Design of Human Detection in Home Security

Wahyudi[1], Kevin Inzaghi Zezha[2], Munawar Agus Riyadi[3], Imam Santoso[4]

[1,2,3,4] *Department of Electrical Engineering, Diponegoro University, Central Java, Indonesia*

ABSTRACT

A home security system is needed to keep objects in the house safe from human interference. This system requires early information on human presence that might endanger security, so the human detection system is one of the most important parts of a home security system. Many people install Closed-Circuit Television (CCTV) at home to maintain home security, so that events that occur around the house can be known. Most of the CCTV installed at home functions as a monitoring and recording of events that occur around the house. With current technological developments, an object such as a human can be recognized and identified through a human object recognition program. By combining a human recognition system with CCTV, as well as a smartphone communication device, a home security system will be produced that can provide information to the smartphone if humans are detected around the house. A human recognition program that is embedded in mini-computer devices such as the Raspberry Pi 4 and gets input in the form of video results from CCTV, then the recorded events around the house can be displayed on a smartphone. In this paper, before the system can recognize human objects, the system first performs the process of detecting human objects using the Single-shot Multibox Detector (SSD) method. The use of the SSD method in detecting human objects has an accuracy rate of more than 80%. By integrating this home security system into a smartphone, users can carry out monitoring functions online and in person and will receive alerts in the form of notifications from applications when video recordings of CCTV detect a human object.

Keyword: - *CCTV, SSD, Smartphone, Raspberry Pi, and Human detection.*

1. INTRODUCTION

The human object detection system is one solution that can be developed to improve the safety and security of the occupants of the house. A camera-based home security system or commonly known as Closed-Circuit Television (CCTV) is a computer system that uses a video camera to display and record an image at the time and place where the device is installed [1]. Along with the development of the times, CCTV has a function not only to carry out surveillance but also to recognize and detect human objects. CCTV must be able to recognize and distinguish between human objects and non-human objects. One of the libraries that performs the human object recognition function is Open Computer Vision (OpenCV) which is written in the Python programming language. OpenCV makes it possible to see and identify like a human. The identification process run by OpenCV can be used to determine between human objects and non-human objects.

Detection of objects is done using the Single-Shot Multibox Detector (SSD) method [2]. The selection of the SSD method as a detection method is due to the high level of accuracy and speed of object detection and can run well on the specifications of the Raspberry Pi so that it does not burden the performance of the Raspberry Pi itself [3]. The Convolution Neural Network (CNN) method is often used in detecting human objects [4], but in its application this method takes a long time to detect, and the level of object detection accuracy produced is almost the same as using the SSD method. The use of the Raspberry Pi as the main engine is inseparable from the specifications and features it has. The high enough system specifications allow the Raspberry Pi to run the SSD detection method, as well as to have hardware equipped with a good cooling system that allows it to work for quite a long time [5].

2. METHOD

The design of the human object detection system is divided into 2 types of designs, namely hardware design and software design. The overall workflow of the system is shown in Fig -1. The system workflow starts with the webcam capturing images and videos, then the results of video and image data from the webcam will be sent to the Raspberry Pi for detection of human objects [6], after the system detects a human object the results of the image capture will be sent by the Raspberry Pi to cloud database via the internet to be stored and displayed on a smartphone, for video data generated by the webcam will be directly displayed on the smartphone application if a request is made to display video via the livestream function that is connected directly via the internet [7].

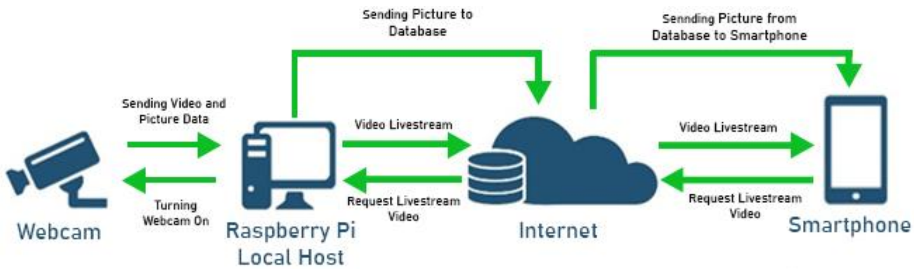

Fig -1: Smart CCTV system block diagram

2.1 Hardware Design

The hardware design in this paper includes the connection between the camera and the Raspberry Pi as shown in Fig -2.

Fig -2: Connection between CCTV and Raspberry

CCTV cameras are equipped with a USB 2.0 connector to communicate with the Raspberry Pi via the USB port on the Raspberry Pi [8]. Besides functioning as a communication medium, USB is also used to power the camera function. The camera turns on by getting 5V voltage directly from the Raspberry via USB.

2.2 Software Design

The detection of human objects in each video frame generated by the CCTV camera will be processed using the SSD method on the Raspberry Pi. Fig -3 is a flow diagram of the human object detection process. The first step in determining a human object is to call the training data file which is written in the program code as follows[9].ap.add_argument("-p","--prototxt",default="MobileNetSSD_deploy.prototxt.txt");ap.add_argument("-m","--model",default="MobileNetSSD_deploy.caffemodel") where the training data file used is a file type in the form of .prototxt.txt and .caffemodel. After the program code calls the training data file, the next step is to adjust the sensitivity level of the system in detecting objects written in the program code. ap.add_argument("-c", "--confidence", type=float, default=0.5), in this paper the value of the sensitivity of the system in detecting human objects is set at 50%. This 50% value means the system will detect the object and classify it as a human object when the system is 50% sure that the detected object is a human object. In designing a home security system, the use of a sensitivity percentage with a value smaller than 100% certainly makes the security system better, because if the sensitivity value used is large, the system takes a long time to detect and classify objects, but if the sensitivity value is large, the system takes a long time to detect and classify objects. If the system is too small, the system will be too fast in classifying objects so that non-human objects will also be detected as humans if the system is sure that the characteristics of the object are like humans [9].

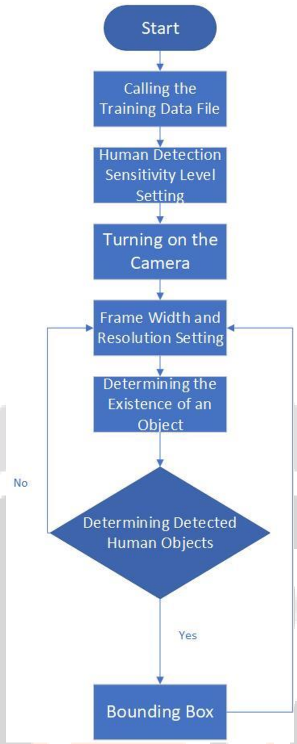

Fig -3: Human object detection block diagram

The next process is to calculate the matrix of an image to determine the presence of human objects in each frame that has been captured by the CCTV camera. Matrix calculations in determining human objects are carried out in a row matrix [10]. After the system detects an object, the next step is to determine the class of the detected object. To determine the class on the object, the system will do a comparison with the training data library that was called before the program was run. In this paper, only one type of class is used, namely "person". So that other objects registered in the class that have been trained will not appear because only objects that are considered human will be detected [11].

2.3 Image Capture Data Processing on Raspberry Pi

Image capture data processing performed using the Raspberry Pi is shown in Fig -4.

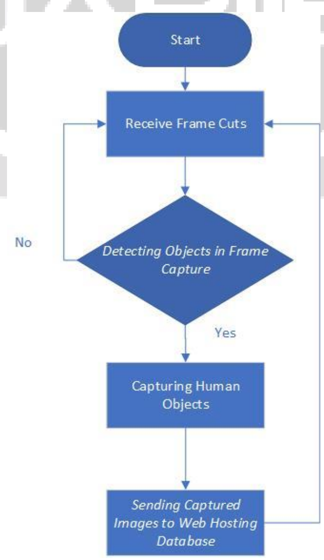

Fig -4: Image data processing block diagram

CCTV camera output in the form of image captures or frames will be processed on the Raspberry Pi. This data processing includes the process of detecting and identifying human objects on each piece of frames and images received from CCTV cameras [12]. The results of this data processing will produce image capture data for detecting human objects. If the image processing detects the presence of a human object, the captured image will be sent directly to the database to be stored and displayed on the smartphone application.

2.4 Video Data Processing on Raspberry Pi

Video data processing is carried out on the Raspberry Pi as shown in Fig -5.

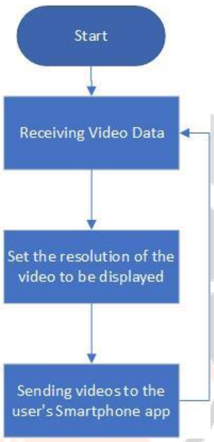

Fig -5: Video data processing block diagram

Processing video data through the process of capturing video data via CCTV, CCTV captures will be sent to the Raspberry Pi to be sent back to the user's smartphone. Before being sent to the Raspberry Pi smartphone, you must first set the video resolution, then send it directly to the user's smartphone application.

3. RESULT AND ANALYSIS

The results and analysis of this paper are testing of hardware connection, human detection, delivering image to database, and livestreaming on smartphone.

3.1 Hardware Connection Testing

Hardware connection testing is carried out to check the connection between hardware devices in the form of a camera and a Raspberry Pi which is connected by a USB 2.0 cable through the Raspberry Pi's USB port. The parameters of the success of this test are indicated by the sending of image capture data and livestream shows on the smartphone. The results of the connection test are shown in Fig -6, which is the capture of a smartphone image and Fig -7, which is the result of a livestream on a smartphone.

Fig -6: Captured image on a smartphone application

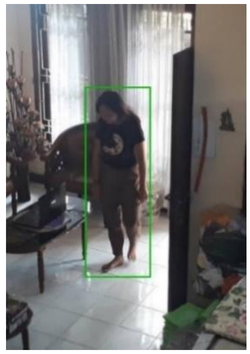

Fig -7: Livestream results on the smartphone application.

The connection between the camera and the Raspberry Pi has worked, this is evidenced by the captured image data and livestream shows that can be displayed on the smartphone application.

3.2 Human Detection Test

Human detection testing is carried out with the aim of knowing the level of accuracy of human detection according to predetermined parameters. Tests carried out in the form of variations in human movement were detected with the first 50 frames as a test sample.

The test uses as many as 200 different sample frames with the number of variations of human objects detected from 1 to 2 people with variations of 100 different frames for each variation in the number of humans. In this system, if a human is detected, the system will display a bounding box on the frame. In this testing process, a confusion matrix is used to see the level of accuracy of the detection system.

In testing the variation of 1 person, the system detects human objects with the detected object being worth 1 human object and the system succeeds in producing a bounding box around the object that shows the detected human object. The results of this test can be categorized in the confusion matrix as True Positive because the system is able to detect human objects and the number of people in the frame and display a bounding box according to the number of people in the frame shown in Fig -8.

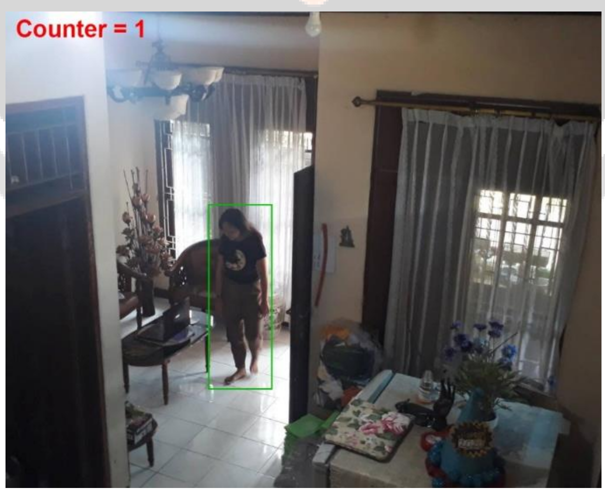

Fig -8: The results of human detection with a variation of 1 person with a True Positive value

In the variation of 2 people, the system succeeded in detecting two human objects in the same 1 frame, and succeeded in displaying a bounding box around the human object with an amount that corresponds to the number of humans in the frame. This test is categorized as True Positive with a variation of two people as shown in Fig -9. In one of the tests sometimes the wrong detection results are obtained. The system succeeded in detecting human objects and succeeded in producing a bounding box around the human object, but there was an error in calculating the number of humans in the frame where the human objects in the frame were two people, but the system detected three human objects in the same 1 frame.

Fig -9: The results of human detection of variations of two people True Positive

This is in the confusion matrix categorized as False Positive where the system successfully detects and displays a bounding box but does not match the number of human objects or detects objects that are not human objects as shown in Fig -10.

Fig. 10. The results of human detection of variations in two people are False Positive.

Testing is also done by not displaying any human objects in the frame. This aims to determine the level of accuracy of system detection if there are no human objects in 1 frame. In this test, the system does not detect human objects in the frame and there are also no human objects in the frame. Because of this condition, the confusion matrix can be categorized as True Negative where the system does not detect any human objects in the frame and the system will not send notifications and image data on smartphone devices as shown in Fig -11. The test is divided into 4 sampling sessions, each session sampling 50 frames on the test video. In the first session of testing with a variation of the number of humans 1 person by taking the number of frames as much as 50 frames, the accuracy value is 86%. Then in the second session of testing carried out using the same number of people with different types of frames, the accuracy value is 78%. Meanwhile, the third and fourth sessions of testing were carried out by placing two people in the same 1 frame with each testing session using 50 different types of frames. The results of the tests carried out in the third session obtained an accuracy value of 80% and in the fourth session testing using 50 different frames with the third session testing an accuracy value of 90% was obtained. From the total test data using 200 different frames, the average accuracy value of the whole system is 83.5% as shown in Table 1.

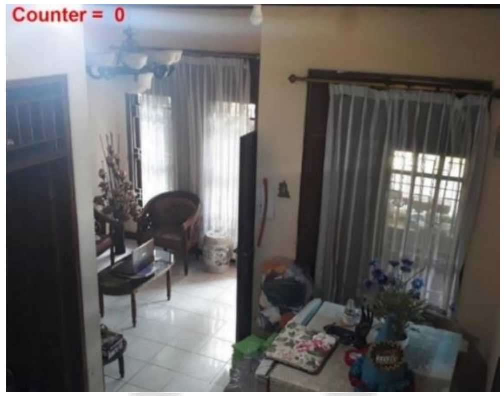

Fig -11: Detection results without human objects that have a True Negative value

Table -1: Human detection test result

Number of Test	Number of Human	Number of Frame	*True Positive*	*True Negative*	*False Positive*	*False Negative*	Accuracy
1	1	50	41	2	7	0	86 %
2	1	50	37	2	11	0	78 %
3	2	50	38	2	10	0	80 %
4	2	50	43	2	5	0	90 %

The confusion matrix method can not only be used to find the level of accuracy of a system but can also be used to find the value of accuracy (precision) and success (recall) of this system. The level of accuracy between the information requested by the user and the results given by the system is 82.81% and the success rate of the system in retrieving information is worth 100% which is shown in Fig -12.

Actual \ Prediction	False	True	Total
False	8	33	41
True	0	159	159
	8	192	200

True Positive = 159
True Negative = 8
False Positive = 33
False Negative = 0

Fig -12: Confusion matrix human object detection system

3.3 Delivering Image to Database Test

Testing of sending image captures to the database is done by viewing the captured images directly from the database as shown in Fig -13 and checking the image capture directly on the smart home application gallery as shown in Fig -14.

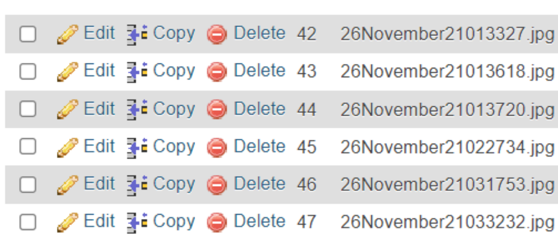

Fig -13: Human detection image capture data in the database

Fig -14: Display of image capture data on the smartphone application

The data displayed in the database is the captured image of the detected object and the captured image is displayed on the smartphone application. Through testing the delivery of image capture in the database, the system can send image data to the database properly.

3.4 Livestreaming on Smartphone Application Test

Testing of livestream shows is done by looking at the results of livestream shows directly on the smartphone application as shown in Fig -15.

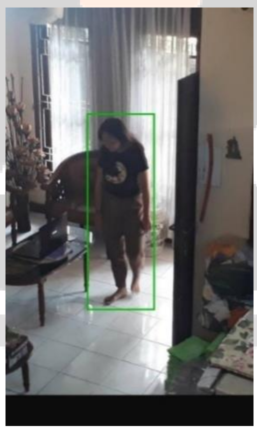

Fig -15: Display livestream on the smartphone application

Through testing livestream shows on smartphone applications, livestream shows can be accessed and functioned on smartphone applications according to the state of the object of human detection.

4. CONCLUSIONS

The proposed human object detection system based on a smartphone shows good results in detecting and alerting users when there are humans caught by CCTV, also showing monitoring data and images of human detection captures. In future research, improvement may be needed if there is no prevention system in place. Other more accurate detection methods can also be applied so that the accuracy of human detection is better.

5. ACKNOWLEDGEMENT

This research was financially supported by The Faculty of Engineering, Diponegoro University, Indonesia through Strategic Research Grand.

6. REFERENCES

[1] B. Welsh, E. Piza, A. Thomas and D. Farrington, "Private Security and ClosedCircuit Television (CCTV) Surveillance: A Systematic Review of Function and Performance", Journal of Contemporary Criminal Justice, vol. 36, no. 1, pp. 56-59, 2020. doi: pdf/10.1177/1043986219890192.

[2] A. Kumar, Z. Zhang and H. Lyu, " Object detection in real time based on improved single shot multi-box detector algorithm", EURASIP Journal on Wireless Communications and Networking, (2020), 2020:204, Doi: 10.1186/s13638-020-01826-x.

[3] Sanjana, Prasad, and Mahalakshmi, "Smart Surveillance Monitoring System Using Raspberry PI and PIR Sensor", International Journal of Computer Science and Information Technologies (IJCSIT), Vol. 5 (6) , 7107-7109, 2014.

[4] J. Angelico and K. Wardani, "Convolutional Neural Network Using Kalman Filter for Human Detection and Tracking on RGB-D Video", Communication & Information Technology Journal (CommIT), pp. 105-110, 2018.

[5] W. Liu, D. Anguelov, D. Erhan and C. Szegedy, " SSD: Single Shot MultiBox Detector," arXiv:1512.02325, 2016, doi: 10.1007/978-3-319-46448-0_2.

[6] Ferm. Oliwer, " Real-time Object Detection on Raspberry Pi 4", Bachelor Degree Project, Civilingenjör Elektroteknik, Mid Sweden University, 2020.

[7] F. Yudianto, F. Susanto, " Design Of Smartschool Webcam Security System Using Raspberry Pi With Motion Detection In Smk Pgri Sukodadi", International Journal of ASRO, vol. 11, no. 20, pp. 202-206, 2020. Doi: ttps://doi.org/10.37875/asro.v11i1.218.

[8] Raspberry Pi Foundation, UK, "Raspberry Pi 4 B", Available: https://www.raspberrypi.org/. Accessed: Jul. 30, 2021.

[9] M. Paul, S. Haque and S. Chakraborty, " Human detection in surveillance videos and its applications", EURASIP Journal on Advances in Signal Processing, 2013:176 ,2013.

[10] A. Trofimova, A. Masciadri, F. Veronese and F. Salice " Indoor Human Detection Based on Thermal Array Sensor Data and Adaptive Background Estimation", Journal of Computer and Communications, Vol. 5, pp. 16-28, 2017.

[11] B. Genç and H. Tunc " Optimal training and test sets design for machine learning", Turkish Journal of Electrical Engineering and Computer Sciences, 27(2), pp. 1534-1545, 2019. Doi: 10.3906/elk-1807-212.

[12] S. Jia, C. Diao, G. Zhang, A. Dun, Y. Sun, X. Li and X. Zhang " Object Detection Based on the Improved Single Shot MultiBox Detector", IOP Conf. Series: Journal of Physics: Conf. Series 1187, 2019. Doi: 10.1088/1742-6596/1187/4/042041.