Classification of Mango (*Mangifera Indica L.*) fruit varieties using Convolutional Neural Network

Sapan Naik
Babu Madhav Institute of Information Technology,
Uka Tarsadia University,
Surat, India.
sapan2307@gmail.com

Hinal Shah
Babu Madhav Institute of Information Technology,
Uka Tarsadia University,
Surat, India.
hinal.vshah@utu.ac.in

***Abstract* – Automation in classification and grading of mango (Mangifera Indica L.) is important for farmers as well as consumers for identifying quality of mango. This paper addresses the issue of classifying mango fruit based on non-destructive method. Fruit classification is the prerequisite stage of fruit grading. Advancement in deep learning and convolution neutral network (CNN) have been proved to be a boon for the image classification and recognition tasks which can be used for fruit recognition. Here in this paper pre-trained CNN is used for mango classification. Expert knowledge has been collected and mango image dataset of 1028 images has been created with seven different categories of mango. Mango categories are Kesar, Rajapuri, Totapuri, Langdo, Aafush, Dahseri and Jamadar. CNN is tuned and trained according to mango dataset. Four modern CNN network architectures are compared namely Inception, Xception, DenseNet and MobileNet. Experiment results show that the MobileNet model is the fastest and DenseNet is the slowest in terms of execution time out of all four models. Xception and DenseNet model give highest accuracy of 91.42%. Accuracy achieved by Inception is 90% and time required to grade a single mango is 9.78 seconds. Mango classification is also performed using traditional feature extraction method with classifier where Histogram of Oriented Gradient (HOG), Scale-Invariant Feature Transform (SIFT) and Chain code methods are used as features extractor and multiclass Support Vector Machine (SVM) as classifier. 80% accuracy is achieved using this method.**

Index Terms—Convolutional Neural Network(CNN), Mango, Classification.

I. Introduction

Agriculture plays a crucial role in the economy of India as it comprises 16.5% of GDP by sector (2016 est.) with approximately 50% of labor force (2014 est.) and 10% of total export. Even in India, agriculture is sole financial source for 70% of the agricultural labor and common man [1]. For developing country like India, post harvesting procedures are bigger issues. Post harvesting phase normally contains processes like cooling, cleaning, sorting, grading and packing. Sorting and grading are important aspects for analyzing fruits. There are some parameters of non-destructive fruit classification and grading like composition, defects, size, shape, strength, flavor and color [2].

Current grading systems have few limitations. Those are time-consuming, laborious, less efficient, monotonous as well as inconsistent while automatic systems provide rapid, economic, hygienic, consistent and objective assessment. This reason motivated us to propose work in field of post harvesting and mainly in fruit grading. Classification being the initial stage of fruit grading, we have considered it.

In this paper, mango (Mangifera Indica L.) classification is performed on seven different varieties of mango fruit as mango is the extraordinary product that substantiates the high-quality standards and ample of nutrients filled in it. There are 1,000 varieties of mango cultivated in India but only small numbers of varieties cultivate commercially all over India or in other countries. With the largest area under mango cultivation, Gujarat is the strapping mango growing state for economic growth stretching from Jamadar, Totapuri, Dahseri, Neelum, Langdo, Kesar, Payri, Alphonso to Rajapuri [3].

Current trend shows the popularity of deep learning and convolution neutral network (CNN) [4]. The development in deep learning and CNN has been led the field of computer vision and specific of image classification since long time. Deep learning instinctively acquires the features of the images and extracts the global features and contextual details, which drastically reduce the errors in the image recognition [5]. It all started when the Hinton's team received the championship of the ImageNet image classification, at that time surveillance of deep learning has been observed [6]. QuocNet, AlexNet, Inception, BN-Inception-v2 are few of the models proposed later and exhibit superior results. The 70% improvement of the results have been observed as Google trained random 10 million images with neural network of 9 layers and classification performed on ImageNet data set of 2000 categories [7]. PASCAL-VOC- the state-of-the-art detection framework [8] consists of two stages.

Color (RGB) and Near-Infrared (NIR) are combined using early and late fusion methods and used in Faster R-CNN model to detect seven different fruits [9]. Pre-trained R-CNN takes four hours to process fully and to train new fruit. Fruit recognition system presented in [10] uses selective search algorithm and fruit image's entropy for selecting fruit's region; which given as input to CNN. Finally voting mechanism is used for classification. K-means feature learning is used as pre-training process with CNN for Weed identification in [11] where 92.89% accuracy achieved and concluded that fine tuning can improve results. For online prediction of food materials, a fast auto-clean CNN model is proposed in [12]. Auto-clean task and multiclass prediction task based adapting learning are used by the model. The proposed work gives precise and fast output. 7 classes of Mixed Crops images mainly oil radish, barley, weed, stump, soil, equipment and unknown are classified using Deep Convolutional Neural Network in [13] where VGG-16's modified version is used for implementation. 79% accuracy is achieved which shows the potential of deep learning.

Multi-class kernel support vector machine (kSVM) with color histogram, texture and shape features are used for fruits classification in [14]. Split-and-merge algorithm is used for segmentation purpose. Principal component analysis (PCA) is used for dimensionality reduction; Winner-Takes-All SVM, Max-Wins-Voting SVM, and Directed Acyclic Graph SVM are used as multiclass SVM. SVM is used with linear kernel Homogeneous Polynomial kernel, and Gaussian Radial Basis kernel. Results conclude that Max-Wins-Voting SVM with Gaussian Radial Basis kernel performs best with 88.2% accuracy and Directed Acyclic Graph SVM is the fastest. Crop and weed plants are discriminated without segmentation in [15] where Random Forest classifier, Markov Random Field and Interpolation methods are used. Experiments perform on organic carrot which give 93.8% average accuracy. 86% classification accuracy is achieved on 15 classes of 2635 images of fruits in [16]. Here color and texture feature are used with minimum distance classifier. Co-occurrence and statistical features are computed from the sub-bands of Wavelet transform.

This paper presents the solution of mango categories classification using four modern CNN architectural models and also using traditional feature extraction method. Paper is organized as; material and methods are discussed in section II. Result discussion of experiments has been done in section III and finally work is concluded with future directions.

II. Material and Methods

For mango classification, we have assembled the data of almost 100 different varieties of mangoes with their features from Navsari Agriculture University, Gujarat and Paria Farm. Seven easily available and more popular mangoes in south Gujarat region have been selected for experiment. These mangoes are Kesar, Rajapuri, Totapuri, Langdo, Aafush, Dahseri and Jamadar. ***Mix*** image dataset for mango classification has been created. Details of ***Mix*** image dataset is given in below fig. 1.

Below section gives the overview of CNN and how to tune, train and implement the CNN. CNN consists of multiple levels where each level consists of multiple training sets. Input and output of each training stage are images or sets (which are known as feature maps) [17].

MIX DATASET

Mango classification and grading

- Modified background in Photoshop cs5.
- Seven different categories of mangos are collected from the farms of
- Navsari Agriculture University and Amalsad Village.
- Captured in day light with clear weather
- Dataset category wise separation is shown in below table

Mango Category	Small	Medium	Big	Unripe	Partially	Ripe	Total
Aafush	19	82	36	14	31	92	137
Dashehari	9	45	9	52	11	0	63
Jamadar	13	76	57	43	41	62	146
Kesar	84	206	31	42	115	164	321
Langdo	21	61	0	62	8	12	82
Rajapuri	55	90	27	42	64	66	172
Totapuri	50	48	63	81	67	13	161

Number of Images	1082	**Camera distance**	30cm
Resolution of images	2448x2448	**Duration**	May-June 2016
Device used	IPhone 5s	**Temperature**	30 degree to 37 degrees
Views	Top	**Grading parameter**	

Sample Images

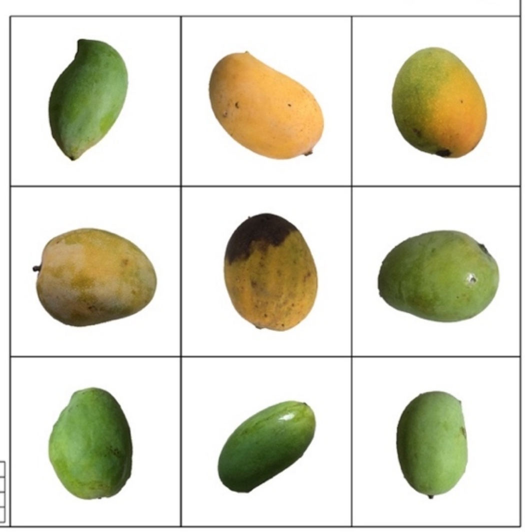

Fig. 1. Mix image dataset - details with sample images

A. *Overview of CNN*

Basic structure of CNN contains four layers namely convolution, non-linearity, pooling and fully connected. The numbers of each layer depend on the structure one has used. Visualization of CNN architecture is available on [18].

1) *Convolution Layer*

First layer of CNN is always Convolutional Layer and input of this layer is always an input image. To understand the working of convolutional layer, let's suppose one image having 32 x 32 x 3 size of pixel values. Assume that one spotlight is shining at the top left corner of an image and shining of spotlight covers the 3 x 3 area. Visualize this spotlight sliding across all the areas of the input image as shown in Fig.2(A).

This spotlight is called a filter or neuron or kernel in machine learning field. The section on which it slides over, is called receptive field. Now this filter is also the array of numbers which are called weights or parameters. Here the filter's dimensions are 3x3x3 because we need to take depth of input image into consideration. Depth of filter is same as that of depth of input image. Initial position of the filter would be at the top left corner of an image. While filter is sliding (called convolving), it is multiplying the values in the filter with the original pixel values of the image. This multiplication is also known as computing element wise multiplications. One single number is computed by summing up all these multiplications. For every location of image, same process gets repeated. Here next step would be moving the filter to the right by 1 pixel or unit, and repeat it. All these moves produce one new number. Once sliding gets over, we will get reduced two-dimensional array of numbers. This two-dimensional array, produced after convolving is called activation map or feature map [19].

What actually happens in Convolutional layer is that, instead of one filter, we use multiple filters of same kind where each filter is used for identifying different feature of image. These filters are actually working as feature identifiers. Small example of curve detector filter is shown in Fig.2(B). As shown in Fig.2(B), if 7x7 filter which represent curve, come across same kind of curve in mango image, it will get multiply and after summing up, produce one big number. This number can be treated as same feature which is available in image. Same way if curve is not present in image, filters produce result, which tends to 0 as shown in Fig.2(C). So, we can say that particular feature is not available in image.

Activation map is the output of the convolutional layer where the activation map illustrates the parts of image; where there are most likely to be feature available. More number of filters tends to give greater depth of the activation map; and this means more information about the input volume one can get [20]. More details about filters and their visualization can be found in [21].

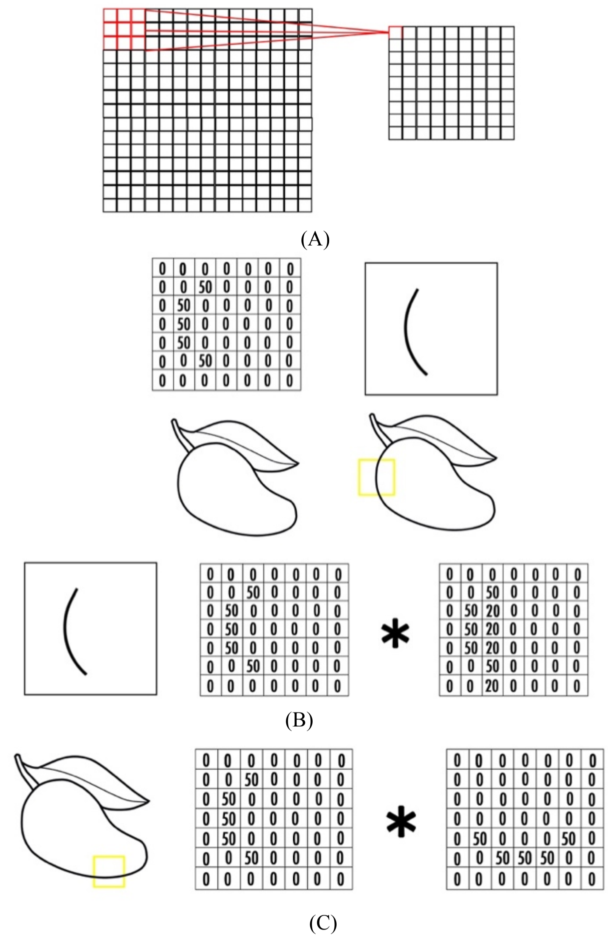

Fig. 2. Operation in convolution layer (A); 3×3 filter representation (B); convolving operation when feature is available (C); convolving operation when feature is not available

2) *Non-linearity Layer*

Various activation functions are applied on this layer. Activation functions are basically Rectified Linear Units (Relu), sigmoid, tanh. Relu is more preferable because the training process gets faster due to it.

3) *Pooling Layer*

Pooling layer comes normally after the convolutional layer to decrease the spatial size. Here the size in terms of height and width gets decreased but not depth. As the number of parameters gets reduced; computation also gets reduced. One more benefit of pooling is, overfitting is being avoided due to less number of parameters. Max pooling is the most common form of pooling. Example shown in Fig.3. Filter of the size m*m with maximum operator is used and apply over the m*m sized part of the image. Instead of maximum operator, if average is used, it will be called as average pooling.

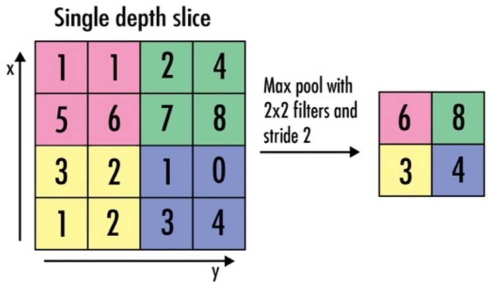

Fig. 3. Example of max pooling in CNN

The common pooling is being done with the filter of size 2*2 and a stride of 2. It basically reduces the half size of input image. After this in flattening step, feature map is converted into single vector because it gives as an input to artificial neural network.

4) *Fully Connected Layer*

In fully connected layer of neural network, input is received by each neuron by previous layer's neurons. Output of this layer is computed using matrix multiplication, followed by bias offset. All neurons in previous layer connects to single neuron to generate specific output.

B. *Tuning CNN Model*

To make the use of CNN, firstly we need to create its model. There are three phases through which tuning of CNN model is done: i) Training, ii) Validate, and iii) Testing.

In Training phase, the network is prepared through which do classification. In Validation phase, the calibration is provided for the network. It corrects the classification performed by the training phase. After all the corrections, model gets ready for testing in Testing phase.

For designing of neural network, one needs to decide many things like arrangement of the layers, types of layers used inside, number of neurons in each layer etc. It is complex to design the architecture of neural network. It is difficult to prepare our own architecture; some standard architectures are available which can be used directly for our work such as AlexNet, GoogleNet, Inception, ResNet, VGG, etc. In the beginning, it is preferable to make the use of standard network architectures [22].

Once the architecture of the network is decided, the next important decision is of weights and biases (the parameters of the network). Backward propagation is used to set parameters in best of manner. Once parameters get finalize and training gets completed all parameters and architecture are saved in binary files. These files are known as model. To test the new input image, this model is loaded and this model will predict the output [22].

Full dataset is not used to train the model. Normally 70% to 80% images are used for training the model and remaining images are used for validation and testing. Suppose, we have total 1600 training images. we split them into small batches of size 16 or 32. This is called batch-size. So, it will take 50 or 100 rounds or iterations to complete full training. This is called epoch. After all, above process gets over, using the same procedure as the training, model will predict the output easily. This time model does not learn from the new input.

C. *Implementation of CNN*

Using transfer learning technique, lots of work get reduced. Here fully trained model is taken which is pre-trained for a set of categories like ImageNet. On this model, the existing weights for new classes can be retrained. All the other layers remain untouched and only the final layer is retrained [23]. This process is faster and does not require graphical processing unit(GPU). Instead of training full new network, this is better alternative and it gives very good results too.

III. Experiments and Result Discussion

Algorithm for digit classification is proposed in [24] using HOG and multiclass SVM. Some more features are found in same reference where SIFT technique is used to extract the features from the image. For classification of mango, shape plays an important role and the chain code is good shape feature extractor [25]. So, we have designed our own chain code for shape feature extraction. Based on the study, basic classification method is implemented by combining features extracted using HOG, SIFT and Chain code techniques; and multiclass SVM as classifier.

Procedure is simple. Inputted image is segmented. Features are extracted from segmented image; and provide extracted features as input to classifier. Due to white background of our dataset, simple thresholding method is used to segment image. Later HOG, SIFT and Chain code features are extracted and provided as input to multiclass SVM.

We took 120 training images of Kesar, Aafush, Rajapuri, Totapuri, Jamadar and Dahseri (Langdo image is not considered, only six categories are experimented). Each category has 20 images. Same way 120 test images are taken. Script has been written in Python for implementation. Experiments give 100% accuracy while giving training images; for testing purpose and 80% accuracy is achieved in case of test images. Time required for classifying single mango is 4.1 seconds for our experiment.

Experiments are performed with the use of MacBook Pro(13-inch,2016) having 2.9GHz Intel Core i5 processor, 8GB 2133 MHz LPDDR3 memory and Intel Iris Graphics 550

1523MB graphics card. For Implementing CNN, we have used TensorFlow. For initial experiments, pre-trained Inception v3 model is used. From this model, old top layer is removed and retrained it with our mango images because none of the mango species are there in the original ImageNet classes. Other than top layer, all the lower layers have been trained for classifying 1000 classes in ImageNet dataset. The weights and biases are directly used to distinguish between new objects recognition tasks. This is the power of transfer learning as discussed before [5].

Initial experiment is done to conclude proper value of training images and epoch for our dataset. We have tested Inception v3 model; by providing training for individual mango category and testing the same. We have used different number of training images (10,20,30 and 40) and epoch values (1500, 2000). Based on this, accuracy for classifying individual mango category is derived. Table I shows the experiment results. The learning rate is set to 0.01; training batch size is taken as 100; testing and validation percentage is kept to 10; for tutoring of deep CNN learning. Based on our initial experiment we have concluded that 60 images/samples of each category for training and 10 images of each category for testing are good enough and selected epoch value for final experiment is 2000.

TABLE I
ACCURACY FOR DIFFERENT TRAINING IMAGES AND EPOCH VALUES ON SEVEN CATEGORIES OF MANGO

		10	20	30	40
Aafush	1500	98	98	96	94
	2000	98	99	97	94
Dahseri	1500	67	68	58	50
	2000	74	75	66	50
Jamadar	1500	99	99	98	97
	2000	99	99	98	97
Kesar	1500	93	82	71	63
	2000	95	86	79	63
Langdo	1500	82	94	87	89
	2000	85	96	90	89
Totapuri	1500	94	89	86	70
	2000	95	91	90	70
Rajapuri	1500	95	97	96	91
	2000	97	98	97	91

Data sets of 1082 images have been prepared which include 137 Aafush, 63 Dahseri, 146 Jamadar, 321 Kesar, 82 Langdo, 172 Rajapuri and 161 Totapuri samples. Photoshop CS5 is used for pre-processing of all image samples. The size of all the images is 512x512x3. As mentioned above, after initial experiment we have used 60 samples (for increasing training set, the images have been randomly transformed to 90 degrees clockwise and anti-clockwise.) for training sets of each category. For testing, 10 images of each category (total 70) are selected. Four CNN architecture models namely Inception v3, Xception, DenseNet and MobileNet are tested for mango classification. For all CNN architecture model same configuration (in terms of training and testing images, epoch value, learning rate, training batch and validation percentage) maintained. The reason for choosing only these four CNN models for experiment is that, accuracy achieved by these models are better compared to other models [26].

Table II and Table III show the experiment results. Values in Table II represents the number of correct prediction out of inputted 10 images for each category of mango for all four models. Table III shows the overall accuracy, error rates and time required for classification of single mango for all four models.

TABLE II
RESULTS FOR NUMBER OF CORRECT PREDICTIONS (INPUTTED IMAGES ARE 10 FOR EACH CATEGORIES)

	Inception v3	Xception	DenseNet	MobileNet
Aafush	9	10	10	10
Jamadar	7	8	10	10
Dahseri	10	9	8	10
Kesar	8	10	9	9
Langdo	9	10	10	10
Rajapuri	10	10	10	6
Totapuri	10	7	7	7

TABLE III
RESULT SHOWING PERFORMANCE OF ALL CNN MODELS

	Accuracy (%)	Error Rate (%)	Time(Seconds)
Inception v3	90	10	9.78
Xception	91.42	8.57	5.10
DenseNet	91.42	8.57	11.52
MobileNet	88.57	11.42	1.09

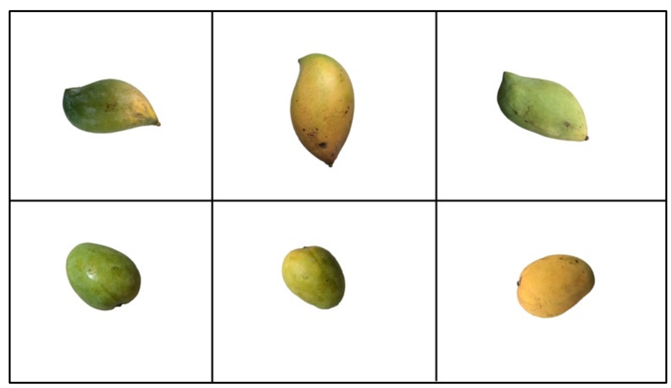

Fig. 4. Sample images of wrongly predicted mangoes

By observing confusion matrix of all four models, we have concluded that misclassification majorly happens with Jamadar and Totapuri category of mango. Figure 4 shows some of the sample images of the wrongly predicted mango by four models. As depicted in samples, the shape, color, size and even texture of these three types are nearly same. MobileNet model has wrongly predicted 4 samples of Rajapuri mango, which is very much unexpected. The shape of Rajapuri is very different from other mango varieties and even other three models have predicted perfectly.

We have not compared mango classification results with other mango classification work because; the datasets used for the experimentation are prepared by us. Other reason is, we have not come across same mango category classification work yet. Though we have compared our results with research work presented in [27] where authors have tried fruit classification using neural network. Table 4 presents the results.

TABLE IV
COMPARISONS OF PROPOSED METHOD WITH DIFFERENT ALGORITHMS

Algorithm	Classification accuracy (%)
GA–FNN	84.8
PSO–FNN	87.9
ABC–FNN	85.4
kSVM	88.2
FSCABC–FNN	89.1
Deep Learning - CNN	91.42 (Xception and DenseNet models)

IV. CONCLUSION AND FUTURE DIRECTIONS

Using deep learning - CNN seven different categories of mango are classified and accuracy achieved using the experiments with Inception v3 is 90% and time required to grade a single mango is 9.78 seconds. Our experiment results show that the MobileNet model is the fastest and DenseNet is the slowest in terms of execution time of all four models. Xception and DenseNet model give highest accuracy of 91.42%. Major misclassification occurs with the Jamadar and Totapuri mango category. Due to their same global features, it is difficult to get accuracy. Local features can be implemented and incorporated to increase classification rate. Proposed methods can be made generalize for fruits of south Gujarat, India. Fine tuning of parameters and combining machine learning methods with CNN can improve the results accuracy. Work can be done to decrease time required to classify a mango. Based on classification, grading and detection of skin disease can be identified.

ACKNOWLEDGMENT

The authors acknowledge the help of Yash Rana, Yash Patel and Purva Desai in implementation, dataset preparation and content writing work.

REFERENCES

[1] Naik, Sapan, and Bankim Patel, "Machine Vision based Fruit Classification and Grading-A Review," *International Journal of Computer Applications*, 170, no. 9, pp. 22-34, July. 2017.

[2] Slaughter, D. C., "Nondestructive Maturity Assessment Methods for Mango," University of California, Davis, pp. 1-18, 2009).

[3] Naik, Sapan, and Bankim Patel, "Thermal imaging with fuzzy classifier for maturity and size based non-destructive mango (Mangifera Indica L.) grading," in *Proc. Emerging Trends & Innovation in ICT (ICEI), 2017 International Conference*, pp. 15-20. IEEE, 2017.

[4] Google Trends, Google. [Online]. Available: trends.google.co.in/trends/. Date:17.06.2018.

[5] TensorFlow Tutorials. [Online]. Available: https://www.tensorflow.org/tutorials. Date:17.06.2018.

[6] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton, "Imagenet classification with deep convolutional neural networks," *In Advances in neural information processing systems*, pp. 1097-1105, 2012.

[7] Le, Quoc V, "Building high-level features using large scale unsupervised learning," *In 2013 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pp. 8595-8598. IEEE, 2013.

[8] Everingham, Mark, SM Ali Eslami, Luc Van Gool, Christopher KI Williams, John Winn, and Andrew Zisserman, "The pascal visual object classes challenge: A retrospective," *International journal of computer vision*, 111, no. 1, pp. 98-136, 2015.

[9] Sa, Inkyu, Zongyuan Ge, Feras Dayoub, Ben Upcroft, Tristan Perez, and Chris McCool, "Deepfruits: A fruit detection system using deep neural networks," *Sensors*, 16, no. 8, pp. 1222, 2016.

[10] Hou, Lei, QingXiang Wu, Qiyan Sun, Heng Yang, and Pengfei Li, "Fruit recognition based on convolution neural network," *In 2016 12th International Conference on Natural Computation, Fuzzy Systems and Knowledge Discovery (ICNC-FSKD)*, pp. 18-22. IEEE, 2016.

[11] Tang, JingLei, Dong Wang, ZhiGuang Zhang, LiJun He, Jing Xin, and Yang Xu, "Weed identification based on K-means feature learning combined with convolutional neural network," *Computers and electronics in agriculture* 135, pp. 63-70, 2017.

[12] Chen, Hao, Jianglong Xu, Guangyi Xiao, Qi Wu, and Shiqin Zhang, "Fast auto-clean CNN model for online prediction of food materials," *Journal of Parallel and Distributed Computing*, 117 pp. 218-227, 2018.

[13] Mortensen, Anders Krogh, Mads Dyrmann, Henrik Karstoft, Rasmus Nyholm Jørgensen, and René Gislum, "Semantic segmentation of mixed crops using deep convolutional neural network," In *International Conference on Agricultural Engineering*, 2016.

[14] Zhang, Yudong, and Lenan Wu, "Classification of fruits using computer vision and a multiclass support vector machine," *Sensors* 12, no. 9, pp.12489-12505, 2012.

[15] Haug, Sebastian, Andreas Michaels, Peter Biber, and Jorn Ostermann, "Plant classification system for crop/weed discrimination without segmentation," In *2014 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pp. 1142-1149, IEEE, 2014.

[16] Arivazhagan, S., R. Newlin Shebiah, S. Selva Nidhyanandhan, and L. Ganesan, "Fruit recognition using color and texture features," *Journal of Emerging Trends in Computing and Information Sciences*, 1, no. 2, pp. 90-94, 2010.

[17] Bhandare, Ashwin, Maithili Bhide, Pranav Gokhale, and Rohan Chandavarkar, "Applications of Convolutional Neural Networks," *International Journal of Computer Science and Information Technologies,* pp. 2206-2215, 2016.

[18] "Convolutional Neural Network," Neural Network - Multi Step Ahead Prediction - MATLAB Answers - MATLAB Central. [Online]. Available: in.mathworks.com/solutions/deep-learning/convolutional-neural-network.html. Date:17.06.2018.

[19] Nielsen, Michael A. "Neural networks and deep learning (2015)," [Online]. Available: http://neuralnetworksanddeeplearning.com (2016).

[20] Deshpande, Adit, "A Beginner's Guide To Understanding Convolutional Neural Networks," A Beginner's Guide To Understanding Convolutional Neural Networks – Adit Deshpande – CS Undergrad at UCLA ('19). [Online]. Available: adeshpande3.github.io/A-Beginner's-Guide-To-Understanding-Convolutional-Neural-Networks/. Date:15/06/2018.

[21] Zeiler, Matthew D., and Rob Fergus, "Visualizing and understanding convolutional networks," *In European conference on computer vision*, pp. 818-833. Springer, Cham, 2014.

[22] "Tensorflow Tutorial 2: Image Classifier Using Convolutional Neural Network," Tensorflow Tutorial 2: Image Classifier Using Convolutional Neural Network – CV-Tricks.com. [Online]. Available. cv-tricks.com/tensorflow-tutorial/training-convolutional-neural-network-for-image-classification/.Date: 15/06/2018.

[23] Donahue, Jeff, Yangqing Jia, Oriol Vinyals, Judy Hoffman, Ning Zhang, Eric Tzeng, and Trevor Darrell, "Decaf: A deep convolutional activation feature for generic visual recognition," *In International conference on machine learning*, pp. 647-655. 2014.

[24] "Digit Classification Using HOG Features," MathWorks. [Online]. Available: https://in.mathworks.com/help/vision/examples/digit-classification-using-hog features.html. Date: 15.07.2108.

[25] Moreda, G. P., M. A. Muñoz, M. Ruiz-Altisent, and A. Perdigones, "Shape determination of horticultural produce using two-dimensional computer vision–A review," *Journal of Food Engineering*, 108, no. 2 pp. 245-261, 2012.

[26] Gogul Ilango, "Using Keras Pre-trained Deep Learning models for your own dataset," [Online]. Available: https://gogul09.github.io/software/flower-recognition-deep-learning, Date:15.07.2018.

[27] Zhang, Yudong, Shuihua Wang, Genlin Ji, and Preetha Phillips, "Fruit classification using computer vision and feedforward neural network," *Journal of Food Engineering*, 143, pp.167-177, 2014.