Pre-trained deep learning-based classification of jujube fruits according to their maturity level

Article *in* Neural Computing and Applications · August 2022

DOI: 10.1007/s00521-022-07213-5

3 authors, including:

Atif Mahmood
SRM Institute of Science and Technology

Neural Computing and Applications
https://doi.org/10.1007/s00521-022-07213-5

ORIGINAL ARTICLE

Pre-trained deep learning-based classification of jujube fruits according to their maturity level

Atif Mahmood[1] · Sanjay Kumar Singh[2] · Amod Kumar Tiwari[1]

Received: 15 December 2021 / Accepted: 29 March 2022

Abstract
Assessment of crop maturity and quality is pivotal in the food industry and for harvesting. The manual classification of crops based on their maturity levels for harvesting and packaging purpose is a tedious process. However, the emergence of machine learning/deep learning techniques has opened up the ways in this direction, but its practical success is still limited. In this research, we examined two convolutional neural network paradigms (i.e., AlexNet and VGG16) utilizing a transfer-learning approach for classifying the jujube fruits based on their maturity level (i.e., unripe, ripe, and over-ripe). The training/testing of models was performed over the collected dataset of around 400 images, which was further augmented to 4398 images collectively for the three maturity classes. The best accuracy achieved for the correct classification of maturity classes with AlexNet and VGG16 for the actual and augmented images are 94.17% & 97.65%, and 98.26% & 99.17% respectively. The examined models were compared with two existing methods for jujube maturity classification and found to be performing better. The significantly improved success rate of VGG16 models over the AlexNet and existing proposed models for jujube classification makes the model recommendable for developing an efficient system for the automated harvesting and sorting of jujube fruits.

Keywords AlexNet · Classification · CNN · Jujube · Maturity · VGG16

List of symbols

ML	Machine learning
DL	Deep learning
CNN	Convolutional neural network
ANN	Artificial neural network
AlexNet	Alex network
VGG16	Visual geometry group (16 Layers)
BPNN	Back-propagation neural network
KNN	K-nearest neighbour
DT	Decision tree
GMM	Gaussian mixture model
SVM	Support vector machine
KPCA	Kernel-principal component Analysis
ELM	Extreme learning machine

✉ Atif Mahmood
atif.mahmood@recsonbhadra.ac.in

1 Department of Computer Science and Engineering, Dr. APJ Abdul Kalam Technical University, Lucknow, India

2 Department of Computer Application, Dr. APJ Abdul Kalam Technical University, Lucknow, India

1 Introduction

Jujube commonly known as Ziziphus is a sweet and delicious fruit of India and grows mostly in the spring seasons. Fruits belong to the “Rhamnaceae” family and are cultivated especially in India, China, the Middle East, and North America [1]. Jujube is referred to with varying names depending upon its cultivation regions. In India it is well known as ‘Ber’, in China it is recognized as ‘Chinese dates’, while in Arabian countries it is popular as ‘Sedra’. Jujube is consumed conventionally as a fresh jujube pulp or functionally edible in a form of jam, jelly, sauce, pickles, compotes, cakes, slices, snacks, and beverages (i.e., wine, juice, vinegar, etc.). Jujube fruits are enriched with bioactive compounds specifically vitamins, minerals, fibers, and natural antioxidant constituents. Therefore the functional products of jujube possess antioxidant, antitumor, antiobesity, and anti-diabetic properties to suppress various diseases [2]. Due to its nutritious and medicinal properties, jujube has attracted the interest of researchers and functional food production industries. The food industry needs to identify and harness the different maturity stages of jujube for preparing the different processed functional

Published online: 18 April 2022

foods. A vital step in ensuring the correct maturity of jujube fruits is its assortment simply based upon the peal color. During the three maturity phases peal color of the jujube looks green at the beginning, then turns to yellow–brown, and finally at the end of maturity turns red-brown. Figure 1 exhibits the three maturity classes of jujube fruits.

At present several destructive and non-destructive approaches were applied for the maturity classification of fruits [3]. The destructive approach is based on ideal laboratory experiments to extract features, which results in the wastage of edibles. Also, destructive approaches are time-consuming and fail to meet the robustness and testing stabilities for practical applications. In contrast, the non-destructive approach ensures a robust system for practical applications and is time-efficient, and achieves better precision as well [4]. Nowadays ML models with computer vision possess a non-destructive approach for the maturity classification of vegetables and fruits. ML models after sufficient training ensure good classification accuracy, optimize the performance, and reduce the system complexity. Survey articles [5] and [6] discuss the grading of different vegetables and fruits using ML and computer vision techniques. Compared with conventional ML techniques, DL models (i.e., CNN) had shown exquisite performance in image classification tasks with sufficient training datasets. CNN models discard the effort of feature extraction and boost the performance with fine-tuning of hyper-parameters.

In this research, we have examined two prevailing CNN architectures- AlexNet [7] and VGG16 [8] for the classification of jujube fruits based on the maturity level. Input to the CNN models is RGB images of jujube fruits belonging to three maturity classes- unripe, ripe, and over-ripe. Both the models possess different sizes and depths and harness the notion of transfer learning for classifying the jujube images into three target classes. Also, the results of examined models were compared with two existing jujube classification methods and an outperformed method was suggested. The upcoming sections of the paper are organized as follows. Sections 2 and 3 discuss the brief review of different methods applied for crop maturity classification and a detailed explanation of the material and methods used to accomplish this paper respectively. The results and discussion are explored in Sect. 4, and Sect. 5 presents the conclusion part.

2 Related work

Implementation of the automated system for crop maturity assessment has gained the valuable interest of farmers and the food industry in recent decades. These automated systems are time-saving and truncate the overall processing cost when compared with manual labor [9]. Several related works reported in the literature are discussed briefly in the current section. In a study [10], the authors proposed a lucid CNN model to classify wine grapes based on their maturity stages and equate the performance with the pre-trained VGG19 model. Two cultivars of grapes, Cabernet Sauvignon and Syrah for the four and three maturity classes respectively experimented collectively for the small dataset of 1008 images. The average accuracy was found to be 72.66% and 93.41% for the Cabernet Sauvignon and Syrah cultivars respectively. In the next study [11], a working model comprised of hardware and software was built for the grading of date fruits into three categories. Hardware includes the convenor, helm, and camera control systems, while software includes two BPNN classifiers. Hidden and output layers of the first and second BPNN classifier possess identical neurons, but the input layers possess 5 and 2 neurons respectively representing features of date fruits. Out of two models, the second model attained the maximum accuracy of 80%. Authors of [12], proposed an ANN-based system to classify the guava fruit into three maturity classes with high accuracy of 97.44%, utilizing extracted features such as R, G, a, and u from the three distinct colors spaces viz. RGB, CIELab, and CIE-Luv. In the work of [13], a non-intrusive approach was developed to assort the fruits according to their ripeness degree. An amalgamation of four distinct ML methods (i.e., KNN, DT, SVM, and ANN) and some color spaces (i.e., L*a*b*, HSV, and RGB) was analyzed. The model built with SVM and L*a*b* color spaces achieved the best accuracy and f-measures. Paper [14] explores the mechanism for watermelon ripeness detection, based on acoustic signals with ELM and KPCA methods. The KPCA extracts the acoustic signal of watermelon belonging to three

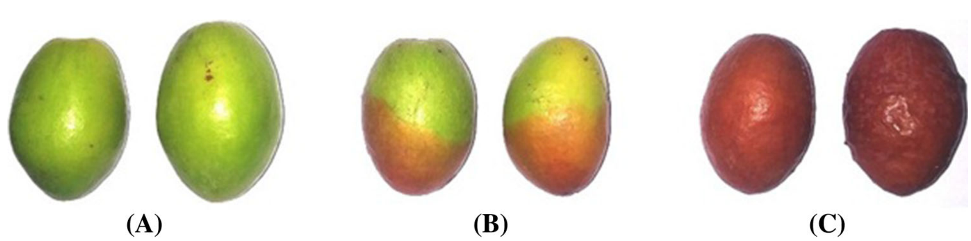


Fig. 1 Jujube fruits maturity classes: (**A**) Unripe, (**B**) Ripe, and (**C**) Over-ripe

maturity classes (i.e., overripe, ripe, and unripe), while the ELM is employed to classify them into respective classes with the accuracy of 92%. In the further two research works, the stratification of mango fruits was presented. In the research [15], a hybrid technique consisting of an odor sensing technique and image processing technique was utilized along with an SVM classifier to classify the mangoes into unripe and ripe classes. The odor sensing and image processing technique were implemented applying array gas sensors and HSV color models respectively to extract the features of mango. The SVM identifies the ripe and unripe classes of mango with a success rate of 94.69%. While in research [16], a machine vision-based system with fuzzy logic was proposed to assort the mango according to their size and maturity level. To extract features from the image samples digital image processing methods were employed, and to examine the features of individual classes for predicting the maturity level GMM was employed. In research [17], the authors proposed the computer vision system employed with ANN to detect the four ripening stages of bananas. The ANN model uses the Tamura statistical texture, brown spots, and color features to grade and classify the banana. Also, the performance of the proposed model was compared with other ML approaches and achieved the highest classification rate of 97.75%. Another recently published article on banana ripeness identification is reported in [18]. The proposed work classifies the four ripening stages of bananas using the CNN model and the results were compared with classical VGG16 and ResNet50 models. The proposed CNN model achieves a high testing accuracy of 96.14%. Paper [19], demonstrates the applicability of a portable four-band sensory device to identify the oil palm maturity level under field conditions. The spectral-reflectance data at four different wavelengths (i.e., 570, 670, 750, 870 nm) represents maturity variation in fruits (i.e., unripe, ripe, and over-ripe). The spectral-reflectance using quadratic discriminant analysis algorithm, and discriminant analysis algorithms with Mahalanobis-distance gives the accuracy above 85%. Authors of literature[20] deployed the enhanced AlexNet architecture for the strawberry ripeness evaluation and compared the performance with classical models. Deployed architecture possesses enhanced convolutional layers with reduced kernel size and uses L2-regularization and BN-layer in place of LRN-layer. Results show that the performance of modified AlexNet architecture enhances by almost 6% compared to original AlexNet architecture. Finally, the authors in the article [21], suggested the different ML and pre-trained CNN models for the papaya maturity classification. The experiment was accomplished for the three maturity classes of papaya with 100 image samples belonging to each class. The ML technique includes three classifiers and three features, where weighted-KNN with histograms of oriented gradients (HOG) achieves the classification accuracy of 100%. And among the seven pre-trained models, VGG19 outperforms with better accuracy of 100%.

3 Materials and methods

The overall mechanism to develop the DL model for jujube fruits classification based on their maturity level is discussed further in detail. The entire process is organized into two vital stages, firstly the materials used in the work, which constitute the jujube image acquisition, pre-processing, image augmentation process, and jujube dataset. Second is the methodology adopted, which constitutes the classification models implementation followed by transfer-learning and dataset train, and test split. Figure 2 displays the overall research framework for the jujube fruits classification.

3.1 Image acquisition

The dataset for the experimental framework, comprising of jujube images was obtained utilizing a color camera (Redmi note 5 pro mobile camera) having 12MP + 5MP dual rear AI-based features capability and pixel size of 1.25 µm. All the jujube images were captured off-field by placing the jujube fruits over a white plain background and capturing it from a distance of 20 cm and angle between 0 and 20 degrees from the top. Other settings of the camera like contrast, saturation, and sharpness were kept in normal mode, camera frameset to 4:3, picture quality set to high, anti-banding was kept at 50 Hz, and auto-exposure was set to center-weighted. To reduce the changing light and shadow effects two daylight bulbs were placed around the object.

3.2 Image pre-processing

The collected dataset contains images of jujube fruits of varying sizes and resolutions. To extract better features, the regions of interest need to be highlighted [22]. Image pre-processing procedure involves manual cropping of images, to highlight the jujube fruits. During the image acquisition phase, images having a resolution above 600 pixels were considered for the final dataset, and the rest were discarded. In that way, vital information for the feature learning in the jujube images was ensured. Furthermore to reduce the training time of CNN models images were rescaled to 300*300 with the help of written code in python, using the OpenCV framework [23].

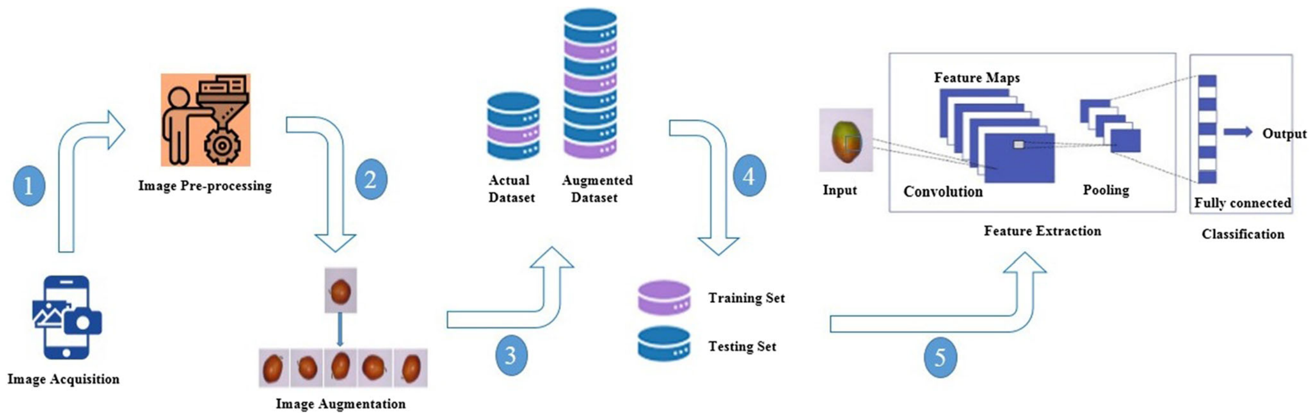


Fig. 2 A research framework for the jujube fruits classification

3.3 Augmentation process

A considerable amount of data is required for the training purpose in the DL models. To enhance the dataset size and acquaint scanty distortion to the images, which assist in diminishing overfitting during the training phase an augmentation process was performed [24]. The augmentation process generates the additional training images through various image processing methods like random rotation, shear, flips, shifts, etc. [25]. To automate the augmentation process for each type of image dataset, python code is written using the OpenCV framework. After the augmentation process, dataset counts for all the three maturity classes of jujube images enhances to almost 10 times.

3.4 Dataset

To train/test, our pre-trained models a small dataset of 400 images was prepared by capturing the actual jujube fruits belonging to three maturity classes. All the collected images were labeled according to their maturity level, such as unripe, ripe, and over-ripe. Unripe jujube seems to be green in color consisting of 146 samples, ripen jujube seems to be yellowish and a bit reddish consisting of 131 samples, while over ripen jujube is completely reddish consisting of 123 samples. After the augmentation process, the volume of the newly formed augmented dataset reaches a total of 4398 images. Where images of unripe, ripe, and over-ripe jujube consist of 1605, 1440, and 1353 images respectively. Table1 shows the actual and augmented dataset count for all three classes of jujube fruits.

3.5 Classification model

3.5.1 AlexNet

AlexNet is quite possibly the most famous deep architecture that depends on CNNs suggested by Alex Krizhevsky in 2012 for the ImageNet Large-Scale Visual Recognition Challenge (ILSVRC) [7]. ILSVRC assesses methods for object recognition and image assortment of 1.2 million images of ImageNet LSVRC-2010 dataset into more than 1000 different classes. The AlexNet model (Fig. 3) comprises of total eight layers, where 5 are convolution layers and the rest 3 are dense layers. Several other layers such as Relu, LRN, pooling, and dropouts are associated with some of the convolution and dense layers. The first convolution layer possesses 96-kernels of size 11*11 with a stride-4 is employed over the input image of size 227*227*3. The second convolution layer possesses 256-kernels of size 3*3 with a single stride. The next two convolution layers and final convolution layers possess 384 and 256-kernels respectively each of size 3*3 with a single stride. Furthermore, the first two dense layers possess 4096 neurons and the final layer possesses 1000 neurons according to ImageNet database classes. In our case, the final dense layer is modified with 3 neurons (i.e., for three maturity classes: unripe, ripe, and over-ripe).

In the AlexNet model, each of the convolution layers and dense layers is linked to the Rectified Linear Unit (ReLu) activation function to append non-linearity for the rapid training of CNN models. Also, a normalization layer (i.e., Local-Response Normalization) is used in the first two convolution layers to normalize the unbounded activation of ReLu. Moreover to extract the vital features and reduce the parameters, and reduce the computation of the network max-pooling layer is concatenated after the first, second, and fifth convolution layers. Finally, the Softmax function

Table 1 Image dataset for maturity classification of jujube fruits

Class	Actual dataset			Augmented Dataset		
	Training dataset	Testing dataset	Training + Testing	Training dataset	Testing dataset	Training + testing
Unripe	102	44	146	1124	481	1605
Ripe	92	39	131	1008	432	1440
Over-ripe	86	37	123	947	406	1353
Total	280	120	400	3079	1319	4398

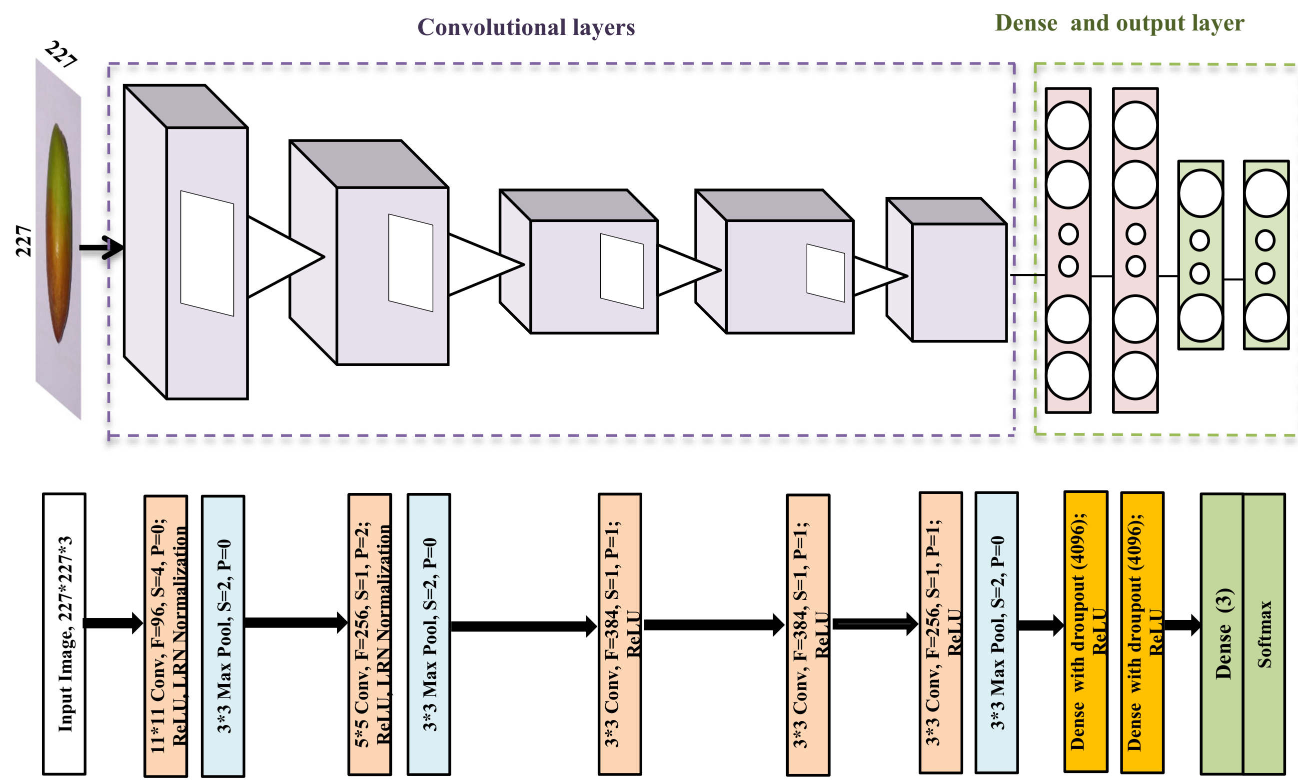


Fig. 3 AlexNet model

is appended to the last dense layer of AlexNet for classification purposes.

3.5.2 VGG16

VGG16 is another popular CNN-based model developed in 2014 by a research group at the University of Oxford, that achieved high accuracy in the ImageNet challenge [8]. The VGG16 model (Fig. 4) comprises of total 16 layers, where the top 13 are convolution layers and the bottom 3 layers are dense. Unlike the AlexNet model, VGG16 doesn't have variations in kernel size, padding, and stride at different convolution and pooling layers. All the convolution layers in the VGG16 possess a fixed filter size of 3*3 with a stride-1 and zero padding, only the number of filters at convolutional layers may vary. The 1st and 2nd convolution layers possess 64 filters, the 3rd, and 4th convolution layers possess 128 filters, the 5th to 7th convolution layers possess 256 filters, and the 8th to 13th convolution layers possess 512 filters. VGG16 operates over the input image of size 224*224*3, whenever an image is passed to the pooling layers image size is reduced to its half due to the fixed size of the filter (i.e., 2*2)and stride (i.e., 2) at the pooling layers. Furthermore, after the convolution layer, the 1st two dense layers possess 4096 neurons and the final dense layer possesses 1000 neurons. Again in our case, the final layer is modified with 3 neurons according to our desired maturity classes. In the VGG16 model, each of the convolution and first two dense layers is linked to the ReLu activation function. And the Softmax function is appended to the last dense layer for classification purposes.

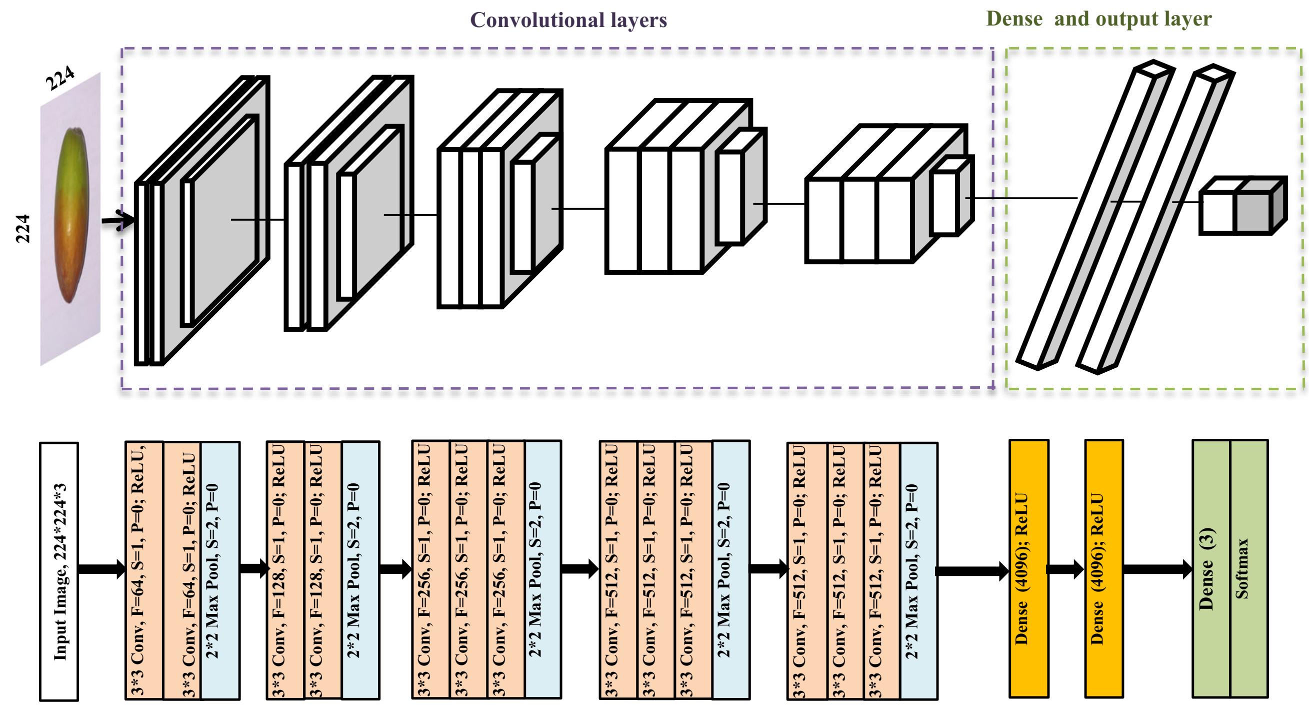


Fig. 4 VGG16 model

3.6 Transfer-learning

The transfer-learning technique utilizes the capabilities of pre-trained CNN models effectively to solve a dissimilar but related problem with the limited dataset. In a transfer-learning-based CNN model, the CNN network can be exercised either as a weight initializer or feature extractor, or both [26, 26]. The CNN models have a vast number of trainable parameters, for an instance, the AlexNet and VGG16 have 60 million and 130 million trainable parameters respectively. Training millions of parameters in an end-to-end CNN model with a limited dataset is a rigorous task, also CNN models can memorize a limited dataset, which gives rise to overfitting. Therefore, to curb the overfitting and avoid the training of millions of parameters transfer-learning was employed in the current study. Furthermore, dataset size was enhanced by applying data augmentation over the training dataset. During the training period, jujube images in each batch were arbitrarily exposed to image augmentation operations such as rotation (with a range of 40), width and height shift (with a range of 0.2), and zoom (with a range of 0.2).

Fine-tuning allows to retrain the weights of dense layers and the feature extractor layers (i.e., convolution and pooling layers), except for the final dense layer. Because the final dense layer needs to be replaced with the desired dataset classes (i.e., three) not the ImageNet classes (i.e., 1000). It's a choice of one to either go for fine-tuning feature extractor layers or keep them unchanged in the course of training [28]. In the current study, we have frozen the feature extractor layers and fine-tuned the dense layers with an adjusted final learning rate of 0.003. We employed dropout layers at the first and second dense layers with a probability of 0.5. The weights of the final dense layer were randomly initialized, and the Softmax layer was appended for the classification task.

3.7 Dataset train and test split

Two types of image datasets for the jujube fruits were used for the training/testing of CNN models. The first dataset contains 400 actual jujube images, and the second dataset contains 4398 augmented jujube images, each belonging to three maturity classes. Table 1 shows the information on three classes of both datasets.

Since there is a small difference in dataset size among each maturity class of both datasets, we can say that the dataset is balanced. And the classification models will possess learning of each maturity class impartially. Therefore, both the datasets are divided into train and test sets separately by randomly bifurcating the 400 and 4398 images, so that 70% constitute the training set and 30% constitute the testing set. The 70/30 split ratio of the train and test dataset is most customarily used in DL applications [29]. Thus for the CNN model training 280 and 3079 images were used from the actual and augmented dataset respectively. While for the CNN models performance testing 120 and 1319 images were retained from the actual and augmented dataset respectively.

4 Results and discussion

The results introduced and talk about in this section are based on the training/testing of the actual and augmented dataset with pre-trained AlexNet and VGG16 models and their comparison with the existing jujube classification methods. The codes for the models were written in python using Keras and TensorFlow libraries. The training/testing processes for the models were performed over the "Tesla K80" GPU with 12 GB RAM as allocated by "Google Colab", which is a cloud-based Jupyter notebook environment. The performance of the models was evaluated with a criterion such as accuracy and average classification time.

4.1 Optimal CNN models

The classification of jujube fruits utilizing pre-trained CNN models as discussed in Sect. 3.5 was trained and tested with the hyper-parameters listed in Table 2. Following several adequate experimentations, these values and algorithms gave the optimal results during training and testing.

The model's performance is evaluated over the actual and augmented testing dataset and achieves good accuracy. Table 3 exhibits the successful classification percentage of jujube fruits during the testing of AlexNet and VGG16 models. The metrics presented in Table 3 are:

(i) The success rate is the percentage of correct classification of jujube classes. This represents the number of accurately classified jujube images over the entire number of jujube images.
(ii) The average loss is the corresponding average error of the models per batch, over the entire batches of the testing dataset.
(iii) The epoch represents the training epoch number at which the best accuracy was achieved.
(iv) The time during the training of models is the average time (in milliseconds) elapsed per epoch.

Table 2 Final model training and testing hyper-parameters

Hyper-parameters	Value/algo
Batch size	32
Optimizer	Adam
Learning rate	0.003
Epochs	50
Dropout	0.5

The results shown in Table 3 indicate that both the models achieved better performance over the augmented dataset of jujube fruits. Firstly AlexNet and VGG16 were trained over the actual dataset for a total of 50 epochs with an average time elapsed per epoch was 5522 ms and 6680 ms respectively. The top accuracy achieved by AlexNet and VGG16 was at epochs 36 and 26 respectively, with the highest success rate of 97.65% (i.e., top-1 error of 2.35%) for VGG16 during testing. Further AlexNet and VGG16 were trained over the augmented dataset and it was observed that the classification success rate for both the models get enhanced, again VGG16 reports the highest success rate of 99.17% (i.e., top-1 error of 0.83%) but this time at epoch 47 with an elapsed time of 65,665 ms per epoch. It has been observed that after 50 epochs no further enhancement in the accuracy was reported till 100 epochs, therefore we limit our training/testing to 50 epochs. Although VGG16 is computationally bit expensive as compared to the Alexnet but it achieved a high classification accuracy which is foremost vital. The training/validation accuracies behavior and training/validation loss behavior for the actual dataset and the augmented dataset are presented in Figs. 5 and 6 respectively.

4.2 Comparision with other existing methods

The performance of the examined classification models is compared with the methods reported in other studies for jujube maturity classification. In the study [30], the authors applied different data mining approaches to classify the actual jujube images into four quality grade classes. A total of 57 features related to shape, texture, and color information was extracted, which was further reduced to 13 features using the co-relation-based feature selection method for the classification of jujube. The validation stage of the four distinct data mining techniques such as Bayesian Networks, DT, SVM, and ANN reported the accuracy of 95.22%, 95.14%, 95.91%, and 98.61% respectively. Here, ANN outperforms the remainder of the data mining methods. In another study [31], the researchers proposed a hyperspectral technique for post-harvest grading of winter jujube maturity. To achieve the jujube visual maturity-sort, the characteristics wavelengths (CW) were obtained and spectral indices (SI) were calculated. Discrimination of jujube maturity classes was performed through three Partial Least Squares Discriminant Analysis (PLS-DA) models, where two based on the CW's determined with successive projection algorithm (SPA) and competitive adaptive reweighted sampling (CARS) and one based on SI's. The discrimination results show that PLS-DA model based on SI's achieves the highest accuracy of 98.18 followed by the

Table 3 Results of pre-trained CNN models for the classification of jujube on the testing dataset

Actual dataset					Augmented dataset			
Model	Success rate (%)	Average loss	Epoch	Time (ms/epoch)	Success rate (%)	Average loss	Epoch	Time (ms/epoch)
AlexNet	94.17	0.1875	36	5522	98.26	0.1551	33	57,590
VGG16	97.65	0.1050	26	6680	99.17	0.1036	47	65,665

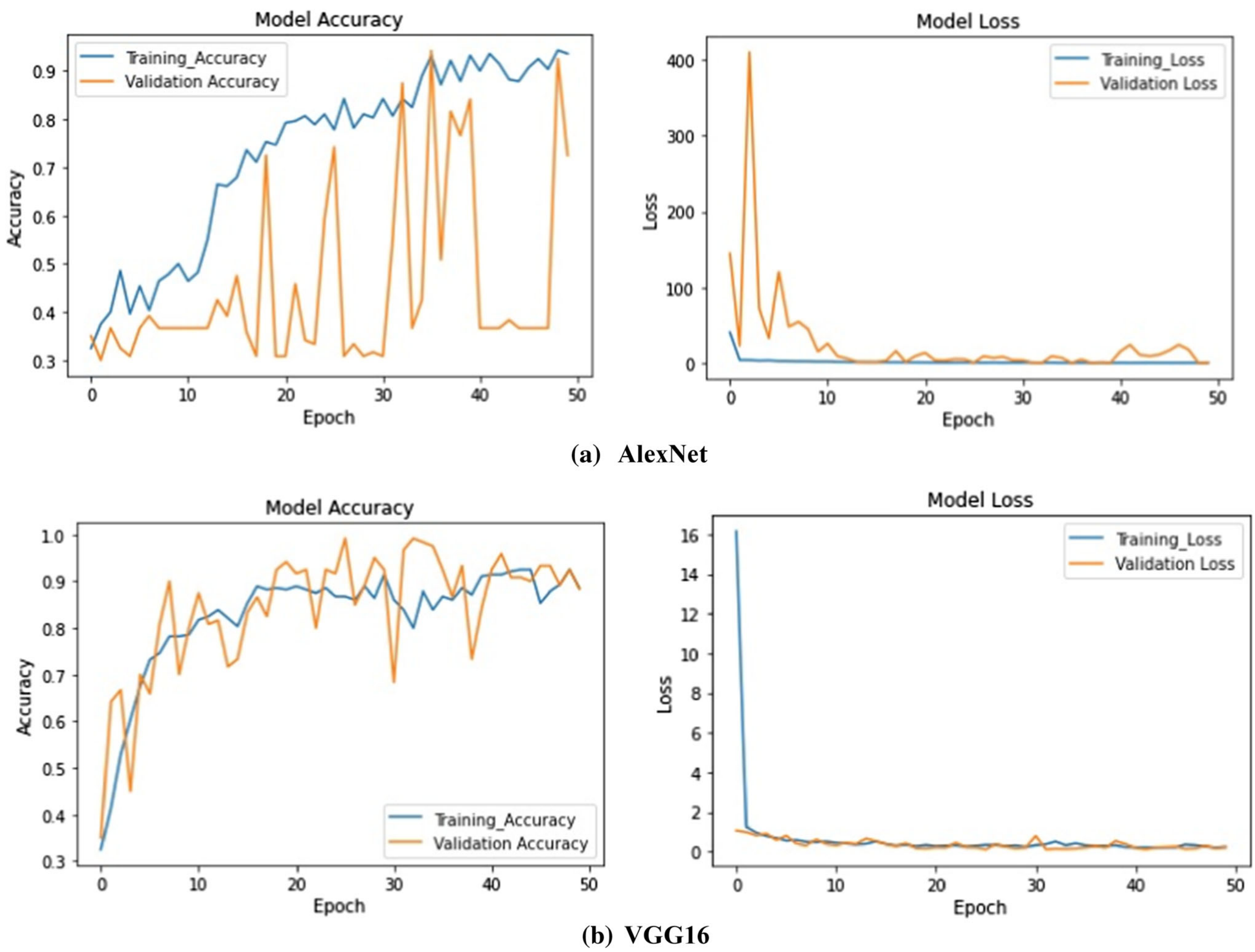
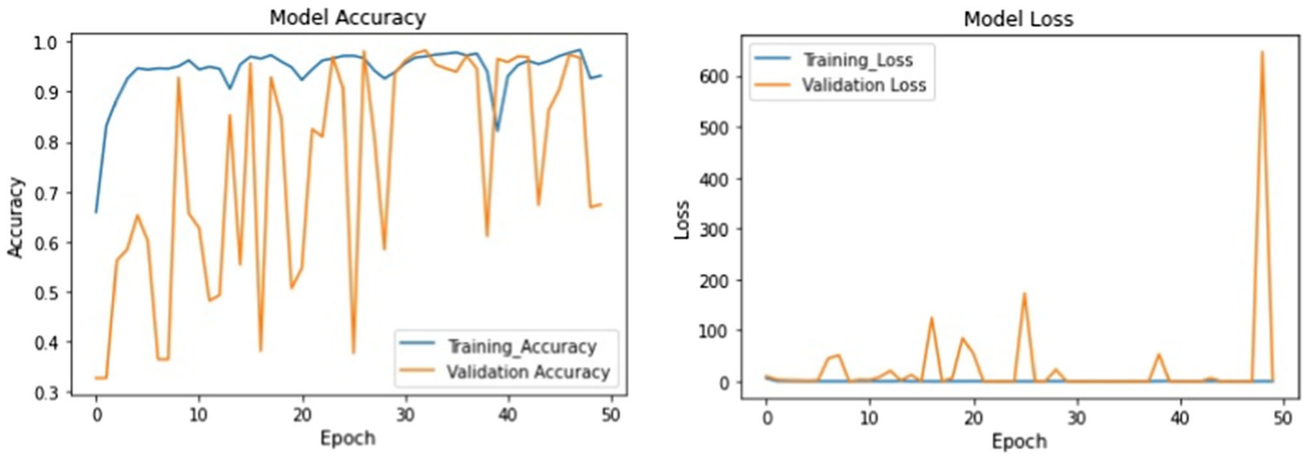


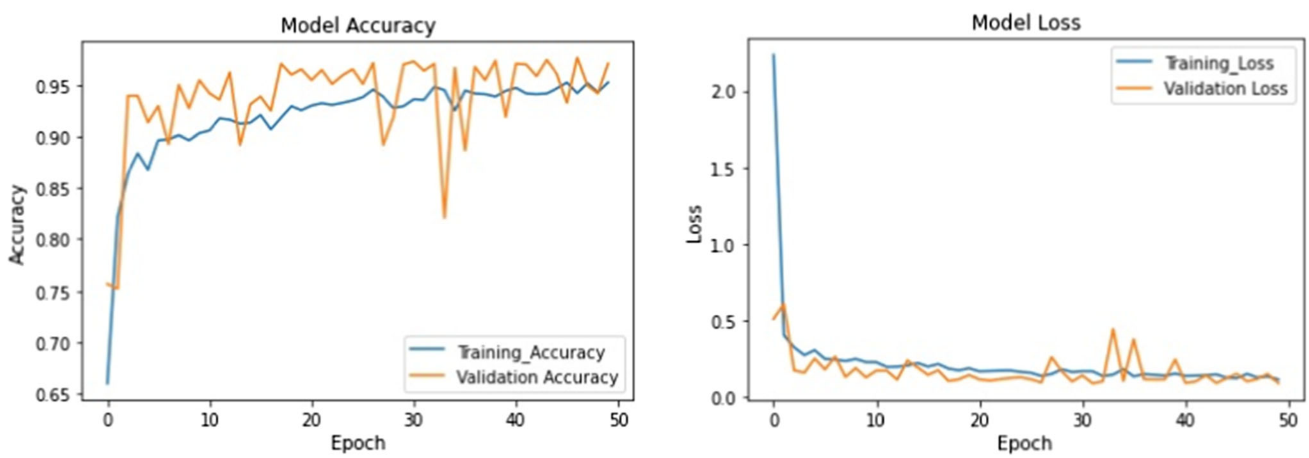
Fig. 5 Training/validation accuracies behavior and training/validation loss behavior of jujube classification models for the actual dataset

PLS-DA model based on CW-SPA and CW-CARS with the accuracy of 97.27% and 95.45% respectively.

The comparison of the current study with the above-discussed studies based on dataset type, dataset size, classes, and accuracy is shown in Table 4. It can be observed that the performance of the AlexNet classification model for the augmented dataset is slightly better than the method reported in [30], but fails to beat the method reported in [31]. While the performance of the VGG16 classification model for the augmented dataset checkmates all the compared classification methods/models by achieving high accuracy of 99.17%. It has been noticed that the performance of AlexNet and VGG16 for the actual dataset falls short of both the compared studies. This might be due to the small number of the actual dataset, as the performance of CNN models relies upon a sufficient training dataset. Also, the standard accuracy of models in the current study and methods reported in other studies were computed and compared. The standard accuracy of both models in the current study is computed through mean accuracy achieved over an actual and augmented dataset. While the standard classification accuracy of the method reported in other



(a) AlexNet



(b) VGG16

Fig. 6 Training/validation accuracies behavior and training/validation loss behavior of jujube classification models for the augmented dataset

Table 4 Performance comparison for jujube classification methods

Method	Dataset type	# of the testing dataset	# of classification classes	Accuracy (%)
AlexNet	Single jujube (actual images)	120	3	94.17
VGG16	Single jujube (actual images)			97.65
AlexNet	Single jujube (augmented images)	1319	3	98.26
VGG16	Single jujube (augmented images)			99.17
[30]	Single jujube (actual images)	40	4	98.61
[31]	Single jujube (actual images)	110	3	98.18

studies is computed by taking all the classification accuracies reported in both the study individually and computing their mean. Figure 7 represents the standard classification accuracy achieved by the examined models and by the methods in [30] and [31]. The VGG16 model reported the highest standard accuracy of 98.41% followed by the method in [31] with 96.96%. While the standard accuracy of AlexNet and the method in [30] is almost the same and falls short of the method in [31]. Hence, the overall performance of the VGG16 model shows that in all the scenarios it achieved a promising outcome with at least 97.61% classification accuracy.

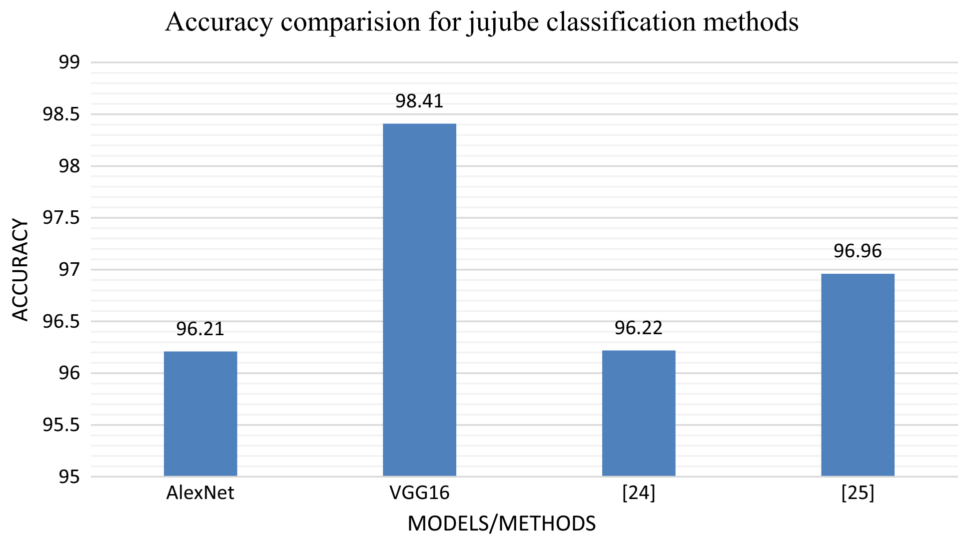


Fig. 7 Accuracy comparison between proposed jujube classification approach and approaches in [30] and [31]

5 Conclusion

In this study, two famous pre-trained DL models (i.e., AlexNet and VGG16) which got great success in the 'ImageNet' challenge were investigated for the classification of jujube fruits according to their maturity level. The idea of transfer learning was employed for the classification of jujube into three maturity classes. The training of the models was firstly performed over the collected dataset of 280 actual jujube images and secondly, over the dataset of 3079 augmented jujube images, each belonging to three different maturity classes. Out of the two examined models, VGG16 achieved a very high classification accuracy of 99.17% (i.e., top-1 error of 0.83%) and 97.65% (i.e., top-1 error of 2.35%) for the augmented and actual dataset respectively. In addition to this, the performance of examined models was compared with the approaches reported in the other two studies. Again VGG16 outperforms the AlexNet and other approaches of the studies in terms of standard classification accuracy by achieving high accuracy of 98.41%. Results of the study demonstrate that utilizing the transfer-learning approach over the enhanced dataset (i.e., augmented dataset) gives better accuracy than that of the small dataset (i.e., actual dataset). The significantly better success rate of VGG16 models over the other models discussed in the study makes the model recommendable for developing an efficient system for the automated harvesting and sorting of jujube fruits.

Furthermore, future work will involve the improvement of the dataset by capturing actual images from different jujube trees and including one more maturity class in the current study. We will also examine some other CNN models to enhance the classification accuracy and minimize the computational complexity.

Acknowledgements This research work received no specified grant funding from any funding agency, commercial, or not-for-profit sectors. The authors express their gratitude in advance to the reviewers for their significant suggestions and comments.

Author contributions A.M contemplated the structure of the article, performs the experimental works, and accomplished the paper writing. AKT and SKS were involved in the suggestions and final revision of the article. All authors have read and approved the manuscript.

Declarations

Conflict of interest The authors declare no conflict of interest.

References

1. Pareek S (2013) Nutritional composition of jujube fruit. Emirates J Food Agric 25(6):463–470. https://doi.org/10.9755/ejfa.v2516.15552
2. Rashwan AK, Karim N, Rezaul M, Shishir I, Bao T, Lu Y (2020) Jujube fruit: apotential nutritious fruit for the development of functional food products. J Funct Foods 75(866):104205. https://doi.org/10.1016/j.jff.2020.104205
3. Magwaza LS, Tesfay SZ (2015) A review of destructive and non-destructive methods for determining avocado fruit maturity. Food Bioprocess Technol 8(10):1995–2011. https://doi.org/10.1007/s11947-015-1568-y
4. Li B, Lecourt J, Bishop G (2018) Advances in non-destructive early assessment of fruit ripeness towards defining optimal time of harvest and yield prediction–a review. Plants 7(1):1–20. https://doi.org/10.3390/plants7010003
5. Bhargava A, Bansal A (2021) Fruits and vegetables quality evaluation using computer vision: a review. J King Saud Univ-Comput Inf Sci 33(3):243–257. https://doi.org/10.1016/j.jksuci.2018.06.002
6. Zhu L, Spachos P, Pensini E, Plataniotis KN (2021) Deep learning and machine vision for food processing: a survey. Curr Res Food Sci 4:233–249. https://doi.org/10.1016/j.crfs.2021.03.009
7. Krizhevsky BA, Sutskever I, Hinton GE, ImageNet classification with deep convolutional neural networks, 2012.

8. Simonyan K, Zisserman A (2015) Very deep convolutional networks for large-scale image recognition. ICLR 2015:1–14
9. Mahendran R, Gc J, Alagusundaram K (2012), Application of computer vision technique on sorting and grading of fruits and vegetables, 1–7, https://doi.org/10.4172/2157-7110.S1-001.
10. Ramos RP et al (2021) Non-invasive setup for grape maturation classification using deep learning. J Sci Food Agric 101(5):2042–2051. https://doi.org/10.1002/jsfa.10824
11. Ohali YA (2011) Computer vision based date fruit grading system: design and implementation. J King Saud Univ-Comput Inf Sci 23(1):29–36. https://doi.org/10.1016/j.jksuci.2010.03.003
12. Lara-espinoza EM, Trejo-duran M, Lizarraga-morales RA (2016) Determination of the ripeness state of guavas using an artificial neural network. Res Comput Sci 121(1):105–111. https://doi.org/10.13053/rcs-121-1-9
13. Bazán K, Avila-george H, Member S (2019), Classification of cape gooseberry fruit according to its level of ripeness using machine learning techniques and different color spaces, 27389–27400, https://doi.org/10.1109/ACCESS.2019.2898223.
14. Xu Z, Yuan P, Zhang Y, Deng X (2019) Watermelon ripeness detection via extreme learning machine with kernel principal component analysis based on acoustic signals. Int J Pattern Recognit Artif Intell. https://doi.org/10.1142/S0218001419510029
15. Mavi MF, Husin Z, Ahmad RB, Yacob YM, Farook RSM, Tan WK (2019) Mango ripeness classification system using hybrid technique. Indones J Electr Eng Comput Sci 14(2):859–868. https://doi.org/10.11591/ijeecs.v14.i2.pp859-868
16. Nandi CS, Tudu B, Koley C, Machine vision based techniques for automatic mango fruit sorting and grading based on maturity level and size, In: *Sensing technology: current status and future trends II,27 smart sensors, Measurement and instrumentation*, no. January, Springer International Publishing Switzerland, 2014.
17. Mazen FMA, Nashat AA (2019) Ripeness classification of bananas using an artificial neural network. Arab J Sci Eng 44:6901–6910. https://doi.org/10.1007/s13369-018-03695-5
18. Saranya N, Srinivasan N, Kumar SKP (2021) Banana ripeness stage identification: a deep learning approach. J Ambient Intell Humaniz Comput. https://doi.org/10.1007/s12652-021-03267-w
19. Mohammed O et al (2012) Classification of oil palm fresh fruit bunches based on their maturity using portable four-band sensor system. Comput Electron Agric 82:55–60. https://doi.org/10.1016/j.compag.2011.12.010
20. Ni J, Gao J, Li J, Yang H, Hao Z, Han Z (2021) E-AlexNet: quality evaluation of strawberry based on machine learning. J Food Meas Charact. https://doi.org/10.1007/s11694-021-01010-9
21. Kumari S, Kumar A, Kumar P (2020), Maturity status classification of papaya fruits based on machine learning and transfer learning approach, *Inf. Process. Agric.*, no. xxxx, https://doi.org/10.1016/j.inpa.2020.05.003.
22. Tabik S and Peralta D (2017), A snapshot of image pre-processing for convolutional neural networks: case study of MNIST, 10, 555–568.
23. Solem JE, *Programming computer vision with python*. 2012.
24. Hawkins DM (2004) The problem of overfitting. J Chem Inf Comput Sci 44(1):1–12. https://doi.org/10.1021/ci0342472
25. Shorten C, Khoshgoftaar TM (2019) A survey on image data augmentation for deep learning. J Big Data. https://doi.org/10.1186/s40537-019-0197-0
26. Sharif A, Hossein R, Josephine A, Stefan S and Royal KTH, CNN features off-the-shelf: an astounding baseline for recognition."
27. Salvador A, Zeppelzauer M, Manch D, Calafell A, Politecnica U and Upc DC, Cultural event recognition with visual convnets and temporal models.
28. B. C. B, V. M. B, O. Beijbom, Hoffman J and Darrell T (2016), Best practices for fine-tuning visual classifiers, pp. 435–442, https://doi.org/10.1007/978-3-319-49409-8.
29. Roshan JV, Akhil V (2021) SPlit: an optimal method for data splitting. Technometrics. https://doi.org/10.1080/00401706.2021.1921037
30. Rahimi A, Banakar A and Zareiforoush H (2014), Classification of jujube fruits using different data mining methods, no. January. 52–61.
31. Cao Xiao-feng SB, Hui-ru REN, Xing-zhi LI, Ke-qiang YU (2018) Discrimination of winter jujube's maturity using hyperspectral technique combined with characteristic wavelength and spectral indices. Spectrosc Spectr Anal 38(7):2175–2182. https://doi.org/10.3964/j.issn.1000-0593(2018)07-2175-08