Journal of Ambient Intelligence and Humanized Computing (2022) 13:4033–4039
https://doi.org/10.1007/s12652-021-03267-w

ORIGINAL RESEARCH

Banana ripeness stage identification: a deep learning approach

N. Saranya[1] · K. Srinivasan[1] · S. K. Pravin Kumar[2]

Received: 15 May 2020 / Accepted: 8 April 2021 / Published online: 3 May 2021

Abstract

In recent days, deep learning has been considered as the state-of-the-art computer vision technique for image classification task. The introduction of Convolutional Neural Network (CNN) made the feature engineering task simple. The classification of various stages of maturity of a fruit is a challenging task using machine learning techniques as it is hard to differentiate the visual feature of the fruits at different maturity stages. In this proposed work, four different ripeness stage of banana were classified using proposed CNN model and compared with the state-of-the-art CNN model using transfer learning. Classification using CNN model requires a huge number of training images to achieve better classification result. The proposed CNN model was trained and tested with both original and augmented images. The CNN model was trained with overall validation accuracy of 96.14%.

Keywords Deep learning · Feature engineering · CNN · Data augmentation

1 Introduction

India is the largest producer of banana in the world and nearly 10% of bananas were used in food processing industry. Banana consumption provides good nutrition to human body which provides high quantity of potassium, serotonin and iron content. Owing to control blood pressure, prevent depression and also highly recommended for anemic patients (Singh et al. 2018). It's also being used as supplementary semisolid food for infants (Bindu et al. 2019). Merely 17 different products are obtained from banana which includes fried chips, candy, puree, pulp, wafers, beer, banana powder, etc.…Ripeness of banana is a great concern in making these products. Right choice of ripeness stage is mandatory in these food processing industries. Quality ensures branding of these products and also their industries. Unripe and partially ripe banana pulp slice were used as raw product in fried bananas chips industry. Ripe and unripe banana were the raw materials for producing banana flour. Ripe bananas were used in preparation of banana jam and banana puree. Over ripe bananas were also used in the preparation of banana jelly (Singh et al. 2018). Grading banana based on its ripeness stage can be done with the help of computer vision based or machine vision system.

Machine vision-based system is one of the best mechanisms to assess the quality of raw materials used in food industry. It is a non-destructive method that can be used in postharvest inspection in tinned food manufacturing industries or packinghouses. Machine vision system was designed for fruit and vegetable inspection to detect defect and grading. It was done based on its size, weight and appearance using the hyper spectral and multispectral images (Cubero et al. 2016).

Digital color images have also been used along with image processing and machine learning algorithm in automatic grading system to detect immaturities. Such systems use image processing techniques for background removal and for identification of essential features (Li et al. 2016). This was the challenging part in designing computer vision-based inspection or grading system. Before the acceleration of Deep Neural Network and Deep Learning techniques, performance of vision based system highly relies on input image, image processing techniques, machine learning

✉ N. Saranya
saranya.pravin@srec.ac.in

K. Srinivasan
hod-eie@srec.ac.in

S. K. Pravin Kumar
skpk.87@gmail.com

1 Sri Ramakrishna Engineering College, Coimbatore, Tamil Nadu, India

2 United Institute of Technology, Coimbatore, Tamil Nadu, India

algorithms and Neuro Fuzzy systems (Athiraja and Vijayakumar 2020; Behera et al. 2020).

In this digital era, computer vision-based applications gain its fame due to advancement in high power GPU systems and also availability of huge amount of digital data. This accelerates the implementation of Deep Learning in computer vision-based applications. It is widely used in the areas like image classification, recognition, segmentation, speech recognition etc. Many research experiments have been conducted to find the ripe state of different fruits using Computer vision, Image processing, Machine Learning and Deep Learning (Kamilaris and Prenafeta-boldú 2018).

Computer vision-based applications could be designed to categories the banana based on its ripeness stage. It will help to minimize the human involvement and also in automating the maintenance process in food industries and cold storage houses. It is one of the non-destructive methods to find the ripeness stage from the image of the fruit. Image acquisition done using digital cameras, laser light backscattering imaging (LLBI) technique (Piedad et al. 2018; Thor 2017; Adebayo et al. 2017a, b; Adebayo et al. 2017a, b). Identifying significant features and its extraction plays a major role in building quality classifiers. Major features like RGB value of the image, shelf life of the banana fruit, Hue, saturation, intensity (HSI) values derived from RGB values were considered as parameters for ripeness prediction (Piedad et al. 2018; Thor 2017). Along with color features some other parameter like tamura statistical texture features were used in combination with image processing techniques like morphological filtering and segmentation. Banana shelf life were also been calculated with the ripeness stage identification along with clustering techniques (Mazen and Nashat 2019). Another challenge in designing computer vision system is the selection of suitable machine learning algorithms. Machine learning algorithms like SVM, PCA-SVM, ANN, Niave-bayes, Decision tree, K-Neighbor and random forest were used as classifiers (Piedad et al. 2018; Thor 2017; Adebayo et al. 2017a, b; Adebayo et al. 2017a, b; Mazen and Nashat 2019; El-Bendary et al. 2015). Ripeness stage classification, elasticity, Soluble Solids Content, and chlorophyll content could be predicted with the help of machine learning classifiers (Adebayo et al. 2017a, b; Mazen and Nashat 2019; El-Bendary et al. 2015).

In traditional machine learning pipeline feature engineering plays a major role. Major research work and experimentation should be done to identify suitable features for different application domain. Using machine learning for image classification involves feature extraction to be pipelined with classifier model. Computer vision related applications need each pixel data to be considered and all the features should be manually analyzed to achieve better performance result. This is time consuming process when compared to deep learning methodologies where feature extraction and classification done as a single process. So Deep learning was used widely for classification and ripeness prediction process. The survey explains different implementation of Deep learning techniques in different agriculture-based applications (Kamilaris and Prenafeta-boldú 2018).

CNN architecture along with triplet loss indicator was used to find the ripeness of the banana. CNN act as both feature extractor and classifier. The accuracy achieved through this architecture is 92.4% after 34 epochs (Zhang et al. 2018a, b). Deep learning model could learn better when there is a huge number of images provided for training. Data augmentation is a technique which is used to increase the number of images in training dataset through various techniques. It duplicates the image by shifting, flipping, rotating, brightening and zooming in and out the training images. Data overfitting in deep learning model could be ruled out using either adding the dropout layer or using data augmentation techniques. Deep learning models could learn features from the image in spite of its position in the image. Huge amount of data could be used to avoid data overfitting. In this, author discussed about various existing augmentation technique and proposed a new technique to generate images of different style using Generative Adversarial Network (GAN) (Perez and Wang 2017).

Data augmentation techniques like horizontal and vertical flipping of images were applied to create Fruit-360 dataset. CNN architecture was designed to classify images of 120 fruits and vegetables with maximum of 95.23% accuracy (Mureşan and Oltean 2018).Data augmentation methods used in this work was random noise and geometric transformation i.e. scaling and rotation. Types of random noise used in this work are pepper, salt and gaussian. Accuracy achieved in this work is 91.9% (Zhang et al. 2018a, b). The data augmentation techniques like flipping, zooming and rotation were randomly applied over the input dataset for fruit classification (Shamim Hossain et al. 2019).

Design of State-of-the-art architectures along with activation function and optimization techniques made deep learning better solution for computer vision-based applications. VGGNET-16consists of 16-convolution layer and works with input image of size $224 \times 224 \times 3$. This network uses stack of convolution layer with a filter of size 3×3 in each layer with padding and stride as 1.ReLU (Rectified Linear Unit) is used as activation function at each convolution layer. Performance of this network is better on image classification and localization task (Simonyan and Zisserman 2014).

Deep learning-based model was used to classify images of 10 different fruit categories. It uses VGG-16 and a simple CNN architecture as classifiers and obtained accuracy of 99.49 and 99.75 respectively (Shamim Hossain et al. 2019). The proposed work uses images of fruit obtained in two forms i.e. color (RGB) and Near infrared (NIR). It intends to find the number of fruits in the tree image. Region based CNN (RCNN) was used to find the fruits from the tree image with annotating bounding box (Sa et al. 2016). A CNN model was to classify the ripening stages of date fruit. Four categories of dates considered are khalal, rutab, tamar and defective. It uses transfer learning i.e. VGG-16 CNN model has been used as a classifier and achieved accuracy of 96.98% (Nasiri et al. 2019). CNN based architecture with 5 convolution, 2 pooling and a fully connected layer to predict the ripeness stage of tomato (Zhang et al. 2018a, b).

2 Dataset

Four different ripeness stage of banana have been considered for classification. They are Stage1: Ripe, Stage2: Partially ripe, Stage3: Ripen and Stage4: Over-ripe. Different ripeness stages of banana were considered from (Mazen and Nashat 2019). There are 104 images in the stage1, 48 images in stage2, 88 images in Satge3 and 33 images in the Stage4 (Fig. 1).

3 Proposed system

The proposed work intends to build a simple CNN model for predicting the ripeness stage of banana. It was compared with state of art modelsVGG16 and ResNet50. The workflow of the proposed work is given in the Fig. 2. Original dataset contains totally 273 images comprising images of four different stages. 80% of the data in each stage were considered for training and 20% of the data were considered for testing. Data augmentation was applied to enlarge the dataset size and the distribution is given in Fig. 3.One hot encoding was considered to indicate the different ripeness stage. Stage1—Unripe was indicated as [1000], stage2—partially

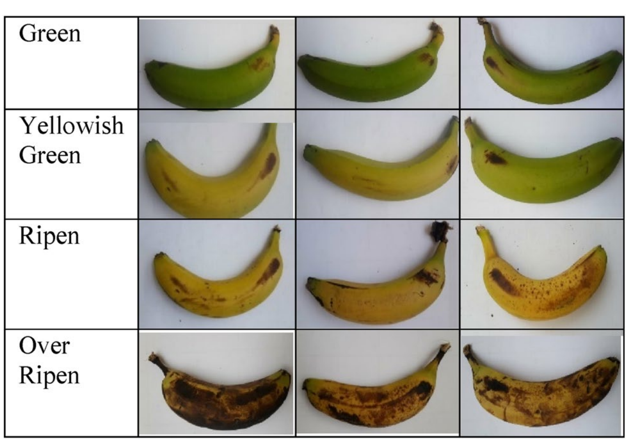


Fig. 1 Sample images in Banana Ripeness Dataset

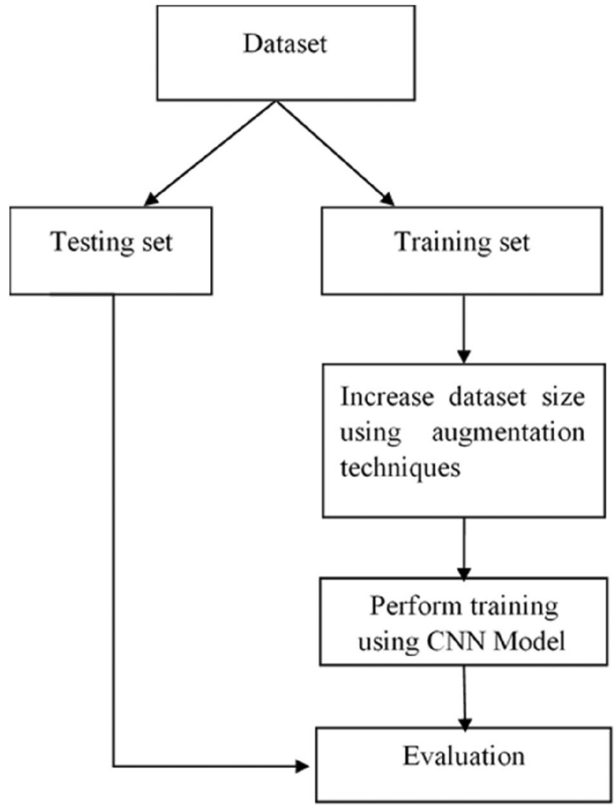


Fig. 2 The workflow of the proposed work

ripe as [0100], stage3—ripe as [0010] and stage4—overripe as [0001].

The data augmentation techniques were used to increase the number of images in training dataset through various techniques. It duplicates the image by shifting, flipping, rotating, brightening and zooming in and out the training images. Sample augmented image of a single original image using rotation, vertical and horizontal flipping technique. Figure 4a represent the sample augmented data using rotation techniques and Fig. 4b represent result of sample augmented data using vertical and horizontal flipping.

A simple CNN architecture was used to classify the ripeness stage at two different scenarios. In our proposed work, a CNN model was designed to categories the different ripeness stages of banana. First experiment was conducted to

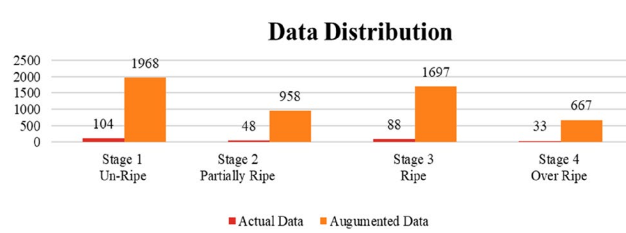


Fig. 3 Data Distribution

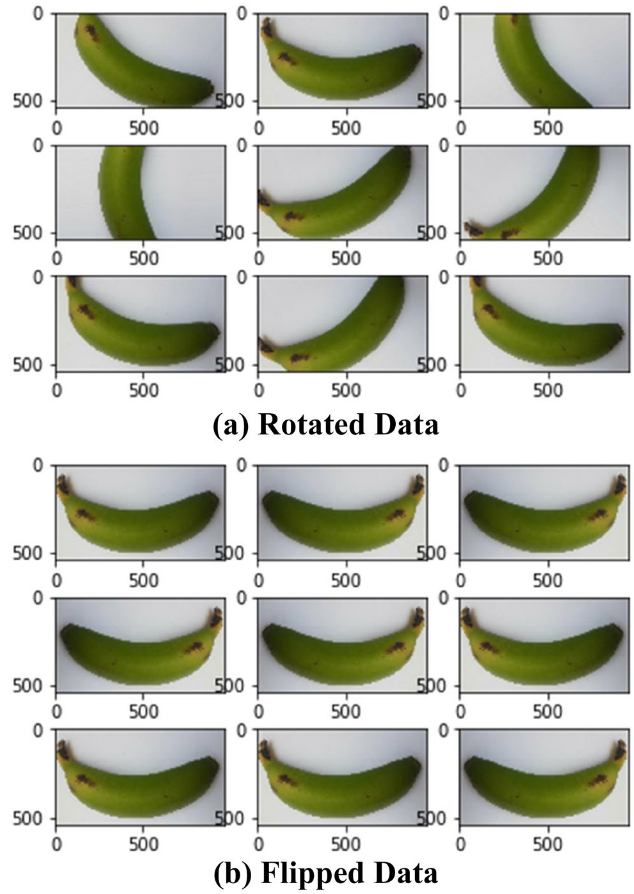


Fig. 4 **a** Rotated data. **b** Flipped data

compare the performance of our simple CNN model with the existing state-of-art architecture, VGGNet16 and ResNet50 as transfer learning. Training accuracy, validation accuracy and time taken to build the model were the performance measures considered to compare the model efficiency. Convolutional layer is used extract features based on the filters across the image. Feature map is created as an output of convolutional layer. Activation functions are used to convert the net input into activations. Most used activation function in CNN are ReLU(Rectified Linear Unit) at hidden layer and softmax at fully connected layer. ReLU is given as

$$f(x) = \max(x, 0) \tag{1}$$

Its output is equal to the x when it is greater than 0, otherwise output is 0. SoftMax activation is used to output categorical probability distribution of the given data. It is given as

$$f(x_i) = \frac{e^{x_i}}{\sum_{j=0}^{k} e^{x_j}} \tag{2}$$

Another experiment was conducted to check the accuracy of the system with original data and augmented data with Adam optimizer for network optimization. System performance is compared based on accuracy and loss percentage throughout the epoch. Evaluation measures used to compare the system performance are accuracy and cross entropy loss with the original and augmented dataset.

$$\text{ACCURACY} = \frac{\text{No of samples predicted correctly}}{\text{Total no of Samples}} \tag{3}$$

$$\mathit{Categorical\ Cross\ Entropy} = -\sum_{i}^{c} t_i \log s_i \tag{4}$$

$$\text{Precision} = \text{TP}/((\text{TP} + \text{FP})) \tag{5}$$

$$\text{Recall} = \text{TP}/((\text{TP} + \text{FN})) \tag{6}$$

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \tag{7}$$

CNN model used in the experiment is specified below.

Layer (type)	Output Shape	Param #
conv2d_1 (Conv2D)	(None, 62, 62, 32)	896
conv2d_2 (Conv2D)	(None, 58, 58, 16)	12816
conv2d_3 (Conv2D)	(None, 54, 54, 16)	6416
max_pooling2d_1 (MaxPooling2	(None, 27, 27, 16)	0
conv2d_4 (Conv2D)	(None, 25, 25, 16)	2320
conv2d_5 (Conv2D)	(None, 21, 21, 16)	6416
max_pooling2d_2 (MaxPooling2	(None, 10, 10, 16)	0
dropout_1 (Dropout)	(None, 10, 10, 16)	0
flatten_1 (Flatten)	(None, 1600)	0
dense_1 (Dense)	(None, 16)	25616
dense_2 (Dense)	(None, 4)	68

4 Experimental result and discussion

In the first experiment, model training was performed with the augmented data over the proposed CNN model and state of the art CNN models. State of art CNN networks considered for this experiment are VGGNet16 and ResNet50. Transfer learning was done to train the model for ripeness stage identification. The different measures considered for comparing the performance of the networks are number of parameters, time taken to train the model, training and validation accuracy. Measures of these parameters during the training process were given in Table 1. Time taken to train the VGGNet16 and ResNet50includes only the training time at fully connected layer. The experiment was conducted using CPU machine with Intel Core i5 processor and 8 GB RAM. Time measure includes only the time taken to train the model (Table 1).

Network was trained for 50 epochs. Results clearly show that (Fig. 5) even though the training accuracy of the VGGNet16 and ResNet16 is high, the validation accuracy of the proposed network outperforms the existing one. State of the art networks has been unsuccessful in classifying the stage 2 and stage 4 ripeness categories. It requires the network to be retrained with new weight values over the training dataset. Re-training the network is time consuming process and it requires many parameters to be trained.

Another experiment was conducted to compare the performance of the proposed model with original and augmented dataset. Results shows that (Fig. 6) validation accuracy of the proposed model over augmented dataset was high when compared to the validation accuracy of the model over original dataset. Detailed comparison of the result was given in the Tables 2, 3, 4 after 50 epochs of training.

5 Conclusion and future work

Four different stages of ripeness were classified using CNN model. Proposed CNN model was trained for 50 epochs and the network performance was compared with of the state-of-the-art models namely VGGNet16 andResNet50. Number of parameters used and time taken to train the proposed network was low. The performance of the CNN model over actual and augmented data was also experimented. Four different ripeness stage of banana was classified using proposed CNN model with better accuracy using augmented data over original data. Time taken to train the proposed model is comparatively low.

This work could be extended to classify different varieties of banana and its maturity stages along with defects detection. Proposed model could be developed as mobile

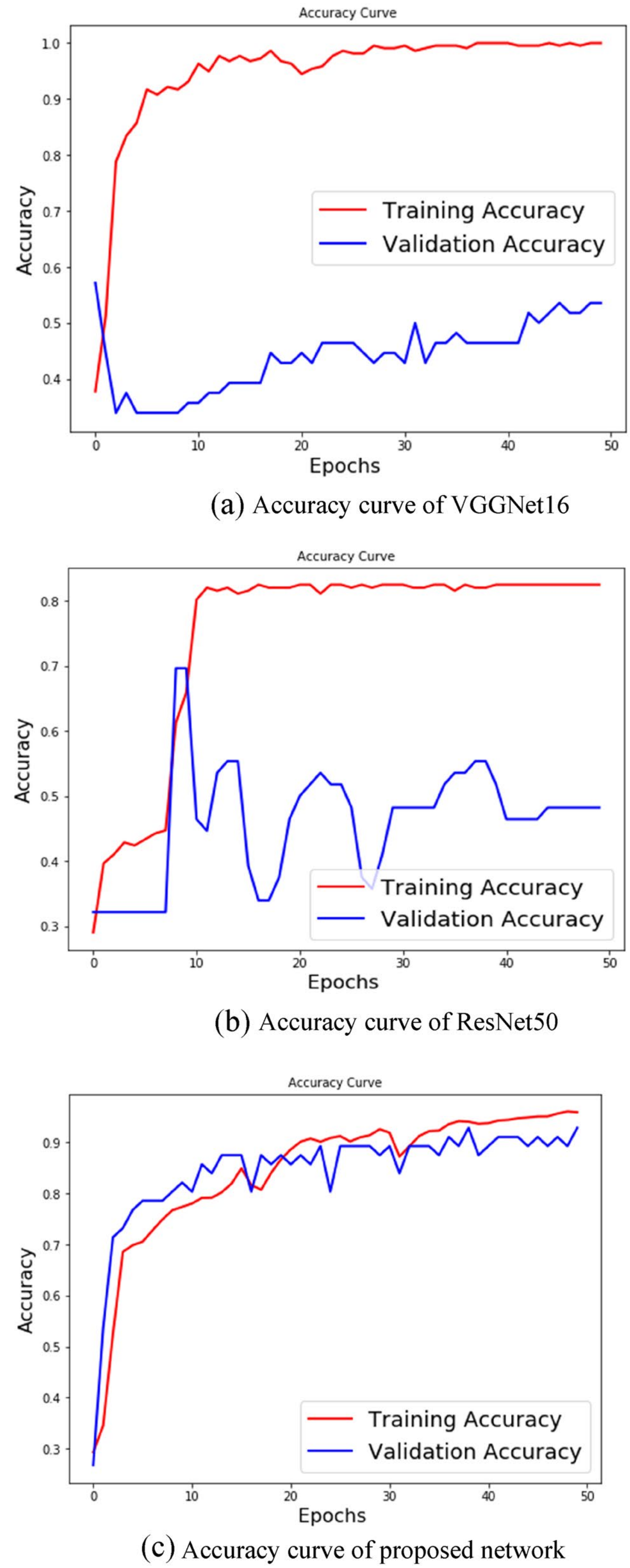
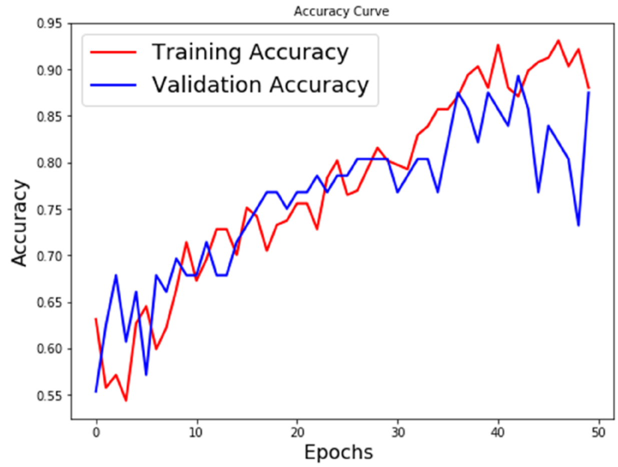
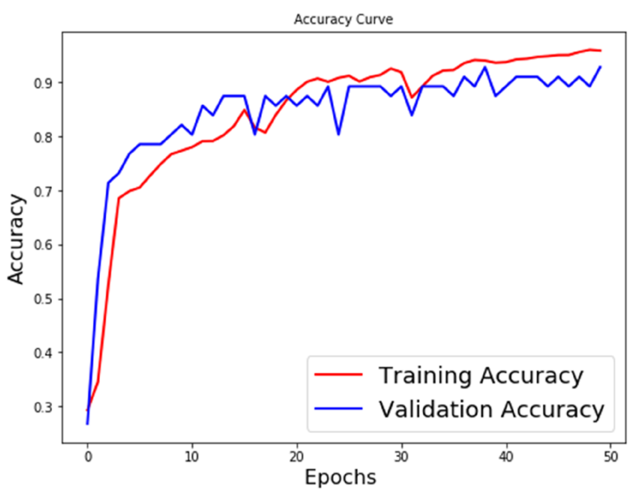


Fig. 5 **a** Accuracy curve of VGGNet16. **b** Accuracy curve of ResNet50. **c** Accuracy curve of proposed network



(a)With Actual Dataset



(b)With Augmented Dataset

Fig. 6 **a** With actual dataset. **b** With augmented dataset

Table 1 Performance measures

Models	1	2	3	4
VGG16	1,48,15,044	6666.57	99.12	53.4438
ResNet50	2,39,89,124	4418.91	82.20	48.2142
Our network	73,764	5140.41	92.8571	96.1436

1. # of Parameters (in million)
2. Time taken totrain the model (in seconds)
3. Training accuracy
4. Validation accuracy

application to identify the ripeness stage from the real time data in the banana field.

Acknowledgement This research work was supported and carried out at the Department of Information Technology, Sri Ramakrishna Engineering College, Coimbatore. We would like to thank our Management, Principal and Head of the Department for supporting us with the infrastructure and learning resource to carry out the research work.

Table 2 Performance comparison

Metrics	Original dataset	Augmented dataset
Training accuracy	94.93	**96.14**
Training loss	13.97	8.37
Validation accuracy	87.5	**92.85**
Validation loss	44.79	17.83
Training time (in seconds)	102	5140

Bold indicates improvement in accuracy

Table 3 Model comparison over actual dataset

Ripeness stage	Actual dataset		
	Precision	Recall	F1 score
Unripe	0.95	1.00	0.98
Partially ripe	0.75	0.60	0.67
Ripe	0.80	0.89	0.84
Over ripe	1.00	0.86	0.92

Table 4 Model comparison over augmented dataset

Ripeness stage	Augmented dataset		
	Precision	Recall	F1 score
Unripe	1.00	1.00	1.00
Partially ripe	0.82	0.90	0.86
Ripe	0.89	0.89	0.89
Over ripe	1.00	0.86	0.92

Data availability Dataset available at https://drive.google.com/drive/folders/1nRWBYAHNRqmL4R0SLrs6dbGQFSWGVY8V.

References

Adebayo SE, Hashim N, Abdan K, Hanafi M, Zude-Sasse M (2017a) Prediction of banana quality attributes and ripeness classification using artificial neural network. Acta Hort 1152:335–343

Adebayo SE, Hashim N, Abdan K, Hanafi M, Zude-Sasse M (2017b) Banana quality attribute prediction and ripeness classification using support vector machine. ETP Int J Food Eng 3(1):42–47

Athiraja A, Vijayakumar P (2020) Banana disease diagnosis using computer vision and machine learning methods. J Ambient Intell Hum Comput 2016

Behera SK, Rath AK, Mahapatra A, Sethy PK (2020) Identification, classification and grading of fruits using machine learning and computer intelligence: a review. J Ambient Intell Hum Comput Kondo 2010

Bindu H, Bhuvaneshwari G, Jagadeesh SL, Ganiger VM (2019) Evaluation of physical and functional properties of weaning food blended with banana, sweet potato and drumstick leaves powder 8(2): 1568–1573

Cubero S, Lee WS, Aleixos N, Albert F, Blasco J (2016) Automated systems based on machine vision for inspecting citrus fruits from the field to postharvest—a review. Food Bioprocess Technol 9(10):1623–1639

El-Bendary N, El Hariri E, Hassanien AE, Badr A (2015) Using machine learning techniques for evaluating tomato ripeness. Expert Syst Appl 42(4):1892–1905

Kamilaris A, Prenafeta-boldú FX (2018) Deep learning in agriculture: a survey. Elsevier Comput Electron Agric vol 147(February): 70–90

Li H, Lee WS, Wang K (2016) Immature green citrus fruit detection and counting based on fast normalized cross correlation (FNCC) using natural outdoor colour images. Precis Agric 17(6):678–697

Mazen FMA, Nashat AA (2019) Ripeness Classification of Bananas Using an Artificial Neural Network. Arab J Sci Eng 44(8):6901–6910

Mureşan H, Oltean M (2018) Fruit recognition from images using deep learning. Acta Univ Sapient Inform 10(1):26–42

Nasiri A, Taheri-Garavand A, Zhang YD (2019) Image-based deep learning automated sorting of date fruit. Postharvest Biol Technol 153(January):133–141

Perez L, Wang J.(2017) The effectiveness of data augmentation in image classification using deep learning. http://arxiv.org/abs/1712.04621

Piedad E, Larada JI, Pojas GJ, Ferrer LVV (2018) Postharvest classification of banana (*Musa acuminata*) using tier-based machine learning. Postharvest Biol Technol 145(June):93–100

Sa I, Ge Z, Dayoub F, Upcroft B, Perez T, McCool C (2016) Deepfruits: a fruit detection system using deep neural networks. Sens (Switzerl) 16(8)

Shamim HM, Al-Hammadi M, Muhammad G (2019) Automatic fruit classification using deep learning for industrial applications. IEEE Trans Ind Inf 15(2):1027–1034

Simonyan K, Zisserman A (2014) Very deep convolutional networks for large-scale image recognition: 1–14. http://arxiv.org/abs/1409.1556

Singh R, Gosewade S, Ravinder Singh C, Kaushik R (2018) Bananas as underutilized fruit having huge potential as raw materials for food and non-food processing industries: a brief review. Pharma Innov J 7(6): 574–580 www.thepharmajournal.com

Thor N (2017) Applying machine learning clustering and classification to predict banana ripeness states and shelf life. Cloud Publi Int J Adv Food Sci Technol 2(1):20–25

Zhang L, Jia J, Gui G, Hao X, Gao W, Wang M (2018a) Deep learning based improved classification system for designing tomato harvesting robot. IEEE Access 6:67940–67950

Zhang Y, Lian J, Fan M, Zheng Y (2018) Deep indicator for fine-grained classification of banana's ripening stages. Eur J Image Video Process 2018(1)