International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

Distance Measurement of Human and Vehicle Objects Using Computer Vision Based on the Combined Method of Triangulation and Pixel Comparison for Autonomous Braking System

Adiratna Ciptaningrum[1], Andhika Putra Widyadharma[2], R. Akbar Nur Apriyanto[3], Imam Junaedi[4], Wahyu Pribadi[5], Edo Zulmi Faikhsan[6], Larissa Kartika Putri[7]
[1, 2, 3, 4, 5, 6]Engineering Department, Madiun State Polytechnic, Madiun, East Java, Indonesia +62

***Abstract*— *The rapid development of artificial intelligence (AI) is increasingly being leveraged in the transportation industry, with a particular focus on railway systems. This research delves into the application of computer vision technologies to detect human and vehicle objects and to estimate their distances from the camera, utilizing human pose estimation and object tracking techniques with triangulation as the core methodology. The system's human object detection achieved a remarkable accuracy rate of 96.8102%, coupled with an optimal detection performance score of 0.9338 at a detection distance of 15 meters. Nevertheless, a minor deviation in distance estimation was noted, with a variance of 0.12738 meters at a 10-meter detection range. Similarly, vehicle object detection exhibited a high confidence score of 0.95, though it also displayed a distance estimation deviation of 0.14223 meters at the same detection distance. These findings underscore the system's capability to perform accurate and reliable object detection for both humans and vehicles, positioning it as a promising candidate for further refinement. The potential for integration into the autonomous braking system of the inspection train at Madiun State Polytechnic is particularly notable, as it could enhance the safety and efficiency of railway operations through precise object detection and distance estimation.*

***Keywords*—*Computer Vision, Human Pose Estimation, Object Tracking, Triangulation, Distance Measurement*

I. INTRODUCTION

Artificial intelligence (AI) technology has become a major driver of change in various aspects of life. AI was developed to make life easier for humans and has brought significant advances in social, business, economic, health, and other fields [1]. AI technology is widely applied in robotics and is now used in automatic control systems in transportation.

One mode of transportation that can be developed using artificial intelligence (AI) technology is trains. A train is a means of railroad that moves on rails, either independently or coupled with other facilities [2]. Based on its function, there are several types of railroad facilities, one of which is the Inspection Train. Inspection Train is a transport vehicle, with or without its own drive, which is used to inspect infrastructure (railroad), carry officers, and load work materials [3].

The development of railway transportation modes using artificial intelligence, one of which is applied to the braking system, namely with a function to measure the distance of objects to the train so that the distance obtained will be forwarded to the braking control system on the inspection train so that an autonomous braking system can be made on the inspection train.

II. RESEARCH METHODOLOGY

This research uses the Research and Development (RnD) method, developing previous research by adding types of objects that can be detected and changing the previous system.

A. Research Stages

The research was carried out with several stages of research so as to produce results that are in accordance with what is expected.

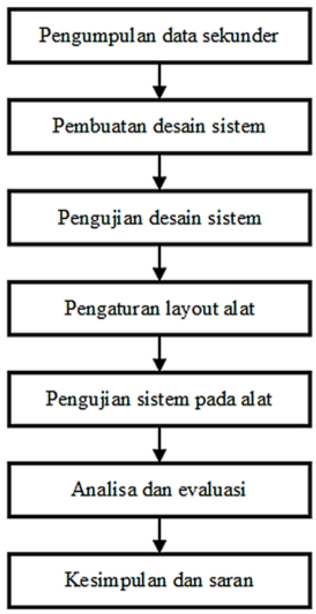

Fig. 1. Research stages

B. Hardware Design

The tool design is prepared to illustrate the appearance and arrangement in installation and application in real situations. The main purpose of making a tool design is to provide guidelines to the author during the manufacturing process so that the tool can be used in accordance with its predetermined functions. The following is the tool design used in the research:

Adiratna Ciptaningrum, Andhika Putra Widyadharma, R. Akbar Nur Apriyanto, Imam Junaedi, Wahyu Pribadi, Edo Zulmi Faikhsan, and Larissa Kartika Putri, "Distance Measurement of Human and Vehicle Objects Using Computer Vision Based on The Combined Method of Triangulation and Pixel Comparison for Autonomous Braking System," *International Research Journal of Advanced Engineering and Science*, Volume 9, Issue 4, pp. 32-36, 2024.

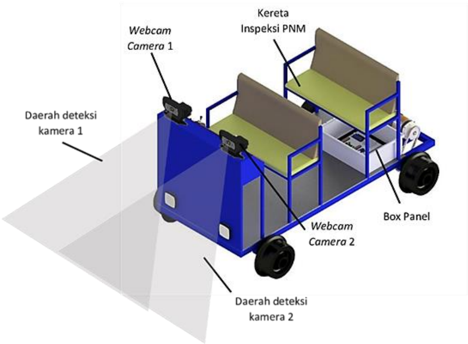

Fig. 2. Hardware Design

Description:

(1) The webcam cameras used are 2 cameras mounted on the dashboard cabin so that the camera can be assumed to be the front end of the inspection train with a distance of 1 meter between cameras.

(2) The hardware components are placed on the Box panel which is placed in a safe position and protected from external interference.

(3) The motor and braking system are adopted from the original design of the PNM inspection cart.

C. *Design System*

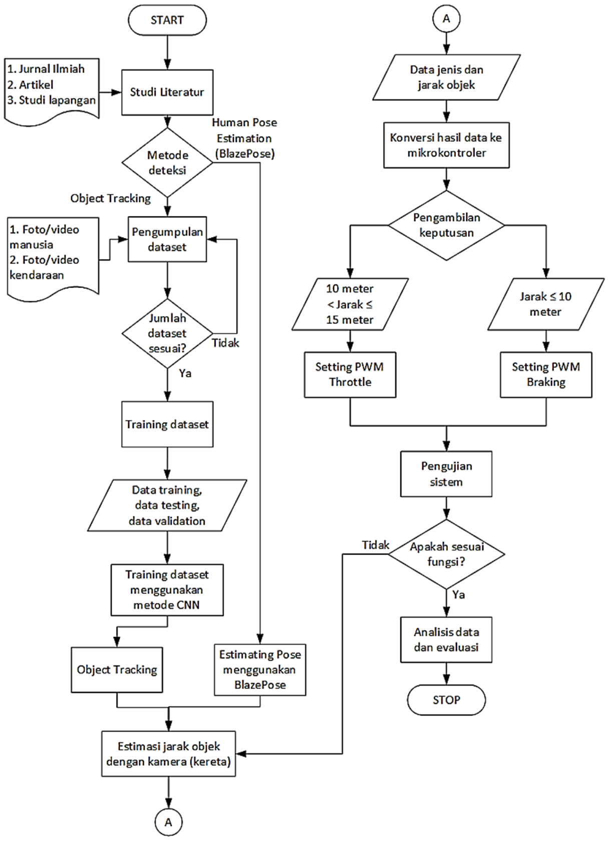

Fig. 3. Flowchart

This research focuses on image processing of human objects and vehicle objects with detection methods using Human Pose Estimation and Object Tracking which then measure the distance of objects with cameras to produce distance data that will be forwarded to the Madiun State Polytechnic inspection train braking control system.

D. *Human Pose Estimation*

The Human Pose Estimation architecture has landmarks totaling 33 keypoints which define key points as well as joint points on the human body [4]. These 33 keypoints cover from head to toe of a balanced human [5]. The following is the architecture of landmarks owned by Human Pose Estimation:

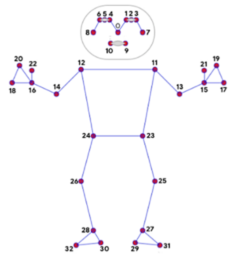

Fig. 4. Human Pose Estimation Architecture

From Fig. 4, 33 keypoints were obtained with the explanation of each keypoint name as follows:

0	Nose	17	Left_pinky
1	Left_eye_inner	18	Right_pinky
2	Left_eye	19	Left_index
3	Left_eye_outer	20	Right_index
4	Right_eye_inner	21	Left_thumb
5	Right_eye	22	Right_thumb
6	Right_eye_outer	23	Left_hip
7	Left_ear	24	Right_hip
8	Right_ear	25	Left_knee
9	Mouth_left	26	Right_knee
10	Mouth_right	27	Left_ankle
11	Left_shoulder	28	Right_ankle
12	Right_shoulder	29	Left_heel
13	Left_elbow	30	Right_heel
14	Right_elbow	31	Left_foot_index
15	Left_wrist	32	Right_foot_indec
16	Right_wrist		

E. *Object Tracking*

Object Tracking is a technique for detecting and tracking a moving object in an image or video with the aim of monitoring the movement of the object from frame to frame [6]. This technique involves continuous identification and tracking of objects, enabling real-time system operation. Some commonly used methods for Object Tracking include Simple Online and Realtime Tracking (SORT), DeepSORT, FairMOT, TransMOT, ByteTrack, and You Only Looks Once (YOLO).

In terms of performance, ByteTrack is an effective method in making Object Tracking programs. ByteTrack is a development of the previous Object Tracking method by optimizing it using Tracklet Interpolation. This method has better performance because it implements an optimized equation with the following algorithm:

$$B_t = B_{t1} + (B_{t2} - B_{t1})\frac{t1-t2}{t2-t1}$$

Adiratna Ciptaningrum, Andhika Putra Widyadharma, R. Akbar Nur Apriyanto, Imam Junaedi, Wahyu Pribadi, Edo Zulmi Faikhsan, and Larissa Kartika Putri, "Distance Measurement of Human and Vehicle Objects Using Computer Vision Based on The Combined Method of Triangulation and Pixel Comparison for Autonomous Braking System," *International Research Journal of Advanced Engineering and Science*, Volume 9, Issue 4, pp. 32-36, 2024.

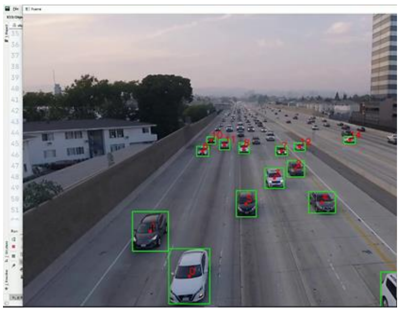

Fig. 5. Object Tracking for Vehicle

F. *Triangulation*

In terms of performance, ByteTrack is an effective method in making Object Tracking programs. ByteTrack is a development of the previous Object Tracking method by optimizing it using Tracklet Interpolation. This method has better performance because it implements an optimized equation with the following algorithm:

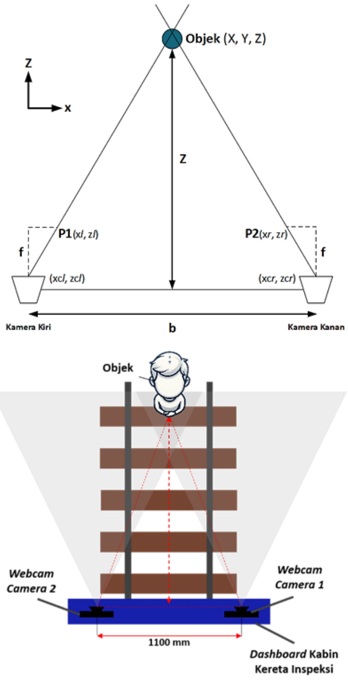

Fig. 6. Triangulation Concept of Camera Mounting

The variables to be used are as follows:

Z	=	Variabel jarak objek dengan kamerayang akan menjadi pembahasan dalam perhitungan
b	=	Jarak *mounting* antara 2 kamera yang akan dipasangkan pada *dashboard* kereta inapeksi
f	=	Titik fokus kamera yang telah dikoneevrsi dari mm menjadi piksel
xl	=	Jarak titik tengah ke arah kanan objek
xr	=	Jarak titik tengah ke arah kiri objek

In the calculation process using the triangulation concept, the required information is the xl and xr values, which are measured from the edge of the image to the object point. The f value from the camera specifications and the b value from the manual measurement of the camera installation on the dashboard of the inspection train are known. The calculation of xl and xr values will be used to obtain the depth value (Disparity) with the following equation:

$$Z = b.\frac{f}{d} \text{ atau } Z = b.\frac{f}{xl - xr}$$

III. Results

A. *Human Detection Using HPE*

The human pose estimation method produces human detection accompanied by 33 keypoints which represent key points or main human joints. Some results of human pose estimation are carried out at several test distances, namely at a distance of 3 meters, 10 meters, and 15 meters. The results of human pose estimation at each distance can be displayed through real-time frame capture results as follows:

Fig. 7. HPE result for distanse 10 meters

Fig. 8. HPE result for distanse 15 meters

From the results shown in Figure 9, Figure 10, and Figure 11, the performance of the Human Pose Estimation method in reconstructing human body landmarks is very good which is proven by still being able to display all landmarks up to a distance of 15 meters from the camera with an overall accuracy percentage value reaching 87.86%.

From testing each sample distance, the coordinate points of each landmarks can be displayed with reference to the world axis coordinate points, so that the coordinate values can be displayed as follows:

TABLE 1. Coordinate of HPE result

Nama Landmark	Coordinate X	Coordinate Y	Coordinate Z
NOSE	0,484570682	0,32476747	0,062780045
LEFT_EYE_INNER	0,481014103	0,314321607	0,032219123
LEFT_EYE	0,479105562	0,314212978	0,03217994
LEFT_EYE_OUTER	0,477933258	0,314178407	0,032100894
RIGHT_EYE_INNER	0,486914575	0,31314382	0,033409268
RIGHT_EYE	0,489103973	0,312111914	0,033410128
RIGHT_EYE_OUTER	0,491515338	0,310931385	0,03332321
LEFT_EAR	0,477923006	0,316209644	-0,047526378
RIGHT_EAR	0,495676517	0,31092611	-0,039888471
MOUTH_LEFT	0,483003229	0,331613809	0,044656001
MOUTH_RIGHT	0,488031328	0,331559151	0,046024628
LEFT_SHOULDER	0,444637001	0,375784755	-0,074849717
RIGHT_SHOULDER	0,538136601	0,37121287	-0,039233603

Adiratna Ciptaningrum, Andhika Putra Widyadharma, R. Akbar Nur Apriyanto, Imam Junaedi, Wahyu Pribadi, Edo Zulmi Faikhsan, and Larissa Kartika Putri, "Distance Measurement of Human and Vehicle Objects Using Computer Vision Based on The Combined Method of Triangulation and Pixel Comparison for Autonomous Braking System," *International Research Journal of Advanced Engineering and Science*, Volume 9, Issue 4, pp. 32-36, 2024.

LEFT_ELBOW	0,440573603	0,466553241	0,058506705
RIGHT_ELBOW	0,560204089	0,444338739	0,056192882
LEFT_WRIST	0,47590661	0,458525509	0,24272339
RIGHT_WRIST	0,54556644	0,453165501	0,208351806
LEFT_PINKY	0,480328411	0,460883737	0,266095579
RIGHT_PINKY	0,546043277	0,454565942	0,230086684
LEFT_INDEX	0,490305126	0,449717909	0,258167058
RIGHT_INDEX	0,534447074	0,447291225	0,233842656
LEFT_THUMB	0,491525114	0,449674278	0,244680688
RIGHT_THUMB	0,533788204	0,446683943	0,213954479
LEFT_HIP	0,476025105	0,559676886	-0,034280945
RIGHT_HIP	0,517970026	0,559381902	0,03414911
LEFT_KNEE	0,486131489	0,692632437	0,074214518
RIGHT_KNEE	0,512220621	0,690477252	0,153033346
LEFT_ANKLE	0,491363436	0,80463022	0,137598351
RIGHT_ANKLE	0,510865569	0,807070792	0,234266832
LEFT_HEEL	0,495917886	0,826310515	0,146494031
RIGHT_HEEL	0,511962891	0,830376685	0,245598793
LEFT_FOOT_INDEX	0,471964121	0,820119858	0,144912452
RIGHT_FOOT_INDEX	0,504499257	0,823541343	0,253106892

B. *Human Distance Result*

The distance results obtained are the results of distance estimation using computer vision with a distance variation of 10 meters and 15 meters measured from the camera. The results obtained are displayed as follows:

TABLE 2. Distance result in 10 meters

Frame	*Accuracy (%)*	*Performance*	*Distance (m)*	*Real Distance (m)*	*Error Distance (m)*
1	100	0,97245	9,91304	10	0,08695
2	100	0,97534	10	10	0
3	100	0,96307	9,73913	10	0,26086
4	100	0,96464	10	10	0
5	100	0,96394	9,82608	10	0,1739
6	100	0,97138	10	10	0
7	100	0,96332	10	10	0
8	100	0,95044	9,91304	10	0,0869
9	100	0,91895	10	10	0
10	100	0,90510	10,3333	10	0,33333
11	90,9090	0,90737	9,82608	10	0,17391
12	90,9090	0,88733	9,82608	10	0,17391
13	90,9090	0,86879	9,82608	10	0,17391
14	90,9090	0,88362	9,73913	10	0,26086
15	90,9090	0,88369	9,91304	10	0,08695
16	90,9090	0,88399	9,91304	10	0,08695
17	81,8181	0,86023	9,82608	10	0,17391
18	81,8181	0,85154	9,91304	10	0,08695
19	81,8181	0,85690	9,73913	10	0,26086
Rata-rata	**94,2583**	**0,91748**	**9,90770**	**10**	**0,12738**

The results obtained show that the highest accuracy level was achieved at a detection distance of 15 meters with an accuracy of 96.8102%, while at a distance of 10 meters, the accuracy was 94.2853%. Similarly, the highest detection performance was achieved at a 15-meter detection distance with a performance score of 0.9338, whereas at 10 meters, the performance score was 0.9174. These outcomes are significantly influenced by environmental conditions, lighting levels, and the color tone of the clothing worn by the individuals. In the distance estimation results, the accuracy was better at a distance of 10 meters with a difference of 0.12738 meters from the actual distance, compared to a difference of 0.28333 meters at 15 meters.

TABLE 3. Distance result in 15 meters

Frame	*Accuracy (%)*	*Performance*	*Distance (m)*	*Real Distance (m)*	*Error Distance (m)*
1	100	0,9697	14,8333	15	0,16667
2	100	0,9640	NaN	15	-
3	100	0,9719	NaN	15	-
4	100	0,9752	NaN	15	-
5	100	0,9641	NaN	15	-
6	100	0,9702	NaN	15	-
7	100	0,9737	15	15	0
8	100	0,9656	14,1667	15	0,833333
9	100	0,9769	14,5	15	0,5
10	100	0,9623	14,8333	15	0,16667
11	100	0,9686	15	15	0
12	100	0,968	14,6666	15	0,33333
13	100	0,9298	NaN	15	-
14	100	0,9198	NaN	15	-
15	84,8484	0,8372	15	15	0
16	100	0,8616	NaN	15	-
17	84,8484	0,8563	NaN	15	-
18	84,8484	0,8777	15	15	0
19	84,8484	0,8291	14,1666	15	0,83333
Rata-rata	**96,8102**	**0,9338**	**14,7167**	**15**	**0,28333**

C. *Vehicle Detection using Objet Tracking*

The object tracking results were obtained by detecting vehicles, specifically motorcycles, crossing the tracks at undefined distances. This was done to determine the confidence score in detecting the motorcycle. A sample of the vehicle detection results using object tracking is as follows:

Fig. 9. Result of Motorcycle detection using Object Tracking

The detection results yielded the lowest confidence score of 0.44 and the highest confidence score of 0.95. These values represent the detection performance on a scale from 0 to 1, where a score closer to 0 indicates poor detection performance, while a score closer to 1 indicates excellent detection performance.

D. *Vehicle Distance Result*

A sample of the vehicle's distance detection was taken at a distance of 10 meters, with the following results:

TABLE 4. Result of vehicle distance in 10 meters

Frame Count	*Class Name*	*Confidence*	*Distance (m)*	*Real Distance (m)*	**Selisih Jarak (m)**
1	*Motorcycle*	0,9273	9,75	10	0,25
2	*Motorcycle*	0,9276	9,75	10	0,25
3	*Motorcycle*	0,9472	9,8333	10	0,16667
4	*Motorcycle*	0,2762	10	10	0
5	*Motorcycle*	0,7669	9,91667	10	0,08333
6	*Motorcycle*	0,5871	9,91667	10	0,08333
7	*Motorcycle*	0,9059	10,1785	10	0,17829
8	*Motorcycle*	0,3780	9,91667	10	0,08333
9	*Motorcycle*	0,3705	9,83333	10	0,16667
10	*Motorcycle*	0,9353	10,1785	10	0,17857
11	*Motorcycle*	0,3085	9,91667	10	0,08333
12	*Motorcycle*	0,9145	9,91667	10	0,08333
13	*Motorcycle*	0,8723	9,91667	10	0,08333
14	*Motorcycle*	0,7992	9,83333	10	0,16667
15	*Motorcycle*	0,5670	9,83333	10	0,16667
16	*Motorcycle*	0,6338	10,1785	10	0,17429
17	*Motorcycle*	0,6261	10	10	0
18	*Motorcycle*	0,5129	9,75	10	0,25
19	*Motorcycle*	0,5351	9,75	10	0,25
Rata-rata	-	**0,6737**	**9,9141**	**10**	**0,14223**

At a detection distance of 10 meters, the object produced a confidence score of 0.6737, which was due to the object being relatively far from the camera. However, at this distance, the estimated distance value had a deviation of 0.14223 meters from the actual measured distance.

IV. CONCLUSION

The conclusion drawn from this research is that Computer Vision functions can be effectively used to measure the distance of human and vehicle objects from the camera by applying the triangulation concept. This is achieved through the pixel comparison method of the objects displayed in the detection frame.

The optimal distance obtained for human object detection was at 10 meters, with a distance deviation of 0.12738 meters from the actual distance. However, the accuracy and detection performance were more optimal at a detection distance of 15 meters, achieving an accuracy of 96.8102% and a detection performance score of 0.9338.

The vehicle detection system successfully detected a motorcycle with a confidence score of 0.95. At a detection distance of 10 meters, the system produced a distance estimation with a deviation of 0.14223 meters. Therefore, this human and vehicle detection system, along with the distance estimation, can be applied to the autonomous braking system of the inspection train of Madiun State Polytechnic.

REFERENCES

[1] Pakpahan, R., Analisa Pengaruh Implementasi Artificial Intelligence Dalam Kehidupan Manusia, JISICOM (Journal of Information System, Informatics and Computing), 5(2), 506-513, 2021.

[2] Kusuma, T., & Mulis, M. T., Perancangan Sistem Monitoring Infus Berbasis Mikrokontroler Wemos D1 R2. Konferensi Nasional Sistem Informasi (KNSI), 1422-1425, 2018.

[3] [3] Atmaja, D. S. & Kurniawan, M. A., Rancang Bangun Prototype Baru Pada Lori Inspeksi Generasi Dua Di Politeknik Perkeretaapian Indoensia Madiun. Jurnal Perkeretaapian Indonesia, 4(1), 21-28, 2020.

[4] Lay, A., & Lina, L. (2022). Pendeteksian Aktivitas Manusia Dengan Human Pose Estimation Dan Convolutional Neural Network. Computatio: Journal of Computer Science and Information Systems, 6(1), 51-60.

[5] Raju, K. (2022). Exercise detection and tracking using MediaPipe BlazePose and Spatial-Temporal Graph Convolutional Neural Network (Doctoral dissertation, Dublin, National College of Ireland).

[6] Zhang, Y. et al (2022). Bytetrack: Multi-Object Tracking By Associating Every Detection Box. In European Conference on Computer Vision (pp. 1-21). Cham: Springer Nature Switzerland.

[7] Hartley, R., & Zisserman, A. (2003). Multiple View Geometry In Computer Vision. Cambridge university press

[8] Hutauruk, J. S. W., Matulatan, T., & Hayaty, N. (2020). Deteksi kendaraan secara real time menggunakan metode YOLO berbasis android. Jurnal Sustainable: Jurnal Hasil Penelitian dan Industri Terapan, 9(1), 8-14.

[9] Mulyawan, Hendy., Samsono, M. Z. H., Setiawardhana. (2011). Identifikasi dan Tracking Objek Berbasis Image Processing Secara Real Time. EEPIS Final Project.

[10] Susanto, A., Mulyono, I. U. W., & Sudaryanto, S. (2024, January). Deteksi Gerak Pada Video Gerak Lansia Berbasis Yolo-V5 Dan Yolo-V7. In Semnas Ristek (Seminar Nasional Riset dan Inovasi Teknologi) (Vol. 8, No. 01).