



POST-HARVEST GRADING OF CARICA PAPAYA FRUIT USING IMAGE SEGMENTATION AND SOFT COMPUTING

Shwetapadma Panda

Department of ECE

Sambalpur University, Burla, India-768019

Prabira Kumar Sethy

Lecturer in Electronics

Sambalpur University, Burla, India-768019

Abstract: In the era of the computer age, with the development of new technologies, the need to compute with accuracy is increasing. The natural approach for detection of the quality of fruits is done by the experts on the basis of the human eye. Automation of quality analysis of fruits is important in order to reduce human efforts and save time. The paper describes the recent development and application of image analysis and soft computing system in quality evaluation of products in the field of agriculture. Soft computing is a rapid, consistent and objective inspection technique, which has expanded into many diverse industries. Image processing can be used to detect the quality of fruit which includes extraction of morphological features. After feature extraction, feature vectors are formed on which K-Means clustering segmentation process is applied to form clusters. In this paper, we present the framework for Carica papaya grading using the Artificial Bee Colony algorithm (ABC) to classify the papaya fruits from digital images. Our initial experiment on the image features indicates that affected area, shape and textures could be used as the parameters for the ABC algorithm for classifying the papaya fruits into its respective grades. Finally, for papaya grading, a comparison between the performance of GUI using support vector machine (SVM), Naive Bayes classifier and fuzzy logic is done. In the classification process, an input papaya is classified into two categories of healthy and defected. In all grading steps, SVM classifier gives an accuracy of 93.5%, naive Bayes classifier gives 92% and fuzzy logic gives 86.04% respectively. Moreover, the accuracy of the proposed optimization algorithms including for different papaya fruits image databases is 94.04% respectively.

Keywords: Post-Harvest grading; Morphological Features; K-MEANS clustering; Soft Computing; Artificial Bee Colony algorithm (ABC).

1. INTRODUCTION

The main purpose of fruits grading is basically for commercialization. Fruit in the same tree differs in quality such as feature, flavor because their growth was affected by many environmental factors. Especially, fruit from different areas differs significantly in size and quality. Grading may not only standardize fruit product but also promote management of the fruit tree in different areas and product quality. Fruit and vegetable are very difficult to grade exactly and rapidly because of their significant difference in features such as size, shape, and color as a result of changeable conditions of nature environment and manual factors. Grading of fruit and vegetable was performed primarily by visual inspecting in many countries. In the case of manual grading, it has lack accuracy and has lower efficiency because of their individual difference in visual inspecting which is affected by human health condition, psychological condition, lightness, fatigue and so on. Grading of fruits in the fresh form for quality is essential, as the people are quality conscious day by day. Quality analysis of fruits after harvesting is an essential step in the post-harvest system. Grading of fruits on basis of physical characteristics like weight, size, color, shape & freedom from diseases depending upon agro-climatic conditions. But for such grading large expertise man power is required. To overcome this it is necessary to have an automatic fruit grading system, to class different grade. It is indispensable to have non-destructive automatic quality detection technology in order to improve fruits quality detection, the system should have grading efficiency and reduce labor requirement. Carica is a scientific name for papaya fruits. Papaya is the source of vitamins and minerals and is one of

the most popular fruit which is used in health care purpose also used as medicine. Papaya not only used in food, but it is also popularly used in cosmetic industry, household, and medical products. India leads the world in the production of papaya in an annual yield of 3 million tons.

In India, processed fruits are grouped mainly into mature and immature fruits [1]. To grade this, it has identified basically three standard grades for papaya fruits. Based on the affected area it is mainly in three grades i.e. grade A is less affected area of fruits, grade B is more affected area and grade C is severely affected area. As quality analysis process continues to improve and contribute to India's economic, it is facing several challenges in the grading of papaya fruits. All this while, the fruit grading process is done semi-automatic, where a commercial color sorter is used to distinguish one grade that is the premium processed. The rest of grades were identified manually using human vision. In this research, we address the issues currently encountered during fruits grading process.

The main motivation of this work is to explore the possibility of improving the Carica papaya fruits grading process using the Artificial Bee Colony (ABC) algorithm. Specifically, we plan to develop a framework for papaya fruits grading using ABC that includes the study of best image features selection, to compare and evaluate ABC with other available classification techniques. Up to date, only size features have been studied. ABC algorithm has some advantages compared to other evolutionary optimization techniques. Some of the advantages of ABC algorithm include strong robustness, fast convergence speed, and high flexibility. A relative study shows that the performance of ABC is competitive with another algorithm such as GUI using SVM, fuzzy logic, naive Bayes classifier. For quality

assessment of fruits these classification algorithm are very much useful in both real time and software [2].

2. LITERATURE REVIEW

Xiaoling Deng *et al.*, have described Citrus greening detection using visible spectrum imaging and C-SVC. This work has given an accuracy of 91.93% [3]. Xuebing Bai *et al.*, have described a fuzzy clustering segmentation method based on neighborhood grayscale information for defining cucumber leaf spot disease images. This analysis has given an accuracy of 98.78% [4].

Vartika Agrawal *et al.*, have illustrated Feature Selection using Artificial Bee Colony Algorithm for Medical Image Classification. Using this classifier it has given an accuracy of 89% [5].

In this section, we review the related work to grading research and image classification using artificial bee colony method. Research in the grading of fruits was described by the work done by Emrah Hancer *et al.* [6]. The aim of this paper is to model an automatic clustering and grading system using centroidal profile analysis of its digital image. The technique involves data acquisition process to obtain the digital image of fruit and its actual weight. The proposed technique has shown the satisfactory result with a classification accuracy of more than 95%.

Lina Yang *et al.*, [7] have described An Agent-Based Artificial Bee Colony (ABC) Algorithm for hyper spectral image end member extraction in parallel. The objective of this work is to measure the weight using results of the image analysis and then classify them according to their grades. The methodology involves measuring the actual volume and weight of fruit samples and capturing their images. The characteristics results derive formulation based on the collected data could be used to estimate the size of a fruit. In the image processing method, the morphological processing and segmentation using excess green color filter allow papaya images to be precisely distinguished from the background and shadow. The identification ability of the proposed system when tested yields above 90% accuracy.

Edward Thomas *et al.*, [8] have described an artificial neural network for real-time hardwood lumber grading. The methodology involves the acquisition of data to collect the sample images and their weights. In this case the RGB sample images were converted to binary images using automatic thresholding based on the Otsu method. The proposed methodology showed the ability to analyze papaya size classification with more than 94% accuracy in this research. These are some of the few work that has been done by various researcher related to papaya fruits grading.

Dattatraya Londhe *et al.*, [9] have presented a survey report related to all types of fruits grading system. They also described various quality analysis methods for various types of fruits and vegetables. Vardhini Sharma *et al.*, [10] have described grading of fruits on the basis of quality using Image processing. Image processing can be used to detect the quality of fruit which includes extraction of morphological features. After feature extraction, feature vectors are formed on which K-Means clustering is applied to form clusters. The formation of different clusters helps in the grading of fruit.

Ashish Kumar *et al.*, [11] have described detection of a defect on fruits using computer vision methods. This paper

presents the three-dimensional color grading technique for sorting the fruits and detecting defects on it. This work describes a novel defect detection of fruits based on color features with K-means clustering segmentation algorithm. The analysis of various algorithms is shown in Table 1.

Table 1 Analysis of Various Algorithms

<i>Authors & Years</i>	<i>Goals</i>	<i>Future Prospective</i>
Dian Rong, <i>et al.</i> , Embedded vision detection of defective orange by fast adaptive lightness correction algorithm (2017).	Quality analysis of orange using adaptive algorithm method by considering its surface color.	Adaptive algorithm technique will be very much useful in other fruits analysis.
Xuebing Bai <i>et al.</i> A fuzzy clustering segmentation method based on neighborhood grayscale information for defining cucumber leaf spot disease images (2017).	Segmentation of fruits using improved image processing technique.	This technique will be useful in the quality analysis for vegetables.
Edward Thomas <i>et al.</i> , An artificial neural network for real-time hardwood lumber grading (2017).	Neural network classifier used for classification of apple fruits.	This classifier will be simple for classification of other fruits and vegetables.
Souraya Benalia <i>et al.</i> , Computer vision for automatic quality inspection of dried figs (<i>Ficus carica L.</i>) in real-time (2016).	This technique has used for papaya fruits quality analysis.	This technique will be useful in the analysis of vegetables.

A. Limitation of existing work:

- The implementation still lacks in the accuracy of the result in some cases. The more advanced technique is needed.
- Database extension is needed in order to achieve more accuracy.
- Very few diseases fruits have been covered. So, work needs to be extended to cover more diseases.

B. Advantages of proposed algorithm are as follows:

- Use of estimators for automatic Initialization of cluster centers so there is no need for user input at the time of segmentation.
- The detection accuracy is enhanced with proposed algorithm.

3. PROPOSED METHODOLOGY

A digital RGB camera or similar devices are used to taking pictures of papaya fruits different types, and then those are used to identify the affected area on the surface of fruits. Then different types of image processing techniques are applied to them. An algorithm written below illustrated the step by step approach for the proposed image segmentation

and grading processes. The total process of the features extraction and segmentation by taking a particular image is shown in Figure 2.

A. Image Acquisition

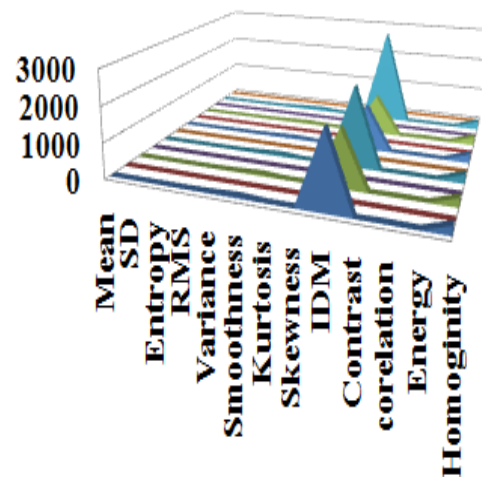
Image acquisition is the initial step of image processing. This process is done by using a digital RGB camera. In this work, we use the digital images of papaya fruits that have been captured in a controlled environment and lighting.

Table 2 Tabulation for all the features of Papaya fruits

Me an	SD	Ent ropy	R MS	V ariance	Smoot hness	Ku rtosis	Ske wness	ID M	Co ntrast	corr elation	En ergy	Homo geneity
0.078876	0.978321	0.762589	0.974878	14.84385	47.81168	1.709878	5.574773	2150.696	1	15.59777	3.632011	255
0.466835	0.865708	0.796721	0.959196	14.15012	48.13958	1.365848	4.313618	1632.216	1	15.7654	3.674427	255
0.367583	0.910197	0.757318	0.964527	16.44411	51.41943	1.667891	5.340374	2305.041	1	13.79264	3.402522	255
0.541238	0.751034	0.538239	0.922202	17.97166	37.66352	2.582884	7.403698	1306.813	1	10.4951	2.588339	255
0.51277	0.710321	0.894702	0.971681	17.1185	35.52045	2.843172	10.45046	1162.225	1	27.60328	4.682011	255
0.697626	0.873892	0.487259	0.910412	31.56037	56.45961	2.982981	8.114045	2844.325	1	4.400836	1.612926	255

B. Image Processing

In the image processing phase, images of various sizes from the samples are standardized to a particular pixel. Resizing images to a standard size is an important task in the Pre-processing phase. Initially the image is segmentation has been done by using K-Means clustering method. The clustering method consists of a pre-defined cluster index, which is a numerical integer number. Then the images were processed by applying various image processing techniques to extract the color, shape and texture features. We extracted the global color features that represent the mean, median, and standard deviation as shown in Table 2. As for the shape feature, we extracted the circularity and solidity of the papaya fruits. We used Gray Level Co-occurrence Matrix (GLCM) for texture extraction. GLCM will be measure by defining the step size in pixels and extract the perimeters of Angular Second Moment, Contrast, Correlation and Inverse Difference, Moment and Entropy. Based on the extracted features, we created a feature vector for each of the image. The graphical representation of all the features parameters are shown in Graph 1.



Graph 1 Feature Extraction Using GLCM Method

C. Grading

In the grading phase, the ABC algorithm will be used to learn and identify the best feature for grading the papaya fruits using ABC optimization algorithm with a classifier.

Upon successful identification of the best feature, ABC is once again used to perform features fusion using its fitness function. The detail of the grading of fruits based on affected area is shown in Table 3. The graph for number of iteration vs. fruits is shown in Graph 2.

D. The ABC Algorithm

Artificial Bee Colony algorithm was proposed for optimizing numerical problems. The performance of ABC is excellent in terms of the local and global optimization and the probability to fall into local optima is very low. It uses only three control parameter such as colony size, the value of limit and the maximum number of the cycle [12]. These three control parameters perform significant roles in ABC algorithm. The number of colony size is based on the number of food sources or the population size of the colony. The algorithm was motivated by the intelligent behavior of honey bees. This algorithm is found simple, robust and population based stochastic optimization algorithm.

In ABC algorithm, the colony consists of three groups of bees which have their own roles, which are:

- Employed bees
- Onlookers
- Scout

The first half of the colony consists of the employed bees and second half include the onlookers. There is only one onlooker bee for a food source; there is only one employed bee. The employed bee of an abandoned food source would eventually become a scout.

The detailed of ABC algorithm pseudocode is given below:

- Initialize a population of search space solution x_{ij} .
- Evaluate the fitness of the population using:

$$f(x) = \sum_{t=1}^{SN} X_2^t \quad (1)$$

- While stopping benchmark is not met, stopping condition is a maximum number of iteration. The cycle=1 and respective.

- Produce new solutions (food source positions) U_{ij} in the neighborhood of X_{ij} for the employed bees using:

$$U_{ij} = x_{ij} + \Phi_{ij} (x_{ij} - x_{kj}) \quad (2)$$

Where k is a solution in the neighborhood of i , Φ is a random number in the range $[-1, 1]$ and evaluate them.

- Apply greedy selection process between x_i and Calculate probability values P_i for the solutions x_i by means of their fitness values using:
- Calculate probability values P_i for the solutions x_i by means of their fitness values using:

$$P_i = \frac{f_i}{\sum_{t=1}^{SN} f_t} \quad (3)$$

To calculate fitness values of solutions, the equation below is applied:

$$f_i = \frac{\xi}{\{1 + abs f(1), i\}} \quad (4)$$

- Till the termination criteria, it will continue the same process.

E. Kappa Coefficient Calculation

Kappa is a quality analysis constant, which is used for quality analysis of any product. In the current work we give the optimum features matrix value into the kappa value. According to the maximum and minimum value or according to the range of value, the grading of fruits has been occurred.

4. STEP BY STEP APPROACH OF PROPOSED ALGORITHM

- In the initial step of the proposed method, initialize the area of fruit position. After initialization, the affected area of fruit has been calculated among total area.
- In the next step, the position of the new fruit area has been determined by the employed bee. Then the amount of affected area has been calculated.
- Then make a condition that all onlooker bees have been distributed. If yes, proceed to the next step and the best fruit has been memorized by the onlooker. If no, the new fruit area has been selected by the onlooker and a neighbor fruit area position has been determined for the onlooker.
- Further new position for the exhausted fruit has been produced. In this situation one condition has applied, if the termination condition satisfied then shifts to the final fruit position. If not satisfied, then again go the initial position of the algorithm and the loop has been continued till it satisfies the termination condition.

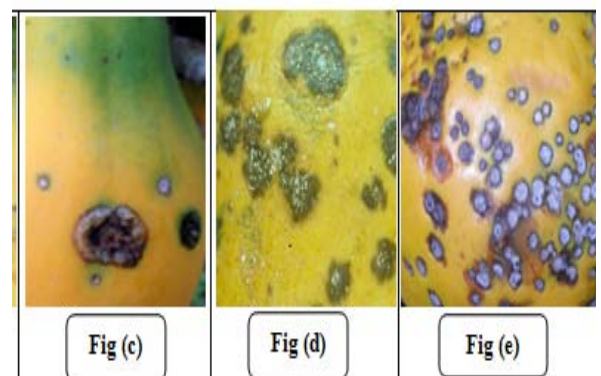
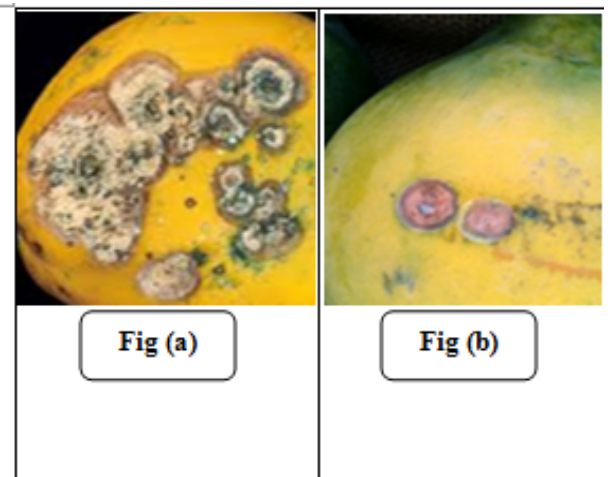
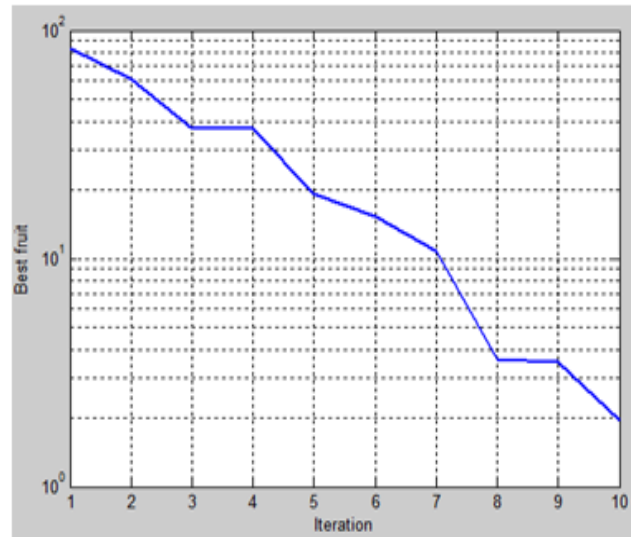
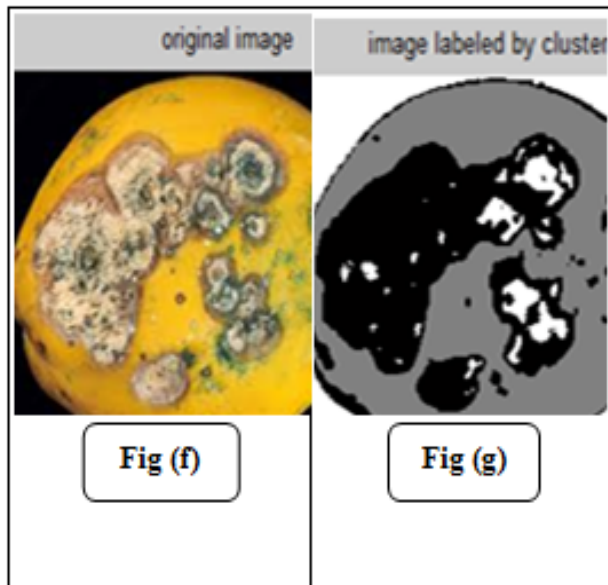


Figure 1 Various Disease Affected Carica Papaya fruit images

5. RESULT & DISCUSSION

In the proposed algorithm, first of all, feature extraction process has occurred. During this process, all features of a particular image have been extracted. By considering those features we can grade the papaya fruits by using any type of soft computing technique or any type of optimization technique. In my work, I have used the bee colony optimization technique for grading of Carica papaya fruits. The algorithm was implemented using Matlab Version 2013. It was applied to a set of 6 images collected from various places. Figure 1 shows the obtained results.



Graph 2 Graphical Representation of No. of Iteration vs. Fruits

Table 3 Grading of Carica papaya fruits in basis of affected Area

Sl. No.	Dimension of Each Images	Accuracy for Grading in Percentage (%)			
		SVM	Fuzzy Logic	Naïve Bayes	ABC+SVM
1	176*287*3	93.71	90.63	89.00	94.00
2	480*356*3	94.89	89.98	89.98	94.58
3	580*380*3	93.89	91.25	90.00	93.89
4	409*276*3	93.00	90.79	88.79	93.90
5	398*250*3	93.00	90.00	89.00	94.00
6	298*290*3	92.97	90.40	88.99	93.98
Overall Accuracy		93.87	90.66	89.44	94.09

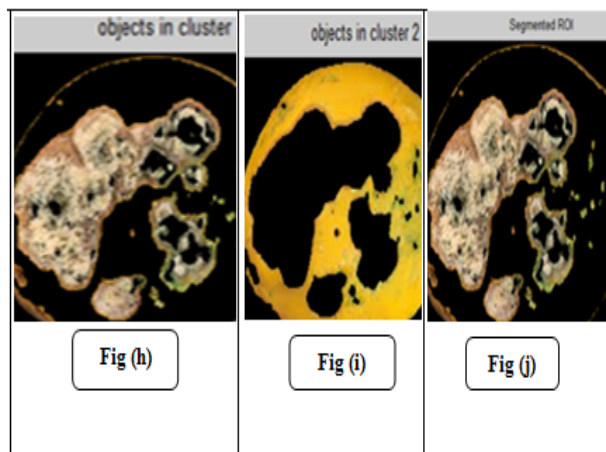


Figure 2 Segmentation of Image by using K-MEANS clustering method. (f) Original Image, (g) Image labeled by cluster Index, (h) Objects in cluster1, (i) Objects in Cluster2, (j) Segmented ROI

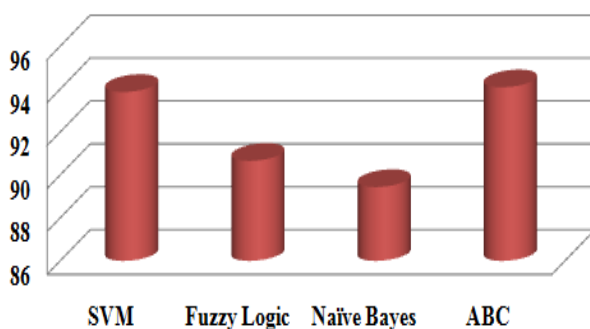
Table 4 Comparison results for Accuracy

Characteristics	Grade Name		
	Grade A	Grade B	Grade C
Total Area of one side of the Fruit Image (X) (In Percentage)	36.39%	30.90%	26.89%
Affected Area or Optimum Value (Y)	16.73%	13.00%	10.37%
(Y/X) *100	54.02%	57.92%	61.43%

6. COMPARATIVE STUDY

The two major differences between the existing algorithm and the proposed algorithm were - accuracy and time complexity. The existing algorithm was only implemented for yield estimation of green Papaya fruits but the proposed algorithm was successfully implemented for non-green and approximate elliptical fruits such as mangoes, watermelon, brinjal, jackfruit, and guava. Proposed algorithm has given similar accuracy as compare with the existing algorithm as the former had better time complexity. Theoretical time complexity analysis of both the algorithms showed that proposed one better than the existing one. All the experiments are performed in MATLAB. For input disease image, samples of papaya fruits images are considered. First, of in all the cases, the grading is first done using the Minimum Distance Criterion with K-Mean Clustering and shows its efficiency with an accuracy of 89.54%. The accuracy is improved to 94.09% by proposed algorithm. In the second phase, classification is done by using the existing algorithm like SVM using graphical user interface, fuzzy logic, and Naïve Bayes classifier as shown in Table 4. Using SVM classifier and shows its efficiency with an accuracy of 93.87%. In the case of the fuzzy logic method, it shows the accuracy of 90.66% and for Naïve Bayes classifier, it gives an accuracy of 89.44% respectively. Further graphical representation of accuracy is shown in Graph 3.

Comparison Result for Accuracy



Graph 3 Graphical Representation of Comparison Result for Accuracy

7. CONCLUSION & FUTURE WORK

This paper presents a new technique for successful grading of Carica papaya fruits. The new technique which has discussed in this paper is artificial bee colony algorithm

(ABC), is an optimization technique. In this paper, we have successfully discussed the technique and also made a comparative analysis between existing soft computing techniques with the proposed one. The soft computing techniques such as SVM using GUI, Fuzzy Logic and Naïve Bayes classifier. The results of comparison have shown in above with tabulation form and also in the graphical method, both for accuracy and time complexity. From all the analysis and comparative study, we have found that the proposed method has better accuracy compare to other soft computing technique. But time complexity is more comparing to another algorithm. So one can develop such an algorithm which will perform an operation with minimum time and with high accuracy.

8. REFERENCES

- [1] Dian Rong et al. "Embedded vision detection of defective orange by fast adaptive lightness correction algorithm" Computers and Electronics in Agriculture, journal homepage: www.elsevier.com/locate/compag.
- [2] Souraya Benalia et al. "Computer vision for automatic quality inspection of dried figs (*Ficus carica* L.) in real-time" Computers and Electronics in Agriculture, journal homepage: www.elsevier.com/locate/compag.
- [3] Xiaoling Deng et al. "Citrus greening detection using visible spectrum imaging and C-SVC" Computers and Electronics in Agriculture, journal homepage: www.elsevier.com/locate/compag.
- [4] Xuebing Bai et al., "A fuzzy clustering segmentation method based on neighborhood grayscale information for defining cucumber leaf spot disease images" Computers and Electronics in Agriculture, journal homepage: www.elsevier.com/locate/compag.
- [5] Vartika Agrawal et al., "Feature Selection using Artificial Bee Colony Algorithm for Medical Image Classification" 978-1-4673-7948-9/15/\$31.00 ©2015 IEEE.
- [6] Emrah Hancer et al., "ARTIFICIAL BEE COLONY BASED IMAGE CLUSTERING METHOD" WCCI 2012 IEEE World Congress on Computational Intelligence June, 10-15, 2012 - Brisbane, Australia.
- [7] Lina Yang et al., "An Agent-Based Artificial Bee Colony (ABC) Algorithm for Hyperspectral Image Endmember Extraction in Parallel" 1939-1404 © 2015 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission. See http://www.ieee.org/publications_standards/publications/rights/index.html for more information.
- [8] Edward Thomas, "An artificial neural network for real-time hardwood lumber grading" Computers and Electronics in Agriculture, journal homepage: www.elsevier.com/locate/compag.
- [9] Dattatraya Londhe et al., "Grader: A review of different methods of grading for fruits and vegetables" Agric Eng Int: CIGR Journal Open access at <http://www.cigrjournal.org> Vol. 15, No.3, September 2013.
- [10] Vardhini Sharma et al., "Grading of fruits on the basis of quality using Image Processing" SSRG International Journal of Computer Science and Engineering (SSRG-IJCSE) – volume 3 Issue 5–May 2016.
- [11] Ashish Kumar et al., "Detection of Defect on Fruit using Computer Vision Technique" International Journal of Research in Advent Technology, Vol.4, No.2, February 2016.
- [12] Dervis Karaboga, "A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm" Published online: 13 April 2007 Springer Science Business Media B.V. 2007.