

Automated Mango Fruit Recognition by Multi-Task Convolutional Neural Networks for Harvest Robot

Nekkalapu Gopi, Rama Koteswara Rao P

Abstract: Efficient and effective mango fruit recognition is viewed as significant for development of a smart agriculture robot (ARo) for yield prediction, pest control, sorting and fruit detection. Several fruit recognition techniques for structuring ARo have been employed during the most recent decades. Recently, ordinary natural fruit identification techniques are lacking progressive response, exactness and extensibility. In this paper, we proposed an improved algorithm of MTCNN (Multi-Task Cascaded Convolutional Network) based on IFD (Intelligence Fruit Detection) technique. This technique has the ability to make the ARo work progressively with high precision. Additionally, in view of the connection between the number of tests on dataset and the boundaries of Neural Networks advancement, this work presents an improved strategy. A method that depends on image clustering is used to improve the identification in this project. The experimental results exhibited that the proposed identification performed significantly, both as far as exactness and time-cost. Besides, the broad trials exhibited that the proposed strategy has the limit and a decent compactness to work with other associated systems.

Keywords: Cascaded Convolutional Neural Networks, Fruit Recognition, Automated Robot, Image Fusion.

I. INTRODUCTION

The natural fruit identification for yield estimation, grading of varieties, disease control and different applications in farming field have been getting fame in the course of recent decades [1]. A few designs have been developed for the Agriculture Smart Robot (ARo), which have prompted impressive improvement in the forming sector [2]. Especially, recognition and characterizing natural fruits as indicated by their quality have been one of the most well-known examination fields attracting the majority of the agriculturists. Product identification is significantly the most important aspect to be considered top to bottom investigations regarding the agriculture matter. Along these lines, numerous scientists have put forth attempts for quite a long time to create powerful algorithms for the fruit recognition [3]. In spite of the fact that, the fruit recognition algorithms have been improved amazingly, they are still differ from practical applications. The essential challenges are growing in fruit recognition framework and over the top situations of plantations in fruit gardens.

These features incorporate various testing undertakings, for example, insufficient or over enlightenment, vague

foundations, substantial impediment by neighborhood trees or foliage, low-contrast, variety of posture, etc.

The negligence of benchmark is another incredible test for mango fruit identification. An adequate measure of test images assumes a significant job in machine learning based models [4]. In this project, we gathered pictures from mango plantation by the advanced cameras. Thereafter, we chose the appropriate ones and marked them to make a dataset. Making a dataset physically is a monotonous and tedious work. So we concentrated another increased strategy dependent on augmented procedure. The inspiration for this combined strategy [5] originated from the rule that the created new examples ought to be close to real pictures. Strengthening measures were made for assorted variety by including fusion method that would help to improve the outcome of this identifier. So as to assess the system whether it could be applied to different sorts of items advantageously, we prepared the identifier on two different natural products species (apple and orange) too.

To sum up, our findings are the following:

1. We proposed an improved algorithm for mango fruit identification called Mango-MTCNN (M-MTCNN) by improving the benchmark model of MTCNN. Also, this identifier has the advantage of high precision and less time-utilization.

2. We proposed a novel enlarged technique called Combined Augment (CA). We create random pictures by including negative patches from test datasets by editing the supplement samples.

3. The proposed approach can be applied to different variety of fruits with a limited quantity training tests.

The remaining part of the paper is organized as follows. In Section II, we survey earlier related work in fruit recognition in agriculture. Parts III, IV, we present strategy implemented in this investigation. Our trials in this experimentation are furnished in Section V. In Section VI, we perform the analysis and discussion on our results and conclude of this experimental work.

II. RELATED WORK

Robotized cultivation is a reasonable solution for some difficulties in forming, for example, the violently expanding worldwide old age public, work cost increment, increased need for of produce, etc [6].

Revised Manuscript Received on July 20, 2020.

* Correspondence Author

Dr P. Rama Koteswara Rao *, Professor, NRI Institute of Technology, Agiripalli, Krishna Dt, AP-521212, prkr74@gmail.com

Mr N Gopi Krishna, M Tech student, NRI Institute of Technology, Agiripalli, Krishna Dt., AP-521212, nekkalapugopi1997@gmail.com

Recognize and getting exact places of mango fruits are the most significant part of the visual framework for agriculture automated robot. Because of this, fruit recognition and identification has been broadly studied for a considerable amount of time [7]. Furthermore, these techniques have been partitioned into three parts by the innovations utilized.

A. IMAGE PROCESSING

A few image processing technologies are being used for agriculture fruit recognition task [8]. For instance, *Bin-hui ZHAN et al* [9] implemented an algorithm on double picture innovation for orchards of apple tree, and investigated the relevance between yield and flowering. To separate branches from pictures, *Anand Koirala et al* [6] changed RGB shading space to $I_1I_2I_3$ with XYZ space by a transition method. Many classification strategies, for example, decision binary trees, K-closest neighbor, and discriminant approach image calculation algorithms are utilized to pick suitable frequencies to characterize features in fruits [10]. To improve the classification accuracy, *Iqbal Zahid et al* [11] proposed an image combination technique by getting effected and natural images at the same time.

B. MACHINE LEARNING

There are many Artificial Intelligence based technologies for fruit detection, those detailed by [8]. To recognize and classify citrus fruits, *Liang-fen YIN et al* [10] removed highlights of neighborhood twofold example (LBP) and identified nearby power maxima around the immature citrus fruit. *Iqbal Zahid et al* [11] built up a framework to improve the quality control and arranging of dried products of *Ficus carica*. These methodologies utilize machine learning strategies, for example, PLS-DA and PCA to detect images and exhibited improved results. *P. Ferentinos et al* [12] introduced an image clustering method dependent on grouping. This procedure was applied to identify the damaged and bacterial spot in tomato fields. All these machine learning based strategies reasonably improved the classification and recognition accuracy of fruits.

C. DEEP LEARNING

In recent years, Artificial Neural Network strategies have gained an impressive impact in numerous fields, Signal Processing, Wireless Networks, Image Processing, Pattern Recognition. Numerous methodologies have been invented in the field of farming too. *Chen S. W et al* [13] introduced a methodology for apple fruit recognition from pictures taken in orchards. They utilized two procedures as Multi-scale Multi-Layered Perceptrons and Convolutional Neural Networks (CNN) to detect the fruits from its images. Their conclusive outcomes indicated the closer to the cutting edge techniques. Advanced RCNN is one of the most developed fruit recognition techniques, has given great outcomes in numerous recognition tasks. Specifically, a Faster-RCNN structure approach was utilized for recognition of apple, mango and almond from orchards. This technique additionally demonstrated that the information can improve the execution and decrease picture samples by more than two-folds [14]. The conclusive outcome indicated that this methodology achieved a momentous location execution for apples and mangoes. Essentially, *Zhang Xihai et al* [15] likewise utilized Faster-RCNN as a benchmark for agriculture product identifier.

III. F-MTCNN DETECTOR

A. MOTIVATION

Generally, the mango fruit includes natural the information such as shape and shading. Besides, it is bound to be overwhelming impediment in the assignments of fruit recognition. Hence, there is no uniform benchmark for fruit segmentation, create image database and explanation are tedious undertakings. Hence, real-time is one of the most significant scenarios for mango fruit recognition. It is one of the reasons that mango fruit identification model is commonly applied to automated systems, for example, yield estimation robot, arranging robot, picking robot, etc. Thus, the system of fruit recognition model, the previously mentioned motivations were considered. In view of all that, we planned a mango fruit detector that can distinguish fruits having various posture, low resolution and impediment. Likewise, it has the ability to check the count of mango fruits. Fig. 1 shows a regular detection of mango fruits showing up at various stances, sizes, separations, shapes and impediments.

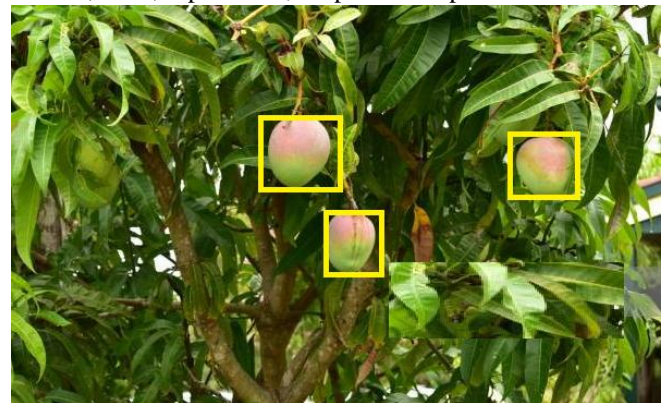


FIGURE 1: This is a sample picture for mango fruit identification, which comprises pose, occlusion illumination and variability.

B. THE SYSTEM ARCHITECTURE

This work improves the structure of F-MTCNN by excluding landmark loss of three multiple networks for mango fruit recognition task. The general architecture of this model includes three phases as appeared in Fig. 2.

The first stage of this model is the proposed network (P-Net). This stage incorporates just three CNN layers for feature extraction and useful for the next two stages. To diminish the highly overlapping candidates, this stage omits the non-max suppression technique before outfitting the result. At that point these candidate windows from the first stage become contributions for the subsequent stage, called refined Network (R-Net). In this stage, a great number of false windows from stage one are dismissed and the up-and-comer bounding boxes are aligned.

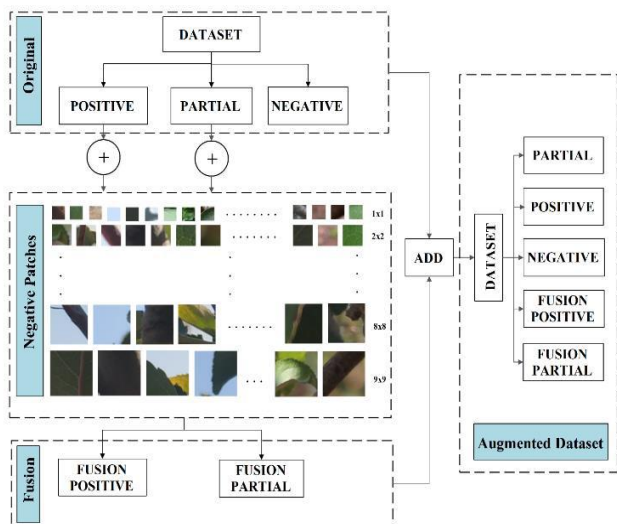


FIGURE 2: Architecture of Recognition Network. Proposed Network, Refined Network, Output Network.

The third stage is the yield Network (O-Net). The contributions to this stage originate from the yield of R-Net. The importance of this stage is that it further rejects false windows and gets exact bounding boxes. These three stages steadily recognize mango fruits against their training and acquire exact bound boxes of these fruits.

C. LOSS FUNCTION

Loss Function for mango fruit recognition comprises of two sections as classify and fruit regression. The training F-MTCNN to obtain the classification of natural fruits or mango fruits, and afterward use regressions box for the fruit location recognition.

1) FRUIT CLASSIFICATION

The fruit classification task is to recognize mangoes from the images, so it tends to be viewed as a two-class detection issue. Therefore, we employ cross-entropy for each object sample as \mathbf{x}_i .

$$L_i^{cls} = -\left(y_i^{cls} \log(p_i) + (1 - y_i^{cls})(1 - \log(p_i))\right) \quad (1)$$

Where, $y_i^{cls} \in \{0, 1\}$ represents ground truth parameter, p_i is the probability of input set \mathbf{x}_i , being a desired fruit.

2) BOUND REGRESSION

The bounding regression is to minimize the data of every candidate key window that is anticipated by the recognizer to the closest ground-truth. Each bounding box incorporates four directions for example height, left, width and top.

$$L_i^{reg} = \|\hat{y}_i^{reg} - y_i^{reg}\|_2^2 \quad (2)$$

Where, \hat{y}_i^{reg} specifies the bound box expected by the identifier and y_i^{reg} is the base truth fruit location.

D. TRAINING

For every section of network, there are mango fruits, mostly non fruit and fruit pictures for data training. The general loss functioning is as per the following in Eq (3).

$$C = \sum_{i=1}^N (\lambda_1 \alpha_i^{cls} L_i^{cls} + \lambda_2 \alpha_i^{reg} L_i^{reg}) \quad (3)$$

Where $\alpha_i^{cls} \in \{0, 1\}$ represents the input object type of \mathbf{x}_i .

We employed $\lambda_1=1, \lambda_2=0.5$ for P-Net, $\lambda_1=1, \lambda_2=0.7$ and for O-Net and R-Net accordingly, to get most accurate bound box.

IV. DATASET DESCRIPTION

A. IMAGE ACQUISITION

In this project, almost 750 pictures were gathered from nearby villages of our institute using high resolution camera. Moreover, somewhere in the range of 140 supplemental pictures from Internet and 120 from Image-Net (an open source database) were used for variety of tests. In particular, each of these pictures contained at least one mango fruit, and the highest number of fruits per picture was 08. All these items were named physically as individual picture datasets.

B. TRAINING DATA

After collection, we separated these pictures into three types as positive, negative and immature data samples. This partition was made based on the Intersection over Union (IoU) esteem with the background, as appeared in Table I.

TABLE I
IOU VALUE FOR THREE TYPE OF SAMPLES DIVISION

Sample	IoU
Positive	≥ 0.7
Partial	$\geq 0.45 \cap < 0.7$
Negative	< 0.2

V. EXPERIMENTS

A. EVALUATION METHOD

In this project, we used True Positive Rate (TPR) and False Positive Rate (FPR) as assessment method for mango fruit recognition. Furthermore, the TPR, FPR can be specified as:

$$TPR = \frac{TP}{TP+FN} \quad FPR = \frac{FP}{FP+FN} \quad (4)$$

Where, TP indicates the number of correlation results, FP is the number of incorrect detections and FN represents the missing fruit objects.

B. ENVIRONMENTS AND TRAINING

We experimented our examinations through a 64-bits PC, furnished with an Intel(R) i5-7500 CPU at 3.20GHz processor having 16 GB-RAM. The execution of this mango fruit recognition employed Tensor-Flow, which is a publicly released profound learning structure created by Google Brain Team.

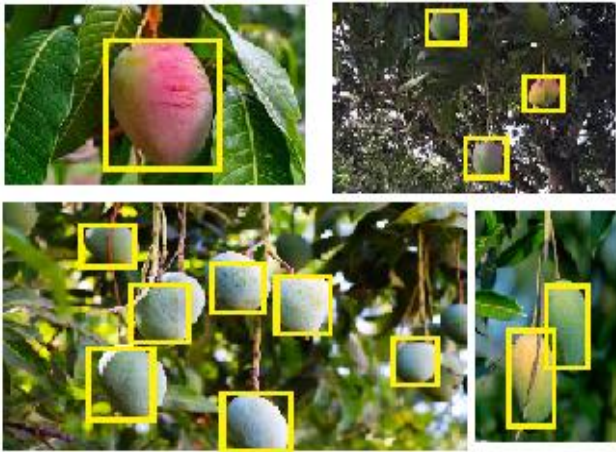


FIGURE 3. Some samples of mango detection results from image data base. The yellow marking indicates the results of the detector.

Other than this, we used Python as the programming language to adapt the structure of Tensor-Flow. The time cost for each layer is appeared in Table II. As evident from the table that the time cost function for this entire structure is about 2.5 hours. This low time consumption is helpful to apply in various other fields also.

TABLE II
EACH STAGE OF NETWORK: TIME COST

Model	Epochs	Time cost (minutes)
PNet	16571	98
RNet	6048	45
ONet	3726	25
Total	26345	168

C. COMPARISON WITH AUGMENTED DATASET

In order to know the capacity of Fusion Augment in detail, we considered a combination of the datasets of R-Net, P-Net, and O-Net. The quantity of every unit has unique dataset and increased by FA as introduced in Table III.

TABLE III
COMBINATIONS OF FUSION AND ORIGINAL DATASET

Name	size	Original	Fusion
Negative	12x12	96337	--
	24x24	96206	--
	48x48	95071	--
Partial	12x12	64682	90554
	24x24	16198	80990
	48x48	16365	65460
Positive	12x12	23003	92012
	24x24	5712	28560
	48x48	5532	33193

From the outcome we can see that the prepared model on augmented neural networks by FA can improve genuine positive pace of 0.05 contrasted with the model prepared on actual dataset. While, the combination of these systems got lower True Positive Rate when False positive examples were

under 100. This has been exhibited when the identifier is prepared on one of these system units (R-Net, P-Net, O-Net) separately, it will possibly influence some portion of the result when utilized on few positive examples. In any case, with the expansion in the quantity of samples, these distinctions decline bit by bit.

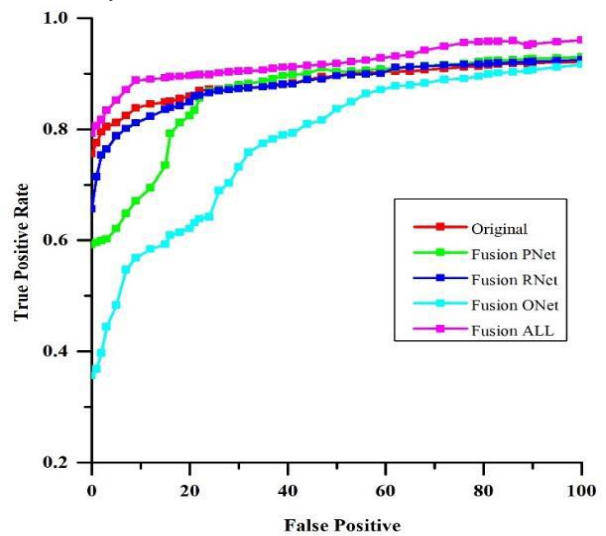


FIGURE 4. Recognition model training based on True Positive Rate for different data sets.

COMPARISON WITH THRESHOLDS

Threshold assumes a key aspect in getting the conclusive outcomes and how to set threshold is important. Along these lines, we set the thresholds as 0.9, 0.8, 0.7, 0.6 and 0.5, and checked the identification on a similar test dataset with the five targets. The quantity of True Positive Samples and False Positive Samples are appeared in Fig. 4.

From the figure, we can see that when the threshold diminishes from 0.9 to 0.5, the quantity of True Positive Samples increments, nonetheless, the quantity of False Positive Samples likewise increments. In this way, we proceeded with further experimentation on the FPR and TPR with similar conditions, as appeared in Fig. 5.

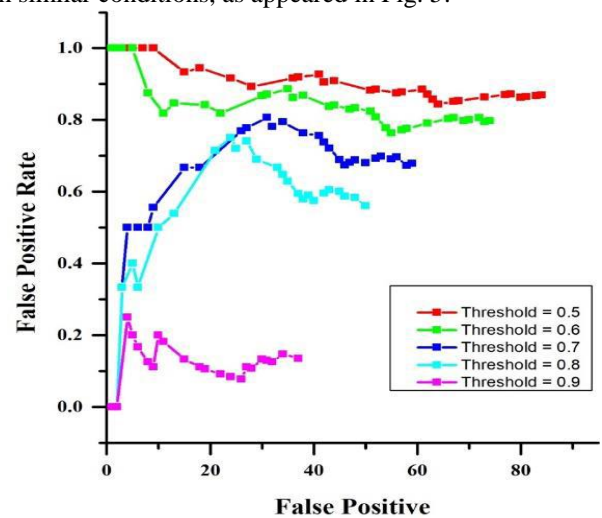


FIGURE 5. Recognition model training based on False Positive Rate for different threshold values.

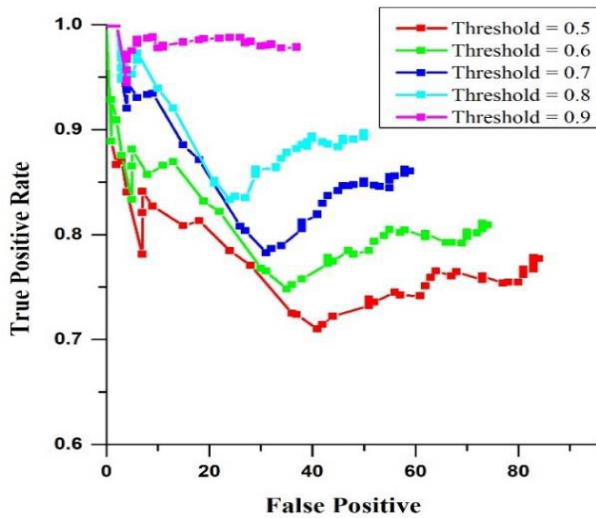


FIGURE 6. Recognition Model with True Positive Rate with various threshold values.

Fig. 5, specifies that the False Positive Samples was under 0.2 when the threshold at 0.9. At the point when the limit diminished to 0.8, the FPR went above 0.5 pointedly. In addition, if the threshold at 0.5, 0.6 or 0.7, the FPR expanded to nearly 0.7. Further, Fig. 6 shows that the TPR was as peak as 0.98 when the threshold at 0.9. The other four layers threshold about qualities below 0.9.

D. TIME COST

Time-Cost is the significant feature for an identifier. This is on the grounds that the automated agriculture system needs to gather and analyze the images and make a selection progressively. We experimented our mechanism on twenty unique images. Each bunch included one thirty pictures. The experimental results are shown in Fig. 7.

Here, the analysis is performed for the identification results introduced in the past segment. As appeared in Fig. 4 and Fig. 5, there are some False Negative and False Positive samples. The smallest dataset is most likely result behind that. In spite of the fact that we proposed a novel expanded technique to improve the results, the smallest number of samples confines the conclusive outcome. Since Deep Learning based identification is our approach,

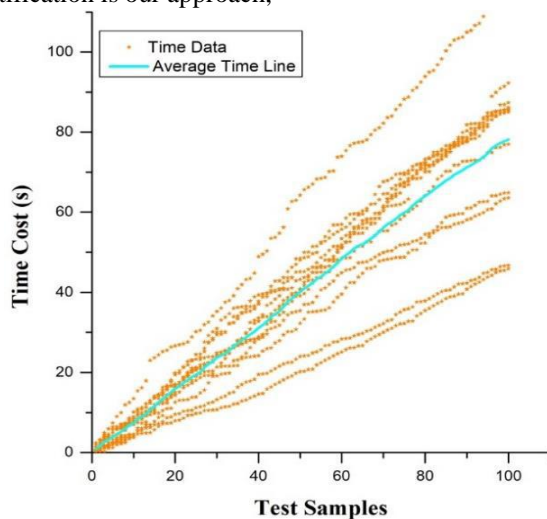


FIGURE 7. Recognizer with time cost for various groups of ten images.

We require adequate samples to affect the parameters of the network. The other reason is the diversity of the Neural Network. The gap between the training set and test is significant; the recognizer may miss few data objects.

VI. CONCLUSION

In this project, we experimented Multi-task Cascaded Convolutional Neural Networks based mango fruit detection. We picked mango for our investigation and gathered more than two hundred pictures from mango plantations and named them. In addition, we included a reasonable number of pictures from web and Image-Net dataset to make a dataset. Moreover, Fusion Augment approach is adopted in this experimentation. The novel investigation results exhibited this expanded strategy can improve the performance. The dataset for experimentation was acquired from Image-Net dataset, which contains many pictures. Our final results indicated that the identifier can efficiently adapt to different fruit recognitions. At outset, we experimented the recognizer on twenty groups of pictures with various resolutions. The mean time-cost of the detector was under 100 seconds for each one hundred pictures, which is extremely close to real-time.

VII. FUTURE SCOPE

We found the proposed Multi-task Cascaded Convolutional Neural Networks based mango fruit identifier has great execution of practicality and exactness to meet the development of image based agriculture robot from the experimental test results. Furthermore, there is significant distance for useful applications and advancement of the agriculture robot. The most significant task is to decide the sorting order of the detected mango fruits. Based on this investigation, we will concentrate on the examination of the human visual consideration when visual saliency recognition and semantic divisions were considered.

REFERENCES

1. Jinhui Li; Yunchao Tang, "Detection of Fruit-Bearing Branches and Localization of Litchi Clusters for Vision-Based Harvesting Robots," IEEE Access, Volume: 8, 2020, 10.1109/ACCESS.2020.3005386.
2. Enrico Bellochio et al., "Combining Domain Adaptation and Spatial Consistency for Unseen Fruits Counting: A Quasi-Unsupervised Approach," IEEE Robotics and Automation Letters, Volume: 5, Issue: 2, April 2020, pp. 1079 - 1086.
3. Chunhua Huet et al., "Automatic Detection of Single Ripe Tomato on Plant Combining Faster R-CNN and Intuitionistic Fuzzy Set IEEE Access, Volume: 7, 2019, 10.1109/ACCESS.2019.2949343.
4. Xiaoyang Liu ; Dean Zhao et al., "A Detection Method for Apple Fruits Based on Color and Shape Features," IEEE Access, Volume: 7, 2019, pp. 67923 - 67933.
5. Steven W. Chen et al., "Monocular Camera Based Fruit Counting and Mapping With Semantic Data Association," IEEE Robotics and Automation Letters, Volume: 4, Issue: 3, July 2019, pp. 2296 - 2303.
6. Anand Koiralaet al., "Deep learning – Method overview and review of use for fruit detection and yield estimation," Electronics and Computers in Agriculture, Volume 162, July 2019, Pages 219-234.
7. F.Abdelghafour et al., "A Bayesian framework for joint structure and colour based pixel-wise classification of grapevine proximal images," Electronics and Computers in Agriculture Volume 158, March 2019, Pages 345-357.

8. Ke-linget al., "Selection for high quality pepper seeds by machine vision and classifiers," Journal of Integrative Agriculture, Volume 17, Issue 9, September 2018, Pages 1999-2006.
9. Bin-hui ZHAN et al., "Detection and characterization of an isolate of Tomato mottle mosaic virus infecting tomato in China," Journal of Integrative Agriculture, Volume 17, Issue 5, May 2018, Pages 1207-1212.
10. Liang-fen YIN et al., "Identification of two Monilia species from apricot in China," Journal of Integrative Agriculture, Volume 16, Issue 11, November 2017, Pages 2496-2503.
11. Iqbal Zahid et al., "An automated detection and classification of citrus plant diseases using image processing techniques: A review," Electronics and Computers in Agriculture, vol. 153, 2018, pp. 12–32.
12. P. Ferentinos and Konstantinos, "Deep learning models for plant disease detection and diagnosis," Electronics and Computers in Agriculture, vol. 145, 2018 pp. 311–318.
13. Chen S. W et al., "Counting Oranges and Apples with Deep Learning: A Data Driven Approach," IEEE Robotics and Automation Letters, vol. 2, April 2017, pp. 781-788.
14. J. Hu, J. Lu, G. Zhao, C. Zhang and F. Mei, "An in-field automated wheat disease analysis system," Computers and Electronics Agriculture, vol. 142, 2017, pp. 369–379.
15. Zhang Xihai et al., "Recognition of Maize Diseases by Improved Deep Convolutional Neural Networks," IEEE ACCESS, doi: 10.1109/ACCESS.2018.28444052018.

AUTHORS PROFILE



Mr Nekkhalapu Gopi is currently pursuing his Master of Technology in the department of Electronics and Communication Engineering in NRI Institute of Technology, Pothavarappadu, Krishna Dt, AP, India. He completed his Graduation in the same Institute in 2018. He has presented many research papers in inter college competitions. He has implemented several working models in the field of Electronics and Communication.



Dr. P. Rama Koteswara Rao is currently working as Professor in ECE Department, NRI Institute of Technology, Pothavarappadu, Krishna Dt, AP, India. He received his Ph.D in Signal Processing from Andhra University College of Engineering, Visakhapatnam. He has twenty five years of experience in teaching undergraduate, postgraduate students and five years of Research experience. His research interests are in the areas of Speech Signal Processing, Pattern Recognition, Image Processing and Embedded Systems. He published more than 16 research papers in the International journals and presented few of them in International Conferences. He also published two patents at Intellectual Property of India, Chennai. He is a life member of ISTE and ISOI.