Accredited Ranking SINTA 2
Decree of the Director General of Higher Education, Research, and Technology, No. 158/E/KPT/2021
Validity period from Volume 5 Number 2 of 2021 to Volume 10 Number 1 of 2026

Published online on: **http://jurnal.iaii.or.id**

JURNAL RESTI
(Rekayasa Sistem dan Teknologi Informasi)
Vol. 6 No. 5 (2022) 728 - 734 ISSN Media Electronic: 2580-0760

Improving AI Text Recognition Accuracy with Enhanced OCR For Automated Guided Vehicle

[1]Florentinus Budi Setiawan*, [2]Farrel Adriantama, [3]Leonardus Heru Pratomo, [4]Slamet Riyadi
[1,2,3,4]Department of Electrical Engineering, Faculty of Engineering, Universitas Katolik Soegijapranata, Indonesia
[1]f.budi.s@unika.ac.id, [2]18f10019@student.unika.ac.id, [3]leonardus@unika.ac.id, [4]riyadi@unika.ac.id

Abstract

AGV (Automated Guided Vehicle) with artificial intelligence (AI) is expected to change the industry's development in Indonesia, this artificial intelligence robot uses a mini-computer to operate it and uses mechanical movement like a four-wheeled vehicle with a 2WD drive system. In this article, a control strategy of the AGV robot will be shown and implemented to detect the location. This research Uses OCR (Optical Character Recognition) for the OpenCV library itself which has been enhanced/modified. This enhanced OCR is the main library used in text recognition. This research produces very accurate text detection compared to the default OCR that was previously used on the AGV robot in our university. After the process of reading this text is passed, it will produce text previously read through the camera which will then provide output in the form of text where the AGV robot is located. After the reading is validated, the AGV robot will move to the next point until it returns to its starting point. Based on hardware implementation through testing in the AGV laboratory with artificial intelligence, it can work according to the algorithm and minimize reading errors with a 95% success rate.

Keywords: *OCR, AGV, AI, Image Processing, Computer Vision*

1. Introduction

Developments in the industrial world in this world are growing quite rapidly artificial intelligence and machine learning are applied to robots for automation in industries in this world, especially in Indonesia [1,2]. One of the robots that is being developed in Indonesia, especially one that focuses on the industry, is the Automated Guided Vehicle (AGV), or a vehicle with automatic control [3]. This AGV is a robot equipped with Artificial Intelligence to follow a line or pattern programmed for implementation in this industry, many of which use laser sensors or LIDAR sensors that utilize reflected light reflections to detect locations to move in the desired direction [4]. This AGV is quite efficient and has quite a tough durability when compared to transportation equipment such as forklifts which still use operators to control it. This AGV can provide a high level of precision, and the operating mode is quite easy, reducing operational costs. This AGV is designed to be as reliable as possible according to industry needs [5].

The types of AGV that are widely used in the industry have several methods for their mode of operation, including using Path Guidance, which in this path guidance method is a simple navigation mode. For this path guidance, it is made to adjust the shape of the destination to be addressed at each point. For the mechanism itself, this AGV robot uses a DC motor as a rear-wheel-drive because this DC motor has a fairly easy way of controlling it compared to other electric motors, and for the steering wheel itself, it uses a servo motor [6,7]. This AGV research is based on a Raspberry PI 4 type B mini-computer. This raspberry can be used as a microcontroller because this raspberry has 40 more GPIO pins when compared to other microcontrollers [8].

This Raspberry also has a clock and RAM that is more advanced when compared to other microcontrollers or it can be said that this Raspberry has specifications similar to computers. The development of this AGV robot focuses on computer vision or image processing which is processed through a mini-computer using the Python programming language and by using the OpenCV library where the library uses OCR, namely Optical Character Recognition [9 - 12]. This OCR is a library that functions to detect a collection of words detected in images previously obtained from the camera which will then be processed analogously to digitally by a mini-computer. The Zernike moment was chosen because it is invariant for rotation and is not noise-

Accepted: 11-07-2022 | Received in revised: 10-09-2022 | Published: 01-10-2022

sensitive [13]. Another advantage of the Zernike moment is the ease of image reconstruction because it's orthogonal With the output produced by the OCR library [14,15]. It will be combined according to the purpose of making this AGV which makes it easier to detect the location of the AGV [16].

This study is a refinement of the AGV robot which was previously made at our university. The development of this AGV focuses on improving text reading/recognition by adding and modifying pre-processing and segmentation to the program where the advantage of using this enhanced pre-processing and segmentation is the presence of text tracking by providing direct reading results on the screen that records directly from the camera and can operate over long distances to recognize the text [17]. The disadvantage of the text reading system used in the previously made AGV robot is that reading text must be done close to the camera and the level of reading precision that can be said is not good enough, and the reading success rate is only 30% while using the operation engine mode (OEM) and page segmentation mode (PSM) from the OCR library will make the results messy when reading text directly and over long distances [18 - 20]. The OEM and PSM in this OCR library can function optimally when it is used to read jpg or png formatted images when used for Livestream purposes directly through the camera there will be many errors and can't even read. Pre-processing and segmentation added to the program will improve the quality of reading in the OCR library and minimize reading errors [21,22]. In the results and discussion section, I will present a comparison between the results of reading text using OEM and PSM from the default OCR and OCR libraries that I have modified to minimize reading errors with a reading success rate of 95% [23].

After reading this text can run smoothly then the next process will be connected to the driver of this vehicle. The drive for this AGV itself uses 2 types of motors, namely a DC motor as the rear-wheel-drive and a servo motor as a steering wheel that will control the AGV robot. This DC motor is used because the control mode is quite simple for programming. Meanwhile, the steering wheel uses a servo motor to control it by changing the angle of rotation by changing the PWM (Pulse Width Modulation) [24]. This study uses the OCR library but is modified in the program section to minimize reading errors and detect over long distances.

2. Research Methods

The flow of the research carried out by the author is shown in Figure 1. Based on Figure 1, this research method includes study literature, Design and Build Prototype, Results and Discussion, and conclusions based on the results of tests that have been carried out literature study was conducted to identify problems and find various solutions to solve these problems. After knowing the problem, The AGV that has been designed is implemented following the solutions to the problems found. the advantages of implementing text reading with the OCR library with the addition of pre-processing for transportation needs as the distribution of goods in the industry. This study also simulates AGV for testing if an error is found. The method used in this paper is as follows. Figure 1.

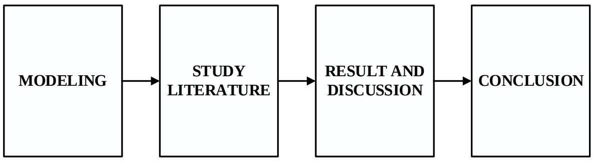


Figure 1. Research Flowchart

2.1 Modelling AGV Robot

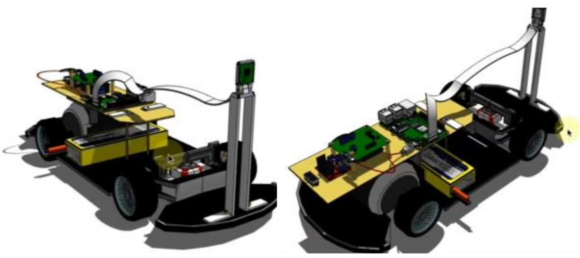


Figure 2. 3D Modelling AGV

In the 3d design that we made shown in Figure 2, we will implement it to the robot that we want to implement where the arrangement of electronic components is arranged on a board, the components needed are Raspberry Pi 4b, Camera for Raspberry, voltage divider Circuit, DC to DC Converter, DC motor, Servo Motor, and Li-Po Battery 11.1 V. AGV is a means of transportation in the form of a robot that is controlled automatically. AGV functions as a cargo transporting robot that uses a navigation system by following a guide or a line on the floor or using a laser reflection to move in a predetermined direction. At this time the AGV that we developed uses a navigation system in the form of text reading.

2.2 Block Diagram of Pattern Recognition

Figure 3 shows a block diagram of the text recognition system. This block diagram is a working sequence of reading text that will be used to detect the location that will be applied to the AGV. This tool is equipped with a cam built-in camera sensor with a raspberry pi that will function for video streams that'll then read predefined text. The power source is obtained from a converted battery using a step-down chopper to control this mini-computer or raspberry. The text is read by the camera when the video stream will be processed by the raspberry that has been previously programmed. After reading this text exactly at the location we want, this

DOI: https://doi.org/10.29207/resti.v6i5.4279

AGV will show where the location of the AGV is, so this raspberry will show it in the form of coordinates (x, y) which will then instruct us to move the motor that was previously programmed to get to that point.

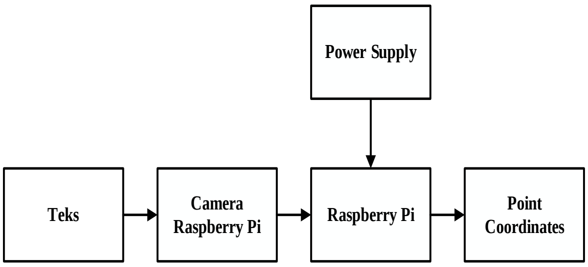


Figure 3. Pattern recognition system block diagram.

2.3. *Invariant Moment and Zernike Moment*

In pattern or object recognition, moment invariant can be used as a method that extracts a pattern or object. This method uses a two-dimensional function discrete, i.e. $f(x,y)$, in the digital form of ordinary moments $(p+q)$. This paper adopts moment invariant like on (1).

$$m_{pq} = \sum x \,(x-\bar{x})^p (y-\bar{y})^q f(x,y) \quad (1)$$

From (1), $\bar{x} = \sum_x x, \bar{y} = \sum_y y$ represents the center mass of the object. The origin of the first digital image $f(x,y)$assumed in translating to the center of mass of an object to achieve invariant translation. Similar to ordinary moments, Zernike moments (m,n) are used for digital images specified as in (2) following.

$$A_{nm} = \frac{n+1}{\pi} \sum x \ \sum_y f(x,y)\, V_{nm}(x,y), x^2+y^2 \leq 1 \quad (2)$$

From(2),$V_{nm}(x,y) = V_{nm}(\rho,\theta) = R_{nm}(\rho)e^{jm}, |m| \leq n$, and $n-|m|$ is an even number. Then $Rnm\,(\rho)$ can be defined As in (3).

$$R_{nm}(\rho) = \sum_{s=0}^{\frac{n-|m|}{2}} (-1)^s \frac{(n-s)!}{s!\left(\frac{n+m}{2}=s\right)!\left(\frac{n-m}{2}=s\right)!} \rho^{n-2s} \quad (3)$$

For real image function $f(x,y)$, only the term $V_{nm}(\rho,\theta)$ which is a complex number as in (4)

$$A_{nm} = \frac{n+1}{\pi} \sum_{xi} \sum_{yj} f(xi,yi) \quad (4)$$
$$[VR_{nm}(x,y) + JVI_{nm}(x,y)]\; xi^2 + yj^2 \leq 1$$

With VR and VI is a real function that represents the real and imaginary components of the function base $V_{nm}(\rho,\theta)$, which is converted into a (x,y) field. The system used by OCR consists of image acquisition, preprocessing, segmentation, feature extraction & selection, and recognition components as shown in Figure 4. Character text or images are obtained using a camera connected to a mini-computer. The image captured by the camera is saved in a grayscale format, but what is seen is processed into a binary image. Pre-processing module converts color/grayscale into binary images, skew detection correction, and digitalization. This process also detects the presence of noise and corrects the noise in the image. This noise can be caused because the quality of camera that is recording does not have good quality.

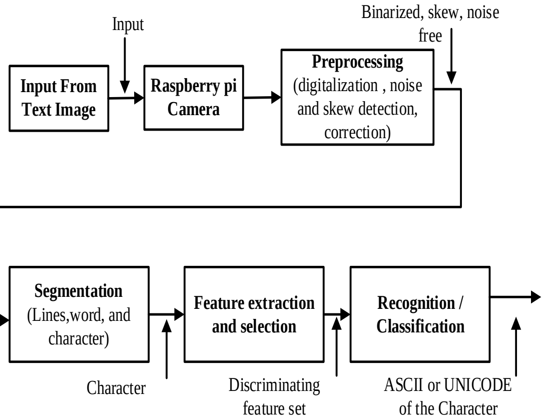


Figure 4. Block diagram of a typical OCR System

2.4 Text Detection Process

Figure 5 shows the flowchart of the text detection system design. After the camera is active for streaming, after the raspberry camera is active in the form of video, the camera will immediately detect the text in front of it. This text is used as a benchmark for the movement of the AGV from each specified point. If text is detected and visible on the monitor, a box will appear on the camera's video stream by generating the text that is checked by this OCR system according to the text read by the camera.

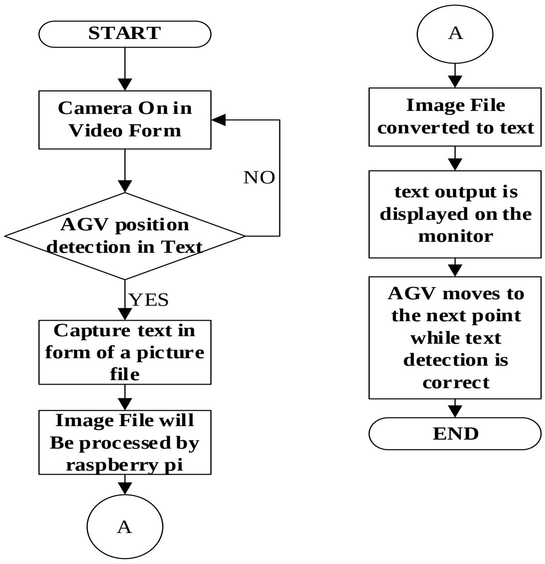


Figure 5. Flowchart of text detection Process

3. Results and Discussions

The prototype is shown in Figure 6. Consists of a raspberry pi 4 model B / mini-computer module that functions as a CPU to process images into text of course

DOI: https://doi.org/10.29207/resti.v6i5.4279

with a program that has been made in it, Camera board 8MP V2 Sony IMX 219 as input for recording directly when the AGV is operated, 11.1V LIPO battery as a source power supply, servo MG996R for steering control, voltage divider circuit using a 15 pin I/O connector cable, and using a DC motor as the rear-wheel drive. This test uses text printed on black paper to detect the location where the AGV is located.

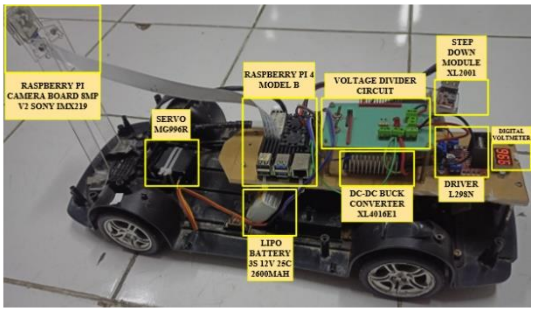


Figure 6. Prototype AGV 2WD

3.1 Enhanced OCR Library Test Result

Among all the phases of OCR preprocessing and segmentation is the most important phase because the accuracy of this OCR system has a considerable influence and is highly dependent on how well the preprocessing and segmentation are performed. So here we modify the preprocessing section and the segmentation section so that OCR readings can be carried out at a considerable distance and reduce read errors by 95%. In this program using an enhanced OCR library, the tracks used in this AGV will be shown in Figure 7. This AGV robot is tested indoors and the lighting is quite bright. When capturing images in the form of HVS paper by using the words “titikA”, “titikB”, “titikC”, “titikD”, “titikE”, “titikF”, “titikG”, and “titikP” are used as destination points by the AGV. Meanwhile, the initial pattern starting with the letter "P" is used as the point of the parking position or the beginning of the AGV.

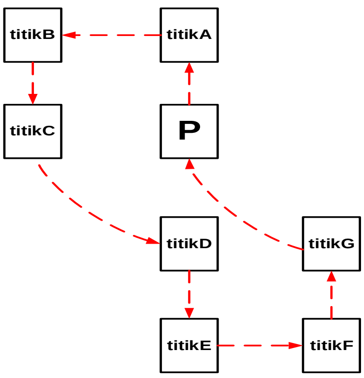


Figure 7. AGV Track / Route

In this discussion, I will compare the reading of the text with the operation engine mode (OEM) and page segmentation mode (PSM) built-in OCR which I will then compare with the OCR that I modified to minimize reading errors. In this comparison, I use single text that reads "titikA" and "titikB" in the default OCR I use OEM-3 and PSM-7 settings which have a function to read a single text.

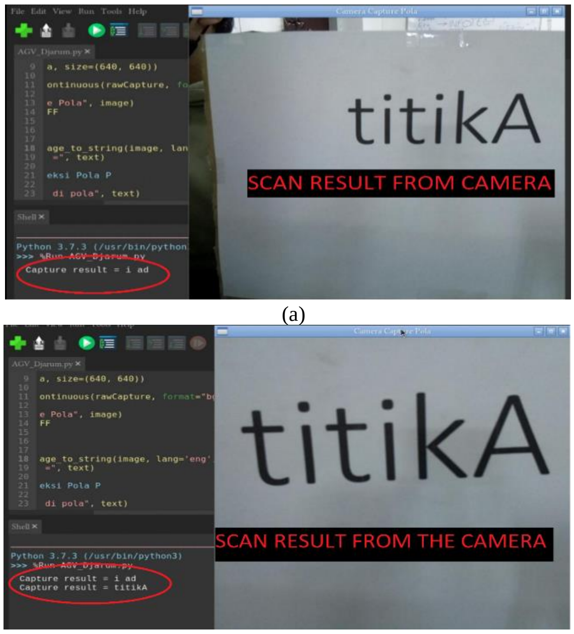


Figure 8. Detecting Point A with Default OCR

As seen in Figure 8, the reading of the text in part (a) the position of the text is taken from a distance away and produces inappropriate words, while in Figure 8 part (b) shows the text taken at a very close distance, if the image is taken at a distance very close, it will produce words that match those in the text taken

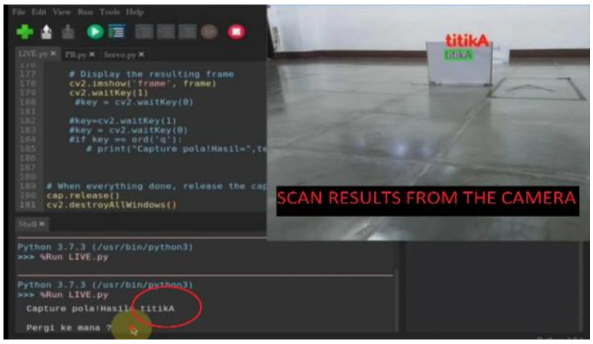
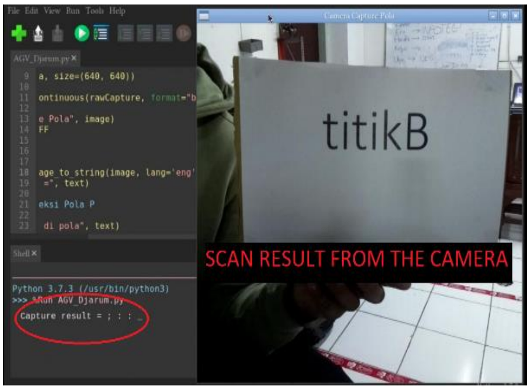


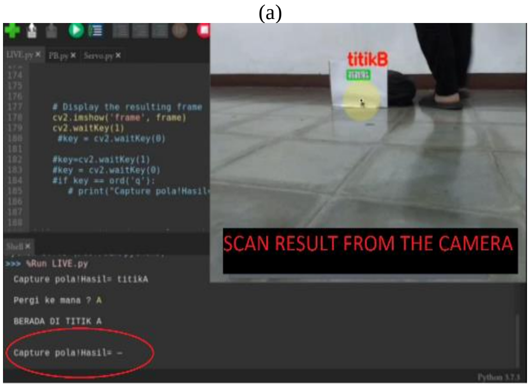
Figure 9. Detecting Point A with enhanced OCR

the readings we get with this enhanced OCR, it can read the text correctly and can be used at a distance of approximately 1 - 2 meters. This enhanced OCR also minimizes reading errors by 95% if the words read are not according to text, they will not be displayed on the camera output screen.

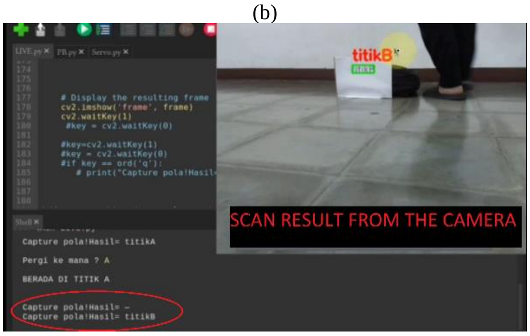
In Figure 10 part (a) is text detection using the default OCR where the results in part (a) cannot produce words that match what is captured by the camera.



(a)



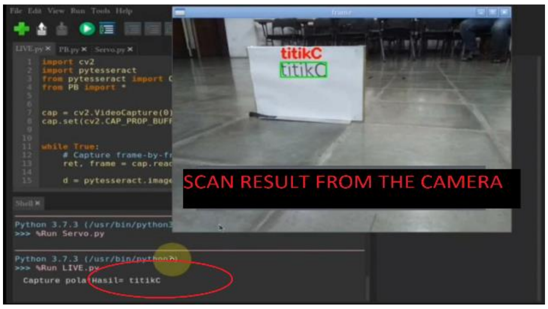
(b)



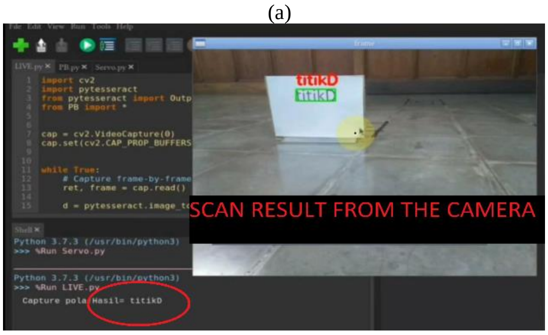
(c)

Figure 10.Comparison Between Default OCR With Enhanced OCR While Detecting Point B

The reading of the text in part (a) is not appropriate because the detection distance of the text can be said to be far enough to cause reading errors. An example can be seen in section (c) using an enhanced OCR and producing a suitable reading that can be detected within a distance of approximately 1 meter. In figure 10 part (b) when it detects the text "titikB" an error occurs, the program in the results produces "-" because it does not match with the text in front of it. If an error occurs like this it will repeat the reading until it is correct so the AGV will not have an error reading when it detects a location with text reading. Next, I will present text recognition using OCR which I have enhanced or modified.



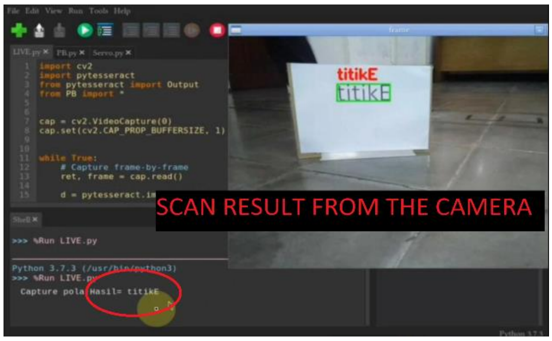
(a)



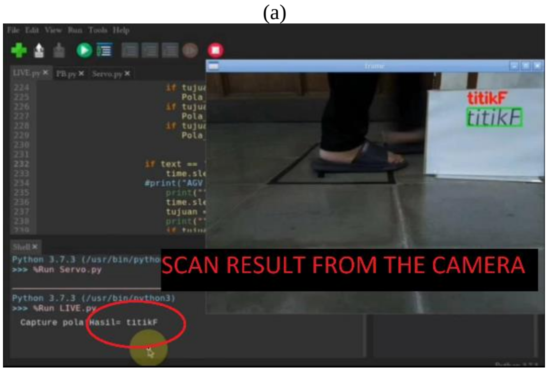
(b)

Figure 11. Detecting Point C and D with Enhanced OCR

As shown in Figure 11, the AGV robot camera detects the text "TitikC" and "TitikD" which will then read and process the image which will produce will be displayed on the monitor.



(a)



(b)

Figure 12. Detecting Points E and F with Enhanced OCR

The point is shown in Figure 12, the AGV robot camera detects the text "TitikE" and "TitikF" which will then read and process the image which will then produce text that will be displayed on the monitor.

DOI: https://doi.org/10.29207/resti.v6i5.4279

Table 1. Text Recognition Distance Test

Text Recognition Distance	Success Rate
0 meter – 0.5 meter	95 %
0.5 meter- 1 meter	95 %
1 meter – 1.5 meter	95 %
1.5 meter – 2 meter	95 %
> 2 meter	depending on lighting conditions and text size

Table 1 shows the success rate of text reading based on distance. Readings based on the table above have a success rate of 95% if within a distance of fewer than 2 meters. Text recognition in the range of 0 to 2 meters can be seen in the results and discussion section Exactly in the Enhanced OCR Library Test Result section. if reading the text from a distance of more than 2 meters this can be successful but must meet several conditions, namely lighting in the room In this room, 6 Tubular lamps are used, the recommended text size is more extensive for the size used is at 250pt, and the color in the room also affects text recognition In this trial, this AGV was tested in a room that has a lot of ornaments.

4. Conclusion

Producing text program works efficiently and can reduce reading errors in the text by up to 95% and works very well when the point reading goes from point A to P. This 95% success rate explains that any text that is read within a distance of fewer than 2 meters is always successful and has minimal reading errors. Of the many experiments that have been carried out by researchers reading this text, the ratio of reading success with reading errors can be said to be 20:1. When testing AGV Robots, it is infrequent to find significant reading errors.The program will show the results of the processing on the monitor screen. This program is also able to successfully display the results of OCR readings. if the enhanced OCR is compared to the default OCR, you will feel the difference in the final result, when using the enhanced OCR it will produce quite accurate results and can detect text at a fairly good distance, which is at a distance of approximately 1-2 meter. A program that has made and researched, it is more accurate and has almost no reading errors if compared with default OCR. In the future, this research will be used for industrial needs which will then be applied to AGVs in the industry.

Acknowledgment

Thank you to the Soegijapranata Catholic University for providing a place to conduct this research. We hope that this research will be a major contribution to the advancement of technology in Indonesia.

Reference

[1] S. Vechet, J. Krejsa, and K. S. Chen, "AGVs mission control support in smart factories by decision networks," *Proc. 2020 19th Int. Conf. Mechatronics - Mechatronika, ME 2020*, no. 1, pp. 3–6, 2020, doi: 10.1109/ME49197.2020.9286465.

[2] J. Zhang and X. Liu-Henke, "Model-based design of the vehicle dynamics control for an omnidirectional automated guided vehicle (AGV)," *15th Int. Conf. Mechatron. Syst. Mater. MSM 2020*, pp. 1–6, 2020, doi: 10.1109/MSM49833.2020.9202248.

[3] H. Song, "The Application of Computer Vision in Responding to the Emergencies of Autonomous Driving," *Proc. - 2020 Int. Conf. Comput. Vision, Image Deep Learn. CVIDL 2020*, no. Cvidl, pp. 1–5, 2020, doi: 10.1109/CVIDL51233.2020.00008.

[4] X. Zhou, T. Chen, and Y. Zhang, "Research on Intelligent AGV Control System," *Proc. 2018 Chinese Autom. Congr. CAC 2018*, no. 1, pp. 58–61, 2019, doi: 10.1109/CAC.2018.8623384.

[5] Y. Li and Y. Zhang, "Application research of computer vision technology in automation," *Proc. - 2020 Int. Conf. Comput. Inf. Big Data Appl. CIBDA 2020*, pp. 374–377, 2020, doi: 10.1109/CIBDA50819.2020.00090.

[6] K. Afsari and M. Saadeh, "Artificial Intelligence Platform for Low-Cost Robotics," *2020 3rd Int. Conf. Signal Process. Inf. Secur. ICSPIS 2020*, pp. 4–7, 2020, doi: 10.1109/ICSPIS51252.2020.9340156.

[7] Y. Ma, B. Jiang, and V. Cocquempot, "Modeling and Adaptive Fault Compensation for Two Physically Linked 2WD Mobile Robots," *IEEE/ASME Trans. Mechatronics*, vol. 26, no. 2, pp. 1161–1171, 2021, doi: 10.1109/TMECH.2021.3052900.

[8] Y. M. Su, H. W. Peng, K. W. Huang, and C. S. Yang, "Image processing technology for text recognition," *Proc. - 2019 Int. Conf. Technol. Appl. Artif. Intell. TAAI 2019*, pp. 1–5, 2019, doi: 10.1109/TAAI48200.2019.8959877.

[9] X. Zhang and S. Xu, "Research on Image Processing Technology of Computer Vision Algorithm," *Proc. - 2020 Int. Conf. Comput. Vision, Image Deep Learn. CVIDL 2020*, no. Cvidl, pp. 122–124, 2020, doi: 10.1109/CVIDL51233.2020.00030.

[10] M. S. Akopyan, O. V. Belyaeva, T. P. Plechov, and D. Y. Turdakov, "Text Recognition on Images from Social Media," *Proc. - 2019 Ivannikov Meml. Work. IVMEM 2019*, pp. 3–6, 2019, doi: 10.1109/IVMEM.2019.00006.

[11] G. Vaidya, K. Vaidya, and K. Bhosale, "Text recognition system for visually impaired using portable camera," *2020 Int. Conf. Converg. to Digit. World - Quo Vadis, ICCDW 2020*, no. Iccdw, pp. 1–4, 2020, doi: 10.1109/ICCDW45521.2020.9318706.

[12] T. T. H. Nguyen, A. Jatowt, M. Coustaty, N. Van Nguyen, and A. Doucet, "Post-OCR error detection by generating plausible candidates," *Proc. Int. Conf. Doc. Anal. Recognition, ICDAR*, pp. 876–881, 2019, doi: 10.1109/ICDAR.2019.00145.

[13] J. Zhang *et al.*, "A perspective transformation method based on computer vision," *Proc. 2020 IEEE Int. Conf. Artif. Intell. Comput. Appl. ICAICA 2020*, pp. 765–768, 2020, doi: 10.1109/ICAICA50127.2020.9182641.

[14] I. Ahmad, I. Moon, and S. J. Shin, "Color-to-grayscale algorithms effect on edge detection - A comparative study," *Int. Conf. Electron. Inf. Commun. ICEIC 2018*, vol. 2018-January, pp. 1–4, 2018, doi: 10.23919/ELINFOCOM.2018.8330719.

[15] Z. Zhong, L. Sun, and Q. Huo, "Improved Localization Accuracy by LocNet for Faster R-CNN Based Text Detection," *Proc. Int. Conf. Doc. Anal. Recognition, ICDAR*, vol. 1, pp. 923–928, 2017, doi: 10.1109/ICDAR.2017.155.

[16] L. Yang, D. Ergu, Y. Cai, F. Liu, and B. Ma, "A review of natural scene text detection methods," *Procedia Comput. Sci.*, vol. 199, pp. 1458–1465, 2021, doi: 10.1016/j.procs.2022.01.185.

[17] X. Li and Y. Shi, "Computer vision imaging based on artificial intelligence," *Proc. - 2018 Int. Conf. Virtual Real. Intell. Syst.*

ICVRIS 2018, pp. 22–25, 2018, doi: 10.1109/ICVRIS.2018.00014.

[18] F. Utaminingrum *et al.*, "Development of computer vision based obstacle detection and human tracking on smart wheelchair for disabled patient," *5th Int. Symp. Comput. Bus. Intell. ISCBI 2017*, pp. 1–5, 2017, doi: 10.1109/ISCBI.2017.8053533.

[19] N. Gupta and N. Goyal, "Machine Learning Tensor Flow Based Platform for Recognition of Hand Written Text," *2021 Int. Conf. Comput. Commun. Informatics, ICCCI 2021*, 2021, doi: 10.1109/ICCCI50826.2021.9402622.

[20] T. T. H. Nguyen, A. Jatowt, M. Coustaty, N. Van Nguyen, and A. Doucet, "Deep statistical analysis of OCR errors for effective post-OCR processing," *Proc. ACM/IEEE Jt. Conf. Digit. Libr.*, vol. 2019-June, no. 1, pp. 29–38, 2019, doi: 10.1109/JCDL.2019.00015.

[21] M. A. Panhwar, K. A. Memon, A. Abro, D. Zhongliang, S. A. Khuhro, and S. Memon, "Signboard detection and text recognition using artificial neural networks," *ICEIEC 2019 - Proc. 2019 IEEE 9th Int. Conf. Electron. Inf. Emerg. Commun.*, pp. 16–19, 2019, doi: 10.1109/ICEIEC.2019.8784625.

[22] Y. Qin and Z. Zhang, "Summary of Scene Text Detection and Recognition," *Proc. 15th IEEE Conf. Ind. Electron. Appl. ICIEA 2020*, pp. 85–89, 2020, doi: 10.1109/ICIEA48937.2020.9248121.

[23] Florentinus Budi Setiawan, F. A. Kurnianingsih, Slamet Riyadi, and Leonardus Heru Pratomo, "Pattern Recognition untuk Deteksi Posisi pada AGV Berbasis Raspberry Pi," *J. Nas. Tek. Elektro dan Teknol. Inf.*, vol. 10, no. 1, pp. 49–56, 2021, doi: 10.22146/jnteti.v10i1.738.

[24] F. B. Setiawan, O. J. Aldo Wijaya, L. H. Pratomo, and S. Riyadi, "Sistem Navigasi Automated Guided Vehicle Berbasis Computer Vision dan Implementasi pada Raspberry Pi," *J. Rekayasa Elektr.*, vol. 17, no. 1, pp. 7–14, 2021, doi: 10.17529/jre.v17i1.18087.

DOI: https://doi.org/10.29207/resti.v6i5.4279