

# Application of Stereo Vision to Control the Movement of the Robot Arm Towards the Position of Red Chilies

I Dewa Made Subrata<sup>1,✉</sup>, Ahmad Dzaky Baiquni<sup>1</sup>

<sup>1</sup> Department of Mechanical and Biosystem Engineering, Faculty of Agricultural Technology, IPB University, Bogor, INDONESIA.

## Article History:

Received : 05 February 2024

Revised : 19 March 2024

Accepted : 28 March 2024

## Keywords:

Red chili fruit,  
Robot arm,  
Stereo vision,  
Three-dimensional position.

Corresponding Author:

✉ [dewamadesubrata.ipb@gmail.com](mailto:dewamadesubrata.ipb@gmail.com)

(I Dewa Made Subrata)

## ABSTRACT

*The trend of decreasing young workers in the agricultural sector needs to be anticipated by developing intelligent machines known as agricultural robots. This research aims to apply a stereo vision system to control the movement of the robot's grip towards the 3D position of the red chili fruit. The stereo vision system installed on the robot waist (joint-2) is used to capture plant images and process them using HSV masking filters and triangulation principal to obtain the 3D center point position of the fruit. The robot joint movement is calculated using geometric based inverse kinematics. The research results show that the average accuracy of the stereo vision system is 93.9 %. The average grip positioning accuracy is 95.6 % to the actual chili fruit position and 98.5 % to the stereo vision calculation value. The average stability of the stereo vision values is 99.5 %, while the average positioning stability of the robot's grip is 99.6 %. Time consumption for image processing is 0.053 s while time consumption for robot grip movement is 9 s. Therefore, the stereo vision system can be used to control robot's grip movement with a good accuracy.*

## 1. INTRODUCTION

Red chili fruit is a horticultural commodity that has high economic value and can be cultivated both in open fields and in greenhouses. The fruit of the chili plant can be sold when it is still green or when the fruit is ripe and red (Nurhafsa *et al.*, 2021). Chili plants are susceptible to disease attacks, especially if cultivated in open fields, therefore it is recommended to cultivate them in a greenhouse or screenhouse. Moekasan *et al.*, (2015) conducted a comparative study between chili cultivated in a screenhouse and cultivated in open fields. The results show that cultivation in a screenhouse can reduce insecticide costs by 73.19 %, and increase production by 106.45 % to 109 %. Delya *et al.*, (2014) cultivates chili plants using a hydroponic system that is controlled automatically in a greenhouse. The research report shows that cultivation using an automatic hydroponic system is better than manual cultivation. Not only cultivation in a greenhouse, increasing production can also be done by selecting superior seeds. Ekowahyuni & Wulansari, (2022) conducted a cultivation comparative study between the IPB University breeding hybrid chili plant and commercial hybrid chilies. The research report shows that hybrid chilies produced by IPB University are superior and provide higher yields compared to commercial hybrid chilies.

Chili plants are an agricultural commodity that requires a lot of labor to cultivate. However, the development of the agricultural sector in both developing and developed countries is facing problems in the availability of agricultural labor, especially the number of young farmers continues to decline, while older farmers are increasing (Susilowati, 2016). To overcome the problem of decreasing young workers in the agricultural sector, efforts need to be made

through the application of agricultural machines that work automatically, known as agricultural robots. Many researches have been carried out on the use of robots for fruit harvesting activities in greenhouses (Onishi *et al.*, 2019; Hua *et al.*, 2019; Barbashov *et al.*, 2022; Yoshida *et al.*, 2022; Wang *et al.*, 2022). In general, agricultural robots are designed in the form of an arm that has a working principle similar to a human arm. In order to be able to work automatically, the robot arm is equipped with at least several basic components including: manipulator, end-effector, visual system, and control system (Kaleem *et al.*, 2023). Autonomous robots are also equipped with a transportation system (Jia *et al.*, 2020). Several types of manipulators have been used in agricultural robots, including: Cartesian type (Subrata & Melania, 2023), articulated type (Feng *et al.*, 2018; Tang *et al.*, 2020), a combination of Cartesian type and articulated type (Raja *et al.*, 2022). Several visual systems have been used in agricultural robots, including: RGB-D cameras (Yoshida *et al.*, 2022), combination of camera and laser sensor (Feng *et al.*, 2018), stereo camera (Sangeetha *et al.*, 2018; Taryudi, 2018; Ibadillah, 2018; Angulo *et al.*, 2022; Erceg *et al.*, 2023; Umam *et al.*, 2023). The aim of this research is to utilize information obtained from stereo vision to control the movement of the robot arm towards the position of red chili fruit hanging on the tree.

## 2. MATERIALS AND METHODS

The materials used in this research were red chili plants of the Tanjung 2 variety with a plant age of 50 days after planting which was cultivated in pots. The tools used include: 2 Logitech type C270 cameras with an image resolution of 640 x 480 pixels and a focal length range of 4 mm, one unit of Mitsubishi Movemaster type RV-M1 articulated manipulator, Asus X200CA notebook, ruler, and millimeter block paper. The research was carried out in the instrumentation and control laboratory, department of mechanical and biosystems engineering, Faculty of agricultural technology, IPB University from August to December 2023.

### 2.1. Stereo Vision

The robot sensing system used in this research is a stereo vision system, namely two Logitech C270 type cameras installed side by side with a separation distance between the two lenses is 8 cm (Figure 1). The image of a chili plant placed in front of the camera will be captured by two cameras (stereo vision) at the same time using the Python and OpenCV programming languages. The resulting RGB (Red Green Blue) image is then converted into an HSV (Hue Saturation Value) image using the function `cv2.cvtColor(image, cv2.COLOR_BGR2HSV)`. Red chili fruit pixels are separated from other pixels such as leaves, stems and unripe fruit using a masking filter with minimum HSV value and maximum HSV value using the `cv2.inRange(HSV_image, HSV_minimum, HSV_maximum)` function. The clustered of red chili pixels were then traced for their contour edges using the `cv2.findContours()` function. After obtaining the contour values, the number of red chilies, the center point of each cluster, and the size of the fruit (length and width) were calculated using the `cv2.moments()` and `cv2.boundingRect()` functions. In this research, all of the OpenCV functions mentioned above are combined for image processing purposes to obtain the 2D center point position of the fruit, then the three-dimensional position is calculated using equations 1 to 3. The flow diagram of the algorithm for determining the center point of red chilies is shown in Figure 2.

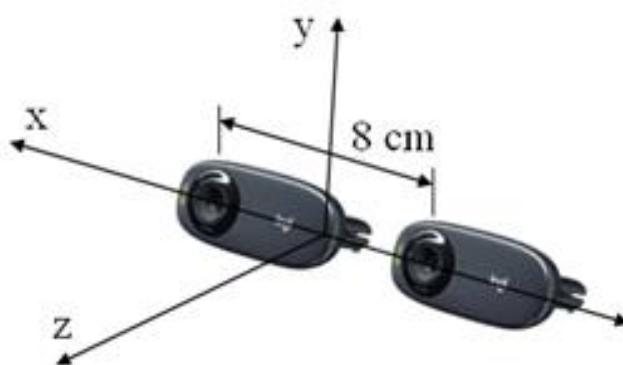


Figure 1. Stereo vision system.

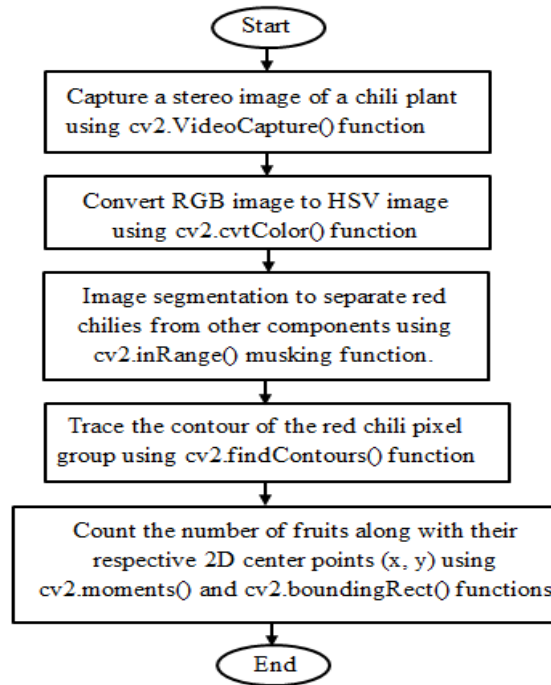


Figure 2. Flow diagram of the algorithm for determining the center point of red chilies.

As explained above, the red chili pixels were separated from pixels of the leaf, stem and green fruit components by taking RGB data as follows: the red chili image as much as 375 points, green chilies about 225 points, stems and leaves about 374 points at random positions in the image. The R, G and B values from that point are converted into H, S and V values and then plotted the relationship between G and R values, B and R value, S and H value, V and H value. The graph that can clearly show the separation between red chili fruit pixels and other plant components are used as reference values to separate red chili fruit pixels from other pixels.

### 2.1.1. Three-Dimensional Positioning with Stereo Vision

The three-dimensional position of the center point of the red chili fruit was calculated using the triangulation principle from images captured by left and right stereo vision cameras (Figure 3). If the lens focal distance from the camera to the sensitive surface of the camera ( $f$ ) is known, then the 3D center point position of the fruit can be calculated using equations 1, 2 and 3.

$$Z = \frac{f \times L}{X_{i2} - X_{i1}} \quad (1)$$

$$X = \frac{X_i \times Z}{f} \quad (2)$$

$$Y = \frac{Y_i \times Z}{f} \quad (3)$$

where  $Z$  is the distance from the camera lens to the object (cm),  $L$  is the distance between the two stereo cameras (cm),  $f$  is the distance of the lens to the sensitive surface of the camera sensor (pixel),  $X_{i2} - X_{i1}$  is shift distance of the center point of red chili fruits in the horizontal direction (pixel),  $X_i$  is the distance from the center point of the fruit to the center point of the left image in the horizontal direction (pixel),  $X$  is the distance from the center of the fruit to the center of stereo vision in the horizontal direction (cm),  $Y_i$  is the distance from the center point of the fruit to the center point of the left image in the vertical direction (pixel), and  $Y$  is the distance from the center of the fruit to the center of stereo vision in the vertical direction (cm).

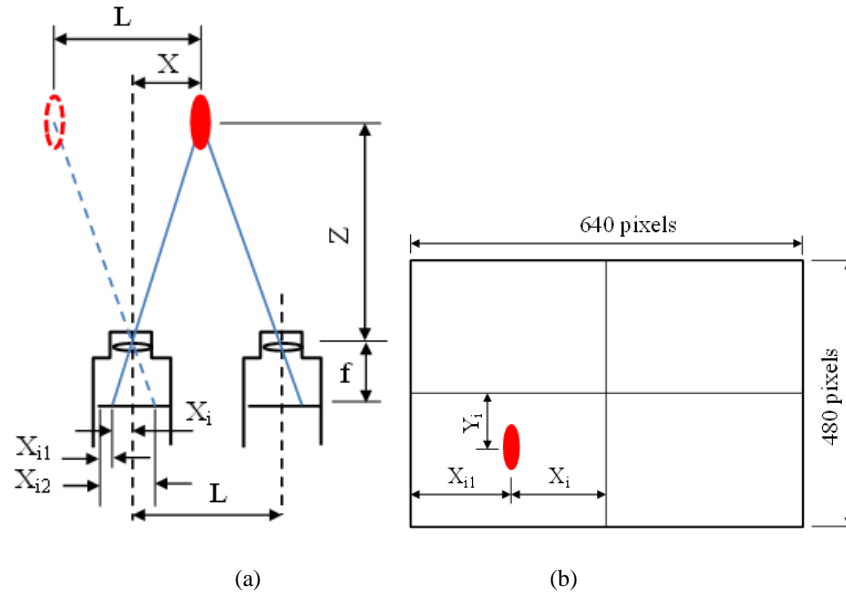


Figure 3. Stereo vision: (a). triangulation principle; (b). left image captured by stereo vision.

The value of the variable  $f$  is found by placing red paper in front of the stereo vision in the  $Z$  axis at distance ranging from 300 mm to 600 mm with increments of 10 mm so that the pixel shifting value ( $X_{i2} - X_{i1}$ ) in the stereo vision can be calculated for each  $Z$  value and ( $X_{i2} - X_{i1}$ ) parameter are then entered into equation 1 to obtain the  $f$  value in pixel units.

## 2.2. Robot Arm

The robot arm used in this research is the Mitsubishi Movemaster RV-M1 type with the physical specifications as shown in the Figure 4. Robot arm weight is 20.2 kgf, robot driver weight is 23 kgf, robot power consumption is 612 watts.

Stereo vision is installed on the robot arm and the 3D position obtained from stereo vision is used to move the robot arm grip towards the center point position of the chili fruit. The robot arm used is an articulated type with five degrees of freedom, all of which are rotational joints (Figure 5)

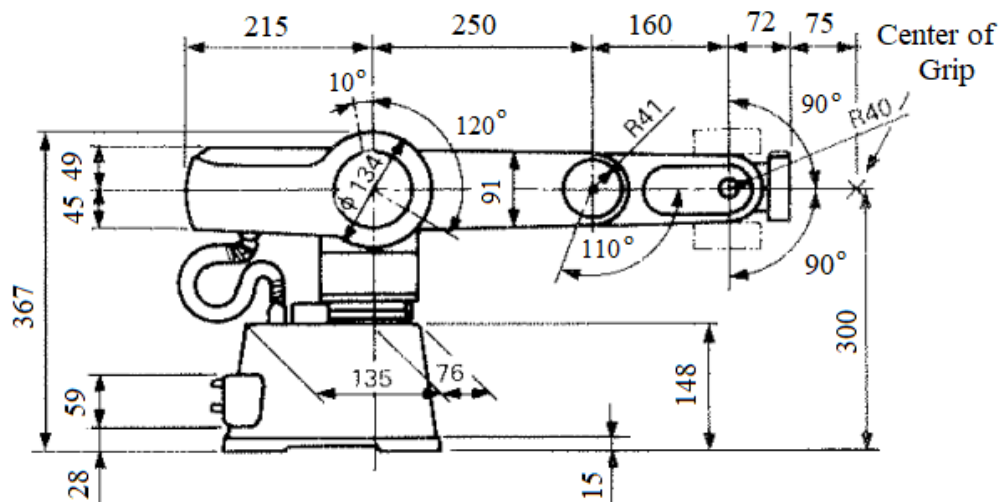


Figure 4. Physical specifications of the robot arm.

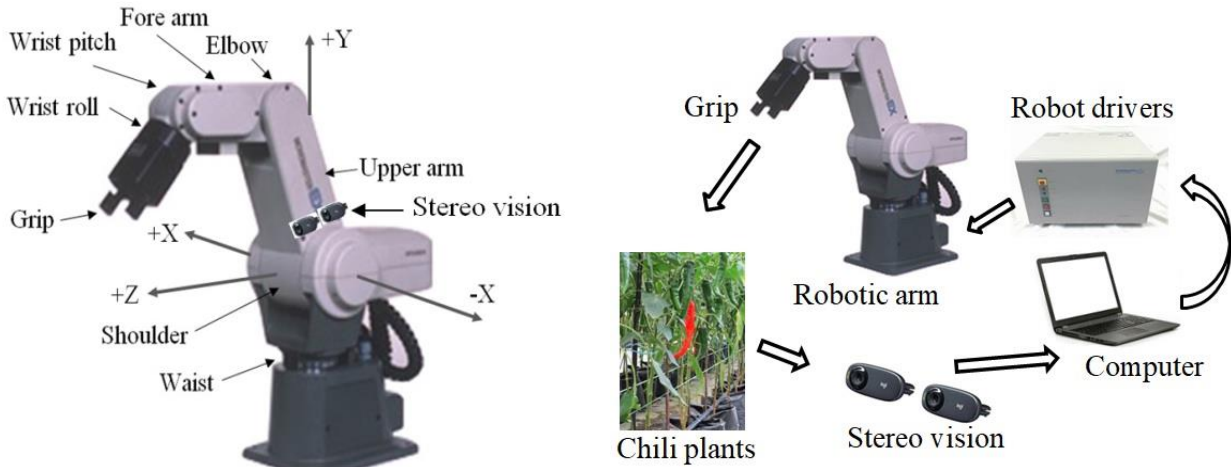


Figure 5. Articulated manipulator (robot arm) with 5 degrees of freedom (left), Grip positioning process flow diagram (right).

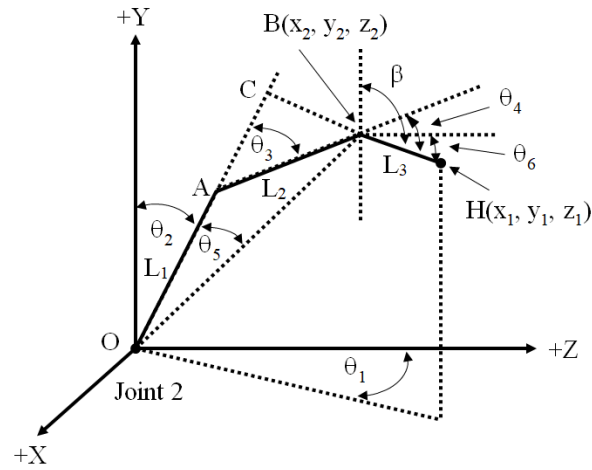


Figure 6. Sketch of the robot arm in Cartesian coordinates.

The maximum rotating ability of each joint is: Waist (joint 1) =  $300^\circ$ , Shoulder (joint 2) =  $130^\circ$ , Elbow (joint 3) =  $110^\circ$ , Wrist pitch (joint 4) =  $180^\circ$ , Wrist roll (joint 5) =  $360^\circ$  with the closest distance when the grip is horizontally oriented is 300 mm and the furthest is 557 mm. To move the grip of the robot arm to the position of the red chili fruit, it is necessary to derive an equation to change the 3D coordinate value of the red chili fruit center point into the absolute value of each joint which is commonly referred to as inverse kinematics. The same as the method used by [Angulo et al., \(2022\)](#) in this research the inverse kinematic calculations is done using geometric principles with the arm sketch in Cartesian Coordinates as shown in Figure 6. In this sketch, the center point of the manipulator's cartesian coordinates is at joint 2 position.

$$y_2 = y_1 - L_3 \sin(\theta_6) \quad (4)$$

$$x_2 = x_1 - L_3 \cos(\theta_6) \sin(\theta_1) \quad (5)$$

$$z_2 = z_1 - L_3 \cos(\theta_6) \cos(\theta_1) \quad (6)$$

$$\theta_1 = \tan^{-1} \left( \frac{x_1}{z_1} \right) \quad (7)$$

$$\theta_3 = 180 - \cos^{-1} \left( \frac{L_1^2 + L_2^2 - x_2^2 - y_2^2 - z_2^2}{2L_1L_2} \right) \quad (8)$$

$$\theta_2 = \cos^{-1}\left(\frac{y_2}{\sqrt{x_2^2 + y_2^2 + z_2^2}}\right) - \sin^{-1}\left(\frac{L_2}{\sqrt{x_2^2 + y_2^2 + z_2^2}} \sin(\theta_3)\right) \quad (9)$$

$$\theta_4 = 90 - \theta_6 - \theta_2 - \theta_3 \quad (10)$$

where  $x_1, y_1, z_1$  are coordinates of the center point of the robot arm grip,  $x_2, y_2, z_2$  are wrist pitch coordinates (joint 4),  $L_1$  is upper arm length (mm),  $L_2$  is fore arm length (mm),  $L_3$  is the length of the link from the center of the wrist pitch to the center of the robot arm grip (mm),  $\theta_1$  to  $\theta_4$  is angle of joint ( $^\circ$ ),  $\theta_6$  is angle of link  $L_3$  to the horizontal plane ( $^\circ$ ).

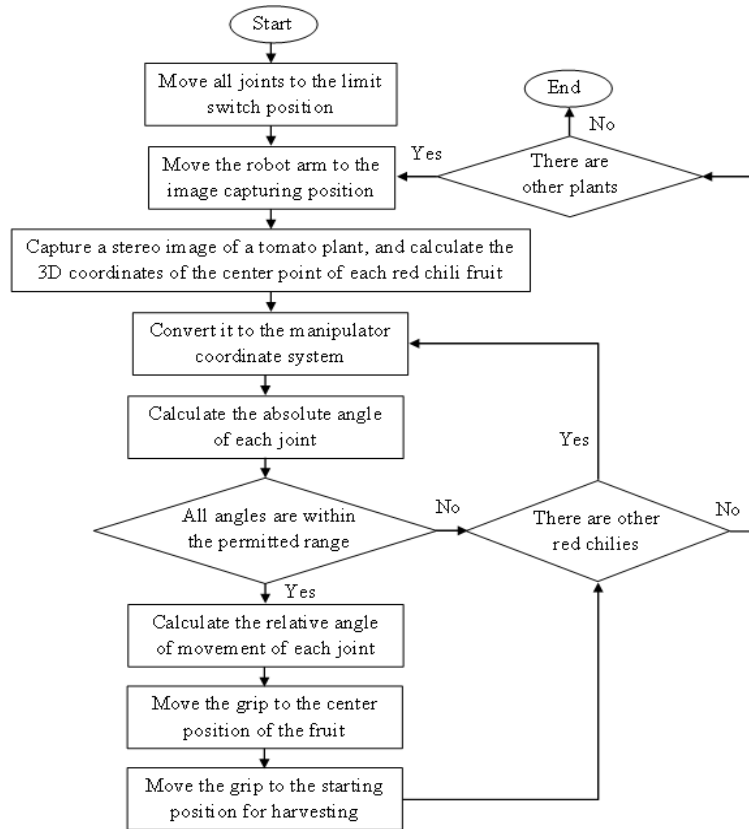


Figure 7. Robot arm motion control algorithm

### 2.2.1. Robotic Arm Control

In this research, stereo vision is installed on the left side of joint 2, therefore the robot arm needs to be moved 90 degrees clockwise so that the stereo vision scanning area directed towards the plant being placed in front of the robot. The computer program then captures and processes the captured image to obtain information on the three-dimensional position of the center point of the red chili fruit. The 3D center point position of the chili fruit with respect to the Cartesian coordinates of the stereo vision is converted into a 3D position of the fruit with respect to the Cartesian coordinates of the robot arm. The 3D center point position of the fruit which become the target point for the grip movement is then converted into an absolute angle for each joint using the inverse kinematic equation. The absolute angle value of each joint is checked for compliance with the permitted angle range. If all joints are within the permitted range, it means the grip can be moved to the position of the center point of the fruit. The movement of each joint is carried out using the relative angle value, namely the absolute value resulting from the inverse kinematic calculation minus the absolute angle value of the current position (before being moved). After being moved to the 3D position of the red chilies fruit then moved again to the initial harvesting position. The complete robot arm motion control algorithm is presented in Figure 7.

### 2.2.2. Data Collection and Processing

Some of the data measured in this research include: the actual 3D center point position of the red chili fruit captured by stereo vision, the 3D center point position of the grip after being moved to the position of the red chili fruit, the time required to take a stereo image until getting the 3D position value of the red chili fruit, the time required for the grip to move to the position of the red chili fruit. The measured data is then processed to determine the 3D positioning accuracy of the stereo vision and the 3D positioning accuracy of the robot arm grip movement.

## 3. RESULTS AND DISCUSSION

### 3.1. Stereo Vision

The stereo vision system has been installed on the left side of joint 2 robot arm using an acrylic plate as in Figure 8 so that the cameras are moving along with the motion of the robot arm. The chili plants cultivated in polybags is placed in front of the robot arm. Before capturing the image, the stereo vision lens needs to be directed towards the chili plants by moving joint 1 and joint 2 to the absolute angle position of  $90^\circ$ , joint 3 to the absolute angle position of  $-90^\circ$ , joint 4 to the absolute angle position of  $80^\circ$ . Images of plants containing red chilies, green chilies, stems and leaves were recorded, then the Red Green Blue (RGB) values from the image of the red chili fruit was taken at 375 points, the image of the green chili fruit was taken at 225 points, and the image of the stem and leaves at 375 points randomly. The R, G and B values from that point are converted into Hue, Saturation and Value (HSV) data and then the relationship between the Green and Red, Blue and Red, Saturation and Hue, Value and Hue values is plotted at Figures 9 and 10.

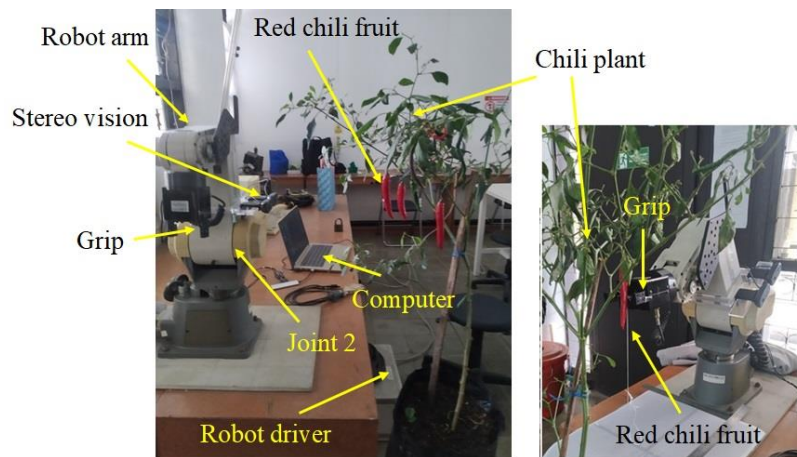


Figure 8. Robot arm while capturing images (left), and movement of the robot grip towards the position of the chili fruit (right).

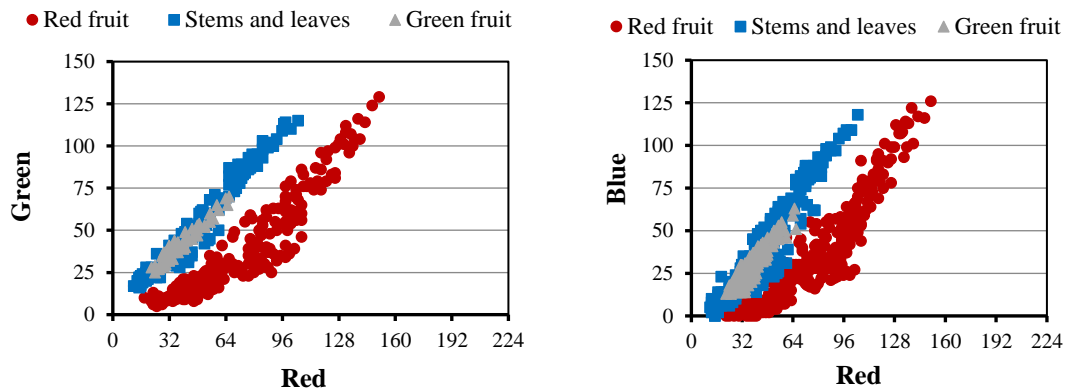


Figure 9. Plots of the RGB values: (left) G and R relationship; (right) B and R relationship.



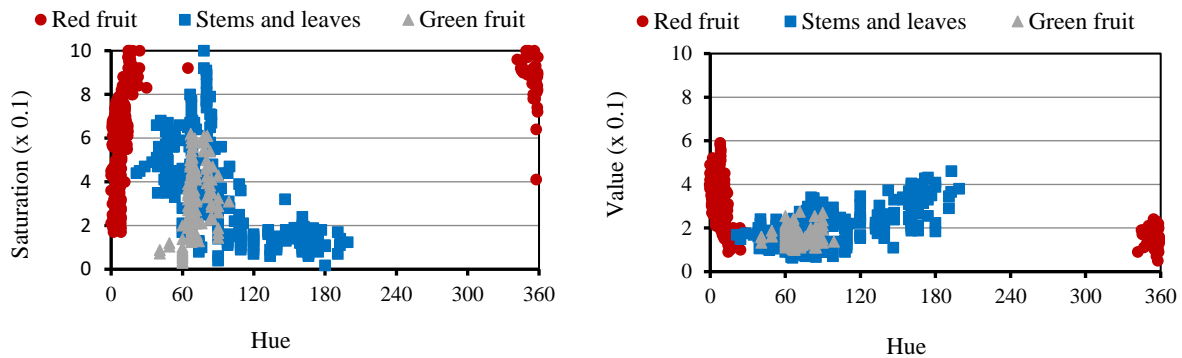


Figure 10. Plot of the HSV values: (left) S and H relationship; (right) V and H relationship.

The red chili pixels in Figure 9 are quite difficult to separate from other components, therefore the RGB values are not used for image segmentation. The red chili fruit pixels in Figure 10 are easier to separate from other components compared to Figure 9, therefore in this study the HSV value was used for image segmentation. The HSV parameter values used are  $351 \leq H \leq 360$ ,  $0 \leq H \leq 22$ ,  $0.17 \leq S \leq 1.00$  and  $0.05 \leq V \leq 0.59$ . Because of the HSV color model used in OpenCV has a value range of 0 to 180 for parameter H, 0 to 255 for parameter S and 0 to 255 for parameter V, therefore the values used to separate red chili pixels are  $175 \leq H \leq 180$ ,  $0 \leq H \leq 11$ ,  $42 \leq S \leq 255$ , and  $14 \leq V \leq 151$ . The HSV color model combined with other methods has been carried out by [Hadinegoro & Rizaldi, \(2021\)](#) to determine the ripeness of cayenne pepper with an accuracy of 93.58 %. The HSV color model combined with the SVM (Support Vector Machine) algorithm was carried out by [Indrabayu et al., \(2019\)](#) for classification of strawberry ripeness with an accuracy of 97%.

### 3.2. Determining the Focus Distance of a Camera Lens

The focal distance of the lens expressed by the variable  $f$  is used to calculate the 3D position of the red chili fruit using the triangulation principle. Therefore, the value of the variable  $f$  used in equations 1 to 3 needs to be calibrated. The calibration process is carried out by placing red chili fruit in front of the stereo vision system starting from a distance of 300 mm to 600 mm with increments of 10 mm. The image captured by the stereo vision system is then segmented into a binary image and then the distance from the center point of the red chili fruit (in pixels) to the left edge of the stereo image is calculated, which is expressed by the variables  $X_{i1}$  and  $X_{i2}$  respectively for the left camera and right camera images. The calibration results using equation 1 show that the average lens focal distance is 822 pixels.

### 3.3. Determination of the Center Point of Fruit

The image of the chili plant placed in front of the robot arm was taken using the left camera as seen in Figure 11a, then segmentation was carried out using HSV masking filter (Figure 11b) and its 2D position was calculated in pixel units (Figure 11c). At almost the same time, the chili plant image was also taken using the right camera (Figure 11d), then segmented using HSV masking filter (Figure 11e) and its 2D position was calculated in pixel units (Figure 11f).

The 2D center point position of the red chilies captured by the left and right cameras is then processed using equations 1-3 to obtain the 3D position of each red chili in stereo vision cartesian coordinates. In this study there were three conditions based on the number of red chili fruit captured in the image, namely only one red chili fruit, three red chili fruits and five red chili fruits. In this study, all the fruits used were oriented straight downward so no further processing was carried out to obtain the real orientation of each chili fruit especially chilies with a sideways orientation. Therefore, the grip orientation is still limited to chilies with a straight downward orientation.

The 3D position of the actual red chili fruit and the 3D position captured by stereo vision are then converted into the Euclidian value using equation 11. The 3D positioning error is also presented in the form of Euclidian value using equation 12.



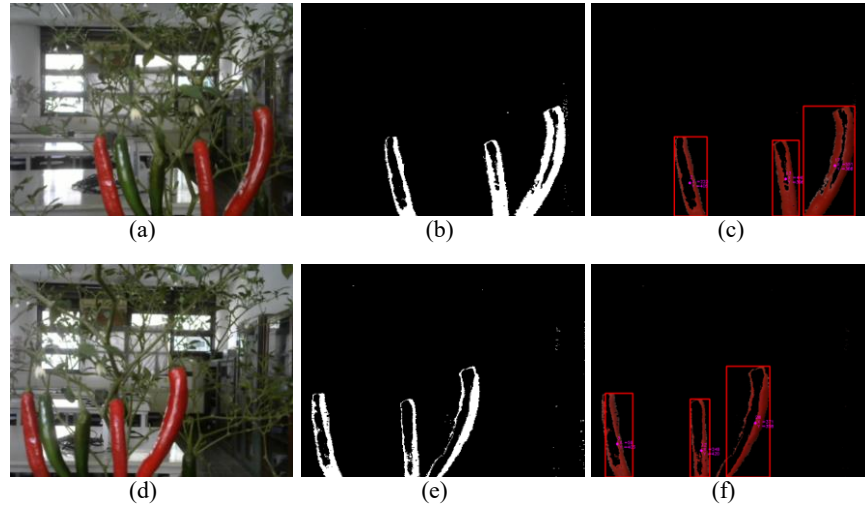


Figure 11. Results of stereo vision image processing: a to c for the left camera, d to f for the right camera.

$$Eu = \sqrt{x^2 + y^2 + z^2} \quad (11)$$

$$Er = \sqrt{E_x^2 + E_y^2 + E_z^2} \quad (12)$$

where  $Eu$  is Euclidian value of 3D position of red chili fruit (mm);  $x, y, z$  is actual or captured 3D position of the red chili fruit (mm);  $Er$  is Euclidian value of the 3D positioning error (mm), and  $E_x, E_y, E_z$  = Error on x-axis, y-axis and z-axis (mm).

The movement pattern of the grip center point towards the position of the chili fruit is starting from the scanning position (position 0), followed by rotation of joint 1 equal to the angle resulting from the inverse kinematic calculation minus the absolute angle at position 0 (movement 1 and the tip of the arrow indicating position 1) as shown in Figure 12. Furthermore, joint 2, 3, 4, 5 are rotated simultaneously so that the grip is moves towards the center of the fruit (Movement 2). In this study, the fruit was not picked but only measured the accuracy of the grip positioning relative to the position of the fruit. After the grip reaches the fruit position, joints 2, 3, 4, 5 are rotated again so that the grip moves to position 1 (movement 3). The movement pattern to the second fruit position is the same as the movement pattern to the first fruit position, except that it starts from position 1 and finishes at position 4 (arrow tip of movement 4 indicating position 4). The movement towards the third fruit starts from position 4 and finishes after movement 9.

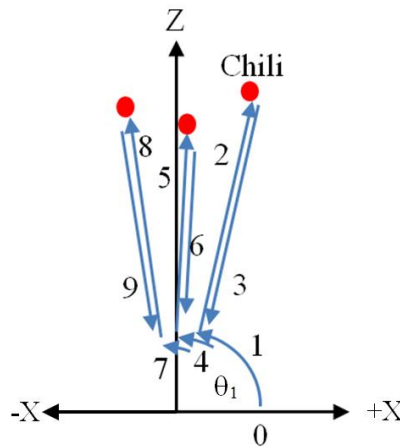


Figure 12. Illustration of the movement pattern of the center grip towards the position of the chili fruit (top view).

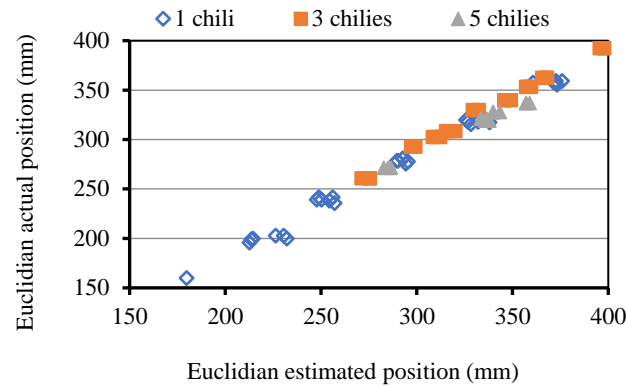


Figure 13. Euclidian 3D position of red chili fruit

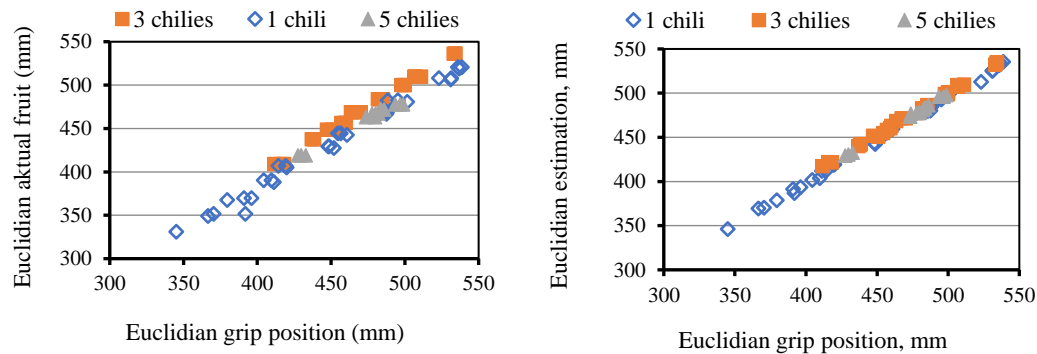


Figure 14. Relationship between Euclidian 3D actual position (left), 3D estimated position (right), and the Euclidian position of the robot arm grip after the robot arm is moved.

The relationship between Euclidian values of the actual 3D position of the red chili fruit and the Euclidian value of the 3D position captured by the stereo vision is then plotted as shown in Figure 13. The relationship between Euclidian values of the actual 3D position of the red chili fruit, the 3D position captured by stereo vision which has been expressed in Cartesian coordinates of the robot arm and the center position of the robot arm grip is plotted sequentially as shown in Figures 14a and 14b.

From Figure 13, it can be seen that the Euclidian value of the 3D position captured by stereo vision is not much different compared to the Euclidian value of the actual 3D position. The average Euclidian error for 1 red chili, 3 red chilies, and 5 red chilies respectively is 19.4 mm, 20.9 mm, and 17 mm or equivalent to 6.6%, 6.4%, and 5.4% respectively. The 3D positioning accuracy of the stereo vision for 3 combinations of the red chili fruit number on image i.e: 1 red chili fruit, 3 red chili fruits, and 5 red chili fruits respectively is 93.4%, 93.6%, and 94.6%. This detection accuracy is the same as that obtained by [Hadinegoro & Rizaldhi, \(2021\)](#) but slightly lower compared to the research results conducted by [Indrabayu et al., \(2019\)](#) which is probably because of the red chili fruit is curved (not straight or round). In Figure 14a, it can be seen that the data values are more spread out compared to Figure 14b because the actual motion of the robot arm grip is calculated based on the 3D position value captured by stereo vision, not based on the actual 3D position of the red chili fruit. The error value shown in Figure 14a is a combination of stereo vision error and robot arm grip positioning error.

On average, the Euclidian error between the 3D position of the robot arm grip and the actual 3D position of the red chili fruit for the three treatments, namely 1 red chili fruit, 3 red chili fruits, 5 chili fruits on the image respectively is 15.4 mm, 24.3 mm, 20.5 mm or equivalent to 3.4 %, 5.2 % and 4.5 % respectively. The positioning error obtained in this study is the same as that obtained by [Sangeetha et al., \(2018\)](#) namely 20 mm. The accuration value between the

3D position of the robot arm grip and the actual 3D position of the red chili fruit for 1 fruit, 3 fruits and 5 fruits on each image respectively is 96.6%, 94.8%, 95.5%. In this research, the inverse kinematic value is calculated based on the 3D position obtained from the stereo vision system, not based on the actual 3D position of the fruit. The average Euclidian error between the 3D position of the robot arm grip and the 3D position of the red chili fruit captured by the stereo vision for three treatments, namely 1 chili, 3 chilies, 5 chilies on each image is 7.6 mm, 6.3 mm, 8.2 mm or equivalent to 1.6 %, 1.3 % and 1.7 % respectively. The accuration value between the 3D position of the robot arm grip and the 3D position of red chili fruit captured by the stereo vision for 1 chili, 3 chilies and 5 chilies on each image respectively is 98.4%, 98.7%, 98.3%.

### 3.4. Stability Measurements of Stereo Vision and Robot Arm Motion

The 3D positioning stability of the stereo vision system used in this research is very important to ensure the accuracy of the data produced by the stereo vision system so that the grip of the robot arm can be moved to the accurate coordinates. The measurement stability of the stereo vision system and the positioning stability of the robot arm grip are presented in Figures 15a and 15b. Figure 15a shows that the measurement stability of the stereo vision system is quite good. In this case, only 20 of the red chili fruits were analysis each with three repetitions. The average Euclidian value of the 3D position measuring error of the stereo vision system is 0.5% or equivalent to 99.5 % stability, while the Euclidian value of the 3D positioning error of the robot arm grip is 0.4% or equivalent to 99.6 % stability (Figure 15b).

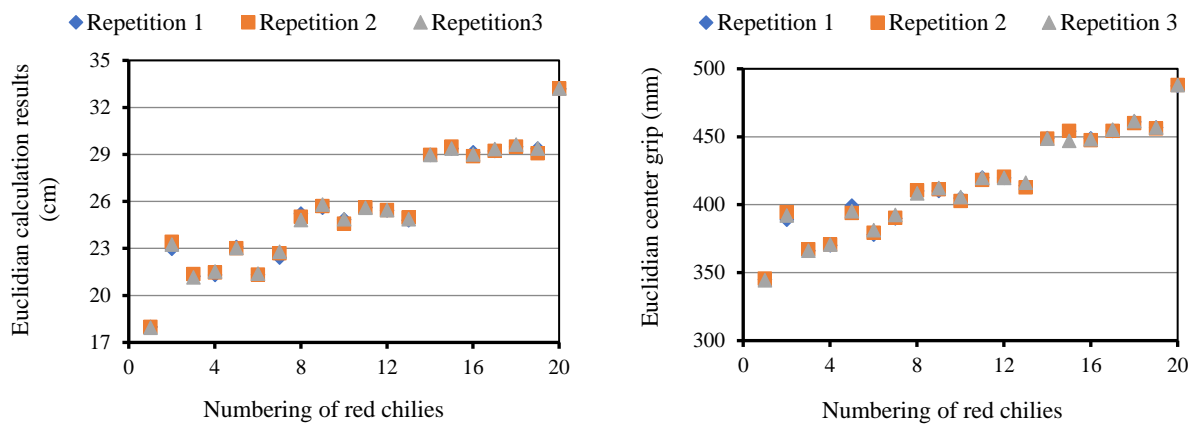


Figure 15. Stability measurement of the stereo vision system (left), and positioning stability of the robot arm grip (right).

### 3.5. Response Time

It is important to know the response time of the stereo vision system and the response time of the robot arm movement to estimate the duration time required to harvest one red chili fruit. In this research, the average scanning time required to obtain the 3D position of a red chili fruit using a stereo vision system for images containing 1 chili, 3 chilies and 5 chilies in each image with 3 repetitions respectively is 0.050 s, 0.052 s, 0.056 s. The duration time required to process images captured by the stereo vision system is faster than the research results conducted by Onishi *et al.*, (2019) which takes 2 seconds. The average duration time required to move the robot arm grip towards the position of the red chili fruit which is in the Euclidian range from 300 mm to 600 mm is not too different, namely around 9 seconds or roughly equivalent to the harvesting time of 18 seconds for each chili fruit using level 7 joint movement speed. The estimated harvesting time in this study is slightly faster than the research results conducted by Yoshida *et al.*, (2022) for one arm which takes 20 s, but slightly slower than the research results conducted by Onishi *et al.*, (2019) which requires a harvesting time of 16 seconds.

#### 4. CONCLUSION

A stereo vision system designed using two Logitech type C270 digital cameras has been installed on the left side of the waist (joint 2) of the Mitsubishi Movemaster RV-M1 robot arm for capturing and calculating the 3D position of red chili fruits. The 3D position value of the red chili fruit captured using stereo vision has been successfully used as an input value to control the positioning movement of the robot arm grip towards the center point of the red chili fruit. The average of the 3D positioning accuracy of red chili fruit images captured using a stereo vision system for 3 treatments is 93.4 %, 93.6 %, and 94.6 % for 1 fruit, 3 fruits, and 5 fruits on each image respectively. The accuration value between the 3D position of the robot arm grip and the actual 3D position of the red chili fruit for 1 chili, 3 chilies and 5 chilies on each image respectively is 96.6%, 94.8%, 95.5%. The accuration value between the 3D position of the robot arm grip and the 3D position of the red chili fruit captured by the stereo vision system for 1 chili, 3 chilies and 5 chilies on each image respectively is 98.4%, 98.7%, 98.3%. The average Euclidian value of the 3D position measuring error of the stereo vision system is 0.5% or equivalent to 99.5 % stability, while the eucledian value of the 3D positioning error of the robot arm grip is 0.4% or equivalent to 99.6 % stability. In this stability analysis, only 20 of the red chili fruits were used with three repetitions. The duration time required for stereo vision to determine the 3D position of the red chili fruits in each image is 0.053s, while the duration time required for the robot arm grip to move towards the center of the red chili fruit is 9s.

#### ACKNOWLEDGMENTS

The authors would like to thank the Department of Mechanical and Biosystem Engineering, Faculty of Agricultural Technology, IPB University, which has facilitated this research to be carried out properly.

#### REFERENCES

- Angulo, A.M., Pinto, L.P., Espinoza, E.S., Vidal, Y.S., & Colquehuanca, E.S. (2022). Assisted operation of a robotic arm based on stereo vision for positioning near an explosive device. *Robotics*, *11*(5), 1–18. <https://doi.org/10.3390/robotics11050100>
- Barbashov, N.N., Shanygin, S.V., & Barkova, A.A. (2022). Agricultural robots for fruit harvesting in horticulture application. *IOP Conference Series: Earth and Environmental Science*, *981*(2022), 032009. <https://doi.org/10.1088/1755-1315/981/3/032009>
- Delya, B., Tusi, A., Lanya, B., & Zulkarnain, I. (2014). Rancang bangun sistem hidroponik pasang surut otomatis untuk budidaya tanaman cabai. *Jurnal Teknik Pertanian Lampung*, *3*(3), 205–212.
- Ekowahyuni, L.P., & Wulansari, L.I. (2022). Yield test of hybrid harapan chili (*Capsicum annum* L.). *Jurnal Teknik Pertanian Lampung*, *11*(1), 35–47. <https://doi.org/10.23960/jtep-l.v11i1.35-47>
- Erceg, D., Knezevvc, B.Z., & Grvhovve, Z. (2023). Stereo vision based robot welding. *Journal of Physics: Conference Series*, *2540*(2023), 012017. <https://doi.org/10.1088/1742-6596/2540/1/012017>
- Feng, Q., Zou, W., Fan, P., Zhang, C., & Wang, X. (2018). Design and test of robotic harvesting system for cherry tomato. *International Journal of Agricultural and Biological Engineering*, *11*(1), 96–100. <https://doi.org/10.25165/j.ijabe.20181101.2853>
- Hadinegoro, A., & Rizaldhi, D.A. (2021). Pengaruh HSV pada pengolahan citra untuk kematangan buah cabai. *Building of Informatics, Technology and Science (BITS)*, *3*(3), 155–163. <https://doi.org/10.47065/bits.v3i3.1020>
- Hua, Y., Zhang, N., Yuan, X., Quan, L., Yang, J., Nagasaka, K., & Zhou, X. G. (2019). Recent advances in intelligent automated fruit harvesting robots. *The Open Agriculture Journal*, *13*(1), 101–106. <https://doi.org/10.2174/1874331501913010101>
- Ibadillah, A.F. (2018). Sistem penjejakan obyek dengan stero vision. *Triac: Jurnal Teknik Elektro dan Komputer*, *5*(1), 1–10. <https://doi.org/10.21107/triac.v5i1.3830>
- Indrabayu, Arifin, N., & Areni, I.S. (2019). Klasifikasi kematangan stroberi berbasis segmentasi warna dengan metode HSV. *Jurnal Penelitian Enjiniring*, *23*(2), 113–116. <https://doi.org/10.25042/jpe.112019.03>
- Jia, W., Zhang, Y., Lian, J., Zheng, Y., Zhao, D., & Li, C. (2020). Apple harvesting robot under information technology: A review. *International Journal of Advanced Robotic Systems*, *17*(3), 1–16. <https://doi.org/10.1177/1729881420925310>

- Kaleem, A., Hussain, S., Aqib, M., Cheema, M.J.M., Saleem, S.R., & Farooq, U. (2023). Development challenges of fruit-harvesting robotic arms: A critical review. *AgriEngineering*, 5(4), 2216–2237. <https://doi.org/10.3390/agriengineering5040136>
- Moekasan, T.K., Gunadi, N., Adiyoga, W., & Sulastrini, I. (2015). Kelayakan teknis dan ekonomi budidaya cabai merah di dalam rumah kaca untuk menanggulangi serangan organisme pengganggu tumbuhan. *J. Hort.*, 25(2), 180–192. <https://doi.org/10.21082/jhort.v25n2.2015.p180-192>
- Nurhafsa, N., Rahmi, H., Andriani, I., & Fitriawaty. (2021). Analisis usahatani cabai di luar musim berdasarkan penerapan komponen budidaya cabai merah di provinsi Sulawesi Barat. *Teknotan*, 15(1), 9–14. <https://doi.org/10.24198/jt.vol15n1.2>
- Onishi, Y., Yoshida, T., Kurita, H., Fukao, T., Arihara, H., & Iwai, A. (2019). An automated fruit harvesting robot by using deep learning. *Robomech Journal*, 6(13), 1–8. <https://doi.org/10.1186/s40648-019-0141-2>
- Raja, V., Bhaskaran, B., Nagaraj, K.K.G., Sampathkumar, J.G., & Senthilkumar, S.R. (2022). Agricultural harvesting using integrated robot system. *Indonesian Journal of Electrical Engineering and Computer Science*, 25(1), 152–158. <https://doi.org/10.11591/ijeecs.v25.i1.pp152-158>
- Sangeetha, G.R., Kumar, N., Hari, P.R., & Sasikumar, S. (2018). Implementation of a stereo vision based system for visual feedback control of robotic arm for space manipulations. *Procedia Computer Science*, 133(2018), 1066–1073. <https://doi.org/10.1016/j.procs.2018.07.031>
- Subrata, I.D.M., & Melania, J. (2023). Design of cartesian type manipulator for automatically capturing plant images inside greenhouse. *Jurnal Teknik Pertanian Lampung*, 12(3), 545–558. <https://doi.org/10.23960/jtep-l.v12i3.545-558>
- Susilowati, S.H. (2016). Fenomena penuaan petani dan berkurangnya tenaga kerja muda serta implikasinya bagi kebijakan pembangunan pertanian. *Forum Penelitian Agro Ekonomi*, 34(1), 35–55.
- Tang, Y., Chen, M., Wang, C., Luo, L., Li, J., Lian, G., & Zou, X. (2020). Recognition and localization methods for vision-based fruit picking robots: A review. *Frontiers in Plant Science*, 11(510), 1–17. <https://doi.org/10.3389/fpls.2020.00510>
- Taryudi. (2018). Rancang bangun vision sensor menggunakan stereo kamera pada sistem manipulasi objek. *Autocracy: Jurnal Otomasi, Kendali Dan Aplikasi Industri*, 5(1), 1–7.
- Umam, F., Fuad, M., Suwarno, I., Ma'arif, A., & Caesarendra, W. (2023). Obstacle avoidance based on stereo vision navigation system for omni-directional robot. *Journal of Robotics and Control (JRC)*, 4(2), 227–242. <https://doi.org/10.18196/jrc.v4i2.17977>
- Wang, Z., Xun, Y., Wang, Y., & Yang, Q. (2022). Review of smart robots for fruit and vegetable picking in agriculture. *International Journal of Agricultural and Biological Engineering*, 15(1), 33–54. <https://doi.org/10.25165/ijabe.20221501.7232>
- Yoshida, T., Onishi, Y., Kawahara, T., & Fukao, T. (2022). Automated harvesting by a dual-arm fruit harvesting robot. *Robomech Journal*, 9(19), 1–14. <https://doi.org/10.1186/s40648-022-00233-9>