Final Project

Machine Learning

A Deep Learning Approach to Pistachio Species Classification: Comparative Analysis of CNN and Transfer Learning Models (VGG16 & VGG19)

By

Somkiad Khenwiset

675380013-9

Present

Dr. Arfat Ahmad Khan

SC348702 Machine Learning

Semester 1/2567

College of Computing

Khon Kaen University

Abstract

This study investigates the use of deep learning techniques to classify pistachio species, specifically Kirmizi and Siirt, aiming to enhance quality control processes. The primary objective was to compare the performance of Convolutional Neural Networks (CNN) built from scratch with transfer learning models (VGG16 and VGG19) in classifying pistachio species. A dataset comprising 2148 images, with 1232 images of Kirmizi and 916 of Siirt, was sourced from an open dataset. Three models were implemented for the classification task: a Convolutional Neural Network (CNN) built from scratch and two transfer learning models, VGG16 and VGG19. To optimize performance, both non-fine-tuned and fine-tuned versions of VGG16 and VGG19 were evaluated. Fine-tuning involved making the last four layers of these pre-trained models trainable to better adapt to the pistachio classification task. Data augmentation techniques, such as shear transformations, zoom adjustments, flips, and translations, were applied to increase model robustness. The models were trained with different configurations, including batch sizes of 16 and 32, and learning rates of 0.0001 and 0.00001. Performance was evaluated using key metrics: accuracy, precision, recall, and F1-score. The fine-tuned VGG16 model achieved the highest accuracy of 98.15%, with a precision of 98.89%, recall of 96.74%, and an F1-score of 97.80%, outperforming the VGG19 and CNN models, which achieved accuracies of 95.83% and 90.74%, respectively. These results highlight the significant advantages of transfer learning and fine-tuning in enhancing classification performance for pistachio species, particularly in moderately sized datasets.

Objectives

The objectives of this study are as follows:

1. To explore Convolutional Neural Networks (CNN) both from scratch and using pre-trained models with transfer learning.
2. To compare CNN models, including those built from scratch and pre-trained models, across different hyperparameter settings to identify the best-performing model.
3. To develop a model capable of classifying pistachio species (Kirmizi and Siirt) through image classification.

Methods

1. Data preprocessing

1.1. The image dataset of pistachio species from Murat Koklu's open-source collection source: https://www.muratkoklu.com/datasets/, consisting of pistachio images 2148 images in dataset have two classes of pistachio species Kirmizi 1232 and Siirt 916 images respectively, splitting the dataset into train, test, validation set in ratio 80:10:10.

1.2. The images were resized to a standard dimension of 224 × 224 pixels, The pixel values of the images were normalized by rescaling them by a factor of 1/255. Augmentation strategies were used: shear transformations were applied with a range of 0.2 and zoom augmentations within

a 20% range. Both horizontal and vertical flipping were implemented. Rotation transformations up to 15 degrees were applied, and slight translations were introduced with width and height shifts of up to 10% of the image dimensions. For areas outside the image, the nearest pixel interpolation method was employed to fill missing values.

2. **Modeling**

This study was conducted on a computer with an Intel i5 11400H 4.40 GHz processor, 32 GB RAM, and a GTX 4060Ti graphics card. Three models were utilized in this study: Convolutional Neural Network (CNN) and two pre-trained models, VGG16 and VGG19. For the pre-trained models, two variations were explored: one without fine-tuning, where the pre-trained layers remained frozen, and another with fine-tuning, where the last four layers were made trainable to improve model adaptability to the pistachio species dataset.

The CNN model consisted of three convolutional layers, each with ReLU activation functions. The convolutional layers applied 32, 64, and 128 filters of size 3 × 3, followed by max-pooling layers with a 2 × 2 pool size. The model included a fully connected layer with 512 neurons using the ReLU activation function.

Both the VGG16 and VGG19 models had fully connected layers with 256 neurons, using the ReLU activation function.

The output layer across all models consisted of a single neuron with a sigmoid activation function, producing a probability for binary classification (Kirmizi or Siirt pistachio species).

Hyperparameters were tuned using batch sizes of 16 and 32, and learning rates of 0.0001 and 0.00001, respectively. All models were trained for 10 epochs. The Adam optimizer was employed to minimize the binary cross-entropy loss function. A dropout rate of 0.5 was applied to mitigate overfitting.

3. **Model Evaluation**

This study evaluated the performance of each model by using key metrics including Accuracy, Precision, Recall, and F1 Score. These metrics provided a comprehensive view of the classification effectiveness, measuring the model's ability to accurately classify the pistachio species across different configurations.

Results

This study developed models for classifying pistachio species (Kirmizi and Siirt) using deep learning techniques on image data. The research was conducted with a focus on applying and evaluating various models to achieve the study's objectives. The following models were utilized in the study:

1. Convolution neural networks built from scratch model
2. VGG16 model
3. VGG19 model

The results of these models are presented below:

1. **Convolution Neural Network (CNN)**

The best performance for the CNN model was observed with a batch size of 32 and a learning rate of 10^{-4}, achieving an accuracy of 90.74%. With a precision of 86.00% and a recall of 93.48%, resulting in an F1 score of 89.58%. The result of all CNN models shown in Table 1.

Table 1. Convolution Neural Network (CNN) model results.

Learning Rate	Batch	Accuracy	Precision	Recall	F1 Score
10^{-4}	16	0.8935	0.8350	0.9348	0.8821
	32	**0.9074**	**0.8600**	**0.9348**	**0.8958**
10^{-5}	16	0.8750	0.8652	0.8370	0.8508
	32	0.8704	0.8556	0.8370	0.8462

Note: bold letter represents to best model.

2. **VGG16**

The best performance for the VGG16 model was observed with fine-tuning, a batch size of 32, and a learning rate of 10^{-5}, achieving an accuracy of 98.15%. With a precision of 98.89% and a recall of 96.74%, this configuration resulted in an F1 score of 97.80%. The results of all VGG16 models are shown in Table 2.

Table 2. VGG16 model results.

Model	Learning Rate	Batch	Accuracy	Precision	Recall	F1 Score
None fine tune	10^{-4}	16	0.9444	0.9444	0.9239	0.9341
		32	0.9537	0.9271	0.9674	0.9468
	10^{-5}	16	0.9167	0.8776	0.9348	0.9053
		32	0.8981	0.8804	0.8804	0.8804
Fine Tune	10^{-4}	16	0.9583	1.0000	0.9022	0.9486
		32	0.9630	0.9884	0.9239	0.9551
	10^{-5}	16	0.9769	1.0000	0.9457	0.9721
		32	**0.9815**	**0.9889**	**0.9674**	**0.9780**

Note: bold letter represents to best model.

3. **VGG19**

The best performance for the VGG19 model was observed with fine-tuning, a batch size of 16, and a learning rate of 10^{-5}, achieving an accuracy of 95.83%. With a precision of 95.60% and a recall of 94.57%, this configuration resulted in an F1 score of 95.08%. The results of all VGG19 models are shown in Table 3.

Table 3. VGG19 model results.

Model	Learning Rate	Batch	Accuracy	Precision	Recall	F1 Score
None fine tune	10^{-4}	16	0.9213	0.8947	0.9239	0.9091
		32	0.9028	0.9494	0.8152	0.8772
	10^{-5}	16	0.8843	0.8941	0.8261	0.8588
		32	0.9120	0.9195	0.8696	0.8939
Fine Tune	10^{-4}	16	0.9491	1.0000	0.8804	0.9364
		32	0.9583	0.9368	0.9674	0.9519
	10^{-5}	**16**	**0.9583**	**0.9560**	**0.9457**	**0.9508**
		32	0.9491	0.9765	0.9022	0.9379

Note: bold letter represents to best model.

Conclusion

This study involved analyzing and comparing three models—CNN from scratch, VGG16, and VGG19—to identify the optimal parameters for each model to effective classification for pistachio species (Kirmizi and Siirt). The best performance achieved by each model is summarized in Table 4.

Table 4. The best performing model of three models.

Model	Learning Rate	Batch Size	Accuracy
CNN from scratch	10^{-4}	32	0.9074
VGG16 (Fine Tune)	10^{-5}	32	0.9815
VGG19 (Fine Tune)	10^{-5}	16	0.9583

The results showed that the fine-tuned VGG16 model achieved the highest overall accuracy at 98.15%, with a precision of 98.89%, a recall of 96.74%, and an F1-score of 97.80%, making it the top-performing model. In comparison, the non-fine-tuned VGG19 model and the CNN model had lower accuracies of 95.83% and 90.74%, respectively. These findings conclude that transfer learning with fine-tuning significantly outperforms both non-fine-tuned pre-trained models and the CNN built from scratch.

Discussion

Using the VGG16 with fine tune model and hyperparameter are learning rate 10^{-5}, batch size 32 for classification pistachio species by images. The corresponding confusion matrix is displayed in Figure 1.

Figure 1 shows that the model predicted Kirmizi species to true positives for 123 images and predicted as false negatives for 1 image, predicted Siirt species to true positives for 89 images and predicted as false negatives for 3 images.

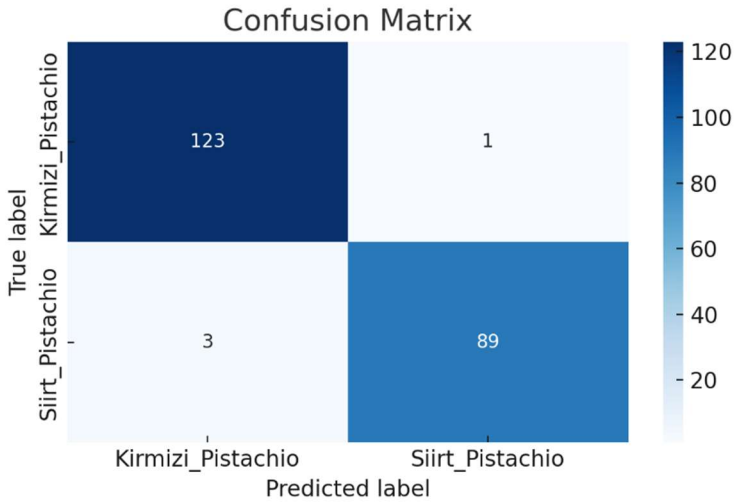

Figure 1. Confusion matrix.

The images that are corrected and incorrectly predicted by the model are displayed in Figures 2 and 3.

Figure 2. Corrected images presented by model.

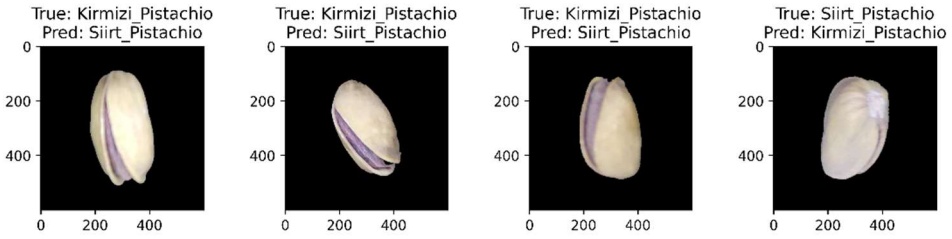

Figure 3. incorrect images presented by model.

This model predictions support the performance of using fine-tuning in transfer learning models. By making some layers of the pre-trained VGG16 network trainable, the model was able to better adapt to the pistachio dataset and improve its classification capability compared to non-fine-tuned models and the custom CNN built from scratch. The precision and recall values of the fine-tuned VGG16 model shows it robustness with precision of 98.89% that means when the model predicted a species as Kirmizi or Siirt, it was corrected 98.89% of time with recall of 96.74% indicates that model can identified many true instances of each species. Also, F1 score of 97.80% indicates that model can reliability across both classes.

Compared to other models, the VGG19 and CNN models revealed lower performance, with accuracies of 95.83% and 90.74%, respectively, the fine tune VGG19 model was performed well, achieved strong precision and recall but still slightly less effective than VGG16, this show that the higher complexity of architecture is unnecessary to the pistachio dataset. While CNN was the lowest performer. These reflected that using transfer learning and pre-trained model is suitable for this specific classification task.

This study shows that fine tuning in transfer learning can help to solve problems for agricultural images classification. Performance of the VGG16 model suggests that transfer learning with fine tuning can enhance model performance in moderate dataset like pistachio species dataset. Which allowed the model to leverage pre-trained knowledge adaptation to new data.

The suggestion, optimization could be explored for example tuning additional hyperparameters, increasing the dataset size, use another pre-trained model like ResNet or EfficientNet for further improvement in classification accuracy.

Reference

1. Pistachio Image Dataset. Available online: https://www.muratkoklu.com/datasets/.