Computer Vision-Assisted Real-Time Bird Eye Chili Classification Using YOLO V5 Framework

K. M. Abubeker, Abhijit, S. Akhil, V. K. Akshat Kumar, and Ben K. Jose

Department of Electronics and Communication Engineering, Amal Jyothi College of Engineering, Kanjirapally, Kerala, India

Abstract: Computer vision (CV)-based classification systems have become increasingly popular in the agricultural industry in recent years. This paper proposes a CV-assisted bird eye chili or "kanthari mulaku" classification framework using the You Only Look Once V5 (YOLO V5) object detection model. Automated sorting systems based on CV can accurately identify and classify chilies based on attributes such as size, shape, color, and texture. The dataset for the research consists of images of bird eye chilies in different positions and backgrounds. The model was trained using this dataset, and it could correctly identify and categorize bird eye chili. The chilies were then picked up by a robot manipulator and sorted by ripeness. Bird eye chili images captured in real agricultural situations have used to assess the effectiveness of the suggested framework. Images of red and green chili were taken from above using a high-resolution Raspberry pi 4B camera attached to a custom-built 3-degrees-of-freedom robot arm. We used public and real-time images to train the YOLO algorithm on photographs of bird-eye chili captured in real-time. As the robot arm goes around the chili plants, this model is connected with the robot's software control system to allow real-time detection and localization of the chili's. By automating bird eye chili crop monitoring and management, this system has the potential to significantly contribute to the growth and viability of the agricultural sector. We got a mAP of 0.94 and an average accuracy of 0.90 with the suggested method. Using a robotic manipulator for chili grading improves productivity and reduces human error compared to traditional methods. To test the robustness of the YOLO V5 framework, it has implemented on the Raspberry pi 4B graphical processing unit computer.

Keywords: artificial intelligence; bird eye chili; computer vision; Raspberry pi 4B; YOLO V5

I. INTRODUCTION

Bird eye chili is a widely used and flavorful ingredient in various cuisines around the globe. Plucking and sorting these tiny chili peppers can be challenging and time-consuming, often requiring a lot of manual effort. Recent advancements in computer vision (CV) technology have enabled automating this process using You Only Look Once (YOLO) algorithms. With the help of this technology, it is possible to pluck and sort bird eye chili more efficiently and accurately, saving valuable time and labor costs for farmers and producers. This paper will discuss how YOLO and advances in CV are causing a seismic shift in the market for bird eye chili.

Capsicum annuum, more often known as bird eye chili, goes by numerous names; kanthari mulaku is one of the more common ones. In South Asian cuisine and traditional plant medicine, it is used as a seasoning and appetite stimulant. It helps with weight loss, clears the airways, stimulates digestion, and lowers blood pressure and cholesterol levels, among many other medicinal benefits [1]. The risk of heart failure, artery blockage, and other cardiovascular disorders is reduced with a faster metabolism and lower cholesterol level [2]. Particularly potent in killing cancer cells without harming healthy ones is the bird eye chili. The strong vitamin B2 concentration of bird eye stew makes it beneficial for eyes. These characteristics elevate bird eye chili above other common green chilies, making it the preferred kind. Picking, arranging, and categorizing bird eye chilies have become more

valuable due to the global premium placed on high-quality products [3]. A unified aesthetic increases the likelihood that consumers will purchase a product. It takes a lot of time and effort to manually choose and arrange chili peppers. One method involves using knowledgeable examiners to choose and choose chilies, although this strategy is prone to prejudice. Bird eye chilies are in high demand on the global market because of their medicinal properties.

One kind of automation in agriculture is smart farming, which uses several automated devices to increase productivity and quality while reducing the amount of human interaction required [4]. Due to their small size, bird eye chilies might be a real pain to harvest mechanically. The grading and sorting of chili peppers, both red and green bird eye, are still done today, but it is now a completely manual operation based on visual examination. Factors such as visual impairments, fatigued witnesses, and varying opinions on quality reduce the method's credibility [5]. There could be a drop in quality if too much time is spent sorting through the chilies. Before the chili can be properly arranged, this must be foreseen by utilizing a procedure and tools for early decision-makers. Alternative technology like image analysis can be used to judge characteristics such as stature, color, material, physical strength, odor, and touch. Mechanization, fast filtering, and decreased human errors are the most often cited advantages of image processing in the workplace and other settings. The proposed sorting technique uses CV and is expected to offer a long-term and efficient solution for classifying red chili.

In conclusion, this chapter has provided valuable insights into the critical roles and current research trends of CV in agricultural

Corresponding author: K. M. Abubeker (e-mail: kmabubeker82@gmail.com).

fields. After the introduction, it emphasizes the research's importance and presents a clear outline of the structure of the paper and the topics that will be addressed in each section. The second chapter comprehensively analyzes the research topic, including the research questions and objectives. The third chapter provides a detailed explanation of the research methodology, as well as the significance of the research. The second last chapter includes information on model development, robotic arm design, and experimental results. The last chapter of the research report is a conclusion that summarizes the significant findings from the investigation.

II. RELATED WORK

Classifiers such as the support vector machine (SVM) and the k-nearest neighbor (k-NN) are just two examples of how machine learning is used outside of agricultural tracking to tackle categorization challenges in processing images and object identification [6]. k-NN and the SVM are just two examples of how machine learning is applied outside of agricultural tracking. In addition, machine learning models have shown remarkable effectiveness in several settings when it comes to accurately classifying objects, which is a testament to their versatility. Even more so, they hold great analytical potential for the agricultural industry. In today's industries, where arranging items can be laborious and time-consuming, using automated technology to boost productivity and cut down on labor expenses is an urgent need. In Abu Salman Shaikat *et al.*'s [7] study, one of the topics discussed was the automatic categorization of objects based on color and height. CV sorts things according to their hue, while the Haar cascade approach sorts things according to height. To do this, a CV system is employed with an artificial limb with six degrees of freedom (DoF).

Using the open-source HI-ROS framework, Mattia Guidolin [8] suggested a real-time human motion recognition method based on the camera acquisition of 3D pictures. Chanchal Gupta *et al.* [9] published new research investigating potential phenotyping and plant growth monitoring technology. In this instance, a real-time image processing technique calculates the height and width of the jalapeño plants in the garden. Open CV modules and the PyCharm Editor are evaluated using the RMSE. Md. Abdullah *et al.* [10] proposed a system that automatically categories objects based on their color, shape, and size data. In this experiment, a PixyCMU video sensor is used to determine the colors of various objects. The size and shape of an item are set through the outline method. An autonomous arm powered by a stepper motor and other essential components has developed to streamline the pick-and-place sorting procedure. Automatic item classification based on color and form is shown in Lennon Fernandes *et al.* [11] using a robotic arm. Outlines are determined using the modified border fill approach, and shapes inside those borders are identified using the Douglas–Peucker algorithm.

Learning that relies on a framework Hasmath channel state information (CSI), a metric for Wi-Fi sensors, was introduced by Farhana Thariq Ahmed [12] to improve identification. Juan Daniel *et al.* [13] developed a technique using convolutional neural network (CNN) designs to enhance the efficacy of quality assurance inspections, particularly for spotting bruising. The system's near-infrared and color CMOS sensors use different types of light. Human error while physically sorting tomatoes by color led to the development of a computerized tomato sorting system. The proposed method uses two TCS3200 RGB color monitors to check whether a tomato is ripe for human consumption. Image processing features like color, shape, and history of gradient may be used to classify produce [14]. Using a CNN, Arnab Rakshit [15] suggested a modified approach that aids the robotic arm in identifying the item using an RGB-D camera, picking the objects, and placing them in the right spot automatically, saving both time and effort. Hiram Ponce [16] developed a similar method of detecting low nutrients in tomato crop by analyzing images of leaf. By his novel CNN-based classifier, CNN+AHN incorporated a set of convolutional layers as feature extraction part, thereby increasing the efficiency from normal CNN methods.

Iman Iraei [17] proposed a novel method to identify the blur elements in a blurred image. For blur length analysis, frequency domain and wavelet-based factor are used, which is used for feature extraction and to identify the range of blur in an image. Detection of images with accuracy is needed for any automated system. Yingming Wang [18], in his study, aimed at maintaining the accuracy of high-resolution aerial images that detect the objects with the help of rotated bounding boxes that help in the process of training and provide better accuracy and detection output.

Since farming accounts for the primary source of income in India, improvements in agricultural technology are particularly pressing there. This improves farm productivity while reducing expenses [19]. Meer Hannan Dairath [20] developed an alternate approach of picking and grading fruits using CV technology where fruits are detected by color scheme conversion, masking of normal skin, masking of defects, and morphological dilation operations. Anand Kumar Pothula [21], in his case study, aims to evaluate the singulating and rotation performance of the sorting system for automatic inspection and grading of apples in terms of conveying speed, fruit size, and the percentage of apples exposure area for imaging. Dataset creation is essential for specific CNN frameworks like YOLO. Yogesh Suryawanshi [22] developed a dataset on four vegetables based on unripe, ripe, old, dried, and damaged. This dataset was developed by capturing real-time images of the vegetables with various possibilities. Md Hafizur Rahman [23] did a similar method for dataset creation and designed a system, which can automatically identify the position of a facemask that can give an alert to the individual before entering a crowded area. This model is trained in the MobileNet model by capturing various real-time images of an individual wearing a facemask.

To overcome sorting algorithm errors, we developed and deployed a YOLO V5 framework on Raspberry pi 4B computer to organize bird eye chilies in real-time. The developed CV-assisted robotic system uses an autonomous tool to inspect and grade chilies as they move along a moving line for differences in hue.

III. METHODOLOGY AND DESIGN

Bird eye chilies have recently been the subject of research in CV and machine learning because of their vast popularity and unique features. The YOLO V5 object recognition model's study of size, shape, and color has been used to classify bird eye chilies. The YOLO paradigm is a state-of-the-art approach to object recognition widely used in academia and industry [24]. Unlike previous object identification algorithms, YOLO was designed for speed and accuracy, making it ideal for real-time applications like autonomous vehicles and surveillance systems [25]. Bird eye chilies may be counted by using CV and machine learning to identify and count individual peppers. This paper explores the possibility of counting

bird eye chilies with the help of the YOLO object identification algorithm.

To begin counting bird eye chilies, the YOLO object identification system requires a set of labeled images. To achieve this, we capture photos of bird eye chilies at different stages of growth and annotate them with boundary frames that identify the chilies in the image. We use this information to teach the YOLO algorithm to recognize bird eye chilies in photographs. A labeled picture collection is required for YOLO to be helpful for bird eye view categorization. The ideal set would have several distinct varieties of chilies, each accurately labeled to indicate its size, color, and shape. Commentary may be made using either human labeling or the currently available annotation systems that can automatically detect and label objects in photos [26]. Once a labeled training sample is available, a deep learning system like TensorFlow may be used to teach the YOLO model. A deep neural network is given the data and trained to recognize the features useful for classifying chilies. This process is continued until a desired level of accuracy is achieved, both on the training data and the confirmation data. The trained algorithm may then be used for the task of classifying new images of chilies. The YOLO model takes an image as input and produces a set of bounding boxes representing the positions of objects in the image as output. The YOLO model can recognize bird eye chilies even when there are other objects in the image's background. This tool is useful for farmers who wish to keep a check on their chilies plants and predict how much they will yield. Growers may save time and effort using the YOLO method instead of labor-intensive hand counting while harvesting bird eye chilies.

A. ARCHITECTURE OF THE PROPOSED SYSTEM

Figure 1 provides a visual representation of the suggested process. The YOLO-powered design of the system allows it to recognize colors and distinguish between green and red chilies using CV technology [27]. The gadget utilizes a motorized tool and a mechanical actuator controlled by a Raspberry pi 4B to do categorization. We propose a self-sufficient agricultural instrument for color-sorting chilies by machine [28]. The suggested system employs a YOLO architecture for fast computer-vision-based red and green bird eye detection. The Raspberry pi 4B, one of a family of small single-board CPUs, is often used for regulating the position of an autonomous tool. It is powered by a Broadcom BCM 2711 quad core Cortex-A72 64-bit SoC running at 1.5 GHz. The algorithm uses a reference set of five images to pinpoint the exact location of each chili in the real-time images.

The labeling and training processes needed to produce datasets are shown in Fig. 2. To train a classification model, we need to collect many labeled images of chili peppers and assign meaningful labels by hand. Each image is converted into a text file with the corresponding category indicated by a red or green chili. Then, these records are taught using the Google Colab platform. A weight_chili. weights file is generated when training is complete.

Seven hundred and thirty-nine images of red and green chilies were utilized to create this database for the algorithm. We used the labeling application to rename each file and give it a new category label. Labeling the photographs will automatically create both the image file and the Word document, which can then be saved in the
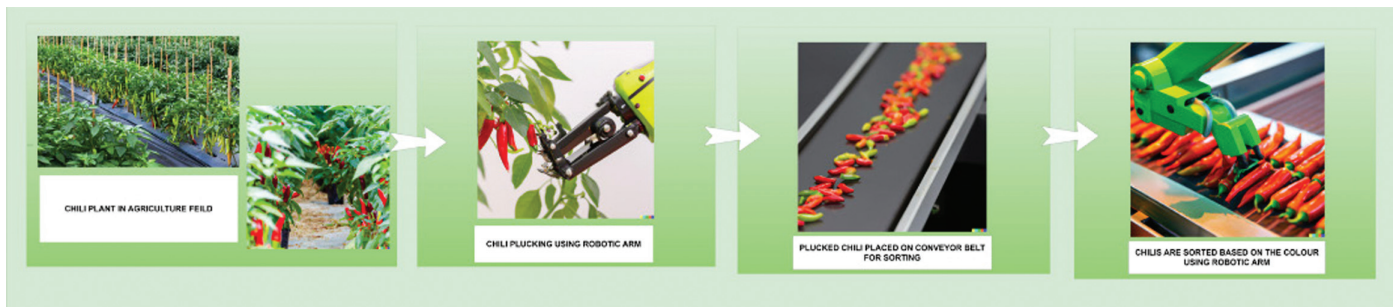
Fig. 1. A system of classification and sorting of bird eye chili, powered by YOLO and aided by computer vision, has been devised. This wondrous contraption doth employs the Raspberry pi 4B computer to great effect.
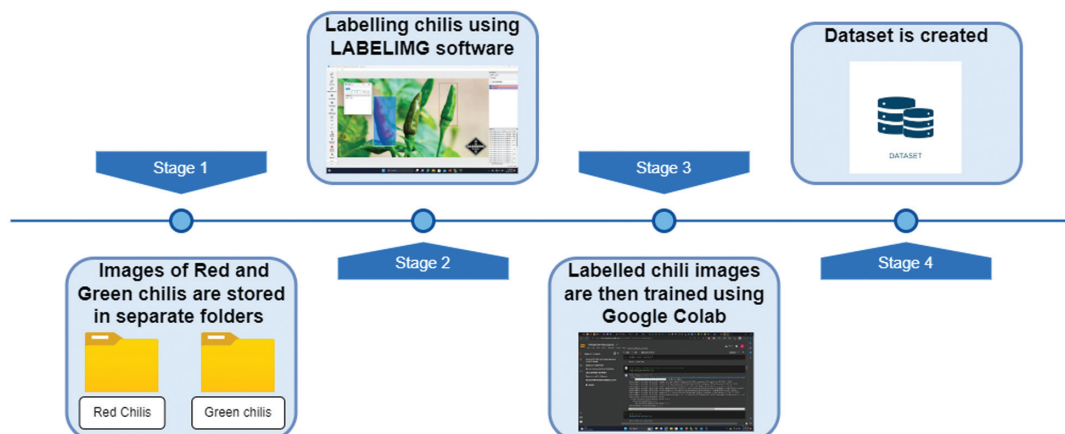
Fig. 2. The logic behind the generation and labeling of red and green chili datasets.

designated location. The Word file and its associated images are saved to a repository. Images are trained using the Google Colab graphical processing unit (GPU) platform after the compressed file is uploaded to a Google Drive subfolder. GPU processing can be enabled by training these photos and importing the necessary tools into Google Colab. The captured picture is utilized as a benchmark against which the learnt dataset may be evaluated.

In conclusion, the YOLO item identification technique can be used to categorize bird eye chilies based on their size, color, and general appearance. Using annotated datasets and deep learning frameworks, the YOLO model might be a helpful tool for tracking food quality and determining how long delicate goods can be stored. It has many potential uses, including in agriculture (particularly vegetable cultivation), processing, pest detection [29], and preserving food [30]

IV. EXPERIMENTAL SETUP

The experimental setup will consist of the following components:

- Conveyor Belt: The chilies will be transported by conveyor belt to the inspection and sorting machinery. For reliable identification and sorting, the belt will be built to keep the chilies moving at a constant velocity.

- CSI Camera: To get clear shots of the chili, a high-resolution camera will be mounted above the conveyor belt.

- Processing System: A Raspberry Pi 4B computer with the latest version of YOLO (V5) installed will serve as the processing system. The system will receive the collected photos of chilies and analyze them for detection and categorization.

- Chili Classification Model: Images of various bird chilies will be used to train a categorization model. The chilies will be detected and categorized in real-time using YOLO V5.

A camera on a tripod is set up above the sorting apparatus. The shaking causes a loss of precision. Red and green chilies are separated using a camera and a Raspberry pi. The Raspberry pi

is the system's central processing unit and data storage device, with its 4GB SD RAM it helps in smoother multitasking and better performance in resource-intensive applications. CSI cameras monitor chili on the conveyor belt for any indications of manipulation. The Raspberry pi 4B acts as the brains of the operation, delivering signals to the sorting motor that sends the red chilies to the left, the green chilies to the right, and the defective chilies and other commodities in a straight line. We may use the same camera to count both the red and green chilies. The quality of the chilies is evaluated in real-time by comparing photos to a database. The quality of chili can be judged based on its size, shape, color, and deformations. Real-time images are classified as red or green based on historical images. Bounding boxes are used to infer the positions of the peppers. The sums of red and green chilies were determined by counting how many fell into their respective border boxes.

A CV camera examines the chili patterns on the moving conveyor belt. A robotic manipulator is now sorting this batch of chilies by color. After a quality assessment of the detection is finished, the previously trained dataset can tell the difference between red and green chilies. The worst chilies are pushed off a conveyor belt, while the better red and green chilies are sorted out.

In Fig. 3, we see the outcome of applying a bounding box detection algorithm on a collection of red chili peppers; the detected peppers are surrounded by colored boxes that match to their respective labels. The acquired real-time image is compared to the trained models, which was fed a total of 1558 images. It lists the total number of chilies in each image and the total number of green and red chilies.

From Table I, we can figure out the overall accuracy of the system. The table shows the number of identified chilies from the total number of chilies present in a frame along with their accuracy percentage for each attempt. The accuracy in detecting both green and red chilies can be understood separately. At the first attempt, all the six chilies present in the frame were able to get identified by the system giving an accuracy of 100%. At the first attempt, the proposed system could identify all chillies present in the frame, giving an accuracy of 100%. Only 17 of the 18 red chillies and 8 of the nine green chillies were identified during the second attempt. At

Fig. 3. The images were detected with bounding boxes containing information regarding the aggregate quantity of chilies present and the number of chilies categorized by their respective colors.

Table I.   Accuracy of bird eye chili detection framework using YOLO V5

No. of attempts	Total no. of red and green chilies in a frame	Total no. of chilies identified	No. of red chilies identified	No. of green chilies identified	Red chili accuracy (%)	Green chili accuracy (%)
1	06/05	11	6	5	100	100
2	08/10	17	8	9	100	90
3	04/03	7	4	3	100	100
4	09/15	19	7	11	77.7	73.33
5	11/6	16	10	6	90.9	100
6	05/13	16	5	11	100	84.61
7	07/07	13	7	7	100	100
8	09/15	21	8	13	80.8	86.66
9	03/05	8	3	5	100	100
10	14/06	16	11	6	78.5	100
Average accuracy					92.79	93.46

the sixth attempt, the total number of chilies in a frame was about 18 and the total number of detected chilies was only 16, where out of the 13 green chilies only 11 were detected. This showed that as the number of chilies increased the percentage of accuracy decreased proportionally. The attempt was repeated for about ten times for which the red chilies gave an average accuracy of 92.79% and for green chilies the average accuracy rate is about 93.46%. We can understand that the accuracy of chilies identified decreases with the total number of chilies present. The decrease in accuracy may be due to the broken off chilies or physically damaged chilies, the blurred images can also be a cause.

The proposed methodology entails the operation of the experimental setup in the following manner:

- The chilis are slated to be positioned onto the conveyor belt.
- The CSI camera shall acquire visual representations of the chili peppers in motion along the conveyor belt.

Fig. 4. The developed working model of the proposed system.

- The acquired visual data will be transmitted to the processing infrastructure implemented in Raspberry pi 4B utilizing YOLO V5.
- The YOLO V5 algorithm can identify and categorize chili peppers based on their respective species.
- The data pertaining to classification shall be duly recorded for analysis.

The developed computer vision-assisted Raspberry Pi 4 B-based system for plucking chilies according to their ripping stage is shown in Figure 4. The system comprises a robotic manipulator, servo motors, and its control unit Raspberry pi 4B. The system works with the help of CV, provided by the camera that sends in visual input to the system. This input is then processed by the Raspberry pi control unit, which is a debit card sized low-cost computer developed with independent processor, memory, and graphics drivers. This control unit can act as an independent computer as it has its own memory and processor that could process the data on its own.

The Raspberry pi unit has been programmed to enable comparison of the captured image with the previously trained models of chili known as the datasets. The YOLO V5 framework, which compares the acquired photos to a set of pretrained data, is used to train the system. Using the labeling software, these datasets are produced by labeling the models or reference photos. Self-identification and the detection of previously trained samples are both made easier by this technology. On identifying these chilies a signal is delivered to the robotic manipulator, which is operated by three Tower pro MG995 servo motors. These metal gear servos are capable of moving arm, base, and gripper portion of the manipulator in desired angles.

The robotic manipulator is moved to perform the sorting operations with the green chilies being sorted to the right and the red chilies being sorted to the left of the conveyor belt based on the information about the coordinates collected from the image. To move the manipulator to exact positions, a mathematical approach called inverse kinematics is used, which determines the angle for the manipulator's movement.

Figure 5 shows three different cases to find angle of the elbow or the middle part with the coordinate (x,y)
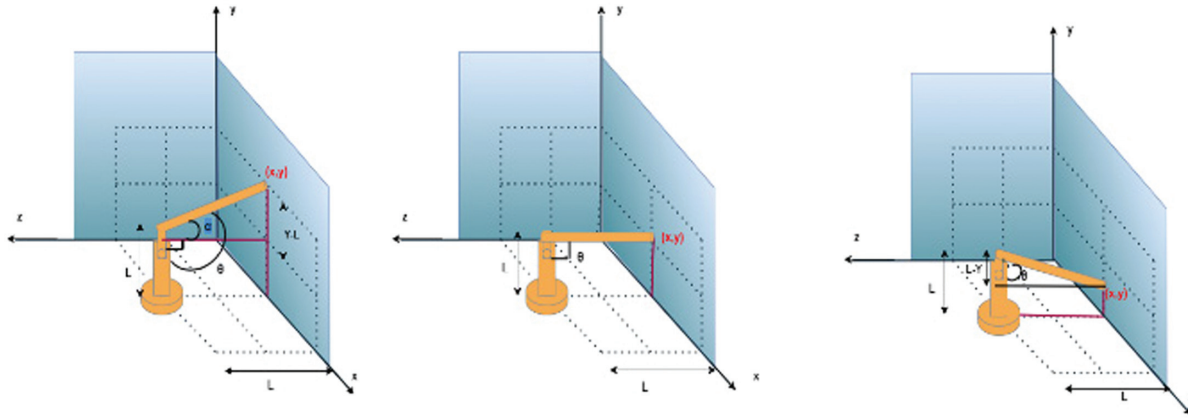
$$\text{If } y > l, \ tan\alpha = \frac{y-l}{l} \tag{1}$$

Fig. 5. Inverse kinematics of elbow part with the coordinate.

$$\theta = 90° + \alpha \qquad (2)$$

$$\text{If } y = l, \ \theta = 90° \qquad (3)$$

$$\text{If } y < l, \ tan\theta = \frac{l}{l - y} \qquad (4)$$

Figure 6 also gives three different cases to find angle of the base with coordinate in (x,y) plane.

$$\text{If } x < l, \ tan\alpha = \frac{l - x}{l} \qquad (5)$$

$$\theta = 90° + \alpha \qquad (6)$$

$$\text{If } x = l, \ \theta = 90° \qquad (7)$$

$$\text{If } x > l, \ tan\theta = \frac{l}{x - l} \qquad (8)$$

Equations 5–8 are used to identify the precise angle for the movement of the robotic manipulator. They are programmed in such a way that the coordinated images are utilized on a real-time basis to identify the angles.

The proposed system along with sorting operation does quality check in a similar manner by comparing the real-time footage captured with pre-trained datasets. The YOLO V5 framework enables a fast and highly accurate trained models. Further, they are a single shot detector that uses a CNN. This facilitates a better

sorting of chilies based on the ripping stage and quality. The control unit of Raspberry pi along with the manipulator performs the sorting task in a fully automated manner. The system is powered by batter source that can be recharged from time to time.

V. CONCLUSION

This paper discussed the process for developing a 3-DoF robotic manipulator guided by CV for use in the agricultural sector. It employed CV technology based on the YOLO V5 architecture to recognize red and green bird eye chilies in real-time. The identification and classification procedures in the proposed system are all managed by a Raspberry pi 4B GPU computer. The YOLO V5 framework compared the detected data in real-time with the trained database. The robotic manipulator used a color comparison method to distinguish between red and green chilies before sorting them. The manipulator's servo motor aids in the sorting process, while the Raspberry pi, which also serves as the system's control unit, uses color recognition to determine how to divide the chilies. The systems used the same technology of chili detection to execute the sorting of red and green chilies to a predetermined location and quality assessment. If put into practice, this technique has the potential to free farmers from the time-consuming chore for many hours. Even if it does not use the Raspberry pi design, this technology might help farmers track the number of chilies they produce. The YOLO V5 framework allows the system to be further developed into a fully automated robotic manipulator capable of
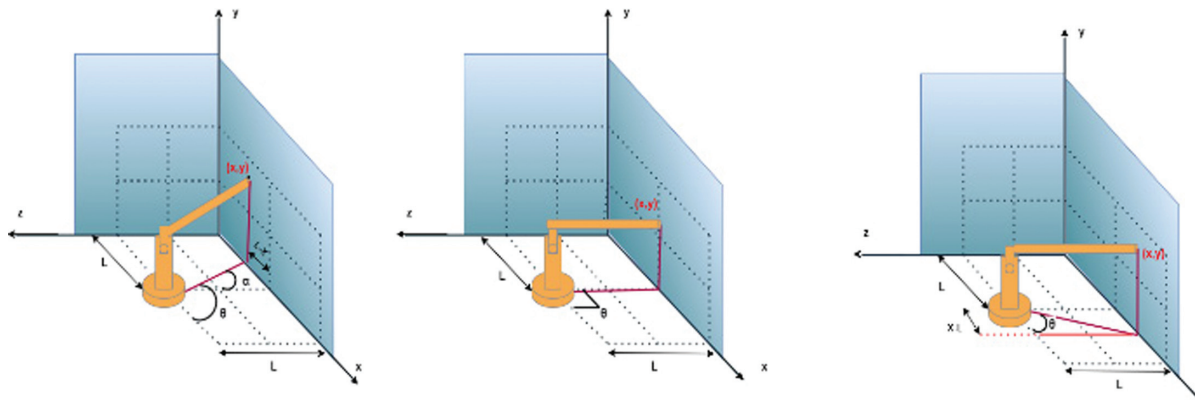

Fig. 6. Inverse kinematics of base part with coordinate.

detecting chilies with CV technology and distinguishing between red and green chilies. The suggested technology has much room to grow into a robot farm hand that can pick chilies with precision.

CONFLICT OF INTEREST STATEMENT

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

REFERENCES

[1] S Kenig, A. Baruca-Arbeiter, N. Mohorko, M. Stubelj, M. Ä. ErneliÄ-Bizjak, D. Bandelj, Z. Jenko-Praanikar, and A. Petelin, "Moderate but not high daily intake of chilies pepper sauce improves serum glucose and cholesterol levels," *J. Funct. Foods*, vol. 44, pp. 209–217, 2018.

[2] N. Yamani, A. Musheer, P. Gosain, S. Sarfraz, H. Qamar, M. M. Waseem, M. S. Arshad, T. Almas, and V. Figueredo, "Meta-analysis evaluating the impact of chilies-pepper intake on all-cause and cardio-vascular mortality: a systematic review," *Ann. Med. Surg.*, vol. 70, 102774, pp. 1–4, 2021.

[3] R. K. Mahendran, S. A. Vino, and S. Anandakumar, , *Fundamentals of Computer Vision System for Sorting and Grading of Food Products, Reference Module in Food Science*: Elsevier eBooks, pp. 1–9, 2016, Doi: https://doi.org/10.1016/b978-0-08-100596-5.03092-4.

[4] T. Ayoub Shaikh, T. Rasool, and F. Rasheed Lone, "Towards leveraging the role of machine learning and artificial intelligence in precision agriculture and smart farming," *Comput. Electron. Agric.*, vol. 198, pp. 1–13, 2022. Doi: https://doi.org/10.1016/j.compag.2022.107119.

[5] T. Brosnan and D.-W. Sun, "Improving quality inspection of food products by computer vision-a review," *J. Food Eng.*, vol. 61, no. 1, pp. 3–16, 2004.

[6] Y. Tace, M. Tabaa, S. Elfilali, C. Leghris, H. Bensag, and E. Renault, "Smart irrigation system based on IoT and machine learning," *Energy Rep.*, vol. 8, no. Supplement 9, pp. 1025–1036, 2022.

[7] A. S. Shaikat, S. Akter, and U. Salma, "Computer vision based industrial robotic arm for sorting objects by color and height," *J. Eng. Adv.*, vol. 1, no. 04, pp. 116–122, 2020.

[8] M. Guidolin, L. Tagliapietra, E. Menegatti, and M. Reggiani, "Hi-ROS: open-source multi-camera sensor fusion for real-time people tracking," *Comput. Vis. Image Underst.*, vol. 232, pp. 1–8, 2023.

[9] C. Gupta, V. Tewari, R. Machavaram, and P. Shrivastava, "An image processing approach for measurement of chilies plant height and width under field conditions," *J. Saudi Soc. Agric. Sci.*, vol. 21, no. 3. pp. 171–179, 2022. Doi: https://doi.org/10.1016/j.jssas.2021.07.007.

[10] M. D. Abdullah-Al-Noman, A. N. Eva, T. B. Yeahyea, and R. Khan, "Computer vision-based robotic arm for object color, shape, and size detection," *J. Robot. Control (JRC)*, vol. 3, pp. 180–186, 2022.

[11] L. Fernandes and B. R. Shivakumar, "Identification and sorting of objects based on shape and color using robotic arm," *Proc. Fourth Int. Conf. Invent. Syst. Control (ICISC 2020)*, Coimbatore, India, vol. 1, pp. 866–871, 2020. Doi: https://doi.org/10.1109/ICISC47916.2020.9171196.

[12] H. F. Thariq Ahmed, H. Ahmad, and C. V. Aravind, "Device free human gesture recognition using Wi-Fi CSI: a survey," *Eng. Appl. Artif. Intell.*, vol. 87, pp. 1–19, 2020.

[13] J. D. Arangoa, B. Staara, A. M. Baig, and M. Freitagac, "Quality control of apples by means of convolutional neural networks – comparison of bruise detection by color images and near-infrared

[14] K. Utai, M. Nagle, S. Hämmerle, W. Spreer, B. Mahayothee, and J. Müller, "Mass estimation of mango fruits (Mangifera indica L., cv. 'Nam Dokmai') by linking image processing and artificial neural network," *Eng. Agric. Environ. Food*, vol. 12, pp. 103–110, 2019.

[15] A. Rakshit, S. Pramanick, A. Bagchi, and S. Bhattacharyya, "Autonomous grasping of 3-D objects by a vision-actuated robot arm using Brain–Computer Interface," *Biomed. Signal Process. Control*, vol. 84, pp. 1–14, 2023.

[16] H. Ponce, C. Cevallos, R. Espinosa, and S. Gutiérrez. "Estimation of low nutrients in tomato crops through the analysis of leaf images using machine learning," *J. Artif. Intell. Technol.*, vol. 1, no. 2, pp. 131–137, 2021.

[17] I. Iraei, M. Sharifi, and Y. Baleghi, "A new approach to enhance the estimation of blur parameters in blurred images," *Optik*, vol. 224, pp. 1–20, 2020.

[18] Y. Wang, L. Wang, H. Lu, and Y. He, "Segmentation based rotated bounding boxes prediction and image synthesizing for object detection of high resolution aerial images," *Neurocomputing*, vol. 388, pp. 202–211, 2020.

[19] M. Javaid, A. Haleem, R. P. Singh, and R. Suman, "Enhancing smart farming through the applications of agriculture 4.0 technologies," *Int. J. Intell. Netw.*, vol. 3, pp. 150–164, 2022.

[20] M. H. Dairath, M. W. Akram, M. A. Mehmood, H. U. Sarwar, M. Z. Akram, M. M. Omar, and M. Faheem, "Computer vision-based prototype robotic picking cum grading system for fruits," *Smart Agric. Technol.*, vol. 4, pp. 1–10, 2023.

[21] A. K. Pothula, Z. Zhang, and R. Lu, "Evaluation of a new apple in-field sorting system for fruit singulation, rotation and imaging," *Comput. Electron. Agric.*, vol. 208, pp. 1–5, 2023.

[22] Y. Suryawanshi, K. Patil, and P. Chumchu, "VegNet: dataset of vegetable quality images for machine learning applications," *Data Brief*, vol. 45, pp. 1–7, 2022.

[23] M. D. H. Rahman, M. K. A. Jannat, M. D. S. Islam, G. Grossi, S. Bursic, and M. D. Aktaruzzaman, "Real-time face mask position recognition system based on MobileNet model," *Smart Health*, vol. 28, pp. 1–12, 2023.

[24] C. Jiang, H. Ren, X. Ye, J. Zhu, H. Zeng, Y. Nan, M. Sun, X. Ren, and H. Huo, "Object detection from UAV thermal infrared images and videos using YOLO models," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 112, pp. 1–17, 2022.

[25] J. Yan and Z. Wang, "YOLO V3 + VGG16-based automatic operations monitoring and analysis in a manufacturing workshop under Industry 4.0," *J. Manuf. Syst.*, vol. 63, pp. 134–142, 2022.

[26] T.-T.-H. Vu, D.-L. Pham, and T.-W. Chang, "A YOLO-based real-time packaging defect detection system," *Proc. Comput. Sci.*, vol. 217, pp. 886–894, 2023.

[27] L. Yu, J. Xiong, X. Fang, Z. Yang, Y. Chen, X. Lin, and S. Chen, "A litchi fruit recognition method in a natural environment using RGB-D images," *Biosyst. Eng.*, vol. 204, pp. 50–63, 2021.

[28] A. Narayan, Elio Tuci, F. Labrosse, and M. H. M. Alkilabi, "A dynamic color perception system for autonomous robot navigation on unmarked roads," *Neurocomputing*, vol. 275, pp. 2251–2263, 2018.

[29] S. Likith, B. R. Reddy, and K. S. Reddy, "A smart system for detection and classification of pests using YOLO and CNN techniques," in *2021 Int. Conf. Comput. Perform. Eval. (ComPE)*, Shillong, India, 2021, pp. 049–052. Doi: https://doi.org/10.1109/ComPE53109.2021.9752185.

[30] L. Zhu and P. Spachos, "Support vector machine and YOLO for a mobile food grading system," *Internet Things*, vol. 13, pp. 1–18, 2021.

images," *Procedia CIRP*, pp. 290–294, 2021. Doi: https://doi.org/10.1016/j.procir.2021.03.043.