Article

Fruit Image Classification Model Based on MobileNetV2 with Deep Transfer Learning Technique

Yonis Gulzar

Special Issue

To the Future: Adoption of Artificial Intelligence and Blockchain in Agriculture and Healthcare from a Sustainability Perspective

Edited by
Dr. Ali A. Alwan Al-juboori and Dr. Yonis Gulzar

Article

Fruit Image Classification Model Based on MobileNetV2 with Deep Transfer Learning Technique

Yonis Gulzar

Department of Management Information Systems, College of Business Administration, King Faisal University, Al-Ahsa 31982, Saudi Arabia; ygulzar@kfu.edu.sa; Tel.: +966-545-719-118

Abstract: Due to the rapid emergence and evolution of AI applications, the utilization of smart imaging devices has increased significantly. Researchers have started using deep learning models, such as CNN, for image classification. Unlike the traditional models, which require a lot of features to perform well, CNN does not require any handcrafted features to perform well. It uses numerous filters, which extract required features from images automatically for classification. One of the issues in the horticulture industry is fruit classification, which requires an expert with a lot of experience. To overcome this issue an automated system is required which can classify different types of fruits without the need for any human effort. In this study, a dataset of a total of 26,149 images of 40 different types of fruits was used for experimentation. The training and test set were randomly recreated and divided into the ratio of 3:1. The experiment introduces a customized head of five different layers into MobileNetV2 architecture. The classification layer of the MobileNetV2 model is replaced by the customized head, which produced the modified version of MobileNetV2 called TL-MobileNetV2. In addition, transfer learning is used to retain the pre-trained model. TL-MobileNetV2 achieves an accuracy of 99%, which is 3% higher than MobileNetV2, and the equal error rate of TL-MobileNetV2 is just 1%. Compared to AlexNet, VGG16, InceptionV3, and ResNet, the accuracy is better by 8, 11, 6, and 10%, respectively. Furthermore, the TL-MobileNetV2 model obtained 99% precision, 99% for recall, and a 99% F1-score. It can be concluded that transfer learning plays a big part in achieving better results, and the dropout technique helps to reduce the overfitting in transfer learning.

Keywords: deep learning; classification; fruits; MobileNetV2; precision agriculture

1. Introduction

Artificial intelligence is becoming a driving force in all aspects of life, and the agricultural and food industries are not far behind. Artificial intelligence has been used in numerous areas, such as healthcare, education, agriculture, and many other areas. In the area of healthcare, AI has been used for diagnosing numerous diseases, such as skin cancer disease [1], identifying different anatomy objects [2], predicting neurodevelopmental disorders in children [3], mental health, and other issues [4,5]. When it comes to agriculture, the world is facing challenges, such as increasing global population, global warming, and other human-caused environmental hazards, which may eventually lead to an increase in demand for food supplies. This is where AI and the computer-vision-driven Agtech industry appears to come to the rescue by speeding up the process of harvesting, quality control, picking and packing, sorting, grading, and other processes [6]. When it comes to fruits, they are very delicate and decay quickly. Around 30–35% of harvested fruits get wasted due to improper and delayed identification, classification, and the grading of fruits undertaken by non-skilled workers. Fruit classification is considered the most difficult and vital process when it comes to selling/purchasing fruit. A person dealing with selling/buying fruits needs to have knowledge of the different varieties of fruit(s) for prizing purposes. So, a person needs to have good knowledge of recognizing the different varieties of fruits.

Many fruit, vegetable, and seed identification, classification, and grading methods have been developed [7]. There have been different classification methods proposed for different classes of fruits. For instance, Altaheri et al. [8] proposed a robotic harvesting model designed to classify five different types of date fruits. This model achieved around 99% accuracy. This model used an in-house dataset for training and testing. The dataset contained overall 8000 images. In another study, Shamim Hossain et al. [9] developed a fruit classification model for industrial applications. They used the publicly available dataset to train and test their model. One of the datasets contained images of fruits which are complex to identify. The proposed model achieved an accuracy of 85%. Gulzar et al. [10] proposed a model for seed classification based on VGG16. They used thirteen types of different seeds, and the model achieved 99% accuracy. On the other hand, Hamid et al. [11] proposed a model based on the same dataset and used MobileNetV2 as the base model. They incorporated a transfer learning technique [12], and the model acquired 99% accuracy. Saranya et al. [13] undertook a comparative study in which they trained different machine learning and deep learning models on a public dataset. This dataset contains images of different fruits, such as apples, bananas, oranges, and pomegranates. They concluded that deep learning-based models outperform machine learning models. Rojas-Aranda et al. [14] developed a model to classify fruits in retail stores using deep learning. The purpose of this study was to improve the checkout process in retail stores. The model showed an accuracy of 95% when the fruits were within the plastic bags, whereas the accuracy was recorded as 93% when the fruits were not covered by plastic. Sridhar et al. [15] proposed a model for 31 different types of fruits using a hybrid approach. They incorporated CNN and an autoencoder to handle these huge data of 31 different fruits. They claim that their model achieved 99% of accuracy. Zhou et al. [16] developed a model to detect the plumpness of the strawberry fruit. They attained around 86% accuracy in terms of detecting strawberries in the greenhouse. They used RGD data while training the proposed model. Mamat et al. [17] proposed a model based on deep learning using with Only Look Once (YOLO) versions and adopted transfer learning for palm oil fruit. The model attained 98.7% accuracy for palm oil fruit.

Some researchers focused on the identification and classification of fruit diseases [18]. This study used VGG19 architecture as a base model. They claimed that their proposed model obtained around 99% accuracy in classifying fruits and their diseases. In another study, Assuncao et al. [19] proposed a deep learning model to operate on mobile devices. This model aims to classify peaches based on their freshness as well as to identify three types of diseases found in peach fruit. The accuracy of the model was recorded as 96%. They incorporated some preprocessing techniques to improve the accuracy of the proposed model.

There have been some studies that focused on the quality of the fruits, such as [20–23]. Garillos-Manliguez et al. [20] proposed a model for the estimation of the maturity of papaya fruit. The unique thing about this model is that it is trained on hyperspectral and visible-light images, unlike other models. These images not only show the external characteristics but also provide details about the inside of the fruit. The model acquired 97% of accuracy in terms of estimating the maturity of papaya fruit. Herman et al. [21] chose oil palm fruit to check its ripeness. The dataset they used contained around seven different types of ripeness levels of the oil palm fruit. They trained two well-known architectures (AlexNet and DenseNet) on this dataset and concluded that DenseNet outperformed AlexNet in terms of accuracy by 8%. Mahmood et al. [23] performed a comparative study on two well-known architectures (AlexNet and VGG16) to check the maturity level of jujube fruit. The dataset contained three different varieties of images in terms of ripeness (unripe, ripe, and over-ripe). They also used some preprocessing techniques, such as data augmentation. They claimed that VGG16 outperformed the AlexNet architecture by achieving an accuracy of 98%.

When it comes to apple fruits, the apple is a Rosaceae family fruit that originated in Asia. It is grown in over 63 countries throughout the world, with China being the main producer. Due to their high-water content, carbohydrates, organic acids, vitamins, minerals, and dietary fibers, apples are regarded as the most nutritious food. The apple, which is the fourth most widely cultivated and consumed fruit on the planet, can be divided into several types depending on its qualitative characteristics [24]. There are around 7500 different varieties of apples found in the world [25]. Different apple varieties have different benefits when it comes to health. For a common person, it is not easy to identify all kinds of apples and other fruits with many varieties. Therefore, there is a need for an approach/model based on deep learning, which can identify different kinds of fruits and solve the problem of being dependent on an expert and improve the efficiency and accuracy in identifying and classifying different fruit types.

In this study, a deep learning approach was proposed for the classification and identification of different kinds of fruits. The proposed model incorporates a transfer learning technique, which helps to solve problems involving issues of insufficient training data. This technique encourages the idea of not training the model from scratch and significantly helps in reducing training time. In this study, a well-known deep learning model, MobileNetV2 [26], was used as the base model but was modified by adding five different layers for improving the accuracy and reducing the error rate during the classification process. The proposed model is trained on a dataset containing 40 varieties of fruits. The results show that the proposed model achieved the highest accuracy rate in identifying different types of fruits. The proposed model can be deployed in a mobile application for practical usage. Further details about it are mentioned in Section 2.

The following points summarize the contributions of this paper:

- A detailed review was conducted to examine the most notable work in fruit classification via machine learning and deep learning.
- A fruit classification problem was re-introduced based on a pre-trained MobileNetV2 CNN model, in which different kinds of fruits were classified.
- A modified model was proposed using advanced deep learning techniques for fruit classification, and different model-tuning techniques were used to reduce the chances of model overfitting, such as dropout and data augmentation techniques.
- An optimization technique was developed to monitor any positive change in the validation accuracy and validation error rate. In case of change, a backup of an optimal model was taken to make sure that the proposed model shows optimal accuracy and the least validation loss.

The remainder of this article is organized as follows: In Section 2, the description of the dataset, model selection, proposed model, model tuning, and experimental settings are reported and discussed. In Section 3, the results and discussions are provided, whereas Section 4 describes the conclusion.

2. Materials and Methods

2.1. Dataset

In this research work, a public dataset [27] was used, containing images of forty different types of fruits, such as Braeburn, Crimson Snow, Golden 1, Golden 2, Golden 3, Granny Smith, Pink Lady, Red 1, Red 2, Red 3, Red Delicious, Red Yellow 1, Red Yellow 2, Apricot, Avocado, Banana, Banana Lady Finger, Banana Red, Beetroot, Blueberry, Cactus Fruit, Cantaloupe 1, Cantaloupe 2, Carambola, Cherry, Cherry Wax Black, Cherry Wax Red, Cherry Wax Yellow, Chestnut, Clementine, Cocos, Grape Blue, Grape White, Grapefruit Pink, Grapefruit White, Guava, Huckleberry, Kaki, Kiwi, and Kumquats. The dataset is divided into two subsets training, and testing datasets. The original number of images per variety of the fruit is shown in Table 1.

Table 1. Details of different types of fruits and their image count.

Fruit Types	No. of Images	Fruit Types	No. of Images
Apple Braeburn	640	Cactus Fruit	654
Apple Crimson Snow	604	Cantaloupe 1	658
Apple Golden 1	644	Cantaloupe 2	656
Apple Golden 2	653	Carambola	738
Apple Golden 3	645	Cherry	656
Apple Granny Smith	646	Cherry Wax Black	656
Apple Pink Lady	620	Cherry Wax Red	645
Apple Red 1	656	Cherry Wax Yellow	658
Apple Red 2	636	Chestnut	616
Apple Red 3	595	Clementine	818
Apple Red Delicious	654	Cocos	654
Apple Red Yellow 1	711	Grape Blue	656
Apple Red Yellow 2	836	Grape White	658
Apricot	635	Grapefruit Pink	656
Avocado	593	Grapefruit White	658
Banana	642	Guava	656
Banana Lady Finger	616	Huckleberry	656
Banana Red	640	Kaki	654
Beetroot	604	Kiwi	622
Blueberry	626	Kumquats	628

The fruits were planted in a rotating motor shaft (3 *rmp*), and a short video of 20 s was taken for each class. The images were taken from that video. The background of the images was removed with an algorithm because of variations in the lighting conditions. Figure 1 shows the sample captured images of fruits.

Figure 1. Sample images of forty types of fruits.

2.2. Model Selection

In the field of image processing, CNN has gained more attention due to its huge economic potential and high accuracy rate. There are well-known CNN architectures, which are very popular in the area of image processing and classification, such as AlexNet [28], VGG16 [29], InceptionV3 [30], ResNet [31], and MobileNetV2 [26]. In any computer vision task, it is the convolution operation(s) which has a high contribution. Nevertheless, due to large and deep network structures, the processing time, as well as cost, increases in most of these networks, such as AlexNet, VGG16, InceptionV3, ResNet, etc. However, when it comes to MobileNetV2, its inverse residual structure and linear bottleneck structure help to reduce the calculation of convolution. Its architecture is preferred over others due to its simple architecture and memory-efficient characteristics. Table 2 presents the precision, recall, and F1-score of AlexNet, VGG16, InceptionV3, ResNet, and MobileNetV2. It is important to note that all the models were trained on the fruit dataset without using any pre-processing techniques. Only the classification layer was modified based on the number of classes present in the said dataset. As mentioned before, MobileNetV2 is memory-efficient and was mainly designed for mobile applications. Due to this, it can be seen from the table that it is outperforming all of the other models in terms of accuracy.

Table 2. Precision, recall, and F1-score of different models while trained on the fruit dataset.

Models	Precision	Recall	F1-Score
AlexNet	0.81	0.79	0.83
VGG16	0.79	0.80	0.80
InceptionV3	0.84	0.82	0.81
ResNet	0.82	0.81	0.82
MobileNetV2	**0.89**	**0.91**	**0.89**

Figure 2 shows the accuracy of different models achieved while training on the fruit dataset containing forty different types of fruits. From the figure, it is evident that MobileNetV2 achieves the highest accuracy with 89%, which is higher than ResNet, InceptionV3, VGG16, and AlexNet by 7, 5, 11, and 7%, respectively.
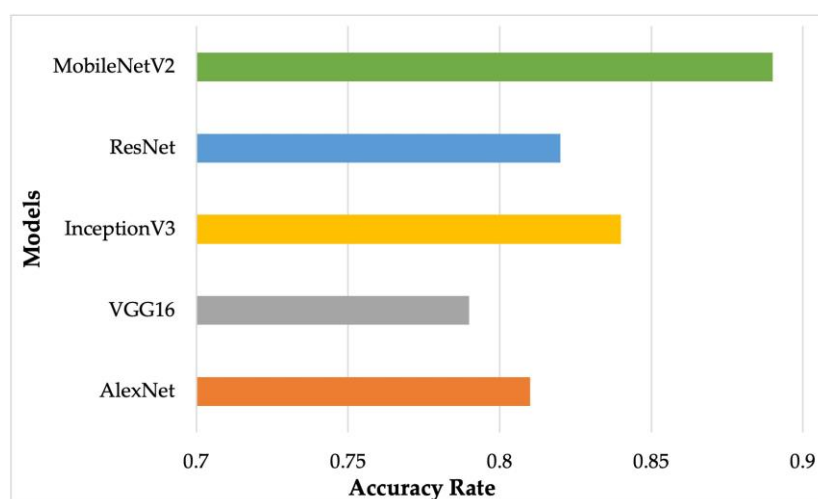

Figure 2. Accuracy rate of different models during model selection.

2.3. TL-MobileNetV2

This subsection focuses on the proposed model based on the fruit classification dataset. The proposed model, TL-MobileNetV2, is based on MobileNetV2 architecture [26]. The said architecture is mainly developed for mobile and resource-constrained environments. The main purpose for adopting this type of architecture lies in its strength in terms of

reducing the memory usage, computational expenses, and compliance of its design to be more focused towards mobile applications. It is important to note that the said architecture has initially around one thousand nodes in its classification layer. Nevertheless, in order to make it matchable with the said problem, it is important to fix the number of nodes in the classification layer. To improve the accuracy, a customized head, containing five different layers was attached to the base model of MobileNetV2 by replacing the classification layer.

The customized head contains (i) average pooling layer, (ii) flatten layer, (iii) dense layer, (iv) dropout layer, and (v) softmax layer, as shown in Figure 3. It is important to mention that the size of the average pooling layer is set to (7×7). In the flatten layer, the flattened neurons were fed to a dense layer with the activation function being ReLU. This was proceeded by setting the probability of a dropout layer with a value of 0.5 and the addition of forty nodes within the classification layer of the model. Carrying out all these changes has produced an improved version of the MobileNetV2 architecture bearing forty different nodes in its last (classification) layer; this is an optimal and befitting model for the said problem addressed in this work.
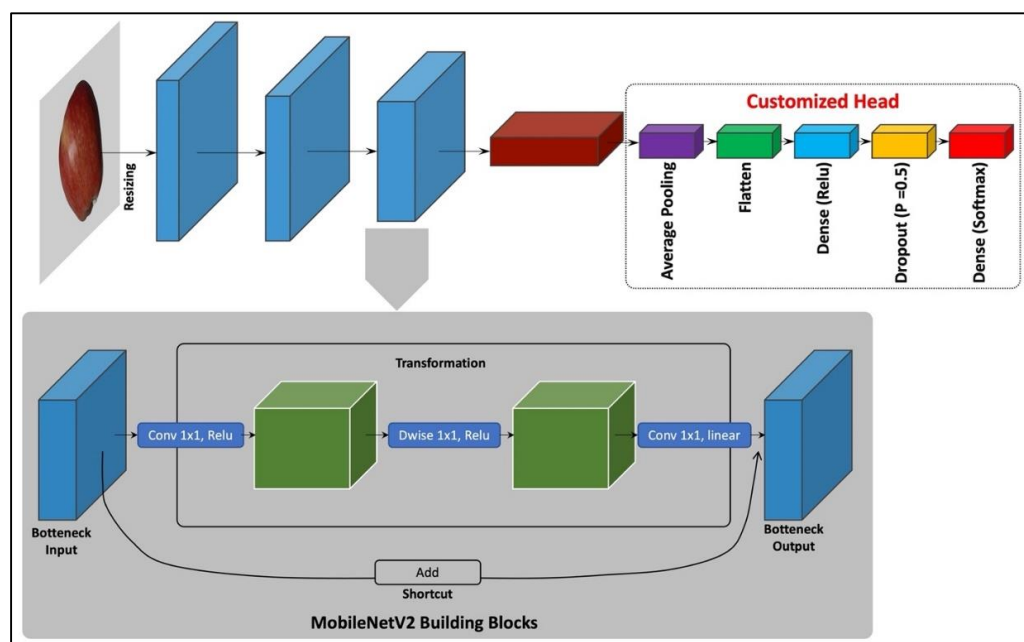

Figure 3. TL-MobileNetV2 model.

From the literature, it is evident that transfer learning [32] has improved the performance of target learners on target domains. As a result, the dependency on a large volume of data has decreased, and the issues related to the unavailability of sufficient data were reduced [33,34]. There are different ways to integrate transfer learning in deep learning models; one way is to train the model totally based on the new dataset. Another way is to train only those layers of the model which were added to the existing model, whereas the existing layers are frozen. In this research work, a hybrid approach was opted, where at the beginning of training the model, only the newly added layers were trained based on the fruit dataset. After the 20th iteration, the existing layers were unfrozen, and a slight weight adjustment was performed to the trained layers of the model based on the said dataset.

2.4. Model Tuning

In this work, we have incorporated different preprocessing/model tuning techniques to prevent the model from overfitting. The following are those techniques briefly explained:

- Data Augmentation: one of the best ways is to overcome the issue of insufficient data is to use the data augmentation technique. This technique helps to create random artificial images from the source data using different methods, such as shifts, shears,

random rotations, and flips. In this work, a built-in function in the Keras library [35] was used to create ten images from each image by randomly zooming 20%, adjusting height by 10%, width shifting by 10%, and rotating by 30%.

- Adaptive Learning Rate: this technique helps to accelerate the training and alleviates the pressure of choosing a learning rate and schedule. In this work, the initial learning rate was set to INIT_LR = 0.001, and then the decay of the form decay = INIT_LT/EPOCHS was used.
- Model Checkpointing: while training the model, a checkpoint is set to monitor whenever there is any positive change in accuracy. The model's weights are saved whenever the accuracy is optimum. In this research work, a model checkpoint of the form checkpoint = Model Checkpoint (fname, monitor = "val_loss", mode = "min", save_best_only = True, verbose = 1) was used. This callback monitors the validation loss of the model and overwrites the trained model only when there is a decrease in the loss compared to the previous best model.
- Dropout: The dropout technique helps avoid the issue of overfitting, and during the training, neurons are randomly chosen and discarded. It implies that the contribution of these neurons is temporally ignored to the activation of downstream neurons, and any weight changes are not implemented on any neuron on the backward pass.

2.5. Experimental Environment Settings and Model Evaluation Indicator

This research aims to propose an optimal model which identifies and classifies different types of fruits. The proposed model was implemented using Python 3.0 on Windows 10 OS, with system configuration using an *i7* processor and 16 GB RAM.

The performance of the model is measured using a wide range of matrices derived from a 2 × 2 confusion matrix, as shown in Figure 4.

Figure 4. Confusion matrix ($C_{i,j}$).

When incorporating machine learning techniques, it is important to evaluate the model to check its performance. For that reason, the matrices, as follows, were used to evaluate the proposed model:

- $Accuracy = \frac{TP+TN}{2TP+FP+TN+FNa}$
- $Precision = \frac{TP}{TP+FP}$
- $Recall = \frac{TP}{TP+FN}$
- Where *TP, TN, FP,* and *FN* are true positive, true negative, false positive, and false negative, respectively
- **Macro Average:** The function to compute F-1 for each label and returns the average without considering the proportion for each label in the dataset.
- **Weighted Average:** The function to compute F-1 for each label and returns the average considering the proportion for each label in the dataset.

The trained model is further assessed using validation accuracy and loss. Moreover, continuous monitoring is used to spot any noticeable deviation in training and validation performance when it comes to validation accuracy and loss.

3. Results

This section presents the performance of the proposed model in terms of training accuracy, training loss, as well as validation accuracy, as shown in Figure 5. The proposed model was trained for 100 iterations. From the figure, it can be noticed that the training accuracy of the proposed model started with 44% from the first iteration, and the accuracy increased dramatically. The accuracy touched 90% within the first 10 iterations. At the 30th iteration, the model reached maximum accuracy (100%). From the 30th iteration onwards, it can be inferred that the accuracy rate of the proposed model has remained at the maximum until the end of the training.
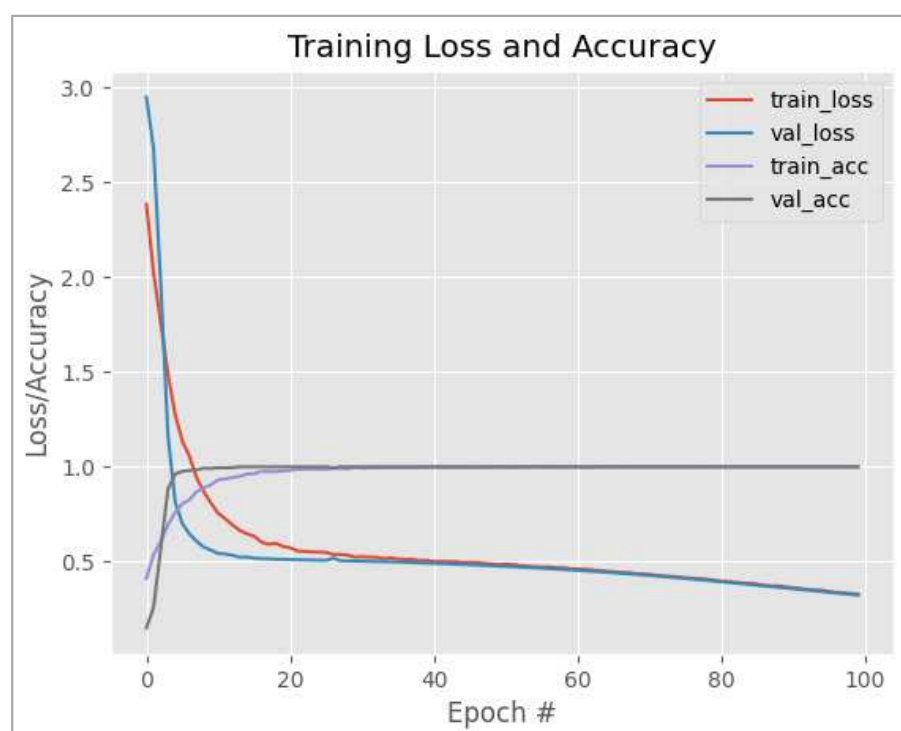

Figure 5. Training accuracy and loss in terms of training and validation of TL-MobileNetV2 model.

Figure 5 also shows the training loss of the proposed model. From the figure, it is depicted that at the beginning of the training phase, the training loss is high as the model has not been exposed to the data. However, gradually the model reads the images and starts to remember them, and eventually the training loss gets reduced. It can be noticed that the training loss reaches 0.6 within the first 20 iterations and gets reduced dramatically by each iteration. By the 100th iteration, the training loss has reached 0.3 which infers the characteristics of a good model.

Generally, most of the models perform well during training. However, they don't perform well during the validation phase. This is due to the fact that the model has only been trained on supervised data. In order to find out its performance, it is important to validate the performance of the model. For this purpose, the proposed model was validated as well. Figure 5 shows the validation accuracy and loss of the proposed model. It depicts that the model starts with 15% of accuracy at the beginning of the process and within 10 iterations, the model reaches 100%. From the 5th iteration onwards, the model's validation accuracy remains steady. When it comes to the validation loss of the proposed model, it can be seen from the figure that the model starts with a very high validation loss,

which is common in all deep learning models. The validation loss of the proposed model dramatically falls and reaches 0.5 within the first 10 iterations. From the 10th iteration onwards until the 50th iteration, the validation loss remains constant and then starts falling again from the 55th iteration. Finally, it reaches 0.35 at the end of the 100th iteration.

During the training and validation process, the model shows stability in its performance. The proposed model shows a very high accuracy rate during training, and it reflected the same during validation. When it comes to training loss and validation loss, the model proves that it does not overfit. This is due to the fact the preprocessing techniques used in the proposed model helped the model to achieve better results without overfitting. Moreover, the data augmentation technique incorporated in the proposed model played a vital role in exposing the model to different variations of the images. Additionally, using of dropout technique helped the model's validation performance by making sure that the model does not deviate much from its training performance.

Table 3 shows the performance of the TL-MobileNetV2 model based on each fruit class in terms of precision, recall, F1-score, and support. The supports show the number of images used for training and validating the model after applying the data augmentation technique. As mentioned earlier, for each instance of the image, the image augmentation technique creates ten different instances of that image, which are used for model training. It can be inferred from the table that the model achieved the maximum value for each class when it comes to precision, recall, and F1-score except for some of fruits, such as Apple Golden 1, Apple Red 3, Apple Red Yellow 2, Banana Lady Finger, Cantaloupe 2, Cherry Wax Red, and Kaki. This is due to the fact that the dataset does not have any variations present in the background of any images. As stated earlier, the background of each instance was removed, which makes this fruit dataset fit for the proposed model. However, the precision of Apple Golden 1, Apple Red 3, and Apple Red Yellow 2 is 0.97, 0.96, and 0.98, respectively. This is due to the fact that Apple Golden 1 resembles Apple Golden 2 in terms of color; likewise, Apple Red 3 resembles Apple Red 2, and Apple Red Yellow 2 resembles Apple Red Yellow 1. The overall accuracy of the proposed model for all classes of fruits during training is 100%, as shown in the figure. It is important to note that applying different preprocessing techniques in the model helped to achieve a high accuracy rate.

Usually, models perform well during training and under supervised data. However, when it comes to real-world data, they usually do not perform well. For that reason, we tested our proposed model's performance by feeding it unseen data during the testing phase. The dataset contains instances of images for testing purposes. It is important to note that the model has not seen testing dataset images before. So, using such a dataset for testing will help to identify the fair performance of the proposed model without any bias. Table 4 presents the testing results of the TL-MobileNetV2 model in terms of precision, recall, F1-score, and support. From the table, it can be inferred that the proposed model has achieved the maximum accuracy in all classes except for a few, such as Apple Golden 1, Apple Red 3, Apple Red Yellow 2, Banana Lady Finger, Cantaloupe 2, Cherry Wax Red, and Kaki. Their precision is recorded 0.92, 0.96, 0.98, 0.92, 0.95, 0.92, and 0.94, respectively, whereas the F1-score of these fruits is 0.94, 0.97, 0.96, 0.94, 0.96, 0.93, and 0.95, respectively. This may be due to the fact that there are chances of false negative predictions between Golden 1 and Golden 2, as well as between Red 1 and Red 3, and Red Yellow 1 and Red Yellow 2 as these classes are somehow similar in terms of color. The model achieved 99% accuracy in both the training and testing phase. This also proves that the model did not overfit as there would have been a difference in the training and testing score of the TL-MobileNetV2 model.

Comparison of Model Performance

The proposed model was further compared with well-known models, as mentioned in Section 2.2 (AlexNet [28], VGG16 [29], InceptionV3 [30], and ResNet [31]). To obtain a fair comparison, the abovementioned model tuning and preprocessing techniques (discussed in Section 2.4) were applied to the mentioned models with the same number of iterations.

Table 5 represents the performance of different models compared with the TL-MobileNetV2 model. It can be inferred that VGG16 performs worse in terms of accuracy, whereas InceptionV3 performs slightly better than AlexNet and ResNet. It can be also noticed that the MobileNetV2 is better than all of the other models, but the modified version of MobileNetV2, TL-MobileNetV2, outperforms all the models mentioned in the table. TL-MobileNetV2 is performing better in all the parameters whether it is precision, recall, or F1-score. This is due to the fact that TL-MobileNetV2 is a modified version of MobileNetV2 with five extra layers, which played a great role in terms of improving the accuracy of the model.

Table 3. Validation results of TL-MobileNetV2 model.

Fruit Types	Precision	Recall	F1-Score	Support
Apple Braeburn	1	1	1	4920
Apple Crimson Snow	1	1	1	4440
Apple Golden 1	0.97	1	0.98	4800
Apple Golden 2	1	1	1	4920
Apple Golden 3	1	1	1	4810
Apple Granny Smith	1	1	1	4920
Apple Pink Lady	1	1	1	4560
Apple Red 1	1	1	1	4920
Apple Red 2	1	1	1	4920
Apple Red 3	0.96	1	0.97	4290
Apple Red Delicious	1	1	1	4900
Apple Red Yellow 1	1	1	1	4920
Apple Red Yellow 2	0.98	1	0.99	6720
Apricot	1	1	1	4920
Avocado	1	1	1	4270
Banana	1	1	1	4900
Banana Lady Finger	0.98	1	0.99	4500
Banana Red	1	1	1	4900
Beetroot	1	1	1	4500
Blueberry	1	1	1	4620
Cactus Fruit	1	1	1	4900
Cantaloupe 1	1	1	1	4920
Cantaloupe 2	0.99	1	1	4920
Carambola	1	1	1	4920
Cherry	1	1	1	4920
Cherry Wax Black	1	1	1	4920
Cherry Wax Red	0.98	1	0.99	4920
Cherry Wax Yellow	1	1	1	4920
Chestnut	1	1	1	4500
Clementine	1	1	1	4900
Cocos	1	1	1	4900
Grape Blue	1	1	1	4900
Grape White	1	1	1	4920
Grapefruit Pink	1	1	1	4900
Grapefruit White	1	1	1	4920
Guava	1	1	1	4900
Huckleberry	1	1	1	4900
Kaki	0.97	1	0.99	4900
Kiwi	1	1	1	4660
Kumquats	1	1	1	4710
Accuracy	-	-	1	194,000
Macro Avg	1	1	1	194,000
Weighted Avg	1	1	1	194,000

Table 4. Testing results of TL-MobileNetV2 model.

Fruit Types	Precision	Recall	F1-Score	Support
Apple Braeburn	1	1	1	148
Apple Crimson Snow	1	1	1	160
Apple Golden 1	0.92	1	0.94	164
Apple Golden 2	1	1	1	161
Apple Golden 3	1	0.96	1	164
Apple Granny Smith	1	1	1	154
Apple Pink Lady	1	1	1	164
Apple Red 1	1	1	1	164
Apple Red 2	1	1	1	144
Apple Red 3	0.96	1	0.97	166
Apple Red Delicious	1	1	1	164
Apple Red Yellow 1	1	1	1	219
Apple Red Yellow 2	0.98	0.94	0.96	164
Apricot	1	1	1	143
Avocado	1	1	1	166
Banana	1	1	1	152
Banana Lady Finger	0.92	1	0.94	166
Banana Red	1	1	1	150
Beetroot	1	1	1	154
Blueberry	1	0.97	1	164
Cactus Fruit	1	1	1	164
Cantaloupe 1	1	1	1	166
Cantaloupe 2	0.95	1	0.96	164
Carambola	1	1	1	246
Cherry	1	1	1	164
Cherry Wax Black	1	1	1	164
Cherry Wax Red	0.92	1	0.93	153
Cherry Wax Yellow	1	1	1	166
Chestnut	1	1	1	166
Clementine	1	1	1	328
Cocos	1	1	1	164
Grape Blue	1	1	1	166
Grape White	1	1	1	166
Grapefruit Pink	1	0.99	1	166
Grapefruit White	1	1	1	166
Guava	1	1	1	166
Huckleberry	1	1	1	166
Kaki	0.94	1	0.95	164
Kiwi	1	1	1	156
Kumquats	1	1	1	157
Accuracy	-	-	0.99	6749
Macro Avg	0.99	0.99	0.99	6749
Weighted Avg	0.99	0.99	0.99	6749

Table 5. Performance of different models compared with the TL-MobileNetV2 model.

Models	Precision	Recall	F1-Score
AlexNet	0.91	0.89	0.92
VGG16	0.88	0.90	0.88
InceptionV3	0.93	0.88	0.91
ResNet	0.89	0.92	0.90
MobileNetV2	0.96	0.97	0.96
TL-MobileNetV2	**0.99**	**0.99**	**0.99**

Furthermore, the proposed model, TL-MobileNetV2, was compared with start-of-the-art models for its qualitative and quantitative features, as mentioned in Table 6. It can

be inferred from the table that different models were proposed for different datasets, for instance, date fruit, pineapple, tomato, strawberry, and even mixed fruit datasets. Some of these models adopted well-known architectures, such as ResNet or MobileNetV2, and some created their models based on CNN. From the table, it can be concluded that our proposed model performed well on the fruit dataset. The modified proposed model based on MobileNetV2 has performed efficiently well when compared with other models and is much faster and takes less space while processing. Additionally, this architecture is best-suited for the collected dataset as it was collected using a mobile camera.

Table 6. Comparison of the TL-MobileNetV2 model with state-of-the-art approaches.

Paper	Dataset	Classes	Method/Model	Classification Type	Accuracy (%)
[36]	Mix Fruit	18	CNN	Identification	95.08
[37]	Date Fruit	8	MobileNetV2	Identification	99.0
[38]	Pineapple	1	CNN	Identification/quality	97.0
[39]	Strawberry	1	ResNet	Quality	84
[40]	Tomato	1	SDF-convnets	Quality	96.5
TL-MobileNetV2	**Mix Fruit**	**40**	**MobileNetV2**	**Identification**	**99.0**

4. Conclusions

Machine learning techniques, particularly those that are suited for computer vision, have started to be widely used in precision agriculture. These techniques are used in various areas, such as fruit classification, quality analysis, yield estimation, and disease prediction. The success of these techniques has encouraged the development of deep learning models for seed classification. In this study, a deep learning model TL-MobileNetV2 was developed based on MobileNetV2 architecture. A dataset of forty types of fruits were used to train and test the proposed model. In the TL-MobileNetV2 model, five different layers were added after removing the classification layer present in the MobileNetV2 architecture to improve the efficiency and accuracy of the model. Along with this, different preprocessing and model-tuning techniques were used to make the TL-MobileNetV2 perform well on the said dataset without overfitting. The experimental results show that the TL-MobileNetV2 has performed well on the fruit dataset by attaining 99% accuracy.

In future work, a mobile-based application will be further enhanced using a larger number of different fruits, which aims to lead to a wider range of fruit classification. This application will help people with limited knowledge to classify different types of fruit and their different varieties. Furthermore, different CNN models will be trained on the dataset, and their results will be compared to identify the best-fit model in terms of accuracy and efficiency.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The dataset used in this study is public [27].

Conflicts of Interest: The author declares that there is no conflict of interest regarding the publication of this paper.

References

1. Gulzar, Y.; Khan, S.A. Skin Lesion Segmentation Based on Vision Transformers and Convolutional Neural Networks—A Comparative Study. *Appl. Sci.* **2022**, *12*, 5990. [CrossRef]
2. Khan, S.A.; Gulzar, Y.; Turaev, S.; Peng, Y.S. A Modified HSIFT Descriptor for Medical Image Classification of Anatomy Objects. *Symmetry* **2021**, *13*, 1987. [CrossRef]
3. Alam, S.; Raja, P.; Gulzar, Y. Investigation of Machine Learning Methods for Early Prediction of Neurodevelopmental Disorders in Children. *Wirel. Commun. Mob. Comput.* **2022**, *2022*, 5766386. [CrossRef]
4. Sahlan, F.; Hamidi, F.; Misrat, M.Z.; Adli, M.H.; Wani, S.; Gulzar, Y. Prediction of Mental Health Among University Students. *Int. J. Perceptive Cogn. Comput.* **2021**, *7*, 85–91.
5. Hanafi, M.F.F.M.; Nasir, M.S.F.M.; Wani, S.; Abdulghafor, R.A.A.; Gulzar, Y.; Hamid, Y. A Real Time Deep Learning Based Driver Monitoring System. *Int. J. Perceptive Cogn. Comput.* **2021**, *7*, 79–84.
6. Kakani, V.; Nguyen, V.H.; Kumar, B.P.; Kim, H.; Pasupuleti, V.R. A Critical Review on Computer Vision and Artificial Intelligence in Food Industry. *J. Agric Food Res.* **2020**, *2*, 100033. [CrossRef]
7. Behera, S.K.; Rath, A.K.; Mahapatra, A.; Sethy, P.K. Identification, Classification & Grading of Fruits Using Machine Learning & Computer Intelligence: A Review. *J. Ambient. Intell. Humaniz. Comput.* **2020**, 1–11. [CrossRef]
8. Altaheri, H.; Alsulaiman, M.; Muhammad, G. Date Fruit Classification for Robotic Harvesting in a Natural Environment Using Deep Learning. *IEEE Access* **2019**, *7*, 117115–117133. [CrossRef]
9. Shamim Hossain, M.; Al-Hammadi, M.; Muhammad, G. Automatic Fruit Classification Using Deep Learning for Industrial Applications. *IEEE Trans. Ind. Inform.* **2019**, *15*, 1027–1034. [CrossRef]
10. Gulzar, Y.; Hamid, Y.; Soomro, A.B.; Alwan, A.A.; Journaux, L. A Convolution Neural Network-Based Seed Classification System. *Symmetry* **2020**, *12*, 2018. [CrossRef]
11. Hamid, Y.; Wani, S.; Soomro, A.B.; Alwan, A.A.; Gulzar, Y. Smart Seed Classification System Based on MobileNetV2 Architecture. In Proceedings of the 2022 2nd International Conference on Computing and Information Technology (ICCIT), Tabuk, Saudi Arabia, 25–27 January 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 217–222.
12. Hamid, Y.; Elyassami, S.; Gulzar, Y.; Balasaraswathi, V.R.; Habuza, T.; Wani, S. An Improvised CNN Model for Fake Image Detection. *Int. J. Inf. Technol.* **2022**, *2022*, 1–11. [CrossRef]
13. Saranya, N.; Srinivasan, K.; Pravin Kumar, S.K.; Rukkumani, V.; Ramya, R. *Fruit Classification Using Traditional Machine Learning and Deep Learning Approach*; Springer: Cham, Switzerland, 2020; Volume 1108, ISBN 9783030372170.
14. Rojas-Aranda, J.L.; Nunez-Varela, J.I.; Cuevas-Tello, J.C.; Rangel-Ramirez, G. *Fruit Classification for Retail Stores Using Deep Learning*; Springer: Cham, Switzerland, 2020; Volume 12088, ISBN 9783030490751.
15. Sridhar, B.; Kiran, K.S.S.; Likhitha, N.; Vardhan, K.P.V.A.; Nikhil, K.V. *Development of an Automatic Fruit Classification Using Hybrid Deep Learning Model for Super Markets*; Springer: Singapore, 2022; Volume 838, ISBN 9789811685491.
16. Zhou, C.; Hu, J.; Xu, Z.; Yue, J.; Ye, H.; Yang, G. A Novel Greenhouse-Based System for the Detection and Plumpness Assessment of Strawberry Using an Improved Deep Learning Technique. *Front. Plant Sci.* **2020**, *11*, 559. [CrossRef] [PubMed]
17. Mamat, N.; Othman, M.F.; Abdulghafor, R.; Alwan, A.A.; Gulzar, Y.; Malaysia, U.T.; Sultan, J.; Petra, Y. Enhancing Image Annotation Technique of Fruit Classification Using a Deep Learning Approach. *Sustainability* **2023**, *15*, 901. [CrossRef]
18. Nasir, I.M.; Bibi, A.; Shah, J.H.; Khan, M.A.; Sharif, M.; Iqbal, K.; Nam, Y.; Kadry, S. Deep Learning-Based Classification of Fruit Diseases: An Application for Precision Agriculture. *Comput. Mater. Contin.* **2020**, *66*, 1949–1962. [CrossRef]
19. Assuncao, E.; Diniz, C.; Gaspar, P.D.; Proenca, H. Decision-Making Support System for Fruit Diseases Classification Using Deep Learning. In Proceedings of the 2020 International Conference on Decision Aid Sciences and Application, Sakheer, Bahrain, 8–9 November 2020; pp. 652–656.
20. Garillos-Manliguez, C.A.; Chiang, J.Y. Multimodal Deep Learning via Late Fusion for Non-Destructive Papaya Fruit Maturity Classification. In Proceedings of the CCE 2021–2021 18th International Conference on Electrical Engineering, Computing Science and Automatic Control, Mexico City, Mexico, 10–12 November 2021.
21. Herman, H.; Cenggoro, T.W.; Susanto, A.; Pardamean, B. Deep Learning for Oil Palm Fruit Ripeness Classification with Densenet. In Proceedings of the 2021 International Conference on Information Management and Technology, Jakarta, Indonesia, 19–20 August 2021; pp. 116–119.
22. Suharjito; Elwirehardja, G.N.; Prayoga, J.S. Oil Palm Fresh Fruit Bunch Ripeness Classification on Mobile Devices Using Deep Learning Approaches. *Comput. Electron. Agric.* **2021**, *188*, 106359. [CrossRef]
23. Mahmood, A.; Singh, S.K.; Tiwari, A.K. Pre-Trained Deep Learning-Based Classification of Jujube Fruits According to Their Maturity Level. *Neural. Comput. Appl.* **2022**, *34*, 13925–13935. [CrossRef]
24. Musacchi, S.; Serra, S. Apple Fruit Quality: Overview on Pre-Harvest Factors. *Sci. Hortic.* **2018**, *234*, 409–430. [CrossRef]
25. SelectHealth.Org. Available online: https://selecthealth.org/blog/2020/02/how-many-types-of-apples-are-there-and-which-is-best (accessed on 1 September 2022).
26. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.-C. Mobilenetv2: Inverted Residuals and Linear Bottlenecks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–22 June 2018; pp. 4510–4520.
27. Oltean, M. Fruits 360 Dataset: A Dataset of Images Containing Fruits and Vegetables 2019. Kaggle, 2020. [Online]. Available online: https://www.kaggle.com/moltean/fruits (accessed on 1 April 2022).

28. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet Classification with Deep Convolutional Neural Networks. *Adv. Neural Inf. Process. Syst.* **2012**, *25*, 84–90. [CrossRef]

29. Simonyan, K.; Zisserman, A. Very Deep Convolutional Networks for Large-Scale Image Recognition. *arXiv* **2014**, arXiv:1409.1556.

30. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the Inception Architecture for Computer Vision. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 2818–2826.

31. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.

32. Ayoub, S.; Gulzar, Y.; Reegu, F.A.; Turaev, S. Generating Image Captions Using Bahdanau Attention Mechanism and Transfer Learning. *Symmetry* **2022**, *14*, 2681. [CrossRef]

33. Zhuang, F.; Qi, Z.; Duan, K.; Xi, D.; Zhu, Y.; Zhu, H.; Xiong, H.; He, Q. A Comprehensive Survey on Transfer Learning. *Proc. IEEE* **2020**, *109*, 43–76. [CrossRef]

34. Tan, C.; Sun, F.; Kong, T.; Zhang, W.; Yang, C.; Liu, C. A Survey on Deep Transfer Learning. In Proceedings of the International Conference on Artificial Neural Networks, Rhodes, Greece, 4–7 October 2018; Springer: Cham, Switzerland, 2018; pp. 270–279.

35. Arnold, T.B. KerasR: R Interface to the Keras Deep Learning Library. *J. Open Source Softw.* **2017**, *2*, 296. [CrossRef]

36. Zhang, Y.; Satapathy, S.C.; Wang, S. Fruit Category Classification by Fractional Fourier Entropy with Rotation Angle Vector Grid and Stacked Sparse Autoencoder. *Expert Syst.* **2022**, *39*, e12701. [CrossRef]

37. Albarrak, K.; Gulzar, Y.; Hamid, Y.; Mehmood, A.; Soomro, A.B. A Deep Learning-Based Model for Date Fruit Classification. *Sustainability* **2022**, *14*, 6339. [CrossRef]

38. Huang, T.-W.; Bhat, S.A.; Huang, N.-F.; Chang, C.-Y.; Chan, P.-C.; Elepano, A.R. Artificial Intelligence-Based Real-Time Pineapple Quality Classification Using Acoustic Spectroscopy. *Agriculture* **2022**, *12*, 129. [CrossRef]

39. Su, Z.; Zhang, C.; Yan, T.; Zhu, J.; Zeng, Y.; Lu, X.; Gao, P.; Feng, L.; He, L.; Fan, L. Application of Hyperspectral Imaging for Maturity and Soluble Solids Content Determination of Strawberry with Deep Learning Approaches. *Front. Plant Sci.* **2021**, *12*, 1897. [CrossRef]

40. Ko, K.; Jang, I.; Choi, J.H.; Lim, J.H.; Lee, D.U. Stochastic Decision Fusion of Convolutional Neural Networks for Tomato Ripeness Detection in Agricultural Sorting Systems. *Sensors* **2021**, *21*, 917. [CrossRef] [PubMed]