseNet201, ResNet152V2, InceptionV3, InceptionResNetV2 and MobileNet, and trained with 5-fold cross-validation. Another contribution of this study to the literature is the promising, highly successful results from these fine-tuned models for other classification problems.

In the Section 2, information about pre-trained CNN models, transfer learning and the background of the proposed method are provided. In the Section 3, the stages of the proposed method, the data set, the transfer learning techniques applied to the models, and the training and validation stages are discussed in details. In the Section 4, the test phase of the study and the results are given by comparing them. The Section 5 discusses the study's place, advantages, and limitations in the literature. Finally, Section 6 overviews the final results and the future place of the study.

2. Deep transfer learning with pretrained CNNs

A CNN [23] is a deep-layered, feed-forward artificial neural network that learns its properties by using convolution in at least one of the layers that can receive images or video as input. The architecture of CNN is presented in Figure 1. A CNN architecture consists of layers of input, convolutional, nonlinear, pooling, flattening, fully connected, and output. The convolution layer's task is to obtain an attribute map using filters (kernel) from the raw image pixels from the input layer. Nonlinear layers determine which neuron becomes active by adding an activation function at the end of the convolution layer or output layer. The Pooling layer has the task of reducing the input image, and the flatten layer has the task of converting 3D tensors to 1D tensors for a fully connected layer. The fully connected layer is the final layer in the classification task.

Deep transfer learning [24] uses a pre-trained CNN model in conjunction with a new dataset to solve a new problem using a large dataset for the image classification problem. The pre-trained model can be used as it is

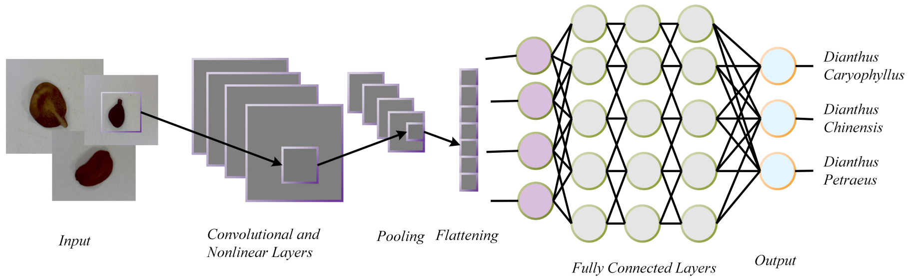


Figure 1: Architecture of CNN

or by being customized to a specific problem (fine-tuning). The first step in the customization process is to add new classifier layers or algorithms to the pre-trained model following the new classification task. Another step is determining how much weight in layers is frozen for property transfer. Depending on the new dataset's compatibility and the previously trained dataset to the model, the convolution base can be fully trained, several layers can be frozen from the top, or the convolution base can be completely frozen.

3. Material and method

3.1. Overview of the proposed method

New models have been obtained by performing fine-tune to VGG16 [25], DenseNet201 [26], ResNet152V2 [27], InceptionV3 [28], InceptionResNetV2 [29] and MobileNet [30] networks from pre-trained CNN models to automatically determining and classifying the species of Dianthus seeds. After the image preprocessing step, these models were trained and classified on the Dianthus seeds dataset. The classification process was carried out in training, 5-fold cross-validation, and testing. During the training and validation stages, real-time data augmentation was applied to the training set to prevent the overfitting of the models. The stages of the proposed method are given in Figure 2.

3.2. Dataset

3.2.1. Image acquisition

The seed images of three species of Dianthus used in this study were obtained from the General Directorate of Seed Registration of the Ministry of Agriculture and Forestry of Turkey. The Dianthus seed dataset contains

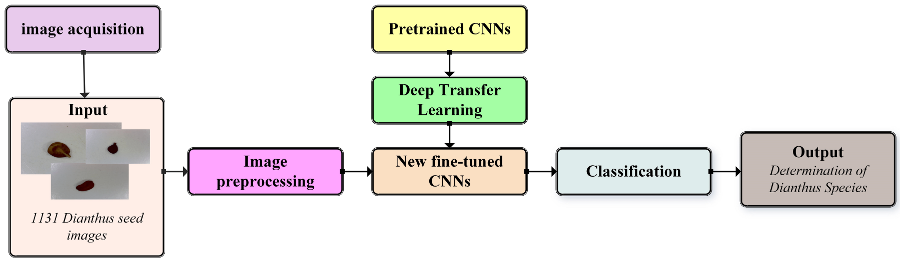


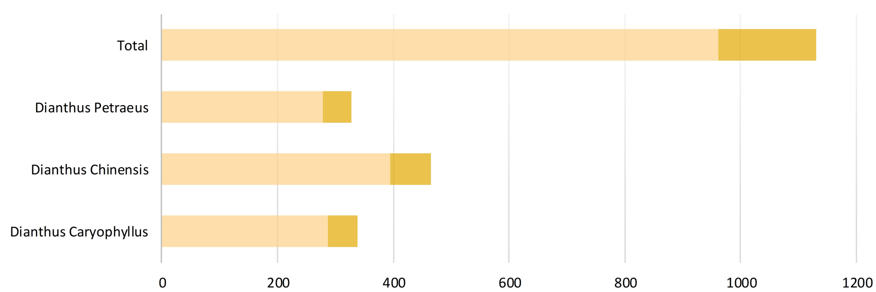
Figure 2: Overview of the proposed method

1131 images, of which 339 images are Dianthus caryophyllus, 465 images are Dianthus chinensis, and 327 images are Dianthus petraeus. Seed images were acquired at a resolution of 100 pixels per millimeter at 1600 $\times$ 1200 pixels in daylight using a digital microscope (Celestron). Image samples of 3 classes from the dataset are given in Figure 3. The image sets used in the current study can be accessed from Dianthus dataset(`https://www.kaggle.com/datasets/btugrul/dianthus`).

Figure 3: Samples from the dataset

3.2.2. Data split

The dataset was divided into training, validation, and test sets. Due to the small number of samples in the data set, 70-15-15% ratio was chosen when dividing the data set. 15% of 1131 images (170 images) were used for the test set. The number of training and test images in each class is given in Figure 4. The sample numbers in the validation set are not given in Figure 4 because of the application of 5-fold cross-validation. Thus, five different validation sets were created during the training phase. The validation set was created by dividing the remaining 961 training set images by 0.177 in each cross-validation fold. The reason for choosing this ratio is to ensure that the validation set has the same number of samples in each fold as the test set. Test and validation sets were created using the Stratified Shuffle Split strategy. In the Stratified Shuffle Split strategy, stratified randomized folds are created by selecting 0.177 random images from each class in each cross-validation fold. The Stratified Shuffle Split strategy was chosen to preserve the sample percentage for each class in each fold and the relatively small number of samples in the dataset.



	Dianthus Caryophyllus	Dianthus Chinensis	Dianthus Petraeus	Total
Train	288	395	278	961
Test	51	70	49	170

Figure 4: Distribution of the dataset

3.2.3. Image preprocessing

Image resizing is an essential preprocessing step, as deep learning models take the input at a fixed size. All images were resized to 150×150 dimensions with bilinear interpolation following the architectures of pre-trained CNN models. All images were set to RGB (red, green, blue) color mode, resulting in 3-channel input size. Image pixels in the 0-255 range were rescaled for normalization to the 0-1 range, which is standard for neural networks.

3.3. Experiments

3.3.1. Data augmentation

Working with large datasets is vital for the performance of the deep learning model. An overfitting problem may occur in a model trained with a small amount of data. Creating new and synthetic data by changing the image with various techniques is a way to avoid the overfitting problem [31]. In the application, real-time data augmentation techniques were applied to the images used only for the training set. The data augmentation techniques and parameters are presented in Table 1.

Table 1: Data augmentation parameters

Data Augmentation Technique	Range
Rotation	30
Width shift	0.2
Height shift	0.2
Shear	0.3
Zoom	0.1
Fill mode	nearest

3.3.2. Details of implementation

In the method used for determining and classifying Dianthus seeds, images are trained with 150 $\times$ 150 $\times$ 3 input sizes, the new VGG16, DenseNet201, ResNet152V2, InceptionV3, InceptionResNetV2 and MobileNet models, which were customized using the same hyperparameters with the same fully connected layers. The final layers of original networks that perform classification were removed and feature extracting convolution bases were used. After the convolution bases, Global Average Pooling layer, Batch Normalization layer, 0.5 Dropout layer, two hidden layers with 1024 and 512 neurons, Batch Normalization layer, 0.5 Dropout layer, and the output layer with three neurons were added, respectively. ReLu [32] is the activation function in the hidden layers, and softmax [33] was used for multi-classification problems in the output layer. In Figure 5, the updated architecture of the models is presented.

Fully connected layers are prone to overfitting. Therefore, classifier layers are designed to prevent overfitting. The Global Average Pooling layer summarizes the feature map from the convolution bases and vectorizes it for fully connected layers. Unlike the flatten layer, averaging pooling in the vectorization process is an advantage in avoiding overfitting. The Global Average Pooling layer has been added due to its advantages, such as the absence of a hyperparameter to optimize, preventing overfitting, summarizing the spatial translations of the input, and retaining the convolutional structure-specific features. Batch Normalization is a normalization process done in mini-batches between layers. Adding a Batch Normalization layer allows for higher learning rates, facilitating learning and speeding up training. The dropout layer ensures that specific randomly selected neurons are ignored at a given rate, preventing overfitting and increasing performance.

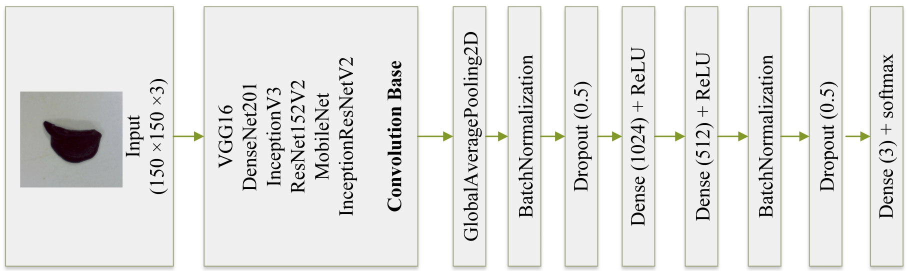


Figure 5: Architectures of fine-tuned models

The ReLU activation function takes the value of 0 on the negative axis and activates some neurons, increasing the training rate. Therefore, it is used in hidden layers and convolutional layers. The softmax activation function in the output layer generates values of 0-1 by interpreting the output probabilistic in multi-classification tasks. In the application, the models' weights were started from the weights of the imagenet. Since the Dianthus seeds dataset is not similar to the imagenet set, the dataset was trained from scratch, and none of the layers was frozen. Transfer learning has only been used to build model architectures and start weights.

In all models, the batch size value is 8, the learning rate value is 2e-3, and the epoch value is 100. Adam [34] optimization algorithm was used for weight update. Categorical cross-entropy was used as a loss function, which calculates the current heap loss for single-label multi-classification problems. The learning rate was controlled with 'Keras ReduceLROnPlateau Callbacks' [35] to avoid overfitting, which monitors the 'validation loss' value during training. With this function, if the validation loss metric does not improve by five epochs during the training and validation phase, the learning rate is multiplied by the factor value of 0.1, and the new learning rate is obtained. The number of epochs for the training phase was set to 100. During training, the training and validation phase was terminated if the 'training accuracy' value did not improve above five epochs by 'Keras EarlyStopping Callbacks' [35]. The values of the network hyperparameters are given in Table 2.

The learning curves show the variation of each fold's training and validation accuracies according to the epoch progression. It can be obtained from a learning curve that the networks undergo overfitting, underfitting, or

Table 2: The values of the network hyperparameters

Hyperparameter	Value
Batch size	8
Learning rate	2.00E-03
Epoch	100
Loss function	categorical_crossentropy
Optimization algorithm	Adam

adapting as they should be during the training phase. At the beginning of the training phase, the network sees a few training samples and the entire validation set, so it generalizes by rote. There is overfitting if the training accuracy is consistently below the validation accuracy after a specific epoch value. The gap between training accuracy and validation accuracy is called the generalization gap. It is typical when the number of training samples is greater than the number of validation samples. There is underfitting if the generalization gap is prominent in a learning curve. In the learning curves given in Figure 6, the fit of each network is typical, and the training phase has been efficient in every bend. When the learning curve of MobileNet is examined, it is observed that the generalization gap is small, but the validation accuracy is above the training accuracy. Although this may suggest overfitting, training data for MobileNet needs to be increased. In addition, it is observed from the learning curves that all models' training and validation accuracy is over 90%.

4. Results

4.1. Performance metrics

The confusion matrix is a performance measure that numerically compares the actual labels with the predicted labels by the model. The meaning of the values given by a confusion matrix is given in Figure 7. The P values represent the number of predictions, green predictions represent correct, and red predictions represent incorrect in Figure 7. For example, P_{Ca-Ca} represents the number of predictions for which Dianthus caryophyllus images were correctly classified. P_{Ch-Ca} represents the number of predictions where Dianthus caryophyllus images were misclassified as Dianthus chinensis. The confusion matrix gives True Positive (TP), False Negative (FN), True Negative (TN), and False Positive (FP) values for each class used to calculate performance measures.

The metrics and mathematical formulas used in the evaluation of the models used in the application are given in Table 3. The accuracy metric is the ratio of the correct predictions to all predictions. The use of accuracy metric alone in datasets where the number of instances per class is unevenly distributed is not a suitable evaluation method. Precision is a positive predictor and is the ratio of true positives to all positives. Precision gives how many of the correct predictions from all classes are correct. Recall measures how accurately the truly correct ones were predicted. The F_1 score is the harmonic mean of the precision and recall metrics and can take values in the $[0, 1]$ range. A Categorical cross-entropy determines only one of many possible categories by measuring the difference between two probability distributions. ROC is a curve that plots True Positive Rate (TPR) versus False Positive Rate (FPR) values against different classification thresholds. The AUC value is the area under the ROC curve and measures the aggregated performance of all possible classification thresholds regardless of which classification threshold is selected.

Table 3: Performance metrics

Performance Metrics	Formulas
Accuracy	$\frac{TP+TN}{TP+TN+FP+FN}$
Precision	$\frac{TP}{TP+FP}$
Recall	$\frac{TP}{TP+FN}$
F_1 Score	$\frac{2*TP}{2*TP+FP+FN}$
Categorical cross entropy	$-\sum_{i=1}^{i} q(y_i)log(p(y_i))$
FPR	$\frac{TP}{TP+FP}$

4.2. Test results

In the testing phase after the training and validation phase, 170 samples with the same preprocessing properties that the models had never seen were used. Confusion matrices of fine-tuned pre-trained networks used in the application are given in Figure 8. Of the 51 test images belonging to the Dianthus caryophyllus class, the VGG16 and InceptionResNetV2 models made the highest predictions by making 51 correct. For the Dianthus caryophyllus class, the MobileNet and InceptionV3 models have the lowest correct prediction, with 49 correct and two incorrect predictions. ResNet152V2 labelled all 70 Dianthus chinensis and 49 Dianthus petraeus test images with correct predictions. The VGG16 and InceptionResNetV2 models have the lowest correct predictions for the Dianthus chinensis class, with 65 correct and five incorrect predictions. For the Dianthus petraeus class, the lowest predictions were made by the InceptionResNetV2 model, with 41 correct and eight incorrect predictions.

Aggregated metrics provide a high-level view of the overall performance of models. In single-label multi-classification problems, micro average batch measurement is the same as accuracy and is essential in problems where a balanced dataset is used. In cases where an unbalanced dataset and all classes are equally important, macro average aggregate measurement should be considered. In cases where the class with a large number of samples is essential, weighted average aggregate measurement should be considered. The test results, which include the values of the macro averaged performance metrics of the models, are presented in Table 4. According to the results, the ResNet152V2, with values 99.45%, 0.9958, 0.9939, 0.9947, and 0.9996, respectively, in the accuracy, precision, recall, F_1 Score, and AUC metrics, is the model that makes the best classification. In addition, the ResNet152V2 has the lowest loss with 0.0373. The DenseNet201 is the second model to achieve the best classification with 0.9705 accuracy after the ResNet152V2 network. Although InceptionV3 and VGG16 models give similar results in loss, accuracy and precision metrics, VGG16 is more successful in other metrics. Although the MobileNet and InceptionResNetV2 models have lower results than other models, the success of all models above 92% proves the reliability of the proposed method. The relatively low success of MobileNet and InceptionNetV2 compared to other models may be due to the selected hyperparameters and the scarcity of data, as well as insufficient compliance with the data set compared to other models.

Since this study uses a dataset with an unbalanced number of samples

Table 4: Test results of pretrained CNNs

Model	Loss	Accuracy	Precision	Recall	F_1 Score	AUC-ROC
VGG16	0.1370	95.88	0.9589	0.9626	0.9601	0.9967
InceptionV3	0.1787	95.29	0.9524	0.9522	0.9519	0.9891
DenseNet201	0.0983	97.05	0.9705	0.9703	0.9700	0.9975
ResNet152V2	**0.0373**	**99.45**	**0.9957**	**0.9939**	**0.9947**	**0.9996**
InceptionResNetV2	0.2514	92.35	0.9307	0.9218	0.9218	0.9889
MobileNet	0.1657	93.52	0.9364	0.9339	0.9339	0.9942

per class, it will not be sufficient to evaluate the accuracy and loss metrics alone. Additionally, it does not matter which of the precision or recall values should be higher for this operation. The important thing is that both metrics are high at relative values. Considering these factors, the F_1 score and AUC, which are the harmonic mean of precision and recall values, are essential for performance evaluation. In Figure 9, the comparison of the performances of the models according to the metrics is visualized. When Figure 9 is examined, it is observed that each model's precision and recall values are almost the same. Moreover, the AUC values of all models are pretty high and close to each other. The highest performance in this study belongs to ResNet152V2 with 0.9996 AUC value. ResNet152V2 exhibited almost perfect discrimination with only one wrong guess.

5. Discussion

When the overview of seed studies in Table 5 is examined, transfer learning techniques are the most used method for seed classification. No study in the literature detects and classifies Dianthus seeds. The size of the dataset, resolution, noise, and variety are significant in the classification task. Therefore, comparing this study with other studies in terms of performance is incorrect. We can refer to the literature comparison to measure the performance of this study in classifying seeds of any plant. When the accuracy values of the studies presented in Table 5 are examined, our study showed superior performance compared to other studies except the study conducted by Xu et al. [12]. The accuracy of the study conducted by Xu et al. [12] reached 99.70% because they used a large dataset containing 8080 samples. Although our study was carried out on a relatively small dataset with 1131 samples compared to other studies in the literature, it reached 99.45% accuracy. Considering these factors, the proposed transfer learning techniques and hyperparameters were more effective than those suggested in the literature. The proposed method can identify and detect any seed. Another critical strategy proposed in this study is Stratified Shuffle Split 5-fold cross-validation compared to the literature. It is more reliable than K-fold and holdout methods because this strategy creates validation and test sets by selecting samples from each class at a given rate. In the sets created with K-fold and holdout applications, biased performance measurements are obtained with the uneven sample distribution and the increased importance given to a class. In some studies in the literature, k-fold cross-validation results are presented, and not performing the test phase is a false assessment that is optimistic about the results. The validation phase is applied to measure the efficiency of the model hyperparameters and to design an efficient model. The implementation of the testing phase is necessary to evaluate the final results and the model's overall performance.

In this study, the convolution layers of pre-trained CNNs were used without any degradation. Classifier layers have been redesigned. Considering that the task of the convolution layer is to extract features, adding the same classification layer to all pre-trained CNNs can lead to the determination of the best feature extractor model. However, since feature extraction and learning actions occur simultaneously, it is best to consider all models as a whole. Determining the best feature extractor for the Dianthus seeds dataset may contribute to future studies. The best performing ResNet152V2 network can be used as a feature extractor, and the performances of hybrid combinations with different machine learning classifiers can be evaluated. On the other hand, the classifier layer we designed by taking overfitting measures significantly affects the models' performance. In the classification layer, learning takes place by extracting class-specific features. The success of over 92% of all models proves the success of the classification layer.

Data-sourced constraints standardized in deep learning applications are also valid for this study. The fact that the Dianthus seeds dataset contains 1131 samples for the three species is a factor that complicates the task. Although this difficulty can be overcome with data augmentation, increasing the number of data is the most crucial factor that will increase the performance of the models. However, another limitation is that our dataset includes only three species, although about 340 species of Dianthus are in nature. The Dianthus seeds dataset and seed classification models, which were brought to the literature with this study, are open to development by other researchers.

Table 5: Overview seed classification studies in the literature

Author(Year)	Type	Classifier	Accuracy
Xu et al.(2022)	Maize	P-ResNet	99.70%
Taheri-Garavand et al. (2021)	Chickpea	Modified VGG16	88.41%
Eryigit and Tugrul (2021)	Grass	DenseNet201	99.42%
Sabanci et al. (2021)	Pepper	CNN-SVM-Cubic	99.02%
Loddo et al. (2021)	Plant	SeedNet	95.65%
Gulzar et al. (2020)	Plant	Modified VGG16	99.00%
This study	Dianthus	Fine-tuned ResNet152V2	99.45%

6. Conclusion

In this study, six different fine-tuned pre-trained CNN models are proposed for the classification of three different Dianthus seed species on a new dataset. The proposed method obtains current models by adding new, fully connected layers to the previously trained models. Adding new fully connected layers to pre-trained models improves performance and enables higher performances using fewer data. As a result of the experiments with the proposed method, we have achieved 99.45% accuracy and 0.9996 AUC with the fine-tuned pre-trained ResNet152V2 model.

The results confirm that the proposed method can be used in seed classification, which is a complicated and costly task for agriculture and plant science. Developing intelligent and automatic applications to increase plant science productivity is crucial. Using the proposed method for seed classification is helpful to experts, reducing time and cost. In addition, it offers the advantages of objective and reliable seed identification. The proposed method can be integrated into mobile applications, seed sorters, or smart farming equipment. Considering the commercial importance of Dianthus in the horticultural and floriculture industry, such practices can prevent the deliberate misclassification of seeds in the seed trade.

References

[1] D. Singh, Y. S. Taspinar, R. Kursun, I. Cinar, M. Koklu, I. A. Ozkan, H.-N. Lee, Classification and analysis of pistachio species with pre-trained deep learning models, Electronics 11 (7) (2022) 981.

[2] M. Seeland, M. Rzanny, N. Alaqraa, J. Wäldchen, P. Mäder, Plant species classification using flower images—a comparative study of local feature representations, PloS One 12 (2) (2017) e0170629.

[3] M. Dyrmann, H. Karstoft, H. S. Midtiby, Plant species classification using deep convolutional neural network, Biosystems Engineering 151 (2016) 72–80.

[4] S. Ramesh, R. Hebbar, M. Niveditha, R. Pooja, N. Shashank, P. Vinod, et al., Plant disease detection using machine learning, in: 2018 International conference on design innovations for 3Cs compute communicate control (ICDI3C), IEEE, 2018, pp. 41–45.

[5] K. P. Ferentinos, Deep learning models for plant disease detection and diagnosis, Computers and Electronics in Agriculture 145 (2018) 311–318.

[6] H. O. Velesaca, P. L. Suárez, R. Mira, A. D. Sappa, Computer vision based food grain classification: A comprehensive survey, Computers and Electronics in Agriculture 187 (2021) 106287.

[7] S. Chen, J. Xiong, W. Guo, R. Bu, Z. Zheng, Y. Chen, Z. Yang, R. Lin, Colored rice quality inspection system using machine vision, Journal of Cereal Science 88 (2019) 87–95.

[8] G. Bernotas, L. C. Scorza, M. F. Hansen, I. J. Hales, K. J. Halliday, L. N. Smith, M. L. Smith, A. J. McCormick, A photometric stereo-based 3d imaging system using computer vision and deep learning for tracking plant growth, GigaScience 8 (5) (2019) giz056.

[9] J. Colmer, C. M. O'Neill, R. Wells, A. Bostrom, D. Reynolds, D. Websdale, G. Shiralagi, W. Lu, Q. Lou, T. Le Cornu, et al., Seedgerm: a cost-effective phenotyping platform for automated seed imaging and machine-learning based phenotypic analysis of crop seed germination, New Phytologist 228 (2) (2020) 778–793.

[10] F. Arias, M. Zambrano, K. Broce, C. Medina, H. Pacheco, Y. Nunez, Hyperspectral imaging for rice cultivation: Applications, methods and challenges, AIMS Agriculture and Food 6 (1) (2021) 273–307.

[11] A. K. Mortensen, R. Gislum, J. R. Jørgensen, B. Boelt, The use of multispectral imaging and single seed and bulk near-infrared spectroscopy

to characterize seed covering structures: Methods and applications in seed testing and research, Agriculture 11 (4) (2021) 301.

[12] P. Xu, Q. Tan, Y. Zhang, X. Zha, S. Yang, R. Yang, Research on maize seed classification and recognition based on machine vision and deep learning, Agriculture 12 (2) (2022) 232.

[13] A. Taheri-Garavand, A. Nasiri, D. Fanourakis, S. Fatahi, M. Omid, N. Nikoloudakis, Automated in situ seed variety identification via deep learning: a case study in chickpea, Plants 10 (7) (2021) 1406.

[14] R. Eryigit, B. Tugrul, Performance of various deep-learning networks in the seed classification problem, Symmetry 13 (10) (2021) 1892.

[15] K. Sabanci, M. F. Aslan, E. Ropelewska, M. F. Unlersen, A convolutional neural network-based comparative study for pepper seed classification: Analysis of selected deep features with support vector machine, Journal of Food Process Engineering (2021) e13955.

[16] A. Loddo, M. Loddo, C. Di Ruberto, A novel deep learning based approach for seed image classification and retrieval, Computers and Electronics in Agriculture 187 (2021) 106269.

[17] Y. Gulzar, Y. Hamid, A. B. Soomro, A. A. Alwan, L. Journaux, A convolution neural network-based seed classification system, Symmetry 12 (12) (2020) 2018.

[18] C. Xia, S. Yang, M. Huang, Q. Zhu, Y. Guo, J. Qin, Maize seed classification using hyperspectral image coupled with multi-linear discriminant analysis, Infrared Physics & Technology 103 (2019) 103077.

[19] S. Zhu, L. Zhou, P. Gao, Y. Bao, Y. He, L. Feng, Near-infrared hyperspectral imaging combined with deep learning to identify cotton seed varieties, Molecules 24 (18) (2019) 3268.

[20] P. Nie, J. Zhang, X. Feng, C. Yu, Y. He, Classification of hybrid seeds using near-infrared hyperspectral imaging technology combined with deep learning, Sensors and Actuators B: Chemical 296 (2019) 126630.

[21] K. Tanase, C. Nishitani, H. Hirakawa, S. Isobe, S. Tabata, A. Ohmiya, T. Onozaki, Transcriptome analysis of carnation (dianthus caryophyllus l.) based on next-generation sequencing technology, BMC Genomics 13 (1) (2012) 1–11.

[22] S. Chandra, D. Rawat, Medicinal plants of the family caryophyllaceae: a review of ethno-medicinal uses and pharmacological properties, Integrative Medicine Research 4 (3) (2015) 123–131.

[23] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436–444.

[24] M. Oquab, L. Bottou, I. Laptev, J. Sivic, Learning and transferring mid-level image representations using convolutional neural networks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2014, pp. 1717–1724.

[25] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv preprint arXiv:1409.1556 (2014).

[26] G. Huang, Z. Liu, L. Van Der Maaten, K. Q. Weinberger, Densely connected convolutional networks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 4700–4708.

[27] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2016, pp. 770–778.

[28] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, A. Rabinovich, Going deeper with convolutions, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2015, pp. 1–9.

[29] C. Szegedy, S. Ioffe, V. Vanhoucke, A. A. Alemi, Inception-v4, inception-resnet and the impact of residual connections on learning, in: Thirty-first AAAI conference on artificial intelligence, 2017, pp. 4278–4284.

[30] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, L.-C. Chen, Mobilenetv2: Inverted residuals and linear bottlenecks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2018, pp. 4510–4520.

[31] C. Shorten, T. M. Khoshgoftaar, A survey on image data augmentation for deep learning, Journal of Big Data 6 (1) (2019) 1–48.

[32] A. F. Agarap, Deep learning using rectified linear units (relu), arXiv preprint arXiv:1803.08375 (2018).

[33] S. Sharma, S. Sharma, A. Athaiya, Activation functions in neural networks, International Journal of Engineering Applied Sciences and Technology 4 (12) (2020) 310–316.

[34] D. P. Kingma, J. Ba, Adam: A method for stochastic optimization, arXiv preprint arXiv:1412.6980 (2014).

[35] F. Chollet, Deep learning with Python, Simon and Schuster, 2021.

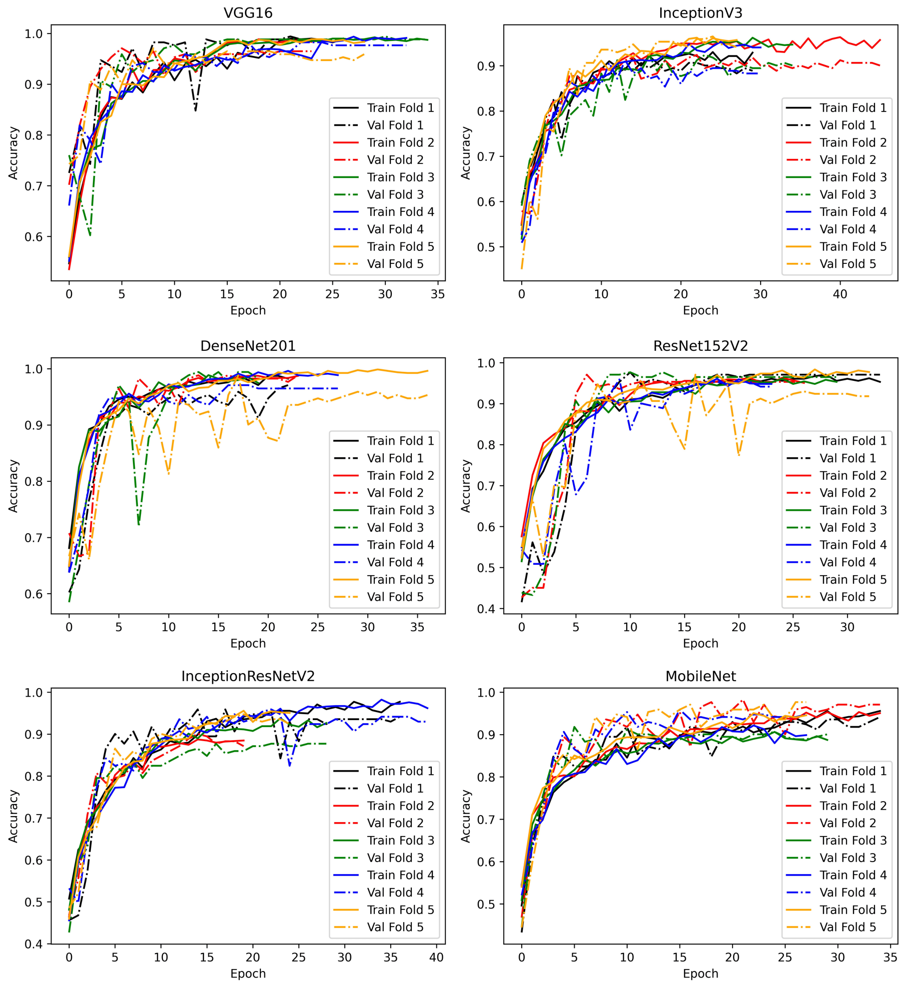


Figure 6: Learning curves of pretrained CNNs

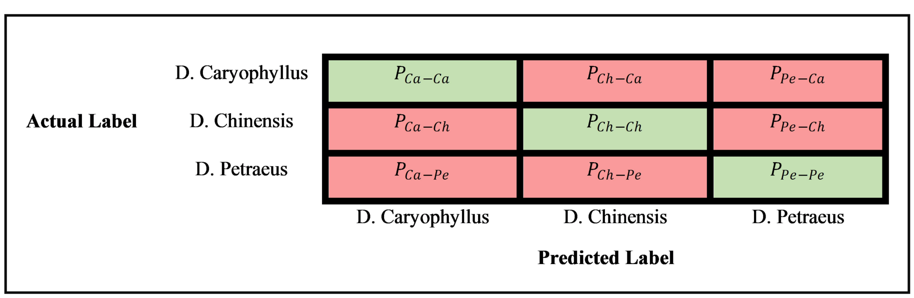


Figure 7: Confusion matrix

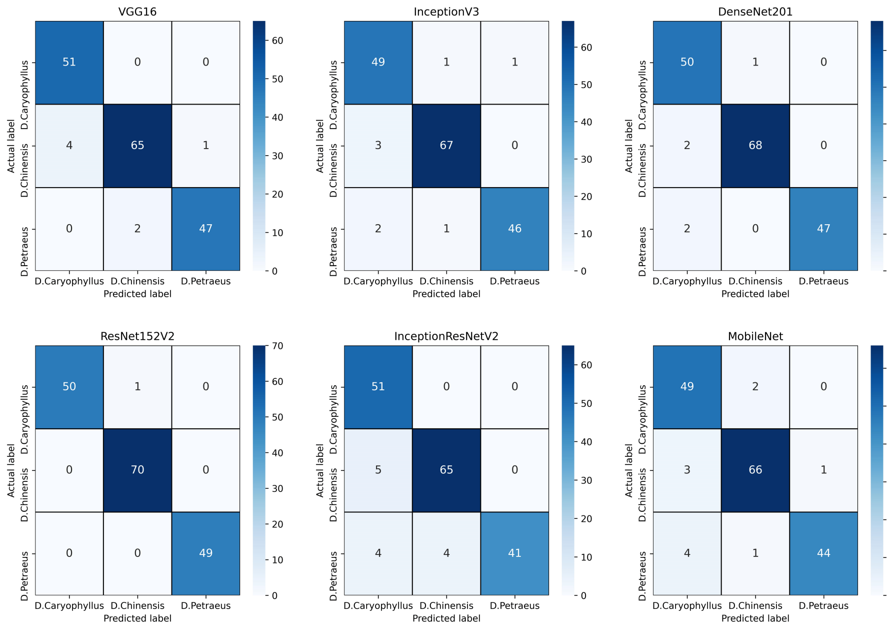


Figure 8: Confusion matrices of pretrained CNNs

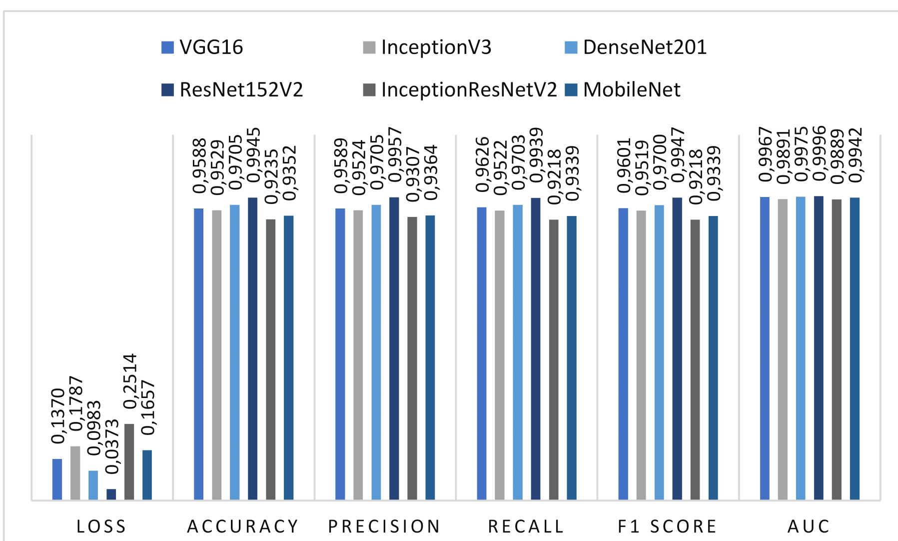


Figure 9: Comparison of performance of models according to metrics