JCSE International Journal of Computer Sciences and Engineering Open Access
Survey Paper Vol.-6, Issue-6, June 2018 E-ISSN: 2347-2693

A Survey on Identification and Detection of Fruits based on Deep Neural Networks

Moksha Lakshmi B N[1*], Vindhya P Malagi[2]

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

*Corresponding Author: meghanabn07@gmail.com

Available online at: www.ijcseonline.org

Accepted: 17/Jun/2018, Published: 30/Jun/2018

Abstract— In Computer vision, object detection has become one of the most popular research fields. Human eyes can distinguish a various number of objects in images with less exertion, despite the fact that the objects in the images differ at various perspectives. This task is as yet a challenge for Computer vision frameworks. The goal is to present an efficient approach for fruit detection which can be used for yield estimation. In this paper, an efficient survey presented for Fruit detection in order to estimate the yield using Convolutional Neural Networks. This approach mainly benefits farmers and also applying automation in the field of agriculture, this helped to create several advancements to the industry. Various methods are surveyed in this paper in order to solve the existing problems of Fruit detection.

Keywords—Convolutional Neural networks, fruit detection, learning methods**.**

I. Introduction

In the agricultural sector, knowing the information on harvesting areas, yield and productivity plays an important role in planning and allocating resources. Reliable and well-timed information acts as a primary input to the farmers and someone who sets the plan pursued by a government or business. Predicting possible yields will enable farmers to define the crop plan, deliver and storage space required. Accurate crop yield estimation is more difficult and also a challenging task in cultivating frameworks that are distinguished by smallholder farms that deliver an extensive variety of various harvests. The information of an accurate number of fruits, flowers, and trees helps farmers to make better decisions on farming, plant leaf disease prevention, pesticides and labor required to harvest. The current practice with regards to yield estimation based on the manual counting of fruits by labors is a very tedious and labor cost is more and it is not suitable for big farms. Fruit Detection in computer vision is a process of extraction of foreground objects from a set of images. Nowadays an increasing number of automated applications went for identifying fruits from images. Despite the fact that different research endeavors have been made in this field, however, it incorporates a few difficulties, such as complex scenes with changing lighting conditions, low contrast amongst fruits and leaves, foreground occlusions and cluttered backgrounds. A large portion of these applications has been to discover the fruits for automatic harvesting [1]. Human eyes can identify objects in images with little effort, but in some cases, problems occurred due to different viewpoints of objects and also in different sizes, even when they are moving. Even though objects are recognized in some instances they are partially obstructed from view, also under varying weather conditions. This task still remains a challenge for computer vision systems in general. In Computer vision, a knowledge base is necessary to find out objects of interest and also efficient algorithms are required for recognizing objects. Finding out an efficient technique for detecting the object of interest in an image based on visual appearance, parts of an object, object features is a challenging and difficult task. Object detection sorting the learning system in order to generate the output value.

The rest of this paper is organized as follows Section II contains the related work of fruit detection using various learning methods, Section III contains datasets and discussions, and Section IV contains Conclusions and future directions.

II. Fruit Detection Using Various Learning Methods

In order to acquire necessary information on the subject of a variety of concepts associated with the current study, the existing literature is studied. Various significant conclusions were made in the course of those that are listed below. Different scientists have handled the issue of fruit detection. Because of high variation occurs in field settings, the problem exists in creating quick and reliable fruit detection framework including shape, texture and color features. Moreover, in the most cases the fruits are partly distracted and lead to persistently conditions of light changes and distorted by an uneven snow surface. An Enormous number of literatures are available in the field of Fruit detection and recognition. An elaborate survey has been conducted to get an in-depth insight of the work.

The authors of [2-7] have mostly adopted pixel-level segmentation approach for object detection. In the paper [2] author has presented an automated method that uses computer vision to identify and count grape berries. This approach distinguishes and includes a number of fruits in images gathered from a camera mounted confronting sideways on a vehicle driven along the lines in a vineyard. The approach is autonomous of shading contrast and can identify berries everything being equal, even those that are similar to leaves color of the background. However, the approach has neglected to include grape bunches early the season, even before berries have formed.

Author Wilshusen K [3] has presented a vision framework that consequently predicts yield in vineyards with high precision and determination. He has additionally done non-destructive yield estimation to find the boundary of neighboring clusters. At the point when the technique is connected to real-world data, accuracy got decreased because grape clusters grow along the same side, which leads to gross errors in estimating cluster size.

Kyosuke Yamamoto [4] performed tomato recognition by performing shading based division. Generally, Pixel-based division was led to segment the pixels of the images into classes made out of fruits, Blob-based segmentation was then directed to dispense with misclassifications, and X-means clustering was connected to identify single fruit cluster. They found that it was hard to identify young fruits in light of their little size and the comparable appearances with that of stems.

Wang [5] exhibited a framework that performs crop yield estimation in an apple orchard with moderately high accuracy. The proposed framework can't incur with plantations with bigger orchards of apples, which will require more exact and more precise to fragment the apple regions inside the images.

Authors Bac C.W and Hemming J [6] have proposed another robust-and-balanced accuracy performance measure PRob for CART pruning and feature selection. The outcomes are deficient to build an exact an accurate obstacle map. Except if, this was the main examination that reports quantitative execution for classification of few plant parts under fluctuating lighting conditions.

In the reference [7] author Calvin Hung presented the utilization of conditional probability for almond segmentation. Author has proposed an approach called five-class segmentation, features are generated using Sparse Auto Encoder (SAE) and these highlights at that point utilized inside a CRF structure and were appeared to overcome past method. Even though good segmentation performance achieved, proper detection has not done because of object occlusion.

Keren Kapach [8] proposed Computer vision for fruit harvesting robots. In that work a comprehensive survey of traditional and best in class machine vision arrangements with special emphasis on the visual cues and machine vision calculations utilized.

Authors Song Y and Glasbey C [9] have done precise fruit detection in controlled glasshouse conditions. The sliding window approach worked great execution when tested on datasets of selected images, can't deal with the fluctuation in scale and appearance of the target objects when deployed in real farm settings.

Deep neural networks have gained significant progress in object classification and detection.

In the paper [10] Author Alex Krizhevsky trained a large, deep convolutional neural network to group the 1.2 million high-determination images in the Image Net LSVRC-2010 challenge into the 1000 distinct classes. Non-saturating neurons and an extremely proficient Graphical Processing Unit execution are utilized for convolution task. As the system goes bigger and prepared it to train for numerous orders of magnitude to go with a specific end goal to coordinate the infero-transient pathway of the human visual framework.

In the paper [11] Koen E. A. van de Sande proposed Object detection using as Selective Search to produce numerous close to the actual, but not completely exact areas more than few object representation.

To separate interested regions from an image they used the deep neural network for classification, authors Lawrence Zitnick and Piotr Dollar [12] propose another technique for producing object bounding box proposals using edges. Simple bounding box objectness scores measure the number of edges that exist in the box and overlapping portion of an outline representing the box's boundary. Region Proposal Networks (RPNs) overcome the problem of edge detection by combining a feature with classification network, in the meantime, framework can anticipate bounded objects at each position where classification has done, the two systems shared same parameters, which brings about a substantially speedier execution, and furthermore appropriate for different applications.

Author Shaoqing Ren [13] presents a Region Proposal Network (RPN) in that obtained features are shared with the classification network. Existing Deep CNN does require a fixed-size input image else it will diminish precision. In the paper [14] author presented a new network structure, called Spatial pyramid pooling network it can produce a fixed-length representation of an image

In the paper [15] proposes Fast R-CNN, performs different sort of developments to enhance preparing and testing speed with high recognition precision. Fast R-CNN trains the simple Deep VGG16 network 9 times speedier than R-CNN, accomplishes a higher precision. In a real dataset, with a specific end goal to recognize fruits under extensive variations in light, incomplete occlusions and distinctive appearances utilizing single shading model are considered as a bottleneck. This impact to the utilization of multi-modular fruit detection framework can give vital data with respect to various parts of the fruits. Deep neural networks have demonstrated incredible work when utilized for multi-modular frameworks automation applications.

In paper [16] author exhibit cross methodology which highlight features for one modality (e.g., video) can be learned if various modalities (e.g., sound and video) are available at include learning time. The author applied a deep learning approach to solving this problem [17], which avoids low processing time and hand-design of features.

In paper [18] present an efficient approach to grayscale and rotation invariant texture based on local binary patterns. The method is robust to handle variations occurred in grayscale against any monotonic transformation. It is computationally simple and compares to rotation invariant patterns, the good performance achieved in uniform patterns.

In paper [19] released an open source implementation called DeCAF, for activation features with associated parameters a wide range of visual concept learning paradigms.

III. DATASETS AND DISCUSSIONS

Deep Neural Network needs a large amount of labeled data for training model, nowadays the most commonly used datasets of object detection are Image Net, PASCAL VOC. PASCAL VOC provides a standard image labeling and evaluation system. PASCAL VOC image dataset includes 20 categories, the dataset has a high-quality and labeled completely image what is very suitable for examining the algorithm performance. For object detection, Image Net provides an important source of data for object detection because of a specific bounding labeling training set. The results of the object detection can be expressed as follow: Input image ***I*** is tested by the detector, obtain the bounding box ***B*** of each object, corresponding category label ***c*** and confidence level ***f***. The evaluation of Multi-Objects detection in the same image is considered as separate detection results, the ground truth boxes ***Bg***. If the predicted bounding box satisfies the following formula,

$$a = \frac{area(B \cap Bg)}{area(B \cup Bg)} \geq a_0 \quad (1)$$

Here, a is an evaluation parameter Intersection over Union (IoU) representing the overlap rate of the ground truth box and object window predicted by the detector. a_0 is a previously set threshold value, the general value is 0.5.

Table 1 shows the comparison of different fruit detection techniques

Technique	Advantage	Disadvantage
Radial symmetry transform, k-nearest neighbor algorithm.	The approach is autonomous of shading contrast and can identify berries everything being equal, even those that are similar to leaves color of the background.	The approach has neglected to include grape bunches early the season, even before berries have formed.
Novel maximal point detection, Radial /Gabor symmetry transforms.	Non-destructive yield estimation to find the boundary of neighboring clusters has done.	Accuracy got decreased because grape clusters grow along the same side, which leads to gross errors in estimating cluster size.
Pixel-wise segmentation, Blob based segmentation.	Pixels of images segmented into classes to identify single fruit cluster.	Hard to identify young fruits in light of their little size. Blob-based segmentation was then directed to dispense with misclassifications
HSV Color Space. k-nearest neighbor algorithm.	Moderately high accuracy has obtained to estimate the yield in apple orchards.	The proposed framework can't incur with plantations with bigger orchards of apples, which will require more exact and more precise to fragment the apple regions inside the images.
Suboptimal Sequential Forward Floating Selection (SFFS), Classification And Regression Trees (CART)	Methods are robust-and-balanced in order to measure performance PRob for CART pruning and feature selection	The outcomes are deficient to build an exact an accurate obstacle map.
Conditional random field,	Good segmentation	Proper detection has not done because of

Kullback Leibler divergence algorithm.	performance achieved.	object occlusion.
Visual cues, spectral reflectance	Traditional comprehensive survey for machine vision has done.	Agricultural robotic systems are complicated integrated systems in which particular modules are often required to address very specific and applicative needs in real time
Novel statistical approach for clusters	The sliding window approach exhibits great execution when tested on datasets of selected images.	The sliding window approach can't deal with the fluctuation in scale and appearance of the target objects when deployed in real farm settings.
Convolutional Neural networks	Non-saturating neurons and an extremely proficient GPU execution are utilized for convolution task and also have gained significant process in object detection.	Non-saturating neurons and an extremely proficient GPU execution are utilized for convolution task.
Selective Search	Produce numerous detections close to the objects.	Failed to extract regions in more object representations.
Fast R-CNN	Performs different sort of developments to enhance training and testing speed with high recognition precision.	Fruits under extensive variations in light, incomplete occlusions and distinctive appearances are not recognized.
Region Proposal Networks (RPNs)	It overcomes the problem of edge detection by combining a feature with classification network.	Requires large dataset and GPU for training.
Multi- model approach	This Methodology highlight features for one modality it can be learned if various modalities are available at include learning time.	Low processing time and hand-design of features are used.

In general, we discuss the advantages and disadvantages of each method that is mentioned above. The results are listed in Table 1. We gave the techniques, advantages and disadvantage of each method. In the future, we suggest researchers to apply deep learning method on fruit detection in order to obtain better performance. Some typical deep learning techniques, such as convolutional neural network, Faster RCNN have been successfully applied in computer vision. The researchers can learn from their applications, and test the feasibility of deep learning in fruit classification.

IV. CONCLUSION AND FUTURE WORK

This study describes several machine learning methods from a series of articles for identify fruit information. The most effective approaches are examined in this paper. Furthermore, pattern recognition methods are applied in other fields has a bright future.

REFERENCES

[1] C. Bac, E. van Henten, J. Hemming and Y. Edan, "Harvesting Robots for High-Value Crops: State-of-the-Art," 2014.

[2] S. Nuske, S. Achar, T. Bates, S. Narasimhan and S. Singh, "Yield Estimation in Vineyards by Visual," September 2011..

[3] S. Nuske, K. Wilshusen, S. Achar and L. Yoder, "Narasimhan, S.; Singh, S. Automated visual yield estimation," 2014.

[4] K. Yamamoto, W. Guo, Y. Yoshioka and S. Ninomiya, "On plant detection of intact tomato fruits using image," vol. 2191–12206., 2014.

[5] Q. Wang, S. Nuske, M. Bergerman and S. Singh, "Automated Crop Yield Estimation for Apple Orchards.," 17–22 June 2012..

[6] C. Bac, J. Hemming and E. van Henten, "Robust pixel-based classification of obstacles for robotic harvesting," Vols. 96, 148–162., 2013.

[7] C. Hung, J. Nieto, Z. Taylor and J. S. Underwood, "Orchard fruit segmentation using multi-spectral," 3–7 November 2013.

[8] K. Kapach, E. Barnea, R. Mairon, Y. Edan and O. Ben-Shahar, "Computer vision for fruit harvesting robots-state," 2012.

[9] Y. Song, C. Glasbey, G. Horgan, G. Polder, J. Dieleman and G. van der Heijden, "Automatic fruit recognition," 2014.

[10] A. Krizhevsky, I. Sutskever and G. Hinton, "Imagenet classification with deep convolutional neural networks.," 8–13 December 2014.

[11] J. Uijlings, K. van de Sande, T. Gevers and A. Smeulders, "Selective search for object recognition," Int. J..

[12] C. Zitnick and P. Dollár, "Edge boxes: Locating object proposals from edges.," 2014.

[13] S. Ren, K. He, R. Girshick and J. Sun, "Faster R-CNN: Towards real-time object detection with region," 7–12 December 2015.

[14] K. He, X. Zhang, S. Ren and J. Sun, "Spatial pyramid pooling in deep convolutional networks for," 2015.

[15] R. Girshick, "Fast r-cnn," 13–16 December 2015.

[16] A. Eitel, J. Springenberg, L. Spinello, M. Riedmiller and W. Burgard, "Multimodal deep learning for robust," 28 September–2 October 2015.

[17] I. Lenz, H. Lee and Saxena, "Deep learning for detecting robotic grasps," 2015.

[18] T. Ojala, M. Pietikainen and Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification," 2002.

[19] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng and T. Darrell, "DeCAF: A Deep Convolutional Activation Feature for Generic Visual Recognition," 2013.

[20] J. E. P. e. a. Orlando, "Convolutional Neural Network Transfer for Automated," 20117.

