Non-destructive mango (Mangifera Indica L., cv. Kesar) grading using Convolutional Neural Network and Support Vector Machine

Sapan Naik

Babu Madhav Institute of Information Technology, Uka Tarsadia University, Gujarat, India

ARTICLE INFO

Article history:
Received 00 December 00
Received in revised form 00 January 00
Accepted 00 February 00

Keywords:
Mango Grading
Convolutional Neural Network
CNN architecture models
Support Vector Machine
Feature Extraction

ABSTRACT

Automation in grading of mango (Mangifera Indica L., cv. Kesar) is important to reach consumer demand for quality mango. This paper addresses this issue of mango grading. In this paper, pre-trained Convolutional Neural Network (CNN) is used as feature extractor and Support Vector Machine (SVM) is used as classifier. Mango grading is performed by considering three parameters namely shape, size and maturity. Two approaches are used for feature extraction using CNN. In first approach, CNN is trained using mango samples labelled as class I, class II, class III and class IV. While in second approach, mango grading is performed in three phases. In first phase CNN is trained for shape parameter using deformed and well-formed labels; in second for size parameter using small, medium and big labels; and finally, for maturity parameter using ripe, partially ripe and unripe labels. Based on these three phases, decision of grading is taken. Four CNN architecture models namely Inception v4, Xception, ResNet and MobileNet are compared and used for experiment. In both the approaches, MobileNet performs excellent with highest accuracy and fastest execution time while ResNet performs poor in both approaches for accuracy as well as for execution. Inception and Xception both performs almost same.

Peer review under responsibility of International Conference on Sustainable Computing in Science, Technology and Management.

1. Introduction

Mango is a tropical fruit native to India, which producing more than 50% of world production of mango. Still, according to National Bank for Agriculture and Rural Development (NABARD) report, India's share is just about 15% in the world mango market (Khoje and Bodhe, 2013). The state of Gujarat in India has richest collections of mango cultivators which include Kesar, Jamadar, Totapuri, Dahseri, Neelum, Langdo, Payri, Alphonso and Rajapuri. All these varieties differ in shape, color and flavor, depending on their exporting countries. 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, maturity, firmness, visual defects, flavor and color (Slaughter, 2009). In Gujarat, the size feature is used for grading the fruit.

Traditionally there are trained inspectors who inspect the fruits but it is a tedious task. Being exporting state for mango, Gujarat is still doing mango grading manually with limitations such as it is a time-consuming process, less efficient and the operation is affected due to non-availability of labors during peak seasons. Labor shortages and a lack of overall consistency in the process resulted in a search for automated solutions. Automatic systems provide rapid, economic, hygienic, consistent and objective assessment. Automatic grading definitely improved manual grading system. Most regions of Gujarat consider the size parameter for mango grading which is done using the naked eye observation, which lead us and motivate for this current study.

Due to advancement in technology, image processing and computer vision systems are better choice for grading. Initially for fruit grading; handcrafted features are used; parameters like shape, size and maturity are considered. Unlike manual methods, machine vision based method is time efficient and doesn't require labor which makes it good to implement. As, convolution neutral network (CNN) is popular and is very effective in object classification and identification, it can also be tested for grading system. Deep learning drastically reduces the errors in the image recognition and classification tasks.

Based on above discussion; pre-trained CNN is used as feature extractor and Support Vector Machine (SVM) is used as classifier for Kesar mango grading in this paper. Paper is organized as follows. Literature review is performed in section 2. Discussion of proposed approaches are done in section 3. In section 4, experiment results are shown with discussion on them. Finally, we have concluded and provide the future direction.

2. Literature review

In last few years, conveyor belt based mechanical setup is used for size based grading. Due to technology advancement, machine vision based systems can be prepared. Hardware configuration of such machine vision based automatic grading systems remain standard which consist of illumination device to test fruit under consideration. Solid-state CCD array camera is used to capture image and given to analog-to-digital convertor for converting image in pixel form. Microprocessor system is used which provide storage and containing software for processing image. Color monitor is sometimes available to visualize the image under consideration and effect of algorithms on image (Sun, 2016).

2.1. Related Work

Quality analysis of pizza, fish, bread and cheese has been done using computer vision methods and same way grain quality is examined using same method. In (Brosnan and Sun, 2004), different elements of computer vision systems are presented and described how they are useful in food industry. Learning(classification) methods like artificial neural network, statistical learning, fuzzy logic, genetic algorithm, and decision tree are important for classification and quality evaluation of food. Review of such methods in the context of food quality evaluation is done in (Du and Sun, 2006). Color, texture, shape and size based image feature extraction techniques are reviewed in the context of food industry in (Zheng and Sun, 2006). Other than that, different techniques are compared. Advantages, disadvantages and their feasibility are discussed. Basic theories, principal components and related processing and analytical methods of computer vision are discussed for food industry in (Zhang, Huang, Li, Zhao, Fan, Wu and Liu, 2014). Here relative introduction, latest developments and computer vision systems' applications are discussed.

Review of non-destructive methods for size and volume determination is presented in (Modera, Ortiz-Canavate, Garcia-Ramos and Ruiz-Altisent, 2009). 3-D multispectral scanning is also discussed. Expansions and applications of Hyperspectral imaging(HSI) in detecting, classifying, and visualizing quality and safety attributes of fruits and vegetables with technical challenges and future trends are discussed in (Pu, Feng and Sun, 2015). Mango grading system based on maturity and size features is proposed in (Nandi, Tudu and Koley, 2013) which uses Gaussian Mixture Model and fuzzy logic. Mass estimation of mango using simple linear regression, multiple linear regression and artificial neural network is calculated in (Schulze, Nagle, Spreer, Mahayothee and Muller, 2015) where they received highest accuracy of 96.7% by artificial neural network. Mango features namely perimeter, area, roundness, and percent defect are extracted in (Ganiron Jr., 2014). Mango size feature is extracted in (Khoje, & Bodhe, 2013), Feed Forward Neural Network and Support Vector Machines are used as classifier. Review of image processing algorithms for fruit skin damage detection for tropical fruits of Maharashtra is done in (Khoje and Bodhe, 2013).

Applications of deep learning (CNN) in field of agriculture with 40 research papers and 18 different applications are discussed in (Kamilaris, & Prenafeta-Boldú, 2018). Pre-processing, segmentation, feature extraction and classification methods for color, texture, size, shape and defects features are reviewed in (Bhargava & Bansal, 2018). Pre-trained ResNet-152 and GoogleNet convolutional neural networks are used to extract features from food image datasets namely Food 5K, Food-11, RawFooT-DB, and Food-101; and used to train machine learning classifiers namely artificial neural network, support vector machine, Random Forest, and Naive Bayes (Aguilar, Bolaños, & Radeva, 2017). Food and non-food classification and food category recognition using CNN is performed in (Singla, Yuan, & Ebrahimi, 2016) where 99.2% accuracy for food/non-food classification and 83.6% on the food category recognition are achieved. Fast auto clean CNN is used for online food material prediction in (Chen, H., Xu, J., Xiao, G., Wu, Q., & Zhang, S. (2018)) where proposed model and algorithm give good accuracy. Food recognition system for mobile is proposed in (Pouladzadeh & Shirmohammadi, 2017) with region mining and CNN on FooDD dataset. accuracy of 94.11% is achieved in this. Fruit detection using R-CNN with RGB and NIR is performed in (Sa, Ge, Dayoub, Upcroft, Perez, & McCool, 2016). Dataset of 26 categories has been prepared for fruit and vegetable classification in (Zeng, 2017). CNN is used and accuracy of 95.6% is achieved. 15 different vegetables are classified using DNN in (Ikeda, Oda, & Barolli, 2017).

2.2. Contribution

Nowadays CNN used widely for classification of different variety of food items but not yet explored on fruit grading. Till now work has been done using handcrafted features and classifier. Here CNN is used as feature extractor with traditional linear classifier. CNN capabilities as feature extractor is utilized with SVM classifier for mango grading.

3. Material and Methods

Here for mango grading, CNN is used as feature extractor and SVM is used as classifier. Mango is graded into four classes namely class I, class II, class III and class IV based on three parameters namely shape, size and maturity. Rules for grading are shown in below table.

Table1. Mango Grading rules used for approach 1 and approach 2

Grade	Shape	Size	Maturity
Class I	Well-shaped	Big	Unripe, Partially ripe
		medium	Unripe
Class II	Well-shaped	Big	Ripe
		Small	Unripe
		Medium	Ripe, Partially ripe
Class III	Well-shaped	Small	Partially ripe, Ripe
Class IV	Deformed	Any	Any

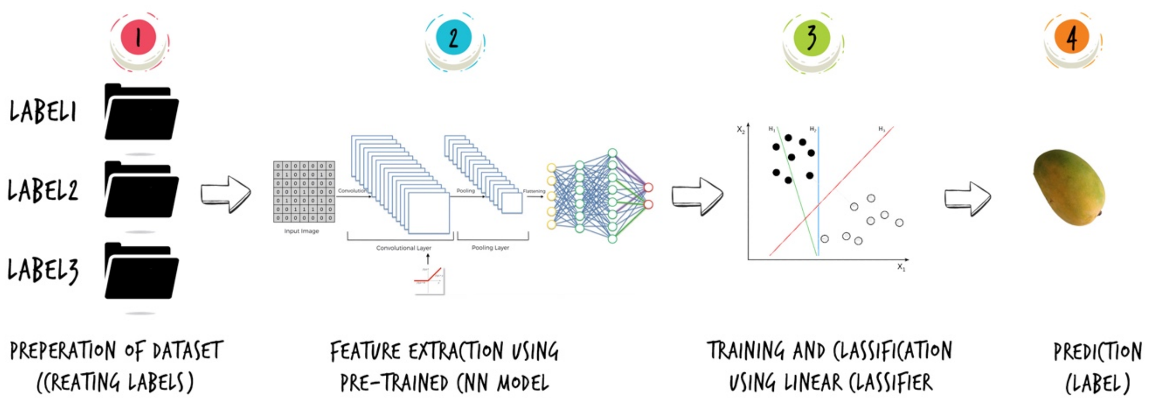

Fig 1. - First approach for mango grading

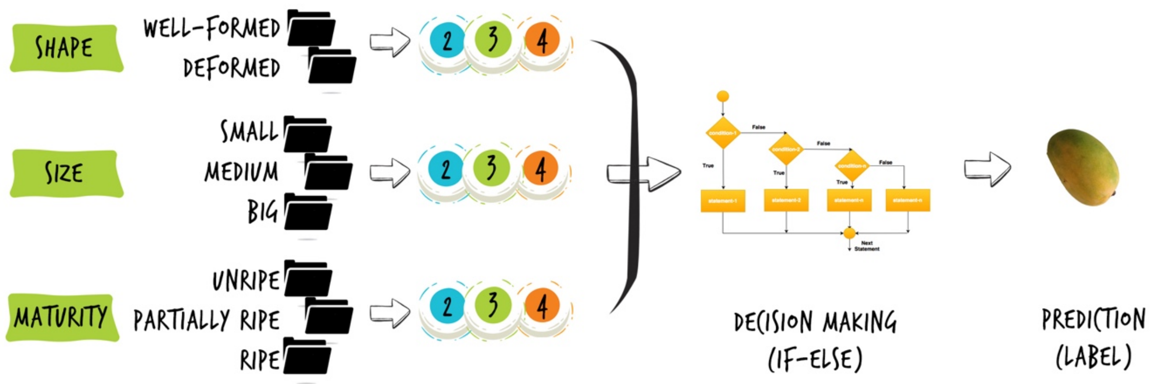

Fig 2. - Second approach for mango grading

Two approaches are used in this paper for mango grading based on above rules. In first approach, CNN is trained using class I, class II, class III and class IV labels While in second approach, mango grading is performed in three phases. In first phase CNN is trained for shape parameter using deformed and well-formed labels; in second for size parameter using small, medium and big labels; and finally, for maturity parameter using ripe, partially ripe and unripe labels. Figure 1 shows graphical representation of first approach and figure 2 is for second approach.

Below in this section, details of dataset are given and data augmentation methods are discussed. Overview of CNN model's training, tuning and implementation is given. Later brief details of CNN architecture models are provided. Lastly in this section SVM is very briefly discussed.

3.1. Dataset preparation and data augmentation

Dataset prepared for the study presented in (Naik, Patel and Pandey, 2015) is used here. Details of this dataset is given in figure 3. The size (resolution) of captured images are 2448×2448 pixels. Images are resized to 224×224 and 299×299 based on CNN model. There are two reasons for resizing. One is, CNN needs fixed size input images and second to reduce computational time.

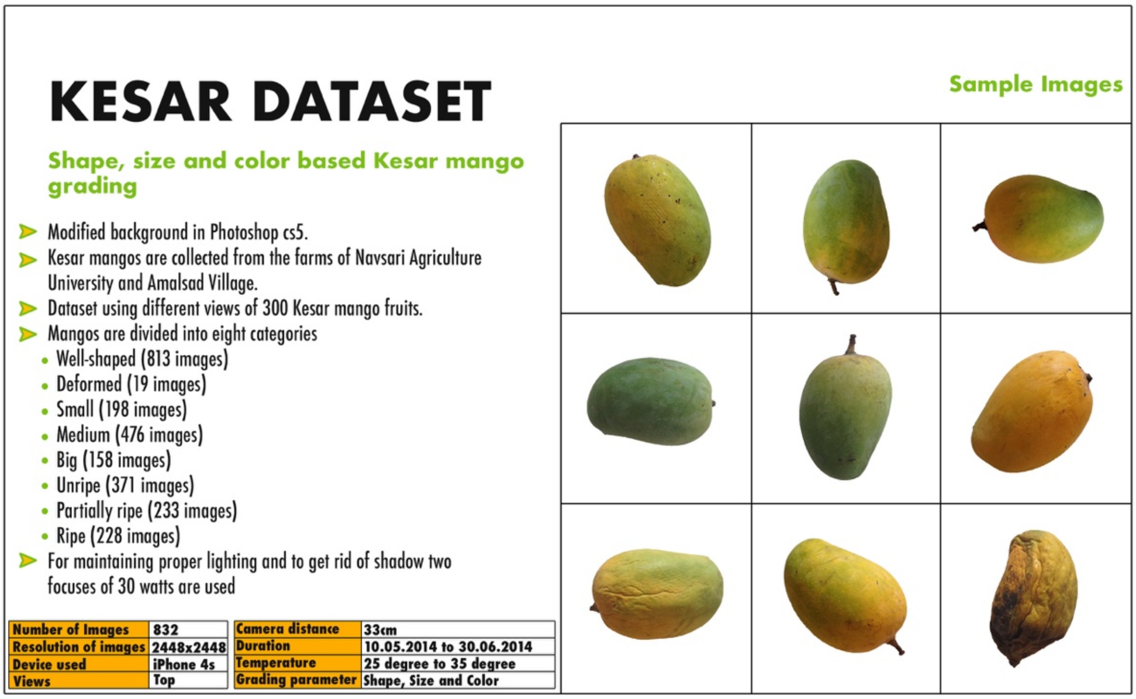

Fig. 3. - Kesar image dataset - details with sample images

For reducing overfitting problem, dataset's size is increased by applying position shifting, rotation and scaling. Even brightness and contrast corrections are performed too, for increasing size of dataset. As CNN doesn't require any pre-processing on images except resizing, no other modifications have been performed on dataset. We have applied data augmentation to 400 randomly selected images and created 1600 new images out of it. This way, we have enlarged the dataset to 2432 images. For testing deformed and well-shaped mangoes, random 500 images are chosen as deformed mangoes.

3.2. Overview of convolution neural network

CNN architecture basically contains four layers, which are convolution, non-linearity, pooling and fully connected. Each of these layers may occur multiple times in architecture. The numbers of each layer are depending on the model under use.

Convolution layer - is the first and top most layer of CNN. Original image is the input to this layer. Filters are applied on input image in this layer. Some literature refer filter as kernel or neuron. Size of filters is normally 3×3×depth, where depth of inputted image and filter are same. Values of filter are multiplied with image pixel value and new array is created as output. This operation is called convolving and output of convolving is two-dimensional array called feature map (activation map) (Nielsen, 2015). In convolution layer, multiple filters are applied where each filter represents one of the features of image like color, curve or intensity. As an example of filter, convolving operation and shape detector filter is shown in figure 4.

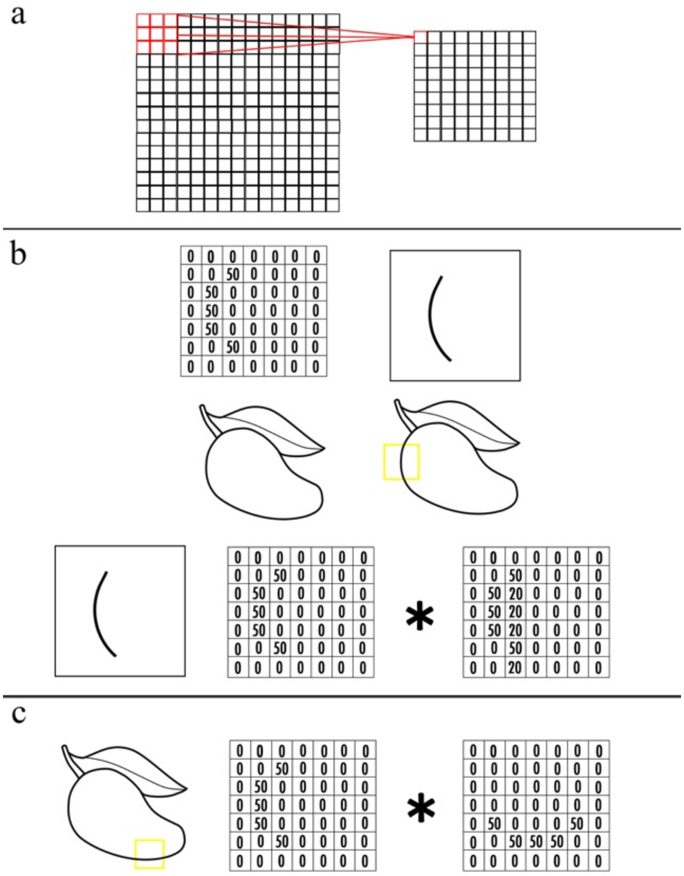

Fig.4. - Operation in convolution layer (A) 3×3 filter representation; (B) convolving operation when feature is available; (C) convolving operation when feature is not available.

Non-linearity layer - different activation functions are applied. Some of these activation functions are Relu, sigmoid and tanh. Due to the advantage of Relu, that it makes the training process faster, it is more popular.

Pooling layer - To reduce the size of activation map, pooling layer is used and even it reduces the chances of overfitting due to less parameter. Common pooling performs with the size of 2×2. MAX pooling is widely used but there is different form of pooling like max, mean and median. Output of pooling layer is given to flattening step where two-dimensional array is converted into single vector. This vector goes as an input to artificial neural network. Below figure 5, shows sample pooling operation.

Fully connected layer - Fully connected neural network is present in this layer where weight and biases are getting set.

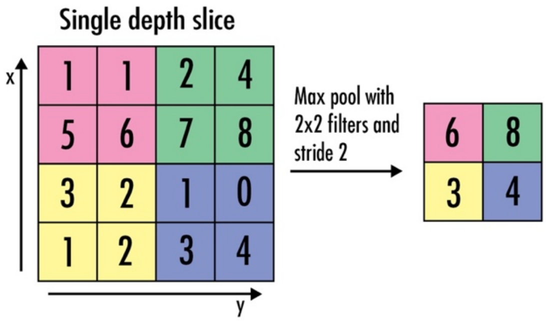

Fig. 5. - Demonstration of max pooling operation.

3.3. Tuning the CNN model

Tuning of CNN contains three phases namely Training, Validating and Testing (Ankit, 2018). In training phase, our model is trained using dataset. After that, model gets validated and finally tested in testing phase using new input for same dataset. Choosing right architecture for CNN or designing the new one is also a part of tuning (Ankit, 2018). Many architectures are available nowadays like GoogleNet, AlexNet, InceptionResNet, VGG, etc. It is preferable to make the use of standard network architecture in the beginning. One needs to decide different parameters of the model like weight, biases, epoch, learning rate, validation and testing percentage during tuning phase. Normally backward propagation method is used to set these parameters. Once training gets completed, all the parameters are saved in binary file known as model. For classifying new image, first the saved model is loaded in same network architecture. Based on the features extracted for new image, it gets classify. This process is called inference (prediction). Training is given in form of epoch (split dataset into batches and give training in iterations).

3.4. Implementation of CNN

Transfer learning technique is used for implementation where fully trained CNN model is directly used with their weights. This model is trained using dataset like ImageNet. The existing weights for new classes get retrained. This method helps in executing (training and testing) CNN on laptop is less time without requiring a GPU. Running time is depend on CNN model under consideration. Though it is not as good as training a full CNN but still it is very effective.

3.5. CNN architecture models

Many architecture models are available for implementing CNN. Here, we have briefly discussed Inception, Xception, ResNet and MobileNet.

Inception - was first proposed by (Szegedy, Liu, Jia, Sermanet, Reed, Anguelov & Rabinovich, 2015) and later in (Szegedy, Vanhoucke, Ioffe, Shlens & Wojna, 2016). This is micro-architecture and initially called GoogLeNet. A weight of Inception requires 96MB of memory space and is smaller compared to VGGNet (Adrian, 2018). Inception uses image distribution, batch normalization and RMSProp methods. It uses 25 million parameters and has ImageNet top 5 error as 5.6%. Inception v5 is the latest version of this model.

ResNet - architecture depends on micro-architecture modules. First, it was proposed by (He, Zhang, Ren & Sun, 2016) where it introduces residual connections. It uses 60 million parameters and has ImageNet top 5 error as 4.5%. In pooling layer, it uses both max and average pooling. The size of this model is 102MB for ResNet50.

Xception - (Chollet, 2017) had proposed this. Xception slightly outperforms Inception v3 model and its size of weight size is 91MB.

MobileNet - normally used for mobile and embedded vision applications where less processing and fast output needed. It was proposed by (Howard, Zhu, Chen, Kalenichenko, Wang, Weyand & Adam, 2017). As mentioned, it is faster in training as well as in execution.

3.6. Support Vector Machine

SVM is powerful classification algorithm, that has shown state of-the-art performance in varieties of classification tasks. SVM is used for both linear and nonlinear data classification. SVM non-linearly maps data to a high-dimensional space using kernel functions. In that high dimensional space, it tries to find the linear optimal hyper plane that separates data with maximum margin. SVM was proposed for only 2-class problems; in multi-class problem SVM is extended using near-against-one or one-against-all strategies. SVM tries to draw a hyper plane between to classes such that the distance between support vectors can be maximized. SVM support vectors, which are nothing but extreme points of both classes that is the reason why SVM is considered special. Basic concepts of SVM working are discussed in (Kirill Eremenko and Hadelin de Ponteves, 2017).

4. Results and Discussion

Experiments are performed on MacBook Pro (13-inch, mid 2012) machine. The machine has 2.5GHz Intel Core i5 processor, 10GB 1333 MHz DDR3 memory and Intel HD Graphics 4000 1536MB graphics card running on macOS High Sierra (version 10.13.6). For implementing CNN and SVM, Keras and TensorFlow libraries are used. Steps used for implementation are listed below.

- step 1. Training dataset is prepared with mango images which contains class I, class II, class III, class IV, deformed, well-shaped, small, medium, big, unripe, partially ripe and ripe labels.
- step 2. Parameters are set in configuration file of CNN (Based on CNN architecture model chosen).
- step 3. Features are extracted from final fully connected layers of pre-trained CNN and get stored. Using transfer learning technique, top layer of CNN model is retrained (Ankit, 2018). Our mango dataset is used for retraining purpose.
- step 4. Machine Learning model (SVM in our case) is get trained for extracted features and labels of step 3.
- step 5. Trained model is evaluated.

Twelve folders are created, each with labels mentioned in step 1. Weights of ImageNet dataset are used for CNN architectures as we are using pre-trained CNN model. Four CNN architecture models namely Inception v4, Xception, ResNet50 and MobileNet are tested for mango grading. The reasons for choosing these four models are; accuracy achieved by these models is good and normally used for CNN performance comparison purpose (Gogul Ilango, 2018).

In approach 1 experiments, 412 images of class I, 440 images of class II, 218 images of class III and 236 images of class IV are randomly chosen for extracting features using CNN model and to provide training to classifier. 20% images of each label are selected for validation and testing purpose. In second approach, CNN is used thrice; for shape, size and maturity feature extraction. For shape, 236 images of deformed and 492 images of well-shaped labels are chosen. For labels small, medium and big; 396,316 and 476 images are chosen respectively. Same way 542, 466 and 456 images are chosen for unripe, partially ripe and ripe labels. All images are chosen randomly. For both approaches, epoch value is set to 1000, learning rate to 0.01, training batch size to 100 and validation percentage to 10.

Table 2 Feature extraction time by CNN models

			Architecture Models			
			Inception v4	**Xception**	**ResNet50**	**MobileNet**
Approach 1		FET	9	11	8	4
		TT	5	1	0.40	1.02
Approach 2	**Shape Parameter**	FET	6	7	5	2
		TT	1.31	0.10	0.18	0.57
	Size Parameter	FET	10	13	10	5
		TT	7.11	0.07	0.07	2.49
	Maturity Parameter	FET	11	14	10	5
		TT	3.53	0.09	0.08	1.58

*FET – Feature Extraction Time (Minutes), TT – Training Time (Minutes)

Features are extracted and stored locally as HDF5 format after configuration. As image size is important while extracting features. Image size of 224×224 pixels are chosen for ResNet50 and MobileNet models and 299×299 pixels are chosen for Inception v4 and Xception models. After feature extraction process gets done, extracted features are provided as input to linear classifier. We have used SVM as classifier. Here using extracted features, SVM gets trained and tested. Feature extraction and training to classifier processes takes different time for different CNN architecture models, as architecture of CNN models are different and size of extracted features are also different. Table 2 shows the time required by each model to extract features.

Experiment results for first and second approach are shown in Table 3, which represents Rank 1 accuracy achieved by each CNN architecture model with SVM classifier. If image under consideration classify with perfect label with highest probability, it is called Rank 1 accuracy; and if it classifies within top 5 probability (may not be first but within first 5 label predicted) than it is called Rank 5 accuracy. As number of class for categorizing mango is less than 5 in our case, our Rank 5 accuracy always remain 100%.

Table 3 Rank-1 and Rank-5 accuracy with training execution time for CNN models

			Inception v4	Xception	ResNet50	MobileNet
Approach 1		**Rank-1 Accuracy (%)**	75.57	77.86	52.67	83.97
Approach 2	**Shape Parameter**	**Rank-1 Accuracy (%)**	94.52	93.15	83.56	91.78
	Size Parameter	**Rank-1 Accuracy (%)**	64.67	65.87	58.68	72.46
	Maturity Parameter	**Rank-1 Accuracy (%)**	79.04	77.25	56.89	82.04

As shown in table 3, MobileNet model gives highest accuracy for approach 1. In second approach, CNN is not providing good results for size parameter. Maximum accuracy for size parameter is again given by MobileNet model. ResNet model is giving minimum accuracy in both approaches. By observing confusion matrix of all models, we have concluded that in approach 1, misclassification majorly happens with class I and class II categories. While in second approach, misclassification majorly occurs with size and maturity parameter. Specific to size, model gives misclassification for medium and big category while in maturity, partially ripe mangoes are categorized either way. We have compared our results with the method proposed in (Naik, Patel and Pandey, 2015) where hand crafted features are combined with fuzzy classifier. Table 4 represents comparison results.

Table 4 Comparisons of leaves classification methods

Manual Grading		**Method Presented in (Naik, Patel and Pandey, 2015)**		**Approach 1**		**Approach 2**	
Accuracy	**Execution Time**	**Accuracy**	**Execution Time**	**Accuracy**	**Execution Time**	**Accuracy**	**Execution Time**
91	3.0	90	2.1	83.97	5.02	70.04	17.48/7.20

Overall highest accuracy achieved in first approach is 83.97% by MobileNet model while in second approach it is 70% by same model. For approach 2, if we do serial implementation, it requires 17.48 minutes for execution while parallel implementation takes 7.20 minutes time for execution. Results show that, CNN is good option for grading. For our experiment, it is also clear that, handcrafted features for size parameter is needed for better grading and result.

5. Conclusion and Future Directions

In India and specific to south Gujarat, automatic non-destructive grading of fruits is the need of current time due to the labour shortage and disadvantages of manual grading. The proposed system is the step towards implementing the same; where shape, size and maturity parameters are considered for automatic mango grading. For implementing the system, CNN is used as feature extractor instead of using handcrafted features and SVM is used as classifier. As discussed, two different approaches are used for feature extraction.

In experiment of mango grading, first approach performs well over second approach. In first approach where four labels are used to train CNN, highest Rank-1 accuracy achieved is 83.97% by MobileNet architecture model and it is fastest too with 5.02 minutes of execution time. MobileNet performs best in both the approaches. The performance of ResNet model is dramatically poor in both approaches. Inception v4 and Xception performs very close to each other in terms of accuracy and execution time. Our experiments show that, majority of the misclassifications occur in second approach and that too in size parameter. This observation leads us and provide direction to use handcrafted features and combined that with CNN extracted features.

For future work, working prototype of the mango grading system can be prepared. Some handcrafted features specific to shape and size parameters can be explored and merge with features extracted by CNN. Even image features like HOG and SIFT can be explored. In current study, SVM works nicely as classifier, but other classifiers like Logistic Regression(LR) and Naïve Bayes can be explored in future. Current study shows the capabilities of CNN for recognition and classification problem, so it can be explored on different varieties of fruits and vegetables, even disease detection in fruits and vegetables can be done using same approach.

REFERENCES

Aguilar, E., Bolaños, M., & Radeva, P. (2017, September). Food recognition using fusion of classifiers based on cnns. In International Conference on Image Analysis and Processing(pp. 213-224). Springer, Cham.

Ankit Sachan, Tensor flow Tutorial 2: Image Classifier Using Convolutional Neural Network, [Online]. Available. cv-tricks.com/tensorflow-tutorial/training-convolutional-neural-network-for-image-classification/. (Accessed on 17/06/2018).

Bhargava, A., & Bansal, A. (2018). Fruits and vegetables quality evaluation using computer vision: A Review. Journal of King Saud University-Computer and Information Sciences.

Brosnan, T., & Sun, D. W. (2004). Improving quality inspection of food products by computer vision—a review. Journal of food engineering, 61(1), 3-16.

Chen, H., Xu, J., Xiao, G., Wu, Q., & Zhang, S. (2018). Fast auto-clean CNN model for online prediction of food materials. Journal of Parallel and Distributed Computing, 117, 218-227.

Chollet, F. (2017). Xception: Deep learning with depthwise separable convolutions. arXiv preprint, 1610-02357.

Du, C. J., & Sun, D. W. (2006). Learning techniques used in computer vision for food quality evaluation: a review. Journal of food engineering, 72(1), 39-55.

Ganiron Jr, T. U. (2014). Size properties of mangoes using image analysis. International Journal of Bio-Science and Bio-Technology, 6(2), 31-42.

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

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 770-778).

Howard, A. G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., ... & Adam, H. (2017). Mobilenets: Efficient convolutional neural networks for mobile vision applications. arXiv preprint arXiv:1704.04861.

Ikeda, M., Oda, T., & Barolli, L. (2017). A vegetable category recognition system: a comparison study for caffe and Chainer DNN frameworks. Soft Computing, 1-8.

Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. Computers and Electronics in Agriculture, 147, 70-90.

Khoje, S. A., & Bodhe, S. K. (2015). A Comprehensive survey of fruit grading systems for tropical fruits of maharashtra. Critical reviews in food science and nutrition, 55(12), 1658-1671.

Khoje, S., & Bodhe, S. (2013). Comparative performance evaluation of size metrics and classifiers in computer vision based automatic mango grading. International Journal of Computer Applications, 61(9).

Kirill Eremenko, Hadelin de Ponteves, Machine Learning A-ZTM: Hands-On Python & R, In Data Science by Udemy, 2017.

Moreda, G. P., Ortiz-Cañavate, J., García-Ramos, F. J., & Ruiz-Altisent, M. (2009). Non-destructive technologies for fruit and vegetable size determination–a review. Journal of Food Engineering, 92(2), 119-136.

Naik, S., Patel, B., & Pandey, R. (2015, July). Shape, size and maturity features extraction with fuzzy classifier for non-destructive mango (Mangifera Indica L., cv. Kesar) grading. In Technological Innovation in ICT for Agriculture and Rural Development (TIAR), 2015 IEEE (pp. 1-7). IEEE.

Nandi, C. S., Tudu, B., & Koley, C. (2014). Machine vision based techniques for automatic mango fruit sorting and grading based on maturity level and size. In Sensing Technology: Current Status and Future Trends II (pp. 27-46). Springer, Cham.

Nielsen, Michael A. Neural networks and deep learning (2015), [Online]. Available: http://neuralnetworksanddeeplearning.com (Accessed on 05/01/2019).

Pouladzadeh, P., & Shirmohammadi, S. (2017). Mobile multi-food recognition using deep learning. ACM Transactions on Multimedia Computing, Communications, and Applications (TOMM), 13(3s), 36.

Pu, Y. Y., Feng, Y. Z., & Sun, D. W. (2015). Recent progress of hyperspectral imaging on quality and safety inspection of fruits and vegetables: a review. Comprehensive Reviews in Food Science and Food Safety, 14(2), 176-188.

Sa, I., Ge, Z., Dayoub, F., Upcroft, B., Perez, T., & McCool, C. (2016). Deepfruits: A fruit detection system using deep neural networks. Sensors, 16(8), 1222.

Schulze, K., Nagle, M., Spreer, W., Mahayothee, B., & Müller, J. (2015). Development and assessment of different modeling approaches for size-mass estimation of mango fruits (Mangifera indica L., cv.'Nam Dokmai'). Computers and Electronics in Agriculture, 114, 269-276.

Singla, A., Yuan, L., & Ebrahimi, T. (2016, October). Food/non-food image classification and food categorization using pre-trained googlenet model. In Proceedings of the 2nd International Workshop on Multimedia Assisted Dietary Management (pp. 3-11). ACM.

Slaughter, D. C. (2009). Nondestructive Maturity Assessment Methods for Mango. University of California, Davis, 1-18.

Sun, D. W. (Ed.). (2016). Computer vision technology for food quality evaluation. Academic Press.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., ... & Rabinovich, A. (2015). Going deeper with convolutions. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 1-9).

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J., & Wojna, Z. (2016). Rethinking the inception architecture for computer vision. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2818-2826).

Zeng, G. (2017, October). Fruit and vegetables classification system using image saliency and convolutional neural network. In Information Technology and Mechatronics Engineering Conference (ITOEC), 2017 IEEE 3rd (pp. 613-617). IEEE.

Zhang, B., Huang, W., Li, J., Zhao, C., Fan, S., Wu, J., & Liu, C. (2014). Principles, developments and applications of computer vision for external quality inspection of fruits and vegetables: A review. Food Research International, 62, 326-343.

Zheng, C., Sun, D. W., & Zheng, L. (2006). Recent developments and applications of image features for food quality evaluation and inspection–a review. Trends in Food Science & Technology, 17(12), 642-655.