CLASSIFICATION OF FRUITS RIPENESS USING CNN WITH MULTIVARIATE ANALYSIS BY SGD

K. Sumathi,[*] *V. Vinod*[*]

Abstract: Ripeness estimation of fruits is an essential process that impact the quality of fruits and its marketing. Nearly 30 % to 35 % get wasted from the harvested fruits due to lack of skilled workers in classification and fruit grading. Although it can be executed by human assessment, it is time consuming, costlier and error prone. Lot of research is carried to automate the quality assessment of fruits. Several hyper-parameters have been considered which have liven up by providing robust convolutional neural network (CNN). This paper has focused on image resizer stochastic gradient descent (SGD) algorithm for computing the loss. It updates the parameter by concentrating channels with respect to red, green, and blue (RGB) to identify and classify the images as ripen and rotten. The real time dataset (6702 images) of oranges, papaya and banana is collected. Using SGD optimizer, learning rate of 0.01 and nearest neighbor interpolation algorithm as resizer, the proposed model has achieved accuracy rate of 96.56 % after 38 epochs in classifying the fruits as ripen and rotten. It is also observed that it is possible to use small dataset on visual geometry group with 16 layer (VGG) with the above specification and good accuracy rate can be achieved.

Key words: *classification, convolution neural network (CNN), visual geometry group with 16 layer (VGG), and stochastic gradient descent (SGD), ripeness, rotten*

Received: August 4, 2022
Revised and accepted: Deecmber 31, 2022

DOI: 10.14311/NNW.2022.32.019

1. Introduction

India is one of the rural nations which has huge atmospheric scope as well as geological conditions. These conditions are most suitable for cultivating various type of plants namely fruits, grains, vegetables etc. There are many applications like object detection, image processing, robotics, facial recognition which has focused on methods to simulate the human brain. In addition, various studies and research experiments are introduced for providing advantages in the food industry's quality assessment by computer applications. During the commencement of industry 4.0

[*]K. Sumathi – Corresponding author; Viji Vinod; Department of Computer Applications, Dr. M.G.R Educational and Research Institute, Chennai, India, E-mail: captksumathi@gmail.com

revolution, the improving trend in human life automations have encouraged the robotic system benefits over food industry [1]. Automation in food packing process is significant for developing the food product production rate as well as minimizing the contamination probability and human contact from the food products. Using vision detection technology, we can reduce the complexity identifying the fruit size, shape and color due to poor processing. Most of the existing detection of fruit quality as well as grading systems may find defects such as low efficiency, less grading speed, complexity and expensive. As a result, it is essential to develop a high-speed, fruit size detection and grading system [2,3].

The policies of recent industries have addressed properties like shape, size and color through visual detections. When the sorting systems got automated based on technology of computer vision, it is proposed for improving production techniques. The product with high quality and the operations completely depends upon classification algorithms which need to be considered either with dissimilar color and shapes or combination of both [4,5]. Even though the classification algorithm has been introduced over several years ago, convolutional neural network CNN is used as main classification algorithm in the area of image processing. The rediscovery of main power from CNN can be identified by the ImageNet [6] whereas the architecture of AlexNet has successfully classified million images with thousands label with accuracy of 85 % which is comparatively higher than traditional algorithm which is 74 % [7]. This is the reason that CNN is always a significant algorithm for image processing classifications. One of the major advantages of using CNN is its robustness over new dataset because CNN does not require manual feature extraction for image preprocessing or classification. CNN technology is not only successful in image classification but got succeeded in applications of climate change detection, speech recognition and text classification.

The process of fruit ripening is closely related on the content of ethylene present in the fruit. The gradual increase of ethylene content level present inside the fruit is used to define the stage of fruit ripeness. The ethylene present has assisted in identifying the change of color, fruit aroma as well as the taste. The physical and chemical properties changes help to detect the fruit ripeness. Change in texture, size, smell and shapes is considered as physical property changes and it can be measured by non-destructive methods. Similarly, the ripeness stage identification can be done using chemical properties and it is measured through destructive methods. The advantage of considering physical properties for measurement while compared to chemical properties are simple, time efficient and less expensive. The disadvantages faced in identifying the physical properties changes is less accuracy in measurement while compared to chemical properties changes. The different levels of ripening banana fruit is illustrated in [8,9]. The general color image representation has been implemented through computer vision system which consist of three primary color combination namely red, green, and blue (RGB). Each primary color consists of triplet values which is generally considered to be a coordinate system with various metric to measure distance such as Euclidean, Hamming and Mahalanobis. Each point in the 3D coordinate system which represent a different color in the visible spectrum. Beyond the RGB color, the other color spaces are generally implemented by providing various 3D representations and it can be classified into three spaces categories namely human-oriented spaces, hardware-oriented spaces and instrumental spaces. The working of human-oriented spaces is based on the principle of inherent color characteristics which is executed through saturation and hue like HSL (hue, saturation and lightness), HSV (hue, saturation and value) and also known as HSB (hue, saturation adn lightness). Similarly, the hardware device properties that utilized for displaying images can be done through hardware-oriented spaces such as RGB, CMYK and YIQ. In the case of instrumental space, the color coordination is same for all the output media such as CIELAB, CIELUV and XYZ.

Hyper parameters are considered as variables for defining the CNN structure and permit to be trained as a model [10]. The various hyper parameters are epochs, batch size, learning rate, optimizer, activation functions and number of layers. These can be adjusted for generating the CNN method with high efficiency. The optimizer utilized in the CNN, stochastic gradient descent (SGD) and adaptive moment estimation (Adam) have better performance in classifying the image [11]. The CNN performance can be affected through batch size and learning rate of image classification. The batch size impact can be assessed with two dissimilar optimizers. This network has been fine-tuned to match the dataset as well as to avoid model training from the scratch and the experimental research focused on determining a better batch size value to understand the available issues by CNN which can be executed through SGD optimizer.

The rest of this paper is organized as follows, the Section 2 discuss about CNN with various ImageNet classification and their performance on fruit ripeness classification ripen. The benefit of visual geometry group (VGG) net classification method is also briefly discussed. Section 3 illustrates material collection, preprocess, feature extraction and training using CNN with VGG algorithm with RGB images of both climacteric and non-climacteric fruit. Section 4 illustrates the training and validation accuracy curve as well as loss curve for CNN-VGG to identify the performance of training model. Finally, Section 5 has concluded the performance of the model in classifying the fruits as ripen and rotten using VGG.

2. Literature review

This section discusses the review of various previous endeavors to use CNN as well as VGG for fruits ripeness classification. Most of the researchers have discussed about color features parameter associated with RGB color spaces extract from the fruits during image processing.

X. Chen et al. [12] has proposed an analysis of detecting watermelon ripeness through wavelet multi resolution decomposition (WMRD). Two kinds of samples are considered for this experimental research namely ripe watermelon and unripe watermelon. The acoustic signals with multi-scale decomposition are used as measuring tool. The discrimination index has been conferred for identifying the WMRD coefficient which is best by accomplishing high accuracy rate in testing. The training model with accuracy curve for training is 91.67 % and testing is 91.76 %.

R. Thakur et al. has discussed about analyzing and sorting of strawberry fruits by an automated vision-based system. He has justified that manual classification may lead to discrete opinion that consume more time. However, the identification of cost efficient and high accuracy performance has become an issue till now. He has proposed automated system with CNN assist to predict the strawberry fruit ripeness level. The suitable features are extracted as succeeded classification in which size, shape and surface color are the essential features of classification. The automated CNN has accomplished with 91.6 % of accuracy [13].

Pooja Kamble et al. has proposed CNN method to identify and classify the ripening stage of fruits in which mango, apple and banana are classified. This proposed system is generalized for only these three fruits and is giving good result. But for other fruits he proposed system is not favorable [14].

Yan Zhang et al. has proposed an architecture of CNN that is designed particularly in classifying the ripeness stage of banana with fine grained. This architecture learns features based on data driven mechanism of fine-grained image set. That allow to study with deep indicator in analyzing ripening stage of banana. The outcome indicator has assisted in differentiate fine grained differences between secondary classes of ripening stage in banana. The resulted ripening stage of banana from 17,311 images illustrate the accomplishment of deep indicator as an accuracy. Therefore, the accuracy determines that modernized computer vision-based system of CNN is better in both rough as well as fine grained ripening stage of classification [15].

Radiuk has studied the effect of batch size on CNN performance for image classification by implementing two datasets namely MNIST and CIFAR-10. The testing batch size is from 2^{16} to 2^{1024}. In the case of MNIST dataset, the LeNet architecture is used but for CIFAR-10 dataset, five layers with convolutional layer as user defined network. The SGD optimizer is utilized for both networks with a learning rate of 0.001 for the MNIST dataset and 0.0001 for the CIFAR-10 dataset. The 2^{1024}-batch size produced the best accuracy for both datasets, and the 2^{16}-batch size produced with the worst results. According to the researcher that the larger the batch size, higher the network accuracy. Hence the researcher concluded that batch size has a significant impact on the performance of CNN [16].

Luschi has investigated the impact of batch sizes ranging from 2^1 to 2^{11} on AlexNet and ResNet topologies using SGD as an optimizer without momentum for preventing the momentum effect while training [17]. D. Masters examined the effect of batch size using three datasets namely ImageNet, CIFAR10, and CIFAR100. The best outcome is accomplished with batch sizes ranging from 2 to 32, and small batch sizes are more robust than high batch sizes [18].

A new method for identifying fruits from images using deep neural networks (DNN) presented in this study has utilized a faster region-based CNN. The goal is to develop a neural network which is utilized through self-driving robots for picking the fruits. RGB and near infrared (NIR) images are used to train the network and the respective models are combined in two different ways namely early and late fusion. The input layer for early fusion contains four channels such as one channel for NIR image and the remaining three channel for RGB images. Late fusion involves the deployment of two independently trained models which have been combined by average the predictions from both models. Thus, the outcome is a multi-modal network that performs significantly better than current networks [19].

In the case of handwritten devanagari characters, Jangid and Srivastava has evaluated the performance of three optimizers. RMSProp, Adam and Adamax are the optimizers considered for the evaluation. The CNN architecture with three convolutional layers and one thick layer with 1000 neurons is used to test each optimizer. RMSProp optimizer is considered to be highly accurate performer for this architecture [20].

Swastika et al. evaluated three optimizers for vehicle classification namely SGD, Adam and Adadelta. Each optimizer is evaluated with three different CNN architectures such as LeNet, shallow network, and MiniVGGNet. The outcome illustrates that Adadelta optimizer is considered to be best for the Mini VGGNet architecture due to high accuracy optimizers [21].

The analysis and review of literature brought forth the fact that there is a need for investigating fruits ripening at different stages. Not much study is carried out from ripen stage to rotten stage. In this paper, the researcher has carried out the first stage of study to establish the facts. This is needed so that overall quality evaluation is established.

3. Research methodology

The goal of this research is to identify and classify the ripeness stage and rotten stage of fruits. The fruits included in this study is both climacteric as well as non-climacteric fruits. The climacteric fruits used are papaya and banana. The non-climacteric fruit considered is orange. This study focused on real time datasets of fruit images created by collecting the fruits in ripe and rotten stages. The dataset creation process is explained in next session followed by proposed architecture using VGG with SGD support and learning rate. The workflow diagram of proposed system is shown in Fig. 1.

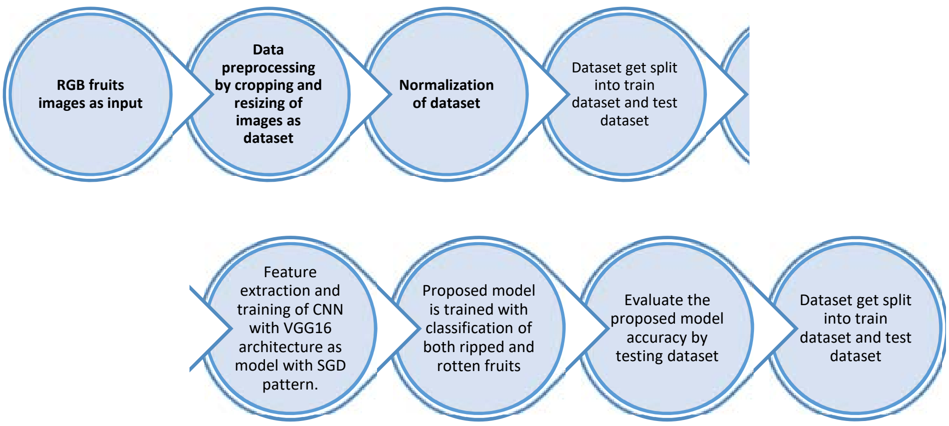

Fig. 1 *Proposed workflow diagram of model.*

3.1 Dataset collection

As a part of research, one of the aims is to create a real time dataset by capturing the ripe and rotten stages of fruits. Initially the researcher collected images from online resources. The total number of images collected were 8562 numbers. It was divided into two parts namely training data and testing data. The collected data from online resource did not meet the requirement and failed to achieve the accuracy rate. The reason being, the images were preprocessed and already resized to 100×100 pixel and it was augmented. When this dataset was used in the proposed algorithm, model could not learn much as it has already lost lot of crucial information during preprocessing. So, this dataset was discarded and author collected new images of fruits. This time the fruits were collected from various retailers in Thiruvallur district, India. Fruits considered in this paper are banana, papaya (climacteric fruit) and oranges (non-climacteric fruit). The data used in this research is RGB Images of ripe and rotten stages of fruits. The total number of images collected were 6702. Out of 6702, 324 images were removed as a part of preprocessing. So, remaining 6378 is considered for training and testing. The collection of fruits with source of collection is segregated and shown in Tab. I. The collected image is pre-processed by cropping, resizing the data as needed which is programmatically done. The aim of the author was to collect different size and shapes so that it will fit the model and can be applied in real life.

Fruit name	Source of collection	Total collected	Removed during preprocessing	Balance
Orange Ripe	Spar super market. RK fruit shop	1115	48	1067
Orange Rotten	Spar super market. RK fruit shop	1120	52	1068
Papaya Ripe	Ponnu super market, reliance fresh	1115	59	1056
Papaya Rotten	Ponnu super market, reliance fresh	1120	50	1070
Banana Ripe	RK fruit shop	1115	60	1055
Banana Rotten	RK fruit shop	1117	55	1062
		6702	324	6378

Tab. I *Summary of ripe & rotten fruit 1.*

Image acquisition: Images are acquired from retailers in Maduravoyal, Thiruvalluvar district. RGB Images are captured using digital still camera of 15.2 mega pixels and complementary metal oxide semiconductor (CMOS) image sensor. It is taken under normal room like condition. The author introduced variation in the dataset, by taking images of the fruits at different positions and rotations. Images are acquired at different time of the day. The sorting and grading expert has labeled the fruits based on the physical outer appearance like color, size, shape, firmness and smell.

Cleaning: As for as cleaning is concerned, it was done by sorting out the images class by class. It was than viewed to find any defects in each image. Initially some setbacks were faced as the images that were considered were in landscape. First it was trained as it is. Later, white region was cropped to make it square.

For base dataset, the author has not used any unique preprocessing technique. In these 6 classes of 3 fruits with ripe and 3 fruits with rotten stages are considered. Actual size of the image is 4000×4000 pixels. The RGB images of fruits dimensions are reduced to square as 2000×2000 pixels with total images of 6378. Once the data preprocessing is completed, the normalization is done which is nothing but dividing each pixel by 255 to ensure that entire pixel is represented in term of 0 and 1. This representation assist the model to learn better. In general, the normalization is the initial process of all neural network that assist in learning invariant feature. This is the common preprocessing technique which is used in every neural network.

Fig. 2 presents few samples of fruit images that were collected. These images are preprocessed and used for training the model.

3.2 VGG architecture as training process

VGG is simple and majorly utilized for very deep convolution network to recognize the large-scale images. A basic convolutional neural network can be seen as a sequence of convolution layers $f(x)$ and pooling layers.

$$f(x) + y = z, \tag{1}$$

where $f(x)$ is the input image of fruits, y is the filter or kernel or features (piece).

When the image goes through them, the important features are kept in the convolution layers, and in pooling layers, these features are intensified and kept over the network, while discarding all the information that doesn't make a difference for the task. The input source of RGB images is converted into the RBG image with fixed size of 224×224. The preprocessing of VGG is done by subtracting the calculated mean value of RGB on the training dataset from every pixel. For this study, features that are considered for real time data set creation are color, shape, firmness, i.e mainly outer look of the fruit. The model consists of 16 layers of convolution and pooling. Convolution extracts the important feature and pooling layer intensifies it. Weights are trained during back propagation. Initially it will be a random number with normal distribution.

$$\text{Weight (Filters)} = (\text{Learning Rate x Loss}) + \text{Previous Weight} \tag{2}$$

Weights (filters) and loss are the learnable parameters.

The hyper parameters considered is learning rate (0.01) and SGD (stochastic gradient descent). The resizer is applicable on the input image where we convert the input image to 224×224 for our model. There are no features considerations as we will bias the model training. Since it is a CNN architecture the feature selection

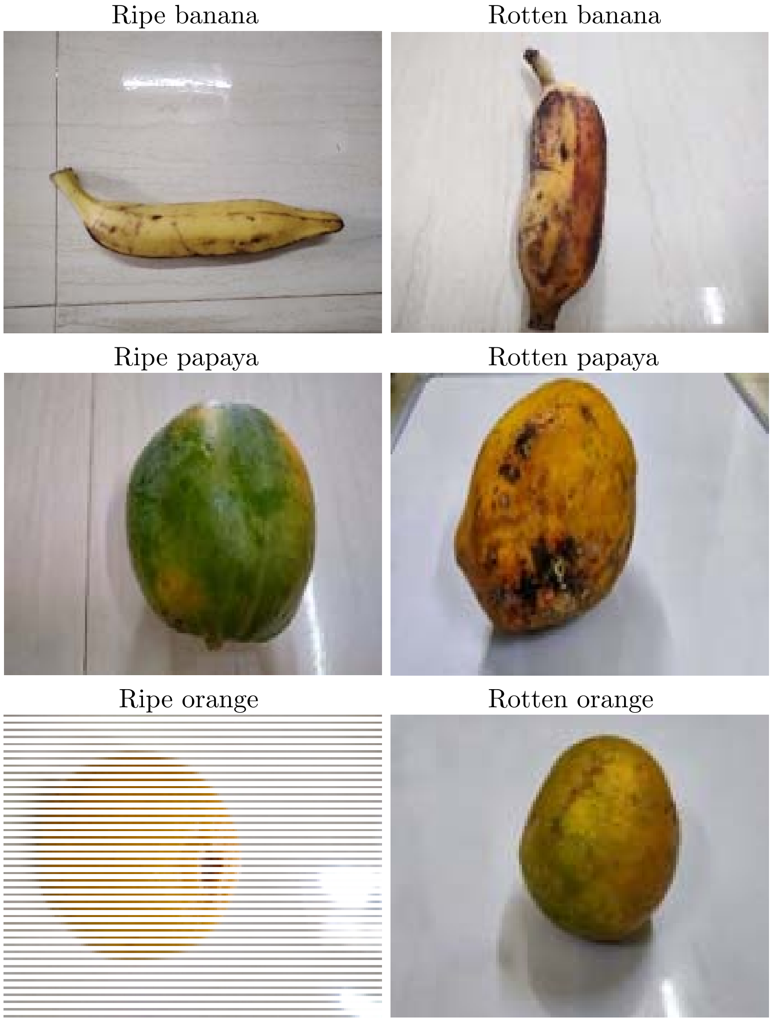

Fig. 2 *Sample of fruit images collected for training.*

is automatically done by the model and that's what the model learns to adjust it by training. The input image of fruits is made to pass through stack of convolution layers and max-pooling layer where the 3×3 filters are applied. There are five numbers of max-pooling layers with 2×2 pool size and 2×2 stride. The output is passed finally through the 3 fully connected dense layer as depicted in Fig. 3.

3.3 SGD optimizer for training process

The training procedure of ConvNet is generally enforced by multinomial logistic regression optimization as an objective through mini-batch gradient descent which is said to be SGD. The setting of batch size is 256 that is 0 to 255, and momentum is up to 0.9. Based on the weight regularization, the sum of squared weight (L2)

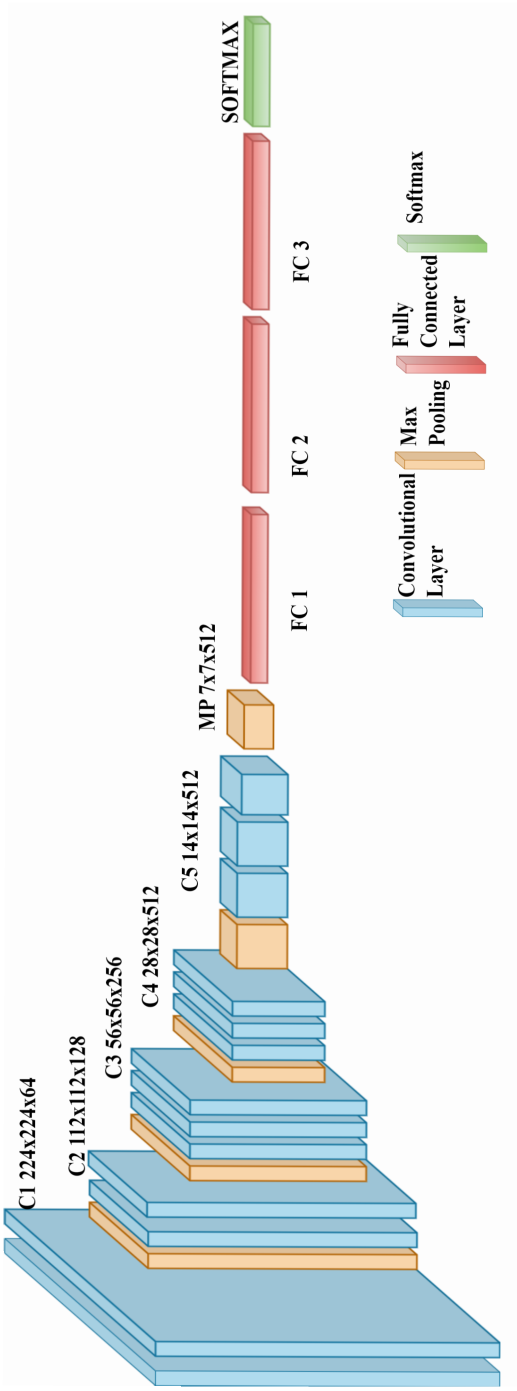

Fig. 3 *VGG architecture.*

penalty multiplier has been set as 5×10^{-4}. Each vector needs a hyper parameter which has to be configured for improving the performance of an overfit deep learning network in python with keras. The regularization of dropout for the setting of first and second fully connected (FC) layers consist of 0.5 as dropout ratio. However, the learning rate is generally set as 0.01 and then further reduced by a factor of 10 while the accuracy of validation set stops growing. Therefore, the learning rate has been lessened by 3 times, and it stops at the iteration of 75 epochs. Thus, it gets estimated instead of the higher parameter amount and the more depth considered in CNN.

Moreover, the network required certain epochs for converging because of pre-initialization by specific layers and implicit regularization executed by smaller convolution filter sizes. Hence, the network weights are significant. In order to avoid these issues, SGD optimizer is introduced in the training configuration that have initialized randomly. Based on the VGG with CNN architecture, first five convolution layers with three FC layer are progressed. Thus, the learning rate for the layer of pre-initialized is not reduced but it may allow to modify while learning. The sample setting of the random initialization of layer progresses the normal distribution weight with mean as 0, variance as 0.01 and bias is initially progressed with 0.

4. Result and discussion

This experimental research has utilized highly performance server along with configuration of an Intel core i7 DMI2 CPU, 12 GB RAM, 100 GB free space and GPU as Quadro K600. The operating system used is Ubuntu 18.04.3 LST for training the image dataset. The optimizer and loss function utilized in this proposed model is SGD and cross entropy. There are three main arguments involved in the constructor of VGG model of CNN is weight, include_top and input_shape. The weight constructor represents the checkpoint weight from the initiated model, include_top has represented to involve the classifier with densely connected on the top of the network and the input shape has represented image tensor shape to the network which enables to the network to process all kind of input sizes. In this research, there are 6378 images involved with the size of 2000×2000 pixel. The dataset was split into 75 % of training data and remaining 25 % as testing data. According to this research, training image dataset is essential which get segregated with 65 % of training image data and 10 % of validating image data for both ripe and rotten fruits of orange, papaya and banana shown in Tab. II.

With stochastic gradient descent optimizer, learning rate of 0.01 against standard learning rate of 0.001 and nearest neighbour interpolation algorithm as resizer, training achieved accuracy achieved 100 % and validation achieved 96.56 % after 38 epochs. It is also observed that it is possible to use small dataset ion VGG with the above specification and good accuracy rate can be achieved in classification.

This research mainly focused to involve VGG rather than considering hyper parameter with large number and the VGG aim to have convolution layers of 3×3 filter with a three stalk with max-pooling each and often utilize the same padding for 2×2 convolution layer filter with max-pooling each for two stalks. This arrangement is sequenced with convolution layer followed by one max-pool layer

Real time fruit dataset	Fruit status	Image count	Training images		Testing images
			Training	Validating	
Orange images	Ripe	1067	693	107	267
	Rotten	1068	694	107	267
Papaya images	Ripe	1056	686	106	264
	Rotten	1070	695	107	268
Banana images	Ripe	1055	686	105	264
	Rotten	1062	691	106	265
Total		6378	4145	638	1595

Tab. II *Training and testing images of real time fruit dataset.*

for each stalk constantly through the entire architecture. Finally, it has 3 dense layers followed by a softmax for output.

The classification of ripe and rotten is defined with the proposed CNN with VGG model whereas the training accuracy and validation accuracy as well as model loss is illustrated in the Fig. 4 and Fig. 5 correspondingly.

The Fig. 4 illustrates the model accuracy whereas the training accuracy is steady after 38 epochs and maintain the constant accuracy with 98.5 %. In the case of validation accuracy, the steady rate of accuracy follows after 38 epochs with 96.56 %. As the epochs increases the learning rate has become better with more iterations and there is a steady rate of accuracy maintained after 38 epochs. This represents the learning rate get modified from 0.01 to 0.1 from the 40 epochs.

Fig. 5 illustrates the model loss in both training as well as validation whereas the loss is more in training because initially the model is not learned better. Once the epoch gets increased, the losses get reduced from 1.75 to 0.19 which likely to define that transfer learning is occurred to minimize the loss. In the case of validation loss curve, the loss in the initial epoch is lesser than training loss but increase after 20 epochs. There is more difference is maintained after 38 epochs where the loss gets increased slightly after that epoch till 50 epochs. Thus, the model is trained well with ripe and rotten fruits in the better way to classify the model with better accuracy of 96.56 %. Tab. III summarizes the specification of model.

5. Conclusion

In this paper, the CNN with VGG classifier has been designed and trained on collected dataset of fruits. The real time dataset is created and has been contributed to the research community. The novel model created is used for classification of 6 classes of fruits as ripe and rotten. The SGD optimizer is utilized for considering the better learning rate of training model by numerous epochs. The classifier training parameters are evolved with VGG convolutional architecture to achieve

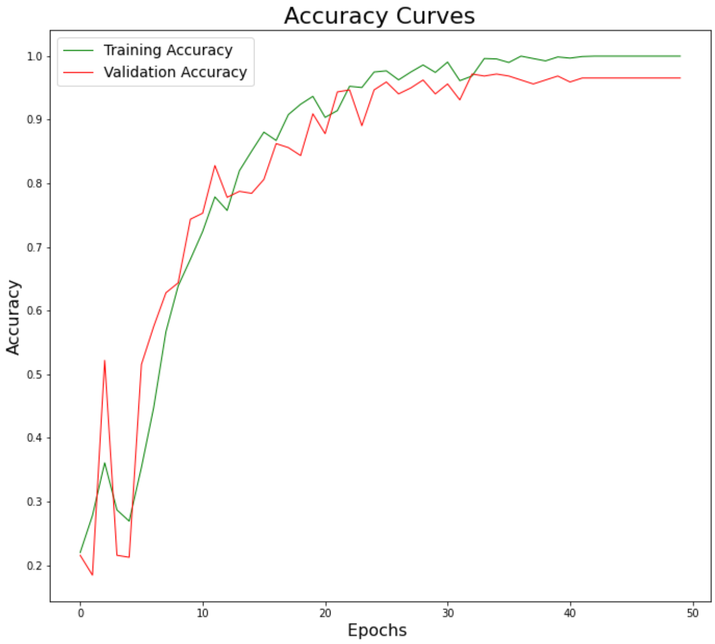

Fig. 4 *Accuracy curve for CNN-VGG method.*

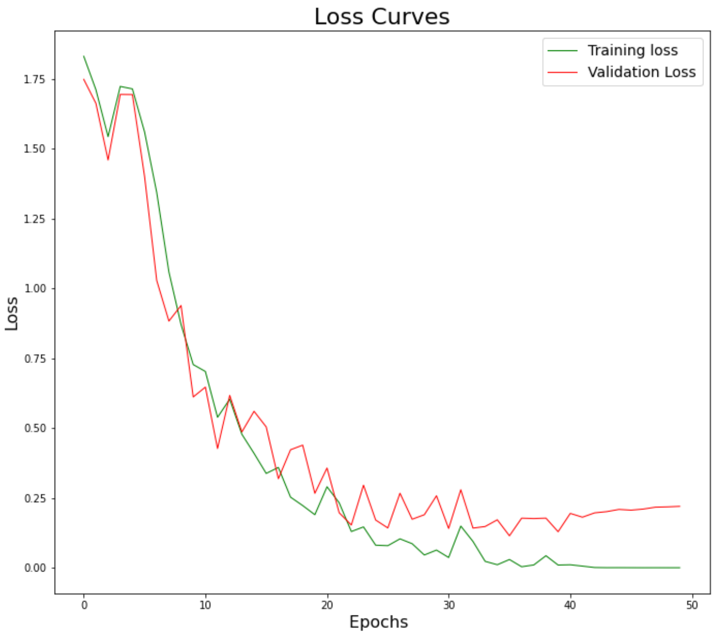

Fig. 5 *Loss curve for CNN-VGG method.*

Automated feature extraction	Total params: 134,285,126 Trainable params: 134,285,126 Non-trainable params: 0
Input image size	(224, 224, 3)
Resizer	Nearest neighbour interpolation
Kernel size	3×3
Optimizer	Stochastic gradient descent
Learning rate current	0.01 (Standard learning rate: 0.001)
Training accuracy	100 %
Validation accuracy	96.56 % after 38 epochs

Tab. III *Specification of model.*

good learning rate. Using stochastic gradient descent optimizer, learning rate of 0.01 against standard learning rate of 0.001 and nearest neighbor interpolation algorithm as resizer, the model has achieved 96.56 % accuracy rate after 38 epochs. It is also observed that it is possible to use small dataset on VGG with the above specification and good accuracy rate can be achieved in classification. Hence, the proposed CNN with VGG has good accuracy in classifying the ripen fruits as well as rotten fruits.

References

[1] BADER F., RAHIMIFARD S. Challenges for industrial robot applications in food manufacturing. In Proceedings of the 2nd International Symposium on Computer Science and Intelligent Control, 2018, 37, pp. 1–8, doi: 10.1145/3284557.3284723.

[2] BANOT S., MAHAJAN P.M. A fruit detecting and grading system based on image processing-review. International Journal Of Innovative Research In Electrical, Electronics, Instrumentation And Control Engineering, 2016, 4(1), pp. 47–52, doi: 10.17148/IJIREEICE.2016.4112.

[3] BEHERA S.K., RATH A.K., MAHAPATRA A., SETHY P.K. Identification, classification & grading of fruits using machine learning & computer intelligence: a review. *Journal of Ambient Intelligence and Humanized Computing,* 2020, pp. 1–11, doi: 10.1007/s12652-020-01865-8.

[4] CASTRO W., OBLITAS J., DE-LA-TORRE M., COTRINA C., BAZÁN K., AVILA-GEORGE H. Classification of cape gooseberry fruit according to its level of ripeness using machine learning techniques and different color spaces. IEEE access, 2019, 7, pp. 27389–27400, doi: 10.1109/ACCESS.2019.2898223.

[5] DE-LA-TORRE M., AVILA-GEORGE H., OBLITAS J., CASTRO W. Selection and fusion of color channels for ripeness classification of cape gooseberry fruits. In Trends and Applications in Software Engineering, Proceedings of the 8th International Conference on Software Process Improvement (CIMPS 2019) Springer International Publishing: Berlin/Heidelberg Germany, 2020, pp. 219–233, doi: 10.1007/978-3-030-33547-2_17.

[6] DENG J., RUSSAKOVSKY O., SU H., KRAUSE J., SATHEESH S., MA S., HUANG Z., KARPATHY A., KHOSLA A., BERNSTEIN M., BERG A.C. Imagenet large scale visual recognition challenge. International journal of computer vision, 2015, 115, pp. 211–252, doi: 10.1007/s11263-015-0816-y.

[7] HINTON G.E., KRIZHEVSKY A., SUTSKEVER I. Imagenet classification with deep convolutional neural networks. Communications of the ACM. 2017, 60(6), pp. 84–90, https://doi.org/10.1145/3065386.

[8] LIAN J., ZHANG Y., FAN M., ZHENG Y. Deep indicator for fine-grained classification of banana's ripening stages. EURASIP Journal on Image and Video Processing. 2018, 1, pp. 1–10. doi: 10.1186/s13640-018-0284-8.

[9] MASTERS D., LUSCHI C. Revisiting small batch training for deep neural networks. 2018, doi: 10.48550/arXiv.1804.07612.

[10] NAHAT A.A., MAZEN F.M, Ripeness classification of bananas using an artificial neural network. Arabian Journal for Science and Engineering, 2019, 44, pp. 6901–6910, doi: 10.1007/s13369-018-03695-5.

[11] NANDI C.S., KOLEY C., TUDU B. A machine vision-based maturity prediction system for sorting of harvested mangoes. IEEE Transactions on Instrumentation and measurement, 2014, 63(7), pp. 1722–1730. doi: 10.1109/TIM.2014.2299527.

[12] PARK S.U., HONG K., HAN J.H., CHOI D.J. Hyperparameter optimization using a genetic algorithm considering verification time in a convolutional neural network. Journal of Electrical Engineering & Technology, 2020, 15, pp. 721–726, doi: 10.1007/s42835-020-00343-7

[13] PATEL H., THAKUR R., SURYAWANSHI G., SANGOI J. An Innovative Approach For Fruit Ripeness Classification. 4th International Conference on Intelligent Computing and Control Systems (ICICCS), 2020, pp. 550–554. doi: 10.1109/ICICCS48265.2020.9121045

[14] PEIPEI Y., CHEN X., DENG X. Watermelon ripeness detection by wavelet multiresolution decomposition of acoustic impulse response signals. Postharvest Biology and Technology, 2018, 142, pp. 135–141, doi: 10.1016/j.postharvbio.2017.08.018

[15] RADIUK P.M. Impact of training set batch size on the performance of convolutional neural networks for diverse datasets. Information Technology and Management Science. 2017, 20(1), pp. 20–24, doi: 10.1515/itms-2017-0003.

[16] RANVARE S.S., MARATHE R.S., JHA S.K., JAYASHREE V.K., POOJA R.K. Development of an effective system to Identify Fruit ripening Stage for Apple, Banana and Mango. International Journal of Advanced Science and Techn, 2020, 29(12), pp. 2766–2772. http://sersc.org/journals/index.php/IJAST/article/view/24972

[17] REN S., HE K., ZHANG X., SUN J. Deep Residual Learning for Image Recognition. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.

[18] SA I., GE Z., DAYOUB F., UPCROFT B., PEREZ T., MCCOOL C. Deepfruits: A fruit detection system using deep neural networks. sensors. 2016, 16(8), doi: 10.3390/s16081222.

[19] SHRIVASTAVA VIMAL K., BERA S. Analysis of various optimizers on deep convolutional neural network model in the application of hyperspectral remote sensing image classification. International Journal of Remote Sensing, 2020, 41(7), pp. 2664–2683. doi: 10.1080/01431161.2019.1694725.

[20] SRIVASTAVA S., JANGID M. Deep convnet with different stochastic optimizations for handwritten devanagari character. In Advances in Computer Communication and Computational Sciences, Proceedings of IC4S 2017, 1, pp. 51–60. Springer Singapore. doi: 10.1007/978-981-13-0341-8_5.

[21] SWASTIKA W., ARIYANTO M.F, SETIAWAN H., IRAWAN P.L. Appropriate CNN architecture and optimizer for vehicle type classification system on the toll road. In: Journal of Physics: Conference Series 2019, 1196(1), doi: 10.1088/1742-6596/1196/1/012044.