



Classification of Ripeness Fruits Using Convolutional Neural Network (CNN) With Multivariate Analysis by SGD in VGG16

***A. Karthikeyan, K. Venkatasalam, S. Andrews, E. Madhorubagan,**

Mahendra Engineering College, Namakkal, Tamil Nadu, India.

*Corresponding author: karthikeyana@mahendra.info

Abstract: The major backbone for a country economic growth is agriculture industry with high yield productivity and development in production. One of the essential process in fruit is estimating of fruit ripeness that influence the fruits quality as well as its marketing. Almost 30% to 35% have been made wasted while fruits harvesting due to lack of skilled workers. The evaluation of automatic fruit ripeness is executed through computer vision arrangement. Image processing has assisted in agricultural industry for improving the estimation of yield, detecting diseases, sorting of fruits and vegetables, irrigation as well as maturity grading. However, these challenges are handled in image processing with several hyperparameters that have lived up by conditional robust Convolutional Neural Network (CNN). The CNN help to obtain accurate classifying images with 16 layers (VGG16) with batch. Therefore, this paper focuses on Stochastic Gradient Descent (SGD) method which has used epoch concept over training dataset for computing the loss. It even update the parameters to identify as well as classify the images of rotten and ripen. The real time dataset is collected with RGB images consists of 6702 images involving three kinds of fruits namely oranges, papaya and banana with six classes such as three classes with the ripped fruits and another three classes with the rotten fruits. Moreover, the evaluation of CNN with VGG16 using SGD algorithm for accurate classification can be determined by parameters like accuracy and loss curve for training dataset as well as validation dataset with different epochs.

Keywords: Stochastic Gradient Descent, Ripped fruit, rotten fruit, Classification, Convolution Neural Network and VGG16

1. Introduction

The concept of precision agriculture has evolved as an innovative approach to modern farming techniques, with the goal of optimizing resource allocation, increasing production, and reducing environmental impact [1]. Precision agriculture relies heavily on the accurate measurement and monitoring of agricultural qualities such as fruit maturity, which is critical for guaranteeing optimal harvest time as well as quality of fruits [2]. Fruit ripeness detection with accurate is critical for agricultural industry, allowing farmers in making educated decisions about harvesting schedule, post harvest treatment, as well as marketing tactics. The significance in determining fruit maturity in precision for agriculture can't be emphasized. Farmers can maximize productivity and reduce waste by assessing fruit maturity on time and accurately [3, 4]. Furthermore, it helps to optimize the supply chain for ensuring customers obtain fruits with the best texture, nutritional and taste value [5]. Detection of fruit ripeness helps with storage as well as distribution operations, reducing spoiling and increasing shelf life. As a result, the capacity to detect fruit ripeness with precision has become vital to agricultural businesses' viability and profitability. Several vision-based approaches for determining fruit maturity have been investigated and developed, including visual signals including color, texture, and form [6, 7].

These techniques have used IP algorithms and ML techniques for extracting significant feature and categorize fruits accordingly with its ripeness level [8].

The current mechanical era, providing good food qualities to a human has key testing duty which is possible that significant labor will be necessary to evaluate the properties of fruit. To complete this project, a programmed framework of fruit-evaluation has required to quality as well as sustenance purposes is crucial. The uncertain automatic quality technique distinguishes the nature of fruit by recognizing innovation while causing no harm to the produce. In any event, faulty procedures have made it complex for determining the nature of fruit depends upon size, shape, and color. To pass this test, ML and computational information approach are utilized in determine the kind of fruits with great accuracy. The number of farmworkers in the fruit industry is decreasing by the day. As a result, it is necessary to adopt labor-saving technology. In this case, IP techniques are the most effective method for identifying fruits, production sorting and packaging easier. In determining the correct market value, the fruits are quantified with respect to its size and quality. Inspecting fruit quality can only be done physically by seeing as well as feeling, which is influenced by a variety of factors like inconsistent and unreliable decision-making. Finally, the IP is the most suitable technique for fruit classification and quantifying [9, 10].

These technologies, which analyze digital photographs of fruits, may deliver non-destructive, scalable and real-time solutions for determining fruits maturity. Previous research has demonstrated an increasing interest in DL approaches for detecting fruit ripeness because of its capacity to effectively learning and extracting complicated features from large scale datasets [11]. SGD is a prominent optimization technique utilized for CNNs in medical images segmentation which operates by update the network's weights using the partial derivations of the loss function when compared to every weight. This method is frequently iteratively across numerous epochs for reducing loss and enhances model accuracy. Specifically in SGD, stochastic optimization method computes gradients and updates weights at each iteration using a fraction of the training data. This strategy may be highly efficient than by the whole training set, particularly for huge datasets. This is distinct of the gradient descent approach in which, rather in computing the cost function gradient across the full dataset which is collected randomly in a single data point at every iteration as well as determines the gradient with respect to data point. Thus, the SGD is significantly quicker as well as highly scalable than normal gradient descent, particularly to large datasets. Furthermore, SGD assist for preventing being stuck with local minima using including unpredictability towards training process. The chosen learning rate has a significant impact on SGD's performance. An excessive learning rate might result in unstable training or divergence, whilst a low learning rate can induce delayed convergence. As results, the optimal learning rate is complex and may need manual adjustment or the usage of flexible learning rate approaches. When working with sparse datasets, where the majority of the features have zero or very few values, SGD may encounter issues.

However, CNNs are a common type of NN for image segmentation tasks because it has learned for extracting features from images related to the task on hand. Hence, it enables them to establish the borders among various parts of the image and generate reliable segmentation mapping. DL models like CNN has shown outstanding performance in a variety of task in computer vision include image classification, object recognition, and other related applications [12, 13]. The benefits of DL methodologies are to create robust as well as reliable models for detecting fruit ripeness, addressing certain drawbacks of classic methods for IP. The remainder of this study is organized as follows, Section 2 discusses CNN with MobileNet classifications and their effectiveness with fruit ripening, as well as the benefits of the MobileNet classification approach. Section 3 shows material collection of both climatic and non-climatic fruit that are preprocessed, as well as feature extraction and training using CNN with the MobileNet for RGB images. Section 4 depicts the training and validation accuracy curves, as well as the loss curve for CNN with MobileNet for determining the training model performance. At last, Section 5 concludes by evaluating the performance of identifying ripe fruit based on physical features via image processing classification utilizing CNN-VGG16 with a 93.39% accuracy.

2. LITERATURE REVIEW

This section reviews earlier attempts to employ CNN and VGG-16 for fruit ripeness classification, and also various researchers' discussions of color characteristics parameters linked with RGB color spacing extracted from fruits in IP. R. Thakur et al. explored the analysis and sorting of strawberries using an automated vision-based method, stating that manual classification may result in separate opinions that take more time. However, identifying cost efficiency and excellent accuracy performance remains a challenge, but the proposed automated system with CNN may help in

predicting the ripeness degree of strawberries. As a result, the appropriate features are retrieved as successful classification, with shape, size, and color being the important elements of classification, and the strawberry fruit surface level determining the ripening level. Thus, the automatic CNN achieved 91.6% accuracy in extracting information from the strawberry fruit surface such as size, color, and shape [14].

PoojaKamble et al. suggested a CNN approach to recognize and categorize the ripening stage for fruits like banana, mango and apple being the three fruits classified and their ripening stages identified, accordingly. This proposed system benefit demonstrates that it is only generalized to these three fruits, but it has a disadvantage for other fruits by defining no ripening as no physical change in the fruits [15]. D.Xu et al. have provided an extensive approach that integrates computer vision approaches and ML methods for detecting and analyze the freshness of jujube fruits. The suggested method has shown capable results with respect to accuracy and efficiency [16]. Moreover, one disadvantage of this study is focused solely on jujube fruits, and the method's application to new fruit varieties remains unknown. More research is required for evaluating its efficiency across various fruit kinds, as well as to determine the suggested method's generalizability.

Fathima N have compared and evaluates the performance of basic optimization algorithms such as Adamax, Rmsprop, Adam, Adadelta, Nadam, Sgd, and Adagrad. These methods had tested on four different datasets autonomously. The major goal is for determining optimization algorithm produces high accuracy with very high efficiency in performance for DNN models. The findings of this study are likely to provide significant awareness to data scientists that assist in making educated judgments when picking the best optimizer for their DNNs [17]. Darwish et al. have utilized the Orthogonal Learning Particle Swarm Optimization (OLPSO) method to effectively optimize many hyperparameters, substituting previous approaches such as human trial and error. Improving the efficient of training and even minimize the local minima during pre-trained ConvNet models, Exponentially Decaying Learning Rate (EDLR) paradigm [18]. H. Zang et al. have proposed a new approach that combines the benefits of neuroevolution and SGD. This fusion enables evolutionary search, concurrent exploration, and a reliable method for discovering optimal deep neural networks (DNNs) [19].

Furthermore, the controlled mechanism of hierarchical cluster has developed in solving the problem of comparable update in weight over individuals. Hence, the population diversity has been increased using suggested approach that applicable for four major DNNs utilizing publically available datasets. The experimental results show that four DNNs optimized using the proposed technique outperforms consistently with optimized exclusively using SGD across all datasets. M.Reyad et al. have created the HN Adam method by combining the ordinary Adam algorithm and the AMSGrad algorithm. Sgd-based algorithms are widely used in DNNs [20]. However, their use raises concerns, such as the occurrence of oscillating learning rates during later training phases in certain adaptive methods, which could lead to non-convergence. S.Wojtowysch as addressed this issue and suggest optimization algorithms based on reduction in variance reduction for improving convergence rate [21]. This presentation examined the various CNN algorithms for predicting fruit repining, as well as the approach optimizer used to classify photos. This evaluation analyzes the RGB image that was trained using the harvesting fruit image in the CNN. However, multiple studies have examined the architecture of SGD using RGB images, which provides greater accuracy and is being implemented in the VGG16 batch dataset. As a result, the VGG16 architecture is used in CNN with SGD assistance to improve accuracy in learning rate, as demonstrated in the following session.

3. RESEARCH METHODOLOGY

Detecting the fruit rotting and ripeness with better accuracy is the motivation of this research and it involve in both climacteric and non-climacteric fruits. This study involves banana, and papaya are involved for climacteric fruits as well as non-climatic fruits as orange. Experimental research focuses on real-time fruit image dataset have been collected, preprocessed and feature gets extracted. Training was carried out using a CNN-VGG16 with SGD algorithm with RGB images. The workflow diagram of proposed system is shown in figure 1.

Dataset Collection: According to this research, the dataset is fruits images which are taken from various retailers in maduravoyal, Thiruvallur district, India. The dataset involved in this research is RGB images for both ripped and rotten fruits. The images of fruits like orange papaya and banana have been captured by digital still camera with the pixel of 15.8 Mega Pixels and image sensor used is Complementary Metal Oxide Semiconductor (CMOS). The fruits

are kept in the room temperature and capturing images with different positions and rotations which are take in different time of a day.

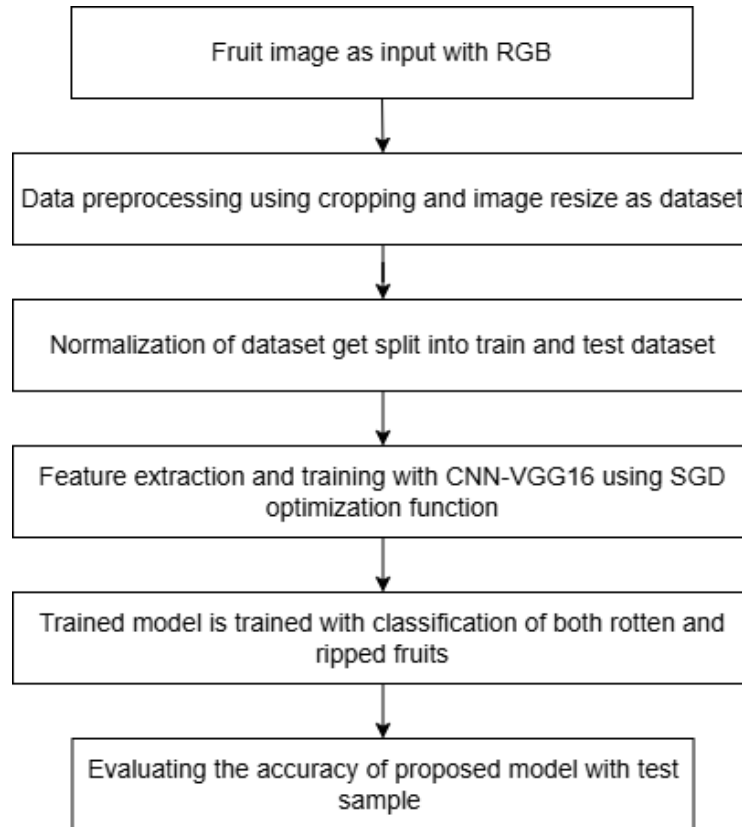


FIGURE 1. Workflow diagram of CNN- VGG16 with SGD architecture

The grade sorted with expert has labeled the fruits with respect to the physical external appearances such as color, firmness, size, shapes and smell. The fruit image with RGB dimensions have minimized to square as 4000 x 4000 pixels with 6378 total images whereas six classes is determined with three dissimilar fruits for ripped and three dissimilar fruits for rotten. The fruits collected images with collection source is illustrated in table 1.

TABLE 1. Collected dataset of rotten and ripped fruit images

Fruit Description	Source collection	Total collection	Removed at preprocessing	Balance
Banana Ripe	Ponnu Super Market, K.P fruit shop	1115	60	1055
Banana Rotten	Ponnu Super Market, K.P fruit shop	1117	55	1062
Papaya Ripe	R.K fruit shop, Ponnu Super Market	1115	59	1056
Papaya Rotten	R.K fruit shop, Ponnu Super Market	1120	50	1070
Orange Ripe	Ponnu Super Market, K.P fruit shop	1115	48	1067
Orange Rotten	Ponnu Super Market, K.P fruit shop	1120	52	1068

After the preparation of data is completed, normalization is performed more than splitting each pixel by 255 and representing the complete pixel in terms of 0 and 1 to help the model learning better. Normalization is the first step in any neural network, and CNNs use deep Forward Neural Networks (FNNs) to acquire invariant feature hierarchy. CNN features are made up of separate properties and invariants that exist in the CNN context and represent the similar transformation which happens at different locations. CNN model has utilized for training by processing data in a spatial relationships or grid topology. The training technique involves using CNN with VGG16.

Working of VGG16 architecture

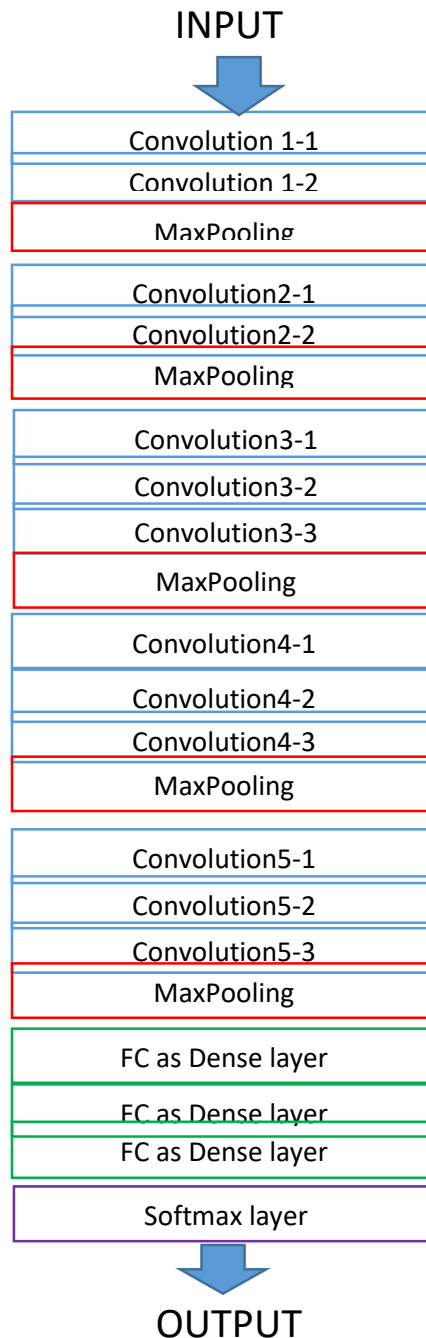


FIGURE 3. VGG16 architecture with CNN model

One of the ease and widely is VGG16 that uses deep convolution network for recognizing large-scale pictures. The RGB input source is transformed into an RBG image with a fixed size of 224 by 224. In general, VGG16 is preprocessed by subtracting the estimated mean value of RGB on the training dataset from each individual pixel.

However, the input image of fruits is produced to be passed via a stack of both convolution layers as well as max pool layer, whilst the filters have performed with a 3 x 3 matrix, which is much smaller in size, to capture the notation from the center, up or down, and left or right. As a result, it is efficient with 7 x 7 stack fields, as well as the convolution and max pool layers utilized in the VGG16 architecture. After the input image has been transformed non-linearly, the 1x1 convolution filter is used to linearly transform the input channel. The convolution stalk and spatial padding layer as input have been preset in a single pixel with 3 x 3 convolutional layers, that decide the spatial resolution subsequent to convolution. There are 5 max-pooling layers, followed by convolutional layers that aid in spatial pooling, and max-pooling occurs in a 2x2 window with stalk 2. Moreover, it consists of Fully Connected (FC) layers, followed by five distinct convolution stacks and a max pooling layer. The first two FC layers are 1x1x4096, demonstrating that each FC layer has 4096 channels, while the final FC layer is 1x1x1000, performing 1000-way ILSVRC classification. The final layer deals with the softmax layer.

SGD optimizer for training process

The ConvNet training technique is often imposed by multinomial logistic regression optimization as a goal via mini-batch gradient descent, also known as SGD. The batch size is set to 256 (0 to 255), and the momentum is up to 0.9. The sum of squared weight (L2) for penalty multiplier is set to 5×10^{-4} in accordance with weight regularization. Each vector requires a hyper parameter, which must be modified to improve the performance of avoiding overfit DL network in Python using Keras. The dropout ratio for the first as well as second FC layers is set to 0.5.

Yet, the learning rate is often set at 0.01 and subsequently dropped by a factor of ten when the validation accuracy stops setting increases. As a result, the learning rate gets reduced by three times, and the iteration ends after 75 epochs. As a result, it is estimated rather than the higher parameter value and depth taken into account by CNN. Furthermore, the network required specific epochs to converge due to pre-initialization by particular layers as well as implicit regularization carried out by reduced convolution filter sizes. As a result, network weights have significance because poor initialization during learning can create gradient destabilization in a deep network. To overcome these concerns, the SGD optimizer was incorporated in the training setup, which was initialized at random. Based on the VGG16 with CNN architecture, the first 5 convolution layers and 3 FC layers are progressed. So, the learning rate associated with the pre-initialized layer remains the same, but it might accommodate for changes while learning. The sample setting of the random layer initialization progresses the normal distribution weight with mean 0, variance 0.01, and bias of 0.

4. RESULT AND DISCUSSION

This experimental research has employed with higher end performing server inclusive of better configuration as an Intel core i7, DMI2 CPU, 100 GB free space with 12GB RAM, and GPU is Quadro K600. The Ubuntu 18.04.3 LST has been utilized as OS for executing the above mentioned GPU DURING image dataset training. The usage of optimizer and loss function for CNN-VGG16 is SGD and cross entropy which is basically pre-packaged with Keras which get import from the module of keras applications. There are 3 major argument considered in the constructor CNN-VGG16 model with weight, input_shape and include_top. The weight constructor represents the weight of checkpoint from the set up model, include_top represents the classifier with numerous connections on the top of the network, and input_shape represents the shape of an image tensor in the network, allowing the network to deal with all types of input sizes. In this study, there are 6378 images at a resolution of 4000 x 4000 pixels. The dataset is split into 70% training data and 30% testing data. In accordance with this research, training image dataset is crucial, which are separated with 60% of training image and 10% of validating image as dataset for the sample fruits of both ripped and rotten are listed in the table 2.

TABLE 2. Training and testing images of realtime fruit dataset

Realtime fruit dataset	Image Count	Fruit status	Training Images		Testing Images
			Training	Validating	
Orange images	1067	Ripped	693	107	267
	1068	Rotten	694	107	267
Banana images	1055	Ripped	686	105	264
	1062	Rotten	691	106	265
Papaya images	1056	Ripped	686	106	264
	1070	Rotten	695	107	268

This study primarily focused on VGG16 rather than hyperparameters with a large number, and VGG16 aims to have convolution layers of 3 x 3 filters with three stalks and maxpooling each, and frequently uses the same padding for 2 x 2 convolution layer filters with maxpooling for two stalks. This configuration is repeated throughout the architecture, beginning with a convolution layer and ending with one max pool layer for each stalk. Finally, it includes three FC layers as a thick layer, followed by a softmax for output. The 16 in VGG16 suggests to the fact that it has 16 layers with weights.

The proposed CNN-VGG16 with SGD model is used to classify rotten items and ripped items whereas the training and validation accuracy, as well as model loss, are shown in Figures 4 and 5, respectively.

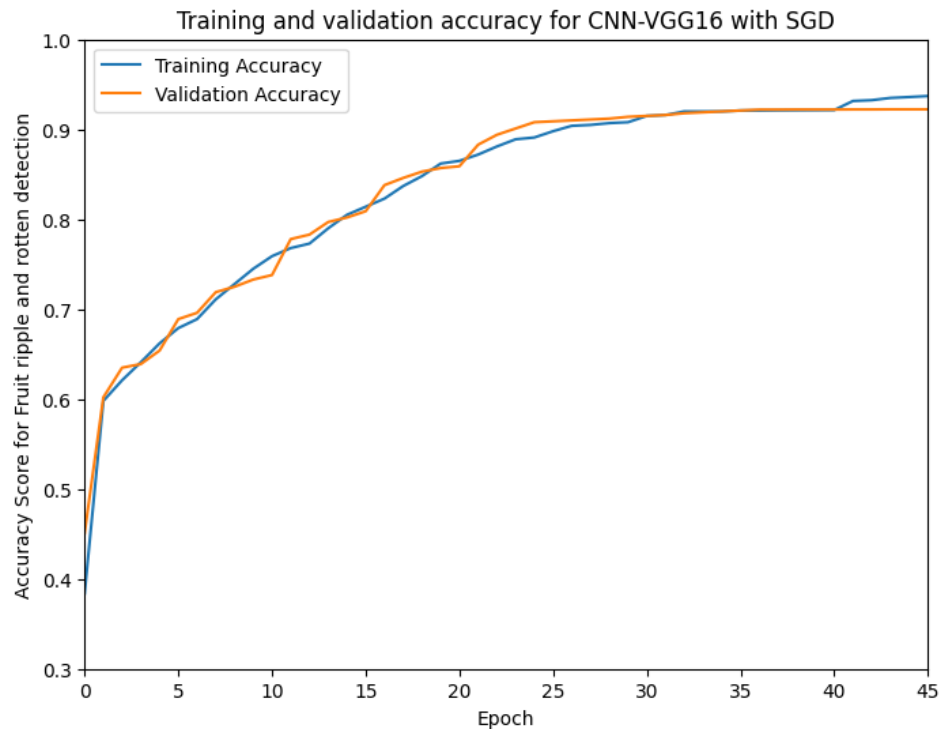
**FIGURE 4.** Accuracy curve for CNN-VGG16 method

Figure 4 shows the model accuracy, whereas the training accuracy remains consistent after 38 epochs, at 92.57%. After 38 epochs, the validation accuracy reaches a stable rate of 94.15%. As the epochs progress, the learning rate improves with each iteration, and after 38 epochs, there is a consistent rate of correctness. This depicts the modification of the learning rate from 0.01 to 0.1 during 45 epochs.

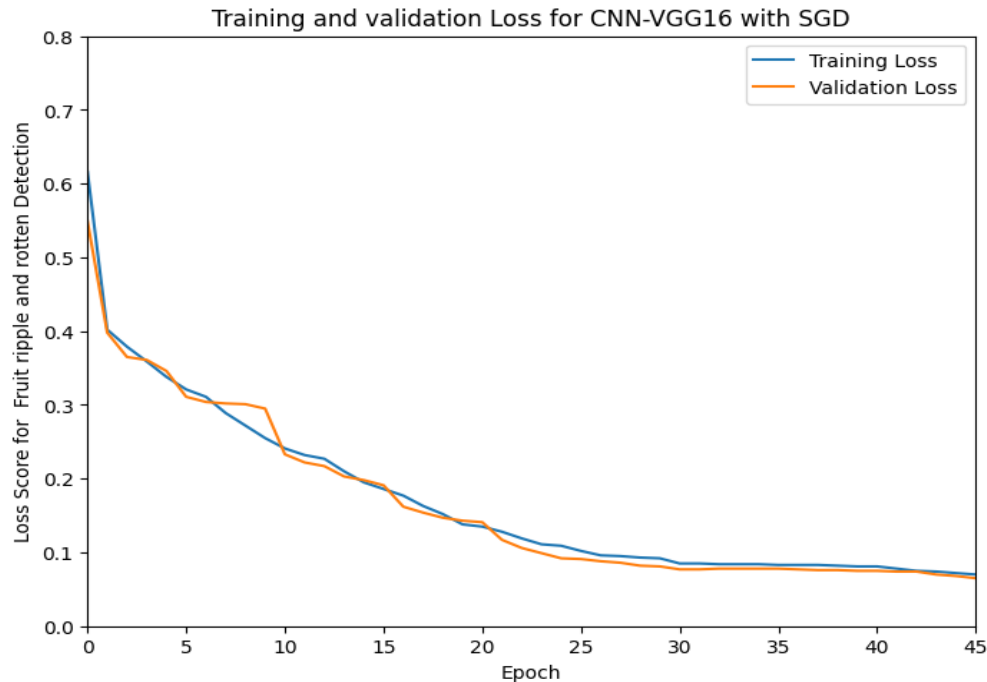


FIGURE 5. Loss curve for CNN-VGG16 method

Figure 5 depicts the model loss in both training and validation, with the loss being greater in training because the model is not initially trained properly. As the epoch increases, the losses decrease from 1.76 to 0.19, indicating that transfer learning occurs to minimize the loss. In the case of a validation loss curve, the loss in the first epoch is lower than the training loss, but it increases after 20 epochs. The difference is more noticeable after 40 epochs; the loss gets decreased to 0.065 significantly till 45 epochs. As a result, the model is well-trained with ripped and rotten fruits, allowing it to classify with 93.39% accuracy.

5. CONCLUSION

In this paper, the CNN-VGG16 classifier has been designed and applied in fruit ripeness and rotten. The proposed approach in two main steps namely VGG16 architecture and the SGD optimizer with learning rate of 0.01. The RGB color image is resized and progressed for training with real-time images of three fruits with six classes to obtain the pre-trained model by training the images with CNN - VGG16 architecture, and the SGD optimizer is used to consider the better learning rate of the training model over multiple epochs. The classifier training settings evolved with VGG16 convolutional architecture resulted in a higher learning rate for ripped and rotting fruit. The suggested CNN with VGG16 achieves higher accuracy (94.15%) and validation accuracy (93.39%), although validation losses are higher than training losses. Hence, the proposed CNN-VGG16 with SGD has better accuracy in classifying the ripen fruits as well as rotten fruits that assist the storekeeper to execute the work with less skill workers and also reduce the expenses by not wasting the fruits by keeping in the store till its became rotten.

REFERENCES

- [1]. Cisternas, I. Velásquez, A. Caro, and A. Rodríguez, "Systematic literature review of implementations of precision agriculture," *Computers and Electronics in Agriculture*, vol. 176, p. 105626, 2020.
- [2]. R. Akhter and S. A. Sofi, "Precision agriculture using IoT data analytics and machine learning," *Journal of King Saud University-Computer and Information Sciences*, vol. 34, no. 8, pp. 5602-5618, 2022.
- [3]. R. Thakur, G. Suryawanshi, H. Patel, and J. Sangoi, "An innovative approach for fruit ripeness classification," in *2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS)*, 2020: IEEE, pp. 550-554.

- [4]. S. Abasi, S. Minaei, B. Jamshidi, and D. Fathi, "Development of an optical smart portable instrument for fruit quality detection," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1-9, 2020.
- [5]. B. N. Gharaghani, H. Maghsoudi, and M. Mohammadi, "Ripeness detection of orange fruit using experimental and finite element modal analysis," *Scientia Horticulturae*, vol. 261, p. 108958, 2020.
- [6]. M. Rizzo, M. Marcuzzo, A. Zangari, A. Gasparetto, and A. Albarelli, "Fruit ripeness classification: A survey," *Artificial Intelligence in Agriculture*, 2023.
- [7]. J. W. Lai, H. R. Ramli, L. I. Ismail, and W. Z. Wan Hasan, "Oil palm fresh fruit bunch ripeness detection methods: a systematic review," *Agriculture*, vol. 13, no. 1, p. 156, 2023.
- [8]. B. Chen, M. Zhang, H. Chen, A. S. Mujumdar, and Z. Guo, "Progress in smart labels for rapid quality detection of fruit and vegetables: A review," *Postharvest Biology and Technology*, vol. 198, p. 112261, 2023.
- [9]. Perumal, Varalakshmi & Narayanan, Vasumathi & Rajasekar, Sakthi. (2020). Detection of COVID-19 using CXR and CT images using Transfer Learning and Haralick features. *Applied Intelligence*. 10.1007/s10489- 020-01831-z.
- [10]. Li L, Qin L, Xu Z, Yin Y, Wang X, Kong B, Bai J, Lu Y, Fang Z, Song Q, Cao K, Liu D, Wang G, Xu Q, Fang X, Zhang S, Xia J, Xia J. Using Artificial Intelligence to Detect COVID-19 and Community-acquired Pneumonia Based on Pulmonary CT: Evaluation of the Diagnostic Accuracy. *Radiology*. 2020 Aug;296(2):E65-E71. doi: 10.1148/radiol.202000905. Epub 2020 Mar 19. PMID: 32191588; PMCID: PMC7233473.
- [11]. B. Chen, M. Zhang, H. Chen, A. S. Mujumdar, and Z. Guo, "Progress in smart labels for rapid quality detection of fruit and vegetables: A review," *Postharvest Biology and Technology*, vol. 198, p. 112261, 2023.
- [12]. A. P. Singh, P. Sahu, A. Chug, and D. Singh, "A Systematic Literature Review of Machine Learning Techniques Deployed in Agriculture: A Case Study of Banana Crop," *IEEE Access*, vol. 10, pp. 87333-87360, 2022.
- [13]. S. Tulli and Yogesh, "Application of Machine Learning for Analysis of Fruit Defect: A Review," *Computational Intelligence: Select Proceedings of InCITE 2022*, pp. 527-537, 2023.
- [14]. R. Thakur, G. Suryawanshi, H. Patel and J. Sangoi, "An Innovative Approach For Fruit Ripeness Classification," 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), 2020, pp. 550-554.
- [15]. PoojaRavindraKamble, Rudrakshi Sanjeev Marathe, Subhankar Kumar Jha, Suraj Sunil Ranvare, Dr. Jayashree V. Katti. (2020). Development of an effective system to Identify Fruit Ripening Stage for Apple, Banana and Mango. *International Journal of Advanced Science and Technology*, 29(12s), 2766-2772.
- [16]. D. Xu, H. Zhao, O. M. Lawal, X. Lu, R. Ren, and S. Zhang, "An Automatic Jujube Fruit Detection and Ripeness Inspection Method in the Natural Environment," *Agronomy*, vol. 13, no. 2, p. 451, 2023.
- [17]. Fatima N. (2020) Enhancing performance of a deep neural network: A comparative analysis of optimization algorithms. *ADCAIJ: Advances in Distributed Computing and Artificial Intelligence Journal*, 9(2):79-90.
- [18]. Darwish A, Ezzat D, Hassanien AE. (2020) An optimized model based on convolutional neural networks and orthogonal learning particle swarm optimization algorithm for plant diseases diagnosis. *Swarm and evolutionary computation*,52:100616.
- [19]. Zhang, H., Hao, K., Gao, L., Wei, B. and Tang, X., (2022). Optimizing deep neural networks through neuroevolution with stochastic gradient descent. *IEEE Transactions on Cognitive and Developmental Systems*, 15(1), pp.111-121.
- [20]. Reyad, M., Sarhan, A.M. and Arafa, M., (2023). A modified Adam algorithm for deep neural network optimization. *Neural Computing and Applications*, pp.1-18.
- [21]. Wojtowysch, S., 2023. Stochastic gradient descent with noise of machine learning type Part I: Discrete time analysis. *Journal of Nonlinear Science*, 33(3), p.45.