

Detection and Localization of Cotton based on Deep Neural Networks

Annapoorna B R¹,

Assistant Professor, Dept. of CSE,
Dayananda Sagar College of Engineering, Bangalore, India

e-mail: br.anu243@gmail.com

Dr. Ramesh Babu D R²

HOD and Vice Principal, Dept. of CSE, Dayananda Sagar College of Engineering, Bangalore, India

e-mail: bobrammysore@gmail.com

Abstract- Cotton detection is the localization and identification of the cotton in an image. It has a wide application in robot harvesting. Various modern algorithms use deep learning techniques for detection of fruits/flowers. As per the survey, the topics travelled include numerous algorithms used, and accuracy obtained on using those algorithms on their data set. The limitations and the advantages in each paper, are also discussed. This paper focuses on various fruit detection algorithms- the Faster RCNN, the RCNN, YOLO. Ultimately, a rigorous survey of many papers related to the detection of objects like fruits/flowers, analysis of the assets and faintness of each paper leads us to understanding the techniques and purpose of algorithms.

Keywords- Faster RCNN, RCNN, YOLO.

I. INTRODUCTION

Cotton detection in an image along with its localization, has been a major challenge in the field of modern agriculture. Cotton Detection has important applications in robot harvesting for plucking the cotton in the field.

This paper analyses the work related to fruits or flower detection and localization. In many papers, they used many deep learning techniques/algorithms for the fruit detection, like Faster R-CNN, has two networks. Region proposals are generated using Region proposal network (RPN) and another network is used to detect the objects using these proposal regions. The core variance of Fast R-CNN, is that, it will use selective search for creating the proposal regions. The time-consumption in producing region proposals is greater in selective search than Region Proposal Network. The two main steps of RCNN algorithm are selective search and extracting CNN features from each region independently for classification. YOLO algorithm will apply a neural network to a whole image. The picture is splits into many $S \times S$ grid with bounding boxes.

II. Related work

The Proposed work describes the detection and localization of fruits/flowers using various methods:

[1] This paper intended in the detection of apples, mangoes and almonds in the orchard. The detection method termed Faster R-CNN was used. This method contains a number of convolutional layers through which the image data is propagated through. The deeper VGG16 net which contains 13 convolutional layers, and the ZF network containing 5 convolutional layers were utilized in the technique proposed in this paper. A total of 300 bounding box were detected during the testing. Higher level task performance such as yield mapping and estimation was estimated using the tiled Faster-RCNN method. The 726 apples images were made available in the dataset. The almond and mango datasets contained 385 and 1154 images respectively. The properties of the ImageNet were enough in performing detection and fine-grained classification in orchards. The F1-score of a value greater than 0.9 was achieved as the result. The limitation of this technique was that it failed to provide a good

accuracy in localizing the fruits in the images captured in various light intensity.

[2] This work presents a fruit counting pipeline that uses deep segmentation, 3D localization and frame to frame tracking to precisely get the count of the fruits, oranges and apples in the collection of images. Fully Convolutional Network (FCN) was implemented using Tensor flow and the COLMAP package was used in the generation of the SFM reconstruction with the preprocessed data images. The visible fruits were counted and labelled to evaluate the performance of the counting system and the efficiency of the localization. The dataset for this method was collected by capturing the plant images during both the daytime and the night-time. The result testified a mean error of 3.3%, L1 error of 322, and standard deviation of 4.1% for the apple images and a mean error of -0.2%, L1 error of 203, and an error standard deviation of 7.8% for the orange image data.

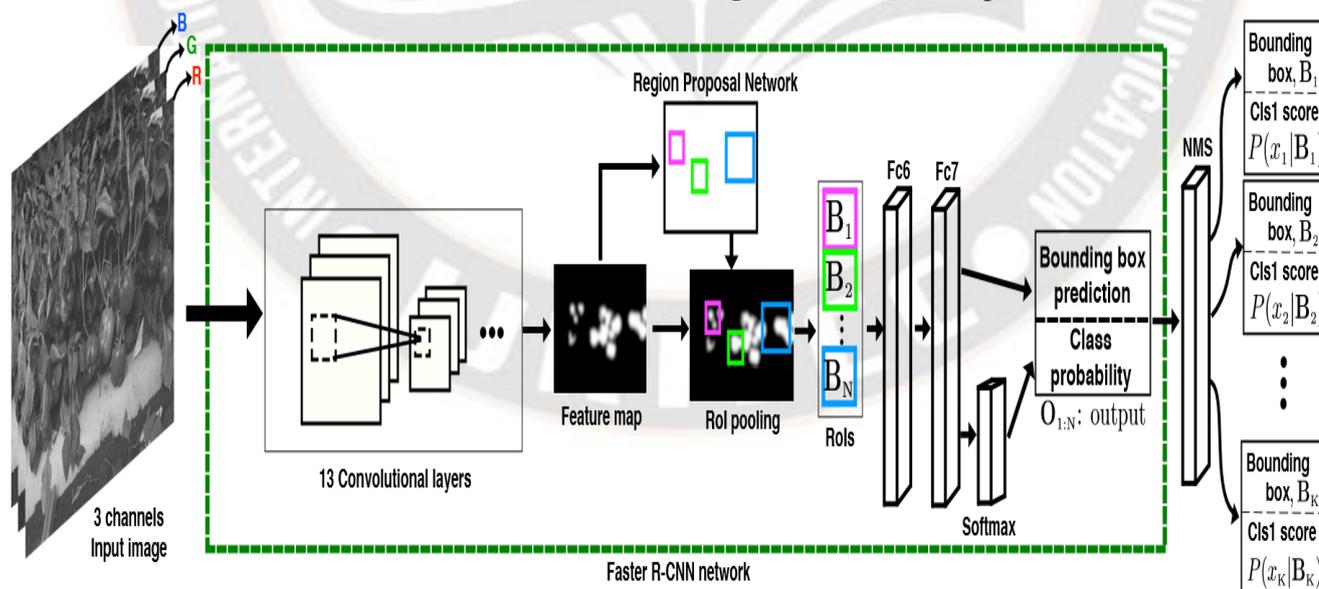
[3] Single shot detectors (YOLO) which is a state-of the art deep neural networks technique being used as a CNN for detecting apple and pears in the tree canopy is the main aim of this work. A state-of-the art convolutional network called YOLO is used for localization and detection. The detection, classification and region extraction are all achieved in a single pass. The input image was divided into 26 X 26 in order to overcome the inability in the detection of smaller fruits, the pooling layer and other convolution layers were also eliminated aiming to increase the speed. A dataset containing 100 images of apples, 50 of pears were used, and combining both the

many layers in the CNN network this method has a higher computation power which is the limitation of this method.

[4] Building a reliable, fast and accurate system that detects the fruits using the deep convolution neural networks was focused in this paper. This system makes use of both the Near-Infrared (NIR) and color (RGB) images. The system described in this paper uses the multi-model Faster R-CNN model to obtain better accuracy. This system was modelled to detect seven types of fruits/vegetables like pepper, strawberry, orange, apple, avocado, mango and rock melon. The system makes use of the DCNN approach. The Faster RCNN method has two parts in its methodology. One is a region proposal and another called the region classifier that makes use of the color images to perform the function of fruit detection as shown in Fig1. The objects of interest reside in the boundary regions proposed in the region proposal step. Whether this region is an object class of interest or not is later decided by the region classifier. PASCAL VOC dataset was used to fine-tune Faster R-CNN. The dataset contained 100 sweet pepper samples, 109 rock melon images. The result showed an accuracy of 80% in the fruit detection. The future work of this paper is based on counting the number of fruits.

[5] Yield estimation is the main concern of this paper. Counting the number of flowers and fruits helps the

Fig1. Faster R-CNN implementation.



dataset a total of more than 5000 images were used for the training of the models. The advantage of this approach is that due to removal of some layers the processing speed is increased while retaining the accuracy. Due to presence of

farmers to make improved decisions on the cultivation practices. This yield estimation also helps farmers in improving the crop quality. This is achieved by using the novel deep learning architecture based on CNN, it is

modified by including a modified version of InceptionResNet layer. The CNN comprises various convolutional and subsampling layers. The model is trained using the simulated data but is tested using the original data. It also detects the occluded fruits and is also effective in various light intensities. The dataset for the training of the model included 24000 images and 2400 images were used to test it. This model provides high performance. The accuracy of this architecture is between 70-100% and the average accuracy for around 100 images is 91.03%. This model is very robust to various environmental conditions. The limitation of this method is that it fails to count the green fruits as the model was not trained with those images.

[6] The objective of this work is to improve the perception and the ability of robot recognition in three-dimensional (3D) space. An automatic recognition method was proposed to achieve apple recognition from point cloud data. First, the pre-processed point clouds were used to extract an enhanced 3D descriptor (Color-FPFH) with the 3D geometry and the fusion of color features. The classification category was then subdivided into leaf, branch and apple in order to provide a more complete perception capability for the system. The three data classes were used to train the classifier that remains optimized using a genetic algorithm and is based on the support vector machine. Finally, a judgement with different 3D descriptors and additional classic classifiers was made to obtain the results of recognition and lateral comparison. This proposed method demonstrated improved performance in obtaining accurate results. The 3D point cloud data provides an accurate spatial characteristics and spatial distribution obtained using RANSAC algorithm of the fruits in the trees for the harvesting robot. This system made use of the SVM classifier. Point cloud data value of 52 was obtained from 8 apple trees, from various views in the natural environment. The GA-SVM model provided a classification accuracy rate of 92.3%.

[7] This paper has presented a data-driven fruit counting pipeline that simplifies across numerous unstructured environments based on deep learning. The data sets are chosen in such a way that they are challenging in unique ways: apple data set consists high colour similarity between foliage and fruit and orange data set consists uncontrolled illumination, more occlusions, depth variation.

They have used label.ag, a crowd sourced label collection platform, along with Support Vector Graphics (SVG) for the need of saving label information. The ability of the pipeline to count the number of apples in the images accurately was demonstrated and presented the linear regression model and count neural network. They got 12 error of 10.5 on the apples and 13.8 on the oranges. They

compared their model against a state-of-the-art texture-based method on their apple data set and verified greater accuracy, lower SD, and l2 error. Along with these, the model proposes accurate counting from a limited data set within lesser time. The disadvantage in the proposed work is that they are vulnerable to errors in the labels.

[8] Image Processing has been a major field of research in automated modern agriculture. This paper talks about an algorithm based on genetic algorithm and artificial neural network for splitting grape clusters from background and leaves. The three major optimized problems of this work are the number of neurons in each hidden layer, supreme colour features and a number of hidden layers. Genetic Algorithm is deployed for selecting supreme colour features and optimizing of artificial neural network structure simultaneously. The results proved that Genetic Algorithm specifies the eight colour features as the accuracy of the artificial neural network in the identification of grape clusters, leaves, and others pixels is 99.07%, 99.50% and 99.61%. The accuracy of the algorithm is 99.40 percent. The specificity is 99.13%, 99.76% and 99.80%. The specificity of the GDA1 algorithm for segmenting the grape clusters and foliage is 88.14 and 21.77% and accuracy was obtained 75.78% and 40.10%.

[9] The aim of this study was to develop a robust algorithm to sum up and localize the immature citrus fruit in images. Every image was taken in low light conditions, and for further analysis, the green elements of the images was used.

In the input images, the presence of the local maxima of green elements were detected after flattening the complete image by Weiner filters and Median. This was classified by a classifier RUSBoost with LBP texture. In the above classified negative and positive samples, the hierarchical outline maps centered at them were extracted and fitted with CHT (Circular Hough Transform). The fruit targets were defined using merged fitted circles. Advantages of detection of immature fruits is that it lends a promising benefit for farmers to estimate their profit and yield and plan application of nutrients. The evaluated test set of twenty-five images, resulted in 80.4% true positive rate and 82.3% precision rate and F measure was 81.3%.

[10] Through the parameter optimization, the Convolutional Neural Network is applied to the tasks of detection and recognition of fruits. Many of the researchers are interested in using convolutional neural network as it provides some advantages such as reduction of computation complexity and less memory requirement and better performance in recognition of object process. A convolutional neural network is a special class of neural

network used to process the visual data. It is a variation of the architecture of a multilayer neural network and includes three layers fully connected layer, sub-sampling layers, and convolution layers. In this paper, they have used Alexnet to perform classification and process of simulation is deployed using Matlab's GPU. For deploying the proposed system, they have built an image database which is very helpful. Then they have applied convolutional neural network to recognize the images and evaluated its performance. The outcome of this test had accuracy of 94% for the thirty classes of 971 images. This indicates that these methods could be used for vision subsystem based control applications.

[11] They have proposed a novel approach for detection of apple flower, which is based on deep learning techniques that represents the state of the art for computer vision applications. In comparison with existing methods, the features extracted by our conventional neural network effectively combines both morphological information and colour, leading to improved performances for considered cases.

This proposed approach is as follows, for the apple flower detection as the convolutional neural network and support vector machine algorithm. The optimal model extracts feature from the convolutional neural network, reduces feature dimensionality to 69, and classification is done using support vector machine.

where the output of multiple classifiers is combined to yield a final classification. This model consists of Fast R-CNN along with Classifier Fusion, a Region Proposal Network and 5 Convolutional layers as shown in Fig2. Objectiveness probability estimation is made by using the Fusion strategy. The dataset obtained from the almond images captured in the three different convolution layers which had 384 almond images and 1574 images with bigger fruits were used for training the model. The validation of the model used 100 images of almond and 360 images of bigger fruits. This model generated a F1-score of 0.8221. This model has obtained 4.71% improvement. Fruit detection using the Faster RCNN technique lacks in detection of the smaller fruits, this model overcomes this limitation.

[13] The image recognition methods usually mistaken tomatoes in adjacent positions to be a single tomato. But this paper combines deep learning with pooling, rectified linear unit (RELU) application and edge contour detection which would help in easy detection of ripe tomatoes. To locate and recognize the candidate ripe tomato regions in complex greenhouse environments Faster R-CNN was used. In order to eliminate the background images from the candidate regions the Gaussian density function was used. A dataset contained maximum number tomato images

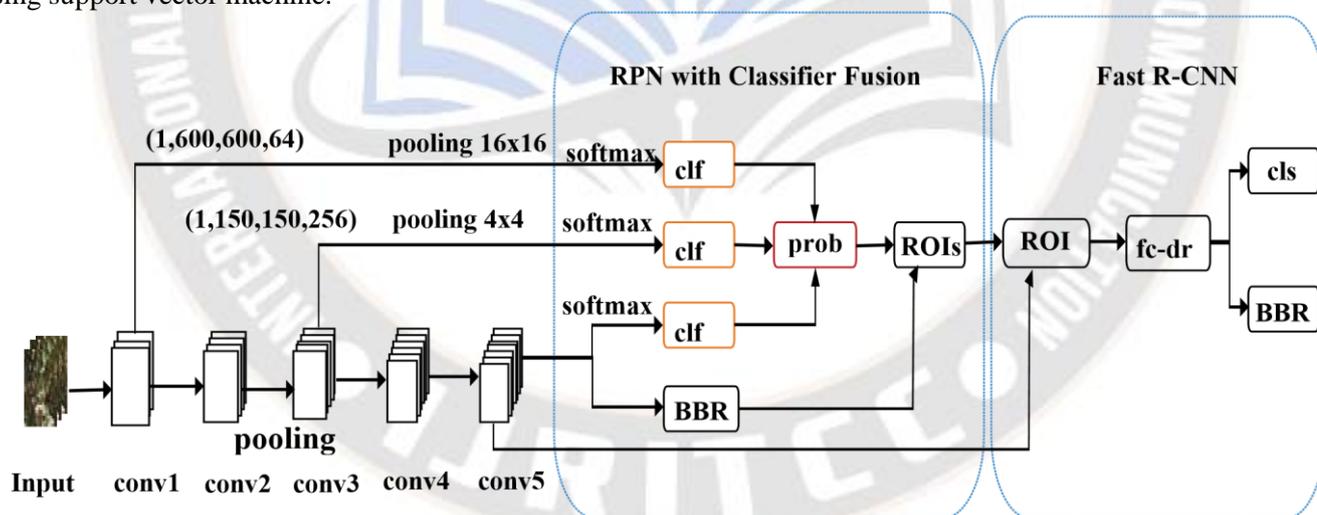


Fig2: Faster R-CNN with classifier fusion.

Tests were conducted on datasets of four different types showed that the proposed convolutional neural network-based model allows precise identification of flower with ideal recall and precision rates of 80% for datasets suggestively different from the training systems.

[12] This paper initiates small fruit detection using the Faster R-CNN network along with multiple classifier fusion strategy. Multiple Classifier fusion is a method

which were taken at different times and from different directions. Of these images, 200 images were used for testing and a total of 600 images were used for training the model. After the training of the model, tomatoes in various states were tested and analyzed. The detection accuracy of separated tomatoes obtained was 78.4%, adjacent tomatoes was 81.9%, overlapped tomatoes was 95.5% and the shaded tomatoes was 93.8%. As a result, tomatoes were detected with greater accuracy in comparison with older theories.

[14] This paper deals with the detection of the blueberry using the hyper spectral imagery. With respect to older studies the redundant information was obtained by analyzing some bands of the images. Therefore, a band selection approach called Kullback–Leibler divergence (KLD) was used to retrieve the most useful bands from the images. The dataset was obtained from the field which included 40 images of blueberry plants. High spatial resolution was an advantage of this band selection method. The KLD approach used 3 classifiers namely CART, RVM and KNN. CART would select the waveband, as the root of the classification tree which had best separating ability. KNN classifies the testing instances into the classes. RVM is a Bayesian sparse kernel method used to obtain the posterior probabilities of the instance. This model and the SVM-based classification both provided similar accuracy but this method had an upper-hand in its faster speed of processing. When the K value was 1 the KLD method provided better performance in its band selection. KNN is the best method for classification of different growth stages as well as the background objects of blueberry fruit.

[15] Deep learning - multi-view object detection approach is used in this paper. The main aim was to improve the performance of detection of small objects using the regression models based on deep learning techniques. YOLO could do this work but since it divides every image into a fixed number of grids, this results in limited number of objects detected in each image. Prior to object detection, region segmentation operation is applied to improve the performance of the method used in paper. The information captured can be increased to improve the performance of this model. This module is called multi-view module. Then an interpolation algorithm was applied for magnifying every region detected by the factor to four. This ensures that the detected objects do not become deformed after being detected. After this a region splicing procedure was implemented to the bounding boxes obtained previously to get the complete bounding box for each object. This test used 160 images consisting of smaller objects. The selected objects were 1/10th of the image in scale value. The average F-measure for calculating the average retrieval results for Multi-view YOLO was

0.729 while for old classical YOLO the F-measure was 0.319. This concluded that the Multiview methods have better retrieval abilities compared to classical methods. 14.3%, 7.4%, and 13.1% are the mAP scores obtained which were higher than the classical methods

[16] In most of the conditions it is very difficult to discriminate some fruits from their backgrounds due to same color and sometimes the high illumination could be a problem as well. In 3D space and image plane the range-

data and RGB are utilized to analyze various shape related properties. Highlight detection of data pruning is the method used in this paper. Symmetry is a process which occurs abundantly in nature. Fruit detection along with symmetry detection was used for better discrimination of fruit from the clusters. Rapid pruning is also included to process the 2D visual data and to find the locations which have fruits. The RGB-D data was taken from the images captured by a simple RGB camera. A total of 88 real images were used to make up the dataset. Dataset was randomly divided into two sets one for training and the other for testing. Results were recorded using sweet pepper dataset with a substantial number of occlusions.

III. CONCLUSION

Based on the thorough research survey, most of the papers have used various Deep Learning models and techniques for the detection and localization of fruits/flowers. Most of them have used Convolutional Neural Networks (CNN), R-CNN, Faster R-CNN and YOLO have proven to give substantial efficiencies. Various training techniques that make these networks easier to train and converge have also been surveyed. As the technology is advancing at tremendous rate, various methods and modifications of CNN are developing, to overcome the limitations. In paper [1], the algorithm failed to provide a good accuracy in localizing the fruits in the images captured in various light intensity. In [5], the algorithm failed to count the green fruits as the model was not trained with those images. And some of the techniques failed to detect occluded images. By taking these as the reference, we are going to overcome these limitations in our future work for the cotton detection using modern techniques.

IV. REFERENCES

- [1] Suchet Bargoti and James Underwood., Deep Fruit Detection in Orchards, 2017.
- [2] Motion Xu Liu, Steven W. Chen, Shreyas Aditya., Robust Fruit Counting: Combining Deep Learning, Tracking, and Structure, 2018.
- [3] Kushtrim Bresilla, Giulio Demetrio Perulli, Alexandra Boini., Single-Shot Convolution Neural Networks for Real-Time Fruit Detection Within the Tree, 2019.
- [4] Inkyu Sa, Zongyuan Ge, Feras Dayoub, Ben Uprocft, Tristan Perez and Chris McCool., Deep Fruits: A Fruit Detection System Using Deep Neural Networks, 2016.
- [5] Maryam Rahnemoonfar and Clay Sheppard., Deep Count: Fruit Counting Based on Deep Simulated Learning, 2017.

- [6] Tao Yongting, Zhou Jun., Automatic apple recognition based on the fusion of color and 3D feature for robotic fruit picking, 2017.
- [7] Steven W Chen¹, Shreyas S. Shivakumar¹, Sandeep Dcunha², Jnaneshwar Das., Counting Apples and Oranges with Deep Learning, 2016.
- [8] Nasser Behroozi-Khazaei, Mohammad Reza Maleki., A robust algorithm based on colour features for grape cluster segmentation, 2017.
- [9] Jun Lu, Won Suk Lee, Hao Gan, Xiuwen Hu., Immature citrus fruit detection based on local binary pattern feature and hierarchical contour analysis, 2018.
- [10] Zaw Min Khaing, Ye Naung, Phyo Hylam Htut., Development of Control System for Fruit Classification Based on Convolutional Neural Network, 2018.
- [11] Philipe A. Dias, Amy Tabb, Henry Medeiros., Apple flower detection using deep convolutional networks, 2017.
- [12] Xiaochun Mai, Hong Zhang and Max Q.-H. Meng., Faster R-CNN with Classifier Fusion for Small Fruit Detection, 2018.
- [13] Chunhua Hu, Xuan Liu., Automatic Detection of Single Ripe Tomato on Plant Combining Faster R-CNN and Intuitionistic Fuzzy Set, 2019.
- [14] Ce Yang, Won Suk Lee, Paul Gader., Hyperspectral band selection using kullback-leibler divergence for blueberry fruit detection, 2017.
- [15] Cong Tang, Yongshun Ling, Xing Yang, Wei Jin and Chao Zheng., Multi-View Object Detection Based on Deep Learning, 2018.
- [16] Ehud Barnea, Rotem Mairon, Ohad Ben-Shahar., Colour-agnostic shape-based 3D fruit detection for crop harvesting robots, 2017.