Agricultural Aid for Tomato Harvesting (AAT)

Dr. D. Nageshwar Rao, HOD
Electronics and Communicational Engineering
TKR College of Engineering and Technology
Hyderabad, INDIA
ece@tkrcet.com

Dr. Girish Kumar, Associate Professor
Electronics and Communicational Engineering
TKR College of Engineering and Technology
Hyderabad, INDIA
mekalagirishkumar@tkrcet.com

Dr. P. Gayatri, Associate Professor
Electronics and Communicational Engineering
TKR College of Engineering and Technology
Hyderabad, INDIA
line 5

Konam Pavan Kartheek, Student
Electronics and Communicational Engineering
TKR College of Engineering and Technology
Hyderabad, INDIA
pavankonam@gmail.com

J. Manichandra, Student
Electronics and Communicational Engineering
TKR College of Engineering and Technology
Hyderabad, INDIA
manichandra858@gmail.com

Garlapati Rohith Reddy, Student
Electronics and Communicational Engineering
TKR College of Engineering and Technology
Hyderabad, INDIA
rohithreddygarlapati@gmail.com

Abstract— **Fruit and vegetable harvesting robots have been widely studied and developed in recent years. However, commercial tomato harvesting robots remain a challenge despite extensive research due to the cost and difficulty. In this paper, we propose an efficient tomato harvesting robot that combines the principle of 3D perception, Manipulation, and an end effector. For this robot, tomatoes are detected based on deep learning, after which the analysis is sent to the user. The user has a live feed of the crop and an analysis of the fruit of whether it is ripe or unripe and the user can specify the robot to harvest the fruit or not. Using technology farming can be done as a game that can be controlled from anywhere. The designed robot consists of a four-wheel independent steering system, a 5-DOF harvesting system, a navigation system, and a binocular stereo vision system. The four-wheel independent steering system was capable of providing a low-speed steering control of the robot based on Ackerman steering geometry. The novelty of this project is that this method is simple, and image processing is kept simple to accommodate the processor's limited computation resources along with using Inception V3 for classification.**

Keywords— Computer Vision, Deep Learning, Harvesting Robot

I. Introduction

Today's world demands professionals that combine scientific and technological competencies with soft and social skills. . The demand for agricultural goods throughout the world is increasing at an unprecedented rate. [1] In the next ten years, an estimated 50% increase in agricultural production will be required to feed and clothe the world's population Increased automation and robotization are critical in the agriculture business to accommodate rising demand and compensate for manpower shortages. Moreover, agricultural tasks are ofter physically demanding and highly repetitive. Employing a robot will ease the task. The robot can give constant information about the farm or land where it is deployed and gives an accurate analysis of the aspects of farming [1]. All the types of robots are possible to be farming robots, such as a drone that can spray pesticides and pick fruits from tall trees effectively [2]. A robot can be effective during harvesting since it can monitor and pick the product at right time based on the input criteria. Arm robot manipulators and the mobile robot can be used as harvester robots. However, picking a robotic arm manipulator is perfect for the job since it can grab any objects.

A wide variety of robots have been developed for harvesting tomatoes in Japan, England, France, Italy, Israel, etc. [1]. Kyoto University developed a tomato harvesting robot with a 5-degree of freedom (DOF) manipulator. Okayama University developed a 7-DOF robot consisting of a moving system, vision, end effector, manipulator, and a control system that detects both red and green tomatoes [2]. Later new tomato harvesting robot consisting of a vision system and a rotating arm was developed. However, the time needed from recognition to pitching was about 15-20s per tomato, with a success rate of 50%-70%. Moreover, the Institute of Agricultural Mechanization Science of Korea also developed a series of tomato harvesting robots [3]. Tomato gathering robots will choose and pluck tomatoes of bound colors in the dark by employing an electric lamp. The robots will work night in pilotless greenhouses beneath these twinkling lights. These styles of robots are being developed by Panasonic [3]. Japan is working greatly in the field of harvesting robots and some Japanese scientists say, "*We started by imitating the work being done by people in farm chores and harvesting*".

The present tomato harvesting robots can pick up the tomatoes in the range of 4 to 5 seconds and cover an area of around 4 to 5 acres of land in about 40 to 45 mins [4]. The robot is capable of picking the fruit over different heights from 10cm to 100cm above the ground. But this robot costs around 3 to 4 lakhs and each robot is different from the other to handle. Any damage to the product inside the robot is difficult to repair or replace and a high sense of knowledge is required to handle the robot. The robot consumes about the power consumed by two air conditioners and not using the robot for a long time can create permanent damage to the robot. The availability these robots is very scarce and manufacturing these robots is far more difficult.

In this paper, we are going to create a robot that compromises the above-mentioned problems. We are going to design a robot with low cost and mutate into the user's needs by a simple procedure. We are going to implement deep learning algorithms to classify the ripe and unripe tomatoes. We are using the 3D printed robotic arm manipulator parts for cost-cutting and to use them reliably. For the movement of the robot, we are going to use DC motors with alloy wheels for easy movement in sand and farming land. For better accuracy, we are using a digital cam that is readily available in the market. We are making use of the easily available equipment to make the robot user-friendly and easy for the user to replace the parts in case of damage. We are designing the robot in a way that can be mainly used in household agriculture. We maintained the control only to the user and there is no automation, this is to increase interest in the field of agriculture for the future generations. We are designing in a way that agriculture is portrayed as a game for children and a healthy habit for the adults to increase the agriculture percentage with the growing population.

II. MATERIALS AND METHODS

This section describes the implemented robotic arm and the proposed methodology that acts in the place of human movements and replaces their harvesting actions on their behalf.

A. Harvesting Robot and Materials

The hardware of the proposed model consists of a robotic arm, a raspberry pi, and the actuators [4]. The selected robotic arm is designed using Fusion 360. The arm is designed with a cutter at the tip to cut the tomato and the arm is lightweight and features low power consumption. The arm is composed of five interlinked segments providing 5DoF with a maximum payload of 1.2 kg. The cutter is underactuated with two fingers like pins with a sharp tip designed for cutting the tomato root for harvesting.

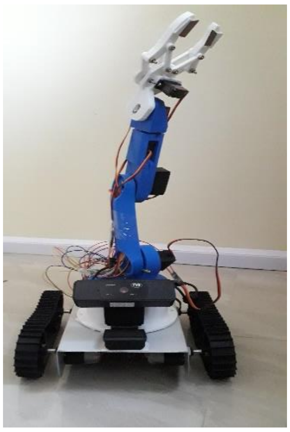

Figure 1 Overview of the designed robot

The vision system consists of a camera, and a TVS WC103, which provides a high-resolution color image. The TVS WC103 has a resolution of 1080 pixels and can capture a video at 30FPS which is superior image quality. The length of the cable made it easier to connect the device and the LEDs made it easy to understand the working of the camera. Any other camera which can capture the images of the required resolution and low cost can replace this device [5].

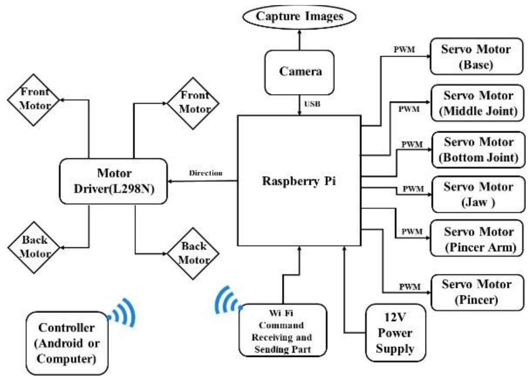

Figure 2 Overview of the proposed hardware architecture

The brain of the system is the Raspberry Pi where we utilized the latest model Raspberry Pi 4 where the features of the board made it easier to connect and control. The ram of the board made the work faster to communicate. All the motors and motor drivers along with the camera are attached to the brain of the system which is Raspberry Pi. The Raspberry Pi creates a web terminal through which the user can connect and control the robot. The functions or operations declared by the user are processed and later the raspberry pi powers the motors according to the user-defined operation. The User will have a live feed through the camera where the image is always processed from the raspberry pi from frame to frame with an FPS (Frames per second) of 10. The user and the raspberry pi should be on the same network for the connection to establish and make sure the internet speed is high to avoid network losses and slow movement or damage to the robot. The DC motors are connected to the motor driver which activates the motor clockwise or anti-clockwise according to the user-defined functions.

B. Method

Figure 3 describes the various steps of the designed and implemented strategy for harvesting robots. Before starting the user has to run the commands of the servo motors to avoid the pulse width modulation in the servo motors which play a vital role in the movement of the arm. Then all the systems are initialized. The main point of the process is the establishment of the connection between the user and the robot which is defined as the IP connection. The Stop function in the flow diagram expresses the stop for all the motors but not the live feed. The stop function is one way similar to the end as it is the final step of the process.

The moving robot deals with the DC motors whether to move in the forward direction or backward or left or right direction which depends on user-defined variables. All the motors are connected to the motor driver and the functions called are processed in Raspberry Pi and those functions are sent to the motor driver as a command to on or off depending on the function called.

For the movement of the robotic arm, there are six servo motors where each has its priority in the project and each motor has two different operations one is to increase the angle and the other to decrease. When the increasing function or decreasing function is called the servo angle changes with a difference of 2 degrees to ensure that even a small movement can be made and observed. Only for the cutter motor, the functions are defined whether to cut or not, and if the user initializes to cut the motors move the jaw closer where the interacted blades to the jaw will cut the tomato root which helps in the harvesting and move back to its original position which maintains at least 5cm gap between the jaws.

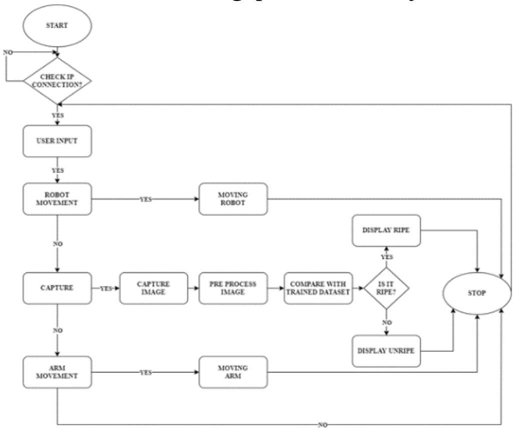

Figure 3 Overview of the main steps involved in the proposed harvesting cycle

III. IMAGE SEGMENTATION AND PLANNED ALGORITHM

Image Segmentation could be a pc vision method that partitions a digital image into multiple regions to facilitate its analysis. Image segmentation is usually wont to find objects and bounds in pictures. This method is trivial for humans; all the same, achieving strong image segmentation remains a challenge in pc vision applications as a result the noise, low distinction, poor illumination, and object boundary irregularities will result in inaccurate results [6], [7]. The techniques normally employed in image segmentation are thresholding-based, gradient-based, region-based, edge-based, and classification based mostly [8]. Inside the classification based mostly on techniques, machine learning and deep learning algorithms play a relevant role by establishing relationships among multiple options to boost. Every instance in each dataset utilized by the educational algorithms is portrayed by a similar set of options. If instances are given best-known labels that represent the corresponding correct outputs, the educational method is named supervised. In distinction, in unsupervised learning, the coaching instances are untagged [9].

In this study, the inputs and the desired outputs of the classification model are known; consequently, the selected learning method is supervised. The first step in supervised learning is to collect the dataset and determine which features are the most informative. In this study, the dataset consists of 553 aubergine samples acquired under different lighting conditions, and the feature used is the colors of the different scene elements. Color is a popular visual cue in machine vision tasks, and it is an appropriate choice for a discriminative feature because vegetables tend to have different reflectance properties than do the foliage and branches around them. However, instead of using the original R, G, and B values directly, we introduce color transformations before applying the segmentation algorithm to reduce its sensitivity to changing illumination conditions. These transformations quantify the intensity differences between the red and green channels(R-G) in the RGB color model and the hues in the HSV(hue saturation value) color model [10]. These images are then used as inputs for the segmentation process.

The proposed image segmentation algorithm is Inception V3 which is a part of the Convolutional Neural Network had given the highest accuracy. We collected the data of various tomatoes and classified them into two classes Plant and Unripe. They are pre-processed in which the images with noise are filtered and the later the features of the plant and unripe are extracted and the extracted features are processed through the algorithm and a model is created and deployed over the weights. These features are stored as weights which the trained model over our system. Later the trained model is used to detect the feature whether it is Plant or Unripe.

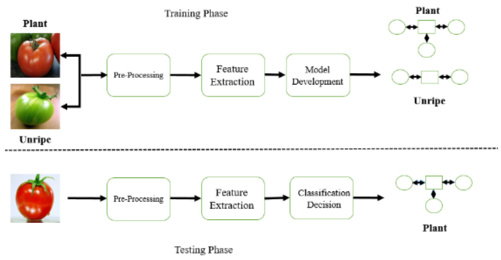

Figure 4 Overview of the proposed software architecture

To make the Inception network and module a reality, the experimenters designed intuitive convnet infrastructures. When multiple deep layers of convolutions were used in a model it redounded in the overfitting of the data [11]. To avoid this from passing the commencement V1 model uses the idea of using multiple pollutants of different sizes in the same position. Therefore, in the commencement models, rather than having deep layers, we've resemblant layers, making our model wider rather than deeper. The Inception model is made up of multiple inception modules. The introductory module of the Inception V1 model is made up of four resemblant layers [12].1×1 convolution, 3×3 convolutions, 5×5 convolutions, and 3×3 maximum pooling. The process of transubstantiating an image by applying a kernel over each pixel and its original neighbors across the entire image is called convolution. Pooling is the process used to reduce the confines of the point chart. There are different types of pooling but the most common bones are uttermost pooling and average pooling. This model is called as Naïve form.

One of the downsides of this naive form is that indeed the 5×5 convolutional subcaste is computationally enough precious i.e. time- consuming and requires high computational power. To overcome this the authors added a 1×1 convolutional subcaste before each convolutional subcaste, which results in reduced confines of the network and brisk calculations. The Inception V3 is just the advanced and optimized interpretation of the Inception V1 model. The Inception V3 model used several ways for optimizing the network for better model adaption. It has advanced effectiveness. It has a deeper network compared to the Inception V1 and V2 models, but its speed is not compromised. It's computationally less precious. It uses supplementary Classifiers as regularizes. The Inception v3 model was released in the time 2015, it has an aggregate of 42 layers and a lower error rate than its forerunners. The major variations done on the Inception V3 model are Factorization into Lower Complications, Spatial Factorization into Asymmetric Complications, Mileage of Auxiliary Classifiers, and Effective Grid Size Reduction [13]. Lower convolutions replacing bigger convolutions with lower convolutions surely leads to briskly training. Say a 5×5 sludge has 25 parameters;

two 3 × 3 pollutants replacing a 5 × 5 convolution has only 18 (3 * 3 3 * 3) parameters rather. A supplementary classifier is a small CNN fitted between layers during training, and the loss incurred is added to the main network loss. In GoogLeNet supplementary classifiers were used for a deeper network, whereas in Inception v3 a supplementary classifier acts as a regularizer. Grid size reductions are performed by max pooling. Based on these operations, the images are classified and trained.

IV. RESULTS

The designed robot model can detect and harvest the tomatoes efficiently and the movement is stable. The robotic arm movement is stabilized and movement is accurate. The robot can move even in the rough sands due to the roller wheels and the belt attached to it. The vision is absolute and the live data feed makes the user easy to control. The battery life is long-lasting as all the devices are less power consumed and the motors are sent to sleep when not initialized where they do not consume any amount of power.

After classifying the images are trained and a model is saved. Training a model over a large number of epochs increases the model accuracy rate and also helps to identify the faults in the dataset model and the training algorithm. We trained this model over 100 epochs and saved the model to the name called rps_model. For the validation part, we partitioned 220 images and tested the trained model. After training and testing the model using the validation the accuracy and losses are given. The accuracy reports of the existing system are near 82% using the SVM methodology but for our system, the highest accuracy was reported to be 95% and the loss is noted to be 15%. The accuracy of the model vs the validation accuracy along with their losses is plotted and shown in Figures 5 and Figure 6.

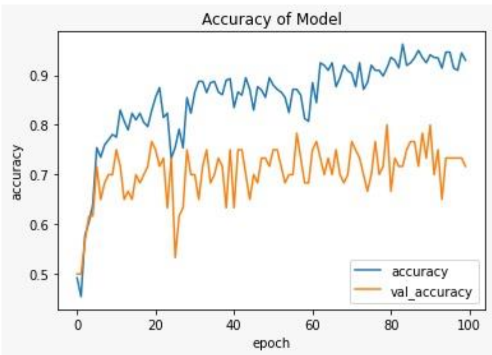

Figure 5 Overview of the accuracy of the implemented model

The results of the trained model are tested with two different methods one is by loading the saved model and the other is to test it by the model function in which the training code is mandatory. However, both the test cases give the exact and accurate output but the implementation of loading a saved model and testing is easier as it need not be run all the training code again.

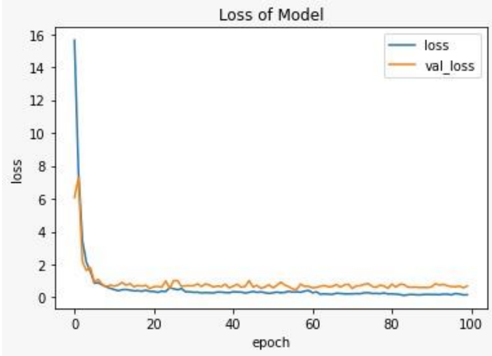

Figure 6 Overview of the loss of the implemented model

V. CONCLUSION

This paper presented a robot that aids in tomato harvesting. The robot is capable of picking the ripe and unripe fruits on command. It has the ability to capture the tomatoes and predict the ripeness level by comparing the array values of the picture to the trained dataset. The robot can move from place to place and gives a live feed of data, along with that it can also pick a tomato using the robotic arm and helps in harvesting. The model for detection of ripeness is trained on the Inception V3 algorithm where the model is trained over large epochs and has greater accuracy. The trained model can be used anywhere in machine learning concepts and can also be used to increase the model accuracy. The robot is capable of cutting the tomato root node without producing damage to the tomato.

ACKNOWLEDGMENT

The preferred spelling of the word “acknowledgment” in America is without an “e” after the “g”. Avoid the stilted expression “one of us (R. B. G.) thanks ...”. Instead, try “R. B. G. thanks...”. Put sponsor acknowledgments in the unnumbered footnote on the first page.

REFERENCES

[1] Selective Harvesting Robotics: Current Research, Trends, and Future Directions | SpringerLink.(n.d.). Retrieved 11 May 2022, from https://link.springer.com/article/10.1007/s43154-020-00034-1.

[2] ‘Which robot is best for harvesting and handling agricultural products like tomato, strawberry cartesian or/articulated or SCARA robot?’, ResearchGate. https://www.researchgate.net/post/Which_robot_is_best_for_harvesting_and_handling_agricultural_products_like_tomato_strawberry_cartesian_or_articulated_or_SCARA_robot.

[3] ‘Introducing AI-equipped Tomato Harvesting Robots to Farms May Help to Create Jobs | Solutions Wherever You Go’, Panasonic Newsroom Global. http://news.panasonic.com/global/stories/2018/57801.html.

[4] Robotics company | Robotic assistive technology | Kinova, (2019). [Online]. Available: https://www.kinovarobotics.com

[5] A. Campeau-Lecours, H. Lamontagne, S. Latour, P. Fauteux,et al. “KinovaModular Robot Arms for Service Robotics Applications,” Int. Journal ofRobotics Applications and Technologies (IJRAT), vol. 5, no. 2, p. 23, 2017.

[6] D. Kaeli, P. Mistry, D. Schaa, D. Zhang, “Case study: Image clustering,” in Heterogeneous Computing with OpenCL 2.0, pp. 213-228, Jan. 2015.

[7] Y. Guo, A. S. Ashour, “Neutrosophic sets in dermoscopic medical image segmentation,” in Neutrosophic Set in Medical Image Analysis. Academic Press, 2019, pp. 229-243.

[8] J. Siswantoro, A. S. Prabuwono, A. Abdullah and B. Idrus, “ Automatic image segmentation using Sobel operator and k-means clustering: A case study in the volume measurement system for food products,” Presented at2015 International Conference on Science in Information Technology(ICSITech), Yogyakarta, 2015, pp. 13-18.

[9] S. Kotsiantis, “Supervised machine learning: A review of classificationtechniques,” Informatica, vol. 31, no.3, pp. 249-268, 2007.

[10] R. Fernández, H. Montes, J. Surdilovic, D. Surdilovic, P. Gonzalez-de-Santos, M. Armada, “Automatic Detection of Field-Grown Cucumbers forRobotic Harvesting,” IEEE Access, vol. 6, pp. 35512-35527, 2018.

[11] Q. Geng, Z. Zhou, X Cao, “Survey of recent progress in semantic image segmentation with CNNs,” Sci. China Inf. Sci. vol. 61, 2018.

[12] F. Lateef, Y. Ruichek, “Survey on semantic segmentation using deep learning techniques,” Neurocomputing, vol. 338, pp 321-348, Apr. 2019.

[13] ‘Deep Learning: Understanding The Inception Module | by Richmond Alake | Towards 12 Data Science’. https://towardsdatascience.com/deep-learning-understand-the-inception-module-56146866e652.

[14] M. Kalakrishnan, S. Chitta, E. Theodorou, P. Pastor, and S. Schaal, “STOMP: Stochastic trajectory optimization for motion planning,” in 2011IEEE International Conference on Robotics and Automation, Shanghai,2011, pp. 4569-4574.

[15] Xiong, Y, Ge, Y, Grimstad, L, From, PJ. “An autonomous strawberry-harvesting robot: Design, development, integration, and field evaluation,” Journal of Field Robotics, vol. 37, no. 2, pp. 202– 224, Aug. 2019 R. Nicole, “Title of paper with only first word capitalized,” J. Name Stand. Abbrev., in press.

[16] S. Hayashi, K. Ganno, Y. Ishii, I. Tanaka, “Robotic Harvesting System forEggplants,” Japan Agricultural Research Quarterly: JARQ, vol. 36, no.3,pp. 163-168, Feb. 2002.