International Journal of
INTELLIGENT SYSTEMS AND APPLICATIONS IN ENGINEERING
ISSN:2147-6799 www.ijisae.org Original Research Paper

Tomato Ripeness Detection and Classification using VGG based CNN Models

Seetha Ram Nagesh Appe[1,4], G. Arulselvi[2], G.N. Balaji[3]

Submitted: 28/10/2022 **Revised:** 15/12/2022 **Accepted:** 04/01/2023

Abstract: Ripening is a normal phase in the maturation process of fruits and vegetables. Computer Vision (CV) along with deep learning models provided several opportunities in the field of the agriculture. One of the important applications of CV is to detect and identify the ripeness of the fruits and vegetables. Also, accurately detecting ripening of vegetables using computer vision aids the farmers in better harvesting. For this purpose, deep learning models are used to extracts in-depth features from the images that consumes less time compared to traditional methods. These deep learning approaches, on the other hand, necessitate a big dataset and longer time for image classification. Many studies recommended using the transfer learning method to solve these issues. This paper proposes a model for the tomato ripeness detection and classification using transfer learning which uses VGG16 model. Further, to improve the efficiency of the method the top layer is replaced by a as Multi-Layer Perceptron (MLP) and employed a strategy of fine-tuning approach. The proposed model with fine-tuning approach gives a better efficacy on the tomato ripeness detection and classification.

Keywords: Computer Vision (CV), Deep Learning (DL), Transfer Learning (TL), Multi-Layer Perceptron.

1. Introduction

In India, agriculture plays a key role, and harvesting is still done by hand with the assistance of labour. Since the manual process is costly and time taking, there is great demand in automating agricultural processes like harvesting, pesticide spraying, classification of fruits and vegetables, detection of fruits and vegetables, grading the vegetables. One of the most essential factors in determining the maturity of a fruit is its quality. To distinguish the various phases of fruit ripeness, a grading system has been adopted. Maturity of fruits can be measured from the texture, shape, size and environmental changes. Fruit ripeness detection and classification via computer vision technology exploits on the fruit texture and colour for visual feature evaluation. Ripening is a normal phase in the maturation process of fruits and vegetables. Tomato is one of the commercial crops which contributes to the economy of our country. It is considered as protective crop because of its nutritive value and gives more yield. As it is short duration of crop and demand increases, the cultivation area increasing day by day.

Ripeness detection and classification has become one of the most difficult challenges in computer vision today. [1] [2]. Various deep learning methods have been developed to automate the work which includes harvest detection [3], automatic fruit recognition [4][5], fruit classification [6], fruit ripeness detection [7], fruit disease detection [8]. In general, fruit recognition systems use the textual, shape and color feature descriptions. The feature descriptor algorithm extracts these features from the feature vector of an image.

The most popular feature descriptor that used in deep learning and computer vision is convolutional neural network (CNN). The layered architecture is followed by the Deep Learning algorithms to extract distinct features. The Convolutional Neural Network is a Deep Learning technique that produces significant results in categorization and recognition of images. However, the traditional DL methods require big dataset and more time for image classification. To overcome these problems, many researchers proposed to use the transfer learning method.

[1], Research Scholar, Department of Computer Science and Engineering, Annamalai University, Chidambaram, nageshf25@gmail.com
[2] Associate Professor, Department of Computer Science and Engineering, Annamalai University, Chidambaram, arulselvidhanasekaran@gmail.com
[3] Associate Professor, School of Computer Science and Engineering, Vellore Institute of Technology, Vellore, balaji.gnb@gmail.com
[4] Senior Assistant Professor, Department of Information Technology, CVR College of Engineering, Hyderabad, nageshf25@gmail.com

Transfer Learning is a technique for reusing a model that has been learned on a machine learning job as a starting point for a new task. It allows us to share the learned features across different learning activities as illustrated in Fig. 1. The main advantages of transfer learning are saving training time, produces better performance and fewer data. So, to reduce the demand of big dataset and computational cost, transfer learning method which takes advantage of pre-trained models is used.

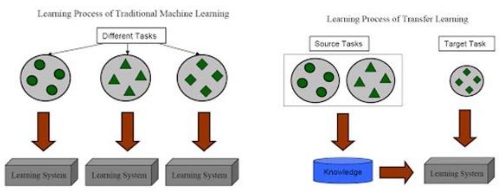


Fig. 1. Traditional Learning vs Transfer Learning

The aim of this paper is to present a model DeepCNN (DCNN) for detecting and classifying tomato ripeness build on VGG16 using transfer learning.

The remainder of the paper is carried out as follows: Section 2 outline the related work on the fruit ripeness detection and classification. Section 3 explains the proposed model of ripeness detection. Section 4 show the experimental results the experimental results as well as an analysis of the outcomes of the proposed methods. Finally, the conclusions are discussed in Section 5.

2. Related work

Over the last several decades, numerous studies have been conducted to assist the assessment of vegetables and fruits quality in industry. Many applications that need visual examination of fruits, such as apples [9], dates [10], mangoes [11], citrus, and pears [12], have employed computer vision-based systems. The authors [13] constructed a classifier for tomato maturity detection with an accuracy of 84 percent based on colour, size, and weight. However, this technique is primarily concerned with size and weight. The degree of ripeness and defects have a big impact on tomato quality.

In [14], a CNN-based model for extracting characteristics from a picture is suggested, as well as a MobileNetV2 model for identifying rotten fruits utilising Max Pooling and Average Pooling. The model was evaluated on a dataset including three distinct types of fruits, and it attained an accuracy of 94.97 percent when using Max Pooling and 93.72 percent when using Average Pooling.

The authors of [15] presented an ANN classifier for determining apple ripeness. The ROI was extracted first, followed by colour features extraction from the segmented image. During the training phase, the classifier classified the image into one of three classes (immature, mature, and mature) based on the input colour features.

In [16], the authors used machine learning to create four classifiers for the classification of mango ripeness stages during harvest. K-means, support vector machine, naive-Bayes, and feed forward neural network are the four classifiers (FANN). The classifiers were trained on mango biochemical properties like titratable acidity and total soluble solids before being tested on physical parameters such as weight and skin colour, as well as electrical properties such as capacitance and voltage The FANN classifier performed well when compared to other classifiers, with an accuracy of 89.6 percent

The machine learning method [17] using CIELab color space and support vector machine for classifying tomato ripeness maturity based on pixel colour produced a mean accuracy of 83.39 percent. The tomato maturity prediction system [18] predicts the maturity level of tomato by depicting the change in surface color of the fruit. The model is entirely dependent on predicting correlations between tomato maturity stages and the temperature. The authors of [19] introduced the FKCNN method by keeping the kernel constant throughout the training process. Instead of considering many features, the method classifies images based on their colour and size. It is appropriate for image classification based on simple visual features.

Deep Learning based novel method is used for classification fruits with CNN architecture [20]. It is used for identification of fruits from 1200 images and achieved the accuracy of 98%. In [21], PCANet deep learning method is used for classification of images by employing multiple filter banks. For indexing and pooling in the model, filters, binary hashing, and block histograms were used. Instead of activation function, a non-linearity is deployed in final layer.

Many researchers now choose to pretrain the model on a large dataset (ex. ImageNet) for feature extraction. However, utilizing a smaller number of examples to train a new classifier dramatically raises the overfitting problem and results in poor observation [22]. Fortunately, numerous techniques such as dropout [23],[24] and data augmentation [25], [26] aids to prevent neural networks from being overfit on which this work concentrates on.

3. Methodology

The overview of the proposed methodology is shown in Fig. 2 consists of four steps: Image acquisition, pre-processing, augmentation, and classification.

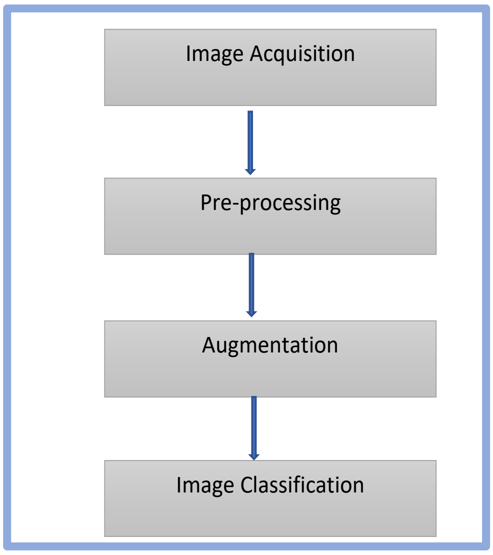
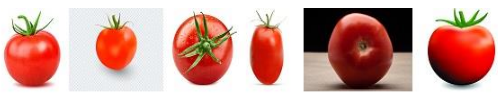


Fig. 2. Steps involved in Proposed methodology

3.1. Image Acquisition

For this work, images of different resolutions and sizes were extracted from Internet Sources and divided into two classes namely ripe as in Fig. 3 (a) and unripe as in Fig. 3 (b). There are 700 ripe and 700 unripe images. All the images are collected in JPEG format and are RGB type where each color channel contains 8 bits per pixel.



(a)

(b)

Fig. 3. Tomato images (a) Ripe and (b) Unripe

3.2. Pre-processing

Initially the pre-processing of images is performed to enhance the features by scaling the pixel values and data augmentation. Images in the dataset are of different sizes, so they need to be resized before being used as input to the model. The images are rescaled to 224 x 224 and given as input to the model.

3.3. Augmentation

Overfitting is sometimes prone to occur when training the model due to the small amount of data gathered, hence small adjustments to the image data are required after pre-processing. In this work the augmentation is performed to increase the dataset through small adjustments such as flips, translations and rotations. The image is translated along the horizontal and vertical axis of the image. The images are flipped horizontally and vertically and rotated randomly by specifying the number of degrees. The result of the augmentation is shown in Fig. 4. Table 1 displays the parameters used during augmentation.

Table 1. Augmentation applied to the Dataset

Parameter	Value
Width & Height Shift	0.5
Rotation	90
Horizontal& Vertical flip	True
Brightness	[0.1,0.7]

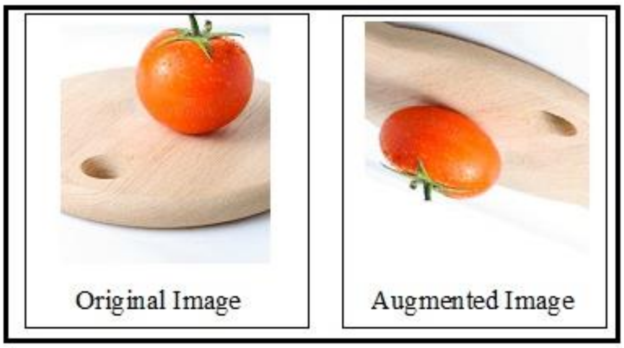


Fig. 4. Image before and after augmentation

3.4. Image Classification

Deep Learning is a subset of the Machine Learning which consists of more layers than Machine Learning. These methods have improved a lot in the field of Image Classification, Speech recognition, Object identification and detection [27].

In this work, a DeepCNN model with transfer learning is proposed without the top layer, utilising a pre-trained model called VGG16. The VGG16 model is trained on a dataset of more than 14 million images having around 1000 classes. VGG16 model consists of 16 layers out of which 13 convolutional layers followed by max pooling and 3 fully connected layers. Fig. 5 shows the VGG16 model with image dimensions mentioned.

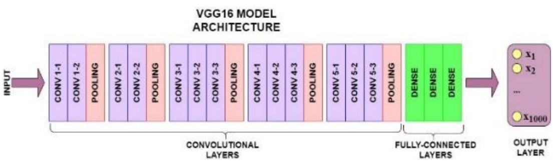


Fig. 5. VGG16 Model Architecture

Fig. 4. illustrates the VGG16 model architecture with the fully-connected layer and generates the 1000 different output labels, whereas the dataset under consideration has only two classes for prediction. Hence, the top-layer of the model is

excluded shown in Fig. 6 and replaced by adding a multilayer perceptron block (MLP) which contains flatten layer, dense layer and dropout.

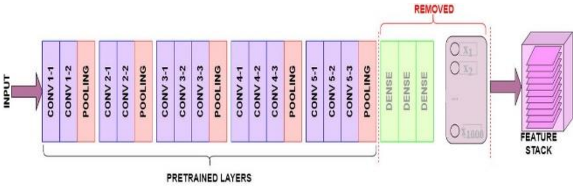


Fig. 6. VGG16 Model without top layer

The proposed DCNN model is built on VGG16 with transfer learning. The images of size 224 x 224 with 3 channels (RGB) are fed into the frozen convolutional layers of the VGG16 model and the visual features are extracted using ImageNet weights. The extracted features are then fed into the unfrozen layers for fine-tuning and next it is connected to the new MLP block which consists of fully connected layer and output layer. The output layer consists of two nodes which represents the two classes (ripe and unripe) with Softmax activation function. The proposed model used for ripeness detection and classification is shown in Fig. 7.

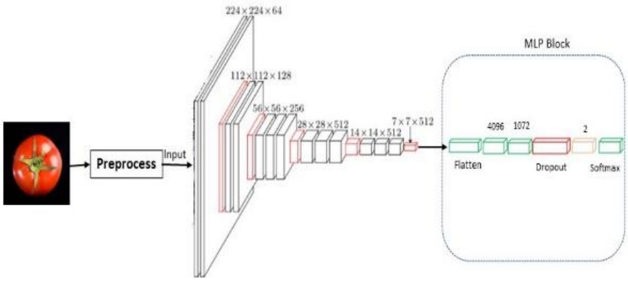


Fig. 7. Proposed DCNN model based on VGG16

As shown in the Fig. 8. the pre-trained convolutional blocks of the proposed model are frozen (non-trainable) to prevent the weights from being updated. By un-freezing (trainable), the last two pre-trained layers are trained on the new dataset to fine-tune the specialized features from the dataset which is called as fine-tuning approach.

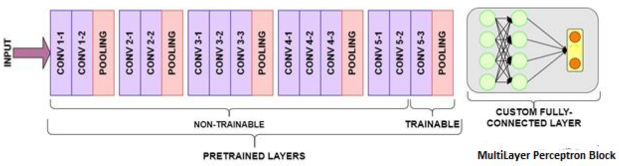


Fig. 8. Layered architecture of proposed DCNN model

4. Results and discussion

The proposed DCNN model is trained by using 70% of dataset for training, 20% for validation and 10% for testing. The images are pre-processed before they are sent into the proposed method. Since the dataset is small, the image augmentation is used to expand the dataset by providing new variations of images at each epoch thereby reducing the effect of overfitting.

Also, to achieve the promising results dropout, early stopping and strategy of fine-tuning is used. The proposed DCNN model is trained for hyper-parameters search and achieved the optimal performance with hyperparameter values as shown in Table 2. During the training process, a 16-batch Stochastic Gradient Descent (SGD) Optimizer with 100 epochs is employed. When the model's performance on the validation dataset starts to deteriorate, the training phase is terminated. After each epoch, the model's performance is assessed by defining a stopping patience of 20 epochs.

Table 2. Hyper parameter values used during training process

Parameter	**Value**
Drop Out	0.3
Total Epochs	100
Early Stopping	20
Optimizer	SGD
BatchSize	16
Initial Learning Rate	1e-3

During the testing process, the performance of the model is assessed by using metrics such as F-Measure, precision, recall and accuracy [30] as mentioned in (1-4).

$$Accuracy(ACC) = \frac{\mathrm{TP_n+TN_n}}{\mathrm{TP_n+TN_n+FP_n+FN_n}} \quad (1)$$

$$Precision(PRE) = \frac{\mathrm{TP_n}}{\mathrm{TP_n + FP_n}} \quad (2)$$

$$Recall(REC) = \frac{\mathrm{TP_n}}{\mathrm{TP_n + FN_n}} \quad (3)$$

$$F1 - Score = 2 * \frac{PRE * REC}{PRE + REC} \quad (4)$$

In the equations (1), (2), (3) and (4), TN_n, TP_n, FN_n, FP_n indicates the true negative, true positive, false negative, false positive respectively.

The process is carried out with and without applying the fine-tuning strategy. The accuracy value of 88.46% is observed when the model is trained by freezing all the pretrained layers and it is shown in the Fig. 8. During the fine-tuning strategy, the training of the model is carried out without freezing the pretrained layers and the performance is increased to 96.66% as illustrated in Fig. 9.

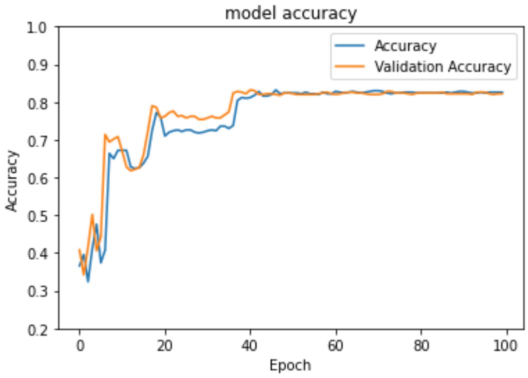


Fig. 8. Training without Fine-tuning

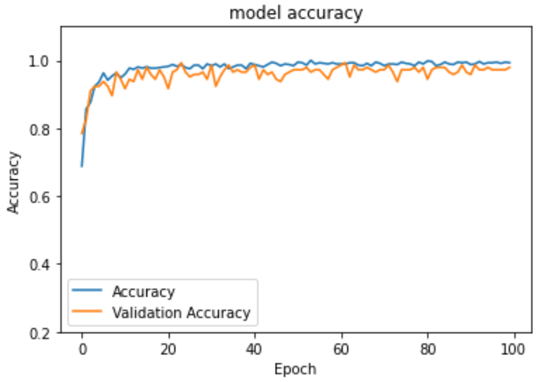


Fig. 9. Training with Fine-tuning

The confusion matrix and classification report of the model without fine-tuning are shown in the Table 3 and Table 4. The precision and recall of the model without fine tuning for ripe and unripe tomato are 88.5%. Furthermore, the model's F1-score is also 88.5 percent, Hence, a fine-tuning strategy is used to increase the model's performance.

Table 3. Confusion matrix using VGG16 without fine-tuning

Confusion Matrix using VGG16 without fine-tuning				
Actual class	**Prediction Class**			**Recall**
		0	1	
	0	66	8	0.89
	1	9	67	0.88
Precision		0.88	0.89	

Table 4. Classification report for transfer learning using VGG16 without fine-tuning

Tomato Classes	**Precision (in %)**	**Recall (in %)**	**F1-score (in %)**
Ripe	0.88	0.89	0.89
Unripe	0.89	0.88	0.89

The confusion matrix and classification report of model with fine-tuning are presented in the Table 5 and Table 6. The F1-score, recall and precision of the proposed model using fine-tuning strategy are higher than those of the model without fine-tuning strategy.

Table 5. Confusion matrix using VGG16 with fine-tuning

Confusion Matrix using VGG16 with fine-tuning				
Actual class	**Prediction Class**			**Recall**
		0	1	
	0	72	2	0.97
	1	3	73	0.96
Precision		0.96	0.97	

Table 6. Classification report for transfer learning using VGG16 with fine-tuning strategy

Tomato Classes	**Precision (in %)**	**Recall (in %)**	**F1-score (in %)**
Ripe	0.96	0.97	0.96
Unripe	0.97	0.96	0.96

5. Conclusion

The DCNN model based on the VGG16 was proposed for detecting and classifying the ripeness of tomato. On providing the dataset as the input to the proposed model, the images are pre-processed, and data augmentation was carried out to improve the efficacy with the help of variations in the images. The results with and without fine tuning approach are compared and it is observed that the proposed model is performing well and having better performance in detecting the ripeness and classifying the tomato fruit. Added the data augmentation process in the dataset and applying the strategy of fine-tuning produced a more robust model. The usage of dropout and data augmentation to reduced overfitting and ensures the robustness of the proposed model. Future study will include using the proposed model to classify real-time tomato images with complicated backgrounds, as well as other fruits and vegetables with larger dataset.

References

[1] Sapan Naik and Bankim Patel. Machine Vision based Fruit Classification and Grading - A Review. *International Journal of Computer Applications* 170(9):22-34, July 2017

[2] Kalia, Parul, Akash Garg, and Amit Kumar. "Fruit quality evaluation using Machine Learning: A review." In 2019 2nd International Conference on Intelligent Computing, Instrumentation and Control Technologies (ICICICT), vol. 1, pp. 952-956. IEEE, 2019.

[3] H. U. Rehman and J. Miura, "Viewpoint Planning for Automated Fruit Harvesting Using Deep Learning," 2021 IEEE/SICE International Symposium on System Integration (SII), 2021, pp. 409-414, doi: 10.1109/IEEECONF49454.2021.9382628.

[4] Nguyen, H.H.C., Luong, A.T., Trinh, T.H., Ho, P.H., Meesad, P., Nguyen, T.T. (2021). Intelligent Fruit Recognition System Using Deep Learning. In: Meesad, P., Sodsee, D.S., Jitsakul, W., Tangwannawit, S. (eds) Recent Advances in Information and Communication Technology 2021. IC2IT 2021. Lecture Notes in Networks and Systems, vol 251. Springer, Cham. https://doi.org/10.1007/978-3-030-79757-7_2

[5] H. Muresan and M. Oltean, "Fruit recognition from images using deep learning", Acta Univ. Sapientiae, Informatica, 2018, pp. 26-42.

[6] Yuesheng, F., Jian, S., Fuxiang, X., Yang, B., Xiang, Z., Peng, G., Zhengtao, W. and Shengqiao, X., 2021. Circular fruit and vegetable classification based on optimized GoogLeNet. IEEE Access, 9, pp.113599-113611.

[7] Saranya, N., Srinivasan, K. and Kumar, S.K., 2021. Banana ripeness stage identification: a deep learning approach. Journal of Ambient Intelligence and Humanized Computing, pp.1-7.

[8] Hakim, L., Kristanto, S.P., Yusuf, D., Shodiq, M.N. and Setiawan, W.A., 2021. Disease detection of dragon fruit stem based on the combined features of color and texture. INTENSIF: Jurnal Ilmiah Penelitian dan Penerapan Teknologi Sistem Informasi, 5(2), pp.161-175.

[9] Tanco, M.M., Tejera, G. and Di Martino, M., 2018, January. Computer Vision based System for Apple Detection in Crops. In VISIGRAPP (4: VISAPP) (pp. 239-249).

[10] Magsi, A., Mahar, J.A. and Danwar, S.H., 2019. Date fruit recognition using feature extraction techniques and deep convolutional neural network. Indian Journal of Science and Technology, 12(32), pp.1-12.

[11] Wu, S.L., Tung, H.Y. and Hsu, Y.L., 2020, December. Deep learning for automatic quality grading of mangoes: methods and insights. In 2020 19th IEEE International Conference on Machine Learning and Applications (ICMLA) (pp. 446-453). IEEE.

[12] Chen, Y., An, X., Gao, S., Li, S. and Kang, H., 2021. A deep learning-based vision system combining detection and tracking for fast on-line citrus sorting. Frontiers in Plant Science, 12, p.622062.

[13] Clement, Javier, Nuria Novas, José-Antonio Gazquez, and Francisco Manzano-Agugliaro. "High speed intelligent classifier of tomatoes by colour, size and weight." Spanish Journal of Agricultural Research 10, no. 2 (2012): 314-325..

[14] Chakraborty, Sovon, FM Javed Mehedi Shamrat, Md Masum Billah, Md Al Jubair, Md Alauddin, and Rumesh Ranjan. "Implementation of deep learning methods to identify rotten fruits." In 2021 5th International Conference on Trends in Electronics and Informatics (ICOEI), pp. 1207-1212. IEEE, 2021.

[15] R. Hamza and M. Chtourou, "Apple Ripeness Estimation Using Artificial Neural Network," 2018 International Conference on High Performance Computing & Simulation (HPCS), 2018, pp. 229-234, doi: 10.1109/HPCS.2018.00049.

[16] Worasawate, Denchai, Panarit Sakunasinha, and Surasak Chiangga. "Automatic Classification of the Ripeness Stage of Mango Fruit Using a Machine Learning Approach." AgriEngineering 4, no. 1 (2022): 32-47.

[17] Garcia, M. B., Ambat, S., & Adao, R. T. (2019, November). Tomayto, tomahto: A machine learning approach for tomato ripening stage identification using pixel-based color image classification. In 2019 IEEE 11th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM) (pp. 1-6). IEEE.

[18] Xiao, Qingmei, Wendi Niu, and Hong Zhang. "Predicting fruit maturity stage dynamically based on fuzzy recognition and color feature." In 2015 6th IEEE International Conference on Software Engineering and Service Science (ICSESS), pp. 944-948. IEEE, 2015.

[19] Shahid, Shamim Ibne, and Md Shahjahan. "A new approach to image classification by convolutional neural network." In 2017 3rd International Conference on Electrical Information and Communication Technology (EICT), pp. 1-5. IEEE, 2017.

[20] Biswas, B., Ghosh, S.K. and Ghosh, A., 2020. A robust multi-label fruit classification based on deep convolution neural network. In Computational Intelligence in Pattern Recognition (pp. 105-115). Springer, Singapore.

[21] T.-H. Chan, K. Jia, S. Gao, J. Lu, Z. Zeng, and Y. Ma, "Pcanet: A simple deep learning baseline for image classification?" IEEE transactions on image processing, vol. 24, no. 12, pp. 5017–5032, 2015.

[22] Sambasivam, G. and Opiyo, G.D., 2021. A predictive machine learning application in agriculture: Cassava disease detection and classification with imbalanced dataset using convolutional neural networks. Egyptian Informatics Journal, 22(1), pp.27-34.

[23] KC, K., Li, R. and Gilany, M., 2021. Joint inference for neural network depth and dropout regularization. Advances in Neural Information Processing Systems, 34, pp.26622-26634.

[24] Kim, H.C. and Kang, M.J., 2020. A comparison of methods to reduce overfitting in neural networks. International journal of advanced smart convergence, 9(2), pp.173-178.

[25] Aquino, N.R., Gutoski, M., Hattori, L.T. and Lopes, H.S., 2017. The effect of data augmentation on the performance of convolutional neural networks. Braz. Soc. Comput. Intell.

[26] Wong, S.C., Gatt, A., Stamatescu, V. and McDonnell, M.D., 2016, November. Understanding data augmentation for classification: when to warp?. In 2016 international conference on digital image computing: techniques and applications (DICTA) (pp. 1-6). IEEE.

[27] Ahmad, J., Farman, H. and Jan, Z., 2019. Deep learning methods and applications. In Deep learning: convergence to big data analytics (pp. 31-42). Springer, Singapore.

[28] Rodrigues, B., Kansara, R., Singh, S., Save, D. and Parihar, S., 2021, May. Ripe-Unripe: Machine Learning based Ripeness Classification. In 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS) (pp. 1-5). IEEE.

[29] Zhang, H., Chen, Y., Liu, X., Huang, Y., Zhan, B. and Luo, W., 2021. Identification of common skin defects and classification of early decayed citrus using hyperspectral imaging technique. Food Analytical Methods, 14(6), pp.1176-1193.

[30] Anand, D., Arulselvi, G., Balaji, G. N., & Chandra, G. R. (2022). A Deep Convolutional Extreme Machine Learning Classification Method to Detect Bone Cancer from Histopathological Images. International Journal of Intelligent Systems and Applications in Engineering, 10(4), 39-47.